paramètres

Clayton Peterson

De la logique des obligations, des permissions et des interdictions

De von Wright à aujourd'hui

Les Presses de l'Université de Montréal


paramètres

Clayton Peterson

DE LA LOGIQUE DES OBLIGATIONS, DES PERMISSIONS ET DES INTERDICTIONS

De von Wright à aujourd'hui

Les Presses de l'Université de Montréal

Ce livre fait partie d'un projet pilote de publication de livres numériques en libre accès, en collaboration avec la Direction des bibliothèques de l'Université de Montréal. Le format PDF de ce livre est téléchargeable gratuitement sur :

www.pum.umontreal.ca/catalogue/de-la-logique-des-obligations

Mise en pages : Clayton Peterson

Catalogage avant publication de Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada

Peterson, Clayton, 1985-

De la logique des obligations, des permissions et des interdictions : de von Wright à aujourd'hui

(Paramètres)

Comprend des références bibliographiques et un index.

ISBN 978-2-7606-3520-3

1. Logique déontique. I. Titre. II. Collection : Paramètres.

BC145.P47 2016 160 C2016-940940-6

Dépôt légal : 3e trimestre 2016

Bibliothèque et Archives nationales du Québec

ISBN (papier) 978-2-7606-3520-3

ISBN (PDF) 978-2-7606-3521-0

Les Presses de l'Université de Montréal remercient de leur soutien financier le Conseil des arts du Canada et la Société de développement des entreprises culturelles du Québec (SODEC).

Financé par le gouvernement du Canada | Canada

IMPRIMÉ AU CANADA

Avant-propos

Le présent ouvrage est le fruit de nombreuses années d'études en logique philosophique et résulte d'une revue de la littérature scientifique faite en vue de l'obtention de grades de second et de troisième cycles. Il s'adresse à un public qui n'est pas nécessairement familier avec la logique philosophique, mais qui désire s'initier à la vaste littérature sur la logique déontique, un domaine en expansion qui attire de plus en plus de chercheurs tant en philosophie qu'en droit. Idéalement, le lecteur aura une formation de base en logique philosophique, couvrant la logique classique et peut-être même le calcul de premier ordre. Une formation en logique modale serait certes idéale, mais j'ai rédigé ce livre de façon à m'adresser à un public aussi large que possible, ce qui inclut les étudiants et les chercheurs en mathématique, en informatique, en science humaine et en droit. Afin de pallier toute difficulté que pourraient poser les présupposés techniques de ce livre, le lecteur trouvera tout au long de la présentation des commentaires concernant les notions nécessaires à la compréhension du propos. En plus de la liste des symboles utilisés et de leur signification, il trouvera en fin de volume un index des concepts et des sujets.

Il existe peu ou pas d'ouvrages qui offrent une synthèse des principaux courants de la logique déontique. Or, c'est dans cet esprit que le présent livre a été rédigé. L'objectif est d'offrir une vue d'ensemble de la discipline, sans favoriser l'une ou l'autre des approches. Cet ouvrage est non seulement une carte conceptuelle permettant de s'orienter dans le vaste domaine de la logique déontique mais est aussi un ouvrage de vulgarisation, expliquant certaines notions nécessaires à la compréhension des différents sujets abordés en logique déontique.

Plusieurs personnes ont contribué à ce livre, que ce soit de près ou de loin, et ce parfois même sans le savoir. Je tiens à remercier Jean-Pierre Marquis qui a su orienter mes travaux et mon parcours universitaire. Le fait de m'avoir si bien « perverti » dès le début de ma formation a sans aucun doute déterminé ma vision de la logique. Merci aussi à François Lepage, Yvon Gauthier et Andrew Irvine pour leurs commentaires sur une version précédente de ce livre. Merci d'avoir participé au projet et d'avoir créé un

environnement stimulant et productif. Je tiens aussi à remercier France Houle pour son soutien et son intérêt envers le projet. En dernier lieu, je tiens à remercier les évaluateurs anonymes, sans qui le livre ne serait pas dans son état actuel.

Du côté personnel, un merci particulier à ma mère, Alberte, sans qui ce livre n'aurait probablement jamais vu le jour. Merci à mon épouse, Sarah-Geneviève, pour son écoute, son soutien et ses conseils. Merci aux membres de ma famille pour leurs éternels encouragements. Finalement, je tiens à remercier mon entourage qui manifeste sans cesse de l'intérêt pour mes travaux.

Cet ouvrage a été financé par le Conseil de recherches en sciences humaines du Canada et par les Fonds de recherche du Québec - Société et Culture.

CLAYTON PETERSON, Ph.D.

Chapitre 1

La logique déontique

La logique déontique se veut d'emblée une application de la logique aux discours éthique et juridique. Le terme « déontique » vient du grec $\delta\epsilon\acute{o}\nu\tau\iota\kappa$, où la racine $\delta\epsilon\acute{o}\nu$ signifie « ce qu'il convient de faire » et le suffixe $\iota\kappa$ signifie « qui porte sur ». Littéralement, l'expression « logique déontique » signifie donc « la logique qui porte sur ce qu'il convient de faire » (McNamara 2010). Par conséquent, la logique déontique est souvent présentée comme la logique qui traite des obligations, des permissions et des interdictions.

Un peu d'histoire

Même si l'étude des modalités déontiques se faisait à l'époque médiévale (Knuuttila 1981), la première tentative de formalisation est apparue avec Ernst Mally (1926).[1] C'est à la suite des travaux de von Wright (1951) que la logique déontique est devenue un domaine d'études en tant que tel.

Plusieurs objections ont été formulées contre l'approche initiale de ce chercheur. La première, avancée par Prior (1954), visait à montrer que sa notion d'*engagement* n'était pas adéquate dans les contextes mettant en jeu des obligations dérivées. Von Wright (1956) prit cette objection au sérieux et présenta la base de ce qu'on nommera plus tard la logique déontique dyadique, où on trouve deux propositions dans la portée de l'opérateur déontique (une qui indique le contexte de l'obligation et l'autre qui indique ce qui devrait être dans ce contexte). Même si la logique déontique dyadique a intéressé plusieurs philosophes à l'époque, notamment Rescher (1958), le système initial de von Wright est demeuré l'objet de plusieurs critiques. En 1958, Prior présenta le paradoxe du bon Samaritain, précurseur du paradoxe du voleur (Nozick et Routley 1962), qui visait à montrer que la relation de conséquence déontique au sein du système de von Wright menait à des résultats indésirables.

[1] Voir Føllesdal et Hilpinen (1970) et McNamara (2010) pour une introduction à la logique déontique.

Cela dit, Chisholm (1963) formula l'objection la plus sérieuse contre l'approche initiale de von Wright et montra que la logique déontique monadique n'est pas assez puissante pour rendre compte des obligations contraires au devoir. À la suite de ces objections, von Wright (1967) introduisit plusieurs systèmes de logique déontique dyadique afin de répondre aux objections touchant aux obligations conditionnelles.

Malgré cela, son approche initiale a néanmoins été reformulée après les avancées faites en logique modale et après l'introduction de la sémantique des mondes possibles, ce qui a donné lieu au fameux *système standard de logique déontique*, à savoir le système modal KD.[2] Par la suite, Forrester (1984) s'éleva contre le système standard afin de montrer que la relation de conséquence ne permettait pas de représenter adéquatement la conséquence déontique que nous trouvons au sein de nos inférences dans la langue naturelle.

L'objet de la logique déontique

D'entrée de jeu, la logique déontique visait la modélisation de la structure des inférences normatives. L'idée de von Wright était de faire un parallèle entre les notions d'obligation et de permission, d'un côté, et les modalités aléthique de nécessité et de possibilité, d'un autre. Ainsi, depuis l'article fondateur de von Wright, la logique déontique est souvent conçue comme une logique modale.

Bien qu'un tel mode de présentation soit réducteur, la logique est souvent présentée comme l'étude de la transmission des valeurs de vérité entre des énoncés.[3] Lorsque la logique est utilisée à des fins d'analyse d'arguments et de raisonnements, les connecteurs sont considérés comme vérifonctionnels et les propositions étudiées sont présupposées être *déclaratives*, c'est-à-dire qu'elles doivent minimalement avoir le potentiel d'être vraies ou fausses (Peterson 2013b, chapitre 3).

Pour déterminer si un énoncé p est déclaratif, il suffit de déterminer si la question « Est-ce que p est vrai ? » a un sens. Si oui, alors l'énoncé a la possiblité, en principe, d'être vrai ou faux. Dans ce cas, l'énoncé est déclaratif. Par exemple, les impératifs (p. ex. « Ouvre la porte ! ») et les énoncés interrogatifs ne sont pas déclaratifs. Parmi les énoncés déclaratifs, on trouve les propositions *descriptives*, qui portent sur comment le monde *est*, et les propositions *normatives*, qui disent comment le monde *devrait être*. Contrairement aux énoncés descriptifs, les propositions normatives

[2] Voir Åqvist (2002). Soulignons que le système initial de von Wright n'est pas équivalent au système standard.

[3] Le concept de vérité n'est pas indispensable en logique. Nous n'avons qu'à penser à la théorie de la preuve ou encore aux logiques d'action afin d'avoir des logiques qui ne portent pas nécessairement sur les énoncés déclaratifs.

sont prescriptives et visent à guider l'action. Notons que la notion d'énoncé déclaratif ne présuppose aucune théorie particulière de la vérité.

La logique déontique est souvent conçue comme une logique modale. Or une modalité $\Box$ est un opérateur qui transforme un énoncé déclaratif en un autre énoncé déclaratif. En ce sens, une modalité $\Box$ est une chose qui influence la valeur de vérité d'une proposition φ. Un exemple de modalité serait la modalité temporelle du *passé*. Au même titre que la proposition « Jean commet l'adultère » n'est pas vraie dans les mêmes conditions que « Jean a commis l'adultère », le premier énoncé n'est pas vrai dans les mêmes conditions que « Jean ne devrait pas commettre l'adultère ». De fait, lorsqu'elle est considérée en tant que logique modale, la logique déontique met en jeu des opérateurs, telles l'obligation, la permission et l'interdiction, qui modifient la valeur de vérité des énoncés déclaratifs.

En logique déontique, l'obligation, la permission et l'interdiction sont respectivement représentées par les opérateurs O, P et F.[4] Il s'agit d'*opérateurs* plutôt que de *prédicats* dans la mesure où les modalités déontiques peuvent être appliquées à des propositions pour former de nouvelles propositions (pouvant être d'un autre type). Initialement, chez von Wright, les opérateurs déontiques étaient appliqués à des *actions* plutôt qu'à des *propositions*. En ce sens, l'approche initiale de von Wright proposait un parallèle entre la logique déontique et la logique modale aléthique, sans pour autant présenter l'opérateur O comme une modalité, c'est-à-dire comme un opérateur qui influence la valeur de vérité d'une proposition déclarative.

Cette remarque nous amène à une distinction importante au sein de la littérature scientifique : les modalités déontiques peuvent être interprétées de deux manières distinctes et mutuellement exclusives (von Wright 1999). D'un côté, il y a l'interprétation de type *ought-to-do*, où les propositions qui se trouvent dans la portée des opérateurs déontiques font référence à des *actions*. Selon cette interprétation, les opérateurs déontiques expriment « ce qui doit être fait ». Considérant que ce sont les *actions* qui sont faites, et non les propositions, l'interprétation *ought-to-do* contraint la définition des énoncés bien formés : seule une action peut se trouver dans la portée d'un opérateur déontique, et, par conséquent, l'itération des opérateurs déontique n'est pas possible. En effet, selon cette interprétation et étant donné une action α, $O\alpha$ est une proposition, et non une action.

À l'opposé, il y a l'interprétation de type *ought-to-be*, où les propositions dans la portée des opérateurs déontiques sont déclaratives. Il s'agit là de l'interprétation *modale* de la logique déontique, où les opérateurs sont

[4] La logique déontique est majoritairement entendue comme une logique modale. Cela dit, soulignons qu'il y a néanmoins plusieurs approches qui ne la traitent pas ainsi. Outre ceux qui traitent les notions déontiques comme des prédicats, on trouve aussi la logique input/output.

vus comme des modalités qui influencent la valeur de vérité des énoncés. Selon cette interprétation, une modalité déontique indique ce qui doit *être* plutôt que ce qui doit être fait. Souvent, l'interprétation de type *ought-to-be* prendra place dans le cadre d'une sémantique des mondes possibles. Par exemple, alors que selon la première interprétation « il est obligatoire de respecter la loi » signifie que l'action *respecter la loi* doit être accomplie, selon la seconde, cela signifie plutôt que la description « la loi est respectée » doit être vraie.

Le dilemme de Jørgensen

La logique déontique, lorsqu'elle est interprétée en tant que logique modale, porte sur les énoncés déclaratifs. Or cela pose problème si on considère le dilemme formulé par Jørgensen (1937), lequel remet en cause la possibilité même d'appliquer l'analyse logique au discours normatif. À la base, le problème soulevé par Jørgensen repose sur la dichotomie sémantique qu'on retrouve entre *faits* et *normes*.[5] En d'autres termes, il s'agit d'une dichotomie sémantique entre les énoncés descriptifs et normatifs. Dans la littérature scientifique, on fait référence à ce problème en tant que thèse de l'être-devoir (*is-ought thesis*). Ce fossé sémantique entre les propositions normatives et descriptives, aussi connu sous le nom de *sophisme naturaliste*, consiste à dire qu'une conclusion normative ne peut pas être la conséquence d'un ensemble de prémisses purement descriptives, et à l'inverse qu'une conclusion descriptive ne peut être la conséquence de prémisses purement normatives. Autrement dit, un argument qui conclut un *devrait* à partir d'un *est* (ou vice versa) est toujours invalide.[6]

Par exemple, ce n'est pas parce qu'*il est interdit de voler* que nous pouvons conclure que *Pierre ne vole pas* : ce n'est pas parce qu'une personne a une obligation qu'elle va nécessairement la respecter ! Dans le même ordre d'idées, une personne peut très bien agir sans pour autant être obligée de le faire : ce n'est pas parce que Paul paie ses taxes qu'il est nécessairement obligatoire qu'il le fasse.

Outre la dichotomie sémantique, la difficulté relative au dilemme de Jørgensen repose sur la thèse réaliste de vérité correspondance. Cette thèse, dont on trouve l'équivalent formel dans un article de Tarski (1944), repose sur l'intuition que la valeur de vérité d'un énoncé dépend de sa correspondance avec le monde, voire de l'état actuel des choses. Selon cette thèse, une proposition comme « le ciel est bleu » est vraie si et seulement si le ciel *est*, dans le monde, effectivement bleu.

[5] Le problème, que Jørgensen analyse de façon plus contemporaine que ses prédécesseurs, apparaît entre autres chez Hume (1993) et Poincaré (1913).

[6] Un argument est valide lorsqu'il est impossible que les prémisses soient vraies alors que la conclusion est fausse (Peterson 2013b, chapitre 7).

Le dilemme de Jørgensen surgit lorsque la thèse réaliste de vérité correspondance est mise en conjonction avec la dichotomie sémantique. Considérant que les énoncés descriptifs sont, d'un point de vue sémantique, différents des propositions normatives, et que la valeur de vérité d'une proposition descriptive dépend de sa correspondance avec la réalité, l'attribution de valeur de vérité aux énoncés normatifs pose problème.

Si les énoncés normatifs n'ont pas pour fonction de décrire le monde et que la valeur de vérité d'une proposition déclarative dépend du fait que son contenu est conforme ou non avec la réalité, alors comment attribuer des valeurs de vérité aux propositions normatives ? Puisque les énoncés normatifs ne sont pas des énoncés qui décrivent la réalité, il s'ensuit que, en vertu de la thèse de vérité correspondance, une proposition normative ne peut pas être vraie ou fausse. Autrement dit, il semble que les énoncés normatifs ne soient pas déclaratifs.

En supposant que la logique traite de la transmission de valeurs de vérité entre les propositions, la dichotomie sémantique, conjointement à la thèse de vérité correspondance, pose problème. Étant fondée sur le postulat de vérifonctionnalité, la logique étudie la validité des raisonnements en observant les conditions de vérité d'un énoncé complexe, lesquelles dépendent de la valeur de vérité des énoncés atomiques qui le composent. En ce sens, la logique propositionnelle étudie la transmission de valeur de vérité entre les propositions et observe dans quelle mesure la valeur de vérité des atomes détermine celle des connecteurs logiques et des énoncés complexes.

Mais si la logique concerne la transmission des valeurs de vérité entre les propositions et que les propositions normatives n'ont pas de valeur de vérité, peut-on réellement faire une analyse logique des inférences normatives ? Si la logique est l'étude de la transmission des valeurs de vérité entre les propositions et que les énoncés normatifs n'ont pas de valeur de vérité, alors l'analyse logique ne s'applique pas aux propositions normatives.[7] Cependant, les inférences normatives semblent valides. Que ce soit dans le cas des raisonnements moraux ou des raisonnements juridiques, il semble y avoir des critères de validité qui permettent de juger de la valeur de nos inférences normatives.

[7] Notons ici que malgré les effets qu'a eus le dilemme au sein de la littérature scientifique, celui-ci ne met pas un point final aux tentatives de formalisation du discours normatif puisque la *vérité*, qu'on utilise comme un outil en logique, est utile mais n'est pas un concept indispensable.

En somme, le problème posé par Jørgensen se résume aux trois propositions suivantes :

1. les propositions normatives n'ont pas de condition de vérité ;
2. la logique traite de la transmission des valeurs de vérité entre les propositions ;
3. les raisonnements normatifs semblent néanmoins valides, notamment dans le cas des arguments légaux et moraux.

En guise de réponse à ce dilemme, on trouve trois pistes de solutions. Dans un premier temps, il est possible d'accepter les trois prémisses du dilemme et d'en conclure que la logique ne s'applique tout simplement pas aux normes. Ce genre de position est notamment défendu par Anderson (1999), qui prétend que nous n'avons besoin que du *sens commun* pour pouvoir évaluer la qualité de nos inférences normatives. Cependant, une telle attitude n'est pas très attrayante : si nous pensons que les inférences normatives peuvent prétendre à la validité, alors il faut dégager les règles selon lesquelles les raisonnements normatifs sont valides. Plusieurs sophismes semblent intuitivement *corrects* sans pour autant que ce soient des raisonnement acceptables. La logique est nécessaire pour juger de ces raisonnements et montrer en quoi ils sont inacceptables.

Une deuxième piste de solution serait de rejeter la première prémisse et de soutenir qu'il est possible d'attribuer des valeurs de vérité aux énoncés normatifs. Même si cette piste peut mener à des formes de réalisme[8], il n'en demeure pas moins que certains anti-réalistes pourraient être tentés de défendre une théorie minimale de la vérité.[9]

Finalement, la troisième option est de rejeter la seconde prémisse et de concevoir la logique d'une autre manière que comme la transmission de valeurs de vérité entre les propositions (p. ex. Alchourrón 1990).

Malgré les répercussions du dilemme au sein de la littérature scientifique, surtout en ce qui concerne l'analyse de la nature des propositions dans la portée des opérateurs déontiques et dans les inférences normatives, notons que celui-ci n'est pas fatal pour la logique déontique. En effet, les logiciens se préoccupent rarement de l'attribution actuelle de valeur de vérité aux propositions, la raison étant que l'analyse des conditions de validité d'un raisonnement en fait abstraction.

[8] Voir par exemple Walter (1996), lequel a notamment été critiqué par Stewart (1997) et Weinberger (1999).

[9] Voir par exemple Timmons (1999) pour une théorie minimale de la vérité, et Volpe (1999) pour une critique de ce genre de position. Notons que les théories minimalistes de la vérité tendent à utiliser le schéma d'équivalence de Tarski (1944) hors de son contexte.

L'étude des conditions de validité des raisonnements ne dépend pas de l'attribution actuelle de valeur de vérité aux propositions. Il est possible d'étudier les règles qui gouvernent la validité des inférences normatives sans pour autant statuer sur les conditions d'attribution de valeur de vérité des propositions normatives. La validité d'un raisonnement dépend de sa forme logique, et non de la valeur de vérité actuelle des propositions.[10] En ce sens, l'analyse de la forme logique des inférences normatives ne présuppose aucune position métaphysique ni ontologique.

En ce qui nous concerne, nous répondons au dilemme en rejetant la première proposition, bien que nous soyons aussi d'avis que l'objet de la logique ne se restreint pas aux énoncés déclaratifs. Quoi qu'il en soit, la vérité au sein des discours normatifs, du moins en droit, est le résultat d'une construction et dépend de ce qui est établi par le législateur et la jurisprudence (Baudoin 2010). En ce sens, la vérité des propositions normatives dépend de certaines normes, lesquelles sont établies par des autorités (Alchourrón et Bulygin 1981).

Pourquoi la logique déontique ?

Avant d'aborder les différents courants en logique déontique, certains pourraient d'emblée se demander en quoi celle-ci est nécessaire à l'analyse du discours.[11] Les outils du calcul propositionnel et ceux de la logique de premier ordre tels qu'ils sont enseignés aux étudiants en philosophie ne sont-ils pas suffisants à l'analyse (de la forme) des inférences ? Cette question, loin d'être naïve, trouve aisément réponse lorsqu'on réfléchit à l'utilité de la logique en vue de l'analyse des raisonnements.

Selon le degré d'abstraction qu'on adopte, certaines choses peuvent se voir comme différentes ou semblables. Pour un logicien, l'étude d'un raisonnement peut être faite à plusieurs niveaux, selon le degré d'abstraction adopté. Le niveau le plus général est celui du calcul propositionnel. Grossièrement, l'idée est de prendre comme bloc de base l'*atome propositionnel*, c'est-à-dire la proposition indivisible, et de construire les propositions complexes à l'aide des atomes et des connecteurs logiques.[12] Les connecteurs logiques de base sont $\neg$, $\wedge$, $\vee$ et $\supset$, qui dénotent respectivement la négation (non), la conjonction (et), la disjonction (ou) et l'implication (si, alors). Alors que la négation est un connecteur unaire, c'est-à-dire

[10] La force d'un raisonnement, quant à elle, dépend de la valeur de vérité des propositions. Voir Peterson (2013b).

[11] Sur ce sujet, le lecteur est invité à consulter Peterson (2013a). De plus, pour une introduction à la logique propositionnelle et à la logique de premier ordre, il peut consulter Tomassi (1999), Lepage (2010) ou encore Arthur (2011).

[12] Le lecteur est invité à consulter Peterson (2013b) quant à l'utilisation de la logique classique et du calcul de premier ordre pour l'analyse des raisonnements.

qui ne s'applique qu'à une seule proposition, la conjonction, la disjonction ainsi que l'implication sont des connecteurs binaires, qui mettent en jeu deux propositions.[13]

Le niveau propositionnel est le plus grossier et se limite à étudier la structure des raisonnements en fonction des connecteurs qui lient les propositions atomiques. Cela dit, il est possible de raffiner l'analyse et d'observer les relations logiques qui se trouvent *à l'intérieur* des propositions atomiques, plutôt que de se limiter aux relations logiques qui se trouvent *à l'extérieur* des atomes, lesquelles s'expriment par les connecteurs logiques usuels (négation, conjonction, disjonction, implication). En allant voir ce qui se passe dans l'atome, on adopte un niveau d'analyse plus fin et on observe les relations entre les concepts. Il s'agit là du point de vue de la logique de premier ordre.

Par exemple, considérons l'argument suivant. Le symbole $\therefore$ est utilisé pour marquer la conclusion.

1. Tous les hommes sont mortels et Socrate est un homme.

$\therefore$ Socrate est mortel.

Selon le point de vue propositionnel, la forme logique de cet argument est la suivante :

$$\begin{array}{ll} 1. & p \wedge q \\ \hline \therefore & r \end{array}$$

Il s'agit d'une prémisse où une conjonction met en jeu deux propositions atomiques et à l'aide de laquelle on conclut une autre proposition atomique. Quiconque est au fait des règles de la logique propositionnelle sera en mesure de démontrer que cette forme logique est invalide. Néanmoins, lorsqu'on raffine l'analyse et qu'on adopte la perspective du calcul des prédicats, on peut voir que ce raisonnement fonctionne. En calcul des prédicats, on augmente le langage en introduisant des prédicats, des variables, des constantes ainsi que les quantificateurs universel $\forall$ et existentiel $\exists$, signifiant respectivement « pour tout » et « il existe ». Selon ce point de vue, la forme logique de l'argument est la suivante :

$$\begin{array}{ll} 1. & \forall x(Hx \supset Mx) \wedge Hs \\ \hline \therefore & Ms \end{array}$$

[13] Les lettres p, q, r, etc. seront utilisées comme variables propositionnelles pour faire référence à des propositions atomiques ; alors que φ, ψ, ρ, etc. seront utilisées comme méta-variables.

Cela dit, le degré d'analyse de la logique du premier ordre possède ses limites et l'étude de certains raisonnements requiert une analyse plus fine. Considérons l'argument suivant :

1. Il est interdit de dépasser la limite de vitesse sur l'autoroute (100 km/h).
2. Nécessairement, si Jean conduit à 120 km/h, alors Jean dépasse la limite de vitesse.

∴ Jean ne doit pas conduire à 120 km/h.

Au niveau propositionnel, la forme de cet argument est :

1. p
2. $q \supset r$

∴ s

Selon le point de vue de la logique de premier ordre, il est cependant moins évident de déterminer la forme logique du raisonnement. Bien que la première prémisse puisse se traduire par un prédicat marquant les actions interdites, la conclusion ne peut pas se traduire dans le langage de premier ordre. En effet, la négation qui s'y trouve ne porte pas sur la proposition en entier mais porte plutôt sur l'action qui se trouve à l'intérieur de la portée du terme « doit ». Pour mettre ce point en évidence, il suffit de considérer que les deux propositions suivantes ne sont pas logiquement équivalentes :

– Jean ne doit pas conduire à 120 km/h.
– Il est faux que Jean doive conduire à 120 km/h.

Ainsi, l'expression « ne doit pas » n'indique pas une négation propositionnelle, et en ce sens la conclusion ne peut être traduite à l'aide d'un prédicat nié.

Or si plutôt que d'interpréter les notions d'obligation, de permission et d'interdiction en tant que prédicats on les considère comme des opérateurs, il devient possible d'introduire des connecteurs logiques dans leur portée. S'il s'agit d'une interprétation *ought-to-do*, alors il s'agira d'une algèbre d'action, sinon il sera question de connecteurs propositionnels. En logique déontique standard augmentée à l'aide d'un opérateur de nécessité $\Box$[14], l'argument susmentionné possède la forme suivante :

1. Fp
2. $\Box(q \supset p)$

∴ $O\neg q$

[14] Sur ce point, voir Peterson et Marquis (2012).

Considérant la définition de l'interdiction par $F\varphi =_{df} O\neg\varphi$, il devient possible de montrer la validité de ce raisonnement. L'utilité de la logique déontique apparaît donc lorsqu'on examine attentivement la forme logique des énoncés et qu'on constate les limites de la logique classique et du calcul des prédicats.

Structure et objectifs

De manière générale, l'application de la logique au discours normatif vise à répondre à certaines questions sémantiques et épistémologiques. Lorsqu'elle est utilisée aux fins de l'analyse du discours normatif, la logique est un outil qui permet d'étudier certaines questions épistémologiques et ontologiques sous un nouvel angle.

Exemple

1. *Quelle(s) relation(s) ont les jugements normatifs avec le monde ?*
2. *Dans quelles conditions est-ce que les propositions normatives sont vraies ?*
3. *Quelles sont les propriétés des concepts normatifs ?*
4. *Quelle est la structure du discours normatif ?*
5. *Dans quelle mesure est-ce que les propositions normatives existent ?*

Ces questions servent de point de départ à l'investigation logico-philosophique. Bien que cela ne soit pas toujours explicite dans la littérature scientifique, la logique déontique peut être utilisée à différentes fins. Comme mentionné, son objectif initial est de déterminer les conditions de validité des inférences normatives en explicitant leur structure. En ce sens, la logique déontique peut être vue comme l'analyse de la structure des inférences normatives, tant éthiques que légales. Cette perspective s'ancre autant dans la logique philosophique que dans la philosophie formelle et la pensée critique. Il s'agit de la perspective que nous adopterons tout au long du présent ouvrage.

Cela dit, au-delà de cet objectif, on peut utiliser la logique déontique pour d'autres buts. En plus de son utilité du point de vue de l'argumentation, elle peut aussi être utilisée pour modéliser les situations où des normes et des contraintes régissent le comportement de plusieurs agents. Selon cet angle d'approche, la logique déontique est un outil visant à étudier l'évolution des systèmes normatifs. Cette perspective est pertinente notamment en informatique et en droit.

Par ailleurs, on peut utiliser la logique déontique pour étudier la structure d'un ensemble de normes, comme dans les cas où on cherche à étudier la structure des lois. Ce genre d'analyse permet la représentation de la connaissance légale et la construction de bases de données.

Finalement, la logique déontique peut servir à la programmation, tant par la modélisation des contraintes (obligations, permissions et interdictions d'un programme informatique) que par le développement de l'intelligence artificielle.[15]

Évidemment, la visée d'un objectif plutôt que d'un autre déterminera en partie les caractéristiques de la logique déontique qu'on doit utiliser. Pareillement, le domaine d'application visé orientera les critiques et les objections qu'on peut faire envers les différents systèmes. Or l'objectif du présent ouvrage est d'introduire le lecteur à la logique déontique du point de vue de l'analyse des inférences normatives. La littérature scientifique sur le sujet étant vaste, l'objectif est de fournir un outil offrant une synthèse des divers courants qu'on y retrouve.

La présentation est faite toujours en ayant en tête l'utilité de la logique déontique pour l'analyse de l'argumentation. Ainsi, nous présenterons des approches ayant des visées différentes mais nous exposerons leurs limites du point de vue de l'analyse du discours. Soulignons que, du point de vue de l'analyse des inférences, notre tâche n'est pas descriptive mais est bien prescriptive : l'objectif n'est pas d'examiner ni de décrire la structure des inférences normatives comme elles se font au quotidien, mais plutôt de déterminer les conditions dans lesquelles les inférences normatives sont valides. En ce sens, lors de l'analyse de la structure d'un raisonnement normatif, l'objectif est de déterminer la structure que *devrait* avoir une inférence normative.

Étant donné le nombre important de publications qu'on trouve sur le sujet, le présent ouvrage est évidemment incomplet et ne couvre pas la totalité des approches existantes. Néanmoins, ce texte a été rédigé avec un souci d'exhaustivité, l'objectif étant de fournir au lecteur le matériel de base pour pouvoir aborder le sujet de front. La logique déontique étant généralement conçue comme une logique modale, nous avons choisi de commencer par la présentation des notions de logique propositionnelle et de logique modale. De plus, considérant que les paradoxes de la logique déontique ont une place centrale au sein de la littérature scientifique, nous avons choisi de les exposer avant de présenter les différents types d'approches, ce qui nous permettra par la suite de positionner ces dernières en fonction des problèmes qu'elles cherchent à résoudre.

Afin de rendre l'écriture plus homogène, nous avons opté pour une notation constante, et de fait nous avons parfois modifié la notation des textes originaux lorsque cela ne prêtait pas à confusion. De plus, soulignons l'usage qui sera fait des termes *cohérence* et *consistance*, lesquels ne signi-

[15] Voir Peterson (2014a) pour une liste plus complète des différentes applications de la logique déontique.

fient pas la même chose en logique philosophique. Tout au long du texte, nous utiliserons le terme *consistance* dans les contextes où il est question de *consistance syntaxique*, c'est-à-dire de non-contradiction formelle ($\nvdash \bot$), où la consistance est la propriété d'un système formel. Nous utiliserons le terme *cohérence* dans des contextes sémantiques où la consistance n'est pas définie, comme dans le cas de la langue naturelle. En droit, par exemple, il y a une présomption de cohérence du discours juridique. En un mot, l'idée est de dire que si un discours est cohérent dans la langue naturelle, alors il y a possibilité de définir un système formel consistant qui permet de le représenter. Si le discours juridique prétend à la cohérence d'un point de vue sémantique, alors il est possible de représenter cette cohérence à l'aide d'un système logique syntaxiquement consistant.

L'idée principale qui a guidé la rédaction de ce livre est d'offrir une lecture critique de la littérature scientifique. L'angle que nous avons adopté pour mener à terme ce projet est celui de la logique philosophique et de la pensée critique, ayant en tête l'utilité de la logique aux fins de l'argumentation. En philosophie, les arguments à l'étude sont des arguments *normatifs*, et non *descriptifs*. En ce sens, le présent ouvrage est une introduction à la logique déontique, qui se veut essentiellement l'étude des discours normatifs.

L'analyse logique d'un discours consiste dans l'examen de sa structure formelle, c'est-à-dire des relations logiques entre les différents concepts d'un domaine particulier. D'un point de vue conceptuel, la logique est un outil puissant qui permet de délimiter le cadre à l'intérieur duquel une théorie opère. La formalisation d'un discours permet non seulement d'étudier les relations entre les concepts, mais permet aussi d'étudier la cohérence du domaine ainsi que la dépendance entre les propositions. La cohérence, et *a fortiori* la logique, est l'un des principaux critères de rationalité. Dans une perspective théorique, l'analyse logique des discours éthique et juridique vise à solidifier leurs fondements, et de fait la pertinence de la logique appliquée aux sphères de l'éthique et du droit est considérable.

Cette revue critique de la littérature scientifique est abordée selon deux objectifs généraux. D'une part, le but est de présenter un aperçu historique de l'origine et de l'évolution de la logique déontique. Pour ce faire, nous avons tâché d'offrir un récit descriptif mettant en évidence les différentes ramifications au sein de la littérature scientifique. D'autre part, l'objectif est de donner une vue d'ensemble de la littérature scientifique, offrant ainsi au lecteur un outil efficace lui permettant d'aborder avec sérénité ce vaste sujet qu'est la logique déontique.

Chapitre 2

Les systèmes standard

Avant d'aborder la logique déontique en tant que telle, voyons d'abord diverses notions élémentaires de logique modale.

La logique modale

Cette dernière est une extension de la logique propositionnelle classique, à laquelle on ajoute une modalité (unaire) vérifonctionnelle $\Box$ dont la valeur de vérité dépend de la proposition φ dans sa portée. La logique propositionnelle classique LPC est construite à partir du langage :

$$\mathcal{L}_{LPC} = \{(,), \supset, \bot, Prop\}$$

La notation $\{\ldots\}$ est utilisée afin de dénoter un ensemble. L'ensemble $Prop = \{p_1, ..., p_n, ...\}$ contient un nombre dénombrable[1] de propositions atomiques. Les énoncés bien formés EBF_{LPC} sont définis récursivement :

$$\varphi := \bot \mid p_i \mid \varphi \supset \psi$$

Cette notation se lit de la manière suivante :

1. $\bot \in EBF_{LPC}$;
2. si $p \in Prop$, alors $p \in EBF_{LPC}$;
3. si $\varphi, \psi \in EBF_{LPC}$, alors $\varphi \supset \psi \in EBF_{LPC}$;

La notation $\in$ est utilisée afin d'indiquer qu'il s'agit d'un élément de l'ensemble. Par exemple, la première clause signifie que $\bot$, qui est une formule qui dénote le faux, est dans l'ensemble des énoncés bien formés de la logique propositionnelle classique. À l'inverse, $\notin$ indique que la proposition n'est pas un élément de l'ensemble.

[1] Équivalent à la cardinalité de l'ensemble des nombres naturels. Il s'agit d'une infinité qu'on peut énumérer.

À partir de l'implication matérielle et de $\bot$, les autres connecteurs peuvent être définis par :

$$\neg\varphi =_{df} \varphi \supset \bot \qquad \text{(df. } \neg\text{)}$$
$$\varphi \wedge \psi =_{df} \neg(\varphi \supset \neg\psi) \qquad \text{(df. } \wedge\text{)}$$
$$\varphi \vee \psi =_{df} \neg\varphi \supset \psi \qquad \text{(df. } \vee\text{)}$$
$$\varphi \equiv \psi =_{df} ((\varphi \supset \psi) \wedge (\psi \supset \varphi)) \qquad \text{(df. } \equiv\text{)}$$

Ici, $\equiv$ est utilisé pour dénoter le biconditionnel, voire l'équivalence. En vertu de la définition de la négation, nous obtenons que φ est faux lorsque φ implique la formule qui dénote le faux.

Le système modal K est une extension de LPC, défini à partir du langage :

$$\mathcal{L}_K = \{(,), \supset, \bot, \Box, Prop\}$$

Les énoncés bien formés EBF_K sont définis récursivement par :

$$\varphi := \bot \mid p_i \mid \varphi \supset \psi \mid \Box\varphi$$

L'opérateur dual de $\Box$ est défini par (df.$\Diamond$) :

$$\Diamond\varphi =_{df} \neg\Box\neg\varphi \qquad \text{(df.}\Diamond\text{)}$$

Dans le cas de la logique déontique, les opérateur $\Box$ et $\Diamond$ sont remplacés respectivement par O et P, où O représente l'obligation et P correspond à la notion de permission *faible* (implicite), par opposition avec la permission *forte* (explicite). L'opérateur F, qui représente l'interdiction, est défini par :

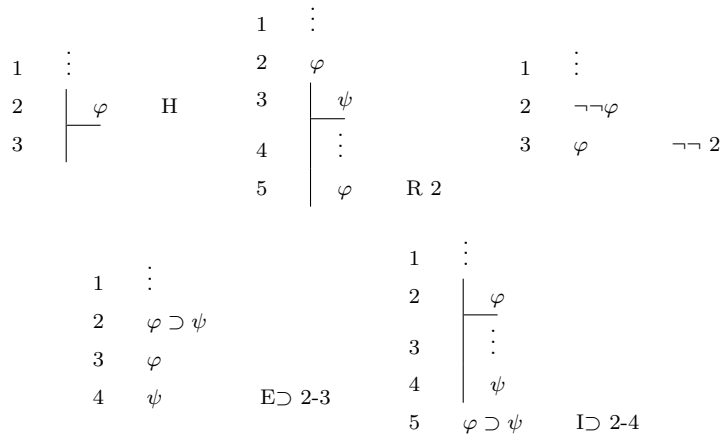
$$F\varphi =_{df} O\neg\varphi \qquad \text{(df. F)}$$

Parmi les caractéristiques notables de K, soulignons que l'itération des opérateurs est possible (p. ex. $\Box\Box\varphi \in EBF_K$) et que les équivalences logiques peuvent être substituées dans la portée de l'opérateur $\Box$.

Déduction naturelle

Suivant Garson (2006), la définition de LPC en déduction naturelle se fait par les règles exposées à la figure 2.1. Ces règles permettent la dérivation des règles usuelles pour les autres connecteurs logiques (figure 2.2).

FIGURE 2.1
Règles de LPC



Exercice 2.1. À l'aide des règles de LPC en déduction naturelle et des définitions des autres connecteurs logiques, prouvez chaque règle dérivée de LPC.

Exercice 2.2. Faites la preuve des théorèmes suivants en déduction naturelle.

$$(\varphi \wedge \psi) \equiv \neg(\neg\varphi \vee \neg\psi) \tag{2.1}$$
$$(\varphi \vee \psi) \equiv \neg(\neg\varphi \wedge \neg\psi) \tag{2.2}$$
$$\neg(\varphi \wedge \psi) \equiv (\neg\varphi \vee \neg\psi) \tag{2.3}$$
$$\neg(\varphi \vee \psi) \equiv (\neg\varphi \wedge \neg\psi) \tag{2.4}$$
$$\bot \supset \psi \tag{2.5}$$

Prises ensembles, les formules 2.3 et 2.4 représentent les dualités de De Morgan.

En déduction naturelle, K s'obtient lorsqu'on ajoute les règles qui gouvernent l'introduction (I$\Box$) et l'élimination (E$\Box$) de l'opérateur $\Box$ (figure 2.3).

FIGURE 2.2
Règles dérivées de LPC

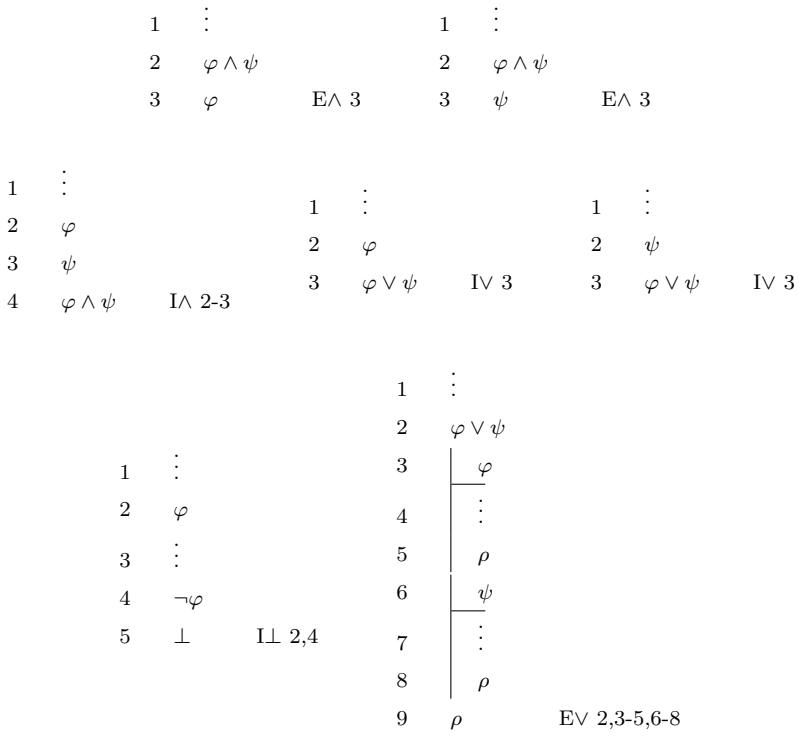
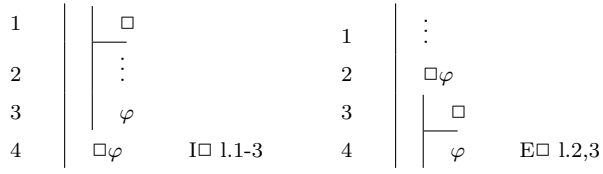


FIGURE 2.3
Règles de K



Système axiomatique

Outre sa présentation en déduction naturelle, la logique modale est aussi parfois définie en tant que système axiomatique. Pour ce faire, on définit LPC à partir de $\mathcal{L}_{LPC}$ comme l'ensemble contenant les axiomes $PL1$, $PL2$ et $PL3$ et étant fermé sous la règle *modus ponens* (MP) :

$$\varphi \supset (\psi \supset \varphi) \tag{$PL1$}$$
$$(\varphi \supset (\psi \supset \rho)) \supset ((\varphi \supset \psi) \supset (\varphi \supset \rho)) \tag{$PL2$}$$
$$(\neg\psi \supset \neg\varphi) \supset ((\neg\psi \supset \varphi) \supset \psi) \tag{$PL3$}$$

$$\frac{\vdash \varphi \supset \psi \qquad \vdash \varphi}{\vdash \psi} \qquad \text{(MP)}$$

La notation $\vdash \varphi$ indique que φ est un théorème (ici un théorème de LPC), c'est-à-dire une formule dérivable à partir des règles et des axiomes. Ce symbole dénote la conséquence syntaxique.

Exercice 2.3. Prouvez que $\neg\neg\varphi \supset \varphi$ est un théorème de LPC.

Afin d'obtenir la présentation axiomatique de K, on ajoute le schéma d'axiome (Dist) pour la distribution de $\Box$ sur l'implication et la règle (Nec) :

$$\frac{\vdash_K \varphi}{\vdash_K \Box\varphi} \qquad \text{(Nec)}$$

$$\vdash_K \Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi) \tag{Dist}$$

Soulignons la différence entre une *règle* et un *schéma d'axiome*. Alors que la règle ne s'applique qu'aux théorèmes, c'est-à-dire aux propositions démontrables sans hypothèse, un schéma d'axiome s'applique à tous les énoncés bien formés. Par exemple, $\vdash_K \Box(p \supset q) \supset (\Box p \supset \Box q)$ est un théorème de K, puisqu'il s'agit d'une instanciation de l'axiome de distribution, mais l'utilisation suivante de la règle n'est pas adéquate, puisque p n'est pas un théorème de LPC ni de K ($\nvdash_K p$).

$$\frac{p}{\Box p}$$

Logique intuitionniste

Les définitions de la logique propositionnelle classique susmentionnées mettent en jeu divers connecteurs logiques définis sur la base de $\supset$ et $\bot$. Cela dit, afin d'expliciter le parallèle entre la logique propositionnelle intuitionniste et la logique propositionnelle classique, il est possible de considérer chaque connecteur logique comme primitif.

Considérons le langage de la logique propositionnelle intuitionniste $\mathcal{L}_{LI} = \{(,), \wedge, \supset, \vee, \bot, Prop\}$. La négation est définie par $\neg\varphi =_{df} \varphi \supset \bot$ et les énoncés bien formés EBF_{IL} sont définis récursivement par :

$$\varphi := \bot \mid p_i \mid \varphi \supset \psi \mid \varphi \wedge \psi \mid \varphi \vee \psi$$

La logique propositionnelle intuitionniste peut être définie de plusieurs manières équivalentes. Notamment, Goldblatt (2006) définit la logique intuitionniste à l'aide de (MP) ainsi que des schémas d'axiomes suivants.

$$\varphi \supset (\varphi \wedge \varphi) \tag{$IL1$}$$
$$(\varphi \wedge \psi) \supset (\psi \wedge \varphi) \tag{$IL2$}$$
$$(\varphi \supset \psi) \supset ((\varphi \wedge \rho) \supset (\psi \wedge \rho)) \tag{$IL3$}$$
$$((\varphi \supset \psi) \wedge (\psi \supset \rho)) \supset (\varphi \supset \rho) \tag{$IL4$}$$
$$\psi \supset (\varphi \supset \psi) \tag{$IL5$}$$
$$(\varphi \wedge (\varphi \supset \psi)) \supset \psi \tag{$IL6$}$$
$$\varphi \supset (\varphi \vee \psi) \tag{$IL7$}$$
$$(\varphi \vee \psi) \supset (\psi \vee \varphi) \tag{$IL8$}$$
$$((\varphi \supset \rho) \wedge (\psi \supset \rho)) \supset ((\varphi \vee \psi) \supset \rho) \tag{$IL9$}$$
$$\neg\varphi \supset (\varphi \supset \psi) \tag{$IL10$}$$
$$((\varphi \supset \psi) \wedge (\varphi \supset \neg\psi)) \supset \neg\varphi \tag{$IL11$}$$

Exercice 2.4. Prouvez que les schémas d'axiomes ($IL1$)-($IL11$) sont dérivables au sein de la présentation de LPC en déduction naturelle.

Exercice 2.5. Prouvez que les schémas d'axiomes ($IL1$)-($IL11$) sont dérivables au sein de la présentation de LPC en système axiomatique.

Par ailleurs, Kleene (1952) définit la logique intuitionniste à l'aide de (MP) et des schémas d'axiomes suivants.

$$\psi \supset (\varphi \supset \psi) \tag{ILa}$$
$$(\varphi \supset \psi) \supset ((\varphi \supset (\psi \supset \rho)) \supset (\varphi \supset \rho)) \tag{ILb}$$
$$\varphi \supset (\psi \supset (\varphi \wedge \psi)) \tag{ILc}$$
$$(\varphi \wedge \psi) \supset \varphi \tag{ILd}$$
$$(\varphi \wedge \psi) \supset \psi \tag{ILe}$$
$$\varphi \supset (\varphi \vee \psi) \tag{ILf}$$
$$\varphi \supset (\varphi \vee \psi) \tag{ILg}$$
$$(\varphi \supset \rho) \supset ((\psi \supset \rho) \supset ((\varphi \vee \psi) \supset \rho)) \tag{ILh}$$
$$(\varphi \supset \psi) \supset ((\varphi \supset \neg\psi) \supset \neg\varphi) \tag{$ILi4$}$$
$$\neg\varphi \supset (\varphi \supset \psi) \tag{ILj}$$

Exercice 2.6. Prouvez que les schémas d'axiomes (ILa)-(ILj) sont dérivables dans la présentation de LPC en déduction naturelle.

Exercice 2.7. Prouvez que les schémas d'axiomes (ILa)-(ILj) sont dérivables dans la présentation axiomatique de LPC.

Exercice 2.8. Prouvez que les définitions de Goldblatt et de Kleene sont équivalentes. Autrement dit, prouvez que les schémas d'axiomes ($IL1$)-($IL11$) sont dérivables à partir de (MP) et de (ILa)-(ILj), et vice versa.

Que ce soit à partir de la présentation de Goldblatt ou de Kleene, la logique propositionnelle classique s'obtient par l'ajout de l'un ou l'autre des schémas d'axiomes suivants.

$$\neg\varphi \vee \varphi \tag{2.6}$$
$$\neg\neg\varphi \supset \varphi \tag{2.7}$$

L'équation 2.6 correspond au principe du tiers exclus. En logique intuitionniste, le tiers exclus est logiquement équivalent à l'élimination de la double négation (équation 2.7).

Extensions de K

Différentes extensions de K s'obtiennent par l'ajout de certains schémas d'axiomes. Parmi les plus connus, on trouve :

$$\Box\varphi \supset \Diamond\varphi \quad (\mathbf{D})$$
$$\Box\varphi \supset \varphi \quad (\mathbf{T})$$
$$\Box\varphi \supset \Box\Box\varphi \quad (\mathbf{4})$$
$$\Diamond\varphi \supset \Box\Diamond\varphi \quad (\mathbf{5})$$
$$\varphi \supset \Box\Diamond\varphi \quad (\mathbf{B})$$

Les systèmes les plus courants sont les suivants :

$$KD = K + (\mathbf{D})$$
$$KT = K + (\mathbf{T})$$
$$K4 = K + (\mathbf{4})$$
$$KB = K + (\mathbf{B})$$
$$K5 = K + (\mathbf{5})$$
$$S4 = K4 + (\mathbf{T})$$
$$S5 = K5 + (\mathbf{T})$$

Sémantique des logiques modales

Ayant débuté dans les années cinquante avec les travaux de Kanger, Hintikka et Montague, Kripke (1959, 1963) a profondément influencé la sémantique des logiques modales, maintenant connue sous le nom de *sémantique de Kripke* ou encore de *sémantique des mondes possibles* (Wolenski 1990). La sémantique des mondes possibles est une sémantique ensembliste où l'attribution de valeur de vérité à une proposition dépend de son appartenance à un ensemble et des relations qui existent entre eux.

Suivant Garson (2006), l'idée est de construire un modèle $\mathcal{M} = \langle U, R, a\rangle$ à partir d'une structure $\mathcal{S} = \langle U, R\rangle$. Au sein de la structure S, U est un univers (non vide, informellement vu comme un ensemble de scénarios ou encore de mondes) et R est un ensemble de relations au sein de U, c'est-à-dire que $R \subseteq U \times U$ est un ensemble de paires ordonnées. L'univers contient des scénarios w[2], lesquels sont des ensembles de propositions. Formellement, les scénarios sont des sous-ensembles de $Prop$, voire des éléments de l'ensemble des parties de $Prop$, c'est-à-dire que $w \in \wp(Prop)$, où $\wp(Prop)$ est l'ensemble de tous les sous-ensembles de $Prop$.

[2] Notons que w est utilisé pour *world*.

Ici, la notation $\times$ indique le produit cartésien entre deux ensembles. Formellement, ayant deux ensembles A et B, le produit cartésien est donné par $A \times B = \{< a, b >: a \in A \text{ et } b \in B\}$. Cette définition doit être comprise de la manière suivante : l'ensemble $A \times B$ contient toutes les paires $< a, b >$ telles que a est élément de A et b est élément de B. Autrement dit, l'ensemble contient toutes les paires ordonnées possibles pouvant être formées à partir des éléments de A et de B. L'ensemble R contient donc des paires de mondes, voire de scénarios possibles.

La notation $\subseteq$ dénote la relation d'inclusion, où $A \subseteq B$ indique que A est un sous-ensemble de B. $A \subseteq B$ signifie que tout élément appartenant à A appartient aussi à B. Formellement, pour tout φ, si $\varphi \in A$, alors $\varphi \in B$. Notons que la relation d'inclusion doit être différenciée de la relation d'inclusion propre. Un sous-ensemble propre est dénoté par $\subset$. Alors que la relation d'inclusion $A \subseteq B$ admet la possibilité que tous les éléments de B soient aussi des éléments de A, $A \subset B$ ne donne pas cette possibilité.

Pour obtenir le modèle $\mathcal{M}$, on ajoute à la structure une fonction $a : Prop \times U \longrightarrow \{\top, \bot\}$ qui attribue des valeurs de vérité aux propositions en fonction des scénarios de U. Une proposition φ est dite *vraie* pour un scénario $w \in U$ à condition qu'elle soit un élément de w :

$$a_w(\varphi) = \top \Leftrightarrow \varphi \in w$$

Cela peut aussi être dénoté par $\models_w \varphi$, analogue de $\vdash$, pour faire référence à la notion de conséquence sémantique (par opposition à syntaxique).

Par le postulat de bivalence, si $\varphi \notin w$, alors $a_w(\neg\varphi) = \top$. La clause sémantique quant à l'opérateur modal est que $\Box\varphi$ est vrai pour un scénario w si et seulement si φ est vrai pour tout scénario v en relation R avec w. Formellement, nous avons :

$$a_w(\Box\varphi) = \top \Leftrightarrow a_v(\varphi) = \top \; \forall v \text{ t.q. } wRv$$

L'interprétation dans la sémantique des mondes possibles tend à considérer w comme le monde *désigné* ou *actuel* et voit les scénarios v comme des *alternatives possibles* à w, c'est-à-dire des scénarios qui pourraient logiquement prendre place à la suite de w.[3]

Cela dit, lorsqu'on y réfléchit en termes d'*alternatives possibles* au monde actuel, l'interprétation normative de la sémantique de Kripke pose un grave problème. En effet, si la valeur de vérité d'une proposition $\Box\varphi$ pour un scénario w dépend du fait que dans toutes les alternatives possibles à w

[3] Une description *totale* du monde *actuel* est problématique. Un scénario w se veut une description qui *correspond* à celui-ci. Par surcroît, la sémantique des mondes possibles n'implique aucun engagement ontologique (Hansson 1969).

la proposition φ est vraie, il s'ensuit que $\Box\varphi$ ne sera jamais vraie puisqu'il est toujours possible pour un agent d'agir à l'encontre de ses obligations. Autrement dit, considérant que des scénarios où des agents enfreindraient leurs obligations sont des alternatives possibles au monde actuel, il s'ensuit qu'il y a toujours des alternatives possibles où la proposition φ n'est pas une description adéquate de la réalité, et de fait φ est fausse dans ces scénarios.

Dès lors, l'interprétation sémantique de la logique déontique en termes de *mondes possibles* requiert qu'on restreigne les alternatives aux alternatives *déontiques*, voire aux mondes possibles *idéaux*, c'est-à-dire aux alternatives où les obligations sont réalisées.[4] En ce sens, une proposition normative $O\varphi$ est dite vraie pour un scénario w dans la mesure où la proposition φ est réalisée (est vraie) dans toutes les alternatives idéales à w. L'interprétation modale s'insère donc dans la catégorie des approches du types *ought-to-be*, lesquelles considèrent que les propositions normatives indiquent des états de choses qui devraient être.

Hansson (2006) a critiqué le concept d'*alternative déontique idéale* et soutient principalement que l'idéalisation ne permet pas de rendre compte de la structure des normes et que les *mondes idéaux* ne nous fournissent pas d'information sur la question de savoir comment nous devrions agir dans le monde actuel.

Par ailleurs, cette notion de perfection a donné lieu à une interprétation de la sémantique des mondes possibles en termes de *préférences*, ce qu'on peut voir notamment dans les travaux de Åqvist (1986), de Hansson (1990) et de van Fraassen (1972). Le point central de ces approches est l'introduction d'une relation $\succcurlyeq$ qui permet d'ordonner les différentes alternatives déontiques possibles, indiquant que certaines alternatives sont préférables à d'autres.

Une caractéristique importante des modèles de Kripke est que les différents schémas d'axiomes induisent des relations précises sur les modèles sémantiques. Par exemple, le système modal KD est fiable et complet relativement à la classe des modèles $\mathcal{M}$ qui mettent en jeu une relation sérielle, c'est-à-dire où, pour tout scénario w, il existe un scénario v tel que wRv. Autrement dit, tout scénario possède une alternative idéale.

Par *fiable*, nous entendons que tout théorème de KD est valide pour la classe des modèles $\mathcal{M}$ qui mettent en jeu une relation sérielle (où *valide* signifie vrai pour tout modèle de cette classe), et par *complet*, nous entendons que toute conséquence valide pour la classe de modèles $\mathcal{M}$ qui mettent en jeu une relation sérielle est dérivable dans KD. Autrement dit, l'ensemble des théorèmes de KD correspond précisément à l'ensemble

[4] Voir par exemple Bonevac (1998), Chellas (1974), Feldman (1986), Jones et Pörn (1985) ou encore Opałek et Woleński (1991)

des conséquences sémantiques valides au sein de la classe des modèles qui mettent en jeu une relation d'accessibilité sérielle.

Le tableau 2.1, inspiré de Garson (2006) et Åqvist (2002), présente les relations induites sur les modèles sémantiques par les axiomes. Le symbole $\Rightarrow$ est utilisé pour marquer une relation d'implication (si, alors) tandis que $\Leftrightarrow$ indique aussi la réciproque (si et seulement si). Le symbole & signifie « et ».

TABLEAU 2.1
Les relations induites sur les modèles sémantiques

	Axiome	Condition sur $\mathcal{M}$	Relation
(**D**)	$\Box\varphi \supset \Diamond\varphi$	$\forall w\ \exists v$ t.q. wRv	Sérielle
(**T**)	$\Box\varphi \supset \varphi$	$\forall w,\ wRw$	Réflexive
(**4**)	$\Box\varphi \supset \Box\Box\varphi$	$\forall w, v, u\ (wRv\ \&\ vRu) \Rightarrow wRu$	Transitive
(**5**)	$\Diamond\varphi \supset \Box\Diamond\varphi$	$\forall w, v, u\ (wRv\ \&\ wRu) \Rightarrow vRu$	Euclidienne
(**B**)	$\varphi \supset \Box\Diamond\varphi$	$\forall w, v\ wRv \Rightarrow vRw$	Symétrique
S5	(**T**) + (**5**)	$\forall w, v\ wRv$	Équivalence
(□**T**)	$\Box(\Box\varphi \supset \varphi)$	$\forall w, v\ wRv \Rightarrow vRv$	Quasi-réflexive
(□**B**)	$\Box(\varphi \supset \Box\Diamond\varphi)$	$\forall w, v, u\ (wRv\ \&\ vRu) \Rightarrow uRv$	Quasi-symétrique
(**CD**)	$\Diamond\varphi \supset \Box\varphi$	$\forall w, v, u\ wRv\ \&\ wRu \Rightarrow v = u$	Unique
(**C4**)	$\Box\Box\varphi \supset \Box\varphi$	$\forall w, v\ wRv \Rightarrow\ \exists u(wRu\ \&\ uRv)$	Dense
(**C**)	$\Diamond\Box\varphi \supset \Box\Diamond\varphi$	$\forall w, v,\ (wRv\ \&\ wRu) \Rightarrow$ $\exists x(vRx\ \&\ uRx)$	Convergente
(**L**)	$\Box(\Box\varphi \supset \psi)\vee$ $\Box((\psi \wedge \Box\psi) \supset \varphi)$	$\forall w, v, u\ (wRv\ \&\ wRu) \Rightarrow$ vRu, uRv ou $v = u$	Antisymétrique

Notons que selon l'approche préférée, la sémantique peut aussi être définie récursivement en faisant référence aux ensembles qui contiennent les scénarios où les propositions sont vraies. En prenant par exemple l'ensemble

$$\|V_\varphi\| = \{w : \models_w \varphi\},$$

on peut reformuler la clause sémantique par

$$a_w(\varphi) = \top \Leftrightarrow\ w \in \|V_\varphi\|,$$

ou encore définir récursivement $\|V_\psi\|$ à partir des atomes et des connecteurs logiques :

$$\begin{aligned}
\|V_p\| &= \{w : \models_w p\} \\
\|V_{\neg\varphi}\| &= \{w : w \in U - \|V_\varphi\|\} \\
\|V_{\varphi\wedge\psi}\| &= \{w : w \in \|V_\varphi\| \cap \|V_\psi\|\} \\
\|V_{\varphi\vee\psi}\| &= \{w : w \in \|V_\varphi\| \cup \|V_\psi\|\} \\
\|V_{\varphi\supset\psi}\| &= \{w : w \in \|V_{\neg\varphi}\| \cup \|V_\psi\|\}
\end{aligned}$$

Ici, le symbole $\cap$ marque l'intersection entre deux ensembles. Formellement, ayant deux ensembles A et B, $A \cap B = \{x : x \in A \text{ et } x \in B\}$. En ce sens, l'intersection de A et B contient les éléments qui se trouvent à la fois dans A et dans B.

Par ailleurs, la notation $A - B$ marque le complément ensembliste, contenant les éléments de A qui ne sont pas éléments de B. Formellement, cet ensemble est défini par $A - B = \{x : x \in A \text{ et } x \notin B\}$.

Finalement, le symbole $\cup$ marque l'union de deux ensembles. Ayant deux ensembles A et B, $A \cup B = \{x : x \in A \text{ ou } x \in B\}$. Il s'agit de l'ensemble qui contient les éléments qui se trouvent soit dans A soit dans B.

Systèmes normaux et fortement normaux

Suivant Åqvist (2002), les systèmes *normaux* de logique déontique sont ceux qui admettent minimalement comme théorèmes ceux du système modal K, et les systèmes *fortement normaux* sont ceux qui admettent minimalement les théorèmes de KD. Un système logique peut être considéré comme un ensemble de propositions contenant certains axiomes et fermé sous certaines règles. En ce sens, un système Δ de logique déontique est normal lorsque $K \subseteq \Delta$, et fortement normal lorsque $KD \subseteq \Delta$.

Les travaux de von Wright (1951), conjointement avec les développements en logique modale, ont mené à ce qu'on nomme aujourd'hui les *systèmes standard de logique déontique*. Un système de logique déontique Δ est dit *standard* lorsque celui-ci appartient à la classe d'équivalence des systèmes dont le représentant est KD. La classe d'équivalence d'un système se veut un ensemble qui contient toutes les logiques permettant de prouver tous et seulement tous les théorèmes du système.[5] Autrement dit, un système Δ est dit *logiquement équivalent* à Σ lorsque Δ permet de dériver toutes les règles et tous les schémas d'axiomes de Σ, et que, vice versa, Σ permet de dériver toutes les règles et tous les schémas d'axiomes de Δ. En ce sens, le système standard est la logique modale KD, qui regroupe tous les autres systèmes qui lui sont syntaxiquement équivalents. Selon cette définition, un système standard est fortement normal.

Cela dit, il faut faire la distinction entre *le* système standard et *les* systèmes standard. Alors que le système standard est le système modal KD, la classe des systèmes dits *standard* correspond à la classe des systèmes fortement normaux.

[5] Une traduction adéquate d'un langage à l'autre peut être nécessaire.

Notons que la définition proposée par Åqvist pour les systèmes normaux est équivalente à celle proposée par Chellas (1980). Soulignons aussi que le système que présente von Wright (1951) n'est pas équivalent au système standard, notamment parce que von Wright rejette les propositions de la forme $O\top$, où $\top$ est une tautologie.[6]

Outre l'axiome (**D**), plusieurs autres peuvent être ajoutés à K pour créer différents systèmes de logique déontique (Åqvist 2002). On trouve notamment 10 systèmes normaux de logique déontique monadique, soit K, OM, $OS4$, OB, $OS5$, KD, OM^+, $OS4^+$, OB^+, $OS5^+$, dont cinq fortement normaux, où $^+$ signifie l'ajout du schéma d'axiome (**D**) :

$$\begin{aligned} OM &= K + (\Box\mathbf{T}) \\ OS4 &= K + OM + (\mathbf{4}) \\ OB &= K + OM + (\Box\mathbf{B}) \\ OS5 &= K + (\mathbf{4}) + (\mathbf{5}) \end{aligned}$$

Les systèmes normaux et fortement noramaux sont des logiques *monadiques*, c'est-à-dire où $\Box$ est unaire. En ce sens, l'opérateur $\Box$ ne met en jeu qu'une proposition dans sa portée.

Hiérarchie des logiques modales

Bien que les logiques modales soient souvent construites sur la base de K, il s'avère que K n'est pas la logique modale la plus faible. En effet, outre les systèmes *normaux*, dont le représentant est le système modal K, on trouve aussi les systèmes classiques, monotones et réguliers (Chellas 1980). Considérons les règles et schémas d'axiomes suivants :

$$\frac{\vdash \varphi \equiv \psi}{\vdash \Box\varphi \equiv \Box\psi}\ (\mathsf{RE}) \qquad \frac{\vdash \varphi \supset \psi}{\vdash \Box\varphi \supset \Box\psi}\ (\mathsf{RM}) \qquad \frac{\vdash (\varphi \wedge \rho) \supset \psi}{\vdash (\Box\varphi \wedge \Box\rho) \supset \Box\psi}\ (\mathsf{RR})$$

$$\frac{\vdash \varphi}{\vdash \Box\varphi}\ (\mathsf{RN}) \qquad \frac{\vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \psi}{\vdash (\Box\varphi_1 \wedge \cdots \wedge \Box\varphi_n) \supset \Box\psi}\ (\mathsf{RK}) \quad \text{avec } n \geq 0$$

$$\Box(\varphi \wedge \psi) \supset (\Box\varphi \wedge \Box\psi) \qquad (\mathsf{M})$$
$$(\Box\varphi \wedge \Box\psi) \supset \Box(\varphi \wedge \psi) \qquad (\mathsf{C})$$
$$\Box(\varphi \wedge \psi) \equiv (\Box\varphi \wedge \Box\psi) \qquad (\mathsf{R})$$
$$\Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi) \qquad (\mathsf{K})$$
$$\Box\top \qquad (\mathsf{N})$$

[6] Il s'agit du principe de contingence normative. Von Wright (1983) acceptera par la suite ce type d'énoncé en adoptant le schéma d'axiome $O(\varphi \vee \neg\varphi)$.

Soulignons que les noms des règles et des axiomes ne sont pas choisis au hasard. La règle (RE) permet la substitution d'équivalences logiques à l'intérieur de $\Box$. La règle (RM) caractérise les systèmes monotones, (RR) les systèmes réguliers, et (RK) permet d'obtenir le système normal K. Quant à (RN), il s'agit de la règle qu'on nomme *necessitation* en anglais.

Soit Δ un système contenant minimalement LPC et (df.$\Diamond$) :

1. Δ est *classique* s'il est fermé sous (RE) ;
2. Δ est *monotone* s'il est fermé sous (RM) ;
3. Δ est *régulier* s'il est fermé sous (RR) ;
4. Δ est *normal* s'il est fermé sous (RK).

Considérons maintenant les systèmes E, M, R et K, étant respectivement les plus petits systèmes classique, monotone, régulier et normal :

$$\begin{aligned} E &= LPC + (\mathsf{RE}) \\ M &= LPC + (\mathsf{RM}) \\ R &= LPC + (\mathsf{RR}) \\ K &= LPC + (\mathsf{RK}) \end{aligned}$$

Conformément à la présentation de Chellas (1980), nous avons les égalités suivantes (i.e., les systèmes sont syntaxiquement équivalents) :

$$\begin{aligned} K &= LPC + (\mathsf{K}) + (\mathsf{RN}) \\ &= LPC + (\mathsf{RK}) \\ &= LPC + (\mathsf{N}) + (\mathsf{RR}) \\ &= LPC + (\mathsf{N}) + (\mathsf{C}) + (\mathsf{RM}) \\ &= LPC + (\mathsf{N}) + (\mathsf{C}) + (\mathsf{M}) + (\mathsf{RE}) \end{aligned}$$

$$\begin{aligned} R &= LPC + (\mathsf{RR}) \\ &= LPC + (\mathsf{C}) + (\mathsf{RM}) \\ &= LPC + (\mathsf{C}) + (\mathsf{M}) + (\mathsf{RE}) \end{aligned}$$

$$\begin{aligned} &= LPC + (\mathsf{K}) + (\mathsf{RM}) \\ &= LPC + (\mathsf{K}) + (\mathsf{M}) + (\mathsf{RE}) \\ &= LPC + (\mathsf{R}) + (\mathsf{RE}) \end{aligned}$$

$$\begin{aligned} M &= LPC + (\mathsf{RM}) \\ &= LPC + (\mathsf{M}) + (\mathsf{RE}) \end{aligned}$$

En vertu de ces équivalences, M, R et K peuvent être formulés de différentes manières :

$$\begin{aligned} K &= R + (\mathsf{N}) \\ &= M + (\mathsf{C}) + (\mathsf{N}) \\ &= E + (\mathsf{M}) + (\mathsf{C}) + (\mathsf{N}) \\ &= M + (\mathsf{K}) + (\mathsf{N}) \\ &= E + (\mathsf{M}) + (\mathsf{K}) + (\mathsf{N}) \\ R &= M + (\mathsf{C}) \\ &= E + (\mathsf{M}) + (\mathsf{C}) \\ &= M + (\mathsf{K}) \\ &= E + (\mathsf{M}) + (\mathsf{K}) \\ &= E + (\mathsf{R}) \\ M &= E + (\mathsf{M}) \end{aligned}$$

Ces diverses formulations permettent de hiérarchiser les différentes logiques modales. Nous avons K extension de R, extension de M et de E :

$$E \subset M \subset R \subset K$$

En fonction des propriétés souhaitées pour l'opérateur $\Box$, ces différents systèmes pourront être utilisés en conjonction avec les axiomes désirés. Du point de vue sémantique, ces systèmes peuvent être interprétés dans des modèles *minimaux*, dont nous discuterons au chapitre 5.

* * *

En conclusion, soulignons l'intérêt des systèmes non réguliers de logique modale pour la logique déontique. Dans un système monotone mais non régulier, le principe d'agrégation, représenté par (C), n'est pas dérivable. Or certains auteurs rendent ce principe responsable du fait que les logiques déontiques normales ne sont pas en mesure de rendre compte des conflits d'obligations.[7]

[7] Voir notamment Horty (1997) et Goble (2004, 2009).

Chapitre 3

Les paradoxes

Historiquement, la majorité des paradoxes sont apparus après la publication de l'article initial de von Wright (1951). Cela dit, les paradoxes de la logique déontique se trouvent au cœur de la littérature scientifique et on a mis au point différentes approches afin de répondre aux problèmes que soulevait l'article de 1951 (et par la suite le système standard). De fait, avant d'aborder les différents courants, nous allons présenter les paradoxes de la logique déontique afin d'avoir une vue d'ensemble des motivations sous-jacentes aux différentes approches.

Dans les faits, chacun de ces paradoxes aurait pu servir de fil conducteur à la rédaction de cet ouvrage considérant qu'il aurait été possible de présenter l'émergence de plusieurs courants en réaction à ceux-ci. En raison de la diversité des paradoxes et des nombreux débats qu'ils ont suscités, nous avons opté pour une présentation synthétique des principaux problèmes de la logique déontique. Par conséquent, nous exposerons les paradoxes et tenterons de montrer la place qu'ils prennent dans la littérature scientifique, sans pour autant faire état de la totalité des discussions qu'ils ont suscitées. Le lecteur intéressé à obtenir plus de détails sur la diversité des paradoxes est invité à consulter McNamara (2010).

Qu'est-ce qu'un paradoxe ?

Selon Åqvist (2002), les paradoxes de la logique déontique résultent d'une inadéquation entre les systèmes formels et la langue naturelle (normative). Un système de logique peut être considéré comme un ensemble de formules fermé sous certaines règles. Soit Δ un système de logique déontique et NDL (pour *natural deontic logic*) un ensemble de propositions formant une représentation intuitive de la langue naturelle. Un paradoxe émerge lorsque Δ n'est pas une représentation adéquate de NDL. Soit $t : NDL \longrightarrow \Delta$ une fonction de traduction de la langue naturelle vers une logique déontique, et $t^{-1} : \Delta \longrightarrow NDL$ l'inverse de t (qui associe une traduction possible dans

NDL à chaque proposition de Δ).[1] La fonction t permet de traduire les énoncés de la langue naturelle dans le langage du système formel Δ. Par exemple, pour « si Pierre vole un vélo, alors Pierre commet un crime », un énoncé de la langue naturelle :

$$t(\text{si Pierre vole un vélo, alors Pierre commet un crime}) = p \supset q$$
$$t^{-1}(p \supset q) = \text{« si Pierre vole un vélo, alors Pierre commet un crime »}$$

Deux situations peuvent donner lieu à un paradoxe :

1. il y a une proposition $t(\varphi) \in \Delta$ telle que $\varphi \notin NDL$ parce qu'intuitivement, φ ne devrait pas être valide ;
2. il y a une proposition $t(\varphi) \notin \Delta$ pour laquelle $\varphi \in NDL$ puisque φ est intuitivement valide.

Åqvist nomme ces deux formes d'inadéquation respectivement une inadéquation de la droite vers la gauche et une inadéquation de la gauche vers la droite. Le sens de l'inadéquation va de la non appartenance $\notin$ vers l'appartenance $\in$. Dans la présentation qui suit, Δ sera utilisé pour de dénoté un système fortement normal (i.e., pour lequel $KD \subseteq \Delta$).

L'obligation dérivée

Le paradoxe de l'obligation dérivée apparaît à la suite de l'article de von Wright (1951). Avancé par Prior (1954), il met en évidence le caractère problématique de la traduction formelle que fait von Wright de la notion d'*engagement* (*commitment*). Ce dernier propose l'analyse suivante : si une action α est obligatoire et que de la poser nous engage à poser l'action β, alors β est aussi obligatoire. Formellement, l'engagement se traduit par $O(\alpha \supset \beta)$ et le raisonnement précédent se représente par :

$$(O\alpha \wedge O(\alpha \supset \beta)) \supset O\beta \tag{E}$$

Dans les mots de l'auteur, si l'implication entre deux actions α et β est obligatoire, alors poser l'action α engage à faire l'action β.[2] Par exemple, puisqu'il est obligatoire que *promettre implique tenir sa promesse*, il s'ensuit que promettre nous engage à tenir notre promesse. Par la suite, von Wright (1963) en viendra à traiter de la notion d'engagement en termes d'obligation dérivée, à savoir que nous sommes engagés à accomplir une obligation

[1] Formellement, t et son inverse associent des classes d'équivalence.

[2] On voit ici un parallèle avec la conséquence déontique telle qu'elle est conçue par Hintikka (1970).

dérivée β dans la mesure où cette action est nécessaire à l'accomplissement d'une action α, qui, elle, est obligatoire.

Cela dit, un théorème du système de von Wright stipule que si une action α est interdite, alors la conjonction de cette action avec n'importe quelle autre action β sera aussi interdite (avec $F\alpha =_{df} \neg P\alpha$) :

$$\neg P\alpha \supset \neg P(\alpha \wedge \neg\beta) \tag{I}$$

Par exemple, s'il est interdit de voler, alors il est aussi interdit de voler tout en conduisant. Toutefois, par la définition de la conjonction, la proposition (I) équivaut à la proposition (I*) :

$$\neg P\alpha \supset \neg P\neg(\alpha \supset \beta) \tag{I*}$$

Étant donné la définition de la permission, la proposition (I*) équivaut à la proposition (I**) :

$$\neg P\alpha \supset O(\alpha \supset \beta) \tag{I**}$$

Cependant, en vertu de la signification de $O(\alpha \supset \beta)$, la proposition (I**) signifie qu'enfreindre une action interdite nous engage à poser n'importe quelle autre action.

Exemple

S'il est interdit de voler et que Paul commet l'acte de voler, alors Paul est engagé à commettre l'adultère.

Alors que la proposition (I) est plausible d'un point de vue déontique, l'analyse de von Wright de la notion d'engagement entraîne le paradoxe de l'obligation dérivée, ce qui indique le caractère problématique de la lecture des propositions déontiques en termes d'*engagement* et d'*action*.

Dans la langue naturelle, cet énoncé est paradoxal pour les systèmes standard considérant que son interprétation est contre-intuitive : s'il n'est pas permis à Paul de voler, alors il est obligatoire que si Paul vole, alors Paul commet l'adultère. D'un point de vue normatif, ce n'est pas parce qu'il n'est pas permis de voler qu'il faut en conclure que voler implique obligatoirement de commettre l'adultère ! Le paradoxe de Prior montre une inadéquation de la droite vers la gauche :

$$\neg P\alpha \supset O(\alpha \supset \beta) \in \Delta$$
$$t^{-1}(\neg P\alpha \supset O(\alpha \supset \beta)) \notin NDL$$

Peu après la parution de l'article de Prior, McLaughlin (1955) a proposé une autre lecture de la traduction de l'engagement de von Wright pour remettre en cause la validité de la proposition (E). Son argument repose principalement sur la distinction entre la valeur de vérité d'une proposition déontique et sa valeur de performance, c'est-à-dire le fait qu'une action soit posée ou non.

Plus précisément, l'auteur soutient que la valeur de vérité d'une proposition déontique est relative à la valeur de performance des atomes propositionnels. Par exemple, afin de remettre en cause la proposition (E), McLaughlin soulève la question suivante : considérant que marcher dans un endroit public nous engage à porter des vêtements et présumant qu'il est obligatoire de marcher dans un endroit public, est-ce qu'il s'ensuit qu'on doive porter des vêtements même si on n'est pas dans un endroit public ?

Autrement dit, étant donné que « marcher dans un endroit public » nous engage à « porter des vêtements », et donc $O(\alpha \supset \beta)$, et qu'il est obligatoire de marcher dans un endroit public, soit $O\alpha$, est-ce qu'il s'ensuit qu'il est obligatoire de porter des vêtements $(O\beta)$? Le point que soulève McLaughlin est que la valeur de vérité de $O\beta$ dépend de la valeur de performance de $O\alpha$, à savoir que l'obligation de porter des vêtements est relative au fait de marcher dans un endroit public. De fait, il se pourrait que $O\alpha$ et $O(\alpha \supset \beta)$ soient vrais, sans pour autant que $O\beta$ le soit aussi.

Dans sa réponse au paradoxe de l'obligation dérivée, von Wright (1956) n'accordera pas vraiment d'importance à l'objection de McLaughlin. Dans la proposition (E), l'obligation dérivée β est conditionnelle à l'obligation α et à l'engagement entre α et β, et donc, il est incorrect de conclure que β est obligatoire, peu importe le contexte. Quoi qu'il en soit, von Wright a su reconnaître la valeur de l'objection de Prior. Il en conclut que la proposition (E) n'est pas une traduction formelle adéquate de la notion d'engagement.

Afin de pallier le paradoxe de l'obligation dérivée, von Wright introduira la logique déontique dyadique, qui stipule les conditions dans lesquelles les actions sont permises ou non et à l'intérieur de laquelle le paradoxe de Prior ne peut être dérivé.[3] Le paradoxe de l'obligation dérivé a historiquement donné lieu à la littérature scientifique sur les obligations conditionnelles mais c'est celui de Chisholm (1963) qui a exposé les limites des systèmes monadiques visant l'analyse des inférences normatives conditionnelles.

On trouve dans Åqvist (2002) une analyse détaillée du paradoxe de l'obligation dérivée. Dans cet ouvrage, l'auteur propose quatre façons de l'entendre du point de vue des systèmes *normaux* de logique déontique, qui sont minimalement des extensions conservatrices du système modal K.

[3] À condition que l'opérateur dyadique soit primitif.

Åqvist propose quatre formulations différentes du paradoxe de l'obligation dérivée[4] :

$$\neg\varphi \supset (\varphi \supset O\psi) \tag{3.1}$$
$$O\psi \supset (\varphi \supset O\psi) \tag{3.2}$$
$$O\neg\varphi \supset O(\varphi \supset \psi) \tag{3.3}$$
$$O\psi \supset O(\varphi \supset \psi) \tag{3.4}$$

Ces quatre propositions, démontrables dans les systèmes monadiques normaux de logique déontique, mettent en jeu la notion d'obligation dérivée. Ces énoncés montrent une inadéquation de la droite vers la gauche.

D'abord, on peut interpréter la proposition 3.1 de la manière suivante : si Paul ne respecte pas les limites de vitesse, alors s'il les respecte, il est dans l'obligation de voler une banque. D'un point de vue normatif, cette proposition est fausse puisque ce n'est pas parce qu'un agent n'accomplit pas une action α que le fait de poser cette action l'engage à poser n'importe quelle autre action.

Par ailleurs, la proposition 3.2, qui résulte de l'application de la même règle que pour la précédente, peut être lue de la manière suivante : Paul doit épouser Marie implique que si Paul tue Marie, alors il doit épouser Marie. Toutefois, en vertu du principe *devoir* implique *pouvoir*, à savoir que nous ne pouvons pas être obligé d'accomplir des actions impossibles (*impossibilium nulla obligatio est*), la proposition 3.2 est fausse dans la langue naturelle. En effet, si Paul tue Marie, alors il ne pourra pas l'épouser, et de fait, il ne sera pas dans l'obligation de le faire.

