

Un modèle de dialogue par les attentes du locuteur

Yannick Fouquet

Laboratoire CLIPS/IMAG, équipe GEOD
385 rue de la Bibliothèque, BP53 – 38041 Grenoble Cedex 9, France
Tél.: ++33 (0)4 76 63 55 91 - Fax: ++33 (0)4 76 63 55 52
mel : Yannick.Fouquet@imag.fr - <http://www-geod.imag.fr/fouquet>

Résumé - Abstract

Dans cet article, nous aborderons la notion d'attentes, vue du côté du locuteur, afin d'améliorer la modélisation du dialogue. Nous présenterons notre définition des attentes ainsi que notre notation, fondée sur une approche pragmatique du dialogue. Nous comparerons deux approches, l'une (uniquement stochastique) fondée sur la prédiction d'actes de parole, l'autre mettant en jeu les attentes du locuteur et leur gestion.

In this paper, we will discuss about a generic dialog model. We will demonstrate the importance of a expectations based approach which consider them from user's side. We will present our definition of expectations and their notation based on a pragmatic approach of dialog. Thus, we will compare this approach (expectations and their management), and another stochastic approach which simply consist of act prediction.

Keywords – Mots Clés

Dialogue, attentes, pragmatique, analyse, statistique, aspects cognitifs
Dialog, expectations, pragmatics, analysis, statistics, cognitive aspects

1 Introduction

Nous nous intéressons dans cet article au dialogue oral, qu'il soit inter-humain ou homme-machine. Nous nous pencherons particulièrement sur l'apport de la modélisation des attentes orientées utilisateur. Nous présenterons, dans une première partie, notre approche fondée sur la modélisation des attentes de l'utilisateur. Puis nous nous appuierons sur l'étude (Fouquet, 2001) du corpus de renseignement touristique (Besacier et al., 2001) pour proposer une première estimation de ce que seront pour nous les attentes en situation, que nous validerons et affinerons en terme de probabilités associées à chaque attente par une analyse plus approfondie. Enfin, nous comparerons cette nouvelle approche avec une approche, uniquement stochastique, fondée sur la prédiction d'actes de parole.

2 Un modèle de dialogue fondé sur les attentes du locuteur

Dans le cadre du dialogue, (Lehuen, 1997), (Lemeunier, 2000) ou (Vilnat, 1997) proposent d'introduire la notion *d'attentes* : pour Lehuen, ces attentes sont "ce qui est attendu par le système au regard de l'état courant de la tâche". Il les définit donc en terme d' "objets du domaine d'application et d'actes de dialogue orientés par la tâche" ; pour Vilnat, ce sont "ce que l'utilisateur est susceptible de dire (ce qui implique qu'il n'est pas obligé de le dire), en raison des buts et des plans en cours de développement" ; pour Lemeunier ce sont "des entités pragmatiques manipulées par le système qui correspondent à ce que l'utilisateur est censé dire et ce à quoi le système est donc censé s'attendre". Ces *attentes* sont donc très liées à la tâche d'une part et considérées du point de vue du système d'autre part.

Pour notre part, nous considérons les attentes du point de vue du locuteur, comme les résultats espérés de ses actes. En effet, un locuteur engagé dans un dialogue entend poursuivre ses buts (Vernant, 1992) et obtenir des effets positifs de ses actes. Il projette donc un certain plan sur son interlocuteur qu'il considère comme coopératif et en attend en retour un certain comportement. Les attentes se construisent au cours du dialogue, en fonction des buts et actions des interlocuteurs ; ce sont essentiellement des hypothèses réflexive construites par chacun des interlocuteurs. Par exemple, si le locuteur B ne se montre pas coopératif, la stratégie du locuteur A, guidée par ses attentes va certainement être davantage directive pour forcer B à adopter les buts de A ; ce qui se formalise par : si B est coopératif alors $F_A(p) \rightarrow F_B(p)$, si on constate $F'_B(p)$ alors B n'est certainement pas coopératif. Notre modélisation s'appuie donc sur l'idée de contexte. En effet, un certain acte (ici $F_A(p)$), devrait fournir, selon le contexte (B coopératif), tel autre acte en réponse ($F_B(p)$) répondant aux attentes de A.

