

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

DOCTORAT EN INFORMATIQUE COGNITIVE

DIC-9411 - PROJET DE THÈSE

CONTRÔLE DES INTERACTIONS ORALES ENTRE HUMAIN ET MACHINE : APPROCHE  
D'APPRENTISSAGE MACHINE.

PAR

MINH-QUANG NGUYEN

DIRECTEURS DE THÈSE

JEAN GUY MEUNIER (UQAM-DÉP. PHILOSOPHIE)

DOUGLAS O'SHAUGHNESSY (INRS-EMT)

FÉVRIER 2007

## RÉSUMÉ

Les techniques de reconnaissance et de synthèse vocale deviennent de plus en plus performantes et robustes, ce qui facilite la création des applications de dialogue oral humain-machine. Ce sont des applications à partir desquelles la machine dialogue avec l'humain en imitant l'acte de communication des humains. Cependant, le dialogue entre l'humain et la machine manque encore de naturel, de souplesse et parfois il reste trop rigide. Alors que les recherches se sont orientées vers l'aspect acoustique et sémantique du signal de la parole, d'autres se sont positionnées au niveau de la stratégie du dialogue, considérée comme étant importante pour contrôler les interactions orales entre l'humain et la machine.

Partant de trois hypothèses de solution : (1) aspect stratégique du dialogue est le cœur des applications de dialogue humain-machine, (2) certaines théories du dialogue humain-humain peuvent être appliquées dans le dialogue humain-machine (i.e. théories des actes de langage, théories de la communication) et (3) apprentissage machine permet à la machine d'acquérir des nouvelles connaissances (i.e. trouver une stratégie optimale du dialogue).

Nous proposons, dans le cadre de notre projet de thèse, une approche d'apprentissage machine par renforcement avec un utilisateur simulé (Pietquin et Beaufort, 2005 ; J. Schatzmann et al., 2006). Notre approche emploie l'algorithme Q-Learning et le formalisme markovien (i.e. état, action, transition) pour trouver une stratégie optimale du dialogue, laquelle sera implantée dans une application de dialogue oral humain-machine pour le domaine de réservation de chambres d'hôtel. Les résultats de notre expérience devraient démontrer la faisabilité de cette approche pour améliorer la performance des applications et pour réduire le coût de modification évolutive de stratégie du dialogue, ce qui contribue à rendre le dialogue oral plus naturel et souple entre l'humain et la machine.

**Mots clés :** contrôle des interactions, stratégie du dialogue, gestion du dialogue, système de dialogue oral humain-machine, apprentissage machine par renforcement, utilisateur simulé, intelligence artificielle, processus de décision de Markov, Q-Learning.

## TABLE DES MATIÈRES

RÉSUMÉ .....	2
LISTE DES TABLEAUX.....	6
LISTE DES FIGURES .....	7
CHAPITRE I .....	8
INTRODUCTION .....	8
I.1 Problématique.....	8
I.2 Hypothèses de réponse .....	8
I.3 Objectif de recherche .....	10
CHAPITRE II .....	11
ÉTAT DE L'ART : DIALOGUE HUMAIN-MACHINE ET APPRENTISSAGE .....	11
II.1 Introduction.....	11
II.2 Dialogue humain humain.....	12
II.3 Dialogue humain machine.....	14
II.3.1 Architecture générale d'un SDOHM.....	14
II.3.1.1 Niveau acoustique.....	15
II.3.1.2 Niveau sémantique .....	17
II.3.1.3 Niveau pragmatique .....	20
II.3.2 Modélisation de la stratégie du dialogue .....	21
II.3.2.1 Degrés d'initiatives .....	21
II.3.2.2 Stratégie de confirmation .....	22
II.4 Apprentissage machine.....	22
II.4.1 Principes d'apprentissage machine.....	23
II.4.2 Apprentissage par renforcement sans utilisateur simulé.....	24
II.4.3 Apprentissage par renforcement avec un utilisateur simulé .....	25
II.4.4 Architecture d'AR avec un utilisateur simulé .....	27
II.4.5 Méthodes d'AR.....	27
II.5 La plateforme informatique.....	29
II.6 Conclusion.....	30
CHAPITRE III .....	31
NOTRE APPROCHE MÉTHODOLOGIQUE .....	31
III.1 Introduction.....	31
III.2 Exemple du système Hotel-Demo de Nuance.....	32