La proposition 3.3 correspond exactement au paradoxe de Prior. La proposition 3.4 est à rejeter pour la même raison que la 3.2.[5]

Soulignons que selon la lecture de la proposition $O\varphi$, le paradoxe de l'obligation dérivée ne mène pas inévitablement à une absurdité dans la langue naturelle. D'abord, il y a une équivalence logique (dans un système standard ou un système fermé sous une règle de substitution pour les équivalences logiques) entre les propositions $O(\varphi \supset \neg\psi)$ et $O\neg(\varphi \wedge \psi)$. Considérons maintenant l'interprétation suivante de $O\varphi$: l'action dénotée par φ est obligatoire. Selon cette lecture, le paradoxe de Prior disparaît : s'il est interdit de faire l'action dénotée par φ, alors il est aussi interdit de faire l'action dénotée par $\varphi \wedge \psi$ pour n'importe quel ψ. Si une action est interdite, alors toute conjonction d'action incluant cette action interdite sera

[4] Par convention, nous utilisons les lettres minuscules du début de l'alphabet grec pour les propositions qui font référence aux actions et les méta-variables φ, ψ, ρ, etc., pour les propositions déclaratives.

[5] Voir Åqvist (2002) pour la résolution du paradoxe dans les systèmes dyadiques.

aussi interdite. Cela est plausible dans la langue naturelle : s'il est interdit de voler, alors il est interdit de voler en conduisant.

Le bon Samaritain

Prior (1958) a formulé le paradoxe du bon Samaritain, qui visait à montrer le caractère inapproprié de la relation de conséquence au sein des inférences normatives. Dans la langue naturelle, le paradoxe prend la forme suivante :

1.	Si le bon Samaritain panse les plaies de Pierre, alors Pierre est blessé.
2.	Le bon Samaritain a l'obligation de panser les plaies de Pierre.
∴	Pierre à l'obligation d'être blessé.

Dans un système normal, le paradoxe se traduit par :

$$\vdash p \supset q \tag{3.5}$$
$$Op \tag{3.6}$$
$$Oq \tag{3.7}$$

Le symbole $\vdash$ est utilisé pour indiquer que $p \supset q$ est une conséquence syntaxique dérivable à l'intérieur du système. Ici, cela signifie que $p \supset q$ est postulé comme axiome. Dans un système normal, le paradoxe découle de l'usage de la règle dérivée (ROM), analogue de (RM) pour l'obligation :

$$\frac{\vdash \varphi \supset \psi}{\vdash O\varphi \supset O\psi}\ (\mathsf{ROM})$$

Formulé ainsi, cependant, le paradoxe du bon Samaritain est insidieux. En effet, $\vdash p \supset q$ n'est pas une traduction disponible dans le langage d'un système normal. En ce sens, le paradoxe ne s'adresse pas à un système normal mais s'adresse plutôt à une théorie normative construite sur la base d'un système normal avec $p \supset q$ comme axiome.[6]

À strictement parler, cette formulation ne met pas en jeu d'inadéquation puisque le raisonnement $(Op \wedge (p \supset q)) \supset Oq$ n'est pas dérivable dans un système normal.

[6] Voir Peterson et Marquis (2012).

Cette formulation est similaire à celle du paradoxe du voleur tel qu'il est avancé par Nozick et Routley (1962) :

1. Si Pierre aide la victime d'un vol, alors la victime a été volée.
2. Il est interdit de voler une victime.

∴ Il est interdit que Pierre aide la victime d'un vol.

Une réponse possible à ces paradoxes est qu'il ne s'agit que d'une autre forme de paradoxe Chisholm, qui met en jeu des obligations qui surviennent dans des contextes de violation. Considérant que l'obligation qu'a le bon Samaritain de panser les plaies de Pierre dépend du fait que Pierre soit blessé, il s'ensuit que le raisonnement repose sur la prémisse « si Pierre est blessé, alors le bon Samaritain doit panser les plaies de Pierre ».

Or, dans son analyse des paradoxes du style *bon Samaritain*, Nowell-Smith (1960) souligne qu'il faut prendre garde à ne pas commettre le *sophisme naturaliste*, qui consiste à dériver une proposition normative à partir d'un ensemble de prémisses purement descriptives. En supposant cette prémisse supplémentaire, il faut prendre garde de ne pas considérer que la proposition normative *le bon Samaritain doit panser les plaies de Pierre* peut être dérivée à partir de la proposition descriptive *Pierre est blessé*. Plutôt, il faut la considérer comme une proposition normative qui permet de conclure *le bon Samaritain doit panser les plaies de Pierre* à partir de *Pierre est blessé*. Autrement dit, l'implication doit être vue comme une proposition normative qui prend une proposition descriptive comme antécédent et une proposition normative comme conséquent.

Par la suite, Åqvist (1967) proposera une version beaucoup plus forte du paradoxe du bon Samaritain, laquelle s'adresse effectivement aux systèmes normaux (contrairement à la formulation initiale de Prior). Åqvist proposera le paradoxe suivant :

1. Pierre ne doit pas voler Paul.
2. Le bon Samaritain doit aider Paul, lequel a été volé par Pierre.

∴ Il est obligatoire que Pierre vole Paul.

Il s'agit d'une inadéquation de la droite vers la gauche pour les systèmes normaux Δ :

$$(O\neg p \wedge O(q \wedge p)) \supset Op \in \Delta$$
$$t^{-1}((O\neg p \wedge O(q \wedge p)) \supset Op) \notin NDL.$$

Dans un système fortement normal, c'est-à-dire une extension conservatrice du système modal KD, le paradoxe prend une autre forme et consiste à dire que les prémisses de l'argument sont inconsistantes.

Démonstration.

1	$O\neg p$	H
2	$O(q \wedge p)$	H
3	$\quad O$	H
4	$\quad q \wedge p$	EO 2,3
5	$\quad p$	E$\wedge$ 4
6	Op	IO 3-5
7	$Op \supset \neg O\neg p$	(**D**)
8	$\neg O\neg p$	E$\supset$ 6,7
9	$\bot$	I$\bot$ 1,8

□

Dans le cas d'un système normal, les lignes 1 à 6 de la dérivation montrent comment obtenir le paradoxe, où les prémisses permettent de conclure que Paul a l'obligation d'être blessé. Par ailleurs, l'utilisation du schéma d'axiome (**D**) montre comment l'inconsistance survient dans un système fortement normal. Dans le cas d'un système fortement normal, il s'agit là d'une seconde inadéquation de la droite vers la gauche :

$$(O\neg p \wedge O(q \wedge p)) \supset \bot \in \Delta$$
$$t^{-1}((O\neg p \wedge O(q \wedge p)) \supset \bot) \notin NDL.$$

Afin de résoudre le paradoxe, Åqvist proposera de traduire la seconde prémisse par la proposition suivante, ce qui permet d'éviter la contradiction :

$$Oq \wedge p$$

Cela dit, cette version du paradoxe du bon Samaritain est comparable au paradoxe de Chisholm puisque l'obligation d'aider la victime d'un vol survient dans un contexte de violation et dépend du fait que Pierre enfreint l'interdiction de voler Paul. La solution que propose Åqvist est de modifier (considérablement) le système standard et d'introduire une distinction entre les obligations primaires (O_1) et les obligations contraires (O_2).

Les obligations violées

L'objection de Chisholm (1963) s'adresse principalement aux systèmes monadiques de logique déontique et tend à montrer que ceux-ci ne sont pas en mesure de rendre compte des obligations qui sont conditionnelles à des contextes où d'autres obligations sont violées.[7] En ce sens, le paradoxe de Chisholm met en jeu des obligations qui surviennent dans des situations sous-optimales (Jones et Pörn 1985). Il s'agit du paradoxe qui a fait couler le plus d'encre et qui est le plus important dans la littérature scientifique.

Le paradoxe repose sur la notion d'obligations qui surviennent dans des contextes de violations, ce que l'auteur nomme des *contrary-to-duty imperatives*. De telles obligations découlent d'un impératif qui dicte ce que nous devons faire dans le cas où nous manquons à certains de nos devoirs. Par exemple, Paul est dans l'obligation de tenir sa promesse à Pierre, mais si Paul ne peut pas la respecter, alors il est dans l'obligation de le dire à Pierre. L'obligation pour Paul de dire à Pierre qu'il n'a pas respecté sa promesse survient dans un contexte où l'obligation initiale de respecter la promesse est violée.

Les obligations qui surviennent dans des contextes de violation sont conditionnelles. Or l'obligation conditionnelle peut être traduite de deux manières en logique déontique monadique, à savoir $O(\varphi \supset \psi)$ ou $\varphi \supset O\psi$. Cependant, en vertu du paradoxe de l'obligation dérivée, la première formulation n'offre pas une traduction adéquate pour les obligations conditionnelles aux contextes de violation. En traduisant les obligations qui surviennent dans des contextes de violation par $O(\varphi \supset \psi)$, on peut conclure $O(\varphi \supset \psi)$ pour n'importe quel ψ dès que φ dénote quelque chose d'interdit. En ce sens, lorsqu'il y a une interdiction, n'importe quoi peut être vu comme une obligation qui survient dans un contexte de violation, ce qui est évidemment indésirable.

Chisholm opte plutôt pour la seconde formulation, où une action particulière implique une obligation particulière. Selon lui, la formule $\varphi \supset O\psi$ est celle qui rend le mieux compte de l'obligation contraire. Selon l'auteur, il est donc nécessaire d'avoir une logique qui admet des formules mixtes (descriptives et normatives) au sein de son ensemble d'énoncés bien formés.[8] Cela dit, Chisholm montre que les logiques monadiques qui admettent des propositions de cette forme ainsi que les (méta)théorèmes exprimant respectivement le détachement déontique (E) et la consistance normative (**D**) mènent à des absurdités.

[7] Il est néanmoins possible de modéliser les obligations conditionnelles et les contextes de violation à l'aide d'un opérateur monadique (Peterson 2015b).

[8] Sur ce sujet, voir Peterson (2014a).

$$(O\varphi \wedge O(\varphi \supset \psi)) \supset O\psi \tag{E}$$
$$\neg(O\varphi \wedge O\neg\varphi) \tag{**D**}$$

L'objection avancée par Chisholm montre que les systèmes monadiques fortement normaux ne peuvent rendre compte des obligations qui émergent dans des contextes où d'autres obligations ont été enfreintes. Le paradoxe se résume à ce qu'on nomme souvent le *Chisholm's set*, soit l'ensemble des propositions suivantes :

1. Paul ne doit pas voler ;
2. cela devrait être le cas que si Paul ne vole pas, alors il ne remet pas l'argent volé ;
3. si Paul vole, alors il doit remettre l'argent volé ;
4. Paul vole.

Dans un tel système, ces énoncés se traduisent par :

$$O\neg p \tag{3.8}$$
$$O(\neg p \supset \neg q) \tag{3.9}$$
$$p \supset Oq \tag{3.10}$$
$$p \tag{3.11}$$

La nécessité d'utiliser les traductions différentes 3.9 et 3.10 pour la seconde et la troisième prémisse apparaît lorsqu'on considère que les énoncés sont indépendants dans la langue naturelle.[9]

Considérons les deux autres traductions possibles :

$$O\neg p \tag{3.12}$$
$$O(\neg p \supset \neg q) \tag{3.13}$$
$$O(p \supset q) \tag{3.14}$$
$$p \tag{3.15}$$

$$O\neg p \tag{3.16}$$
$$\neg p \supset O\neg q \tag{3.17}$$
$$p \supset Oq \tag{3.18}$$
$$p \tag{3.19}$$

[9] Voir aussi Åqvist (1967), Decew (1981), Tomberlin (1981, 1983) et Hansen (1999).

Ces traductions sont inadéquates puisqu'elles ne préservent pas l'indépendance entre les énoncés dans la langue naturelle. En effet, 3.14 est la conséquence de 3.12 et 3.17 est la conséquence de 3.19.

Le paradoxe de Chisholm montre une inadéquation de la droite vers la gauche. En effet, celui-ci met en évidence que les quatre propositions susmentionnées, intuitivement cohérentes dans la langue naturelle, sont inconsistantes lorsqu'elles sont traduites dans le langage d'une logique déontique fortement normale. Soit Φ la conjonction des énoncés traduits dans le langage de Δ[10] :

$$\Phi = O\neg p \wedge (O(\neg p \supset \neg q) \wedge ((p \supset Oq) \wedge p))$$

$$\Phi \supset \bot \in \Delta$$
$$t^{-1}(\Phi \supset \bot) \notin NDL.$$

Le paradoxe de Chisholm donnera lieu à la littérature scientifique sur les systèmes dyadiques de logique déontique (Hansson 1969), mais celui-ci sera aussi une motivation pour les systèmes non monotones et adaptatifs. Chisholm a cerné un problème qui met en évidence le fait que certains systèmes de logique déontique ne sont pas en mesure de rendre compte du caractère conditionnel des obligations contraires (Decew 1981). Certains, notamment Hansen (1999), concluent que l'implication matérielle n'est pas adéquate pour la représentation des obligations conditionnelles. Cela suggère la nécessité d'utiliser un opérateur primitif pour en rendre compte (Chellas 1974, 1980).[11]

Certains ont tenté de rendre compte de la notion d'obligation contraire à l'intérieur de systèmes monadiques. Après l'article de Chisholm, Mott (1973) a proposé d'augmenter le système standard de logique déontique en lui ajoutant l'opérateur introduit par Lewis représentant la conditionnelle contrefactuelle. En réaction à la position de Mott, Decew (1981) soutient que l'introduction de la contrefactuelle ne permet pas de résoudre le paradoxe de Chisholm puisque la conditionalité des obligations contraires met en jeu un paramètre temporel, dont le système proposé par Mott ne parvient pas à rendre compte. En définitive, à la suite de l'objection de Decew, Niles (1997) a tenté de préserver l'approche contrefactuelle en introduisant une quantification sur les actions. Cela dit, notons que Prakken et Sergot (1996) ont proposé une formulation du paradoxe de Chisholm qui ne met

[10] Voir Sellars (1967) pour la formalisation du paradoxe de Chisholm.

[11] Peterson (2015b) a récemment montré que cette intuition est erronée et qu'il est possible de modéliser les raisonnements normatifs conditionnels sans opérateur primitif pour l'obligation conditionnelle.

pas en jeu de dimension temporelle, et donc la temporélisation des logiques déontiques n'y offre pas une solution définitive (Straßer 2011).

En insistant sur l'aspect conditionnel des obligations contraires, le paradoxe de Chisholm aura principalement eu des répercussions sur les approches conditionnelles, temporelles, contextuelles et non monotones.

Pour conclure, soulignons un caractère problématique de la modélisation des obligations conditionnelles au sein d'un système monadique. Dans la formulation du paradoxe, Chiholms opte pour $\varphi \supset O\psi$ pour modéliser les obligations qui surviennent dans un contexte de violation de $O\neg\varphi$. Or l'énoncé $\varphi \supset O\psi$ est une proposition mixte, qui met en jeu un antécédent descriptif et un conséquent normatif. Cela dit, comme nous le verrons lors de l'analyse du problème du détachement factuel, la vérité de l'antécédent descriptif ne suffit pas à conclure à la vérité du conséquent normatif. En ce sens, il semble que l'implication matérielle $\supset$ ne soit pas adéquate pour modéliser la transmission des valeurs de vérité entre un énoncé descriptif et un énoncé normatif, et *a fortiori* pour représenter les obligations conditionnelles.

Le meurtre gentil

À la suite du paradoxe du bon Samaritain, Forrester (1984) a avancé le paradoxe du meurtre gentil afin de mettre en évidence le caractère inadéquat de la relation de conséquence implicite aux systèmes normaux. Cette relation de conséquence normative peut être représentée par la règle dérivée (ROM) :

$$\frac{\vdash_\Delta \varphi \supset \psi}{\vdash_\Delta O\varphi \supset O\psi} \qquad \text{(ROM)}$$

Alors que l'ensemble des théorèmes d'un système monadique normal est fermé sous la règle (ROM), Forrester reprend le raisonnement sous-jacent au paradoxe du bon Samaritain de façon à montrer que cette règle ne représente pas le comportement de l'obligation légale. Supposons un système légal qui adopte deux normes, représentées par les prémisses 1 et 2 :

1. Il est interdit que Pierre tue Jean.
2. Si Pierre tue Jean, alors il est obligatoire qu'il le fasse gentiment.
3. Pierre tue Jean.
4. Si Pierre tue Jean gentiment, alors Pierre tue Jean.

$\therefore$ Il est obligatoire que Pierre tue Jean.

La seconde prémisse équivaut à supposer que si Pierre tue Jean, il doit le faire de façon à causer le moins de souffrances possible. Le paradoxe de Forrester fonctionne en deux temps. D'abord, il consiste à dire que dans un système normal, il y a inadéquation de la droite vers la gauche puisque les prémisses permettent de conclure que Pierre a l'obligation de tuer Jean, ce qui ne semble pas être dérivable dans la langue naturelle.

Dans un système normal, le paradoxe se traduit par :

$$O\neg p \tag{3.20}$$
$$p \supset Oq \tag{3.21}$$
$$p \tag{3.22}$$
$$q \supset p \tag{3.23}$$
$$Oq \tag{3.24}$$

Soit Φ la conjonction des prémisses 2, 3 et 4 :

$$\Phi = p \wedge ((q \supset p) \wedge (p \supset Oq))$$

Le paradoxe vise à montrer une inadéquation de la droite vers la gauche pour un système normal Δ' :

$$\Phi \supset Op \in \Delta'$$
$$t^{-1}(\Phi \supset Op) \notin NDL.$$

Soulignons que dans sa formulation, le paradoxe de Forrester commet la même faute que la formulation initiale du paradoxe du bon Samaritain de Prior. À strictement parler, $\Phi \supset Oq$ n'est pas une conséquence pour un système normal Δ. Il s'agit d'une conséquence pour une théorie normative construite sur la base de Δ à l'aide de l'axiome $q \supset p$. À ce sujet, le lecteur est invité à consulter Peterson et Marquis (2012) pour une analyse détaillée.

En supposant une extension $\Delta' = \Delta \cup \{q \supset p\}$, Op est dérivable dans la mesure où 3.21 et 3.22 entraînent Oq, et puisque par (ROM) il est possible de dériver $Oq \supset Op$ en vertu de l'axiome $q \supset p$, Oq entraîne Op.

Le second niveau auquel le paradoxe de Forrester fonctionne concerne les systèmes fortement normaux. Considérons la conjonction de toutes les prémisses Ψ :

$$\Psi = O\neg p \wedge \Phi$$

Soit une extension $\Delta' = \Delta \cup \{q \supset p\}$ d'un système fortement normal Δ. Le paradoxe montre une inadéquation de la droite vers la gauche dans la mesure où les prémisses, qui semblent cohérentes dans la langue naturelle, sont inconsistantes dans Δ' :

$$\Psi \supset \bot \in \Delta'$$
$$t^{-1}(\Psi \supset \bot) \notin NDL.$$

En réponse à l'article de Forrester, Sinnott-Armstrong (1985) a proposé une analyse différente de la proposition 3.21 afin d'éviter le paradoxe. Selon lui, l'introduction de quantificateurs fait s'effondrer le paradoxe, puisque cela permet de traduire 3.21 par la proposition suivante :

$$(\exists x)(Mxpj) \supset (\exists x)(Mxpj \wedge OGx) \tag{3.25}$$

Autrement dit, s'il existe une action x telle que x est le meurtre de Jean par Pierre, alors il existe une action x telle que x est le meurtre de Jean par Pierre et il est obligatoire que x soit faite gentiment. Ainsi, dans l'éventualité où le meurtre était commis, l'obligation porterait sur le fait que l'action soit commise gentiment.

Remarquons au passage que cette traduction n'est qu'une ébauche et que ce type de formalisation pose quelques problèmes. En effet, Sinnott-Armstrong semble vouloir traiter l'action à la fois comme une constante et comme un prédicat. La quantification sur la variable x se fait sur un domaine $A = \{a_1, ..., a_n, ...\}$ de constantes d'actions. Or le prédicat $Mxpj$ signifie que l'action x est le meurtre de Jean par Pierre, ce qui engendre une confusion : l'action est-elle un prédicat ou une constante ? Sur quoi est-ce qu'on quantifie exactement ? Si l'action est aussi un prédicat, alors nous tombons dans une logique de deuxième ordre ! En ce sens, la solution de Sinnott-Armstrong, qui, à la base, semble plutôt simple et convenable, laisse place à plusieurs éléments sur lesquels il serait important de statuer.

Pour rester en premier ordre, il faudrait quantifier à la fois sur les actions et sur les individus dans un système typé qui admet l'identité. Ainsi, la traduction de 3.21 par la proposition suivante serait peut-être moins problématique :

$$(\forall x : P)(\forall y : A)((xFy \wedge y = m) \supset OGy) \tag{3.26}$$

Ici, $\forall x : P$ et $\forall y : A$ signifient respectivement « pour tout x de type P (pour *personne*) » et « pour tout y de type A (pour *action*) ». De fait, l'énoncé signifierait que pour toute personne x et pour toute action y, si x pose l'action y et que l'action y est celle de commettre un meurtre, alors il est obligatoire que le meurtre soit gentil. Néanmoins, cette formulation

pose problème dans la mesure où on met un prédicat G dans la portée d'un autre prédicat O, ce qui nous ferait demeurer en deuxième ordre.

La traduction des propositions normatives en termes de quantificateurs n'est pas si évidente et la proposition de Sinnott-Armstrong, même si elle est originale, est loin d'être complète. En réaction à Sinnott-Armstrong, Goble (1991) en est venu à la conclusion que peu importe la traduction, le paradoxe de Forrester persiste. Selon lui, la solution consiste plutôt à rejeter la règle (ROM).

Dans un autre ordre d'idées, Jacquette (1986) soutient que le paradoxe de Forrester ne dépend que du fait que l'ensemble de normes est mal construit, et donc que la contradiction ne provient pas des règles de la logique déontique, mais bien de l'inconsistance des normes de Forrester. Selon Jacquette, la proposition 3.21 indique l'obligation conditionnelle de commettre un meurtre d'une manière particulière.

Cependant, l'idée derrière la seconde prémisse est que, considérant qu'un meurtre gentil est préférable à un meurtre cruel, une personne ne doit pas commettre de meurtre, et *a fortiori* ne doit pas commettre de meurtre de façon cruelle (non gentille). En ce sens, la seconde prémisse est plutôt « si Pierre tue Jean, alors il est obligatoire qu'il ne le tue pas de manière cruelle ».

Cette prémisse est beaucoup plus plausible d'un point de vue normatif. En traduisant l'idée véhiculée par la seconde prémisse de la sorte, le paradoxe s'évapore puisqu'il n'est plus possible de conclure qu'il est obligatoire de commettre un meurtre : il est seulement possibe de conclure qu'il est obligatoire de ne pas commetre de meurtre cruel. Ce ne sont donc pas les lois de la logique déontique qui sont paradoxales, mais bien les normes de Forrester.

À la suite de la position de Jacquette, Robinson (1986) a préféré soutenir que ce n'est pas la traduction de la seconde prémisse qui cause le paradoxe de Forrester, mais bien son utilisation de la proposition 3.27[12] :

$$O(p \supset q) \supset (p \supset Oq) \tag{3.27}$$

Cette proposition est invalide dans les systèmes fortement normaux, et de fait Robinson reproche à Forrester de critiquer ces systèmes à l'aide d'une proposition que ceux-ci n'admettent pas. Néanmoins, soulignons que le paradoxe de Forrester peut être formulé sans l'aide de la proposition 3.27.

[12] Nous avons laissé de côté une étape non nécessaire lors de la présentation du paradoxe. Dans son article, Forrester stipule que la norme est en fait « il est obligatoire que si Pierre tue Jean, alors il le fait gentiment », et qu'en vertu du principe $O(p \supset q) \supset (p \supset Oq)$, qu'il pose en hypothèse, on obtient la proposition « si Pierre tue Jean, alors il doit le tuer gentiment ». Forrester fait ce détour afin d'éviter un débat quant à la portée de l'opérateur O au sein de la proposition.

Toutes choses considérées, le paradoxe de Forrester exemplifie la tendance en logique déontique à identifier les problèmes aux règles et aux schémas d'axiomes qui gouvernent le comportement des opérateurs déontiques. Cela dit, il est important de souligner que la plupart des problèmes sont en fait imputables aux propriétés de la logique propositionnelle sous-jacente aux systèmes classiques. Sur ce sujet, le lecteur est invité a consulter Peterson (2014b).

La conséquence déontique

L'analyse des paradoxes du bon Samaritain et du meurtre gentil met en évidence un paradoxe qui résulte plutôt d'une inadéquation de la gauche vers la droite. Peterson et Marquis (2012) ont proposé ce paradoxe, qui montre que les systèmes normaux et fortement normaux ne parviennent pas à modéliser adéquatement certaines inférences qui semblent intuitivement valides dans la langue naturelle. Considérons le raisonnement suivant :

1. Il est interdit que Pierre vole.
2. Si Pierre prend la bicyclette de Paul sans son consentement, alors il Pierre vole.

$\therefore$ Il est interdit que Pierre prenne la bicyclette de Paul sans son consentement.

Dans la langue naturelle, ce raisonnement est évidemment valide : il serait absurde de permettre (de ne pas interdire) une action qui viole une interdiction. Dans un système normal ou fortement normal, ces énoncés se traduisent par :

$$O\neg p \tag{3.28}$$
$$q \supset p \tag{3.29}$$
$$O\neg q \tag{3.30}$$

Or il s'avère que $O\neg q$ n'est pas une conséquence $O\neg p \wedge (q \supset p)$ ni pour un système normal, ni pour un système fortement normal. En ce sens, il y a une inadéquation de la gauche vers la droite :

$$(O\neg\varphi \wedge (\psi \supset \varphi)) \supset O\neg\psi \notin \Delta$$
$$t^{-1}((O\neg\varphi \wedge (\psi \supset \varphi)) \supset O\neg\psi) \in NDL.$$

L'explosion déontique

Outre le paradoxe de Chisholm, les logiques déontiques non monotones sont aussi motivées par ce qu'on nomme l'explosition déontique. Dans un système normal, il est possible de dériver que $(O\varphi \wedge O\neg\varphi) \supset O\psi$ pour n'importe quel ψ. En effet, il suffit de remarquer que la formule $(O\varphi \wedge O\neg\varphi) \supset O(\varphi \wedge \neg\varphi)$ est dérivable, et que les propositions logiquement équivalentes peuvent être substituées à l'intérieur de la portée de l'opérateur O. En ce sens, un conflit d'obligations entraîne $O\bot$, et puisque $\bot \supset \psi$ pour n'importe quel ψ, il s'ensuit que $O\bot \supset O\psi$.

Dans un système fortement normal, l'explosion déontique vient du fait que $(O\varphi \wedge O\neg\varphi) \supset \bot$ et que $\bot \supset \psi$ pour n'importe quel ψ.

L'explosion déontique consiste donc à dire que lorsqu'il y a un conflit d'obligations, alors n'importe quoi est obligatoire. Cela montre une inadéquation de la droite vers la gauche pour les systèmes normaux et fortement normaux :

$$(O\varphi \wedge O\neg\varphi) \supset O\psi \in \Delta$$
$$t^{-1}((O\varphi \wedge O\neg\varphi) \supset O\psi) \notin NDL.$$

De même, cela montre aussi que les systèmes normaux et fortement normaux sont incapables de modéliser les conflits d'obligations, qui semblent pourtant possibles dans la langue naturelle.

Le paradoxe de Ross

Dans sa réponse au dilemme avancé par Jørgensen, Alf Ross (1941), en cherchant à déterminer si les impératifs peuvent faire partie des inférences logiques, en est venu à formuler ce qu'on nomme aujourd'hui le paradoxe de Ross.[13] Ce paradoxe émerge lors de l'analyse de la validité et de la satisfaction des impératifs disjonctifs.

D'un point de vue propositionnel, si φ est satisfait (ou valide), alors $\varphi \vee \psi$ est aussi satisfait (ou valide). Or Ross prend cette analyse et la transpose sur le plan des impératifs afin de montrer que ceux-ci ne sont pas sujets aux lois de la logique propositionnelle. En effet, son point principal est que la règle « si l'impératif φ est satisfait (ou valide), alors l'impératif $\varphi \vee \psi$ est aussi satisfait (ou valide) » est inadéquate. Le contre-exemple qu'il propose est le suivant : si l'impératif *poste cette lettre !* est satisfait (ou valide), alors l'impératif *poste cette lettre ou brûle-la !* est aussi satisfait (ou valide).

[13] L'article initial a été republié dans Ross (1944).

En logique déontique, l'impératif « poste la lettre ! » s'interprète par l'énoncé (déclaratif) normatif « tu dois poster la lettre », et donc cet impératif se traduit formellement par Op. Le paradoxe de Ross en logique déontique repose sur le fait que les systèmes normaux admettent les théorèmes de la forme $O\varphi \supset O(\varphi \vee \psi)$. En ce sens, il y a donc une inadéquation de la droite vers la gauche :

$$O\varphi \supset O(\varphi \vee \psi) \in \Delta$$
$$t^{-1}(O\varphi \supset O(\varphi \vee \psi)) \notin NDL.$$

Le paradoxe de Ross prend toute sa signification lorsqu'on considère l'équivalence entre la disjonction et l'implication :

$$O(p \vee q) \equiv O(\neg p \supset q)$$

Dans la langue naturelle, cela signifie que si Paul doit poster la lettre, alors s'il ne la poste pas, il doit la brûler, ce qui est évidemment inadéquat.

Cela dit, bien que ceux qui s'initient à la logique déontique insistent sur ce paradoxe, le paradoxe de Ross est le moins important pour la logique déontique. En fait, cela n'est pas plus paradoxal que $\varphi \supset (\varphi \vee \psi)$ en logique propositionnelle. Dès qu'on comprend les conditions de vérité de la disjonction, le paradoxe s'évapore, d'autant plus que cela ne permet pas de conclure que dans les faits, si Paul ne poste pas la lettre, alors il doit la brûler. Soulignons que Oq n'est pas dérivable à partir de Op, $O(p \vee q)$ et $\neg p$ dans un système normal ou fortement normal.

Par ailleurs, le paradoxe de Ross peut aussi être résolu similairement au paradoxe de l'obligation dérivée. Considérons le cas d'une interdiction $O\neg p$. En vertu de l'équivalence entre $O(\neg p \vee \neg q)$ et $O\neg(p \wedge q)$, le paradoxe de Ross est une instance de $O\neg p \supset O\neg(p \wedge q)$, ce qui indique que si une action est interdite, alors la conjonction de n'importe quelle autre action avec cette action interdite sera aussi interdite. Du point de vue du discours normatif, cela est plausible.

* * *

Somme toute, le cadre d'analyse que propose Åqvist (2002) permet une représentation conceptuel claire des paradoxes de la logique déontique. Si la logique déontique vise la modélisation du discours normatif, alors celle-ci doit fournir une représentation adéquate de la forme des inférences dans la langue naturelle. Par conséquent, un système de logique déontique ne doit pas satisfaire des énoncés qui semblent intuitivement inadéquats.

Comme le mentionne Nowell-Smith (1960), les paradoxes de la logique déontique n'ont pas pour objectif de montrer l'inconsistance des systèmes syntaxiques, contrairement aux antinomies de Russell. Plutôt, ils visent à montrer que la syntaxe des systèmes ne représente pas adéquatement la sémantique du discours normatif. En ce sens, un paradoxe surgit lorsque soit un système valide un raisonnement qui ne semble pas dérivable dans la langue naturelle (i.e., une inadéquation de la doite vers la gauche), soit lorsqu'il ne valide pas une inférence qui semble adéquate dans la langue naturelle (i.e., une inadéquation de la gauche vers la droite).

Historiquement, les paradoxes ont une place extrêmement importante au sein de la littérature scientifique puisque les différentes approches se sont développées pour répondre à l'une ou à l'autre des objections contre les systèmes normaux et fortement normaux. Cela dit, malgré l'importance de chaque paradoxe, l'objection de Chisholm demeure celle qui a le plus marqué la littérature scientifique.

Chapitre 4

Les systèmes monadiques

Outre les avancées de Mally (1926), l'article fondateur de von Wright (1951) aura eu énormément de répercussions sur la logique déontique. Dans cet article, l'auteur cherchait à rendre compte des modalités déontiques comme l'obligation, la permission et l'interdiction en traçant un parallèle avec les modalités aléthiques de nécessité et de possibilité. Le système initial de von Wright est une logique monadique, c'est-à-dire une logique où il n'y a qu'une seule proposition dans la portée d'un opérateur. En ce sens, les logiques monadiques traitent de l'obligation en tant qu'opérateur unaire.

Le système initial de von Wright

Le système initial de von Wright est une logique du type *ought-to-do*, différente de la majorité des approches en logique modale et de la position que lui-même adoptera au cours de sa carrière. Les propositions à l'intérieur de ce système font référence à des catégories générales d'actions (des propriétés) plutôt qu'à des actes individuels. En ce sens, le système initial porte sur des propositions de la forme $O\alpha$, où α est une proposition (atomique ou complexe) qui dénote une action (générale) qui doit être faite par un individu (ou un groupe d'individus).

Par ailleurs, l'itération des opérateurs n'est pas permise dans le système initial de von Wright, c'est-à-dire qu'une expression qui contient des opérateurs déontiques ne peut pas être dans la portée d'un opérateur déontique. De fait, $O(O\alpha \supset P\beta)$ n'est pas un énoncé bien formé du système initial de von Wright.

D'emblée, l'auteur se sert de l'obligation pour définir l'interdiction et la permission, à savoir qu'un acte non permis est interdit et qu'une action est interdite seulement si la négation de cette action est obligatoire. Ici, F signifie *forbidden* alors que P signifie *permitted*. Ces définitions perdureront dans la littérature scientifique.

$$F\alpha =_{df} O\neg\alpha \qquad \text{(df. F)}$$
$$P\alpha =_{df} \neg O\neg\alpha \qquad \text{(df. P)}$$

Le système initial de von Wright se veut essentiellement une approche syntaxique, laquelle est principalement fondée sur des intuitions sémantiques. En effet, l'auteur se sert de principes intuitivement plausibles en vue de créer son système formel. Par exemple, von Wright prendra le principe de distribution de la permission, à savoir qu'une disjonction d'action est permise si et seulement si l'une des deux actions est permise, comme axiome pour son système[1] :

$$P(\alpha \vee \beta) \equiv (P\alpha \vee P\beta) \tag{4.1}$$

En vertu des dualités de De Morgan et de la définition de la permission, le principe de distribution est équivalent au principe d'agrégation complète pour l'obligation :

$$O(\alpha \wedge \beta) \equiv (O\alpha \wedge O\beta) \tag{4.2}$$

À ce principe il ajoutera celui de permission, qui stipule que pour une action α, soit α est permise soit sa négation l'est :

$$P\alpha \vee P\neg\alpha \tag{4.3}$$

Ce principe, conforme à l'intuition telle qu'un système normatif ne peut pas dicter des obligations contradictoires, est équivalent d'un point de vue propositionnel à l'axiome (**D**). Il représente la notion de consistance normative qui représente la cohérence normative présupposée dans la langue naturelle :

$$O\alpha \supset P\alpha \tag{**D**}$$

En effet, il suffit de remarquer que la définition de la permission engendre que l'axiome 4.3 est propositionnellement équivalent à $\neg(O\alpha \wedge O\neg\alpha)$. Autrement dit, $\neg(O\alpha \wedge O\neg\alpha)$ est dérivable à partir de $O\alpha \supset P\alpha$ et vice versa.

[1] Soulignons que l'intuition sémantique derrière le principe de distribution ne stipule pas les conditions dans lesquelles une proposition normative est vraie, mais propose seulement une condition de validité pour les raisonnements déontiques.

Cette équivalence dépend uniquement des règles propositionnelles :

1	$O\alpha \supset P\alpha$	H
2	$O\alpha \supset \neg O\neg\alpha$	df. P 1
3	$O\alpha \wedge O\neg\alpha$	H
4	$O\neg\alpha$	E$\wedge$ 3
5	$O\alpha$	E$\wedge$ 3
6	$\neg O\neg\alpha$	E$\supset$ 1,2
7	$\perp$	I$\perp$ 4,6
8	$(O\alpha \wedge O\neg\alpha) \supset \perp$	I$\supset$ 3-7
9	$\neg(O\alpha \wedge O\neg\alpha)$	df. $\neg$ 8

1	$\neg(O\alpha \wedge O\neg\alpha)$	H
2	$O\alpha$	H
3	$O\neg\alpha$	H
4	$O\alpha \wedge O\neg\alpha$	I$\wedge$ 2,3
5	$\perp$	I$\perp$ 1,4
6	$O\neg\alpha \supset \perp$	I$\supset$ 3-5
7	$\neg O\neg\alpha$	df. $\neg$ 6
8	$O\alpha \supset \neg O\neg\alpha$	I$\supset$ 2-7

Par ailleurs, von Wright adoptera le principe de contingence, à savoir qu'une tautologie n'est pas nécessairement obligatoire et qu'une contradiction n'est pas nécessairement interdite. Ce principe sera notamment repris par Chellas (1974) et Jones et Pörn (1985).

Essentiellement, le système de von Wright se résume à une extension de la logique classique, à laquelle on ajoute trois axiomes (von Wright 1967) :

$$P\alpha \equiv \neg O\neg\alpha \tag{A1}$$

$$P(\alpha \vee \beta) \equiv (P\alpha \vee P\beta) \tag{A2}$$

$$P\alpha \vee P\neg\alpha \tag{A3}$$

À ces axiomes s'ajoutent trois règles d'inférence, soit une règle de détachement (*modus ponens*), une règle de substitution pour les variables propositionnelles et une règle d'extensionnalité qui stipule que les propositions appartenant à la même classe d'équivalence sont interchangeables dans la portée des opérateurs.

Par la suite, l'auteur ajoutera un quatrième axiome au système (von Wright 1983), ce qui le rendra syntaxiquement équivalent au système standard. L'ajout de cet axiome équivaut au rejet du principe de contingence :

$$O(\alpha \vee \neg\alpha) \tag{A4}$$

La validité des raisonnements déontiques est évaluée en fonction du principe distribution :

$$a(P(\alpha \vee \beta) \equiv (P\alpha \vee P\beta)) = \top$$

Ayant cette hypothèse en main, il est possible de construire des tables de vérités pour les propositions déontiques. Si on suppose que (A2) est vrai, il s'ensuit que (A3) l'est aussi, et ainsi von Wright se sert uniquement du principe de distribution afin d'évaluer la validité des inférences normatives. Cela dit, le système initial de von Wright fait preuve de (très) peu de considérations sémantiques et celui-ci aura été fortement critiqué par les paradoxes. Néanmoins, ce système notable est à la base de ce qu'on connaît aujourd'hui sous le nom du *système standard de logique déontique*, dont le représentant est le système modal KD.

Peter Schotch et Ray Jennings

Schotch et Jennings (1981) critiquent l'interprétation de la logique déontique dans le cadre d'une sémantique de Kripke (1963). Cette critique, qui a pour objectif d'invalider l'interprétation modale de la logique déontique, vise une propriété fondamentale du système K, à savoir la distributivité de l'opérateur O sur la conjonction.

Leur argument consiste à dire que le système K n'est pas adéquat quant à la formalisation de *ce qui doit être* (*ought-to-be*), puisque la distributivité de O sur la conjonction fait s'effondrer la distinction déontique entre le schéma d'axiome (**D**) et le théorème $\neg O\bot$. L'effondrement de cette distinction empêche KD d'être en mesure de rendre compte des conflits d'obligations.[2]

D'entrée de jeu, la prémisse sur laquelle repose l'argument de Schotch et Jennings est qu'il existe des conflits insolubles d'obligations morales, comme dans le cas où un agent ferait plusieurs promesses incompatibles.[3]

Exemple

Il se pourrait très bien que Paul promette à sa soeur d'être chez elle à 13h30 le lundi 17 mai 2010 et qu'il promette à sa mère la même chose (étant entendu que la mère et la sœur n'habitent pas au même endroit). Considérant que si Paul est chez sa sœur, alors il n'est pas chez sa mère, il s'ensuit que dans une telle situation, Paul serait dans l'obligation d'être chez sa mère et de ne pas être chez sa mère. Autrement dit, il se pourrait très bien que Paul soit dans l'obligation de faire deux actions qui s'excluent mutuellement, donc qu'il se retrouve face à un conflit d'obligations insoluble.

En acceptant une telle prémisse et la possibilité d'une telle situation, il s'ensuit que le système adopté pour formaliser le discours normatif doit tenir compte de ce genre de conflit d'obligations. Or selon Schotch et Jen-

[2] Soulignons que les systèmes normaux ne satisfont pas nécessairement (**D**) et que ceux-ci sont aussi incapables de modéliser les conflits d'obligations.

[3] Sur ce point, voir le paradoxe de Lemmon (1962, 1965).

nings, KD ne permet pas de rendre compte d'une telle situation puisque le théorème (**A**) entraîne le paradoxe de l'agrégation complète, et par le fait même rend impossible le conflit d'obligations :

$$\vdash_K (O\varphi \wedge O\psi) \equiv O(\varphi \wedge \psi) \qquad (\mathbf{A})$$

Démonstration.

1	$O\varphi \wedge O\psi$	H
2	$O\varphi$	E$\wedge$ 1
3	$O\psi$	E$\wedge$ 1
4	$\quad O$	H
5	$\quad \varphi$	EO 2,4
6	$\quad \psi$	EO 3,4
7	$\quad \varphi \wedge \psi$	I$\wedge$ 5,6
8	$O(\varphi \wedge \psi)$	IO 4,7

1	$O(\varphi \wedge \psi)$	H
2	$\quad O$	H
3	$\quad \varphi \wedge \psi$	EO 1,2
4	$\quad \varphi$	E$\wedge$ 3
5	$O\varphi$	IO 2,4
6	$\quad O$	H
7	$\quad \varphi \wedge \psi$	EO 1,6
8	$\quad \psi$	E$\wedge$ 7
9	$O\psi$	IO 6,8
10	$O\varphi \wedge O\psi$	I$\wedge$ 5,9

□

En vertu de ce théorème, KD ne rend pas compte de la situation susmentionnée. En effet, dans l'exemple précédent, Paul a l'obligation d'être chez sa mère et l'obligation de ne pas être chez sa mère. Cependant, une telle situation n'est pas possible au sein de KD :

$$\vdash_{KD} \neg(O\varphi \wedge O\neg\varphi) \qquad (4.4)$$

Démonstration.

1	$O\varphi \wedge O\neg\varphi$	H
2	$O(\varphi \wedge \neg\varphi)$	(**A**)
3	$\quad O$	H
4	$\quad \varphi \wedge \neg\varphi$	EO 2,3
5	$\quad \bot$	I$\bot$ 4
6	$O\bot$	IO 3-5
7	$\neg O\bot$	thm 4.5
8	$\bot$	I$\bot$ 6,7
9	$(O\varphi \wedge O\neg\varphi) \supset \bot$	I$\supset$ 1-8
10	$\neg(O\varphi \wedge O\neg\varphi)$	df. $\neg$ 9

□

En ce sens, KD ne permet pas de rendre compte du conflit d'obligations puisque le système rend impossible une situation où un agent aurait deux obligations contradictoires, alors qu'une telle situation est tout à fait plausible.

Par surcroît, les auteurs soutiennent que le paradoxe de l'agrégation complète entraîne l'effondrement de la distinction entre (**D**) et $\neg O\bot$. D'une part, la signification de (**D**) est que si une action est obligatoire, alors elle est permise. Mais le sens de *permis* va ici plus loin que la signification déontique « ce qui est permis n'est pas interdit ». En effet, les auteurs interprètent l'axiome (**D**) dans le sens suivant : « si φ est une obligation, alors φ est non seulement permise, mais il est aussi possible d'accomplir φ ». D'autre part, la proposition $\neg O\bot$ est un théorème de KD :

$$\vdash_{KD} \neg O\bot \tag{4.5}$$

Démonstration.

1	O	H
2	$\quad\bot$	H
3	$\quad\bot$	R 2
4	$\bot \supset \bot$	I$\supset$ 2-3
5	$\neg\bot$	df. $\neg$ 4
6	$O\neg\bot$	IO 1,5
7	$O\bot \supset P\bot$	(**D**)
8	$O\bot$	H
9	$\neg O\neg\bot$	E$\supset$ 7,8
10	$\bot$	I$\bot$ 6,9
11	$O\bot \supset \bot$	I$\supset$ 8-10
12	$\neg O\bot$	df. $\neg$ 11

□

Bref, l'argument de Schotch et Jennings est que le paradoxe de l'agrégation, en empêchant la possibilité de conflits d'obligations, empêche par le fait même la possibilité de différencier ce théorème de l'axiome (**D**), puisque les deux deviennent un critère de consistance.

D'un côté, le théorème 4.5 stipule que l'absurde (voire l'*impossible*) ne peut pas être obligatoire, ce qui équivaut au principe *devoir* implique *pouvoir*.[4] Toutefois, cela implique aussi en vertu du théorème (**A**) que les obligations doivent nécessairement être cohérentes entre elles.

Dans le même ordre d'idées, l'axiome (**D**) est aussi un critère de consistance puisqu'il empêche la possibilité d'avoir deux obligations contradictoires. De fait, la distributivité de l'opérateur O au sein du système K, qui est par extension une propriété de KD, entraîne l'impossibilité d'avoir un conflit d'obligations et par le fait même réduit l'interprétation de l'axiome (**D**) et du théorème susmentionné à un critère de cohérence.

Selon Schotch et Jennings, chaque personne souscrit à différents systèmes d'obligations, lesquels peuvent entrer en contradiction les uns avec les autres. De fait, même si chaque système d'obligations pris individuellement doit être cohérent, il n'en demeure pas moins que le système total engendré par la conjonction des différents systèmes auxquels l'agent souscrit puisse contenir des obligations incompatibles.

Il est possible, par exemple, que les obligations légales d'une personne soient incompatibles avec ses obligations morales ou religieuses. En rejetant l'agrégation, le système total d'obligations ne sera pas incohérent puisque l'agrégation ne pourra pas être faite à l'intérieur du système. Autrement dit, même si deux obligations sont incompatibles, le système total reste cohérent puisqu'on ne peut pas conclure que la conjonction des obligations est obligatoire. En ce sens, il est donc possible qu'il y ait au sein du système total des conflits d'obligations, sans pour autant qu'il y ait d'incohérence. Puisque KD ne permet pas de telles situations, Schotch et Jennings en concluent que ce système n'est pas adéquat pour formaliser le concept d'obligation.

La solution qu'ils proposent en retour est d'élargir la sémantique afin de rendre l'opérateur O ambigu, plutôt que de la restreindre et d'ajouter des clauses sémantiques. Pour ce faire, Schotch et Jennings définissent une structure ${}^n\mathcal{F}$ à l'intérieur de laquelle il y a n relations :

$$ {}^n\mathcal{F} = < U, R_1, \ldots, R_n > \text{ où } U \neq \varnothing $$

À l'intérieur de cette structure, chaque relation R_i représente une relation entre une proposition et un ensemble d'obligations, et U est un ensemble non vide. Considérons par exemple la structure suivante.

[4] Soulignons ici que l'interprétation de $\neg O\bot$ en termes de *devoir implique pouvoir* est insidieuse. En effet, $\bot$ fait référence à une contradiction, c'est-à-dire à une impossibilité *logique*. Or le principe devoir implique pouvoir ne fait pas référence uniquement à la notion de possibilité logique, mais fait surtout référence à la notion de possibilité factuelle.

Exemple

$$^{3}\mathcal{F} = < U, R_1, R_2, R_3 >$$

Cette structure pourrait représenter une situation où une personne souscrit à un ensemble d'obligations légales, religieuses et familiales.

Dans cet exemple, la clause sémantique est que $O\varphi$ est vrai pour w si et seulement si φ est vrai pour tout v en relation R_1 avec w, ou φ est vrai pour tout v en relation R_2 avec w, ou φ est vrai pour tout v en relation R_3 avec w. Autrement dit, φ est une obligation dans w si et seulement si φ est vrai pour toutes les alternatives déontiques d'un des ensembles d'obligations auxquels souscrit l'agent. Formellement, la clause sémantique de Schotch et Jennings est la suivante :

$$\models_w O\varphi \Leftrightarrow \forall v,\ wR_1v \Rightarrow \models_v \varphi, \text{ ou } \ldots, \text{ ou } wR_nv \Rightarrow \models_v \varphi$$

En mots, cela équivaut à dire que $O\varphi$ est vrai pour w si et seulement s'il y a au minimum une relation R_i telle que φ est vrai pour tout v en relation R_i avec w. Une telle sémantique permet d'éliminer le paradoxe de l'agrégation complète puisque ce n'est pas parce que $O\varphi$ et $O\psi$ sont vrais dans w que $O(\varphi \wedge \psi)$ l'est aussi. En effet, il est possible que φ soit vrai en fonction d'une relation R_i alors que ψ le soit en fonction d'une relation R_j (où i $\neq$ j) sans pour autant que la conjonction de φ et ψ soit vraie dans un scénario en relation avec w.

L'interprétation sémantique de Schotch et Jennings permet donc d'éviter le paradoxe de l'agrégation complète, et le conflit d'obligations devient possible. En effet, il est possible d'avoir deux obligations incompatibles φ et ψ sans pour autant que leur conjonction, qui sera alors contradictoire (en supposant $\varphi \equiv \neg\psi$), soit aussi une obligation. La conséquence d'une telle approche est qu'il est possible de formaliser des conflits d'obligations, ce qui permet de mieux refléter l'utilisation du discours normatif.

Les paradoxes mettent en évidence que les systèmes monadiques standard ne sont pas en mesure d'offrir un langage assez riche pour traiter de l'obligation conditionnelle et des conflits d'obligations. Dans le cas de Schotch et Jennings, ces auteurs offrent surtout des modifications sémantiques relativement au système standard. Malgré cela, certains auteurs, à l'instar de Castañeda et Jones, tentent de surmonter les paradoxes tout en restant dans le cadre d'une logique déontique monadique en enrichissant son langage.

Hector-Neri Castañeda

L'approche de Castañeda (1981)[5], qui repose principalement sur la distinction entre la nature des propositions qui se trouvent dans la portée des opérateurs déontiques, s'insère dans le cadre des logiques du premier ordre.[6] D'emblée, son objectif est de présenter un système de logique déontique qui soit en mesure de rendre compte de la structure logique du discours normatif ordinaire, c'est-à-dire tel qu'il est utilisé. En effet, cet auteur cherche à représenter l'usage grammatical du discours normatif, de façon à résoudre les paradoxes et à formaliser certains raisonnements où un type particulier de proposition peut être sorti ou inséré dans la portée des opérateurs déontiques sans en changer le sens.

Considérons par exemple l'énoncé suivant : « s'il est obligatoire lorsqu'on conduit de ne pas boire d'alcool, alors lorsqu'on conduit, il est obligatoire de ne pas boire d'alcool ». Selon Castañeda, cet exemple met en lumière un élément fondamental à prendre en considération lors de la formalisation du discours déontique, à savoir qu'il est nécessaire de faire la distinction entre les propositions à l'infinitif et celles à l'indicatif dans la portée des opérateurs déontiques.

Alors que la proposition indicative *lorsqu'on conduit* décrit les conditions dans lesquelles l'action *ne pas boire d'alcool* est obligatoire, la proposition infinitive *ne pas boire d'alcool* est ce sur quoi porte réellement l'opérateur déontique.[7] Autrement dit, Castañeda différencie l'action en tant que circonstance et en tant qu'objet d'une proposition déontique. Une proposition qui dénote une action peut servir à décrire un contexte dans lequel une action (en tant qu'objet d'un opérateur) est considérée comme obligatoire.

Un aspect intéressant de la position de Castañeda est que celui-ci introduit une distinction sémantique et syntaxique entre les différents types d'obligations. En effet, considérant qu'un même type d'obligation ne peut être itéré, il introduit un indice $_i$ pour faire la distinction entre les différents types d'obligations.

[5] On trouve dans cet ouvrage l'apogée de l'approche de Castañeda, dont certaines idées similaires peuvent être trouvées notamment dans Castañeda (1959, 1968, 1970, 1977). Il est cependant à noter que les idées de l'auteur ont considérablement évolué depuis 1959. Même si la distinction de base entre les propositions considérées en tant que prescriptions et descriptions demeure, on constate un changement du côté de l'axiomatisation.

[6] Cette distinction entre les actions considérées *déontiquement* et celles considérées en tant que *circonstances* est au coeur de l'approche de Castañeda (1983).

[7] L'auteur utilise aussi *practitive* pour signifier que la proposition dénote l'action en tant que telle, et non simplement une description du contexte où l'action est exécutée.

Du côté de la syntaxe, Castañeda utilise comme symboles primitifs $\{\neg, \wedge, [,], (,), \forall, O_i, =\}$, où les crochets [ ] permettent d'identifier les propositions qui sont les objets des propositions déontiques et où les parenthèses identifient les propositions qui décrivent les contextes.[8]

Les autres connecteurs sont définis comme à l'habitude. Il utilise les variables propositionnelles $\{p, q, r, ...\}$ pour dénoter les propositions indicatives (de contexte) et $\{A, B, C, ...\}$ pour dénoter celles qui sont les objets des opérateurs déontiques (il utilise $p*$ comme méta-variable pour référer aux deux types de propositions).[9] Finalement, il utilise un ensemble de prédicats à n-places, dont le représentant est $C(1, ..., n)$ (ou $C[1, ..., n]$), et un ensemble de variables $\{x, y, z, ...\}$.

Cela fait, il construit un nombre dénombrable de structures déontiques D_i^* $(i \in \mathbb{N})$ où D_1^* représente l'obligation primordiale (*the overriding ought*)[10] et où les autres sont des obligations *prima facie*. L'ensemble des énoncés bien formés de D_i^* est défini de la manière suivante :

1. $C(x_1, ..., x_n), C[x_1, ..., x_n] \in EBF$ (où $x_1, ..., x_n$ sont des variables ou des constantes) ;
2. $x = y \in EBF$ (où x, y sont des variables ou des constantes) ;
3. si $p, q, \in EBF$, alors $\neg p, p \wedge q, (\forall x)p \in EBF$;
4. si $A, B, p \in EBF$, alors $\neg A, A \wedge B, p \wedge A, A \wedge p, (\forall x)A \in EBF$;
5. $O_i A \in EBF$ (où A est une méta-variable qui dénote l'objet d'un opérateur O_i).

En ce qui concerne les axiomes, Castañeda utilise les schémas (A0)–(A5) :

$$O_i A \supset C_i \tag{A0}$$

$$p* \text{ pour les tautologies du calcul propositionnel} \tag{A1}$$

$$O_i A \supset \neg O_i \neg A \tag{A2}$$

$$O_1 A \supset A \tag{A3}$$

$$(\forall x)p* \supset p*(x||y) \tag{A4}$$

$$(\exists y)(x = y) \tag{A5}$$

[8] Notons que l'auteur n'inclut pas le quantificateur universel dans ses symboles primitifs, ce qui est probalement un oubli. De fait, nous l'avons inclus dans la syntaxe.

[9] Notons que l'auteur utilise $p, q, r, ...$ et $A, B, C, ...$ autant comme variables que méta-variables propositionnelles.

[10] L'obligation primordiale est, en quelques sortes, l'obligation actuelle. Il s'agit du type d'obligation qui peut être utilisé à l'aide des clauses *toutes choses considérées* et *par-dessus tout*, à savoir que ce type d'obligation est celui qui prédomine dans un certain contexte en cas de conflit (Tomberlin 1983).

L'axiome (A0) indique que toute obligation O_iA va de pair avec les conditions qui lui sont nécessaires, (A3) stipule que les obligations prioritaires sont faites (dans les alternatives déontiques parfaites) et (A4) signifie que lors d'une quantification, il est possible de substituer toutes les variables x liées de $p*$ par n'importe quel objet du domaine. À ces axiomes s'ajoutent les règles (R0)–(R3)[11] :

$$\frac{p*, p* \supset q*}{q*} \qquad \text{(R0)}$$

$$\frac{\vdash_i (p* \wedge A_1 \wedge ... \wedge A_n) \supset B}{\vdash_i (\forall x)(p* \wedge O_iA_1 \wedge ... \wedge O_iA_n) \supset (O_i(\forall x)B \wedge (\forall x)O_iB)} \qquad \text{(R1)}$$

$$\frac{\vdash_i p* \supset q*}{\vdash_i p* \supset (\forall x)q*} \qquad \text{(R2) } (p* \text{ sans occurrence libre de } x)$$

$$\frac{\vdash_i p* \supset q*}{\vdash_i (x = y) \supset (p* \supset q*(x|y))} \qquad \text{(R3)}$$

La première règle est le *modus ponens* du calcul propositionnel, la seconde est une modification de la règle de généralisation pour le calcul du premier ordre, la troisième règle correspond à la conséquence des axiomes (4) et (5) du calcul des prédicats[12] et finalement la dernière règle indique qu'il est possible de substituer les termes identiques à l'intérieur des propositions. La règle (R1) est restreinte aux théorèmes qui ne dépendent pas de (A3) afin d'éviter d'obtenir des propositions de la forme $O_1A \supset O_iA$, ce qui empêcherait les conflits entre les différents types d'obligations.

Ce système permet d'obtenir la règle dérivée suivante :

$$\frac{\vdash_i p* \supset A}{\vdash_i p* \supset O_iA} \qquad \text{(R4)}$$

Or cette règle dérivée n'est certainement pas adéquate, ce qui peut être mis en évidence à l'aide du contre-exemple suivant. Si Paul prend l'argent de Pierre sans sa permission, alors Paul vole. Donc, si Paul prend l'argent de Pierre sans sa permission, alors obligatoirement Paul vole. Clairement, ce n'est pas parce que Paul prend l'argent de Pierre sans sa permission que Paul a l'obligation de voler !

[11] Certaines parenthèses sont omises pour faciliter la lecture.

[12] L'axiome 4 est $(\forall x)(B(x) \supset B(t))$ avec t libre pour x dans B et 5 est $(\forall x)(B \supset C) \supset (B \supset (\forall x)C)$ où x n'est pas libre dans B.

Du côté de la sémantique, Castañeda considère le sous-système D_i^{c*} qui n'inclut pas les quantificateurs, et donc qui repose sur les axiomes (A1)–(A2) et les règles (R0)–(R1'), où (R1') ne contient pas de quantificateur.

Un modèle $M = < W_0, W, a >$ pour D_i^{c*} contient un ensemble W de scénarios déontiques possibles, un scénario $W_0 \in W$ qui représente le monde actuel et une fonction $a : W_i \longrightarrow \{\bot, \top\}$ qui assigne des valeurs de vérité aux propositions dans W_i. Cette fonction assure la bivalence et la vérifonctionnalité habituelle, à l'exception de la quatrième clause :

1. pour tout $p* \in EBF$, $a_{W_j}(p*) = \top$ ou $a_{W_j}(p*) = \bot$;
2. $a_{W_j}(\neg p*) = \top \Leftrightarrow a_{W_j}(p*) = \bot$;
3. $a_{W_j}(p* \wedge q*) = \top \Leftrightarrow a_{W_j} I(p*) = \top$ et $a_{W_j}(q*) = \top$;
4. s'il y a un scénario W_j tel que $a_{W_j}(p) = \top$, alors $a_{W_h}(p) = \top$ pour tout $W_h \in W$.

La quatrième clause est utilisée afin de ne pas avoir à prendre en considération la temporalité. Considérant que W_0 est le monde actuel, tout ce qui est vrai dans W_0 sera vrai dans les autres scénarios. Autrement dit, tout scénario partage le même présent et le même passé, les alternatives déontiques étant caractérisées par les différences qui se trouvent dans le futur. En ce sens, toute alternative déontique possible possède comme sous-ensemble le monde actuel, c'est-à-dire que toute alternative déontique possible est une alternative du monde actuel.

Cela dit, les clauses sémantiques quant à l'attribution de valeur de vérité aux propositions déontiques sont :

5. $a_{W_0}(O_i A) = \top \Leftrightarrow a_{W_j}(A) = \top$ pour tout W_j différent de W_0 ;
6. $a_{W_0}(O_1 A) = \top \Leftrightarrow a_{W_j}(A) = \top$ pour tout $W_j \in W$.[13]

À partir de ce modèle, Castañeda construit une extension $M = < W, W_0, D, P, V, \pi, a >$, où D est un domaine de personnes et d'objets, V une fonction qui assigne des membres de D aux termes primitifs du langage D^*, π une fonction qui assigne des membres de P aux prédicats de D^* et I une fonction qui assigne des valeurs de vérité, à laquelle on ajoute une clause pour les formules quantifiées.[14]

[13] Cette clause entraîne une confusion du point de vue de l'obligation : tout ce qui est vrai dans le scénario actuel est une obligation primordiale.

[14] La clause sémantique est que $I(C(x_1, ..., x_n), W_j) = \top$ si et seulement si $C(x_1, ..., x_n)$ est satisfait par la suite $< x_1, ...x_n, ... > \in D$.

Andrew J. I. Jones et Ingmar Pörn

Les travaux de Jones et Pörn ont eu des répercussions considérables du point de vue de l'application de la logique déontique au discours juridique, notamment en ce qui à trait à la représentation de la connaissance légale en informatique. Leur motivation initiale dans le développement du système DL était d'offrir une solution au paradoxe de Chisholm (Jones 1991). En effet, ces auteurs voulaient offrir un système en mesure de rendre compte des obligations qui surgissent dans des contextes de violation.

L'approche de Jones et Pörn utilise une sémantique des mondes possibles, à laquelle ils ajoutent une clause visant à rendre compte des situations quasi-parfaites ou encore sous-optimales. Ainsi, les scénarios accessibles à un scénario w comprennent des alternatives idéales, où toutes les obligations sont remplies, et des alternatives sous-optimales, où certaines obligations sont violées.

Le système proposé est une extension du système standard de logique déontique, et donc un système fortement normal. D'entrée de jeu, Jones et Pörn (1986) voient trois raisons pour modifier KD. Premièrement, ce système ne parvient pas à rendre compte du fait que les obligations sont *violables*, à savoir qu'il est toujours possible pour un agent d'agir à l'encontre de ses obligations. Autrement dit, KD ne satisfait pas le principe de contingence normative. Deuxièmement, KD ne permet pas de faire la distinction entre les termes *ought* et *must*, au même titre qu'il ne parvient pas à différencier les obligations *prima facie* des obligations actuelles. Finalement, les paradoxes, surtout celui de Chisholm, montrent que KD ne rend pas compte adéquatement du discours normatif.

Le système DL est une extension de KD[15] auquel on ajoute l'opérateur O' (Jones et Pörn 1985). Cet opérateur, qui se comporte exactement comme l'opérateur O, diffère par sa clause sémantique : $O'\varphi$ est vrai pour w si et seulement si φ est vrai pour toutes les alternatives sous-optimales v de w :

$$\models_w O'\varphi \Leftrightarrow \models_v \varphi \; \forall v \text{ t.q. } wR_{O'}v$$

Cette clause sémantique vise à rendre compte du fait que certains scéarios sont des alternatives au monde actuel bien que certaines obligations ne soient pas respectées. En ce sens, $R_{O'}$ caractérise les alternatives sous-optimales à un scénario, où certaines obligations sont violées.

[15] Rappelons au passage que le modèle sémantique de KD est caractérisé par une relation sérielle, et donc que pour tout scénario w, il existe au moins une alternative idéale (i.e., une alternative déontique parfaite).

Le modèle de KD est donc enrichi par la relation $R_{O'}$, restreinte par deux conditions :

$$R \cap R_{O'} = \varnothing$$
$$\Delta_U \subseteq R \cup R_{O'} \qquad (\text{où } \Delta_U = \{(u,u) : u \in U\})$$

Le symbole $\varnothing$ dénote l'ensemble vide. Ces deux restrictions, qui stipulent respectivement qu'aucune alternative idéale ne peut être aussi une alternative sous-optimale et que tout scénario u est soit une alternative idéale soit une alternative sous-optimale à lui-même, sont accompagnées du côté de la syntaxe par les axiomes (OP') et (OO')[16] :

$$Op \wedge O'\neg p \qquad \text{(OP')}$$
$$(O\varphi \wedge O'\varphi) \supset \varphi \qquad \text{(OO')}$$

Le premier schéma d'axiome porte sur des variables propositionnelles atomiques alors que le second porte sur des méta-variables. Alors que le premier axiome équivaut à soutenir qu'il est faux qu'une proposition vraie dans toutes les alternatives idéales possibles le soit aussi dans toutes les alternatives sous-optimales possibles, le second signifie que pour tout scénario w, soit φ n'est pas vrai pour toutes les alternatives idéales à w, soit φ n'est pas vrai pour toutes les alternatives sous-optimales à w, soit φ est vrai pour w. Autrement dit, tout w étant une alternative idéale ou sous-optimale à lui-même, il s'ensuit que si φ est vrai pour toutes ses alternatives idéales et sous-optimales, φ est vrai pour w.

Les énoncés bien formés sont similaires à ceux de KD, où on ajoute la condition suivante[17] :

$$\varphi \in EBF \Rightarrow O'\varphi \in EBF$$

Soulignons que l'interprétation dans la langue naturelle d'une proposition de la forme $O\varphi$ est différente dans DL. En effet, $O\varphi$ signifie simplement que, pour un scénario donné, φ est vrai pour toutes ses alternatives idéales, ce qui est une lecture littérale de la clause sémantique de O.

Afin de rendre compte du fait que les obligations peuvent être violées, Jones et Pörn formalisent la notion d'*obligation* (ou de *devoir*) de la manière suivante :

$$Ought\ \varphi \ =_{df}\ O\varphi \wedge P'\neg\varphi \qquad \text{(df. } Ought\text{)}$$

[16] Dans leurs textes, Jones et Pörn utilisent les variables propositionnelles à titre de méta-variables. Nous changeons la notation de façon à éviter la confusion.

[17] L'itération des opérateurs déontiques fait cependant changer leur signification.

En mots, *il est obligatoire que* φ signifie que φ est vrai dans toutes les alternatives idéales au monde actuel et qu'il existe une alternative sous-optimale où φ est faux.

Par la suite, Jones et Pörn (1986) ajouteront l'opérateur suivant :

$$N_d\varphi =_{df} O\varphi \wedge O'\varphi \qquad \text{(df. } N_d\text{)}$$

Cet ajout vise à rendre compte de la notion de *nécessité déontique*, qui permet d'exprimer les obligations qui tiennent dans toutes les alternatives (idéales et sous-optimales) d'une situation donnée. Principalement, l'ajout de cette notion sert à formaliser la modalité *Must*, laquelle est représentée par l'expression N_d *Ought* φ.

La notion de nécessité déontique sert principalement à exprimer les choses qui ne peuvent pas être changées dans les alternatives (idéales ou sous-optimales) d'une situation. Ainsi, conformément à la définition de N_d et à l'axiome (OO'), toute nécessité déontique est vraie dans le monde actuel. La modalité *Must* vise à représenter les obligations actuelles (primordiales) d'un agent, contrairement à *Ought*, qui exprime les obligations *prima facie*.

Par la suite, Jones (1991) introduira l'opérateur O_δ pour mieux rendre compte du détachement des obligations conditionnelles. L'ajout de cet opérateur à DL, qui donne le système DL_{min}, amène Jones à redéfinir la sémantique par un modèle minimal (Chellas 1980).

Considérant que les obligations conditionnelles $\psi \supset$ *Ought* φ indiquent des obligations qui peuvent varier d'un contexte à l'autre, où l'ajout de certaines conditions ψ' peut influencer le caractère obligatoire de φ, un énoncé $O_\delta\varphi$ est vrai pour w si et seulement si φ est vrai pour toutes les alternatives sous-optimales *normales* de w.[18]

Cet opérateur entraîne une redéfinition du modèle puisque sa clause sémantique s'exprime à l'aide d'un modèle minimal, qui permet de définir la vérité en fonction d'un sous-ensemble de $R_{O'}$. Même si Jones (1991) redéfinit le modèle, nous allons exprimer la sémantique de O_δ à l'aide des concepts que nous avons utilisés jusqu'à maintenant.

Un énoncé de la forme $O_\delta\varphi$ est vrai pour w si et seulement si φ est vrai pour toute alternative sous-optimale normale de w, où une alternative sous-optimale normale appartient à l'ensemble des alternatives sous-optimales dont aucune condition ψ' ne change le caractère obligatoire de φ. Formellement, nous avons la clause suivante :

$$a_w(O_\delta\varphi) = \top \Leftrightarrow a_v(\varphi) = \top \; \forall v \text{ t.q. } wR_\delta v$$

[18] Jones utilise l'expression *typical/unexceptional sub-ideal version of w.*

En d'autres termes, l'idée est de définir une sous-relation R_δ de $R_{O'}$ (c'est-à-dire $R_\delta \subset R_{O'}$) qui permet d'identifier les alternatives sous-optimales normales de w. À l'aide de ce nouvel opérateur, Jones définit une notion plus faible de nécessité déontique :

$$N_\delta\varphi =_{df} O\varphi \wedge O_\delta\varphi \qquad \text{(df. } N_\delta\text{)}$$

Cela permet de restreindre la nécessité déontique à certaines classes d'alternatives sous-optimales normales. Il formalisera l'obligation conditionnelle par la proposition $N_\delta(\psi \supset Ought\ \varphi)$. Par définition, cela signifie que $\psi \supset Ought\ \varphi$ est vrai pour w dans toutes les alternatives idéales et dans toutes les alternatives sous-optimales normales à w, c'est-à-dire celles où aucune condition ψ' n'influence le caractère obligatoire de φ.

* * *

Tout compte fait, il existe plusieurs autres approches de type monadique, dont certaines seront abordées dans les chapitres qui suivent. Malgré le raffinement des systèmes, il n'en demeure pas moins que les approches monadiques ont de la difficulté à modéliser les obligations conditionnelles. Afin de résoudre ces problèmes, certains auteurs ont suggéré de modéliser les raisonnements normatifs conditionnels à l'aide d'opérateurs stipulant les conditions dans lesquelles les obligations tiennent. Nous allons maintenant aborder ces approches.

Chapitre 5

Les systèmes dyadiques

Le présent chapitre ne prétend pas fournir une revue exhaustive de la littérature scientifique sur l'obligation conditionnelle et les logiques déontiques dyadique, mais vise plutôt à présenter son origine, son évolution et les principaux problèmes auxquels elle fait face. Tomberlin (1981) considère les systèmes de Al-Hibri (1978), Chellas (1974), Mott (1973) et van Fraassen (1972) comme les meilleurs propositions qui traitent de l'obligation conditionnelle. Nous n'examinerons ici que celles de Chellas et van Fraassen, notamment en raison du fait que Chellas propose une bonne analyse des problèmes attribuables à la définition de l'opérateur dyadique et que van Fraassen utilise une sémantique des préférences. Nous aborderons cependant d'abord von Wright et Rescher de façon à montrer les débuts de la logique déontique dyadique, et nous conclurons avec les problèmes du détachement et de l'augmentation.

Georg Henrik von Wright

Les bases de la logique déontique dyadique sont apparues pour la première fois dans l'article de von Wright (1956), rédigé en réaction au paradoxe de Prior (1954). L'objectif de von Wright était de formaliser la notion d'*engagement* de façon à éviter le paradoxe de l'obligation dérivée et à rendre compte de l'aspect conditionnel de certaines obligations.

Von Wright propose alors de formaliser l'obligation conditionnelle par $O(\varphi/\psi)$, qui se lit « dans les conditions ψ, φ est obligatoire ». Certains auteurs auront préféré par la suite écrire $O_\psi\varphi$.

La première tentative de von Wright s'est traduite par les deux axiomes suivants :

$$P(\varphi/\psi) \vee P(\neg\varphi/\psi) \tag{A1}$$

$$P(\varphi \wedge \rho/\psi) \equiv (P(\varphi/\psi) \wedge P(\rho/\psi \wedge \varphi)) \tag{A2}$$

En mots, cette axiomatisation consiste à dire que pour toute circonstance ψ, soit φ est permis ou $\neg\varphi$ l'est, et que si deux actes φ et ρ sont permis dans des circonstances ψ, alors φ est permis dans les circonstances ψ et ρ est permis dans les circonstances où l'acte φ est performé dans les circonstances ψ.

Cette première tentative n'était cependant qu'une ébauche visant à jeter les bases à partir desquelles l'obligation dérivée et l'engagement pourraient être formalisés. Plus tard, en réponse à Rescher (1958), von Wright (1967) en viendra à développer un système axiomatique différent qui, en fonction des différentes extensions construites à partir de ce système, vise à rendre compte de différents degrés de permissibilité.

Le système dyadique de base consiste en une extension de la logique propositionnelle, à laquelle on ajoute trois axiomes :

$$P(\varphi/\psi) \equiv \neg O(\neg\varphi/\psi) \tag{A1}$$

$$P(\varphi \vee \rho/\psi \vee \tau) \equiv [P(\varphi/\psi) \vee P(\varphi/\tau) \vee P(\rho/\psi) \vee P(\rho/\tau)] \tag{A2}$$

$$P(\varphi \vee \neg\varphi/\rho \vee \neg\rho) \tag{A3}$$

Autrement dit :

1. si φ est permis dans les conditions ψ, alors $\neg\varphi$ n'est pas obligatoire dans les mêmes conditions ;
2. si une disjonction d'actes est permise dans une disjonction de contextes, alors il y a au moins un acte de permis dans un des contextes ;
3. toute tautologie est permise dans des contextes tautologiques.

L'ensemble des énoncés bien formés, qui n'admet pas l'itération des opérateurs déontiques, est défini récursivement par :

1. si $\varphi, \psi \in EBF_{LPC}$, alors $O(\varphi/\psi) \in EBF$;
2. si $\varphi, \psi \in EBF$, alors $\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi, \varphi \supset \psi, \varphi \equiv \psi \in EBF$.

À l'aide de ces nouveaux schémas d'axiomes, von Wright développe quatre extensions qui se différencient principalement de par l'axiome de distributivité (A2).[1] Ces quatre systèmes ont pour objectif de rendre compte de quatre interprétations différentes de la notion de *permission*.

[1] Nous présenterons seulement un système et une interprétation de la permission.

L'interprétation sémantique que fait l'auteur de ces systèmes prend place dans le cadre d'une sémantique des mondes possibles. En effet, son interprétation de la permission $P_1(\varphi/\psi)$ est que parmi certains mondes possibles où se trouve le contexte ψ, il y a certains mondes possibles permis où φ est vrai. Autrement dit, certains mondes possibles qui contiennent φ sont des alternatives permises à certains mondes possibles qui contiennent ψ :

$$\exists w,\ \exists v \text{ t.q. } a_w(\psi) = \top, a_v(\varphi) = \top \text{ et } v \text{ est permis} \qquad (P_1)$$

Bref, il existe certains mondes possibles dans lesquels il y a le contexte ψ et où l'action φ est permise. Cette interprétation de la permission, que von Wright considère comme la plus *faible*, donne place à une notion *forte* d'obligation : $O_1(\varphi/\psi)$ signifie que dans tous les mondes possibles où ψ est vrai, tous les mondes possibles où φ n'est pas vrai sont interdits :

$$\forall w \text{ et } \forall v,\ \text{si } a_w(\psi) = \top \text{ et } a_v(\varphi) = \bot,\ \text{alors } v \text{ est interdit} \qquad (O_1)$$

En d'autres termes, tout monde possible qui découle du fait que le contexte ψ est vrai et où φ est vrai est obligatoire, c'est-à-dire que si un agent se trouve dans une situation w où le contexte ψ est vrai, alors il est dans l'obligation d'opter pour un scénario v où φ est vrai.

L'axiome de distributivité propre à P_1 est l'axiome (A2) susmentionné. Cette interprétation, qui n'est pas adéquate pour la formalisation des obligations *prima facie*, est celle d'une obligation actuelle (ou primordiale) dont aucune autre condition ne peut venir en affecter le caractère obligatoire.

Outre la contribution de von Wright pour les débuts de la logique déontique dyadique, les travaux de Rescher ont aussi eu une place importante quant à la formalisation de la notion de *permission conditionnelle.*

Nicholas Rescher

À la suite des bases jetées par von Wright (1956), Rescher (1958) a proposé un système de logique déontique dyadique visant à rendre compte de la permission conditionnelle. Le système de Rescher diffère de celui de von Wright (1967) de par son axiomatisation, mais aussi dans la mesure où il admet comme formules bien formées des énoncés purement propositionnels (p. ex. A3).[2]

[2] Soulignons au passage le caractère inapproprié de l'axiome (A6) : ce n'est pas parce qu'une action est permise dans des circonstances tautologiques qu'elle est nécessairement permise dans n'importe quelles circonstances. Il est permis de prendre une bière dans des circonstances où il fait soleil ou il ne fait pas soleil, sans pour autant qu'il soit permis de prendre une bière en voiture !

$$P(\varphi \vee \neg\varphi/\psi) \tag{A1}$$
$$P(\varphi \vee \rho/\psi) \equiv (P(\varphi/\psi) \vee P(\rho/\psi)) \tag{A2}$$
$$(\varphi \supset \rho) \supset (P(\varphi/\psi) \supset P(\rho/\psi)) \tag{A3}$$
$$P(\varphi \wedge \rho/\psi) \supset P(\varphi/\psi \wedge \rho) \tag{A4}$$
$$(P(\varphi/\psi) \wedge P(\rho/\psi \wedge \varphi)) \supset P(\varphi \wedge \rho/\psi) \tag{A5}$$
$$P(\varphi/\psi \vee \neg\psi) \supset P(\varphi/\tau) \tag{A6}$$
$$P(\varphi/\tau) \supset P(\varphi/\psi \wedge \neg\psi) \tag{A7}$$

En mots, cela équivaut à soutenir que :

1. dans toute circonstance, soit une action est permise soit sa négation l'est ;
2. si une disjonction d'actions est permise dans certaines circonstances, alors au moins une des deux actions est permise dans ces circonstances ;
3. si φ implique ρ, alors si φ est permis dans les circonstances ψ, ρ l'est aussi ;
4. si une conjonction d'actions φ et ρ est permise dans des circonstances ψ, alors l'action φ est permise dans les circonstances ψ où l'action ρ est posée ;
5. si une action φ est permise dans des circonstances ψ et que l'action ρ est permise dans les mêmes circonstances et où l'action φ est posée, alors la conjonction de φ et ρ est permise dans les circonstances ψ ;
6. une action permise dans des circonstances tautologiques est permise dans n'importe quelles circonstances ;
7. une action permise dans certaines circonstances est aussi permise dans des circonstances contradictoires.

Par la suite, Rescher (1962) en viendra à modifier l'axiome (A5) pour répondre à l'objection soulevée par Anderson (1959), qui consiste à dire que l'axiome permet de dériver un théorème non acceptable dans une interprétation déontique.[3] En effet, (A5) permet de dériver la proposition :

$$(P(\varphi/\psi) \wedge P(\neg\varphi/\psi \wedge \varphi)) \supset P(\rho/\psi) \tag{5.1}$$

Autrement dit, si deux actions contraires sont permises dans certaines circonstances, alors n'importe quoi est permis. Or cela est inacceptable

[3] Rescher (1962), en répondant à l'objection de Anderson (1959), en profitera aussi pour critiquer sa position. Anderson (1962) répondra à ces objections. Plus tard, Robison (1967) proposera une autre critique de l'axiomatisation de Rescher.

dans une perspective normative puisqu'il est possible qu'une action et son contraire soient permis, sans pour autant que n'importe quoi le soit. Ce n'est pas parce qu'il est permis de fumer ou permis de ne pas fumer dans un endroit fumeur qu'il est permis de voler ceux qui se trouvent à cet endroit !

En réponse à l'objection, Rescher modifie l'axiome (A5) en introduisant une modalité aléthique :

$$(P(\varphi/\psi) \wedge P(\rho/\psi \wedge \varphi) \wedge \Diamond(\varphi \wedge \rho)) \supset P(\varphi \wedge \rho/\psi) \qquad \text{(A5*)}$$

L'introduction de l'opérateur $\Diamond$, qui signifie la possibilité, permet à l'auteur de restreindre l'utilisation de (A5) et d'éviter l'objection. En effet, (A5*) permet d'éviter l'objection puisque la clause $\Diamond(\varphi \wedge \rho)$ indique qu'il doit être possible de faire la conjonction des deux actions. Puisqu'il est impossible de faire à la fois une action et son contraire, par exemple marcher tout en ne marchant pas, il s'ensuit que (A5*) ne peut pas être utilisé avec φ et $\neg\varphi$.

Finalement, Rescher (1967) proposera une sémantique pour la permission conditionnelle. D'entrée de jeu, l'interprétation de Rescher s'insère dans le cadre d'une sémantique des mondes possibles. Suivant la tradition, l'auteur définit un ensemble non vide de mondes possibles $W = \{w_1, \ldots, w_n, \ldots\}$, où $\{p\}$ et $[p]$ représentent respectivement l'ensemble des mondes possibles où p est vrai et l'ensemble des mondes possibles où p est permis. Les conditions sont les suivantes :

$$\{\neg p\} = W - \{p\} \tag{5.2}$$
$$\{p \wedge q\} = \{p\} \cap \{q\} \tag{5.3}$$
$$\{p \vee q\} = \{p\} \cup \{q\} \tag{5.4}$$
$$\text{si } p \supset q, \text{ alors } \{p\} \subset \{q\} \tag{5.5}$$

En mots, cela équivaut à dire que l'ensemble des mondes possibles où $\neg p$ est vrai équivaut au complément dans W de l'ensemble des mondes possibles où p est vrai (l'univers du discours se divise en deux classes d'ensembles, celle où p est vrai et celle où p ne l'est pas). La seconde condition indique que la classe des mondes possibles où la conjonction $p \wedge q$ est vraie est l'intersection des classes où p est vrai et où q est vrai. La troisième indique que la classe des mondes possibles où la disjonction $p \vee q$ est vraie est l'union de celles où p est vrai et où q est vrai, et la quatrième que si $p \supset q$, alors la classe des mondes possibles où p est vrai est un sous-ensemble de celle où q l'est.

