

DIALOGUE HOMME-MACHINE ET INTERACTION VERBALE : une logique dialogique

J. Caelen



CLIPS

**Communication Langagière et
Interaction Personne-Système**

CNRS - INPG - UJF

BP 53 - 38041 Grenoble Cedex 9 - France

PLAN

INTRODUCTION.....	3
1. ACTION ET COMMUNICATION.....	3
2. LE DIALOGUE HOMME-MACHINE	6
3. UN CADRE PRAGMATIQUE.....	8
3.1. LES MODES	9
3.2. LES INTERLOCUTEURS	9
3.3. LES CONNAISSANCES	9
<i>Axiomes.....</i>	<i>11</i>
<i>Représentation du monde en machine.....</i>	<i>11</i>
<i>Apprentissage.....</i>	<i>14</i>
3.4. CONTEXTE OPERATIF	14
<i>Gestion des buts</i>	<i>15</i>
<i>Activité, actions.....</i>	<i>16</i>
<i>Axiomes.....</i>	<i>16</i>
<i>Apprentissage.....</i>	<i>17</i>
4. UN MODELE DE DIALOGUE.....	18
4.1. LES ACTES DE DIALOGUE.....	18
4.2. LES STRATEGIES DE DIALOGUE.....	19
<i>Stratégies non inférentielles.....</i>	<i>19</i>
<i>Stratégies inférentielles</i>	<i>20</i>
<i>Axiomes.....</i>	<i>21</i>
5. LE CONTROLE DU DIALOGUE	23
<i>Axiomes.....</i>	<i>25</i>
<i>Au départ</i>	<i>25</i>
6. EXEMPLE DE DIALOGUE HOMME-MACHINE	25
CONCLUSION	26
BIBLIOGRAPHIE.....	27

Introduction

De nombreuses disciplines se sont intéressées au dialogue humain. Les cognisciences retiennent de la communication les aspects liés à la perception, à la planification et au raisonnement du point de vue de *l'individu* ; la philosophie s'intéresse à cet *individu* placé en situation de communication, sur un plan *intentionnel* et *actionnel*. L'éthnoscience 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. Quant aux technosciences, elles visent à intégrer la machine dans un univers de "communication" homme-machine, soit comme un canal de régulation entre des interlocuteurs humains, soit comme un assistant dans la tâche. Dans ce dernier cas, la machine — pseudo-interlocuteur humain — est assujettie à *comprendre* pour participer et collaborer au mieux à la tâche de l'utilisateur : c'est dans ce cadre que se situe la plupart des travaux en DHM (Dialogue Homme-Machine). Ce cadre entretient toutefois une grande ambiguïté : quel rôle assigne-t-on exactement à la machine ? Quelles capacités doit-elle posséder *in fine* ? Quels sont les objectifs réels de la DHM ?

Le postulat implicite que le dialogue est une activité rationnelle entretient l'espoir que la machine peut la simuler utilement : le dialogue s'appuierait alors sur des principes logiques et ne se distinguerait pas du raisonnement et autres activités humaines. Le DHM relèverait dès lors du domaine de l'I.A. (Intelligence Artificielle). Il y aurait des buts à atteindre, une activité à planifier selon des contraintes données par la situation, des connaissances à utiliser et des données à manipuler. Le dialogue ne serait alors qu'une forme d'interaction permettant à une personne A d'utiliser son interlocuteur B comme participant à son propre plan (faire-faire ou faire-faire-savoir) en vue d'un certain but. Il est de fait que A préférerait dans ce cas que B soit une machine qui n'offrirait pas autant de résistance qu'un humain à sa volonté, à condition toutefois que son niveau de compréhension soit suffisant. Mais en est-il réellement ainsi ? Voudrait-on ne parler qu'à des machines ?

Il est clair qu'il n'y a aucun intérêt à vouloir *convaincre*, *influencer*, *apitoyer*, etc. une machine. Il y aura toujours une limite *perlocutoire* au DHM : "haut les mains" sera toujours un acte sans effet pour une machine, aussi bien que "les femmes et les enfants d'abord". Il n'y aura guère d'intérêt non plus à utiliser les actes indirects ou les actes expressifs. Si donc le DHM peut offrir une situation artificielle de dialogue, ce sera toujours un dialogue à fonction purement *illocutoire* — limitée au surplus aux assertifs et directifs voire aux déclaratifs.

1. Action et communication

La plupart des études sur la communication humaine nous viennent des sciences qui se sont intéressées au langage : ce n'est donc peut-être pas avec ce parti pris fortement tourné vers le langage qu'il faut examiner la question de la communication homme-machine, il n'est pas évident non plus que le dialogue soit la meilleure forme d'interaction de l'homme et de la machine : on connaît depuis peu la manipulation directe

offerte par la souris sur des objets graphiques ou des menus déroulants.

Presque implicitement, dès qu'on évoque la communication entre personnes humaines (disposant de la parole), on pense à la communication verbale ; celle-ci sous-tend une communication de type langagier. Pourtant il y a de nombreuses autres formes d'interaction entre l'homme et son environnement : l'humain ne semble marquer une préférence pour le langage que lorsqu'il communique avec son semblable (parfois aussi avec certains animaux domestiques !). Par la langue il construit des ordres, transmet des informations, formule des raisonnements, coordonne ses actions et rythme ses efforts (chants de marin par exemple). Il dispose de plusieurs modalités sensorielles qu'il utilise de manière redondante ou complémentaire. La multimodalité, naturelle chez l'homme, semble donc être au service *d'un* langage plutôt que *du* langage.

L'action s'accompagne de nombreuses formes de langage — Vernant va jusqu'à penser que le langage doit être analysé à travers une logique de l'action — geste, parole, écriture, etc. Ces formes sont plus ou moins adaptées aux situations d'usage et structurées dans un (des) code(s) dialogique(s). Cela pose la finalité de la communication dans la satisfaction de buts actionnels et conversationnels. Communiquer revient alors utiliser une rhétorique du faire-faire et du faire-faire-savoir adaptée à son interlocuteur. Posée en des termes analogues, la CHM (Communication Homme-Machine) a peut-être plus de chances de succès, même si cette rhétorique est entièrement artificielle et doit être apprise.

De quelles sources la CHM peut-elle s'inspirer ?

- Les éthnométhodologues [Garfinkel, 1967], [Heritage, 72] 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 chaque individu. En d'autres termes pour les éthnométhodologues, 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 " [Bange, 1987], " un système de comportements de différents partenaires qui *s'influencent* réciproquement dans des actions concrètes " [Cranach, 1980]. 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 de 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 [Schütz, 1962]. 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.

Même en supposant que la machine soit un outil *socialisé*, il n'est guère envisageable de la considérer à travers ce type de rhétorique de la *moralité*. On n'entretient guère d'attitudes morales avec une machine que les formes de politesse laissent complètement indifférente.

- Certains linguistes voient 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 veulent créer 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 " [Moeschler, 1987]. 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.

L'approche entièrement ascendante est souvent utilisée en linguistique computationnelle et en TALN (Traitement Automatique du Langage Naturel). Il se trouve que de nombreux phénomènes (ellipses, anaphores, déictiques, etc.) ne peuvent être pris en compte sans référence à la situation. Les interdire conduit à des rejets du dialogue homme-machine par les utilisateurs. Il y a donc toujours tôt ou tard un niveau de résolution des références qui prend en compte la situation.

Dérivés des modèles structurels de dialogue [Roulet, 85], de nombreux modèles de DHM ont été implémentés [Bilange, 92], [Sabah, 90] pour des applications de renseignement. Ces modèles montrent rapidement leurs limites dans des tâches de conception qui nécessitent de nombreuses incidences, reprises, mises en attente, dans le dialogue. Le plan d'action ne peut pas être représenté en extension dans ce type de tâche, il en résulte de nombreux inattendus dialogiques, des ruptures, qui déstructurent le dialogue.

- Jugeant des insuffisances d'une sémantique générale, certains philosophes du langage comme Austin et Searle [Austin, 62], [Searle, 69] mettent l'accent sur la pragmatique. Searle replace la communication 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.

En (c) on peut représenter tout acte de langage par $F(p)$, 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"). A ce niveau Searle ne prend pas encore en compte le rôle du destinataire ni celui du destinataire qu'il ne fait apparaître qu'en (d). Ceci lui a été reproché puisque cela implique jusqu'à (c), l'élimination du locuteur parlant au profit d'un locuteur abstrait et donc une difficulté pour l'interprétation des actes indirects qui se font par énonciation non littérale. Searle propose 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. Récanati propose d'en rester pour le niveau (c) à des *potentialités* illocutoires qui ne seront instanciées qu'au niveau (d) après la prise en compte des sujets parlants.

Formulée ainsi, il semble que cette théorie propose un lieu d'articulation de l'illocutoire ou perlocutoire intéressant pour le DHM. En effet il semble bien que les niveaux (a), (b) et (c) soient réellement présents en DHM (résultats d'expériences menées en technique Magicien d'Oz) — avec un appauvrissement du niveau (c) — et absent pour le niveau (d). Cela supposerait qu'un sujet puisse mettre en œuvre ses capacités linguistiques sans égard pour son destinataire dans un premier temps, puis que des processus d'ajustement soient appelés dans un second temps pour contraindre les formes de productions langagières selon les contraintes de la communication. Il semble bien qu'il puisse en être ainsi puisqu'on constate qu'un utilisateur humain formule ses commandes vis-à-vis d'une machine en utilisant toutes ses compétences linguistiques.

- Pour Sperber et Wilson [Sperber et Wilson, 89] — à la suite de Grice [Grice, 75] —, la communication est un " échange d'indices qui orientent (ou réorientent) les processus inférentiels des interlocuteurs en présence ". 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 du code" 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 $\text{intention}(A) \subset S$)
- (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 $\text{intention}(A) \subset S$) 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 $\text{intention}(\text{intention}(A) \subset S) \subset S'$ 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 : c'est cela qui a amené Grice à énoncer le principe des *maximes*. On ne voit pas comment pourrait s'arrêter 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.

Prêter des capacités cognitives à une machine pour tenter d'intégrer cette théorie au DHM ne semble guère raisonnable. Par contre il est peut-être envisageable de tenir compte de l'usager dans la phase de génération des réponses de la machine. A ce moment c'est effectivement une certaine pertinence qui doit servir de guide pour la production des réponses de la machine. Un biais peut d'introduire cependant : c'est de faire croire que la machine a plus de capacités qu'elle n'en a vraiment ; cela conduit à des mauvaises interprétations de la part de l'usager.

