

Extraire et modéliser des préférences à partir d'un dialogue

Nicholas Asher¹, Elise Bonzon²

¹ IRIT, Université Paul Sabatier, Toulouse, asher@irit.fr

² CRIP5, Université Paris Descartes, elise.bonzon@mi.parisdescartes.fr

Résumé : Dans de nombreux dialogues, les participants s'engagent sur certaines préférences, puis les développent ou les révisent. Ce travail, encore préliminaire, propose une méthode basée sur la structure discursive du dialogue pour extraire et modéliser ces préférences.

Mots-clés : Représentation de préférences, Dialogues, CP-nets.

1 Introduction

Il est en général bien accepté que les dialogues sont structurés par les actes effectués par les participants. Chaque participant répond aux questions posées par d'autres acteurs du dialogue, puis pose les questions qui en découlent. Ils expliquent leurs annonces, classent les événements temporellement et spatialement. Ils doivent élaborer et défendre leurs revendications. La plupart de ces actes de langage effectués par un agent *A* peuvent affecter la façon dont les autres acteurs du dialogue visualisent les préférences de *A*. Modéliser les préférences d'un agent est important pour pouvoir planifier ses propres coups de conversation, que ce soit dans un cadre coopératif ou stratégique. C'est également important pour de nombreuses applications pratiques qui utilisent des interfaces en langage naturel (le dialogue des systèmes de gestion, etc.). Nous allons modéliser de telles préférences en utilisant des CP-nets, qui permettent d'exploiter les dépendances existant entre les variables au sein d'un dialogue de façon compacte et intuitive. Puis, nous allons élaborer des règles permettant d'utiliser ces coups de dialogue pour créer et transformer ces CP-nets.

Nous illustrons notre approche sur un dialogue similaire à ceux trouvés dans le corpus Verbmobil (Wahlster (2000)) :

- (1)
- π_1 *A* : Shall we meet sometime in the next week ?
 - π_2 *A* : What days are good for you ?
 - π_3 *B* : Well, I have some free time on almost every day except Fridays.
 - π_4 *B* : Fridays are bad.
 - π_5 *B* : In fact, I'm busy on Thursday too.
 - π_6 *A* : Well next week I am out of town Tuesday, Wednesday and Thursday.
 - π_7 *A* : So perhaps Monday ?

La question posée par A en π_1 lui permet d'exprimer sa préférence pour une rencontre la semaine prochaine. Il pose ensuite en π_2 une question élaborative dont la réponse lui permettra de savoir comment satisfaire au mieux ses préférences. B répond à π_2 en π_3 , et explique sa réponse en π_4 . En répondant à la question élaborative de A , B montre qu'il souhaite aussi rencontrer A la semaine prochaine. B précise quelles sont ses préférences pour ce rendez-vous – il préfère du lundi au jeudi. En π_5 , B corrige les préférences sur lesquelles il s'est engagé, en éliminant le jeudi. A doit donc modifier son modèle des préférences de B . A précise ses préférences en π_6 , et propose ensuite une date en π_7 .

2 Dialogue

Une théorie du dialogue doit lier l'interprétation du discours à certains principes généraux de rationalité et de coopérativité (Grice (1975)). Les approches mentalistes considèrent les dialogues comme étant fonction des attitudes des agents, habituellement formalisées par des logiques BDI (e.g., Grosz & Sidner (1990)). Les locuteurs expriment alors leurs croyances, à partir desquelles il est possible de déduire leurs intentions et leurs désirs. Comme soulevé par Hamblin (1987), Asher & Lascarides (2003) soutiennent que les locuteurs peuvent mentir, ou ne pas connaître leurs croyances ou préférences. Lorsqu'un locuteur donne une information sur un état du monde ou sur ses préférences, il s'engage à ce que cette information, ou ces préférences, soient vraies. Nous nous intéressons ici à ces engagements, que nous traitons dans la suite comme étant de simples préférences.

Pour cela, nous partons d'une théorie de la structure et de l'interprétation du discours, la *Segmented Discourse Representation Theory*, ou encore SDRT (Asher (1993); Asher & Lascarides (2003)). Comme de nombreuses autres théories sur l'interprétation du discours (Hobbs *et al.* (1993); Mann & Thompson (1987)), la SDRT construit des structures discursives récursivement, en partant des unités de discours élémentaires (UDE) qui sont liés par des *relations rhétoriques*, comme par exemple *Narration*, *Explication*, *Q-Elab*, etc. La définition formelle d'une SDRS (*Segmented Discourse Representation Structure*) se base sur la description des UDE, sur des prédicats permettant de décrire les relations rhétoriques, ainsi que sur un ensemble d'étiquettes π_i qui correspondent aux éléments des SDRSs.

- Soit ϕ une formule du langage des UDE. ϕ est une formule SDRS.
- Soit π_1, \dots, π_n des étiquettes et soit R une relation rhétorique n -aire. $R(\pi_1, \dots, \pi_n)$ est une formule SDRS.
- Soit ϕ, ϕ' deux formules SDRS. $(\phi \wedge \phi')$, $\neg\phi$ sont des formules SDRS.

