

# Modèles formels du dialogue

Jean Caelen

Laboratoire CLIPS-IMAG  
Domaine universitaire, BP 53  
38041 Grenoble Cedex 9  
Jean.Caelen@imag.fr

**Résumé.** Cet article présente un état de l'art, quelques modèles formels du dialogue et détaille plus particulièrement une approche fondée sur une logique de l'action utilisant les concepts de but et de stratégie. Le dialogue est considérée dans le cadre d'une interaction verbale homme-machine à propos d'une tâche. Les modèles présentés tirent parti de différentes théories et sont classés en quatre grandes catégories. Ils sont discutés et leurs limites sont énoncées.

## 1 INTRODUCTION

Beaucoup d'ouvrages ont paru en français sur le dialogue homme-machine (DHM) ces dernières années. Cela note l'intérêt de ce domaine auprès des chercheurs. C'est en effet un sujet fascinant dans la mesure où il tente de répondre au fameux défi d'A. Turing [1] sur l'intelligence des machines "les machines seront intelligentes, le jour où elles pourront dialoguer". Ce jour n'est manifestement pas encore arrivé, mais on en perçoit mieux aujourd'hui les difficultés, les limites et ce qu'il est raisonnable d'espérer en matière de modélisation du dialogue.

Parallèlement (et peut-être conséquemment), le besoin en systèmes interactifs ergonomiques ne cesse de s'amplifier. Après la vague des interfaces graphiques qui ont introduit l'écran graphique et la souris ainsi que la possibilité de *manipulation directe*, on s'interroge sur les meilleures formes d'interaction, les plus utilisables, les plus adaptées à la tâche et celles qui nécessitent le moins d'apprentissage pour l'utilisateur. Ce besoin, d'une autre nature que celui de relever le défi de Turing,

appelle des solutions concrètes pour les systèmes à composante orale : *interagir en dialoguant* avec des machines.

C'est de ce besoin dont il est question dans cet article : montrer l'état de l'art et des solutions en modélisation du DHM (dialogue homme-machine). Nous prendrons le point de vue *interactionnel* dans lequel le langage prend place comme modalité d'expression et qui emprunte aux modèles de conversation et de dialogues humains ses paradigmes fondateurs. On ne trouvera donc pas dans cet article de considérations traitant du langage naturel mais au contraire des modèles suffisamment généraux sur le dialogue et pouvant être mis à profit au-delà du cadre de l'usage de la langue, pas exemple en interaction homme-machine.

## **2 LES THEORIES FONDATRICES PAR APPROCHES DISCIPLINAIRES**

De nombreuses disciplines se sont intéressées au dialogue humain :

- (a) l'éthnométhodologie (branche de l'ethnoscience) qui pose la communication dans une perspective *sociale* : les individus agissent dans un cadre normalisé selon des règles et des conventions qui sont socio-culturellement bien définies,
- (b) la philosophie du langage (notamment la philosophie analytique) qui s'intéresse à un *individu* placé en situation de communication, sur un plan *intentionnel* et *actionnel*,
- (c) les cognisciences qui retiennent de la communication les aspects liés à la perception, à la planification et au raisonnement cognitifs,
- (d) la linguistique qui étudie la structure et la fonction du langage dans le dialogue,
- (e) enfin l'intelligence artificielle qui utilise préférentiellement la notion d'agent rationnel et/ou de planification pour résoudre le problème de la coordination d'actions conjointes faites à travers et par le dialogue.

### **2.1 L'éthnométhodologie**

Les éthnométhodologistes [2], [3] ont développé leur approche autour de l'étude des capacités de raisonnement de sens commun que manifestent des individus, membres d'une même culture, pour produire et reconnaître des actions intelligibles. Cette approche est plus descriptive — parce que naturaliste — que prédictive : elle met l'accent sur une forme de

rhétorique de la moralité à travers le principe d'identité qui proclame que les méthodes de raisonnement sont partagées entre les individus et apparaissent à la surface de la vie sociale parce que les règles d'interaction et d'action sociales sont profondément inscrites chez chacun. En d'autres termes pour les éthnométhodologistes, le raisonnement humain est de nature normative. En particulier, il est lisible dans une conversation qui devient pour eux le lieu d'une recherche fonctionnaliste sur les actions humaines. Les pragmatolinguistes se trouvent également dans la même ligne de pensée : "parler n'est pas simplement la mise en fonctionnement d'un système linguistique, mais une forme essentielle d'action sociale" [4], "un système de comportements de différents partenaires qui s'influencent réciproquement dans des actions concrètes" [5]. Pour eux l'interaction fonctionne selon les principes de la réciprocité des perspectives et de la réciprocité des motivations. Ces principes s'appuient sur la notion d'intercompréhension qui définit le projet d'action du locuteur A (ou intention) à travers la réaction qu'il attend de son partenaire B, comme moyen de réaliser son but. Le principe de réciprocité des motivations est l'anticipation par A que son projet, une fois compris, sera accepté par B comme la raison et la motivation « à cause de » du projet et de l'action de B [6]. Si, pour eux, ce principe suffit à régler les niveaux locaux de l'interaction (tours de parole fondés sur le concept des paires adjacentes), le deuxième principe, celui de la réciprocité des perspectives est nécessaire pour régler les niveaux supérieurs d'organisation de l'interaction. Ces niveaux sont liés à une conception hiérarchique de l'action dans laquelle ce principe fonde la complémentarité ou la symétrie des rôles des partenaires pour le guidage des niveaux d'exécution. De lui résultera la stratégie utilisée dans l'interaction (négociation, coopération, etc.) issue d'un accord entre les partenaires.

Goffman [7] va plus loin, il introduit les notions de rituel et de face : l'interaction sociale est guidée par le souci de ne pas "perdre la face". La notion de face renvoie à celles de rôle, de statut, c'est-à-dire de la position d'où on parle et que l'on doit conserver — car "en parlant on construit une image de soi". Au-delà des tours de parole et des paires adjacentes, il y a d'autres mécanismes de régulation comme la mise en scène des places de laquelle on attend des "retours" comme l'acquiescement (Mmm...), la surprise (ah oui ?...), les marques de sympathie, le rire, etc.

Plus récemment, Suchman [8] a introduit la notion d'action située, qui donne une grande importance à la prise de décision en situation (en d'autres termes au contexte) et qui montre que du fait de la grande dynamique du dialogue, les locuteurs restent en état d'adaptation et

d'ajustement permanent. Dès lors, il est difficile pour un observateur de prédire les actes de parole puisqu'ils résultent d'une double interprétation, du locuteur et de l'allocataire, et que les effets n'en sont pas, de ce fait, clairement prévisibles. Elle fonde alors sa description de la conversation sur des stratégies opportunistes, qui ne sont *instanciables* que localement et en tous cas non planifiables à long terme. Par cela elle remet en cause un grand nombre de théories fondées de près ou de loin sur la logique des intentions. Malheureusement son apport très riche par ailleurs, reste plus critique que constructif dans la perspective de poser un modèle formalisable pour le dialogue.

## 2.2 La philosophie analytique

Jugeant des insuffisances d'une sémantique générale en linguistique, certains philosophes du langage comme Austin [9] et Searle [10] mettent l'accent sur la pragmatique en partant du principe que tout énoncé est un acte de langage. Austin, s'intéressant aux verbes performatifs, remarque qu'ils contiennent l'action qu'ils dénotent. Ainsi, dire "je te baptise Queen Elisabeth" c'est faire l'action du baptême pour peu qu'on ait la légitimité pour le faire ; de même "la séance est levée" ne donne pas seulement le signe de fin mais produit une action dans le monde. Pour Austin tout énoncé devient un acte de parole qui contient trois composantes : le locutoire (le dire), l'illocutoire (le faire), le perlocutoire (l'effet produit sur son interlocuteur). Ainsi l'énoncé "les canots à la mer" est-il chargé de sens pour l'équipage ! "Haut les mains" vise à neutraliser autant qu'à effrayer la victime. "je te promets de venir demain" est une forme engageante même si elle ne garantit pas totalement que cet événement va se produire : il peut y avoir une série d'empêchements imprévisibles, ou il se peut que l'intention soit plutôt de rassurer que de s'engager vraiment. Ainsi à travers un même énoncé, plusieurs buts sont poursuivis, le but illocutoire (effet espéré de l'action, par exemple neutraliser la personne) et le but perlocutoire (effet espéré sur l'allocataire, lui faire peur).