Inspiré de (Austin, 1962) et (Searle, 1972) sur les actes de langage, (Vanderveken, 1990) définit la force illocutoire $F(p)$ comme un acte de force F s'appuyant sur le locutoire représenté par le contenu propositionnel p . Celui-ci découpe ses actes illocutoires en cinq catégories (déclaratif, directif, assertif, promissif, expressif). Nous pensons que ce découpage peut être affiné à l'aide des modaux sémiotiques de (Greimas, 1966). Ainsi, un acte sera pour nous une combinaison de faire, savoir, croire, devoir, pouvoir. Soulignant (Austin, 1962), chaque acte débute par un faire, la précision de la force étant fournie en exposant ; $F^{fs}(p)$ pour un faire-faire-savoir de contenu p . Nous aurons donc des actes de la forme $F(p)$, $F^f(p)$, $F^s(p)$, $F^c(p)$, $F^{fs}(p)$, $F^d(p)$, $F^p(p)$, $F^{fc}(p)$, $F^{fd}(p)$, $F^{fp}(p)$.

Cette notation pragmatique en actes permet, selon nous, de mieux appréhender les irrégularités soulevées dans (Fouquet, 2001) et leurs effets (ruptures, réparations, demande de précision ; Bilange, 1992). Souhaitant tenir compte des aspects intentionnel (Colineau, 1997), cognitif (Sperber, Wilson, 1989) et surtout social (Goffman, 1974) (souvent mis de côté), les attentes sont donc, pour nous, codées par l'état mental (d'un locuteur) qui contient les buts (et les engagements), les connaissances (sur le monde et sur la tâche), les règles de conversation et les rôles des participants.

Ainsi, lorsque le locuteur A propose un acte de faire-faire-savoir ($F^{fs}_A(p)$), les attentes seront, par exemple, une réponse de l'allocutaire ($F^s_B(p)$) ou une demande de précision de celui-ci ($F^{fs}_B(p)$) ; $F^{fs}_A(p) \rightarrow F^s_B(p) \mid F^{fs}_B(q)$. Les attentes correspondraient alors dans ce cadre aux actes de B, $\{F'^n_B(p)\}$ possibles en réponse à l'acte du locuteur A, $F_A(p)$. Parmi ces actes possibles (les attentes) se trouvera l'acte que l'allocutaire choisira effectivement : $F'^i_B(q)$.

3 Actes et attentes (vers une approche stochastique)

Notre but, ici, sera la comparaison de deux approches possibles tenant compte des actes en situation, l'une consiste à prédire l'acte suivant à partir de l'acte en cours tandis que l'autre, plus fine, se fonde sur les attentes et leur gestion pour déterminer l'acte suivant.

3.1 Prédiction des actes

Nous adoptons ici la démarche stochastique qui, à partir des actes annotés sur notre corpus, nous permet d'associer des probabilités aux bigrammes {Acte-courant, Acte-suivant}, par la formule de probabilité conditionnelle de Bayes : $P(A_{n+1} | A_1, \dots, A_n) = P(A_1, \dots, A_n, A_{n+1}) / P(A_1, \dots, A_n)$. Nous employons en revanche, dans notre étude, uniquement l'acte courant comme historique (n=1) :

$$P(\text{Acte-suivant} | \text{Acte-courant}) = P(\text{Acte-courant}, \text{Acte-suivant}) / P(\text{Acte-courant})$$

Nous avons donc regroupé les 30 premiers dialogues du corpus afin de définir la fréquence de cooccurrence de chaque acte. Nous avons obtenu 4315 actes dont 3582 actes en réponse (actes formulés par l'allocutaire en réponse au locuteur). La matrice de cooccurrence des bigrammes {Acte-en-cours, Acte-suivant} nous fournit la probabilité de cooccurrence $P(\text{Acte-courant}, \text{Acte-suivant})$ et celle d'apparition des actes $P(\text{Acte-courant})$ de l'agent et du client. Par la formule de Bayes, nous obtenons donc la probabilité d'apparition de chaque acte en contexte, $P(\text{Acte-suivant} | \text{Acte-courant})$ que nous observons sur la (Figure 1) présentant la fréquence d'apparition de chaque Acte-suivant (en colonne) en rapport à l'acte énoncé (en ligne).