Une SDRS est un triplet $\langle \Pi, \mathcal{F}, \text{Last} \rangle$, où Π est un ensemble d'étiquettes; Last est une étiquette de Π (intuitivement, Last représente la dernière clause du discours); et $\mathcal{F} : \Pi \longrightarrow$ formules SDRS.

Une SDRT modélise un dialogue comme étant une suite de paires de SDRSs, chaque paire constituant un tour dans le dialogue (c'est-à-dire un échange minimal entre les participants du dialogue). On appelle une telle structure une DSDRS (*Dialogue Segmented Discourse Representation Structure*). Chaque SDRS dans une DSDRS représente les engagements publics d'un agent. Chaque agent construit une SDRS pour représenter ses propres engagements, et une seconde pour ceux des autres agents.

Tour	SDRS de A	SDRS de B
1	$\pi_{1A} : Q\text{-Elab}(\pi_1, \pi_2)$	\emptyset
2	$\pi_{1A} : Q\text{-Elab}(\pi_1, \pi_2)$	$\pi_{2B} : Q\text{-Elab}(\pi_1, \pi_2) \wedge QAP(\pi_2, \pi)$ $\pi : \text{Correction}(\pi', \pi_5)$ $\pi' : \text{Explication}(\pi_3, \pi_4)$
3	$\pi_{3A} : Q\text{-Elab}(\pi_1, \pi_2) \wedge QAP(\pi_2, \pi) \wedge$ $\text{Elab}(\pi_1, \pi_6) \wedge \text{Elab}(\pi_1, \pi_7) \wedge$ $\text{Résultat}(\pi_6, \pi_7)$	$\pi_{2B} : Q\text{-Elab}(\pi_1, \pi_2) \wedge QAP(\pi_2, \pi)$ $\pi : \text{Correction}(\pi', \pi_5)$ $\pi' : \text{Explication}(\pi_3, \pi_4)$

TAB. 1 – La DSDRS du dialogue (1).

Dans la suite, nous nous sommes limités à l'étude d'un sous-ensemble de relations dans les SDRT : *QAP*, *Correction*, *Q-Elab*, *Elab*, *Explication intentionnelle*, *Commit*, et *Alternatives*. Par exemple, la DSDRS du dialogue (1) est représentée Table 1¹. $\pi :_S \phi$, ou plus simplement $\pi : \phi$ lorsqu'il n'y a pas d'ambiguïté sur la SDRS, signifie que \mathcal{F}_S assigne la formule ϕ à l'étiquette π dans S .

La représentation du dialogue (1) montre qu'à la fin du premier tour, A s'est engagé à s'intéresser aux réponses données à π_1 et π_2 . Nous interprétons π_1 comme étant un engagement de la part de A à ce qu'il préfère une rencontre cette semaine. La sémantique de la relation *Q-Elab*, selon laquelle toute réponse à la question doit préciser π_1 (Asher & Lascarides (2003)), permet d'interpréter la question π_2 par "Quels jours vous conviennent la semaine prochaine?". *Q-Elab* ne modifie pas les préférences de A. Au second tour, B répond à π_2 avec la relation *QAP*. Lascarides & Asher (2008) montre que B s'est ainsi engagé sur la contribution illocutoire de π_2 : B s'engage à vouloir une rencontre au cours de la semaine suivante. Il précise son engagement en répondant à π_2 , et explique sa réponse : il préfère une rencontre tous les jours sauf le vendredi. Il corrige ensuite cette réponse avec π_5 , en précisant qu'il n'est finalement pas disponible le jeudi. Le troisième tour permet à A de proposer une date à partir des réponses de B. La relation *Elab* l'engage publiquement à résoudre la question π_2 (grâce à la sémantique de *QAP*). *Elab* lui permet de préciser ses préférences : il n'est pas disponible du mardi au jeudi. Il suggère ensuite en π_7 une date respectant toutes les contraintes.

3 CP-nets

Nous allons à présent présenter un langage de représentation compacte des préférences : les CP-nets. Ce langage graphique, introduit dans Boutilier *et al.* (2004), utilise l'indépendance préférentielle conditionnelle pour structurer les préférences d'un agent sous l'hypothèse *Ceteris Paribus*.

Nous considérons ici des CP-nets "propositionnels" pour lesquels les variables sont binaires. Nous allons également introduire une relation d'indifférence dans ces CP-nets, qui autorisera un agent à être indifférent entre deux instanciations d'une variable.

¹Lascarides & Asher (2008) explique comment construire cette DSDRS, ce que nous ne détaillons pas ici par manque de place.

Définition 1

Soit V un ensemble de variables propositionnelles, et $\{X, Y, Z\}$ une partition de V . X est **conditionnellement préférentiellement indépendant** de Y selon Z ssi $\forall z \in 2^Z$, $\forall x_1, x_2 \in 2^X$ et $\forall y_1, y_2 \in 2^Y$ on a : $x_1 y_1 z \succeq x_2 y_1 z$ ssi $x_1 y_2 z \succeq x_2 y_2 z$.