III.3	Cheminement méthodologique.....	34
III.4	Dialogue humain-humain .....	35
III.5	Dialogue humain-machine .....	35
III.5.1	Degré d'initiatives.....	35
III.5.2	Stratégie de confirmation .....	36
III.5.3	Dialogue comme processus de décision de Markov .....	36
III.5.3.1	Les actions.....	37
III.5.3.2	Les états .....	37
III.5.3.3	Les stratégies.....	39
III.6	Apprentissage.....	39
III.6.1	Paramètres pour l'apprentissage par renforcement .....	40
III.6.1.1	Propriété de Markov .....	40
III.6.1.2	Valeur de la fonction de récompense.....	40
III.6.1.3	Fonction objective .....	41
III.7	Plateforme informatique .....	42
III.8	Conclusion.....	42
	CHAPITRE IV .....	43
	ÉVALUATION.....	43
IV.1	Méthode d'évaluation .....	43
IV.2	Conclusion.....	45
	CHAPITRE V .....	46
	DISCUSSION .....	46
V.1	Introduction.....	46
V.2	Conclusion préliminaire .....	46
V.3	Plan d'avancement.....	48
	BIBLIOGRAPHIE .....	49

## LISTE DES TABLEAUX

Tableau	page
Tableau 3.1 Exemple d'un dialogue extrait du système Hotel-Demo développé par Nuance (2006)	33
Tableau 3.2 Les paramètres des états de dialogue et leurs valeurs.	39
Tableau 3.3 Valeurs définies pour la fonction de récompense.	41
Tableau 5.1 Exemple de dialogue contenant des anaphores et ellipses	45

## LISTE DES FIGURES

Figure		page
Figure 2.1	Vue générale des 5 modules d'un SDOHM (adaptée de Pietquin, 2004).	15
Figure 2.2	L'utilisateur simulé apprend des stratégies optimales du dialogue à la machine (d'après Schatzmann et al., 2006)	27

## CHAPITRE I

### INTRODUCTION

#### I.1 Problématique

Dans un dialogue oral entre deux personnes, l'aspect interactif est important, par exemple un client appelle le centre de réservation d'un hôtel. Le préposé, après l'avoir écouté, doit comprendre, réfléchir, interpréter la demande du client, ensuite il doit chercher dans différents dossiers pour trouver l'information correspondant aux critères demandés du client. Le préposé peut éventuellement demander au client de préciser quel type de chambre il cherche, quelle date lui convient et combien de personnes occuperont la chambre, etc. L'interaction entre ces interlocuteurs ne se fait pas par hasard, mais ces échanges se réalisent dans un ordre bien contrôlé selon des règles sociales de communication. De manière générale, un dialogue peut être décrit comme une suite d'actions se succédant ainsi : à l'ouverture du dialogue, le client ou le préposé commence le dialogue par une formule de politesse (i.e. *bonjour, je vous appelle pour...*), vient ensuite l'objet de l'échange où les deux interlocuteurs s'interrogent, négocient, partagent des idées (i.e. *avez-vous des chambres disponibles pour ce week-end ?, je voudrais une chambre moins chère*, etc.) Enfin le dialogue se termine par une formule de politesse marquant la fermeture du dialogue (i.e. *merci, au revoir*). Le dialogue peut être réussi sans que les deux interlocuteurs soient d'accord sur une idée commune ou sur un but précis.

En s'inspirant de l'exemple du dialogue humain-humain, on se demande si ces principes peuvent être appliqués dans un dialogue humain-machine. Quelles sont les théories en dialogue humain-humain qui pourraient être applicables dans le dialogue humain-machine ? Est-ce que la machine peut apprendre d'elle-même ces principes ? Quel support informatique permettra de construire une telle machine ?

#### I.2 Hypothèses de réponse



Comme dans un dialogue humain-humain, l'aspect interactif est important, nous choisissons, dans notre thèse, d'étudier l'interaction orale entre humain et machine. Deux hypothèses semblent se dégager :

Selon la première hypothèse, il n'existe aucune machine<sup>1</sup> qui rivalise avec le comportement cognitif de l'homme (i.e. perception, raisonnement, apprentissage, généralisation, etc.) Aucune machine n'a passé avec succès le test de Turing [TUR50]. Au lieu de construire une machine qui dialogue avec l'humain en inspirant les principes de dialogue entre humains, on peut imaginer une autre façon de dialoguer. Par exemple, en travaillant sur la puissance de la machine et sur sa capacité de stockage des données, nous pouvons enregistrer toutes les paroles et ensuite concevoir et implanter un algorithme qui fasse le lien entre les paroles (i.e. questions/réponses) pour reproduire un dialogue cohérent. Mais, le problème est de trouver cet algorithme.