Pour calculer les valeurs de la permission, l'auteur doit introduire un principe de conséquence :

$$\text{si } p \supset q, \text{ alors } [p] \subset [q] \tag{5.6}$$

Cela permet d'obtenir les deux règles pour calculer les valeurs de vérité de la permission :

$$[p \wedge q] \subset ([p] \cap [q]) \tag{5.7}$$
$$([p] \cup [q]) \subset [p \vee q] \tag{5.8}$$

Après avoir défini la machinerie de base, Rescher énonce quatre interprétations possibles de $P(\varphi/\psi)$:

1. φ est permis dans tous les mondes possibles où ψ est vrai ;
2. φ est permis dans la majorité des mondes possibles où ψ est vrai ;
3. φ est permis dans presque tous les mondes possibles où ψ est vrai ;
4. φ est permis dans tous les cas probables où ψ est vrai.

Les interprétations 1 et 3 sont celles qu'il retiendra.

Bas C. van Fraassen

En réaction au paradoxe de Chisholm, van Fraassen (1972)[4] a proposé une lecture axiologique de la logique déontique dyadique. Partant de l'intuition que φ *est obligatoire dans les conditions* ψ équivaut à soutenir qu'une situation où se trouvent $\varphi \wedge \psi$ est préférable à une ou se trouverait $\neg\varphi \wedge \psi$[5], l'auteur en vient à formuler le système CD de la manière suivante :

$$\text{Schémas d'axiomes de } LPC \tag{A1}$$
$$O(\varphi/\psi) \supset \neg O(\neg\varphi/\psi) \tag{A2}$$
$$O(\varphi \supset \rho/\psi) \supset (O(\varphi/\psi) \supset O(\rho/\psi)) \tag{A3}$$
$$O(\varphi/\psi) \supset O(\varphi \wedge \psi/\psi) \tag{A4}$$
$$B(\varphi/\rho) \supset (B(\rho/\psi) \supset B(\varphi/\psi)) \tag{A5}$$

[4] Nous avons choisi le texte de van Fraassen (1972) comme paradigme de la logique dyadique axée sur une sémantique des préférences. Le lecteur intéressé par ce type d'approche peut consulter notamment Åqvist (1986), Danielsson (1968) et Hansson (1990, 1997).

[5] Formellement, $O(\varphi/\psi) \Leftrightarrow B(\varphi \wedge \psi/\neg\varphi \wedge \psi)$.

$$\neg B(\varphi/\rho) \supset (B(\varphi/\psi) \supset B(\rho/\psi)) \quad \text{(A6)}$$
$$\neg B(\varphi/\rho) \supset (B(\psi/\rho) \supset B(\psi/\varphi)) \quad \text{(A7)}$$

Alors que les axiomes (A1)–(A4) sont proprement déontiques, (A5)–(A7) visent à représenter la transitivité de la préférence, où $B(\varphi/\psi)$ se lit φ est préférable à ψ. La préférence est définie par :

$$B(\varphi/\psi) =_{df} O(\neg\psi/\varphi \vee \psi) \quad \text{(df. } B\text{)}$$

De façon équivalente :

$$O(\varphi/\psi) =_{df} B(\varphi \wedge \psi/\neg\varphi \wedge \psi)$$

Autrement dit, si φ est obligatoire dans les conditions ψ, alors le scénario où φ et ψ sont vrais est préférable à celui où φ est faux, mais ψ est vrai.

Cela dit, notons qu'il n'est pas si évident que la définition de l'obligation en termes de préférences soit adéquate. Considérons l'exemple suivant.

Exemple

1. *Il est obligatoire de ne pas conduire dans des conditions où soit on boit de l'alcool, soit on conduit.*

$\therefore$ *Il est préférable de boire de l'alcool plutôt que de conduire.*

À ces sept axiomes s'ajoutent les règles (R1)–(R4), qui représentent respectivement le *modus ponens*, l'équivalent dyadique de (ROM) (Chellas 1980), l'équivalent dyadique de (Nec), et finalement une règle de substitution qui stipule que si φ est obligatoire dans les conditions ψ, et que les conditions ρ sont équivalentes à ψ, alors φ est aussi obligatoire dans les conditions ρ.

$$\frac{\vdash \varphi \supset \psi \qquad \vdash \varphi}{\vdash \psi} \text{ (R1)} \qquad \frac{\vdash \varphi \supset \psi}{\vdash O(\varphi/\rho) \supset O(\psi/\rho)} \text{ (R2)}$$

$$\frac{\vdash \varphi}{\vdash O(\varphi/\varphi)} \text{ (R3)} \qquad \frac{\vdash \psi \equiv \rho}{\vdash O(\varphi/\psi) \equiv O(\varphi/\rho)} \text{ (R4)}$$

En ce qui à trait aux énoncés bien formés, van Fraassen admet autant les propositions mixtes que celles où se trouvent des opérateurs déontiques itérés. L'ensemble des énoncés bien formés est défini récursivement à partir d'un ensemble dénombrable d'atomes propositionnels $Prop = \{p_1, ..., p_n, ...\}$ et du langage $\mathcal{L} = \{Prop, (,), \neg, \supset, O(/)\}$:

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid O(\varphi/\psi)$$

En vertu de la prémisse axiologique de l'auteur, celui-ci en vient à construire une sémantique qui requiert un peu plus de matériel que celle des systèmes habituels. Dans un premier temps, van Fraassen définit un modèle $\mathcal{M} = \langle K, V, R, f \rangle$, où $K, V \neq \varnothing$, R est une relation d'ordre antisymétrique, transitive et connexe sur un sous-ensemble non vide de V et f est un morphisme de K vers les sous-ensembles ordonnés de V.[6]

Autrement dit, K est un ensemble de scénarios possibles, V est un ensemble ordonné par R de scénarios possibles et f est une fonction qui permet de hiérarchiser les scénarios possibles de K en fonction de V, et donc de déterminer quels scénarios sont préférables par rapport à d'autres.

Cela dit, van Fraassen en vient à définir les conditions dans lesquelles les propositions à l'intérieur du modèle sont vraies. Outre les clauses sémantiques propres à la logique propositionnelle classique (incluant donc le postulat de bivalence), l'innovation de l'auteur concerne l'attribution de valeur de vérité aux propositions déontiques, laquelle est conforme à la clause suivante :

$$\begin{aligned} a_\alpha(O(\varphi/\psi)) = \top \Leftrightarrow \exists \beta, w \in K \text{ t.q.} \\ & a_\beta(\varphi \wedge \psi) = \top; \\ & w \in f_\alpha(\beta); \\ & \forall u \in f_\alpha(\gamma) \text{ et } \forall \gamma \text{ t.q. } a_\gamma(\neg\varphi \wedge \psi) = \top, wRu \end{aligned}$$

Cela équivaut à soutenir que $O(\varphi/\psi)$ est vrai pour α si et seulement si $\varphi \wedge \psi$ est vrai pour un β tel que w est un élément de l'ensemble des suites ordonnées incluant β et que w est préférable à tout scénario u qui appartient à l'ensemble des suites ordonnées incluant le scénario γ et où $\neg\varphi \wedge \psi$ est vrai pour γ.

Autrement dit, $O(\varphi/\psi)$ est vrai pour α si et seulement s'il existe un scénario β, pour lequel $\varphi \wedge \psi$ est vrai, qui fait partie d'une suite d'alternatives possibles où se trouve au moins un scénario w préférable à toute alternative γ où φ est faux, mais ψ est vrai.

En ce sens, cette clause sémantique est conforme à l'intuition de départ à savoir qu'il est vrai que φ est obligatoire dans les conditions ψ à condition qu'il y ait au moins un scénario où $\varphi \wedge \psi$ est préférable à toute alternative qui inclut $\neg\varphi \wedge \psi$.

La sémantique des préférences a notamment été introduite afin de rendre compte du caractère conditionnel de l'obligation. En effet, la hiérarchisation de l'ensemble des alternatives déontiques possibles vise à rendre compte du fait que dans certaines situations où se trouve un ensemble de

[6] Une relation connexe, dans ce contexte, est telle que pour toute paire w, v, soit wRv ou vRw.

conditions ψ, la réalisation d'une obligation φ est préférable à toute alternative où φ n'est pas accomplie. Ainsi, il devient possible de rendre compte des situations où certaines obligations ne sont pas réalisées puisqu'il est toujours possible de trouver une alternative préférable à une autre, même dans un cas où certaines obligations sont enfreintes. Cela dit, il est néanmoins possible d'obtenir le même genre de résultat sans pour autant utiliser un modèle fondé sur une sémantique des préférences.

Brian F. Chellas

Chellas, par exemple, développe une sémantique qui permet de capter les alternatives déontiques attribuables à certaines conditions particulières. Dans son introduction aux logiques modales, Chellas (1980) relève quelques difficultés relativement à la définition de l'opérateur dyadique $O(/)$ en termes monadiques. Les définitions potentielles de $O(\varphi/\psi)$ sont :

$$\psi \supset O\varphi \tag{1}$$

$$O(\psi \supset \varphi) \tag{2}$$

$$\Box(\psi \supset O\varphi) \tag{3}$$

L'opérateur $\Box$ représente une notion de nécessité. Ces définitions, cependant, sont à rejeter. D'abord, (1) rend $O(\varphi/\psi)$ vrai dès que ψ est faux en vertu des règles du calcul propositionnel :

$$\neg\psi \supset (\psi \supset O\varphi)$$

Par ailleurs, (2) rend $O(\varphi/\psi)$ vrai dès que $O\neg\psi$ ou $O\varphi$ l'est.

Exemples

1. *La définition (1) signifie que dans une situation où Pierre ne paie pas ses taxes, il est possible de conclure que Pierre a l'obligation de commettre l'adultère dans les conditions où il paie ses taxes.*
2. *La définition (2) implique que dans une situation où Pierre a l'obligation de porter assistance à son prochain, il a aussi l'obligation de porter assistance à son prochain même dans des conditions où cela mettrait sa vie en danger (ce qui est évidemment faux).*

Finalement, (1)–(3) impliquent que si φ est obligatoire dans les conditions ψ, alors φ est obligatoire dans les conditions $\psi \wedge \rho$ pour n'importe quel ρ. Or, les obligations conditionnelles étant caractérisées par le fait que l'ajout de certaines conditions peut influencer le caractère obligatoire d'une action, cela n'est pas représentatif du discours normatif. De fait, dans les trois cas, la définition de l'obligation conditionnelle en termes monadiques mène à des conséquences indésirables.

La difficulté de définir l'obligation conditionnelle à l'aide d'un opérateur monadique se comprend aussi en fonction des problèmes du détachement et de l'augmentation. D'une part, le problème du détachement de l'obligation conditionnelle peut s'apercevoir sous différents angles.

D'abord, il y a distinction entre le détachement *factuel* et le détachement *déontique* des obligations conditionnelles. Nute et Yu (1997), par exemple, argumentent que le paradoxe de Chisholm survient parce que les deux formes de détachement sont disponibles au sein du système standard. En effet, dans le paradoxe de Chisholm, l'obligation conditionnelle Oq est factuellement détachée de la norme qui dicte ce qui doit être fait dans le contexte d'une violation $p \supset Oq$ et le contexte p, et l'obligation $O\neg q$ est déontiquement détachée de $O\neg p$ et $O(\neg p \supset \neg q)$.

Par ailleurs, le détachement factuel pose problème dans la mesure où celui-ci n'est désirable que dans certains contextes (van Eck 1982a). Par exemple, dans un contexte de violation, il semble que l'obligation Oq puisse être factuellement détachée, et en ce sens, que le détachement factuel soit souhaitable. Cependant, cela ne peut être érigé en principe général. En effet, d'autres conditions peuvent bloquer le détachement factuel. Par exemple, si Pierre est en danger, alors Paul doit lui porter assistance, mais si cela met la vie de Paul en danger aussi, alors l'obligation ne peut être factuellement détachée du contexte. Par conséquent, il semble que le détachement factuel soit souhaitable, mais seulement lorsqu'on peut assurer que d'autres faits ne le bloqueront pas.

Le détachement factuel non restreint est donc à mettre de côté, mais le détachement factuel restreint est souhaitable, d'où le problème du détachement factuel. Le lecteur est invité à consulter Loewer et Belzer (1983) et Vorobej (1986) pour une discussion sur le détachement, ainsi que Straßer (2011) et Peterson (2015b,a) pour une analyse plus récente.

D'autre part, l'obligation conditionnelle fait face au problème de l'augmentation (Jones 1991), voire du renforcement de l'antécédent d'une obligation conditionnelle (Alchourrón 1996). Ce problème n'est pas sans lien avec celui du détachement factuel (Peterson 2015b). En un mot, cela consiste à dire qu'un conditionnel normatif $\psi \supset O\varphi$ ne doit pas entraîner $(\psi \wedge \rho) \supset O\varphi$. Autrement dit, l'antécédent d'un conditionnel normatif ne peut pas être augmenté puisque ce n'est pas parce que φ est obligatoire dans un contexte ψ que φ est aussi obligatoire dans un contexte où d'autres conditions ρ sont ajoutées.

Chellas (1974) propose de définir l'opérateur $O(/)$ à l'aide d'un nouvel opérateur (dyadique) $\Rightarrow$ pour les conditionnels. D'abord, Chellas rejette les axiomes (ON) et (OK)[7].

[7] Correspondant aux axiomes (A2) et (A4) du système initial de von Wright.

$$O\top \tag{ON}$$

$$(O\varphi \wedge O\psi) \supset O(\varphi \wedge \psi) \tag{OK}$$

Le rejet de (OK) équivaut au rejet de l'agrégation des obligations, et le rejet de (ON) équivaut à l'acceptation du principe de contingence normative.

Chellas n'adopte que la règle (ROM) et l'axiome (OD), ou leur équivalent dyadique en ce qui à trait à l'obligation conditionnelle :

$$\frac{\varphi \supset \psi}{O\varphi \supset O\psi} \qquad \text{(ROM)}$$

$$\neg O\bot \tag{OD}$$

Remarquons au passage que la lecture de Chellas (1974) peut porter à croire qu'un ensemble d'obligations est fermé sous l'implication, c'est-à-dire que (ROM) peut être appliquée autant à des implications matérielles qu'à des théorèmes. Cela dit, comme il le mentionne dans son introduction aux logiques modales, il est important de noter que les règles dérivées des logiques modales, incluant (ROM), ne peuvent être appliquées qu'à des théorèmes du système. Autrement dit, ce n'est pas l'ensemble des implications matérielles qui est fermé sous (ROM) mais bien l'ensemble des théorèmes, c'est-à-dire des énoncés dont il est possible de faire la preuve sans hypothèse.

Le rejet de (ON) et de (OK) empêche l'équivalence entre $\neg O\bot$ et l'axiome (**D**) du système standard, ce qui empêche l'agrégation et rend ainsi possible les conflits d'obligations. Chellas définit l'obligation conditionnelle par :

$$O(\varphi/\psi) =_{df} \psi \Rightarrow O\varphi \tag{df. O}$$

La clause sémantique de $\Rightarrow$ permet de restreindre la classe des mondes accessibles à une classe unique d'alternatives déontiques. La clause sémantique est que $\psi \Rightarrow \varphi$ est vrai pour w si et seulement si φ est vrai pour tout v appartenant à une classe (unique) de scénarios déterminée par les conditions ψ de w. En définissant l'obligation conditionnelle ainsi et en utilisant les règles (RCK) et (RCEA) pour l'opérateur $\Rightarrow$ (tout cela étant axé sur les règles du calcul propositionnel), Chellas forme le système (assez faible) CK, où seule la version dyadique de (ROM) tient :

$$\frac{(\psi_1 \wedge ... \wedge \psi_n) \supset \psi}{((\varphi \Rightarrow \psi_1) \wedge ... \wedge (\varphi \Rightarrow \psi_n)) \supset (\varphi \Rightarrow \psi)} \qquad \text{(RCK)}$$

$$\frac{\varphi \equiv \psi}{(\varphi \Rightarrow \rho) \equiv (\psi \Rightarrow \rho)} \qquad \text{(RCEA)}$$

Cela dit, considérant que l'opérateur $\Rightarrow$ signifie que la classe des mondes possibles déterminés par une condition est vide seulement si la condition l'est, Chellas ajoute l'axiome (CD) à CK pour créer le système minimal de logique déontique dyadique CD[8] :

$$\Diamond\varphi \supset \neg(\varphi \Rightarrow \bot) \qquad \text{(CD)}$$

Chellas utilise la définition usuelle de $\Diamond$ par $\neg\Box\neg$ et son modèle sémantique vise à rendre compte d'une logique déontique monadique capable de traiter de l'obligation conditionnelle (laquelle se traduit par $\psi \Rightarrow O\varphi$). Le modèle $\mathcal{M} = \langle W, R, f, P, a\rangle$ est construit de la manière suivante. W est l'univers du discours ($W \neq \varnothing$), R est une relation sérielle dans W, f est une fonction qui permet de déterminer la classe (unique) d'alternatives déontiques à w pour certaines conditions. P est un morphisme qui permet de déterminer la classe des scénarios possibles dans lesquels les propositions atomiques $\{p_1, ..., p_n, ...\}$ sont vraies et a est une fonction qui assigne des valeurs de vérité aux propositions dans W.[9] Les clauses sémantiques sont les suivantes, où $Y = \{y : a_y(\psi) = \top\}$ et $Z = \{z : a_z(\varphi) = \top\}$:

$$a_w(\varphi) = \top \Leftrightarrow \varphi \in w \tag{5.9}$$
$$a_w(\varphi \supset \psi) = \top \Leftrightarrow a_w(\varphi) = \bot \text{ ou } a_w(\psi) = \top \tag{5.10}$$
$$a_w(\Box\varphi) = \top \Leftrightarrow a_v(\varphi) = \top \; \forall v \in W \tag{5.11}$$
$$a_w(O\varphi) = \top \Leftrightarrow \exists X \subseteq W \text{ t.q. } \forall v \in X, wRv \text{ et } a_v(\varphi) = \top \tag{5.12}$$
$$a_w(\psi \Rightarrow \varphi) = \top \Leftrightarrow f(w, Y) \subseteq Z \tag{5.13}$$

La clause 5.11 admet un modèle pour l'opérateur $\Box$ équivalent à celui du système modale $S5$, à l'intérieur duquel se trouve une relation universelle. La clause 5.12 fait que l'opérateur O se comporte de la même manière qu'à l'intérieur du système standard KD, où la relation sérielle permet d'isoler une classe d'alternatives idéales à w, et 5.13 signifie que l'ensemble des alternatives idéales à w où ψ est vrai est un sous-ensemble de l'ensemble des mondes possibles où φ est vrai. Autrement dit, la classe des scénarios déterminés par f est telle que ses éléments sont des alternatives idéales à w où les conditions ψ sont satisfaites.

[8] Ou de façon équivalente $(\varphi \Rightarrow \bot) \supset \Box\neg\varphi$.

[9] P sert à déterminer la classe des mondes possibles où certaines propositions sont vraies. $\varphi \in w$ à la clause 5.9 peut donc être exprimé par $w \in P(\varphi)$.

En ce sens, 5.13 permet d'isoler la classe des alternatives déontiques où les conditions ψ sont remplies, laquelle fait partie de l'ensemble des mondes possibles où φ est vrai. Cette sémantique permet à Chellas de traiter de l'obligation conditionnelle dans la mesure où $\psi \Rightarrow O\varphi$ est vrai pour w à condition que $O\varphi$ soit vrai pour toute alternative déontique v où les conditions ψ sont remplies, ce qui implique que φ sera aussi vrai pour ces alternatives. La valeur de vérité de la proposition $O\varphi$ est donc restreinte à la classe des alternatives déontiques où les conditions ψ sont remplies.[10]

* * *

La logique déontique dyadique est apparue afin de rendre compte de l'aspect conditionnel de certaines obligations. Même si, à l'origine, von Wright cherchait à formaliser la notion d'*engagement* en réaction au paradoxe de Prior, il n'en demeure pas moins que le paradoxe de Chisholm est celui qui a eu le plus de répercussions concernant l'émergence du courant dyadique. Or plusieurs défendent que, en plus de mettre en évidence l'aspect conditionnel de certaines obligations, le paradoxe de Chisholm met aussi en lumière son caractère temporel. En effet, les obligations qui surviennent dans des contextes de violation ne le font qu'à un moment t_i où une obligation est enfreinte. De fait, le paradoxe de Chisholm a non seulement donné lieu à la littérature scientifique portant sur l'obligation conditionnelle, mais a aussi laissé place à celle qui traite de la temporalité des propositions déontiques, à laquelle nous allons maintenant porter notre attention.

[10] On trouve dans un article de Bonevac (1983) une analyse du traitement que fait Chellas de l'obligation conditionnelle.

Chapitre 6

La logique temporelle

La logique déontique temporelle est principalement apparue en réaction au paradoxe de Chisholm (1963). Même si les logiques dyadiques permettent de résoudre le paradoxe en bloquant soit le détachement factuel soit le déontique, celles-ci ne parviennent pas à rendre compte du caractère temporel mis en évidence par ce paradoxe. En effet, les obligations qui surgissent dans des contextes de violation peuvent être vues comme implicitement temporelles puisque celles-ci surviennent seulement à partir du moment où un agent n'agit pas conformément à ce qu'il devrait faire.

Étant donné le caractère temporel des obligations conditionnelles, certains en sont venus à formaliser la notion d'obligation dans le cadre des logiques temporelles. On trouve une diversité considérable d'approches temporelles en logique déontique (p. ex. Chellas 1969 ; McKinney 1977 ; Åqvist et Hoepelman 1981 ; Bailhache 1991, 1993), et, par conséquent, le présent chapitre ne prétend pas être exhaustif. Nous nous limiterons à présenter l'approche de van Eck (1982b), qui, comme le mentionne Åqvist (2002) est une des plus riches en ce qui concerne les logiques déontiques temporelles, ce qui sera suivi de la proposition de Thomason (1981).

Job van Eck

D'emblée, la motivation de van Eck (1982a,b) est de fournir un système de logique déontique en mesure d'indiquer ce que les agents doivent, au sens moral, faire dans certaines situations. En effet, son objectif est de développer un langage formel qui aide les agents à déterminer ce qu'ils devraient faire, c'est-à-dire de déterminer leurs obligations actuelles.[1]

[1] Voir Åqvist (2002) pour un résumé des problèmes que van Eck avance concernant la nécessité d'une logique déontique temporelle.

Alors que van Eck considère que les systèmes monadiques ne permettent pas de représenter les obligations qui surgissent dans des contextes de violation, cet auteur rejette les systèmes dyadiques en vertu du fait qu'ils ne parviennent pas à rendre compte de la temporalité implicite à l'obligation.

Dans un premier temps, l'auteur indique que les obligations d'un agent peuvent changer en fonction de ses actions, comme le montre le paradoxe de Chisholm. Alors que les obligations actuelles sont celles du moment *présent*, les obligations *prima facie* admettent un laps de temps entre le moment où elles surviennent et celui où elles sont remplies.

Dans le cas du paradoxe de Chisholm, van Eck voit là une indication du caractère temporel des obligations puisqu'il y a un passage de l'obligation *prima facie Paul doit remettre l'argent volé (si Paul vole)* à l'obligation actuelle *Paul doit remettre l'argent volé* lorsque Paul a effectivement volé. Par ailleurs, les obligations *prima facie* peuvent changer dans certains contextes, notamment lorsque d'autres obligations plus fortes surviennent. De fait, les obligations d'un agent dépendent de certains moments dans le temps, et donc la logique déontique doit être construite sur la base d'une logique temporelle.

Van Eck opte pour une logique déontique du premier ordre axée sur une logique modale temporelle relative puisque le principe kantien *devoir implique pouvoir* requiert une interprétation temporelle de la possibilité. En effet, la notion de *possibilité* sous-jacente à ce principe va au-delà de la simple possibilité logique, puisqu'elle concerne les choix qu'un agent *peut* faire dans certaines situations.

En ce sens, la formule $O\varphi \supset \Diamond\varphi$ n'indique pas seulement que φ n'est pas une contradiction, ce qui peut être représenté par $\neg O\bot$, mais indique aussi que φ peut être accomplie dans un certain contexte. Autrement dit, la formule *devoir implique pouvoir* signifie qu'une action obligatoire *peut* être faite dans un certain contexte, et cela dépend du temps.

Par exemple, si Paul a l'obligation d'aider son prochain à un moment t' différent du présent t, alors le moment t' doit être accessible à Paul à partir du moment t. La possibilité implicite au principe kantien est donc plus large que la simple possibilité logique et inclut une dimension temporelle.

Remarquons qu'une contradiction $p \wedge \neg p$ est une impossibilité logique. L'impossibilité pour Paul d'être en Chine à 3h de l'après-midi le 7 mai 2011 alors qu'il est au Canada cette journée même à 2h58 n'est pas une impossibilité logique, mais plutôt une impossibilité temporelle (voire physique).

Un autre exemple serait l'impossibilité imputable à la (présumée) linéarité du temps. Il serait absurde d'obliger Paul à faire quelque chose dans le passé puisqu'il lui serait *impossible* de le faire, et cette impossibilité est d'un autre ordre que logique (voir Peterson 2013b, chapitre 5 pour une discussion des différentes notions de possibilité).

Le système déontique que van Eck (1982b) propose est une extension du système modal quantifié temporel $QMTL$ (*Quantified Modal Temporal Logic*). L'intuition de base derrière l'approche de van Eck est qu'une situation w_i peut être vue comme une suite temporelle d'événements où à chaque moment se trouve un éventail de possibilités :

$$... \longrightarrow t_{-2} \longrightarrow t_{-1} \longrightarrow t \longrightarrow t_{+1} \longrightarrow ...$$

Ainsi, à partir d'une situation w_i, il est possible de construire plusieurs scénarios différents qui incluent w_i, et donc avec un passé commun. La syntaxe de $QMTL$ contient un ensemble de prédicats à n-places (où $n \geq 2$), des variables et des constantes ontologiques et temporelles, $<$ (une relation d'ordre temporelle), $=$ (une relation d'identité pour les objets du domaines), $\neg$, $\supset$, $\forall$ et $\Box$.[2] L'ensemble des formules bien formées est défini usuellement[3] :

1. $P_z(a_1, ..., a_n) \in EBF$ pour tout prédicat où z est un terme temporel et $a_1, ..., a_n$ sont des termes ontologiques ;
2. $a_1 = a_2, t_1 < t_2 \in EBF$ si a_1, a_2 et t_1, t_2 sont respectivement des termes ontologiques et temporels ;
3. $\neg\varphi, \varphi \supset \psi, (\forall u)\varphi, (\forall^z x)\varphi, \Box_z\varphi \in EBF$ si u est une variable (ontologique ou temporelle), z est un terme temporel et x une variable ontologique.

En ce qui concerne la sémantique, van Eck définit une structure $\mathcal{S} = \langle T, \ll, D, W\rangle$ où T est un ensemble non vide de points dans le temps, $\ll$ une relation d'ordre dans T, D est un ensemble non vide d'objets (le domaine) et $W =< w^1, w^2, w^3, w^4 >$ un ensemble de mondes possibles représentés par les fonctions w^1, w^2, w^3 et w^4, où w^1 assigne des constantes temporelles aux moments dans T, w^2 assigne des constantes ontologiques aux objets du domaine, w^3 assigne à chaque moment un sous-ensemble de l'univers du discours (elle lui assigne un *passé*) et w^4 assigne des objets aux prédicats.

[2] Ce système n'admet que des prédicats binaires ou plus étant donné que le premier terme du prédicat est temporel.

[3] Les définitions pour les autres connecteurs, à savoir $\wedge, \vee, \equiv, \Diamond$ et $\exists$, sont aussi usuelles.

Afin d'obtenir le système de logique déontique $QDTL$ (*Quantified Deontic Temporal Logic*), van Eck ajoute à $QMTL$ l'opérateur déontique $\psi O_t \varphi$, qui équivaut à un opérateur dyadique temporalisé et signifie que dans les conditions ψ au temps t, φ est obligatoire. Conformément à l'ajout de cet opérateur, la clause suivante est ajoutée pour redéfinir l'ensemble des énoncés bien formés :

4. si $\psi, \varphi \in EBF$ et z est un terme temporel, alors $\psi O_z \varphi \in EBF$

Poursuivant dans la tradition dyadique, van Eck définit l'obligation inconditionnelle et la permission de façon habituelle :

$$O_z \varphi =_{df} (\psi \supset \psi) O_z \varphi \qquad \text{(df. O)}$$
$$\psi P_z \varphi =_{df} \neg(\psi O_z \neg \varphi) \qquad \text{(df. P)}$$

Du point de vue de la sémantique, l'auteur ajoute à la structure une fonction $Q(w, t, Y) : W \times T \times \wp(W) \longrightarrow \wp(W)$ qui permet de déterminer les scénarios $v \subset Y$ accessibles possibles à w au temps t. Ces alternatives peuvent être idéales ou sous-optimales.

Cette fonction est caractérisée par le fait que $Q(w, t, Y) = \varnothing$ seulement s'il n'y a aucun scénario accessible à w au temps t, que la meilleure alternative pour w au temps t doit être accessible[4] et par le fait que deux scénarios qui ont le même passé jusqu'à un temps t avaient aussi pour tout temps $t' < t$ le même ensemble d'alternatives déontiques possibles. Finalement, $\psi O_z \varphi$ est vrai pour un scénario w si et seulement si φ est vrai pour tout $u \in Q(w, t, \{v \in W :\models_v \psi\})$, c'est-à-dire pour toute alternative déontique accessible à w au temps t où la proposition φ est vraie et où les conditions ψ sont remplies. De manière équivalente, $O_t \varphi$ est vrai pour w si et seulement si φ est vrai pour toute alternative déontique u accessible à w au temps t (van Eck 1986).

On peut voir un parallèle entre les approches de van Eck, Jones et Pörn et celle de Chellas considérant que les trois formalisent la sémantique de l'obligation conditionnelle en fonction d'un sous-ensemble d'alternatives déontiques accessibles où certaines conditions sont remplies. Cela dit, même si la temporalisation de la logique déontique n'est pas nécessaire à la résolution du paradoxe de Chisholm, van Eck développe un langage qui est extrêmement riche et qui permet d'exprimer avec une précision considérable certaines subtilités dont les systèmes monadiques et dyadiques sont

[4] On voit ici l'effet de $QMTL$ pour $QDTL$. En effet, il s'agit là de l'interprétation temporelle de la nécessité, conformément au principe kantien *devoir implique pouvoir*. La relation d'accessibilité détermine les scénarios qui *peuvent* être accessibles à partir d'un temps t. Un scénario accessible à un temps t doit avoir le même domaine ainsi que le même passé jusqu'au temps t.

souvent incapables. Néanmoins, ce ne sont pas toutes les logiques déontiques temporelles qui reposent sur une logique du premier ordre. Dans une autre perspective, Thomason propose une interprétation temporelle moins riche que celle de van Eck, laquelle repose sur une logique déontique monadique standard.

Richmond Thomason

Selon Thomason (1981), la logique déontique doit rendre compte de l'interaction fondamentale entre le temps et l'obligation. Il considère que l'obligation participe de la temporalité sous deux aspects.

D'une part, une obligation dépend du fait qu'à un certain moment dans le temps, un agent s'engage à des actions particulières. Par exemple, quelqu'un sera dans l'obligation de remplir sa promesse seulement à partir du moment où il promet. En ce sens, la phrase déontique se comporte comme toute phrase dépendant du temps, et donc la valeur de vérité d'une proposition comme $O\varphi$ dépendra du temps.

D'autre part, l'obligation dépend du temps puisque la valeur de vérité des énoncés déontiques dépend d'un ensemble de choix, voire d'actions possibles futures, qui varie en fonction du temps. Afin de soutenir ces deux points et en vue de justifier son approche, Thomason présente deux exemples qui, selon lui, mettent en lumière certaines considérations temporelles relatives à l'obligation.

Le premier exemple, qui a pour objectif de montrer que l'obligation est sujette à la temporalité, consiste à donner une situation où une personne a la possibilité de promettre quelque chose, du moment que cela n'est pas interdit.

Exemple

Supposons que Pierre possède un permis de conduire. Alors, Pierre a la permission de conduire ses enfants au parc (p) et celle de ne pas les conduire au parc (¬p). Autrement dit, Pierre est libre de faire p et il est libre de faire ¬p : $Pp \wedge P\neg p$. Supposons maintenant que Pierre promette à ses enfants de les conduire au parc. À ce moment, Pierre a l'obligation de conduire ses enfants au parc. Dès lors, il est permis que Pierre soit dans l'obligation de conduire ses enfants au parc : POp.

Or l'argument de Thomason est qu'il faut prendre en compte certaines considérations temporelles pour valider ce raisonnement. En effet, l'auteur soutient qu'un tel raisonnement doit plutôt être formalisé à l'aide d'un

opérateur temporel ($\vec{F}$) représentant le futur[5] : il est permis que Pierre soit, à un moment donné, dans l'obligation de conduire ses enfants au parc, c'est-à-dire $P\vec{F}Op$.

Autrement dit, la représentation adéquate du fait que Pierre ait la permission d'être dans l'obligation d'amener ses enfants au parc se fait par un opérateur temporel. Le point que souligne Thomason est que les obligations dépendent du temps, à savoir qu'elles surgissent à certains moments dans le temps. De fait, puisque l'obligation dépend du temps, il s'ensuit qu'il faut la représenter formellement à l'aide d'une logique temporelle.

Le second exemple vise à montrer qu'une obligation dépend d'un ensemble de choix, lequel varie en fonction du temps.

Exemple

*Supposons que Pierre promette à son enfant de lui acheter de la crème glacée d'ici 15h (*p*). Si Pierre tient sa promesse, alors il s'ensuit qu'il devra payer la crème glacée d'ici 15h (*q*). Donc, Pierre a l'obligation de devoir payer la crème glacée d'ici 15h : $Op \wedge OOq$. Maintenant, supposons que Pierre a oublié son portefeuille. Dès lors, il n'aura pas l'obligation de payer pour la crème glacée puisqu'il ne l'achètera pas : $\neg Oq$.*

Cela semble être en contradiction avec OOq. Or la solution de Thomason est d'introduire l'opérateur $\vec{F}$ pour rendre le raisonnement valide. Selon lui, la situation se représente plutôt ainsi :

$$O\vec{F}p \wedge OO\vec{F}q$$
$$\neg O\vec{F}q$$

En somme, les obligations de Pierre varient en fonction du temps puisque, lorsqu'il aura remarqué qu'il ne peut pas payer la crème glacée, il s'ensuivra qu'il n'aura pas l'obligation de la payer étant donné qu'il ne pourra pas le faire. Considérant qu'il est possible que Pierre ne soit finalement pas dans l'obligation de payer la crème glacée, Thomason en conclut qu'il faut introduire l'opérateur $\vec{F}$ pour rendre compte de ce fait. Si Pierre achète la crème glacée, alors il sera dans l'obligation de la payer. Toutefois, si Pierre n'achète pas la crème glacée, alors il est permis qu'il ne la paie pas.

Considérant le caractère temporel intrinsèque à l'obligation, Thomason soutient que la logique déontique doit rendre compte formellement de cette propriété. De fait, il propose un modèle sémantique axé en partie sur la logique temporelle de Prior. D'emblée, il définit une structure temporelle

[5] Thomason utilise F, mais nous avons opté pour $\vec{F}$ pour ne pas confondre avec l'interdiction.

$\langle \mathcal{K}, < \rangle$, où $\mathcal{K}$ est un ensemble non vide et où $<$ est une relation à l'intérieur de $\mathcal{K}$ telle que pour tout $\alpha, \beta, \gamma \in \mathcal{K}$, si $\beta < \alpha$ et $\gamma < \alpha$, alors soit $\beta < \gamma$, $\gamma < \beta$ ou $\beta = \gamma$. Autrement dit, dans une structure temporelle linéaire, si les moments β et γ viennent avant le moment α, alors soit β vient avant γ, soit l'inverse ou il s'agit du même moment.

Il définit ensuit une histoire h, laquelle représente une suite maximale de tous les moments sur une ligne temporelle, et $\mathcal{H}_\alpha$ l'ensemble de toutes les histoires contenant le moment α. Ayant défini la structure temporelle, Thomason est maintenant en mesure d'établir la clause sémantique quant à l'attribution de valeur de vérité aux propositions de la forme $\vec{F}\varphi$. La clause est la suivante : $\vec{F}\varphi$ est vrai à un moment α faisant partie d'une histoire h si et seulement si φ est vrai à un moment β de h tel que $\alpha<\beta$. Formellement, nous avons :

$$a^h_\alpha(\vec{F}\varphi) = \top \;\Leftrightarrow\; a^h_\beta(\varphi) = \top \text{ pour un } \beta \in h \text{ tel que } \alpha < \beta$$

Cela dit, avant de définir la clause sémantique de l'obligation, Thomason introduit des *lacunes* de valeur de vérité[6] conformément à la méthode de van Fraassen :

$$a_\alpha(\varphi) = \top \;\Leftrightarrow\; a^h_\alpha(\varphi) = \top \text{ pour tout } \alpha \in h$$
$$a_\alpha(\varphi) = \bot \;\Leftrightarrow\; a^h_\alpha(\varphi) = \bot \text{ pour tout } \alpha \in h$$
$$\text{Sinon, } a_\alpha(\varphi) \text{ n'a pas de valeur de vérité.}$$

Considérant que Thomason fait la distinction entre les obligations relatives à un contexte délibératif (p. ex. j'ai promis, donc je dois tenir ma promesse) et celles qui surviennent dans des contextes de jugements ou de souhaits (p. ex. je ne peux pas tenir ma promesse, donc je dois dire à x que je ne peux pas tenir ma promesse), l'auteur développe deux clauses sémantiques, chacune visant à rendre compte d'un type d'obligation.

D'une part, la clause sémantique concernant les obligations qui surgissent dans des contextes délibératifs nécessite une structure différente.[7] À partir de la structure temporelle $\langle \mathcal{K}, < \rangle$, Thomason définit une structure $\langle \mathcal{K}, <, \mathcal{L} \rangle$, où $\mathcal{L}$ est une relation entre les moments et les histoires, de telle sorte que si $\alpha \mathcal{L} h$, alors $h \in \mathcal{H}_\alpha$, où $\alpha \mathcal{L} h$ représente une suite d'événements ayant une conséquence morale acceptable. La clause sémantique est que φ

[6] Traduction libre de « *truth-value gaps* ».

[7] Le contexte délibératif est caractérisé par le fait que l'agent raisonne (délibère) par rapport à ses obligations. De fait, il prend en compte les suites d'actions possibles dans l'évaluation de ses obligations. L'obligation qui surgit dans un contexte de jugement est plus faible et dépend des principes que la personne est prête à accepter.

est une obligation à un moment α d'une histoire h si et seulement si φ est vrai au moment α d'une histoire g pour toute histoire g telle que $\alpha \mathcal{L} g$:

$$a^h_\alpha(O\varphi) = \top \Leftrightarrow a^g_\alpha(\varphi) = \top \text{ pour tout } g \text{ tel que } \alpha \mathcal{L} g$$

Bref, il est vrai que φ est obligatoire à un moment α si et seulement si φ est vrai au moment α pour toute suite moralement acceptable d'événements qui inclut ce moment.

Par ailleurs, afin de chercher à rendre compte de l'obligation dans un contexte de jugement, Thomason redéfinit la relation $\mathcal{L}$ afin d'évaluer la valeur de vérité d'une proposition en fonction d'un moment qui détermine le futur possible. Autrement dit, l'auteur soutient qu'il faut évaluer la valeur de vérité d'une obligation à un certain moment en fonction d'un moment précédent qui le détermine :

$${}^\beta a^h_\alpha(O\varphi) \text{ avec } \beta \text{ tel que } \beta < \alpha$$

La relation $\mathcal{L}$ permet, à partir d'un moment α, de trouver un moment différent préférable à α relativement à β. Pour ce faire, $\mathcal{L}$ se décompose en deux fonctions $\mathcal{L}_1$ et $\mathcal{L}_2$, où $\mathcal{L}_1$ est un ensemble d'instants et $\mathcal{L}_2$ un ensemble d'histoires qui incluent γ pour $\gamma \in \mathcal{L}_1$. La clause sémantique devient donc :

$$\begin{gathered} {}^\beta a^h_\alpha(O\varphi) = \top \Leftrightarrow \text{ pour tout } \gamma \in \mathcal{L}_1, \\ {}^\gamma a^g_\gamma(\varphi) = \top \text{ pour tout } g \in \mathcal{L}_2 \end{gathered}$$

L'énoncé $O\varphi$ est vrai pour un moment α appartenant à une histoire h relativement à un instant β si et seulement si pour tout γ appartenant à l'ensemble d'instants $\mathcal{L}_1$, φ est vrai pour le moment γ de l'histoire g pour tout g appartenant à l'ensemble des histoires qui contiennent γ. En termes simples, $O\varphi$ est vrai à un moment α si et seulement si φ est vrai pour toutes les suites d'actions qui incluent ce moment et qui sont moralement acceptables.

Soulignons au passage que ce type de sémantique ne convient pas à la représentation de l'obligation légale. En effet, Thomason soutient qu'une proposition $O\varphi$ est vraie pour un moment α d'une histoire h si et seulement si l'action dénotée par φ est vraie au moment α de toute suite d'événements g moralement acceptable. Autrement dit, il est moralement obligatoire de poser une action moralement acceptable pour toute suite d'événements qui inclut l'action.[8]

[8] Notons aussi que cette clause sémantique est circulaire.

Or la valeur de vérité d'une obligation légale ne dépend en rien de la suite d'événements moralement acceptable (ou non) qui inclut la réalisation de l'action. Il est vrai qu'une personne a l'obligation de ne pas voler, peu importe la suite moralement acceptable d'événements où le vol pourrait avoir lieu. La valeur de vérité d'une proposition déontique ne dépend pas de ce qui est *moralement acceptable*, mais plutôt des normes auxquelles l'agent doit se soumettre et des principes qu'il est prêt à accepter.

De plus, la sémantique de Thomason implique qu'une proposition vraie pour toute suite moralement acceptable d'événements qui inclut un moment α est obligatoire. Or considérant que toute suite d'événements moralement acceptable accessible à α possède le même passé, il s'ensuit que toute proposition qui décrit le passé dans α est vraie pour toute suite d'événements moralement acceptable, et donc que la proposition est obligatoire. En ce sens, en supposant que Paul a commis un vol hier, il s'ensuit qu'il est obligatoire que Paul ait commis un vol hier. Toutefois, il est évidemment faux de soutenir qu'il est légalement obligatoire que Paul ait commis un vol simplement parce qu'il a effectivement volé !

* * *

Somme toute, les approches temporelles en logique déontique offrent un raffinement des modèles sémantiques. La notion d'alternative idéale pouvant parfois sembler contre intuitive, notamment dans les cas où ces alternatives ne sont simplement pas réalisables, les logiques déontiques temporelles redonnent de la force à cette notion dans la mesure où elles restreignent les alternatives à celles accessibles à partir d'une situation donnée. En ce sens, les approches temporelles enlèvent le caractère purement *idéal* aux alternatives et en offrent une conception plus réaliste, à savoir qu'une alternative représente la meilleure série d'actions qui peut être entreprise à partir d'un certain moment.

Cela dit, les logiques déontiques temporelles sont principalement motivées par le paradoxe de Chisholm, sous prétexte qu'il incorpore une dimension temporelle implicite (Decew 1981). Or il s'avère que la temporalité n'est pas une caractéristique nécessaire ou intrinsèque à l'obligation, comme ont montré Prakken et Sergot (1996) à l'aide d'une version du paradoxe de Chisholm qui ne dépend aucunement de la temporalité (Straßer 2011).

La logique déontique temporelle n'offre donc pas de solution définitive au paradoxe de Chisholm. Les logiques non monotones peuvent ainsi être vues comme des alternatives aux cadres de travail dyadique et temporel en vue de la modélisation des obligations conditionnelles.

Chapitre 7

La logique non monotone

La logique déontique non monotone constitue une section importante de la logique déontique. Ses principales applications se trouvent en informatique théorique et en programmation, plus précisément en intelligence artificielle, où on cherche à modéliser le raisonnement et l'apprentissage. La logique non monotone est pertinente en intelligence artificielle considérant qu'elle permet de modéliser les raisonnements qui se produisent à la lumière d'une information spécifique, la conclusion du raisonnement pouvant changer dans l'éventualité où l'information change.

On trouve une importante diversité d'approches en logique déontique non monotone, et le lecteur intéressé par ce cadre de travail est invité à consulter l'ouvrage édité par Nute (1997) en guise d'introduction. Considérant la diversité des approches et des méthodes, le présent chapitre ne prétend pas couvrir l'ensemble des systèmes non monotones, mais vise simplement à introduire le lecteur et à le familiariser avec ce type d'approche.

Il sera d'abord question des principales motivations en faveur des logiques déontiques non monotones, où les notions de monotonie et de non monotonie seront expliquées. Par la suite, l'approche de Horty (1997) est présentée à titre de cas paradigmatique. Ce choix a été réalisé en fonction du fait que l'approche de Horty est très claire et est moins axée vers ses applications en informatique, ce qui rend le texte plus abordable pour le lecteur qui provient des sciences humaines plutôt que de l'informatique théorique. Finalement, différentes notions de défaisabilité sont présentées à la dernière section.

Motivations

La logique non monotone vise la formalisation des inférences quotidiennes (ou ordinaires), qui se font à la lumière d'une information plus souvent qu'autrement incomplète. Ce type d'approche a une incidence directe en intelligence artificielle, où l'objectif est de programmer des machines capables

de raisonner à l'instar de l'homme. Ainsi, la logique non monotone permet la représentation formelle des inférences où les conclusions sont adaptées en fonction de l'information qu'on obtient. Dans un certain sens, la logique non monotone offre à une machine la capacité de « changer d'idée ».

En termes d'*input* et d'*output*, la logique non monotone permet de faire varier l'*output* en fonction de l'*input*, ce qui n'est pas possible dans les cadres de travail monotones. Outre ses applications en informatique et en intelligence artificielle, la logique déontique non monotone vise aussi la représentation des inférences légales, où certaines conclusions sont parfois renversées, notamment lorsqu'elles sont prises à partir d'une information incomplète.

Grossièrement, l'idée de base derrière les cadres de travail non monotones est de fournir des systèmes qui permettent de modifier les conclusions auxquelles on arrive en fonction de l'information qu'on possède. Une relation de conséquence est dite *monotone* lorsque l'ajout de certaines prémisses à une inférence valide ne change pas la conclusion. Autrement dit, une relation de conséquence monotone respecte la règle structurale d'affaiblissement (en anglais *weakening*) (Gentzen 1934) :

$$\frac{\varphi \vdash \psi}{\varphi, \rho \vdash \psi} \quad (\mathsf{wk})$$

Si ψ est une conséquence de φ et que la relation de conséquence est monotone, alors ψ sera aussi la conséquence d'un ensemble de prémisses qui contient φ. Une relation de conséquence est donc monotone lorsque celle-ci est stable et produit toujours le même résultat à partir d'un ensemble de prémisses. Lorsque la relation de conséquence est monotone, l'ajout d'autres prémisses ne change pas la conclusion.

Certains raisonnements quotidiens ne respectent cependant pas ce schéma d'inférence. En effet, certaines inférences peuvent être *défaites*, voire *déconstruites*. Dans certains cas, l'ajout d'information change la conclusion à laquelle on parvient. Cela peut être mis en évidence à l'aide des deux exemples suivants.

Exemples

1. *Considérons les énoncés suivants :*

$p =$ *Il pleut.*
$q =$ *Paul va au cinéma.*
$r =$ *Marie appelle Paul.*
$s =$ *Paul va chez Marie.*

Supposons que Paul planifie sa journée et qu'il en arrive à la conclusion que « s'il pleut, alors il va au cinéma, mais si Marie appelle, alors il ira chez elle ». Supposons maintenant que cette information soit transmise à Pierre, et qu'en plus, on lui dise que cette journée-là, il pleuvait. On demandera à Pierre ce qu'a fait Paul. Son raisonnement sera alors :

$$p \supset q, r \supset s, p \vdash q$$

Pierre conclura que Paul est allé au cinéma. Cela dit, si, par la suite, on apprend à Pierre que Marie a appelé Paul, alors son raisonnement sera :

$$p \supset q, r \supset s, p, r \vdash s$$

Pierre conclura plutôt que Paul est allé chez Marie. De plus, si on lui demande si Paul est allé au cinéma, Pierre répondra que non puisque Paul est allé chez Marie. De fait, l'ajout de l'information r à l'ensemble de prémisses $\{p \supset q, r \supset s, p\}$ fait changer la conclusion. En ce sens, l'inférence de Pierre n'est pas monotone considérant que la règle (wk) n'est pas respectée.

$$p \supset q, r \supset s, p \vdash q$$
$$p \supset q, r \supset s, p, r \nvdash q$$

2. *Soit les énoncés suivants* [1] *:*

   $p =$ *Paul est québécois.*
   $q =$ *La langue maternelle de Paul est le français.*
   $r =$ *Paul est né en France.*

   Supposons qu'on transmet l'information suivante à Pierre :

   (a) La langue maternelle de la plupart des Québécois est le français.
   (b) La plupart des Français sont nés en France.

   Pierre acceptera donc comme prémisses que $p \supset q$ et que $q \supset r$. Toutefois, ce n'est pas parce que Pierre accepte ces deux prémisses qu'il acceptera aussi nécessairement que $p \supset r$. Bien que, de manière générale, si une personne est québécoise, alors sa langue maternelle est le français, et que, dans la plupart des cas, si la langue maternelle d'un individu est le français, alors celui-ci est né en France, cela n'implique pas que si un individu est québécois, alors celui-ci est né en France.

[1] L'exemple est adapté d'Antonelli (2012).

Le second exemple vise à montrer que certaines inférences dans la langue naturelle ne respectent pas la règle de coupure (cut), qui représente la transitivité de la relation de conséquence :

$$\frac{\varphi \vdash \psi \qquad \psi \vdash \rho}{\varphi \vdash \rho} \quad (\mathsf{cut})$$

Cette règle indique que lorsque ψ est une conséquence de φ et que ρ en est une de ψ, alors, nécessairement, il en résulte que ρ est une conséquence de φ.[2] Cependant, certaines inférences quotidiennes ne respectent pas cette règle. Soulignons cependant que ce type d'argument n'est pas présenté en faveur de la non monotonie des inférences normatives.

Dans le cas de la logique déontique, l'introduction d'un cadre de travail non monotone peut être justifiée de deux manières. Alors que certains utilisent la logique déontique non monotone pour représenter les conflits d'obligations, comme Ryu et Lee (1997) et Horty (1997), d'autres l'utilisent plutôt pour représenter formellement les obligations conditionnelles, par exemple van der Torre et Tan (1997).

Prenons le système standard à titre d'exemple. Considérant le schéma d'axiome (**D**), propositionnellement équivalent à 7.1, il en résulte que dans une situation où il y a un conflit d'obligations, alors n'importe quelle action est obligatoire, puisque KD satisfait 7.2. Il s'agit là de l'explosion déontique :

$$\neg(O\varphi \wedge O\neg\varphi) \tag{7.1}$$
$$(O\varphi \wedge O\neg\varphi) \supset O\psi \tag{7.2}$$

Exemple
Si Paul a le devoir d'aider sa patrie en allant à la guerre, mais qu'il a aussi le devoir moral de rester à la maison pour aider sa mère paraplégique, alors Paul est dans une situation où il est dans l'obligation d'aller à la guerre et de ne pas aller à la guerre. Au sein de KD, cela entraîne que Paul est dans l'obligation de voler une banque, ce qui, évidemment, n'est pas un résultat désiré.

Alors que le premier argument en faveur de la non monotonie des inférences normatives concerne la modélisation des conflits d'obligations, le second consiste à dire que les inférences normatives ne satisfont pas le schéma

[2] Il est à noter que cette version de la règle de coupure est différente de la règle usuelle qu'on retrouve en calcul des séquents.

d'inférence (wk). Ayant maintenant vu les principales motivations pour l'introduction de la logique déontique non monotone, passons à l'approche de Horty (1997), qui nous servira à titre de paradigme afin d'exemplifier la stratégie derrière les cadres de travail non monotones.

John Horty

L'introduction d'une relation de conséquence non monotone vise la représentation des inférences *ordinaires* ou *quotidiennes*, où les conclusions auxquelles on parvient varient en fonction de l'information que nous avons. En plus de viser à représenter comment les individus réfléchissent au quotidien, ce type de relation de conséquence est aussi appliqué dans le domaine de l'intelligence artificielle, où l'objectif est de faire le tri dans l'information fournie afin d'obtenir les conclusions désirées.

Ainsi, on s'assure que les machines soient en mesure d'atteindre certains de leurs buts même lorsque que ceux-ci sont contradictoires. Un des principaux objectifs qui a motivé l'introduction des logiques non monotones était de développer une logique capable de rendre compte des inférences humaines pour servir à des fins de programmation en intelligence artificielle.

En logique déontique, les logiques non monotones sont pertinentes lorsqu'on cherche à formaliser les conflits d'obligations ou les obligations conditionnelles. Le texte de Horty (1997) offre une introduction quant aux motivations qui guident les développements de la logique déontique non monotone (voir aussi Horty 1994). D'emblée, cet auteur considère la logique non monotone pertinente lorsqu'on cherche à formaliser les conflits d'obligations. Un conflit d'obligations survient lorsque deux normes dictent des actions incompatibles, c'est-à-dire qui ne peuvent être réalisées simultanément.

Afin de mettre en évidence l'importance des fondements non monotones pour une logique déontique, Horty insiste sur le fait que les conflits d'obligations ne sont pas possibles au sein du système standard. En vertu de l'axiome (**D**), il est impossible d'avoir une situation telle que $O\varphi \wedge O\neg\varphi$ est vrai. Par conséquent, le système standard ne permet pas de rendre compte des conflits d'obligations. Considérant que les conflits d'obligations sont courants, Horty en conclut qu'une logique déontique digne de ce nom doit rendre compte de telles situations.

Comme l'auteur le mentionne, on trouve déjà des approches monotones au sein de la littérature scientifique qui permettent de rendre compte des conflits d'obligations. Chellas (1974), par exemple, a proposé un système qui permet la représentation des conflits d'obligations. La proposition de Chellas est de modéliser les obligations conditionnelles à l'aide d'une logique modale non normale plus faible que le système K, construite comme une

extension de la logique propositionnelle à laquelle on ajoute la règle d'inférence (ROM) et l'axiome $\neg O\bot$. Les conflits d'obligations sont possibles au sein d'un tel système dans la mesure où, contrairement au système standard, il n'y a pas d'équivalence entre (**D**) et $\neg O\bot$, ce qui résulte du fait que l'agrégation $(O\varphi \wedge O\psi) \supset O(\varphi \wedge \psi)$ n'est pas dérivable.

Or Horty soutient que le rejet complet de l'agrégation rend le système de Chellas *trop* faible. En effet, bien qu'il soit souhaitable de rejeter $(O\varphi \wedge O\neg\varphi) \supset O(\varphi \wedge \neg\varphi)$, sans quoi nous obtiendrions $(O\varphi \wedge O\neg\varphi) \supset O\psi$ puisque $O(\varphi \wedge \neg\varphi) \supset O\psi$ en vertu de (ROM) et du fait que $(\varphi \wedge \neg\varphi) \supset \psi$, l'agrégation est parfois désirable, notamment lorsqu'il n'y a pas de conflit.

Afin de mettre ce point en évidence, il propose l'argument suivant, qui, pour être validé dans une approche comme celle de Chellas, doit utiliser une instance du principe d'agrégation.

Exemple

Supposons qu'un agent est dans l'obligation de payer ses taxes en un ou deux versements. Supposons maintenant que cet agent est dans l'obligation de ne pas payer ses taxes en un seul paiement en raison de contraintes financières familiales.

Dans une telle situation, nous voulons conclure que l'agent est dans l'obligation de payer ses taxes en deux versements. Cependant, afin que ce raisonnement soit valide dans le système de Chellas, il faudrait être en mesure de passer de $O(p \vee q) \wedge O\neg p$ à $O((p \vee q) \wedge \neg p)$ à l'aide de l'agrégation pour ensuite utiliser (ROM) et le fait que $((p \vee q) \wedge \neg p) \supset q$ pour conclure que Oq. Cependant, l'agrégation n'est pas disponible au sein du système de Chellas. Par conséquent, ce système n'est pas en mesure de valider l'argument susmentionné.

La conclusion à laquelle parvient Horty est que le principe d'agrégation ne doit pas permettre de conclure les conflits d'obligations, mais qu'il doit néanmoins être possible d'utiliser ce principe lorsqu'il n'engendre pas de conflit. L'auteur propose une solution inspirée de la définition de la conséquence déontique chez van Fraassen (1973). Soit Γ un ensemble de propositions qui contient des énoncés du type $O\varphi$ et $\mathcal{M}$ un modèle de la logique classique. Le *pointage* d'un modèle $\mathcal{M}$ relativement à un ensemble Γ est défini par :

$$score_\Gamma(\mathcal{M}) = \{O\varphi \in \Gamma : \models_{\mathcal{M}} \varphi\}$$

Autrement dit, l'ensemble $score_\Gamma(\mathcal{M})$ contient les propositions $O\varphi$ membres de Γ pour lesquelles φ est vrai dans $\mathcal{M}$.

Exemple
Prenons l'ensemble Γ *:*

$$\Gamma = \{Op, O\neg q, O(p \supset q)\}$$

Alors il y a :

$$\begin{aligned}
\mathcal{M}_1 &= \{p\} \\
\mathcal{M}_2 &= \{p, \neg q\} \\
\mathcal{M}_3 &= \{p, p \supset q\} \\
\mathcal{M}_4 &= \{\neg q\} \\
\mathcal{M}_5 &= \{\neg q, p \supset q\} \\
\mathcal{M}_6 &= \{p \supset q\}
\end{aligned}$$

Et :

$$\begin{aligned}
score_\Gamma(\mathcal{M}_1) &= \{Op\} \\
score_\Gamma(\mathcal{M}_2) &= \{Op, O\neg q\} \\
score_\Gamma(\mathcal{M}_3) &= \{Op, O(p \supset q)\} \\
score_\Gamma(\mathcal{M}_4) &= \{O\neg q\} \\
score_\Gamma(\mathcal{M}_5) &= \{O\neg q, O(p \supset q)\} \\
score_\Gamma(\mathcal{M}_6) &= \{O(p \supset q)\}
\end{aligned}$$

Notons cependant qu'aucun modèle $\mathcal{M}$ ne satisfait à la fois p, $\neg q$ et $p \supset q$.

Maintenant, soit $|\varphi| = \{\mathcal{M} : \models_{\mathcal{M}} \varphi\}$ l'ensemble qui contient les modèles pour lesquels φ est vrai.

Exemple

$$|p| = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3\}$$

La conséquence déontique $\vdash_F$ est définie par :

$$\begin{aligned}
\Gamma \vdash_F O\varphi \Leftrightarrow \text{il y a } \mathcal{M}_1 \in |\varphi| \text{ pour lequel il n'y a pas de} \\
\mathcal{M}_2 \in |\neg\varphi| \text{ t.q. } score_\Gamma(\mathcal{M}_1) \subseteq score_\Gamma(\mathcal{M}_2)
\end{aligned}$$

Conformément à l'exemple précédent, nous avons :

$$\begin{aligned}
score_\Gamma(\mathcal{M}_1) &\subseteq score_\Gamma(\mathcal{M}_2) \\
score_\Gamma(\mathcal{M}_1) &\subseteq score_\Gamma(\mathcal{M}_3)
\end{aligned}$$

$$score_\Gamma(\mathcal{M}_4) \subseteq score_\Gamma(\mathcal{M}_2)$$
$$score_\Gamma(\mathcal{M}_4) \subseteq score_\Gamma(\mathcal{M}_5)$$
$$score_\Gamma(\mathcal{M}_6) \subseteq score_\Gamma(\mathcal{M}_3)$$
$$score_\Gamma(\mathcal{M}_6) \subseteq score_\Gamma(\mathcal{M}_5)$$

Et :

$$|q| = \{\mathcal{M}_3\}$$
$$|\neg q| = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_4, \mathcal{M}_5, \mathcal{M}_6\}$$

Selon cet exemple, on voit aisément que $\Gamma \vdash_F O\neg q$ puisque $\mathcal{M}_2 \in |\neg q|$ et $score_\Gamma(\mathcal{M}_2) \not\subseteq score_\Gamma(\mathcal{M}_3)$. De même, nous pouvons conclure que $\Gamma \vdash_F Oq$ puisque nous savons qu'il n'y a pas de modèle dans $|\neg q|$ tel que $score_\Gamma(\mathcal{M}_3) \subseteq score_\Gamma(\mathcal{M}_i)$ pour $i \in \{1, 2, 4, 5, 6\}$. Cependant, nous avons $\Gamma \nvdash_F O(q \wedge \neg q)$ puisque $|q \wedge \neg q| = \varnothing$ et donc il n'y a pas de $\mathcal{M} \in |q \wedge \neg q|$ qui satisfait la condition.

L'objectif de Horty est de représenter ces intuitions dans le cadre d'une logique non monotone. Pour ce faire, il utilise la logique des conditions de base (*default logic*) de Reiter (1980). Ici, l'idée est d'indexer une inférence de φ vers ψ à une condition ρ. Autrement dit, plutôt que de représenter une inférence comme une paire (φ, ψ), où φ est une prémisse et ψ une conclusion, l'inférence est représentée par un triplet $(\varphi, \psi|\rho)$ où la conclusion ψ ne peut être inférée à partir de φ que lorsque cela est consistant avec ρ.

Une *théorie de base* (*a default theory*) $\Delta = \langle \mathcal{W}, \mathcal{D} \rangle$ est définie comme un ensemble de propositions $\mathcal{W}$ auquel on ajoute un ensemble de règles $\mathcal{D}$ qui guident les inférences. Ayant une théorie de base à notre disposition, l'objectif est de définir ce qu'est une extension $\mathcal{E}$ pour une théorie Δ. Une extension $\mathcal{E}$ pour une théorie $\Delta = \langle \mathcal{W}, \mathcal{D} \rangle$ est définie par les trois conditions suivantes :

$$\mathcal{W} \subseteq \mathcal{E} \tag{7.3}$$
$$\mathcal{E} = Cn(\mathcal{E}) \tag{7.4}$$
$$\text{pour tout } (\varphi, \psi|\rho) \in \mathcal{D},\ \varphi \in \mathcal{E} \text{ et } \neg\rho \notin \mathcal{E} \Rightarrow \psi \in \mathcal{E} \tag{7.5}$$

Alors que la première condition stipule que $\mathcal{E}$ est une extension de $\mathcal{W}$, c'est-à-dire que toute formule de $\mathcal{W}$ est contenue dans $\mathcal{E}$, la seconde indique que l'extension contient l'ensemble de ses conséquences logiques. La troisième condition assure que l'extension contient les conclusions qui peuvent être dérivées à partir des règles de base (*default rules*), et donc pour chaque règle de la forme *ψ est dérivable à partir de φ à condition que cela est consistant avec ρ*, on conclut que ψ est dans l'extension lorsque φ s'y trouve et que rien ne permet de contredire ρ.

Cela fait, Horty en vient à tracer le parallèle entre la relation de conséquence $\vdash_F$ et le cadre de travail non monotone. En définissant la théorie de base Δ_Γ pour un ensemble d'obligations Γ par $\mathcal{W} = \varnothing$ et l'ensemble de règles de base par (df. $\mathcal{D}$), Horty obtient que $\Gamma \vdash_F O\varphi$ si et seulement si $\varphi \in \mathcal{E}$ pour une extension $\mathcal{E}$ de Δ_Γ :

$$\mathcal{D} = \{(\top, \varphi|\varphi) : O\varphi \in \Gamma\} \qquad \text{(df. } \mathcal{D})$$

En mots, ce résultat signifie qu'une obligation $O\varphi$ peut être inférée d'un ensemble d'obligations lorsque φ est consistante avec les autres propositions au sein d'une extension de la théorie de base. Afin d'illustrer cela, considérons l'exemple susmentionné.

Exemple
L'ensemble de règles de base de Δ_Γ *est :*

$$\mathcal{D} = \{(\top, p|p), (\top, \neg q|\neg q), (\top, p \supset q|p \supset q)\}$$

Cette théorie possède plusieurs extensions, selon l'ordre dans lequel elles sont construites.

Par exemple, si on prend d'abord $(\top, p|p)$*, alors nous avons* $\top \in \mathcal{E}_1$ *puisque* $\mathcal{W} = \varnothing$ *et* $\top \in Cn(\varnothing)$*, et puisque* $\neg p \notin \mathcal{E}_1$ *nous obtenons* $p \in \mathcal{E}$*.*

Ensuite, il y a deux possibilités. Si on considère d'abord $(\top, p \supset q|p \supset q)$*, alors on obtient* $p \supset q \in \mathcal{E}_1$ *par le même raisonnement et puisque* $\mathcal{E}_1 = Cn(\varnothing \cup \{p, p \supset q\})$*, on obtient que* $q \in \mathcal{E}_1$*, voire que* $\neg\neg q \in \mathcal{E}_1$*. Dans cette optique, la condition (3) ne permet pas de conclure que* $\neg q \in \mathcal{E}_1$ *étant donné que* $\neg\neg q \in \mathcal{E}$ *(pour que ce soit le cas, il faudrait que* $\neg\neg q \notin \mathcal{E}_1$*).*

Par ailleurs, si nous avions plutôt opté pour l'ajout de $(\top, \neg q|\neg q)$*, alors nous aurions obtenu que* $\neg q \in \mathcal{E}_2$ *puisque* $\top \in \mathcal{E}_2$ *et* $\neg\neg q \notin \mathcal{E}_2$*, et par conséquent* $\neg(p \supset q) \in \mathcal{E}_2$ *puisque* $\neg(p \supset q) \in Cn(\varnothing \cup \{p, \neg q\})$*.*

Finalement, si on débute par considérer $(\top, p \supset q|p \supset q)$ *et ensuite* $(\top, \neg q|\neg q)$*, on obtient une extension* $\mathcal{E}_3 = Cn(\varnothing \cup \{p \supset q, \neg q\})$ *pour laquelle* $\neg p \in \mathcal{E}_3$*.*

En ce sens, différentes extensions pour une même théorie peuvent être obtenues en fonction de l'ordre dans lequel les propositions sont considérées :

$$\mathcal{E}_1 = Cn(\varnothing \cup \{p, p \supset q\})$$
$$\mathcal{E}_2 = Cn(\varnothing \cup \{p, \neg q\})$$
$$\mathcal{E}_3 = Cn(\varnothing \cup \{p \supset q, \neg q\})$$

Certaines théories peuvent n'avoir aucune, ou avoir une ou plusieurs extensions. Lorsqu'il y a plusieurs extensions disponibles, les deux stratégies possibles sont la stratégie *crédule*, qui consiste à prendre une extension au hasard et à accepter comme conclusion ce qui se trouve dans cette extension, ou la *sceptique*, qui consiste à n'accepter comme conclusion que ce qui fait partie de chaque extension.

Le fait qu'une théorie de base peut posséder plusieurs extensions amène Horty à définir une relation de conséquence déontique *sceptique* $\vdash_S$, pour laquelle une obligation $O\varphi$ peut être dérivée d'un ensemble d'obligations Γ seulement lorsque φ appartient à toutes les extensions $\mathcal{E}$ de Δ_Γ. Autrement dit, la relation de conséquence est définie par $\Gamma \vdash_S O\varphi$ si et seulement si $\varphi \in \mathcal{E}$ pour tout $\mathcal{E}$ extension de Δ_Γ. Selon cette définition et conformément à l'exemple susmentionné, cela entraîne que :

$$\Gamma \vdash_F Op$$
$$\Gamma \vdash_F O\neg q$$
$$\Gamma \vdash_F O(p \supset q)$$

Cependant, nous avons les conséquences suivantes puisque $\mathcal{E}_1 \cap (\mathcal{E}_2 \cap \mathcal{E}_3) = Cn(\varnothing)$:

$$\Gamma \nvdash_S Op$$
$$\Gamma \nvdash_S O\neg q$$
$$\Gamma \nvdash_S O(p \supset q)$$

Cela dit, si nous avions plutôt $\Gamma' = \Gamma \cup \{Or\}$, alors nous aurions les extensions suivantes, et donc $\Gamma' \vdash_S Or$:

$$\mathcal{E}'_1 = Cn(\varnothing \cup \{p, p \supset q, r\})$$
$$\mathcal{E}'_2 = Cn(\varnothing \cup \{p, \neg q, r\})$$
$$\mathcal{E}'_3 = Cn(\varnothing \cup \{p \supset q, \neg q, r\})$$

Outre la relation de conséquence déontique sceptique, Horty discute des relations entre le système de Chellas, KD et la relation de conséquence $\vdash_F$. Par la suite, cet auteur aborde aussi la question des obligations conditionnelles, ce qui ne sera pas présenté considérant que les développements qu'il propose ne sont que préliminaires et sujets à plusieurs problèmes.

Pour conclure, soulignons que même si les approches non monotones peuvent sembler souhaitables, celles-ci laissent néanmoins place à des résultats indésirables. Dans le cas de Horty, par exemple, autant $\vdash_F$ que $\vdash_S$ nous laissent dans une impasse.

D'un côté, l'interprétation de $\vdash_F$ nous laisse le choix entre les différentes extensions qui s'offrent à nous. Ainsi, conformément à l'exemple susmentionné, on peut autant choisir entre $\Gamma \vdash_F Op$ et $\Gamma \vdash_F O(p \supset q)$ que $\Gamma \vdash_F O\neg q$. Cependant, certaines obligations au sein d'un conflit sont prioritaires à d'autres, et, par conséquent, il est plausible que, plutôt que d'avoir un choix à faire entre deux obligations, l'une domine l'autre.

En ce sens, si l'objectif de la logique déontique non monotone est de permettre de déterminer les obligations en place lors d'un conflit, la stratégie crédule ne permet pas d'atteindre cet objectif puisqu'elle laisse le choix à l'agent. Dans le même ordre d'idées, la relation de conséquence déontique sceptique élimine simplement toutes les obligations liées au conflit et ne garde que celles qui n'y sont pas rattachées. Dans l'exemple précédent, nous n'avions ni $\Gamma \vdash_S Op$, ni $\Gamma \vdash_S O\neg q$, ni $\Gamma \vdash_S O(p \supset q)$.

Or l'intérêt de pouvoir modéliser les conflits d'obligations est d'être en mesure de déterminer les obligations actuelles d'un agent. De manière générale, un conflit d'obligations ne se résoud pas simplement en se débarrassant des obligations conflictuelles. Lorsqu'un conflit d'obligations surgit, il est plutôt plausible que l'une des obligations domine l'autre. Certains, notamment Royakkers et Dignum (1997) et van der Torre et Tan (1997), répondent à cette intuition en introduisant une relation de préférence (de hiérarchie) entre les obligations. Ce type de solution est inspiré des travaux de Alchourrón et Makinson (1981).

Il n'en demeure pas moins qu'il est possible d'avoir une situation où les deux obligations sont incommensurables, par exemple demander à un parent qui il préfère sauver entre ses deux enfants. Dans une telle situation, même si on accepte une hiérarchie entre les obligations, la relation de conséquence ne permettra pas de conclure ce qui devrait être fait.

En ce sens, même si l'approche de Horty permet de répondre à certaines objections qui peuvent être soulevées contre la logique déontique standard, elle ne parvient pas à modéliser adéquatement les conflits d'obligations. Lorsqu'un tel conflit prend place, la solution ne consiste pas simplement à *tout* rejeter, mais plutôt, lorsque possible, à sélectionner l'obligation qui devrait être actuelle.

Différents types de défaisabilité

Ces considérations nous amènent aux différentes manières de défaire un argument.[3] Dans un article sur le traitement de l'obligation conditionnelle, van der Torre et Tan (1997) font la distinction entre trois types de défaisabilité : la *défaisabilité factuelle* (*factual defeasability*), la *défaisabilité de priorité forte* (*strong overridden defeasability*) et la *défaisabilité de priorité faible* (*weak overridden defeasability*).

La défaisabilité factuelle survient lorsqu'un fait contredit une règle de base. Prenons le fameux exemple de Tweety, qui est un pingouin.

Exemple

1.	*Les oiseaux volent.*
2.	*Tweety est un oiseau.*
∴	*Tweety vole.*

Dans ce raisonnement, la règle de base « les oiseaux volent » peut être défaite par l'ajout de l'information « Tweety ne vole pas ». Il s'agit donc d'un cas de défaisabilité factuelle dans la mesure où la règle de base est rejetée en vertu d'un fait.

La défaisabilité de priorité quant à elle survient lorsqu'une règle annule une autre. Dans le cas de Kant et du nazi, par exemple, la règle « il faut toujours dire la vérité » peut être surpassée par la règle « il faut mentir si cela permet de sauver la vie de quelqu'un ».

Exemple

1.	*Il faut toujours dire la vérité.*
∴	*Paul doit avouer au nazi qu'il cache une famille juive dans son grenier.*

En ce sens, le raisonnement précédent peut être défait en accordant priorité à la règle « il faut mentir si cela permet de sauver la vie de quelqu'un ». La défaisabilité de priorité est pertinente dans les cas de conflits d'obligations où l'une doit être retenue en faveur de l'autre.[4]

[3] Faute d'une meilleure traduction, nous allons utiliser les termes défaire, défaisable ainsi que défaisabilité pour référer aux concepts de *defeasible reasoning* et de *defeasibility* en logique non monotone.

[4] Pour une discussion sur les différents types de défaisabilité, voir Peterson (2015b).

La différence entre la défaisabilité de priorité *faible* et *forte* peut s'apercevoir à l'aide de la distinction entre l'*annulation* d'une obligation et son *surpassement contextuel* (*overshadowing*). Alors que la défaisabilité de priorité forte survient lorsqu'il y a nécessairement annulation (rejet), la défaisabilité de priorité est faible lorsqu'il y a possibilité de surpassement contextuel.

Dans l'exemple susmentionné, l'obligation de dire la vérité n'est pas complètement rejetée, mais est seulement mise de côté en faveur d'une autre obligation qui a préséance sur celle-ci. En ce sens, il s'agit d'un cas de défaisabilité de priorité faible puisque le surpassement de l'obligation « il faut mentir pour sauver la vie de quelqu'un » sur « il faut dire la vérité » dépend du contexte dans lequel Paul se trouve.

Un cas de défaisabilité de priorité forte survient lorsqu'une règle est abandonnée en faveur d'une autre. Cela peut se produire lorsque, dans un certain contexte, une obligation n'est simplement plus applicable.

La conclusion à laquelle parviennent les auteurs est que chaque type de défaisabilité s'applique à une situation précise. Alors que la défaisabilité factuelle permet de représenter les cas où certaines obligations sont violées, la défaisabilité de priorité forte permet de représenter les cas où certaines obligations sont plus spécifiques et la défaisabilité faible représente les obligations *prima facie*.

Selon van der Torre et Tan, les obligations contraires mettent en jeu des violations et doivent être modélisées à l'aide de la défaisabilité factuelle. Par exemple, dans le cas du paradoxe de Chisholm, nous avons :

$$O\neg p$$
$$O(\neg p \supset \neg q)$$
$$p \supset Oq$$
$$p$$

Or, dans le système standard, cela permet de dériver $O\neg q$ et Oq. Dans cette situation, la défaisabilité de l'argument en faveur de Oq vient du *fait* que l'obligation $O\neg p$ a été violée. Ainsi, la défaisabilité du paradoxe de Chisholm en faveur de Oq vient du fait que p.

Les conflits d'obligations qui ne découlent pas d'obligations survenant dans des contextes de violations sont mieux représentés par la défaisabilité de priorité. La défaisabilité de priorité faible représente les conflits d'obligations entre des obligations *actuelles* (toutes choses considérées) et des obligations *prima facie* dans la mesure où, même si

les obligations actuelles ont priorité dans un contexte donnée, les obligations *prima facie* demeurent et s'appliquent. Dans l'exemple du nazi, l'obligation de ne pas mentir est toujours présente, même si on opte plutôt pour mentir afin de sauver la famille juive.

Finalement, la défaisabilité de priorité forte s'applique aux cas où les conflits d'obligations proviennent d'un conflit de normes. En effet, ce type de défaisabilité permet de représenter les cas où certaines règles sont rejetées en faveur d'autres plus spécifiques. Afin de mettre ce point en évidence, van der Torre et Tan utilisent l'exemple de la clôture, tel qu'il est exposé dans Prakken et Sergot (1996).

Exemple
Imaginons un règlement municipal :

1) *il ne doit pas y avoir de clôture autour de la villa ;*
2) *si la villa est au bord de la mer, alors il peut y avoir une clôture autour de la villa.*

En supposant que la villa se trouve au bord de la mer, cela est inconsistant au sein du système standard dans la mesure où nous obtenons $O\neg c$ *et* $\neg O\neg c$ *en vertu de :*

$$O\neg c \tag{7.6}$$
$$v \supset \neg O\neg c \tag{7.7}$$
$$v \tag{7.8}$$

Ce raisonnement peut cependant être défait par une priorité forte dans la mesure où la règle (7.7) a priorité sur la règle (7.6) puisqu'elle est plus spécifique. Dans une telle situation, la règle (7.7) *annule* la règle (7.6), c'est-à-dire que, contrairement à la défaisabilité de priorité faible, l'obligation qu'il n'y ait pas de clôture ne vaut plus.

Pour un résumé des exemples utilisés en logique non monotone et pour modéliser les obligations conditionnelles, le lecteur est invité à consulter van der Torre et Tan (1999).

Lou Goble

Goble (2009) a proposé une analyse approfondie des principes en jeu lors des conflits d'obligations. D'emblée, son analyse repose sur la base de deux arguments :

Argument 1. L'agrégation d'obligations, en conjonction avec le principe *devoir implique pouvoir*, mène à l'impossiblité des conflits normatifs.

Argument 2. Le principe de distribution « si, nécessairement, α implique β, alors $O\alpha$ implique $O\beta$ », en conjonction avec le schéma d'axiome pour la consistence normative, mène à l'impossiblité des conflits normatifs.

Le premier argument est attribué à Lemmon (1962, 1965), pour lequel le principe *devoir implique pouvoir*, en conjonction avec l'agrégation et le schéma d'axiome pour la consistence normative, mènent à l'impossiblité de représenter adéquatement les conflits normatifs. Ces arguments reposent sur les principes suivants :

(Agg) si α est obligatoire et β est obligatoire, alors leur conjonction est obligatoire ;
(Can) si α est obligatoire, alors il est possible d'accomplir α ;
(Dist) si α implique β, alors si α est obligatoire, β est obligatoire ;
(D) si α est obligatoire, alors non-α ne l'est pas.

Malgré le deuxième argument, Goble soutient que (Dist) ne doit pas être rejeté puisqu'il s'agit d'un principe fondamental, permettant d'inférer à partir d'un ensemble d'obligations d'autres obligations pouvant en être dérivées. De fait, le principe de distribution est nécessaire afin d'exprimer la relation de conséquence déontique, à savoir que si α implique β, alors $O\alpha$ implique $O\beta$ (voir aussi Castañeda 1968). Cela permet aussi de représenter la distinction entre les obligations *fixes* et les *dérivées* (Alchourrón et Bulygin 1981).

Dans son article, Goble analyse comment ces principes sont liés et quelles sont les modifications pouvant leur être apportées. Cela dit, Goble en conclut qu'une logique voulant modéliser adéquatement les conflits d'obligations doit satisfaire quatre conditions :

1. la consistance (un conflit ne doit pas entraîner de contradiction) ;
2. la non-trivialité (ne doit pas satisfaire l'explosion déontique) ;
3. le système doit satisfaire le schéma d'inférence « s'il est obligatoire que soit α ou β et qu'il est obligatoire que non-α, alors il est obligatoire que β » (il s'agit du syllogisme disjonctif) ;
4. le système doit satisfaire le schéma d'inférence « il est obligatoire que α et β implique qu'il est obligatoire que α ».

La conclusion de l'auteur est que ni (Can) ni (Agg) sont problématiques, mais que la solution consiste plutôt à utililiser une version restreinte du principe de distribution (Dist). Soulignons cependant que Goble considère que (D) doit être rejeté puisqu'il s'agit d'un principe interdisant les conflits. Pour une analyse détaillée de son argumentaire, le lecteur est invité à consulter Peterson (2014a).

* * *

Les arguments en faveur de fondations non monotones pour la logique déontique agissent de deux manières : on insiste soit sur le fait que les systèmes standard ne permettent pas de rendre compte des conflits d'obligations, soit sur le fait qu'ils ne sont pas en mesure de modéliser adéquatement les obligations conditionnelles et les conflits potentiels qui peuvent en résulter. Hormis les approches qui mettent l'accent sur le caractère non monotone des inférences quotidiennes, on trouve aussi des approches multi-modales qui visent à répondre aux mêmes problèmes, et que nous allons maintenant aborder.