Pour une valeur de Z fixée, la relation de préférence sur les instanciations de X est la même quelles que soient les valeurs des instanciations de Y .

Pour chaque variable X , l'agent spécifie un ensemble de *variables parents* $Pa(X)$ qui influent sur ses préférences entre les différentes valeurs de X . Formellement, X est conditionnellement préférentiellement indépendant de $V \setminus (\{X\} \cup Pa(X))$ selon $Pa(X)$. Le CP-net est créé à partir de tables de préférences conditionnelles, qui décrivent les préférences de chaque agent sur les valeurs de la variable X , étant données toutes les combinaisons des valeurs des variables parents.

Définition 2

$\mathcal{N} = \langle \mathcal{G}, \mathcal{T} \rangle$ est un **CP-net** sur V , \mathcal{G} étant un graphe orienté sur V , et \mathcal{T} étant un ensemble de tables de préférences conditionnelles $CPT(X_i)$ pour chaque $X_i \in V$. Chaque $CPT(X_i)$ est associée à un préordre total \succ_p^i sur les valeurs de X_i , selon chaque instanciation $p \in 2^{Pa(X_i)}$.

Informellement, un CP-net \mathcal{N} est satisfait par \succ si \succ satisfait chacune des préférences conditionnelles exprimées dans les CPTs de \mathcal{N} sous l'interprétation *ceteris paribus*.

Chercher l'unique résultat optimal d'un CP-net \mathcal{N} acyclique, consiste intuitivement à parcourir le graphe de haut en bas (c'est-à-dire des parents aux descendants), en instanciant chaque variable à sa valeur préférée selon l'instanciation de ses parents (*forward sweep procedure*, Boutilier et al. (2004)).

Si les CP-nets sont extrêmement intuitifs, ils ne permettent pas de représenter toutes les préférences : il est impossible de représenter des relations de préférences qui comparent 2 à 2 des états n'étant pas identiques "toutes choses étant égales par ailleurs". Pourtant, le pouvoir d'expression de ce langage est suffisant pour les exemples que nous avons étudiés du corpus Verbmobil. La notion de dialogue implique une certaine dépendance entre les variables, qui s'apparente à celle exploitée dans les CP-nets.

4 Du discours aux préférences

Pour raisonner sur les engagements des agents sur leurs préférences, nous devons être capables d'extraire ces préférences à partir de ce que les agents disent, et de ce que leurs actes de langages impliquent. Cette extraction se fait en deux étapes : il faut tout d'abord extraire les préférences à partir des UDES, puis ensuite appliquer des règles qui modifient ces préférences en fonction de la structure discursive. Les SDRTs utilisent une logique de description, la *glue logic* (GL) pour trouver les relations discursives à partir des unités du discours (Asher & Lascarides (2003)). Les formules de GL décrivent les propriétés des étiquettes, et nous permettent de détecter les UDES qui dérivent de clauses de la forme *J'aimerais* ϕ , *Nous voulons* ϕ , etc. Pour spécifier de telles préférences, nous définissons une fonction de transfert F qui transforme des étiquettes en formules de la logique propositionnelle. Si une UDE π d'un agent A est de l'une de ces formes, F

retournera une formule propositionnelle atomique X_π , qui représentera une préférence pour ϕ . Cette variable permettra d'introduire les préférences suivantes dans le CP-net de A : $X_\pi \succ_A \bar{X}_\pi$.²

Il existe différents autres moyens de traduire les préférences. Il est par exemple possible d'utiliser les verbes qui expriment une opinion ou un sentiment. L'utilisation de questions orientées comme³ *Ne devrions nous pas rentrer maintenant ?* exprime la préférence du locuteur pour rentrer avec ses auditeurs. De même, les formules de politesse peuvent souvent être des moyens détournés pour exprimer des préférences.

Les UDES peuvent également spécifier des combinaisons booléennes de préférences. La négation d'une préférence pour ϕ est immédiate. Nous considérerons les préférences disjonctives comme étant non exclusives. Dans le cas contraire, nous introduirons des contraintes permettant d'exclure tout état satisfaisant les deux préférences. Nous pensons que les conjonctions sont ambiguës lorsqu'il s'agit de préférences, mais qu'il est possible de résoudre cette ambiguïté dans certains cas. Une des interprétations possibles pour X et Y est que l'état préféré de l'agent est celui qui satisfait X et Y , mais que cet agent préfère avoir l'une de ces deux variables satisfaite plutôt qu'aucune. On représente ce type de conjonction avec le symbole $\&$. Un agent peut également préférer la "fusion" de X et Y , et ne pas vouloir X ou Y séparément. Nous devons donc spécifier deux formes logiques distinctes pour les UDES du type *Je veux ϕ et ψ* : soit *Je veux ϕ & je veux ψ* , ou *Je veux $\phi \wedge je veux \psi$* . Cette dernière forme logique nous amènera à introduire une co-dépendance entre les variables de la conjonction. Un dernier cas à étudier est celui des questions. Si toutes les questions n'engagent pas les préférences de leur auteur, la majorité le font. En effet, si A demande *peut-on se rencontrer la semaine prochaine ?*, il exprime ses préférences pour une rencontre. Nous pourrions décrire cette question avec $?(rencontre\ moi\ la\ semaine\ prochaine, \pi)$ dans une SRDT. Cette implication est encore plus forte lorsque les questions sont négatives ou ouvertes.

