Révision de systèmes d'argumentation : changement minimal du statut des arguments

Sylvie Coste-Marquis Sébastien Konieczny Jean-Guy Mailly Pierre Marquis
CRIL
Univ. Artois – CNRS UMR 8188
Lens, France
{coste,konieczny,mailly,marquis}@cril.fr

Résumé

Dans cet article, on étudie le problème de la révision des systèmes d'argumentation à la Dung. On se concentre sur la révision en tant que changement minimal du statut des arguments. Contrairement à la plupart des travaux précédents sur ce sujet, l'ajout de nouveaux arguments n'est pas permis dans le processus de révision, ainsi le système révisé est obtenu uniquement par une modification de la relation d'attaque. On introduit un langage de formules de révision qui est suffisamment expressif pour permettre la représentation de conditions complexes sur l'acceptabilité des arguments dans le système révisé. On montre ensuite comment les postulats de révision du cadre AGM peuvent être adaptés au cas des systèmes d'argumentation. On donne un théorème de représentation correspondant en terme de changement minimal du statut des arguments. Enfin, plusieurs opérateurs de révision basés sur des distances, satisfaisant les postulats, sont mis en évidence.

Abstract

In this paper, we investigate the revision issue for argumentation systems à la Dung. We focus on revision as minimal change of the arguments status. Contrarily to most of the previous works on the topic, the addition of new arguments is not allowed in the revision process, so that the revised system has to be obtained by modifying the attack relation, only. We introduce a language of revision formulae which is expressive enough for enabling the representation of complex conditions on the acceptability of arguments in the revised system. We show how AGM belief revision postulates can be adapted to the case of argumentation systems. We provide a corresponding representation theorem in terms of minimal change of the arguments status. Several distance-based revision operators satisfying the postulates are also pointed out.

1 Introduction

Dans ce papier, on s'intéresse au problème de la révision des systèmes d'argumentation abstraits à la Dung [11]. Les systèmes d'argumentation sont des graphes orientés où les nœuds correspondent aux arguments et les arcs aux attaques entre arguments. Dans de tels systèmes, le statut (c'est-à-dire l'acceptation ou non) d'un argument dépend de la sémantique d'acceptabilité choisie (de base, préférée, stable – entre autres).

Dans son livre [12], Gärdenfors introduit de façon abstraite le changement de croyances comme une opération qui permet de modifier le statut épistémique d'une information selon l'état épistémique d'un agent. Il y a trois statuts possibles : accepté, refusé ou indéfini. La révision, la contraction et l'expansion sont définies comme étant les transitions possibles entre ces statuts, comme indiqué à la figure 1.

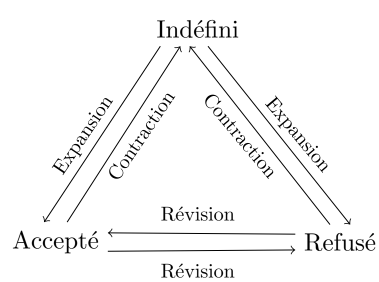


FIGURE 1 – Les transitions épistémiques de Gärdenfors.

Ensuite Gärdenfors instancie cette définition générale dans un cadre logique. Il est intéressant de noter qu'avant de définir les opérateurs de changement de croyances, on doit préciser le cadre logique dans lequel les statuts « accepté », « refusé » et « indéfini » sont définis.

De façon similaire, si on veut instancier la définition générale de changement de croyances, donnée par Gärdenfors, dans le cadre de la théorie de l'argumentation de Dung, il est nécessaire de définir en premier lieu ce que sont les informations disponibles et ce que ces statuts signifient. Dans la théorie de l'argumentation de Dung, les informations de base sont les arguments du système, et leur statut dépend de la sémantique d'acceptabilité prise en compte.

Ainsi, cela n'a aucun sens d'étudier la révision des systèmes d'argumentation directement sur le graphe d'attaque, indépendemment d'une sémantique. Autrement dit, la révision d'un système d'argumentation donné sous deux sémantiques différentes peut conduire à deux résultats différents. Par exemple, dans le cas du système d'argumentation donné à la figure 2, on constate que sous les sémantiques stable et préférée, d appartient à toutes les extensions, tandis que d n'appartient à aucune extension dans le cas de la sémantique de base. Dans ce cas, réviser pour faire accepter d ne nécessite pas de changement pour les deux premières sémantiques, tandis qu'il faut obligatoirement un changement pour la troisième sémantique.

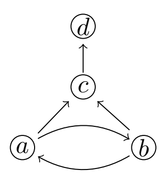


FIGURE 2 – Le changement à apporter pour réviser ce système n'est pas le même selon la sémantique.

Dans cet article, on se concentre sur la révision en tant que changement minimal du statut des arguments. Précisément, sous une sémantique choisie, et étant donné un système d'argumentation et une formule de révision qui exprime de quelle façon le statut de certains arguments doit changer, on souhaite dériver un ou plusieurs systèmes d'argumentation qui satisfont la formule de révision, et qui sont tels que les extensions correspondantes sont aussi proches que possible des extensions du système de départ. Contrairement à la plupart des précédents travaux sur le sujet, l'ajout de nouveaux arguments n'est pas permis dans le processus de révision, ainsi le système révisé est obtenu en modifiant uniquement la relation d'attaque. En particulier, la formule de révision n'indique pas *pourquoi* le statut des arguments a changé.

Le changement minimal du graphe d'attaque peut être considéré comme un critère pour définir les systèmes révisés, mais pas comme le principal. En effet, comme on l'a expliqué ci-dessus, le statut d'acceptation des arguments peut être considéré comme une information plus fondamentale. Dans ce cas, assurer un changement minimal de ces statuts est plus important que (et différent de) la garantie d'un changement minimal du graphe d'attaque.

De tels opérateurs de révision, où le changement minimal porte sur le statut des arguments, peuvent être très utiles pour des applications d'argumentation sur l'opinion publique : supposons qu'un système d'argumentation représente les opinions d'un groupe d'individus, où une attaque entre arguments existe si la majorité du groupe est d'accord avec cette attaque. Si un meneur d'opinion veut modifier le statut de certains arguments, en effectuant une révision sur le système de départ, il obtient dans les systèmes résultant les attaques sur lesquelles faire changer l'opinion de la majorité pour obtenir les statuts d'arguments voulus.