La deuxième hypothèse consiste à penser au dialogue oral entre une personne et une machine. Si on veut que le dialogue soit naturel et souple, il faut que la machine ait la même capacité cognitive que l'humain ou au moins elle doit être en mesure de simuler le comportement social de l'humain et qu'elle contrôle les interactions de dialogue comme les humains, c'est-à-dire coordonner les tours de parole, interroger, deviner l'intention de l'utilisateur à travers l'énoncé pour donner les réponses cohérentes. En partant de cette hypothèse, nous proposons d'examiner quatre aspects du problème de contrôle des interactions orales entre l'humain et la machine. D'abord, nous examinons les théories en **dialogue humain-humain** pour comprendre les principes et les règles de ce contrôle. Comment un dialogue s'établit-il entre humains? Comment ces règles de contrôle peuvent-elles être appliquées dans un **dialogue humain-machine**? C'est-à-dire du côté de la machine, comment ces règles peuvent-elles être représentées et traitées par la machine? Ensuite, il faut voir comment la machine apprend ces règles. Cet **apprentissage** permet à la machine d'acquérir les nouvelles connaissances et optimiser sa performance. Concrètement, on implante une méthode d'apprentissage dans la machine (i.e. apprentissage à partir des exemples d'observation ou apprentissage supervisé/non supervisé, apprentissage par renforcement, etc.), après une période d'apprentissage, la machine acquièrera des nouvelles capacités de contrôle et réajustera son comportement face à une situation déjà apprise. Pour

---

<sup>1</sup> Dans ce qui suit, les termes la machine, l'agent machine ou le système désignent le même objet.

implémenter une telle méthode dans la machine, il faut prévoir une **plateforme informatique flexible et adaptative** afin de rendre la machine exploitable, évolutive, performante et assurer ainsi que le dialogue entre humain-machine sera souple et naturel.

### **I.3 Objectif de recherche**

L'objectif de cette thèse est d'apporter une contribution significative dans le contrôle des interactions orales entre l'humain et la machine et dans la méthode d'apprentissage de ce contrôle. Ce contrôle implique une gestion d'interaction c'est-à-dire une manière de gérer efficacement les interactions entre deux interlocuteurs (humain et machine). Nous appliquons notre recherche dans le domaine de la réservation des chambres d'hôtel. Cette contribution aura un effet sur la performance du système de dialogue. Elle permettra de réduire le temps et le coût de développement de nouvelles applications et aidera les non experts en dialogue à créer des applications de dialogue oral humain-machine. Nous découpons cet objectif en deux sous objectifs : informatique et cognition.

**Informatique.** Développer un outil qui incorpore une architecture d'apprentissage par renforcement pour optimiser automatiquement les contrôles des interactions.

**Cognition.** Appliquer des règles d'interactions orales humaines à la machine et simuler le comportement d'apprentissage humain à la machine.

## CHAPITRE II

### ÉTAT DE L'ART : DIALOGUE HUMAIN-MACHINE ET APPRENTISSAGE

#### II.1 Introduction

Des progrès importants en traitement du signal de la parole<sup>2</sup> et en traitement automatique des langues naturelles (TALN) depuis ce début de XXI<sup>ème</sup> siècle ont permis de commercialiser plusieurs applications en dialogue oral humain-machine. Parmi ces applications, on peut citer quelques unes très récentes comme par exemple l'application NJFun de (Singh et al., 2002) [SIN02] née d'une collaboration entre les universités de Pennsylvania, Pittsburgh et du laboratoire de recherche d'AT&T, l'application EMILY<sup>3</sup> (2003) développée par *Nuance Communications Inc.* pour *Bell Canada* ou le browser CLARISSA (2005) [RAY05] construit par la même compagnie pour *International Space Station* (ISS), ou des outils pour développer les applications de dialogue comme l'outil 'Dialog Studio' de CSLU<sup>4</sup> utilisé dans le projet TALK<sup>5</sup> (2006), un projet Européen regroupant plusieurs universités européennes (Saarland University, The Univesiry of Cambridge, etc.) et les partenariats industriels comme BMW, BOSCH. Ces applications ont montré qu'il est possible aujourd'hui d'utiliser le signal vocal comme un mode de communication efficace entre l'humain et la machine. Cependant, Il reste encore beaucoup de problèmes pour maîtriser la communication entre humain et machine. Un des problèmes importants dans les applications de

---

<sup>2</sup> Au point de vue physique, la parole est assimilée à une suite de données acoustiques. Le traitement du signal de la parole dans ce contexte se réfère au traitement du signal acoustique sans tenir compte des théories linguistiques (sémantique, pragmatique).

<sup>3</sup> EMILY. Disponible sur: [http://www.cipa.com/press\\_room/success\\_stories\\_04/BellCanada-Emily.pdf](http://www.cipa.com/press_room/success_stories_04/BellCanada-Emily.pdf) [consulté le 2 nov 2005].