Les axiomes de GL sont les suivants :

1. $F(\pi) = X_\pi$ pour π atomique
2. $\text{Not}(\pi_1, \pi) \rightarrow F(\pi) = \neg F(\pi_1)$
3. $\text{Or}(\pi_1, \pi_2, \pi) \rightarrow F(\pi) = F(\pi_1) \vee F(\pi_2)$
4. $\&(\pi_1, \pi_2, \pi) \rightarrow F(\pi) = F(\pi_1) \& F(\pi_2)$
5. $\wedge(\pi_1, \pi_2, \pi) \rightarrow F(\pi) = F(\pi_1) \wedge F(\pi_2)$
6. $?(\pi_1, \pi) \rightarrow F(\pi) = F(\pi_1)$
7. $?(\neg \pi_1, \pi) \rightarrow F(\pi) = F(\pi_1)$

5 Les règles

Nous pouvons à présent décrire les règles qui nous permettent de modifier les préférences des agents à partir d'une structure discursive. Nous étudions les relations discursives suivantes, qui sont prépondérantes dans le corpus Verbmobil : *QAP* (paire

²GL nous informe également si de telles UDE contiennent une négation qui modifie les préférences.

³Pour une étude des questions orientées et de leur implications, voir Reese (in preparation).

de question-réponse), *Correction*, *Q-Elab* (question élaborative), *Explication intentionnelle* (explication permettant de comprendre pour quelles raisons un agent veut quelque chose), *Commit* (engagement à des préférences), *Alternatives* (questions alternatives) et *Elab* (élaboration).

Chacune de ces règles modifie les tables de préférences conditionnelles d'un ou plusieurs agents. Le CP-net associé représente alors l'engagement de l'agent sur ces préférences. Dans la suite, X dénote un littéral, et $Var(X)$ la variable issue de ce littéral.

Commit : $Commit(\pi)$ représente un engagement de préférences sur les variables de π .

1. $F(\pi) = X$. L'agent préfère que le littéral X soit satisfait. Nous avons $X \succ \bar{X}$.
2. $F(\pi) = X \wedge Y$. L'agent préfère que les littéraux X et Y soient tous deux satisfaits, mais est indifférent s'il ne peut avoir les deux. Nous avons :
 - $Var(X) \in Pa(Var(Y))$ et $X : Y \succ \bar{Y}, \bar{X} : Y \sim \bar{Y}$.
 - $Var(Y) \in Pa(Var(X))$ et $Y : X \succ \bar{X}, \bar{Y} : X \sim \bar{X}$.
3. $F(\pi) = X \& Y$. L'agent préfère avoir les deux littéraux X et Y satisfaits mais est content si au moins l'un des deux l'est. Nous avons : $Y \succ \bar{Y}$ et $X \succ \bar{X}$.
4. $F(\pi) = X \vee Y$. L'agent préfère avoir au moins l'un des deux littéraux X et Y satisfaits. Nous avons :
 - $Var(X) \in Pa(Var(Y))$ et $X : Y \sim \bar{Y}, \bar{X} : Y \succ \bar{Y}$.
 - $Var(Y) \in Pa(Var(X))$ et $Y : X \sim \bar{X}, \bar{Y} : X \succ \bar{X}$.
5. $F(\pi) = \Phi$. Nous pouvons appliquer les règles 1-4 en décomposant Φ .

Par manque de place, nous ne décrivons pas dans la suite les règles pour les cas \vee , qui n'apparaissent pas dans notre exemple illustratif.

Notre second ensemble de règles traite les *Explications intentionnelles* ou *Iexplications*.

- (2) π_1 Je veux aller au supermarché
 π_2 pour manger quelque chose

Dans cet exemple, l'agent veut manger quelque chose, et c'est pourquoi il veut aller au supermarché. π_1 dépend donc causalement de π_2 .

Explication intentionnelle : Dans $Iexplication(\pi_i, \pi_j)$, π_j sont les préférences permettant d'expliquer les préférences exprimées dans π_i .