Le reste du papier est organisé comme suit. Après une brève introduction à la théorie de l'argumentation abstraite de Dung, à la section 2, on introduit un langage de formules de révision qui est assez expressif pour permettre la représentation de conditions complexes sur l'acceptabilité des arguments dans le système révision (section 3). À la section 4, on montre comment les postulats de la révision de croyances AGM peuvent être adaptés au cas des systèmes d'argumentation, et on fournit un théorème de représentation correspondant, de façon à garantir le changement minimal sur le statut des arguments. À la section 5, plusieurs opérateurs de révision basés sur des distances, satisfaisant les postulats, sont mis en évidence. Ensuite la section 6 traite de la façon d'associer des systèmes d'argumentation aux ensembles d'extensions obtenus, et discute du problème du changement minimal sur le graphe d'attaque. La section 7 présente quelques travaux connexes, et la section 8 conclut le papier.

2 Préliminaires

On débute par une très courte introduction à la théorie de l'argumentation de Dung (voir [11] pour plus de détails). Un système d'argumentation (fini) est un couple $AF = \langle A, R \rangle$ où A est un ensemble (fini) d'arguments et R est une relation binaire sur A (c'est-à-dire un sous-ensemble de $A \times A$). Dans la suite, A est supposé contenir au moins deux éléments et la relation d'attaque R est supposée irréflexive, c'est-à-dire que les arguments qui s'attaquent eux-mêmes sont rejetés. On note AFs_A l'ensemble de tous les systèmes construits sur l'ensemble d'arguments A.

Un argument $a \in A$ est acceptable selon un ensemble d'arguments $S \subseteq A$ quand il est défendu par cet ensemble, c'est-à-dire quel que soit $b \in A$ tel que $(b, a) \in R$, il existe $c \in S$ tel que $(c, b) \in R$. On dit qu'un sous-ensemble S de A est sans conflit si et seulement si pour tout $a, b \in S$, on a $(a, b) \notin R$. Un sous-ensemble S de A est admissible pour AF si et seulement si S est sans conflit et acceptable selon S. Les « solutions » d'un système d'argumentation sont les ensembles S d'arguments qui peuvent être acceptés conjointement. Plusieurs sémantiques (en particulier, la sémantique complète, la sémantique préférée, la sémantique stable et la sémantique de base) peuvent être utilisées pour capturer formellement cette notion, et chacune d'elles conduit à une notion spécifique d'extensions. Par exemple :

– S est une extension complète de AF si et seulement si c'est un ensemble admissible et chaque argument qui est acceptable selon S appartient à S,
– S est une extension préférée de AF si et seulement si c'est un élément maximal (selon l'inclusion ensembliste) de l'ensemble des ensembles admissibles pour AF,
– S est une extension stable de AF si et seulement si S est sans conflit et $\forall a \in A \setminus S$, $\exists b \in S$ tel que $(b, a) \in R$,
– S est l'(unique) extension de base de AF si et seulement si c'est le plus petit élément (selon l'inclusion ensembliste) de l'ensemble des extensions complètes.

On note $Ext_\sigma(AF)$ l'ensemble des extensions de AF pour la sémantique σ (appelées aussi σ-extensions). Le statut épistémique d'un argument $a \in A$ selon l'état épistémique représenté par AF est donné par : a est accepté si a appartient à tout $\varepsilon \in Ext_\sigma(AF)$, a est rejeté si a n'appartient à aucun $\varepsilon \in Ext_\sigma(AF)$, et a est indéfini dans le cas restant.

On introduit de plus une notation concernant les éléments minimaux d'un ensemble. Tout d'abord, on note $<$ la partie stricte de $\leq$ et $\simeq$ la relation d'indifférence associée. Étant donnés un ensemble E et un pré-ordre $\leq$, on définit les éléments minimaux de E selon ce pré-ordre par $\min(E, \leq) = \{e \in E | \nexists e' \in E, e' < e\}$.

3 Les formules de révision

On veut définir un cadre de révision pour les systèmes d'argumentation de Dung dans lequel des formules de révision sophistiquées peuvent être prises en compte, et pas seulement le fait qu'un argument donné doit être accepté ou refusé. Dans ce but, on considère un langage logique $\mathcal{L}_A$, où la négation est utilisée pour noter le fait qu'un argument donné doit être rejeté, et les formules peuvent être connectées en utilisant la conjonction et la disjonction.

Définition 1. Étant donné $A = \{\alpha_1, \ldots, \alpha_k\}$ un ensemble d'arguments, $\mathcal{L}_A$ est le langage généré par la grammaire hors-contexte en BNF suivante :

$$\begin{array}{lll} arg & ::= & \alpha_1 | \ldots | \alpha_k \\ \Phi & ::= & arg | \neg\Phi | (\Phi \wedge \Phi) | (\Phi \vee \Phi) \end{array}$$

Par exemple, $\varphi_1 = (a \wedge ((\neg b \vee c) \wedge (b \vee \neg c)))$ exprime que dans l'état épistémique révisé on veut que a soit accepté et que b et c soient tous les deux acceptés ou tous les deux refusés. Le sens d'une telle formule φ_1 de $\mathcal{L}_A$ dans un système d'argumentation $AF \in \mathsf{AFs}_A$ pour une sémantique donnée σ est donné par :

Définition 2. Soit $\varepsilon \subseteq A$ et $\varphi \in \mathcal{L}_A$. Le concept de *satisfaction* de φ par ε, noté $\varepsilon \mid\!\sim \varphi$, est défini inductivement comme suit :

– Si $\varphi = a \in A$, alors $\varepsilon \mid\!\sim \varphi$ ssi $a \in \varepsilon$,
– Si $\varphi = (\psi_1 \wedge \psi_2)$, $\varepsilon \mid\!\sim \varphi$ ssi $\varepsilon \mid\!\sim \psi_1$ et $\varepsilon \mid\!\sim \psi_2$,
– Si $\varphi = (\psi_1 \vee \psi_2)$, $\varepsilon \mid\!\sim \varphi$ ssi $\varepsilon \mid\!\sim \psi_1$ ou $\varepsilon \mid\!\sim \psi_2$,
– Si $\varphi = \neg\psi$, $\varepsilon \mid\!\sim \varphi$ ssi $\varepsilon \not\mid\!\sim \psi$.