Chapitre 8

Les logiques multi-modales

À l'instar du chapitre sur les systèmes monadiques, qui recoupe d'autres approches comme certaines logiques déontiques algébriques, le présent chapitre n'abordera pas toutes les approches qui tombent sous la bannière des logiques multi-modales. Notamment, plusieurs approches en logique STIT, comme celle de Horty (2001), de Pacheco et Carmo (2003) et de Broersen (2011a), et en logique dynamique, par exemple Segerberg (2009) et van der Meyden (1996), ne seront pas traitées ici, mais plus loin.

Les systèmes normatifs

La logique déontique est parfois utilisée comme un outil d'analyse d'un *système normatif* (*normative system*, NS), ou, plus précisément, de *système normatif avec plusieurs agents* (*multiagent normative system*, NMAS). Carmo et Jones (2002) ont proposé la définition usuelle d'un *système normatif*[1] : un ensemble d'agents (humains ou artificiels), dont les interactions sont gouvernées par des normes, où les normes prescrivent comment les agents devraient idéalement agir ou non en spécifiant ce qu'ils peuvent faire, et ce qu'ils ont le droit de faire.

Autant dans le cas de l'analyse des NS que dans celui des NMAS, l'objectif est de voir les répercussions des normes sur le comportement d'un agent, la différence entre les deux étant que, du côté du NMAS, on considère les interactions entre les agents, alors que pour le système normatif (entendu en un sens plus général), on ne considère les répercussions des normes que sur un seul agent. Bien que l'objectif soit de raisonner sur ce que les agents doivent faire, il n'est pas de modéliser l'inférence, mais bien d'étudier l'influence des normes sur le comportement d'un individu.

[1] Il s'agit de la définition courante d'un NMAS. Voir aussi Pacheco et Santos (2004). On trouve aussi d'autres définitions pour les NS, comme celle de Føllesdal et Hilpinen (1970, p.16) : « by a 'normative system' we understand simpy any set of normative sentences closed under deduction ».

Dans ce qui suit, nous verrons d'abord l'approche de Carmo et Jones (2002), qui utilisent une myriade de modalités pour pallier les problèmes et les paradoxes de la logique déontique standard, surtout ceux soulevés par Chisholm. Par la suite, nous aborderons brièvement Dellunde (2008), qui propose des fondations communes pour les différentes logiques déontiques multi-modales.

José Carmo et Andrew Jones

À la suite des nombreuses objections contre la logique déontique standard, mais surtout en réaction au paradoxe de Chisholm (1963), les approches en logique déontique se sont complexifiées afin de modéliser certains paramètres mis en lumière par ces problèmes. Même si, d'emblée, la logique déontique visait l'analyse de l'inférence normative, les travaux des chercheurs ayant un intérêt en informatique et en droit ont fait que la logique déontique est parfois considérée comme un outil pertinent à la modélisation d'un système normatif. Par *système normatif*, Carmo et Jones (2002) entendent un ensemble d'agents dont le comportement et les interactions sont contraints par des normes. Les auteurs considèrent la logique déontique comme un outil permettant l'analyse d'un système normatif, plutôt que d'un NMAS, dans la mesure où leur approche ne vise pas à formaliser explicitement les répercussions des normes sur les agents individuels.

L'objectif de Carmo et Jones est de modéliser l'interaction entre un ensemble de normes et le comportement d'un agent. Outre la dimension temporelle implicite à l'évolution des obligations d'un agent, pouvant être exposée par le paradoxe de Chisholm, ce paradoxe met aussi en évidence qu'une logique déontique qui traite d'un système normatif doit être en mesure de rendre compte du fait que certaines obligations surgissent dans des contextes de violation. Afin de rendre compte de ce phénomène mais en laissant de côté la dimension temporelle, Jones et Pörn (1985) ont introduit la notion d'alternative sous-optimale, où certaines obligations ne sont pas réalisées.

En plus de mettre en évidence le fait que les obligations d'un agent évoluent dans le temps et varient lorsque certaines d'entre elles sont violées, le paradoxe de Chisholm montre que certaines obligations sont conditionnelles. Cependant, les obligations conditionnelles font face au problème du détachement, sans compter les problèmes relatifs à la définition de l'opérateur dyadique : L'opérateur $O(/)$ est-il primitif ou peut-il être défini en termes d'obligation actuelle ? Quel(s) lien(s) se trouve(nt) entre les obligations conditionnelles et les actuelles ? Quelles règles gouvernent le passage d'une obligation conditionnelle à une actuelle ?

Afin de répondre à ces problèmes, Carmo et Jones (2002) ont développé une logique déontique multi-modale conformément à la notion d'alternative sous-optimale de Jones et Pörn (1985).[2] Leur système vise à rendre compte de sept principales caractéristiques mises en évidence par le paradoxe de Chisholm. Parmi les caractéristiques mises de l'avant par Carmo et Jones, les plus importantes sont :

1. le paradoxe de Chisholm est cohérent dans la langue naturelle ;
2. les propositions sont indépendantes dans la langue naturelle ;
3. il est possible de dériver des obligations actuelles ;
4. il est possible de dériver des obligations idéales ;
5. les obligations sont parfois violées, ce qui donne lieu à alternatives sous-optimales.

Principalement en raison du fait que le paradoxe peut survenir même dans un contexte où le temps n'a aucune répercussion, les auteurs proposent une logique multi-modale sans modalité temporelle. À la base, l'idée est de faire la distinction entre les obligations *idéales* et les obligations *actuelles* d'un agent, ce qui a une incidence sur la distinction entre les contextes qui *dépendent* des agents et ceux qui n'en *dépendent pas*.

Du côté de la syntaxe, on suppose le langage suivant :

$$\mathcal{L} = \{Prop, (,), \neg, \supset, \wedge, \vee, \equiv, O(/), O_a, O_i, \Box, \blacksquare\}$$

Le langage contient un ensemble dénombrable de variables propositionnelles, les connecteurs usuels pour la négation, l'implication, la conjonction, la disjonction et l'équivalence ainsi que les opérateurs monadiques O_a pour l'obligation actuelle, O_i pour les obligations idéales, $\Box$ et $\blacksquare$ pour la nécessité, avec l'opérateur dyadique $O(/)$ pour l'obligation conditionnelle.

Les opérateurs $\Diamond$ et $\blacklozenge$ sont respectivement les opérateurs duaux de $\Box$ et $\blacksquare$:

$$\Diamond\varphi =_{df} \neg\Box\neg\varphi \qquad \text{(df. } \Diamond\text{)}$$
$$\blacklozenge\varphi =_{df} \neg\blacksquare\neg\varphi \qquad \text{(df. } \blacklozenge\text{)}$$

[2] Voir aussi Jones et Pörn (1986) et Jones (1991).

L'ensemble des énoncés bien formés (EBF) n'est pas explicitement exposé, mais les auteurs mentionnent que celui-ci est défini de manière habituelle. Considérant que l'axiome (A9) admet la combinaison de différentes modalités, cela suggère que l'itération et la combinaison des opérateurs est possible, donc que l'ensemble peut être défini récursivement ainsi :

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \equiv \psi \mid O(\varphi/\psi) \mid O_a\varphi \mid O_i\varphi \mid \Box\varphi \mid \blacksquare\varphi$$

En supposant les axiomes de la logique propositionnelle classique, les schémas d'axiomes proposés par Carmo et Jones sont :

KT pour $\Box$	(A1)
KD pour $\blacksquare$	(A2)
$\Box\varphi \supset \blacksquare\varphi$	(A3)
$\neg O(\bot/\psi)$	(A4)
$(O(\varphi/\psi) \wedge O(\rho/\psi)) \supset O(\varphi \wedge \rho/\psi)$	(A5)
$O(\varphi/\psi) \supset O(\varphi/\varphi \wedge \psi)$	(A6)
$\Diamond O(\varphi/\psi) \supset \Box O(\varphi/\psi)$	(A7)
$(\Diamond(\varphi \wedge (\rho \wedge \psi)) \wedge O(\varphi/\psi)) \supset O(\varphi/\psi \wedge \rho)$	(A8)
$(O_i\varphi \wedge O_i\psi) \supset O_i(\varphi \wedge \psi)$	(A9)
$(O_a\varphi \wedge O_a\psi) \supset O_a(\varphi \wedge \psi)$	(A10)
$\blacksquare\varphi \supset (\neg O_a\varphi \wedge \neg O_a\neg\varphi)$	(A11)
$\Box\varphi \supset (\neg O_i\varphi \wedge \neg O_i\neg\varphi)$	(A12)
$\blacksquare(\varphi \equiv \psi) \supset (O_a\varphi \equiv O_a\psi)$	(A13)
$\Box(\varphi \equiv \psi) \supset (O_i\varphi \equiv O_i\psi)$	(A14)
$[O(\varphi/\psi) \wedge (\blacksquare C \wedge (\blacklozenge\varphi \wedge \blacklozenge\neg\varphi))] \supset O_a\varphi$	(A15)
$[O(\varphi/\psi) \wedge (\Box\psi \wedge (\Diamond\varphi \wedge \Diamond\neg\varphi))] \supset O_i\varphi$	(A16)
$O(\varphi/\psi) \wedge (\blacklozenge(\varphi \wedge \psi) \wedge (\blacklozenge(\neg\varphi \wedge \psi))) \supset O_a(\psi \supset \varphi)$	(A17)
$O(\varphi/\psi) \wedge (\Diamond(\varphi \wedge \psi) \wedge (\Diamond(\neg\varphi \wedge \psi))) \supset O_i(\psi \supset \varphi)$	(A18)

Les axiomes (A1) et (A2) signifient que les modalités $\Box$ et $\blacksquare$ sont axiomatisées respectivement comme une KT- et une KD-modalité. L'opérateur $\Box$ représente une forme de nécessité qu'un agent ne peut éviter. Autrement dit, $\Box\varphi$ signifie que φ est nécessairement vrai, indépendamment des choix ou des actions d'un agent. De la même manière, $\Diamond\varphi$ signifie que φ a le potentiel d'être réalisé (φ n'est pas fixé au moment présent).

L'opérateur ■ représente ce qui est nécessaire relativement aux choix et aux actions posées par l'agent. Il s'agit de ce qui est nécessaire en vertu de ce qui est fixé au moment présent. Quant à ⧫, cela représente la possibilité à un moment donné. Par exemple, il peut être physiquement possible pour un agent de sauver une personne de la noyade ($\Diamond\varphi$) sans pour autant qu'il soit effectivement possible de le faire ($\neg\blacklozenge\varphi$), comme dans un cas où l'agent ne sait pas nager. Informellement, l'opérateur ■ restreint l'ensemble de choix possibles pour un agent relativement au contexte et à ses capacités.

En vertu de l'axiomatisation, $\Box\varphi$ implique que φ est vrai dans le scénario actuel alors que $\blacksquare\varphi$ signifie que φ sera vrai dans tous les scénarios accessibles, mais pas nécessairement dans l'actuel. En ce sens, le second opérateur permet de représenter les choses nécessairement vraies en vertu des actions des agents et dont la vérité est établie au prochain état du système.

L'axiome (A3) signifie que ce qui est nécessaire l'est aussi au moment présent. La contraposée de cet axiome, représentée par (A3*), permet de mieux en saisir le sens : ce qui est possible d'être fait au moment présent est possible (d'être fait). La notion de possibilité représentée par $\Diamond$ est donc plus large que celle exprimée par ⧫. S'il est effectivement possible de faire φ, alors il est logiquement, physiquement et temporellement possible de faire φ.

$$\blacklozenge\varphi \supset \Diamond\varphi \qquad \text{(A3*)}$$

Dans le même ordre d'idées, l'opérateur O_i représente ce qui est *idéalement* obligatoire, c'est-à-dire ce qui est obligatoire dans toutes les alternatives idéales (parfaites) à un scénario. De l'autre côté, O_a représente ce qui est *effectivement* obligatoire, et qui peut être le fruit de la violation d'une obligation idéale.

L'opérateur dyadique $O(\varphi/\psi)$ se lit « dans les circonstances ψ, φ est obligatoire », et l'axiome (A4) signifie que, dans toute circonstance ψ, il est faux que l'absurde (l'impossible) est obligatoire. (A5) est le principe d'agrégation pour les obligations attribuables à un même contexte : si φ est obligatoire dans le contexte ψ et que ρ l'est aussi, alors la conjonction de φ et ρ est aussi obligatoire dans ce contexte. (A6) stipule que si φ est obligatoire conditionnellement au contexte ψ, alors φ est aussi obligatoire dans un contexte où la conjonction $\psi \wedge \varphi$ est vraie.

L'axiome (A7) gère la relation entre la nécessité et l'obligation conditionnelle. Sémantiquement, l'axiome indique que s'il est possible que φ soit obligatoire dans le contexte C pour un scénario w, alors il est vrai que φ est une obligation conditionnelle à C dans les alternatives possibles à w. En d'autres termes, si $O(\varphi/\psi)$ est vrai pour un scénario accessible, alors $O(\varphi/\psi)$ est vrai pour tous les scénarios accessibles. Quant à l'axiome (A8),

cela signifie que s'il est possible que trois événements φ, ρ et ψ se produisent conjointement, alors si φ est obligatoire dans les conditions ψ, φ l'est aussi dans les conditions $\psi \wedge \rho$.

Les axiomes (A9) et (A10) représentent respectivement l'agrégation pour les obligations idéales et les actuelles. Les axiomes (A11) et (A12) posent que la nécessité influence ce qui est obligatoire. (A11) indique que si φ est effectivement nécessaire, alors ni φ ni $\neg\varphi$ ne sont effectivement obligatoires. De même, pour l'obligation idéale : si φ est nécessaire, alors ni φ ni $\neg\varphi$ n'est idéalement obligatoire. Ces axiomes indiquent qu'il doit être possible pour un agent d'agir à l'encontre de ses obligations.

(A13) et (A14) indiquent que les équivalences peuvent être substituées au sein des obligations. (A13) signifie que si φ et ψ sont équivalent dans un contexte, alors φ est effectivement obligatoire si et seulement si ψ l'est aussi.

Exemple

Si dans un contexte « Paul sauve la vie de Pierre si et seulement si Paul conduit à 150 km/h », alors si Paul, dans ce contexte, a l'obligation actuelle de sauver la vie de Pierre, il aura aussi l'obligation actuelle de conduire à 150 km/h.

La même interprétation s'applique pour (A14) : s'il est nécessaire que φ est vrai si et seulement si ψ est vrai, alors φ est une obligation idéale si et seulement si ψ en est aussi une. Notons que la nécessité actuelle rend *nécessairement effectivement* vraies des propositions pas *nécessairement* vraies. L'opérateur $\blacksquare$ restreint les interprétations où une proposition est vraie à un certain contexte.

L'axiome (A15), quant à lui, signifie que si φ est une obligation conditionnelle à un contexte ψ, alors si ψ est nécessairement effectivement vrai (les interprétations sont restreintes aux scénarios où ψ est vrai), qu'il est effectivement possible que φ soit vrai et qu'il est effectivement possible que φ soit faux, alors φ est effectivement obligatoire. Autrement dit, si φ est une obligation conditionnelle à ψ et que le contexte ψ est nécessairement vrai dans les circonstances actuelles (il est effectivement impossible que ψ soit faux), alors si un agent a la possibilité actuelle de faire φ ou de ne pas faire φ, et donc que l'agent n'est pas contraint dans le contexte à ne pas faire φ, alors φ est effectivement obligatoire. La condition $\blacklozenge(\neg\varphi \wedge \psi)$ vise à assurer que l'agent a la possibilité d'agir à l'encontre de ses obligations. Il s'agit là d'une version du principe de contingence. La même lecture s'applique pour l'obligation idéale et la nécessité pour (A18).

La présentation axiomatique est accompagnée de quelques règles d'inférences usuelles, incluant le *modus ponens* et les deux règles de substitution suivantes, qui permettent de rendre compte des équivalences logiques du point de vue de l'antécédent (REA) et du conséquent (REC) :

$$\frac{\vdash \rho \equiv \psi}{\vdash O(\varphi/\rho) \equiv O(\varphi/\psi)}\ (\text{REA}) \qquad \frac{\vdash \psi \supset (\varphi \equiv \rho)}{\vdash O(\varphi/\psi) \equiv O(\rho/\psi)}\ (\text{REC})$$

En ce qui à trait à la sémantique, Carmo et Jones définissent un modèle $\mathcal{M} = \langle W, av, pv, ob, V\rangle$, où $W \neq \varnothing$ est un ensemble (non vide) de scénarios possibles et $V : Prop \longrightarrow \wp(W)$ une fonction qui détermine l'ensemble des scénarios pour lesquels une proposition atomique est vraie. La fonction $av : W \longrightarrow \wp(W)$, où $av \neq \varnothing$, détermine l'ensemble des alternatives actuelles à un scénario w, lesquelles prennent en compte le contexte.

Dans le même ordre d'idées, la fonction $pv : W \longrightarrow \wp(W)$ détermine les alternatives potentielles à w. Par exemple, soit w un scénario où Pierre menace Paul avec un couteau afin d'obtenir son portefeuille. Un scénario v où Paul ne donne pas son portefeuille à Pierre est une alternative *potentielle* au scénario w, mais n'est pas une alternative *actuelle*. La fonction pv est restreinte par les clauses $av(w) \subseteq pv(w)$ et $w \in pv(w)$. Autrement dit, toute alternative actuelle est une alternative potentielle (A3) et tout scénario w est une alternative potentielle à lui-même (en vertu de l'axiomatisation par KT). Cependant, w n'est pas nécessairement une alternative actuelle à lui-même puisque $\blacksquare$ est axiomatisé par KD, et, de fait, la relation induite sur le modèle sémantique n'est pas réflexive pour cet opérateur.

Finalement, la fonction $ob : \wp(W) \longrightarrow \wp(\wp(W))$ sélectionne les scénarios dans lesquels une proposition est obligatoire relativement à un certain contexte. Elle prend un ensemble de scénarios et l'associe à un ensemble d'ensembles de scénarios. Si $Y \in ob(X)$, alors Y est un ensemble de scénarios obligatoires dans le contexte X. La fonction est restreinte par quatre conditions, où $X, Y, Z \in \wp(W)$:

1. $\varnothing \notin ob(X)$;
2. si $Y \cap X = Z \cap X$, alors $Y \in ob(X)$ si et seulement si $Z \in ob(X)$;
3. si $Y, Z \in ob(X)$, alors $Y \cap Z \in ob(X)$;
4. si $Y \subseteq X$, $Y \in ob(X)$ et $X \subseteq Z$, alors $((Z - X) \cup Y) \in ob(Z)$.

Carmo et Jones mentionnent que la première condition indique qu'une contradiction ne peut pas être obligatoire (axiome A4) : l'absurde n'est pas obligatoire dans un contexte exprimé par un scénario membre de X. En fait, la condition indique que la relation pour l'obligation conditionnelle est sérielle (rappelons-nous que dans KD il y a une équivalence entre **D** et

$\neg O\bot$) : pour n'importe quel contexte exprimé par un scénario membre de X, il y a minimalement une alternative obligatoire, c'est-à-dire que l'ensemble qui contient les ensembles qui contiennent les scénarios obligatoires dans le contexte X n'est pas vide.

La seconde condition vise à assurer que si deux propositions sont équivalentes dans un contexte, alors l'une est obligatoire dans le contexte X si et seulement si l'autre l'est aussi. Cela se comprend par le fait qu'un scénario est un ensemble de propositions, lesquelles décrivent un état de choses dans le monde. En supposant que l'intersection des ensembles X et Y est égale à celle de X et Z, et donc que chacune des intersections *décrit la même chose*, alors la description faite par Y est membre de l'ensemble qui contient les scénarios obligatoires relativement au contexte X si et seulement si celle faite par Z l'est aussi.

La troisième condition correspond au principe d'agrégation et signifie que si deux propositions sont obligatoires dans un contexte X, alors leur conjonction l'est aussi (A5, A9 et A10). Finalement, la dernière condition signifie que si Y est sous-ensemble de X, que X est sous-ensemble d'un contexte Z et que Y est obligatoire dans le contexte X, alors ce qui est obligatoire dans le contexte Z se trouve soit dans Y soit dans la partie de Z qui ne contient pas X. En quelque sorte, l'idée est d'isoler l'ensemble de scénarios obligatoires Y relativement à un contexte plus large X, où le complément de Y par rapport à X n'est pas nécessairement obligatoire, et de dire que dans un contexte encore plus large Z, ce qui est obligatoire dans le contexte Z est soit dans Y soit dans le complément de X par rapport à Z.

La vérité dans le modèle sémantique est définie récursivement de manière usuelle. Soit $\|\varphi\|$ l'ensemble qui contient les scénarios où la proposition φ est vraie pour le modèle $\mathcal{M}$[3] :

$$\|\varphi\| = \{w \in W : \models_{\mathcal{M},w} \varphi\}$$

$$\models_{\mathcal{M},w} p \Leftrightarrow w \in V(p) \tag{1}$$
$$\models_{\mathcal{M},w} \neg\varphi \Leftrightarrow \not\models_{\mathcal{M},w} \varphi \tag{2}$$
$$\models_{\mathcal{M},w} \varphi \wedge \psi \Leftrightarrow \models_{\mathcal{M},w} \varphi \text{ et } \models_{\mathcal{M},w} \psi \tag{3}$$
$$\models_{\mathcal{M},w} \varphi \vee \psi \Leftrightarrow \models_{\mathcal{M},w} \varphi \text{ ou } \models_{\mathcal{M},w} \psi \tag{4}$$
$$\models_{\mathcal{M},w} \varphi \supset \psi \Leftrightarrow \not\models_{\mathcal{M},w} \varphi \text{ ou } \models_{\mathcal{M},w} \psi \tag{5}$$
$$\models_{\mathcal{M},w} \varphi \equiv \psi \Leftrightarrow (\models_{\mathcal{M},w} \varphi \text{ et } \models_{\mathcal{M},w} \psi) \text{ ou } (\not\models_{\mathcal{M},w} \varphi \text{ et } \not\models_{\mathcal{M},w} \psi) \tag{6}$$
$$\models_{\mathcal{M},w} \blacksquare\varphi \Leftrightarrow av(w) \subseteq \|\varphi\| \tag{7}$$

[3] Nous avons explicité les conditions sémantiques des connecteurs propositionnels, laissés implicites chez Carmo et Jones.

$$\models_{\mathcal{M},w} \Box\varphi \Leftrightarrow pv(w) \subseteq \|\varphi\| \tag{8}$$

$$\models_{\mathcal{M},w} O(\varphi/\psi) \Leftrightarrow \|\varphi\| \cap \|\psi\| \neq \varnothing \text{ et}$$
$$X \subseteq \|\psi\| \text{ et } X \cap \|\varphi\| \neq \varnothing \Rightarrow \|\varphi\| \in ob(X) \tag{9}$$

$$\models_{\mathcal{M},w} O_a\varphi \Leftrightarrow \|\varphi\| \in ob(av(w)) \text{ et } av(w) \cap \|\neg\varphi\| \neq \varnothing \tag{10}$$

$$\models_{\mathcal{M},w} O_i\varphi \Leftrightarrow \|\varphi\| \in ob(pv(w)) \text{ et } pv(w) \cap \|\neg\varphi\| \neq \varnothing \tag{11}$$

Comme d'habitude, φ est vrai pour un modèle $\mathcal{M}$ à condition que $\models_{\mathcal{M},w} \varphi$ pour tout w, et φ est valide lorsque φ est vrai pour tout modèle.

Les six premières clauses sont les conditions habituelles pour les connecteurs classiques. Dans la première clause, V est une fonction qui assigne un ensemble de valeurs de vérité aux propositions atomiques.

Les clauses (7) et (8) signifient respectivement que $\blacksquare\varphi$ est vrai à condition que l'ensemble qui contient les alternatives *actuelles* à w est inclus dans l'ensemble qui contient les scénarios où φ est vrai, et $\Box\varphi$ est vrai à condition que l'ensemble qui contient les alternatives *potentielles* à w est inclus dans l'ensemble qui contient les scénarios où φ est vrai.

La clause (9) signifie que la proposition « φ est obligatoire dans les conditions ψ » est vraie si et seulement s'il y a au moins un scénario où à la fois φ est vrai et ψ est vrai et que si X est un ensemble de scénarios où ψ est vrai et qu'il y a au moins un scénario où à la fois les propositions membres des scénarios de X sont vraies et φ est vrai, alors l'ensemble qui contient les scénarios où φ est vrai est obligatoire dans les conditions X. Autrement dit, $O(\varphi/\psi)$ est vrai pour w à condition qu'il y ait au moins un scénario v tel que $\varphi, \psi \in v$ et que s'il y a au moins un scénario v' tel que $\varphi \in v'$ et $v' \in X$ (où X contient des scénarios où la condition ψ est vraie), alors tout scénario où φ est vrai est obligatoire dans un contexte où les propositions membres des scénarios de X sont vraies. En isolant un ensemble $X \subseteq \|\psi\|$, on détermine un ensemble où la proposition ψ est vraie, mais où il n'y a pas d'autres conditions ψ' qui s'ajoutent au scénario et qui pourraient faire changer la valeur de vérité de l'obligation actuelle O_a.

La première partie de la clause (10) indique que φ est effectivement obligatoire pour un scénario w si tout scénario où φ est vrai est obligatoire dans les conditions v, où v est une alternative actuelle à w. La seconde indique qu'il y a au moins un scénario v' qui est une alternative actuelle à w, mais où φ est faux pour v'. En ce sens, ce qui est obligatoire n'est pas inévitable.

Enfin, la clause (11) stipule qu'il est vrai que φ est idéalement obligatoire pour un scénario w si tout scénario où φ est vrai est obligatoire dans les conditions v, où v est une alternative potentielle à w, et qu'il y a au moins un scénario v' tel que v' est une alternative potentielle à w et φ

est faux pour v'. La seconde partie de cette clause indique, à l'instar de la clause (10), que ce qui est idéalement obligatoire n'est pas inévitable. La seconde partie des clauses (10) et (11) se comprend en fonction de l'argument de Jones et Pörn (1985), pour qui les tautologies ne sont pas obligatoires, puisqu'il est impossible d'agir à l'encontre d'une tautologie.

Somme toute, le système proposé par Carmo et Jones vise à rendre compte des obligations conditionnelles dans un cadre qui admet des situations non idéales, où certaines obligations sont enfreintes. L'idée de base est d'être en mesure de déterminer ce qu'un agent devrait faire, même dans des situations où son comportement n'est pas conforme à celui prescrit par les normes. Le lecteur intéressé par l'approche de Carmo et Jones trouvera une critique ainsi qu'une réduction de leur système dans l'article de Gabbay et Schlechta (2010), ainsi que la preuve de complétude et de la décidabilité du système dans Carmo et Jones (2013).

Pilar Dellunde

La majorité des approches qui visent à modéliser l'interaction entre un système normatif et un système multi-agents utilisent une variante ou une autre des logiques modales. Cela dit, trois logiques importantes utilisées pour modéliser l'évolution d'un système dans le temps, de par la représentation du passage d'un état à un autre, sont les logiques dynamiques, ATL (*Alternating-Time Temporal Logic*) et CTL *Computational Tree Logic* (Dellunde 2008).

L'approche de Dellunde (2008) vise à construire une logique *générale* qui permet de rendre compte de plusieurs logiques différentes (p. ex. PD_eL, ATL, CTL, etc.). Ainsi, plutôt que de définir une logique déontique spécifique, Dellunde présente un cadre de travail général pouvant servir de base pour la construction de différentes logiques multi-modales. Son objectif est de définir un cadre de travail général permettant d'englober différentes approches avec des caractéristiques communes. En ce sens, elle définit un système qui englobe une classe de systèmes spécifiques, lesquels varieront en fonction des axiomes (ou règles) ajoutés.

Plutôt que de présenter l'analyse proposée par Dellunde en détail, nous allons nous limiter à présenter le matériel nécessaire à la compréhension du système général permettant la représentation et la comparaison de différentes logiques multi-modales, qui constitue l'essentiel de son approche. Nous avons jugé bon de présenter l'approche de Dellunde considérant, d'une part, que cela offre un cadre commun à partir duquel différentes logiques peuvent être analysées et que, d'autre part, l'auteur propose une interprétation d'un système normatif différente de ce qu'on retrouve usuellement dans la littérature scientifique.

L'objectif de Dellunde est de définir un cadre de travail commun pour les logiques qui visent à formaliser la structure d'un système normatif. Cependant, contrairement à la définition usuelle d'un système normatif, c'est-à-dire l'interaction entre un ensemble de normes et un ensemble d'agents, Dellunde conçoit un système normatif η comme un *ensemble de contraintes* restreignant la transition d'un état à l'autre du système.

Plus précisément, dans le cas d'une structure $\langle U, R\rangle$, où U est un ensemble d'états et R un ensemble de paires ordonnées (voire une relation, $R \subseteq U \times U$), $\eta \subseteq R$ contient les transitions interdites pour un système donné. En ce sens, si $(w, v) \in \eta$ (ou encore $wR_\eta v$), alors la transition de l'état w à l'état v est interdite. Soulignons que dans les approches plutôt axées vers l'informatique, nous parlons d'*états* (statiques) plutôt que de *scénarios* considérant qu'il y a toujours une forme ou une autre de logique temporelle en arrière-plan. En supposant une logique L visant à modéliser un MAS (*multi-agent system*), Dellunde considère la logique d'un système normatif L_η comme l'augmentation de L à l'aide d'un ensemble de systèmes normatifs (i.e., un ensemble d'ensembles de contraintes).

Même si le langage développé par Dellunde pouvait, dans une certaine mesure, être compris comme un *méta-langage* étant donné qu'il s'agit d'un langage général qui permet de parler de plusieurs autres langages plus spécifiques, il faut garder à l'esprit que l'objectif est de développer un cadre de travail abstrait permettant de voir sous le même angle des systèmes fondamentalement différents.

Soit $\tau = \langle F, \rho\rangle$ un type qui contient un ensemble de modalités F pour lesquelles $\rho : F \longrightarrow \mathbb{N}$ assigne un nombre naturel à chaque modalité (et T l'ensemble de tous les types). Un langage $\mathcal{L}$ pour un type τ est composé des connecteurs logiques (propositionnels classiques) usuels, des symboles $\top$ et $\bot$ (qui représentent respectivement une tautologie et une contradiction arbitraire) et un nombre fini de modalités $\Box_1, ..., \Box_n$.

Rappelons-nous qu'une logique peut être définie comme « le plus petit ensemble de formules bien formées étant fermé sous les règles $r_1, ..., r_k$ et contenant les axiomes (A1),...,(A_n) ». Lorsqu'elle est conçue ainsi, on considère un ensemble d'axiomes Γ ainsi qu'une relation de conséquence $\vdash$ (ou encore Cn) qui répond aux règles d'inférences désirées, et on définit la logique (i.e., l'ensemble des théorèmes) comme $Cn(\Gamma)$, soit l'ensemble qui contient toutes les conséquences logiques de Γ. Autrement dit, $Cn(\Gamma)$ est la logique ayant comme axiomes les formules $\varphi \in \Gamma$ et comme règles d'inférences les conditions sur Cn.

L'objectif de Dellunde est de proposer une logique contenant le plus petit ensemble d'axiomes (communs à toute logique modale) possible. En ce sens, Dellunde définit une relation de conséquence $\vdash_\tau$ relativement à un

type τ (i.e., relativement à un langage ; un langage contient un nombre fini de modalités et Dellunde définit la logique pour tout langage modal). La relation de conséquence se résume ainsi :

1. $\vdash_\tau$ contient les axiomes de la logique propositionnelle classique ;
2. $\vdash_\tau$ est fermée sous le *modus ponens* ;
3. $\vdash_\tau$ est compact ;
4. $\vdash_\tau$ est transitive ;
5. $\vdash_\tau$ est monotone ;
6. $\vdash_\tau$ est fermée sous une règle de substitution pour les propositions équivalentes ;
7. $\vdash_\tau$ est fermée sous la règle (RE) pour chaque modalité $\Box_i$ du langage ;
8. toute proposition membre d'un ensemble de propositions est la conséquence de cet ensemble.

Bien que Dellunde, faisant appel à l'ouvrage de Gabbay (1994), nomme les logiques qui répondent à ces critères des *logiques modales classiques*, il s'agit en fait de la définition proposée par Chellas (1980). Notons que la définition est plus succincte chez Chellas : un système *classique* est fermé sous (RE) et possède la définition de l'opérateur dual par $\neg\Box\neg\varphi \equiv \Diamond\varphi$ (cette dernière condition est introduite par la suite chez Dellunde).

Notons qu'un système classique ne doit pas être confondu avec un système *normal*. En général, le système K est la logique modale utilisée comme point de référence dans la littérature scientifique pour quiconque veut modéliser un phénomène avec une logique modale. Cela dit, K est un système *normal* (Chellas 1980). Un système *normal*, plutôt que d'être fermé uniquement sous (RE), l'est sous (RK).

Au même titre que KD est une extension de K, les systèmes normaux sont des extensions (conservatrices) des systèmes classiques. En ce sens, le système classique est plus *faible* que le système normal, c'est-à-dire qu'il permet de prouver moins de théorèmes.

Les conditions 2–6 représentent la notion de conséquence pour la logique propositionnelle classique, qu'on obtient en ajoutant la condition 1. La condition 3 (compacité) signifie que si une proposition est la conséquence d'un ensemble de propositions Γ, alors elle est la conséquence d'un sous-ensemble fini de Γ. Autrement dit, s'il existe une dérivation de φ à partir de Γ, alors il existe une dérivation de φ à partir d'un sous-ensemble fini de Γ (ou de manière similaire, si φ est démontrable à partir de Γ, alors il existe une dérivation de longueur finie de φ à partir de Γ). La condition 7 fait appel à la règle (RE_i).

$$\frac{\vdash \varphi \equiv \psi}{\vdash \Box_i \varphi \equiv \Box_i \psi} \qquad (\mathsf{RE}_i)$$

Finalement, 5 signifie que si φ est la conséquence de Γ et que $\Gamma \subseteq \Delta$, alors φ est aussi la conséquence de Δ.

Nous avons donc la logique $L = \{L_\tau : \tau \in T\}$ où $L_\tau = \{\varphi : \vdash_\tau \varphi\}$. En d'autres termes, L est la logique multi-modale qui contient toutes les logiques multi-modales de type τ.

En supposant un type τ et une structure $S = \langle U, R\rangle$ (avec $U \neq \varnothing$ et $R = \{R_f : f \in F\}$ un ensemble de relations d'accessibilité pour chaque modalité), Dellunde définit un système normatif η comme un sous-ensemble de l'ensemble $U*$:

$$U* = \{(w_1, ..., w_n) : \text{pour tout } i < n \text{ il y a } f \text{ t.q. } w_i R_f w_{i+1}\}$$

L'ensemble $U*$ est un élément de $\wp(U)$, qui contient des séquences finies d'états pour lesquelles chaque w_{i+1} est accessible à son prédécesseur w_i par le biais d'une relation R_f pour une modalité f donnée. L'ensemble $U*$ contient donc des séquences $w_1, \ldots, w_n$ où chaque état w_{i+1} est un successeur à son prédecesseur w_i en vertu d'une modalité (et de la clause sémantique qui l'accompagne) du langage.

Ainsi, étant un sous-ensemble de $U*$, un système normatif isole les séquences finies où chaque transition de w_i à w_{i+1} est interdite. Le système normatif stipule qu'une séquence d'accessibilité est interdite, peu importe le chemin par lequel on passe de w_1 à w_n. En ce sens, $U*$ contient les séquences $(w_1, ..., w_n)$ pour lesquelles toute sous-séquence binaire (w_i, w_{i+1}) est interdite. Un système normatif est un ensemble qui contient des séquences finies à l'intérieur desquelles chaque w_{i+1} est accessible à l'état qui le précède w_i de par une relation R_f quelconque, mais où le passage de w_i à w_{i+1} est interdit.

À partir d'un type τ, le langage est augmenté afin d'obtenir un type τ^η, où $\eta \in N$ est élément d'un ensemble de symboles dénotant des systèmes normatifs. Le langage est alors augmenté par les modalités $\Box_1^\eta, ..., \Box_n^\eta$, où $\Box_i^\eta \varphi$ se lit « φ est obligatoire selon le système normatif η ». En ce sens, φ appartient à tout scénario accessible à w, c'est-à-dire que les séquences qui admettent la transition d'un scénario w à un scénario où φ est faux sont interdites.

En un mot, l'idée est de prendre une logique multi-modale L et de créer une extension conservatrice L^η à l'intérieur de laquelle la relation d'accessibilité $R_f^\eta \subseteq R_f$ détermine les transitions interdites, où le langage de L^η est celui de L augmenté par τ^η. Autrement dit, pour chaque $L_\tau = \{\varphi : \vdash_\tau \varphi\}$ on défini $L_{\tau^\eta} = \{\varphi : \vdash_{\tau^\eta} \varphi\}$ afin d'obtenir $L^\eta = \{L_{\tau^\eta} : \tau^\eta \in T^\eta\}$, où T^η est l'ensemble qui contient tous les types

et types augmentés. La relation de conséquence pour $\vdash_{\tau^\eta}$ est définie de la même manière, seul le langage est augmenté. De fait, nous obtenons une logique multi-modale pour les systèmes normatifs capable de rendre compte d'une diversité de logiques spécifiques qui seront obtenues en ajoutant des conditions sur Cn et des axiomes.

Dans la suite de son article, Dellunde concentre son attention sur les systèmes normatifs *élémentaires*, à savoir ceux dont la structure (particulièrement les relations d'accessibilité) peut être décrite par une logique de premier ordre monadique. Le lecteur intéressé par les formules de Sahqlvist trouvera chez Dellunde (2008) une définition précise.

* * *

Tout compte fait, les logiques multi-modales offrent un cadre de travail riche permettant une analyse détaillée du discours. Bien qu'on puisse trouver une panoplie d'approches multi-modales en logique déontique, dans chaque cas, l'idée est de représenter sémantiquement une modalité en définissant les relations qui permettent d'isoler les scénarios accessibles à la suite de son instanciation. Comme nous l'avons mentionné au début de ce chapitre, il existe plusieurs autres approches multi-modales au sein de la littérature scientifique. Cependant, elles se regroupent sous une autre bannière, celle des logiques de l'action, que nous allons maintenant aborder.

Chapitre 9

La logique stit

L'analyse des systèmes normatifs et des systèmes normatifs à multi-agents a amené certains auteurs à se concentrer sur la notion d'*action*. En effet, dans un système normatif, le comportement des agents est non seulement contraint par des normes, mais les actions de ces agents, qui peuvent aller à l'encontre des normes en place, détermineront l'évolution du système. En conséquence, certains auteurs en sont venus à traiter la logique déontique dans le cadre d'une logique de l'action.

Suivant Segerberg *et al.* (2009), les logiques de l'action qu'on retrouve en philosophie se divisent en deux catégories, à savoir les logiques STIT et les logiques dynamiques. Alors que le présent chapitre porte sur les logiques de l'action de type STIT, les logiques déontiques dynamiques seront considérées au chapitre suivant. En intégrant la logique déontique à une forme ou une autre de logique de l'action, il devient possible de modéliser le fait que les obligations s'adressent à des agents en particulier et que les actions des agents influencent les normes en place pour une situation donnée.

Pour une analyse philosophique détaillée de la notion d'*action humaine*, le lecteur est invité à consulter Peterson (2015c).

Origines

En considérant que la logique déontique est un outil pouvant servir à l'analyse d'un NMAS, lequel est défini comme un ensemble de normes qui gouvernent le comportement d'un ensemble d'agents, l'intérêt d'intégrer une logique de l'action à la logique déontique devient évident. Les logiques STIT, dont l'acronyme signifie *seeing to it that*, visent à rendre compte de la réalisation de certaines actions par des agents.[1]

[1] Le lecteur intéressé par les logiques de l'action est invité à consulter Segerberg (1992) et Segerberg *et al.* (2009).

Bien que l'origine de la logique STIT soit généralement attribuée à Belnap et Perloff (1988), qui ont introduit l'opérateur [*i stit* : φ], l'origine des logiques de type STIT peut être retracée aux travaux de Kanger (1957)[2], qui a été le premier à traiter des notions déontiques à l'aide d'une logique de l'action (Kanger 1972).

Quoique Chellas (1969) ait été le premier à proposer une sémantique bien définie pour les logiques de type STIT (Segerberg 1992), l'opérateur E_i indexé à un agent i est apparu avec Pörn (1970) (Carmo et Pacheco 2001), où l'auteur cherchait à rendre compte des relations d'influence et des relations normatives qui se trouvent entre certains agents.

L'opérateur E_i, qui, en français, peut être lu « *i* réalise... », se comprend mieux en fonction de sa formulation anglaise, où l'opérateur signifie « *i sees to it that...* » ou encore « *i brings it about that...* ». C'est d'ailleurs en raison de cette première formulation que Belnap et Perloff (1988) ont utilisé l'acronyme *stit* pour définir l'opérateur [*i stit* : φ].

Belnap et Perloff (1988) ont développé une logique STIT qui inclut une dimension temporelle, conformément aux travaux de Chellas (1969). En ce qui nous concerne, nous regroupons sous la bannière des logiques de type STIT celles qui utilisent un opérateur du type « *i* réalise... ».

Plusieurs auteurs utilisent des logiques de type STIT, et faire une revue complète de cette littérature scientifique dépasserait la portée de notre propos. Le lecteur intéressé est invité à consulter Xu (1995) pour le développement formel de la logique STIT, et Horty et Belnap (1995) pour une introduction détaillée à la logique déontique de type STIT.

Réduction Anderson-Kanger

La réduction de type Anderson-Kanger n'est pas propre à la logique STIT, mais est surtout utilisée dans le cadre des logiques déontiques axées sur des logiques d'action. Une réduction de type Anderson-Kanger consiste à ne pas considerer les opérateurs déontiques comme primitifs, mais plutôt à les définir en fonction de certaines constantes.

Alors qu'Anderson (1958a,b, 1967) a proposé de définir l'obligation en fonction d'une constante qui représente ce qui est *mauvais*, voire une constante de violation ou de sanction, Kanger (1957, 1971) a plutôt proposé de définir l'obligation en fonction de ce qui est *désirable*. Ainsi, chez Anderson, une action ou un état de choses est obligatoire lorsque ne pas réaliser cette action ou cet état de choses mène à quelque chose de mauvais. À l'inverse, Kanger considère qu'une action ou un état de choses est

[2] Le texte de Kanger (1957) a été réimprimé dans Kanger (1971) et ses travaux se sont développés dans Kanger et Kanger (1966) et Kanger (1972).

obligatoire lorsque cette action ou cet état de choses mène à quelque chose de désirable.

Bref, une réduction de type Anderson-Kanger consiste à réduire l'opérateur O à une constante qui indique quelque chose de bon ou de mauvais.[3] Cela dit, il faut être prudent avec une telle réduction puisque cela nous ramène au problème posé par Jørgensen. En effet, définir ce qui devrait être ou ce qui devrait être fait en fonction d'une constante qui décrit un état de choses se veut, dans une certaine mesure, une façon de réduire les énoncés normatifs à des énoncés descriptifs. Ainsi, il faut prendre garde à ne pas tenter de définir ce qui devrait être, ce qui devrait être fait ou encore ce qui est bon ou mauvais en fonction d'états de choses ou de propriétés purement descriptives (Moore 1971).

Comme nous le verrons, les réductions de type Anderson-Kanger sont surtout utilisées en logique STIT et en logique dynamique.

Ingmar Pörn

Dans son ouvrage intitulé *The Logic of Power*, Pörn (1970) visait à développer une logique capable de rendre compte des relations de pouvoir coercitif et de pouvoir d'influence entre certains agents.[4] Plus précisément, son objectif était de développer un langage formel capable de rendre compte de plusieurs nuances de la langue naturelle (normative), telles les notions de *droit*, de *devoir*, d'*obligation*, de *tolérance* ou encore de *contrôle*.

Suivant les travaux de Hohfeld (1917), le langage développé a pour but de rendre compte de certaines paires de concepts normatifs, comme *droit/devoir*, *pouvoir/responsabilité* et *immunité/imputabilité*. Les travaux de Pörn visaient à rendre compte de ces concepts de manière générale, en développant une logique qui facilite l'analyse de ces notions. L'ouvrage de Pörn est général dans la mesure où il s'intéresse aux concept de *droit* ou de *responsabilité*, par exemple, sans s'intéresser à la logique sous-jacente aux *droits* ou à la *responsabilité*. En ce sens, l'approche de Pörn possède ses lacunes : son système ne permet pas de rendre compte des notions d'*intention* ou de *responsabilité* du point de vue de l'action, ce qui est crucial dans la mesure où on cherche à formaliser une logique de l'*imputabilité*. Néanmoins, les travaux de Pörn se trouvent d'un point de vue historique au début du courant qui traite de la logique de l'action avec un opérateur de type STIT, d'où l'importance de l'exposer.

[3] Voir McNamara (2010) et Straßer et Beirlaen (2012).

[4] Pörn définit les relations (binaires) de pouvoir entre les agents comme des ensembles qui contiennent des formules d'un certain type. Voir Pörn (1970) pour les relations d'influence et les relations normatives.

Pörn (1970) a développé une logique de l'action axée sur un langage similaire au calcul de premier ordre augmenté à l'aide de deux opérateurs (modaux) visant à rendre compte de l'action.[5] Du côté de la syntaxe, on utilise le langage $\mathcal{L}$, qui contient les connecteurs usuels ainsi que le quantificateur universel, à partir desquels les autres connecteurs sont définis.

$$\mathcal{L} = \{(,), \neg, \supset, E_i, C_i, \forall, =, B, \mathbb{P}, \mathbb{CV}\}$$

Les opérateurs $E_i\varphi$ et $C_i\varphi$ signifient respectivement « l'agent i réalise φ » et « φ est compatible avec les actions posées par i ». Les symboles $=$ et B sont des prédicats qui représentent respectivement l'identité et le fait qu'un agent subisse quelque chose de mauvais (B est utilisé pour *bad*). Autrement dit, Bi signifie que « i subit quelque chose de mauvais » ce qui sera utilisé pour définir l'obligation. Même si cette définition ressemble à celle proposée par Anderson (1958a), elle n'y est pas équivalente.

Les ensembles $\mathbb{P} = \{P_1^1, ..., P_m^n, ...\}$, $\mathbb{C} = \{Ag_c, Obj_c\}$ et $\mathbb{V} = \{Ag_v, Obj_v\}$ sont respectivement des ensembles de prédicats (où P^1 est un prédicat unaire, P^2 binaire, etc.), de constantes d'agents et d'objets et de variables d'agents et d'objets. L'ensemble des énoncés bien formés est défini récursivement :

1. si $s_1, s_2, ..., s_n \in \mathbb{C} \cup \mathbb{V}$ et $P_i^n \in \mathbb{P}$, alors $s_1 = s_2, P_i^n(s_1, ..., s_n) \in EBF$;
2. si $\varphi, \psi \in EBF$, $x \in \mathbb{V}$, alors $\neg\varphi, \varphi \supset \psi, (\forall x)\varphi \in EBF$;
3. si $\varphi \in EBF$ et $i \in Ag_c \cup Ag_v$, alors $E_i\varphi, C_i\varphi, Bi \in EBF$.

La sémantique est définie suivant les travaux de Hintikka, lequel fait la distinction entre un *ensemble modèle* (*model set*) et un *système de modèle* (*model system*). Cela dit, ces notions peuvent être comprises en termes de *scénario* et de *modèle* dans le cadre d'une sémantique des mondes possibles. Un scénario w doit satisfaire les conditions suivantes :

$$\varphi \in w \Rightarrow \neg\varphi \notin w \tag{1}$$
$$\neg\neg\varphi \in w \Rightarrow \varphi \in w \tag{2}$$
$$\varphi \supset \psi \in w \Rightarrow \neg\varphi \in w \text{ ou } \psi \in w \tag{3}$$
$$\neg(\varphi \supset \psi) \in w \Rightarrow \varphi \in w \text{ et } \neg\psi \in w \tag{4}$$
$$(\forall x)\varphi \in w \Rightarrow \varphi(s/x) \in w \text{ pour tout } s \in \mathbb{C} \tag{5}$$
$$(\forall x)\varphi \in w \Rightarrow \varphi(s/x) \in w \text{ pour au moins un } s \in \mathbb{C} \tag{6}$$

[5] Le symbolisme utilisé par Pörn a été modifié de façon à ce que la présentation de son approche soit plus concise et cohérente avec la notation utilisée dans le présent ouvrage. Nous avons d'ailleurs formulé la sémantique qu'il propose dans le cadre d'une sémantique de Kripke plutôt que la sémantique développée par Hintikka.

$$\neg(\forall x)\varphi \in w \Rightarrow \neg\varphi(s/x) \in w \text{ pour au moins un } s \in \mathbb{C} \tag{7}$$
$$\varphi, s_1 = s_2 \in w \text{ et } \varphi(x/s_1) = \psi(x/s_2) \Rightarrow \psi \in w \tag{8}$$
$$s_1 \neq s_1 \notin w \text{ pour tout } s_1 \in \mathbb{C} \cup \mathbb{V} \tag{9}$$

Les conditions 1 et 2 assurent la consistance de w ainsi que la bivalence classique. Les clauses 3 et 4 représentent les conditions sémantiques habituelles pour l'implication matérielle. En un mot, les conditions 1–4 assurent que les scénarios, qui sont des ensembles de propositions, sont fermés d'un point de vue syntaxique sous les règles et axiomes de la logique propositionnelle classique.

La condition 5 indique que la quantification se fait sur l'ensemble du domaine de w, représenté par les constantes de $\mathbb{C}$, où $\varphi(s/x)$ est la proposition obtenue après avoir substitué x par s au sein de φ (avec s élément du domaine). La clause 6 indique qu'une quantification se fait sur un domaine non vide, au même titre que 7 indique que la négation d'une quantification sur φ est vraie à condition qu'il y ait au moins un objet du domaine pour lequel φ est faux. La clause 8 signifie que pour deux objets s_1 et s_2 identiques, si φ et ψ sont similaires, c'est-à-dire que lorsque toute occurrence de s_1 ou de s_2 est remplacée par x dans chaque énoncé entraîne que $\varphi(x/s_1) = \psi(x/s_2)$, alors la proposition ψ est aussi vraie. Autrement dit, les énoncés obtenus après substitution d'objets identiques dans la portée des prédicats se trouvent aussi dans le scénario. Finalement, 9 assure que tout énoncé d'identité est vrai pour un ensemble modèle, et donc que dans le scénario, aucun objet (ou aucune constante) n'est différent de lui-même.

Le modèle est défini comme une paire ordonnée (U, R), où U est un ensemble de scénarios et $R \subseteq U \times U$ une relation binaire définie pour tout agent i. Notons que, à strictement parler, il s'agit d'une *structure* et non d'un *modèle*. Néanmoins, une fonction d'attribution de valeur de vérité peut aisément être définie. Cette paire ordonnée doit satisfaire les conditions suivantes :

$$E_i\varphi \in w \text{ et } wR_iv \Rightarrow \varphi \in v \tag{10}$$
$$\neg E_i\varphi \in w \Rightarrow C_i\neg\varphi \in w \tag{11}$$
$$C_i\varphi \in w \Rightarrow \text{ il y a au moins un } v \text{ t.q. } wR_iv \text{ et } \varphi \in v \tag{12}$$
$$\neg C_i\varphi \in w \Rightarrow E_i\neg\varphi \in w \tag{13}$$
$$wR_iw \text{ pour tout } w \in U \tag{14}$$

La clause 10 correspond à la condition usuelle pour la valeur de vérité de $\Box$ dans le système modal K et 11 signifie que s'il est faux que i réalise φ, alors les actions de i sont compatibles avec l'état de choses décrit par $\neg\varphi$. La condition 12 indique que si les actions de i sont compatibles avec φ pour un scénario w, alors φ est vrai dans au moins une alternative v en relation R_i avec w. Autrement dit, φ n'est pas contradictoire. La condition 13 signifie que s'il est faux que les actions de i sont compatibles avec φ, alors i réalise $\neg\varphi$. Notons que cela implique que tout agent réalise toute tautologie. Finalement, la dernière condition indique que la relation R_i est réflexive. Soulignons au passage que même si les constantes (de la syntaxe) sont utilisées pour référer aux objets du domaine (de la sémantique), une distinction nette s'opère entre les deux.

Bien qu'on ne trouve pas d'axiomatisation du système proposé dans l'ouvrage de Pörn, les restrictions du modèle sémantique nous donnent une bonne idée de ce à quoi le système axiomatique ressemblerait. Les clauses 10 et 14 donnent à l'opérateur E_i le schéma d'axiome (**T**). Les clauses 11 et 13 font que, d'un point de vue syntaxique, C_i est l'opérateur dual de E_i, et donc le système aurait pu être axiomatisé à l'aide de l'équivalence $C_i \equiv \neg E_i \neg$. Prendre chaque opérateur comme primitif s'explique probablement du fait que l'interprétation de « φ est compatible avec les actions de i » en termes de « il est faux que i réalise $\neg\varphi$ » ne rendait pas compte adéquatement de la signification de C_i. Notons simplement que le système aurait pu être axiomatisé comme une extension modale du calcul de premier ordre satisfaisant KT.[6]

Pörn utilise le prédicat B, appliqué à un agent, pour définir les énoncés normatifs. Il propose la définition suivante :

$$B_i E_j \varphi =_{df} E_i(E_j \varphi \supset Bj),$$

Cela signifie « i fait advenir que si j réalise φ, alors j subit quelque chose de mauvais ». La formulation anglaise rend mieux compte du sens de l'énoncé : *i sees to it that if j brings it about that* φ*, then something bad happens to j*. Soulignons la différence entre B_i, où l'agent est en indice, et Bj : seulement le deuxième cas signifie que j subit quelque chose de mauvais (*bad*).

D'emblée, cette définition permet de rendre compte de la relation de pouvoir coercitif considérant que l'autorité i contraint j à agir d'une certaine manière (à réaliser φ), sans quoi j fera face à des conséquences négatives. Cette formule est utilisée pour représenter l'interdiction. En effet, $B_i E_j \varphi$ signifie que i interdit à j de faire φ. Similairement, *il est permis*

[6] Le lecteur familier avec les logiques modales reconnaîtra d'ailleurs plusieurs théorèmes et axiomes de KT dans la liste des énoncés valides énoncés par Pörn.

que j réalise φ signifie qu'il est faux que i s'assure que j subit une conséquence négative s'il réalise φ. En ce sens, la permission est représentée par $\neg B_i E_j \varphi =_{df} \neg E_i(E_j\varphi \supset Bj)$. Finalement, l'obligation est modélisée par $B_i \neg E_j \varphi =_{df} E_i(\neg E_j \varphi \supset Bj)$, où *il est obligatoire que j fasse* φ (ou encore *j doit faire* φ) signifie qu'une autorité i s'assure que si j ne réalise pas φ, alors j subit quelque chose de mauvais. De fait, s'il est faux que φ est une description vraie du monde à la suite des actions posées par j, alors j subira une conséquence négative. Bref, les opérateurs déontiques sont définis de la manière suivante :

$$E_i(\neg E_j \varphi \supset Bj) \qquad (O)$$
$$E_i(E_j \varphi \supset Bj) \qquad (F)$$
$$\neg E_i(E_j \varphi \supset Bj) \qquad (P)$$

À l'aide de ces définitions, il est possible de montrer que $O\varphi$ implique $F\neg\varphi$. Voici la preuve avec les règles de KT en déduction naturelle (restreintes à un opérateur indexé).

Démonstration.

1	$E_i(\neg E_j \varphi \supset Bj)$	H
2	E_i	H
3	$\neg E_j \varphi \supset Bj$	EE_i 1,2
4	$E_j \neg \varphi$	H
5	$E_j \varphi$	H
6	$E_j \neg \varphi \supset \neg \varphi$	(**T**)
7	$\neg \varphi$	E$\supset$ 4,6
8	$E_j \varphi \supset \varphi$	(**T**)
9	φ	E$\supset$ 5,8
10	$\perp$	I$\perp$ 7,9
11	$E_j \varphi \supset \perp$	I$\supset$ 5-10
12	$\neg E_j \varphi$	df. $\neg$ 11
13	Bj	E$\supset$ 3,12
14	$E_j \neg \varphi \supset Bj$	I$\supset$ 4-13
15	$E_i(E_j \neg \varphi \supset Bj)$	IE_i 2-14

□

Cela dit, considérant que $\neg E_j \varphi \not\supset E_j \neg \varphi$, il est faux que F implique O, et de fait nous n'avons pas l'équivalence habituelle entre $O\varphi$ et $F\neg\varphi$. Dans la langue naturelle, cela signifie qu'il est vrai que si φ est obligatoire, alors $\neg\varphi$ est interdit, mais qu'il est faux que si $\neg\varphi$ est interdit, alors φ est obligatoire.

Par exemple, cela implique que s'il est obligatoire de respecter la loi, alors il est interdit de ne pas respecter la loi, mais qu'il est faux que s'il est interdit de voler, alors il est obligatoire de ne pas voler. Le second exemple peut être reformulé de la manière suivante : il est logiquement possible qu'il soit interdit de voler, mais qu'il ne soit pas obligatoire de ne pas voler. Notons que cela implique que $P\varphi$ ne soit pas équivalent à $\neg O\neg\varphi$. Soulignons aussi au passage que l'équivalence entre $\neg F\varphi$ et $P\varphi$ est triviale :

$$\neg E_i(E_j\varphi \supset Bj) \qquad (\neg F)$$
$$\neg E_i(E_j\varphi \supset Bj) \qquad (P)$$

Cette équivalence exemplifie la relation entre O et P. Considérant que $O\varphi$ implique $F\neg\varphi$ et que $F\neg\varphi$ implique $\neg P\neg\varphi$, cela nous donne $O\varphi$ implique $\neg P\neg\varphi$. La réciproque n'est cependant pas vraie. En ce sens, la permission ne se résume pas à ce dont la négation n'est pas obligatoire. En fait, la permission se résume à ce qui n'est pas explicitement interdit.

Du côté de la sémantique, Pörn ajoute trois conditions pour rendre compte de la cohérence d'un système normatif (entendu comme un ensemble de relations normatives entre des agents)[7] :

$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_i(E_j\varphi \supset Bj) \notin w \qquad (C1)$$
$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_k(E_j\varphi \supset Bj) \notin w \qquad (C2)$$
$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_i(E_k\neg\varphi \supset Bk) \in w \qquad (C3)$$

Autrement dit, (C1) stipule que si φ est obligatoire selon une autorité i, alors φ n'est pas interdit par cette même autorité. Plus encore, (C2) indique que si φ est obligatoire en vertu d'une autorité i, alors φ n'est pas interdit par une autre autorité k. Alors que la première clause vise à assurer une cohérence horizontale, la seconde assure une cohérence verticale. Finalement, la clause (C3) signifie que si i rend φ obligatoire pour j, alors i interdit à tout autre agent k de réaliser $\neg\varphi$.[8] Cette clause vise à assurer que j est en mesure de remplir ses obligations.

[7] Pörn utilise en fait une méta-méta-variable $\phi(j)$ pour les clauses C1 et C2, laquelle renvoie à un énoncé de la forme $E_j\psi$, où ψ est un énoncé de premier ordre (atomique ou complexe, donc qui ne contient pas d'opérateur E_i), ou encore à un énoncé complexe composé à partir d'énoncés de cette forme. En ce qui nous concerne, nous allons nous restreindre à la présentation de la clause pour les énoncés de la forme $E_j\psi$.

[8] Notons que (C3) est exprimée négativement chez Pörn par la formulation équivalente $E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ \neg E_i(E_k\neg\varphi \supset Bk) \notin w$.

Somme toute, Pörn propose d'analyser les énoncés normatifs comme des relations entre des agents et des propositions, un système normatif étant un ensemble d'agents liés par des relations normatives. Un énoncé d'obligation, par exemple, marque une relation normative entre une autorité i, un subordonné j et une proposition φ, où l'obligation pour j de faire φ signifie que i s'assure de punir j dans l'éventualité où il ne réalise pas son obligation φ. La logique de l'action, à elle seule, ne permet pas de rendre compte adéquatement de l'imputabilité et de l'intention puisque celle-ci n'inclut pas de modalités épistémique et temporelle.

Il s'agit là d'une limite des logiques de type STIT, lesquelles ne parviennent pas à faire la distinction entre le résultat de la réalisation de l'action, un état de choses qui prend place à la suite de l'action ou encore la conséquence de l'action. Les approches présentées dans les sections qui suivent tentent de pallier ces limites. Pacheco et Carmo cherchent à rendre compte de la notion de responsabilité en introduisant le concept d'*agir dans un rôle*, et Broersen introduit des modalités épistémiques et temporelles pour rendre compte de l'intentionnalité. Avant d'aborder ces approches, voyons en premier lieu celle de Horty.

John Horty

Le livre de Horty (2001), intitulé *Agency and Deontic Logic*, est à ce jour probablement l'ouvrage le plus complet qui propose une analyse de la logique déontique dans le cadre des logiques STIT. Ce dernier propose plutôt une analyse de la notion d'agentivité (*agency*) dans un cadre utilitariste. Nous avons jugé bon de présenter sa position considérant que celle-ci offre une excellente vue d'ensemble de l'intérêt d'utiliser une logique STIT pour l'analyse du discours normatif. Bien que l'approche de Horty vise l'analyse de ce que les agents *doivent faire* (*ought-to-do*), l'analyse qu'il propose prend place dans le cadre plus large de ce qui *doit être* (*ought-to-be*), ce qui se justifie par le fait que ce qu'un agent *doit faire* peut être identifié par ce qui *doit être* en vertu des actions de cet agent.

En ce sens, ce qu'un agent *doit faire* est subsumé par ce qui *doit être* dans le monde. Ainsi, l'objectif de Horty est de formuler l'idée que ce que les agents doivent faire dépend de ce qui doit être à la suite des actions des agents.

Bien que le système syntaxique ne soit pas explicitement défini dans l'ouvrage, nous allons tenter de présenter la syntaxe implicite à l'approche de Horty. Le langage $\mathcal{L}$ comprend un ensemble dénombrable de propositions atomiques ainsi que les connecteurs représentant la négation et la conjonction, à partir desquels la disjonction, l'implication et l'équivalence sont définies :

$$\mathcal{L} = \{(,), Prop, \neg, \wedge, \Box, [a\ stit], O, \odot\}$$

L'opérateur modal $\Box$ représente la nécessité historique. Considérant qu'un même scénario peut faire partie de plusieurs histoires différentes, la nécessité historique est à comprendre en termes de ce qui est *fixé pour toute histoire*. Autrement dit, ce qui est historiquement nécessairement vrai pour un scénario est ce qui est vrai pour toute histoire qui passe par ce scénario. Reformulons cette idée : soit une *histoire* une suite de scénarios $w_i, \ldots, w_j, \ldots, w_k$. La notion de nécessité historique signifie que ce qui est nécessairement vrai pour w_j correspond à ce qui est vrai pour toute histoire incluant w_j.

L'opérateur $[a\ stit]\varphi$, plus précisément $[\{a\}\ stit]\varphi$ comme nous le verrons bientôt, signifie *l'agent a réalise* φ, où φ est un état du monde qui advient à la suite des actions ou des choix de l'agent a.[9] Ainsi, lorsque $[a\ stit]\varphi$ est vrai, la description φ est vraie en vertu des actions de l'agent a, et de fait φ est vrai pour tous les scénarios accessibles à la suite des actions de a.

L'opérateur O représente *ce qui doit être*, alors que $\odot$ dénote *ce qui doit être fait*, où $\odot$ ne s'applique qu'aux propositions de la forme $[a\ stit]\varphi$ et signifie a doit réaliser φ.

En vertu de la clause sémantique (4) définie plus bas, $\Box$ peut être axiomatisé comme une modalité de type $S5$. En effet, le schéma (**T**) est trivial en vertu de la condition (4) et le schéma (**5**) peut aisément être démontré.

Proposition 1. $\Box$ est une modalité de type $S5$.

Démonstration. Supposons que $\models_{\mathcal{M}_{w,h}} \Diamond\varphi$, mais que $\models_{\mathcal{M}_{w,h}} \neg\Box\Diamond\varphi$. D'une part, cela signifie que $\not\models_{\mathcal{M}_{w,h}} \Box\neg\varphi$, et donc qu'il y a au moins une histoire $h' \in H_w$ pour laquelle $\models_{\mathcal{M}_{w,h'}} \varphi$. D'autre part, cela signifie que $\not\models_{\mathcal{M}_{w,h}} \Box\Diamond\varphi$, et donc qu'il y a un $k \in H_w$ tel que $\models_{\mathcal{M}_{w,k}} \neg\Diamond\varphi$. Or cela implique que $\models_{\mathcal{M}_{w,k}} \Box\neg\varphi$, et donc que $\models_{\mathcal{M}_{w,h'}} \neg\varphi$ pour tout $h' \in H_w$, ce qui contredit l'hypothèse de départ. Le schéma d'axiome (**5**) $\Diamond\varphi \supset \Box\Diamond\varphi$ est donc valide, ce qui, conjointement à (**T**), signifie que $\Box$ est une modalité axiomatisable par $S5$. □

Par ailleurs, $[a\ stit]$ est aussi un opérateur de type $S5$, alors que O est de type $KD45$.[10] L'opérateur $\odot$ quant à lui est une modalité de type KD. Notons aussi que $\odot$ n'est pas une modalité de type $KD45$ considérant

[9] À l'instar de Broersen, nous utilisons $[a\ stit]\varphi$ plutôt que $[a\ stit : \varphi]$ afin d'insister sur le fait que l'opérateur $[a\ stit]$ est une modalité.

[10] Soulignons qu'une modalité de type $KD45$ est plus faible qu'une de type $S5$.

que l'itération de l'opérateur $\odot$ n'est pas permise, $\odot$ étant une modalité de type *ougth-to-do*. Prenant tout cela en compte, l'approche de Horty peut être axiomatisée à l'aide des schémas d'axiomes suivants et du système modal K.

$$\Box\varphi \supset \varphi \qquad (\mathbf{T}_{\Box})$$
$$\Diamond\varphi \supset \Box\Diamond\varphi \qquad (\mathbf{5}_{\Box})$$
$$[a\ stit]\varphi \supset \varphi \qquad (\mathbf{T}_{\text{STIT}})$$
$$\neg[a\ stit]\neg\varphi \supset [a\ stit]\neg[a\ stit]\neg\varphi \qquad (\mathbf{5}_{\text{STIT}})$$
$$\neg(O\varphi \wedge O\neg\varphi) \qquad (\mathbf{D}_{O})$$
$$\neg O\neg\varphi \supset O\neg O\neg\varphi \qquad (\mathbf{5}_{O})$$
$$O\varphi \supset OO\varphi \qquad (\mathbf{4}_{O})$$
$$\neg(\odot[a\ stit]\varphi \wedge \odot[a\ stit]\neg\varphi) \qquad (\mathbf{D}_{\odot})$$

Le principe *devoir implique pouvoir* est représenté par les axiomes $\Diamond O$ et $\Diamond\odot$ pour les modalités de types *ought-to-be* et *ought-to-do* :

$$O\varphi \supset \Diamond\varphi \qquad (\Diamond O)$$
$$\odot\,[a\ stit]\varphi \supset \Diamond[a\ stit]\varphi \qquad (\Diamond\odot)$$

Finalement, la cohérence entre ce qui *doit être* et ce qui *doit être fait* est exprimée par le schéma d'axiome suivant :

$$\neg(\odot[a\ stit]\varphi \wedge O[a\ stit]\neg\varphi) \qquad (\mathbf{D}_{O\odot})$$

L'ensemble des énoncés bien formés peut être défini récursivement de la manière suivante :

$$\varphi := \ \mid p_i \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi \mid [a\ stit]\varphi \mid O\varphi$$

La dernière condition, spécifique à $\odot$, est que si φ est bien formé et ne contient pas de $\odot$, alors $\odot[a\ stit]\varphi$ est bien formé.

L'approche de Horty repose sur les logiques temporelles (*theory of branching time*) proposées par Prior (1967) et Thomason (1970, 2002). Ces logiques, qui reposent sur la notion d'*arbre temporel*, conçoivent le passé comme *déterminé*, mais l'avenir comme *incertain* (*indéterminé*).

Une structure temporelle $\mathcal{S} = \langle U, < \rangle$ est composée d'un univers U non vide de *moments*, lesquels peuvent aussi être considérés comme des scénarios. À l'intérieur de U, on trouve une relation d'ordre $<$ transitive et antiréflexive (pas réflexive). Dans une structure temporelle, une histoire h est une suite temporelle maximale de moments liés par la relation $<$, et H_w représente l'ensemble des histoires qui contiennent le moment w :

$$H_w = \{h : w \in h\}$$

En ce sens, H_w représente l'ensemble des suites d'événements possibles pouvant passer par w. Un modèle temporel $\mathcal{M} = \langle U, <, a \rangle$ est défini à partir de la structure $\mathcal{S}$, à laquelle on ajoute une fonction a qui détermine les paires moment/histoire auxquelles les propositions dans U appartiennent. Autrement dit :

$$a(\varphi) = \{(w, h) : \varphi \in w \text{ et } w \in h\}$$

La vérité dans le modèle est définie récursivement[11] :

$$\models_{\mathcal{M}_{(w,h)}} p \Leftrightarrow (w, h) \in a(p) \tag{1}$$

$$\models_{\mathcal{M}_{(w,h)}} \neg\varphi \Leftrightarrow \not\models_{\mathcal{M}_{(w,h)}} \varphi \tag{2}$$

$$\models_{\mathcal{M}_{(w,h)}} \varphi \wedge \psi \Leftrightarrow \models_{\mathcal{M}_{(w,h)}} \varphi \text{ et } \models_{\mathcal{M}_{(w,h)}} \psi \tag{3}$$

$$\models_{\mathcal{M}_{(w,h)}} \Box\varphi \Leftrightarrow \models_{\mathcal{M}_{(w,h')}} \varphi \text{ pour toute histoire } h' \in H_w \tag{4}$$

La vérité et la validité sont définies comme à l'habitude, où *valide* signifie vrai pour toute interprétation $\mathcal{M}_{(w,h)}$, c'est-à-dire pour tout modèle $\mathcal{M}$ indexé à une paire moment/histoire.

Jusqu'à présent, seul un modèle *temporel* a été défini. Afin de pouvoir définir un modèle STIT, Horty introduit l'ensemble :

$$\|\varphi\|_w^{\mathcal{M}} = \{h \in H_w : \models_{\mathcal{M}_{(w,h)}} \varphi\}$$

Cet ensemble contient toutes les histoires h où la proposition φ est vraie selon le modèle $\mathcal{M}$ au moment w. Clairement, nous avons $\|\varphi\|_w^{\mathcal{M}} \subseteq H_w$, c'est-à-dire que l'ensemble qui contient les histoires où φ est vrai pour w est un sous-ensemble de l'ensemble qui contient toutes les histoires qui passent par le moment w. Il est aussi évident que $\|\neg\varphi\|_w^{\mathcal{M}} = H_w - \|\varphi\|_w^{\mathcal{M}}$, à savoir que l'ensemble qui contient les histoires

[11] Le modèle tel qu'il est défini dans le livre inclut des clauses pour les opérateurs représentant le *passé* et le *futur*. Cela dit, ces opérateurs ne seront pas réutilisés dans l'ouvrage, et de fait ils ont été laissés de côté.

où φ est faux est le complément relatif à H_w de l'ensemble qui contient les histoires où φ est vrai. Un modèle STIT est obtenu en ajoutant un ensemble non vide d'agent(s) Ag et une fonction de choix Ch au modèle temporel :

$$\mathcal{M} = \langle U, <, Ag, Ch, a \rangle$$

La fonction de choix Ch_a^w est indexée à un moment w et à un agent (ou groupe d'agents) a. Informellement, la fonction Ch_a^w isole les histoires possibles qui s'offrent à a au moment w. En ce sens, Ch_a^w est un ensemble d'histoires. L'auteur utilise l'expression $Ch_a^w(h)$ pour référer à l'action que a pose au moment w afin d'obtenir l'histoire h. Autrement dit, alors que Ch_a^w offre l'ensemble des alternatives possibles, $Ch_a^w(h)$ dénote le choix de a au moment w. Il s'agit d'un ensemble d'histoires (une classe d'équivalence) sélectionnées par les actions de a. La clause sémantique ajoutée pour obtenir un modèle STIT est la suivante :

$$\models_{\mathcal{M}_{(w,h)}} [a\ stit]\varphi \Leftrightarrow\ Ch_a^w(h) \subseteq \|\varphi\|_w^{\mathcal{M}} \tag{5}$$

La justification de cette clause repose sur l'intuition que si un agent a réalise φ (où φ est une description du monde), alors φ est vrai à la suite des actions posées par a. En ce sens, la clause signifie que s'il est vrai que a réalise φ, alors l'ensemble des choix possibles que a peut faire au moment w et qui mènent à l'histoire h est un sous-ensemble de l'ensemble qui contient les histoires où la proposition φ est vraie au moment w.

Notons au passage que cela marque la principale distinction entre l'approche de Broersen, que nous verrons plus loin et qui repose sur une logique XSTIT, et la logique STIT. Dans le modèle proposé par Horty, l'effet de l'action est immédiat et prend place dans le même moment (dans le même scénario), alors que chez Broersen, l'effet de l'action prend place dans le scénario suivant. La principale distinction entre XSTIT et STIT est que XSTIT différencie entre les scénarios et les moments alors que pour STIT, il s'agit de la même chose.

Soulignons aussi que Horty utilise en fait l'opérateur [a *cstit*], qui représente le « Chellas STIT », inspiré des travaux de Chellas (1969) et auquel nous avons référé par [a *stit*]. Horty met en opposition l'opérateur [a *cstit*] avec un opérateur *délibératif* [a *dstit*], lequel diffère du premier de par l'ajout d'une condition négative sur la clause sémantique :

$$\models_{\mathcal{M}_{(w,h)}} [a\ dstit]\varphi \Leftrightarrow\ Ch_a^w(h) \subseteq \|\varphi\|_w^{\mathcal{M}} \text{ et } \|\varphi\|_w^{\mathcal{M}} \neq H_w$$

Cette condition négative vise à rendre compte du fait que les états du monde réalisés par un agent après délibération n'étaient pas *inévitables*. Autrement dit, l'opérateur [a *dstit*] est caractérisé par le fait que si a

réalise φ après délibération, alors φ n'est ni une tautologie ni une nécessité historique.[12] En effet, si $\|\varphi\|_w^{\mathcal{M}} = H_w$, alors φ est vrai pour toute histoire passant par le moment w, et donc par la clause (4) φ est une nécessité historique. Cela dit, après une discussion sur les vertus de chacun des opérateurs, Horty en vient à montrer que les opérateurs $[a\ cstit]$ et $[a\ dstit]$ sont inter définissables, et donc qu'il suffit de prendre le « Chellas STIT » comme primitif.

$$[a\ dstit]\varphi =_{df} [a\ cstit]\varphi \wedge \neg\Box\varphi \qquad \text{(df. } dstit\text{)}$$

En plus de développer un modèle qui permet de rendre compte de l'action d'un individu, Horty adapte le modèle afin de rendre compte des actions faites en groupe (*group agency*). Pour ce faire, il introduit dans le modèle un ensemble Se de fonctions de sélection s. Soit $s : Ag \longrightarrow H_w$ une fonction de sélection qui associe une histoire contenant w à un agent a. Une fonction de sélection s peut être vue comme un sous-ensemble de $Ag \times H_w$ qui contient des paires ordonnées $< a, h >$. Considérons l'ensemble de toutes les fonctions de sélection possibles. L'ensemble Se est tel que pour chaque paire $< a, h >$, h se trouve dans les choix accessibles à l'agent a. Formellement, l'ensemble est défini de la manière suivante :

$$Se_w = \{s : s(a) \in Ch_a^w\}$$

En ce sens, Se contient les fonctions de sélection telles que pour toute paire $< a, h > \in s$, $h \in Ch_a^w$. Informellement, les fonctions de sélection isolent les histoires compatibles avec les choix de tous les agents. En effet, cette fonction doit satisfaire la condition (I), qui stipule que l'intersection de tous les $s(a)$ pour tout $a \in Ag$ n'est pas vide :

$$\bigcap_{a \in Ag} s(a) \neq \varnothing \qquad \text{(I)}$$

Cette condition, qui marque l'indépendance des actions des agents, assure que les actions qui s'offrent aux agents se croisent minimalement à un moment (un scénario) donnée. À l'aide de cet ensemble, Horty redéfinit la fonction Ch_Γ^w relativement à un groupe d'agent $\Gamma \subseteq Ag$:

$$Ch_\Gamma^w = \{\bigcap_{a \in \Gamma} s(a) : s \in Se_w\}$$

[12] Notons qu'une tautologie *est* une nécessité historique par définition.

Cela fait, Horty modifie la clause sémantique pour l'opérateur [Γ *stit*] par (5') et propose la définition suivante pour l'opérateur STIT :

$$\models_{\mathcal{M}_{(w,h)}} [\Gamma\ stit]\varphi \Leftrightarrow\ Ch^w_\Gamma(h) \subseteq \|\varphi\|^{\mathcal{M}}_w \qquad (5')$$

$$[a\ stit]\varphi =_{df} [\{a\}\ stit]\varphi \qquad (\text{df. } [a\ stit])$$

En mots, (5') exprime qu'un groupe d'agents réalise φ à condition que l'ensemble des histoires accessibles au groupe soit sous-ensemble de l'ensemble qui contient les histoires où φ est vrai.

La conception que Horty se fait de l'ensemble des choix d'un groupe s'explique par le fait qu'il considère qu'un groupe d'agents se résume aux agents qui le composent. Soulignons que cette conception serait critiquée par Carmo et Pacheco (2001), lesquels soutiennent qu'un groupe ne se réduit pas aux agents individuels qui le composent.

Ayant exposé sa logique de l'action, l'auteur en vient à considérer les opérateurs déontiques[13] et à développer un modèle utilitariste en introduisant une fonction de valeur V et un ordre partiel $\sqsubseteq$. Le modèle utilitariste général est défini à partir d'un modèle STIT :

$$\mathcal{M} = \langle U, <, Ag, Ch, a, V, \sqsubseteq\rangle.$$

Notons que le modèle utilitariste ne présuppose aucune théorie morale autre qu'une forme d'utilitarisme en son sens large. En effet, l'auteur ne précise pas quels critères permettent de déterminer l'ordre partiel sur V, mais ne fait que présupposer une sémantique des préférences qui présume que certains scénarios sont meilleurs que d'autres, nonobstant ce que signifie la relation *être meilleur que*. Informellement, la fonction $V_w(h)$ détermine la valeur d'un moment w relativement à une histoire h. La relation d'ordre partiel $\sqsubseteq$ est une relation réflexive, transitive et antisymétrique, et représente informellement la *préférence*.[14] Le modèle doit satisfaire deux conditions :

1. $V : U \times \wp(U) \longrightarrow \mathbb{R}$ est une fonction qui associe à chaque paire (w, h) une valeur dans les nombres réels ;
2. l'assignation de valeur pour une histoire ne varie pas d'un moment à l'autre, c'est-à-dire que $V_w(h) = V_{w'}(h)$ pour tout moment $w \in h$ et $w' \in h$.

[13] Par souci d'économie, nous ne considérerons que les opérateurs de type *ought-to-be* et *ought-to-do*, laissant de côté les obligations conditionnelles analysées à l'aide d'une relation de préférence.

[14] Dans ce contexte, une relation antisymétrique est telle que si $w \sqsubseteq v$ et $v \sqsubseteq w$, alors $w = v$.

La première condition vise simplement à fournir un moyen de quantifier la valeur accordée à une histoire alors que la seconde signifie que la valeur d'une histoire repose sur l'histoire dans son ensemble. La clause sémantique pour O est représentée par (6) :

$$\begin{aligned} \models_{\mathcal{M}_{(w,h)}} O\varphi \Leftrightarrow\ & \text{il y a } h' \in H_w \text{ t.q.} \\ & 1.\ \models_{\mathcal{M}_{(w,h')}} \varphi \text{ et} \\ & 2.\ \models_{\mathcal{M}_{(w,h'')}} \varphi \text{ pour tout } h'' \in H_w \text{ t.q. } V_w(h') \sqsubseteq V_w(h'') \end{aligned} \tag{6}$$

Autrement dit, il est vrai que φ *devrait être* pour un moment w dans la mesure où φ est vrai dans au moins une histoire h' qui passe par w et où φ est vrai pour toute histoire avec une plus grande valeur que h'. La première sous-condition assure qu'il est possible de réaliser φ alors que la seconde implique que toute histoire où φ est vrai a plus de valeur qu'une où φ est faux.

Bien que Horty propose une analyse de ce qui *doit être fait* en termes de ce qui *doit être*, cet auteur en vient à définir un opérateur primitif $\odot$ pour représenter ce qui *doit être fait* plutôt que de le définir en fonction de O. La principale raison pour cette manœuvre est que la réduction de ce qui *doit être fait* à ce qui *doit être* pose problème dans un cadre utilitariste.