2. Le dialogue homme-machine

Le terme " dialogue homme-machine " semble abusif relativement au sens dévolu au dialogue humain. La machine n'est pas un être *social*, n'a pas d'intention ni de culture. Selon les communautés scientifiques dans lequel il est employé, ce terme ne contient pas les mêmes concepts : pour les concepteurs d'interfaces graphiques il se réduit presque exclusivement à la manipulation directe d'objets et au déroulement de menus, pour les chercheurs en LN (Langage Naturel) il recouvre certains présupposés du dialogue entre humains, du moins les capacités de compréhension du langage. Personne ne sait exactement ce que doit être le dialogue homme-machine puisque aucune machine ne permet de véritable dialogue — seules des expériences en Magicien d'Oz ont pu en donner quelques contours.

Il faut donc définir une *sémiotique* claire et non ambiguë de l'interface pour offrir un cadre fécond à la *communication homme-machine*. Il est évident que les capacités limitées de la machine ne peuvent conduire à une sémiotique générale, puisqu'il serait nécessaire que la machine soit un système non seulement intelligent mais aussi social (inséré dans le monde). Les performances de la machine en reconnaissance de la parole, reconnaissance du geste, vision par ordinateur, etc., semblent intéressantes *a priori* bien qu'encore limitées ; en utilisant ces nouveaux médias on peut escompter ainsi améliorer l'efficacité de l'interaction (entrées de plusieurs commandes simultanément), sa fiabilité (utilisation de la redondance), sa souplesse (choix des modes de communication les

mieux adaptés à la tâche), en un mot son ergonomie. Il faut cependant prendre garde au rôle dissymétrique que joue la machine dans la communication avec un humain et du biais que l'on peut introduire par une sémiotique mal définie : l'utilisateur, face à la machine, doit suppléer le manque de naturel, les limitations de compréhension, se contrôler dans ses productions langagières, inférer les réactions de la machine, adapter son comportement, etc. En fin de compte, l'attrait des interfaces multimodales pourrait se retourner contre elles si le concepteur complique inutilement les schémas de planification et les modèles de représentation de la tâche en machine.

Vis-à-vis de l'humain la machine est (et doit rester) un outil qui doit être le plus *ergonomique* possible, c'est-à-dire adapté et efficace : tout en reconnaissant que l'utilisateur humain garde ses habitudes linguistiques face à la machine, la communication homme-machine ne peut pas être un artefact de la communication humaine. Le DHM n'a semble-t-il, d'utilité que dans un cadre opératoire, c'est-à-dire pour effectuer des tâches coordonnées (résoudre des problèmes, renseigner, aider à la conception, assister l'enseignement, etc.). La communication se situe donc dans une relation opérateur/tâche où la machine a un rôle collaboratif [Falzon, 92]. Elle doit par conséquent avoir des capacités qui lui permettent de coordonner les processus actionnels en fonction des buts de l'usager. Cette relation opérateur/tâche place le dialogue homme-machine dans un *cadre actionnel*. A travers et par le dialogue la machine doit également apprendre de nouvelles actions ou optimiser son comportement face à de nouvelles situations. Pour cela elle doit pouvoir inférer et gérer les buts de l'utilisateur, comprendre ses actes de langage, être capable de les interpréter en fonction de la situation pour finalement générer et effectuer l'action ou le plan d'action le plus adéquat.

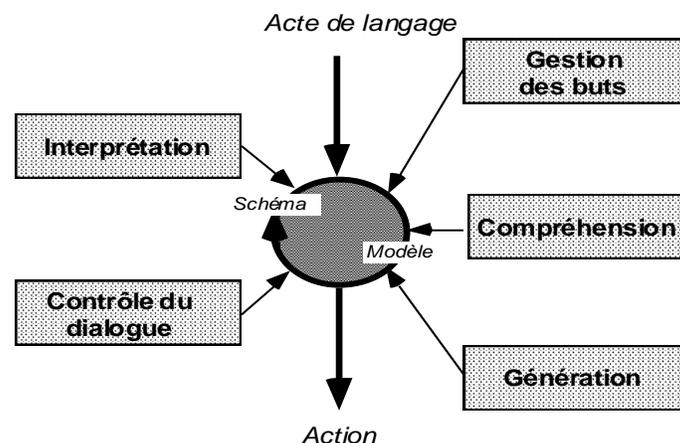


Fig. 1 : les processus inférentiels que doit posséder la machine. Un acte de langage est interprété en regard de la situation, essentiellement le plan d'action, but de la communication avec la machine. Cet acte est ensuite projeté dans un modèle par rapport auquel il est défini en "compréhension" pour finalement provoquer une action selon les performances de la machine manifestées dans une composante de "génération".

Nous définissons les schéma, modèle et action par :

Schéma = structure de données communes représentant les actes de langage, les connaissances et le plan d'action induit des actes de langage analysés sous l'angle de la compétence linguistique et du contexte actionnel,

Modèle = script sélectionné ou appris en fonction des connaissances sur l'usager et des connaissances pragmatiques (mondes d'arrière-plan, situation, etc.).

Action = réponse de la machine (éventuellement multimodale) en terme de changement d'état dans la situation et dans les connaissances.

En logique dialogique (LD) on suppose que les interlocuteurs construisent leurs dialogues de manière rationnelle — c'est-à-dire autour d'actions coordonnées — en respectant des conventions (sociales) 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 LD s'appuie sur la logique modale qui traduit des hypothèses sur les *états mentaux* des conversants. Ces états mentaux résument l'état de connaissance (savoir, croire), le contexte de l'action (faire, vouloir), les choix et engagements (pouvoir, devoir). 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.) [Cohen et al., 90]. Bien que d'inspiration mentaliste ce modèle peut être utile en DHM (dialogue homme-machine) : la machine prête des états mentaux à son interlocuteur ce qui ne signifie pas que l'interlocuteur a effectivement ces états mentaux ni même que ces derniers existent vraiment, ni encore moins que la machine a des états mentaux.

L'approche générale de cette théorie est fondée sur la rationalité qui ne peut être mise en œuvre en fait, qu'à travers la conscience. 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, etc.) et respectent les conventions sociales à travers les rôles qu'ils jouent (et dont ils ont aussi conscience). Dans le cadre du dialogue humain, on peut trouver une telle théorie très limitante : 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 à des états mentaux et donc à une composante cognitive dans un dialogue pseudo-naturel.

Dans le contexte de dialogue homme-machine, le contrôleur du dialogue implémenté en machine a 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. Cependant nous désignons — à la fois au sens de *design* et au sens d'interlocuteur — la machine comme un être individualisé et nous lui prêtons des états mentaux logiques par symétrie avec l'interlocuteur humain (usager).

3. Un cadre pragmatique

Le cadre pragmatique en dialogue homme-machine, peut se formaliser à travers (a) la pragmatique actionnelle et (b) en définissant un modèle de dialogue fondé sur les actes

de dialogue.

Formellement, et réduite à ce propos, la pragmatique se réduit à :

Pragmatique = {modes, interlocuteurs, monde, contexte_opératif}

avec :

modes = μ = {parole, geste, etc.}
interlocuteurs = {usager = U, machine = M}
monde = m = {mondes d'arrière-plan m_i , situation $\xi(t)$ }
contexte_opératif = {tâche T, activité A}

et :

U = {compétence = (langagière, perceptive, motrice)
performance = (expertise, habiletés)}

Le dialogue se situe au niveau de la coordination des actes des interlocuteurs. Ces actes modifient la situation, c'est-à-dire font passer à un instant donné la situation $\xi(t)$ à une nouvelle situation $\xi(t+1)$. L'usager doit effectuer une tâche T, c'est-à-dire doit atteindre un certain but B_0 caractérisé par la situation ξ_{B_0} . On suppose que ce but motive le dialogue.

3.1. Les modes

Dans le dialogue homme-machine nous appelons *langue* la composante langagière parlée, doublée d'une gestuelle déictique ou expressive. Elle peut être formalisée par l'intermédiaire d'une grammaire multimodale [Caelen, 91]. Elle ne sera pas détaillée ici. Toute intervention d'un locuteur est un *acte* (langagier ou non).

3.2. Les interlocuteurs

Les locuteurs mis en jeu ici sont typiquement l'usager (U) et la machine (M). Il est utile de classer les usagers en types pour déclencher les stratégies de dialogue et de sous-dialogues et particulièrement, les aides et les guides. Par souci de simplification on considère souvent uniquement les trois catégories d'usager suivantes :

{expert, occasionnel, naïf}

catégories qui résument des types plus complexes décrivant des compétences et performances données :

Usager = {compétence = (langagière, opérative, connaissances du domaine)
performance = (comportementale, perceptive, habiletés motrices)}

3.3. Les connaissances

En IA on appelle souvent, par abus de langage, *connaissances*, des données stockées en machine, pouvant évoluer et présentant un certain degré d'incertitude.

On distingue en logique floue :

a) les croyances, notées $c(P,t)$ où P est un prédicat, par exemple $Uc(P,t)$: U croit P dans la situation $\xi(t)$ avec une certitude $0 < p < 1$ (cette notation est abusive, en fait il s'agit de la plausibilité que l'on attribue à U d'avoir une certaine connaissance P) ce qui ne préjuge rien sur la vérité effective de P ni sur la certitude de P ,

b) les savoirs ou connaissances sûres, notées $s(P) = c(P)$ avec $p=1$.

- Pour la machine $Ms(P)$: M sait P , $Ms(P)$ signifie seulement que P est représentée dans M .

- Pour l'usager $Us(P,t) = Uc(P,t) \wedge P$. En DHM ce savoir résulte de la mise en commun de connaissances au cours du dialogue, U partage la connaissance P avec M quand $Us(P,t) \Leftrightarrow Ms(P) \wedge Uc(P,t)$, car en effet il suffit que P soit représentée dans M pour qu'elle soit vraie, $Ms(P) \Rightarrow P$.

