

Aspects statiques et dynamiques des croyances

S. Djeffal, B. Chaib-draa
Département d'informatique, Faculté des Sciences, Université Laval,
Sainte-Foy, PQ, Canada, G1K 7P4
e-mail : djeffal,chaib@ift.ulaval.ca

1 Introduction

Il est généralement admis, en intelligence artificielle (IA) et en intelligence artificielle distribuée (IAD) tout particulièrement, que l'attribution d'attitudes mentales aux agents artificiels est légitime et d'une grande utilité. McCarthy, l'un des pionniers de cette approche, soutient que :

Attribuer des attitudes mentales à une machine est légitime lorsqu'une telle attribution exprime la même information à propos de la machine qu'elle exprime à propos d'une personne. Cette attribution est utile quand elle nous permet de comprendre la structure de la machine, son comportement ancien ou futur, ou comment la réparer ou l'améliorer. . . . exprimer raisonnablement et brièvement ce qui est connu à propos de l'état d'une machine dans une situation particulière peut nécessiter l'attribution de qualités mentales. [McC79, traduction libre]

Ceci dit, il est donc avantageux d'avoir un modèle pour ce *niveau mental*.

Ce modèle nous donne la possibilité de décrire le comportement d'un système sans avoir recours à une description détaillée de son implémentation en machine ou de ses composants physiques. Nous pouvons, aussi, l'utiliser pour évaluer un système en jugeant si ses attitudes mentales ont un sens. C'est pour ces raisons, entre autres, que la modélisation formelle des attitudes mentales connaît en IA, depuis plusieurs années, un intérêt, sans cesse, grandissant. Par modèles formels, nous faisons référence aux modèles basés sur la logique. Les attitudes mentales qui ont été les plus étudiées sont les croyances, les connaissances, les intentions, les buts, les désirs, les obligations et les engagements. Parmi les thèmes de recherche où ces attitudes mentales ont un rôle central, on peut citer :

- l'interaction entre les connaissances et l'action [Moo80];
- la compréhension du langage naturel et la théorie des actes de langage [App85, LC90, LA87, Gal88, AP80, CL90, CP79];
- la formalisation d'agents rationnels basée sur une architecture BDI (croyance–désir–intention) [RG91];
- la programmation orientée agent [Sho93, Tho93];
- les logiques pour la spécification et la vérification des systèmes multi-agents [WF92, Woo94];
- les logiques pour la formalisation d'agents à rationalité limitée [Hua94];

- spécification et raisonnement sur les connaissances dans les systèmes distribués [FHV92, FHMV95];
- la robotique cognitive [L⁺94, Les91].

1.1 Statique vs. dynamique

Comme toutes les attitudes mentales, les croyances présentent deux aspects distincts : la statique et la dynamique. Le but de la statique est de définir les plus importantes propriétés que les croyances devraient avoir. La consistance et l'introspection sont, par exemple, des propriétés des croyances d'un agent rationnel. Le modèle des mondes possibles est le modèle le plus utilisé en IA/IAD pour modéliser formellement les croyances d'un point de vue statique.

Une propriété fondamentale de la croyance est, sans aucun doute, sa sensibilité au changement. C'est la dynamique qui se préoccupe de la modélisation du changement de croyances. La théorie formelle de changement des croyances la plus influente et la mieux développée est la théorie AGM, développée par Alchourrón, Gärdenfors et Makinson [MG91, Gär88, AGM85, MG88]. AGM propose un ensemble de postulats que toute opération de révision des croyances doit satisfaire et définit des constructions explicites pour ces opérations. AGM est, jusqu'à présent, un passage recommandé pour la compréhension de la dynamique des croyances en général.

1.2 Révision vs. mise à jour

La révision des croyances ne représente pas, à elle seule, la dynamique des croyances. Certes, la révision est un aspect important de la dynamique des croyances, mais n'est, en fait, qu'un des aspects de la dynamique des croyances. La mise à jour constitue un autre aspect de la dynamique des croyances.

La révision et la mise à jour diffèrent par rapport aux préoccupations considérées. En effet, la révision se focalise sur la manière avec laquelle un agent change ses croyances en adoptant de nouvelles croyances. Par contre, la mise à jour se focalise, plutôt, sur la manière qu'un agent devrait changer ses croyances quand il réalise que son monde a changé. En effet, les croyances d'un agent peuvent devenir inexacts à cause des changements intervenus dans son monde. Chaque fois que des événements surviennent ou que d'autres agents réalisent des actions, certains faits deviennent vrais et d'autres faux. Un agent observant de tels processus ou utilisant leurs résultats doit s'assurer que son état épistémique reflète ces changements.

La révision et la mise à jour diffèrent aussi par rapport aux buts poursuivis. En effet, le but essentiel de la révision est l'amélioration de la qualité globale des croyances sur une situation statique mais imparfaitement appréhendée. Par contre, le but de la mise à jour est de faire de telle sorte que les croyances qui décrivent une situation évolutive reflètent, le plus fidèlement que possible, la dynamique de la situation décrite.

Pour bien illustrer la distinction entre révision et mise à jour, prenons l'exemple suivant : supposons qu'un agent croit qu'une salle contient exclusivement des chaises ou des tables. Maintenant, supposons que cet agent apprenne, d'une manière ou une autre, qu'il n'y a pas de chaises dans la salle. En révisant ses croyances, l'agent peut déduire que la salle contient des tables. Dans une autre interprétation, on peut raisonnablement dire que le fait qu'il n'y a pas de chaises reflète la situation où il y avait des chaises dans la salle et que maintenant il n'en reste plus (quelqu'un les a retirées, un robot par exemple). Dans ce cas, l'agent ne peut déduire que la salle contient des tables. La seule chose que l'agent peut raisonnablement croire c'est qu'il n'y a pas de chaises dans

la salle (il peut y avoir des tables comme il peut ne pas y en avoir). Dans le premier cas, l'agent a fait une révision. Le changement de croyances dû à la nouvelle information permet à l'agent d'améliorer ses croyances par rapport à une situation considérée statique et imparfaitement appréhendée. Dans le second cas, l'agent a fait une mise à jour. Pour que le changement de croyances dû à la nouvelle information reflète le plus fidèlement que possible la dynamique de la situation décrite par ses croyances, l'agent doit raisonner sur tous les cas possibles. En quelque sorte, il doit chercher les explications possibles au changement décrit par la nouvelle information. Pour l'exemple précédent, il est clair que la mise à jour est le changement de croyances le plus raisonnable. Katsuno et Mendelzon [KM91a] proposent pour la mise à jour des croyances de nouveaux postulats et une autre modélisation qui diffèrent des postulats et de la modélisation de la révision des croyances proposés par AGM.

1.3 "Fondements" vs. "cohérence"

Il y a deux principales approches pour aborder la révision : l'approche "fondements" et l'approche "cohérence". L'approche "fondements" exige que toutes les croyances d'un agent rationnel doivent être justifiées c'est à dire que l'agent doit avoir pour chacune de ses croyances une justification ou une raison satisfaisante. Cependant, certaines croyances, dites primitives, peuvent ne pas avoir de justifications : elles s'auto-justifient. Selon cette approche, il y a changement de croyances chaque fois qu'un agent adopte ou abandonne des justifications. Les systèmes de maintenance de la cohérence, dont le plus connu est TMS de Doyle [Doy79], s'inscrivent dans cette approche.

À l'inverse, l'approche "cohérence" n'exige pas de justifications pour les croyances. Un agent peut adopter une croyance tant que celle-ci est logiquement cohérente avec ses autres croyances. La théorie de révision des croyances AGM d'Alchourrón, Gärdenfors et Makinson [MG91, Gär88, AGM85, MG88], que nous présenterons avec plus de détails ultérieurement, est considérée comme l'exemple typique de l'approche "cohérence".

1.4 Types de révision

Il y a plusieurs types de changements de croyances. Le type le plus familier a lieu lors de l'acquisition de nouvelles connaissances par perception ou par acceptation d'informations fournies par d'autres personnes. Ce type de changement est appelé *expansion* d'un état épistémique¹.

Le second type de changement a lieu quand de nouvelles évidences viennent contredire des croyances faussement acceptées et que nous devons par conséquent réviser. Ce changement est appelé *révision*² d'un état épistémique.

Le troisième type de changement a lieu quand nous découvrons que les raisons de certaines croyances sont invalides et que nous devons de ce fait les rejeter. Ce type de changement est appelé *contraction*.

1.5 Problèmes de la révision des croyances

Nous avons vu que réviser des croyances consiste à faire l'une des trois opérations suivantes : l'expansion, la révision ou la contraction. Si faire une expansion est une opération simple, la révision et la contraction posent des problèmes. Pour illustrer d'une façon simple la complexité de ces deux

¹L'"état épistémique" est utilisée par AGM pour faire référence aux croyances d'un agent.

²Il faut distinguer la révision des croyances en tant qu'aspect de la dynamique et révision comme opération. Cette remarque est valable pour la mise à jour.

opérations, nous allons présenter un exemple de Gärdenfors [Gär92]: soient p , q , r et s quatre énoncés représentant nos croyances :

- p : tous les cygnes d'Europe sont blancs.
- q : l'oiseau attrapé est un cygne.
- r : l'oiseau attrapé vient de Suède.
- s : La Suède fait partie de l'Europe.

Par inférence logique, nous pouvons déduire à partir de p - s , l'énoncé suivant :

- t : l'oiseau attrapé est blanc.

Maintenant, supposons que nous nous apercevons que l'oiseau attrapé n'est pas blanc. Si nous acceptons cet énoncé qui n'est autre que $\neg t$, nos croyances deviennent inconsistantes du fait que nous avons deux croyances contradictoires. Si nous voulons maintenir la consistance de nos croyances, nous devons rejeter certaines d'entre elles. Ce rejet ne peut se faire de n'importe quelle manière. On ne peut, entre autres, rejeter inutilement des informations valables. Pour notre exemple, nous avons à choisir entre p , q , r , et s . Les considérations logiques, seules, ne peuvent nous aider à choisir. Ce qui rend le choix, encore, plus compliqué, c'est la prise en considération des conséquences logiques: en rejetant une croyance, nous avons aussi à décider lesquelles, parmi ses conséquences logiques, devons-nous retenir et lesquelles devons-nous rejeter. Si nous choisissons p , p a pour conséquences logiques, entre autres, les énoncés suivants :

- p' : tous les cygnes d'Europe à l'exception de celui attrapé sont blancs.
- p'' : tous les cygnes d'Europe à l'exception de certains provenant de Suède sont blancs.

La question qui se pose à nous, maintenant, est: “ Devons-nous rejeter p' et p'' , retenir l'un et rejeter l'autre ou bien les maintenir tous les deux? ”.

Il existe d'autres questions à considérer pour la révision des croyances, les plus importantes sont les deux suivantes :

1. *Comment les croyances sont représentées?* Il s'avère que le modèle de représentation influence la dynamique des croyances en général. Comme le souligne Gärdenfors [Gär92], la révision des croyances est sensitive au formalisme choisi pour représenter les croyances. Pour bien montrer cette sensitivité, prenons l'exemple [Som92] suivant: supposons que nous utilisons un radar pour repérer sur un écran la position d'un bateau. L'écran est divisé en quatre quadrants: nord-est, nord-ouest, sud-est et sud-ouest. Le radar indique que le bateau est en sud-ouest. Après un certain moment, on apprend que le bateau n'est pas en ouest, nous devons donc faire une révision car la nouvelle information contredit notre croyance de départ. Si nous adoptons le langage de représentation $L = \{\text{sud, ouest, nord, est}\}$, notre état épistémique est $\{\text{sud, ouest}\}$ et après révision devient $\{\text{sud, } \neg \text{ouest}\}$ c'est-à-dire que nous croyons que le bateau est en sud-est. Maintenant si $L = \{\text{sud-est, sud-ouest, nord-ouest, nord-est}\}$, notre état épistémique est $\{\text{sud-ouest}\}$, et après révision par la nouvelle information qui se traduit dans ce langage par $(\text{nord-est} \vee \text{sud-est})$, notre état épistémique devient $\{\text{nord-est} \vee \text{sud-est}\}$ c'est à dire que nous croyons que le bateau est en nord-est ou en sud-est. Il est clair que $\{\text{nord-est} \vee \text{sud-est}\}$ est différent de $\{\text{sud, } \neg \text{ouest}\}$. De cet exemple, on voit bien que la révision est sensiblement affectée par le langage de représentation adoptée.

2. *Quels sont les critères de choix qu'il convient d'adopter pour rejeter des croyances?* Il est clair que la logique, seule, est insuffisante pour déterminer quelles croyances retenir et lesquelles rejeter. Des facteurs extra-logiques devraient être utilisés dans la prise de décision. Cependant, ces facteurs ne sont pas bien définis et il semble peu probable qu'ils soient constants. L'idée directrice consiste à faire le moins de changements possibles. Il est utile d'avoir une évaluation des croyances : certaines croyances seraient considérées plus importantes que d'autres. La révision consisterait alors à rejeter les croyances les moins importantes. Ces facteurs extra-logiques peuvent jouer un rôle important dans la révision des croyances dans un environnement multi-agent.

2 La Statique

2.1 Approches formelles

Pour aborder les croyances d'un agent d'un point de vue statique, nous avons besoin d'un modèle de représentation. Par modèle nous faisons référence à une caractérisation abstraite de l'objet en considération. Dans notre cas, cet objet est le concept de croyance. Un modèle est une abstraction parce qu'il n'a pas toutes les propriétés de l'objet à modéliser. Il est incontestablement reconnu que les modèles sont très utiles pour comprendre et analyser des objets ou des concepts. Ils le sont davantage quand ils ne retiennent que les propriétés pertinentes du concept à modéliser.

Les modèles formels, en plus d'être utiles, ont l'avantage d'être concrets dans le sens où il est possible de prouver rigoureusement les propriétés du modèle et les prédictions qu'il fait. Maintenant, en adoptant la logique comme modèle formel pour la représentation des croyances d'un agent, on se donne l'accès aux outils et aux résultats de l'une des théories formelles les plus connues et les mieux définies. Avec son langage muni d'une syntaxe rigoureuse, sa sémantique claire permettant d'établir fidèlement et sans ambiguïté une correspondance entre ce qu'on veut exprimer et la manière avec laquelle il est représenté, et ses techniques bien établies pour la formalisation et la validation du raisonnement, la logique est un modèle formel puissant et fiable. Adopter la logique comme modèle de représentation des croyances est donc un choix judicieux. Cela ne veut pas dire que c'est l'unique modèle ou que c'est le modèle parfait.

Un modèle logique pour la représentation des croyances peut être caractérisé en termes de deux facteurs indépendants : *le modèle* et *le langage de formulation* (formalisme) de ce modèle [Kon86]. Deux principales approches peuvent être utilisées pour la caractérisation du modèle. La première consiste à adopter *le modèle d'attitude propositionnelle*. Dans ce cas, "croire" est considéré comme une attitude propositionnelle c'est-à-dire une relation entre un agent et des propositions abstraites concernant le monde. Les croyances de l'agent sont caractérisées par un ensemble de mondes possibles reliés entre eux par une relation d'accessibilité. La seconde approche consiste à adopter *le modèle syntaxique*³. Dans ce cas, les croyances d'un agent sont considérées comme des formules symboliques exprimées dans un langage interne et représentées explicitement dans une structure de données associée à cet agent. Pour expliciter d'une manière simple la différence entre les deux approches, considérons le concept "croire" dans "l'agent i croit p "⁴. Dans le modèle d'attitude propositionnelle cela signifie que *dans tous les mondes possibles accessibles à l'agent i , p*

³correspond à l'appellation "symbol-processing" ou "sentential model" de Konolige [Kon86, p. 83].

⁴Normalement, on doit dire "croit que p " à la place de "croit p " puisque c'est une attitude propositionnelle. La seconde expression est utilisée pour des raisons de simplifications et a le sens de la première.

est vraie. Dans le modèle syntaxique cela signifie que p exprimé dans un langage interne est inclus dans la structure de données associée à l'agent i .

En ce qui concerne le formalisme, il y a deux possibilités. La première consiste à utiliser un langage modal contenant des opérateurs modaux s'appliquant sur des formules. La seconde consiste à utiliser un métalangage : un langage de premier ordre contenant des termes faisant référence à des formules d'un autre langage, le langage objet. Le métalangage correspond au langage externe c'est-à-dire le langage utilisé par un observateur externe pour faire référence aux croyances d'un agent observé. Le langage objet quant à lui correspond au langage interne c'est-à-dire le langage utilisé par l'agent pour formaliser ses croyances. Si on considère que les croyances d'un agent i sont un ensemble d'énoncés formalisés dans un langage propositionnel, ce dernier est le langage interne. Le langage externe doit avoir des termes faisant référence aux éléments du langage interne et un prédicat *Croire*. $Croire(i, p)$ a pour signification "*l'agent i croit l'énoncé p* ". Ce langage externe est un métalangage puisqu'il fait référence à un autre langage, le langage objet.