Horty introduit l'opérateur $\odot$, qui ne porte que sur des énoncés bien formés de type $[a\ stit]\varphi$ et signifie que l'agent a doit réaliser φ. Informellement, la clause sémantique pour cet opérateur est que a doit réaliser φ à condition que pour toute action accessible à a pour un moment w qui ne garantit pas la vérité de φ, il y a une autre action préférable qui la garantit. Afin de bien saisir cette clause, il faut introduire quelques notions. Considérons d'abord le cas de la préférence entre les propositions. Une proposition φ est dite *préférable* à une proposition ψ pour un moment w à condition que toute histoire où φ est vrai pour w soit préférable à toute histoire où ψ est vrai pour w. Formellement :

$$\begin{aligned} \psi \sqsubseteq \varphi \Leftrightarrow\ & \|\psi\|_w^{\mathcal{M}} \sqsubseteq \|\varphi\|_w^{\mathcal{M}} \\ \Leftrightarrow\ & V(h) \sqsubseteq V(h') \text{ pour tout } h \in \|\psi\|_w^{\mathcal{M}} \text{ et } h' \in \|\varphi\|_w^{\mathcal{M}} \end{aligned}$$

Soit $S_a^w = Ch_{Ag-\{a\}}^w$ l'ensemble qui comprend les choix accessibles à tous les agents excluant ceux accessibles à l'agent a. Chaque $s \in S_a^w$ représente une histoire accessible à un agent autre que a. L'expression $k \cap s$ représente l'intersection entre deux histoires à un moment w. Autrement dit, $k \cap s$ représente ce qui est vrai conjointement de k et de s au moment w. En ce sens, $k \cap s$ peut être vue comme une proposition complexe qui décrit ce qui est vrai de w à la fois dans l'histoire k et l'histoire s. Ayant cela en tête, Horty introduit une relation $\preceq$ de *dominance* entre les choix (actions) des agents.

Informellement, cette relation indique que l'histoire réalisée par un agent (en vertu de ses actions ou de ses choix) est *préférable* à une autre dans la mesure où les actions de l'agent mènent à une histoire avec une plus grande valeur peu importe les actions des autres agents. Autrement dit, si on considère que $k \cap s$ représente la suite des événements qui prend place après les actions de a avec celles d'un autre agent b, alors la relation de dominance indique que k est préférable à toute histoire k' accessible à a peu importe le choix de n'importe quel agent b.[15] Formellement, nous avons :

$$k \preceq k' \Leftrightarrow\ k \cap s \sqsubseteq k' \cap s \text{ pour tout } s \in S_a^w$$

Autrement dit, k domine k' à condition que k soit préférable à k' nonobstant les actions des autres agents. Ayant tout cela en tête, la clause sémantique pour $\odot$ s'exprime formellement par la condition suivante.[16]

$$\models_{\mathcal{M}_{(w,h)}} \odot[a\ stit]\varphi \Leftrightarrow \text{ pour tout } k \in Ch_a^w \text{ t.q. } k \notin \|\varphi\|_w^{\mathcal{M}} \qquad (7)$$
il y a $k' \in Ch_a^w$ t.q.
1. $k \prec k'$
2. $k' \in \|\varphi\|_w^{\mathcal{M}}$
3. $k'' \in \|\varphi\|_w^{\mathcal{M}}$ pour tout $k'' \in Ch_a^w$ t.q. $k' \preceq k''$

En mots, cela signifie qu'il est vrai que a doit réaliser φ au moment w à condition que pour toute histoire k où φ est faux, il y ait une histoire k' où φ est vrai strictement préférable à k et pour laquelle toute histoire k'' où φ est vrai est préférable à k'. Autrement dit, la première condition assure que k' est préférable (considérée conjointement aux suites d'événements possibles pouvant être engendrées par les autres agents) à toute histoire k où φ est faux. La seconde signifie que φ est vrai dans cette histoire, et la troisième implique que φ est vrai dans toute histoire qui y est préférable.

Notre présentation de Horty (2001) s'est finalement limitée à ses chapitres 2, 3 et 4 principalement en raison du fait qu'ils constituent le cœur de son approche. L'analyse qu'il propose dans les chapitres 5 et 6 vise certains problèmes spécifiques à la logique déontique, comme l'introduction d'un opérateur dyadique pour la formalisation des obligations qui surgissent dans

[15] On trouve une coquille dans la notation de Horty sur la relation de dominance. En effet, relativement à la condition (7), cet auteur écrit $k \subseteq \|\varphi\|_w^{\mathcal{M}}$ plutôt que $k \in \|\varphi\|_w^{\mathcal{M}}$. Or k est une histoire et $\|\varphi\|_w^{\mathcal{M}}$ est un ensemble d'histoire, et de fait k peut être un membre de $\|\varphi\|_w^{\mathcal{M}}$ mais n'est pas un sous-ensemble.

[16] Notons que Ch_a^w doit être fini afin de s'assurer qu'il y ait au moins une suite d'événement préférable.

des contextes de violations et la notion d'obligation de groupe. Le chapitre 7 parle des considérations relatives aux stratégies employées en vue de certaines fins, prenant en compte la capacité des agents à accomplir certaines actions.[17] L'ouvrage de Horty occupe une place considérable : son langage est riche et permet de rendre compte de la complexité du discours normatif quotidien. Outre l'analyse de Horty, qui s'intéresse plutôt au discours éthique, d'autres approches de type STIT se sont penchées explicitement sur la formalisation du discours légal. Voyons-en maintenant deux exemples.

Olga Pacheco et José Carmo

Certaines approches utilisant les logiques STIT cherchent à modéliser la structure d'un NMAS par la notion de *rôle*. Pacheco et Carmo (2003), par exemple, introduisent cette notion à partir de l'analyse de la loi.

Le premier élément qu'ils mettent de l'avant est qu'un agent n'est pas nécessairement une personne *physique* et qu'il existe des agents *artificiels*. Cela est mis en évidence par le fait que, d'un point de vue légal, certaines institutions (ou organisations) possèdent la *personnalité juridique* (p. ex. une entreprise), et sont donc sujettes à des obligations, des droits, des permissions, etc.

Notons que cette conception de l'*agent* n'est pas surprenante d'un point de vue informatique, perspective dans laquelle s'insère l'approche de Pacheco et Carmo : un agent peut très bien être un programme qui pose des actions et dont le comportement est contraint par des règles (des normes, des instructions). Cela dit, un agent peut être artificiel, mais aussi être *complexe*, voire *collectif*. Au même titre qu'un programme informatique peut être composé de sous-programmes, une organisation qui possède la personnalité juridique est composée de membres, et peut-être de sous-organisations.

L'analyse du discours légal de Pacheco et Carmo met en lumière certaines caractéristiques d'une organisation, qui est un agent collectif :

1. l'organisation est un agent artificiel qui agit *indirectement* par le truchement des actions posées par les agents individuels qui la composent ;
2. elle possède un ensemble de rôles déterminés, un ensemble de normes qui caractérisent ces rôles et une structure stable (indépendante des agents qui remplissent les rôles) ;
3. certains rôles peuvent être délégués, c'est-à-dire qu'un agent peut représenter un autre agent.

[17] Voir aussi Kooi et Tamminga (2006) pour l'analyse des actions de groupe.

L'élément fondamental que soulignent les auteurs et sur lequel repose leur approche est qu'un agent agit toujours dans un certain *rôle*. Cela se comprend notamment en raison du fait que le comportement d'un agent est évalué en fonction du *rôle* dans lequel l'agent a posé une action (Pacheco et Santos 2004).

Exemple
Un policier qui utilise son arme dans le cadre de ses fonctions sera jugé en tant que policier, et non en tant que simple citoyen.

En un mot, tous les rôles ne jouissent pas des mêmes privilèges, des mêmes droits, et ne subissent pas les mêmes restrictions.

Le caractère innovateur de Pacheco et Carmo est de modifier la logique de type STIT, inspirée des travaux de Pörn, afin d'indexer l'action d'un agent à un rôle. L'opérateur $E_{a:r}\varphi$ se lit « l'agent a, qui agit dans le rôle r, réalise φ ». Dans le même ordre d'idées, ils indexent les opérateurs $O_{a:r}$, $P_{a:r}$ et $F_{a:r}$ à des agents qui agissent dans certains rôles.

Notons que les opérateurs sont une abréviation de $OE_{a:r}$, $PE_{a:r}$ et $FE_{a:r}$, lesquels se lisent respectivement « il est obligatoire que a dans le rôle r réalise... », « il est permis que a dans le rôle r réalise... » et « il est interdit que a dans le rôle r réalise... ».

Contrairement aux systèmes standard, l'opérateur $P_{a:r}$ est pris en tant que primitif et l'opérateur $F_{a:r}$ se définit de la manière suivante :

$$F_{a:r}\varphi =_{df} \neg P_{a:r}\varphi \qquad \text{(df. } F_{a:r}\text{)}$$

L'argument en faveur de ne pas définir l'interdiction et la permissions en matière d'obligation repose sur le fait que les énoncés $OE_{a:r}\neg\varphi$ et $O\neg E_{a:r}\varphi$ ne sont pas équivalents (Carmo et Pacheco 2000).

D'une part, le premier énoncé signifie que l'agent a dans le rôle r a l'obligation de réaliser l'état de choses décrit par la proposition $\neg\varphi$. Autrement dit, $\neg\varphi$ est le résultat des actions de a. D'autre part, $O\neg E_{a:r}\varphi$ signifie qu'il est obligatoire que a dans le rôle r ne réalise pas l'état de choses décrit par φ, ce qui peut résulter autant d'une omission que d'une abstention. De fait, $O_{a:r}$ et $P_{a:r}$ sont tous deux pris en tant que primitifs afin d'avoir :

$$O_{a:r}\neg\varphi \equiv OE_{a:r}\neg\varphi \not\equiv O\neg E_{a:r}\varphi \equiv F_{a:r}\varphi$$

Cela est nécessaire puisque, lorsqu'elles sont combinées avec un opérateur STIT, l'obligation et l'interdiction ne signifient pas la même chose.

La syntaxe de leur approche est définie de la manière suivante. Le langage $\mathcal{L}_{\mathcal{DA}}$ contient un nombre fini $(s_1, ..., s_n)$ de types de variables et de constantes, incluant Ag (les variables/constantes d'agents), R (les variables/constantes de rôle) et AgR (les variables/constantes d'agents dans des rôles). Ces types contiennent quant à eux un nombre dénombrable de variables et au moins une constante. De plus, $\mathcal{L}_{\mathcal{DA}}$ contient un ensemble P de prédicats, incluant rg (un générateur de rôle), $is - rg$ un prédicat de qualification, où $is - rg(a, t_1, ..., t_n) = qual(a : rg(t_1, ..., t_n))$ et signifie que a est qualifié pour le rôle $rg(t_1, ..., t_n)$. De façon équivalente, cela signifie que le rôle de a est décrit par $r(t_1, ..., t_n)$. Notons que rg peut référer à différents rôles, auquel cas rg est substitué par un mot décrivant le rôle de manière plus appropriée. Par exemple, $is - membre(k, i)$ pourrait signifier que k est un membre de l'organisation i. En ce sens, un générateur de rôle associe un prédicat à un rôle.

Afin de définir l'ensemble des énoncés bien formés, les auteurs doivent d'abord définir ce qu'est un *terme* :

1. Si rg est de type $(s_1, ..., s_n, R)$ et que $t_1, ..., t_n$ sont des variables de type $(s_1, ..., s_n)$, alors $rg(t_1, ..., t_n)$ est un terme (de type R).
2. Si t est un terme de type Ag et $rg(t_1, ..., t_n)$ un terme de type R, alors $t : rg(t_1, ..., t_n)$ est un terme de type AgR.

En un mot, un terme d'un certain type est construit à partir de variables ou de constantes du même type. De manière plus générale, les constantes et les variables sont des termes. Cela nous amène à la définition de l'ensemble des énoncés bien formés :

1. si $p \in P$ et $t_a, ..., t_n$ sont des termes du même type que p, alors $p(t_1, ..., t_n) \in EBF$;
2. si $\varphi, \psi \in EBF$, alors $\neg\varphi, \varphi \wedge \psi \in EBF$;
3. si $\varphi \in EBF$ et x^s est une variable de type s, alors $(\forall x^s)\varphi \in EBF$;
4. si $\varphi \in EBF$ et $a : r$ est un terme de type AgR, alors $E_{a:r}\varphi$, $O_{a:r}\varphi$, $P_{a:r}\varphi \in EBF$.

Les autres connecteurs logiques $\vee$, $\rightarrow$, $\leftrightarrow$ et $\exists$ sont définis usuellement à partir de $\neg$, $\wedge$ et $\forall$. Considérant qu'un agent agit à l'intérieur d'un rôle donné, une obligation φ dans un rôle r signifie que tout agent a qualifié au rôle r doit réaliser φ, c'est-à-dire que $O_r\varphi$ signifie :

$$(\forall x)(qual(x : r) \rightarrow O_{x:r}\varphi)$$

Similairement, si φ est permis pour un rôle R, alors φ est permis à tout agent qualifié pour le rôle r, et donc $P_r\varphi$ signifie :

$$(\forall x)(qual(x : r) \rightarrow P_{x:r}\varphi)$$

Finalement, les auteurs présentent leur système sous forme axiomatique, dont les axiomes et règles d'inférence se divisent en trois groupes[18], sans compter les axiomes de la logique propositionnelle classique :

1. Les axiomes et règles d'inférence de premier ordre :

$$\vdash \varphi \Rightarrow \vdash (\forall x)\varphi \qquad (\mathsf{Gen})$$
$$(\forall x)(\varphi \to \psi) \to ((\forall x)\varphi \to (\forall x)\psi) \qquad (\mathrm{A1})$$
$$\varphi \to (\forall x)\varphi \qquad (x \text{ n'est pas libre pour } \varphi) \qquad (\mathrm{A2})$$
$$(\forall x)\varphi \to \varphi(x/t) \qquad (\mathrm{A3})$$

$\varphi(x/t)$ est obtenu en remplaçant les occurrences libres de x par t.

Notons que (Gen) est la règle d'inférence qu'on retrouve dans la présentation classique du calcul de premier ordre.[19] (A1) est un axiome de distributivité et (A3) est aussi un axiome du calcul des prédicats. (A1) conjointement à (A2) permet de dériver l'axiome usuel du calcul des prédicats (où φ ne contient pas d'occurence libre de x).

$$(\forall x)(\varphi \to \psi) \to (\varphi \to (\forall x)\psi)$$

En plus des axiomes de premier ordre, on trouve les axiomes propres à la logique de l'action, qui forment le second groupe.

2. Les axiomes et règles d'inférence STIT :

$$E_{a:r}\varphi \to \varphi \qquad (\mathrm{A4})$$
$$(E_{a:r}\varphi \wedge E_{a:r}\psi) \to E_{a:r}(\varphi \wedge \psi) \qquad (\mathrm{A5})$$
$$E_{a:r}\varphi \to qual(x : r) \qquad (\mathrm{A6})$$
$$(\forall x)qual(x : itself) \qquad (\mathrm{A7})$$
$$\vdash \varphi \leftrightarrow \psi \Rightarrow \vdash E_{a:r}\varphi \leftrightarrow E_{a:r}\psi \qquad (\mathsf{RE}_E)$$

L'axiome (A4) correspond à l'axiome (**T**), qui représente la nécessité en logique aléthique. En un mot, cet axiome signifie que s'il est vrai que a dans le rôle r réalise φ, alors φ est vrai. L'axiome (A5) est un axiome d'agrégation qui stipule que si a réalise φ et réalise ψ (dans le rôle r), alors a réalise la conjonction φ et ψ.[20] Il s'agit de l'équivalent du schéma d'axiome (C).

[18] Voir Pacheco et Santos (2004) pour une présentation plus concise.

[19] Voir par exemple Kleene (1952) ou Mendelson (2009).

[20] L'auteur souligne en conclusion que ce système gagnerait à être augmenté à l'aide d'une logique temporelle. Soulignons que, en effet, sans temporalité, cet axiome peut amener à des conséquences étranges. Imaginons, par exemple, qu'à un temps t_1, a ne chante pas ($E_{a:r}\neg c$) et qu'à un temps t_2, a chante ($E_{a:r}c$). Dans une telle situation, l'agrégation permet de dire que a chante et ne chante pas !

L'axiome (A6) quant à lui précise que si un agent réalise une action dans le cadre d'un certain rôle, alors cet agent est qualifié pour le rôle. Cet axiome est, en quelque sorte, un axiome de légitimité : un agent ne peut pas réaliser une action dans le cadre d'un rôle qu'il ne joue pas. En ce sens, un agent ne peut pas agir de manière illégitime, voire frauduleuse : il est impossible que Pierre donne une contravention à Paul dans le cadre de son rôle de policier, mais que Pierre ne soit pas un policier qualifié. Autrement dit, un agent ne peut pas faire semblant d'agir dans un rôle.[21]

L'axiome (A7) signifie que tout agent est qualifié à agir dans son propre rôle, où $itself$ est une forme de rôle sans convention.[22] Finalement, la règle d'inférence (RE_E) est une règle de substitution qui implique que si φ et ψ sont des propositions équivalentes dans $\mathcal{L}_{\mathcal{DA}}$, alors $E_{a:r}\varphi$ et $E_{a:r}\psi$ le sont aussi.

En dernier lieu, l'axiomatisation comprend les règles qui gouvernent les opérateurs déontiques et leurs relations avec l'opérateur STIT.

3. Les axiomes et règles d'inférence déontiques :

$$(O_{a:r}\varphi \wedge O_{a:r}\psi) \rightarrow O_{a:r}(\varphi \wedge \psi) \tag{A8}$$
$$O_{a:r}\varphi \rightarrow P_{a:r}\varphi \tag{A9}$$
$$O_{a:r}\varphi \rightarrow \neg P_{a:r}\neg\varphi \tag{A10}$$
$$(O_{a:r}\varphi \wedge P_{a:r}\psi) \rightarrow P_{a:r}(\varphi \wedge \psi) \tag{A11}$$
$$\vdash \varphi \leftrightarrow \psi \Rightarrow \vdash O_{a:r}\varphi \leftrightarrow O_{a:r}\psi \tag{RE_O}$$
$$\vdash \varphi \rightarrow \psi \Rightarrow \vdash P_{a:r}\varphi \rightarrow P_{a:r}\psi \tag{RM_P}$$
$$\vdash E_{a_1:r_1}\varphi \rightarrow E_{a_2:r_2}\psi \Rightarrow \vdash P_{a_1:r_1}\varphi \rightarrow P_{a_2:r_2}\psi \tag{RM_{EP}}$$

L'axiome (A8) est, au même titre que (A5), un axiome d'agrégation pour les obligations. L'axiome (A9) est un analogue de l'axiome (**D**) des logiques déontiques standard et stipule que ce qui est obligatoire est aussi permis. L'axiome (A10) quant à lui vise à définir la relation entre $O_{a:r}$ et $P_{a:r}$. Considérant que $P_{a:r}$ est pris en tant que primitif (donc, contrairement à la définition usuelle, $\neg O_{a:r}\neg \neq_{df} P_{a:r}$), l'axiome indique que si a doit, dans le rôle r, faire φ, alors il est faux qu'il est permis à a (dans le rôle r) de ne pas faire φ. Le dernier axiome (A11) assure une certaine cohésion entre les obligations et les permissions qui caractérisent un rôle. En effet, il assure que ce qui est permis peut être fait en conjonction avec ce qui est obligatoire.

[21] Il s'agit d'une limite de leur approche, qui ne peut rendre compte des agents qui agissent de manière illégitime.

[22] Voir Pacheco et Carmo (2003) pour l'analyse des relations entre le rôle $itself$ et les autres rôles.

En ce qui concerne les règles d'inférence, (RE_O) est une règle de substitution à l'instar de (RE_E), alors que (RM_P) et (RM_{EP}) sont des analogues de la règle dérivée (RM). Axiomatisés ainsi, l'obligation et l'action opèrent dans des systèmes classiques alors que la permission s'exerce dans un système monotone.

Du côté de la sémantique, Carmo et Pacheco définissent un modèle $\mathcal{M} = \langle W, I, f_e, f_o \rangle$ à l'aide de la sémantique des mondes possibles, où $W \neq \varnothing$ est un ensemble de scénarios possibles.[23] I est une interprétation du langage de premier ordre pour chaque $w \in W$ telle que :

1. le domaine associé à chaque type s est le même pour tout scénario et est dénoté par $D_s = I(w, s)$;
2. si g est une fonction de type $(s_1, ..., s_n, R)$, alors $I(s, g) = D_{s_1} \times ... \times D_{s_n} \longrightarrow D_R$;
3. si p est un prédicat de type $(s_1, ..., s_n)$, alors $I(w, s) \subseteq D_{s_1} \times ... \times D_{s_n}$;
4. si c est une constante, alors $I(w, c) = I(w_1, c)$;
5. $I(w, itself) = D_{Ag}$.

Autrement dit, la première condition stipule que le domaine est stable, donc qu'il ne varie pas d'un scénario à l'autre. Cette contrainte assure que les scénarios possèdent tous les mêmes types de prédicats avec des extensions équivalentes. La seconde condition ne fait qu'assurer que le domaine d'un rôle est du même type que les parties dont il est composé. La troisième implique que le domaine qui contient un prédicat d'un certain type est inclus dans l'ensemble qui contient les domaines de chacun des types. La quatrième clause signifie que les individus restent les mêmes d'un scénario à l'autre, et finalement la cinquième condition stipule que le domaine des objets qui peuvent remplir le rôle *itself* pour un scénario donné est l'ensemble de tous les agents.

À cela s'ajoute une définition récursive des domaines de type R et AgR :

1. si r est un rôle de type $(s_1, ..., s_n, R)$, que $x_j \in D_{s_j}$(où $j = 1, ..., n$) et que $a \in D_{Ag}$, alors $r(x_1, ..., x_n) \in D_R$ et $a : r(x_1, ..., x_n) \in D_{AgR}$;

[23] Pacheco et Santos (2004) utilisent le système syntaxique développé par Pacheco et Carmo (2003). Cela dit, dans les deux cas, les auteurs renvoient à la sémantique développée dans Carmo et Pacheco (2000, 2001). Ici, nous proposons une synthèse de l'approche proposée dans l'article de 2003, utilisant la sémantique des articles de 2000 et 2001. Cela dit, la position des auteurs a évolué depuis, et la sémantique qu'ils ont proposée ne rend pas pleinement compte de la syntaxe développée par la suite. Cette dichotomie vient principalement du fait que leur système a évolué en fonction de leurs objectifs. Nous renvoyons le lecteur au dernier paragraphe de cette section.

2. si r est un rôle de type $(AgR, s_1, ..., s_n, R)$, $x_j \in D_{s_j}$(où $j = 1, ..., n$), $a \in D_{Ag}$ et $b \in D_{AgR}$, alors $r(b, x_1, ..., x_n) \in D_R$ et $a : r(b, x_1, ..., x_n) \in D_{AgR}$.

La première condition implique que si r est un rôle dont le type est s_j, que le prédicat x_j est du même type et que a est un agent, alors le rôle caractérisé par le prédicat est un élément du domaine, de même que la combinaison *agent : rôle*. La seconde stipule que les rôles peuvent se combiner. En ce sens, si r est un rôle d'un type s_j et qu'un prédicat x_j est du même type, alors si b est un rôle que peut prendre un agent, alors le rôle caractérisé par b, x_j est aussi élément du domaine.

Les deux autres fonctions servent à isoler les scénarios dans lesquels les propositions $E_{a:r}B$ et $O_{a:r}B$ sont vraies. D'une part, f_e est définie de façon à ce que si $a \in D_{Ag}$ et que Z est un sous-ensemble de W, alors $f_e(a : r, Z)$ dénote l'ensemble qui contient tous les scénarios où a réalise Z dans le rôle r, Z étant une description du scénario qui est vraie à la suite des actions posées par a.

$$f_e : (D_{Ag} \cup D_{AgR}) \times 2^W \longrightarrow 2^W$$

D'autre part, f_o isole les scénarios où un état de choses Z est obligatoire : si $Z \subseteq W$, alors $f_o(Z)$ est l'ensemble qui contient tous les scénarios où Z est obligatoire. Cela est similaire aux clauses sémantiques proposées par Chellas (1974) et Jones (1991) pour rendre compte des obligations conditionnelles, ce qui s'explique par le fait que le modèle est minimal. Notons au passage qu'en vertu de cette définition, cette approche s'insère dans le cadre des logiques de type *ought-to-be* : l'agent n'a pas l'obligation de *faire* une action, mais plutôt de faire advenir un *état de choses* (dans le monde).

$$f_o : 2^W \longrightarrow 2^W$$

La vérité dans $\mathcal{M}$ est définie de manière usuelle. Carmo et Pacheco introduisent une fonction d'évaluation v qui assigne pour tout scénario et à tout terme t de type s un objet membre du domaine D_s. La fonction est définie de manière récursive et attribue ultimement aux constantes l'objet qui leur correspond à l'intérieur du domaine. Les conditions de vérité sont les suivantes :

$\mathcal{M} \models_{w,v} p(t_1, ..., t_n) \Leftrightarrow v(w, t_1), ..., v(w, t_n) \in I(w, p)$ (1)

$\mathcal{M} \models_{w,v} \neg\varphi \Leftrightarrow \mathcal{M} \not\models_{w,v} \varphi$ (2)

$\mathcal{M} \models_{w,v} \varphi \wedge \psi \Leftrightarrow \mathcal{M} \models_{w,v} \varphi$ et $\mathcal{M} \models_{w,v} \psi$ (3)

$\mathcal{M} \models_{w,v} (\forall x^s)\varphi \Leftrightarrow \mathcal{M} \models_{w,v_1} \varphi$ pour tout v_1 qui assigne les mêmes valeurs que v aux variables différentes de x^s, y compris les variables d'autres types (4)

$$\mathcal{M} \models_{w,v} O\varphi \Leftrightarrow\ w \in f_o(\|\varphi\|_v) \text{ où } \|\varphi\|_v = \{w : \mathcal{M} \models_{w,v} \varphi\} \tag{5}$$

Finalement, si t est un terme de type Ag ou AgR, alors

$$\mathcal{M} \models_{w,v} E_t\varphi \Leftrightarrow\ w \in f_{e_{v(t)}}(\|\varphi\|_v) \tag{6}$$

La clause (1) signifie qu'une proposition de premier ordre est vraie dans la mesure où les objets dans la portée du prédicat sont dans l'extension du concept pour un scénario donné. Les conditions (2)–(3) sont les conditions habituelles pour la négation et la conjonction (classique). (4) assure que si une quantification universelle est vraie, alors la proposition dans la portée du quantificateur est vraie pour toute évaluation qui assigne les mêmes valeurs aux variables des différents types. En d'autres termes, (4) assure une certaine stabilité, qui garde la quantification dans les scénarios qui ont le même domaine et dont les prédicats identiques ont des extensions identiques.

Notons que la quantification est indexée à un type. Elle ne porte donc pas sur le domaine en entier, mais seulement sur la portion D_s. La condition (5) stipule qu'une proposition de la forme « obligatoirement φ » est vraie pour un scénario w à condition que ce scénario soit membre de l'ensemble qui contient tous les scénarios où l'ensemble qui contient tous les scénarios pour lesquels φ est vrai selon l'évaluation v est obligatoire.

Cette dernière condition mérite d'être reformulée : l'ensemble $\|\varphi\|_v$ contient tous les scénarios où la proposition φ est vraie pour l'évaluation v. L'énoncé $O\varphi$ est vrai pour w et l'évaluation v à condition que tout scénario x où φ est vrai pour la même évaluation est obligatoire. Autrement dit, $f_o(\|\varphi\|_v)$ dénote les scénarios obligatoires pour lesquels φ est vrai (selon l'évaluation v), et $O\varphi$ est vrai pour w à condition que w fasse partie de cet ensemble. Cette clause présume l'idéalité du modèle, liant la vérité de la proposition normative $O\varphi$ à celle de la proposition descriptive φ dans la portée de l'opérateur déontique.

En dernier lieu, la clause (6) rend compte de l'axiome (A4) et signifie que si la proposition « x réalise φ dans la rôle r » est vraie pour w, alors w est membre de l'ensemble qui contient tous les scénarios où φ est vrai (et où l'état de choses φ résulte des actions de x dans le rôle r).

Une proposition est dite vraie pour un modèle $\mathcal{M}$ à condition qu'elle soit vraie pour tout scénario et toute interprétation dans $\mathcal{M}$, et elle est dite valide lorsqu'elle est vraie pour tout modèle.

Afin de rendre compte de la structure représentée par les axiomes, les auteurs ajoutent des clauses supplémentaires au modèle sémantique. D'abord, Carmo et Pacheco (2001) ajoutent la clause suivante, où $X \subseteq W$:

$$f_{e_{a:r}}(X) \subseteq f_{e_a}(X) \subseteq X$$

En conjonction avec la clause (6) susmentionnée, cela vise à assurer la validité de l'axiome (A4). En mots, cela signifie que l'ensemble qui contient les scénarios où il est vrai qu'un agent réalise X dans un rôle est sous-ensemble de l'ensemble où a réalise X, lequel est sous-ensemble de X (présupposé vrai). Autrement dit, s'il est vrai qu'un agent (dans un rôle ou non) réalise X, alors X est vrai : mais ce n'est pas parce que X est vrai que l'agent a réalisé X.

Les deux conditions suivantes visent à valider respectivement les axiomes (A5) et (A8). Notons que la seconde clause n'est pas retenue dans le texte de 2001 considérant qu'à ce moment, les auteurs rejetaient l'agrégation pour l'obligation. De plus, soulignons aussi que nous avons ajouté l'indexation de f_o à b afin d'être conforme au texte de 2003 :

$$f_{e_b}(X) \cap f_{e_b}(Y) \subseteq f_{e_b}(X \cap Y)$$
$$f_{o_b}(X) \cap f_{o_b}(Y) \subseteq f_o(X \cap Y)$$

Dans chacun des cas, la clause est conforme à la conception ensembliste de la conjonction : l'ensemble qui contient les scénarios où φ est vrai et ψ est vrai est inclus dans l'ensemble qui contient les scénarios où φ et ψ est vrai. Cela dit, considérant que $O_{a:r}$ est une abréviation pour $OE_{a:r}$, la clause sémantique pour l'axiome (A8) est mieux représentée par :

$$f_o(f_{e_b}(X)) \cap f_o(f_{e_b}(Y)) \subseteq f_o(f_{e_b}(X \cap Y))$$

L'axiome (A9), analogue à (**D**), va de pair avec la clause suivante :

$$f_{o_b}(X) \cap f_{o_b}(W - X) = \varnothing$$

Cela signifie que l'ensemble qui contient les scénarios où X est obligatoire ne partage aucun membre avec celui qui contient les scénarios où le complément de X dans W est obligatoire. En ce sens, les scénarios où X est vrai qui sont obligatoires ne peuvent pas se retrouver à l'intérieur d'autres scénarios obligatoires où X est faux. L'axiome (A11) quant à lui est représenté sémantiquement par la clause suivante :

$$(f_o(f_{e_b}(X)) - f_o(W - f_{e_b}(Y))) \cap f_o(W - f_{e_b}(X \cap Y)) = \varnothing$$

Cette clause mérite quelques explications. Allons-y étape par étape :

1. $W - f_{e_b}(Y)$ est le complément de l'ensemble qui contient les scénarios où Y (réalisé par b) est vrai par rapport à W. Il s'agit donc de l'ensemble qui contient les scénarios où Y (réalisé par b) est faux ;
2. $f_o(W - f_{e_b}(Y))$ est l'ensemble qui contient les scénarios où « Y (réalisé par b) est faux » est obligatoire ;

3. $(f_o(f_{e_b}(X)) - f_o(W - f_{e_b}(Y))$ est le complément de l'ensemble qui contient les scénarios où « Y (réalisé par b) est faux » est obligatoire par rapport à l'ensemble qui contient les scénarios où « X (réalisé par b) est vrai » est obligatoire. De fait, l'ensemble contient « X (réalisé par b) est vrai » est obligatoire et « Y (réalisé par b) est vrai » est obligatoire ;
4. $f_o(W - f_{e_b}(X \cap Y))$ est l'ensemble contenant les scénarios obligatoires où soit X (réalisé par b) est faux soit Y (réalisé par b) est faux.

En ce sens, l'ensemble qui contient « X (réalisé par b) est vrai » est obligatoire et « Y (réalisé par b) est vrai » est obligatoire ne partage aucun membre avec celui qui contient les scénarios où « X et Y » (réalisé par b) est faux est obligatoire.

Nous n'avons pas encore discuté de tous les axiomes proposés. Le cas des axiomes et règles d'inférence de premier ordre est trivial, et de fait nous n'en discuterons pas davantage. En ce qui concerne les axiomes STIT, nous n'avons pas encore discuté de (A6), (A7) et (RE_E). Carmo et Pacheco mentionnent au passage (en une ligne) que l'axiome (A7) est d'emblée valide dans leur sémantique, la raison étant que sinon, il faudrait un scénario w et une évaluation v tels que $is - itself(x)$ est faux, et donc que x n'est pas membre de l'extension du concept *être x* (où *être x* regroupe les choses qui satisfont à la définition de x).[24]

L'axiome (A6) quant à lui est validé par l'ajout de la condition suivante :

$$w \in f_{e_{a:r}}(X) \Rightarrow \ qual(w, a : r)$$

Si w est un scénario membre de l'ensemble contenant les scénarios où a réalise X dans le rôle r, alors a est qualifié pour le rôle r dans w.[25] La validité de la règle (RE_E) est triviale, à présumer la complétude (dont la preuve n'est pas fournie par les auteurs), considérant la clause de vérité (6).

Cela nous amène aux axiomes et règles d'inférence déontiques. Jusqu'à présent, nous avons traité des axiomes (A8), (A9) et (A11). La règle (RE_O) se comprend dans la même mesure que (RE_E). Les autres cas sont moins évidents et quelques remarques s'imposent. Le système syntaxique est présenté dans l'article de Pacheco et Carmo (2003), où les auteurs justifient le choix des axiomes et des règles d'inférence, et analysent en profondeur les notions d'*action commune*, de *contrat*, d'*organisation* et de *rôle*. L'article de Pacheco et Santos (2004) reprend la syntaxe proposée par Pacheco et Carmo (2003) afin d'analyser la délégation de rôle, d'obligation et de responsabilité au sein d'une organisation (un NMAS).

[24] On ne trouve d'ailleurs pas plus d'explication dans Carmo et Pacheco (2000).
[25] Le lecteur est renvoyé au texte pour la définition récursive de *qual*.

De l'autre côté, la sémantique est présentée dans les articles de Carmo et Pacheco (2000, 2001), où le texte de 2001 est en fait une reformulation de celui de 2000. Cela dit, Pacheco et Santos (2004) renvoient à Pacheco et Carmo (2003) et Carmo et Pacheco (2001) pour la discussion sur les opérateurs déontiques primitifs, où Pacheco et Carmo (2003) ne font que résumer l'analyse proposée dans le texte de 2001.

Par surcroît, Pacheco et Carmo (2003) renvoient aux textes de Carmo et Pacheco (2000, 2001) pour les détails relatifs à la sémantique. Or le fait est que la syntaxe proposée dans les textes de 2004 et 2003 ne se trouve pas complètement dans ceux de 2001 et 2000. Autrement dit, les considérations sémantiques proposées dans le texte de 2001 ne couvrent pas la totalité du système syntaxique proposé en 2003. Le lecteur ne trouvera donc pas la clause sémantique de (A10) et la discussion proposée dans Carmo et Pacheco (2001) ne traite pas des règles (RM_P) et (RM_{EP}).[26]

Somme toute, Pacheco Carmo proposent une façon intéressante de combiner les logiques de premier ordre, STIT et déontique. Cette approche vise à rendre compte de phénomènes complexes comme la délégation d'obligations au sein d'une organisation, l'interaction entre les obligations, les agents et les rôles au sein d'une organisation et, de façon plus générale, des obligations collectives. Un aspect original de leur approche est qu'ils ne considèrent pas que l'*institution* se réduit aux agents individuels qui la composent (Carmo et Pacheco 2001).[27] L'institution va au-delà de ses membres. Conformément à leur analyse de la loi, ils introduisent la notion d'agent artificiel pour rendre compte des organisations en tant qu'agents complexes, caractérisés par les rôles qui les composent et par les normes qui régissent ces rôles.

Jan Broersen

Conformément à la tradition engendrée par Belnap et Perloff (1988), la logique STIT a pour objectif la modélisation des NMAS, et de fait inclut généralement une dimension temporelle. En plus de cette dimension, qui sert à modéliser l'évolution des obligations, la logique STIT peut être augmentée de façon à rendre compte de phénomènes plus complexes, comme l'interaction entre les modalités temporelles, déontiques et épistémiques.[28]

[26] Pour de plus amples détails sur les modèles minimaux, le lecteur est invité à consulter Chellas (1980).

[27] Voir aussi Pacheco et Carmo (2003) et Pacheco et Santos (2004).

[28] Pour une introduction à la logique épistémique et à l'épistémologie formelle, le lecteur est invité à consulter Hendricks (2006).

Broersen (2011a), par exemple, propose une logique multi-modale fondée sur une logique déontique STIT et augmentée par une modalité épistémique K_a.[29] D'emblée, l'approche de Broersen vise la représentation de l'action intentionnelle. En effet, cet auteur propose une nouvelle interprétation de l'opérateur K_a lorsque celui-ci est combiné avec l'opérateur $[A\ stit]$: la proposition $K_a[a\ stit]\varphi$ signifie « l'agent a réalise φ *sciemment* ».[30]

La motivation principale de l'auteur est de formaliser certaines notions de droit criminel, notamment la responsabilité d'un accusé. En *common law*, le fait qu'un accusé soit considéré comme responsable ou non d'un crime dépend de deux facteurs, l'*actus reus* et la *mens rea*, qui renvoient à l'acte et à l'état mental d'un individu au moment du crime. Afin qu'un agent soit considéré comme responsable d'un crime, il faut établir, d'une part, que le crime a effectivement été commis (l'*actus reus*) et, d'autre part, que l'individu maîtrisait ses capacités intellectuelles.[31]

Alors que plusieurs systèmes de droit criminel font la distinction entre différents *modes* de *mens rea*, par exemple *faire en vue d'un but*, *faire en connaissance de cause* ou encore *faire négligemment*, le système proposé vise la représentation formelle de ces différents modes de *mens rea*, de l'*actus reus* et, de manière générale, de la culpabilité.

Même si le système de Broersen s'insère dans le cadre des logiques de type STIT, celui-ci utilise en fait une logique STIT augmentée. En effet, Broersen utilise la logique XSTIT, qui se différencie de la logique STIT dans la mesure où les actions n'ont pas d'effet *immédiat* : l'effet d'une action prend place dans un état subséquent.[32] Usuellement, l'opérateur STIT satisfait le schéma d'axiome (**T**) : s'il est vrai que a réalise φ dans le scénario w, alors φ est vrai pour w. Or avec XSTIT, φ sera vrai dans l'état subséquent. À présumer que (dans une conception linéaire du temps) le monde se trouve dans un certain *état* à chaque moment, l'effet d'une action posée au temps t_i prend place au moment t_{i+1}, où t_{i+1} est le successeur immédiat de t_i.

Le langage $\mathcal{L}_{\text{XSTIT}}$ contient un ensemble dénombrable de propositions atomiques $Prop$, un ensemble fini $Ags = \{a_1, ..., a_j\}$ d'agents (où $A \subseteq Ags$ est un ensemble d'agents), les connecteurs pour la négation et la conjonction ainsi que les opérateurs pour la nécessité, l'état subséquent et la connaissance.[33]

[29] L'article de Broersen (2011a) provient d'un acte du colloque DEON'08 (Broersen 2008). Il a aussi traité (différemment) de ce sujet dans Broersen (2009).

[30] Traduction libre de « *the agent a knowingly sees to it that* φ ».

[31] Pour une analyse détaillée de ces notions, le lecteur peut consulter Parent (2008, 2007).

[32] Nous utilisons l'expression « état subséquent » comme traduction de « *next state* ».

[33] Notons que l'auteur écrit $[A\ xstit]\varphi$ plutôt que $[A\ xstit\ \varphi]$, comme on le retrouve chez Belnap et Perloff (1988), afin d'insister sur le fait qu'il s'agit d'une modalité. De plus, $[a\ xstit]$ est une abréviation de $[\{a\}\ xstit]$.

$$\mathcal{L}_{\text{XSTIT}} = \{Prop, Ags, (,), \neg, \wedge, \Box, [\ xstit], X, K_a\}$$

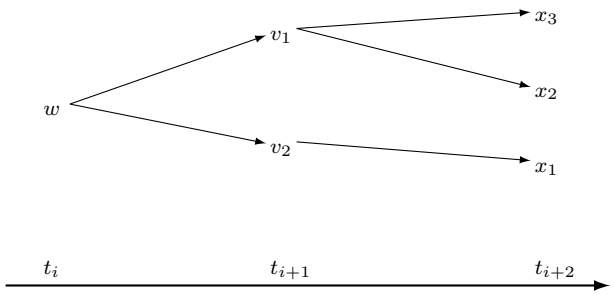
L'ensemble des énoncés bien formés (EBF) est défini récursivement[34] :

$$\varphi := p_i \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi \mid [A\ xstit]\varphi \mid X\varphi \mid K_a\varphi$$

L'opérateur $\Box$ représente la nécessité historique et $X\varphi$ exprime la transition d'un état (statique) à un autre (son successeur). Les connecteurs $\rightarrow$, $\leftrightarrow$, $\vee$ et l'opérateur dual $\Diamond$ sont définis de manière usuelle et $\langle A\ xstit\rangle$ est l'abréviation de $\neg[A\ xstit]\neg$.

La notion de nécessité *historique* vise la représentation de ce qui est vrai pour tout scénario accessible (Thomason 2002). Considérons par exemple la figure suivante :

FIGURE 9.1
La nécessité historique



Les scénarios v_1 et v_2 sont accessibles au scénario w au temps t_i, où v_1 et v_2 sont les deux états subséquents possibles à w. De la même manière, x_3 et x_2 sont les deux scénarios subséquents possibles à v_1 au temps t_{i+1}. Une *histoire* peut informellement être considérée comme un « chemin » qui passe par différents scénarios, c'est-à-dire une suite. En ce sens, nous avons dans cet exemple trois histoires (ou suite d'événements) possibles : (w, v_1, x_3), (w, v_1, x_2) et (w, v_2, x_1). La nécessité historique est ponctuelle et est déterminée à un certain moment t_j. Ce qui est nécessairement (historiquement) vrai à t_i, par exemple, est ce qui est vrai pour toute histoire passant par w. Cela dit, ce qui est historiquement nécessairement vrai peut changer dans le futur.

[34] Cette définition permet l'itération des opérateurs.

Le système axiomatique est construit à partir d'une axiomatisation de la logique propositionnelle classique ainsi que des règles d'inférence usuelles pour les opérateurs modaux. Autrement dit, pour $\star$ une modalité de XSTIT, la règle d'inférence suivante ainsi que le schéma d'axiome s'appliquent. Ainsi, nous avons affaire à des modalités normales :

$$\frac{\vdash \varphi}{\vdash \star\varphi} \qquad (\text{Nec}\star)$$

$$\vdash \star(\varphi \to \psi) \to (\star\varphi \to \star\psi) \qquad (\text{Dist}\star)$$

L'axiomatisation se fait par les schémas d'axiomes suivants :

$\Box\varphi \to \varphi$ (**T**)
$\Diamond\varphi \to \Box\Diamond\varphi$ (**5**)
$[A\ xstit]\varphi \to \langle A\ xstit\rangle\varphi$ (**D**)
$\neg X\neg\varphi \to X\varphi$ (A1)
$[\varnothing\ xstit]\varphi \leftrightarrow \Box X\varphi$ (A2)
$[Ags\ xstit]\varphi \leftrightarrow X\Box\varphi$ (A3)
$[A\ xstit]\varphi \to [A \cup B\ xstit]\varphi$ (A4)
$(\Diamond[A\ xstit]\varphi \wedge \Diamond[B\ xstit]\psi) \to \Diamond([A\ xstit]\varphi \wedge [B\ xstit]\psi)$ (A5)
(où $A \cap B = \varnothing$)

Les axiomes (**T**) et (**5**) donnent à la nécessité historique une structure de type $S5$ et l'axiome (**D**) donne à l'opérateur $[A\ xstit]$ une structure de type KD (Garson 2006). (**T**) signifie que ce qui est nécessairement vrai pour un scénario est vrai. L'axiome (**5**) indique que s'il est possible que φ soit vrai pour un scénario v au moment t_i, qui est un état subséquent d'un scénario w à t_{i-1}, alors φ est aussi vrai pour les autres alternatives v' de w. (**D**) stipule que si A réalise φ, alors il est faux que A réalise $\neg\varphi$.

L'axiome (A1) signifie que s'il est faux que $\neg\varphi$ est le cas à un état subséquent, alors φ est le cas pour cet état. Il s'agit d'une instance de l'axiome (**CD**) et indique que l'état subséquent est unique. (A2) stipule que si φ advient « naturellement », c'est-à-dire que φ n'est réalisé par aucun agent, alors c'est une nécessité historique que φ prenne place dans l'état subséquent. Autrement dit, un changement qui n'est causé par aucun agent est une nécessité historique pour l'état subséquent. (A3) signifie que si φ est réalisé par tous les agents, alors, dans l'état subséquent, φ est une nécessité historique. En ce sens, les propositions qui ne sont pas des nécessités historiques dans les états subséquents représentent des états de choses qui peuvent être réalisés à la suite des actions et des choix des agents.

L'axiome (A4) indique que si φ est réalisé par A, alors il est aussi réalisé conjointement par un ensemble qui contient A.[35] Soulignons que cet axiome implique que les actions d'un groupe se réduisent aux actions de chaque membre. L'axiome (A5) quant à lui est restreint par la clause $A \cap B = \varnothing$ et signifie que s'il est possible que A réalise φ et qu'il est possible que B réalise ψ, alors il est possible que A réalise φ et que B réalise ψ.

Broersen propose ensuite d'ajouter trois axiomes épistémiques afin de modéliser adéquatement la notion de « réaliser sciemment ».

$$K_a X\varphi \rightarrow K_a[a\ xstit]\varphi \tag{A6}$$

$$K_a[a\ xstit]\varphi \rightarrow XK_a\varphi \tag{A7}$$

$$\Diamond K_a\varphi \rightarrow K_a\Diamond\varphi \tag{A8}$$

(A6) indique que si a sait que φ sera le cas dans l'état subséquent, alors a réalise sciemment φ.[36] (A7) signifie que si a réalise sciemment φ, alors dans l'état subséquent a sait que φ. Finalement, (A8) stipule que s'il est possible que a sache que φ, alors a sait que φ est possible.[37]

Cette axiomatisation se traduit littéralement sur le plan de la sémantique. Suivant les résultats de Sahlqvist (1975), qui a montré que les axiomes modaux induisent des relations précises sur les modèles sémantiques (Garson 2006), Broersen présente le modèle sémantique en explicitant avec quelle condition chaque axiome est lié.

Le modèle sémantique $\mathcal{M} = \langle \mathcal{F}, \pi \rangle$ est défini à partir d'une structure $\mathcal{F}$ et d'une fonction $\pi : P \rightarrow 2^{S \times H}$ qui attribue à chaque proposition l'ensemble des états dynamiques dans lesquels la proposition est vraie.

$$\mathcal{F} = \langle S, H, R_X, R_\Box, R_A, \sim_a \rangle$$

Considérons d'abord la structure $\mathcal{F}$. S est un ensemble dénombrable d'états statiques et $H \neq \varnothing$ est un ensemble non dénombrable d'histoires h, où h est un ensemble dénombrable ordonné de membres de S. Autrement dit, H contient toutes les combinaisons dénombrables possibles qu'on peut faire à partir des membres de S. Une histoire h est ordonnée par R_X, une relation d'état subséquent. Cette relation permet d'ordonner les états dynamiques $\langle s, h \rangle$.

[35] Cela ne veut toutefois pas dire que $[B - A\ xstit]\varphi$.

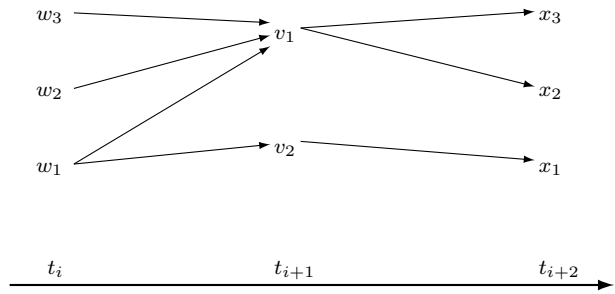
[36] Cet axiome est discutable. Imaginons que nous sommes le 27 août 2013, et que Paul sait que le lendemain, la situation en Égypte ne se sera pas améliorée. Donc Paul réalise sciemment que la situation en Égypte ne s'améliore pas ?

[37] Cet axiome aussi est discutable. Il est possible que petit Jules, qui a quatre ans, sache les principes de la mécanique quantique, sans nécessairement que Jules sache que les principes de la mécanique quantique sont possibles.

Un état dynamique est une paire ordonnée qui contient un état (statique) et une histoire. Un état s' est subséquent à s dans une histoire h si et seulement si $\langle s, h\rangle R_X \langle s', h\rangle$. La relation R_X est sérielle et déterministe, ce qui est représenté respectivement par les axiomes (**D**) et (A1).

Par surcroît, la relation R_X est contextuelle à une histoire. En effet, si $\langle s, h\rangle R_X \langle s', h'\rangle$, alors $h = h'$, et donc dans l'état subséquent s', les agents demeurent dans la même histoire. $R_\Box$ est une relation de nécessité historique telle que $\langle s, h\rangle R_\Box \langle s', h'\rangle$ si et seulement si $s = s'$. En ce sens, la relation de nécessité historique est ponctuelle et s'évalue à un moment donné et pour un scénario donné (elle évalue ce qui est vrai pour toute histoire à un moment donné). Considérons la figure 9.2.

FIGURE 9.2
La nécessité historique (2)



Les différentes histoires visent à représenter les choix que les agents peuvent poser pour faire advenir des états subséquents spécifiques. À l'état w_1 par exemple, un agent peut faire advenir v_1 ou v_2. Cela dit, la nécessité historique évalue ce qui est commun à chaque histoire pour un moment et un état donné. Par exemple, si on fixe le moment à t_{i+2} et l'état à x_3, la nécessité historique évalue ce qui est commun à (w_3, v_1, x_3), (w_2, v_1, x_3) et (w_1, v_1, x_3). Quant aux relations R_A, où $A \subseteq Ags$, ce sont des relations *efficaces*[38] telles que :

1. $R_\varnothing = R_\Box \circ R_X$;
2. $R_{Ags} = R_X \circ R_\Box$;
3. si $B \subset A$, alors $R_A \subseteq R_B$;
4. si $A \cap B = \varnothing$, $\langle s_1, h_1\rangle R_\Box \langle s_2, h_2\rangle$ et $\langle s_1, h_1\rangle R_\Box \langle s_3, h_3\rangle$, alors
   (a) il existe $\langle s_4, h_4\rangle$ tel que $\langle s_1, h_1\rangle R_\Box \langle s_4, h_4\rangle$;

[38] La relation *effectivity* est utilisée en théorie du choix social et signifie que les choix des individus et que les actions qu'ils posent en fonction de ces choix réalisent l'effet désiré (Pauly 2002).

(b) si $\langle s_4, h_4\rangle R_A \langle s_5, h_5\rangle$, alors $\langle s_2, h_2\rangle R_A \langle s_5, h_5\rangle$;
(c) si $\langle s_4, h_4\rangle R_B \langle s_6, h_6\rangle$, alors $\langle s_3, h_3\rangle R_B \langle s_6, h_6\rangle$.

Avant d'expliquer en mots ce que ces conditions signifient, notons que dans chacun des cas (R_X, $R_\Box$ et R_A), les relations sont des sous-ensembles de $(S\times H)\times(S\times H)$ qui respectent certaines conditions. De fait, ces relations peuvent être composées à l'aide de $\circ$.

Soit $R_1, R_2 \subseteq S \times S$ et (a, b) t.q. $a, b \in S$. La composition de relations est donnée par $R_1 \circ R_2 = \{(a, b) :$ il y a $x \in S$ t.q. aR_1x et $xR_2b\}$ (Harzheim 2005).

Voyons d'abord ce que la relation composée $R_\Box \circ R_X$ signifie. Ayant $\langle s_1, h_1\rangle R_X \langle s_2, h_2\rangle$ et $\langle s_2, h_2\rangle R_\Box \langle s_3, h_3\rangle$, cette relation donne $\langle s_1, h_1\rangle R_\Box \circ R_X \langle s_3, h_3\rangle$. Or en vertu des définitions de R_X et $R_\Box$, cela implique que $h_1 = h_2$ et que $s_2 = s_3$. De fait, la première condition signifie que les « actions » du groupe vide d'agent $\varnothing$ déterminent ce qui est nécessairement historiquement vrai pour tout scénario subséquent. R_X isole les scénarios subséquents possibles, puis $R_\Box$ détermine ce qui est vrai dans chaque scénario subséquent possible. Cette condition va de pair avec (A2).

La seconde condition accompagne (A3) et signifie que les actions du groupe Ags déterminent les états subséquents possibles et fixent, pour chacun, ce qui y sera historiquement nécessairement vrai (en vertu des actions de chaque agent). (A4) et signifie que si B est un sous-groupe de A, alors A est minimalement autant efficace que B (ou que les choix possibles d'un agent sont inclus dans les choix possibles pour la collectivité).

La dernière condition représente l'axiome (A5). Conformément à la clause susmentionnée pour la relation $R_\Box$, deux états dynamiques sont en relation de nécessité historique à condition que l'état statique soit le même. En ce sens, dans la condition 4, nous pouvons conclure qu'il s'agit du même état $s = s_1 = s_2 = s_3$. Si A et B sont deux groupes d'agents qui ne partagent aucun membre et que $\langle s, h_1\rangle$, $\langle s, h_2\rangle$ et $\langle s, h_3\rangle$ sont en relation de nécessité historique, 4(a) signifie qu'il y a un autre état dynamique $\langle s_4, h_4\rangle$ en relation de nécessité historique avec $\langle s, h_1\rangle$ (et donc $s_4 = s_1 = s$).

Informellement, h_1 peut être considérée comme l'histoire où A et B n'agissent pas, h_2 comme celle où A agit, h_3 celle où B agit et h_4 celle où les deux agissent. Dans l'éventualité où les deux agissent, les conditions 4(b) et 4(c) marquent l'indépendance des actions de A et de B : 4(b) signifie que si l'état dynamique $\langle s_5, h_5\rangle$ prend place après $\langle s, h_4\rangle$ et est réalisé par A, alors il prend place après $\langle s, h_2\rangle$, et (c) signifie que si l'état dynamique $\langle s_6, h_6\rangle$ prend place après $\langle s, h_4\rangle$ et est réalisé par B, alors il prend place après $\langle s, h_3\rangle$. De fait, il est possible que $h_5 = h_6$, mais que néanmoins A et B soient responsables de leurs actions respectives, dans la mesure où l'histoire résulte de deux choix différents.

Finalement, $\sim_a$ (avec $a \in Ags$) sont des relations d'équivalences épistémiques par rapport aux états dynamiques, où deux états sont en relations $\sim_a$ d'équivalence épistémique lorsque les connaissances de a sont les mêmes dans chaque état.

La vérité dans un modèle $\mathcal{M}$ relativement à un état dynamique $\langle s, h\rangle$, notée $\models_{\mathcal{M}_{\langle s,h\rangle}}$, est définie récursivement :

$$\models_{\mathcal{M}_{\langle s,h\rangle}} p \Leftrightarrow \langle s,h\rangle \in \pi(p) \tag{1}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \neg\varphi \Leftrightarrow \not\models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \tag{2}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \wedge \psi \Leftrightarrow \models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \text{ et } \models_{\mathcal{M}_{\langle s,h\rangle}} \psi \tag{3}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \Box\varphi \Leftrightarrow \langle s,h\rangle R_\Box \langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{4}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} [A\ xstit]\varphi \Leftrightarrow \langle s,h\rangle R_A \langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{5}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} X\varphi \Leftrightarrow \langle s,h\rangle R_X \langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{6}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} K_a\varphi \Leftrightarrow \langle s,h\rangle \sim_a \langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{7}$$

La clause (1) stipule qu'un atome propositionnel p est vrai dans un état dynamique à condition que cet état soit membre de l'ensemble qui contient tous les états dynamiques où p est vrai et la condition (2) est la bivalence classique, à savoir que ce qui n'est pas vrai est faux, et vice versa. (3) est la condition sémantique usuelle pour la conjonction. La clause (4) signifie qu'une proposition φ est une nécessité historique pour un état dynamique à condition que cette proposition soit vraie pour tout autre état dynamique en relation de nécessité historique avec lui. (5) implique que A réalise φ est vrai pour un état dynamique dans la mesure où φ est vrai pour tout état dynamique subséquent réalisé par A. La condition (6) indique que φ est vrai dans le prochain état à condition que φ soit vrai pour tout état dynamique subséquent possible. Finalement, (7) signifie qu'il est vrai que a sait φ pour un état dynamique $\langle s, h\rangle$ à condition que φ soit vrai pour tout autre état dynamique en relation d'équivalence épistémique avec $\langle s, h\rangle$.

En ce qui concerne les trois axiomes proposés pour rendre compte des interactions entre les modalités $[A\ xstit]$ et K_a, les restrictions respectives sur le modèle sémantique sont :

$$\sim_a \circ R_a \subseteq \sim_a \circ R_X \tag{8}$$
$$R_X \circ \sim_a \subseteq \sim_a \circ R_a \tag{9}$$
$$\langle s_1,h_1\rangle R_\Box \langle s_2,h_2\rangle \text{ et } \langle s_1,h_1\rangle \sim_a \langle s_3,h_3\rangle \Rightarrow \exists s_4,h_4 \text{ t.q. } \langle s_2,h_2\rangle \sim_a \langle s_4,h_4\rangle \text{ et } \langle s_3,h_3\rangle R_\Box \langle s_4,h_4\rangle \tag{10}$$

Cela nous amène aux considérations quant aux opérateurs déontiques. À l'instar de la tendance en logique dynamique, inspirée par les travaux de Kanger (1957, 1971) et Anderson (1958a), Broersen ne prend pas l'opérateur déontique comme primitif, mais le définit plutôt à l'aide de l'opérateur $[A\ xstit]$ et d'une constante de violation V.

L'auteur propose deux définitions possibles pour l'obligation :

$$O[a\ xstit]\varphi =_{df} \Box(\neg[a\ xstit]\varphi \rightarrow [a\ xstit]V) \qquad \text{(df. } O\text{)}$$
$$O'[a\ xstit]\varphi =_{df} \Box([a\ xstit]\neg\varphi \rightarrow [a\ xstit]V) \qquad \text{(df. } O'\text{)}$$

Ces deux définitions jouent sur une nuance que le langage de XSTIT est capable d'exprimer.[39] Alors que le premier énoncé est dérivable dans le langage de XSTIT, le second ne l'est pas :

$$[A\ xstit]\neg\varphi \rightarrow \neg[A\ xstit]\varphi \qquad (9.1)$$
$$\neg[A\ xstit]\varphi \rightarrow [A\ xstit]\neg\varphi \qquad (9.2)$$

Autrement dit, il est vrai que si A réalise $\neg\varphi$, alors il est faux que A réalise φ. Cependant, il n'est pas vrai que s'il est faux que A réalise φ (i.e. A ne réalise pas φ), alors A réalise $\neg\varphi$. Cette proposition est en fait satisfiable. Par exemple, ce n'est pas parce que Paul ne chante pas qu'il fait exprès de ne pas chanter. Ainsi, « ne pas réaliser φ » n'est pas équivalent à « réaliser la négation φ ».

Conformément à cette distinction, l'opérateur O est plus sévère que O'. En effet, dans le premier cas « il est obligatoire que a réalise φ » signifie que pour toute histoire (au moment t_i et pour l'état statique s), si a ne réalise pas φ, alors a réalise une violation V. Dans le second cas, la définition stipule que « il est obligatoire que a réalise φ » indique que pour toute histoire (au moment t_i et pour l'état statique s), si a réalise $\neg\varphi$, alors a réalise une violation V.

La première définition est plus sévère dans la mesure où l'absence d'action (la non-réalisation) d'un agent peut mener au fait qu'il réalise néanmoins une violation, alors que dans le second cas, la violation est réalisée seulement lorsque $\neg\varphi$ est réalisé. Autrement dit, la première définition implique que pour remplir ses obligations, un agent doit s'assurer que ses obligations sont remplies. En ce sens, cela exclut la possibilité de respecter ses obligations en omettant d'agir.

[39] Puisque KT est une extension de KD, cette nuance peut aussi être exprimée dans un langage STIT.

L'auteur propose aussi d'autres définitions en conjonction avec l'opérateur K_a pour exprimer différents modes de *mens rea*. Par la suite, Broersen (2011b) introduira un opérateur d'intention pour rendre compte de certains modes de *mens rea* où un agent agit non seulement en connaissance de cause, mais aussi en vue d'accomplir un certain but.

* * *

La caractéristique principale des logiques de type STIT est de rendre compte de l'action en modélisant l'évolution d'un scénario, où l'action est associée à une description du monde. En plus des logiques STIT, on trouve aussi les logiques dynamiques, qui présument une ontologie de l'action plutôt que de considérer les actions en tant que propositions déclaratives. Passons maintenant à ces approches.

Chapitre 10

Les logiques dynamiques

Bien qu'il existe plusieurs logiques de l'action dans la littérature scientifique, celles qu'on retrouve principalement en philosophie sont les logiques STIT et les logiques dynamiques (Segerberg *et al.* 2009). Malgré les avantages de la logique STIT, comme la capacité de faire la distinction entre ne pas agir ou encore omettre ($\neg E_i\varphi$), et s'abstenir d'agir ($E_i\neg\varphi$), cette logique n'est toutefois pas en mesure de rendre compte du fait que ce sont les *actions* qui sont obligatoires. En effet, elle est de type *ought-to-be*, où les descriptions des états sont obligatoires. Le principal avantage des logiques dynamiques par rapport aux logiques STIT est leur capacité à faire la distinction entre les *actions* et les *propositions*, ce qui requiert cependant un cadre de travail à deux niveaux.

Selon Anglberger (2009), une logique déontique est dite *réductible* lorsqu'aucun opérateur déontique n'est pris en tant que primitif, donc que toute proposition qui contient un opérateur déontique est équivalente à une proposition qui n'en contient pas. De manière générale, les logiques déontiques dynamiques utilisent des réductions de type Anderson-Kanger. En effet, en logique déontique dynamique, l'obligation est usuellement définie à l'aide d'une constante de violation V, une action étant interdite lorsque sa performance entraîne un état où la constante de violation est vraie.

Préliminaires

Avant d'aborder explicitement les logiques dynamiques qui traitent de l'obligation et de la permission, voyons d'abord brièvement la formulation initiale de la logique dynamique. Suivant Segerberg (1992), la logique dynamique a été développée par Pratt (1976, 1980) et visait la représentation de l'évolution d'un scénario à la suite des actions d'un programme informatique.[1]

[1] Le lecteur peut consulter Harel *et al.* (2000) pour une introduction à la logique dynamique.

La logique dynamique est une logique modale, où l'opérateur $\Box$ est modifié afin d'inclure une action. Ainsi, l'énoncé $[\alpha]\varphi$ signifie qu'à la suite de l'exécution de α (qui peut être un programme informatique ou encore une action), la proposition φ est vraie, c'est-à-dire que c'est une description adéquate du scénario (ou de l'état dans lequel se trouve le scénario).

Sur la base d'un langage $\mathcal{L} = \{Prop, Act, (,), \cup, ;, ^*, \neg, \supset, [\,]\}$, où $Prop$ et Act sont respectivement des ensembles dénombrables d'atomes propositionnels et d'actions atomiques (le dernier incluant 0 et 1), les énoncés bien formés sont définis récursivement ainsi :

$$\varphi := \; p_i \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi$$

Les actions complexes sont composées des connecteurs $\cup$, ; et *, qui dénotent respectivement le choix, la séquence et l'itération, les trois éléments de base pour une logique visant à la programmation (Pratt 1991). Les énoncés d'action bien formés sont définis récursivement ainsi :

$$\alpha := \; a_i \mid \alpha \cup \beta \mid \alpha; \beta \mid \alpha^*$$

L'opérateur $[\,]$ est une modalité normale de type K et est axiomatisée de manière usuelle. La structure des actions qui prennent place au sein de l'opérateur $[\,]$ s'exprime par une algèbre de Kleene (Harel *et al.* 2000). Une algèbre de Kleene est un semi-anneau idempotent qui satisfait les axiomes suivants pour l'itération :

$$1 \cup (\alpha; \alpha^*) = 1 \cup (\alpha^*; \alpha) = \alpha^* \tag{K*1}$$
$$\text{si } \beta \cup (\alpha; \gamma) \leq \gamma, \text{ alors } \alpha^*; \beta \leq \gamma \tag{K*2}$$
$$\text{si } \beta \cup (\gamma; \alpha) \leq \gamma, \text{ alors } \beta; \alpha^* \leq \gamma \tag{K*3}$$

Un semi-anneau est une structure avec une opération ; associative et ayant pour unité 1, une opération $\cup$ associative et commutative où ; distribue sur $\cup$ des deux côtés et où $0; \alpha = 0 = \alpha; 0$ (Peterson 2015c). Formellement, une algèbre de Kleene $\mathcal{KA} = (K, \cup, ;, 0, 1, *)$ se définit à l'aide des axiomes suivants (en plus des trois axiomes susmentionnés) :

$$\alpha \cup (\beta \cup \gamma) = (\alpha \cup \beta) \cup \gamma \tag{K1}$$
$$\alpha \cup \beta = \beta \cup \alpha \tag{K2}$$
$$\alpha \cup 0 = \alpha = \alpha \cup \alpha \tag{K3}$$
$$\alpha; (\beta; \gamma) = (\alpha; \beta); \gamma \tag{K4}$$
$$\alpha; 1 = \alpha = 1; \alpha \tag{K5}$$
$$\alpha; 0 = 0 = 0; \alpha \tag{K6}$$
$$\alpha; (\beta \cup \gamma) = (\alpha; \beta) \cup (\alpha; \gamma) \tag{K7}$$

$$(\alpha \cup \beta); \gamma = (\alpha; \gamma) \cup (\beta; \gamma) \tag{K8}$$

Ici, (K1) et K4) représentent respectivement l'associativité de $\cup$ et ;, (K2) rend $\cup$ commutatif et (K3) le rend idempotent avec 0 comme unité. Similairement, (K5) fait que 1 est l'unité de ; et les équations (K7) et (K8) font que ; se distribue sur $\cup$, (K7) indiquant que 0 absorbe ;.

John-Jules Meyer

À la suite des avancées de Pratt, Meyer (1987, 1988) a analysé la logique déontique dans le cadre de la logique dynamique afin de résoudre le paradoxe de Forrester. Meyer modifie le langage de la logique dynamique en mettant de côté l'itération et en introduisant des connecteurs pour la négation d'action $\bar{\alpha}$, l'action conjointe $\alpha \cap \beta$ et l'action conditionnelle (ou alternative) $\varphi \rightarrow \alpha/\beta$, formant ainsi le langage de la logique déontique dynamique propositionnelle PD_eL.

Plutôt que d'utiliser 0 et 1, Meyer introduit les actions $\varnothing$ et U pour dénoter respectivement l'action qui échoue et l'action universelle. À partir du langage $\mathcal{L} = \{Prop, Act, (,), \bar{}, \cup, \cap, ;, \neg, \supset, [\ ], \rightarrow\}$, les énoncés bien formés sont définis récursivement de la manière suivante :

$$\varphi := \ p_i \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi \mid \varphi \rightarrow \alpha/\beta$$

Les actions complexes sont définies comme suit :

$$\alpha := \ a_i \mid \bar{\alpha} \mid \alpha \cap \beta \mid \alpha \cup \beta \mid \alpha; \beta$$

À l'instar de la logique dynamique, l'opérateur $[\alpha]$ est axiomatisé comme une modalité normale de type K avec les définitions usuelles pour les autres connecteurs.

Les axiomes gouvernant les actions sont les suivants :

$$[\bar{\bar{\alpha}}]\varphi \equiv [\alpha]\varphi \tag{A1}$$
$$[\alpha; \beta]\varphi \equiv [\alpha]([\beta]\varphi) \tag{A2}$$
$$[\overline{\alpha; \beta}]\varphi \equiv [\bar{\alpha}]\varphi \wedge [\alpha]([\bar{\beta}]\varphi) \tag{A3}$$
$$[\alpha \cup \beta]\varphi \equiv [\alpha]\varphi \wedge [\beta]\varphi \tag{A4}$$
$$[\bar{\alpha}]\varphi \vee [\bar{\beta}]\varphi \supset [\overline{\alpha \cup \beta}]\varphi \tag{A5}$$
$$[\alpha]\varphi \vee [\beta]\varphi \supset [\alpha \cap \beta]\varphi \tag{A6}$$
$$[\bar{\alpha}]\varphi \wedge [\bar{\beta}]\varphi \equiv [\overline{\alpha \cap \beta}]\varphi \tag{A7}$$
$$[\varnothing]\varphi \tag{A8}$$
$$[\varphi \rightarrow \alpha/\beta]\psi \equiv (\varphi \supset [\alpha]\psi) \wedge (\neg\varphi \supset [\beta]\psi) \tag{A9}$$

$$[\overline{\varphi \rightarrow \alpha/\beta}]\psi \equiv (\varphi \supset [\overline{\alpha}]\psi) \wedge (\neg\varphi \supset [\overline{\beta}]\psi) \qquad \text{(A10)}$$

L'axiome (A1) indique la bivalence du point de vue de l'action, à savoir qu'une action est équivalente à la négation de sa négation. (A2) signifie que l'expression « l'action α suivie de l'action β entraîne φ » est équivalente à « si l'action α est réalisée, alors si l'action β est réalisée, alors φ est une description vraie du monde ». Soulignons que cet axiome implique que le connecteur pour la séquence d'actions n'est pas primitif, et que, par conséquent, celui-ci aurait pu être omis.

L'axiome (A3) indique que si la négation de l'action « α suivie de β » entraîne φ, alors la négation de α entraîne φ et α entraîne que la négation de β entraîne φ. Autrement dit, si φ est la conséquence de la négation de l'action combinée $\alpha;\beta$, alors φ est réalisé à la suite de non α, et, à la suite de α, φ est réalisé à la suite de non β.

(A4) signifie que si φ est vrai après un choix entre α ou β, alors φ est vrai après l'exécution de α et après celle de β. (A5) indique que si la négation de α mène à φ ou la négation de β mène à φ, alors la négation de l'action disjointe $\alpha \cup \beta$ mène aussi à φ. De la même manière, (A6) signifie que si α mène à φ ou β mène à φ, alors l'action conjointe de α et β y mène aussi. Les axiomes (A5) et (A6) doivent cependant mettre en jeu des actions qui ont la même durée dans le temps.

(A7) stipule que si la négation de α mène à φ, de même que la négation de β, alors la négation de l'action conjointe de α et β y mènera aussi (et vice versa). L'action conditionnelle quant à elle n'est pas à être comprise dans les même termes que pour une implication matérielle. En termes simples, $\varphi \rightarrow \alpha/\beta$ signifie que φ entraîne α et que $\neg\varphi$ entraîne β. En ce sens, si cela mène à une description ψ, alors φ entraîne que α mène à ψ et $\neg\varphi$ entraîne que β y mène aussi. Il en va de même pour la négation de l'action conditionnelle et l'interprétation de (A9). L'expression $\overline{\varphi \rightarrow \alpha/\beta}$ signifie que φ entraîne non α et que $\neg\varphi$ entraîne non β.

Finalement, l'axiome (A10) signifie qu'une action impossible mène à n'importe quel état. Puisqu'une action impossible entraîne à la fois φ et $\neg\varphi$, son exécution entraîne n'importe quel état (*ex falso sequitur quodlibet*).

Meyer propose une réduction des opérateurs déontiques à une constante de violation V :

$$F\alpha =_{df} [\alpha]V \qquad \text{(df. } F\text{)}$$
$$O\alpha =_{df} F\bar{\alpha} \qquad \text{(df. } O\text{)}$$
$$P\alpha =_{df} \neg F\alpha \qquad \text{(df. } P\text{)}$$

Une action α est interdite si sa performance implique une violation V, elle est obligatoire si sa négation est interdite, et elle est permise lorsqu'elle n'est pas interdite. Dans le langage de la logique dynamique, nous avons $O\alpha \equiv [\bar{\alpha}]V$ et $P\alpha \equiv \neg[\alpha]V$.

Du côté de la sémantique, Meyer introduit les notions d'*ensemble de simultanéité* (*synchronicity set*) et de *trace de simultanéité*. Les ensembles de simultanéité $S_1, ..., S_n, ...$, sont tels que S_i est fini, $S_i \neq \varnothing$ et $S_i \subset Act$. Une trace de simultanéité est une suite t_i d'ensembles de simultanéité (laquelle peut être finie ou non). Ces notions sont introduites pour rendre compte formellement de la sémantique qui gouverne les opérateurs d'action. Informellement, un ensemble de simultanéité correspond à un scénario où les actions (atomiques) sont faites au même moment (ou dans la même période de temps) et une trace de simultanéité correspond à une histoire.

L'idée derrière ces notions est qu'une action fait référence aux scénarios dans lesquels elle est posée. Afin d'isoler certaines parties des différentes traces de simultanéité, Meyer introduit une fonction de pertinence, qui attribue 1 à un ensemble de simultanéité lorsque celui-ci est pertinent et 0 lorsqu'il ne l'est pas, ce qui est noté $S^{(i)}$.

Soit $T_1, ..., T_n$ des ensembles de traces de simultanéité. Meyer définit le domaine du modèle comme la collection $\mathcal{C}$ d'ensembles qui contiennent des traces infinies, lesquelles contiennent une partie finie pertinente suivie d'une partie infinie non pertinente. Par exemple, T_2 peut contenir $t_1 = S_1^{(1)}, S_4^{(1)}, S_5^{(0)}...$, où S_1 et S_4 forment la partie pertinente de t_1, suivie d'une partie non pertinente infinie. Afin de rendre compte des connecteurs pour l'action conjointe et l'action niée, Meyer définit un opérateur $\Cap$ pour l'intersection entre les ensembles de simultanéité. $S_1^{(i_1)} \Cap S_2^{(i_2)} = S^{(j)}$ si $S_1 = S_2 = S$ (où $j = i_1$ si $i_2 \leq i_1$, sinon $j = i_2$), sinon $S_1^{(i_1)} \Cap S_2^{(i_2)} = \varnothing$ lorsque $S_1 \neq S_2$.

L'intersection définie par $T_1 \Cap T_2$ est similaire à l'intersection $T_1 \cap T_2$ à la différence que la première prend en compte les traces pertinentes. L'intersection entre deux traces isole les ensembles de simultanéité qui appartiennent aux deux traces en plus d'y associer le plus haut niveau de pertinence. En plus de cette nouvelle opération, Meyer introduit une opération pour rendre compte sémantiquement de l'action niée, où $\wp(Act) - S$ est le complément de S par rapport à l'ensemble des parties de Act :

$$\sim S^{(i)} = (\wp(Act) - S)^{(i)}$$

Cela fait, Meyer introduit les conditions sémantiques pour les opérateurs d'action. Soit $\|\alpha\|$ l'ensemble qui contient toutes les traces possibles où α peut avoir lieu. Cet ensemble est construit dans le même esprit que $\|V_\varphi\|$ (chapitre 2) : il isole tous les scénarios où α peut être posée (et donc où « α est réalisée » est vrai). Les conditions sémantiques sont définies récursivement, où $S_1 \circ S_2$ est la composition (voire la concaténation) d'ensembles de simultanéité et où $cut()$ isole la sous-trace pertinente d'une trace infinie :

$$\|a\| = \{S_i : a \in S_i\}^{(1)} \circ (\wp(Act)^{(0)})^\omega \tag{1}$$
$$\|\overline{\alpha}\| = \sim \|\alpha\| \tag{2}$$
$$\|\alpha; \beta\| = cut(\|\alpha\|) \circ \|\beta\| \tag{3}$$
$$\|\alpha \cup \beta\| = \|\alpha\| \cup \|\beta\| \tag{4}$$
$$\|\alpha \cap \beta\| = \|\alpha\| \Cap \|\beta\| \tag{5}$$
$$\|\varnothing\| = \varnothing \tag{6}$$
$$\|U\| = \wp(Act)^{(1)} \circ (\wp(Act)^{(0)})^\omega \tag{7}$$

En mots, (1) signifie que l'ensemble qui contient toutes les traces possibles où une action atomique a est réalisée correspond à un ensemble de traces composé de tous les ensembles de simultanéité pertinents qui contiennent a suivi de toutes les combinaisons non pertinentes possibles. (2) signifie que l'ensemble qui contient les traces possibles où l'action niée $\overline{\alpha}$ est réalisée équivaut au complément de l'ensemble qui contient toutes les traces possibles où α est réalisée dans $\wp(Act)$.

La troisième condition indique que l'ensemble qui contient toutes les traces possibles où l'action α suivie de β est réalisée est la composition de l'ensemble qui contient les traces pertinentes possibles où α est réalisée avec celui qui contient les traces possibles où β l'est. (4) indique que l'ensemble qui contient les traces possibles où l'action disjointe $\alpha \cup \beta$ est réalisée correspond à l'union des ensembles qui contiennent respectivement les traces possibles où α et β sont réalisées.

La condition (5) stipule que l'ensemble qui contient les traces possibles où l'action conjointe $\alpha \cap \beta$ est réalisée correspond à l'intersection qui contient les traces pertinentes possibles où respectivement α et β sont réalisées. (6) indique que l'ensemble qui contient les traces possibles où l'action impossible est réalisée est vide, et finalement (7) stipule que l'ensemble qui contient les traces possibles où l'action universelle est réalisée équivaut à l'ensemble des parties pertinentes de Act composé avec toute les combinaisons non pertinentes possibles.

Ayant maintenant à sa disposition une sémantique pour les actions, Meyer doit faire le pont entre les traces d'actions possibles et les scénarios possibles qui peuvent en résulter. Soit $W \neq \varnothing$ l'univers du discours qui contient des scénarios possibles et $r : \wp(Act) \times W \longrightarrow W$ une fonction qui décrit l'état $r(S_i, w)$ subséquent à w lorsque les actions de S_i sont accomplies. La fonction $\mathcal{R}$ est définie récursivement à partir de r :

$$\mathcal{R}(S_i, w) = r(S_i, w)$$
$$\mathcal{R}(t_i \circ t_j, w) = \mathcal{R}(t_j, \mathcal{R}(t_i, w))$$

Considérant qu'une trace t_i est une suite de S_j, la seconde condition stipule que l'état subséquent à w lorsque les actions membres de la combinaison de traces t_i et t_j sont réalisées équivaut à l'état subséquent de « l'état subséquent de w lorsque t_i est réalisé » lorsque t_j est réalisé. Autrement dit, l'état subséquent pour une combinaison $t_i \circ t_j$ se comprend de la manière suivante : soit v l'état subséquent de w lorsque t_i est réalisé, i.e. $\mathcal{R}(t_i, w) = v$, alors l'état subséquent de $t_i \circ t_j$ pour w est l'état subséquent de t_j pour v, i.e. $\mathcal{R}(t_i \circ t_j, w) = \mathcal{R}(t_j, v)$.

Étant donné T_i un ensemble de traces, $\mathcal{R}(T_i, w)$ correspond à l'ensemble qui contient les scénarios v pour lesquels v est l'état subséquent de w pour une trace élément de T_i.

$$\mathcal{R}(T_i, w) = \{v : v = \mathcal{R}(t, w) \text{ pour un } t \in T_i\}$$

À l'aide de cette fonction, l'auteur définit la fonction $\|\alpha\|_R(w)$ qui permet d'isoler l'état subséquent à w par rapport à l'ensemble qui contient les traces pertinentes possibles qui contiennent α. Autrement dit, $cut(\|\alpha\|)$ isole les traces pertinentes à l'intérieur de $\|\alpha\|$ et $\|\alpha\|_R(w)$ isole les états subséquents possibles relativement aux traces pertinentes de $\|\alpha\|$.

$$\|\alpha\|_R(w) = \mathcal{R}(cut(\|\alpha\|), w) \qquad \text{(df. } \|\|_R\text{)}$$

Cela fait, il est maintenant possible de redéfinir $\|\alpha\|$ de façon à prendre en compte l'état w dans lequel on se trouve :

$$\|a\|(w) = \{S_i : a \in S_i\}^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \qquad (1')$$
$$\|\overline{\alpha}\|(w) = \sim \|\alpha\|(w) \qquad (2')$$
$$\|\alpha; \beta\|(w) = cut(\|\alpha\|(w)) \circ \|\beta\|(\|\alpha\|_R(w)) \qquad (3')$$
$$\|\alpha \cup \beta\|(w) = \|\alpha\|(w) \cup \|\beta\|(w) \qquad (4')$$
$$\|\alpha \cap \beta\|(w) = \|\alpha\|(w) \Cap \|\beta\|(w) \qquad (5')$$
$$\|\varnothing\|(w) = \varnothing \qquad (6')$$
$$\|U\|(w) = \wp(Act)^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \qquad (7')$$

Seule la signification de la clause (3) change : l'ensemble qui contient toutes les traces possibles où l'action α suivie de β est réalisée dans w est la composition de l'ensemble qui contient les traces pertinentes possibles où α est réalisée dans w avec celui qui contient les traces possibles où β est réalisé dans l'état subséquent à w relativement à l'ensemble qui contient les traces pertinentes qui contiennent α. Autrement dit, considérant que l'action $\alpha;\beta$ marque la réalisation de deux actions consécutives, l'état dans lequel β est réalisée est l'état subséquent de celui dans lequel α est réalisée. Dans les autres cas, il suffit d'ajouter un « relativement à l'état w » à la lecture que nous avons proposée des clauses (1)–(7).