Searle [11] replace la langue dans la théorie des actes de langage et plus généralement dans une théorie des intentions. Pour A, énoncer une proposition résulte de l'intention de la produire ; pour B, la comprendre c'est interpréter l'intention qui l'a sous-tendue, dans le contexte où elle a été produite. Dans cette théorie, la communication se situe dans la pragmatique, et la pragmatique des actes de langage s'inscrit à son tour dans une théorie du langage et une théorie de l'action selon deux perspectives : la description des actes de langage et leur régulation selon le principe d'exprimabilité, c'est-à-dire,

- (a) énoncer des mots = effectuer des actes d'énonciation,
- (b) référer et prédiquer = effectuer des actes propositionnels (locutoires),
- (c) affirmer, ordonner, promettre, etc. = effectuer des actes illocutoires,
- (d) effectuer des actes perlocutoires = agir sur son interlocuteur.

Jusqu'en (c) on peut représenter un acte de langage par  $Fp$ ,  $p$  = contenu propositionnel et  $F$  = force illocutoire, (ex. : "je te promets que je viendrai",  $F$  marquée par "je te promets", force promissive, et  $p$  marqué par "je viendrai"). Vanderveken [12] a formalisé une logique illocutoire du discours, essentiellement monologique, poursuivant ainsi les travaux de Searle.

Jusqu'à donc, Searle comme Vanderveken, ne prennent pas encore en compte le rôle du locuteur ni celui de l'allocutaire qu'il ne font apparaître qu'en (d). Ceci leur a été reproché puisque cela implique jusqu'en (c), l'élimination du locuteur parlant au profit d'un locuteur abstrait ce qui les oblige à introduire la notion de sens littéral. Cela introduit une difficulté pour l'interprétation des actes indirects qui se font par énonciation non littérale ("le sucre s'il te plaît !" ne peut pas être analysé sans l'action qui la sous-tend qui prend encore un sens différent si c'est un diabétique qui la prononce). Searle propose alors pour l'interprétation de ce type d'acte une stratégie inférentielle, qui examine toutes les conditions de réalisation de l'acte (situation, monde, arrière-plan, etc.), le pourquoi, les intentions du demandeur, le but poursuivi, etc.

La taxonomie des actes de langage est la suivante :

**Acte assertif** : la composante illocutoire décrit un état de fait existant. Le locuteur dit comment sont les choses. Le but est de rendre le contenu propositionnel (qui est une proposition) conforme au monde. L'acte assertif révèle les croyances du locuteur. Nous notons cet acte  $F^S$  (faire savoir).

**Acte directif** : le but illocutoire est de mettre l'interlocuteur (qui est ici le locuteur lui-même) dans l'obligation de réaliser une action future. Le locuteur essaie de faire faire les choses. Le but est de rendre le monde conforme au contenu propositionnel (qui contient l'action future de l'interlocuteur). L'acte directif exprime les désirs et la volonté du locuteur. Nous notons cet acte  $F^F$  (faire faire une action) ou  $F^D$  (faire devoir) lorsque l'obligation est forte ou  $F^{FS}$  (faire faire savoir) pour une demande d'information.

**Acte promissif** : il s'agit d'une obligation contractée par le locuteur lui-même de réaliser une action future. Le locuteur s'engage à faire quelque chose. Le but est de rendre le monde conforme au contenu propositionnel (qui contient l'action future de l'interlocuteur). L'acte promissif révèle l'intention du locuteur. Nous notons cet acte  $F^P$  (faire pouvoir).

**Acte expressif** : le but illocutoire de l'acte expressif est d'exprimer l'état psychologique qui lui est associé. La direction d'ajustement n'est pas de rendre le monde conforme aux mots ou vice versa. La proposition exprimée est présupposée : on se réjouit ou on déplore qu'elle soit vraie. Cet acte est très peu présent en DHM, nous le notons  $F^S\emptyset$ .

**Acte déclaratif** : le but illocutoire de l'acte déclaratif est de rendre effectif son contenu. Le locuteur provoque des changements effectifs dans le monde par ses déclarations. Cet acte a simultanément deux directions d'ajustement entre le langage et le monde. Il faut qu'il soit accompli dans une certaine institution extra linguistique qui confère au locuteur les pouvoirs de provoquer de nouveaux faits institutionnels par le seul accomplissement approprié d'actes de langage. Nous notons cet acte  $F^A$ .

Malgré ses qualités, cette approche est essentiellement monologique, c'est-à-dire qu'elle ne prend pas en compte les énoncés dans la dynamique du dialogue. Une extension est donc nécessaire, pour aborder les problèmes du dialogue et de l'intersubjectivité.

### 2.3 La psychologie cognitive

Parmi les théories les plus fécondes pour le dialogue, nous retiendrons les approches de Grice puis celles de Sperber et Wilson.

(a) Pour Grice [13], un échange dialogique suit une certaine logique fondée sur un principe général de coopérativité ; les interlocuteurs reconnaissent chez leurs partenaires participant à un échange, un ou plusieurs buts communs dans une direction acceptée par tous, ce que Grice formule sous forme de maximes à l'intention des conversants eux-mêmes (à la manière des classiques français, La Rochefoucauld notamment) : "que votre contribution à la conversation soit, au moment où elle intervient, telle que le requiert l'objectif ou la direction de l'échange verbal dans lequel vous êtes engagés". Il exprime plus précisément cette maxime à l'aide de deux autres et définit la notion d'implicature comme l'ensemble des conséquences de l'application ou de la non application de ces maximes par les conversants.

- maxime de qualité : “que votre contribution soit véridique” qui se décompose en, “n’affirmez pas ce que vous croyez être faux” et “n’affirmez pas ce pour quoi vous manquez de preuves”.

- maxime de quantité : “que votre contribution contienne autant d’informations qu’il est requis (pour les visées conjoncturelles de l’échange)” et “que votre contribution ne contienne pas plus d’informations qu’il n’est requis”

La maxime de qualité semble limitative voire même quelque peu naïve car elle semble exclure l’ironie, le mensonge ou la dissimulation. En réalité il n’en est rien parce que précisément, une règle violée a autant d’implications qu’une règle respectée par le fait qu’il y a eu à moment donné un besoin de la transgresser. Ainsi pour Grice, l’ironie fonctionne parfaitement sous ce principe car elle est intentionnellement codée par l’intonation de la voix ou des gestes significatifs. De même le mensonge finit par se détecter à travers l’incohérence du menteur qui ne respectant pas la maxime “n’affirmez pas ce pour quoi vous manquez de preuves”, se trouvera tôt ou tard face à ses contradictions.

La maxime de quantité peut se réduire à la seule maxime de pertinence “soyez pertinent”. Cette maxime cache à elle seule toute une théorie de la pertinence que nous développerons plus loin.

Enfin, Grice complète ses maximes sur le contenu (quoi dire) par des maximes de modalité, c’est-à-dire sur le comment dire: “soyez clair”, “soyez bref”, “soyez méthodique”, “évittez d’être ambigu”, qui visent à une meilleure efficacité du dialogue.

Par le terme implicature, Grice veut distinguer certaines conclusions que l’on peut tirer des énoncés, des véritables implications logiques. En effet les implicatures sont fondées sur l’usage du langage d’une part et prennent en compte les non dits, implicites, ellipses, etc., d’autre part. Le calcul des implicatures ne devient donc possible qu’en appliquant les maximes, par exemple “j’ai faim” suggère que d’une part ce fait est vrai et d’autre part que je réclame à manger. Cela peut suggérer également qu’il s’est déroulé un certain temps depuis le dernier repas, etc.. Les implicatures dépendent fortement du contexte et du sujet parlant.