Au modèle d'attitude propositionnelle est souvent associée une logique modale normale comme formalisme. Cette approche est la plus répandue dans la littérature concernant les croyances en IA. Au modèle syntaxique, c'est souvent un métalangage (logique du premier ordre) qui lui est associé comme formalisme. Le modèle de déduction de Konolige [Kon86] est une exception : Konolige utilise une approche syntaxique à laquelle est associée une logique modale comme formalisme. Pour plus de détails, une étude exhaustive et comparative des deux modèles et de leurs formalismes associés est présentée dans [Kon86]. Un des plus importants résultats de cette étude est qu'une correspondance entre les deux modèles est possible selon certaines conditions. Nous nous en tiendrons à cela.

Dans ce qui suit, nous allons donner une présentation détaillée de la logique épistémique⁵ qui repose sémantiquement sur le modèle des mondes possibles et syntaxiquement sur un langage modal. Du fait que la logique épistémique se fonde sur la logique modale normale, nous introduirons en premier lieu cette dernière. Nous avons choisi de présenter cette approche pour deux raisons : la première est que c'est la plus utilisée en IA pour la représentation des croyances et la deuxième c'est qu'elle nous permet de présenter clairement les plus importantes propriétés des croyances.

2.2 Logiques modales normales

La logique modale est venue répondre à l'inaptitude de la logique classique (logique des propositions et logique des prédicats) à modéliser et à résoudre certains problèmes, spécialement ceux faisant intervenir des concepts intentionnels. La logique modale est une extension de la logique classique : ses langages et ses théorèmes généralisent ceux de la logique classique. En plus des opérateurs logiques classiques, la logique modale utilise en plus des opérateurs logiques classiques, deux opérateurs supplémentaires \Box et \Diamond . \Box est appelé opérateur modal universel et \Diamond est appelé opérateur modal existentiel.

2.2.1 Significations des opérateurs

La première signification de \Box est "*il est nécessaire*" et celle de \Diamond est "*il est possible*". $\Box p$ signifie " p est nécessairement vrai" et $\Diamond p$ signifie " p est possiblement vrai". La vérité exprimée par \Box est une vérité nécessaire c'est-à-dire, dans aucun cas, p ne peut être faux. Par contre, la vérité exprimée par \Diamond est une vérité contingente c'est-à-dire la fausseté de p peut être envisagée.

⁵Le terme "épistémique" fait souvent référence aux connaissances, en ce qui nous concerne, il est utilisé pour faire référence aux croyances.

La forte expressivité de la logique modale est une propriété remarquable : on peut donner à chacun des deux opérateurs plusieurs significations en fonction de la modalité qu'on voudrait exprimer à la seule condition de respecter la relation de dualité qui lie les deux opérateurs :

$$\diamond \stackrel{\text{def}}{=} \neg \Box \neg$$

La syntaxe

La syntaxe de la logique modale se définit comme suit :

- Toutes les règles syntaxiques de la logique des propositions sont aussi des règles de la logique modale;
- si ϕ est une formule alors $\Box\phi$ et $\diamond\phi$ sont aussi des formules.

Il est à noter qu'on peut utiliser un seul opérateur, le deuxième peut être déduit de la relation de dualité.

La sémantique

La sémantique utilisée est la sémantique des mondes possibles de Kripke [Kri63]. L'univers est considéré comme un ensemble de mondes reliés par une relation d'accessibilité. Chaque monde représente une vue de ce que pourrait possiblement être le monde. L'ensemble des mondes possibles est un modèle complet pour le langage sans opérateurs modaux. En effet, chaque monde possible spécifie la vérité ou la fausseté de chaque formule d'une manière consistante. Le monde réel représente un monde particulier. Les modalités sont incorporées dans la relation d'accessibilité. Cette vision des choses permet de définir plus clairement les notions de vérité nécessaire et vérité contingente. Une proposition ou une formule est une vérité nécessaire si elle est vraie dans tous les mondes possibles et c'est une vérité contingente si elle ne l'est que dans certains mondes. Nous dirons qu'un monde w_2 est accessible à partir d'un monde w_1 , si et seulement si, il existe une relation d'accessibilité entre w_1 et w_2 . D'une manière plus formelle, nous appellerons structure le couple $\langle W, R \rangle$ et modèle le triplet $\langle W, R, \pi \rangle$ avec :

- W : ensemble non-vide des mondes possibles;
- R : relation binaire sur W , $R \subseteq W \times W$;
- π : fonction de valuation,
 $\pi : W \times Prop \rightarrow \{vrai, faux\}$.

La sémantique de la logique modale est définie à partir de \models_w qui n'est autre que la relation de satisfaction définie par rapport au monde w :

- $M \models_w vrai$.
- $M \models_w p \in Prop$ ssi $\pi(w, p) = vrai$.
- $M \models_w \neg\phi$ ssi $M \not\models_w \phi$.
- $M \models_w \phi \vee \psi$ ssi $M \models_w \phi$ ou $M \models_w \psi$.

- $M \models_w \phi \wedge \psi$ ssi $M \models_w \phi$ et $M \models_w \psi$.
- $M \models_w \phi \rightarrow \psi$ ssi $M \not\models_w \phi$ ou $M \models_w \psi$.
- $M \models_w \phi \equiv \psi$ ssi $(M \models_w \phi \text{ et } M \models_w \psi)$ ou $(M \not\models_w \phi \text{ et } M \not\models_w \psi)$.
- $M \models_w \Box\phi$ ssi $\forall w'$ si wRw' alors $M \models_{w'} \phi$.
- $M \models_w \Diamond\phi$ ssi $\exists w'$ tel que wRw' et $M \models_{w'} \phi$.

Les axiomes

Toute axiomatisation de la logique modale normale comporte :

- tous les axiomes de la logique des propositions et la règle du modus ponens;
- le schéma d'axiome de distribution noté K : $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$
(L'axiome K est présent dans tout système logique modal du fait qu'il est valide (vrai dans toute structure). $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$ est équivalent à $\Box\phi \wedge \Box(\phi \rightarrow \psi) \rightarrow \Box\psi$.);
- la règle d'inférence modale de nécessité : si $\vdash \phi$ alors $\vdash \Box\phi$.

Il y a plusieurs autres axiomes mais qui ne sont vrais que dans certaines structures. Chellas [Che80] présente une longue liste de ces axiomes. On se limitera à citer ceux qui ont de l'intérêt pour les croyances :

- l'axiome T : $\Box\phi \rightarrow \phi$.
- l'axiome D : $\Box\phi \rightarrow \Diamond\phi$.
- l'axiome 4 : $\Box\phi \rightarrow \Box\Box\phi$.
- l'axiome 5 : $\Diamond\phi \rightarrow \Box\Diamond\phi$.

Une logique modale est dite normale si elle contient le schéma d'axiome K et si elle est dotée de la règle de nécessité.

2.2.2 La théorie de correspondance

Comme toute relation binaire, la relation d'accessibilité possède plusieurs propriétés. Chaque propriété permet de définir une structure particulière. À part l'axiome K , tous les autres axiomes ne sont vrais que dans des structures particulières. D'un autre côté, chaque axiome caractérise une certaine propriété de R (cependant certaines propriétés de R ne correspondent à aucun axiome : la non-réflexivité, l'antisymétrie et l'asymétrie.). Par exemple, l'axiome T n'est vrai que dans une structure réflexive. La structure d'un système logique modal qui admet T , D , 4 et 5 comme axiomes est nécessairement munie d'une relation d'équivalence. Cette façon de faire correspondre axiomes et propriétés de la relation d'accessibilité R est appelée théorie de correspondance. Cette caractéristique fondamentale est pour beaucoup dans le succès rencontré par la sémantique des mondes possibles.

2.2.3 Systèmes logiques modaux

Un système logique est un ensemble de formules valides dans une classe de modèles. Une formule qui appartient à cet ensemble est un théorème de cette logique. Le plus petit système logique normal contenant les schémas d'axiomes $\Sigma_1 \dots \Sigma_n$ est noté par :

$$K\Sigma_1 \dots \Sigma_n.$$

Du fait que chaque axiome définit une propriété particulière de R , pour les axiomes T , D , 4 et 5 on peut obtenir $2^4 = 16$ systèmes logiques (16 structures). Réellement, on n'obtient que 11 puisque certains systèmes sont équivalents : ils contiennent les mêmes théorèmes [Kon86].

2.3 Logique épistémique

La logique modale offre un cadre formel intéressant pour la modélisation des connaissances et des croyances d'un agent. Comme on l'a déjà mentionné, les opérateurs modaux peuvent avoir plusieurs significations en fonction des modalités qu'on voudrait représenter. Pour la connaissance, l'opérateur modal universel signifie "*il est connu*" et on le note K (Knowledge). Kp se lit "*p est connu*". Pour les croyances, l'opérateur modal universel signifie "*il est cru*" et on le note B (Belief). Bp se lit "*p est cru*". La relation R définit les liens entre les mondes compatibles avec la connaissance ou la croyance d'un agent. Concentrons nous juste sur les croyances tout en ayant présent à l'esprit que "connaître p " équivaut à "croire p et p est vrai" :

$$Kp \stackrel{\text{def}}{=} Bp \wedge p$$

Tout d'abord, montrons ce que les mondes possibles représentent pour la modélisation des croyances d'un agent. Considérons que *l'agent i croit p* . Imaginons, maintenant, que nous le prenons tel qu'il est dans un monde quelconque w et montrons lui tous les autres mondes un à la suite de l'autre. Si l'agent i trouve dans un monde qu'une proposition est incompatible avec ce qu'il croit, alors ce monde-là ne lui est pas accessible. Par contre, si tout correspond avec ce qu'il croit, ce monde lui est accessible. Tout ce qu'un agent croit vrai doit obligatoirement être vrai dans tous les mondes accessibles et tout ce qu'il croit faux doit être faux. Toute proposition pour laquelle il n'a aucune croyance particulière peut être vraie dans certains mondes et fautive dans d'autres. Si on désigne par W' l'ensemble des mondes compatibles avec les croyances de i , on a :

- i croit p si, et seulement si, pour tout monde $w' \in W'$, p est vrai dans w' ;
- i ne croit pas p si, et seulement si, il existe un monde $w' \in W'$ dans lequel p est faux;
- tout monde $w \in W'$ peut être le monde réel.

La syntaxe et la sémantique de la logique épistémique sont identiques à celles de la logique modale normale en remplaçant \Box par B . Voyons, maintenant, ce que les axiomes définissent comme propriétés pour les croyances d'un agent. L'axiome K définit la fermeture sur les conséquences logiques d'une formule. Si un agent croit ϕ alors il croit toutes les conséquences logiques de ϕ . L'axiome T signifie que tout ce qu'un agent croit est vrai. Cet axiome est valable pour la modélisation des connaissances mais ne l'est pas pour les croyances. En effet, les croyances d'un agent peuvent s'avérer fausses. L'axiome D indique que les croyances d'un agent ne sont pas contradictoires : il ne peut croire à la fois à une formule et à sa négation. L'axiome 4 munit l'agent modélisé de la faculté

d'*introspection positive*: s'il croit à une formule alors il croit qu'il la croit. L'axiome 5 définit la faculté d'*introspection négative*: si l'agent ne croit pas à une formule alors il croit qu'il ne la croit pas. La règle de nécessité signifie que toute formule valide est une croyance de l'agent.

2.3.1 Choix d'un système modal

Le choix d'un système modal dépend des propriétés à attribuer aux croyances d'un agent. Si on veut par exemple modéliser les croyances d'un agent idéalement rationnel ayant une faculté d'introspection parfaite de ce qu'il croit et ne croit pas, il faut choisir le système *faible S5* (*KD45*). Si on veut que la faculté d'introspection soit réduite à ce que l'agent croit alors il faut choisir le système *faible S4* (*KD4*). Généralement, on utilise le système *S5* (*KDT45*) pour la modélisation des connaissances et le système *faible S5* pour la modélisation des croyances. Les systèmes *faibleS5* et *faible S4* sont adéquats et complets pour leur sémantique respective. Il est à signaler qu'indépendamment du système modal choisi, l'agent modélisé croit nécessairement toutes les conséquences logiques de ce qu'il croit puisque *K* fait partie de n'importe quel système modal.

2.3.2 Logique épistémique multi-agent

La logique modale présentée précédemment ne permet de formaliser que les croyances d'un seul agent. Pour modéliser un système multi-agent, nous avons besoin d'une logique multimodale. Une logique multimodale est une logique modale utilisant plus d'un opérateur modal. Nous utiliserons l'ensemble $\{B_i \mid i \in \{1, \dots, n\}\}$ comme opérateurs universels. $B_i\phi$ signifie "l'agent *i* croit ϕ ".

La syntaxe

- Toutes les règles syntaxiques de la logique des propositions sont aussi des règles syntaxiques de la logique épistémique multi-agent;
- si ϕ est une formule alors $B_i\phi$ est une formule.

La sémantique

Une structure pour une logique épistémique multi-agent est le couple $\langle W, \{R_i\} \rangle$ et un modèle est le triplet $\langle W, \{R_i\}, \pi \rangle$. $\{R_i\}$ est l'ensemble de relations d'accessibilité, R_i est la relation d'accessibilité de l'agent *i*. W et π ont les significations habituelles (voir subsection 2.3.3). La sémantique de la logique épistémique multi-agent est identique à celle de la logique épistémique normale sauf pour la sémantique des opérateurs modaux :

- $M \models_w B_i\phi$ ssi $\forall w' \in W$ si wR_iw' alors $M \models_{w'} \phi$.

À chaque axiome *T*, *D*, 4, 5 correspond respectivement un ensemble d'axiomes $\{T_i\}$, $\{D_i\}$, $\{4_i\}$, $\{5_i\}$ pour la logique épistémique multi-agent et à tout système modal Σ correspond un système épistémique multi-agent Σ_I avec $I = \{1, \dots, n\}$. Le plus petit système est K_I et le plus large est $S5_I$.

Les axiomes

Toute axiomatisation de la logique épistémique multi-agent comporte :

- tous les axiomes de la logique des propositions et la règle du modus ponens;
- les schémas d'axiome de distribution $K_i : B_i(\phi \rightarrow \psi) \rightarrow (B_i\phi \rightarrow B_i\psi)$;
 $(B_i(\phi \rightarrow \psi) \rightarrow (B_i\phi \rightarrow B_i\psi) \equiv B_i\phi \wedge B_i(\phi \rightarrow \psi) \rightarrow B_i\psi)$;
- les règles d'inférence modales de nécessité: si $\vdash \phi$ alors $\vdash B_i\phi$.

D'autres axiomes ne sont vrais que dans certaines structures :

- les axiomes $D_i : B_i\phi \rightarrow \neg B_i\neg\phi$;
- les axiomes $4_i : B_i\phi \rightarrow B_iB_i\phi$;
- les axiome $5_i : \neg B_i\phi \rightarrow B_i\neg B_i\phi$.

On a enlevé les T_i parce qu'ils ne s'appliquent pas aux croyances.

2.3.3 Conséquences du modèle des mondes possibles

Pour illustrer les conséquences du modèle des mondes possibles sur les croyances d'un agent i , considérons les cas suivants :

1. i croit ϕ , et ψ est une conséquence logique de ϕ . Dans ce cas, ψ est présente dans tout monde possible où ϕ est présente. Du fait que i croit ϕ , ϕ est présente dans tous les mondes qui sont accessibles à l'agent i . Puisque ψ est présente là où ϕ l'est, elle doit être, entre autres, présente dans les mondes accessibles à l'agent i . Par conséquent, i croit ψ . C'est ce qu'exige l'axiome K .
2. soit ϕ une formule valide. ϕ est donc présente dans tous les mondes possibles parmi lesquels se trouvent nécessairement les mondes accessibles à l'agent i . Par conséquent, i croit ϕ . C'est ce qu'exige la règle de nécessité.
3. soient ϕ une formule valide et ψ une conséquence logique de ϕ . Par 1 et 2, i croit ψ ;
4. soient ϕ et ψ deux formules inconsistantes. Cela signifie que les deux formules ne peuvent être présentes dans un même monde possible. Par définition, nous savons qu'une croyance d'un agent doit être nécessairement présente dans l'ensemble des mondes accessibles par l'agent. Par conséquent, l'agent ne peut croire ϕ et ψ .
5. i croit ϕ et ϕ est équivalente à ψ . Cela signifie que dans chaque monde où ϕ est présent ψ l'est aussi. Comme ϕ est présente dans tous les mondes accessibles à l'agent i , nous pouvons, donc, déduire que ψ est aussi présente dans ces mondes accessibles. Par conséquent, l'agent i croit ψ .
6. i croit ϕ et croit ψ . Dans ce cas, $\phi \wedge \psi$ est vraie dans tous les mondes compatibles avec les croyances de l'agent. Par conséquent, l'agent croit $\phi \wedge \psi$.
7. i croit ϕ . Dans ce cas pour n'importe quelle proposition ψ , $\phi \vee \psi$ est vrai dans tous les mondes compatibles avec les croyances de i , puisque ϕ est vraie. Par conséquent, i croit $\phi \vee \psi$.