Ayant fait le pont entre la sémantique relative aux actions et celle des propositions, Meyer est en mesure d'exprimer la condition pour l'opérateur d'action conditionnelle $\rightarrow /$.

$$\|\varphi \rightarrow \alpha/\beta\|(w) = \begin{cases} \|\alpha\|(w) \text{ si } \models_w \varphi \\ \|\beta\|(w) \text{ si } \not\models_w \varphi \end{cases} \tag{8'}$$

À cela s'ajoute la clause (9'), qui est une légère modification de (df. $\|\|\|_R$) où on indexe $\|\alpha\|$ à w :

$$\|\alpha\|_R(w) = \mathcal{R}(cut(\|\alpha\|(w)), w) \tag{9'}$$

Tous ces matériaux nous permettent d'en arriver à la condition de vérité pour $[\alpha]\varphi$:

$$\models_w [\alpha]\varphi \Leftrightarrow \models_v \varphi \text{ pour tout } v \in \|\alpha\|_R(w)$$

En mots, cela signifie qu'il est vrai que réaliser l'action α entraîne φ pour un scénario w à condition que φ soit vrai pour toute alternative subséquente v résultant de l'action α. Bien que le modèle ne soit pas explicitement défini chez Meyer, il est possible de représenter la sémantique à l'aide d'un modèle $\mathcal{M} = \langle W, Act, \|\|\|(), r, V\rangle$, où W est l'univers (non vide) du discours, *Act* est un ensemble d'actions atomiques, $\|\|\|()$ est une fonction qui détermine l'ensemble des traces pertinentes possibles relativement à un scénario, r une fonction qui détermine les états subséquents possibles et $V(\varphi)$ une fonction qui détermine l'ensemble qui contient tous les scénarios où une proposition φ est vraie. Les clauses pour les autres connecteurs sont définies de manière habituelle.

Meyer proposera ultérieurement d'augmenter PD_eL avec un axiome de cohérence pour les obligations similaire à l'axiome (**D**), et il proposera l'introduction de plusieurs constantes de violation pour rendre compte des conflits normatifs.

La structure de la logique d'action inhérente à la logique déontique dynamique de Meyer n'est pas explicitement définie par l'auteur. Cela dit, notons que la proposition de Meyer, en mettant de côté le connecteur pour les actions conditionnelles (qui ne sera d'ailleurs pas réutilisé par la suite), est réductible à une algèbre de Boole (Peterson 2015c).

Meyer ne définit pas explicitement l'algèbre d'actions inhérente à PD_eL mais indique que les propositions suivantes sont satisfaites :

$$\overline{\overline{\alpha}} = \alpha \tag{DL1}$$
$$\overline{\alpha \cup \beta} = \overline{\alpha} \cap \overline{\beta} \tag{DL2}$$
$$\overline{\alpha \cap \beta} = \overline{\alpha} \cup \overline{\beta} \tag{DL3}$$
$$\alpha \cup \varnothing = \alpha \tag{DL4}$$
$$\alpha \cup \alpha = \alpha \tag{DL5}$$
$$\alpha \cap \alpha = \alpha \tag{DL6}$$
$$\alpha \cap \overline{\alpha} = \varnothing \tag{DL7}$$
$$\alpha \cap (\beta \cup \gamma) = (\alpha \cap \beta) \cup (\alpha \cap \gamma) \tag{DL8}$$
$$\alpha \cup (\beta \cap \gamma) = (\alpha \cup \beta) \cap (\alpha \cup \gamma) \tag{DL9}$$

Considérons ces équations à la lumière de la définition d'une algèbre de Boole, c'est-à-dire d'un treillis distributif complémenté, défini par les équations suivantes :

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma \tag{L1}$$
$$\alpha \cup (\beta \cup \gamma) = (\alpha \cup \beta) \cup \gamma \tag{L2}$$
$$\alpha \cdot \beta = \beta \cdot \alpha \tag{L3}$$
$$\alpha \cup \beta = \beta \cup \alpha \tag{L4}$$
$$\alpha \cdot \alpha = \alpha \tag{L5}$$
$$\alpha \cup \alpha = \alpha \tag{L6}$$
$$\alpha \cdot 1 = \alpha \tag{L7}$$
$$\alpha \cup 0 = \alpha \tag{L8}$$
$$\alpha \cup (\alpha \cdot \beta) = \alpha \tag{L9}$$
$$\alpha \cdot (\alpha \cup \beta) = \alpha \tag{L10}$$
$$\alpha \cup (\beta \cdot \gamma) = (\alpha \cup \beta) \cdot (\alpha \cup \gamma) \tag{L11}$$
$$\alpha \cdot (\beta \cup \gamma) = (\alpha \cdot \beta) \cup (\alpha \cdot \gamma) \tag{L12}$$
$$\alpha \cdot \bar{\alpha} = 0 \tag{L13}$$
$$\alpha \cup \bar{\alpha} = 1 \tag{L14}$$

Les équations (L1)-(L10) définissent un treillis, (L11) et (L12) le rendent distributif et (L13) et (L14) complémenté.

On voit aisément que les équations (L5), (L6), (L8), (L11), (L12) et (L13) sont satisfaites (où $\cdot$ représente $\cap$). En vertu de l'axiome (A4), on peut aussi retrouver la commutativité et l'associativité de $\cup$ considérant les propriétés de la conjonction, et ainsi, (L2) et (L4) peuvent aussi être obtenus. L'associativité et la commutativité de $\cap$ sont aussi recouvertes par les propriétés de $\cup$ et les équations (DL2) et (DL3), donnant ainsi (L1) et (L3).

(L7) s'obtient par (DL1), (DL2) et (DL4) sachant que $\overline{\alpha} = \overline{\alpha} \cup 0$, où $1 = \overline{0}$.

$$\alpha = \overline{\overline{\alpha}} = \overline{\overline{\alpha} \cup 0} = \overline{\overline{\alpha}} \cdot \overline{0} = \alpha \cdot \overline{0}$$

Le résultat suivant, qui s'obtient à partir de (L8), (L13), (L1)-(L3), (L13), (L1), (L5) et (L13), permet de dériver (L10) à l'aide de (L8), (L11) et (L8) :

$$0 = 0 \cdot 0 = 0 \cdot (\alpha \cdot \overline{\alpha}) = \alpha \cdot (0 \cdot \overline{\alpha}) = \alpha \cdot ((\alpha \cdot \overline{\alpha}) \cdot \overline{\alpha}) =$$
$$\alpha \cdot (\alpha \cdot (\overline{\alpha} \cdot \overline{\alpha})) = \alpha \cdot (\alpha \cdot \overline{\alpha}) = \alpha \cdot 0$$

$$\alpha = \alpha \cup 0 = \alpha \cup (0 \cdot \beta) = (\alpha \cup 0) \cdot (\alpha \cup \beta) = \alpha \cdot (\alpha \cup \beta)$$

(L9) s'obtient à partir de (L12) et (L5) :

$$\alpha \cdot (\alpha \cup \beta) = (\alpha \cdot \alpha) \cup (\alpha \cdot \beta) = \alpha \cup (\alpha \cdot \beta)$$

(L14) résulte de (DL7), (DL3) et (DL1) :

$$\overline{0} = \overline{\alpha \cdot \overline{\alpha}} = \overline{\alpha} \cup \overline{\overline{\alpha}} = \overline{\alpha} \cup \alpha$$

En somme, la structure des actions au sein de la logique dynamique de Meyer s'exprime à l'aide d'une algèbre de Boole. Pour une analyse détaillée, le lecteur est invité à consulter Peterson (2015c).

Lambèr M.M. Royakkers

Malgré l'avantage de la logique déontique dynamique développée par Meyer (1988) pour appréhender la dichotomie entre les actions et les propositions, il n'en demeure pas moins qu'elle ne permet pas de faire la distinction entre les obligations de groupe et les individuelles. De fait, PD_eL ne permet pas de spécifier *à qui* la norme est destinée. Puisque les normes légales sont destinées à des individus spécifiques, Royakkers (1998) est d'avis qu'une logique déontique visant à formaliser le discours légal doit être en mesure de spécifier les agents à qui les normes se destinent.

L'ouvrage de Royakkers (1998), intitulé *Extending deontic logic for the formalisation of legal rules*, est une version retravaillée de la thèse de doctorat de l'auteur, dont l'objectif est de formaliser les contradictions inhérentes au discours légal. La consistance *normative*, par opposition à la consistance *logique*, où (φ est logiquement consistant s'il n'y a pas d'énoncé ψ pour lequel ψ et $\neg\psi$ se trouvent à la fois dans l'ensemble des conséquences logiques de φ), est telle qu'on ne trouve pas d'énoncé ψ pour lequel à la fois $O\psi$ et $O\neg\psi$ sont dans l'ensemble des conséquences logiques de φ. La consistance normative s'exprime généralement à l'aide du schéma d'axiome (**D**) de la logique déontique standard, lequel est propositionnellement équivalent à l'énoncé suivant :

$$\neg(O\varphi \wedge O\neg\varphi)$$

Soulignons que si une logique admet cet axiome, alors un énoncé normativement inconsistant sera aussi par le fait-même logiquement inconsistant. En effet, si φ est normativement inconsistant, alors il y a ψ tel que $O\psi$ et $O\neg\psi$ se trouvent dans l'ensemble des conséquences de φ, et donc $O\psi \wedge O\neg\psi$ et $\neg(O\psi \wedge O\neg\psi)$ se trouveront aussi dans l'ensemble des conséquences de φ. Par conséquent, l'ajout de (**D**) permet de réduire l'inconsistance normative à l'inconsistance logique.

La principale innovation de Royakkers est d'augmenter la logique déontique afin de rendre compte des individus visés par les normes. Considérant que la logique déontique standard et que la logique déontique dynamique permettent respectivement de représenter les énoncés du type *ought-to-be* et *ought-to-do*, Royakkers propose d'augmenter chacun de ces systèmes. Dans ce qui suit, nous avons choisi de présenter brièvement son analyse de la logique déontique standard étant donné que cela nous amènera à expliciter certaines considérations relatives aux différentes interprétations de l'opérateur O. Notre présentation se restreindra aux innovations de sa proposition, principalement afin de nous concentrer sur son système de logique déontique dynamique ainsi que sur les applications qu'il en tire.

Dans le chapitre 4 de son livre, Royakkers propose d'indexer l'opérateur O de la logique déontique standard à un agent $_i$, lequel peut être considéré autant comme une constante qu'une variable sur laquelle on quantifie. L'auteur augmente la logique déontique standard en indexant l'opérateur à un agent $_i$, ce qui se traduit du côté de la sémantique par l'indexation de la relation R sur le modèle sémantique à un agent $_i$ et à l'augmentation du modèle par un ensemble d'agents Ag.[2]

[2] Cela est similaire à ce qu'on trouve en logique épistémique, où $\Box$ est indexé à un agent pour représenter la connaissance individuelle.

Cet opérateur représente l'obligation personnelle et permet de définir quatre autres opérateurs, ayant chacun une interprétation bien spécifique :

$$O^+(\varphi) =_{df} \forall_{i \in Ag} O_i(\varphi) \qquad \text{(df. } O^+\text{)}$$
$$P^+(\varphi) =_{df} \exists_{i \in Ag} P_i(\varphi) \qquad \text{(df. } P^+\text{)}$$
$$O^-(\varphi) =_{df} \exists_{i \in Ag} O_i(\varphi) \qquad \text{(df. } O^-\text{)}$$
$$P^-(\varphi) =_{df} \forall_{i \in Ag} P_i(\varphi) \qquad \text{(df. } P^-\text{)}$$

Ces opérateurs représentent respectivement l'obligation générale, la permission non spécifique, l'obligation non spécifique et la permission générale. L'obligation (ou la permission) générale est valable pour tous les agents, alors que l'obligation (ou la permission) non spécifique ne s'adresse qu'à au moins un individu (voir le texte pour les relations entre ces opérateurs).

Du côté sémantique, chaque opérateur sera représenté par une restriction. Par exemple, alors que nous avons $O\varphi$ vrai pour w si et seulement si φ est vrai pour tout v t.q. wRv, nous aurons $O_i\varphi$ vrai pour w si et seulement si φ est vrai pour tout v t.q. wR_iv. Dans le cas de O^+, nous obtenons que $O^+(\varphi)$ est vrai pour w si et seulement si pour tout i, φ est vrai pour tout v t.q. wR_iv. Pour O^-, il suffira d'avoir au moins un i et un v pour lequel wR_iv.

L'intérêt de l'approche de Royakkers est l'introduction de *groupes* d'agents. Il propose deux interprétations de la notion d'obligation collective, à savoir l'interprétation *stricte* et l'interprétation *faible*. L'interprétation stricte s'adresse à tout le groupe, alors que l'interprétation faible ne s'adresse qu'à un sous-groupe. Nous ne présenterons que la première. L'augmentation du modèle standard pour rendre compte des groupes d'agents se fait de manière similaire à ce que nous avons déjà décrit. En un mot, il suffit d'indexer la relation R à un groupe d'agents $_X$.

Alors que dans le cas de l'obligation personnelle O_i l'agent i est membre de l'ensemble d'agents Ag, l'obligation collective O_X met en jeu un groupe membre de l'ensemble des parties de Ag, c'est-à-dire que $X \in \wp(Ag)$ et donc $X \subseteq Ag$. Le modèle sémantique $\mathcal{M} = \langle W, Ag, R_{Ag}, a\rangle$ est défini par un ensemble d'agents $Ag \neq \varnothing$ et $R_{Ag} = \{R_X : X \in \wp(Ag)\}$ un ensemble de relations indexées à des groupes d'agents. W est un ensemble non vide de scénarios possibles et a est une fonction qui attribue des valeurs de vérité aux propositions dans W. La clause sémantique pour O_X est simplement que $O_X\varphi$ est vrai pour w si et seulement si φ est vrai pour tout v t.q. wR_Xv. Dans le cas de l'obligation collective stricte, Royakkers admet le schéma d'axiome (**D**), et donc R_X est sérielle.

Cela fait, Royakkers introduit les obligations collectives *fortes* ($O^{\oplus}$) et *faibles* ($O^{\ominus}$). Alors que l'obligation $O^{+}(\varphi)$ signifie que tous les agents ont l'obligation φ, $O^{\oplus}(\varphi)$ indique que tout groupe X possède une obligation *collective*. Ainsi, l'obligation s'applique à tout le groupe, pas à chaque individu. À partir de l'opérateur O indexé à un groupe d'agents $_X$, l'auteur définit quatre opérateurs d'obligations collectives, à l'instar des obligations personnelles générales et non spécifiques :

$$O^{\oplus}(\varphi) =_{df} \forall_{X \in \wp(Ag)} O_X(\varphi) \qquad \text{(df. } O^{\oplus}\text{)}$$
$$O^{\ominus}(\varphi) =_{df} \exists_{X \in \wp(Ag)} O_X(\varphi) \qquad \text{(df. } O^{\ominus}\text{)}$$
$$P^{\ominus}(\varphi) =_{df} \forall_{X \in \wp(Ag)} P_X(\varphi) \qquad \text{(df. } P^{\ominus}\text{)}$$
$$P^{\oplus}(\varphi) =_{df} \exists_{X \in \wp(Ag)} P_X(\varphi) \qquad \text{(df. } P^{\oplus}\text{)}$$

Les clauses sémantiques sont pareillement définies. Par exemple, $O^{\oplus}(\varphi)$ est vrai pour w si et seulement si pour tout X, φ est vrai pour tout v tel que wR_Xv. Le lecteur est invité à consulter le texte de Royakkers concernant les relations entre les différents opérateurs déontiques et les deux interprétations de l'obligation collective.

Passons maintenant à l'analyse de la logique déontique dynamique. Nous allons suivre la présentation de l'auteur, introduisant d'abord la logique déontique dynamique et la modifiant par la suite pour rendre compte des agents et des groupes d'agents. Royakkers utilise essentiellement l'approche proposée par Meyer (1988), qu'il modifiera afin de mieux servir son propos. Soit le langage suivant :

$$\mathcal{L} = \{(,), Act, Prop, \bar{}, \&, \cup, \neg, \supset, \wedge, \vee, [\,], \mathbf{any}, \mathbf{fail}, \mathbf{skip}, \mathbf{change}\}$$

D'emblée, on remarque que Royakkers ne s'intéresse qu'aux actions et exclut les séquences d'actions et les actions conditionnelles, représentées chez Meyer par les connecteurs ; et $\rightarrow$ /.

L'ensemble $Act = \{\mathbf{any}, \mathbf{change}, a_1, ..., a_n, ...\}$ contient un nombre dénombrable d'actions atomiques ainsi que les actions **any** (l'action universelle, « peu importe laquelle ») et **change** (« peu importe laquelle sauf **skip** »). L'ensemble $Act*$ d'actions complexes est défini récursivement à l'instar de chez Meyer et la lecture des connecteurs est similaire :

$$\alpha := a_i \mid \mathbf{fail} \mid \mathbf{skip} \mid \mathbf{any} \mid \mathbf{change} \mid \overline{\alpha} \mid \alpha \& \beta \mid \alpha \cup \beta$$

Alors que l'action **skip** n'apporte aucun changement au scénario, **fail** représente l'action qui échoue et après laquelle aucun scénario n'est accessible. Les énoncés bien formés sont définis récursivement :

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid [\alpha]\varphi$$

L'axiomatisation de la logique dynamique chez Royakkers se fait à l'instar de celle de PD_eL, où $[\alpha]$ est une modalité normale de type K. On ajoute une règle de substitution d'équivalences pour les actions (Subst) :

$$\frac{\alpha = \beta}{[\alpha]\varphi \equiv [\beta]\varphi} \qquad \text{(Subst)}$$

À cela s'ajoutent les axiomes pour les actions :

$$[\alpha \cup \beta]\varphi \equiv ([\alpha]\varphi \wedge [\beta]\varphi) \tag{A1}$$
$$([\alpha]\varphi \vee [\beta]\varphi) \supset [\alpha \& \beta]\varphi \tag{A2}$$
$$[\mathbf{skip}]\varphi \equiv \varphi \tag{A3}$$
$$[\mathbf{fail}]\varphi \tag{A4}$$

L'axiome (A1) indique que si un choix entre deux actions mène à une description φ, alors chaque action individuellement y mène. Le deuxième axiome signifie que si deux actions mène individuellement à une description φ, alors les deux actions faites conjointement y mènent aussi. L'axiome (A3) stipule que faire l'action **skip** équivaut à ne rien faire. Finalement, (A4) signifie que n'importe quoi est vrai lorsque le système échoue.

Du côté de la sémantique, Royakkers utilise, à l'instar de Meyer, la notion d'ensemble de *simultanéité* (*synchroniticity set*). D'entrée de jeu, $\{\mathbf{fail}\}$, $\{\mathbf{skip}\}$ et tout sous-ensemble non vide de *Act* sont des ensembles de simultanéité. Soit $S_1, ..., S_n, ...$ des ensembles de simultanéité et $T_1, ..., T_n, ...$ des ensembles d'ensembles de simultanéité. Maintenant, soit $T^{\mathbf{fail}}$ une opération sur les ensembles telle que :

$$T^{\mathbf{fail}} = \begin{cases} T - \{\{\mathbf{fail}\}\} \text{ s'il y a un } S_i \in T \text{ t.q. } S_i \neq \{\mathbf{fail}\} \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}$$

Autrement dit, $T^{\mathbf{fail}}$ est l'ensemble qui contient les ensembles de simultanéité qui ne contiennent pas l'action **fail** lorsqu'il y a au moins un ensemble de simultanéité différent de l'ensemble qui contient **fail**, et si aucun ensemble de simultanéité dans T est différent de celui qui contient **fail**, alors T est l'ensemble qui contient l'ensemble qui contient **fail**. Notons que $T - \{\{\mathbf{fail}\}\}$ est le complément de $\{\{\mathbf{fail}\}\}$ relativement à T. Les opérations sémantiques ensemblistes représentant respectivement les actions négatives, disjointes et conjointes sont données par $\sim, \sqcup$ et $\sqcap$:

$$\sim S = (\wp(Act) \cup \{skip\}) - S$$
$$\sim T = \sqcap\{\sim S : S \in T\}$$

$$T_1 \sqcup T_2 = (T_1 \cup T_2)^{\mathbf{fail}}$$
$$T_1 \sqcap T_2 = \begin{cases} T_1 \cap T_2 \text{ si } T_1 \cap T_2 \neq \varnothing \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}$$

En mots, l'opération $\sqcap$ est similaire à l'intersection ensembliste usuelle dans la mesure où l'intersection n'est pas vide, sans quoi $T_1 \sqcap T_2$ se résume à l'ensemble qui contient l'ensemble de simultanéité qui contient **fail**. L'opération $\sqcup$ est similaire à l'union ensembliste à l'exception qu'elle s'assure d'exclure l'ensemble de simultanéité **fail** lorsqu'au moins un ensemble de simultanéité de T_1 ou T_2 y est différent. Finalement, $\sim S$ pour un ensemble de simultanéité S est le complément de S relativement à l'ensemble des parties de Act joint avec l'ensemble qui contient **skip**, et $\sim T$ pour un ensemble d'ensembles de simultanéité équivaut à l'intersection $\sqcap$ de tous les $\sim S$ pour lesquels S est élément de T.

Au moyen de ces opérations, Royakkers définit récursivement la sémantique pour les connecteurs d'actions :

$$\begin{aligned} \|\alpha\| &= \{S_i : \alpha \in S_i\} \text{ pour un atome } \alpha \\ \|\overline{\alpha}\| &= \sim \|\alpha\| \\ \|\alpha \cup \beta\| &= \|\alpha\| \sqcup \|\beta\| \\ \|\alpha \& \beta\| &= \|\alpha\| \sqcap \|\beta\| \\ \|\mathbf{skip}\| &= \{\{\mathbf{skip}\}\} \\ \|\mathbf{fail}\| &= \{\{\mathbf{fail}\}\} \\ \|\mathbf{any}\| &= \wp(Act) \cup \{\{\mathbf{skip}\}\} \\ \|\mathbf{change}\| &= \wp(Act) \end{aligned}$$

Cela fait, on ajoute une fonction qui permet de déterminer les scénarios accessibles après l'accomplissement de certaines actions. Soit l'ensemble $\Omega = \wp(Act) \cup \{\{\mathbf{fail}\}\} \cup \{\{\mathbf{skip}\}\}$, W un ensemble de scénarios et $W*$ l'ensemble qui contient les sous-ensembles de W qui eux contiennent au plus un élément. La définition de $W*$ a pour objectif de pouvoir spécifier qu'aucun scénario n'est accessible après avoir posé l'action qui échoue. Royakkers introduit une fonction $\rho : (\Omega \times W) \longrightarrow W*$ respectant les conditions suivantes :

$$\rho(\{\mathbf{fail}\})(w) = \varnothing$$
$$\rho(\{\mathbf{skip}\})(w) = \{w\}$$

Autrement dit, il n'y a aucun monde possible accessible à w lorsque l'action **fail** est posée et le monde accessible à w après l'action **skip** est w lui-même.[3] À présumer que les ensembles de scénarios ne sont pas vides, l'auteur utilise ρ pour définir la fonction $R : \wp(W) \longrightarrow \wp(W)$ telle que pour $W_0 \subseteq W$ et $S_i \in \Omega$:

$$R(S_i)(W_0) = \bigcup_{w \in W_0} \rho(S_i)(w)$$
$$R(T_i)(W_0) = \bigcup_{S_i \in T_i} R(S_i)(W_0)$$

En ce sens, la fonction R permet de déterminer pour un ensemble de simultanéité et un ensemble de mondes possibles l'ensemble de tous les scénarios accessibles à la suite des actions posées dans S_i, de même que pour un ensemble d'ensembles de simultanéité T_i. À l'aide de cette fonction, Royakkers définit récursivement une fonction sémantique $\|\|\|_R$:

$$\|\alpha\|_R(w) = R(\|\alpha\|_R)(w)$$
$$\|\alpha\|_R(W_0) = \bigcup_{w \in W_0} \|\alpha\|_R(w)$$

Le lecteur aura certainement remarqué que cette approche est fort similaire à celle de Meyer. L'idée de base est la même : déterminer les ensembles qui indiquent les actions posées et définir une fonction qui permet d'isoler l'ensemble des alternatives possibles à un scénario (ou un ensemble de scénarios) relativement à un ensemble d'actions posées. Le modèle sémantique $\mathcal{M} = \langle W, Act, \|\|\|_R, a\rangle$ est défini habituellement avec W, un ensemble de scénarios, Act un ensemble d'actions atomiques, $\|\|\|_R$ une fonction qui isole les scénarios accessibles et a, une fonction qui attribue des valeurs de vérité aux propositions.[4] La clause sémantique pour l'opérateur dynamique est la même que chez Meyer, à savoir que $[\alpha]\varphi$ est vrai pour w si et seulement si φ est vrai pour tout scénario v accessible à w lorsque α est posée. Les autres clauses sémantiques sont définies de manière usuelle :

$$\models_w [\alpha]\varphi \Leftrightarrow \models_v \varphi \text{ pour tout } v \in \|\alpha\|_R(w)$$

[3] Notons qu'il y a une coquille dans la notation utilisée par Royakkers. Ce dernier écrit $\rho(\{\mathbf{skip}\})(w) = w$ plutôt que $\rho(\{\mathbf{skip}\})(w) = \{w\}$. Or ρ a pour image $W*$, qui est un ensemble d'ensembles.

[4] Précisément, le modèle est construit à partir de $Act \cup \{\mathbf{fail}, \mathbf{skip}\}$.

L'originalité de Royakkers est d'augmenter la logique déontique dynamique pour rendre compte des agents et groupes d'agents. Pour ce faire, il modifie la sémantique en introduisant de nouvelles définitions pour les ensembles de simultanéité. D'abord, il augmente le langage $\mathcal{L}$ à l'aide des connecteurs $\cup^\star$ et $\&^\star$ pour les actions individuelles et $\cup'$ et $\&'$ pour les actions collectives. Soit Ag un ensemble d'agents. Royakkers définit récursivement un ensemble Ev d'événements par la condition suivante :

si $i \in Ag$, alors

(a) si $\alpha \in Act*$, alors $[i : \alpha] \in Ev$;
(b) si $\alpha, \beta \in Ev$, alors $\overline{\alpha}, \alpha \cup^\star \beta, \alpha\&^\star\beta \in Ev$.

Notons une certaine ambiguïté quant à la notion d'*événement nié* $\overline{\alpha}$. En effet, Royakkers considère que $\overline{[i : \beta]} = [i : \overline{\beta}]$, et donc la négation se comporte de la même manière pour les événements et les actions, mais les autres connecteurs diffèrent. Par surcroît, cela signifie que « il est faux que i fait l'action β » est équivalent à « i fait l'action $\overline{\beta}$ », ce qui est contestable sur le plan philosophique. Cela va à l'encontre de la distinction entre *ne pas faire* (omettre) et *ne pas faire délibérément* (s'abstenir).

La même augmentation s'opère pour les événements collectifs Ev' (où $X \in \wp(Act)$ et Act est un ensemble non vide d'agents) :

1. si $\alpha \in Act*$, alors $[X : \alpha] \in Ev'$;
2. si $\alpha, \beta \in Ev'$, alors $\overline{\alpha}, \alpha \cup' \beta, \alpha\&'\beta \in Ev'$.

Dans chacun des cas, les ensembles $EBF^\star$ et EBF' sont définis respectivement par l'ajout des clauses suivantes :

1. si $\alpha \in Ev$, alors $[\alpha]\varphi \in EBF^\star$;
2. si $\alpha \in Ev'$, alors $[\alpha]\varphi \in EBF'$.

Au niveau axiomatique, l'auteur ne fait qu'ajouter les axiomes (A2*) et (A3*) pour les événements individuels et (A2') et (A3') pour les événements collectifs. Les autres schémas d'axiomes et règles s'appliquent aux modalités $[i : \alpha]$ et $[X : \alpha]$. Ces différents opérateurs d'action permettent de construire trois logiques dynamiques distinctes DDL, $DDL^\star$ et DDL' :

$$[\alpha \cup^\star \beta]\varphi \equiv ([\alpha]\varphi \wedge [\beta]\varphi) \qquad \text{(A2}^\star\text{)}$$
$$([\alpha]\varphi \vee [\beta]\varphi) \supset [\alpha\&^\star\beta]\varphi \qquad \text{(A3}^\star\text{)}$$
$$[\alpha \cup' \beta]\varphi \equiv ([\alpha]\varphi \wedge [\beta]\varphi) \qquad \text{(A2')}$$
$$([\alpha]\varphi \vee [\beta]\varphi) \supset [\alpha\&'\beta]\varphi \qquad \text{(A3')}$$

À l'instar des opérations $\sim$, $\sqcap$ et $\sqcup$, il définit les opérations $\sim^\star$, $\sqcap^\star$ et $\sqcup^\star$ pour les opérateurs individuels et $\sim'$, $\sqcap'$ et $\sqcup'$ pour les opérateurs collectifs. Dans le cas des opérateurs individuels, il faut d'abord reconsidérer les ensembles de simultanéité. Soit S_i^j un ensemble de simultanéité pour un agent i qui contient des événements (relatifs à i). En supposant une énumération de tous les ensembles de simultanéité possibles pour un agent i, l'exposant ^{j} permet de référer à l'un de ces ensembles. Soit S_i^* l'ensemble qui contient tous les ensembles de simultanéité pour l'agent i (incluant $[i : \mathbf{skip}]$). Soit Σ l'ensemble qui contient le produit cartésien de tous les S_i^* (où $i \in Ag$). Une *trace* t est un élément de Σ, c'est-à-dire une suite d'ensembles de simultanéité (qui peuvent mettre en jeu différents agents).

Tout cela peut informellement être considéré de la manière suivante. On considère d'abord tous les ensembles de simultanéité possibles pour un agent i. En faisant le produit cartésien de tous les ensembles de simultanéité pour tous les agents i, on obtient un ensemble de n-uplets (où Ag contient n agents) qui propose toutes les combinaisons d'actions possibles pour tous les agents. En ce sens, une trace t correspond à un n-uplet et représente les actions posées par chaque agent. Cela permettra d'évaluer les scénarios accessibles à la suite des actions de tous les agents.

Maintenant, soit T_i un ensemble de traces.[5] Les opérations sont définies à l'instar des opérations susmentionnées pour la logique déontique dynamique, à l'exception qu'on considère des traces plutôt que des ensembles de simultanéité :

$$T_i^{\mathbf{fail}\star} = \begin{cases} T_i - \{\{\mathbf{fail}\}\} \text{ s'il y a un } t \in T_i \text{ t.q. } t \neq \{\mathbf{fail}\} \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}$$

$$\begin{aligned} \sim^\star t &= \Sigma - \{t\} \\ \sim^\star T_i &= \sqcap^\star \{\sim^\star t : t \in T_i\} \\ T_1 \sqcup^\star T_2 &= (T_1 \cup T_2)^{\mathbf{fail}} \\ T_1 \sqcap^\star T_2 &= \begin{cases} T_1 \cap T_2 \text{ si } T_1 \cap T_2 \neq \varnothing \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases} \end{aligned}$$

[5] Les autres conditions pour les ensembles de simultanéité restent les mêmes, c'est-à-dire que $\{\mathbf{fail}\}$, $\{\mathbf{skip}\}$ et $[i : A]$ pour tout sous-ensemble non vide A de $Act*$ sont des ensembles de simultanéité.

Royakkers définit récursivement la sémantique pour les connecteurs d'actions et d'événements individuels :

$$\begin{aligned}
\|i : \alpha\| &= \{t \in \Sigma : \alpha \in S_i^j \text{ et } S_i^j \text{ est dans } t\} \text{ pour un atome } \alpha \\
\|\overline{\alpha}\| &= \sim^\star \|\alpha\| \\
\|i : \alpha \cup \beta\| &= \|i : \alpha\| \sqcup^\star \|i : \beta\| \\
\|i : \alpha \& \beta\| &= \|i : \alpha\| \sqcap^\star \|i : \beta\| \\
\|i : \overline{\alpha}\| &= \sim^\star \|i : \alpha\| \\
\|\alpha \cup^\star \beta\| &= \|\alpha\| \sqcup^\star \|\beta\| \\
\|\alpha \&^\star \beta\| &= \|\alpha\| \sqcap^\star \|\beta\| \\
\|i : \mathbf{skip}\| &= \{t \in \Sigma : \mathbf{skip} \in S_i^j \text{ et } S_i^j \text{ est dans } t\} \\
\|\mathbf{fail}\| &= \{\{\mathbf{fail}\}\} \\
\|\mathbf{any}\| &= \Sigma \\
\|\mathbf{change}\| &= \Sigma - \|i : \mathbf{skip}\|
\end{aligned}$$

Encore une fois, la signification de ces conditions est similaire à ce qu'on trouve en logique dynamique sans agent. La première clause indique que l'ensemble qui contient les événements où il est vrai que i pose l'action atomique α est l'ensemble qui contient les traces dans lesquelles se trouvent les ensembles de simultanéité qui contiennent α pour i. L'événement nié $\overline{\alpha}$ est représenté par le complément de α relativement à Σ. Les événements qui mettent en jeu des actions disjointes et conjointes s'interprètent de la même manière, à l'exception qu'on parle de trace d'ensemble de simultanéité plutôt que d'ensembles de simultanéité. Les événements disjoints et conjoints se représentent sémantiquement par les opérations ensemblistes définies plus haut, et finalement, les constantes d'actions s'interprètent trivialement selon leur signification respective.

En somme, du point de vue syntaxique, la notation change, mais l'interprétation sémantique requiert simplement qu'on prenne en compte des combinaisons d'actions possibles pour tous les agents, par opposition avec de simples ensembles de simultanéité. En un certain sens, nous aurions pu parler d'*ensembles collectifs de simultanéité*, qui représenteraient toutes les actions posées par tous les agents à un moment donné et équivaudraient à la notion de trace.

On ajoute au modèle sémantique $\mathcal{M} = \langle W, Act, Ag, \|\|_R, a\rangle$ un ensemble non vide d'agents, où $\|\|_R$ est redéfinie en termes de traces :

$$\begin{aligned}
\rho(\{\mathbf{fail}\})(w) &= \varnothing \\
R(t)(w) &= \rho(t)(w) \\
R(T_i)(w) &= \{v : v = R(t)(w) \text{ pour tous les } t \in T_i\}
\end{aligned}$$

$$\|\alpha\|_R(w) = R(\|\alpha\|)(w)$$

La fonction ρ identifie *le* scénario accessible à w lorsque toutes les actions (atomiques) d'une trace t sont réalisées par les agents. Évidemment, l'action qui échoue ne mène à aucun scénario possible. Lorsqu'une trace met en jeu des événements complexes, l'ensemble des scénarios accessibles à w est donné par $R(T_i)(w)$, qui contient les scénarios accessibles à chaque trace t de T_i. La clause sémantique pour $[i : \alpha]\varphi$ est identique à celle de $[\alpha]\varphi$, à l'exception des remarques susmentionnées par rapport à $\|\alpha\|_R(w)$. Il est vrai pour w que i pose l'action α, ce qui rend la description φ vraie, à condition que φ le soit pour tout scénario accessible à w après que i a posé l'action α.

Dans le cas de la logique dynamique mettant en jeu des groupes d'agent, la sémantique de DDL' est identique à celle de $DDL^\star$, à l'exception du fait qu'on redéfinit les opérations sur des traces collectives plutôt que sur des traces. Que ce soit pour DDL, $DDL^\star$ ou DDL', la définition de l'interdiction se fait à l'instar de chez Meyer par une réduction de type Anderson-Kanger, où V est une constante de violation et $[\,]$ la modalité appropriée :

$$F(\alpha) =_{df} [\alpha]V \qquad \text{(df. } F\text{)}$$

Notons au passage que Royakkers propose la même analyse pour la logique déontique dynamique que pour la logique déontique standard concernant les obligations générales, non spécifiques, collectives fortes et collectives faibles. À ce sujet, le lecteur peut consulter le chapitre 5 de l'ouvrage de Royakkers.

En guise de conclusion, notons que les travaux de Royakkers sont fortement inspirés de ceux de Dignum *et al.* (1996), qui utilisent une logique dynamique augmentée de certaines modalités temporelles, mais ne formalise pas les agents et groupes d'agents. Par ailleurs, Hughes et Royakkers (2008) utilisent la logique déontique dynamique pour rendre compte des obligations qui perdurent et qui n'ont pas de « date d'expiration ». Les auteurs introduisent une constante I de *dette* et définissent $O\alpha$ comme $[\overline{\alpha}]I$, c'est-à-dire que tant que α n'est pas réalisée, l'agent reste en *dette*. Le lecteur peut consulter l'article de Royakkers (2000) pour une critique de la logique STIT en faveur de la logique dynamique.[6]

[6] Dans cet article, l'auteur intègre le connecteur $\Rightarrow_s$ de Jones et Sergot (1996) à la logique dynamique pour actions collectives.

Kirster Segerberg

Krister Segerberg est un logicien qui a surtout travaillé en logique modale, principalement en logique dynamique épistémique ainsi qu'en logique de l'action. Il offre une logique déontique dynamique enrichie beaucoup plus subtile que celle proposée par Meyer. L'originalité de son approche est d'augmenter la logique dynamique avec des modalités temporelles, l'objectif étant de fournir une représentation formelle adéquate de l'*action*. Nous présenterons ici sa sémantique de 2009. Le lecteur peut aussi consulter ses écrits de 2007 et 2012 pour l'essentiel de sa position en logique déontique dynamique.

Segerberg propose un modèle qui comprend un univers U (non vide), lequel contient des *points*. Les suites non vides (finies ou infinies) de points sont nommées des *traces* (ou chemins, en anglais *path*), auxquelles l'auteur fait référence par les lettres p, q, r, etc.[7]

Deux traces p et q peuvent être combinées en une trace pq à condition que le dernier point de p soit le premier point de q. L'ensemble Ev est un ensemble d'*événements*[8], lequel est aussi considéré comme un ensemble d'actions (i.e. un ensemble de choses réalisées). Un événement $a \in Ev$ (où $Ev \subseteq \wp(U)$) est tel que les événements de Ev ne contiennent que des traces finies. En ce sens, U contient des points, $\wp(U)$ contient des traces et $\wp(\wp(U))$ contient des événements.

Notons que Segerberg ne fait pas la distinction pas entre une action et un événement et accorde la même signification à ces deux termes. Rappelons-nous qu'en logique dynamique, une action est généralement associée à un ensemble qui contient toutes les manières de la réaliser. Ici, une action, voire un événement, correspond à un ensemble de traces, où une trace est une manière de réaliser une action. L'ensemble Ev correspond donc à un ensemble d'actions.

Informellement, une action a peut être réalisée (advenir) de plusieurs manières. Par exemple, si la porte est ouverte, alors Paul peut ouvrir la porte, la garder ouverte avec son pied, mettre un objet pour l'empêcher de se fermer, etc. Plusieurs traces différentes peuvent donc faire advenir le même événement (la même action). L'événement est formellement conçu comme un ensemble qui contient toutes les traces (finies, sans quoi l'événement n'adviendrait jamais) qui peuvent y mener.

Les événements *trivial* et *impossible* correspondent respectivement à **any** et **fail** chez Royakkers, ou encore à U et $\varnothing$ chez Meyer.

[7] Attention à ne pas confondre les traces avec les variables propositionnelles.

[8] À ne pas confondre avec la notion d'événement chez Royakkers, qui associe un agent ou groupe d'agents à une action.

L'ensemble H est un ensemble d'*histoires* maximales (à comprendre de la même manière que dans le cas des logiques temporelles), où une histoire $hg = (h, g)$ est une trace infinie pour laquelle le dernier point de h est le premier point de g et représente le présent (et donc h et g représentent respectivement le passé et le futur). L'ensemble $cont(h) = \{g : hg \in H\}$ dénote les continuations possibles (les futurs possibles) de h.

Le langage contient un ensemble dénombrable de propositions atomiques $Prop = \{P_1, ..., Pn, ...\}$ (en majuscule pour ne pas confondre avec les traces) et un ensemble dénombrable de termes $T = \{e_1, ..., e_n, ...\}$ qui représentent des événements atomiques. Le modèle $\mathcal{M} = \langle U, Ev, H, V\rangle$ contient les ensembles susmentionnés (l'univers U, un ensemble d'événements Ev et un ensemble d'histoires H) ainsi qu'une fonction d'évaluation V qui assigne des valeurs de vérité aux propositions.

Formellement, $V : (Prop \cup T) \longrightarrow \wp(U)$ déterminent les ensembles à l'intérieur desquelles se trouvent (respectivement) les atomes propositionnels et les termes. Informellement, V détermine les scénarios où les atomes propositionnels sont vrais et les termes sont vrais, ce qui est dénoté par $V(\varphi) = \|\varphi\|$. La fonction est définie récursivement de façon habituelle pour les propositions (où $\varphi, \psi \in EBF_{prop}$, les énoncés bien formés étant définis usuellement) :

$$\begin{aligned} V(P_i) &= \|P_i\| \\ \|\neg\varphi\| &= U - \|\varphi\| \\ \|\varphi \vee \psi\| &= \|\varphi\| \cup \|\psi\| \\ \|\varphi \wedge \psi\| &= \|\varphi\| \cap \|\psi\| \end{aligned}$$

Quant aux actions (où $\alpha, \beta \in EBF_T$ pour un ensemble de termes bien formés)[9], nous avons la définition récursive suivante :

$$\begin{aligned} V(e_i) &= \|e_i\| \\ \|\alpha + \beta\| &= \|\alpha\| \cup \|\beta\| \\ \|\alpha; \beta\| &= \|\alpha\| \uparrow \|\beta\| \\ \|\alpha;; \beta\| &= \|\alpha\| \Uparrow \|\beta\| \end{aligned}$$

[9] La définition des énoncés bien formés n'est pas explicite chez Segerberg. Néanmoins, considérant que celui-ci introduira des opérateurs dynamiques, on peut en comprendre que ceux-ci sont définis de manière usuelle en logique dynamique.

Celle-ci utilise les ensembles définis de la manière suivante (où A et B contiennent des traces finies) :

$A \uparrow B = \{pq : p \in A, q \in B \text{ et } pq \text{ est une combinaison}\}$
$A \Uparrow B = \{prq : p, r \in A, q \in B \text{ et } pr \text{ et } rq \text{ sont des combinaisons}\}$

Le connecteur + permet d'ouvrir des choix possibles pour les suites d'événements et les connecteurs ; et ; ; permettent de concaténer respectivement des traces finies et infinies. En ce sens, + représente l'action disjonctive et ; la séquence d'actions.

Cela fait, Segerberg introduit des opérateurs (temporels) modaux pour représenter le passé $\overleftarrow{P}$, le futur $\vec{F}$ et la nécessité historique $\Box$.[10] Un scénario peut être représenté par une histoire (h, g). Soit w le point (présent) qui constitue le dernier point de h et le premier point de g. Les clauses sémantiques pour ces opérateurs sont :

$$\models_{(h,g)} \varphi \Leftrightarrow w \in \|\varphi\| \tag{1}$$
$$\models_{(h,g)} \vec{F}\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } v, h', g' \text{ t.q. } hv = h' \text{ et } vg' = g \tag{2}$$
$$\models_{(h,g)} \overleftarrow{P}\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } v, h', g' \text{ t.q. } h'v = h \text{ et } g' = vg \tag{3}$$
$$\models_{(h,g)} \Box\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } g' \in cont(h) \tag{4}$$

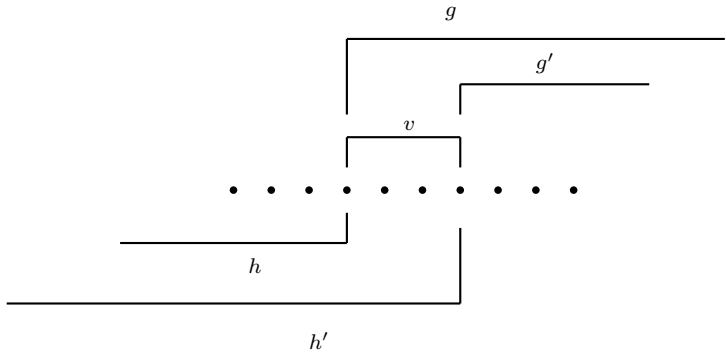
Les clauses (1) et (4) vont de soi : une proposition est vraie pour un scénario lorsque ce scénario fait partie de l'ensemble qui contient les scénarios où la proposition est vraie et une proposition est historiquement nécessairement vraie pour un scénario w relativement à un passé h à condition que la proposition soit vraie pour tout futur possible accessible à w relativement à h. Les clauses (3) et (4) quant à elles se comprendront mieux à la lumière de la figure 10.1.

En un mot, l'idée est de dire qu'une proposition $\vec{F}\varphi$ est vraie pour un scénario w d'une histoire hg à condition que φ soit vrai pour tout scénario v de la même histoire qui vient après w. Soulignons qu'un scénario est un ensemble de points, et qu'une histoire peut être vue comme la représentation de l'évolution d'un scénario (global) où chaque point représente un état statique, voire une description, du scénario à un moment donné.

Dans la figure 10.1, on voit que $hv = h'$ et que $g = vg'$. Ainsi, une proposition $\vec{F}\varphi$ est vraie lorsque φ est vrai dans tous les développements possibles d'une histoire. Notons qu'une proposition est vraie pour un futur relativement à un passé spécifique. La proposition $\vec{F}\varphi$ est vraie lorsque φ est vrai pour tous les déploiements possibles de w.

[10] Ici, nous avons modifié la notation de Segerberg afin de ne pas confondre avec les opérateurs dynamiques.

FIGURE 10.1
Futur



La clause pour le passé se comprend dans le même ordre d'idées. Il est vrai à un moment w pour une histoire hg que φ était vrai à condition que φ soit vrai pour tout moment qui vient avant w sur hg.

En plus de ces opérateurs temporels, Segerberg définit un opérateur binaire $\vec{U}$ qui signifie *jusqu'à* (*until*). L'énoncé $\vec{U}_{\varphi}\psi$ signifie ψ *jusqu'à ce que* φ. L'interprétation sémantique de cet énoncé est similaire à celle du *à moins que* en logique propositionnelle :

$$\begin{aligned} \models_{(h,g)} \vec{U}_{\varphi}\psi \Leftrightarrow\ & \text{soit} \\ & 1.\ \not\models_{(h',g')} \varphi \text{ pour tout } h', g' \text{ t.q. } h'g' = hg \text{ ou} \\ & 2.\ \models_{(h',g')} \varphi \text{ pour un } v \text{ t.q. } h' = hv \text{ et } g = vg' \text{ et} \\ & \quad \models_{(h'',g'')} \psi \text{ pour tout } x \text{ t.q. } h'' = hx \text{ et } g = xg'' \end{aligned} \tag{5}$$

Autrement dit, il est vrai pour w relativement à hg que ψ *jusqu'à ce que* φ à condition que soit 1. φ soit faux pour tout scénario v de hg ou 2. φ est vrai pour un v de hg qui suit w et que ψ est vrai pour tout x de hg qui vient avant w.

Cela nous amène aux opérateurs dynamiques. Alors qu'en logique dynamique propositionnelle, on peut simplement exprimer qu'une description est vraie après qu'une action est réalisée, l'approche de Segerberg permet de rendre compte de beaucoup plus de subtilités. En effet, cet auteur introduit les opérateurs **does**, **realises**, **occurs**, **done**, **realised** et **occured**.[11] Pour des raisons d'économie (et parce que leur interprétation sémantique est triviale), nous ne présenterons pas en détail la sémantique de ces opérateurs. Soulignons simplement que leur signification est univoque et se

[11] À ce sujet, soulignons que la signification de ces opérateurs est beaucoup plus claire dans Segerberg (2012) que dans Segerberg (2009).

traduit directement à l'aide des modalités temporelles développées jusqu'à présent. La différence entre **does** et **realises** est que le premier s'applique aux actions alors que le second s'applique aux propositions, au même titre que **occurs** et **occured** s'appliquent seulement à des actions.

En ce qui concerne les notions déontiques, Segerberg mentionne que son approche, reposant sur une logique dynamique, s'insère dans le cadre des approches de type *ougth-to-do*. D'abord, l'auteur considère les normes *simples*, c'est-à-dire les normes qui assignent à un passé plusieurs futurs idéaux possibles. Étant donné que $cont(h)$ correspond à l'ensemble des futurs possibles à la suite du passé h, $N(h) \subseteq cont(h)$ isole les futurs accessibles idéaux. Cet ensemble doit satisfaire à la condition suivante, qui indique la cohérence normative :

$$g = vg' \Rightarrow [g' \in N(hv) \Rightarrow g \in N(h)]$$

En d'autres termes, si g' est un futur idéal accessible à l'histoire hv, alors g est un futur accessible à l'histoire h lorsque g est vg'. Segerberg adopte aussi la condition de sérialité (i.e. $N(h) \neq \varnothing$), mais souligne que celle-ci pourrait très bien ne pas être incluse. De fait, le schéma d'axiome (**D**) est valide. Cela fait, l'auteur est en mesure d'introduire quatre modalités déontiques *locales* et quatre modalités déontiques *normales*. Les quatre modalités sont l'*obligation*, la *permission*, l'*interdiction* et l'*omission*. Commençons d'abord par les modalités locales. Soit les quatre conditions sémantiques suivantes :

$\models_{(h,g)} O\alpha \Leftrightarrow$ (6)

1. pour tout $g' \in N(h)$ il y a $v \in \|\alpha\|$ (v sous-trace de g')
2. pour tout $v \in \|\alpha\|$ il y a $g' \in N(h)$(v sous-trace de g')

$\models_{(h,g)} P\alpha \Leftrightarrow$ (7)

2. pour tout $v \in \|\alpha\|$ il y a $g' \in N(h)$(v sous-trace de g')

$\models_{(h,g)} F\alpha \Leftrightarrow$ (8)

3. pour tout $g' \in N(h)$ et pour tout $v \in \|\alpha\|$
   v n'est pas sous-trace de g'

$\models_{(h,g)} OMMI\alpha \Leftrightarrow$ (9)

4. il y a un $g' \in N(h)$ t.q. pour tout $v \in \|\alpha\|$
   v n'est pas sous-trace de g'

En mots, il est vrai que α est obligatoire relativement à (h, g) si et seulement si 1. tout futur accessible à partir de h contient au moins une partie v membre de l'ensemble qui contient les scénarios où α est réalisé et 2. pour tout scénario v qui est membre de l'ensemble qui contient les scénarios où α est réalisé il y a au moins une trace g' membre de l'ensemble des futurs idéaux à partir de h et dont v fait partie.

Il est vrai que α est permis si et seulement si pour tout scénario v qui est membre de l'ensemble qui contient les scénarios où α est réalisé il y a au moins une trace g' membre de l'ensemble des futurs idéaux à partir de h et dont v fait partie.

Il est vrai que α est interdit si et seulement si pour tout futur idéal accessible à h et pour tout scénario v membre de l'ensemble qui contient les scénarios où α est réalisé, v ne fait pas partie de g'. Autrement dit, si α est interdit, alors chaque futur idéal accessible ne contient pas de traces qui font partie de l'ensemble qui contient les traces où α est réalisé.

Finalement, α peut être omis si et seulement s'il y a au moins un futur accessible idéal qui ne contient pas de trace qui fait partie de l'ensemble qui contient les traces où α est réalisé.

Dans le cas des modalités normales, nous avons besoin de la condition suivante, à laquelle nous ferons référence par (C) : « pour toute trace finie v telle que hv est une combinaison (donc, le dernier point de h est le premier point de v), s'il n'existe pas de trace x élément de $\|\alpha\|$ telle que x est une sous-trace de v, alors ... » :

$\models_{(h,g)} O^{\star}\alpha \Leftrightarrow (C)$ (6')

1. pour tout $g' \in N(hv)$ il y a $x \in \|\alpha\|$ (x sous-trace de g')
2. pour tout $x \in \|\alpha\|$ il y a $g' \in N(hv)$(x sous-trace de g')

$\models_{(h,g)} P^{\star}\alpha \Leftrightarrow (C)$ (7')

2. pour tout $x \in \|\alpha\|$ il y a $g' \in N(hv)$(x sous-trace de g')

$\models_{(h,g)} F^{\star}\alpha \Leftrightarrow (C)$ (8')

3. pour tout $g' \in N(hv)$ et pour tout $x \in \|\alpha\|$
   x n'est pas sous-trace de g'

$\models_{(h,g)} OMMI^{\star}\alpha \Leftrightarrow (C)$ (9')

4. il y a un $g' \in N(hv)$ t.q. pour tout $x \in \|\alpha\|$
   x n'est pas sous-trace de g'

Chaque clause utilise la condition (C), laquelle indique que s'il n'y a pas de sous-trace x incluse dans v pour hv (pour toute trace v), alors les conditions 1–4 doivent être remplies selon les cas. Autrement dit, s'il n'y a pas de v tel que x est une sous-trace de v et x est membre de l'ensemble des traces où α est réalisé, alors : dans le cas de l'obligation normale, cela implique que tout futur accessible à hv contient au moins une trace qui fait partie de l'ensemble des traces où α est réalisé et que toute trace qui fait partie de l'ensemble des traces où α est réalisé fait partie d'au moins un futur accessible à hv ; pour la permission, cela implique que toute trace qui fait partie de l'ensemble des traces où α est réalisé fait partie d'au moins un futur accessible à hv ; pour l'interdiction, cela signifie que les traces qui font partie de l'ensemble qui contient les traces où α est réalisé ne sont pas des sous-traces des futurs accessibles à hv ; finalement, dans le cas de l'omission, cela implique qu'au moins un futur accessible à hv ne contient aucune trace dans laquelle α est réalisé.

Après avoir considéré les normes *simples*, Segerberg en vient à considérer les normes *complètes*, la principale distinction étant que la dernière ne sélectionne pas l'ensemble des futurs idéaux accessibles, mais plutôt un sous-ensemble (préférable) de l'ensemble des futurs idéaux accessibles. Le formalisme est similaire à celui des normes simples, à l'exception du fait qu'on considère une histoire (h, g) relativement à un $S \subseteq N(h)$ où tout futur accessible membre de S est préférable aux futurs accessibles pas membres de S. Cela se fait notamment en introduisant un opérateur qui permet de sélectionner un ensemble de scénarios où une certaine proposition est vraie.

Tout compte fait, l'approche de Segerberg concernant la logique déontique dynamique a ses mérites, incluant le fait d'introduire plusieurs modalités dynamiques pour rendre compte des modes d'actions. En plus de la richesse de son langage, un aspect intéressant de l'approche de Segerberg est que celui-ci propose une sémantique pour les logiques dynamiques différente de celle de Meyer. Plutôt que de définir des ensembles de simultanéité et des mondes possibles, Segerberg fait d'une pierre deux coups en considérant les mondes possibles comme des suites d'événements.

Outre ses travaux en logique dynamique, Segerberg a aussi proposé une formalisation de la logique déontique reposant sur une logique de l'action axée sur une algèbre de Boole. Nous reviendrons sur ses travaux dans le chapitre 12. Ils ont notamment été repris par Demolombe (2014), lequel supplémente le cadre de travail de Segerberg pour rendre comte des obligations avec une date de tombée.[12]

[12] Concernant les obligations sujettes à certains *deadlines*, voir Broersen *et al.* (2004).

Ron van der Meyden

Alors que la plupart des logiciens se concentrent sur la formalisation de la notion d'*obligation*, certains sont plutôt portés à analyser la *permission*. Van der Meyden (1996), par exemple, a proposé un cadre de travail différent de celui de Meyer pour rendre compte autant des permissions de *choix* (*free-choice permission*) [13] que des permissions qui résultent de ce qui n'est pas interdit.[14]

D'emblée, il est important de mentionner que van der Meyden utilise l'opérateur dynamique $\langle\alpha\rangle$ comme primitif plutôt que $[\alpha]$, ce qui se comprend, considérant qu'il insiste sur la permission plutôt que sur l'obligation (la permission étant souvent définie comme l'opérateur dual de l'obligation). Il est aussi important de souligner que cette approche provient de l'informatique plutôt que de la logique philosophique. En effet, l'*action* est plutôt considérée comme un *programme* qui, à la suite de son exécution, fait que le système est dans un certain état φ (φ étant la description de l'état dans lequel le système se trouve à la suite de l'exécution du programme).

Le langage $\mathcal{L}$ contient un ensemble dénombrable d'actions atomiques $Act = \{a_1, ..., a_n, ...\}$, un ensemble dénombrable de propositions atomiques $Prop = \{p_1, ..., p_n, ...\}$ et les connecteurs d'actions représentant respectivement le choix, la séquence d'actions et l'itération d'action. Ces connecteurs, contrairement à ceux de Meyer, sont dans la lignée de la définition initiale de la logique dynamique. L'itération représente l'action *répétée*, où a^* est l'action obtenue en répétant a un nombre fini de fois.

$$\mathcal{L} = \{(,), Act, Prop, \cup, ;, ^*, \vee, \neg, \langle\,\rangle, \diamond, \pi\}$$

Outre les connecteurs pour les actions, on trouve les connecteurs propositionnels usuels, l'opérateur dynamique $\langle\alpha\rangle\varphi$, qui se lit « φ est vrai après certaines exécutions de α » et les opérateurs déontiques $\diamond$ et π, où $\diamond(\alpha, \varphi)$ et $\pi(\alpha, \varphi)$ signifient respectivement que φ est vrai à la suite de certaines exécutions de α qui ne sont pas interdites et que toute exécution de α qui mène φ est permise. L'opérateur dynamique $[\alpha]\varphi$ est défini par $\neg\langle\alpha\rangle\neg\varphi$ et signifie que φ est vrai dans tous les scénarios possibles à la suite de l'exécution de α. Finalement, $O(\alpha, \varphi)$ est utilisé comme abréviation de $\neg\diamond(\alpha, \neg\varphi)$ et signifie que toute exécution permise (non interdite) de α mène à un état décrit par φ. Les connecteurs propositionnels sont définis normalement.

Les énoncés bien formés d'action et de proposition sont respectivement définis récursivement de la manière suivante :

$$\alpha := a_i \mid \alpha \cup \beta \mid \alpha; \beta \mid \alpha^*$$

[13] Pour une analyse de la notion de *free-choice permission*, le lecteur est invité à consulter Anglberger *et al.* (2014) et Anglberger *et al.* (2015).

[14] La seconde notion renvoie à la permission faible.

$$\varphi := p_i \mid \neg\varphi \mid \varphi \vee \psi \mid \langle\alpha\rangle\varphi \mid \diamond(\alpha, \varphi) \mid \pi(\alpha, \varphi)$$

L'axiomatisation se fait en quatre temps. D'abord, on introduit les axiomes pour la modalité dynamique. Pour ce faire, on utilise le symbole $\perp$ pour référer à une contradiction arbitraire (p. ex. $p \wedge \neg p$) :

$$\langle\alpha\rangle\perp \equiv \perp \tag{1}$$
$$\langle\alpha;\beta\rangle\varphi \equiv \langle\alpha\rangle\langle\beta\rangle\varphi \tag{2}$$
$$\langle\alpha \cup \beta\rangle\varphi \equiv (\langle\alpha\rangle\varphi \vee \langle\beta\rangle\varphi) \tag{3}$$
$$\langle\alpha^*\rangle\varphi \equiv (\varphi \vee \langle\alpha;\alpha^*\rangle\varphi) \tag{4}$$
$$\langle\alpha\rangle(\varphi \vee \psi) \equiv (\langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi) \tag{5}$$
$$(\varphi \wedge [\alpha^*](\varphi \supset [\alpha]\varphi)) \supset [\alpha^*]\varphi \tag{6}$$

Le premier axiome signifie qu'une contradiction prend place à la suite de α si et seulement s'il y a déjà une contradiction dans le système. Comme l'auteur le souligne, les axiomes (2)–(5) permettent d'éliminer les actions complexes de la portée de l'opérateur dynamique.[15] L'axiome (2), comme chez Meyer, signifie que φ est vrai à la suite de l'action complexe « α suivie de β » si et seulement si « φ est vrai à la suite de β » est vrai à la suite de α. (3) stipule qu'un choix entre α ou β mène à φ à condition que chaque action y mène individuellement. Le quatrième axiome signifie que si φ est vrai à la suite d'une itération finie de α, alors soit φ est vrai soit φ est le résultat de α suivie d'une itération finie de α. (5) indique que α mène à φ ou ψ si et seulement si α mène à φ ou α mène à ψ, et finalement (6) signifie que si φ est vrai et que « φ implique que φ est le résultat de toute exécution de α » est le résultat d'une séquence finie d'exécutions de α, alors φ est le résultat d'une séquence finie d'exécutions de α.

Dans un deuxième temps, van der Meyden propose d'axiomatiser la modalité *non interdit*, c'est-à-dire la permission faible, de la même manière que *après* (i.e. $\langle\,\rangle$) :

$$\diamond(\alpha, \varphi) \supset \langle\alpha\rangle\varphi \tag{7}$$
$$\diamond(\alpha;\beta, \varphi) \equiv \diamond(\alpha, \diamond(\beta, \varphi)) \tag{8}$$
$$\diamond(\alpha \cup \beta, \varphi) \equiv \diamond(\alpha, \varphi) \vee \diamond(\beta, \varphi) \tag{9}$$
$$\diamond(\alpha^*, \varphi) \equiv \varphi \vee \diamond(\alpha; \alpha^*, \varphi) \tag{10}$$
$$\diamond(\alpha, \varphi \vee \psi) \equiv \diamond(\alpha, \varphi) \vee \diamond(\alpha, \psi) \tag{11}$$
$$(\varphi \wedge O(\alpha^*, \varphi \supset O(\alpha, \varphi))) \supset O(\alpha^*, \varphi) \tag{12}$$

[15] Sur cet aspect, il n'est pas clair pour nous de voir l'utilité d'introduire des actions complexes dans le langage considérant qu'elles peuvent être éliminées par l'utilisation de schémas d'axiomes appropriés. Autrement dit, le langage pourrait très bien exclure les connecteurs pour actions.

(7) signifie que s'il n'est pas interdit que φ soit le résultat de (d'une certaine exécution de) α, alors φ est le résultat d'une certaine exécution de α. L'axiome (8) indique que s'il n'est pas interdit que φ soit le résultat de l'exécution de la séquence « α suivie de β », alors il n'est pas interdit que « il n'est pas interdit que φ soit le résultat de β » soit le résultat de α. (9) stipule que s'il n'est pas interdit que φ soit le résultat du choix entre α ou β, alors soit il n'est pas interdit que φ soit le résultat de α soit il n'est pas interdit que φ soit le résultat de β. Le dixième axiome signifie que s'il n'est pas interdit que φ soit le résultat d'une suite d'exécutions de α, alors soit φ est d'emblée vrai soit il n'est pas interdit que φ soit le résultat de α suivie d'une suite d'exécutions de α. L'axiome (11) indique que s'il n'est pas interdit que φ ou ψ soit vrai à la suite de α, alors soit il n'est pas interdit que φ soit le résultat de α soit il n'est pas interdit que ψ soit le résultat de α. En dernier lieu, (12) signifie que si φ est vrai et que toute séquence d'exécutions non interdite de α mène à un scénario où la proposition « si φ est vrai, alors toute exécution non interdite de α mène à φ » est vraie, alors toute séquence d'exécutions non interdite de α mène à un scénario où φ est vrai.

En troisième lieu, l'axiomatisation de la permission se fait aussi de manière similaire à celle de l'opérateur dynamique :

$$[\alpha]\neg\varphi \supset \pi(\alpha, \varphi) \tag{13}$$
$$\pi(\alpha; \beta, \varphi) \equiv \pi(\alpha, \langle\beta\rangle\varphi) \wedge [\alpha]\pi(\beta, \varphi) \tag{14}$$
$$\pi(\alpha \cup \beta, \varphi) \equiv \pi(\alpha, \varphi) \wedge \pi(\beta, \varphi) \tag{15}$$
$$\pi(\alpha^*, \varphi) \equiv \pi(\alpha; \alpha^*, \varphi) \tag{16}$$
$$\pi(\alpha, \varphi \vee \psi) \equiv \pi(\alpha, \varphi) \wedge \pi(\alpha, \psi) \tag{17}$$
$$[\alpha^*]\pi(\alpha, \langle\alpha^*\rangle\varphi) \supset \pi(\alpha^*, \varphi) \tag{18}$$

L'axiome (13) signifie que si toute exécution de α mène à $\neg\varphi$, alors toute exécution de α qui mène à φ est permise, ce qui ne pose pas problème puisque par hypothèse, aucune exécution de α ne mènera à φ. (14) indique que si toute exécution de « α suivie de β » qui mène à φ est permise, alors toute exécution de α qui mène à « certaines exécutions de β mènent à φ » est permise et toute exécution de α mène à un scénario où toute exécution où β mène à φ est permise. (15) signifie que si toute exécution de α ou β qui mène à φ est permise, alors toute exécution de α qui mène à φ est permise et toute exécution de β qui mène à φ est permise. (16) stipule que si toute séquence d'exécutions de α qui mène à φ est permise, alors toute action α suivie d'une séquence d'exécutions de α qui mène à φ est permise. L'axiome (17) signifie que si toute exécution de α qui mène à φ ou ψ est permise, alors toute exécution de α qui mène à φ est permise et toute exécution de α qui mène à ψ est permise. Finalement, (18) signifie que si après toute exécution

d'une séquence d'exécutions de α il en résulte que toute exécution de « α qui mène à un scénario où certaines exécutions de séquences d'exécutions de α mènent à φ » est permise, alors toute séquence d'exécutions de α qui mène à φ est permise.

La quatrième étape consiste à établir des liens entre les modalités :

$$(\pi(\alpha, \varphi) \wedge \langle\alpha\rangle\varphi) \supset \diamond(\alpha, \varphi) \tag{19}$$

$$(\diamond(\alpha, \varphi) \wedge [\alpha](\varphi \supset \psi)) \supset \diamond(\alpha, \psi) \tag{20}$$

$$(\pi(\alpha, \psi) \wedge [\alpha](\varphi \supset \psi)) \supset \pi(\alpha, \varphi) \tag{21}$$

L'axiome (19) signifie que si toute exécution de α qui mène à φ est permise et que certaines exécutions de α mènent à φ, alors certaines exécutions de α qui mènent à φ ne sont pas interdites. (20) signifie que si certaines exécutions de α qui mènent à φ ne sont pas interdites et que toute exécution de α mène à un scénario où φ implique ψ, alors certaines exécutions de α qui mènent à ψ ne sont pas interdites. L'axiome (21) quant à lui signifie que si toute exécution de α qui mène à ψ est permise et que toute exécution de α mène à un scénario où φ implique ψ, alors toute exécution de α qui mène à φ est permise.[16]

La sémantique proposée par van der Meyden est dans la lignée de celle de Segerberg plutôt que de celle de Meyer. En effet, van der Meyden propose d'interpréter une action comme faisant référence à une suite d'états statiques. Par exemple, si on présume que w est un scénario dynamique, alors w_1, w_2 peuvent être utilisés pour décrire certains états statique du scénario w. Au même titre que Segerberg considère un univers rempli de *points*, van der Meyden considère un ensemble d'*états*, et une action fait référence à un ensemble de suites d'états. Autrement dit, une action fait référence à un ensemble qui contient les suites d'états $e_1, ..., e_n$, où e_1 est l'état actuel, e_n est l'état final (le résultat de l'action) et « ... » est ce qui mène à l'état final. En ce sens, considérant que plusieurs suites d'événements peuvent mener au même résultat, une action est considérée comme l'ensemble des suites d'événements possibles pouvant mener à ce résultat. Soit $U \neq \varnothing$ un univers d'états et $U^+ \subset \wp(U)$ l'ensemble de toutes les suites d'états de longueur finie $n \geq 1$.

Avant d'aller plus loin, voyons d'abord un exemple informel pour bien mettre en évidence la conception de l'action que se fait van der Meyden. Son approche repose sur une logique *dynamique*, et cet aspect est crucial pour quiconque veut bien saisir l'effet d'une action sur le monde.[17] Mettons-la en opposition avec Meyer et Royakkers. Chez ces auteurs, l'exécution de

[16] Réitérons que dans les explications précédentes, l'opérateur dynamique *après* doit être compris au sens de *certaines exécutions de ... mènent à ....*

[17] Van der Meyden n'aborde toutefois pas la question de la causalité.

certaines actions restreint l'ensemble des scénarios accessibles au scénario actuel : il y a un ensemble d'actions, un ensemble de scénarios et une fonction qui permet de déterminer quels scénarios sont accessibles à la suite de l'exécution de certaines actions.

Bien que l'idée soit similaire chez van der Meyden (et Segerberg, l'approche de van der Meyden repose sur une logique dynamique pour programmation telle qu'elle est exposée dans Segerberg 1982a), il y a une différence relativement à la conception qu'il se fait de l'action. En un certain sens, une action chez van der Meyden fait référence à une (un ensemble de) transition(s) dans le monde. Autrement dit, une action correspond à un (ensemble d') état(s) dynamique(s).

Soit w le monde actuel. Chez van der Meyden, l'univers U du modèle sémantique n'est pas un ensemble de *scénarios*, mais plutôt un ensemble de *points*, voire de *tranches de scénarios*. En effet, chaque point est considéré comme la description d'un scénario à un moment précis. D'un point de vue formel, la différence est mineure, mais sur le plan de l'interprétation, elle est considérable. Tentons une analogie : le monde actuel w est une pomme. Les points w_1, w_2, w_3 peuvent être vus comme des *tranches* de cette pomme (i.e. de w), où chaque tranche correspond à la description de w à un temps fixe. L'état dynamique w, soumis au changement, peut être vu comme un ensemble d'états statiques qui décrivent l'état dynamique à chaque moment.

Dans un telle perspective, l'action est considérée comme une suite d'états statiques (plus précisément un ensemble de suites d'états statiques). L'action « Paul vole une bicyclette », par exemple, peut être vue comme la suite s_1, s_2, s_3 où à s_1 Paul met la main sur la bicyclette, à s_2 il l'enfourche et à s_3 il part avec elle. Évidemment, cette conception pose problème si on postule la densité de l'action et du temps, où s_1, s_2 et s_3 peuvent eux-mêmes être considérés comme des ensembles d'états (i.e. mettre la main sur le vélo peut être considéré au $\frac{1}{10^{10}}$ ième de seconde et ainsi de suite). Néanmoins, l'intuition est assez claire : le vol du vélo peut être vu comme la suite s_1, s_2, s_3. Plus encore, considérant qu'il y a plusieurs manières de voler la bicyclette, l'action « Paul vole une bicyclette » peut être considérée comme l'ensemble de toutes les suites d'états possibles qui peuvent être interprétées comme le vol d'une bicyclette par Paul. Dans une certaine mesure, l'action fait référence à toutes les manières de la réaliser.

Le modèle $\mathcal{M} = \langle U, P, \tau, a \rangle$ est constitué d'un univers $U \neq \varnothing$ d'états, d'un ensemble $P \subseteq U \times U$ de paires d'états, d'une fonction $\tau : Act \longrightarrow \wp(U^+)$ qui associe à chaque action atomique un ensemble de suites d'états et $a : Prop \times U \longrightarrow \{\bot, \top\}$ une fonction qui attribue

des valeurs de vérité aux propositions relativement à certains états. Informellement, l'ensemble P est un ensemble de transitions permises. Pour bien comprendre la sémantique de van der Meyden, il faut voir que ce n'est pas un *état* permis ou non, mais bien la *transition d'un état à un autre*. En ce sens, si $< w_1, w_2 > \in P$, alors la transition de w_1 à w_2 est permise.

Afin de rendre compte sémantiquement des actions, van der Meyden doit définir de nouvelles opérations sur les ensembles pour modéliser les connecteurs ; et *. Soit $\sigma_j, \sigma_k \in U^+$ des suites de longueur respective n et m. Autrement dit, $\sigma_j = (s_1, ..., s_n)$ et $\sigma_k = (t_1, ..., t_m)$. Laissons $\sigma[i]$ et $\sigma[f]$ dénoter respectivement l'élément *initial* et l'élément *final* de la suite σ. La composition de suite $\sigma_k; \sigma_j$ n'est définie que lorsque $\sigma_j[i] = \sigma_k[f]$. Ayant cela en main, van der Meyden définit les opérations suivantes pour $A, B \in \wp(U^+)$ des ensembles de suites :

$$A; B = \{(\sigma_k; \sigma_j) : \sigma_k \in A, \sigma_j \in B \text{ et } \sigma_j[i] = \sigma_k[f]\}$$
$$A^* = \{w : w \in U\} \cup A \cup (A; A) \cup (A; A; A) \cup \ldots$$

Alors que la première définition indique que $A; B$ est l'ensemble de toutes les combinaisons possibles (définies) des suites de A et B, la seconde dépend du nombre d'itération(s). En enlevant les '...' de la définition précédente, cela donnerait par exemple A itéré 3 fois. Selon le nombre d'itérations, l'ensemble A^* contiendra tous les scénarios possibles (en effet, $\{w : w \in U\}$ est égal à U !) ainsi que toutes les suites de A, toutes les combinaisons des suites de A, toutes les combinaisons de séquences de suites de A avec les suites de A, etc. De fait, pour les connecteurs d'actions, nous avons :

$$\tau(\alpha; \beta) = \tau(\alpha); \tau(\beta)$$
$$\tau(\alpha \cup \beta) = \tau(\alpha) \cup \tau(\beta)$$
$$\tau(\alpha^*) = \tau(\alpha)^*$$

En ce sens, la séquence d'actions $\alpha; \beta$ dénote l'ensemble qui contient toutes les combinaisons possibles de α et de β (n'oublions pas qu'une action est considérée comme faisant référence à un ensemble qui contient des suites possibles). Le choix entre α et β est équivalent à l'union qui contient les suites auxquelles α et β font référence alors que l'action itérée fait référence à l'ensemble itéré.

Maintenant, soit $\tau_w(\alpha)$ l'ensemble qui contient toutes les séquences σ_j de $\tau(\alpha)$ pour lesquelles $\sigma_j[i] = w$, c'est-à-dire pour lesquelles l'état (initial) w est le point de départ. La vérité est définie récursivement.[18] Soit $\|p\|$ l'ensemble qui contient les états pour lesquels p est vrai :

$$\models_w p \Leftrightarrow w \in \|p\|$$
$$\models_w \neg\varphi \Leftrightarrow \not\models_w \varphi$$
$$\models_w \varphi \vee \psi \Leftrightarrow \models_w \varphi \text{ ou } \models_w \psi$$
$$\models_w \langle\alpha\rangle\varphi \Leftrightarrow \models_{\sigma_j[f]} \varphi \text{ pour un } \sigma_j \in \tau(\alpha)$$
$$\models_w \diamond(\alpha, \varphi) \Leftrightarrow \models_{\sigma_j[f]} \varphi \text{ pour un } \sigma_j \in \tau(\alpha) \text{ vert}$$
$$\models_w \pi(\alpha, \varphi) \Leftrightarrow \text{pour tout } \sigma_j \in \tau(\alpha), \text{ si } \models_{\sigma_j[f]} \varphi, \text{ alors } \sigma_j \text{ est vert}$$

Les quatre premières clauses sont usuelles en logique dynamique, les trois premières étant les clauses propositionnelles classiques et la quatrième celle pour l'opérateur *après*, laquelle stipule qu'il est vrai que φ est la conséquence de α pour un état w à condition que φ soit vrai pour l'état final d'une suite d'états membre de l'ensemble qui contient les suites auxquelles l'action α fait référence. Autrement dit, considérant que α fait référence à un ensemble de suites, il est vrai que φ est le cas après l'action α à condition que φ soit vrai pour au moins un état qui est l'état final d'une suite à laquelle l'action α fait référence.

Les cinquième et sixième clauses se comprennent à l'aide de la notion de suite *verte*. (Cela n'a rien d'écologique, car il s'agit d'une analogie avec le code de la route !) Soit $\sigma_j = (s_1, ..., s_n)$ une suite où chaque s_i, s_{i+1} représente un état transitoire (dynamique entre deux états statiques). On dit que σ_j est une suite *verte* lorsque pour tout état transitoire s_i, s_{i+1} de σ_j, $< s_i, s_{i+1} > \in P$. La cinquième clause signifie qu'il est vrai qu'il n'est pas interdit que l'action α entraîne φ pour l'état w à condition que φ soit vrai pour l'état final d'au moins une suite verte membre de l'ensemble qui contient les suites auxquelles l'action α fait référence. En d'autres termes, soit σ_j, une suite verte d'états à laquelle l'action α fait référence. Il est vrai qu'il n'est pas interdit d'exécuter α et d'obtenir φ à condition que φ soit vrai pour l'état final de cette suite verte (pour au moins une suite verte).

Quant à la sixième clause, elle signifie qu'il est vrai qu'il est permis que l'exécution de α mène à φ à condition que toutes les suites auxquelles l'action α fait référence et pour lesquelles φ est vrai pour le dernier état soient vertes. Autrement dit, l'action α fait seulement référence à des suites vertes et φ est vrai pour l'état final de chaque suite.

[18] L'auteur définit la sémantique avec une implication de la droite vers la gauche. Nous avons cependant défini la sémantique avec une bi-conditionnelle conformément à ce qu'on trouve usuellement en logique dynamique.

L'intérêt de l'approche de van der Meyden vient du fait qu'il insiste sur la permission plutôt que sur l'obligation. Soulignons que van der Meyden propose cette approche sous prétexte que celle de Meyer ne rend pas compte adéquatement de la notion de permission *légale*. Par ailleurs, même si son approche permet de rendre compte de différents types de permissions, celle-ci n'est toutefois pas en mesure de rendre compte de la notion d'*obligation*, laquelle est fondamentale au discours légal. Plus encore, van der Meyden ne traite pas explicitement de la notion d'*interdiction*, ce qui se comprend en fonction du fait qu'il ne présente pas de sémantique pour l'action *niée*. Cela est d'ailleurs plutôt problématique : même si la formalisation de l'action négative pose problème, il n'en demeure pas moins qu'elle est nécessaire dans la mesure où on veut rendre compte de l'abstention.[19]

Notamment, Broersen (2004) a étudié l'action négative, ce que nous aborderons dans la prochaine section. Outre les suites *vertes* introduites par van der Meyden, ses travaux ont été repris par Csirmaz (1990), qui a introduit d'autres *couleurs* (et d'autres ensembles de paires d'états) pour rendre compte de différents degrés de permission. Van der Meyden (1991) a aussi travaillé sur des systèmes où les actions réalisées changent l'ensemble des suites vertes.

Jan Broersen

Dans un article où il cherche à fournir une formalisation adéquate du concept d'*action négative*, Broersen (2004) reprend une partie des travaux faits pour sa thèse de Ph.D. (Broersen 2003). Puisque nous avons jusqu'à présent donné un bon aperçu de ce que sont les logiques déontiques dynamiques, nous n'exposerons pas l'approche de Broersen en détail, mais présenterons seulement son analyse de l'action négative dans le cadre de la logique dynamique.[20]

D'emblée, l'objectif de Broersen est de proposer un formalisme adéquat à la représentation de l'action négative $-\alpha$, interprétée comme une abstention (délibérée), par opposition à une simple omission. Autrement dit, $-\alpha$ (lorsqu'elle est considérée avec un opérateur dynamique) signifie *s'abstenir de faire* α (*to refrain from doing* α). Le formalisme doit être en mesure de rendre compte de trois caractéristiques :

1. l'interprétation doit être plausible lorsque la négation est combinée avec les autres connecteurs pour actions ;
2. n'introduit pas de restriction *ad hoc* sur la séquence et l'itération ;
3. l'interprétation doit être plausible dans un contexte normatif.

[19] Sur ce sujet, voir Peterson (2015c).

[20] Broersen (2004) utilise une logique dynamique plus complexe que celle proposée par Meyer, où l'itération, l'action inverse et l'action *test* sont considérées.

Broersen refuse l'interdéfiniabilité des opérateurs O, F et P, et ne fait que poser quelques conditions de cohérence normative pour axiomes.[21] Plutôt, il propose d'introduire trois constantes de violations V_P, V_O et V_F pour rendre compte individuellement de chaque opérateur déontique :

$$P\alpha =_{df} [\alpha]\neg V_P \qquad \text{(df. } P\text{)}$$
$$O\alpha =_{df} P\alpha \wedge [\wr^I \alpha]\neg V_O \qquad \text{(df. } O\text{)}$$
$$F\alpha =_{df} \langle\alpha\rangle V_F \qquad \text{(df. } F\text{)}$$

La définition de la permission signifie que α est permis pour w lorsque tous les scénarios accessibles à w à la suite de l'exécution de α ne contiennent pas de violation V_P. L'action α est interdite pour w lorsque son exécution mène à une violation V_F dans certains scénarios accessibles à w à la suite de l'exécution de α. La définition de O se comprend à la lumière de l'interprétation de Broersen de l'action négative. Comme il le mentionne, le connecteur pour l'action négative s'interprète différemment en fonction des autres connecteurs avec lesquels il peut être jumelé. En ce sens, la logique du connecteur $\wr^I$ dépend de l'ensemble de connecteurs pour actions considérés.[22]

Brièvement, Broersen identifie les ensembles de connecteurs suivants, où l'interprétation sémantique de la négation varie d'un ensemble à l'autre :

$$(\cup, \wr^K)$$
$$(^{\leftarrow}, \cup, \wr^B)$$
$$(;, \cup, \wr^4)$$
$$(^*, ;, \cup, \wr^{S4})$$
$$(^{\leftarrow}, ;, \cup, \wr^{B4})$$
$$(^{\leftarrow}, ^*, ;, \cup, \wr^{S5})$$

Autrement dit, lorsqu'elle est considérée avec l'action disjointe, la relation sémantique sur le modèle pour l'action négative est celle de K, elle est symétrique lorsqu'on ajoute le connecteur pour l'action inverse, elle est transitive lorsqu'on considère l'action disjointe et la séquence d'actions, elle est transitive et réflexive lorsqu'on ajoute à cela l'itération d'actions, elle

[21] Voir le texte de 2004 pour les conditions. Notons que la condition (8) $F(\alpha \cap \beta) \supset F(\alpha) \wedge F(\beta)$ est trivialement fausse si on exclut les considérations temporelles : ce n'est pas parce qu'il est interdit de « boire de l'alcool et conduire » qu'il est nécessairement interdit de boire et interdit de conduire.