(b) Pour Sperber et Wilson [14] — qui critiquent Grice sur la notion d’implicature —, la communication est un “échange d’indices qui orientent (ou ré-orientent) les processus inférentiels des interlocuteurs en présence”. Ils rejettent la théorie codique trop simpliste ; les pensées ne sont pas codées par A, transportées puis décodées par B : les pensées de A orientent seulement celles de B. Sperber et Wilson réfutent ainsi le

modèle codique qui stipule une phase de codage et de décodage des informations linguistiques et extra linguistiques (signes dans un cadre sémiotique plus large) selon un ensemble de conventions, de savoirs partagés par les deux interlocuteurs et de suppositions mutuelles de l'un sur l'autre. Cette vision, fondée sur une approche psychologique, se distingue de celle de Searle pour lequel existent des intentions mais aussi, implicitement, un code. Sperber et Wilson nient l'utilité de la théorie des actes de parole qu'ils trouvent classificatoire mais non opératoire. Ils affirment qu'il faut remonter aux attitudes cognitives à travers les processus inférentiels pour comprendre les mécanismes de la communication. Ils posent la signification au plan de la pertinence : les indices échangés doivent être pertinents pour pouvoir servir à retrouver l'intention du destinataire ; signifier, c'est rendre signifiant quelque chose à quelqu'un ; signifier c'est vouloir dire c'est-à-dire communiquer ses intentions ou obtenir un succès dans la compréhension de ses intentions par autrui. Plus formellement, le succès de (A veut dire S à B) est obtenu si :

- (a) B produit une réponse R car
- (b) B reconnaît l'intention de A contenue dans S (notée  $\text{intention}(A) \subseteq S$ )
- (c) et la réponse R résulte en partie de (b)

Cela permet de distinguer deux sortes d'intentions : les intentions informatives (A informe B de  $\text{intention}(A) \subseteq S$ ) c'est-à-dire A rend manifeste au destinataire B un ensemble d'indices ou d'hypothèses) et les intentions communicatives (A informe B de  $\text{intention}(\text{intention}(A) \subseteq S) \subseteq S$  c'est-à-dire A rend manifeste au destinataire qu'il a une intention informative). Cette information se fait au travers d'indices pertinents. De ce fait A et B doivent se montrer coopérants pour que la communication réussisse. La coopération est donc une résultante cognitive plus qu'une résultante sociale.

Une critique essentielle de cette théorie repose sur la régression à l'infini, engendrée par la récursivité sur l'intention que A a de produire un stimulus qui rende manifeste à B que A veut lui rendre manifeste un ensemble d'hypothèses, etc. Cela semble un processus cognitif peu réaliste. Le concept de pertinence est quant à lui plus fécond. Pratiquement, la pertinence repose sur deux facteurs : l'effet cognitif produit chez le destinataire et l'effort qu'il a fallu pour le produire. La pertinence est donc le rapport de ces deux quantités, c'est le juste équilibre du coût de traitement chez les deux interlocuteurs.

## 2.4 La linguistique

Les linguistes de l'école de Genève [15], [16], envisagent le discours comme cadre de structuration d'échanges linguistiques. Ils s'efforcent d'utiliser le moins possible les contraintes ou les données extérieures au discours comme les intentions, les présuppositions, etc., qu'ils jugent trop subjectives. Ils proposent une théorie de la cohérence et non une théorie de l'interprétation : "le problème principal est de formuler un ensemble d'unités conversationnelles, un ensemble de relations entre ces unités, un ensemble de principes gouvernant la composition des unités simples en unités complexes, bref la formulation de règles de bonne formation" [17]. Cette approche n'est pas prédictive ; elle reste essentiellement une description fonctionnaliste et/ou structurelle des énoncés de A et de B rendue possible une fois le discours achevé et l'action accomplie.

Les phases du dialogue que retient l'école de Genève sont les suivantes : l'ouverture (initiative, évaluative, réactive), la continuation (initiative, évaluative, réactive), la clôture (initiative, évaluative, réactive), l'incidence (abandon temporaire d'une activité en cours de réalisation pour la reprendre par la suite) et la rupture (abandon définitif d'une activité en cours de réalisation). Un dialogue commence par une ouverture et s'achève par une clôture. La phase non marquée est la continuation à caractère initiatif.

Luzzati [18] développe de son côté un modèle de dialogue à deux axes : l'axe régissant horizontal correspond à la recherche d'un accord (système questions réponses principales et secondaires) et l'axe incident vertical répond aux exigences de clarté et d'explication. La mise en œuvre s'appuie ensuite sur des variables d'avancement, de profondeur, d'incidence et d'écartement.

## 3 LES MODELES FORMELS

Ces théories ont inspiré divers modèles computationnels que l'on peut classer en quatre groupes :

- (a) les grammaires structurelles,
- (b) les plans et intentions,
- (c) les logiques mentalistes,
- (d) les jeux.

### 3.1 Les grammaires structurelles

Le modèle genevois est à la source de la plupart des modèles structurels. Il est souvent utilisé en DHM car il est d'une implémentation claire et commode. Il se prête bien à une formalisation de type grammaire de dialogue. On peut citer par exemple [18], [19], [20], [21], [22] en France.

Le modèle est de nature hiérarchique et se développe sur plusieurs axes (l'axe régissant et l'axe incident chez Luzzati [23], directeur et subordonné chez Bilange [19]). Il se caractérise essentiellement par l'existence d'une structure à l'intérieur de laquelle le dialogue peut se tisser. Cette structure se décrit à l'aide d'une grammaire hors-contexte. Par exemple Bilange [19] propose la grammaire suivante pour un dialogue finalisé (de type application de réservation aérienne) :

$$\begin{aligned} E &\rightarrow I.\{I\}^* \\ I &\rightarrow CD \mid CS.CD \mid CD.CS \\ CD &\rightarrow \{A \mid I\}D \\ CS &\rightarrow \{E \mid I \mid A\}S \end{aligned}$$

Avec,

Catégories discursives : E = échange, I = intervention, A = acte de langage

Fonctions : S = subordonné, D = directeur

Constituants : CS constituant subordonné, CD constituant directeurs

Une expression telle que  $CD \rightarrow \{A \mid I\}D$  se lit « un constituant directeur est formé d'une suite d'actes directeurs et/ou d'une intervention directrice ».

A ces règles structurelles qui régissent l'enchaînement possible des constituants du dialogue s'ajoutent des fonctions que le modèle prévoit (fonctions illocutoires essentiellement). Par exemple les fonctions d'initiative, de réactive et d'évaluative, parfois de conclusive (pour la clôture).

Une initiative ouvre un échange (ou réintroduit un échange momentanément mis en attente), ce qui met l'allocutaire en position de réaction. La plupart du temps cet effet attendu l'amène à produire une réactive, à laquelle le locuteur réagit à son tour par une évaluative ou par une clôture, lorsqu'il a obtenu ce qu'il désirait. Notons que la clôture approbative n'est pas toujours marquée explicitement. Quant à l'intervention évaluative, elle peut être positive (elle marque l'avancée vers le but) ou négative (écartement du but) : il s'agit d'actes qui

marquent la position (et parfois les raisons de) du locuteur par rapport à son interlocuteur.

Un échange, pris dans sa globalité, peut être lui-même initiatif (échange de politesses par ex.) ou réactif ou évaluatif ou conclusif. Dans un échange complet il peut y avoir tout ou partie de ces divers types d'échanges.