Ensuite, quel que soit AF dans AFs_A, et pour toute sémantique σ, on dit que :

– φ est *acceptée* pour AF, noté $AF \mid\!\sim_\sigma \varphi$, si $\varepsilon \mid\!\sim \varphi$ pour tout $\varepsilon \in Ext_\sigma(AF)$,
– φ est *refusée* pour AF, noté $AF \mid\!\sim_\sigma \neg\varphi$, si $\varepsilon \mid\!\sim \varphi$ pour aucun $\varepsilon \in Ext_\sigma(AF)$,
– φ est *indéfinie* pour AF dans le cas restant.

À partir de maintenant on appelle *candidat*[1] tout sous-ensemble ε de A. Les candidats peuvent être interprétés comme des modèles ou contre-modèles des formules de révision. Reprenant l'exemple précédent, si $A = \{a, b, c\}$, $\varphi_1 = (a \wedge ((\neg b \vee c) \wedge (b \vee \neg c)))$ est satisfaite par les candidats de $\{\{a\}, \{a, b, c\}\}$. Ainsi, pour la sémantique de base, φ_1 est acceptée pour AF_1 avec $R_1 = \{(b, c), (c, b)\}$ mais est refusée pour AF_2 avec $R_2 = \{(a, b), (b, a)\}$.

De façon assez évidente, quel que soit l'ensemble $M = \{\varepsilon_1, \ldots, \varepsilon_k\}$ de candidats, il existe une formule $\varphi \in \mathcal{L}_A$ telle que $M = \mathcal{A}_\varphi$, où $\mathcal{A}_\varphi = \{\varepsilon \subseteq A \mid \varepsilon \mid\!\sim \varphi\}$ est l'ensemble des candidats qui satisfont φ. Cependant, dans le cas général, $\mathcal{A}_\varphi$ n'est pas l'ensemble de toutes les σ-extensions d'un AF dans AFs_A. Considérons par exemple $A = \{a, b, c\}$ et $\varphi_1 = (a \wedge b \wedge c) \vee (a \wedge \neg b \wedge \neg c)$[2]. $\{a\}$ et $\{a, b, c\}$ sont les deux candidats qui satisfont φ_1, mais il n'y a aucun AF dans AFs_A tel que $Ext_\sigma(AF) = \{\{a\}, \{a, b, c\}\}$ pour $\sigma =$ de base, $\sigma =$ préférée ou $\sigma =$ stable.

Un concept de σ-cohérence peut être défini comme suit :

Définition 3. Une formule $\varphi \in \mathcal{L}_A$ est σ-*cohérente*[3]

1. C'est-à-dire ensemble candidat à être une extension.
2. Équivalente à $(a \wedge ((\neg b \vee c) \wedge (b \vee \neg c)))$.
3. ou simplement *cohérente* si la sémantique est fixée.

ssi il existe un ensemble $\mathcal{S}$ de systèmes d'argumentation AF dans AFs_A tels que $\mathcal{A}_\varphi = \bigcup_{AF \in \mathcal{S}} Ext_\sigma(AF)$.

Quand $A = \{a, b, c\}$, $\varphi_1 = (a \wedge ((\neg b \vee c) \wedge (b \vee \neg c)))$ est base-cohérente, préférée-cohérente et stable-cohérente car $\{\{a\}, \{a, b, c\}\} = Ext_\sigma(AF_3) \cup Ext_\sigma(AF_4)$ où $R_3 = \{(a, b), (a, c)\}$ et $R_4 = \emptyset$. À l'opposé, $\varphi_2 = \neg a \wedge \neg b \wedge \neg c$ est base-cohérente et préférée-cohérente, mais n'est pas stable-cohérente. $\varphi_3 = a \wedge \neg a$ n'est ni base-cohérente, ni préférée-cohérente, mais est stable-cohérent (il suffit de considérer AF_5 tel que $R_5 = \{(a, b), (b, c), (c, a)\}$).

4 Révision de systèmes d'argumentation

Un opérateur de révision de systèmes d'argumentation est une application qui associe un ensemble de systèmes d'argumentation à un système et une formule de révision donnés en entrée :

Définition 4. Étant donné un ensemble d'arguments A, un *opérateur de révision* de systèmes d'argumentation $\star$ est une application de $\mathsf{AFs}_A \times \mathcal{L}_A$ vers 2^{AFs_A}.

Clairement, le résultat de la révision d'un système d'argumentation n'est pas un unique système dans le cas général, mais un ensemble de systèmes d'argumentation. La raison en est assez simple : il peut y avoir plusieurs résultats possibles qui ont exactement la même plausibilité. Dans ce cas, il n'y a pas de raison *a priori* de choisir l'un d'entre eux (on reviendra sur ce point plus loin).

Pour définir des opérateurs de révision, notre approche suit un processus en deux étapes. Intuitivement, la première étape sélectionne parmi les candidats satisfaisant la formule de révision φ ceux qui sont « le plus près » possible des σ-extensions de AF. Ensuite, on engendre des systèmes d'argumentation tels que l'union de leurs σ-extensions coïncide précisément avec les candidats sélectionnés.

Bien sûr, chaque application de $\mathsf{AFs}_A \times \mathcal{L}_A$ vers 2^{AFs_A} n'est pas un opérateur de révision raisonnable. Par exemple, l'opérateur constant défini par $AF \star \varphi = \emptyset$ doit être éliminé. Pour identifier les opérateurs de révision intéressants, on doit identifier les propriétés logiques qui garantissent un comportement rationnel.

Une telle approche axiomatique est standard en logique, et les postulats AGM [1, 13] ont été mis en évidence pour leur caractérisation de « bons » opérateurs de révision dans un cadre logique. Comme dans [14], on peut revisiter ces postulats dans un cadre ensembliste, ici adapté au cas de l'argumentation. Soit $\mathcal{S}$ un ensemble de systèmes d'argumentation AF dans AFs_A. Pour toute sémantique σ, on définit l'ensemble $Ext_\sigma(\mathcal{S})$ des σ-extensions de $\mathcal{S}$ comme $\bigcup_{AF \in \mathcal{S}} Ext_\sigma(AF)$. La contrepartie des postulats AGM dans le cas de l'argumentation est donné par :

(AE1) $Ext_\sigma(AF \star \varphi) \subseteq \mathcal{A}_\varphi$

(AE2) Si $Ext_\sigma(AF) \cap \mathcal{A}_\varphi \neq \emptyset$,
alors $Ext_\sigma(AF \star \varphi) = Ext_\sigma(AF) \cap \mathcal{A}_\varphi$