Actes	F _A (p)	F ^d _A (p)	F ^f _A (p)	F rd _A (p)	F ^{is} _A (p)	F ^p _A (p)	F ^s _A (p)	F _C (p)	F ^d _C (p)	F ^f _C (p)	F rd _C (p)	F ^{ip} _C (p)	F ^{is} _C (p)	F ^p _C (p)	F ^s _C (p)	Debut	fin
F _A (p)	0	0	9,20	1,15	1,15	1,15	9,20	2,30	1,15	2,30	0,00	0	8,05	0	52,87	0	11,49
F ^d _A (p)	1,39	0	0	0	1,39	0	16,67	0	0	1,39	2,78	0	8,33	4,17	63,89	0	0
F ^f _A (p)	2,90	0	0	7,25	1,45	1,45	18,84	52,17	0	0	1,45	0	0,00	0	14,49	0	0
F rd _A (p)	0	0	0	0	0	0	3,09	1,03	85,57	0	0	0	3,09	0	7,22	0	0
F ^{is} _A (p)	0	0	0	0	3,03	3,03	3,64	0	0,61	0,61	0	0	4,24	1,21	83,64	0	0
F ^p _A (p)	0,65	0	0	0,65	1,30	0,65	7,14	9,09	1,30	0,65	1,30	0	5,84	28,57	42,21	0	0,65
F ^s _A (p)	1,63	0,20	1,29	1,23	3,34	1,97	6,06	0,54	2,18	0,82	1,84	0,07	10,48	1,02	67,26	0	0,07
F _C (p)	9,09	0	7,58	15,15	1,52	4,55	31,82	0	1,52	0	0	0	0	1,52	4,55	0	22,73
F ^d _C (p)	4,00	0	2,40	4,80	11,20	2,40	60,00	0	0	0	0,80	0	1,60	1,60	11,20	0	0
F ^f _C (p)	68,00	0	0	0	8,00	0	16,00	0	0	0	0	0	0	0	8,00	0	0
F rd _C (p)	0	89,36	0	0	0	2,13	2,13	0	0	0	0	0	2,13	2,13	0	0	2,13
F ^{ip} _C (p)	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
F ^{is} _C (p)	0,31	0,31	0,31	3,38	6,77	2,15	76,92	0	0	0	0,62	0	6,46	0,92	1,85	0	0
F ^p _C (p)	11,34	2,06	0	2,06	1,03	35,05	21,65	0	0	0	2,06	0	13,40	2,06	7,22	0	2,06
F ^s _C (p)	1,30	1,65	2,27	2,95	4,53	3,43	65,52	0,34	0,34	0,55	0,62	0	6,80	1,24	8,45	0	0
debut	0	0	0	0	0	61,29	3,23	0	0	0	3,23	0	6,45	19,35	3,23	3,23	0
fin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0

Figure 1 : pourcentages d'actes suivant par rapport à chaque acte

Dès lors, nous pouvons proposer un modèle simple de prédiction d'acte à partir de l'acte énoncé. Ainsi, l'acte prédit est simplement celui qui maximise la probabilité (acte suivant le plus probable) : $\text{Acte-prédit} = \text{argmax}_{\text{acte} \in \{\text{Ens. Actes}\}} P(\text{Acte-suivant} | \text{Acte-courant})$. La

prédiction sera, par exemple, $F^s_c(p)$ si l'on a $F^{fs}_A(p)$ car cet acte obtient 83,64% de probabilité de succéder à l'acte courant.