Usager et machine ont des croyances et des savoirs à la fois sur le monde représenté à travers la situation, sur les mondes d'arrière-plan et sur eux-mêmes (fig. 2). L'ensemble de ces connaissances est le suivant :

- croyances de U sur le monde, $Uc\{m\}$, ces croyances s'expriment à travers le dialogue,
 - croyances de U sur les connaissances de M , $UcMs\{m\}$, ces croyances ne s'expriment généralement pas, sauf en cas d'incompréhension où elles deviennent explicites,
 - croyances de U sur M , UcM , manifestent les limites de M ,
 - croyances de U sur lui-même, UcU , sont des croyances de U sur ses compétences,
 - représentations du monde dans M , $Ms\{m\}$, sont toutes les connaissances et données utiles à la tâche,
 - croyances de M sur U , McU , sont des croyances de M sur les compétences de U ,
 - représentations de M d'elle-même, MsM , sont des connaissances qui permettent à M de s'adapter et d'évoluer.
- d'où il résulte des connaissances partagées par U et M : $Uc\{m\} \otimes Ms\{m\}$

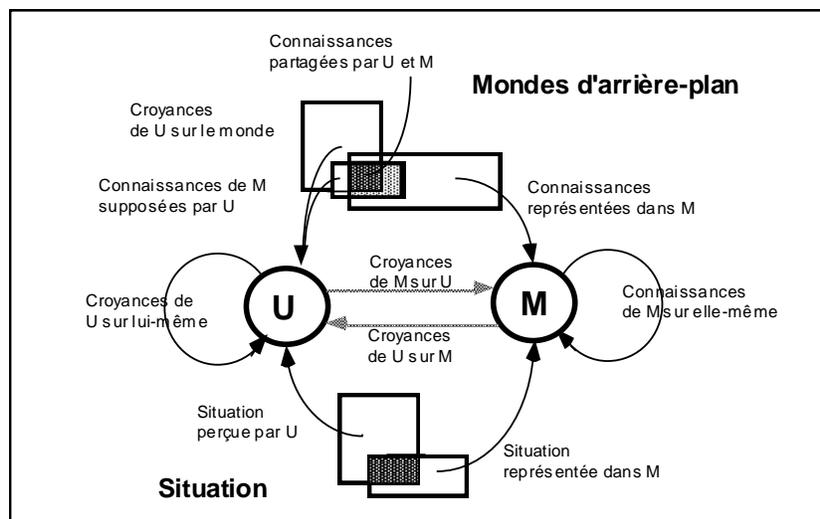


Fig. 2 : Processus inférentiel de l'usager U et de la machine M mis en jeu au cours du dialogue.

La machine doit avoir des capacités d'adaptation, c'est-à-dire doit être capable de modifier ses connaissances et faire évoluer ses croyances en fonction de la situation et de l'usager. C'est là le principe fondamental qui peut rendre le dialogue homme-machine

plus proche du dialogue humain¹ .

Axiomes

- principes de consistance des croyances de U :

$Us(P) \Rightarrow Uc(P)$	la croyance forte entraîne la croyance faible,
$Uc(P) \wedge Uc(P \Rightarrow Q) \Rightarrow Uc(Q)$	U croit aux conséquences de ses croyances,
$Uc(P) \Rightarrow U\text{-}c(\neg P)$	U ne croit pas le contraire de ce qu'il croit,
$Uc(P) \Rightarrow Uc(Uc(P))$	U croit ses propres croyances,
$\neg Uc(P) \Rightarrow Uc(U\text{-}c(P))$	U croit qu'il ne croit pas ce qu'il ne croit pas.

- principes de non monotonie des croyances de U et M :

$\models \exists \tau, Uc_{\tau}(P) \wedge Uc_{\tau}(\neg P)$	les croyances de U peuvent changer dans le temps.
$\models \exists \tau, Ms_{\tau}(P) \wedge Ms_{\tau}(\neg P)$	les connaissances de M peuvent changer dans le temps.

- le savoir de M fait loi :

$Ms(P) \Rightarrow P$	la réalité est celle qui est représentée,
$Uc(P) \wedge Ms(P) \Rightarrow Uc(P) \wedge P = Us(P)$	la croyance mutuelle est un savoir partagé.

- principe d'acquisition de connaissances par M :

$M\text{-}s(P) \wedge Ufs(P) \wedge (\neg \exists Q : P = \neg Q) \Rightarrow Ms(P)$ toute connaissance nouvelle non contradictoire avec une autre connaissance est acquise par la machine sans négociation.

Représentation du monde en machine

Le *monde* est celui qui est représenté en machine. Il comprend la situation courante définie en extension $\xi(t)$, et les mondes d'arrière-plan ou ensemble de connaissances nécessaires à l'interprétation d'un acte par la machine. Il peut y avoir plusieurs processus en machine qui s'exécutent simultanément si plusieurs fils de dialogue sont menés en parallèle. Il y a dans ce cas autant de "mondes" que de fils distincts. Notons que l'historique de la situation est $\{\xi(t_0), \xi(t_1), \dots, \xi(t_n)\} = H(t_n)$.

Monde = {Situation(je, tu, ici, maintenant), Mondes d'arrière-plan m_i }

```

Situation  $\xi(t)$ 
|   Interlocuteurs : (U,M)
|   Date : t
|   Fil : {m}
|   Etat du monde
|       Objets : {Instances O}
|       Relations entre objets
|           Spatiales : F(O1, O2,... On)
|           Temporelles : X(O1, O2,... Om)
└──
    
```

Ce schéma décrit la situation courante des objets instanciés par la machine. Cette situation peut être différente de la situation *perçue* par l'utilisateur car (a) tout n'est pas représentable sur l'écran et perceptible par l'utilisateur et (b) l'utilisateur a un niveau d'interprétation qui restera toujours opaque à la machine.

¹ La machine ne pourra pas offrir d'artefact pour simuler le dialogue humain car l'être humain ne peut pas se prendre à un jeu immotivé. La machine pourra tout au plus servir d'outil pour modéliser une certaine logique dialogique en reproduisant des dialogues humains ou en explicitant des raisons d'agir.

Pour représenter les objets, on utilise le formalisme des frames [Minsky, 75] qui sont des hiérarchies arborescentes d'objets possédant des attributs hérités des classes dominantes. Les éléments situés sur les feuilles de l'arbre sont des instances concrètes représentant les objets du monde considéré.

Chaque objet est représenté comme suit :

```
Objet : sorte_de Classe_Objct
|      attributs (de taille, de couleur, de forme, etc.)
|      propriétés (perceptives, physiques, etc.)
|      fonctions (vis-à-vis des autres objets du domaine)
|      comportements (actions pouvant opérer sur l'objet)
|_     liens_sémantiques (vers les mondes d'arrière-plan)
```

De même pour les instances (sauf pour le lien *est_un* à la place de *sorte_de*) et avec les conventions habituelles en matière d'héritage des attributs (les liens sémantiques n'étant pas hérités).

Par exemple :

```
Objet : sorte_de Classe_Objct
|      attributs {taille = fixe, orientation = fixe}
|      propriétés {état = solide}
|_     fonction = élément_plan
```

```
Objet_a_priori : sorte_de Objct
|      fonction = objet_atomique
|_     comportement = déplaçable
```

```
Objet_appris : sorte_de Objct
```

```
Pièce : sorte_de Objct_a_priori
|      attributs {forme = carré}
|      fonction {contenant, lieu_de_circulation}
|_     liens-sémantiques {composé_de = carré, synonyme_de = espace}
```

```
Cuisine 1 : est_un Pièce
|      attributs {position = zone_de_travail, visible}
|      fonction {repas}
|_     liens_sémantiques {synonyme_de = coin repas}
```

Classiquement, on accède à un objet d'une base de connaissances et aux informations qu'il contient par son nom ou à la suite d'inférences depuis une ou plusieurs propriétés distinctives — elles-mêmes déterminées à l'aide d'un calcul sur des valeurs d'attributs. Par exemple on peut définir des prédicats pour différencier des modes d'accès relativement à des relations spatiales :

DA(arg) : désignation par inférence d'attribut(s),

DN(arg) : désignation par le nom.

par exemple :

“ un lieu surveillé ”	DA(Pièce)
“ la cuisine ”	DN(cuisine)

Remarquons qu'ici nous ne nous intéressons pas aux figures linguistiques utilisées pour désigner le référent. Outre la connaissance des objets, il faut également expliciter les relations entre objets dans la situation. A cette fin, on formalise deux types de prédicats,

les prédicats spatiaux et les prédicats temporels. Le premier argument d'un prédicat est toujours l'objet désigné. Les attributs contiennent l'information nécessaire et suffisante pour définir complètement une relation (dans le contexte limité de nos expériences en désignation). Les prédicats spatiaux sont définis de la manière suivante :

F(arg1, arg2,... argn)
{relation ensembliste : intérieur, extérieur, etc.
distance en x et en y entre les centres des arguments
direction : haut, bas, gauche, droite, etc.
référentiel : ego, pattern}

où :

- F est un nom de prédicat à n arguments et quatre attributs,
- la relation ensembliste indique si les arguments sont distincts ou (partiellement) contenus l'un dans l'autre,
- le référentiel indique le système de coordonnées spatiales de la direction : la position du sujet (face à l'écran) ou la direction du dernier déplacement dans la figure.

De la même façon, pour les prédicats temporels nous avons :

X(arg1, arg2,... argm)
{ distance temporelle,
direction : présent, passé, futur
référentiel : début de tâche, moment de l'énonciation}

Par exemple, la relation,
" un carré à gauche du trait " sera représentée par le prédicat :

PR(carré, trait)
{ extérieur,
 $d(\text{carré}, \text{trait})(x, y) = (1/2(\text{Long}(\text{carré}) + \text{Long}(\text{trait})), 0)$,
gauche,
ego}

qui indique un positionnement spatial relatif (PR), dont le premier argument (carré) est désigné par rapport au second (trait), ces deux objets ayant été déjà instanciés par ailleurs. Les modalités particulières prises par ce prédicat dans le cadre de la tâche sont décrites par les attributs, à savoir ici : (1) le carré est extérieur au trait, (2) le carré est adjacent au trait, (3) à sa gauche, (4) dans un référentiel situé par rapport au sujet.

Il est évident que les habitudes des usagers les amènent à se comporter avec une machine en utilisant des mots (et donc des objets) référant à leurs mondes d'arrière-plan. Il n'est évidemment pas question de représenter tous ces mondes en machine. Le problème est pour la machine de repérer à quel monde de référence un objet peut appartenir et en cas d'échec, d'instaurer un dialogue d'éclaircissement avec l'utilisateur. Une solution, issue des définitions de "mondes possibles", est de lier les objets à travers des bases de connaissances permettant à l'aide de liens sémantiques. Par exemple, la fig. 3 montre les liens à propos des mots cercle, soleil, ballon, qui conduisent à la même représentation. De fait la commande " dessine un soleil " doit produire un "cercle" comme objet perçu dans un monde métaphorique.