Pour résumer ces diverses considérations, nous dirons qu'en prenant les mondes possibles comme modèle de représentation des croyances d'un agent, on fait face aux conséquences problématiques suivantes :

1. la fermeture sur l'implication matérielle: un agent croit toutes les conséquences logiques de ses croyances;
2. la fermeture sur la règle de nécessité: un agent croit toute formule valide
3. la fermeture sur l'implication valide: un agent croit toutes les conséquences de toute formule valide;
4. l'impossibilité d'avoir des croyances inconsistantes;
5. l'impossibilité de discerner des propositions équivalentes: des propositions équivalentes sont pour un agent des croyances identiques;
6. la fermeture sur la conjonction: un agent croit toutes les conjonctions de ses croyances;
7. la fermeture sur la disjonction: un agent croit toute disjonction dont au moins un composant est une croyance.

Toutes ces considérations forment ce que Hintikka [Hin62] a appelé *le problème de l'omniscience logique*.

2.4 L'omniscience logique

La logique épistémique telle qu'on vient de la présenter est reconnue depuis qu'elle a été proposée par Hintikka [Hin62] comme étant une version idéalisée des connaissances et des croyances. Cette idéalisation se manifeste dans l'omniscience logique attribuée à l'agent par rapport à ses croyances. Cela provient, comme on l'a déjà montré, du modèle des mondes possibles et de ses conséquences. Les formes de l'omniscience les plus problématiques sont la fermeture sur les conséquences logiques due à l'axiome K et l'adoption de toutes les formules valides comme croyances due à la règle de nécessité. Plusieurs tentatives ont été faites pour régler le problème de l'omniscience tout en gardant la nature fondamentale du modèle des mondes possibles. Les plus connues de ces tentatives sont celles de Levesque [Lev84] et de Fagin et Halpern [FH88].

2.4.1 Logique de Levesque

Levesque [Lev84] a proposé pour éviter l'omniscience logique de faire la distinction entre croyances *explicites* et *implicites*. Les croyances explicites représentent ce qu'un agent croit explicitement et les croyances implicites sont toutes les croyances obtenues à partir des croyances explicites par dérivation.

Syntaxiquement, la logique de Levesque est une logique propositionnelle classique étendue par deux opérateurs modaux : B pour croyance explicite et L pour croyance implicite (L correspond au B classique). Si ϕ est une formule alors $B\phi$ et $L\phi$ sont aussi des formules. L'enchâssement des opérateurs modaux n'est pas permis et le cas multi-agent n'est pas considéré.

Sémantiquement, les croyances explicites et implicites sont définies à partir de *situations*. Les situations peuvent être considérées comme des mondes possibles à la seule différence que dans une

situation, à toute proposition peuvent être assignées les valeurs vrai, faux, les deux à la fois ou aucune des deux.

La logique de Levesque a pour modèle $M = \langle \mathcal{S}, \mathcal{B}, \mathcal{T}, \mathcal{F} \rangle$:

- \mathcal{S} : ensemble de situations;
- $\mathcal{B} \subseteq \mathcal{S}$ ensemble des situations compatibles avec les croyances explicites d'un agent;
- $\mathcal{T} : Prop \longrightarrow \mathcal{S}$ est une fonction qui, pour une proposition primitive, donne comme résultat l'ensemble des situations où cette proposition est vraie;
- $\mathcal{F} : Prop \longrightarrow \mathcal{S}$ est une fonction qui, pour une proposition primitive, donne comme résultat l'ensemble des situations où cette proposition est fausse.

$\mathcal{T}(p)$ représente toutes les situations où p est vraie, et $\mathcal{F}(p)$ représente toutes les situations où p est fausse. Une situation s est :

- *partielle*, si elle ne supporte ni la vérité ni la fausseté d'au moins une proposition p ;
- *incohérente*, si elle supporte à la fois la vérité et la fausseté d'au moins une proposition p ;
- *complète*, si elle définit la vérité ou la fausseté pour toute proposition p : s appartient exactement à $\mathcal{T}(p)$ ou $\mathcal{F}(p)$ pour toute proposition p ;
- *compatible* avec une situation s' , si s et s' concordent partout où s' est définie : si $s' \in \mathcal{T}(p)$ alors $s \in \mathcal{T}(p)$ et si $s' \in \mathcal{F}(p)$ alors $s \in \mathcal{F}(p)$ pour toute proposition p .

Soit \mathcal{B}' l'ensemble de toutes les situations complètes de \mathcal{S} qui sont compatibles avec des situations de \mathcal{B} .

Levesque définit deux relations de satisfaction $\models_{\mathcal{T}}$ et $\models_{\mathcal{F}}$. La relation $\models_{\mathcal{T}}$ supporte la vérité et la relation $\models_{\mathcal{F}}$ supporte la fausseté. Les règles sémantiques se définissent à partir de ces deux relations de la manière suivante :

- $M, s \models_{\mathcal{T}} p$ ssi $s \in \mathcal{T}(p)$,
- $M, s \models_{\mathcal{F}} p$ ssi $s \in \mathcal{F}(p)$;
- $M, s \models_{\mathcal{T}} \neg\phi$ ssi $M, s \not\models_{\mathcal{F}} \phi$,
- $M, s \models_{\mathcal{F}} \neg\phi$ ssi $M, s \models_{\mathcal{T}} \phi$;
- $M, s \models_{\mathcal{T}} \phi_1 \wedge \phi_2$ ssi $M, s \models_{\mathcal{T}} \phi_1$ et $M, s \models_{\mathcal{T}} \phi_2$,
- $M, s \models_{\mathcal{F}} \phi_1 \wedge \phi_2$ ssi $M, s \models_{\mathcal{F}} \phi_1$ ou $M, s \models_{\mathcal{F}} \phi_2$;
- $M, s \models_{\mathcal{T}} B\phi$ ssi $\forall t \in \mathcal{B}, M, t \models_{\mathcal{T}} \phi$,
- $M, s \models_{\mathcal{F}} B\phi$ ssi $M, s \not\models_{\mathcal{T}} B\phi$;
- $M, s \models_{\mathcal{T}} L\phi$ ssi $\forall t \in \mathcal{B}', M, t \models_{\mathcal{T}} \phi$,
- $M, s \models_{\mathcal{F}} L\phi$ ssi $M, s \not\models_{\mathcal{T}} L\phi$.

La relation entre B et L est définie par $B\phi \rightarrow L\phi$, si ϕ est une croyance explicite alors c'est aussi une croyance implicite. L se comporte comme un opérateur modal classique et l'opérateur B a les propriétés suivantes :

- les croyances explicites ne sont pas fermées sur l'implication :
 $Bp \wedge B(p \rightarrow q) \wedge \neg Bq$ peut être satisfaite;
- les formules valides ne sont pas nécessairement des croyances explicites :
 $\neg B(p \vee \neg p)$ peut être satisfaite;
- les croyances explicites ne sont pas fermées sur l'implication valide :
 $\models p \rightarrow (p \wedge (q \vee \neg q))$, $Bp \wedge \neg B(p \wedge (q \vee \neg q))$ peut être satisfaite.

Bien que la logique de Levesque règle certains problèmes de l'omniscience (1,2 et 3 de la subsection 2.4.3), elle a été l'objet de beaucoup de critiques. La plus importante concerne le concept de situation incohérente. Ce concept n'est finalement qu'une prouesse technique qui ne règle le problème qu'en apparence. Il est généralement admis que l'incohérence se situe dans la pensée ou dans le discours, mais pas dans les choses. Ce qu'il faut retenir de la logique de Levesque est la pertinente distinction entre croyance explicite et croyance implicite.

2.4.2 Logique générale de conscience

En se basant sur la distinction entre croyances explicites et croyances implicites et sans recourir aux situations incohérentes, Fagin et Halpern[FH88] ont proposé la logique générale de la conscience (l'appellation originale est *logic of general awareness* et la traduction est de Gochet [Goc92]). En plus des opérateurs B et L de Levesque, Fagin et Halpern ont introduit dans leur logique un troisième opérateur noté A . $A\phi$ a pour interprétation "*l'agent est conscient de ϕ* ". Cette logique considère le cas du multi-agent et permet l'enchâssement des opérateurs modaux. Pour chaque agent, on a B_i pour la croyance explicite, L_i pour la croyance implicite et A_i pour la "*conscience*". L'opérateur de "*conscience*" joue essentiellement le rôle d'un filtre syntaxique dans chaque monde possible.

La logique générale de la "*conscience*" a pour modèle $M = \langle S, \pi, \{\mathcal{A}_i\}, \{\mathcal{B}_i\} \rangle$:

- S : ensemble de situations;
- $\pi: Prop \times S \rightarrow \{vrai, faux\}$;
- $\mathcal{B}_i: \mathcal{B}_i \subseteq S \times S$ ensemble des situations compatibles avec les croyances explicites de l'agent i ;
- \mathcal{A}_i : fonction qui associe à une situation un ensemble de formules.

$\mathcal{A}_i(s)$ se compose des propositions dont l'agent i est conscient dans la situation s . Les formules qui sont dans $\mathcal{A}_i(s)$ sont celles dont l'agent i est conscient et ne sont pas nécessairement celles qu'il croit. Les règles sémantiques sont définies par la relation de satisfaction classique \models de la manière suivante :

- $M, s \models vrai$,
- $M, s \models p$ ssi $\pi(s, p) = vrai$,
- $M, s \models \neg\phi$ ssi $M, s \not\models \phi$,

- $M, s \models \phi_1 \wedge \phi_2$ ssi $M, s \models \phi_1$ et $M, s \models \phi_2$,
- $M, s \models A_i\phi$ ssi $\phi \in \mathcal{A}_i(s)$,
- $M, s \models L_i\phi$ ssi $M, t \models \phi, \forall t$ tel que $(s, t) \in \mathcal{B}_i$,
- $M, s \models B_i\phi$ ssi $\phi \in \mathcal{A}_i(s)$ et $M, t \models \phi, \forall t$ tel que $(s, t) \in \mathcal{B}_i$.

Dans cette logique, un agent i croit explicitement ϕ si :

- l'agent i croit implicitement ϕ (ϕ est vraie dans tous les mondes qu'il considère possibles) et
- l'agent i est "conscient" de ϕ .

Ce qui donne : $B_i\phi \equiv L_i\phi \wedge A_i\phi$. Cela signifie que pour qu'un agent croie explicitement une formule il doit en être "conscient". La règle de nécessité est affaiblie : pour qu'une formule valide soit crue par un agent il faut que l'agent soit conscient de cette formule. Ces modifications permettent de régler plusieurs formes de l'omniscience logique. La plus importante est que l'agent cesse de croire toutes les conséquences logiques de ses croyances. En effet, $B_i p \wedge B_i(p \rightarrow q) \wedge \neg B_i q$ peut être satisfaite puisque l'agent peut être conscient de p et de $p \rightarrow q$ sans toutefois être conscient de q . Un agent peut aussi croire deux formules sans toutefois croire à leur conjonction si l'agent n'en est pas "conscient".

La logique générale de la conscience arrive à résoudre plusieurs problèmes de l'omniscience. Cependant, sa définition de la croyance explicite comme conjonction de la croyance implicite et de la "conscience" a été l'objet de plusieurs critiques. La définition considère la "conscience" comme un "réservoir" d'énoncés : être conscient d'une formule signifie avoir cette formule présente dans une structure représentant la "conscience". Stalnaker [Sta91] critique une telle conception de la "conscience". Selon lui, un modèle du type "réservoir" offre peu d'espoir pour clarifier le problème de l'omniscience logique.

D'un point de vue modèle, l'opérateur de conscience A est un concept syntaxique puisque son argument est un énoncé et non une proposition. Cela ne permet pas de faire de connections entre lui et la relation d'accessibilité. Par conséquent, la relation entre relation d'accessibilité et croyance explicite est affaiblie. Plusieurs propriétés obtenues à partir du modèle des mondes possibles ne peuvent plus être obtenues dans la logique générale de conscience [Kon86, p. 110].

Enfin, la logique générale de la conscience attribue à l'agent une immunité par rapport aux croyances incohérentes [Goc92]. Si l'agent peut ne pas croire une formule valide par affaiblissement de la règle de la nécessité, il ne peut, par contre, croire une contradiction. Certes l'agent peut être conscient à la fois de p et de $\neg p$ mais il ne peut croire p et croire $\neg p$. Cette impossibilité est liée à la définition de la croyance explicite comme la conjonction de la conscience et de la croyance implicite. En effet, si $B_i p \equiv L_i p \wedge A_i p$ on a alors $B_i p \rightarrow L_i p$. Or, l'agent i ne peut croire implicitement une contradiction (axiome D) c'est-à-dire que $L_i p \rightarrow \neg L_i \neg p$. Donc, i ne peut croire explicitement une contradiction (par modus tollens⁶). C'est pour cette raison que Fagin et Halpern ont proposé la logique de raisonnement local. Dans cette logique, un agent peut considérer une pluralité d'ensembles de mondes possibles correspondant chacun à un amas de croyances. Ces amas de croyances constituent les *états d'esprit*⁷ de l'agent. Avec ces considérations, $B_i\phi \wedge B_i\neg\phi$ peut être satisfaite puisqu'un agent peut dans un "état d'esprit" croire ϕ et dans un autre croire $\neg\phi$. Pour plus de détails voir [FH88, p. 58].

⁶Si $\vdash p \rightarrow q, \vdash \neg q$ alors $\vdash \neg p$.

⁷"frames of mind" à la Minsky.

2.4.3 Discussion

La logique épistémique multi-agent permet de représenter et de raisonner sur les croyances des agents. Une telle logique est généralement basée sur la sémantique des mondes possibles. Cette dernière permet de définir les propriétés des croyances des agents d'une manière fort élégante par l'intermédiaire de la théorie de la correspondance. Cependant, les conséquences problématiques qu'engendre ce modèle doivent être prises en considération. Celles qui posent le plus de problèmes sont la fermeture sur l'implication matérielle et la fermeture sur la règle de nécessité. Ces deux fermetures font que l'agent modélisé est doté implicitement de la propriété de rationalité idéale. Cette propriété ne peut être envisageable pour un agent à ressources limitées qui ne peut ni croire toutes les conséquences de ses croyances ni croire toutes les vérités du monde.

En dépit de ces considérations négatives, la logique épistémique multi-agent nous a permis de bien comprendre les croyances d'un point de vue statique. Cela répond aux objectifs qu'on s'est fixés dès le début. Mais, ce qu'on retiendra le plus de cette modélisation formelle est que l'axiome T ne s'applique aux croyances d'un agent rationnel. Si *un agent i croit p* cela n'implique pas que p est vrai. Ce constat nous permet d'énoncer une propriété fondamentale des croyances : *Les croyances sont révisables*. Elles sont révisables dans le sens où si l'agent i croit p et qu'il découvre, d'une manière ou une autre, la fausseté de p , i doit cesser de croire p . Cette phase d'analyse fait partie de la *dynamique des croyances*. La suite de notre étude lui sera totalement consacrée.

3 La dynamique

3.1 Révision des croyances : AGM

Dans AGM, l'état épistémique (les croyances) d'un agent rationnel est représenté par un ensemble d'énoncés fermé sur les conséquences logiques. Cet ensemble, noté K , représente les énoncés acceptés comme croyances. Ces énoncés sont exprimés dans un langage propositionnel standard. Cet ensemble est une théorie partielle. Théorie parce que $K = Cn(K)$ et partielle parce que pour certaines propositions p , il se peut que ni p ni $\neg p$ ne soient dans K . Cn est l'opération des conséquences logiques, $Cn(K) = \{p : K \vdash p\}$. Cn a les propriétés suivantes :

- $K \subseteq Cn(K)$ (inclusion)
- $Cn(K) \subseteq Cn(Cn(K))$ (itération)
- $Cn(K) \subseteq Cn(K')$ si $K \subseteq K'$ (monotonie)

Les attitudes épistémiques par rapport à une proposition p sont :

- $p \in K$: p est acceptée (acceptation);
- $\neg p \in K$: p n'est pas acceptée (non-acceptation);
- $p \notin K$ et $\neg p \notin K$: p est indéterminée (indétermination).

Les changements possibles, changements d'attitude par rapport aux propositions, dûs aux entrées épistémiques (nouvelles informations) peuvent être classés comme suit :

1. *Expansion*

- de l'indétermination à l'acceptation;
- de l'indétermination à la non-acceptation.

2. Révision

- de l'acceptation à la non-acceptation;
- de la non-acceptation à l'acceptation.