3.2 Génération et gestion des attentes

Considérant l'approche de prédiction d'acte suivant insuffisante en dialogue, nous avons affiné ce modèle avec notre notion d'attentes. En effet, les actes qui posent des buts n'ont pas la même importance que les autres. Ils ont, pour nous, un lien très fort avec l'acte fournit en réponse (cf. Figure 1 ci-dessus ; bigrammes $\{F_A(p), F_B(p)\}$ et $\{F_B(p), F_A(p)\}$). Lorsque je demande une information, je souhaite que l'allocutaire me fournisse la réponse. Inversement, lorsque je donne une information, je n'ai pas d'attentes vis-à-vis de mon allocutaire mais je répons davantage aux siennes. Nous formulons donc l'hypothèse que lorsque l'acte courant pose un but (donc des attentes) à savoir $F^f(p)$, $F^{fs}(p)$, $F^{fd}(p)$ ou $F^{fp}(p)$, l'acte immédiat de l'allocutaire doit être, dans un dialogue coopératif, un acte parmi les attentes du locuteur (respectant ainsi les maxims de Grice).

Acte/su	$F_A(p)$	$F^d_A(p)$	$F^f_A(p)$	$F^{fd}_A(p)$	$F^{fs}_A(p)$	$F^p_A(p)$	$F^s_A(p)$	$F_c(p)$	$F^d_c(p)$	$F^f_c(p)$	$F^{fd}_c(p)$	$F^{fp}_c(p)$	$F^{fs}_c(p)$	$F^p_c(p)$	$F^s_c(p)$
$F^f_A(p)$								76,60	0,00	0,00	2,13	0,00	0,00	0,00	21,28
$F^{fd}_A(p)$								1,06	88,30	0,00	0,00	0,00	3,19	0,00	7,45
$F^{fs}_A(p)$								0,00	0,67	0,67	0,00	0,00	4,70	1,34	92,62
$F^f_c(p)$	73,91	0,00	0,00	0,00	8,70	0,00	17,39								
$F^{fd}_c(p)$	0,00	95,45	0,00	0,00	0,00	2,27	2,27								
$F^{fs}_c(p)$	0,34	0,34	0,34	3,75	7,51	2,39	85,32								
debut	0,00	0,00	0,00	0,00	0,00	63,33	3,33	0,00	0,00	0,00	3,33	0,00	6,67	20,00	3,33

Figure 2 : pourcentages d'attentes par rapport à chaque acte posant un but

A partir de cette génération d'attentes (Figure 2), nous avons adopté un mécanisme de gestion des attentes suivant celui des buts. Ainsi, chaque acte qui pose un but, pose ses attentes associées dans une pile. Par exemple, lorsque je propose un $F^{fs}_A(p)$, mes attentes sont, si je suis un agent $\{F^s_B(p), F^{fs}_B(p), F^p_B(p), F^f_B(p), F^{fd}_B(p)\}$. Lorsque le but est satisfait ou abandonné, les attentes sont dépilées. Ces attentes seront donc dépilées lorsque un des actes attendus est proposé (avec, l'exemple précédent, $F^s_B(p)$ permet de dépiler simplement les attentes). Ainsi, les $F(p)$, $F^s(p)$, $F^d(p)$ et $F^p(p)$ ne génèrent-ils pas d'attentes mais permettent d'avantage de les dépiler. Nous pouvons voir un exemple de règles de gestion simplifiées ci-dessous (? pour poser un but, + pour l'atteindre, ++ pour le satisfaire et @ pour l'abandonner ; s pour un sous but) :