<sup>4</sup> Center for Spoken Language Understanding. Disponible sur: <http://www.cslu.ogi.edu> [consulté le 1 oct 2006].

<sup>5</sup> Projet TALK. Disponible sur : (<http://www.talk-project.org/>) [consulté le 1 oct 2006].

dialogue est le contrôle des interactions orales entre humain et machine. On cherche à contrôler les échanges entre deux interlocuteurs. Ces échanges ne se font pas par hasard, mais ils suivent des règles précises de communication. De ce fait, comprendre les principes des interactions orales et les règles de contrôle de ces interactions, nous permet de construire des applications de dialogue oral humain-machine performantes.

Nous présenterons les travaux réalisés sur le dialogue humain-humain, le dialogue humain-machine, l'apprentissage machine et la plateforme informatique. Ces travaux permettent de bien cerner les aspects essentiels en gestion de dialogue du côté machine.

## II.2 Dialogue humain humain

Le terme **dialogue humain-humain** désigne un acte de communication et d'interaction entre deux agents humains. Chaque acte produit un énoncé. La séquence des énoncés émis successivement par deux interlocuteurs et régis par une certaine règle constitue un dialogue. Un dialogue n'est pas une séquence aléatoire d'énoncés, mais un échange qui se réalise, où chaque agent, à tour de rôle, prend alternativement le rôle du locuteur et de l'auditeur. L'échange pourrait être concrétisé sous forme de dialogue oral, écrit, gestuel ou multimodal (à la fois parole, écrit et gestuel). L'étude de dialogue oral ou dialogue<sup>6</sup> relève des théories issues de la linguistique, de la philosophie, de la communication et de la psycholinguistique. Parmi ces études théoriques, il faut citer l'analyse des actes de langage qui étudie l'acte de parole (Searles, 1976) et l'analyse de la conversation qui considère le phénomène du dialogue comme une interaction sociale [LEV83]. Ces théories contribuent, plus ou moins, dans le développement du système de dialogue humain-machine, par exemple, au point de vue psycholinguistique, l'analyse du dialogue pose le problème du contexte et de l'interprétation pragmatique [LEV83], c'est-à-dire, comment la connaissance du monde peut être intégrée, comprise et traduite par les deux interlocuteurs [CHO69], [SAU74], [SEA76]. Alors qu'au point de vue de l'approche communicative, le dialogue peut être vu comme une théorie de jeux où chaque interlocuteur essaie de deviner le but que l'autre poursuit par son action (i.e. donner une réponse ou poser une question) (Wittgenstein, 1957).

---

<sup>6</sup> Le terme dialogue est sous-entendu dialogue oral.

**Actes de langage** : depuis les travaux en philosophie d'Austin (1962) [AUS62] et de Searle (1969) [SEA69] et de Grice (1989) [GRI89], on distingue dans l'acte de langage trois composantes : l'acte **locutoire** (1) qui est défini comme l'action d'articulation et de combinaison des sons, un simple mécanisme biologique que nous possédons tous, dès l'enfance et que nous faisons naturellement bien. L'acte locutoire concerne le mécanisme de production de la parole à l'aide d'une structure grammaticale acquise (syntactique, lexicale et sémantique). L'acte **illocutoire** (2) concerne l'intention de l'interlocuteur qui produit un énoncé avec une intention. Il est caractérisé par une force illocutoire (phrase affirmative, interrogative, déclarative, etc.) Selon Meunier [MEU06], « *il est certain qu'un ordinateur ne peut identifier le véritable contenu signifiant de la plus petite des conversations entre lui et un utilisateur sans mettre en jeu ces dimensions illocutoires* ». Il nous semble que la théorie philosophique des actes de langage est omniprésente dans les traitements de langue et aussi dans le dialogue humain-machine puisque c'est par les traitements syntaxiques et grammaticaux que la machine essaie de décortiquer l'intention de l'autre à travers l'acte illocutoire. Enfin, l'acte **perlocutoire** (3) concerne les effets que les actes illocutoires peuvent avoir sur l'interlocuteur (faire changer sa croyance, ses actions ou son jugement), par exemple, convaincre une personne de sa croyance ou effrayer les gens par des propos choquants sont des actes perlocutoires.