3. Contraction

- de l'acceptation à l'indétermination;
- de la non-acceptation à l'indétermination.

3.1.1 Expansion

L'expansion est le moyen le plus simple pour modéliser le changement dû à une nouvelle information. L'attitude épistémique envers une proposition p passe de l'indétermination vers l'acceptation. Formellement, l'expansion, notée $+$, est une fonction qui prend pour arguments un ensemble de croyances K et une proposition p et donne comme résultat un nouvel ensemble de croyances noté K_p^+ .

Postulats de l'expansion

- (+1) K_p^+ est un ensemble de croyances.
Le résultat d'une expansion est un ensemble de croyances.
- (+2) $p \in K_p^+$
Le changement survenu sur K est décrit par le fait que p est acceptée dans K_p^+ .
- (+3) $K \subseteq K_p^+$
En faisant une expansion, on retient tout ce qu'il y avait dans K .
- (+4) Si $p \in K$ alors $K_p^+ = K$
L'acceptation d'une proposition déjà acceptée n'a aucun effet.
- (+5) Si $K \subseteq H$ alors $K_p^+ \subseteq H_p^+$
Si dans l'état épistémique H on connaît plus que dans l'état épistémique K , l'expansion de H et de K par la même proposition maintient la relation d'inclusion.
- (+6) Pour tous les ensembles K et toutes les formules p , K_p^+ est le plus petit ensemble qui satisfait (+1)–(+5).
 K_p^+ ne contient que les croyances requises par les postulats (+1)–(+5). La possibilité d'inclure des croyances qui n'ont aucune relation avec p est exclue.

De ces postulats, il est possible de tirer les conséquences suivantes :

- Si $q \in K_p^+$ alors $K_q^+ \subseteq K_p^+$
[Gär88, conséquence 3.2]
Si $q \notin K$ (le cas où $q \in K$ n'est pas important) et $q \in K_p^+$ alors q est une conséquence de l'expansion de K par p . Par conséquent, si on fait une expansion de K par q , tout ce qui est accepté dans K_q^+ l'est aussi dans K_p^+ .

- $K_p^+ = K_q^+$ ssi $q \in K_p^+$ et $p \in K_q^+$
[Gär88, conséq. 3.3]

Cette propriété est une conséquence directe de la propriété précédente.

En prenant compte la fermeture de l'ensemble des croyances sur les conséquences logiques, d'autres conséquences peuvent encore être déduites :

- Si $\models p \equiv q$ alors $K_p^+ = K_q^+$
[Gär88, conséq. 3.4]
Puisqu'on a affaire au contenu d'un énoncé (proposition) et non à sa formulation linguistique, deux propositions logiquement équivalentes donnent le même résultat par expansion sur le même ensemble de croyances.
- $(K_p^+)_q^+ = K_{p \wedge q}^+$ [Gär88, conséq. 3.5]
Accepter p puis accepter q est équivalent à les accepter toutes les deux simultanément puisque $K_{p \wedge q}^+ = K_{q \wedge p}^+$.
- $(K_p^+)_q^+ = (K_q^+)_p^+$ [Gär88, conséq. 3.6]
Cela vient du fait que $K_{p \wedge q}^+$ est identique à $K_{q \wedge p}^+$. L'ordre des informations n'est pas pertinent pour le processus d'expansion. L'expansion est donc commutative.
- Si $\neg p \in K$ alors $K_p^+ = K_\perp$
[Gär88, conséq. 3.7]
 K_\perp est l'ensemble absurde (l'ensemble de tous les énoncés du langage). Si $\neg p$ est acceptée dans K et si on veut accepter p , l'expansion ne peut être utilisée puisqu'elle mène à une inconsistance (K va contenir p et $\neg p$). En plus, une expansion ne peut lever l'inconsistance de K_\perp du fait que $(K_\perp)_p^+ = K_\perp$ pour tout p .
- $(K \cap H)_p^+ = K_p^+ \cap H_p^+$ [Gär88, conséq. 3.9].

Du fait que (+6) est équivalent à l'exigence $K_p^+ \subseteq Cn(K \cup \{p\})$ et que les cinq autres postulats pris conjointement sont équivalents à $Cn(K \cup \{p\}) \subseteq K_p^+$, l'inclusion peut être réduite à une identité. La fonction d'expansion $+$ définie par cette identité satisfait (+1)–(+6) [Gär88, théorème 3.1]. Par conséquent,

$$K_p^+ = Cn(K \cup \{p\})$$

est une définition explicite du processus d'expansion.

3.1.2 Révision

Souvent, de nouvelles informations viennent contredire des croyances antérieurement acceptées. Dans ce cas, pour pouvoir accepter ces nouvelles croyances, une révision est nécessaire pour le maintien de la consistance de l'ensemble des croyances. Certaines croyances doivent être rejetées pour permettre l'acceptation de la nouvelle croyance d'une manière consistante. La révision est donc non-monotone du fait que l'ajout d'une croyance engendre le rejet d'anciennes croyances. Tout processus de révision doit donner lieu à un changement minimal. Ceci provient du fait que perdre gratuitement de l'information est une action irrationnelle. Formellement, la révision est une fonction, notée \star , qui prend pour arguments un ensemble de croyances et une proposition et donne comme résultat un nouvel ensemble de croyances consistant. D'une manière générale, on peut dire que le problème central de la révision réside dans la détermination des croyances à rejeter.

Postulats de la révision

- (★1) K_p^* est un ensemble de croyances.
Le résultat d'une révision est un état épistémique.
- (★2) $p \in K_p^*$
 p est acceptée dans K_p^* .
- (★3) $K_p^* \subseteq K_p^+$
Si $\neg p \in K$, $K_p^+ = K_\perp$ et donc l'inclusion est vérifiée.
- (★4) Si $\neg p \notin K$ alors $K_p^+ \subseteq K_p^*$
L'expansion n'est autre qu'un type de révision.
- (★5) $K_p^* = K_\perp$ ssi $\models \neg p$
 K_p^* est un ensemble consistant sauf dans le cas où $\neg p$ est logiquement valide.
- (★6) Si $\models p \equiv q$ alors $K_p^* = K_q^*$
Des propositions logiquement équivalentes donnent lieu par révision à des changements identiques.
- (★7) $K_{p \wedge q}^* \subseteq (K_p^*)_q^+$
- (★8) Si $\neg q \notin K_p^*$ alors $(K_p^*)_q^+ \subseteq K_{p \wedge q}^*$
Le changement minimal de K pour l'acceptation de p et q ($K_{p \wedge q}^*$) doit être le même que celui engendré par l'expansion de K_p^* par q tant que q ne contredit pas les croyances de K . Ce postulat est divisé en ces deux axiomes (★7) et (★8) pour des raisons techniques, puisque si $\neg q \in K_p^*$ on a $(K_p^*)_q^+ = K_\perp$.

Les postulats (★1)– (★6) sont dits postulats de base. À partir des postulats de la révision et des postulats de l'expansion, les propriétés suivantes peuvent être déduites :

- si $p \in K$ alors $K_p^* = K$ [Gär88, conséq. 3.11];
- $K_p^* = (K \cap K_p^*)_p^+$ [Gär88, conséq. 3.12];
- $K_p^* \cap K_q^* \subseteq K_{p \vee q}^*$ [Gär88, conséq. 3.14];
- si $\neg q \notin K_{p \vee q}^*$ alors $K_{p \vee q}^* \subseteq K_p^*$ [Gär88, conséq. 3.15];
- $K_{p \vee q}^* = K_p^*$ ou $K_{p \vee q}^* = K_q^*$ ou $K_{p \vee q}^* = K_p^* \cap K_q^*$ [Gär88, conséq. 3.16].

3.1.3 Contraction

Une contraction a lieu quand des croyances sont rejetées sans qu'il y ait d'acceptation de nouvelles croyances. Le principal problème de la contraction provient du fait que lorsqu'on rejette une croyance p d'un ensemble de croyances K il peut y avoir des croyances qui individuellement ou conjointement impliquent p . Puisque K est fermé sur les conséquences logiques, il est nécessaire de rejeter ces croyances aussi. Supposons que p soit acceptée dans K du fait que c'est une conséquence logique de q et de r qui sont tous deux dans K . Dans ce cas, en plus de p , q ou r ou les deux doivent être rejetées aussi. Le problème serait alors de déterminer qui des deux doit être rejeté. Comme

pour la révision, aucune croyance ne doit être rejetée gratuitement. La contraction, notée $-$, est une fonction qui prend pour arguments un ensemble de croyances K et une proposition p et donne comme résultat un nouvel ensemble de croyances K_p^- .

Postulats de la contraction

- (-1) K_p^- est un ensemble de croyances.
 K_p^- découle de K par l'élimination de certaines croyances, K_p^- est donc un ensemble de croyances.
- (-2) $K_p^- \subseteq K$
 K_p^- découle de K par l'élimination de certaines croyances, K_p^- ne peut contenir aucune autre croyance qui n'est pas dans K .
- (-3) Si $p \notin K$ alors $K_p^- = K$
 Quand $p \notin K$, la contraction n'a aucun effet.
- (-4) Si $\not\models p$ alors $p \notin K_p^-$
 Si on fait la contraction de p alors p n'est pas contenu dans K_p^- sauf si p est logiquement valide.
- (-5) Si $p \in K$ alors $K \subseteq (K_p^-)_p^+$
 En contractant p de K puis en faisant une expansion de p on retrouve l'ensemble K . Ce postulat assure le recouvrement.
- (-6) Si $\models p \equiv q$ alors $K_p^- = K_q^-$
 Deux propositions logiquement équivalentes donnent lieu par contraction à des changements identiques.
- (-7) $K_p^- \cap K_q^- \subseteq K_{p \wedge q}^-$
 Les croyances qui sont à la fois dans K_p^- et K_q^- sont aussi dans $K_{p \wedge q}^-$. Pour expliquer ce postulat prenons l'exemple suivant : soit un propriétaire de voiture qui voudrait prendre le train du matin. Pour se rendre à la gare, située à une distance qu'il ne peut faire à pied, il doit utiliser sa voiture. Supposons que son état épistémique contienne les propositions suivantes :
 p : le train du matin viendra à l'heure.
 q : la voiture fonctionnera le matin.
 r : je (le propriétaire) ne manquerai pas le train.
 Si le propriétaire fait la contraction de p , il peut continuer à croire r . Maintenant, s'il fait la contraction de $p \wedge q$ (puisque $p \in K$ et $q \in K$) il doit rejeter p ou q . Du fait qu'il fait plus confiance au train qu'à sa voiture, il rejette q . Dans ce cas, il doit aussi rejeter r puisque si sa voiture ne fonctionne pas il manquera le train. On a donc $r \in K_p^-$, $r \notin K_q^-$ et $r \notin K_{p \wedge q}^-$.
- (-8) Si $p \notin K_{p \wedge q}^-$ alors $K_{p \wedge q}^- \subseteq K_p^-$
 Si on fait la contraction de $p \wedge q$, il faut rejeter p ou q ou les deux doivent être rejetés. Dans le cas où p est rejeté, le changement minimal nécessaire pour la contraction de $p \wedge q$ est relié au changement minimal pour le rejet de p .

Les postulats (-1)-(-6) sont dits postulats de base. À partir des postulats(-1)-(-5) et des postulats de l'expansion, on peut déduire la propriété suivante :

- $K_p^- = K \cap (K_p^-)_\neg p^+$ [Gär88, conséq. 3.22].

3.1.4 Relations entre contraction et révision

Lévi [Lev80] soutient que les seules formes légitimes de changement des croyances sont l'expansion et la contraction. Il montre que la révision peut être analysée comme une séquence de contraction et d'expansion. Pour le cas des ensembles de croyances cette idée peut être énoncée comme suit : l'ensemble des croyances K_p^* résultant de la révision d'un ensemble de croyances K par la proposition p consiste en une contraction de K par $\neg p$ suivie d'une expansion de $K_{\neg p}^-$ par p . D'une façon schématique, on pourrait dire que pour ajouter une proposition p à un ensemble de croyances K contenant $\neg p$, il faut, en quelque sorte, "préparer le terrain" en éliminant d'abord $\neg p$ et toutes les croyances qui l'entraînent logiquement. Cela donne lieu à l'identité suivante :

$$(déf L) \quad K_p^* = (K_{\neg p}^-)_p^+$$

Cette identité est appelée identité de Lévi et permet de définir la révision en termes de contraction et d'expansion. En effet, le changement minimal pour ajouter p à K quand $\neg p \in K$ nécessite la contraction de $\neg p$, $K_{\neg p}^-$ est le changement minimal pour enlever $\neg p$. Et puisqu'il faut, d'une manière ou d'une autre, ajouter p , l'expansion de $K_{\neg p}^-$ par p est le changement minimal nécessaire pour l'ajout de p .

Le problème inverse, c'est-à-dire définir la contraction en fonction de la révision, est traité par Harper [Har77]. L'idée proposée est qu'une proposition q est acceptée dans K_p^- si et seulement si q est acceptée dans K et $K_{\neg p}^*$. Cette analyse s'appuie sur le fait que $K_{\neg p}^*$ est le changement minimal de K pour accepter $\neg p$, $K_{\neg p}^*$ contient le plus possible de croyances de K n'entraînant pas p . Du fait que $K_{\neg p}^*$ contient aussi $\neg p$ et ses conséquences et que K contient p et ses conséquences, les croyances que K et $K_{\neg p}^*$ ont en commun sont celles qui appartiennent à K et n'entraînent pas p . Cela donne lieu à l'identité suivante :

$$(déf H) \quad K_p^- = K \cap K_{\neg p}^*$$

Cette identité, appelée identité d'Harper, permet de définir la contraction à partir de la révision.

L'interchangeabilité est vérifiée dans les deux sens et par conséquent la contraction et la révision sont interchangeables par le biais de l'identité de Lévi et de l'identité d'Harper. Ceci montre qu'en considérant uniquement les postulats de base de la révision et ceux de la contraction, chacun des deux processus peut être défini par l'intermédiaire de l'autre à l'aide de l'identité de Lévi ou l'identité d'Harper. Cela permet aussi, lors de la construction explicite des fonctions de révision et de contraction, de ne considérer que l'une des deux fonctions.

3.1.5 Constructions explicites

Les postulats présentés auparavant définissent les contraintes que toute fonction rationnelle de révision ou de contraction doit satisfaire. Il reste maintenant à construire explicitement ces fonctions qui répondent aux exigences définies par les postulats. L'idée consiste à utiliser $(K \downarrow p)$ (se lit K sans p) l'ensemble de tous les sous-ensembles maximaux de K qui n'entraînent pas logiquement p . En effet, lorsqu'on fait la contraction de K par p , l'ensemble obtenu K_p^- contient, par le principe de la minimalité du changement, le plus possible de ce qu'il y avait dans K et n'entraînait pas p . Par conséquent, l'ensemble des sous-ensembles maximaux de K qui ne contiennent pas p peut

être utilisé pour la construction des fonctions de contraction. Formellement, cet ensemble peut être présenté par :

$$(K \downarrow p) = \{K' \subseteq K \mid K' \not\vdash p \text{ et si } (K' \subseteq K'' \subseteq K) \text{ alors } K'' \vdash p\}.$$

Il est clair que si K est une théorie, les éléments de $(K \downarrow p)$ le sont aussi. Seules les fonctions de contraction sont considérées du fait que la révision peut se définir par la contraction à l'aide de l'identité de Lévi.

Fonction de contraction “choix maximum”

Le but est de définir K_p^- l'ensemble résultant de la contraction d'un ensemble de croyances K et une proposition p . L'idée générale est de se donner un moyen de choisir les propositions à rejeter de K de telle façon que K_p^- ne contienne pas p comme conséquence logique. En respect du critère de minimalité du changement, K_p^- doit être aussi large que possible c'est-à-dire on doit rejeter le minimum de propositions de K . La fonction “choix maximum” consiste à choisir un élément de $(K \downarrow p)$.

Un sous-ensemble de croyances K' est un élément de $(K \downarrow p)$ si et seulement si :

1. K' est un sous-ensemble de K
2. $p \notin K'$
3. $\forall q$ tel que $q \in K$ et $q \notin K'$ on a $(q \rightarrow p) \in K'$

La troisième condition signifie que si on fait une expansion de K' par q on va obtenir p . C'est pour cela que K' ne peut contenir q si $(q \rightarrow p)$ est dans K (donc dans K' aussi, puisque $K' \subseteq K$).