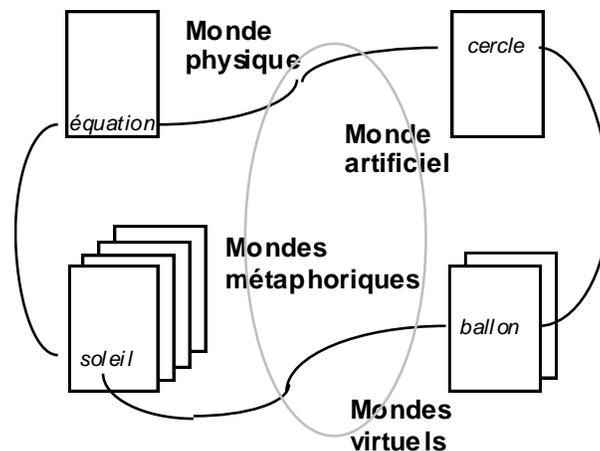


Fig. 3 : liens sémantiques entre objets appartenant à des mondes possibles différents.

Les mondes d'arrière-plan se divisent en autant de mondes qu'il est possible d'y référer différemment un même objet. On peut par exemple parler de "soleil" ou de "ballon" en regardant un "cercle" à l'écran ; ainsi, cet "objet" prendra son sens dans l'un de ces mondes. Le monde réel est le monde objectivable, le monde métaphorique est celui qu'est censée représenter la machine pour la tâche, le monde virtuel est celui de l'utilisateur et le monde artificiel est celui des objets abstraits de la machine.

Chacun de ces mondes est lié aux autres par un réseau sémantique muni des relations utiles à l'univers de l'application par exemple composé_de, partie_de, synonyme_de, forme_de, etc., pour une application de dessin ou de conception de plans architecturaux.

Exemple :

```
Pyramide_1 : est_un Objet
|_ liens_sémantiques : forme_de (Grand_Triangle)
```

Apprentissage

Le *monde* ainsi représenté peut-être "appris" au cours du dialogue puisqu'il devient possible d'acquérir de nouveaux attributs, propriétés, fonctions, comportements et significations. L'apprentissage porte à l'heure actuelle de grandes potentialités en DHM.

Cet apprentissage peut être mené par des sous-dialogues adaptés.

Exemple :

```
U : dessine une salle de bain.
M : qu'est-ce qu'une salle de bain ?
U : une pièce.
M : de quoi est-elle composée ?
etc..
```

Ces sous dialogues sont dirigés par les buts de la machine.

3.4. Contexte opératif

Le dialogue homme-machine se situe clairement dans un contexte opératif pour lequel on distingue généralement la tâche de l'activité : (a) la tâche est une représentation abstraite des opérations logiquement possibles pour atteindre un objectif, (b) l'activité est la séquence d'actes réellement effectués dans le contexte d'une tâche donnée. Classiquement il existe deux manières de représenter les tâches *a priori* :

(a) sous forme explicite, la tâche est hiérarchisée de manière arborescente (arbre ET/OU) en plan et/ou sous-plans ou sous tâches, scénarios, etc., jusqu'aux scripts quiinstancient les actions élémentaires. Le script étant la structure atomique de base constituée d'instructions élémentaires. Ce mode de représentation est statique et décrit la combinatoire des actions qu'il est potentiellement possible d'enchaîner pour exécuter une tâche avec succès. Cette représentation convient pour les tâches dites de routine.

(b) sous forme implicite où l'on ne représente que les buts à atteindre, éventuellement des contraintes, et les scripts de base. L'avantage est que le plan se construit au fur et à mesure de l'avancée du dialogue à partir des buts poursuivis par l'utilisateur. Ceux-ci sont inférés à partir des actes de dialogue. Le problème se ramène donc à une *gestion de buts*. Un autre avantage dû à cette économie de représentation est la possibilité d'apprendre de nouvelles tâches. Cela se paie évidemment par une plus grande lourdeur de traitement.

On appelle *but* un état abstrait du monde que désire atteindre l'usager. Cet état concerne aussi bien un état mental de ce dernier (par exemple connaître une information, acquérir un savoir-faire) qu'une situation définie. On suppose qu'on peut toujours représenter ce but à l'aide d'un prédicat B, par exemple :

- $B = (\exists \text{ Cercle}) \wedge \text{Attr}(\text{Cercle} = \text{rouge}) \wedge \text{Sur}(\text{Cercle}, \text{Carré})$ un cercle rouge sur le carré
- $B = (\exists x : x = \text{Cercle}) \wedge \text{Us}(x = \text{ballon})$ savoir qu'un des cercles représente un ballon

On note :

- un nouveau but : ζB , ce but vient d'être exprimé par l'usager,
- un but atteint : $\dagger B$, l'état de la situation est telle que le prédicat B est vrai,
- un but satisfait : $\ddagger B$, l'usager manifeste son accord explicitement ou implicitement,
- un but mis en attente : $-B$, l'usager ou la machine résout temporairement un autre problème,
- un but réparé : B' , à la suite d'une incompréhension le but est modifié,
- un but déplacé : B' , à la suite d'une négociation le but est modifié,
- un but réduit à un ou plusieurs sous-buts : sB , le problème est décomposé en sous-problèmes,
- un but abandonné : $@B$, à la suite d'un échec et d'un souhait d'abandon de l'usager.

Gestion des buts

La gestion des buts se fait de manière opportuniste, à l'aide des scripts appris au fur et à mesure des sessions d'utilisation du logiciel applicatif. Au départ, on suppose qu'on dispose d'actions élémentaires bien définies $\{A_i\}$. Le principe général est celui de la programmation par l'exemple : une fois un but posé ζB par l'usager, il s'agit pour la machine (A) d'observer puis (B) d'associer les actions montrées par l'usager pour satisfaire B.

A- Phase d'observation :

Il s'agit dans cette phase d'observation de reconnaître les actions qui sont en rapport avec le but B. On peut distinguer en effet, entre les actions, les relations de :

- Succession dans le temps de deux actions A_i et A_j (fortuite ou nécessaire),
 $A_i \leq A_j$ (se lit A_j suit fortuitement A_i) (par exemple : planter \leq pleuvoir)
 $A_i < A_j$ (se lit A_j suit nécessairement A_i) (par exemple : fleurir $<$ faner)
- Finalité (motivé par le but),
 $A_i \Rightarrow A_j$ (se lit A_j est la raison de A_i) (par ex. : planter \Rightarrow récolter)
- Conditionnalité (motivé par la situation),
 $A_j \perp A_i$ (se lit A_j accompli pour réaliser A_i) (par ex. : faire un trou \perp planter)
- Causalité (provoqué par les lois du monde ou par volonté),
 $A_i \bullet A_j$ (se lit A_j est causé par A_i) (par ex. : faner \bullet tomber)

Par exemple pour dessiner une maison on aura observé :

(dessiner(carré) \leq dessiner(triangle)) \Rightarrow dessiner(maison)

avec (activer(carré) $<$ poser(carré)) \perp dessiner(carré)

et (sélectionner(triangle) \perp dupliquer(triangle)) $<$ déplacer(triangle) \Rightarrow dessiner(triangle)

B- Phase d'association :

Il s'agit dans cette phase d'association d'attacher des buts aux séquences d'actions déduites des observations relativement à leurs cadres d'emploi.

début : $\exists B$ dans la situation $\xi(t)$
 suite d'actions observées

fin : $\exists B$ dans la situation $\xi(t+1)$

on en déduit $B = \{\xi(t+1) \cap \xi(t)\}$

Activité, actions

Au cours du dialogue l'utilisateur et la machine produisent des actions. On peut les classer-en :

$Uf(A,t), Mf(A,t)$: U ou M fait (ou montre) une action A à l'instant t,	(déclaratif)
$Uff(A,t), Mff(A,t)$: U fait faire à M une action A (ou réciproquement),	(directif)
$Ufs(X,t), Mfs(X,t)$: U fait savoir (énonce, répond) X à M à l'instant t (ou réciproquement),	(assertif)
$Uffs(X,t), Mffs(X,t)$: U demande X à M à l'instant t (ou réciproquement),	(directif)
$Mfp(.,t)$: M donne un choix à U.	(commissif+délég.)

Seules les actions de U sont intentionnelles. On note :

$Uv(B,t)$ = U veut atteindre le but B à l'instant t.

Axiomes

• attitudes de U devant les buts

$Uv(B1,t) \wedge Uv(B2,t) \Rightarrow Uc(B1 \leftrightarrow B2)$

U est cohérent : il ne désire qu'un but à la fois

$\dagger B \Rightarrow U \neg v(B,t)$

U ne maintient pas un but déjà atteint

$Uv(B,t) \wedge \neg B \Rightarrow Uv(B,t+1)$

U maintient son intention (persistance des buts de U)

$Uv(B1,t) \wedge Uv(B2,t+1) \wedge \neg \dagger B1 \Rightarrow (B2 = sB1)$ U est consistant dans le temps

• attitudes de U devant la machine

$Ufs(B,t) \Rightarrow Uv(B,t)$

U a des intentions sincères

$Ufs(X,t) \Rightarrow Ms(X,t+1)$

U est sincère et M modifie ses connaissances en conséquence

$U \neg s(X,t) \wedge Mfs(X,t) \Rightarrow Uc(X,t+1)$

U fait confiance à M

Uffs(X,t) \Rightarrow U \neg s(X,t)
 Uf(A,t) \vee Uff(A,t) \Rightarrow Uv(B,t)

les demandes de U sont motivées par son ignorance
 les actions de U sont motivées par le but

• obligations de la machine

Uff(A,t) \Rightarrow Mf(A,t+1) \vee Mfs(A,t+1)
 Uffs(X,t) \Rightarrow Mfs(X,t+1)
 Uv(B,t) \Rightarrow Mfs(B,t+1) \vee Mf(A•B,t+1)
 Mfp(.,t) \Rightarrow Uv(.,t)

M doit effectuer une action demandée par U (si elle est possible)
 M doit répondre à une demande de U
 M doit contribuer à la résolution du but ou expliquer les raisons de
 M donne un choix à U

• règles sur les buts

(Ufs(B,t) \vee Mfs(B,t)) \wedge \neg B \Rightarrow ζ B

un nouveau but est posé et devient candidat à la résolution
 le but est atteint

B est vrai dans $\xi(t) \Rightarrow \dagger B$

Ufs(\dagger B,t) \vee (\dagger B \wedge Ufs(Q,t) \wedge (Q \neq B)) $\Rightarrow \dagger B$

le but est satisfait après accord explicite ou implicite de U

(Uv(B,t) \wedge Mfs(B',t) \wedge B'= s B) $\Rightarrow \zeta$ B' \wedge \neg B

le but est déplacé par M pour des raisons de planification

(Uv(B,t) \wedge Mfs(B',t) \wedge B'= \neg sB) $\Rightarrow \zeta$ B' \wedge @B

le but est déplacé par M pour des raisons motivées par la situation (erreur par exemple)

M \neg s(B) \wedge \neg B \Rightarrow @B

le but est abandonné car M ne sait pas le résoudre

U \neg v(B) \vee \dagger B \Rightarrow @B

le but est abandonné si U change d'avis ou s'il est satisfait

• règles sur les effets des actes

On suppose que les effets des actes sur la situation sont observables.