Exemple (cité de [19]) :

- |                                |                                   |
|--------------------------------|-----------------------------------|
| L : Quand voulez-vous partir ? | (1) Intervention initiative       |
| A : Le 13 novembre             | (2) Intervention réactive         |
| L : Le 13 novembre...          | (3) Intervention évaluative       |
| à quelle heure ?               | (4) Intervention initiative       |
| A : Non, le 20 novembre !      | (5) Intervention évaluative à (3) |
| à 10 heures                    | (6) Intervention réactive à (4)   |

Dans cet échange complet, (3)+(5) est un échange évaluatif et (4)+(6) un échange réactif. Ces deux échanges sont imbriqués. La clôture est ici implicite, le locuteur ayant obtenu les renseignements qu'il demandait.

Les limites de ce modèle proviennent de la difficulté d'interpréter les fonctions des actes : (6) est une réaction à (4) car (4) est une demande et (6) une réponse cohérente dans le cadre de la tâche sous-tendue par le dialogue : une demande de renseignement d'horaire de train. On ne peut donc interpréter correctement la séquence qu'au regard de la tâche d'une part, mais aussi au regard du but poursuivi par le demandeur (il s'agit pour lui de fournir un billet en consultant une base de données et non de partir lui-même en train).

Luzzati [23] a tenté de dépasser cette rigidité structurelle en proposant un modèle dynamique évoluant sur deux axes, l'axe régissant qui est celui de l'avancée dans la tâche et l'axe incident ou axe de clarification, aidant à la compréhension pour une meilleure avancée sur l'axe régissant. Il distingue, pour une application de renseignement, les questions et réponses principales et secondaires sur l'axe régissant et les questions et réponses incidentes. Ce modèle vise avant tout à éviter les erreurs d'incompréhension à répétition qui risquent de bloquer le dialogue. Il contrôle des variables interactionnelles, qui mesurent les écarts à l'axe régissant, et force le retour sur cet axe en cas de trop fort éloignement.

### 3.2 Les plans et intentions

Avec la gestion de dialogue à l'aide de *plans*, on ne considère plus seulement la structure et la fonction des actes de dialogue mais également les intentions des interlocuteurs. Les locuteurs dialoguent pour réaliser un

but (c'est particulièrement vrai dans les dialogues finalisés), ils se servent pour cela de plans et de schémas [24], [25], [26], [27], [28], [29], [30], [31]. Cette approche se situe donc dans le cadre plus général de la théorie de la planification [32]. Par analogie aux actions d'un robot pour atteindre un but, les actes de langage sont des actions faites pour modifier les mondes de connaissance des interlocuteurs et le monde de la tâche. Lorsqu'il s'agit d'un dialogue homme-machine, le rôle de la machine est donc de comprendre le plan de l'utilisateur pour l'aider dans sa tâche. Avant de comprendre le plan du locuteur il s'agit bien sûr de le reconnaître puis d'identifier les buts restant à atteindre et correspondant aux intentions de l'utilisateur. L'approche est donc schématiquement la suivante :

- **si** l'on suppose que le locuteur a des buts, qu'il planifie sa tâche et le dialogue en produisant des actes de langage,
- **alors** la machine doit reconnaître le plan à travers les actes de langage de manière à déduire les buts du locuteur.

Cette approche nécessite une modélisation précise des plans et des buts des interlocuteurs. Les buts sont modélisés par des *opérateurs de croyance* sur les connaissances (par ex. objets de la transaction dans le cas de la demande de renseignements). Les plans sont modélisés comme des suites d'actions mettant en relation un état initial du monde et un état but. Il s'agit alors d'activer les plans pour atteindre cet état but. Les méthodes mises en œuvre relèvent de la planification en Intelligence Artificielle [33], [34], [35], [36]. On utilise souvent des représentations d'actions et de plans fondées sur des *schémas* comme par exemple :

PRENDRE-TRAIN(Passager, Train)  
définition : appartient(Voiture, Train)  
corps : embarquer(Passager, Voiture)  
pré-conditions : possède(Passager, Billet)  
sur(Passager, Quai)  
effets : dans(Passager, Voiture)  
contraintes : accoster(Train, Quai)

Les limites d'une telle représentation sont évidentes : on ne peut traiter que des dialogues qui sont dirigées par la tâche et dont le but est connu. Les incidences *hors tâche* du dialogue sont quasiment impossibles à traiter (on peut ici répondre à des questions comme "que faut-il faire avant d'embarquer ?", réponses : "être en possession d'un billet et attendre que

le train ait accosté au quai”, mais on ne peut pas répondre à une question comme “combien de temps faut-il attendre avant d’embarquer ?”).

Pour dépasser ces limites, Cohen, Allen et Perrault dans leurs travaux pionniers [24], [25], ont appliqué cette technique de planification à l’activité de dialogue elle-même. Par exemple, l’acte d’informer est un plan tel que :

INFORM (Speaker, Hearer, P)

Preconditions : speaker knows P (P is true & speaker believes P)

Body : hearer believes speaker wants (hearer knows P)

Effet : hearer knows P

La puissance d’un tel modèle est évidente : il permet de modéliser le dialogue comme une activité en-soi et de la distinguer de l’activité déployée pour la tâche. Mais les difficultés demeurent de reconnaître les intentions et les buts des locuteurs : que se passe-t-il pour les buts implicites ? Comment traiter les énoncés complexes ciblant plusieurs buts à la fois ? Comment rendre robuste un tel système plus sensible aux erreurs de reconnaissance des buts ? Comment savoir qu’un locuteur veut savoir quelque chose ? Pour répondre à ces problèmes Cohen, Perrault et Allen introduisent finalement des heuristiques et des techniques de reconnaissance de plans non indépendantes des plans de la tâche.

Dans la continuation de ces travaux, Litman [34] a présenté un modèle qui sépare mieux le discours et le domaine dit *de sens commun*. Pour modéliser le discours elle adopte l’approche de Cohen & all. tandis que pour le second elle adopte des représentations structurelles fondées sur la notion de *focus*. C’est ce deuxième niveau, dépendant du domaine, qui contrôle le dialogue. Les plans du discours sont organisés en trois classes :

- la classe Continue, qui contient les plans de tâches non-linguistiques, et qui sont en relation avec les plans du domaine,
- la classe Clarification, qui contient les plans évoqués en cas d’incompréhension, d’impasse, etc.
- la classe Topic, qui contient des plans d’étape du dialogue, comme *Introduce-Plan*, pour la phase d’ouverture par exemple.

Son approche est alors à deux niveaux : plans et méta-plans.

### 3.3 Les approches mentalistes

Les modèles « mentaux » sont une extension des modèles précédents. Ils font une place prépondérante à la logique dialogique plus qu'à la logique de la tâche. Les initiateurs de ce courant sont Cohen, Perrault et Allen [24], [25] déjà cités précédemment. C'est pourquoi, il n'y a pas réellement de rupture de paradigme entre les modèles dits de planification et les modèles mentaux.

Grosz et Sidner [29] ont proposé une théorie du discours articulée autour de la reconnaissance des intentions et du focus d'attention un peu à la manière de Litman. Un discours est composé de trois éléments, (a) sa structure linguistique, modélisée en termes de segments de discours, (b) sa structure intentionnelle – plus précisément l'ensemble des buts des participants et (c) sa structure attentionnelle ou focalisation sur le discours. Cette dernière structure permet surtout de résoudre les problèmes de référence pragmatique et discursive. Les buts sont de nature discursive (par exemple informer) ou de nature pragmatique (par exemple faire une action, identifier un objet, demander d'identifier un objet). La focalisation sur le discours, l'apport le plus original du travail, permet de faire ressortir les éléments les plus pertinents du moment. Pour cela l'espace est organisé en sous-espaces (nœuds) liés entre eux (arcs). On associe un sous-espace à un segment de discours que l'on charge dans une pile, elle-même gérée par le module de reconnaissance des intentions. Il n'est pas clair cependant comment tout cela s'articule en pratique ni comment les intentions sont reconnues. On ne voit pas non plus comment sont traitées les incidences au discours principal.