(AE3) Si φ est σ-cohérente, alors $Ext_\sigma(AF \star \varphi) \neq \emptyset$

(AE4) $Ext_\sigma(AF \star \varphi) \cap \mathcal{A}_\psi \subseteq Ext_\sigma(AF \star (\varphi \wedge \psi))$

(AE5) Si $Ext_\sigma(AF \star \varphi) \cap \mathcal{A}_\psi \neq \emptyset$,
alors $Ext_\sigma(AF \star (\varphi \wedge \psi)) \subseteq Ext_\sigma(AF \star \varphi) \cap \mathcal{A}_\psi$

(AE1) indique que les σ-extensions de l'ensemble de systèmes résultant de la révision doivent appartenir aux candidats qui satisfont φ. **(AE2)** demande que si certaines σ-extensions du système de départ satisfont φ, alors les σ-extensions du résultat doivent coïncider avec elles. **(AE3)** exige que l'ensemble de σ-extensions du résultat soit non vide quand φ est σ-cohérente. Les deux derniers postulats **(AE4)** et **(AE5)** expriment un principe de changement minimal sur le statut des arguments : on s'attend à effectuer le moins de changements possibles sur le statut des arguments du système de départ. Le lecteur familier avec les postulats sur la révision peut remarquer l'absence de traduction du postulat d'indépendance à la syntaxe. En effet, dans le cadre de l'argumentation, la syntaxe est la structure du graphe, donc elle est clairement à prendre en compte.

Il est intéressant de noter que, comme dans le cas logique, on peut dériver un théorème de représentation qui caractérise exactement les opérateurs de révision qui satisfont les postulats. Pour cela, on doit d'abord étendre la notion d'assignement fidèle [13] :

Définition 5. Un *assignement fidèle* est une application qui associe tout système d'argumentation $AF = \langle A, R \rangle$ sous une sémantique σ à un pré-ordre total $\leq^\sigma_{AF}$ sur l'ensemble des candidats tel que :

1. si $\varepsilon_1 \in Ext_\sigma(AF)$ et $\varepsilon_2 \in Ext_\sigma(AF)$, alors $\varepsilon_1 \simeq^\sigma_{AF} \varepsilon_2$,

2. si $\varepsilon_1 \in Ext_\sigma(AF)$ et $\varepsilon_2 \notin Ext_\sigma(AF)$, alors $\varepsilon_1 <^\sigma_{AF} \varepsilon_2$.

Le théorème de représentation peut être énoncé ainsi :

Proposition 1. *Étant donnée une sémantique σ, un opérateur de révision $\star$ satisfait les postulats de rationalité* ***(AE1)*** *-* ***(AE5)*** *ssi il existe un assignement fidèle qui associe tout système $AF = \langle A, R \rangle$ à un pré-ordre total $\leq^\sigma_{AF}$ tel que*

$$Ext_\sigma(AF \star \varphi) = \min(\mathcal{A}_\varphi, \leq^\sigma_{AF})$$

Ce théorème est utile pour définir des opérateurs satisfaisant les postulats de rationalité, tels que ceux présentés à la section suivante.

5 Révision à base de distances

Présentons maintenant des opérateurs de révision, basés sur des distances, satisfaisant les postulats de rationalité **(AE1) - (AE5)**.

Soit d une distance sur 2^A, par exemple la distance de Hamming donnée par $d_H(\varepsilon_1, \varepsilon_2) = |(\varepsilon_1 \setminus \varepsilon_2) \cup (\varepsilon_2 \setminus \varepsilon_1)|$. Pour un $\varepsilon \in 2^A$ donné et $\mathcal{E} \subseteq 2^A$, d peut être étendue en une « distance » entre ε et $\mathcal{E}$, en posant $d(\varepsilon, \mathcal{E}) = \min_{\varepsilon' \in \mathcal{E}} d(\varepsilon, \varepsilon')$. Pour tout système d'argumentation $AF \in \mathsf{AFs}_A$, cette distance induit un pré-ordre total entre candidats $\varepsilon_1, \varepsilon_2 \in 2^A$ donné par

$$\varepsilon_1 \leq^{\sigma,d}_{AF} \varepsilon_2 \text{ ssi } d(\varepsilon_1, Ext_\sigma(AF)) \leq d(\varepsilon_2, Ext_\sigma(AF)).$$

De là, on peut définir des opérateurs de révision :

Définition 6. Soit σ une sémantique. Un *opérateur de révision basé sur une distance* $\star^d$ est un opérateur de révision tel qu'il existe une distance d sur 2^A telle que pour tout AF et pour tout φ, on a $Ext_\sigma(AF \star^d \varphi) = \min(A_\varphi, \leq^{\sigma,d}_{AF})$.

Proposition 2. *Soit σ une sémantique. Tout opérateur de révision basé sur une distance $\star^d$ satisfait les postulats de rationalité **(AE1) - (AE5)**.*

Définissons à présent une autre famille d'opérateurs à base de distance, qui exploite les *labellings*. Rappelons d'abord qu'au lieu d'utiliser les extensions, les solutions d'un système d'argumentation peuvent être exprimées en utilisant le concept de *labelling* [8]. Formellement, un *labelling* est une application L qui associe une étiquette *in*, *undec* ou *out* à chaque argument de l'ensemble A. La notion plus forte de *reinstatement labelling* dépend de la relation d'attaque R : un argument a est étiqueté *in* ssi tout argument attaquant a est *out* ; un argument a est *out* ssi il existe un argument *in* attaquant a ; un argument est *undec* ssi il n'est ni *in* ni *out*. Ces *reinstatement labellings* correspondent aux extensions complètes de Dung de façon bijective. Ainsi, pour toute extension complète ε, le *reinstatement labelling* associé est tel que chaque argument $a \in \varepsilon$ est *in*, tout argument attaqué par un argument *in* est *out*, tout autre argument est *undec*. On introduit la notation $L_\sigma(AF)$ pour désigner l'ensemble des *labellings* associés aux σ-extensions de AF. À l'inverse, pour tout *reinstatement labelling* L, l'extension correspondante $E(L)$ est l'ensemble des arguments étiquetés *in* par ce *labelling*. On introduit quelques notations : L_φ est l'ensemble des *labellings* L tels que $E(L) \in \mathcal{A}_\varphi$. Étant donné un ensemble de *labellings* Lab, $E(Lab) = \{E(L) | L \in Lab\}$.