Mf(A,t) \Rightarrow Uc(P ξ ,t+1)

une action de M a des effets observables pour U (l'inverse est vrai)

Apprentissage

A l'issue de l'apprentissage le modèle de tâches devient explicite. Ce modèle utilise des représentations classiques (graphes de scripts) et des processus d'apprentissage mettant en œuvre :

- la généralisation de buts,
 - la spécialisation de buts,
 - la factorisation de tâches,
 - etc.,
- processus bien maîtrisés en IA.

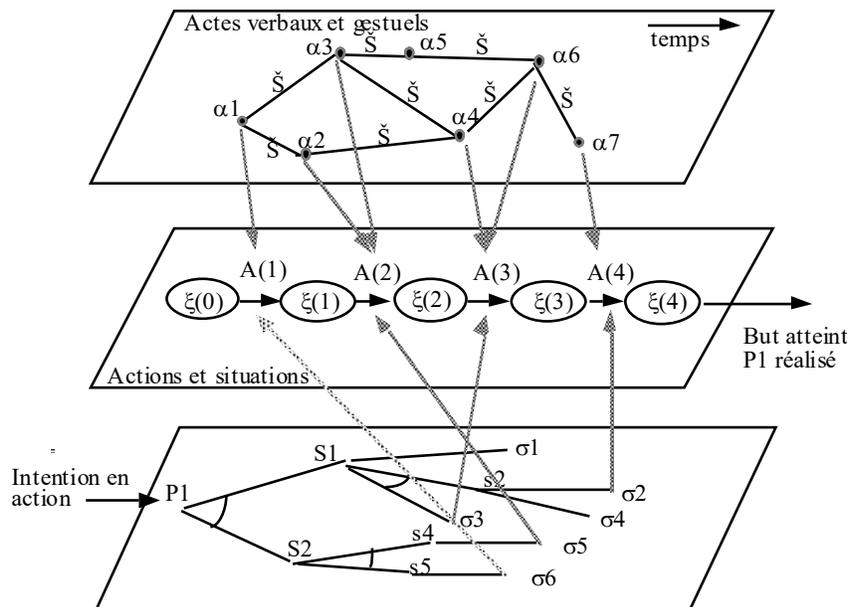


Fig. 4 : planification et organisation de scripts. Dans cet exemple le plan est réalisé par instanciation des scripts $\sigma_6, \sigma_5, \sigma_3, \sigma_2$.

Les structures formelles de représentation des scripts et actions sont :

Script : σ

- | appartient-à {tâche T}
- | liens de séquentialité {précède (<), suit (>), simultané (=)}
- | liens de finalité (\Rightarrow)
- | liens de conditionnalité (\perp)
- | liens de causalité (\bullet)
- | Corps : {Actions}

Action : A

- | appartient-à {scripts σ }
- | liens de séquentialité {précède (<), suit (>), simultané (=)}
- | liens de finalité (\Rightarrow)
- | liens de conditionnalité (\perp)
- | liens de causalité (\bullet)
- | Corps : {processus élémentaires}

4. Un modèle de dialogue

Le modèle de dialogue définit et précise les deux composantes :

- a) les actes de dialogue autorisés,
- b) les stratégies.

4.1. Les actes de dialogue

Le modèle des actes de dialogue précise ce que l'utilisateur a droit ou non de dire et de faire. Il précise le vocabulaire (lexique), la syntaxe et le type d'actes autorisés. Chaque interlocuteur peut énoncer des séquences de mots et/ou des séquences de gestes. Ces séquences correspondent à un ou plusieurs actes de langage. Un acte de dialogue doit donc être segmenté en actes de langage puis reconnu avant d'être interprété. Nous distinguons donc un *acte* d'une *action* dont nous avons parlé jusqu'ici. Un acte peut contenir plusieurs actions et réciproquement. Un acte est désigné par α .

Un acte est décrit par ses 3 composantes, locutoire, illocutoire et perlocutoire. En DHM la composante perlocutoire n'a pas d'intérêt (cf. §1). La composante locutoire est traitée par des modules de TALN (Traitement Automatique du Langage Naturel) (non détaillées ici) qui fournissent une structure sémantique de la forme :

α : nom de l'acte (verbe), actants

- | objet (quoi) : O
- | destinataire (à qui)
- | agent (qui)
- | manière (comment)
- | instrument (avec quoi)
- | lieu (où)
- | but (pour)
- | condition (si)
- | concession (malgré)

objet O : est_un Objet

- | attributs (de taille, de couleur, de forme, etc.)
- | propriétés (perceptives, physiques, etc.)
- | fonctions (vis-à-vis des autres objets du domaine)
- | comportements (actions pouvant opérer sur l'objet)
- | liens_sémantiques (vers les mondes d'arrière-plan)

| restriction (sauf)
| destination (vers), etc.

Le type d'acte (f, ff, fs, ffs, fp) est calculé par la machine en fonction des formes linguistiques autorisées [Colineau, 94] et de la situation.

4.2. Les stratégies de dialogue

Une stratégie de dialogue est la manière de gérer les échanges entre interlocuteurs pour aboutir à un but : c'est une rhétorique de l'inter-action. Cela consiste comme nous l'avons vu plus haut, à maintenir un but, le déplacer, faire converger les interlocuteurs, proposer un nouveau but, le mettre en attente, etc., de la manière la plus pertinente possible. Un critère de pertinence concernant la *justesse* d'une action dans une situation donnée, est de choisir cette action pour minimiser la distance globale au but (il peut y avoir des écartements locaux). On note Δ cette distance. Notons par ailleurs B1 le but de l'interlocuteur U1 (ici M) et B2 le but de l'interlocuteur U2 (ici U). Dans le dialogue humain il existe des stratégies très sophistiquées pour atteindre un but.

Par souci de clarté je présenterai uniquement deux grands types de stratégies :

- les stratégies non inférentielles,
- les stratégies inférentielles.

Stratégies non inférentielles

Ce type de stratégie ne met pas en jeu la compréhension du but de l'interlocuteur. M agit sans tenir compte du but B2 de l'utilisateur. On distingue deux stratégies de ce type.

— stratégie directive

l'initiative reste toujours du côté de la machine, l'utilisateur doit répondre strictement aux questions ou exécuter les ordres demandés. Il y a, en général, réduction progressive du focus :

- M impose son but B1 en ignorant B2
- M impose à U une réponse réactive ou négociée
- M garde l'initiative

Commentaire : Ce mode est utile en début de session de dialogue ou pour guider un utilisateur occasionnel ou en cas d'incompréhension.

— stratégie réactive

dans ce mode chaque interlocuteur (mais particulièrement la machine) réagit le plus complètement possible au dernier échange. S'il s'agit d'une commande, celle-ci est toujours interprétée et exécutée (prise de décision par défaut). Ce mode est particulièrement bien adapté à la manipulation directe, aux actions réversibles et immédiates (dessin par exemple).

- Il y a maintien du focus
- M "accepte" le but d'autrui a priori (mais ne le comprend pas) B1<-B2
- M laisse l'initiative à U

Commentaire : la machine répond par une action à l'initiative de l'utilisateur par échanges réactifs jusqu'à la clôture. Si les choix par défaut sont opportuns cette stratégie peut être très efficace (minimum d'échanges pour aboutir au succès de la commande).

Stratégies inférentielles

Ce type de stratégie passe par la recherche du but de l'interlocuteur et met en relation les connaissances supposées et partagées.

— stratégie coopérative

ici la machine se fait obligation de fournir un maximum d'informations pour aider et orienter son interlocuteur. Mais fournir trop d'informations augmente sa charge de travail et n'allège celle de son interlocuteur que si cette information est pertinente. La règle est donc de fournir l'information la plus pertinente eu égard à la situation (principe de la pertinence de Sperber & Wilson, maximes de coopération de Grice). Proposer une (ou des) solution qui convient au mieux à l'interlocuteur revient à : évaluer la situation, présenter une explication, des exemples, des aides ou des arguments pertinents et offrir un choix fermé (parce que facile cognitivement). Le mode coopératif procède par recherche d'un optimum dans un espace d'hypothèses. La coopérativité c'est accompagner l'interlocuteur jusqu'à la solution.

Il y a en général élargissement du focus

M fait sien le but de U, $B1=B2$

M ouvre toutes les stratégies pour U2 ainsi que les incidences

M laisse l'initiative à U

Commentaire : la véritable difficulté de ce mode est de donner toutes les informations nécessaires et suffisantes à la réalisation du but, en maximisant les concessions. Ce mode peut-être parfois un peu lourd mais il semble bien adapté à l'interrogation de bases de données car l'utilisateur ne sait pas toujours ce qu'il cherche ou comment l'obtenir.

— stratégie de négociation

par opposition au mode précédent, cette stratégie est plutôt celle de deux antagonistes : elle consiste à minimiser l'espace de concession accordé à son partenaire. La règle est ici l'argumentation et la réfutation. Dès qu'il n'y a plus de réfutation possible, l'action est exécutée s'il s'agit d'une commande ou le problème est résolu s'il s'agit de trouver une preuve.