Pour Pollack [37] les plans doivent modéliser des attitudes mentales complexes, en particulier à côté des plans nécessaires à l'organisation de la tâche et du discours, il doit y avoir des plans pour organiser les intentions. Pollack avance que de tels *plans d'intentions* sont utiles pour conduire le dialogue à un niveau plus profond. Cela permet aussi de réparer ou de corriger les erreurs ou les engagements dans des mauvais plans par manque de connaissance des locuteurs. Par exemple le dialogue :

A : je voudrais parler à Kathy, avez-vous son numéro de téléphone ?

B : je suis désolé, mais Kathy est sortie de l'hôpital.

montre que A s'engage sur un mauvais plan par manque de connaissance.

Pollack modélise les intentions à l'aide de la notion de connaissance mutuelle. Cette connaissance mutuelle est une pré-condition pour activer

un état mental (ou état intentionnel), et un état mental est un plan. Nous avons donc :

ETAT-MENTAL(Intention)

Pré-conditions : croyances et connaissances mutuelles

Effets : plan d'action

Avec ce niveau de représentation il est facile pour une intention donnée, de vérifier les pré-conditions pour activer ou non un plan d'action. C'est ce qui permet à B de donner le bon niveau de réponse à A dans l'exemple ci-dessus plutôt que de poursuivre un plan en impasse (comme tenter de rechercher le numéro de téléphone en ayant peut-être demandé à A d'épeler le prénom par exemple). Cette modélisation est très sophistiquée mais se heurte à la difficulté de description de tous les cas possibles pour des applications pratiques. Il est pratiquement impossible d'avoir une bonne couverture et une bonne efficacité pour développer de nouvelles applications.

Plus récemment Sadek [38] a mis au point un type de modèle pour des applications pratiques (projet ARTIMIS) en définissant les principes d'un *agent rationnel dialoguant*. Cet agent a des capacités de négociation permettant une très grande flexibilité dans l'interaction : il suit le dialogue en autorisant les incidences, reprises, explications, demandes d'aide, etc. tout en ne perdant pas de vue le but final. Il doit donc à la fois avoir des réactions coopératives et prendre des initiatives dans la tâche, "un système de dialogue intelligent doit en fait, être un système intelligent qui dialogue". Sadek rajoute : "ce système doit avoir un comportement rationnel" [39]. Cela l'amène à définir un cadre logique complet apte à rendre compte d'*états mentaux*. Nous ne détaillerons pas ce modèle ici, dans la mesure où il a été déjà exposé aux journées du GdR I<sup>3</sup>.

### 3.4 Les jeux de langage

Un autre courant en modélisation du dialogue tend à se développer autour de la théorie des jeux [40]. Wittgenstein le premier a considéré qu'une conversation est un *jeu de langage*. Peut-être un peu abusivement peut-on aussi ranger dans cette catégorie les modèles argumentatifs comme [41], [42].

En règle générale, on suppose que chacun des interlocuteurs est engagé dans un jeu dont les tours de parole représentent des *coups*. Ces coups sont plus ou moins pertinents (c'est-à-dire qu'ils atteignent plus ou moins

leur but). On avance dans le jeu en suivant les règles et en tentant de maximiser ses gains ou de réduire ses pertes (par exemple le gain à la sortie du jeu sera d'avoir obtenu un renseignement, ou d'avoir résisté aux arguments de son interlocuteur, etc.). On distingue plusieurs catégories de modèles : les modèles logiques et les modèles stratégiques. Dans les modèles logiques l'accent est mis sur les gains (convaincre son interlocuteur par exemple) sans s'intéresser outre mesure à la manière dont est obtenu le résultat, tandis que dans les modèles stratégiques c'est l'inverse, il s'agit d'atteindre un résultat de manière optimale.

### L'argumentation logique

Les modèles logiques utilisent les principes de l'argumentation. Le *coup* est gagné pour le locuteur lorsqu'il arrive à réfuter les arguments de l'allocutaire ou à le convaincre des siens. Le principe général est illustré par l'exemple suivant : supposons que le locuteur A affirme  $p \wedge q$ , l'allocutaire B peut réfuter p (en prouvant  $\neg p$ ) ou réfuter q ou demander à A de prouver p et de prouver q. Supposons que la stratégie de B soit de demander à A la preuve de p et q. A peut alors répondre : p parce que  $a \supset p$  et q parce que  $p \wedge b \supset q$ . Ce qui poussera ensuite B à demander à A de prouver a et b ou de prouver lui-même que  $\neg a$  ou  $\neg b$ . Et ainsi de suite jusqu'à l'arrêt du processus du jeu. Comme on le voit le problème se pose en termes de logique d'une part mais aussi en termes de stratégie d'autre part, puisque B peut perdre s'il fait le mauvais choix de stratégie en n'attaquant pas le bon argument.

Pour Baker [41] le jeu de langage est une négociation, dans laquelle les locuteurs ont des attitudes sur des propositions. Les attitudes de base du modèle sont l'offre, l'acceptation, la ratification. Le dialogue se déroule à partir de séquences types :

Séquence de base	Séquence étendue
OFFER(A, p)	OFFER(A, p)
ACCEPT(B, p)	OFFER(B, q)
RATIFY(A, p)	ACCEPT(A, TF(p))
	Etc... RATIFY(X, z)

Dans ces séquences, par exemple la séquence étendue, le locuteur A présente un prédicat p et l'allocutaire B le sien q. Ils essaient de se mettre d'accord en transformant progressivement et à tour de rôle, les prédicats p et/ou q. Les TF (transformations fonctionnelles) dont ils disposent sont de quatre types : les expansions logiques, les contractions, les neutres et les fondamentales. Par exemple on aura un dialogue comme :

A : hier, je suis allé à Grenoble  
 B : non ce n'était pas hier, c'était la semaine dernière  
 A : ah oui, pardon, la semaine dernière.  
 Qui se traduit par la séquence :  
 OFFER(A,  $p_1 \wedge p_2$ )  
 ACCEPT(B,  $p_2$ ) et OFFER(B,  $p'_1 = TF(p_1)$ )  
 RATIFY(A,  $p'_1$ )

Pour Dessalles [42] l'enchaînement du dialogue s'opère sur trois modalités, IMPR = improbable, IND = indésirable et FAUX = paradoxal. La conversation débute par l'exposition d'un fait le plus inhabituel possible comme « tu sais que Jérôme a acheté une voiture cette semaine ? » qui présuppose qu'il est étonnant que Jérôme ait fait cet achat (il est peut-être écologiste) ou qu'il a l'habitude de déléguer les achats à sa femme, etc. A cet événement IMPR la réplique attendue est soit de marquer son étonnement (« ah bon »), soit de le banaliser (« mais non c'est pour son frère »). De même pour les événements IND ou FAUX. Le déroulement peut alors être décrit par une grammaire :

Exposé-fait  $\rightarrow$  Co-étonnement  $\cup$  Demande-clarification  
 Exposé-fait  $\rightarrow$  Réaction-antagoniste  $\cup$  Banalisation  
 Réaction-antagoniste  $\rightarrow$  Exposé-fait(opposé)  
 Banalisation  $\rightarrow$  Exposé-fait(modérateur)  
 Co-étonnement  $\rightarrow$  fin-dialogue  $\cup$  Exposé-fait(nouveau)

La demande de clarification est considérée comme une incidence dans ce modèle.

Ces deux modèles reposent sur une « bonne » interprétation des prédicats par les interlocuteurs. Or ce problème mériterait un long développement : c'est le problème des implicatures conversationnelles sur lequel déjà Grice avait été critiqué. Rappelons que les implicatures conversationnelles sont l'ensemble des inférences que l'on peut faire à partir des énoncés d'une conversation. En effet le langage naturel a une puissance d'expression qui dépasse celle des logiques standards, notamment par l'utilisation de l'implicite, de l'ellipse, de la référence multiple, de la sémantique polyvalente, etc.