[22] Voir Broersen (2004) pour une brève revue des différentes formalisations de l'action négative.

est symétrique et transitive lorsqu'on considère l'action inverse, la séquence et l'action disjointe. Finalement, elle est universelle lorsqu'on ajoute l'itération. Prenant cela en considération, la définition de l'obligation signifie que α est obligatoire pour w lorsque α est permis et que dans tout scénario accessible à la suite de l'exécution de $\wr^I\alpha$ (où $\wr^I$ est modélisé en fonction de l'ensemble des connecteurs pour actions de la logique dynamique utilisée) il n'y a pas de violation V_O.

Le lecteur intéressé par cette approche est invité à consulter Broersen qui, dans ses écrits de 2004, prend certaines des suppositions avancées en 2003 pour acquis.

* * *

Ceux qui désirent poursuivre la réflexion sur la logique déontique dynamique peuvent aussi consulter Anglberger (2008), qui montre comment elle peut donner lieu à des paradoxes indésirables, même si elle peut éviter la majorité des paradoxes de la logique déontique standard. Par ailleurs, Anglberger (2009) fait un bref survol des différentes réductions proposées pour la logique dynamique (réduction d'un opérateur déontique à une proposition qui ne contient pas d'opérateur déontique) et d'une réduction alternative, qui introduit plusieurs constantes de violations, sur lesquelles on impose un ordre partiel pour rendre compte des scénarios alternatifs préférables à d'autres. Étant donné que la logique dynamique provient de l'informatique et que plusieurs informaticiens s'intéressent à la logique déontique, le lecteur ne sera pas surpris de trouver beaucoup d'applications de la logique déontique dymamique à la programmation. Herzig *et al.* (2011), par exemple, augmente le langage de la logique dynamique afin de pouvoir programmer des NMAS.

Chapitre 11

La logique Input/Output

Le présent chapitre aborde les *Input/Output Logics*, aussi appelées logiques I/O. Il regroupe trois approches, dont la présentation des logiques I/O de Makinson et van der Torre (2000). Même si la seconde ne porte pas sur la logique I/O, nous présentons néanmoins le texte de Jones et Sergot (1996) puisque leur approche a été incorporée à la logique I/O par Boella et van der Torre (2006c), dont nous aborderons les théories en troisième lieu.

David Makinson et Leendert van der Torre

Les logiques I/O offrent un cadre de travail non modal et sont apparues chez Makinson et van der Torre (2000). Plutôt que de tenter de développer une logique déontique qui peut rendre compte de l'inférence normative dans sa totalité, les chercheurs introduisent les logiques I/O en guise d'*assistants* qui peuvent être utilisés afin d'aider un processus de transformation. On doit comprendre les logiques I/O en termes d'*entrée* et de *sortie* : on cherche à déterminer certaines règles qui permettent de gouverner le processus de transformation qui part d'un *input* et donne un *output*.

D'entrée de jeu, on considère un langage propositionnel booléen (qui obéit aux règles de LPC) à l'aide duquel on définit des paires ordonnées d'énoncés $(a, x) \in EBF \times EBF$, lesquelles sont interprétées comme des règles qui déterminent l'*output* x relativement à l'input a (Makinson et van der Torre 2001). Même si le parallèle peut être tracé entre une paire (a, x) et un conditionnel $a \supset x$, la paire ne doit pas être entendue au sens de l'implication matérielle. Cela se comprend notamment par le fait qu'une norme, exprimée par une paire, est sans valeur de vérité (Makinson et van der Torre 2003b). En ce sens, il s'agit d'une solution au problème posé par Jørgensen.

Un *ensemble générateur* $G \subseteq EBF \times EBF$ est un ensemble de paires (a, x) et l'ensemble $G(A)$ contient l'*output* x pour certains a éléments d'un ensemble de propositions A relativement à l'ensemble générateur G. Autrement dit, G est un ensemble de règles (de normes) et $G(A)$ est un ensemble de propositions contenant les *outputs* x pour lesquels $a \in A$ et $(a, x) \in G$.

$$G(A) = \{x : (a, x) \in G \text{ pour certains } a \in A\}$$

Makinson et van der Torre introduisent l'opération $out(G, A)$, qui, à partir d'un ensemble générateur et d'un ensemble d'*inputs*, donne un ensemble d'*outputs*. Cette opération est utilisée pour définir quatre opérateurs, qui, à leur tour, donneront quatre logiques I/O. Soit $Cn(A)$ l'ensemble des conséquences classiques de A. Soulignons que $\top \in out(G, A)$ pour toute tautologie classique. Les opérateurs (qui représentent des opérations sur des ensembles) sont les suivants[1] :

$$out_1(G, A) =_{df} Cn(G(Cn(A)))$$

$$out_2(G, A) =_{df} \bigcap_V Cn(G(V)) \text{ t.q. } A \subseteq V \text{ et } V \text{ est complet}$$

$$out_3(G, A) =_{df} \bigcap_B Cn(G(B)) \text{ t.q. } B = Cn(B), A \subseteq B \text{ et } G(B) \subseteq Cn(B)$$

$$out_4(G, A) =_{df} \bigcap_V Cn(G(V)) \text{ t.q. } A \subseteq V, G(V) \subseteq V \text{ et } V \text{ est complet}$$

La première définition explique que l'ensemble qui contient les *outputs* pour les paires de G formées à partir des *inputs* de A est celui qui contient les conséquences logiques de l'ensemble qui contient les paires ayant pour *inputs* les conséquences logiques de A. Reprenons cette définition à l'envers : $Cn(A)$ représente l'ensemble des conséquences de l'ensemble de propositions A.

$$Cn(A) = \{x : a \in A \text{ et } a \vdash_{LPC} x\}$$

L'ensemble $G(Cn(A))$ correspond à l'ensemble qui contient les *outputs* pour lesquels l'ensemble qui contient les conséquences de A sont des *inputs*. Si, par exemple, $G = \{(a, x), (a \vee b, y)\}$, alors $G(a) = \{x\}$ et $G(Cn(a)) = \{x, y\}$.[2] L'*output* de type 1 est, par définition, l'ensemble

[1] À partir de ces définitions, les auteurs introduisent aussi $out_i^+(G, A) = out_i(G^+, A)$ avec $G^+ = G \cup I$ et $I = \{(a, a) : a \in EBF\}$. Nous laissons ces définitions de côté par souci d'économie. Les auteurs notent que $out_2^+ = out_4^+$ et que cela redonne la logique classique.

[2] Afin de faciliter la lecture, les auteurs écrivent $G(a)$ ou $Cn(a)$ plutôt que $G(\{a\})$ ou $Cn(\{a\})$ lorsqu'il s'agit d'un singleton.

qui contient les conséquences logiques de $G(Cn(A))$. Informellement, on entre l'ensemble A comme *input*, on considère ses conséquences, ensuite on considère l'*output* des conséquences de A (relativement à G) et, enfin, les conséquences de l'*output* des conséquences de A.

La seconde définition requiert l'introduction de la notion d'ensemble *complet*. Un ensemble est *complet* lorsqu'il est soit *maximalement consistant* (i.e., Γ tel que pour tout $\varphi \in EBF$, soit $\varphi \in \Gamma$ ou $\neg\varphi \in \Gamma$ et $\bot \notin \Gamma$) soit qu'il est équivalent à l'ensemble de toutes les propositions du langage. Conformément au lemme de Lindenbaum, à savoir que tout ensemble consistant possède une extension maximalement consistante, il est possible de construire un nombre dénombrable d'extensions maximalement consistantes pour un ensemble (consistant) d'*inputs* A.

Un ensemble complet est un ensemble auquel on ne peut rien ajouter de plus. Dans le cas d'un ensemble maximalement consistant, on ne peut rien ajouter de plus sans briser sa consistance. Lorsqu'il s'agit de l'ensemble de tous les énoncés du langage, alors il est clair que rien ne peut y être ajouté. Dans l'éventualité où l'ensemble correspond à tous les énoncés du langage, celui-ci est inconsistant.

La définition de out_2 considère l'intersection de tous les ensembles des conséquences des *outputs* qui ont V comme *input*, où V est une extension complète de A. Par opposition avec out_1 qui ne fait que considérer l'ensemble d'*inputs*, out_2 permet de prendre en considération tout ce qui est commun avec toute extension complète de l'ensemble d'*inputs*. En plus de considérer les *outputs* de A relativement à G (ainsi que leurs conséquences), on considère aussi les *outputs* des extensions maximalement consistantes de A. Par exemple, si G contient des paires pour lesquelles l'*input* n'est pas élément de A, alors out_2 permet d'isoler les *outputs* consistants avec ceux de A.

La troisième définition (qui ressemble à la deuxième) fait appel aux conditions $B = Cn(B)$ et $A \subseteq B$, ce qui peut sembler étrange à première vue. En fait, il suffit de remarquer que $EBF = Cn(EBF)$ et que $V = Cn(V)$ pour un ensemble V maximalement consistant. De surcroît, il est trivial que $Cn(Cn(A)) = Cn(A)$ puisque la conséquence est transitive (i.e., les conséquences des conséquences de A sont des conséquences de A). En ce sens, il est possible de considérer toute extension B de A sans pour autant que B soit complète. Il suffit d'avoir $B = Cn(X)$ avec $A \subseteq X$, ce qui donne $A \subseteq B$.

La troisième condition de la définition, soit que $G(B) \subseteq Cn(B)$, signifie que l'ensemble des *outputs* qui ont B comme *input* (relativement à G) est sous-ensemble de l'ensemble des conséquences de B. De fait, les *outputs* peuvent être réutilisés comme *inputs*. Sachant que $G(B) \subseteq Cn(B)$, cela implique que $Cn(G(B)) \subseteq Cn(B)$: B contient toutes ses conséquences

et contient $G(B)$, et donc contient toutes les conséquences de $G(B)$. Mais puisque $Cn(G(B)) \subseteq Cn(B)$, il s'ensuit que l'ensemble des conséquences de l'ensemble qui contient les *outputs* qui ont B comme *input* fait partie de l'ensemble des conséquences de B. Étant donné que B est considéré comme *input* dans $G(B)$, il s'ensuit que les *outputs* de $G(B)$ sont aussi d'emblée considérés en tant qu'*inputs*. La définition de out_3 donne l'intersection des ensembles de conséquences d'*outputs* qui ont pour *inputs* B (relativement à G) lorsque B équivaut à l'ensemble de ses conséquences, A est sous-ensemble de B et l'ensemble des *outputs* qui ont pour *input* B fait partie de B. Cette dernière condition amène la possibilité de réutiliser les *outputs* comme *inputs*.

Une façon peut-être un peu plus simple de voir la chose serait de définir $out_3(G, A) = Cn(G(A*))$ où $A*$ est le plus petit ensemble fermé sous Cn et G tel que $A \subseteq A*$ (Makinson et van der Torre 2000). Autrement dit, $A*$ est le plus petit ensemble pour lequel A est un sous-ensemble et est fermé sous la relation de conséquence et G.

En dernier lieu, la définition de out_4 est à l'image de celle de out_2, à l'exception qu'on permet de reprendre les *outputs* comme *inputs* à l'instar de out_3. Il s'agit de prendre l'intersection des conséquences de toute extension complète V de A où l'ensemble des *outputs* de V est sous-ensemble de V.

Évidemment, les *outputs* varient en fonction de l'ensemble générateur G. Les auteurs nomment ces définitions respectivement *simple minded output*, *basic output*, *reusable simple minded output* et *reusable basic output*. Les points cruciaux à souligner sont que les *inputs* ne se retrouvent pas dans les *outputs* (sauf pour les out_i^+) et que la contraposition ne s'applique pas dans chaque cas.

Le lecteur qui consulte les articles sur les logiques I/O ne verra peut-être pas tout de suite ce qui différencie les quatre types d'*outputs* du point de vue de la sémantique, sans compter que les auteurs donnent certains exemples qui engendrent la confusion. Notamment, Makinson et van der Torre donnent cet exemple.

Exemple

Soit $G = \{(\top, \neg a), (a, x)\}$. *Dans ce cas,* $out_i(G, a) = Cn(\{\neg a, x\})$ *pour* $i \in \{1, 2, 3, 4\}$.

Considérons chacun des outputs *pour* $A = \{a\}$ *:*

1. $out_1(G, a) = Cn(G(Cn(a)))$ *et puisque* $\top, a \in Cn(A)$ *il s'ensuit que* $G(Cn(a)) = \{\neg a, x\}$ *et donc que* $out_1(G, a) = Cn(\{\neg a, x\})$.
2. *Si on prend en compte tout* $G(V)$ *pour toute extension complète* V *de* a*, nous obtiendrons encore* $G(V) = \{\neg a, x\}$ *pour chaque* V*, et de fait* $out_2(G, a) = \bigcap Cn(G(V)) = Cn(\{\neg a, x\})$.

3. $G(B)$ *où* $B = Cn(B)$ *et* $\{a\} \subseteq B$ *donne aussi seulement* $G(B) = \{\neg a, x\}$*, donc* $out_3(G, a) = \bigcap Cn(G(B)) = Cn(\{\neg a, x\})$.

4. Le même raisonnement s'applique que pour 2 et 3.

Mais alors, qu'est-ce qui différencie chaque *output* ? Il faut insister sur le fait que l'intersection dans les définitions susmentionnées ne porte pas sur V, mais bien sur $Cn(G(V))$. Voyons deux exemples pour mettre en lumière les différences entre le *simple minded* et le *basic output*, d'un côté, et le *simple minded* et le *reusable simple minded output* de l'autre.

Exemple

Considérons l'ensemble $G = \{(a, x), (b, x)\}$*. Quels sont* $out_1(G, a \vee b)$ *et* $out_2(G, a \vee b)$ *? Alors que* $out_1(G, a \vee b) = Cn(G(Cn(a \vee b))) = Cn(\varnothing)$*,* $out_2(G, a \vee b) = \bigcap Cn(G(V)) = Cn(x)$*. En effet, toute extension complète de* $\{a \vee b\}$ *contient soit* a*, soit* b *soit* a *et* b*. De fait, dans chaque cas,* $G(V) = \{x\}$*. En ce sens, la sémantique de* out_2 *permet d'aller chercher ce qui est commun à toutes les extensions complètes de* $a \vee b$ *du point de vue de leur* output *relativement à* G*. La distinction entre* out_1 *et* out_2 *se transpose à* out_3 *et* out_4*.*

Quant à la différence entre out_1 et out_4 (ou encore out_2 et out_3), considérons l'ensemble $G = \{(a, x), (a \wedge x, y)\}$.

Exemple

Soit $G = \{(a, x), (a \wedge x, y)\}$*. Quels sont* $out_1(G, a)$ *et* $out_4(G, a)$ *? D'un côté,* $out_1(G, a) = Cn(G(Cn(a))) = Cn(x)$*. De l'autre côté,* $out_4(G, a) = \bigcap Cn(G(V))$*. Puisque* $a \in V$*, il s'ensuit que* $x \in G(V)$*. Or considérant que* $G(V) \subseteq V$*, il en résulte que* $x \in V$*, et donc que* $y \in G(V)$*. De fait,* $out_4(G, a) = Cn(\{x, y\})$.[3]

Les définitions sémantiques des différents types d'*outputs* s'accompagnent du côté de la syntaxe par les règles suivantes :

(SI) $\dfrac{(a, x)}{(b, x)}$ (lorsque $b \vdash_{LPC} a$) (AND) $\dfrac{(a, x)\ (a, y)}{(a, x \wedge y)}$ (OR) $\dfrac{(a, x)\ (b, x)}{(a \vee b, x)}$

(WO) $\dfrac{(a, x)}{(a, y)}$ (lorsque $x \vdash_{LPC} y$) (CT) $\dfrac{(a, x)\ (a \wedge x, y)}{(a, y)}$

[3] Le lecteur est invité à consulter les figures dans Makinson et van der Torre (2000) pour mieux visualiser les différents types d'*outputs* ainsi que pour un résumé des propriétés des logiques I/O.

(SI) signifie que si x est l'*output* de a et que a est une conséquence logique de b, alors x est l'*output* de b. (AND) signifie que si x et y sont les *outputs* de a, alors $x \wedge y$ est l'*output* de a. (WO) stipule que si y est la conséquence de x et que x est l'*output* de a, alors y est l'*output* de a. (OR) indique que si x est l'*output* de a et est l'*output* de b, alors c'est aussi l'*output* de $a \vee b$. (CT) signifie que si x est l'*output* de a et que y est l'*output* de $a \wedge x$, alors y est l'*output* de a.

Conjointement à chaque définition d'*output* se joignent des règles de dérivation spécifiques. Soit $deriv_i(G)$ le plus petit ensemble qui contient G, les paires $(\top, \top)$ où $\top$ est une tautologie, et fermé sous les règles faisant partie de R_i. Les notions de dérivabilité $deriv_1(G) - deriv_4(G)$ correspondent respectivement aux *outputs* $out_1 - out_4$[4] :

$$R_1 = \{(SI), (AND), (WO)\}$$
$$R_2 = \{(SI), (AND), (WO), (OR)\}$$
$$R_3 = \{(SI), (AND), (WO), (CT)\}$$
$$R_4 = \{(SI), (AND), (WO), (OR), (CT)\}$$

Les logiques I/O ont été développées par Makinson et van der Torre afin d'êtres appliquées à la logique déontique. Jusqu'à présent, nous avons vu que le discours normatif met en jeu deux formes d'inconsistances, à savoir la propositionnelle et la normative. Le paradoxe de Chisholm met en évidence le fait que tout système qui veut rendre compte des obligations contraires doit être capable d'exprimer (dans une certaine mesure) les inconsistances normatives. Or les logiques I/O permettent d'emblée de faire la distinction entre les deux types d'inconsistance. Makinson et van der Torre font la distinction entre l'inconsistance de l'*output*, qui, en définitive, est une inconsistance propositionnelle, et l'inconsistance entre les *inputs* et les *outputs*, une forme d'inconsistance normative. Alors qu'un *output* de type i est dit inconsistant lorsque $\bot \in out_i(G, A)$, l'inconsistance entre les *inputs* et les *outputs* advient lorsque $\bot \in Cn(out_i(G, A) \cup A)$. Voici l'exemple des auteurs pour mettre en évidence la différence entre les deux types d'inconsistance.

Exemple

Soit $G = \{(\top, \neg a), (a, x)\}$. *Informellement,* G *affirme l'existence d'une obligation conditionnelle : en général l'*output $\neg a$ *est préféré, mais s'il advient que* a *est donné, alors l'*output x *sera désiré. Comme susmentionné, en supposant* $A = \{a\}$, $out_i(G, a) = Cn(\{\neg a, x\})$ *pour* $i \in \{1, 2, 3, 4\}$. *L'*output *est consistant puisque* $\bot \notin Cn(\{\neg a, x\})$. *Cependant, l'*output *est inconsistant avec l'*input, *car* $\bot \in Cn(\{\neg a, x\} \cup \{a\})$.

[4] Afin d'obtenir les *outputs* out_i^+, il suffit d'ajouter la règle (ID) qui permet d'ajouter (a, a) à tout moment.

Cela dit, les logiques I/O permettent d'arriver à un résultat étrange : considérant que $a \vdash_{LC} \top$, la règle (SI) donne dans l'exemple précédent que $(\top, \neg a)$ entraîne $(a, \neg a)$.[5] Ce résultat semble peu probable, surtout si on considère que les obligations contraires surviennent à un certain moment dans le temps. Par exemple, si a représente « Paul vole de l'argent » et x signifie « Paul remet l'argent volé », alors on obtient que dans toutes circonstances idéales, Paul ne vole pas, que dans une circonstance sous-idéale où Paul vole, il remet l'argent volé, mais que dans une circonstance où Paul vole, il ne vole pas. Il ne s'agit pas d'une contradiction à proprement parler puisqu'il s'agit d'une paire *input/output*, mais cela pose néanmoins problème dans un contexte déontique. Plus précisément, une fois que l'action qui ne devait pas être posée a été accomplie, on veut mettre de côté les scénarios où l'action n'était pas faite et se concentrer uniquement sur ceux où elle est accomplie.

La stratégie pour parvenir à ce résultat est d'utiliser une technique propre aux logiques épistémiques. Pour ce faire, les auteurs introduisent des *contraintes* sur les *outputs*. Soit C un ensemble de contraintes qui contient des propositions. L'ensemble $maxfamily(G, A, C)$ est défini comme l'ensemble de tous les plus grands $H \subseteq G$ pour lesquels $out(H, A)$ est consistant avec C (i.e., $\bot \notin Cn(out(H, A) \cup C)$). L'ensemble $outfamily(G, A, C)$ est l'ensemble des *outputs* générés par $maxfamily(G, A, C)$, à savoir que $outfamily(G, A, C) = \{out(H, A) : H \in maxfamily(G, A, C)\}$.

Conformément à l'exemple susmentionné, pour restreindre l'ensemble des normes G en excluant celle qui stipule que Paul ne doit pas voler, on ajoute la contrainte $C = \{a\}$. Alors que dans le cas sans contrainte, on obtient $maxfamily(G, a, \varnothing) = \{G\}$ et $outfamily(G, a, \varnothing) = \{Cn(\{\neg a, x\})\}$, l'ajout de la contrainte $C = \{a\}$ permet d'isoler l'ensemble $maxfamily(G, a, a) = \{\{(a, x)\}\}$ puisque $\bot \in Cn(out(G, a) \cup \{a\})$ où $Cn(out(G, a) \cup \{a\}) = Cn(Cn(\{\neg a, x\}) \cup \{a\})$ et $outfamily(G, a, a) = \{Cn(\{x\})\}$.[6]

À l'aide de l'imposition d'une contrainte sur les différents types d'*outputs*, Makinson et van der Torre introduisent la notion de *contrainte* I/O (*input/output constraint*). Pour un *output* $out(G, A)$, l'ajout de la contrainte I/O correspond à la restriction de l'ensemble des *outputs* à l'ensemble d'*inputs* $maxfamily(G, A, C)$ pour $C = A$. Ainsi, on s'assure que l'ensemble d'*outputs* est consistant et qu'il est consistant avec l'ensemble des *inputs*.

[5] Un exemple plus complexe est donné dans Makinson et van der Torre (2003b).
[6] Pour d'autres exemples, voir le texte de 2001.

Les logiques I/O contraintes s'obtiennent d'un point de vue syntaxique en ajoutant la notion de *dérivation contrainte*. Conformément aux quatre types d'*outputs* ainsi qu'à leur notion conjointe de dérivation, les *outputs contraints* s'obtiennent de par leur notion respective de dérivation contrainte, où une dérivation est *contrainte* lorsque $(a, \neg a) \notin deriv_i(L)$ pour tout $L \subseteq G$. Le lecteur est invité à consulter Makinson et van der Torre (2001) pour un résumé des résultats sur les logiques I/O contraintes.

Jusqu'à présent, les logiques I/O offrent une façon de formuler des obligations qui tiennent dans tous les contextes, p. ex. $(\top, a)$, et des obligations qui surviennent dans des contextes de violation, p. ex. $\{(\top, a), (\neg a, x)\}$. Traité de la sorte, un ensemble G est considéré comme un ensemble de normes, lesquelles dictent des obligations (des *outputs*) dans certains contextes (des *inputs*). Ainsi, la paire $(\top, a)$ signifie que dans toutes circonstances, a doit être alors que $\{(\top, a), (\neg a, x)\}$ signifie que a doit être, mais que, si jamais a n'est pas le cas, alors x doit être.

Mais qu'en est-il de la permission ? D'emblée, Makinson et van der Torre (2003a) font la distinction entre les permissions *explicites* et *implicites*, ce à quoi on fait référence aussi avec des permissions *fortes* et *faibles*. Une permission *implicite*, voire une permission *faible*, correspond à quelque chose qui n'est pas interdit par un ensemble de normes. Cela dit, le pendant de la permission implicite est la permission *explicite* voire *forte* : une action est explicitement permise lorsqu'un ensemble de normes le mentionne. Par exemple, il est explicitement permis par la législation d'arrêter son véhicule sur la voie d'accotement d'une autoroute en cas d'urgence, notamment si le véhicule tombe en panne. Une autre façon de nommer ces types de permissions serait de les appeler les permissions *négatives* et *positives*.

D'abord, les auteurs considèrent la permission implicite conditionnelle. Une action est dite implicitement permise par un ensemble de normes dans un certain contexte lorsque cet ensemble n'interdit pas cette action dans les mêmes conditions. Dans le langage des logiques I/O, $x \in negperm(G, a)$ si et seulement si $\neg x \notin out_i(G, a)$ pour $i \in \{1, 2, 3, 4\}$. Considérant que G est un ensemble de normes qui exprime des obligations conditionnelles, quelque chose est implicitement permis dans un certain contexte relativement à G lorsque la négation de ce quelque chose n'est pas obligatoire dans ce contexte. Défini ainsi, $negperm(G, a)$ n'est pas nécessairement consistant.

Les auteurs adaptent aussi leur approche pour rendre compte de la permission explicite. Toutefois, ils font la distinction entre deux types de permission explicite : la *statique* et la *dynamique*. Alors que la permission statique détermine explicitement le cadre à l'intérieur duquel un agent peut agir, la permission dynamique guide le législateur, qui doit s'assurer de ne pas légiférer à l'encontre des actions permises statiquement.

Considérons d'abord la permission statique. Soit G, un ensemble générateur d'obligations, et P, un ensemble générateur de permissions explicites. La définition de la permission statique est telle que $x \in statperm_i(P, G, a)$ si et seulement si $x \in out_i(G \cup \{(c, z)\}, a)$ pour $i \in \{1, 2, 3, 4\}$ et une paire $(c, z) \in P$. Prenons par exemple le cas de out_1, où $G*$ est la fonction qui isole les *outputs* potentiels de $G \cup \{(c, z)\}$:

$$\begin{aligned} x \in statperm_1(P, G, a) &\Leftrightarrow x \in out_1(G \cup \{(c, z)\}, a) \\ &\Leftrightarrow x \in Cn(G * (Cn(a))) \end{aligned}$$

Soit $G = \{(b, y), (a \vee b, u)\}$ et $P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}$, alors nous obtenons (en répétant l'opération pour chaque paire de P) $G * (Cn(a)) = \{u, v, z\}$, donc $statperm_1(P, G, a) = Cn(\{u, v, z\})$. Le lecteur aura remarqué que tout ce qui est (explicitement) obligatoire est par le fait-même explicitement statiquement permis. Soulignons aussi que dans le cas où $P = \varnothing$, l'ensemble de permissions statiques se réduit aux *outputs* de a relativement à G. La définition est telle que l'ensemble des permissions statiques se construit en observant une à une les paires de P et le contexte (l'*input*) dans la permission, qui est un élément de P, peut être une conséquence de a. En ce sens, l'ensemble de permissions statiques offre à un agent toutes les permissions explicites qu'il a dans un certain contexte relativement aux ensembles de normes G et P.

Toutefois, ce type de permission pourrait très bien mener à une impasse. Par exemple, posons $G = \{(b, y), (a \vee b, u), (a, \neg v)\}$ et $P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}$. En ajoutant $\neg v$ obligatoire dans le contexte a, on contrevient à une permission explicite statique, ce qui engendre $statperm_1(P, G, a) = Cn(\{u, v, \neg v, z\})$. Même si un ensemble de permissions peut être inconsistant (une action et sa négation peuvent toutes les deux être permises, voire *indifférentes*), il y a un problème dans le cas de la permission *explicite* puisque si on suppose que v est explicitement permis dans le contexte a, alors dans le même contexte v ne doit pas être interdit, sans quoi on obtient une inconsistance normative.

C'est là qu'entre en jeu la notion de permission explicite *dynamique*. Formellement, $x \in dynperm_i(P, G, a)$ si et seulement si $\neg z \in out_i(G \cup \{(a, \neg x)\}, c)$ pour $i \in \{1, 2, 3, 4\}$ et un $z \in statperm_i(P, G, c)$ et où c n'est pas contradictoire. En ce sens, l'ensemble des permissions dynamiques correspond aux actions qu'on ne peut interdire sans engendrer d'inconsistance normative. Autrement dit, il s'agit d'un ensemble de permissions qui, si elles étaient interdites, engendreraient des inconsistances normatives. Reprenons l'exemple susmentionné :

$$\begin{aligned} G &= \{(b, y), (a \vee b, u)\} \\ G' &= G \cup \{(a, \neg v)\} \end{aligned}$$

$$P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}$$
$$statperm_1(P, G, a) = Cn(\{u, v, z\})$$

Dans ce cas, il y a un $v \in statperm_1(P, G, a)$ pour lequel $\neg v \in out_1(G', a)$, et donc $v \in dynperm(P, G, a)$. Autrement dit, dans le contexte a, v est une permission dynamique, c'est-à-dire une permission qui ne peut être interdite sans engendrer d'inconsistance normative. Évidemment, si on veut absolument interdire v, il y a toujours une possibilité de redéfinir P.[7]

En laissant de côté l'aspect « vérité » des propositions, les logiques I/O offrent une solution intéressante au dilemme de Jørgensen (1937). Elles sont utilisées dans le domaine de l'informatique pour rendre compte de la structure des NMAS. Avant de voir comment cela fonctionne, il nous faut d'abord présenter l'approche de Jones et Sergot (1996).

Andrew Jones et Marek Sergot

Cette approche vise à rendre compte de la notion de *pouvoir institutionnalisé*, c'est-à-dire le pouvoir accordé à un agent par une institution (p. ex. le pouvoir d'un policier de donner une contravention). Jones et Sergot (1996) introduisent un opérateur $\Rightarrow_s$ en conjonction avec une logique de type STIT. La raison pour laquelle nous n'avons pas placé l'approche de Jones et Sergot dans le chapitre 9 est qu'elle est utilisée en conjonction avec les logiques I/O par Boella et van der Torre (2003b, 2004, 2006a,c,b) afin d'analyser la structure d'un NMAS.

Du côté de la syntaxe, Jones et Sergot utilisent un langage qui comprend les connecteurs habituels et permet l'itération de l'opérateur d'action E_x. Cependant, même si leur approche est de type STIT, celle-ci est axiomatisée de manière différente. Plutôt que de l'être comme une modalité de type K, l'opérateur E_x répond à la règle (RE) de substitution d'équivalences logiques dans la portée de l'opérateur E, au schéma d'axiome (**T**) de la logique aléthique ainsi qu'à la règle (non Nec) qui permet de représenter le fait qu'un agent ne réalise pas une tautologie :

$$E_x\varphi \supset \varphi \qquad (\mathbf{T})$$

$$\frac{\vdash \varphi \equiv \psi}{\vdash E_x\varphi \equiv E_x\psi} \quad (\mathsf{RE}) \qquad \frac{\vdash \varphi}{\vdash \neg E_x\varphi} \quad (\text{non } \mathsf{Nec})$$

[7] Le lecteur qui désire plus de détails formels concernant les propriété des ensembles de permissions est invité à consulter Makinson et van der Torre (2003a).

Les auteurs mentionnent qu'on pourrait ajouter le schéma d'axiome (C), sans pour autant l'adopter. Cependant, on doit rejeter le schéma d'axiome (M) afin de ne pas obtenir (RM). Alors que (non Nec) implique que les tautologies ne sont pas réalisées, (RM) impliquerait que si un agent réalise φ, alors il réalise aussi une tautologie puisque $\vdash_{LPC} \varphi \supset \top$. L'opérateur E_x est donc axiomatisé à l'aide d'une logique modale classique.

$$(E_x\varphi \wedge E_x\psi) \supset E_x(\varphi \wedge \psi) \qquad \text{(C)}$$
$$E_x(\varphi \wedge \psi) \supset (E_x\varphi \wedge E_x\psi) \qquad \text{(M)}$$

En plus de la logique de l'action sur laquelle repose leur approche, Jones et Sergot introduisent le connecteur conditionnel $\Rightarrow_s$. En présumant les règles de la logique propositionnelle classique, ce connecteur est régi par les règles et schémas d'axiomes suivants :

$$\frac{\vdash \psi \equiv \psi'}{\vdash (\varphi \Rightarrow_s \psi) \equiv (\varphi \Rightarrow_s \psi')} \qquad \text{(EqC)}$$

$$\frac{\vdash \varphi \equiv \varphi'}{\vdash (\varphi \Rightarrow_s \psi) \equiv (\varphi' \Rightarrow_s \psi)} \qquad \text{(EqA)}$$

$$((\varphi \Rightarrow_s \psi) \wedge (\varphi \Rightarrow_s \rho)) \supset (\varphi \Rightarrow_s (\psi \wedge \rho)) \qquad (\text{Agg}_s)$$
$$((\varphi \Rightarrow_s \psi) \wedge (\rho \Rightarrow_s \psi)) \supset ((\varphi \vee \rho) \Rightarrow_s \psi) \qquad \text{(Disj)}$$

Les règles (EqC) et (EqA) permettent de substituer les équivalences logiques dans le conséquent et l'antécédent du connecteur conditionnel. Le connecteur $\varphi \Rightarrow_s \psi$ s'interprète comme φ *compte pour* ψ *relativement à* s. En ce sens, le schéma d'axiome (Agg_s) indique que si φ compte pour ψ et φ compte pour ρ, alors φ compte pour ψ et ρ. Dans le même ordre d'idées, (Disj) indique que si φ compte pour ψ et que ρ compte aussi pour ψ, alors la disjonction φ ou ρ compte pour ψ.

L'action étant axiomatisée par un système classique, la sémantique de l'opérateur E_x se traduit en matière de *modèle minimal*. Bien que Jones et Sergot renvoient le lecteur au chapitre 7 de l'ouvrage de Chellas (1980) pour la définition du modèle minimal, nous allons l'expliciter brièvement. La particularité d'un modèle minimal est qu'il permet de rejeter la règle (RN) ainsi que les schémas (M) et (C).

Un modèle minimal $\mathcal{M} = \langle W, N, a\rangle$ est constitué d'un univers non vide W et d'une fonction qui assigne des valeurs de vérité aux propositions relativement aux scénarios. La principale différence est que N est une collection de sous-ensembles de W (Chellas 1980), par opposition avec la relation R

dans les modèles normaux, qui est un ensemble de paires ordonnées. Informellement, N_w est un ensemble de propositions *nécessaires* au moment w. L'attention est portée sur la notion de nécessité puisqu'il ne s'agit pas d'une nécessité *logique*, comme dans le cas des modèles normaux. Plutôt, cette notion doit être entendue comme une notion arbitraire de nécessité : il s'agit de ce qui est postulé comme nécessaire dans un scénario donné. De fait, il est possible d'avoir $N_w = \varnothing$, et donc $\top \notin N_w$ pour $\top$ une tautologie du système.

Le modèle sémantique $\mathcal{M} = \langle W, N, f_s, V\rangle$ contient un univers $W \neq \varnothing$, un ensemble N contenant différentes collections de sous-ensembles N_w où certaines propositions sont jugées *nécessaires*, une fonction f_s qui détermine les scénarios dans lesquels « φ compte pour ψ » est vrai relativement à une institution s, et finalement V un ensemble d'assignations $\|\varphi\|_{\mathcal{M}} = \{w : \varphi \in w\}$, où $\|\varphi\|_{\mathcal{M}}$ est l'ensemble qui contient les scénarios où φ est vrai (relativement au modèle $\mathcal{M}$).

$$\models_w p \Leftrightarrow\ w \in \|p\|_{\mathcal{M}} \tag{1}$$
$$\models_w \neg\varphi \Leftrightarrow \not\models_w \varphi \tag{2}$$
$$\models_w \varphi \supset \psi \Leftrightarrow \not\models_w \varphi \text{ ou } \models_w \psi \tag{3}$$
$$\models_w E_x\varphi \Leftrightarrow\ \|\varphi\|_{\mathcal{M}} \in N_w \tag{4}$$
$$\models_w \varphi \Rightarrow_s \psi \Leftrightarrow\ \|\psi\|_{\mathcal{M}} \in f_s(w, \|\varphi\|_{\mathcal{M}}) \tag{5}$$

Les clauses (1)–(3) sont les clauses usuelles et la (4) indique que « x réalise φ » est vrai pour w si et seulement si l'ensemble qui contient les scénarios où φ est vrai est élément de la collection d'ensembles qui contiennent les propositions nécessairement vraies pour w. La clause (5) signifie que « φ compte pour ψ » est vrai relativement à l'institution s pour le scénario w si et seulement si l'ensemble qui contient les scénarios où ψ est vrai est élément de l'ensemble qui contient les scénarios où φ est vrai. Autrement dit, f_s isole les scénarios accessibles à w où φ est vrai. En ce sens, « φ compte pour ψ » est vrai si et seulement si tout scénario où ψ est vrai est une alternative à w dans laquelle φ est vrai.

À cela s'ajoutent les conditions (C1) et (C2) pour assurer la validité de (Agg_s) et (Disj) respectivement :

$$Y \in f_s(w, X) \text{ et } Z \in f_s(w, X) \Rightarrow\ Y \cap Z \in f_s(w, X) \tag{C1}$$
$$X \in f_s(w, Y) \text{ et } X \in f_s(w, Z) \Rightarrow\ X \in f_s(w, Y \cup Z) \tag{C2}$$

En mots, (C1) signifie que si les ensembles de propositions Y et Z sont vrais dans les scénarios accessibles à w et où X est vrai, alors l'intersection de Y et Z est aussi vrai dans les alternatives à w où X est vrai. Quant à (C2), cela signifie que si X est vrai dans les alternatives à w où Y est vrai et que X est vrai dans les alternatives à w où Z est vrai, alors X est vrai dans les alternatives à w où l'union de Y et Z est vraie (autant les scénarios où Y est vrai que ceux où Z l'est mènent à des alternatives où X est vrai).

L'introduction de l'opérateur $\Rightarrow_s$ amène Jones et Sergot à considérer la modalité (de contrainte) D_s : si φ compte pour ψ dans l'institution s, alors φ implique ψ est une contrainte pour l'institution s. Autrement dit, à partir du moment où on accepte que φ compte pour ψ relativement à une institution s, alors cette institution est contrainte par le fait que $\varphi \supset \psi$ est vrai. En ce sens, « φ et non ψ » n'est pas acceptable pour s. Cela se traduit syntaxiquement par le schéma d'axiome (D_s). Outre ce schéma d'axiome, la modalité D_s est axiomatisée par KD :

$$(\varphi \Rightarrow_s \psi) \supset D_s(\varphi \supset \psi) \qquad (D_s)$$

Du côté de la sémantique, le modèle $\mathcal{M} = \langle W, E, f_s, d_s, V\rangle$ est augmenté par une relation d'accessibilité d_s, laquelle détermine les scénarios accessibles à w relativement à une institution s. La relation d_s est sérielle (en vertu de KD) et doit satisfaire la condition (C3) :

$$Y \in f_s(w, X) \Rightarrow\ d_s(w) \cap X \subseteq Y \qquad \text{(C3)}$$

En d'autres termes, (C3), qui vise à assurer la validité du schéma (D_s), signifie que si Y est vrai dans les scénarios accessibles à w où X est vrai, alors les scénarios accessibles à w où X est vrai sont sous-ensembles de ceux où Y est vrai. La clause sémantique est la suivante :

$$\models_w D_s A \Leftrightarrow\ d_s(w) \subseteq \|A\|_{\mathcal{M}} \qquad (6)$$

S'il est vrai que φ est une contrainte pour l'institution s, alors les scénarios accessibles à w sont parmi ceux où φ est vrai.

À la fin de leur article, Jones et Sergot soulignent l'intérêt de leur approche du point de vue du discours normatif, sans pour autant ajouter une extension déontique à leur système. L'approche présentée jusqu'à présent vise à formaliser la notion de *pouvoir institutionnel*. Nous avons dû introduire le connecteur $\Rightarrow_s$ pour représenter le fait que certaines actions *comptent pour* d'autres actions dans le cadre d'une institution. Par exemple, *Paul prend l'argent de Pierre sans sa permission* compte pour un vol du point de vue légal. Même si l'approche de Jones et Sergot ne fait pas partie des logiques I/O, celle-ci a été reprise ultérieurement par des logiciens qui ont travaillé sur la représentation informatique d'un NMAS.

Guido Boella et Leendert van der Torre

Une bonne partie de la littérature scientifique sur la logique déontique porte sur ses applications en informatique et en programmation. L'approche de Boella et van der Torre (2006c), qui jumèle les logiques I/O de Makinson et van der Torre (2000, 2001, 2003a) avec l'analyse de Jones et Sergot (1996), s'insère dans cette catégorie.[8] En effet, leur objectif est de développer des outils formels pertinents à la modélisation de l'architecture d'un système normatif, facilitant ainsi l'automatisation des raisonnements. Dans ce qui suit, l'idée est de définir récursivement un système normatif par des opérations *input/output*. L'obectif est de débuter à partir de définitions de base (p. ex. permission, obligation, contraintes institutionnelles, etc.), puis de déterminer les *outputs combinés* de ces opérations, l'objectif ultime étant d'obtenir *une* opération qui représente le système global. En ce sens, l'objectif est de définir un système normatif comme *une seule* opération *input/output*, où on « entre φ et on obtient ψ ».

L'approche de Boella et van der Torre (2006c) repose notamment sur ce qui a été présenté dans Boella et van der Torre (2006a), où les auteurs introduisent l'architecture d'un système normatif. D'entrée de jeu, les auteurs utilisent l'approche de Jones et Sergot (1996), à laquelle ils ajoutent un axiome pour la transitivité (Tran) ainsi que la condition (C4) sur le modèle sémantique. Cette dernière signifie simplement que si Y est vrai pour les alternatives à w où X est vrai et que Z est vrai pour les alternatives à w où Y est vrai, alors Z est vrai pour les alternatives à w où X est vrai[9] :

$$((\varphi \Rightarrow_s \psi) \wedge (\psi \Rightarrow_s \rho)) \supset (\varphi \Rightarrow_s \rho) \qquad \text{(Tran)}$$

$$Y \in f_s(w, X) \text{ et } Z \in f_s(w, Y) \Rightarrow Z \in f_s(w, X) \qquad \text{(C4)}$$

Cela dit, Boella et van der Torre reformulent la logique des *counts-as conditionals* de Jones et Sergot en matière de logique I/O. Alors que chez Makinson et van der Torre on parlait d'ensemble générateur G, Boella et van der Torre parlent d'ensemble de paires CA (pour *counts-as*), lequel s'évalue relativement à un système normatif ou à une institution s.

L'ensemble CA est fonction de l'ensemble de normes qui caractérise le système normatif. L'opération $out_{CA}(CA, s, a)$ est régie par les règles (AND), (OR) et (T), la dernière représentant la transitivité :

$$\frac{(a, x)(x, y)}{(a, y)} \qquad \text{(T)}$$

[8] Voir aussi Boella et van der Torre (2003b, 2004, 2006a,c,b).

[9] Dans leur article, Jones et Sergot (1996) acceptent l'axiome par défaut simplement parce qu'ils n'ont pas de contre-exemple pour le rejeter. Nous l'avons mis de côté dans la présentation de leur approche puisqu'ils n'insistent pas sur celui-ci.

Empruntant l'idée de *contrainte institutionnelle*, représentée par l'opérateur modal D_s de Jones et Sergot, Boella et van der Torre introduisent l'ensemble de paires IC (*institutional constraints*), dont l'*output* $out_{IC}(IC, s, a)$ est régi par les règles (SI), (WO) et (Id).

$$\frac{}{(a, a)} \qquad \text{(Id)}$$

La définition de out_{IC} est plutôt simple :

$$out_{IC}(IC, s, a) =_{df} Cn(\{a\} \cup \{a \supset y : (a, y) \in IC\}) \qquad \text{(df. } out_{IC}\text{)}$$

Autrement dit, l'*output* des contraintes institutionnelles pour un *input* a correspond à l'ensemble des conséquences de a avec $a \supset y$ pour toute paire $(a, y) \in IC$.

L'axiome (D_s) de Jones et Sergot est remplacé par la règle suivante sur les *outputs* :

$$\frac{x \in out_{CA}(CA, s, a)}{x \in out_{IC+CA}(IC, CA, s, a)} \qquad (out_{IC+CA})$$

L'*output* out_{IC+CA} répond aux règles (AND), (OR), (T), (SI), (WO) et (Id) et est sémantiquement défini récursivement de la manière suivante :

$$\begin{aligned} out_{IC+CA}(IC, CA, s, X) & \qquad \text{(df. } out_{IC+CA}\text{)} \\ &=_{df} out_{IC}(IC \cup CA, s, X) \\ &=_{df} Cn(X \cup \{x \supset y : x \in X \text{ et } (x, y) \in IC \cup CA\}) \end{aligned}$$

En mots, cela signifie que out_{IC+CA} est par définition l'ensemble des *outputs* d'IC en union avec CA relativement à une institution s et à un *input* X. Tout ce qui est un *output* out_{CA} est un *output* out_{IC+CA}, mais certaines contraintes institutionnelles peuvent ne pas être des *counts-as*. L'*output* $out_{IC+CA}(IC, CA, s, x)$ équivaut donc à l'ensemble des conséquences de tous les $\{x, x \supset y\}$ pour lesquels $x \supset y$ est dans IC ou dans CA.

Cela fait, les auteurs introduisent une opération pour les obligations conditionnelles, qui se fait exactement dans le même esprit que la définition de l'ensemble générateur G chez Makinson et van der Torre. En effet, Boella et van der Torre présentent l'ensemble de paires O, où chaque paire de O donne une obligation (*output*) conditionnelle à un contexte (*input*). L'opération out^i_O est définie au même titre que les *outputs* 1–4 chez Makinson et van der Torre. Ensuite, les auteurs introduisent l'*output* relatif aux contraintes institutionnelles, à l'ensemble CA ainsi qu'à l'ensemble d'obligations conditionnelles.

L'opération $out_{IC+CA+O}$ est donnée par la règle suivante (nous avons omis le i) :

$$\frac{x \in out_{IC+CA}(IC, CA, s, a), y \in out_O(O, s, x)}{y \in out_{IC+CA+O}(IC, CA, O, s, a)} \qquad (out_{IC+CA+O})$$

Cette règle est accompagnée par la définition sémantique suivante :

$$\begin{aligned} out_{IC+CA+O}(IC, CA, O, s, a) & \qquad (\text{df. } out_{IC+CA+O}) \\ &=_{df} out_O(O, s, out_{IC+CA}(IC, CA, s, a)) \\ &=_{df} out_O(O, s, out_{IC}(IC \cup CA, s, a)) \\ &=_{df} out_O(O, s, Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})) \end{aligned}$$

En d'autres termes, l'*output* conjoint $out_{IC+CA+O}$ ayant a comme *input* correspond à l'*output* out_O (relativement à l'ensemble générateur O) lorsque l'*input* est l'ensemble des conséquences logiques de l'ensemble $\{a, a \supset y\}$ qui contient tous les $a \supset y$ pour lesquels (a, y) est une contrainte institutionnelle ou un *counts-as*.

Conformément à ce qu'on trouve dans Makinson et van der Torre (2001), Boella et van der Torre introduisent aussi les contraintes sur les *outputs* afin de mieux rendre compte des obligations conditionnelles. Toutefois, ces derniers utilisent une notion laissée de côté chez Makinson et van der Torre, à savoir celle de *meet*, voire de *geatest lower bound*. Tel que nous l'avons mentionné au début de ce chapitre, l'ensemble *outfamily* est obtenu après avoir mis comme *input* l'ensemble *maxfamily*, lequel permet de déterminer tous les sous-ensembles de l'ensemble générateur consistants avec l'*output*. Cela dit, *outfamily* est une collection d'ensembles, laquelle contient tous les *outputs* des sous-ensembles de l'ensemble générateur qui, lorsqu'il est pris en tant qu'*input*, sont consistants avec l'*output*.

Deux opérations possibles sur *outfamily* ont été laissées de côté dans la présentation de Makinson et van der Torre, principalement en raison du fait que ces notions, bien que mentionnées, n'étaient pas à l'étude dans leur article. Ces deux opérations sont celles du *meet* et du *join*, voire du *greatest lower bound* et du *least upper bound*, représentées respectivement par :

$$\bigcap outfamily(G, A, C)$$
$$\bigcup outfamily(G, A, C)$$

En un mot, le *meet* identifie le plus petit ensemble commun à tous les membres d'*outfamily* (l'intersection) alors que le *join* sélectionne le plus grand (l'union). À l'aide de la définition du *meet*, Boella et van der Torre définissent un nouvel *output*, soit $out^{\cap}_{O}$, où C est un ensemble de contraintes (à ne pas confondre avec l'ensemble des contraintes institutionnelles IC) :

$$out^{\cap}_{O}(O, s, a, C) =_{df} \bigcap outfamily(O, s, a, C) \qquad \text{(df. } out^{\cap}_{O})$$

Le i a encore été laissé de côté pour ne pas encombrer l'écriture. Alors que l'ensemble $outfamily(O, s, a, C)$ est la famille des *outputs* générés par les éléments $maxfamily(O, s, a, C)$, la définition de cet ensemble est paramétrée par une institution s qui détermine les contraintes.

Ayant cette opération en main, Boella et van der Torre définissent récursivement $out^{\cap}_{IC+CA+O}$ de la manière suivante, où $\Gamma = Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})$.

$$\begin{aligned}
&out^{\cap}_{IC+CA+O}(IC, CA, O, s, a) && \text{(df. } out^{\cap}_{IC+CA+O}) \\
&\quad =_{df} out^{\cap}_{O}(O, s, out_{IC+CA}(IC, CA, s, a), out_{IC+CA}(IC, CA, s, a)) \\
&\quad =_{df} \bigcap outfamily(O, s, out_{IC+CA}(IC, CA, s, a), out_{IC+CA}(IC, CA, s, a)) \\
&\quad =_{df} \bigcap outfamily(O, s, \Gamma, \Gamma)
\end{aligned}$$

Autrement dit, l'opération $out^{\cap}_{IC+CA+O}$ garantit la consistance de l'*output* des obligations avec les contraintes institutionnelles en isolant le plus petit ensemble commun à tous les membres d'*outfamily*.

Ayant traité des obligations conditionnelles et de l'ajout de contraintes au système normatif afin que les obligations soient consistantes avec les contraintes institutionnelles de s, Boella et van der Torre introduisent un ensemble de paires P représentant des permissions conditionnelles. Conjointement à cet ensemble, ils définissent l'opération (out_P) de la manière suivante :

$$\begin{aligned}
&out_P(P, s, A) && \text{(df. } out_P) \\
&\quad =_{df} \{Cn(x) : (a_i \wedge ... \wedge a_n, x) \in P \text{ pour certains } a_i, ..., a_n \in A\}
\end{aligned}$$

Cette définition est univoque : l'*output* out_P relativement à l'input A correspond à l'ensemble qui contient les ensembles des conséquences de x pour lesquels x est la permission conditionnelle à la conjonction de certaines propositions membres de A. Autrement dit, on observe toutes les combinaisons possibles des propositions de A en matière de conjonction, on isole les x qui sont les permissions conditionnelles à ces combinaisons et on détermine l'ensemble qui contient les ensembles des conséquences de chaque x.

En plus des permissions explicites introduites par l'*output* out_P, les auteurs définissent l'opération *merge* pour déterminer l'ensemble qui contient les permissions explicites, mais aussi les permissions qui découlent du fait que quelque chose est obligatoire (ce qui valide le schéma d'axiome (**D**) des systèmes standard).

$$merge(X, Y) = \{Cn(x \cup Y) : x \in X\} \qquad \text{(df. Merge)}$$

L'opération *merge* détermine l'ensemble qui contient tous les ensembles des conséquences de $x \cup Y$, où x est un ensemble de permissions et Y un ensemble d'obligations. Dans cette définition, X est donc un ensemble de sous-ensembles de propositions, alors que Y est un ensemble de propositions, conformément à la définition de out_P.

Les dernières étapes consistent à définir les *outputs* du système global. Pour ce faire, les auteurs doivent d'abord définir l'*output* conjoint $out_{IC+CA+P}$, défini de la même manière que $out_{IC+CA+O}$.

$$\begin{aligned} out_{IC+CA+P}(IC, CA, P, s, a) & \qquad \text{(df. } out_{IC+CA+P}\text{)} \\ &=_{df} out_P(P, s, out_{IC+CA}(IC, CA, s, a)) \\ &=_{df} out_P(P, s, out_{IC}(IC \cup CA, s, a)) \\ &=_{df} out_P(P, s, Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})) \end{aligned}$$

Cet *output* est régi par la règle suivante :

$$\frac{x \in out_{IC+CA}(IC, CA, s, a), y \in out_P(P, s, x)}{y \in out_{IC+CA+P}(IC, CA, P, s, a)} \qquad (out_{IC+CA+P})$$

Il reste maintenant deux opérations à définir. En premier lieu, l'opération $out^{\cap}_{IC+CA+P+O}$ s'assure que l'*output* conjoint d'IC, CA, P et O est consistant avec les contraintes institutionnelles. Dans la définition qui suit, nous avons écrit out_{IC+CA} plutôt que $out_{IC+CA}(IC, CA, s, a)$ pour faciliter la lecture :

$$\begin{aligned} out^{\cap}_{IC+CA+P+O}(IC, CA, P, O, s, a) & \qquad \text{(df. } out^{\cap}_{IC+CA+P+O}\text{)} \\ &=_{df} out^{\cap}_{O}(O, merge(out_{IC+CA+P}, out_{IC+CA}), s, out_{IC+CA}) \\ &=_{df} \bigcap outfamily(O, s, merge(out_{IC+CA+P}, out_{IC+CA}), out_{IC+CA}) \end{aligned}$$

Autrement dit, on prend le plus petit ensemble commun à tous les *outputs* d'O consistants avec les *outputs* de IC et CA et qui ont pour *input* l'ensemble qui contient toutes les combinaisons possibles entre les permissions (prises en considération avec les contraintes institutionnelles) et les contraintes institutionnelles. On prend donc comme *input* l'ensemble de tout ce qui est permis (au sens large, et non seulement explicite) en fonction des contraintes institutionnelles et des *counts-as* pour déterminer le plus petit *output* consistant avec les contraintes institutionnelles.

Finalement, l'opération $out_{IC+CA+P+O+P}$ détermine l'*output* du système normatif global. Cette opération est définie de la manière suivante (nous avons encore laissé de côté ce qui se trouve dans la portée des opérations) :

$$out_{IC+CA+P+O+P}(IC, CA, P, O, s, a) \qquad \text{(df. } out_{IC+CA+P+O+P}\text{)}$$
$$=_{df} merge(out_{IC+CA+P}, out^{\cap}_{IC+CA+P+O})$$

Somme toute, l'idée de Boella et van der Torre est de définir récursivement un système normatif à l'aide d'opérations *input/output*. Ayant défini étape par étape les différents types d'*outputs* ainsi que leurs combinaisons possibles, les auteurs parviennent à la définition d'un système normatif en termes d'entrée et de sortie.

Voici une reconstruction de l'explication à partir de la dernière définition. Prenons out_1 pour les définitions d'out_P et d'out_O, avec $P(A)$ et $O(A)$ définis dans le même esprit que $G(A)$ au début de ce chapitre. De plus, supposons que l'*input* de départ est a. L'*output* d'un système normatif est considéré comme le *merge* d'$out_{IC+CA+P}$ et d'$out^{\cap}_{IC+CA+P+O}$. Le *merge* est l'ensemble qui contient l'ensemble des conséquences de $X \cup out^{\cap}_{IC+CA+P+O}$ pour tout $X \in out_{IC+CA+P}$. Un $X \in out_{IC+CA+P}$ est un élément d'out_P, où $out_P = Cn(P(Cn(\Gamma)))$, Γ étant l'ensemble qui contient les conséquences des unions de $\{a\}$ avec $\{a \supset y\}$ pour tout y tel que $(a, y) \in IC \cup CA$.

En ce sens, un $X \in out_{IC+CA+P}$ est un ensemble qui contient les conséquences de $\{a\} \cup \{a \supset y\}$ pour un y tel que $(a, y) \in IC \cup CA$. De fait, le *merge* consiste à prendre les conséquences de $X \cup out^{\cap}_{IC+CA+P+O}$ pour chaque X qui contient les conséquences de $\{a\} \cup \{a \supset y\}$ pour un y tel que $(a, y) \in IC \cup CA$.

De manière informelle, $out_{IC+CA+P}$ équivaut à considérer l'out_P ayant comme *input* toutes les paires d'$IC \cup CA$ qui partagent le même *output*. On prend cela, on observe les conséquences, on regarde ce que cela donne comme *output* relativement à P lorsqu'on met ces conséquences en *input* et on observe les conséquences du tout.

Quant à $out^{\cap}_{IC+CA+P+O}$, il s'agit du plus petit sous-ensemble partagé par tous les *outputs* consistants avec out_{IC+CA} et ayant comme *input* des sous-ensembles de $merge(out_{IC+CA+P}, out_{IC+CA})$. En ce sens, on prend l'ensemble qui contient les conséquences de $X \cup out_{IC+CA}$ pour tout $X \in out_{IC+CA+P}$ (où out_{IC+CA} contient les conséquences de tout ce que a implique dans $IC \cup CA$), puis on détermine l'intersection d'$out_O(O, s, \Gamma)$ consistant avec $IC \cup CA$ pour tout Γ tel que :

$$\Gamma \subseteq merge(out_{IC+CA+P}, out_{IC+CA})$$

L'approche de Boella et van der Torre est réellement complexe et il est évident que de tenter de la résumer en quelques phrases simples nous éloignerait de la subtilité du formalisme. Néanmoins, par souci pédagogique, voici comment nous pourrions tenter de le faire.

Le système normatif est considéré comme une entité qui se comprend en termes d'*inputs* et d'*outputs*. On entre des données dans la « boîte », on brasse, puis on observe ce qui sort. La « boîte » du système normatif est divisée en « sous-boîtes ». Dans le cas des obligations conditionnelles, on entre l'*input*, on ajoute certaines contraintes pour s'assurer que le résultat est cohérent (et ainsi éviter le paradoxe de Chisholm), on brasse et on observe ce qui sort (c'est-à-dire les obligations qui s'appliquent). Il est aussi possible de regarder la « sous-boîte » de la permission de la même manière.

Au total, on peut créer une « boîte » constituée d'une ou de plusieurs « sous-boîte(s) » afin d'observer ce qui se passe lorsqu'on augmente l'information qu'on prend en compte. Au final, le système normatif peut être compris ainsi (en présumant qu'à chaque étape, on brasse la boîte !) : on entre un *input* a, on regarde les permissions contextuelles à a relativement à l'ensemble de permissions explicites et l'ensemble de contraintes institutionnelles (i.e., tout ce que a implique dans le cadre du système normatif s), on observe les conséquences de cela en rapport avec toutes les obligations qui tiennent dans ce contexte, et on brasse à nouveau la boîte afin d'observer les conséquences que cela nous donne.

En un mot, l'opération *input/output* d'un système normatif consiste à prendre en compte tout ce que le système présuppose (contraintes institutionnelles, implications considérées vraies dans le système, équivalences entre les actions et les descriptions de faits, *counts-as*, etc.), toutes les permissions explicites et obligations relatives à l'*input* et, finalement, à considérer les conséquences logiques du tout.

Les auteurs présument qu'un système légal se réduit à une opération *inputs/outputs* bien définie, ce que conteste la majorité des juristes, mais ce qui va aussi à l'encontre de l'esprit derrière les principes d'interprétation des lois (Côté 2006). Au moment de la rédaction des lois, le législateur ne pouvait pas prévoir toutes les façons possibles dont un système légal peut

évoluer. Par conséquent, si, *a posteriori*, il faut réévaluer cette législation à la lumière de l'information actuelle, il faut se questionner sur l'intention du législateur et sur l'esprit de la loi, ce qui n'est pas définissable d'emblée par un ensemble de contraintes institutionnelles et de *counts-as*.

Pour aller plus loin, le lecteur pourra consulter Boella et van der Torre (2006b), qui rendent compte autant des systèmes normatifs que des NMAS.

* * *

Les logiques I/O offrent une alternative aux logiques modales, fournissant un cadre de travail permettant d'éviter le dilemme de Jørgensen. Les logiques I/O permettent l'analyse des inférences normatives dans une sémantique autre que celle des mondes possibles. Passons maintenant aux approches algébriques.

Chapitre 12

Les algèbres déontiques

Outre les logiques I/O et les logiques modales, on trouve plusieurs approches algébriques au sein de la littérature scientifique. Les logiques modales étant souvent de type *ought-to-be*, certains préfèrent utiliser des algèbres d'action afin d'insister sur le caractère *ought-to-do* des opérateurs déontiques. Nous présenterons ici l'approche de Lindahl et Odelstad (2000) qui considèrent les systèmes normatifs comme des ensembles sujets à certaines opérations, puis quelques approches algébriques insistant sur la structure des actions, incluant celle de Segerberg (1982b), qui influencera celle de Trypuz et Kulicki (2009). Finalement, il sera question de Castro et Maibaum (2009) et de leur logique dynamique fondée sur une logique de l'action. En dernier lieu, nous considérerons les travaux de Lucas, qui propose une analyse de la logique de l'action de von Wright dans le cadre de la théorie des catégories.

Lars Lindahl et Jan Odelstad

L'article de Lindahl et Odelstad (2000), où les auteurs proposent une analyse algébrique des systèmes normatifs, est très fortement inspiré du texte d'Odelstad et Lindahl (2000). Outre la similarité entre ces deux articles, l'approche de Lindahl et Odelstad (2000) s'insère dans la lignée des travaux d'Alchourrón et Bulygin (1971). Sans proposer de définition précise, les auteurs entendent par *système normatif* un ensemble de lois, que les auteurs tentent de représenter par la jonction de deux (ou plusieurs) fragments d'une algèbre de Boole. Plus précisément, leur objectif est d'être en mesure d'exprimer formellement la jonction d'un fragment *normatif* à un fragment *descriptif*.

Dans le même esprit que celui des logiques I/O, Lindahl et Odelstad conçoivent un système normatif comme un processus *input/output* qui permet de déduire à partir de certains faits les conséquences légales applicables dans une situation donnée.

Essentiellement, Lindahl et Odelstad considèrent un système normatif comme un pré-ordre booléen. Cette notion repose sur celle d'une *algèbre de Boole*, que Goldblatt (2006) décrit comme un *treillis distributif complémenté*. Un *treillis* est un ensemble partiellement ordonné (en anglais *poset*, pour *Partially Ordered SET*) pour lequel chaque paire d'éléments possède une borne inférieure (en anglais notée g.l.b. pour *greatest lower bound*) et une borne supérieure (notée l.u.b. pour *least upper bound*). Un ensemble partiellement ordonné $\mathbf{P}$ est une structure $\langle P, \sqsubseteq \rangle$, où P est un ensemble et $\sqsubseteq$ est une relation d'ordre partiel. Un *ordre partiel* est un pré-ordre antisymétrique, c'est-à-dire une relation réflexive, transitive et antisymétrique. Soulignons qu'un ordre partiel n'est pas nécessairement linéaire. La borne inférieure, ici dénotée par $p \sqcap q$, est telle que :

1. $p \sqcap q \sqsubseteq p$ et $p \sqcap q \sqsubseteq q$;
2. si $c \sqsubseteq p$ et $c \sqsubseteq q$, alors $c \sqsubseteq p \sqcap q$.

La borne supérieure, ici dénotée par $p \sqcup q$, est telle que :

1. $p \sqsubseteq p \sqcup q$ et $q \sqsubseteq p \sqcup q$;
2. si $p \sqsubseteq c$ et $q \sqsubseteq c$, alors $p \sqcup q \sqsubseteq c$.

Un treillis est donc un ensemble partiellement ordonné pour lequel chaque paire d'éléments possède une borne inférieure et une borne supérieure. Un treillis est *distributif* lorsque les deux lois suivantes sont respectées pour tout x, y et z :

1. $x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z)$;
2. $x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z)$.

Un treillis distributif correspond à une algèbre de Heyting. Finalement, un treillis est *complémenté* lorsque chaque élément possède un complément, c'est-à-dire que pour tout x il y a un y tel que :

1. $x \sqcup y = 1$;
2. $x \sqcap y = 0$.

Ici, 0 est un objet initial et 1 est un objet terminal, c'est-à-dire des objets pour lesquels (respectivement) $0 \sqsubseteq p$ et $p \sqsubseteq 1$ pour tout p. Cette définition d'une algèbre de Boole est équivalente à celle proposée au chapitre 10.

Mais quel est le rapport entre un système normatif et un treillis distributif complémenté ? D'emblée, il faut voir que la structure de la logique propositionnelle classique est équivalente à celle d'une algèbre de Boole (Peterson 2015b,c). En effet, en logique classique, la borne inférieure est la conjonction, la borne supérieure la disjonction, l'objet initial est le falsum $\bot$ et l'objet terminal est le verum $\top$. De plus, nous pouvons aisément prouver que :

1. $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$;
2. $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$;
3. $p \vee \neg p \equiv \top$;
4. $p \wedge \neg p \equiv \bot$.

Le symbole $\equiv$ renvoie à l'équivalence logique et l'ordre partiel à la conséquence logique.[1]

C'est dans cet esprit que Lindahl et Odelstad parlent d'un système normatif en tant que pré-ordre booléen. Leur objectif est de déterminer les propriétés de la relation R d'un système normatif qui permettent d'interpréter aRb comme $a \supset b$.

Une *condition* peut être une proposition atomique au sens du calcul des propositions ou une proposition atomique au sens du calcul des prédicats. Bien que cela ne soit pas présenté ainsi, voici comment comprendre la notion de *condition*. Soit C_i un ensemble dénombrable de conditions qui ont i variable(s) dans leur portée. Si $i = 0$, alors nous obtenons l'ensemble des conditions qui ne portent sur aucune variable et sont représentées par des atomes propositionnels. Lorsque $i = 1$, on obtient les conditions représentées par des prédicats unaires (p. ex. x est majeur), $i = 2$ donne des prédicats binaires (p. ex. x est le père de y), $i = 3$ des prédicats ternaires, etc. Lorsque la condition a_i est définie, la condition $a'_i(x_1, ..., x_i)$ représente la négation de la condition $a_i(x_1, ..., x_i)$, $(a_i \wedge b_i)(x_1, ..., x_i)$ la conjonction $a_i(x_1, ..., x_i)$ et $b_i(x_1, ..., x_i)$, et la condition $(a_i \vee b_i)(x_1, ..., x_i)$ représente la disjonction $a_i(x_1, ..., x_i)$ ou $b_i(x_1, ..., x_i)$, où $a \vee b =_{df} (a' \wedge b')'$.

[1] Cependant, lorsqu'on traite la logique classique dans un cadre algébrique, il faut considérer un ordre partiel entre des classes d'équivalences de propositions.

Soit C, l'ensemble de toutes les conditions atomiques possibles (portant sur le nombre de variables i, avec $C_i = \{c_i^j : j \in \mathbb{N}\}$).

$$C = \bigcup_{i \in \mathbb{N}} C_i$$

L'ensemble des énoncés bien formés pour les conditions est défini récursivement à partir de l'ensemble C de toutes les conditions atomiques[2] :

1. si $a \in C$, alors $a \in EBF$;
2. si $\varphi, \psi \in EBF$, alors $\varphi', \varphi \wedge \psi \in EBF$.

Ainsi, en posant $\mathbf{EBF} = \langle EBF, \sqsubseteq \rangle$, on obtient une algèbre de Boole $\langle \mathbf{EBF}, \wedge, ' \rangle$, où $\bot$ et $\top$ sont respectivement l'objet initial et l'objet terminal.

À partir de cela, les auteurs définissent un *système normatif* $\mathcal{S} = \langle \mathcal{N}, \wedge, ', R \rangle$ (où $\mathcal{N} \subseteq EBF$ est un ensemble partiellement ordonné) lorsque pour tout $a, b \in \mathcal{N}$, aRb si et seulement si a implique b selon $\mathcal{S}$. Un parallèle mérite d'être tracé entre la conception d'un système normatif de Lindahl et Odelstad, et l'ensemble des contraintes institutionnelles de Boella et van der Torre. Au même titre que l'ensemble des contraintes institutionnelles contient les paires (a, x) pour lesquelles a compte pour x au sein du système normatif (i.e., pour lesquelles $a \supset x$ est vrai), un système normatif est conçu chez Lindahl et Odelstad comme une structure qui détermine un ensemble de paires $(a, b) \in R$ pour lesquelles $a \supset b$ est vrai dans le système normatif. La principale différence est que chez Lindahl et Odelstad, $\mathcal{S}$ contient à la fois les contraintes institutionnelles et les ensembles (générateurs) d'obligations et de permissions. Cela dit, dans les cas où a est descriptif et que b est normatif, donc que (a, b) est plutôt dans un ensemble générateur d'obligations ou de permissions chez Boella et van der Torre, Lindahl et Odelstad parleront de *corrélations normatives*.

Afin de formaliser la notion de système normatif, Lindahl et Odelstad introduisent la notion de *pré-ordre booléen*, une algèbre de Boole sur laquelle on impose un pré-ordre R, c'est-à-dire une relation réflexive et transitive, qui doit satisfaire les conditions suivantes :

1. si aRb et aRc, alors $aR(b \wedge c)$;
2. si aRb, alors $b'Ra'$;
3. $(a \wedge b)Ra$;
4. non $\top R \bot$.

[2] Nous avons modifié la formulation de la définition récursive des énoncés bien formés afin que cela soit cohérent avec la suite de leur texte ainsi qu'avec la position d'Odelstad et Lindahl (2000).

Le lecteur familiarisé avec la logique propositionnelle aura remarqué que cela se résume aux conditions suivantes :

1. si a implique respectivement b et c, alors a implique la conjonction de b et c ;
2. si a implique b, alors non a implique non b (contraposition) ;
3. une conjonction implique chacun de ses membres ;
4. le vrai n'implique pas le faux.

Soulignons au passage que Lindah et Odestad utilisent un pré-ordre booléen plutôt qu'une algèbre de Boole pour décrire l'implication matérielle par la relation R et ainsi éviter de traiter des classes d'équivalences. En utilisant un ordre partiel plutôt qu'un pré-ordre, on obtient que aRb et bRa implique que $a = b$, ce qui requiert qu'on parle d'identité sur les classes d'équivalences des propositions plutôt que sur les propositions elles-mêmes. Les auteurs préfèrent utiliser un pré-ordre afin de pouvoir faire la distinction entre deux propositions avec une signification différente bien que aRb et bRa.

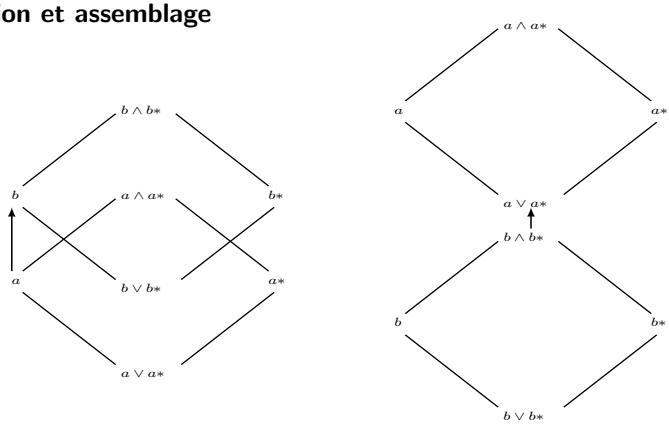
Dans le but de rendre compte des corrélations normatives, c'est-à-dire des conditionnels qui mettent en jeu un antécédent descriptif et un conséquent normatif, Lindahl et Odelstad représentent un système normatif $\mathcal{S}$ comme la jonction de deux (ou plusieurs) sous-structures $\mathcal{N}_1$ et $\mathcal{N}_2$. Dans un tel contexte, $\mathcal{S}_1 = \langle \mathcal{N}_1, \wedge, ', R_1 \rangle$ et $\mathcal{S}_2 = \langle \mathcal{N}_2, \wedge, ', R_2 \rangle$ sont des *fragments* de $\mathcal{N}$, c'est-à-dire des sous-structures pour lesquelles $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{N}$ et $R_i \subset R$ est une sous-relation qui isole les implications pertinentes à $\mathcal{N}_i$. Deux fragments peuvent être joints de deux manières, en l'occurrence la *connexion* et l'*assemblage* (*coupling*).

Dans le cas de deux fragments $\mathcal{S}_1$ et $\mathcal{S}_2$, une *connexion* a lieu lorsqu'il y a $a \in \mathcal{N}_1$ et $b \in \mathcal{N}_2$ tels que aRb. Un *assemblage* est tel qu'il y a une seule connexion entre $\mathcal{S}_1$ et $\mathcal{S}_2$. En fait, l'*assemblage* se différencie de la connexion dans la mesure où $\mathcal{S}_1$ et $\mathcal{S}_2$ sont uniquement connectés de par la borne supérieure de $\mathcal{S}_1$ et la borne inférieure de $\mathcal{S}_2$. En ce sens, dans le cas d'une *connexion*, il y a des liens directs entre les conditions de $\mathcal{S}_1$ et celles de $\mathcal{S}_2$, alors que dans le cas de l'*assemblage*, toute relation est indirecte et requiert le passage par un intermédiaire.

L'assemblage permet de mettre l'ensemble $\mathcal{S}_1$ en relation avec $\mathcal{S}_2$, alors que la connexion ne met qu'une partie de $\mathcal{S}_1$ en relation avec $\mathcal{S}_2$. Considérons la figure 12.1, où a, $a*$ et b, $b*$ représentent respectivement des membres de $\mathcal{S}_1$ et $\mathcal{S}_2$. À gauche, la flèche entre a et b représente une connexion à partir de laquelle il est seulement possible de lier $a \vee a*$ et a à $b \wedge b*$. En ce sens, la connexion ne permet de lier qu'une partie des deux fragments. Notons qu'il pourrait très bien y avoir une autre connexion de $a*$

à $b*$. Dans le cas de droite, la flèche représente plutôt un assemblage puisqu'il s'agit de la seule connexion, laquelle permet de lier chaque membre de $\mathcal{S}_1$ avec ceux de $\mathcal{S}_2$.

FIGURE 12.1
Connexion et assemblage



L'idée de Lindahl et Odelstad est donc de représenter un système normatif complexe, par exemple la législation canadienne, par un pré-ordre booléen $\mathcal{S}$ qu'on peut diviser en fragments, comme le Code criminel, le Code civil, la Charte des droits et libertés, etc., lesquels peuvent aussi être divisés en fragments, et ainsi de suite. Cela dit, chaque fragment peut être divisé en deux fragments, l'un descriptif l'autre normatif. En ce sens, la législation canadienne consiste en la jonction d'un fragment descriptif et d'un normatif, de même pour le Code criminel, le Code civil, etc. Selon cette conception, un système normatif peut donc être divisé de plusieurs manières afin d'obtenir différents fragments selon ce qu'on cherche à étudier. Dans l'ensemble, un système normatif correspond à la jonction d'un fragment qui contient des « vérités » descriptives (ou du moins des conventions, ce qui est pris pour acquis) avec un fragment qui contient des normes.

Il s'agit là d'un portrait brossé à gros traits. Pour rendre justice à l'approche de Lindahl et Odelstad, il faudrait plutôt comprendre un système normatif comme la jonction d'un fragment descriptif, d'un fragment de corrélations normatives (qui font le pont entre le descriptif et le normatif) et d'un fragment purement normatif. Dans la suite de leur article, les auteurs utilisent leur conception d'un système normatif comme pré-ordre booléen afin d'étudier les notions de *concept intermédiaire* et de *conséquence légale*, lesquels s'interprètent en fonction de la relation R transitive qui permet de passer d'un fragment à l'autre. La notion de concept intermédiaire se comprend en parallèle avec la notion de *counts-as* chez Jones et Sergot ou chez Boella et van der Torre (chapitre 11).

À titre d'exemple de concept intermédiaire, considérons l'énoncé suivant : tout citoyen canadien a le droit de vote. Autrement dit, si x est un citoyen canadien, alors x a le droit de vote. Mais qu'est-ce qui *compte pour* (*counts-as*) un citoyen canadien ? Notamment, le fait que x soit né au Canada de parents ayant la citoyenneté canadienne et qu'il réside au Canada compte pour « x est un citoyen canadien ». Dans ce cas, « citoyen canadien » est un concept intermédiaire entre les conditions et le fait que x ait le droit de vote.

Dans leur texte de 2002, les normes sont conçues comme des corrélations normatives, c'est-à-dire comme les connexions entre les fragments purement descriptifs et ceux purement normatifs. Dans celui de 2003, ils utilisent les pré-ordres booléens afin d'étudier les changements de normes ainsi que la préservation de la cohérence du système normatif dans les cas d'obligations contraires.

Le lecteur intéressé trouvera dans Lindahl et Odelstad (2004) une combinaison des pré-ordres booléens avec une forme de logique de l'action, où ils reprennent l'analyse des concepts intermédiaires dans le cadre algébrique pour la vulgariser et rendre le tout pertinent pour les juristes.

Lindahl et Odelstad (2006) reprennent ce qui a été présenté en 2000 par rapport aux concepts intermédiaires et un exemple légal plus complexe est donné dans l'article de 2008b. Les travaux publiés en 2006 et en 2008b ont aussi été repris dans les textes de 2008a et 2011.

Krister Segerberg

En plus de ses travaux en logique déontique dynamique, Segerberg (1982b) a élaboré la notion de *modèle urne*, axée sur une algèbre déontique d'action. À l'instar de son approche en logique déontique dynamique, une action est considérée en parallèle avec l'ensemble des séquences d'états qui mènent à un certain résultat.

Le langage $\mathcal{L} = \{Act, \smile, \frown, \bar{}, =, P, F, \neg, \supset\}$ fait la distinction entre les *termes*, qui font référence aux actions, et les *propositions*. Le langage contient un ensemble dénombrable d'actions atomiques (d'événements) ainsi que les constantes $\mathbf{0}$ et $\mathbf{1}$, qui sont respectivement les actions impossible et universelle. Le langage contient les symboles $\smile$, $\frown$ et $\bar{}$, qui représentent respectivement l'union, la conjonction et la négation d'action. Le symbole $=$ renvoit à l'identité sur les actions et les opérateurs P et F tiennent pour la permission et l'interdiction. Les autres connecteurs sont définis usuellement.

Les termes bien formés sont définis récursivement :

$$\alpha := a_i \mid \mathbf{0} \mid \mathbf{1} \mid \overline{\alpha} \mid \alpha \smile \beta \mid \alpha \frown \beta$$

L'ensemble des énoncés bien formés est défini récursivement de la manière suivante :

$$\varphi := P\alpha \mid F\alpha \mid \alpha = \beta \mid \neg\varphi \mid \varphi \supset \psi$$

Une structure $\mathcal{S} = \{\mathbf{A}, \frown, \smile, -, 0, 1, F, P\}$ ($\mathbf{A} \subseteq Act$) est qualifiée d'*algèbre déontique d'action* lorsque celle-ci respecte les trois conditions suivantes :

1. $\mathcal{S} = \{\mathbf{A}, \frown, \smile, -, 0, 1\}$ est une algèbre de Boole ;
2. F et P sont des ensembles d'alternatives idéales ;
3. $F \cap P = \{\mathbf{0}\}$ (i.e., rien n'est à la fois permis et interdit).

Les conditions sémantiques sont représentées par la fonction $v : Act \longrightarrow \{0, 1\}$ qui attribue des valeurs de vérité aux actions atomiques. Les clauses sémantiques sont les suivantes :

$$\begin{aligned}
v(\mathbf{0}) &= 0 \\
v(\mathbf{1}) &= 1 \\
v(\alpha \frown \beta) &= v(\alpha) \frown v(\beta) \\
v(\alpha \smile \beta) &= v(\alpha) \smile v(\beta) \\
v(\overline{\alpha}) &= -v(\alpha) \\
v(\alpha = \beta) &= \begin{cases} \top & \text{si } v(\alpha) = v(\beta) \\ \bot & \text{sinon} \end{cases} \\
v(F\alpha) &= \begin{cases} \top & \text{si } v(\alpha) \in F \\ \bot & \text{sinon} \end{cases} \\
v(P\alpha) &= \begin{cases} \top & \text{si } v(\alpha) \in P \\ \bot & \text{sinon} \end{cases}
\end{aligned}$$

Les deux premières clauses sont relativement simples : pour toute algèbre déontique d'action, $v(\mathbf{0})$ est l'objet initial 0 et $v(\mathbf{1})$ est l'objet terminal 1. En ce sens, considérant qu'il s'agit d'une algèbre de Boole, nous obtenons à la fois $\alpha \frown \overline{\alpha} = 0$ et $\alpha \frown \overline{\alpha} = \mathbf{0}$, d'une part dans la sémantique, de l'autre dans la syntaxe. Côté syntaxique, cela indique que la conjonction d'actions α et $\overline{\alpha}$ équivaut à l'action impossible $\mathbf{0}$.