Du fait que K est un ensemble de croyances et en supposant que \vdash est compacte⁸, il existe toujours des sous-ensembles maximaux de K qui n'impliquent pas p sauf si p est valide. Cependant, il existe généralement plusieurs sous-ensembles maximaux. La contraction est déterminée par une fonction de choix C qui choisit un seul élément de $(K \downarrow p)$. Formellement, K_p^- est définie par :

$$K_p^- = \begin{cases} C(K \downarrow p) & \text{si } \not\vdash p \\ K & \text{autrement} \end{cases}$$

K_p^- est un certain sous-ensemble maximal de K qui n'implique pas p . Une fonction de contraction déterminée par une telle fonction de choix est appelée fonction de contraction “choix maximum”. Cette fonction, notée $\overset{m}{-}$, a les propriétés suivantes :

- $\overset{m}{-}$ satisfait les postulats (-1)-(-6) [Gär88, lemme 4.1].
- $\overset{m}{-}$ satisfait la condition $(-F)$ ⁹ :
 $(-F)$: Si $q \in K$ et $q \notin K_p^-$ on a $q \rightarrow p \in K_p^-$ pour tout K .
- Toute fonction de contraction qui satisfait (-1)-(-6) et $(-F)$ peut être générée par $\overset{m}{-}$ [Gär88, théorème 4.2].

⁸si $X \vdash p$ alors $\exists X'$ tel que $X' \subseteq X$ et $X' \vdash p$, X' doit être un ensemble fini.

⁹C'est la troisième condition d'un sous-ensemble maximal.

La fonction $\overset{m}{-}$ ne satisfait pas toujours (-7) et (-8). Pour ce faire, $\overset{m}{-}$ doit être contrainte par une relation d'ordre. La contrainte proposée par Alchourròn et Makinson [AM85] se présente comme suit : une fonction $\overset{m}{-}$ sera dite ordonnée si et seulement si, il existe un ordre partiel \leq_E , tel que $K' \leq K_p^-$ pour toute proposition p et tout $K' \in (K \downarrow p)$. La fonction $\overset{m}{-}$ est ordonnée si K_p^- est un des meilleurs éléments de $(K \downarrow p)$ en accord avec \leq_E . Par le biais de cette contrainte, $\overset{m}{-}$ satisfait (-7) et (-8) [Gär88, lemme 4.3]. C'est peine perdue car $\overset{m}{-}$ présente un grand inconvénient : l'ensemble obtenu est trop grand. Ceci vient du fait qu'un ensemble K est maximal si pour toute proposition q du langage on a $q \in K$ ou $\neg q \in K$. En effet, si p appartient à K et K_p^- est définie par une fonction "choix maximum", pour toute proposition q , on aura, dans K_p^- , $p \vee q$ ou $p \vee \neg q$ [Gär88, lemme 4.5]. Maintenant, en ce qui concerne la révision, il s'avère qu'une fonction de révision définie à partir de la fonction de contraction "choix maximum" à l'aide de l'identité de Lévi est, elle aussi, maximale [Gär88, corollaire 4.6].

Nous pouvons conclure que l'ensemble K_p^* est trop grand puisque les fonctions "choix maximum" donnent comme résultats des ensembles de croyances maximaux. Dans ce cas K_p^* contient q ou $\neg q$ pour toute proposition q en dépit du fait qu'avant la révision K pouvait ne contenir ni q ni $\neg q$. Ce résultat est sans aucun doute contre-intuitif. La fonction "choix maximum" n'est pas une construction acceptable pour une révision rationnelle des croyances.

Fonction de contraction "intersection complète"

La seconde idée pour construire la fonction de contraction serait que K_p^- ne contienne que les propositions communes à tous les éléments de $(K \downarrow p)$. K_p^- est définie comme suit :

$$K_p^- = \begin{cases} \bigcap (K \downarrow p) & \text{si } \not\vdash p \\ K & \text{autrement} \end{cases}$$

Donc, une proposition q appartient à K_p^- si et seulement si q appartient à tout $K' \in (K \downarrow p)$. Cette fonction de contraction est appelée fonction de contraction d'intersection complète. Elle est notée $\overset{f}{-}$ et a les propriétés suivantes :

- $\overset{f}{-}$ satisfait (-1)-(-6) [Gär88, lemme 4.7].
- $\overset{f}{-}$ satisfait aussi la condition d'intersection :
 $(-I)$: pour tout p et q , $K_{p \wedge q}^- = K_p^- \cap K_q^-$.
- Toute fonction de contraction qui satisfait (-1)-(-6) et $(-I)$ peut être générée par $\overset{f}{-}$ [Gär88, Théorème 4.8].

Ces propriétés auraient pu justifier l'utilisation de $\overset{f}{-}$ comme modèle de la fonction de contraction si $\overset{f}{-}$ ne présentait pas un inconvénient de taille qui ne peut être ignoré : $\overset{f}{-}$ donne comme résultat un ensemble trop petit. En effet, Alchouron et Makinson [AM85] ont montré que si K_p^- est défini par $\overset{f}{-}$ alors une proposition q est dans K_p^- si, et seulement si, q est dans K et q est une conséquence logique de $\neg p$ [Gär88, lemme 4.9]. Contrairement à $\overset{m}{-}$, $\overset{f}{-}$ donne comme résultat des ensembles de

croyances trop petits. En rejetant p , on ne retient que les propositions qui sont des conséquences logiques de $\neg p$.

Une fonction de révision définie à partir de $\frac{f}{-}$ par l'identité de Lévi donne aussi comme résultat un ensemble de croyances trop petit. En effet, K_p^* ne va contenir que p et ses conséquences logiques [Gär88, corollaire 4.10].

La fonction de contraction d'"intersection complète" transgresse le critère de minimalité du changement et par conséquent ne peut être envisagée comme représentation du processus de révision. Il n'est pas normal qu'en révisant avec une proposition on finisse par ne croire que cette proposition et ses conséquences logiques.

Fonction de contraction "intersection partielle"

Du fait que pour construire K_p^- , l'utilisation d'un seul élément ou de l'intersection de tous les éléments de $(K \downarrow p)$ donnent respectivement comme résultat un ensemble trop grand ou trop petit, une solution moyenne consiste à utiliser l'intersection de certains éléments de $(K \downarrow p)$. La fonction de contraction ainsi obtenue est dite fonction "d'intersection partielle". Elle est notée $\frac{p}{-}$ et définie comme suit :

$$K_p^- = \begin{cases} \bigcap K & \text{si } \not\vdash p \\ K & \text{autrement} \end{cases}$$

La fonction S choisit les meilleurs éléments de $(K \downarrow p)$. Donc une proposition q appartient à K_p^- si et seulement si q est un élément de tous les sous-ensembles de K choisis. Cette fonction, notée $\frac{p}{-}$, a les propriétés suivantes :

- $\frac{p}{-}$ satisfait les postulats (-1)-(-6) [Gär88, lemme 4.12].
- toute fonction de contraction $-$ qui satisfait (-1)-(-6) est une fonction d'intersection partielle $\frac{p}{-}$ [Gär88, théorème 4.13].

À partir de ces deux propriétés, $\frac{p}{-}$ peut être considérée comme un modèle raisonnable pour la contraction d'un ensemble de croyances.

Il reste maintenant à définir comment déterminer les meilleurs éléments de $(K \downarrow p)$. La solution proposée par Gärdenfors [Gär88] consiste à supposer l'existence d'une relation \sqsubseteq sur tous les sous-ensembles maximaux de K qui peuvent être utilisés pour choisir les éléments de $(K \downarrow p)$ aux rangs les plus élevés. La relation \sqsubseteq est indépendante de la proposition à rejeter. À partir de \sqsubseteq , la fonction de sélection S est définie comme suit :

$$S(K \downarrow p) = \{K' \in (K \downarrow p) \mid \forall K'' \in (K \downarrow p) K'' \sqsubseteq K'\}$$

Une fonction générée à partir d'une telle fonction de sélection est dite fonction de contraction d'"intersection partielle relationnelle" et satisfait (-7) [Gär88, lemme 4.14]. Cette façon de déterminer la fonction de sélection engendre plusieurs conséquences sur les propriétés de la fonction de contraction.

Dans le cas où \sqsubseteq est transitive, la fonction de sélection correspondante est transitivement relationnelle et la fonction de contraction générée par une telle fonction de sélection est dite fonction

de contraction d’’intersection partielle transitivement relationnelle’’. Cette fonction de contraction satisfait tous les postulats de la contraction [Gär88, théorème 4.16].

On peut donc conclure que d’un point de vue théorique la fonction de contraction d’’intersection partielle transitivement relationnelle’’ est la représentation appropriée de la fonction de contraction et par conséquent, de la fonction de révision aussi. Maintenant, si on se préoccupe aussi du côté pratique d’une telle construction, le coût excessif qu’engendre la détermination des sous-ensembles maximaux de K présente une forte raison contre son utilisation [Gär92].

3.1.6 Enracinement épistémique

Le problème principal du changement de croyances réside dans la détermination des croyances à rejeter. Nous avons, aussi, indiqué qu’une forme de mesure est nécessaire pour déterminer les croyances à rejeter en respectant la minimalité du changement. En effet, on n’accorde pas une égale considération à toutes nos croyances : nous considérons certaines d’elles plus importantes que d’autres. Lévi soutient cette thèse en déclarant :

En évaluant les stratégies de contraction, on est amené à faire la distinction, dans le corpus à contracter, entre les items qui sont plus vulnérables à la suppression et ceux qui ne le sont pas. Dans ce sens, on peut parler des différences dans un corpus en respect des degrés de correction. [Lev83, p. 165]

Le corpus et les items dont parle Lévi correspondent respectivement à l’état épistémique et aux croyances.

En ce qui concerne la signification de la mesure, la première idée qui vient à l’esprit est de lier la mesure aux degrés de probabilités des croyances. Cette idée manque de fondements. Lévi présente contre elle une sérieuse objection :

Il serait tentant de faire correspondre les degrés de correction aux degrés de probabilité ou de certitude des croyances. Je soutiens que ce serait une erreur de le faire. Tous les items qui sont dans un corpus initial L à contracter sont certains et infailliblement vrais. Ils ont tous 1 comme probabilité. [Lev83, p. 165]

Il faut donc chercher ailleurs une signification à la mesure. Gärdenfors [Gär88, p. 87–88] avance l’idée selon laquelle cette mesure reflète l’utilité des croyances dans la recherche et la délibération. Il appelle *enracinement épistémique* la mesure ainsi définie. L’enracinement épistémique d’une croyance peut être lié à sa puissance d’explication et à sa valeur informationnelle. Par exemple, les lois sont épistémiquement plus enracinées que les généralisations basées sur quelques observations.

Voyons maintenant les implications de ce concept sur la révision des croyances. La première idée qui vient à l’esprit consiste à considérer que la croyance la plus difficile à rejeter est celle qui est épistémiquement la plus enracinée. Cela rejoint la notion de la vulnérabilité à la suppression de Lévi. Supposons que l’on fasse une contraction de $p \wedge q$ d’un état épistémique K . Cela veut dire que K ne peut contenir à la fois p et q . Le rejet de p est une indication que q est au moins aussi épistémiquement enracinée que p . L’enracinement épistémique est défini comme étant un ordre \leq_E permettant de comparer des propositions. Si p et q sont deux propositions, $p \leq_E q$ signifie “ q est au moins aussi épistémiquement enracinée que p ”.

Étant donné qu’une mesure quantitative des degrés d’enracinement épistémique n’est ni simple ni toujours significative, Gärdenfors propose un ensemble de postulats qui représentent les propriétés qualitatives de l’enracinement épistémique. Ces postulats se définissent comme suit :

- (EE1) : Si $p \leq_E q$ et $q \leq_E r$ alors $p \leq_E r$ (transitivité)
- (EE2) : Si $p \vdash q$ alors $p \leq_E q$ (dominance)
- (EE3) : Pour tout p et q , $p \leq_E p \wedge q$ ou $q \leq_E p \wedge q$ (conjonction)
- (EE4) : Si K est consistant, $p \notin K$ ssi $p \leq_E q$ pour tout q (minimalité)
- (EE5) : Si $q \leq_E p$ pour tout q alors $\models p$ (maximalité)

(EE2) signifie que si p entraîne logiquement q et que si p ou q doit être rejetée de K alors le changement sera minimal en rejetant p et en gardant q du fait que si on rejette q on devra aussi rejeter p . La justification de (EE3) est que si on veut rejeter $p \wedge q$ ceci peut se faire en rejetant p ou q et par conséquent, la perte d'information engendrée par le rejet de $p \wedge q$ est égale à la perte engendrée par le rejet de p ou de q . Une conséquence de (EE1)–(EE3) est que \leq_E est connectée : pour tout p et q , $p \leq_E q$ ou $y \leq_E p$. (EE4) signifie que les propositions qui ne sont pas dans K ont un enracinement épistémique minimal par rapport à K . (EE5) signifie que les propositions valides ont un enracinement épistémique maximal.

Il est à noter que la relation \leq_E est définie par rapport à K . À des ensembles de croyances différents sont associés des ordres différents. Les connections entre l'enracinement épistémique et les fonctions de contraction et de révision présentées précédemment se définissent par deux conditions. La première permet de déterminer l'ordre d'enracinement épistémique à partir d'une fonction de contraction donnée. Cette condition, notée $(C \leq_E)$, est définie comme suit :

- $(C \leq_E)$: $p \leq_E q$ ssi $p \notin K_{p \wedge q}^-$

Cette définition signifie que si on fait une contraction de $p \wedge q$ on doit rejeter p ou q (ou les deux) alors p est rejetée seulement dans le cas où q est au moins aussi enracinée que p .

La deuxième condition permet quant à elle de déterminer explicitement la fonction de contraction à partir de la relation d'enracinement épistémique. Cette condition, notée $(C-)$, est définie comme suit :

- $(C-)$: $q \in K_p^-$ ssi $q \in K$ et $(p <_E p \vee q$ ou $\models p)$.

$(C-)$ signifie qu'une proposition q est dans K_p^- si, et seulement si, q est dans K avant la contraction par p et $p \vee q$ est strictement plus enracinée que p . Dans le cas où q était dans K et p est logiquement valide ($\vdash p$), par le postulat (-4) on a $q \in K_p^-$. La comparaison $p <_E p \vee q$ a été adoptée pour des raisons techniques [MG91, p. 89-90]. En fin de compte, $(C-)$ veut exprimer le fait que si on fait la contraction de K par p , les propositions à retenir dans K_p^- sont celles qui sont les plus enracinées épistémiquement que p . La condition $(C-)$ est donc un moyen approprié pour définir explicitement une fonction de contraction en termes de l'ordre d'enracinement. Gärdenfors montre que si un ordre \leq_E satisfait (EE1)–(EE5) alors la fonction de contraction définie par $(C-)$ satisfait (-1) – (-8) et $(C \leq_E)$. La révision construite par $(C-)$ à l'aide de l'identité de Lévi satisfait (1) – (8) [MG88, théorème4]. Ces résultats permettent de réduire le problème de la construction des fonctions de contraction et de révision à un problème de détermination d'un ordre d'enracinement épistémique approprié.

Il est clair que la comparaison $p <_E p \vee q$ dans $(C-)$ est contre-intuitive. Pour y remédier, Rott [Rot92] propose une version plus naturelle de cette condition qui se définit, maintenant, par :

- $(C-R)$: $y \in K_p^-$ ssi $y \in K$ et $(p <_E q$ ou $\models p)$.

Rott montre que si la relation d'enracinement épistémique satisfait $(EE1)-(EE5)$, la fonction de contraction définie à partir de $(C-R)$ satisfait tous les postulats de la contraction sauf le postulat (-5) . Par conséquent, la contraction définie par (C_R) est différente de celle définie par $(C-)$. Cependant, la révision définie par $(C-R)$ à partir de l'identité de Lévi est identique à la révision définie par $(C-)$. Il est donc beaucoup plus simple d'utiliser $(C-R)$ pour modéliser la fonction de révision que de se compliquer la vie avec la comparaison contre-intuitive de $(C-)$. Même pour la fonction de contraction, il nous semble plus intéressant de sacrifier le postulat (-5) , d'ailleurs très contesté [LR89, Nie89], pour profiter de la simplicité du test de $(C-R)$ que de garder (-5) et s'imposer ainsi les complications de $(C-)$.

3.1.7 Révision des bases de croyances

La révision des croyances telle qu'elle est définie par AGM ne prend pas en considération les problèmes d'implémentation. Elle est plutôt dédiée à la modélisation du processus de révision de croyances d'un agent idéalement rationnel et indépendant du monde réel.

Le plus grand handicap pour une utilisation de cette théorie dans un contexte informatique provient du fait que les ensembles de croyances fermées sur la déduction sont très grands. Une façon de surmonter cet handicap consiste à utiliser des bases de croyances à la place d'ensembles de croyances, c'est ce que propose Nebel [Neb90]. Ces bases sont plus faciles à manipuler puisque ce sont des structures explicitement finies. Une base de croyances B est un sous-ensemble fini de croyances d'un ensemble de croyances K et tel que $Cn(B) = K$. Ceci permet de déterminer des fonctions de révision et de contraction sur des bases au lieu des ensembles. Ces fonctions seront appelées respectivement révision et contraction d'une base de croyances.