Il y a en général maintien ou déplacement du focus

M impose son but ou accepte un compromis $B1 \neq B2$

M et U prennent l'initiative à tour de rôle

— stratégie dirigée par les intentions

c'est une sorte de stratégie coopérative fondée intensivement sur le but. Cela consiste à comprendre et interpréter les intentions de son interlocuteur. Il s'agit pour la machine d'un processus inférentiel pour déduire les objectifs de l'utilisateur à travers la succession de ses actes relativement au contexte de la situation. Les grands traits de cette stratégie sont :

M comprend (ou s'informe sur) le but de U pour continuer le dialogue $B1 < B2$

M ouvre toutes les stratégies à U sauf les incidences

M garde l'initiative

Commentaire : un tel mode ne peut être opérant en dialogue homme-machine que s'il est sous-tendu par une action dirigée par des objectifs. Comme dans le mode coopératif, il y a tentative de compréhension de la part de la machine dans un échange, mais les initiatives ne sont pas alternées.

— stratégie constructive

le principe de ce mode est d'apporter des informations nouvelles dans le focus du discours ou dans un focus voisin, dans le but de provoquer une rupture dialogique si possible enrichissante vis-à-vis des connaissances partagées. Elle peut procéder par présentation d'exemples.

- Il y a déplacement systématique du focus
- M déplace temporairement le but de U, B1 <- B2+
- M ouvre toutes les stratégies à U

Evidemment on a intérêt en DHM à varier autant que faire se peut la stratégie de dialogue au cours d'une session de travail pour en augmenter l'efficacité et pour en rompre la monotonie et donc maintenir un niveau d'attention ou de vigilance que réclament certaines tâches.

Axiomes

Le calcul d'une stratégie δ se fait à l'aide de règles qui tiennent compte de la complétude d'un acte de dialogue (s'il est incomplet il faut peut-être tenter de le compléter en posant une question adéquate avant d'agir), de la "vérité" d'un but (est-il possible de l'atteindre ?), des conditions de réussite de l'acte (la situation est-elle favorable ?), de l'état du dialogue (un autre but est-il en cours de traitement ou en attente ?), des attentes de l'utilisateur (quel but poursuit-il ?), de ses compétences (est-il expert ou non ?), de la stratégie précédemment adoptée (on a peut-être pas intérêt à changer trop souvent de stratégie pour ne pas égarer l'utilisateur).

Dialogue réactif

Règle : Un acte est traité en mode **réactif** si le nombre de tours de parole π depuis la précédente action (f) dépasse un certain seuil. Ce mode est également activé en cas de refus d'une autre stratégie ou pour conclure un dialogue. Les conditions de complétude, de vérité et de réussite ne sont pas prises en compte.

- $(\pi > \pi_0) \vee (Ufs(\delta = \text{réactif}, t)) \vee Ufs(\text{fin}, t) \Rightarrow (\delta = \text{réactif})$

Posons $P = P_\xi$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P, t)$

$Uf(\alpha, t) \Rightarrow Ms(P \cup \{Effets(\alpha)\}, t+1)$	U fait un acte, M enregistre les effets (RAJ)
$Uff(\alpha, t) \wedge Cond(\alpha) \Rightarrow Mf(\alpha, t+1) \wedge Ms(P \cup \{Effets(\alpha)\}, t+1)$ $(\exists attr(\alpha_i) = \emptyset) \Rightarrow \text{défaut}(attr(\alpha_i))$	U fait faire un acte, M exécute et RAJ si acte incomplet M complète par défaut
$(\forall attr(\alpha_i) \neq \emptyset) \wedge \alpha \bullet Plan) \wedge Mc(Upf(\alpha)) \Rightarrow Cond(\alpha)$	l'acte doit pouvoir déclencher un Plan
$Ufs(Q, t+1) \wedge Cond(Q) \Rightarrow Ms(P \cup Q, t+1)$ $(\exists attr(Q_i) \neq \emptyset) \Rightarrow Cond(Q)$	U donne une information, M l'enregistre
$Uffs(Q, t) \wedge (Q \subseteq P) \Rightarrow Mfs(Q, t+1)$	U pose une question, M y répond

Dialogue directif

Règle : Au début l'initiative est à la machine pour lui permettre de "se" présenter et de connaître son interlocuteur. Elle doit être pour cela en mode **directif**. Elle revient à ce mode dès qu'une incompréhension surgit (pour éviter le risque de bouclage ou d'impasse).

- $(\pi = 0) \vee (Ufs(\delta = \text{directif}, t)) \vee (Mfs(\text{erreur}, t)) \Rightarrow (\delta = \text{directif})$

Posons $P = P_\xi$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P, t)$

$Mff(\alpha, t) \Rightarrow Uf(\alpha, t+1) \wedge Ms(P \cup \{Effets(\alpha)\}, t+1)$	M fait faire un acte, U exécute et RAJ
$Mfs(Q, t) \Rightarrow Ms(Uc(Q, t+1))$	M donne une info. et suppose que U l'accepte
$Mffs(Q, t) \Rightarrow Ufs(Q, t+1) \vee Uffs(Q, t+1)$	M pose une question, U y répond
$Ufs(Q, t) \wedge Cond(Q) \Rightarrow Ms(P \cup Q, t+1)$ $(\exists attr(Q_i) \neq \emptyset) \Rightarrow Cond(Q)$	U donne une information, M l'enregistre
$Ufs(Q, t) \wedge Cond(Q = \text{contestation}) \Rightarrow \delta = \text{négocié}$	si U conteste il y a changement de stratégie

$$Ufs(Q,t) \wedge (Q \subseteq P) \Rightarrow Mfs(Q,t+1)$$

U pose une question incidente, M y répond

Dialogue intentionnel

Règle : Pour un expert et chaque fois que cela est possible, mais surtout en fin d'échange le dialogue s'établit en mode **intentionnel** dans le but pour M de se ramener à une situation prototypique facilement identifiable. Lorsqu'une situation prototypique (plan) est rencontrée, acceptée comme telle par U et éventuellement complétée par M, la phase de sous-dialogue est achevée, le plan peut être exécuté.

$$\bullet (Ufs(\delta = \text{intentionnel}, t) \vee (\dagger B \wedge Ms(U = \text{expert})) \Rightarrow (\delta = \text{intentionnel}))$$

Posons $P = P_{\xi}$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P,t)$

$$Mff(\alpha,t) \Rightarrow Uf(\alpha,t+1) \wedge Ms(P \cup \{Effets(\alpha)\}, t+1)$$

M fait faire un acte, U exécute et RAJ

$$Mfs(Q,t+1)$$

M donne une information

$$Mffs(B,t) \Rightarrow Ufs(B,t+1)$$

M pose une question sur le but, U y répond

$$Ufs(B,t) \wedge Cond(B) \Rightarrow Ms(B,t+1)$$

U donne une information sur le but, M l'enregistre pour le reconnaître

$$(\exists \text{attr}(B) \neq \emptyset) \Rightarrow Cond(B)$$

$$\dagger B \wedge B \bullet \text{Plan} \Rightarrow Mf(\text{Plan}, t+1)$$

M exécute le plan dès qu'il en a trouvé un

$$Ufs(B,t) \wedge (B \subseteq \{B\}) \Rightarrow Mfs(B,t+1)$$

U pose une question, M y répond

Dialogue négocié

Règle : Une action supporte la **négociation** si elle est incomplètement spécifiée et si le nombre de tours de parole depuis la précédente action n'est pas trop élevée. La négociation est menée jusqu'à son terme, elle ne peut être rompue par la machine.

$$\bullet ((\pi < \pi_0) \vee (Ufs(\delta = \text{négocié}, t) \wedge (Ufs(\alpha,t) \wedge (\exists \text{attr}(\alpha) = \emptyset))) \Rightarrow (\delta = \text{négocié}))$$

Posons $P = P_{\xi}$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P,t)$

$$Uf(\alpha,t) \Rightarrow Mfs(Effets(\alpha), t+1) \wedge (Ms(P \cup \{Effets(\alpha)\}, t+1) \cup \text{fait un acte, M enregistre les effets (RAJ)})$$

$$Uff(\alpha,t) \wedge Cond(\alpha) \Rightarrow Mf(\alpha,t+1) \wedge Ms(P \cup \{Effets(\alpha)\}, t+1)$$

U fait faire un acte, M exécute et RAJ

$$(\exists \text{attr}(\alpha) = \emptyset) \Rightarrow Mffs(\alpha,t+1)$$

si acte incomplet M négocie

$$Mffs(\alpha,t) \Rightarrow Ufs(\alpha,t+1) \vee (Ufs(\text{contestation}, t+1) \wedge Ufs(\alpha,t+1))$$

$$(\forall \text{attr}(\alpha) \neq \emptyset) \wedge \alpha \bullet \text{Plan} \wedge Ms(U = \text{expert}) \Rightarrow Cond(\alpha)$$

l'acte doit pouvoir déclencher un Plan

$$\neg Cond(\alpha) \Rightarrow Mffs(\alpha)$$

si l'acte n'est pas exécutable, M négocie

$$Ufs(Q,t) \wedge Cond(Q) \Rightarrow Ms(P \cup Q, t+1)$$

U donne une information, M l'enregistre si d'accord

$$(\exists \text{attr}(Q) \neq \emptyset) \Rightarrow Cond(Q)$$

$$\neg Cond(Q) \Rightarrow Mfs(\text{contestation}, t+1) \wedge Mffs(Q, t+1)$$

si l'information est erronée, M négocie

$$Mffs(Q,t) \Rightarrow Ufs(Q,t+1) \vee (Ufs(\text{contestation}, t+1) \wedge Ufs(Q,t+1))$$

$$Ufs(Q,t) \wedge (Q \subseteq P) \Rightarrow Mfs(Q,t+1)$$

U pose une question, M y répond

Dialogue constructif

Règle : La stratégie **constructive** sert surtout à alerter l'utilisateur ou à attirer son attention sur des sujets voisins de son propos. Elle peut être utilisée si le thème du dialogue n'a pas évolué depuis un certain temps et qu'aucune action n'a été faite.

$$\bullet ((\pi > \pi_0) \vee (Ufs(\delta = \text{constructif}, t) \wedge (Ufs(Q,t) \wedge (\text{focus}(Q) = \text{ct}))) \Rightarrow (\delta = \text{constructif}))$$

Posons $P = P_{\xi}$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P,t)$

$$Uf(\alpha,t) \vee Uff(\alpha,t) \Rightarrow Mfs(Effets(\alpha), t+1)$$

U fait ou fait faire un acte, M commente les effets

$$(Ufs(Q,t) \vee Mffs(Q,t)) \wedge Cond(Q) \Rightarrow Mfs(P \neq Q, t+1)$$

M donne une information en dehors du focus

$$(\exists \tau, \text{focus}(Q,t) = \text{focus}(Q,t-\tau)) \Rightarrow Cond(Q)$$