**Actes de dialogue** : l'acte de dialogue désigne une action de communication entre les interlocuteurs. Bunt (1996) distingue trois aspects de l'acte de dialogue : la forme de l'énoncé, sa fonction communicative et son contenu sémantique. Prenons, par exemple, la forme de l'énoncé « *qui nous rendra visite ce soir ?* » qui est une représentation de surface de la proposition « *quelqu'un nous rendra visite ce soir* ». Sa fonction communicative est une question. Le contenu sémantique s'exprime dans l'intention du locuteur qui veut savoir si la proposition « *quelqu'un nous rendra visite ce soir* » est vraie ou fausse. La réponse à cette proposition est justifiée par la croyance de l'auditeur en la question « *qui nous rendra visite ce soir ?* » Autrement dit, le dialogue est régi par des propriétés spécifiques.

**Théorie de jeux et de décision** : la communication peut être vue comme une partie de jeu où chaque participant prend son tour de parole. L'approche de modélisation de la communication par des modèles de jeux consiste à concevoir le dialogue comme un jeu de langage (Wittgenstein,

1957). Selon Caelen (2005)<sup>7</sup>, “*les stratégies de dialogue sont des manières d’atteindre un but à travers le dialogue considéré comme une activité conjointe (Vernant, 1992).*” Cette stratégie se manifeste sous forme d’un jeu de conversation dans lequel les objectifs des interlocuteurs peuvent être différents, mais ils ont quand même une motivation commune, celle de l’intention de communiquer pour partager leur but. La décision de partager ou non leur but relève du comportement individuel des interlocuteurs. Mais la prise de décision est souvent obtenue en fonction des événements (action, situation) historiques, de la croyance et des buts communs ou différents des interlocuteurs.

### II.3 Dialogue humain machine

Pour présenter le dialogue humain-machine, il est essentiel de bien définir les termes employés dans ce domaine. Le terme agent humain ou simplement humain désigne une personne physique en pleine possession de ses capacités cognitives (perception, adaptation, interaction, évolution). Un agent machine ou machine est une entité à qui l’on attribue une capacité de simulation de certaines de nos capacités cognitives (i.e. la simulation, la perception du signal de la parole, le raisonnement, etc.) L’agent machine représente une entité autonome qui interagit avec l’agent humain. Nous employons aussi le terme *système* pour désigner l’agent machine. De cette définition, un **système de dialogue oral humain-machine** (SDOHM) est un agent machine qui réalise un ensemble des tâches, telles que la reconnaissance et l’interprétation de l’énoncé oral, la gestion du dialogue, la génération et la synthèse du signal de la parole. L’agent machine joue un rôle déterminant dans le succès du dialogue, puisque par la fonction qu’on lui assigne, l’agent machine doit accomplir certaines tâches particulières pour satisfaire l’agent humain sinon il n’y a pas d’intérêt pour un tel dialogue. Le terme dialogue oral humain-machine sous-entend un dialogue finalisé, c’est-à-dire l’agent machine doit, par exemple, chercher à satisfaire le but de l’agent humain en devinant son intention (but) à travers l’expression orale. Il doit guider et contrôler le dialogue pour s’assurer du bon déroulement de l’échange.

#### II.3.1 Architecture générale d’un SDOHM

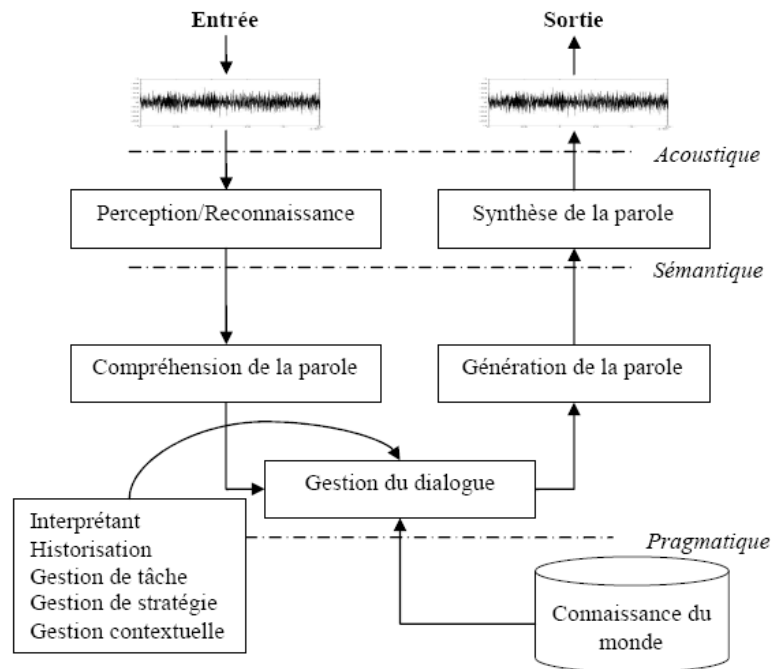
Après avoir donné les définitions des différents termes employés dans un système de dialogue,