Un autre handicap provient, quant à lui, du concept de l'enracinement épistémique et spécialement de l'axiome de dominance $(EE2)$. En effet, cet axiome considère les conséquences d'une croyance plus enracinées qu'elle, cela signifie que pour déterminer laquelle de deux croyances est plus enracinée épistémiquement que l'autre revient à déterminer laquelle des deux entraîne logiquement l'autre. Nebel propose un autre concept appelé *pertinence épistémique*, noté \leq , et qui, à la différence de l'enracinement épistémique, est un ordre arbitraire ne prenant pas en considération les dépendances logiques entre propositions. Dans ce cas, $p \leq q$ a pour signification: " q est au moins aussi pertinente que p ".

L'idée essentielle de la pertinence épistémique est de considérer les croyances contenues explicitement dans une base de croyances B plus importantes que n'importe quelle autre croyance qui n'est pas contenue explicitement dans B . De cette façon, la révision de croyances d'une base revient finalement à minimiser la perte de croyances épistémiquement pertinentes.

Contraction et révision d'une base

La base de croyances résultant de la contraction d'une base de croyances B par p est notée B_p^\sim et définie par :

$$B_p^\sim = \begin{cases} \bigvee_{C \in (B \downarrow p)} C \wedge (B \vee \neg p) & \text{si } \not\vdash p \\ B & \text{autrement} \end{cases}$$

La base de croyances résultant de la révision de B par p est notée B_p^\ddagger et définie en utilisant l'identité de Lévi par :

$$B_p^\ddagger = (B_p^\sim) \wedge p$$

En identifiant K à $Cn(B)$ et K_p^- à $Cn(B_p^\sim)$, la contraction d'une base peut être jugée en termes de postulats de la contraction d'un ensemble de croyances. On obtient le résultat suivant : \sim satisfait (–1)–(–6) [Neb90, lemme 11].

En utilisant une fonction de sélection \mathcal{S}_B définie par :

$$\begin{aligned} \mathcal{S}_B(Cn(B) \downarrow p) &= \{C \in (Cn(B) \downarrow p) \mid \\ \forall C' \in (Cn(B) \downarrow p), C' \cap B \not\subseteq C \cap B\} \end{aligned}$$

Nebel définit une fonction de contraction “intersection partielle” qui, sur des ensembles de croyances, a la propriété d'être identique à \sim .

La connection entre la contraction d'une base et la contraction d'un ensemble se définit de la manière suivante [Neb90, théorème 14] : la contraction d'ensembles finis de prémisses B en utilisant \sim est identique (en respect de \vdash) à une contraction “intersection partielle” $\overset{p}{-}$ définie à partir de la fonction \mathcal{S}_B ,

$$Cn(B)_p^\sim = Cn(B)_p^{\overset{p}{-}}.$$

En introduisant une relation \sqsubseteq_B sur les ensembles de croyances définie par :

$$X \sqsubseteq_B Y \text{ si } X \cap B \not\subseteq Y \cap B$$

Nebel prouve qu'une contraction “intersection partielle” utilisant la fonction de sélection \mathcal{S}_B satisfait (–1)–(–7) [Neb90, théorème 15]. Le postulat (–8) n'est pas satisfait du fait que \sqsubseteq_B n'est pas, généralement, transitive.

En élargissant la portée de la pertinence épistémique \leq d'une simple distinction binaire à un préordre total \leq_ρ sur les propositions de B , \leq_ρ peut être utilisée pour définir \sqsubseteq_ρ , un ordre total sur des sous-ensembles de B :

$$X \sqsubseteq_\rho Y \text{ ssi } \forall p \in (X - Y) \exists q \in (Y - X) : p \leq_\rho q$$

Cet ordre classe le sous-ensemble X moins pertinent que Y juste dans le cas où les éléments les plus pertinents de X sont moins pertinents que les plus pertinents éléments de Y sans considérer les éléments qu'ils ont en commun. Une nouvelle fonction de sélection est définie par Nebel en terme de \sqsubseteq_ρ :

$$\begin{aligned} \mathcal{S}_{B, \leq_\rho}(Cn(B) \downarrow p) &= \{C \in (Cn(B) \downarrow p) \mid \\ \forall C' \in (Cn(B) \downarrow p) C' \cap B \sqsubseteq_\rho C \cap B\} \end{aligned}$$

Ainsi, une contraction “intersection partielle” définie à partir de $\mathcal{S}_{B, \leq_\rho}$ satisfait les postulats (–1)–(–8) [Neb90, théorème 16]. Si \leq_ρ est un ordre linéaire alors chaque sous-ensemble a un élément considéré le plus pertinent et de ce fait \sqsubseteq_ρ se définit par :

$$X \sqsubseteq_\rho Y \text{ ssi } \max(X - Y) \leq_\rho \max(Y - X).$$

Puisque \sqsubseteq_ρ choisit un sous-ensemble unique, le plus pertinent de $(B \downarrow p)$, $\mathcal{S}_{B, \leq_\rho}$ est similaire à une fonction “choix maximum” [Neb90, lemme 17]. Ce résultat permet de conclure que la fonction de contraction “choix maximum” appliquée sur des bases de croyances n'a pas les mêmes effets indésirables que sur des ensembles de croyances. Cette conclusion donne un autre exemple de l'influence du modèle de représentation sur la révision des croyances.

3.1.8 Discussion

Pour AGM, les principes des changements de croyances rationnels ont pour but de ressortir les contraintes logiques influençant la contraction et la révision. Les postulats concernent principalement les relations logiques entre les croyances. Des contraintes d'ordre pragmatique sont aussi prises en considération, mais indirectement. Comme le souligne Doyle [Doy91], le principe de conserver autant de croyances que possible ne tient en aucun lieu à la logique. En effet, le principe de minimalité du changement n'a rien de commun ni avec la consistance ou la complétude des croyances ni avec l'adéquation des inférences qui sont les seules propriétés d'intérêt pour la logique. Gärdenfors [Gär88] souligne que ce principe a aussi des motivations d'ordre économique : les croyances sont d'une grande utilité pour l'action et leurs acquisitions sont coûteuses. Rejeter inutilement une croyance est irrationnel. Cependant, la motivation économique de ce principe n'est pas purement pragmatique car elle n'est pas spécifique à une situation bien particulière.

L'enracinement épistémique est sans aucun doute un concept puissant et d'une grande utilité pour la révision des croyances. Néanmoins, on peut faire plusieurs critiques sur la manière de définir ce concept dans AGM :

- La première concerne le postulat (*EE2*). Ce postulat est, d'une certaine manière, contradictoire avec l'essence même de l'enracinement épistémique tel qu'il a été défini. Si une proposition q est une conséquence logique d'une proposition p , on s'attend naturellement que p soit considérée plus importante d'un point de vue informationnel que q . Dans ce cas, si on devait faire une contraction de $p \wedge q$, légitimement on garderait p et on rejetterait q . C'est tout à fait compatible avec le sens même de l'enracinement épistémique. Donc, normalement, on a $q \leq_E p$. (*EE2*) stipule tout à fait le contraire c'est-à-dire $p \leq_E q$. Gärdenfors lui-même reconnaît, sans toutefois s'y arrêter, cette ambiguïté [Gär88, page 89]. La seule motivation qui justifie (*EE2*) est sa compatibilité avec la minimalité du changement. Comme nous l'avons déjà expliqué, si on rejette q on doit aussi rejeter p sinon on a encore q (puis que $p \vdash q$). Par contre, si on rejette p on peut garder q et donc on fait moins de changements que la première fois.
- La seconde critique concerne l'équivalence de deux croyances en termes d'enracinement épistémique. Il n'y a aucune façon de prendre en considération l'équivalence épistémique de p et q comme un moyen autorisant le rejet d'une seule proposition choisie indifféremment. Si on fait la contraction de $p \wedge q$ on rejette p si on considère $p \leq_E q$ ou q si on considère $q \leq_E p$. Dans le cas où $p \simeq_E q$, on doit rejeter p et q [Gär88, p. 65].
- La troisième critique concerne le rejet de toutes les croyances les moins enracinées épistémiquement. Nous avons vu que si on faisait une contraction d'une proposition p on devait par la condition (*C-*) rejeter toute proposition q moins enracinée épistémiquement que p . Si nous considérons le cas où q n'est reliée d'aucune façon à p , le rejet de q n'a aucun sens. Par exemple, si on est appelé à faire la contraction d'une de nos croyances les plus enracinées épistémiquement, disons "la terre tourne autour du soleil"¹⁰, on doit aussi rejeter toutes les croyances moins enracinées épistémiquement, même celles qui n'ont rien à voir avec le fait que la terre tourne ou ne tourne pas autour du soleil. En quelque sorte, il ne nous resterait pas grand chose à croire.

¹⁰L'exemple est de Lindstrom et Rabinowisc [LR89].

- La quatrième critique concerne le fait que toutes les croyances sont comparables puisque \leq_E est un ordre total. Une conséquence directe de cette considération est que la révision des croyances est un processus déterministe. Cette considération est très forte car elle exige qu'on soit capable de tout comparer et que le tout soit comparable.

Considérons maintenant la pertinence épistémique. Telle que définie par Nebel, elle rend certainement la révision des croyances plus pratique. Néanmoins, son utilisation peut avoir des effets secondaires coûteux [Doy91]. La pertinence épistémique de Nebel classe chaque alternative en fonction de la proposition retenue la plus pertinente sans donner aucune considération aux autres propositions retenues. En appliquant \sqsubseteq_ρ aux éléments de $(B \downarrow p)$, la sélection de l'élément maximal revient finalement à choisir le sous-ensemble qui ne retient pas les propositions les moins pertinentes. Prenons l'exemple donné par Doyle [Doy91], soient $X = \{p_1, \dots, p_{100}\}$ et $Y = \{p_{101}\}$, \leq_ρ range les propositions par l'ordre croissant de l'indice, donc on a $X \sqsubseteq_\rho Y$. La révision $(\{p_1, \dots, p_{101}\} \dot{+} (p_{101} \wedge (p_1 \vee \dots \vee p_{100})) = p_{101})$ choisit entre deux alternatives X et Y , rejette la première et garde la seconde. Maintenant, considérons le cas où les propositions ont approximativement la même valeur mais chacune diffère légèrement des autres. Choisir Y revient à abandonner cent bonnes propositions en faveur d'une seule qui, en fait, n'est que de très peu meilleure que celles qu'on rejette. Ceci n'est certainement pas raisonnable et provient du fait que l'ordre linéaire ne permet pas de faire des comparaisons pondérées c'est-à-dire que même si chacune des cent propositions (p_1, \dots, p_{100}) est individuellement moins pertinente que p_{101} , les cent propositions prises ensemble sont plus pertinentes que p_{101} seule.

3.2 La mise à jour des croyances : KM

La dynamique des croyances ne se résume pas seulement à la révision des croyances. La mise à jour est un autre aspect de la dynamique des croyances. Nous avons vu, précédemment, que la révision des croyances se préoccupe de la manière avec laquelle un agent change ses croyances en adoptant de nouvelles croyances. Cependant, la mise à jour se préoccupe plutôt de la manière qu'un agent devrait changer ses croyances quand il réalise que le monde a changé. En effet, les croyances d'un agent peuvent devenir inexactes à cause des changements intervenus dans le monde. Chaque fois que des événements surviennent ou que d'autres agents réalisent des actions, certains faits deviennent vrais et d'autres faux. Un agent observant de tels processus ou utilisant leurs résultats doit s'assurer que son état épistémique reflète ces changements.

Pour bien illustrer la distinction entre révision et mise à jour, prenons l'exemple suivant : supposons qu'un agent croit qu'une salle contient exclusivement des chaises ou des tables. L'état de croyance de cet agent est $K = \{(\neg c \wedge t) \vee (c \wedge \neg t)\}$. Maintenant, l'agent apprend qu'il n'y a pas de chaises c'est à dire $\neg c$. En révisant, l'état de croyance devient $K' = \{(\neg c \wedge t)\}$ et l'agent peut déduire que la salle contient des tables. Dans une autre interprétation, on peut dire que le fait qu'il n'y a pas de chaises reflète la situation où il y avait des chaises dans la salle et que maintenant il n'en reste plus (quelqu'un les a retirées de la salle). Dans ce cas, l'agent ne peut pas déduire que la salle contient des tables (t). L'état de croyance est dans ce cas $K'' = \{\neg c\}$ et signifie que l'agent croit seulement qu'il n'y a pas de chaises (il peut y avoir des tables comme il peut ne pas y en avoir). Dans le premier cas, l'agent a fait une révision. Dans le second, il a fait une mise à jour.

En s'appuyant sur cette distinction, Katsuno et Mendelzon [KM91a] soutiennent qu'AGM est appropriée pour le changement de croyances dû à des informations sur un monde statique mais ne l'est pas pour la mise à jour. Pour cette dernière, ils proposent de nouveaux postulats et une caractérisation différente. Nous utiliserons l'appellation KM pour faire référence à la théorie proposée

par Katsuno et Mendelzon.

3.2.1 Mise à jour

Les croyances d'un agent sont représentées dans KM par une formule. Cette formule représente la base de croyances d'un agent. L'opérateur de mise à jour est noté Δ . ϕ_μ^Δ représente la base de croyances obtenue par mise à jour de la base de croyances ϕ par la formule μ .

Soit L un langage d'une logique propositionnelle et $Prop$ l'ensemble des propositions primitives de L . Une interprétation de L est une fonction I de $Prop$ vers l'ensemble des valeurs de vérité $\{vrai, faux\}$. Un modèle pour une formule ϕ est une interprétation I tel que $I(\phi) = vrai$. $mod(\phi)$ représente tous les modèles de ϕ . Dans le cas où ϕ est inconsistante $mod(\phi) = \{\}$. ϕ est complète si et seulement si pour toute formule propositionnelle μ , ϕ implique μ ou $\neg\mu$.

Pour la modélisation des croyances, une version des mondes possibles est utilisée. Les opérateurs de mise à jour sont définis par des préordres partiels sur les interprétations. Si on veut mettre à jour une base de croyances ϕ par une formule μ alors l'opérateur de mise à jour Δ sélectionne pour chaque modèle I de ϕ les modèles qui lui sont les plus "proches" parmi tous les modèles de μ . L'ensemble des modèles de la nouvelle base de croyances ϕ_μ^Δ est défini par :

$$mod(\phi_\mu^\Delta) = \bigcup_{I \in mod(\phi)} min(mod(\mu), \leq_I).$$

La proximité entre deux interprétations I et J est mesurée par $diff(I, J)$, l'ensemble des propositions primitives ayant des valeurs différentes dans I et J . Pour toutes interprétations J_1 et J_2 , J_1 est plus "proche" que J_2 de I si, et seulement si, $diff(I, J_1)$ est un sous-ensemble de $diff(I, J_2)$.

En termes de modèles, la révision diffère de la mise à jour. Pour montrer cette différence, assignons à chaque ensemble de croyances K un préordre \leq_K sur l'ensemble \mathcal{I} des interprétations de L ; $<_K$ est le préordre strict correspondant. Un préordre est total, si pour tout J et $J' \in \mathcal{I}$ on a $J \leq J'$ ou $J' \leq J$. \leq_K est persistant si les conditions suivantes sont satisfaites :

- si $J, J' \in mod(K)$ alors $J \not<_K J'$;
- si $J \in mod(K)$ et $J \notin mod(K)$ alors $J <_K J'$.

Un modèle pour un ensemble K est une interprétation I telle que $I(\psi) = vrai$ pour tout $\psi \in K$, $mod(K)$ représente l'ensemble des modèles de K . Soit M un sous-ensemble de \mathcal{I} . Une interprétation $J \in M$ est minimale s'il n'y a pas une interprétation $J' \in M$ telle que $J' <_K J$. Soit $min(M, \leq_K)$ l'ensemble de tous les $J \in M$ tel que J est minimale par rapport à \leq_K . K_μ^* peut être déterminée à partir de \leq_K comme étant l'ensemble de croyances ayant exactement $min(mod(\mu), \leq_K)$ comme modèles. Ceci donne lieu au résultat suivant : une fonction de révision \star satisfait $(\star 1)$ – $(\star 8)$ seulement s'il existe un préordre \leq_K tel que $mod(K_\mu^\star) = min(mod(\mu), \leq_K)$ [KM91a, théorème 5].

Ce résultat donne une représentation de la révision des croyances en termes de modèles utilisée par KM et nous permet de voir que

$$mod(K_\mu^\star) = min(mod(\mu), \leq_K)$$

alors que

$$mod(K_\mu^\Delta) = \bigcup_{I \in mod(K)} min(mod(\mu), \leq_I)$$

La révision de K par μ qui satisfait ($\star 1$)–($\star 8$) choisit parmi les modèles de μ ceux qui sont les plus “proches” de K . La mise à jour de K par μ choisit pour tout modèle I de K , l’ensemble des modèles de μ qui sont les plus proches de I . La mise à jour est alors caractérisée par l’union de tous ces modèles.