Les *labellings*, qui portent une information plus riche que les extensions, peuvent être utilisés pour définir des opérateurs à base de distances. En effet, considérons la notion de distance d'édition ci-dessous :

Définition 7. Soient m, n, o trois entiers, soient L_1 et L_2 deux *labellings*, et soient X et Y deux étiquettes parmi *in*, *out* et *undec*.
Une *distance d'édition* $d_{(m,n,o)}$ entre *labellings* est définie comme :

$$d_{(m,n,o)}(L_1, L_2) = \sum_{a \in A} ad(L_1(a), L_2(a))$$

où
- $ad(in, out) = m$
- $ad(in, undec) = n$
- $ad(out, undec) = o$
- $ad(X, Y) = ad(Y, X)$
- $ad(X, X) = 0$

Proposition 3. *Soient m, n, o trois entiers. $d_{(m,n,o)}$ est une distance.*

Il est intéressant de noter que ces distances d'édition ne sont pas nécessairement neutres ou symétriques. On appelle neutre une distance telle que $ad(in, undec) + ad(undec, out) = ad(in, out)$ et symétrique une distance telle que $ad(in, undec) = ad(undec, out)$. Définir des distances non symétriques est un moyen, par exemple, de favoriser l'acceptation d'arguments plutôt que le rejet, en choisissant $ad(in, undec) = 1$, $ad(out, undec) = 9$, et $ad(in, out) = 10$.

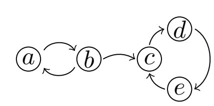
Pour toute distance $d_\mathcal{L}$ entre *labellings*, on peut définir un pré-ordre $\leq^{\sigma,d_\mathcal{L}}_{AF}$ entre *labellings* comme on l'a fait pour les candidats :

$$L_1 \leq^{\sigma,d_\mathcal{L}}_{AF} L_2 \text{ ssi } d_\mathcal{L}(L_1, L_\sigma(AF)) \leq d_\mathcal{L}(L_2, L_\sigma(AF)).$$

Définition 8. Soit σ une sémantique. Un *opérateur de révision basé sur une distance entre labellings* $\star_{d_\mathcal{L}}$ est un opérateur de révision tel qu'il existe une distance $d_\mathcal{L} = d_{(m,n,o)}$ sur 2^A telle que pour tout AF et pour tout φ, on a $Ext_\sigma(AF \star_{d_\mathcal{L}} \varphi) = E(\min(L_\varphi, \leq^{\sigma,d_\mathcal{L}}_{AF}))$.

L'exemple suivant illustre l'impact de la distance choisie sur le système révisé :

Exemple 1. Soit σ la sémantique préférée. On révise le système AF_6 ci-dessous par la formule $\varphi = c \vee \neg d$.



$Ext_\sigma(AF_6) = \{\{a\}, \{b, d\}\}$. Quand on révise AF_6 par φ avec l'opérateur à base de distance induit par la distance $d_{(10,9,1)}$ sur les *labellings*, le résultat est un ensemble de systèmes auxquels on peut associer un unique *labelling* $\{\{(a, in), (b, out), (c, undec), (d, out), (e, undec)\}\}$, qui correspond à l'ensemble d'extensions $\{\{a\}\}$. Quand la distance $d_{(10,1,9)}$ est utilisée, on obtient $\{\{(a, in), (b, out), (c, in), (d, undec), (e, undec)\}\}$, et l'ensemble des extensions correspondantes est à présent $\{\{a, c\}\}$. La première distance,

$d_{(10,9,1)}$, considère qu'il est moins « coûteux » de changer un argument *undec* en *out* qu'en *in*. Une telle distance permet donc de choisir des candidats qui acceptent moins d'arguments. Au contraire, la distance $d_{(10,1,9)}$ permet de choisir des candidats qui acceptent plus d'arguments.
Le choix d'une distance particulière permet donc d'influer sur le contenu des extensions révisées.

Comme les opérateurs à base d'extensions, les opérateurs utilisant les *labellings* exhibent de bonnes propriétés logiques :

Proposition 4. *Soit σ une sémantique. Tout opérateur de révision basé sur une distance entre labellings $\star_{d^{\mathcal{L}}}$ satisfait les postulats de rationalité* ***(AE1)*** *-* ***(AE5)****.*

6 Révision au niveau des systèmes

Les opérateurs définis dans la partie précédente se concentrent sur les candidats qui sont aussi près que possible des extensions du système en entrée. Cependant, ils n'indiquent pas comment engendrer les systèmes d'argumentation correspondants, c'est-à-dire les systèmes d'argumentation tels que l'union de leurs ensembles d'extensions coïncide avec les candidats sélectionnés.

Dans le but d'effectuer cette tâche, on considère une application $\mathcal{AF}_\sigma$ de $2^{2^{\mathcal{A}}}$ vers $2^{\mathsf{AFs}_{\mathcal{A}}}$, appelée opérateur de génération, qui associe à chaque ensemble $\mathcal{C}$ de candidats un ensemble de systèmes d'argumentation tels que $Ext_\sigma(\mathcal{AF}_\sigma(\mathcal{C})) = \mathcal{C}$.

Observons que, quelle que soit la sémantique σ, une telle application $\mathcal{AF}_\sigma$ existe. Cela découle de :

Proposition 5. *Quelle que soit la sémantique σ, pour tout ensemble $\mathcal{C}$ non vide de candidats de 2^A, tel que $\emptyset \notin \mathcal{C}$, il existe un ensemble fini $\mathcal{S} \subseteq \mathsf{AFs}_A$ tel que $\mathcal{C} = \bigcup_{AF \in \mathcal{S}} Ext_\sigma(AF)$.*

Le fait est que chaque candidat C peut être associé à un système d'argumentation AF tel que C en est l'unique σ extension. Par exemple, si $A = \{a, b, c\}$ and $C = \{a, b\}$, AF donné par $R = \{(a, c), (b, c)\}$ remplit ce rôle, quelle que soit la semantique σ. Cette méthode permet de montrer qu'un ensemble de systèmes correspondant aux candidats existe. Cependant, on peut s'intéresser au problème de la génération d'un ensemble minimal de systèmes couvrant l'ensemble de candidats. Une première approche traitant de ce problème est évoquée dans la suite.

De là, des opérateurs de révision peuvent être définis comme suit :

Définition 9. Étant donnés une sémantique σ, un assignement fidèle qui associe tout système d'argumentation à un pré-ordre total $\leq^\sigma_{AF}$, et un opérateur de génération $\mathcal{AF}_\sigma$, l'*opérateur de révision basique* correspondant $\star$ est défini par : $AF \star \varphi = \mathcal{AF}_\sigma(\min(\mathcal{A}_\varphi, \leq^\sigma_{AF}))$.