Dialogue coopératif

Règle : Une action est menée en **coopération** si elle est incomplètement spécifiée ET si le nombre de tours de parole depuis la précédente action n'est pas trop élevée ET si U n'est pas un expert. En situation de non-action, M propose une continuation au dialogue dans ce mode (relance).

$$\bullet ((\pi < \pi_0) \vee (Ufs(\delta = \text{coopératif}, t) \wedge (Ufs(\alpha,t) \wedge (\exists \text{attr}(\alpha) = \emptyset) \wedge (Ms(U = \text{expert}))) \Rightarrow (\delta = \text{coopératif}))$$

$$\bullet (U \neg f \wedge U \neg ff \wedge U \neg fs \wedge U \neg ffs) \Rightarrow (\delta = \text{coopératif})$$

Posons $P = P_{\xi}$ l'ensemble des prédicats qui décrit la situation $\xi(t)$ alors $Ms(P,t)$

$$Uf(\alpha,t) \Rightarrow Mfs(Effets(\alpha), t+1) \wedge (Ms(P \cup \{effets(\alpha)\}, t+1) \cup \text{fait un acte, M enregistre les effets (RAJ)})$$

$\begin{aligned} & \text{Uff}(\alpha, t) \wedge \text{Cond}(\alpha) \Rightarrow \text{Mf}(\alpha, t+1) \wedge \text{Mfs}(P \cup \{\text{effets}(\alpha)\}, t+1) \\ & (\exists \text{attr}(\alpha) = \emptyset) \Rightarrow \text{Mfs}(\alpha, t+1) \wedge \text{Mffs}(\alpha, t+1) \\ & \text{Mffs}(\alpha, t) \Rightarrow \text{Ufs}(\alpha, t+1) \vee \text{Ufs}(\text{rectification}, t+1) \\ & (\forall \text{attr}(\alpha) \neq \emptyset) \wedge \alpha \bullet \text{Plan} \Rightarrow \text{Cond}(\alpha) \\ & \neg \text{Cond}(\alpha) \Rightarrow \text{Mfs}(\alpha, t+1) \wedge \text{Mffs}(B', t+1) \end{aligned}$	<p>U fait faire un acte, M exécute et commente les effets si acte incomplet M l'explique et questionne</p> <p>l'acte doit pouvoir déclencher un Plan si l'acte est erroné, M propose un but voisin</p>
$\begin{aligned} & \text{Ufs}(Q, t+1) \wedge \text{Cond}(Q) \Rightarrow \text{Ms}(P \cup Q, t+1) \wedge \text{Mfs}(\text{effets}(Q), t+1) \\ & (\exists \text{attr}(Q) \neq \emptyset) \Rightarrow \text{Cond}(Q) \end{aligned}$	<p>U donne une information, M l'enregistre</p>
$\begin{aligned} & \text{Uffs}(Q, t) \wedge (Q \subseteq P) \Rightarrow \text{Mfs}(Q, t+1) \wedge \text{Mfs}(\text{effets}(Q), t+1) \\ & (U \neg f \wedge U \neg ff \wedge U \neg fs \wedge U \neg ffs) \Rightarrow \text{Mfp}(\cdot, t+1) \end{aligned}$	<p>U pose une question, M y répond en cas de non-action, M propose des choix à U</p>

5. Le contrôle du dialogue

Le DHM aide à la construction d'un but commun à partir des objectifs des interlocuteurs et à l'enrichissement des connaissances mutuelles. Il est souvent coopératif (les acteurs coordonnent leurs actions pour aboutir à un but commun) mais peut prendre des aspects de négociation par argumentation (contrôle d'un système expert, EIAO par exemple) ou rester entièrement réactif comme dans la plupart des logiciels à manipulation directe. Il est peut-être intéressant d'aboutir à des dialogues utilisant des stratégies variées.

Les rôles du dialogue dans un logiciel sont :

- la construction d'un univers commun (mondes),
- la réparation des erreurs de communication,
- la structuration et organisation de la communication,
- la gestion des buts,
- la conduite de l'action,
- la mise en contexte,
- l'apprentissage.

Le dialogue se déroule simultanément sur 2 plans :

- actionnel (f) et transactionnel (ff, fp)
- communicationnel (fs, ffs).

Il procède de deux mécanismes distincts contrôlés à long terme (mise en perspective des buts et planification de l'activité) et à court terme, ajustement local des tours de parole (planification du dialogue).

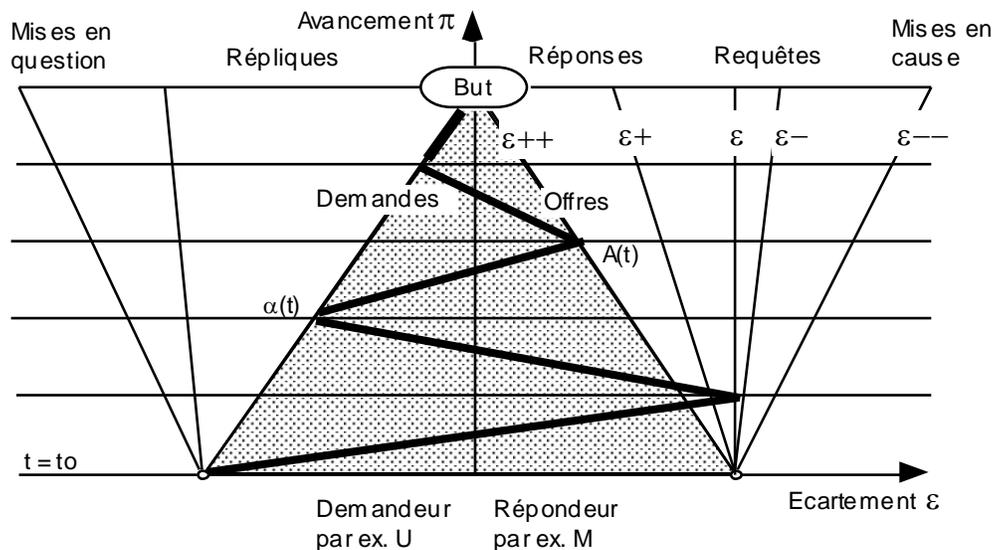


Fig. 5 : Convergence du dialogue au cours du temps. Ce schéma – qui représente un Echange – montre les axes convergents (Demandes, Offres) et les axes divergents (Mises en question, Mises en cause, Répliques et Requêtes) du dialogue.

Nous appelons Echange, une suite de tours de parole pendant lesquels un but est maintenu. On sort d'un échange lorsque le but est satisfait ou par une rupture (voir définitions ci-après). C'est l'utilisateur qui prend l'initiative de passer d'un échange au suivant. Un dialogue est donc une succession d'échanges interrompue éventuellement par des incidences (échanges secondaires) [Luzzati, 89]. Vernant [Vernant, 92] replace la conduite d'un échange dans une grille de dialogue dont 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 sur des variables d'écartement ε et de profondeur π .

Le contrôleur du dialogue gère les tours de parole c'est-à-dire la réponse à l'acte $\alpha(t)$: il interprète $\alpha(t)$ vis-à-vis du contexte, puis génère une réponse et, éventuellement, prédit l'acte suivant $\alpha(t+1)$. Il fonctionne selon un cycle qui peut être schématisé de la manière suivante :

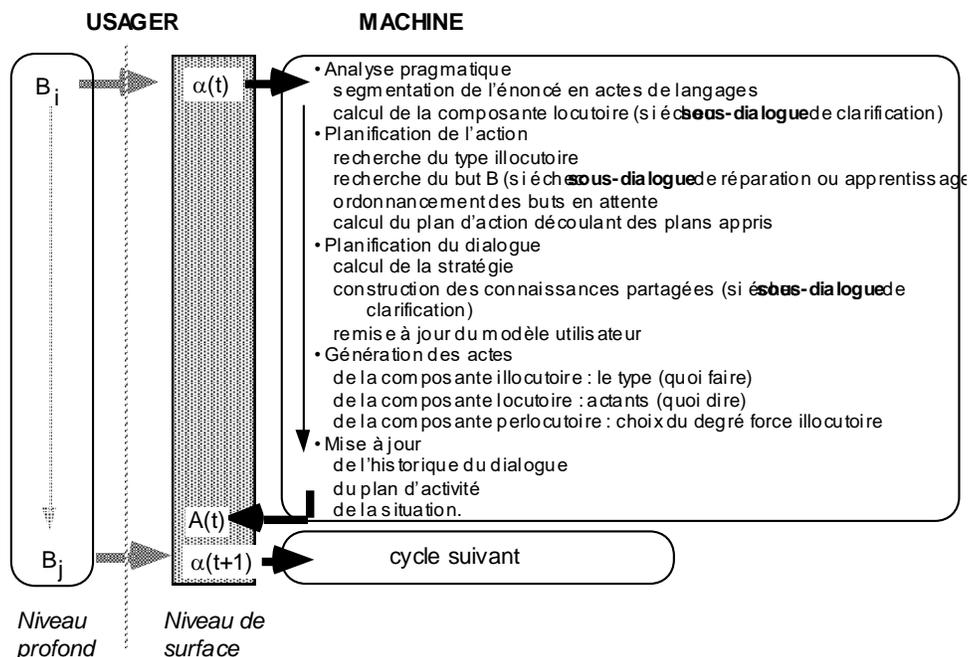


Fig. 6 : le cycle du contrôleur de dialogue. Dans ce schéma, le locuteur passe du souhait de réaliser le but B_i ($Uv(B_i)$) au souhait de réaliser le but B_j (en supposant que B_i a été satisfait entre-temps par l'action $A(t)$ de la machine) en produisant les deux actes $\alpha(t)$ et $\alpha(t+1)$ observables par la machine. Celle-ci doit dans l'intervalle effectuer les traitements indiqués dans l'encart de droite, dans le but de produire l'action (ou le plan d'action) $A(t)$ qui est la meilleure réponse attendue par l'utilisateur pour résoudre le problème posé à travers le but B_j .

Axiomes

Une incidence : se produit sur un acte qui change de thème. Cela provoque l'initialisation d'un sous-dialogue mais ne remet pas en cause le but de l'échange, qui est mis en attente pour être repris plus tard.