Pour bien illustrer cette différence en termes de modèles prenons l’exemple de Winslett [Win88] : supposons que tout ce qu’on sait sur une chambre est qu’elle contient une table et deux objets A et B et qu’un seul des deux est sur la table. Dans ce cas $K = \mathcal{Cn}((p \wedge \neg q) \vee (\neg p \wedge q))$, avec p : “l’objet A est sur la table” et q : “l’objet B est sur la table”. Si on apprend que l’objet A sur la table, on doit réviser K avec p . En accord avec le postulat ($\star 4$) puisque p est consistant avec K , $K_p^- = \mathcal{Cn}(p \wedge \neg q)$. En apprenant p , par révision on conclut que l’objet A est sur la table et l’objet B n’est pas sur la table. La première conclusion est logique du fait qu’on a appris p mais rien nous autorise à faire la seconde. Ce n’est pas parce que l’objet A est sur la table que l’objet B ne serait pas sur la table.

Faisons, maintenant, une mise à jour au lieu de la révision. Dans ce cas, en adoptant le formalisme de KM, on a :

$$\phi \equiv (p \wedge \neg q) \vee (\neg p \wedge q)$$

$$\mu \equiv p$$

$$\text{mod}(\phi) = \{I_1, I_2\} \text{ avec}$$

$$I_1 = \langle \text{vrai}(\text{correspond à } P), \text{faux}(\text{correspond à } q) \rangle \text{ et } I_2 = \langle \text{faux}, \text{vrai} \rangle$$

$$\text{mod}(\mu) = \{J_1, J_2\} \text{ avec } J_1 = \langle \text{vrai}, \text{vrai} \rangle \text{ et } J_2 = \langle \text{vrai}, \text{faux} \rangle \text{ (} p \text{ est vrai et } q \text{ peut être vrai ou faux).}$$

Pour le modèle I_1 , on a $(J_2 \leq_{I_1} J_1)$ puisque $\text{diff}(I_1, J_2) = \{\}$ est un sous-ensemble de $\text{diff}(I_1, J_1) = \{q\}$. Pour le modèle I_2 , on a $(J_1 \leq_{I_2} J_2)$ puisque $\text{diff}(I_2, J_1) = \{p\}$ est un sous-ensemble de $\text{diff}(I_2, J_2) = \{p, q\}$. On obtient donc $\phi_\mu^\Delta \equiv p$ puisque :

$$\begin{aligned} \text{mod}(\phi_\mu^\Delta) &= \bigcup_{I \in \text{mod}(\phi)} \min(\text{mod}(\mu), \leq_I) \\ &= \bigcup (\{p\}, \{\}) = \{p\} \end{aligned}$$

Par la mise à jour on conclut que l’objet A est sur la table et on ne peut rien conclure par rapport à l’objet B . Par rapport à cet exemple on voit bien que la mise à jour et la révision donnent des résultats différents et que la première est plus appropriée que la seconde.

Postulats de la mise à jour

Les postulats représentent les propriétés que devrait avoir tout opérateur de mise à jour Δ .

- ($\Delta 1$) : $\phi_\mu^\Delta \vdash \mu$
- ($\Delta 2$) : Si $\phi \vdash \mu$ alors $\phi_\mu^\Delta \equiv \phi$
- ($\Delta 3$) : Si ϕ et μ sont satisfaites alors ϕ_μ^Δ est aussi satisfaite.
- ($\Delta 4$) : Si $\phi_1 \equiv \phi_2$ et $\mu_1 \equiv \mu_2$ alors $\phi_{1\mu_1}^\Delta \equiv \phi_{2\mu_2}^\Delta$
- ($\Delta 5$) : $\phi_{\mu_1}^\Delta \wedge \mu_2 \vdash \phi_{\mu_1 \wedge \mu_2}^\Delta$
- ($\Delta 6$) : Si $\phi_{\mu_1}^\Delta \vdash \mu_2$ et $\phi_{\mu_2}^\Delta \vdash \mu_1$ alors $\phi_{\mu_1}^\Delta \equiv \phi_{\mu_2}^\Delta$
- ($\Delta 7$) : Si ϕ est complète alors $\phi_{\mu_1}^\Delta \wedge \phi_{\mu_2}^\Delta \vdash \phi_{\mu_1 \vee \mu_2}^\Delta$

- $(\Delta 8): (\phi_1 \vee \phi_2)_{\mu}^{\Delta} \equiv \phi_{1_{\mu}}^{\Delta} \vee \phi_{2_{\mu}}^{\Delta}$

Les postulats $(\Delta 1)$ – $(\Delta 5)$ correspondent directement aux postulats de la révision d'AGM.

Le postulat $(\Delta 2)$ est plus faible que (-2) et signifie que si une formule μ est dérivable d'une formule ϕ alors la mise à jour de ϕ par μ n'a aucun effet. Dans le cas où ϕ est inconsistante, la mise à jour de ϕ par n'importe quelle formule μ donne comme résultat une formule inconsistante. Cette propriété a pour conséquence que si une inconsistance est introduite dans une base de croyances, il n'y a aucun moyen de l'éliminer par une mise à jour. La révision n'a pas ce comportement : le résultat d'une révision est toujours consistant.

Le postulat $(\Delta 6)$ signifie que si la mise à jour de ϕ par μ_1 implique μ_2 et la mise à jour de ϕ par μ_2 implique μ_1 alors les deux mises à jour sont équivalentes.

Le postulat $(\Delta 7)$ s'applique uniquement aux bases de croyances complètes dans lesquelles il n'y a pas d'incertitude sur les mondes possibles et signifie que si le même monde possible résulte de la mise à jour d'une base de croyances complète par μ_1 ou par μ_2 alors il sera aussi le résultat de la mise à jour de la base de croyances par $\mu_1 \vee \mu_2$.

Le postulat $(\Delta 8)$ signifie que le résultat de mise à jour de ϕ_1 ou de ϕ_2 par μ est équivalent à la mise à jour de ϕ_1 par μ ou à celle de ϕ_2 par μ . Ce postulat définit syntaxiquement la condition qui consiste à donner égale considération à chacun des mondes possibles.

3.2.2 Effacement

L'effacement est à la contraction ce que la mise à jour est à la révision. En termes de modèles, *effacer* une formule μ d'une base de croyances ϕ revient à ajouter des modèles à ϕ . Pour chaque modèle I de ϕ , on ajoute les modèles pour lesquels μ est fausse et qui sont les plus "proches" de I . *Contracter* une formule μ d'une base de croyances ϕ revient, par contre, à ajuster la description des mondes possibles à la possibilité que μ soit fausse. ϕ_{μ}^{∇} représente la base de croyances obtenue par l'effacement de μ de la base de croyances initiale ϕ .

Pour bien illustrer la différence entre l'effacement et la contraction, reprenons l'exemple de la chambre avec les deux objets A et B . Soient les propositions p : "l'objet A est sur la table", q : "l'objet B est sur la table" et $\phi = (p \wedge \neg q) \vee (\neg p \wedge q)$. Si on fait la contraction de ϕ par $\mu \equiv p$ (on ne croit plus que l'objet A est sur la table), on obtient : $\phi_{\mu}^{-} \equiv \phi$ puisque $\phi \not\vdash p$ (postulat -3). Par contre si on efface p de ϕ , on obtient :

$$\begin{aligned} \phi_{\mu}^{\nabla} &\equiv (p \wedge \neg q) \vee (\neg p \wedge \neg q) \vee (\neg p \wedge q) \\ &\equiv (p \wedge \neg q) \vee \neg q \end{aligned}$$

À partir de cet exemple on peut faire ressortir une explication intuitive de la différence entre l'effacement et la contraction. La contraction de μ signifie que rien n'a changé dans la chambre puisque ϕ n'implique pas μ . La contraction n'a pas eu d'effets. L'effacement de μ signifie que l'état de la chambre a changé de telle manière que si l'objet A était avant sur la table, il aurait maintenant changé de place d'une façon incertaine. L'effacement n'affecte que les mondes possibles dans lesquels A était sur la table.

Postulats de l'effacement

Les postulats représentent les propriétés que devrait avoir tout opérateur d'effacement ∇ .

- $(\nabla 1) \phi \vdash \phi_{\mu}^{\nabla}$

- ($\nabla 2$) Si $\phi \vdash \neg\mu$ alors $\phi_\mu^\nabla \equiv \phi$
- ($\nabla 3$) Si ϕ est satisfaite et $\not\vdash \mu$ alors $\phi_\mu^\nabla \not\vdash \mu$
- ($\nabla 4$) Si $\phi_1 \equiv \phi_2$ et $\mu_1 \equiv \mu_2$ alors $\phi_{1\mu_1}^\nabla \equiv \phi_{2\mu_2}^\nabla$
- ($\nabla 5$) $\phi_\mu^\nabla \wedge \mu \vdash \phi$
- ($\nabla 6$) $(\phi_1 \vee \phi_2)_\mu^\nabla \equiv \phi_{1\mu}^\nabla \vee \phi_{2\mu}^\nabla$

Nous pouvons constater qu'en termes de postulats il y a deux différences essentielles entre l'effacement et la contraction. La première réside dans le fait qu'une contraction d'une formule μ n'influence pas une base de croyances ϕ si ϕ n'implique pas μ (postulat (-3)) et que l'effacement de μ modifie ϕ si ϕ n'implique pas $\neg\mu$ (postulat ($\nabla 2$)). La seconde est due au postulat ($\nabla 6$). Ce postulat n'est pas requis pour la contraction.

3.2.3 Correspondance entre mise à jour et effacement

L'opérateur d'effacement peut se définir à partir de l'opérateur de mise à jour par :

$$\phi_\mu^\nabla \equiv \phi \vee (\phi_{\neg\mu}^\Delta)$$

L'opérateur de mise à jour peut se définir à partir de l'opérateur d'effacement par :

$$\phi_\mu^\Delta \equiv \phi_\mu^\nabla \wedge \mu$$

Similairement à la révision et à la contraction, l'effacement et la mise à jour sont interchangeables.

3.3 Différence entre révision et mise à jour

En termes de modèles, on peut noter deux différences importantes entre la révision d'AGM et la mise à jour de KM :

- une inconsistance peut être corrigée par une révision en ajoutant de nouvelles informations qui éliminent cette inconsistance ou par contraction en rejetant l'information contradictoire. On ne peut, par contre, remédier à une inconsistance par une mise à jour puisqu'elle spécifie un changement dans le monde [KM91a, lemme 3.1]. S'il n'y a pas de monde qui s'accorde à la description courante, on n'a aucun moyen de refléter le changement dans le monde réel;
- pour la mise à jour, un ordre différent est induit pour chaque modèle de K , alors que pour la révision l'ordre est induit pour tous les modèles de K . Le comportement local de la mise à jour en contraste avec le comportement global de la révision est la différence essentielle entre les deux opérateurs.

3.4 Discussion

Si la révision et la mise à jour sont tous deux des aspects de la dynamique des croyances, leurs objectifs sont, par contre, différents. Le principal objectif de la révision est l'amélioration de la qualité globale des croyances sur une situation statique et imparfaitement appréhendée. L'objectif de la mise à jour est de faire en sorte que les croyances qui décrivent une situation évolutive

reflètent le plus fidèlement possible l'évolution de la situation décrite. Avant d'entamer la dynamique des croyances d'un agent évoluant dans un environnement multi-agent, nous allons présenter une synthèse sommaire de la dynamique telle que nous l'avions présentée dans les précédents chapitres. La dynamique des croyances présente deux facettes : la révision et la mise à jour. Dans la première, un agent révisé ses croyances par rapport à des aspects statiques du monde. Dans la deuxième, il les met à jour en réponse à des changements intervenus dans le monde. Mais, que ce soit dans l'une ou dans l'autre, le changement d'état épistémique, c'est-à-dire la transition d'un état épistémique vers un autre, répond aux exigences suivantes :

- *la préservation de la consistance* : l'état épistémique résultant du changement doit nécessairement être consistant. C'est ce qu'exigent les postulats d'AGM pour la révision et ceux de KM pour la mise à jour.
- *la minimalité du changement* : l'état épistémique résultant du changement doit être aussi peu que possible différent de l'état épistémique précédant le changement. Exprimé autrement, cela signifie que l'état épistémique résultant du changement doit contenir autant qu'il est possible des croyances contenues dans l'état épistémique initial. Dans AGM, cela se manifeste par plusieurs postulats de l'expansion, de la contraction et de la révision, par l'utilisation des sous-ensembles maximaux pour la construction des fonctions de révision et de contraction et par le postulat de l'enracinement épistémique (*EE2*). Dans KM, cela se manifeste dans le choix des modèles présentant un changement minimal au niveau des formules atomiques par rapport aux modèles initiaux.
- *la primauté de l'information nouvelle* : une nouvelle information (entrée épistémique) déclenche automatiquement un changement d'état épistémique. Cette exigence est définie par les postulats (+2), (\star 2) et (-2) pour AGM et par (\diamond 1) et (\diamond 1) pour KM. Cette propriété est appropriée lorsque la nouvelle information reflète un changement dans le monde parce que l'agent concerné doit nécessairement mettre à jour ses croyances pour qu'il ait une image exacte du monde et de son évolution. Par contre, elle ne l'est pas toujours lorsque la nouvelle information porte sur un aspect statique du monde. Les postulats (+2), (\star 2) et (-2) d'AGM doivent être transgressés si on veut considérer des agents autonomes communiquant entre eux [Gal88].

Concentrons nous, maintenant, sur la révision des croyances telle qu'elle est définie par AGM et voyons quelles sont les propriétés des opérations de révision de croyances c'est-à-dire l'expansion, la contraction et la révision? Ces opérations ont trois propriétés principales :

- *la réalisation* : l'expansion, la contraction et la révision sont réalisables. En effet, si une information n'est pas contenue dans l'ensemble de croyances, elle y est simplement incorporée par expansion. Par contre, si elle est incohérente avec une croyance, cette dernière est automatiquement rejetée par contraction. Par révision, non seulement la croyance incohérente est rejetée, mais, en plus, la nouvelle information est incorporée dans l'ensemble des croyances. La propriété de réalisation des opérations de révision de croyances est une conséquence directe de la primauté de l'information nouvelle.
- *l'idempotence* : l'expansion, la contraction et la révision sont idempotentes. Cette propriété est définie par les postulats (+4) pour l'expansion, (-3) pour la contraction et par (+4) et (-3) pour la révision¹¹.

¹¹puisque la fonction de révision se définit par l'identité de Lévi à partir de l'expansion et de la contraction.

- *le déterminisme* : l'état épistémique résultant d'une expansion, d'une contraction ou d'une révision est unique. Cela découle directement du fait que, dans AGM, l'expansion, la contraction et la révision sont considérées comme des fonctions.

4 Aspects multi-agents de la dynamique

La révision des croyances telle qu'elle est définie dans AGM concerne un agent isolé. Cet isolement ne signifie pas l'absence de contact avec le monde car les entrées épistémiques ne sont, en fait, que les résultats de ce contact. L'isolement dont on parle se traduit uniquement par le fait qu'aucun facteur externe, à part les entrées épistémiques, n'est pris en considération dans la révision des croyances : la structure logique de l'état épistémique d'un agent ne dépend pas de ses contacts avec le monde [Gär88, p. 19]. Cela provient, en quelque sorte, du fait qu'aucune typologie n'est proposée par AGM pour rendre compte des différents types d'entrées épistémiques. Dans AGM, ces dernières ne sont définies qu'en vertu des effets qu'elles engendrent sur l'état épistémique. Ceci ne permet pas à l'agent concerné par la révision des croyances de distinguer, par exemple, les informations obtenues par ses propres observations des informations reçues des autres et concernant leurs observations. Un autre point, qui nous paraît important, a été négligé par AGM. Ce point concerne le fait qu'AGM définit les opérations de changement des croyances qui permettent à un agent de réviser ses croyances d'une manière rationnelle mais ne donne aucune réponse concrète à la manière de choisir le type d'opération de changement appropriée à la situation de l'agent. Prenons l'exemple suivant : supposons qu'un agent reçoit une information nouvelle qui est incohérente avec son état épistémique. Dans ce cas, l'agent peut faire une révision ou une contraction. Mais laquelle des deux doit-il faire ? AGM ne donne aucun moyen pour choisir entre les deux. Dans un contexte multi-agent, cette vision n'est pas tout à fait convenable. Les facteurs externes ne peuvent être ignorés et un agent doit non seulement savoir comment réviser ses croyances mais savoir aussi quelle opération de changement de croyances appliquer selon sa situation. Pour ce faire, d'autres aspects de la révision des croyances, non-considérés par AGM, doivent être pris en compte.