Un des résultats clés de l'article est :

Proposition 6. *Tout opérateur de révision basique $\star$ satisfait les postulats* ***(AE1)-(AE5)****.*

Les opérateurs basiques traitent seulement du problème du changement du statut des arguments. En effet, les postulats de rationalité demandent la préservation autant que possible du statut des arguments par rapport au système de départ : cette garantie et celle que la formule de révision est satisfaite n'assure pas en général un changement minimal sur la relation d'attaque (et vice-versa).

Pour illustrer ces propos, considérons les systèmes d'argumentation AF_7, AF_8, AF_9, et AF_{10} représentés respectivement par les figures 3.a, 3.b, 3.c, 3.d.

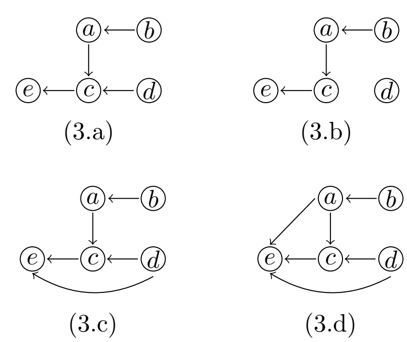


FIGURE 3 – Différents changements minimaux.

Supposons que notre but est de rejeter e, c'est-à-dire d'obtenir un système tel que e n'apparait dans aucune extension. On considère la formule de révision $\varphi = \neg e$. Un changement minimal de la relation d'attaque de AF_7 conduit à AF_8 ou AF_9 : chacun d'eux diffère de AF_7 sur une seule attaque. Ce n'est pas le cas avec AF_{10}, car le changement sur la relation d'attaque pour aller de AF_7 à AF_{10} est strictement plus grand que le changement sur la relation d'attaque nécessaire pour aller de AF_7 à AF_9. Chacun de ces systèmes a une unique extension quelle que soit la sémantique : $\{b, d, e\}$ pour AF_7, $\{b, c, d\}$ pour AF_8, et $\{b, d\}$ pour AF_9 et AF_{10}. Ainsi, le changement du statut des arguments effectué quand on passe de AF_7 à AF_9 où AF_{10} est strictement plus petit que le changement du statut des arguments effectué quand on passe de AF_7 à AF_8.

Un opérateur de génération simple $\mathcal{AF}_\sigma$ se calcule en engendrant les ensembles $\mathcal{S}$ par cardinal croissant

n jusqu'à ce que la condition $\mathcal{C} = \bigcup_{AF \in \mathcal{S}} Ext_\sigma(AF)$ soit satisfaite et ensuite en retournant l'union de tous les $\mathcal{S}$ de taille minimale qui la satisfont. Cependant, utiliser cet opérateur simple n'assure pas d'obtenir un ensemble minimal de systèmes d'argumentation, que l'optimalité soit interprétée comme la minimalité du nombre de systèmes d'argumentation ou comme un changement minimal sur la relation d'attaque.

La minimisation du changement sur la relation d'attaque peut néanmoins être prise en compte dans la définition de l'opérateur de génération (comme un second critère, avec une plus faible priorité que le critère qui concerne le changement minimal du statut des arguments). Cela peut être facilement effectué en considérant une distance dg sur les graphes. Une distance simple est la distance de Hamming, donnée par $dg_H(AF_1, AF_2) = |(\mathcal{R}_1 \setminus \mathcal{R}_2) \cup (\mathcal{R}_2 \setminus \mathcal{R}_1)|$. Mais on peut aussi choisir des distances d'édition plus élaborées telles que celles données dans [10]. La distance dg_H entre deux systèmes d'argumentation correspond au nombre d'attaques qui doivent être ajoutées ou retirées pour les rendre identiques. Chaque distance dg induit un pré-ordre entre systèmes d'argumentation, défini par $AF_1 \leq^{dg}_{AF} AF_2$ ssi $dg(AF_1, AF) \leq dg(AF_2, AF)$. Ce pré-ordre peut être utilisé pour filtrer les systèmes d'argumentation superflus produits par l'opérateur de génération. Cependant, une simple minimisation selon $\leq^{dg}_{AF}$ sur la sortie de l'opérateur de génération n'est pas suffisante en général, comme on le montre dans l'exemple suivant :

Exemple 2. Supposons que les candidats sont $\mathcal{C} = \{C_1, C_2\}$, et que l'opérateur de génération $\mathcal{AF}_\sigma$ appliqué sur $\mathcal{C}$ donne $\mathcal{AF}_\sigma(\mathcal{C}) = \{AF_{11}, AF_{12}\}$ tel que, pour la sémantique considérée σ, $Ext_\sigma(AF_{11}) = \{C_1\}$ et $Ext_\sigma(AF_{12}) = \{C_2\}$. Si AF_{11} et AF_{12} ne sont pas à égale distance du système de départ AF (c'est-à-dire $AF_{11} <_{AF} AF_{12}$ ou $AF_{12} <_{AF} AF_{11}$), un des systèmes n'est pas minimal et doit être retiré. Par conséquent une extension est perdue.

Pour aborder ce problème, on considère des fonctions de sélection qui retirent des systèmes d'argumentation quand ce retrait ne change rien à l'ensemble d'extensions :

Définition 10. Étant donnés une sémantique σ, un ensemble de candidats $\mathcal{C}$, un pré-ordre $\leq_{AF}$ et un opérateur de génération $\mathcal{AF}_\sigma$, on définit la *fonction de sélection de $\mathcal{C}$-couverture minimale* $\gamma_{\mathcal{C}}$ par :
- $\gamma_{\mathcal{C}}(\mathcal{AF}_\sigma(\mathcal{C})) \subseteq \mathcal{AF}_\sigma(\mathcal{C})$
- $Ext_\sigma(\gamma_{\mathcal{C}}(\mathcal{AF}_\sigma(\mathcal{C}))) = Ext_\sigma(\mathcal{AF}_\sigma(\mathcal{C}))$
- $AF \in \mathcal{AF}_\sigma(\mathcal{C})$ et $AF \notin \gamma_{\mathcal{C}}(\mathcal{AF}_\sigma(\mathcal{C}))$ seulement si $\exists AF' \in \mathcal{AF}_\sigma(\mathcal{C})$ tel que $AF' <_{AF} AF$