6. $F(\pi_i) = X, F(\pi_j) = Y$. Cela signifie que l'agent explique ses préférences sur X par Y . On sait donc que Y est préférée à $\neg Y$, et que si aucune préférence sur X n'est déjà définie, Y est une raison de vouloir X . Dans le cas contraire, l'agent préfère X si Y est vrai, mais ne modifie pas ses préférences sinon.
 - $Y \succ \bar{Y}, Var(Y) \in Pa(Var(X))$,
 - si $Pa(Var(X)) = \{Var(Y)\}$ alors $Y : X \succ \bar{X}, \bar{Y} : \bar{X} \sim X$. Sinon, si \succ_X représente la relation de préférences associée à $CPT(X)$, nous avons $\succ_{X,Y} = X \succ \bar{X}$, $\succ_{X,\bar{Y}} = \succ_X$.⁴
7. $F(\pi_i) = X \wedge Z$ et $F(\pi_j) = Y$. L'agent explique ses préférences sur $X \wedge Z$ par Y : il veut satisfaire X et Z si Y est satisfait, mais ne veut pas avoir X ou Z séparément.
 - $Y \succ \bar{Y}, Var(Y), Var(Z) \in Pa(Var(X)), Var(Y), Var(X) \in Pa(Var(Z))$
 - \succ_X dénote la relation de préférences associée à $CPT(X)$. Si \succ_X n'est pas encore définie, nous avons : $Y \wedge Z : X \succ \bar{X}, \bar{Y} \vee \bar{Z} : \bar{X} \succ X$. Sinon, $\succ_{X,Y,Z} = X \succ \bar{X}$, $\succ_{X,Y,\bar{Z}} = \bar{X} \succ X$, $\succ_{X,\bar{Y},Z} = \succ_{X,\bar{Y},Z} = \succ_X$,

⁴Si nous avons \succ_X tel que $Z : \bar{X} \succ X, \bar{Z} : X \succ \bar{X}$, $\succ_{X,Y}$ représente les préférences définies par $Z \wedge Y$ et $\bar{Z} \wedge Y$, tandis que $\succ_{X,\bar{Y}}$ représente les préférences définies par $Z \wedge \bar{Y}$ et $\bar{Z} \wedge \bar{Y}$.

- $CPT(Z)$ est défini de la même façon que $CPT(X)$ en inversant X et Z .
- 8. $F(\pi_i) = X \wedge Z$ et $F(\pi_j) = Y \wedge W$.
 - $Y \succ \bar{Y}, W \succ \bar{W}$,
 - $Var(Y), Var(W), Var(Z) \in Pa(Var(X)), Var(Y), Var(W), Var(X) \in Pa(Var(Z))$
 - Si \succ_X représente la relation de préférences associée à $CPT(X)$ (déjà définie ou pas), alors : $\succ_{X,Y,W,Z} = X \succ \bar{X}, \succ_{X,Y,W,\bar{Z}} = \succ_{X,Y,\bar{W},Z} = \succ_{X,Y,\bar{W},\bar{Z}} = \succ_{X,\bar{Y},W,Z} = \succ_{X,\bar{Y},\bar{W},\bar{Z}} = \succ_{X,\bar{Y},\bar{W},Z} = \bar{X} \succ X$,
 - $CPT(Z)$ est défini de la même façon que $CPT(X)$ en inversant X et Z .
- 9. $F(\pi_i) = X \wedge Z$ et $F(\pi_j) = Y \& W$.
 - $Y \succ \bar{Y}, W \succ \bar{W}$,
 - $Var(Y), Var(W), Var(Z) \in Pa(Var(X)), Var(Y), Var(W), Var(X) \in Pa(Var(Z))$,
 - \succ_X dénote la relation de préférences associée à $CPT(X)$. Si \succ_X n'est pas déjà définie, nous avons $(Y \vee W) \wedge Z : X \succ \bar{X}, (\bar{Y} \wedge \bar{W}) \vee \bar{Z} : \bar{X} \succ X$. Sinon, $\succ_{X,(Y \vee W) \wedge Z} = X \succ \bar{X}, \succ_{X,(Y \vee W) \wedge \bar{Z}} = \bar{X} \succ X, \succ_{X,(\bar{Y} \wedge \bar{W}),Z} = \succ_X$.
 - $CPT(Z)$ est défini de la même façon que $CPT(X)$ en inversant X et Z .
- 10. $F(\pi_i) = X \& Z$ et $F(\pi_j) = \Phi$. L'agent explique ses préférences sur $X \& Z$ par Φ . Donc Φ peut expliquer X et peut expliquer Z . Nous appliquons 2 fois la règle 6.
- 11. $F(\pi_i) = X$ et $F(\pi_j) = Y \wedge Z$. L'agent explique ses préférences sur X par $Y \wedge Z$: Y et Z doivent être satisfaits pour expliquer X .
 - $Y \succ \bar{Y}, Z \succ \bar{Z}, Var(Y), Var(Z) \in Pa(Var(X))$,
 - Si $Pa(Var(X)) = \{Var(Y), Var(Z)\}$ alors $Y \wedge Z : X \succ \bar{X}, \bar{Y} \vee \bar{Z} : \bar{X} \succ X$. Sinon, si \succ_X dénote la relation de préférences associée à $CPT(X)$, nous avons $\succ_{X,(Y \wedge Z)} = X \succ \bar{X}, \succ_{X,(\bar{Y} \vee \bar{Z})} = \succ_X$.
- 12. $F(\pi_i) = X$ et $F(\pi_j) = Y \& Z$. Y ou Z doivent être satisfaits pour expliquer X .
 - $Y \succ \bar{Y}, Z \succ \bar{Z}, Var(Y), Var(Z) \in Pa(Var(X))$,
 - Si $Pa(Var(X)) = \{Var(Y), Var(Z)\}$ alors $Y \& Z : X \succ \bar{X}, \bar{Y} \wedge \bar{Z} : \bar{X} \succ X$. Sinon, si \succ_X dénote la relation de préférences associée à $CPT(X)$, nous avons $\succ_{X,(Y \vee Z)} = X \succ \bar{X}, \succ_{X,(\bar{Y} \wedge \bar{Z})} = \succ_X$.
- 13. $F(\pi_i) = \Phi$ et $F(\pi_j) = \Psi$. Nous pouvons appliquer les règles 6-12 en décomposant Φ et Ψ .