On distingue les implicatures directes et les implicatures indirectes. Dans l'exemple :

A : As-tu invité Jean et Pierre pour ce soir ?  
 B : Oui, j'ai invité Pierre.  
 on peut déduire directement que invité(Pierre), mais la réponse laisse à penser aussi que  $\neg$ invité(Jean) (principe d'omission) ou que même

–vouloir(B, inviter(Jean)) ou préférer(B, inviter(Pierre)), etc. Il est clair que ces implicatures indirectes ne peuvent pas toujours se faire sans une connaissance du contexte et des conversants. Dans l'exemple :

A : As-tu fait les courses ?

B : Ma voiture est tombée en panne.

les connaissances de sens commun peuvent suffire sans connaître nécessairement les conversants : il est hautement probable que la réponse de B signifie « non », bien que la voiture ait pu tomber en panne après avoir fait les courses. Pour interpréter la réponse de B il faut donc analyser les attentes de A contenues dans sa question.

Green et Carberry [43] proposent des plans par défaut pour résoudre ces problèmes, comme Use-contrast(), Use-obstacle(), Use-otherwise(), etc. Ces plans sont indépendants des conversants. Hinkleman [44] inclut un niveau syntaxique et sémantique dans l'interprétation de tels actes en affirmant que les plans ne peuvent résoudre convenablement le problème sans introduire le rôle des conversants. Enfin, Zuckerman [45] propose d'examiner les effets des implicatures sur les connaissances mutuelles des conversants. Pour cela il génère toutes les implicatures possibles au fil du dialogue. Des heuristiques sont alors nécessaires pour limiter la combinatoire.

On le voit donc, un jeu logique ne peut se développer que si les prédicats (ou les modalités) sont bien posés et bien compris tout au long du jeu. Il peut y avoir rapidement incompréhension ou blocage si les inférences effectuées par les locuteurs ne sont pas les bonnes, dues à leur manque de connaissances mutuelles ou à leur appréciation différente de la situation. C'est le point fragile de ces modèles qui ont un niveau de traitement complexe des implicatures.

### **Les modèles stratégiques**

Les modèles stratégiques [48] restent centrés sur les buts comme dans les modèles de planification, mais le jeu consiste à optimiser la stratégie pour les atteindre. Nous supposons qu'il y a deux locuteurs qui dialoguent et qu'au départ chacun vise un certain but dans l'arrière-plan [46]. Nous noterons A pour locuteur et B pour allocutaire. Leurs buts seront notés,  $b_A$  et  $b_B$ , l'un des deux pouvant être vide.

On définit :

But initial : l'état du monde ou l'état mental que l'un des deux locuteurs veut atteindre, soit pour lui-même (obtenir un renseignement, acquérir un savoir-faire, etc.), soit pour son partenaire (lui donner une information, lui faire-faire quelque chose, lui donner un conseil, etc.).

Echange : une suite de tours de parole pendant laquelle un but est maintenu. Le début d'un échange est marqué par l'apparition d'un nouveau but, ce but se transforme éventuellement au cours de l'échange (il peut s'affiner par exemple ou se décomposer en sous-buts) et devient un but final irréductible sur lequel l'échange se termine par un succès ou par un échec. Le succès obéit à la double condition d'être un *but atteint* et un *but satisfait* [47].

But de l'échange : celui qui est maintenu en jeu pendant l'échange.

But final : l'état du monde ou de la situation à la fin d'un échange (cela se termine toujours, au moins par l'accord des deux interlocuteurs sur le fait qu'il y a échec quand il y a échec : « les syndicats et le patronat se sont séparés sur un constat d'échec »). Le but final n'est pas toujours prévisible au départ.

Incidence : un acte de parole qui a pour effet de mettre un but en attente ou en question (par changement de thème, demande de clarification, demande de précision, etc.) mais ne remet pas en cause le but conversationnel de l'échange.

Stratégie de dialogue : la manière de gérer les tours de parole entre interlocuteurs pour conduire un échange ou une incidence. La stratégie vise à choisir la meilleure direction d'ajustement des buts à moment donné.

Direction d'ajustement : il y a 5 directions d'ajustement possibles des buts qui conduisent à 5 types de stratégies :

- B abandonne son but au profit de celui de A (stratégie réactive), en d'autres termes B ajuste son but sur celui de A (en abrégé  $b_B \rightarrow b_A$ )
- B impose son but au détriment de celui de A (stratégie directive), en d'autres termes il force B à adopter son but (en abrégé  $b_B \leftarrow b_A$ )
- A et B gardent chacun leur but (stratégie de négociation), en d'autres termes ils n'essaient pas d'ajuster leurs buts *a priori* (en abrégé  $b_A \leftarrow b' \rightarrow b_B$ ) même si à l'issue de la négociation un compromis  $b'$  est trouvé
- A et B tiennent compte du but de l'autre (stratégie de coopération), en d'autres termes ils essaient de les ajuster l'un à l'autre (en abrégé  $b_A \leftrightarrow b_B$ )
- A et B abandonnent leurs buts au profit d'un troisième (stratégie constructive), en d'autres termes ils font un détour constructif (en abrégé  $b_A \rightarrow b' \leftarrow b_B$ )

On peut alors définir les types de stratégies suivants [48] (on se place dans la suite, du point de vue de l'allocutaire B) :

*1. Stratégie réactive*

Consiste pour B à déléguer l'initiative à A soit en lui faisant endosser son but (cas de demande d'aide ou d'assistance), soit en adoptant son but (cas du serviteur). Le déroulement du dialogue se fait :

- en maintenant le but de l'échange, mais sans prendre d'initiative,
- en abandonnant son propre but  $b_B$  ou en le faisant passer sous la dépendance de  $b_A$ .

Le sens d'ajustement est alors  $b_B \rightarrow b_A$

*2. Stratégie directive*

Consiste pour B à garder l'initiative pour conduire le dialogue :

- en maintenant le but de l'échange et en gardant l'initiative,
- en imposant son but  $b_B$ , (donc on cherche à ce que  $b_I = b_B$ )
- en ignorant éventuellement celui du locuteur  $b_A$ , qui est donc en quelque sorte considéré comme inexistant

Cela a pour conséquence d'imposer une réponse réactive ou négociée à A, et de limiter ainsi la variété de ses stratégies.

*3. Stratégie constructive (ou du détour)*

Consiste à déplacer le but courant momentanément afin de provoquer un détour (supposé constructif) qui n'est pas nécessairement une incidence, par exemple pour faire remarquer un oubli, une erreur, faire une citation, rappeler un fait ancien, une expérience, etc. :

- le but courant est mis en attente, ainsi que les buts initiaux,
- un nouveau but  $b'$  est posé,
- l'initiative peut être partagée.

Contrairement à une incidence, un détour ne ramène pas nécessairement à l'échange initial, il peut laisser la conversation en suspens ou conduire à un autre détour<sup>1</sup>.

---

<sup>1</sup> Cette forme de dialogue était très utilisée dans la Chine ancienne.

#### 4. Stratégie de coopération

Consiste à tenir compte du but de son interlocuteur en lui proposant une (ou des) solution(s) qui les amènent tous deux à atteindre leurs buts, si ces derniers ne sont pas incompatibles :

- cela amène à dérouler un processus complexe — évaluer la situation, présenter une explication, éventuellement des exemples, des aides ou des arguments pertinents et offrir un choix fermé (parce que plus facile au plan cognitif pour la prise de décision), en maximisant l'espace de concession,
- en procédant par recherche d'un optimum dans un espace de possibles,
- en accompagnant l'interlocuteur jusqu'à la solution,
- en élargissant le but conversationnel si nécessaire,

#### 5. Stratégie de négociation

La négociation peut se produire dans une situation où les buts sont incompatibles et que les interlocuteurs veulent minimiser les concessions. La négociation procède sur un schéma assez classique, par des séquences argumentatives (argumentation/réfutation) avec proposition d'une solution sous-optimale jusqu'à convergence ou constat d'échec. La tactique locale est de :

- tenter d'imposer son but ou accepter un compromis,
- maintenir le but conversationnel,
- pousser la négociation le plus loin possible jusqu'à un but acceptable  $b_f$ ,

Efficacité de la stratégie : une stratégie est efficace si la vitesse de convergence des actes de parole vers le but final est optimale. Le critère d'efficacité qui s'en déduit, est de réduire au maximum à chaque tour de parole, la distance au but final (il peut cependant y avoir des écartements locaux pour éviter de tomber dans des impasses ou provoquer des échecs).