Du point de vue de l'action, α est une action *tautologique* (*universelle*) lorsque $v(\alpha) = 1$ et α est une action *contradictoire* (*impossible*) lorsque $v(\alpha) = 0$. Dans les autres cas, il s'agit d'actions *satisfiables*, qui ont le potentiel d'être réalisées ou non.

Le lecteur aura remarqué que $v(\alpha)$ n'est pas défini pour les actions atomiques. C'est qu'il ne faut pas y réfléchir en termes d'ensembles : dans les ensembles, nous dirions que $v(\alpha)$ correspond à l'ensemble qui contient les scénarios où l'action α est accomplie. Ici, ce n'est pas tout à fait juste. Un parallèle qui peut être tracé (et qui sera explicité sous peu lorsque nous aborderons les modèles déontiques d'action) est que l'action est considérée comme un *événement*, et donc que celle-ci est informellement comprise comme faisant référence à une multitude de séquences de descriptions qui peuvent être interprétées comme « l'action α ».

Toutefois, dans le cas d'une algèbre déontique d'action, la clause sémantique ne dit pas simplement que $v(\alpha) = 1$ lorsque $\alpha \in \mathbf{A}$. Plutôt, il faut concevoir (informellement) une infinité de valeurs intermédiaires : $v(\alpha)$ est égal à *quelque chose*, ce quelque chose n'étant pas réductible à 0 ou 1 dans le cas des actions atomiques. Toujours informellement, ce quelque chose peut être entendu comme un ensemble de regroupements de points (où chaque regroupement constitue une exécution de α). En ce sens, l'attribution faite par v à :

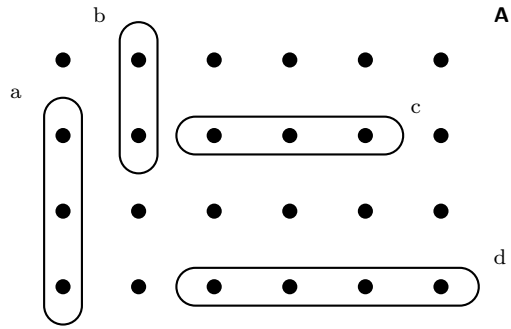
1. l'action impossible est 0 ;
2. l'action universelle est 1 ;
3. la conjonction d'actions correspond aux regroupements de points conjoints de chaque action ;
4. la disjonction d'actions correspond à au moins un regroupement de points d'une action ;
5. la négation d'action correspond aux regroupements de points autres que ceux qui correspondent à α.

Soulignons que cette explication sous forme de regroupements de points est informelle et a une visée purement pédagogique. Il ne s'agit pas de ce qui est présenté dans l'article de Segerberg. En fait, il s'agit d'une interprétation de ce qui est présenté, principalement motivée par le fait qu'aucune explication relativement à $v(\alpha)$ pour une action atomique n'y est présentée. Cette interprétation provient notamment du parallèle entre l'algèbre déontique d'action et le modèle déontique d'action.

Pour l'instant, poursuivons l'analogie avec les regroupements de points. Considérons la figure 12.2. Une action atomique α peut être vue comme dénotant les regroupements a,b,c et d. Aux fins de notre exemple, imaginons que α correspond aux regroupements a et b, et que β correspond à c et d. La conjonction d'actions $\alpha \frown \beta$ correspond à l'accomplissement simultané d'un regroupement (p. ex. a et b et c et d). La disjonction d'actions $\alpha \smile \beta$ correspond à l'accomplissement de l'un des regroupements pour une

des actions (p. ex. a ou b, ou c ou d). Quant à la négation $\overline{\alpha}$, il s'agit de tous les regroupements autres que a et b. En ce sens, $v(\alpha)$ pour une action atomique correspond à un ensemble de regroupements de points dans **A**.

FIGURE 12.2
Action et regroupements de points



Alors que dans le cas de l'attribution faite par v pour les actions, on ne peut pas réellement parler de *valeur de vérité*, v assigne les valeurs $\bot$ (*falsum*) et $\top$ (*verum*) pour les propositions atomiques. La sixième clause signifie qu'il est vrai que $\alpha = \beta$ à condition que $v(\alpha) = v(\beta)$. Par conséquent, deux actions sont identiques lorsqu'elles renvoient exactement aux mêmes regroupements de points. En ce sens, deux actions différentes peuvent partager certaines suites de points.

Les clauses pour la permission et l'interdiction s'interprètent de la même manière. À supposer que F et P renvoient à des ensembles d'actions interdites et permises, $F\alpha$ (respectivement, $P\alpha$) est vrai lorsque $v(\alpha)$ est membre de F (respectivement, de P).

De fait, on doit comprendre v comme une fonction qui, sur le plan des actions, détermine des ensembles de points (lesquels représentent des actions) et qui, sur le plan des propositions, attribue des valeurs de vérité. Il est vrai que α est interdite à condition que l'ensemble qui contient les regroupements de points qui correspondent à des exécutions de α, c'est-à-dire $v(\alpha)$, est membre d'un ensemble qui contient des regroupements de points interdits. Dans le cas des propositions complexes formées par des connecteurs propositionnels, l'attribution faite par v se fait usuellement.

Une proposition φ est dite *vraie* pour $\mathcal{S}$ lorsque $v(\varphi) = \top$ et elle est *valide* lorsqu'elle est vraie pour tout $\mathcal{S}$ et tout v.

Outre l'interprétation algébrique proposée par Segerberg, cet auteur construit aussi un modèle en termes ensemblistes. Soit $\mathbb{S} = \langle U, \mathcal{I}ll, \mathcal{L}eg \rangle$ où $U \neq \varnothing$ est un ensemble de points (d'états) et $\mathcal{I}ll$ et $\mathcal{L}eg$ sont des sous-ensembles de U qui ne partagent aucun membre. Informellement, les sous-

ensembles de $\mathcal{I}ll$ représentent des ensembles illégaux de points alors que ceux de $\mathcal{L}eg$ représentent des ensemble légaux de points. Soulignons que $\mathcal{L}eg$ et $\mathcal{I}ll$ ne sont pas le complément l'un de l'autre dans U.

Un modèle $\mathbb{M} = \langle U, \mathcal{I}ll, \mathcal{L}eg, V\rangle$ est construit à l'aide d'une fonction d'attribution de valeurs de vérité V qui se comporte dans le même esprit que v pour les algèbres déontiques d'actions. C'est d'ailleurs ici que le parallèle susmentionné devient évident relativement à l'interprétation informelle de ce qu'est une action : une action est considérée comme un *événement*, c'est-à-dire comme un ensemble de *résultats* (*outcome*) qui prennent place à la suite d'une séquence d'états subséquents. La fonction V est définie de manière similaire à v :

$$
\begin{aligned}
V(\mathbf{0}) &= \varnothing \\
V(\mathbf{1}) &= U \\
V(\alpha \frown \beta) &= V(\alpha) \cap V(\beta) \\
V(\alpha \smile \beta) &= V(\alpha) \cup V(\beta) \\
V(\overline{\alpha}) &= U - V(\alpha) \\
V(\alpha = \beta) &= \begin{cases} \top & \text{si } V(\alpha) = V(\beta) \\ \bot & \text{sinon} \end{cases} \\
V(F\alpha) &= \begin{cases} \top & \text{si } V(\alpha) \subseteq \mathcal{I}ll \\ \bot & \text{sinon} \end{cases} \\
V(P\alpha) &= \begin{cases} \top & \text{si } V(\alpha) \subseteq \mathcal{L}eg \\ \bot & \text{sinon} \end{cases}
\end{aligned}
$$

Autrement dit, l'action impossible et l'action universelle sont représentées respectivement par l'ensemble vide et l'univers en entier. La conjonction d'actions se représente par l'intersection entre les résultats possibles pour chaque action, c'est-à-dire que la conjonction d'actions correspond aux séquences qui mènent à des résultats communs à chaque action. La disjonction d'actions correspond à l'union des ensembles qui représentent la réalisation de chaque action. La négation d'une action équivaut au complément dans U de l'action elle-même.

Les clauses pour les propositions atomiques sont similaires à celles de v : deux actions sont identiques si elles renvoient exactement aux mêmes ensembles de points, une action est interdite lorsque les ensembles de points auxquels elle renvoie sont illégaux et une action est permise lorsque les ensembles de points auxquels elle renvoie sont légaux. Les conditions pour les connecteurs propositionnels se font comme à l'habitude. Tout comme dans le cas de v, un énoncé est dit vrai pour un modèle lorsque $V(A) = \top$ et valide lorsqu'il est vrai pour tout modèle.

La syntaxe du système axiomatique proposé est assez direct. Segerberg n'utilise que la règle *modus ponens*, les schémas d'axiomes de la logique propositionnelle classique, les axiomes qui correspondent à une algèbre de Boole pour les actions (distributivité et complémentation), les axiomes usuels pour l'identité :

$$\alpha = \alpha \tag{Id}$$
$$\alpha = \beta \supset (\varphi \supset \varphi(\alpha|\beta)) \tag{Subst}$$

Le premier exprime l'identité de chaque action avec elle-même et le second indique que si deux actions sont identiques, alors elles peuvent être substituées dans une proposition. Finalement, les axiomes pour les opérateurs déontiques sont les suivants :

$$F(\alpha \smile \beta) \equiv (F\alpha \wedge F\beta) \tag{1}$$
$$P(\alpha \smile \beta) \equiv (P\alpha \wedge P\beta) \tag{2}$$
$$a = \mathbf{0} \equiv (F\alpha \wedge P\alpha) \tag{3}$$

L'axiome (1) dit que si une disjonction d'actions est interdite, alors chaque action est interdite individuellement. L'axiome (2) représente la notion de permission libre de choix (*free choice permission*), à savoir que si une disjonction d'actions est permise, alors chaque action prise individuellement est permise. Finalement, (3) signifie que si une action est impossible, alors elle est à la fois permise et interdite.

Dans la suite de son article, Segerberg étudie les systèmes clos où $\mathcal{I}ll$ est le complément de $\mathcal{L}eg$ dans U (et donc où toute action est explicitement permise ou interdite). De plus, cet auteur propose deux définitions distinctes sur la base d'un seul opérateur primitif. Par exemple, il propose de définir une forme d'obligation et de permission sur la base de l'opérateur F :

$$O_F\alpha =_{df} F\overline{\alpha} \tag{df. O_F}$$
$$P_F\alpha =_{df} \neg F\alpha \tag{df. P_F}$$

Lorsqu'ils sont définis ainsi, les opérateurs se comportent comme les opérateurs usuels en logique déontique standard. Contrairement à la permission définie par P qui exprime un choix libre entre deux actions, la permission définie par P_F est une permission faible (implicite, négative, ce qui n'est pas interdit).

Robert Trypuz et Piotr Kulicki

Les travaux de Segerberg (1982b) ont notamment influencé ceux de Trypuz et Kulicki (2009, 2010), dont l'objectif est de construire diverses extensions et de comparer ainsi les travaux de Segerberg avec l'approche plus récente de Castro et Maibaum (2009). Dans ce qui suit, nous nous limiterons à la syntaxe ainsi qu'à la conception de l'action atomique de Trypuz et Kulicki (2009).

Conformément aux travaux de Segerberg, Trypuz et Kulicki font la distinction entre les *actions* et les *propositions*. Soit le langage $\mathcal{L} = \{(,), GA, 0, 1, \bar{}, \sqcap, \sqcup, \neg, \wedge, \bot, P, F\}$. La logique de l'action est formalisée par une algèbre de Boole et est construite à partir des connecteurs usuels. Comme nous le verrons, GA est un ensemble (fini) de générateurs d'actions a_i. L'ensemble des énoncés bien formés pour la logique de l'action est défini récursivement :

$$\alpha := a_i \mid 0 \mid 1 \mid \overline{\alpha} \mid \alpha \sqcap \beta \mid \alpha \sqcup \beta$$

On interprète habituellement les connecteurs comme la négation (le complément), la conjonction et la disjonction d'action (l'ordre partiel est défini par $\alpha \sqsubseteq \beta =_{df} (\alpha \sqcap \beta) = \alpha$).

Cette logique de l'action se combine à une logique propositionnelle construite à partir de l'ensemble d'énoncés bien formés défini récursivement de la manière suivante :

$$\varphi := \top \mid \alpha = \beta \mid P(\alpha) \mid F(\alpha) \mid \neg\varphi \mid \varphi \wedge \psi$$

Les autres connecteurs propositionnels sont définis de manière usuelle. L'opérateur P représente la permission forte, et F représente l'interdiction. La permission faible est définie par $P_w(\alpha) =_{df} \neg F(\alpha)$.

Le système $\mathcal{DAL}^0$ est construit à partir de $\mathcal{L}$, des axiomes du calcul propositionnel, des axiomes d'identités (AI1) et (AI2), et des axiomes (D1), (D2) et (D3) pour les opérateurs déontiques :

$$\alpha = \alpha \tag{AI1}$$
$$(\alpha = \beta) \supset (\varphi \supset \varphi(\alpha/\beta)) \tag{AI2}$$
$$P(\alpha \sqcup \beta) \equiv P(\alpha) \wedge P(\beta) \tag{D1}$$
$$F(\alpha \sqcup \beta) \equiv F(\alpha) \wedge F(\beta) \tag{D2}$$
$$(\alpha = 0) \equiv P(\alpha) \wedge F(\alpha) \tag{D3}$$

Plusieurs extensions se construisent sur la base de $\mathcal{DAL}^0$ et la sémantique s'inspire des travaux de Segerberg.

L'intérêt de cette approche se révèle lorsqu'on considère la conception que les auteurs se font de l'action atomique. Trypuz et Kulicki font la distinction entre une *action atomique* et un *générateur d'actions*. Alors qu'un *générateur d'actions* est membre de l'ensemble fini $GA = \{a_1, \ldots a_n\}$, une *action atomique* est composée de générateurs d'actions. En effet, il s'agit d'une composition de la forme $\delta_1 \sqcap \cdots \sqcap \delta_n$ où δ_k est soit un générateur soit son complément. Cette distinction est importante dans la mesure où les auteurs soutiennent qu'une action est la description complète de ce qu'un agent accomplit dans une situation donnée.

Bien que cette perspective puisse être utile en informatique (l'utilisation d'un ensemble fini de générateurs permet la description finie d'un programme), celle-ci pose cependant certains problèmes conceptuels d'un point de vue philosophique lorsqu'on considère l'action humaine. Cela se constate sur deux plans.

D'une part, il est impossible de décrire *complètement* (description finie) ce qu'un agent accomplit dans la mesure où il y a un nombre dénombrable d'actions que nous n'accomplissons pas, même si on accepte un ensemble fini de générateurs. Par exemple, prenons une action a_i qui n'est pas accomplie. Si a_i n'est pas accomplie, alors la conjonction de a_i avec a_i ne l'est pas non plus, comme la conjonction de a_i avec a_i avec a_i, etc. Par ce raisonnement, on peut aisément montrer qu'un nombre dénombrable d'actions ne sont pas accomplies. Par conséquent, si la description complète de ce qu'un agent accomplit requiert la description complète de ce qu'il n'accomplit pas, alors une description complète (finie) de ce qu'un agent accomplit est impossible puisqu'il y a un nombre dénombrable d'actions que l'agent n'accomplit pas.

D'autre part, cette restriction est discutable : bien qu'un agent ne puisse accomplir qu'un nombre fini d'actions, pourquoi présumer qu'il n'y a qu'un nombre fini de générateurs d'actions ? Cela n'est pas explicitement mentionné chez Trypuz et Kulicki, mais il est probable que cette hypothèse vise l'application de leur système en informatique. En supposant un nombre fini de générateurs d'actions, on peut faire une description complète de chaque action atomique pouvant être accomplie par le programme. Ainsi, une description complète du programme, incluant ce qu'il peut faire, ce qu'il doit faire et ce qu'il ne doit pas faire, est possible.

Cela dit, du point de vue de l'action humaine, cette hypothèse est superflue dans la mesure où le nombre d'actions possibles pouvant être accomplies par un agent est tellement grand qu'une description totale de ce que l'agent fait est impossible. Prenons simplement l'ensemble des endroits possibles dans l'univers où quelqu'un peut se trouver. La

description complète de « Paul est dans son bureau » requiert qu'on spécifie partout où Paul ne se trouve pas, ce qui, lorsqu'on considère l'univers en entier, entraîne beaucoup de possibilités ! De surcroît, il faudra préciser tout ce que Paul ne fait pas à chacune de ces positions. Par conséquent, l'hypothèse d'un nombre fini de générateurs d'actions perd son utilité si on considère que la description totale ne pourra pas être accomplie.

Par ailleurs, les auteurs affirment qu'on peut interpréter un générateur d'actions comme l'ensemble des états possibles qu'il est susceptible d'engendrer. Il s'agit là de la même idée qu'on retrouve chez Segerberg : une action renvoie à l'ensemble des résultats possibles qui peuvent en découler. Bien qu'elle soit anodine en elle-même, cette conception peut cependant difficilement être jumelée à l'autre hypothèse de Trypuz et Kulicki, à savoir qu'une action peut être décrite dans sa totalité.

En effet, si l'action se résume à l'ensemble de résultats possibles et qu'elle peut être décrite complètement, alors l'ensemble des résultats possibles peut, lui aussi, être décrit complètement. Or l'ensemble des résultats possibles est une totalité, et une totalité peut être décrite complètement que si chacune de ses parties l'est aussi, sans quoi une partie de la totalité ne serait pas décrite complètement, et de fait, la totalité ne le serait pas non plus. Par conséquent, si une action peut être décrite complètement, alors chaque état peut aussi être décrit complètement.

Rappelons-nous ici qu'un *état*, dans l'interprétation ensembliste, correspond à un scénario et que l'action est interprétée comme un regroupement d'états (de points). Le problème avec ce raisonnement est que, comme le souligne Hansson (1969), une description *totale* d'un scénario est au mieux problématique, sinon peu plausible.

Un état correspond à une description de l'univers (U) à un certain moment. Il s'agit, en un certain sens, d'une *photo* : en un clic, on fige l'univers dans une certaine description. Or, même si la description complète de U dans le cas d'une application informatique peut être faite, elle pose problème dans des systèmes plus complexes.

Prenons le Canada. Est-il possible d'en faire une description complète au moment présent ? Logiquement, c'est possible, certes, mais, pour reprendre le raisonnement de Hempel (1972), c'est irréaliste d'un point de vue pratique : il faudrait décrire l'emplacement de chaque grain de sable, de chaque roche, de chaque insecte, de chaque bactérie etc. Dans le meilleur cas, nous ne pouvons avoir qu'une description partielle d'un état du système. Comme l'utilité de l'hypothèse (les actions atomiques sont en nombre fini) vise la possibilité (réelle) de faire une description totale du système et qu'il est impossible (dans les faits, mais non pas logiquement) de le faire, il en résulte que cette hypothèse est inutile à la modélisation de l'action humaine.

En dernier lieu, soulignons un problème avec la conception de l'action atomique de Trypuz et Kulicki et, de manière générale, de Segerberg et des autres approches en logique dynamique où on considère une action comme une suite d'événements. Considérons la notion de générateur d'actions, entendue comme une brique à partir de laquelle une action atomique est construite ; les auteurs donnent l'exemple des générateurs *être assis* (*sit*) et *fumer* (*smoke*). Les quatre actions atomiques possibles construites à partir de ces générateurs sont :

$$sit \sqcap smoke$$
$$sit \sqcap \overline{smoke}$$
$$\overline{sit} \sqcap smoke$$
$$\overline{sit} \sqcap \overline{smoke}$$

Dans le cas du générateur *sit*, un « fondement à partir duquel il est possible de construire des actions atomiques »[3], on remarque que ce fondement n'est pas fondamental et que le critère qui permet de différencier un générateur de ce qui n'est pas un générateur n'est pas tellement évident. En fait, un générateur d'actions peut toujours être divisé en sous-générateurs. L'action de boire un verre d'eau, par exemple, se résume à tendre le bras, serrer les doigts autour du verre, le lever, ouvrir la bouche, pencher le verre afin d'y faire couler l'eau, l'avaler, etc.

Existe-t-il un générateur d'actions qui, lui-même, n'est pas réductible à des sous-générateurs ? L'intuition derrière ce raisonnement est que, considérant qu'une action implique un mouvement, le seul bloc indivisible possible d'une action est une description statique de chaque moment de l'action dans le temps. Pensons à l'action de serrer les doigts autour du verre d'eau : au temps 1, les doigts sont immobiles, au temps 2, ils ont bougé un peu, au temps 3, un peu plus, et ainsi de suite. Mais si on considère que l'action est un mouvement, ce qui implique le passage d'un moment à un autre, alors, conceptuellement, il est toujours possible de penser à une subdivision de ce mouvement en intervalles plus petits.

Le problème ici est que l'action, en tant que totalité, n'est pas réductible à la somme de ses descriptions. Une action peut être décrite, certes, mais cette description ne capture pas conceptuellement ce que signifie pour une action d'être une action. Revenons au générateur *sit*. Est-ce réellement un bloc indivisible ? Si on pense seulement au *résultat* de l'action de s'asseoir, alors *être assis*, pris en tant que description d'un état statique, est effectivement un bloc indivisible. Cela dit, une action est quelque chose

[3] Soulignons que l'emploi de l'expression action *atomique* est discutable considérant qu'une action atomique est divisible en générateurs d'actions.

de dynamique : une action est quelque chose qui est *fait*, et la notion de *faire quelque chose* requiert une évolution dans le temps. En ce sens, l'action *être assis* ne se résume pas à une description statique : cette action est faite par un agent d'un temps t_1 à un temps t_n. L'action d'*être assis* nécessite plusieurs sous-actions. Par exemple, les muscles du dos doivent travailler afin de maintenir la posture, ceux des jambes travailleront afin de les garder dans une certaine position, sans compter toutes les autres actions fondamentales à l'accomplissement de l'action *sit*.

L'action d'un muscle elle-même peut se concevoir par une suite d'actions différentes : la transformation énergétique des cellules, la génération d'influx nerveux, les sous-mouvements nécessaires à l'accomplissement du mouvement total, etc. La conception d'une action atomique en logique dynamique et dans les approches algébriques peut se comprendre de façon intuitive, mais il n'en demeure pas moins que cette notion est vague et mal définie. Où trace-t-on la ligne pour déterminer les blocs fondamentaux ? L'action *sit*, prise comme bloc de départ pour construire des actions atomiques, peut elle-même être conçue comme action atomique construite à partir d'autres générateurs d'actions plus précis.

Ces arguments seraient probablement rejetés par la communauté informatique puisque la réduction d'une action à une séquence de descriptions (une suite d'états du système) est essentielle à la programmation. Mais cela ne fait qu'exemplifier les différences entre les applications de la logique déontique. La logique déontique vise-t-elle l'analyse du discours ou la programmation ?

En informatique, la logique déontique a une visée *pratique*. On veut des résultats et des applications. On veut être en mesure de déterminer le comportement d'une banque de données lorsqu'elle est sujette à certaines contraintes (p. ex. Carmo et Jones 1996; Carmo *et al.* 2001), de formaliser les ententes contractuelles (p. ex. Boella et van der Torre 2004; Brown 2005; Prisacariu et Schneider 2012), de gérer le comportement et l'évolution d'un système informatique (p. ex. Boella et van der Torre 2003a), de gérer l'évolution des obligations sujettes à des échéances (p. ex. Dignum et Kuiper 1997, 1998; Dignum *et al.* 2005; Demolombe 2014), ou encore de construire des robots capables d'agir conformément à certaines normes (p. ex. Bringsjord *et al.* 2011; Bringsjord et Taylor 2012). Certains veulent même formaliser un système légal de façon à pouvoir déterminer sans équivoque le résultat normatif d'une situation lorsqu'on *input* certains faits avec certaines normes.

Le lecteur est invité à consulter Wieringa et Meyer (1993) pour un survol des applications de la logique déontique en informatique.

Pablo Castro et Tom Maibaum

L'approche de Castro et Maibaum (2009) présente une logique dynamique construite sur la base d'une logique de l'action booléenne. Soit $Prop$, un ensemble dénombrable de propositions atomiques, et $Act = \{a_1, \ldots, a_n\}$, un ensemble fini d'actions atomiques. Soit le langage suivant, avec les définitions usuelles pour les autres connecteurs propositionnels :

$$\mathcal{L} = \{(,), Prop, Act, 0, 1, \bar{}, \sqcap, \sqcup, =, \neg, \supset, [\,], P, P_w, \top, \bot\}$$

Les énoncés d'action bien formés sont définis récursivement :

$$\alpha := a_i \mid 0 \mid 1 \mid \overline{\alpha} \mid \alpha \sqcap \beta \mid \alpha \sqcup \beta$$

De même pour les énoncés bien formées :

$$\varphi := p_i \mid \bot \mid \top \mid \alpha = \beta \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi \mid P(\alpha) \mid P_w(\alpha)$$

L'obligation est définie de la manière suivante :

$$O(\alpha) =_{df} P(\alpha) \wedge \neg P_w(\overline{\alpha})$$

P_w représente la permission faible et P la permission forte. Une action est obligatoire lorsqu'elle est explicitement permise et que sa négation n'est pas faiblement permise (non interdite).

Le système axiomatique est construit à partir des axiomes du calcul propositionnel et d'une algèbre de Boole pour les actions. L'opérateur $[\,]$ est axiomatisé de manière usuelle par une modalité normale de type K, et les axiomes pour les autres opérateurs sont les suivants :

$$[0]\varphi \tag{A1}$$
$$[\alpha \sqcup \beta]\varphi \equiv [\alpha]\varphi \wedge [\beta]\varphi \tag{A2}$$
$$[\alpha]\varphi \supset [\alpha \sqcap \beta]\varphi \tag{A3}$$
$$P(0) \tag{A4}$$
$$P(\alpha \sqcup \beta) \equiv P(\alpha) \wedge P(\beta) \tag{A5}$$
$$P(\alpha) \vee P(\beta) \supset P(\alpha \sqcap \beta) \tag{A6}$$
$$\neg P_w(0) \tag{A7}$$
$$P_w(\alpha \sqcup \beta) \equiv P_w(\alpha) \vee P_w(\beta) \tag{A8}$$
$$P_w(\alpha \sqcap \beta) \supset P_w(\alpha) \wedge P_w(\beta) \tag{A9}$$
$$(P(\alpha) \wedge \alpha \neq 0) \supset P_w(\alpha) \tag{A10}$$

On trouve aussi des règles de substitution pour les égalités et des contraintes sur l'élément terminal.

(A1) stipule qu'après une action impossible, toute proposition est vraie et (A2) assure que lorsqu'une action disjointe mène à un état, alors chaque membre de l'action mène individuellement l'état. (A3) implique que si un état est le résultat d'une action, alors ce sera aussi le résultat de la combinaison de cette action avec une autre. (A4) stipule que l'action impossible est fortement permise, et (A7) indique qu'elle n'est pas faiblement permise (d'où l'impossibilité de l'accomplir, voir les remarques de Castro et Mailbaum sur ce point). (A5) indique qu'une disjonction d'action est fortement permise lorsque chaque action l'est individuellement et (A6) signifie que lorsque soit α soit β est fortement permise, alors leur conjonction l'est aussi. (A8) et (A9) signifient respectivement qu'une disjonction d'action est faiblement permise lorsque l'une ou l'autre des actions l'est et que si une conjonction est faiblement permise, alors chaque action prise individuellement l'est aussi. Finalement, (A10) stipule que si une action est fortement permise et qu'elle n'est pas impossible, alors elle est aussi faiblement permise.

Le lecteur est invité à consulter Castro et Maibaum (2009) pour la sémantique, construite dans les ensembles de manière usuelle à l'aide du complément, de l'intersection et de l'union pour les opérateurs, en prenant l'ensemble des scénarios où une proposition est vraie. Les auteurs ont repris leur approche de 2009 dans leur analyse de 2012.

Thierry Lucas

Thierry Lucas est un des rares logiciens à avoir appliqué la théorie des catégories (Mac Lane 1971) à la logique déontique.[4] La logique qu'il propose s'inspire des travaux de von Wright et est construite à partir du langage $\mathcal{L} = \{Prop, (,), \neg, \wedge, Ac(-,-,-), \Box\}$. $Prop$ est un ensemble dénombrable d'atomes propositionnels descriptifs et $\neg$ et $\wedge$ sont des connecteurs booléens (Lucas 2006).

Une action atomique $Ac(\varphi, \chi, \psi)$ est la description de trois états : l'état présent φ, l'état χ qui résulterait des actions d'un agent et l'état ψ qui en résulterait si l'agent n'agissait pas. Les énoncés dans la portée de $Ac(\varphi, \chi, \psi)$ sont uniquement propositionnels et appartiennent à l'ensemble défini récursivement de la manière suivante :

$$\varphi := p_i \mid \neg\varphi \mid \varphi \wedge \psi$$

[4] Voir aussi Peterson (2014a).

L'ensemble des énoncés bien formés est défini par :

$$\varphi := \ p_i \mid a \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi$$

Les autres connecteurs sont définis de manière usuelle (avec a une action atomique). Les axiomes pour l'action sont les suivants :

$$Ac(\varphi_1 \vee \varphi_2, \chi, \psi) \equiv Ac(\varphi_1, \chi, \psi) \vee Ac(\varphi_2, \chi, \psi) \quad \text{(A1)}$$
$$Ac(\varphi, \chi_1 \vee \chi_2, \psi) \equiv Ac(\varphi, \chi_1, \psi) \vee Ac(\varphi, \chi_2, \psi) \quad \text{(A2)}$$
$$Ac(\varphi, \chi, \psi_1 \vee \psi_2) \equiv Ac(\varphi, \chi, \psi_1) \vee Ac(\varphi, \chi, \psi_2) \quad \text{(A3)}$$
$$Ac(\varphi_1 \wedge \varphi_2, \chi_1 \wedge \chi_2, \psi_1 \wedge \psi_2) \equiv Ac(\varphi_1, \chi_1, \psi_1) \wedge Ac(\varphi_2, \chi_2, \psi_2) \quad \text{(A4)}$$
$$\varphi \equiv Ac(\varphi, \top, \top) \quad \text{(A5)}$$
$$\neg Ac(\varphi, \bot, \psi) \quad \text{(A6)}$$
$$\neg Ac(\varphi, \chi, \bot) \quad \text{(A7)}$$

L'opérateur $\Box$ est un opérateur de nécessité axiomatisé par KT et les actions respectent la règle suivante :

$$\frac{\varphi \equiv \varphi', \chi \equiv \chi', \psi \equiv \psi'}{Ac(\varphi, \chi, \psi) \equiv Ac(\varphi', \chi', \psi')}$$

L'axiome (A1) indique que si l'action est telle que i) la disjonction $\varphi_1 \vee \varphi_2$ est vraie pour l'état présent, ii) χ sera vrai si l'action est accomplie et iii) ψ sera vrai si elle ne l'est pas, alors soit $Ac(\varphi_1, \chi, \psi)$ est vrai soit $Ac(\varphi_2, \chi, \psi)$ l'est. (A2) indique que si l'action est telle que i) φ est vrai, ii) $\chi_1 \vee \chi_2$ serait vrai advenant que l'action soit accomplie, et iii) ψ le serait si elle ne l'était pas, alors soit $Ac(\varphi, \chi_1, \psi)$ est vrai soit $Ac(\varphi, \chi_2, \psi)$ l'est. Dans le même ordre d'idées, (A3) stipule que si l'action est telle que i) φ est vrai, ii) χ serait vrai si l'action est accomplie et iii) $\psi_1 \vee \psi_2$ serait vrai si elle ne l'était pas, alors soit $Ac(\varphi, \chi, \psi_1)$ est vrai soit $Ac(\varphi, \chi, \psi_2)$ l'est.

Le quatrième axiome permet d'éliminer les conjonctions à l'intérieur de la portée de Ac, et le cinquième indique qu'une description φ équivaut à l'action décrite par $Ac(\varphi, \top, \top)$. Finalement, (A6) indique qu'il est faux qu'une action telle que φ soit vrai, mais que $\bot$ serait réalisé si l'action était accomplie. Donc, il est impossible de faire une action contradictoire, au même titre que (A7) indique qu'il est faux qu'une action telle que $\bot$ soit réalisée si l'action n'était pas accomplie. La règle permet la substitution d'équivalences.

Lucas définit un modèle sémantique $\mathcal{M} = \langle U, Ag, R, M \rangle$.

$$a : U \longrightarrow U$$
$$n : U \longrightarrow U$$
$$R \subseteq U \times U$$
$$M : U \times Prop \longrightarrow \{0, 1\}$$

La fonction a représente l'action d'un agent, à savoir la transition d'un état à un autre, tandis que la fonction n représente les « actions » de la *nature*, c'est-à-dire la transition d'un état à un autre advenant l'absence d'actions de l'agent. La relation d'accessibilité R entre les scénarios est réflexive et M est une assignation de valeurs de vérité qui donne à chaque paire <scénario, proposition> une valeur de vérité. Les conditions de vérité pour les énoncés propositionnels sont définies usuellement et la condition de vérité pour les actions est la suivante :

$$\models_w Ac(\varphi, \chi, \psi) \Leftrightarrow \models_w \varphi \text{ et } \models_{a(w)} \chi \text{ et } \models_{n(w)} \psi$$

Autrement dit, l'action $Ac(\varphi, \chi, \psi)$ est vraie pour w si et seulement si φ est vrai pour w, χ est vrai pour les scénarios accessibles à w qui correspondent aux scénarios qui résultent des actions de l'agent, et ψ est vrai pour les scénarios accessibles à w qui correspondent aux actions qui résultent de la nature.

L'originalité de l'approche de Lucas consiste dans sa vision des actions comme *morphismes* dans le cadre de la théorie des catégories. En effet, il décrit les actions comme des flèches allant d'un ensemble de conditions initiales à un ensemble de résultats.

Lucas (2008) définit une algèbre d'action comme un triplet $\langle B, \mathcal{ACT}, C \rangle$ avec B et C, deux algèbres de Boole (pouvant être différentes l'une de l'autre) et $\mathcal{ACT} = \langle A, 0, 1, \neg, \wedge, \vee, C_0, \rangle$ un ensemble partiellement ordonné se comportant comme une logique bi-intuitionniste (Lucas 2006). La notion de co-support C_0 est introduite pour représenter les conditions dans lesquelles on obtient la négation d'une action. En supposant que $\mathcal{F}_1$ et $\mathcal{F}_2$ sont respectivement les ensembles d'énoncés bien formés de B et C, Lucas conçoit une action comme un morphisme $a : \Sigma \longrightarrow \mathcal{F}_2$, avec $\Sigma \subseteq \mathcal{F}_1$.

En ce sens, une action est vue comme un morphisme d'un ensemble de conditions à un ensemble de résultats. Les actions doivent satisfaire une condition de cohérence telle que $a(\sigma) = a(\sigma')$ pour $\sigma, \sigma' \in \Sigma$, c'est-à-dire que l'action a donne le même résultat pour toute sous-condition $\sigma \in \Sigma$. En d'autres termes, les sous-conditions sont cohérentes entres elles, c'est-à-dire que deux sous-conditions d'un ensemble Σ donnent les mêmes résultats.

Les actions sont préordonnées par $\leq$, 1 est l'action vide et 0 l'action nulle. L'action vide correspond à ne rien faire, alors qu'il est impossible d'accomplir l'action nulle.

Sur la base de $(A, 0, 1, \wedge, \vee)$, Lucas définit $(A, 0, 1, \wedge, \vee, \sim, \rightarrow)$ comme un système déductif intuitionniste avec $\rightarrow$ l'adjoint de $\wedge$ et $\sim$ la négation intuitionniste, définie par $\sim \alpha = \alpha \rightarrow 0$. Cela dit, il ajoute aussi un adjoint à la disjonction, en plus d'une négation v. L'adjoint à la disjonction est gouverné par (cl') et la négation est définie par :

$$v\alpha = 1 \backslash \alpha$$

$$\frac{\gamma \leq \alpha \vee \beta}{\gamma \backslash \beta \leq \alpha} \qquad \text{(cl')}$$

De fait, $\mathcal{ACT}$ est une logique bi-intuitionniste, c'est-à-dire une algèbre de bi-Heyting (un treillis distributif qui comprend une algèbre de Heyting et une algèbre de co-Heyting).

Lucas utilise une troisième négation pour dualiser les connecteurs, en introduisant les connecteurs duaux $(-)^*$:

$$\begin{aligned}
&\alpha \leq^* \beta \Leftrightarrow \neg\beta \leq \neg\alpha \\
&\alpha \leq \beta \Leftrightarrow \neg\beta \leq^* \neg\alpha \\
&\alpha \wedge^* \beta = \neg(\neg\alpha \vee \neg\beta) \\
&\alpha \vee^* \beta = \neg(\neg\alpha \wedge \neg\beta) \\
&1^* = \neg 0 \\
&0^* = \neg 1 \\
&\alpha \backslash^* \beta = \neg(\neg\beta \rightarrow \neg\alpha) \\
&v^*\alpha = \neg \sim \neg\alpha \\
&\alpha \rightarrow^* \beta = \neg(\neg\beta \backslash \neg\alpha) \\
&\sim^* \alpha = \neg v \neg\alpha
\end{aligned}$$

Ainsi, $\wedge$ et $\wedge^*$ se distribuent sur $\vee$. L'interprétation de ces connecteurs n'est cependant pas explicitée.

Outre les avancées qu'il a faites en 2006, et qu'il a reprises en 2007, Lucas (2008) propose d'utiliser sa logique de l'action pour construire une logique déontique. L'idée est de définir l'obligation de α comme la nécessité que la description des conditions entraîne les résultats. Autrement dit, une action est obligatoire lorsque les conditions entraînent nécessairement les résultats. Cette conception s'interprète de la même façon qu'en logique

modale, où la relation induite sur le modèle sémantique par l'axiomatisation de l'obligation permet d'isoler les scénarios où l'obligation est réalisée. En ce sens, il s'agit d'une sémantique idéale : on met de côté les scénarios où l'obligation n'est pas réalisée.

De même, chez Lucas, la conception de l'obligation est telle que la vérité des conditions entraîne nécessairement la vérité du résultat. Sa conception de l'obligation permet de dériver plusieurs conséquences du système standard, notamment l'agrégation, la distribution de O sur l'implication, (ROM), et le paradoxe de Ross.

Un aspect intéressant de son approche est qu'il adopte la distinction courante en logique dynamique entre la logique de l'action et la logique propositionnelle. En effet, son idée fondamentale est que la structure algébrique des actions est différente de la structure algébrique des propositions utilisée pour parler d'elles.

* * *

Ce livre visait à offrir au lecteur une revue concise et précise de la littérature scientifique sur la logique déontique, afin qu'il s'oriente rapidement dans ses recherches. Vu l'étendue de cette littérature, il est évidemment impossible d'en couvrir l'entièreté, et c'est pourquoi ce livre n'aborde pas toutes les subtilités et les distinctions inhérentes aux approches que nous avons décrites, dans le meilleur des cas, très brièvement. Nous espérons néanmoins qu'il saura donner à notre lecteur les repères nécessaires pour pouvoir naviguer dans ce monde complexe ainsi que le goût de continuer le voyage.

Bibliographie

Al-Hibri, A. (1978). *Deontic logic : A comprehensive appraisal and a new proposal*. University Press of America.

Alchourrón, C. (1990). Logic without truth. *Ratio Juris*, 3(1):46–67.

Alchourrón, C. (1996). Detachment and defeasibility in deontic logic. *Studia Logica*, 57(1):5–18.

Alchourrón, C. et Bulygin, E. (1971). *Normative systems*. Springer.

Alchourrón, C. et Bulygin, E. (1981). The expressive conception of norms. *In* Hilpinen, R., éditeur : *New Studies in Deontic Logic*, pages 95–124. D. Reidel Publishing Company.

Alchourrón, C. et Makinson, D. (1981). Hierarchies of regulations and their logic. *In* Hilpinen, R., éditeur : *New Studies in Deontic Logic*, pages 125–148. D. Reidel Publishing Company.

Anderson, A. R. (1958a). A reduction of deontic logic to alethic modal logic. *Mind*, 67(265):100–103.

Anderson, A. R. (1958b). The logic of norms. *Logique et Analyse*, 1(2):84–91.

Anderson, A. R. (1959). On the logic of commitment. *Philosophical Studies*, 10(2):23–27.

Anderson, A. R. (1962). Reply to Mr Rescher's conditional permission in deontic logic. *Philosophical Studies*, 13(1):6–8.

Anderson, A. R. (1967). Some nasty problems in the formal logic of ethics. *Noûs*, 1(4):345–360.

Anderson, B. (1999). A comment on Walter's response to Jørgensen's dilemma : Common sense and scientific attitudes. *Ratio Juris*, 12(1):100–107.

Anglberger, A. J. J. (2008). Dynamic deontic logic and its paradoxes. *Studia Logica*, 89(3):427–435.

Anglberger, A. J. J. (2009). Alternative reductions for dynamic deontic logics. *In* Hieke, A. et Leitgeb, H., éditeurs : *Reduction-Abstraction-Analysis*, volume 11, pages 179–191. Ontos Verlag.

Anglberger, A. J. J., Dong, H. et Roy, O. (2014). Open reading without free choice. *In* Cariani, F., Grossi, D., Meheus, J. et Parent, X., éditeurs : *Deontic Logic and Normative Systems*, volume 8554 de *Lecture Notes in Computer Science*, pages 19–32. Springer.

Anglberger, A. J. J., Gratzl, N. et Roy, O. (2015). Obligation, free choice and the logic of weakest permissions. *The Review of Symbolic Logic*, 8(4):807–827.

Antonelli, G. A. (2012). Non-monotonic logic. *In* Zalta, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*.

ÅQVIST, L. (1967). Good Samaritans, contrary-to-duty imperatives, and epistemic obligations. *Noûs*, 1(4):361–379.

ÅQVIST, L. (1986). Some results on dyadic logic and the logic of preference. *Synthese*, 66:95–110.

ÅQVIST, L. (2002). Deontic logic. *In* GABBAY, D. M. et GUENTHNER, F., éditeurs : *Handbook of Philosophical Logic*, volume 8, pages 147–264. Kluwer Academic Publishers, 2^e édition.

ÅQVIST, L. et HOEPELMAN, J. (1981). Some theorems about a "tree" system of deontic tense logic. *In* HILPINEN, R., éditeur : *New Studies in Deontic Logic*, pages 187–221. D. Reidel Publishing Company.

ARTHUR, R. T. W. (2011). *Natural Deduction : An introduction to logic with real arguments, a little history and some humour*. Broadview Press.

BAILHACHE, P. (1991). *Essai de logique déontique*. Vrin.

BAILHACHE, P. (1993). The deontic branching time : Two related conceptions. *Logique et Analyse*, 36(141-142):159–175.

BAUDOIN, J.-L. (2010). *Droit et vérité*. Éditions Thémis.

BELNAP, N. et PERLOFF, M. (1988). Seeing to it that : A canonical form for agentives. *Theoria*, 54(3):175–199.

BOELLA, G. et van der TORRE, L. (2003a). Local policies for the control of virtual communities. *In Proceedings of IEEE/WIC web intelligence conference*, pages 161–167. IEEE Press.

BOELLA, G. et van der TORRE, L. (2003b). Permissions and obligations in hierarchical normative systems. *In Proceedings of the 9th International Conference on Artificial Intelligence and Law*, ICAIL '03, pages 109–118. ACM.

BOELLA, G. et van der TORRE, L. (2004). Contracts as legal institutions in organizations of autonomous agents. *In 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, pages 948–955. IEEE Computer Society.

BOELLA, G. et van der TORRE, L. (2006a). An architecture of a normative system : Counts-as conditionals, obligations and permissions. *In* NAKASHIMA, H., WELLMAN, M. P., WEISS, G. et STONE, P., éditeurs : *5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*, pages 229–231. ACM.

BOELLA, G. et van der TORRE, L. (2006b). Delegation of power in normative multiagent systems. *In* GOBLE, L. et MEYER, J.-J. C., éditeurs : *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 36–52. Springer.

BOELLA, G. et van der TORRE, L. (2006c). A logical architecture of a normative system. *In* GOBLE, L. et MEYER, J.-J. C., éditeurs : *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 24–35. Springer.

BONEVAC, D. (1983). Chellas on conditional obligation. *Philosophical Studies*, 44:247–256.

BONEVAC, D. (1998). Against conditional obligation. *Noûs*, 32(1):37–53.

BRINGSJORD, S. et TAYLOR, J. (2012). The divine-command approach to robot ethics. *In* LIN, P., ABNEY, G. et ABNEY, K., éditeurs : *Robot Ethics : The Ethical and Social Implications of Robotics*, pages 85–108. MIT Press.

BRINGSJORD, S., TAYLOR, J., HOUSTEN, T., van HEUVELN, B., ARKOUDAS, K. et CLARK, M. (2011). Robotethics via category theory : Moving beyond mere formal operations to engineer robots whose decisions are guaranteed to be ethically correct. *In* ANDERSON, M. et ANDERSON, S., éditeurs : *Machine Ethics*, pages 361–374. Cambridge University Press.

Broersen, J. (2003). *Modal action logics for reasoning about reactive systems*. Thèse de doctorat, Vrije Universiteit Amsterdam.

Broersen, J. (2004). Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, 2(1):153–168.

Broersen, J. (2008). A logical analysis of the interaction between 'obligation-to-do' and 'knowingly doing'. *In* van der Meyden, R. et van der Torre, L., éditeurs : *DEON 2008*, volume 5076 de *Lecture Notes in Computer Science*, pages 140–154. Springer.

Broersen, J. (2009). A complete stit logic for knowledge and action, and some of its applications. *In* Baldoni, M., Son, T. C., van Riemsdijk, M. et Winikoff, M., éditeurs : *Declarative Agent Languages and Technologies VI*, volume 5397 de *Lecture Notes in Computer Science*, pages 47–59.

Broersen, J. (2011a). Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of Applied Logic*, 9(2):137–152.

Broersen, J. (2011b). Making a start with the stit logic analysis of intentional action. *Journal of Philosophical Logic*, 40(4):499–530.

Broersen, J., Dignum, F., Dignum, V. et Meyer, J.-J. C. (2004). Desinging a deontic logic of deadlines. *In* Lomuscio, A. et Nute, D., éditeurs : *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 43–56. Springer.

Brown, M. (2005). Obligation, contracts and negociation : Outlining an approach. *Journal of Applied Logic*, 3(3-4):371–395.

Carmo, J., Demolombe, R. et Jones, A. J. I. (2001). An application of deontic logic to information system constraints. *Fundamenta Informaticae*, 48(2-3):165–181.

Carmo, J. et Jones, A. J. I. (1996). Deontic database constraints, violation and recovery. *Studia Logica*, 57(1):139–165.

Carmo, J. et Jones, A. J. I. (2002). Deontic logic and contrary-to-duties. *In* Gabbay, D. M. et Guenthner, F., éditeurs : *Handbook of Philosophical Logic*, volume 8, pages 265–343. Kluwer Academic Publishers, 2^e édition.

Carmo, J. et Jones, A. J. I. (2013). Completeness and decidability results for a logic of contrary-to-duty conditionals. *Journal of Logic and Computation*, 23(3):585–626.

Carmo, J. et Pacheco, O. (2000). Deontic and action logics for collective agency and roles. *In* Demolombe, R. et Hilpinen, R., éditeurs : *Proceedings of the Fifth International Workshop on Deontic Logic in Computer Science (DEON'00)*, pages 93–124.

Carmo, J. et Pacheco, O. (2001). Deontic and action logics for organized collective agency, modeled through institutionalized agents and roles. *Fundamenta Informaticae*, 48(2-3):129–163.

Castañeda, H.-N. (1959). The logic of obligation. *Philosophical Studies*, 10(2):17–23.

Castañeda, H.-N. (1968). Acts, the logic of obligation, and deontic calculi. *Philosophical Studies*, 19(1):13–26.

Castañeda, H.-N. (1970). On the semantics of the ought-to-do. *Synthese*, 21(3):449–468.

Castañeda, H.-N. (1977). Ought, time and the deontic paradoxes. *The Journal of Philosophy*, 74(12):775–791.

Castañeda, H.-N. (1981). The paradoxes of deontic logic. *In* Hilpinen, R., éditeur : *New Studies in Deontic Logic*, pages 37–85. D. Reidel Publishing Company.

Castañeda, H.-N. (1983). Obligation and conditionals. *In* Tomberlin, J., éditeur : *Agent, Language and the Structure of the World*, pages 441–448. Hackett.

Castro, P. F. et Maibaum, T. S. E. (2009). Deontic action logic, atomic Boolean algebras and fault-tolerance. *Journal of Applied Logic*, 7(4):441–466.

Castro, P. F. et Maibaum, T. S. E. (2012). Towards a first-order deontic action lo-

gic. *In* MOSSAKOWSKI, T. et KREOWSKI, H.-J., éditeurs : *Recent Trends in Algebraic Development Techniques*, volume 7137 de *Lecture Notes in Computer Science*, pages 61–75. Springer.

CHELLAS, B. F. (1969). *The logical form of imperatives.* Perry Lane Press.

CHELLAS, B. F. (1974). Conditional obligation. *In* STENLUND, S., éditeur : *Logical Theory and Semantic Analysis*, pages 23–33. D. Reidel Publishing Company.

CHELLAS, B. F. (1980). *Modal logic : An introduction.* Cambridge University Press.

CHISHOLM, R. (1963). Contrary-to-duty imperatives and deontic logic. *Analysis*, 24(2): 33–36.

CSIRMAZ, L. (1990). Multi-level permission. Rapport technique 90-25, Rutgers University.

CÔTÉ, P.-A. (2006). *Interprétation des lois.* Éditions Thémis, 3^e édition.

DANIELSSON, S. (1968). *Preference and obligation : Studies in the logic of ethics.* University of Uppsala.

DECEW, J. (1981). Conditional obligation and counterfactuals. *Journal of Philosophical Logic*, 10(1):55–72.

DELLUNDE, P. (2008). On the multimodal logic of normative systems. *In* SICHMAN, J. S., PADGET, J. A., OSSOWSKI, S. et NORIEGA, P., éditeurs : *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, pages 261–274. Springer.

DEMOLOMBE, R. (2014). Obligations with deadlines : A formalization in dynamic deontic logic. *Journal of Logic and Computation*, 24(1):1–17.

DIGNUM, F., BROERSEN, J., DIGNUM, V. et MEYER, J.-J. C. (2005). Meeting the deadline : Why, when and how. *In* HINCHEY, M. G., RASH, J. L., TRUSZKOWSKI, W. F. et ROUFF, C. A., éditeurs : *Formal Approaches to Agent-Based Systems*, volume 3228 de *Lecture Notes in Computer Science*, pages 30–40. Springer.

DIGNUM, F. et KUIPER, R. (1997). Combining dynamic deontic logic and temporal logic for the specification of deadlines. *In System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference*, volume 5, pages 336–346. IEEE.

DIGNUM, F. et KUIPER, R. (1998). Obligations and dense time for specifying deadlines. *In System Sciences, 1998, Proceedings of the Thirty-First Hawaii International Conference*, volume 5, pages 186–195. IEEE.

DIGNUM, F., MEYER, J.-J. C., WIERINGA, R. et KUIPER, R. (1996). A modal approach to intentions, commitments and obligations : Intention plus commitment yields obligation. *In* BROWN, M. et CARMO, J., éditeurs : *Deontic Logic, Agency and Normative Systems*, pages 80–97. Springer.

FELDMAN, F. (1986). *Doing the best we can : An essay in informal deontic logic.* D. Reidel Publishing Company.

FØLLESDAL, D. et HILPINEN, R. (1970). Deontic logic : An introduction. *In* HILPINEN, R., éditeur : *Deontic Logic : Introductory and Systematic Readings*, pages 1–35. D. Reidel Publishing Company.

FORRESTER, J. (1984). Gentle murder, or the adverbial Samaritan. *The Journal of Philosophy*, 81(4):193–197.

GABBAY, D. M. (1994). *What is a Logical System ?* Clarendon Press.

GABBAY, D. M. et SCHLECHTA, K. (2010). Critical analysis of the Carmo-Jones system of contrary-to-duty obligations. *ArXiv e-prints*, 1002.3021.

GARSON, J. (2006). *Modal logic for philosophers.* Cambridge University Press.

GENTZEN, G. (1934). Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210 ;405–431.

GOBLE, L. (1991). Murder most gentle : The paradox deepens. *Philosophical Studies*, 64(2):217–227.

GOBLE, L. (2004). A proposal for dealing with deontic dilemmas. *In* LOMUSCIO, A. et NUTE, D., éditeurs : *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 74–113. Springer.

GOBLE, L. (2009). Normative conflicts and the logic of 'ought'. *Noûs*, 43(3):450–489.

GOLDBLATT, R. (2006). *Topoi : The categorical analysis of logic*. Dover Publications.

HANSEN, J. (1999). Paradoxes of commitment. *In* MEGGLE, G., éditeur : *Actions, Norms, Values*, pages 255–263. Walter de Gruyter.

HANSSON, B. (1969). An analysis of some deontic logics. *Noûs*, 3(4):373–398.

HANSSON, S. (1990). Preference-based deontic logic (PDL). *Journal of Philosophical Logic*, 19(1):75–93.

HANSSON, S. (1997). Situationist deontic logic. *Journal of Philosophical Logic*, 26(4):423–448.

HANSSON, S. (2006). Ideal worlds – wishful thinking in deontic logic. *Studia Logica*, 82(3):329.

HAREL, D., KOZEN, D. et TIURYN, J. (2000). *Dynamic Logic*. MIT Press.

HARZHEIM, E. (2005). *Ordered Sets*, volume 7 de *Advances in Mathematics*. Springer.

HEMPEL, C. G. (1972). *Éléments d'épistémologie*. Armand Colin.

HENDRICKS, V. (2006). *Mainstream and formal epistemology*. Cambridge University Press.

HERZIG, A., LORINI, E., MOISAN, F. et TROQUARD, N. (2011). A dynamic logic of normative systems. *In* WALSH, T., éditeur : *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pages 228–233. AAAI Press.

HINTIKKA, J. (1970). Some main problems of deontic logic. *In* HILPINEN, R., éditeur : *Deontic Logic : Introductory and Systematic Readings*, pages 59–104. D. Reidel Publishing Company.

HOHFELD, W. N. (1917). Fundamental legal conceptions as applied to judicial reasoning. *The Yale Law Journal*, 26(8):710–770.

HORTY, J. (1994). Moral dilemmas and non-monotonic logic. *Journal of Philosophical Logic*, 23(1):35–65.

HORTY, J. (1997). Non-monotonic foundations for deontic logic. *In* NUTE, D., éditeur : *Defeasible Deontic Logic*, pages 17–44. Kluwer Academic Publishers.

HORTY, J. (2001). *Agency and deontic logic*. Oxford University Press.

HORTY, J. et BELNAP, N. (1995). The deliberative stit : A study of action, omission, ability and obligation. *Journal of Philosophical Logic*, 24(6):583–644.

HUGHES, J. et ROYAKKERS, L. (2008). Don't ever do that ! : Long-term duties in PD_eL. *Studia Logica*, 89(1):59–79.

HUME, D. (1993). *Traité de la nature humaine [1740]*, volume 3. Flammarion.

JACQUETTE, D. (1986). Forrester's paradox. *Dialogue : Canadian Philosophical Review/Revue Canadienne de Philosophie*, 25(4):761–763.

JONES, A. J. I. (1991). On the logic of deontic conditionals. *Ratio Juris*, 4(3):355–366.

JONES, A. J. I. et PÖRN, I. (1985). Ideality, sub-ideality and deontic logic. *Synthese*, 65(2):275–290.

JONES, A. J. I. et PÖRN, I. (1986). 'Ought' and 'must'. *Synthese*, 66(1):89–93.

JONES, A. J. I. et SERGOT, M. (1996). A formal characterisation of institutionalised

power. *Logic Journal of the IGPL*, 4(3):427–443.

Jørgensen, J. (1937). Imperatives and logic. *Erkenntnis*, 7(1):288–296.

Kanger, S. (1957). *New foundations for ethical theory*. Stockholm.

Kanger, S. (1971). New foundations for ethical theory. *In* Hilpinen, R., éditeur : *Deontic Logic : Introductory and Systematic Readings*, pages 36–58. D. Reidel Publishing Company.

Kanger, S. (1972). Law and logic. *Theoria*, 38(3):105–132.

Kanger, S. et Kanger, H. (1966). Rights and parliamentarism. *Theoria*, 32(2):85–115.

Kleene, S. (1952). *Introduction to metamathematics*. ISHI Press International.

Knuuttila, S. (1981). The emergence of deontic logic in the fourteenth century. *In* Hilpinen, R., éditeur : *New Studies in Deontic Logic*, pages 225–248. D. Reidel Publishing Company.

Kooi, B. P. et Tamminga, A. M. (2006). Conflicting obligations in multi-agent deontic logic. *In* Goble, L. et Meyer, J.-J. C., éditeurs : *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 175–186. Springer.

Kripke, S. (1959). A completeness theorem in modal logic. *The Journal of symbolic logic*, 24:1–14.

Kripke, S. (1963). Semantical analysis of modal logic I : Normal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9:67–96.

Lemmon, E. J. (1962). Moral dilemmas. *The Philosophical Review*, 71(2):139–158.

Lemmon, E. J. (1965). Deontic logic and the logic of imperatives. *Logique et Analyse*, 8:39–71.

Lepage, F. (2010). *Éléments de logique contemporaine*. Presses de l'Université de Montréal.

Lindahl, L. et Odelstad, J. (2000). An algebraic analysis of normative systems. *Ratio Juris*, 13(3):261–278.

Lindahl, L. et Odelstad, J. (2002). The role of connections as minimal norms in normative systems. *In* Bench-Capon, T., Daskalopulu, A. et Winkels, R. G. F., éditeurs : *Legal Knowledge and Information Systems. Jurix 2002 : The Fifteenth Annual Conference*, pages 31–40. IOS Press.

Lindahl, L. et Odelstad, J. (2003). Normative systems and their revision : An algebraic approach. *Artificial Intelligence and Law*, 11(2-3):81–104.

Lindahl, L. et Odelstad, J. (2004). Normative positions within an algebraic approach to normative systems. *Journal of Applied Logic*, 2(1):63–91.

Lindahl, L. et Odelstad, J. (2006). Intermediate concepts in normative systems. *In* Goble, L. et Meyer, J.-J. C., éditeurs : *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 187–200. Springer.

Lindahl, L. et Odelstad, J. (2008a). Intermediaries and intervenients in normative systems. *Journal of Applied Logic*, 6(2):229–250.

Lindahl, L. et Odelstad, J. (2008b). Strata of intervenient concepts in normative systems. *In* van der Meyden, R. et van der Torre, L., éditeurs : *DEON 2008*, volume 5076 de *Lecture Notes in Computer Science*, pages 203–217. Springer.

Lindahl, L. et Odelstad, J. (2011). Stratification of normative systems with intermediaries. *Journal of Applied Logic*, 9(2):113–136.

Loewer, B. et Belzer, M. (1983). Dyadic deontic detachment. *Synthese*, 54(2):295–318.

Lucas, T. (2006). Von Wright's action revisited : Actions as morphisms. *Logique et Analyse*, 49(193):85–115.

Lucas, T. (2007). Axioms for action. *Logique et Analyse*, 50(200):103–123.

Lucas, T. (2008). Deontic algebras of actions. *Logique et Analyse*, 51(202):367–389.

Mac Lane, S. (1971). *Categories for the working mathematician*. Springer, 2^e édition.

Makinson, D. et van der Torre, L. (2000). Input/output logics. *Journal of Philosophical Logic*, 29(4):383–408.

Makinson, D. et van der Torre, L. (2001). Constraints for input/output logics. *Journal of Philosophical Logic*, 30(2):155–185.

Makinson, D. et van der Torre, L. (2003a). Permission from an input/output perspective. *Journal of Philosophical Logic*, 32(4):391–416.

Makinson, D. et van der Torre, L. (2003b). What is input/output logic ? *In Foundations of the Formal Sciences II : Applications of Mathematical Logic in Philosophy and Linguistics*, volume 17 de *Trends in Logic Series*, pages 163–174. Kluwer Academic Publishers.

Mally, E. (1926). Grundgesetze des Sollens. Elemente der Logik des Willens. *In* Wolf, K. et Weingartner, P., éditeurs : *Logische Schriften : Großes Logikfragment, Grundgesetze des Sollens*, page 227–324. D. Reidel Publishing Company.

McKinney, A. M. (1977). *Conditional obligation and temporally dependent necessity*. Thèse de doctorat, University of Pennsylvania.

McLaughlin, R. N. (1955). Further problems of derived obligations. *Mind*, 64:400–402.

McNamara, P. (2010). Deontic logic. *In* Zalta, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*.

Mendelson, E. (2009). *Introduction to mathematical logic*. CRC Press, 5^e édition.

Meyer, J.-J. C. (1987). A simple solution to the "deepest" paradox in deontic logic. *Logique et Analyse*, 30(117-118):81–90.

Meyer, J.-J. C. (1988). A different approach to deontic logic : Deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136.

Moore, G. E. (1971). *Principia Ethica*. Cambridge University Press, London.

Mott, P. (1973). On Chisholm's paradox. *Journal of Philosophical Logic*, 2(2):197–211.

Niles, I. (1997). Rescuing the counterfactual solution to Chisholm's paradox. *Philosophia*, 25(1-4):351–371.

Nowell-Smith, P. H. (1960). Escapism : The logical basis of ethics. *Mind*, 69:289–300.

Nozick, R. et Routley, R. (1962). Escaping the good Samaritan paradox. *Mind*, 71(283):377–382.

Nute, D., éditeur (1997). *Defeasible deontic logic*. Kluwer Academic Publishers.

Nute, D. et Yu, X. (1997). Introduction. *In* Nute, D., éditeur : *Defeasible Deontic Logic*, pages 1–16. Kluwer Academic Publishers.

Odelstad, J. et Lindahl, L. (2000). Normative systems represented by Boolean quasi-orderings. *Nordic Journal of Philosophical Logic*, 5(2):161–174.

Opałek, K. et Woleński, J. (1991). Normative systems, permission and deontic logic. *Ratio Juris*, 4(3):334–348.

Pacheco, O. et Carmo, J. (2003). A role based model for the normative specification of organized collective agency and agents interaction. *Autonomous Agents and Multi-Agent Systems*, 6(2):145–184.

Pacheco, O. et Santos, F. (2004). Delegation in a role-based organization. *In* Lomuscio, A. et Nute, D., éditeurs : *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 209–227. Springer.

PARENT, H. (2007). *Traité de droit criminel, Tome II - La culpabilité (actus reus et mens rea)*. Éditions Thémis.

PARENT, H. (2008). *Traité de droit criminel, Tome I - L'imputabilité*. Éditions Thémis.

PAULY, M. (2002). A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166.

PETERSON, C. (2011). *La logique déontique : Une application de la logique à l'éthique et au discours juridique*. Mémoire de maîtrise, Université de Montréal.

PETERSON, C. (2013a). Normative inferences and validity : A heuristic. *Cogency*, 5(1):85–105.

PETERSON, C. (2013b). *Pensée rationnelle et argumentation*. Presses de l'Université de Montréal.

PETERSON, C. (2014a). The categorical imperative : Category theory as a foundation for deontic logic. *Journal of Applied Logic*, 12(4):417–461.

PETERSON, C. (2014b). Monoidal logics : How to avoid paradoxes. *In* LIETO, A., RADICIONI, D. P. et CRUCIANI, M., éditeurs : *Proceedings of the International Workshop on Artificial Intelligence and Cognition (AIC 2014)*, volume 1315, pages 122–133. CEUR Workshop Proceedings.

PETERSON, C. (2015a). Conditional reasoning with string diagrams. *South American Journal of Logic*, 1(2):401–434.

PETERSON, C. (2015b). Contrary-to-duty reasoning : A categorical approach. *Logica Universalis*, 9(1):47–92.

PETERSON, C. (2015c). A logic for human actions. *In* PAYETTE, G. et URBANIAK, R., éditeurs : *Applications of Formal Philosophy*. Springer. Sous presse.

PETERSON, C. et MARQUIS, J.-P. (2012). A note on Forrester's paradox. *Polish Journal of Philosophy*, VI(2):53–70.

POINCARÉ, H. (1913). *Dernières pensées*. Kessinger Publishing.

PRAKKEN, H. et SERGOT, M. (1996). Contrary-to-duty obligations. *Studia Logica*, 57(1): 91–115.

PRATT, V. (1976). Semantical considerations of Floyd-Hoare Logic. Rapport technique MIT/LCS/TR-168.

PRATT, V. (1980). Application of modal logic to programming. *Studia Logica*, 39(2-3):257–274.

PRATT, V. (1991). Action logic and pure induction. *In* EIJCK, J., éditeur : *Logics in AI*, volume 478 de *Lecture Notes in Computer Science*, pages 97–120. Springer.

PRIOR, A. (1954). The paradoxes of derived obligation. *Mind*, 63(249):64–65.

PRIOR, A. (1958). Escapism : The logical basis of ethics. *In* MELDEN, A. I., éditeur : *Essays in Moral Philosophy*, pages 135–146. University of Washington Press.

PRIOR, A. (1967). *Past, present and future*. Oxford University Press.

PRISACARIU, C. et SCHNEIDER, G. (2012). A dynamic deontic logic for complex contracts. *Journal of Logic and Algebraic Programming*, 81(4):458–490.

PÖRN, I. (1970). *The logic of power*. Basil Blackwell.

REITER, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13:81–132.

RESCHER, N. (1958). An axiom system for deontic logic. *Philosophical Studies*, 9(1):24–30.

RESCHER, N. (1962). Conditional permission in deontic logic. *Philosophical Studies*, 13(1):1–6.

RESCHER, N. (1967). Semantic foundations for conditional permission. *Philosophical Studies*, 18(4):56–61.

ROBINSON, R. (1986). Reply to Jacquette's Forrester's paradox. *Dialogue : Canadian Philosophical Review/Revue Canadienne de Philosophie*, 25:765–768.

ROBISON, J. (1967). Further difficulties for conditional permission in deontic logic. *Philosophical Studies*, 18(1):27–30.

ROSS, A. (1941). Imperatives and logic. *Theoria*, 7(1):53–71.

ROSS, A. (1944). Imperatives and logic. *Philosophy of Science*, 11(1):30–46.

ROYAKKERS, L. (1998). *Extending deontic logic for the formalisation of legal rules*. Kluwer Academic Publishers.

ROYAKKERS, L. (2000). Combining deontic and action logics for collective agency. *In* BREUKER, J., LEENES, R. et WINKELS, R., éditeurs : *Legal Knowledge and Information Systems. Jurix 2000 : The Thirteenth Annual Conference*, pages 135–146. IOS Press.

ROYAKKERS, L. et DIGNUM, F. (1997). Defeasible reasoning with legal rules. *In* NUTE, D., éditeur : *Defeasible Deontic Logic*, pages 263–286. Kluwer Academic Publishers.

RYU, Y. U. et LEE, R. M. (1997). Deontic logic viewed as defeasible reasoning. *In* NUTE, D., éditeur : *Defeasible Deontic Logic*, pages 123–137. Kluwer Academic Publishers.

SAHLQVIST, H. (1975). Completeness and correspondence in first and second order semantics for modal logic. *In* KANGER, S., éditeur : *Proceedings of the Thrid Scandanavian Logic Symposium*, pages 110–143.

SCHOTCH, P. et JENNINGS, R. E. (1981). Non-Kripkean deontic logic. *In* HILPINEN, R., éditeur : *New Studies in Deontic Logic*, pages 149–162. D. Reidel Publishing Company.

SEGERBERG, K. (1982a). A completeness theorem in the modal logic of programs. *In* TRACZYK, T., éditeur : *Universal Algebra and Applications*, volume 9, pages 31–46. Polish Scientific Publishers.

SEGERBERG, K. (1982b). A deontic logic of action. *Studia Logica*, 41(2):269–282.

SEGERBERG, K. (1992). Getting started : Beginnings in the logic of action. *Studia Logica*, 51(3):347–378.

SEGERBERG, K. (2007). A blueprint for deontic logic in three (not necessarily easy) steps. *In* BONANNO, G., DELGRANDE, J. P., LANG, J. et ROTT, H., éditeurs : *Formal Models of Belief Change in Rational Agents*. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI).

SEGERBERG, K. (2009). Blueprint for a dynamic deontic logic. *Journal of Applied Logic*, 7(4):388–402.

SEGERBERG, K. (2012). DΔL : A dynamic deontic logic. *Synthese*, 185(1):1–17.

SEGERBERG, K., MEYER, J.-J. et KRACHT, M. (2009). The logic of action. *In* ZALTA, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*.

SELLARS, W. (1967). Reflections on contrary-to-duty imperatives. *Noûs*, 1:303–344.

SINNOTT-ARMSTRONG, W. (1985). A solution to Forrester's paradox of gentle murder. *The Journal of Philosophy*, 82:162–168.

STEWART, I. (1997). Facing Walter's dilemma. *Ratio Juris*, 10(4):397–402.

STRASSER, C. (2011). A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80.

STRASSER, C. et BEIRLAEN, M. (2012). An andersonian deontic logic with contextualized sanctions. *In* ÅGOTNES, T., BROERSEN, J. et ELGESEM, D., éditeurs : *Deontic Logic in Computer Science*, volume 7393 de *Lecture Notes in Computer Science*, pages 151–169. Springer.

TARSKI, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):341–376.

THOMASON, R. (1970). Indeterminist time and truth-value gaps. *Theoria*, 36(3):264–281.

THOMASON, R. (1981). Deontic logic as founded on tense logic. *In* HILPINEN, R., éditeur : *New Studies in Deontic Logic*, pages 165–176. D. Reidel Publishing Company.

THOMASON, R. (2002). Combinations of tense and modality. *In* GABBAY, D. M. et GUENTHNER, F., éditeurs : *Handbook of Philosophical Logic*, volume 7, pages 205–234. Kluwer Academic Publishers, 2^e édition.

TIMMONS, M. (1999). *Morality without foundations : A defense of ethical contextualism*. Oxford University Press.

TOMASSI, P. (1999). *Logic*. Routledge.

TOMBERLIN, J. (1981). Contrary-to-duty imperatives and conditional obligations. *Noûs*, 15(3):357–375.

TOMBERLIN, J. (1983). Contrary-to-duty imperatives and Castañeda's system of deontic logic. *In* TOMBERLIN, J., éditeur : *Agent, Language and the Structure of the World*, pages 231–249. Hackett.

TRYPUZ, R. et KULICKI, P. (2009). A systematics of deontic action logics based on Boolean algebra. *Logic and Logical Philosophy*, 18:253–270.

TRYPUZ, R. et KULICKI, P. (2010). Towards metalogical systematisation of deontic action logics based on Boolean algebra. *In* GOVERNATORI, G. et SARTOR, G., éditeurs : *DEON 2010*, volume 6181 de *Lecture Notes in Computer Science*, pages 132–147. Springer.

van der MEYDEN, R. (1991). A clausal logic for deontic action specification. *In* SARASWAT, V. et UEDA, K., éditeurs : *Logic Programming, Proceedings of the 1991 International Symposium*, pages 221–238. MIT Press.

van der MEYDEN, R. (1996). The dynamic logic of permission. *Journal of Logic and Computation*, 6(3):465–479.

van der TORRE, L. et TAN, Y.-H. (1997). The many faces of defeasibility in defeasible deontic logic. *In* NUTE, D., éditeur : *Defeasible Deontic Logic*, pages 79–121. Kluwer Academic Publishers.

van der TORRE, L. et TAN, Y.-H. (1999). Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, 27(1-4):49–78.

van ECK, J. (1982a). A system of temporally relative modal and deontic predicate logic and it's philosophical applications. *Logique et Analyse*, 25(99):249–290.

van ECK, J. (1982b). A system of temporally relative modal and deontic predicate logic and it's philosophical applications (2). *Logique et Analyse*, 25(100):339–381.

van ECK, J. (1986). In defense of temprally relative deontic logic : A reply to Hector-Neri Castañeda's The paradoxes of deontic logic. *Logique et Analyse*, 29(115):335–348.

van FRAASSEN, B. C. (1972). The logic of conditional obligation. *Journal of Philosophical Logic*, 1(3):417–438.

van FRAASSEN, B. C. (1973). Values and the heart's command. *The Journal of Philosophy*, 70(1):5–19.

VOLPE, G. (1999). A minimalist solution to Jørgensen's dilemma. *Ratio Juris*, 12(1):59–79.

von WRIGHT, G. H. (1951). Deontic logic. *Mind*, 60(237):1–15.

von WRIGHT, G. H. (1956). A note on deontic logic and derived obligation. *Mind*, 65(260):507–509.

von WRIGHT, G. H. (1963). Practical inference. *The Philosophical Review*, 72:159–179.

von WRIGHT, G. H. (1967). Deontic logics. *American Philosophical Quarterly*, 4(2):136–143.

von WRIGHT, G. H. (1983). Norms of higher order. *Studia Logica*, 42(2):119–127.

von WRIGHT, G. H. (1999). Ought to be–ought to do. *In* MEGGLE, G., éditeur : *Actions, Norms, Values*, pages 3–10. Walter de Gruyter.

VOROBEJ, M. (1986). Conditional obligation and detachement. *Canadian Journal of Philosophy*, 16(1):11–26.

WALTER, R. (1996). Jørgensen's dilemma and how to face it. *Ratio Juris*, 9(2):168–171.

WEINBERGER, O. (1999). Against the ontologization of logic : A critical comment on Robert Walter's tackling Jørgensen's dilemma. *Ratio Juris*, 12(1):96–99.

WIERINGA, R. et MEYER, J.-J. C. (1993). Applications of deontic logic in computer science : A concise overview. *In Deontic Logic in Computer Science : Normative System Specification*, pages 17–40. John Wiley & Sons.

WOLENSKI, J. (1990). Deontic logic and possible world semantics : A historical sketch. *Studia Logica*, 49(2):273–282.

XU, M. (1995). On the basic logic of stit with a single agent. *The Journal of Symbolic Logic*, 60(2):459–483.

Abréviations et notations

df	définition
EBF	énoncé bien formé
MAS	*multi-agent system*
NMAS	*normative multi-agent system*
NS	*normative system*
$Prop$	ensemble de propositions atomiques
IL	logique propositionnelle intuitionniste
LPC	logique propositionnelle classique
t.q.	tel que
$\overline{\alpha}$	algèbre (complément, négation)
$\cap, \frown, \sqcap, \cdot, ;$	algèbre (conjonction)
$\cup, \smile, \sqcup, +$	algèbre (disjonction)
$*$	algèbre (étoile de Kleene pour la récursion)
$1, \mathbf{1}$	algèbre (unité de la conjonction)
$0, \mathbf{0}$	algèbre (unité de la disjonction)
$\neq$	différence
$=$	égalité
$\therefore$	en conséquence, donc
$a_w(\varphi)$	logique (assignation de valeur de vérité)
$\wedge$	logique (conjonction)
$\models$	logique (conséquence sémantique)
$\vdash$	logique (conséquence syntaxique)
$\vee$	logique (disjonction)
$\|\|\varphi\|\|$	logique (ensemble contenant les scénarios où φ est vrai)
$\equiv, \leftrightarrow$	logique (équivalence, biconditionnel)
$\bot$	logique (*falsum*, faux, contradiction)
$\supset, \rightarrow$	logique (implication)
$\mathcal{L}$	logique (langage)
$\Box$	logique (modalité)

$[\alpha]$	logique (modalité dynamique)
$[i :\ stit]$	logique (modalité *stit*)
$\Diamond$	logique (modalité duale)
$\langle\alpha\rangle$	logique (modalité duale dynamique)
$\langle i :\ stit\rangle$	logique (modalité duale *stit*)
$\neg, \sim$	logique (négation)
$\not\models$	logique (non conséquence sémantique)
$\nvdash$	logique (non conséquence syntaxique)
$\exists$	logique (quantificateur existentiel)
$\forall$	logique (quantificateur universel)
$\top$	(*verum*, vrai, tautologie)
O	opérateur déontique (obligation)
$O(\varphi/\psi), O_\psi\varphi$	opérateur déontique (obligation conditionnelle)
P	opérateur déontique (permission)
F	opérateur déontique (interdiction)
$<, \prec, \sqsubset, \ll$	relation d'ordre (plus petit que)
$\leq, \preceq, \sqsubseteq$	relation d'ordre (plus petit ou égal)
$\Rightarrow$	si, alors
$\Leftrightarrow$	si et seulement si
$+$	somme
$\in$	théorie des ensembles (appartenance)
$-$	théorie des ensembles (complément)
$\circ$	théorie des ensembles (composition de fonctions, de relations)
$\{,\}$	théorie des ensembles (ensemble)
$\mathbb{N}$	théorie des ensembles (ensemble des nombres naturels)
$\mathbb{R}$	théorie des ensembles (ensemble des nombres réels)
$\wp$	théorie des ensembles (ensemble des parties)
$\varnothing$	théorie des ensembles (ensemble vide)
$f : A \longrightarrow B$	théorie des ensembles (fonction)
$\cap$	théorie des ensembles (intersection)
$\notin$	théorie des ensembles (non appartenance)
$< a, b >$	théorie des ensembles (paire ordonnée)
$\times$	théorie des ensembles (produit cartésien)
$\subseteq$	théorie des ensembles (sous-ensemble)
$\subset$	théorie des ensembles (sous-ensemble propre)
$\langle,\rangle$	théorie des ensembles (structure)
$\cup$	théorie des ensembles (union)

Index

Table des matières

EXTRAIT DU CATALOGUE

Âme et iPad • Maurizio Ferraris

L'Amérique selon Sartre. Littérature, philosophie, politique • Yan Hamel

Approches critiques de la pensée japonaise du XXe siècle • Livia Monnet

Confucius du profane au sacré • Herbert Fingarette

Le cosmopolitisme. Enjeux et débats contemporains • Sous la direction de Ryoa Chung et Geneviève Nootens

La culture de la mémoire ou comment se débarrasser du passé ? • Éric Méchoulan

De ce qui agit à ce qui voit • Nishida Kitaro, traduction et appareil critique de Jacynthe Tremblay

Le devenir-juif du poème. Double envoi : Celan et Derrida • Danielle Cohen-Levinas

Éléments de logique contemporaine. Troisième édition revue et augmentée avec exercices et corrigés • François Lepage

Entre science et culture. Introduction à la philosophie des sciences • Yvon Gauthier

L'esprit et la nature • Daniel Laurier

Introduction à la métaphysique • Jean Grondin

Justice et démocratie. Une introduction à la philosophie politique • Christian Nadeau

Kulturkritik et philosophie thérapeutique chez le jeune Nietzsche • Martine Béland

Le laboratoire de la dialectique de la raison. Discussions, notes et fragments inédits • Max Horkheimer et Theodor W. Adorno

Lectures nietzschéennes. Sources et réceptions • Sous la direction de Martine Béland

Leibniz et Diderot. Rencontres et transformations • Christian Leduc, François Pépin, Anne-Lise Rey et Mitia Rioux-Beaulne

Leibniz et l'individualité organique • Jeanne Roland

Les limites du soi. Immunologie et identité biologique • Thomas Pradeu

Mythe et philosophie à l'aube de la Chine impériale. Études sur le Huainan zi • Sous la direction de Charles Le Blanc et Rémi Mathieu

Pensée rationnelle et argumentation • Clayton Peterson

Philosophes japonais contemporains • Jacynthe Tremblay

Profession éthicien • Daniel M. Weinstock

Profession philosophe • Michel Seymour

Raconter et mourir. Aux sources narratives de l'imaginaire occidental (nouvelle édition) • Thierry Hentsch

Raisonnement et pensée critique. Introduction à la logique informelle • Martin Montminy

Substance, individu et connaissance chez Leibniz • Christian Leduc

Sur le seuil du temps. Essais sur la photographie • Siegfried Kracauer

Le temps aboli. L'Occident et ses grands récits • Thierry Hentsch

paramètres

La logique déontique s'insère dans la tradition de la philosophie analytique et s'applique aux discours éthique et juridique. Elle s'attache à ce qu'il convient de faire dans une situation donnée en regard de ce qui est obligatoire, permis ou interdit et peut être vue comme l'analyse de la structure des inférences normatives, c'est-à-dire des relations entre les énoncés et les concepts.

Fruit d'une revue exhaustive de la littérature scientifique, ce livre présente les travaux les plus importants du domaine regroupés par sujets, méthodes et objectifs. On y aborde les problèmes et les paradoxes habituels, les notions de base formelles, ainsi que les divers courants en les comparant entre eux. L'auteur s'adresse à un public intéressé par l'analyse des raisonnements éthique et juridique ainsi qu'aux étudiants en philosophie, en mathématiques, en informatique, en sciences humaines et en droit.

Clayton Peterson est titulaire d'un doctorat en philosophie de l'Université de Montréal. Il a effectué un stage postdoctoral au Munich Center for Mathematical Philosophy (LMU-Munich) et a publié aux PUM *Pensée rationnelle et argumentation* (2013). Ses travaux portent sur la philosophie formelle et celle des sciences.

34,95 $ • 31 €

Couverture: ©Tanor/Canstockphoto.com

également disponible en version numérique
www.pum.umontreal.ca

isbn 978-2-7606-3520-3

9 782760 635203