La dépendance causale en oeuvre dans l'explication est très proche de la dépendance logique que l'on trouve dans l'exemple suivant :

- (3) π_1 J'aimerais boire du vin
 π_2 J'aimerais boire du vin blanc

Les préférences de l'agent pour du vin blanc dépendent de ses préférences pour du vin. Cette relation discursive entre π_1 et π_2 est appelée *Elaboration*. Ses effets sur les préférences sont similaires à ceux d'explication.

Elab $Elab(\pi_i, \pi_j) \equiv Iexplication(\pi_j, \pi_i)$ en ce qui concerne les modifications de préférences.

On en arrive à présent aux questions. La façon dont ces dernières affectent les préférences dépend des réponses qui leur sont données dans le discours. Etudions tout d'abord les questions fermées.

Q-Elab-ON, QAP $QElab-ON_A(\pi_i, \pi_j), QAP_B(\pi_j, \pi_k)$ représente la réponse de B (de la forme "oui" ou "non") à une question élaborative posée par A .

Dans ce cas, nous savons que A s'engage sur $F(\pi_i) : \text{Commit}_A(\pi_i)$.

Ensuite, nous devons modifier les préférences de B en fonction de sa réponse :

Oui $\text{QElab-ON}_A(\pi_i, \pi_j)$, $\text{QAP}_B(\pi_j, \text{oui})$. Les préférences de B sur π_i et π_j sont définies en appliquant les règles : $\text{Commit}_B(\pi_i)$; $\text{Elab}_B(\pi_i, \pi_j)$.

Non $\text{QElab-ON}_A(\pi_i, \pi_j)$, $\text{QAP}_B(\pi_j, \text{non})$. Les préférences de B sur π_i et π_j sont définies en appliquant les règles⁵ : $\text{Commit}_B(\pi_i)$; $\text{Elab}_B(\pi_i, \pi_j^*)$, avec $\pi_j^* = \neg(\text{Core}(\pi_j, \pi))$.

Etudions à présent les questions ouvertes.

Q-Elab-Wh, QAP : $\text{QElab-Wh}_A(\pi_i, \pi_j)$, $\text{QAP}_B(\pi_j, \pi_k)$ représente la réponse de B à une question élaborative ouverte posée par A . La sémantique de Q-Elab impose que la question ouverte porte sur les variables définies dans π_i : on sait que π_j ne contient aucune variable.

Comme précédemment, A s'engage sur $F(\pi_i) : \text{Commit}_A(\pi_i)$.

Les préférences de B sur les variables de π_i et π_k sont identiques à celles définies pour une question Q-Elab-ON à laquelle on aurait répondu oui : les variables dans π_k raffinent les préférences sur les variables de π_i . Les préférences de B sont donc définies en appliquant les règles : $\text{Commit}_B(\pi_i)$; $\text{Elab}_B(\pi_i, \pi_j)$.

La dernière, et plus complexe, catégorie de questions est celle des *alternatives*. Ces questions sont de la forme *préféreriez vous un restaurant traditionnel ou une pizzeria ?*.

Alternatives L'agent A pose une question contenant n alternatives à un agent B . La réponse de B nous donne des informations sur ses préférences. Soit π la question alternative posée par A , et π^* la réponse de B à la question π , avec $\text{QAP}_B(\pi, \pi^*)$ et $\&(X_i, \dots, X_n, \pi^*)$. Cette description de la réponse de B prévoit plusieurs réponses toutes les unes aussi bonnes que les autres : pour tout $i \leq n$, B veut satisfaire le littéral X_i . On ajoute donc les préférences suivantes pour chacun des X_i , ou on modifie les préférences existantes le cas échéant, avec $\text{Pa}(\text{Var}(X_i)) = \emptyset$, et $X_i \succ \bar{X}_i$.

Correction : Les études linguistiques faites sur la Correction (Asher & Lascarides (2003)) ont montré que les révisions impliquées dans une relation discursive sont très restrictives. Leur sémantique nécessite essentiellement une opération "cherche et remplace", l'élément devant être remplacé étant signalé par une marque lexicale.

Au niveau des préférences, $\text{Correction}(\pi_i, \pi_j)$ signifie donc que certains littéraux de π_i sont remplacés par des littéraux de π_j . Nous avons un ensemble de règles de la forme $X \leftarrow (Y_1, \dots, Y_m)$, signifiant que le littéral X de π_i est remplacé par l'ensemble de littéraux (Y_1, \dots, Y_m) de π_j . Nous supposons que X ne peut pas dépendre de (Y_1, \dots, Y_m) avant la correction.