Basés sur cette notion, les opérateurs de révision complets peuvent être définis comme suit :

Définition 11. Étant donnés une sémantique σ, un assignement fidèle $\leq^{\sigma,d}_{AF}$, un opérateur de génération $\mathcal{AF}_\sigma$, un pré-ordre $\leq^{dg}_{AF}$ sur AFs_A et une fonction de sélection de $\mathcal{C}$-couverture minimale γ, l'*opérateur de révision complet* correspondant $\circ^{\langle \leq^{\sigma,d}_{AF}, \leq^{dg}_{AF}, \gamma \rangle}$ est défini par :

$$AF \circ^{\langle \leq^{\sigma,d}_{AF}, \leq^{dg}_{AF}, \gamma \rangle} \varphi = \gamma_{\min(A_\varphi, \leq^{\sigma,d}_{AF})}(\mathcal{AF}_\sigma(\min(A_\varphi, \leq^{\sigma,d}_{AF})))$$

Clairement les opérateurs de révision complets sont des opérateurs de révision basique.

Comme la minimisation du changement sur la relation d'attaque est faite sans modifier les candidats sélectionnés, on a :

Proposition 7. *Tout opérateur de révision complet $\circ^{\langle \leq^{\sigma,d}_{AF}, \leq^{dg}_{AF}, \gamma \rangle}$ satisfait les postulats **(AE1)-(AE5)**.*

À titre d'illustration, présentons maintenant quelques exemples d'opérateurs de révision. Dans ces exemples, on utilise la fonction de sélection de $\mathcal{C}$-couverture minimale $\gamma_1^{\leq^{dg}_{AF}}$:
- V est une liste ordonnée définie à partir de $\mathcal{AF}_\sigma(\mathcal{C})$ en triant dans l'ordre décroissant [4] selon $\leq^{dg}_{AF}$.
- pour i de 1 à n faire
  - si $\exists V_j \in V$ tel que $V_j <^{dg}_{AF} V_i$ et $Ext_\sigma(V \setminus V_i) = Ext_\sigma(V)$, alors $V := V \setminus V_i$.
- $\gamma_1^{\leq^{dg}_{AF}} := V$.

Avec cette méthode, les systèmes retirés de V sont ceux qui sont les plus éloignés du système en entrée AF.

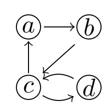


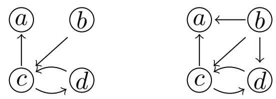
FIGURE 4 – Le système AF_{13}.

Illustrons à présent ces concepts en considérant la révision du système d'argumentation AF_{13} donné à la figure 4, en utilisant l'opérateur de révision complet défini à partir de la distance de Hamming entre candidats, la distance de Hamming entre graphes et la fonction de sélection $\gamma_1^{\leq^{dg}_{AF}}$ définie précédemment.

L'ensemble des extensions de AF_{13} est $\{\{a, d\}\}$ pour la sémantique préférée et la sémantique stable, et son extension de base est $\emptyset$. Donc, quelle que soit la sémantique choisie, on a $AF_{13} \mid\!\sim \neg b$. Supposons à présent que l'on veut que b soit accepté, on effectue alors une révision par $\varphi = b$. Pour la sémantique préférée et la sémantique de base, on construit d'abord l'ensemble des

4. Il existe plusieurs listes ordonnées possibles selon la façon de ranger les éléments égaux pour le pré-ordre. On choisit donc arbitrairement une de ces listes.

candidats qui satisfont b et tels que la distance avec $\{\{a,d\}\}$ est minimale. Cet ensemble est $\{\{a,b,d\}\}$. Ensuite on construit les systèmes d'argumentation qui couvrent ces candidats, en ce concentrant sur ceux qui sont à une distance minimale de AF_{13}. Le résultat est AF_{14}, donné à la figure 5.a.



(5.a) (5.b)

FIGURE 5 – Le résultat de la révision de AF_{13} par b.

On procède de façon similaire pour la sémantique de base : l'ensemble d'extensions $\{\{b\}\}$ est calculé, et ensuite l'unique système, AF_{15}, donné à la figure 5.b est généré.

Un dernier point dont on souhaite discuter est le fait qu'un opérateur de révision basique $\star$ donne en sortie un ensemble de systèmes d'argumentation, et pas un unique système, dans le cas général. En fait, c'est une conséquence de l'expressivité du langage de formules de révision que l'on considère. Pour illustrer ceci, considérons $A = \{a,b,c,d\}$, et AF_{16} donné à la figure 6.

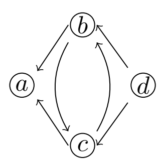
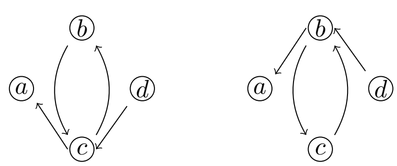


FIGURE 6 – Le système AF_{16}.



(7.a) (7.b)

FIGURE 7 – Revision of AF_{16}.

Les extensions de AF_{16} sont les mêmes pour les sémantiques de base, stable et préférée, $Ext(AF_{16}) = \{\{a,d\}\}$. Soit $\varphi = b \vee c$. Observons que b et c jouent des rôles symétriques, aussi bien dans AF_{16} que dans φ. Quand on calcule le résultat de la révision avec l'opérateur de révision complet basé sur la distance de Hamming entre candidats, la distance de Hamming entre graphes et la fonction de sélection $\gamma_1^{\leq_{AF}^{dg}}$, on obtient deux candidats $\{a,b,d\}$ et $\{a,c,d\}$, et deux systèmes d'argumentation correspondants AF_{17} et AF_{18} (donnés respectivement à la figure 7.a et à la figure 7.b). Choisir un de ces systèmes nécessiterait d'accepter un certain arbitraire étant donnés les rôles symétriques de b et c.

Finalement, quelques mots à propos du problème de l'itération de la révision. Étant donné que le résultat d'un opérateur de révision n'est pas un unique système dans le cas général, mais un ensemble de systèmes, la révision ne peut pas être directement itérée. Néanmoins, il est facile d'étendre les opérateurs de révision de façon à ce qu'un ensemble de systèmes d'argumentation puisse être accepté en entrée. Basiquement, $\mathcal{S} = \{AF_1, \ldots, AF_n\} \star \varphi$ peut être défini comme $\bigcup_{AF_i \in \mathcal{S}} \bigcup_{AF \in AF_i \star \varphi} AF$.