$$\begin{aligned} \text{But} = \emptyset \wedge \text{Acte} = F^{fs}_Y(p) &\Rightarrow ?S_X(p) \wedge \text{Empile}(F^s_X(p) \vee F^{fs}_X(q)) \\ \text{But} = S_X(p) \wedge \text{Acte} = F^s_Y(p) &\Rightarrow +S_X(p) \wedge \text{Empile}(F^s_X(+ S_X(p)) \vee (F^{fs}_X(q) \wedge S_X(q) \neq S_X(p))) \\ \text{But} = S_X(p) \wedge \text{Acte} = F^{fs}_Y(q) \wedge S_Y(q) = sS_X(p) &\Rightarrow -S_X(p) \wedge ?S_Y(q) \wedge \text{Empile}(F^s_X(q) \vee F^{fs}_X(q)) \\ \text{But} = S_X(p) \wedge \text{Acte} = F^{fs}_Y(q) \wedge S_Y(q) \neq sS_X(p) &\Rightarrow @S_X(p) \wedge ?S_Y(q) \wedge \text{Dépile} \wedge \text{Empile}(F^s_X(q) \vee F^{fs}_X(q)) \\ \text{But} = S_X(p) \wedge \text{Acte} = F^s_Y(q) \wedge q = +S_X(p) &\Rightarrow ++S_X(p) \wedge \text{Dépile} \\ \text{But} = S_X(p) \wedge \text{Acte} = F^{fs}_Y(q) \wedge +S_X(p) \wedge S_X(q) \neq S_X(p) &\Rightarrow ++S_X(p) \wedge ?S_X(q) \wedge \text{Dépile} \wedge \text{Empile}(F^s_Y(q) \vee F^{fs}_Y(q)) \end{aligned}$$

Empiler un but (?) revient donc à empiler ses attentes associées, et inversement, le dépiler (++ ou @) revient à dépiler les attentes de leur satisfaction. Atteindre un but permet d'empiler les attentes liées à la satisfaction du but. Selon ce mécanisme et à partir de notre analyse du corpus, nous formulons des règles de la forme $F^{fs}_A(p) \rightarrow F^s_B(p)(85,96\%) \mid F^{fs}_B(q)(10,53\%) \mid F^{fd}_B(q)(3,51\%)$; lorsque l'acte $F^{fs}_A(p)$ apparaît, les attentes du locuteur A sont $F^s_B(p)$ à 85,96% ou $F^{fs}_B(q)$ à 10,53% ou encore $F^{fd}_B(q)$ à 3,51%. Ainsi, nous n'empilons pas seulement

les attentes mais aussi leur probabilité d'apparition. Cette pondération (qui reflète les rôles des participants) permet de tenir compte de l'aspect social du dialogue. Dès lors nous pouvons proposer un modèle opportuniste qui, considérant l'acte généré, propose les attentes associées.

3.3 Test : Comparaison des deux approches

Nous avons émis l'hypothèse que le deuxième système devrait offrir de meilleures performances que le premier, considérant que la compréhension est plus fine. Nous avons donc analysé le 31^{ème} dialogue (considéré donc comme dialogue de test) selon le même protocole que précédemment.

Actes	F _A (p)	F ^d _A (p)	F ^f _A (p)	F ^{fd} _A (p)	F ^{fs} _A (p)	F ^p _A (p)	F ^s _A (p)	F _C (p)	F ^d _C (p)	F ^f _C (p)	F ^{fd} _C (p)	F ^{fp} _C (p)	F ^{fs} _C (p)	A	C	A+C
F _A (p)													1	1		1
F ^d _A (p)							1				1		1	3		3
F ^f _A (p)								2						2		2
F ^{fd} _A (p)									2					2		2
F ^{fs} _A (p)													4	4		4
F ^p _A (p)												3		3		3
F ^s _A (p)	1		1		1	1	2						2	14	22	22
F _C (p)				2											2	2
F ^d _C (p)							1						1	2		2
F ^{fd} _C (p)		2				1								3		3
F ^{fs} _C (p)							4							4		4
F ^p _C (p)		1				1	1							3		3
F ^s _C (p)			1		3		13			2	1	1	2	23		23
debut										1						1
														37	37	75

Figure 3 : matrice de cooccurrence des bigrammes {Acte-courant, Acte-suivant} dans le 31^{ème} dialogue de notre corpus