La pertinence d'un acte de dialogue énoncé à l'instant  $t$  par A (et donc de la stratégie associée), est relative au but poursuivi par A mais aussi relative pour B, au but que B poursuit de son côté. On distingue donc deux cas, (a) *partage* et (b) *concurrence* :

- dans le cas où le but est partagé, la pertinence d'un acte  $\alpha_A$  de A à l'adresse de B, doit amener A et B dans une situation de convergence (ou les maintenir dans cette situation s'ils y étaient déjà) et contribuer à les rapprocher du but final,

• dans le cas où le but n'est pas partagé, on doit distinguer une série de scénarios possibles :

(a) soit A et B engagent une négociation qui peut réussir ou échouer, du point de vue de l'un ou de l'autre (ou des deux),

(b) soit A et B restent sur leurs positions car l'un des deux a abandonné son but au profit de l'autre ou a fait un détour ;

dans les deux cas, la stratégie de A vis-à-vis de B est alors d'arriver à  $b_A$  ou d'empêcher que B n'arrive à  $b_B$ . La pertinence des actions de A peut prendre alors un sens négatif pour B. Nous appelons cette pertinence *pertinence transactionnelle*. Elle prend une valeur comprise entre 0 et 1 et peut être définie comme suit :

• pertinence constructive :

$$P_C\{\alpha_A\} = \exp\{-d[b_A, \text{effet}(\alpha_A)]\}$$

où  $d$  est une distance<sup>2</sup> qui mesure l'écart entre le but recherché par A et le but atteint par l'effet de l'acte  $\alpha_A$  ; il suffit de considérer les effets potentiels de l'acte  $\alpha_A$  pour juger de la pertinence et non de son exécution réelle, puisqu'elle est relative au but de l'énonciateur seul.

• pertinence obstructive :

$$P_{\sim C}\{\alpha_A\} = 1 - \exp\{-d[b_B, \text{effet}(\alpha_A)]\}$$

ici A s'oppose à la réalisation du but  $b_B$  de B.  $d$  est une distance qui mesure l'écart entre le but recherché par B et le but atteint par l'effet de l'acte  $\alpha_A$  ; si cette distance est grande, l'acte  $\alpha_A$  aura été pertinent puisqu'il éloigne B de son but.

La conduite du jeu se fait alors à chaque instant pour A en maximisant  $P_C\{\alpha_A\}$  et en maximisant  $P_{\sim C}\{\alpha_A\}$ . Les deux objectifs étant rarement atteints simultanément il faut chercher un point d'équilibre (dit équilibre de Nash).

## 4 DISCUSSION GENERALE

Nous avons parcouru un ensemble de modèles, des plus rigides aux plus souples. Les plus « rigides » s'appuient sur la structure du dialogue et appellent des traitements descendants, les plus « souples » tentent de s'adapter au plus bas niveau des actes dialogiques, considérant le dialogue

---

<sup>2</sup> La distance peut se mesurer par exemple en nombre de tours de parole pour atteindre le but.

comme un jeu ayant ses règles et ses stratégies. Entre les deux se trouvent des modèles qui empruntent à des méta-structures comme les plans et qui projettent les actes sur d'autres représentations comme les états mentaux ou les intentions, pour trouver un appui fonctionnel.

Les méthodes structurelles sont des méthodes essentiellement descriptives et fonctionnelles, elles modélisent le comment dialoguer mais non le pourquoi des intentions. Il n'est pas possible de savoir quel but poursuit le locuteur, ni quelle stratégie mettre en place pour améliorer l'interaction. Il est difficile également de modéliser des dialogues de négociation, nécessitant un niveau d'argumentation évolué. Cependant la réalisation d'applications de DHM est relativement aisée avec ces méthodes qui peuvent être implémentées par des automates et utiliser des outils généraux existants à base de grammaires. De nombreux modèles de DHM ont été implémentés avec cette technique pour des applications de renseignement. Mais ils montrent rapidement leurs limites dans des tâches dites *innovatives* qui nécessitent de nombreuses incidences, reprises, mises en attente, et négociations. Le plan d'action ne peut pas être représenté en extension dans ce type de tâche, il en résulte de nombreux inattendus dialogiques et des ruptures qui dé-structurent le dialogue.

Avec les modèles de planification et les modèles mentaux, on suppose que les interlocuteurs construisent leurs dialogues autour d'actions coordonnées, en respectant des conventions normalisées. On suppose en général que le dialogue est à la fois constructif (il conduit à la construction d'un but à partir des objectifs des interlocuteurs et à l'enrichissement des connaissances mutuelles) et co-interactif (les acteurs coordonnent leurs actions pour aboutir à un certain but). La logique de dialogue développée est double, actionnelle et épistémique pour traduire les hypothèses sur les *états mentaux* des conversants en terme de *croyances* et d'*engagement*. Les états mentaux résument l'état de connaissance (croire), le contexte de l'action (faire), les choix (pouvoir, devoir) et les engagements (vouloir). A chaque instant du déroulement du dialogue, le modèle fait l'hypothèse que le dialogue est dirigé par les états mentaux qui sous-tendent les actes (intentions, choix, engagements, etc.).

L'approche générale de ces modèles est fondée sur la rationalité. On suppose que les conversants sont conscients des buts qu'ils poursuivent (leurs désirs sont rationalisés), maîtrisent les connaissances dont ils disposent (ils savent les évaluer, les faire partager, les remettre en question, les rendre mutuelles, etc.) et respectent les conventions sociales [4], [5] à travers les rôles qu'ils jouent (et dont ils ont aussi conscience). Dans le cadre du dialogue humain, on peut trouver une telle modélisation très limitative : elle ne fait pas, en effet, la place aux inattendus (la

planification et l'intention prennent une place dominante) et place les désirs sous le contrôle de la rationalité (ces désirs sont non seulement rationalisés mais aussi conscients). Mais en DHM cette réduction peut être intéressante dans la mesure où, malgré tout, elle assigne une place au concept d'états mentaux et donc à une composante cognitive dans un dialogue pseudo-naturel. Le contrôleur du dialogue a alors pour mission de gérer les actes de l'interlocuteur reflétant ses états mentaux supposés. Cette modélisation ne vise pas à donner un comportement de type humain à la machine mais seulement à lui fournir des éléments logiques de choix et des raisons d'agir. On ne cherche pas à donner des "états mentaux" à la machine mais à modéliser les raisons (à cause de) d'action de l'interlocuteur.

Avec les modèles de jeu, on utilise une métaphore reprise à la fois des jeux de langage de Wittgenstein [49] mais aussi de la théorie des jeux. Le problème est de trouver la bonne métaphore pour un cadre de dialogue donné. Faut-il considérer l'interlocuteur comme un être logique ou un être de désir ou un être social ? Nous avons vu trois modèles différents à ce propos, il en existe évidemment d'autres [50]. La question se pose aussi pour la généralité de ces modèles : ne risque-t-il pas d'y avoir un trop grand nombre de jeux différents et partant autant de modèles ? Au-delà de ces questions, l'intérêt de tels modèles métaphoriques est qu'ils peuvent apporter plus de précision pour la métaphore ou le domaine ciblé - au prix peut-être de leur généralité - que les autres modèles.

## 5 CONCLUSION

Nous avons montré que toutes les théories de départ empruntées au monde pluridisciplinaire des sciences du langage et de la communication ont été également fécondes pour modéliser le dialogue. Les chercheurs en IA ont tenté de combiner ces théories pour obtenir des modèles efficaces en empruntant des techniques et des paradigmes très divers allant des automates à la logique en passant par la planification.