---

<sup>7</sup> Rapport d’activités 2002-2005 de l’équipe GEOD (Groupe d’Étude sur l’Oral et le Dialogue). Disponible sur : [http://www-clips.imag.fr/geod/doc/Rapport\\_Activites\\_2002-2005.pdf](http://www-clips.imag.fr/geod/doc/Rapport_Activites_2002-2005.pdf) [consulté le 22 fev 2006]

nous sommes en mesure d'expliquer le cœur du système. Ainsi, un SDOHM peut être découpé en cinq modules : un module de reconnaissance du signal de la parole, un module de compréhension de la parole, un module de gestion du dialogue, un module de génération de la parole, et enfin un module de production du signal de la parole.

Pietquin [PIE04] a distingué trois niveaux de traitement dans un système de dialogue : un niveau acoustique, un niveau sémantique et un niveau pragmatique. Nous détaillerons chaque niveau de traitement ci-dessous.



**Figure 2.1** Vue générale des 5 modules d'un SDOHM (adaptée de Pietquin, 2004).

### II.3.1.1 Niveau acoustique

#### *Reconnaissance du signal de la parole*

Le niveau acoustique est le niveau le plus bas dans la chaîne de communication parlée. Le processus de reconnaissance de la parole traite le signal de parole comme une suite de séquences de sons, laquelle subit une série de processus particuliers, avant d'être transcrite sous forme textuelle, pour enfin être perçue comme un message linguistique et éventuellement compréhensible. Pour O'Shaughnessy [OSH00], ces processus particuliers font partie de la perception de la parole. La perception de la parole comporte plusieurs étapes d'analyse : auditive,

phonétique, phonologique, lexicale (mots), syntaxique et sémantique. Pour McQueen et Culter (1997), les processus de perception et de reconnaissance ne peuvent être traités comme des processus indépendants et séquentiels, car ils se réalisent en parallèle et s'activent mutuellement et de façon dynamique selon la qualité de l'information en entrant. Cependant, quelques étapes doivent se réaliser en parallèle ou en simultané par rétroaction (feedback), par exemple, les étapes de traitement auditif, phonétique et lexical se réalisent simultanément. Lorsque l'information de bas niveau (phonétique acoustique) est manquante, le processus de reconnaissance de mots (lexical) utilise la connaissance de haut niveau (phrastique ou contextuel) pour corriger ce manque d'information (réalisation en parallèle) et ainsi permettre au processus de bas niveau de retarder la décision phonétique jusqu'à la réception complète de l'information. L'exemple typique de ce processus simultané se manifeste lorsque le signal acoustique représentant un mot contient du bruit ou est en cours de se produire, l'agent humain est capable de percevoir ce mot avant que le son du mot soit entièrement sorti de l'appareil phonatoire. Les étapes de la perception évoquées ci-dessus nous donnent des indications quant à la nature des informations extraites du signal sonore, afin de le percevoir comme de la parole. La question qui se pose est celle de savoir quand et comment ces informations peuvent être extraites et intégrées dans la connaissance [MCQ97]. Pour McQueen et Culter (1997), l'un des aspects importants, sinon le plus important dans la perception de la parole est la reconnaissance des mots. En effet, McQueen et Cutler (1997) renvoient le problème de la perception de la parole au problème du lexique mental (dictionnaire mental), qui est le point essentiel dans tous les modèles de compréhension de la parole et qui se situe au niveau sémantique. Donc, le module de reconnaissance se situe à la limite entre le niveau acoustique et sémantique. Au point de vue de réalisation informatique, il est du ressort du concepteur de décider quelle façon est la mieux adaptée pour intégrer les modèles acoustiques et linguistiques dans un système de dialogue. Mais, quoiqu'il en soit, la perception et la reconnaissance emploient, presque les mêmes processus de traitement du signal et ces processus ne sont pas tous situés au niveau acoustique comme le soulignent O'Shaughnessy, (2000) [OSH00], McQueen et Culter, (1997) [MCQ97].

### *Synthèse de la parole*

La synthèse de la parole consiste à transformer les séquences de caractères en sons digitalisés. La synthèse de la parole pourrait très bien se placer au niveau sémantique, car certains traitements comme les recherches d'homophones, allophones font appel aux modèles



linguistiques. Mais, puisque la synthèse de la parole est le dernier module situé dans un système de dialogue qui produit directement une suite sonore. Il nous paraît légitime de le classer au niveau acoustique.