7 Travaux connexes

Différents travaux ont déjà examiné le problème du changement dans les systèmes d'argumentation à la Dung.

Ainsi, Boella, Kaci et Van der Torre [7, 6] ont étudié les principes d'abstraction et de raffinement. Une abstraction est une réduction de la relation d'attaque ou de l'ensemble d'arguments, tandis qu'un raffinement est l'ajout d'attaques ou d'arguments au système. Les auteurs se sont concentrés sur l'étude de sémantiques qui assurent l'existence d'une unique extension (par exemple, l'extension de base), et ont formulé des principes de la forme « si on change le système *de cette façon particulière*, alors l'extension du résultat est *comme ça* ». Ils ont identifié des principes satisfaits par la sémantique de base.

Cayrol, Dupin de Saint-Cyr et Lagasquie-Schiex [9] ont étudié l'ajout d'un argument à un système d'argumentation. Elles ont donné certaines propriétés qui peuvent être satisfaites quand un changement se produit dans un système d'argumentation, et mis en évidence celles satisfaites (et sous quelles conditions) quand un argument (et les attaques qui le concernent) est ajouté au graphe. Avec Bisquert, une étude similaire a été réalisée à propos de la suppression d'un argument [5].

Amgoud et Vesic [2] ont fait une étude similaire dans le cas des systèmes de décision basés sur l'argumentation. La principale différence avec notre travail est la même que pour les travaux précédents : la question qui est posée dans cet article est « que se passe-t-il si on ajoute un argument ? » (et en particulier qu'est-ce que ça change en ce qui concerne le choix d'une option), alors que la révision que nous avons définie n'ajoute pas d'argument, et effectue un processus inverse : on sait ce qu'on veut comme conclusions du (des) nouveau(x) système(s) (c'est la satisfaction de la formule de révision), et on veut savoir comment modifier la relation d'attaque pour y arriver.

Baumann [3] a aussi étudié le problème du changement minimal dans le cadre de l'argumentation abstraite. Il a déterminé des bornes au nombre de modification à apporter à un système d'argumentation pour faire d'un ensemble d'arguments une extension. Ces bornes dépendent de la sémantique et des types de changements autorisés.

Tous ces travaux concernent la modification de la relation d'attaque. Ansi, ils diffèrent d'une façon significative du problème de changement étudié dans cet article, où le changement minimal concerne d'abord le statut des arguments (et ensuite seulement la relation d'attaque).

8 Conclusion

Dans cet article, on s'est intéressé au problème de la révision des systèmes d'argumentation abstraits à la Dung. On s'est concentré sur la révision en tant que changement minimal du statut des arguments. On a introduit un langage de formules de révision qui est assez expressif pour permettre la représentation de conditions complexes sur l'acceptabilité des arguments dans le système révisé. On a montré comment les postulats sur la révision de croyances AGM peuvent être adaptés au cas des systèmes d'argumentation. On a de plus fourni un théorème de représentation correspondant à ces postulats sur le changement minimal du statut des arguments, et mis en évidence plusieurs opérateurs de révision, basés sur des distances, qui satisfont les postulats.

Comme perspective, on envisage dans un premier temps de coder nos opérateurs de révision en représentant les systèmes d'argumentation avec des contraintes logiques (d'une façon similaire aux travaux de Besnard et Doutre [4]), afin de pouvoir bénéficier de la puissance des solveurs de contraintes pour calculer les systèmes révisés. L'étude d'autres opérations de changements sur les systèmes d'argumentation est une autre perspective pour de futures recherches.

De plus, la question de l'association d'un ensemble minimal de systèmes d'argumentation correspondant à un ensemble de candidats présente un intérêt certain, y compris en dehors du cadre de la révision que l'on a évoqué ici.

Références

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change : Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50 :510–530, 1985.

[2] Leila Amgoud and Srdjan Vesic. Revising option status in argument-based decision systems. *Journal of Logic and Computation*, 22(5) :1019–1058, 2012.

[3] Ringo Baumann. What does it take to enforce an argument ? minimal change in abstract argumentation. In *Proceedings of the European Conference on Artificial Intelligence (ECAI'12)*, pages 127–132, 2012.

[4] Philippe Besnard and Sylvie Doutre. Checking the acceptability of a set of arguments. In *Proceedings of the 10th International Workshop on Non-Monotonic Reasoning (NMR'04)*, pages 59–64, 2004.

[5] Pierre Bisquert, Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex. Change in argumentation systems : Exploring the interest of removing an argument. In *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'11)*, volume 6929 of *Lecture Notes in Computer Science*, pages 275–288. Springer, 2011.

[6] Guido Boella, Souhila Kaci, and Leendert van der Torre. Dynamics in argumentation with single extensions : Abstraction principles and the grounded extension. In *Proceedings of the European Conferences on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, volume 5590 of *Lecture Notes in Computer Science*, pages 107–118. Springer, 2009.

[7] Guido Boella, Souhila Kaci, and Leendert van der Torre. Dynamics in argumentation with single extensions : attack refinement and the grounded extension. In *Proceedings of the International Conference on Autonomous Agents and Multiagents Systems (AAMAS'09)*, pages 1213–1214, 2009.

[8] Martin Caminada. On the issue of reinstatement in argumentation. In *Proceedings of the European Conference on Logics in Artificial Intelligence (JELIA'06)*, pages 111–123. Springer, 2006.

[9] Claudette Cayrol, Florence Dupin de Saint-Cyr, and Marie-Christine Lagasquie-Schiex. Change in abstract argumentation frameworks : Adding an argument. *Journal of Artificial Intelligence Research*, 38 :49–84, 2010.

[10] Sylvie Coste-Marquis, Caroline Devred, Sébastien Konieczny, M-C Lagasquie-Schiex, and Pierre Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, 171 :730–753, 2007.

[11] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person

games. *Artificial Intelligence*, 77(2) :321–357, 1995.

[12] Peter Gärdenfors. *Knowledge In Flux*. Cambridge University Press, Cambridge, UK, 1988.

[13] Hirofumi Katsuno and Alberto O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52 :263–294, 1991.

[14] Guilin Qi, Weiru Liu, and David A. Bell. Knowledge base revision in description logics. In *Proceedings of the European Conference on Logics in Artificial Intelligence (JELIA'06)*, volume 4160 of *Lecture Notes in Computer Science*, pages 386–398. Springer, 2006.