Propriétés des agents

Dans un système multi-agents, les agents considérés ont généralement les propriétés suivantes :

- *la sociabilité* : chaque agent interagit avec les autres agents de son environnement. Cette interaction peut avoir plusieurs formes. Celle qui nous importe le plus est la communication. Chaque agent peut envoyer de l'information aux autres agents de son environnement comme il peut en recevoir d'eux. Ces informations peuvent affecter les croyances des agents qui les reçoivent.
- *l'autonomie* : chaque agent a ses propres croyances. Celles-ci définissent la représentation que se fait l'agent de son environnement. En plus d'être privée, cette représentation peut être tout à fait différente de la réalité de l'environnement et des représentations des autres agents. En plus, elle ne cerne pas tous les aspects de l'environnement. En d'autres termes, elle est incomplète et imparfaite, et par conséquent sujette à révision. Cependant, chaque agent a le contrôle total sur ses croyances. Aucun autre agent ne peut intervenir dans le déroulement du processus de révision ni entraver son accomplissement.

Propriétés de la dynamique

Pour des agents sociables et autonomes, nous pensons que la dynamique des croyances doit satisfaire, entre autres, les considérations suivantes :

- *la consistance* : chaque agent doit avoir un état épistémique consistant. La consistance du système multi-agent est considérée à un niveau local. Le système multi-agent est consistant si chacun des agents qui le compose est consistant. Les agents peuvent avoir des points de vue différents par rapport à leur environnement.
- *la minimalité du changement* : en révisant ses croyances, chaque agent fait le minimum de changements possibles. La minimalité du changement est considérée d'un point de vue qualitatif et elle est relative à chaque agent.
- *la persistance* : un agent n'est pas obligé d'intégrer impérativement les nouvelles informations. Il peut persister sur ses attitudes épistémiques initiales. Ceci est valable, spécialement, pour les informations qu'il reçoit des autres agents. La persistance est une conséquence directe de l'autonomie. Les propriétés de la primauté de la nouvelle information et de la réalisation des opérations de changement, définis dans AGM, ne sont plus valables.
- *la considération de la pertinence* : chaque agent considère certaines de ses croyances plus pertinentes que d'autres. La pertinence est relative à chaque agent : ce qui est pertinent pour un agent peut ne pas l'être pour un autre.
- la distinction entre deux principaux types d'entrées :
 - *les entrées internes* : ces entrées correspondent aux résultats obtenus par l'agent de sa propre "expérience". Le contact entre l'agent et son environnement est direct. Par exemple, le résultat d'une perception est une entrée interne.
 - *les entrées externes* : ces entrées correspondent aux informations reçues des autres agents du système. Dans ce cas, le contact entre l'agent et son environnement est indirect : il se fait à travers les autres agents.

Cette distinction est primordiale car elle nous permet de différencier deux types de croyances :

- les croyances privées : croyances acquises par l'agent à partir des entrées internes.
- les croyances communiquées : croyances acquises à partir des entrées externes.
- *la considération de la crédibilité des informateurs* : un agent n'accepte pas toute information provenant de n'importe quel agent de son environnement. Il considère certains d'entre eux plus crédibles que d'autres. Les jugements de crédibilité sont relatifs : chaque agent a ses propres jugements. Un agent i peut considérer un agent j crédible alors qu'un agent k peut le considérer non-crédible.
- L'indexation des croyances par leurs sources qui permet, entre autres, de :
 - *choisir la stratégie de changement de croyances adéquate*. Pour expliquer ce point considérons le cas suivant : un agent i reçoit l'information p de l'agent j considéré crédible en p par i . L'agent i fait une expansion par p . Après cela, i reçoit $\neg p$ de l'agent k . Dans le cas où i considère k plus crédible en p que lui, i révisé par $\neg p$ et par conséquent croit

- $\neg p$ à la place de p . En fait, la comparaison devrait se faire entre j et k et non entre i et k puisque p provient de j . Même si k est plus crédible que i , ce dernier peut persister à croire p tant que j est plus crédible que k . L'agent i ne peut le faire car il n'a aucune indication que p provient de j .
- *propager les effets de la remise en cause de la crédibilité d'un agent*. En effet, l'agent récepteur peut changer d'avis, pour une raison ou une autre, sur la crédibilité d'un autre agent de qui, des informations ont été, auparavant, acceptées. Ainsi, un agent considéré, auparavant, comme crédible peut cesser de l'être. Dans ce cas, les informations acceptées de lui doivent être révisées. Pour ce faire, l'agent récepteur doit être capable de les identifier. La propagation peut aussi avoir lieu dans le cas où l'agent informateur revient, pour une raison ou une autre, sur les informations qu'il a précédemment communiquées.

5 Conclusion

Nous avons étudié, tout au long de cet article, les croyances d'un agent dans ses deux aspects : statique et dynamique.

D'un point de vue statique, nous avons présenté une logique épistémique multi-agent qui permet de représenter et de raisonner sur les croyances des agents. Cette logique est basée sur le modèle des mondes possibles. Ce modèle permet de définir les propriétés des croyances des agents d'une manière simple et concise par l'intermédiaire de la théorie de correspondance. Les plus importantes propriétés des croyances sont :

- les croyances d'un agent doivent être consistantes;
- les croyances d'un agent ne sont pas nécessairement vraies et, par conséquent, sont révisables.

Cependant, le modèle des mondes possibles présente des conséquences problématiques pour la modélisation des croyances des agents. Ces conséquences forment le problème de l'omniscience logique. Mais, en dépit de ces considérations, la logique épistémique multi-agent nous a permis de bien comprendre l'aspect statique des croyances. Cela répond aux objectifs qu'on s'est fixés dès le début concernant la statique des croyances.

La dynamique des croyances, quant à elle, présente deux aspects distincts : la révision et la mise à jour. Le principal objectif de la révision est l'amélioration de la qualité globale des croyances sur une situation statique et imparfaitement appréhendée. L'objectif de la mise à jour est de faire en sorte que les croyances qui décrivent une situation évolutive reflètent le plus fidèlement possible l'évolution de la situation décrite.

La révision consiste à faire l'une des trois opérations : l'expansion, la révision ou la contraction.

Le processus de mise à jour consiste à faire deux opérations : la mise à jour et l'effacement. Pour chacune de ces deux opérations, la théorie KM, développée par Katsuno et Mendelzon, définit un ensemble de postulats et une caractérisation en termes de modèles. L'opération de mise à jour consiste à sélectionner les modèles présentant un changement minimal au niveau des formules atomiques par rapport aux modèles initiaux. L'effacement consiste à ajouter des modèles à la base de croyances.

Dans un environnement multi-agent où les agents sont considérés autonomes et sociables, la dynamique des croyances doit satisfaire de nouvelles considérations non considérées par AGM et KM. En plus de la consistance et de la persistance, viennent s'ajouter, entre autres, la possibilité de la persistance et la considération de la crédibilité des agents informateurs.

Finalement, cet article montre d'une part l'importance de la dynamique, souvent oubliée, dans la modélisation des attitudes mentales et d'une autre part la complexité d'une telle prise en compte.

Références

- [AGM85] C. E. Alchourròn, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50: 510–530, 1985.
- [All84] J. F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2): 123–154, 1984.
- [AM85] C. E. Alchourròn and D. Makinson. On the logic of theory change: Safe contraction. *Studia Logica*, 44: 405–422, 1985.
- [AP80] J. F. Allen and C. R. Perrault. Analysing intentions in dialogues. *Artificial Intelligence*, 15: 143–178, 1980.
- [App85] D. E. Appelt. *Planing English Sentences*. Cambridge University Press, 1985.
- [Bou92] C. Boutilier. *Conditional Logics for Default Reasoning and Belief Revision*. PhD thesis, University of British Columbia, 1992.
- [Bou95] C. Boutilier. Generalized update: Belief change in dynamics settings. In *International Joint Conference on Artificial Intelligence (IJCAI 95)*, volume 2, pages 1550–1556, Montreal, 1995.
- [Bra90] M. E. Bratman. What is intention? In M. E. Pollack P. R. Cohen, J.L. Morgan, editor, *Intentions in Communication*, pages 15–32. MIT Press, 1990.
- [Che80] B. F. Chellas. *Modal Logic*. Cambridge University Press, Cambridge, UK, 1980.
- [CL90] P. R. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42: 213–261, 1990.
- [CP79] P. R. Cohen and C. R. Perrault. Elements of plan-based theory of speech acts. *Cognitive Science*, 3: 177–212, 1979.
- [Dal88] M. Dalal. Investigations into a theory of knowledge base revision: preliminary report. In *Proc. of the 7th National Conf. in Artificial Intelligence*, pages 475–479, 1988.
- [dK86] J. de Kleer. An assumption-based truth maintenance systems. *Artificial Intelligence*, 28: 127–162, 1986.
- [Doy79] J. Doyle. A truth maintenace system. *Artificial Intelligence*, 1979.
- [Doy91] J. Doyle. Rational belief revision: Preliminary report. In E. Sandewall R. Fikes, editor, *Principles of Knowledge Representation and Reasoning*, pages 163–174. Morgan Kaufmann, 1991.

- [Dub86] D. Dubois. Belief structures, possibility theory and decomposable confidence. *Computers and Artificial Intelligence*, 5: 403–416, 1986.
- [FH88] R. Fagin and I. Y. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34: 39–76, 1988.
- [FH94a] N. Friedman and J. Y. Halpern. A knowledge-based framework for belief revision. part i: Foundations. In R. Fagin, editor, *Theoretical Aspects of Reasoning about Knowledge: Proc. Fifth Conference*, pages 44–64, 1994.
- [FH94b] N. Friedman and J. Y. Halpern. A knowledge-based framework for belief revision. part ii: Revision and update. In *Principles of knowledge representation and reasoning: Proc. Fourth International Conference*, 1994.
- [FHMV95] R. Fagin, I. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.
- [FHV92] R. Fagin, I. Y. Halpern, and M. Y. Vardi. What can machines know? on the properties of knowledge in distributed systems. *Journal of the ACM*, 39(2): 328–376, 1992.
- [FR89] N. Foo and A. Rao. Formal theories of belief revision. In *Proc. of the 1st Inter. Conf. on Principles of Knowledge Representation and Reasoning*, pages 369–380, 1989.
- [Gal88] J. R. Galliers. *A theoretical framework for computer models of cooperative dialogue*. PhD thesis, Open University: UK, 1988.
- [Gär88] P. Gärdenfors. *Knowledge in Flux*. The MIT Press, 1988.
- [Gär90] P. Gärdenfors. The dynamics of belief revision: Foundations vs coherence theories. *Revue Internationale de philosophie*, (172): 24–46, 1990.
- [Gär92] P. Gärdenfors. Belief revision: an introduction. In P. Gärdenfors, editor, *Belief Revision*. Cambridge University Press, 1992.
- [GN87] Michael R. Genesereth and Nils J. Nilsson. *Logical foundations of Artificial Intelligence*. Morgan Kaufmann Publishers, INC, Los Altos CA, 1987.
- [Goc92] P. Gochet. Problèmes de logique de connaissance. Montréal, 1992. colloque Intelligence Artificielle et Logiques non classiques: le point de vue philosophique.
- [Gro88] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, (17): 157–170, 1988.
- [Haa86] A. Haas. A syntactic theory of belief and knowledge. *Artificial Intelligence*, 28(3), 1986.
- [Har77] W. L. Harper. Rational conceptual change. *Philosophy of science association*, 2: 462–494, 1977.
- [Hin62] J. Hintikka. *Knowledge and belief*. Cornell University Press, Ithaca, NY, 1962.
- [Hua94] Z. Huang. *Logics For Agents With Bounded Rationality*. PhD thesis, ILLC, Université d’Amsterdam, 1994.

- [KL88] S. Kraus and D. J. Lehmann. Knowledge, belief, and time. *Theoretical Computer Science*, 58: 155–174, 1988.
- [KM91a] H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In *Principles of Knowledge Representation and Reasoning*, pages 387–394, 1991.
- [KM91b] H. Katsuno and A. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3): 263–294, 1991.
- [Kon86] K. Konolige. *A Deduction Model of Belief*. Pitman/Morgan Kaufmann, London, England/Los Altos, CA, 1986.
- [Kri63] S.A. Kripke. Semantical analysis of modal logic. *Acta Philosophica Fennica*, 1963.
- [L⁺94] Y. Lesperance et al. A logical approach to high-level programming – a progress report. In *Control of the Physical World by Intelligent Systems*. 1994.
- [LA87] D. J. Litman and J. F. Allen. A plan recognition model for subdialogues in conversations. *Cognitive Science*, 11(163-200), 1987.
- [LC90] H. J. Levesque and P.R. Cohen. Rational interaction as the basis for communication. In M. E. Pollack P. R. Cohen, J. Morgan, editor, *Intentions in Communications*, pages 221–256. MIT Press, 1990.
- [Les91] Y. Lesperance. *A Formal Theory of Indexical Knowledge and Action*. PhD thesis, University of Toronto, 1991.
- [Lev80] I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, 1980.
- [Lev83] I. Levi. Truth, fallibility and growth of knowledge. *Boston Studies in the Philosophy of Science*, 31: 153–174, 1983.
- [Lev84] H. A. Levesque. A logic of implicit and explicit belief. In *Proceedings AAAI-84*, pages 198–202, 1984.
- [LR89] S. Lindstrom and W. Rabinowicz. Epistemic entrenchment with incomparabilities and relational belief revision. In M. Morreau A. Fuhrmann, editor, *The Logic of Theory Change*, 1989.
- [McC79] J. McCarthy. Ascribing mental qualities to machines. In M. Ringle, editor, *Philosophical Perspectives in Artificial Intelligence*. NJ. Humanities Press, Atlantic Highlands, 1979.
- [MG88] D. Makinson and P. Gärdenfors. Revision of knowledge systems using epistemic entrenchment. In M. Vardi, editor, *Proceedings of the second Conference on Theoretical Aspects of Reasoning about Knowledge*. Los Altos, CA: Morgan Kaufmann, 1988.
- [MG91] D. Makinson and P. Gärdenfors. Revision of knowledge systems using epistemic entrenchment. In *Principles of Knowledge Representation and Reasoning*, pages 83–95. Morgan Kaufmann, 1991.

- [MH91] J. McCarthy and P.J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In M.L.Ginsberg, editor, *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann Publishers, INC, 1991.
- [Moo80] R. Moore. *Reasoning about Knowledge and Action*. PhD thesis, MIT, 1980.
- [MS88] J. P. Martins and S. C. Shapiro. A model for belief revision. *Artificial Intelligence*, 35: 25–79, 1988.
- [Neb89] B. Nebel. A knowledge level analysis of belief revision. In *Proc. of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 301–311, 1989.
- [Neb90] B. Nebel. Lecture notes in computer science. In *Reasoning and Revision in Hybrid Representation Systems*, volume 422. Springer Verlag, Berlin, 1990.
- [Nie89] R. Niederée. Multiple contraction. a further case against gardenfors. In M. Morreau A. Fuhrmann, editor, *The Logic of Theory Change*, 1989.
- [RG91] A. Rao and M. P. Georgeff. Modeling rational agents within a bdi-architecture. In E. Sandewall R. Fikes, editor, *Proceedings of knowledege representation and reasoning*, pages 473–484. Morgan Kaufmann Publishers, 1991.
- [Rot91] H. Rott. Two methods of constructing contractions and revisions of knowledge systems. *Journal of Philosophical Logic*, 21, 1991.
- [Rot92] H. Rott. Preferential belief change using generalized epistemic entrenchment. *Journal of Logic, Language and Information*, 1992.
- [SA91] V. Sperschneider and G. Antoniou. *A Foundation For Computer Science*. Addison-Wesley Publishing Company, 1991.
- [Sho88] Y. Shoham. *Reasoning about change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, 1988.
- [Sho93] Y. Shoham. Agent-oriented programming. *Artificial Intelligence*, 60(1): 51–92, 1993.
- [Sin91] M. P. Singh. Towards a formal theory of communication in multi-agents systems. In *Proceedings of the Second European Joint Conference of Artificial Intelligence*, Sydney, Australia, 1991.
- [Som92] Léa Sombé. Révision de bases de croyances. In *PRC-GDR Intelligence Artificielle*, pages 207–238. 1992.
- [Sta91] R. Stalnaker. The problem of logical omniscience. *Synthèse*, 89: 425–440, 1991.
- [Tho93] S. R. Thomas. *PLACA, an Agent Oriented Programming Language*. PhD thesis, Computer Science department, Standford University, Standford, 1993.
- [WF92] M. Wooldridge and M. Fisher. A first-order branching time logic of multi-agent systems. In *Proceedings of the 10th European Conference on Artificial Intelligence (ECAI'92)*, pages 234–238, 1992.

- [Win88] M. Winslett. Reasoning about action using a possible models approach. In *Proceedings of the seventh National Conference on Artificial Intelligence*, 1988.
- [Woo94] M. Wooldridge. *The Logical Modeling of Computational Multi-Agents Systems*. PhD thesis, The Manchester Metropolitan University, 1994.