Il y a plusieurs techniques pour reproduire des sons de la parole. Nous ne développerons pas en détail cette partie, mais nous citerons seulement ces techniques. Pour de plus amples informations sur la synthèse de la parole, les lecteurs peuvent consulter la référence de [OSH00]. Ces techniques sont : la synthèse par prédiction linéaire, la synthèse par concaténation (phonème, diphone), la synthèse par règle, etc. Bien qu'il y ait toujours des recherches et des découvertes à faire dans ce domaine, la synthèse de la parole est la partie la mieux maîtrisée dans l'ensemble du traitement du signal acoustique de parole (reconnaissance, compréhension et génération de la parole). Ceci est dû en partie à l'implication moins importante de la sémantique dans le processus de conversion du texte au son de la parole.

### II.3.1.2 Niveau sémantique

#### *Compréhension de la parole*

À ce niveau, le traitement de la compréhension des langues naturelles cherche à extraire le sens des séquences de caractères issues de la reconnaissance de la parole. En dialogue oral spontané entre les interlocuteurs, la sémantique dépend du contexte dans lequel le dialogue est engagé, par exemple, la phrase « *le chat prend la souris et la traîne vers la fenêtre* » décrit un comportement instinctif de l'animal (on peut sous-entendre le verbe *prendre* comme *manger*), par contre la phrase « *je prends la souris et la déplace vers la fenêtre X* » décrit une situation où le mot « *souris* » a pris un autre sens. Par abus du langage et à force d'employer cette expression dans le contexte informatique, on finit par lui donner un autre sens. Ainsi, l'extraction de la sémantique d'une suite de caractères se fait en analysant la syntaxe, la sémantique et l'interprétation contextuelle, parce qu'au point de vue d'implémentation informatique, les modules du système de dialogue fonctionnent de façon séquentielle (le module reconnaissance se trouve en amont du module de compréhension), ce qui veut dire que l'extraction sémantique dépend beaucoup de la performance du module de reconnaissance. A l'issue du module de compréhension, une représentation non linguistique de l'énoncé est créée sous forme de schéma structurel. Cette représentation peut être vue comme un transfert de l'information de l'agent humain vers l'agent machine (Pietquin, 2004) [PIE04]. Elle est aussi appelée « concept ». Par exemple, en supposant

que le nombre de chambres et la date de réservation sont deux critères suffisants pour réserver une chambre d'hôtel, alors la phrase « *je voudrais réserver une chambre pour ce vendredi* » pourrait être conceptualisée sous une forme générale telle que :

<Concept>

Réservation (Nbc : 1 ; Dte : vendredi)

<Fin>

Où Nbc est le nombre de chambres et Dte contient la date de réservation. Le but d'une telle description est de représenter les phrases par des concepts, des objets et des règles à travers lesquels le système interagit avec le monde extérieur. Le niveau sémantique se réfère à la théorie de l'usage de la langue. Les interlocuteurs essaient de reconnaître (et faire reconnaître) leur but et leur croyance dans l'objectif de parvenir à satisfaire leur demande. Ce qui suppose qu'il doit y avoir une coopération, un partage entre deux interlocuteurs. Le système COLLAGEN (Rich et al., 1998, 2001) [RIC98], [RIC01], par exemple, est construit à partir du modèle collaboratif d'Allwood (Allwood, 1979) [ALL79].

### *Gestion du dialogue*

Contrôler les interactions orales entre humain et machine signifie gérer les échanges entre ces deux agents. C'est un des rôles du module de gestion du dialogue qui se situe entre le module de compréhension et de génération de la parole. Pour bien gérer ces échanges, il faut avoir une **stratégie** de gestion des échanges. C'est le rôle du gestionnaire de dialogue dans le module de gestion du dialogue. Avoir un bon gestionnaire de dialogue garantira l'efficacité et la performance du SDOHM. Minker (2000), inspiré par d'autres chercheurs (Béroule et Néel, 1984; Pierrel, 1981,1987), a distingué deux catégories de connaissances requises pour bien gérer l'acte de dialogue : des connaissances statiques et dynamiques. Les connaissances statiques sont celles du langage, du dialogue, de la tâche et de l'utilisateur. Les connaissances dynamiques concernent l'historisation du dialogue, le contexte de la tâche et le contexte système lui-même.