$Uv(B,t-1) \wedge \neg B \wedge Ufs(P,t) \wedge \neg focus(P,B)$	incidence délibérée provoquée par U
$Uff(\alpha,t) \cdot (Mfs(erreur(\alpha),t+1) \wedge Mffs(\alpha,t+1))$	clarification sur l'acte α provoquée par M
$Uv(B,t) \cdot ((Mfs(erreur(B),t) \vee Mffs(erreur(plan),t)) \wedge Mffs(B,t))$	réparation sur le but ou sur le plan provoquée par M
$Ufs(Q,t) \cdot Mffs(erreur(Q),t)$	incompréhension de M

Une rupture : se produit par changement inopiné de but

$$Uv(B,t-1) \wedge U\nu(B,t) \wedge \neg B \quad \text{est une rupture (elle viole le principe de persistance des buts)}$$

Au départ

$\models Uv(Bo,t_0)$ on suppose que la session de dialogue est motivée par un but Bo .

6. Exemple de dialogue homme-machine

La recherche des règles adéquates pour le dialogue doit se faire à partir de corpus acquis selon des protocoles expérimentaux bien maîtrisés [Fréchet et al., 92], [Ozkan, 94] puis doit être validée sur des plates-formes de test possédant des interfaces homme-machine les mettant en œuvre [Bourguet et al., 92]. Voici ce que cela peut donner sur un exemple

avec les règles de stratégie adaptées au domaine du dessin assisté par ordinateur et les règles décrites ci-dessus.

Dialogue	Acte	δ	But	Finalité
M1 : Bienvenue sur ICPtalk	fp			
Connaissez-vous le système ?	ffs	D	$B1=Ms(U)$	
U : oui, un peu	fs		$\dagger B1=Ms(U=occas)$	
M2 : je vous présente ses principales fonctions	fp	C	$\ddagger B1, B2=(\forall B \in Liste) \wedge fp(B)$	$\Rightarrow \zeta B$
1. l'édition graphique...	fp			
U : stop	ff		$\dagger B2$	
je connais	fs		$\ddagger B2$	
je veux ICPdraw	ff		$B3=B=Run(ICPdraw)$	
M3 : action (activation ICPdraw)	f	R	$\dagger B3$	
que voulez-vous faire ?	ffs	I	$\ddagger B3, B4=Ms(B)$	$\Rightarrow \zeta B$
j'attends votre intention générale	fs			
principe d'obligation violé				
U : ()	-fs		$B4 \text{ échec} \Rightarrow M \neg s(B)$	
M fait une incidence				
M4 : j'attends une commande	fp	D	$B5=fp(B') \wedge (B' \subset B)$	$\bullet B' \Rightarrow B$
U : dessine un cercle rouge	ff		$\dagger B5, B'=B6=(\exists Cercle1) \wedge Attr(rouge)$	
M5 : action (dessin cercle)	f	R	$\dagger B6$	
U : un autre, noir	ff		$\ddagger B6, B7=(\exists Cercle2) \wedge Attr(noir)$	
M6 : où ? à droite du premier ?	ffs	N	$B7'=B7 \wedge A_droite(Cercle2, Cercle1)$	
U : non, à gauche	fs		$B7''=B7 \wedge A_gauche(Cercle2, Cercle1)$	
M7 : identique au premier ?	ffs		$B7'''=B7'' \wedge (Taille(Cercle2)=Taille(Cercle1))$	
U : oui	fs		$\dagger B7'''$	
M8 : action	f		$\ddagger B7''' \Rightarrow \ddagger B7$	
que représentent-ils ? des roues ?	ffs	I	$B8=Ms(B)$	$\Rightarrow \zeta B$
U fait une incidence (sur ffs et focus)				
U : dessine un carré vert	ff		$B8 \text{ échec}, B9=(\exists Carré1) \wedge Attr(vert)$	
M9 : action	f	R	$\dagger B9$	
U : et deux autres sur les cercles	ff		$\ddagger B9, B10=(\exists Carré2) \wedge (\exists Carré3) \wedge \dots$	
M10 : sur chacun des cercles...	ffs	N	$B10'=B10 \wedge Sur(Carré2, Cercle1) \wedge \dots$	
U : non, superposés	fs		$B10''=B10' \wedge (?)$	
M fait une incidence				
@M11 : je ne comprends pas ce mot	fs-s	D	$B11=Ms(Réf(superposés))$	
attitude coopérative fs-s \Rightarrow ffs				
U : l'un sur l'autre	fs		$\dagger B11$	
M12 : bon	fs		$\ddagger B11$	
action (dessin)	f		$\ddagger B10''=B10$	
oui, mais en vue de quelle figure ?	ffs	I	$B8=Ms(B)$	$\Rightarrow B$
U : une sorte d'organigramme	fc		$\dagger B8, B=B12=Dessin(Orga)$	
M13 : ah ! je comprends	fs		$\ddagger B8, \dagger B12$	
j'aligne donc les figures	f			
j'équilibre les espaces	f			
U : oui, merci, c'est ça	fs		$\ddagger B12$	
Sauve ça maintenant	ff		$B13=Sauver(Orga)$	
M14 : voilà	f	R	$\dagger B13$	
Est-ce que je fais autre chose ? Par ex...	fp		$\ddagger B13, B14=Réf(fichier)$	
U : non, ça va	fs-f		$B14 \text{ abandon}$	
M15 : bon, au revoir	fs-f	R	$B14=fin, \ddagger B14$	

Conclusion

Le dialogue homme-machine (DHM) est fondamentalement différent du dialogue humain. La logique que nous avons décrite ne sert pas à simuler un dialogue humain avec une

machine mais à représenter et formaliser les mécanismes du dialogue homme-machine lui-même. Cette logique s'appuie sur les croyances et les actions et contient implicitement une logique déontique (obligations des partenaires). En faisant varier les stratégies dialogiques on aboutit à des dialogues relativement naturels. Nous avons volontairement considéré que le DHM est un mécanisme de gestion de buts à composante langagière : il doit mettre en correspondance un usager avec sa tâche, avec l'assistance de la machine. C'est donc à travers un cadre opératif et actionnel que nous fondons notre démarche.

Bibliographie

- [Bilange, 92] E. Bilange, Dialogue personne/machine, Hermès éd., Paris, 1992
- [Bourguet, 92] M.L. Bourguet, Conception et réalisation d'une interface de dialogue personne/machine multimodale. Thèse SIP, INPG, Grenoble, 1992
- [Bratman, 87] M. Bratman, Intentions, plans and practical reason. Cambridge MIT Press, 1987.
- [Caelen, 92] J. Caelen, J. Coutaz, Interaction homme-machine multimodale : quelques problèmes. Bulletin de la Communication Parlée, n° 2, pp. 125-140, 1992.
- [Caelen, 94] J. Caelen, Multimodal Human-Computer Interface, in Fundamentals of Speech Synthesis and Speech Recognition, E. Keller éd., J. Wiley & Sons, Ltd, Chichester, England, pp. 339-373, 1994.
- [Carberry, 90] S. Carberry, Plan Recognition in Natural Language Dialogue, MIT Press, Cambridge, 1990.
- [Clark, 82] H.H. Clark, T.B. Carlson, Hearers and speech acts, Language, 58, p. 332-373, 1982.
- [Cohen, 79] Ph. R. Cohen, C.R. Perrault, Elements of a Plan-Based Theory of Speech Acts. Cognitive Science 3, p. 177-212, 1979.
- [Cohen, 90] P.R. Cohen, J.L. Morgan, M.E. Pollack (eds.), Intentions in communication. Cambridge MIT Press, 1990.
- [Ducrot, 84] O. Ducrot, Le dire et le dit. Editions de Minuit, Paris, 1987.
- [Fréchet, 92] A.L. Fréchet, Analyse linguistique d'un corpus de dialogue oral homme-machine. Thèse Linguistique, Paris-Sorbonne, 1992.
- [Grice, 75] H.P. Grice, Logic and conversation. in Syntax and Semantic 3; Speech Acts, Coles and Morgan, New York Academic Press, pp. 41-58, 1975.
- [Grosz, 86] B.J. Grosz, C.L. Sidner, Attention, Intention in the Structure of Discourse, Computational Linguistics 12(3), p. 175-204, 1986.
- [Kerbrat-Orecchioni, 90] C. Kerbrat-Orecchioni, Les interactions verbales. Tome 1, Armand Colin éd., Paris, 1990.
- [Litman, 87] D.J. Litman, J.F. Allen, A Plan Recognition Model for Subdialogues, Cognitive Science, 11, p. 163-200, 1987.
- [Luzzati, 89] D. Luzzati, Recherches sur le dialogue Homme-Machine : modèles linguistiques et traitement automatique. Thèse d'Etat, Paris III.
- [Moeschler, 89] J. Moeschler, Modélisation du dialogue, représentation de l'inférence argumentative. Hermès éd., 1989.
- [Ozkan, 93] N. Ozkan, J. Caelen, Towards an adaptive dialogue model, Actes du symposium "Interfaces des mondes réels et virtuels", Montpellier, pp. 77-84, 1993.
- [Ozkan, 94] N. Ozkan, Analyses communicationnelles de dialogues finalisés. Thèse SC, INPG, Grenoble, 1994.
- [Pierrel, 87] J.M. Pierrel, Dialogue oral homme-machine. Hermès éd., Paris, 1987.
- [Sabah, 88] G. Sabah, L'I.A. et le langage. 2 tomes, Hermès éd., 1988 et 1989
- [Sabah, 90] G. Sabah, A model for interaction between cognitive processes. Proc. of COLING'90, vol. 3, Helsinki, p. 446-448, 1990.
- [Sadek, 91] 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.
- [Searle, 69] J.R. Searle, Speech acts. Cambridge University Press (traduit sous le titre Actes de langage, Hermann éd., Paris, 1972), 1969.
- [Searle, 85] J.R. Searle, D. Vanderveken, Foundations of illocutionary logic. Cambridge University Press, 1985.
- [Shegloff, 77] E.A. Shegloff, G. Jefferson, H. Sacks, The preference for self-correction in the organisation of repair in conversation. Language, 53, p. 361-382, 1977.
- [Sperber, 86] D. Sperber, D. Wilson, Relevance : communication and cognition. Basil Blackwell éd., Oxford, 1986.
- [Suchman, 87] L.S. Suchman, Plans and Situated Actions, Cambridge University Press, 1987
- [Vanderveken, 91] D. Vanderveken, Meaning and speech acts, Vol. I & II, Cambridge University Press, 1991
- [Vernant, 92] D. Vernant, Du dialogue, Vrin éd., Paris, 1992