La (Figure 3) nous permet de comparer l'acte effectivement produit après chaque acte et celui prédit. En sélectionnant l'acte suivant le plus probable, nous obtenons 45 actes correctement prédits contre 30 moins probables. Le taux d'erreur de 40% semble montrer que l'approche par prédiction d'acte n'est peut-être pas la plus pertinente et qu'elle mériterait d'être affinée. Pour tester nos attentes, nous ne conserverons que les actes qui posent des attentes : F^f(p), F^{fd}(p), F^{fp}(p) et F^{fs}(p). L'attente la plus probable est donc respectivement F(p), F^d(p), F^p(p) et F^s(p). Nous obtenons alors 14 actes corrects contre 1 moins probable. Avec un taux d'erreur de 6,67%, cette deuxième méthode, associée au mécanisme de gestion des attentes, reflétant plus finement le déroulement du dialogue, offre donc des résultats plus pertinents (l'efficacité de la gestion des attentes étant plutôt liée à celle de la gestion des buts). En outre, nous ne conservons ici que l'acte le plus probable mais éviter d'écarter les interventions moins prévisibles permet d'améliorer la prédiction des interventions non par la restriction des actes possibles mais par une pondération de cette possibilité.

4 Conclusion

Dans cet article, nous nous sommes intéressé aux dialogues coopératifs de renseignement touristique avec des rôles non équivalents. Dès lors, l'acte suivant correspond le plus souvent

à l'attente la plus probable. Dans le cadre du dialogue homme-machine, cet acte serait, si possible, choisi, engendrant parfois quelques erreurs. Associé à un mécanisme de gestion des attentes, nous disposons en revanche d'une plus grande flexibilité dans la compréhension de l'énoncé. Dès lors, les actes moins probables sont envisageables, le premier n'étant pas toujours possible (nous ne comprenons pas tout). En dialogue interhumains médiatisé, nous pouvons proposer toutes les attentes avec leur probabilité d'apparition et affiner rétroactivement le contexte (intentionnel, cognitif et social) avec l'acte effectivement choisi. Cette approche nous permet ainsi de modéliser le dialogue de manière plus 'humaine'.

Références

Austin J.L. (1962), *How To Do Things With Words*, Oxford U.P.

Besacier L., Blanchon H., Fouquet Y., Guilbaud J.P., Helme S., Mazonot S., Moraru D., Vaufraydaz D. (2001), *Speech Translation for French in the NESPOLE ! European Project*, Actes de Eurospeech.

Bilange E. (1992) *Dialogue Personne-Machine : Modélisation et réalisation informatique*, Paris, Hermès.

Colineau N. (1997), *Étude Des Marqueurs Discursifs Dans Le Dialogue Finalisé*, Thèse De Doctorat De Sciences Cognitives, Université Grenoble 1.

Fouquet Y. (2001), *De l'étude de dialogues oraux unilingues dans une langue non maternelle*, Acte de *Quatrième Rencontre Jeunes Chercheurs – Parole*, 42- 44

Goffman E. (1974), *Les rites d'interaction*, Éditions de Minuit, Paris.

Greimas A. J. (1966), *Sémantique structurale*, Paris, Seuil.

Lehuen J. (1997), *Un Modèle De Dialogue Dynamique Et Générique Intégrant L'acquisition De Sa Compétence : Le Système Coala*. Thèse De Doctorat, Université De Caen

Lemeunier T. (2000), *L'intentionnalité communicative dans le dialogue homme-machine en langue naturelle*, Thèse de doctorat, Université du Mans.

Searle J.R (1972). *Les Actes De Langage : Essai De Philosophie Du Langage* (Trad. Française Par H. Pauchard), Paris, Hermann.

Sperber D., Wilson. D. (1989), *La pertinence, communication et cognition*, Paris, Les Éditions de Minuit.

Vanderveken D. (1990), *La logique illocutoire*, Bruxelles, Mandarga éd..

Vernant D. (1992), *Approche Actionnelle Et Modèle Projectif Du Dialogue Informatif. Du Dialogue, Recherches Sur La Philosophie Du Langage*, Paris, Vrin N° 14, 295-314.

Vernant D. (1997), *Du Discours A L'action : Etudes Pragmatiques*, Paris, Presses Universitaires De France.

Vilnat A. (1997), *Quels Processus Pour Les Dialogues Homme-Machine ?* In Sabah, G., Vivier, J., Vilnat, A., Pierrel, J.M., Romary, L., Nicolle, A., *Machine, Langage Et Dialogue, Figures De L'interaction*, Paris, L'harmattan.