## 6 RÉFÉRENCES

- [1] A. Turing, *Computing Machinery and Intelligence*. Mind Vol. LIX, 1950.
- [2] H. Garfinkel. *Studies in ethnomethodology*. New Jersey, Prentice Hall, Englewood Cliffs, 1967.

- [3] H. Garfinkel, H. Sacks, On formal structures of practical actions, in Mc Kinney, J.C., Tiryakian, E.A. (Ed.) *Theoretical Sociology. Perspectives and Development*, New York, Appleton Century Crofts, 337-66, 1970.
- [4] H. Sacks, Everybody has to lie, in *Sociological dimensions of language use*, M.Sanchez et B.Blount (Ed.), USA, New York, Academic Press, 1975.
- [5] E.A. Schegloff, G. Jefferson, H. Sacks, The preference for self-correction in the organisation of repair in conversation. *Language*, 53, p. 361-382, 1977.
- [6] E.A. Schegloff, Sequencing in conversational openings, *American anthropologist*, N°70, p.1075-95, 1968.
- [7] E. Goffman, *Forms of talk*. Oxford, Basil Blackwell ed., 1981.
- [8] L.S. Suchman, *Plans and Situated Actions*, Cambridge University Press, 1987.
- [9] J.L. Austin, *How to do Things with Words*. Oxford University Press, 1962. Version française : *Quand dire, c'est faire*. Edition du Seuil, Paris, 1970.
- [10] J. Searle, *Speech Acts*. Cambridge University Press, UK., 1969. traduit de l'anglais *Actes de langage*, Hermann ed., Paris, 1972.
- [11] J Searle, *Intentionality*. Cambridge: Cambridge University Press, 1983.
- [12] J. Searle, D. Vanderveken, *Foundations in illocutionary logic*. Cambridge University Press, 1985.
- [13] H.P. Grice, Logic and Conversation. in P. Cole and J.L. Morgan eds. *Syntax and Semantics*, vol. 3, Academic Press, p. 41-58,1975.
- [14] D. Sperber, D. Wilson, *Relevance : communication and cognition*. Basil Blackwell ed., Oxford, 1986.
- [15] E. Roulet, *L'articulation du discours en français contemporain*. Peter Lang éd., Berne, 1985.
- [16] J. Moeschler, *Modélisation du dialogue, représentation de l'inférence argumentative*. Hermès éd., 1989.
- [17] J. Moeschler, A. Reboul, *Dictionnaire encyclopédique de pragmatique*, éditions du Seuil, Paris, 1994.
- [18] D. Luzzati, *Recherches sur le dialogue Homme-Machine: modèles linguistiques et traitement automatique*. Thèse d'Etat, Paris III, 1989.
- [19] E. Bilange, *Dialogue personne-machine*, éd. Hermès, Paris, 1992.
- [20] G. Sabah, *l'IA et le langage*, éd. Hermès, Paris, 1989.

- [21] A. Vilnat, L. Nicaud, Un système de dialogue homme-machine : STANDIA. Actes du séminaire Dialogue du GdR-PRC Communication homme-machine, Dourdan, 1992.
- [22] J. Lehuen Un modèle de dialogue dynamique et générique intégrant l'acquisition de sa compétence linguistique. Thèse de doctorat de l'université de Caen, 1997.
- [23] D. Luzzati, *Le dialogue verbal homme-machine*. Masson éd., Paris, 1995.
- [24] J.F. Allen, CR Perrault. Participating in Dialogues Understanding via Plan Deduction. In: Proceedings of the 2nd National Conference of the Canadian Society for Computational Studies of Intelligence, Toronto, Canada, 1978.
- [25] P.R. Cohen, On knowing what to say : Planning speech acts. Ph.D. Thesis, Technical Report n°118, Department of Computer Science, University of Toronto, January 1978.
- [26] J.F. Allen, C.R. Perrault. Analyzing intentions in utterances. *Artificial Intelligence*, 15(3), p. 143-178, 1980.
- [27] J. Allen. Recognizing Intentions From Natural Language Utterances. In: Computational Models of Discourse Chapter 2 (M Brady and R Berwick eds.), Cambridge, MA: MIT Press, 1982.
- [28] M. Bratman, *Intentions, plans and practical reason*. Cambridge MIT Press, 1987.
- [29] B Grosz, C Sidner. *Plans for Discourse*. In: Intentions in Communication (J Morgan, P Cohen, M Pollack eds.), Cambridge, MA: MIT Press, 1989.
- [30] M. Guyomard, P. Nerzic, J. Siroux. Plans métaplans et dialogue. Actes de la 4ème école d'été sur le Traitement des langues naturelles. Lannion, 1993.
- [31] D.R. Traum, JF Allen, PA Heeman, CH Hwang, N Martin, M Poesio, LK Schubert. Integrating Natural Language Understanding and Plan Reasoning in the TRAINS-93 Conversation System. Unpublished Manuscript, Computer Science Department, University of Rochester, 1996.
- [32] H.J. Nilsson. Principles of artificial intelligence. Tioga ed., Palo Alto, CA, USA, 1980.
- [33] R. Wilensky. *Planning and Understanding. A computational approach to human reasoning*. Reading, Addison-Wesley Publishing Company, 1983
- [34] D.J. Litman, J.F. Allen. Discourse processing and commonsense plans. in P.R. Cohen, J.L. Morgan, M.E. Pollack (eds.), *Intentions in communication*. Cambridge MIT Press, 1990.
- [35] S. Carberry, *Plan Recognition in Natural Language Dialogue*, MIT Press, Cambridge, 1990.

- [36] P. Nerzic. Erreurs et échec dans le dialogue oral homme-machine: détection et réparation. *Thèse de doctorat de l'université de Rennes*, 1993.
- [37] M.E. Pollack, Plans as complex mental attitudes. In *Intentions in Communications*. Bradford/MIT Press, Cambridge, Mass. 1990.
- [38] D. Sadek, Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication. Thèse de doctorat de l'université de Rennes I (IRISA), 1991.
- [39] D. Sadek. Le dialogue homme-machine : de l'ergonomie des interfaces à l'agent intelligent dialoguant. OFTA, ARAGO 18 Nouvelles interfaces homme-Machine, Lavoisier éd., Paris, p. 277-321, 1996.
- [40] J. Levin, J. Moore, Dialogue-games : meta-communication structure for natural language interaction. *Cognitive sciences*, 1(4), p. 395-420, 1980.
- [41] M. Baker. A model for negociation in teaching-learning dialogues. *Journal of Artificial Intelligence in Education*, 5(2), p. 199-254, 1994.
- [42] J.L. Dessalles. Modèle cognitif de la conversation spontanée appliqué à l'apprentissage. Thèse de doctorat de l'ENST, Paris, 1993.
- [43] N. Green, S. Carberry, Conversational implicatures in indirect replies. Proc. of the Annual Meeting of the ACL, p. 64-71, 1992.
- [44] E.A. Hinkleman, J.F. Allen, Two constraints of speech act interpretation. Proc. of the ACL, p. 212-219, 1989.
- [45] I. Zuckerman, Content planning based on a model of a user's beliefs and inferences. Proc. of the *Third international Workshop on User Modeling*, p. 162-173, 1992.
- [46] D. Vernant, Modèle projectif et structure actionnelle du dialogue informatif. in *Du dialogue, Recherches sur la philosophie du langage*, Vrin éd., Paris, n°14, p. 295-314, 1992.
- [47] D. Vanderveken, *La logique illocutoire*. Mandarga éd. Bruxelles, 1990.
- [48] J. Caelen, A.L. Fréchet. Attitudes cognitives et actes de langage. in *Du dialogue, Recherches sur la philosophie du langage*, Vrin éd., Paris, n°14, p. 19-48, 1992.
- [49] L. Wittgenstein, *Philosophical Investigations*. Oxford University Pres, London, England, 1953.

- [50] N. Maudet, Modéliser les conventions des interactions langagières : la contribution des jeux de dialogue. Thèse de doctorat de l'université Paul Sabatier, Toulouse, 2001.