14. Si $\text{Pa}(\text{Var}(X)) = \emptyset$, on ajoute pour tous $k \in \{1, \dots, m\}$, $Y_k \succ \bar{Y}_k$. Sinon, si \succ_X dénote la relation de préférences associée à $\text{CPT}(X)$, nous avons pour tous $k \in \{1, \dots, m\}$, $\succ_{Y_k} = \succ_X$.

15. Pour tout W tel que $\text{Var}(X) \in \text{Pa}(\text{Var}(W))$, si \succ_W dénote la relation de préférences associée à $\text{CPT}(W)$, nous définissons \succ'_W comme : $\succ'_{\{W \setminus \{X, \bar{X}\}\}, \bigwedge_{1 \leq i \leq m} Y_i} = \succ_{\{W \setminus \{X, \bar{X}\}\}, X}$, et $\succ'_{\{W \setminus \{X, \bar{X}\}\}, \bigvee_{1 \leq i \leq m} \bar{Y}_i} = \succ_{\{W \setminus \{X, \bar{X}\}\}, \bar{X}}$.

Nous supprimons ensuite X du CP-net.

⁵ $\text{Core}(\pi_1, \pi) = \pi_1$, $\text{Core}(\neg\pi_1, \pi) = \pi_1$.

6 Traitement de notre exemple

Appliquons à présent ces règles au dialogue (1).

π_1 A : Shall we meet sometime in the next week ?

Seule la règle Commit s'applique sur π_1 . Nous avons $\text{Commit}_A(\pi_1)$, $F(\pi_1) = M$, où M signifie Meet. On obtient :

$$CPT_A(M) = M \succ \bar{M}$$

π_2 A : What days are good for you ?

Q-Elab-(π_1 , π_2). A continue à s'engager pour M en π_2 .

π_3 B : Well, I have some free time on almost every day except Fridays.

π_4 B : Fridays are bad.

π_4 explique π_3 , mais n'a aucun impact sur les préférences de B. Nous avons Q-Elab-Wh_A(π_1 , π_2), QAP_B(π_2 , π_3), avec $\&(J_1, J_2, J_3, J_4, \neg J_5, \pi_3)$ où J_1 signifie Monday, J_2 Tuesday, J_3 Wednesday, J_4 Thursday et J_5 Friday. Nous avons $F(\pi_3) = F(J_1) \& F(J_2) \& F(J_3) \& F(J_4) \& \neg F(J_5) = J_1 \& J_2 \& J_3 \& J_4 \& \neg J_5$.

On applique $\text{Commit}_B(\pi_1)$, $\text{Elab}_B(\pi_1, \pi_3)$, *i.e.* $\text{Iexplication}_B(\pi_3, \pi_1)$. On obtient :

$$\begin{array}{lll} CPT_B(M) = M \succ \bar{M} & CPT_B(J_1) = \frac{M : J_1 \succ \bar{J}_1}{\bar{M} : \bar{J}_1 \sim J_1} & CPT_B(J_2) = \frac{M : J_2 \succ \bar{J}_2}{\bar{M} : \bar{J}_2 \sim J_2} \\ CPT_B(J_3) = \frac{M : J_3 \succ \bar{J}_3}{\bar{M} : \bar{J}_3 \sim J_3} & CPT_B(J_4) = \frac{M : J_4 \succ \bar{J}_4}{\bar{M} : \bar{J}_4 \sim J_4} & CPT_B(J_5) = \frac{M : \bar{J}_5 \succ J_5}{\bar{M} : J_5 \sim \bar{J}_5} \end{array}$$

π_5 B : In fact, I'm busy on Thursday too.

On a $\text{Correction}_B(\pi_3, \pi_5)$, avec $F(\pi_5) = \neg J_4$, et donc $J_4 \leftarrow \neg J_4$. On obtient :

$$\begin{array}{lll} CPT_B(M) = M \succ \bar{M} & CPT_B(J_1) = \frac{M : J_1 \succ \bar{J}_1}{\bar{M} : \bar{J}_1 \sim J_1} & CPT_B(J_2) = \frac{M : J_2 \succ \bar{J}_2}{\bar{M} : \bar{J}_2 \sim J_2} \\ CPT_B(J_3) = \frac{M : J_3 \succ \bar{J}_3}{\bar{M} : \bar{J}_3 \sim J_3} & CPT_B(J_4) = \frac{M : \bar{J}_4 \succ J_4}{\bar{M} : J_4 \sim \bar{J}_4} & CPT_B(J_5) = \frac{M : \bar{J}_5 \succ J_5}{\bar{M} : J_5 \sim \bar{J}_5} \end{array}$$

π_6 A : Well next week I am out of town Tuesday, Wednesday and Thursday.

On a $\text{Elab}_A(\pi_1, \pi_6)$, et donc $\&(\neg J_2, \neg J_3, \neg J_4, \pi_6)$. $F(\pi_6) = \neg J_2 \& \neg J_3 \& \neg J_4$. On applique $\text{Iexplication}_A(\pi_6, \pi_1)$, et on obtient :

$$\begin{array}{ll} CPT_A(M) = M \succ \bar{M} & CPT_A(J_2) = \frac{M : \bar{J}_2 \succ J_2}{\bar{M} : J_2 \sim \bar{J}_2} \\ CPT_A(J_3) = \frac{M : \bar{J}_3 \succ J_3}{\bar{M} : J_3 \sim \bar{J}_3} & CPT_A(J_4) = \frac{M : \bar{J}_4 \succ J_4}{\bar{M} : J_4 \sim \bar{J}_4} \end{array}$$

π_7 A : So perhaps Monday ?

$\text{Elab}_A(\pi_1, \pi_7)$, et $\text{Résultat}_A(\pi_6, \pi_7)$ qui n'affecte pas les préférences. On a $F(\pi_7) = J_1$, c'est-à-dire $\text{Iexplication}_A(\pi_7, \pi_1)$. On obtient :

$$\begin{array}{lll} CPT_A(M) = M \succ \bar{M} & CPT_A(J_1) = \frac{M : J_1 \succ \bar{J}_1}{\bar{M} : \bar{J}_1 \sim J_1} & CPT_A(J_2) = \frac{M : \bar{J}_2 \succ J_2}{\bar{M} : J_2 \sim \bar{J}_2} \\ CPT_A(J_3) = \frac{M : \bar{J}_3 \succ J_3}{\bar{M} : J_3 \sim \bar{J}_3} & CPT_A(J_4) = \frac{M : \bar{J}_4 \succ J_4}{\bar{M} : J_4 \sim \bar{J}_4} & \end{array}$$

A ce point du dialogue, on peut voir que les issues préférées de B sont MJ_1 , MJ_2 , MJ_3 , $M\neg J_4$, $M\neg J_5$; tandis que celles de A sont MJ_1 , $M\neg J_2$, $M\neg J_3$ et $M\neg J_4$. Une date leur convient à tous deux pour un rendez-vous : lundi.

Cet exemple est assez représentatif des dialogues du corpus Vermobil. Nous avons également analysé quelques dialogues de tourisme. Toutes nos règles ne sont pas appliquées dans cet exemple, mais toutes sont utiles dans d'autres dialogues du corpus Vermobil ou du tourisme. Nous espérons donner plus d'illustration dans une version plus complète de ce papier, et automatiser l'application de ces règles sur des corpus annotés.

7 Conclusion

Modéliser les préférences exprimées dans des textes est important pour beaucoup d'applications en traitement automatique du langage. Nous avons montré comment utiliser les CP-nets et les modèles des structure discursives pour accomplir cette tâche formellement. Nos règles permettant de modéliser les préférences sont simples et intuitives. Il est en suite possible d'utiliser des techniques de la théorie des jeux pour raisonner sur ces préférences. Bien sûr, tout ceci dépend de l'extraction des structures discursives à partir d'un texte, ce qui est une tâche difficile. Néanmoins, Baldridge & Lascarides (2005) ont montré comment extraire automatiquement cette structure discursive à partir de dialogues du corpus Vermobil. Notre prochain objectif est d'automatiser l'application de nos règles sur les dialogues de ces corpus.

Références

- ASHER N. (1993). *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers.
- ASHER N. & LASCARIDES A. (2003). *Logics of Conversation*. Cambridge University Press.
- BALDRIDGE J. & LASCARIDES A. (2005). Probabilistic head-driven parsing for discourse structure. In *CoNLL'05*.
- BOUTILIER C., BRAFMAN R. I., DOMSHLAK C., HOOS H. H. & POOLE D. (2004). CP-nets : A Tool for Representing and Reasoning with Conditional *Ceteris Paribus* Preference Statements. *Journal of Artificial Intelligence Research*, **21**, 135–191.
- GRICE H. P. (1975). Logic and conversation. In *Syntax and Semantics Volume 3 : Speech Acts*, p. 41–58. Academic Press.
- GROSZ B. & SIDNER C. (1990). Plans for discourse. In *Intentions in Communication*, p. 365–388. MIT Press.
- HAMBLIN C. (1987). *Imperatives*. Blackwells.
- HOBBS J. R., STICKEL M., APPELT D. & MARTIN P. (1993). Interpretation as abduction. *Artificial Intelligence*, **63**(1–2), 69–142.
- LASCARIDES A. & ASHER N. (2008). Grounding and correcting commitments in dialogue. submitted to SIGDIAL.
- MANN W. C. & THOMPSON S. A. (1987). Rhetorical structure theory : A framework for the analysis of texts. *International Pragmatics Association Papers in Pragmatics*, **1**, 79–105.
- REESE B. (in preparation). *Bias in Questions*. PhD thesis, University of Texas at Austin.
- W. WAHLSTER, Ed. (2000). *Vermobil : Foundations of Speech-to-Speech Translation*. Springer.