On a considéré, depuis longtemps, la recherche d'une stratégie optimale du dialogue comme un art plutôt qu'une science (Jones et Galliers, 1996) [SCH06]. Bien que le terme stratégie ait été employé par plusieurs auteurs, il n'existe pas de définition claire d'une stratégie efficace de dialogue (Schatzmann et al., 2006; Pietquin, 2005) [SCH06], [PIE05]. Une **stratégie de dialogue**

est dite efficace ou **optimale**, si les actions exécutées satisfont le but recherché par l'agent humain. Il est donc nécessaire que les deux agents partagent un but commun et que la stratégie optimale consiste à trouver les actions appropriées qui orientent le dialogue vers ce but. Plus précisément, trouver ces actions revient à déterminer les règles de transitions permettant de passer d'un état à un autre. En situation dialogique, une stratégie optimale doit être définie en termes de degré d'initiatives (système-initiative, utilisateur-initiative ou mixte-initiative)<sup>8</sup> et de stratégie de confirmation (explicite ou implicite par rapport au taux d'erreurs de reconnaissance et/ou au domaine et au type d'applications). Nous employons le terme *stratégie de dialogue* pour désigner une action appropriée à une instance<sup>9</sup> de dialogue. Cette action, lorsque exécutée, permet de passer d'une situation à une autre.

Plusieurs approches de modélisation de la stratégie de dialogue ont été proposées dans la littérature, nous résumons les principales approches récentes :

**Approche basée sur les états sous forme de réseau** : la structure du dialogue est représentée sous forme de réseau où les nœuds représentent des questions et où les transitions entre les nœuds représentent les actes de dialogue (Mc Tear, 1998, Pietquin, 2004) [PIE04]. Cette approche a été appliquée avec le réseau Bayésien.

**Approche basée sur les propriétés markoviennes (action, état, transition)** : la structure de dialogue est basée sur les propriétés de Markov. Les actions du système de dialogue représentent les possibilités des réponses et des questions que le système peut fournir à l'utilisateur. Les états représentent les connaissances du système (ressources internes et externes qui entretiennent en relation avec le système). La transition entre les états détermine quelle action doit être choisie pour aller à l'état suivant. Une stratégie optimale de dialogue consiste à choisir le bon couplage entre action et état. Cette approche a été utilisée dans plusieurs travaux de (Singh et al., 2002 ; Scheffler et Young, 2002 ; Pietquin, 2004, 2006).

---

<sup>8</sup> On utilisera les termes *système-initiative*, *utilisateur-initiative* ou mixte-initiative pour désigner une initiative du système, une initiative de l'utilisateur ou une initiative mixte (soit le système soit l'utilisateur qui prend le contrôle).

<sup>9</sup> Une instance, un stade, un état ou une situation de dialogue ont la même signification.

**Approche basée sur les cadres** : cette structure est la plus simple à gérer en termes de tâches à accomplir, car le dialogue est défini comme un cadre contenant l'ensemble des attributs variables –valeurs. Le dialogue est guidé par les questions/réponses du système. Les réponses de l'utilisateur sont les valeurs des variables. Cette approche permet d'utiliser la stratégie utilisateur-initiative qui laissant l'initiative aux utilisateurs de donner l'information, alors que le système associe les différentes réponses données aux différentes variables-valeurs (Goddeau et al., 1996 ; Chu-Carroll, 1999 ; Pieraccini et al., 2001) [GOD96], [CHU99], [PIER01].

**Approche basée sur la mise à jour de données** : cette architecture est un dérivé de l'architecture basée sur les cadres avec plus de complexité dans la construction de stratégie et pour des domaines plus complexes (Bohus et al., 2003) [BOH03].

**Approche basée sur l'agent collaboratif** : cette approche est utilisée dans COLLAGEN (Rich et al., 2001) [RIC01], elle modélise le comportement collaboratif des agents (humains) sur des tâches devant être partagées entre deux agents humains. La gestion du dialogue de ce système est guidé par la croyance, la tâche et le but de l'agent humain (Rich et al., 1998, Asai et al., 2005; Chu-Carroll et al., 1998,2000) [RIC98], [ASA05], [CHU00b], [CHU98].

### *Génération de la parole*

Le module de génération de la parole (GL) est situé en aval du module de synthèse et en amont du module de gestion du dialogue. Le processus de GL consiste à transformer les concepts (formés par le module de compréhension) en une forme textuelle compréhensible dans un langage humain. Pietquin (2004) distingue trois niveaux de GL dans ce processus qui va de la signification de mot jusqu'à la forme du mot : niveau sémantique, niveau syntactique et niveau lexical [PIE04].

#### II.3.1.3 Niveau pragmatique

##### *La connaissance du monde*

Au point de vue système informatique, l'information provenant de l'environnement extérieur est souvent représentée dans une base de données dans laquelle on essaie de rassembler, autant que possible, les éléments associés aux connaissances du monde. La connaissance du monde représente la connaissance, dont l'agent machine, lors de son apprentissage, est supposé être en mesure d'acquérir, plus la base de connaissance du monde est grande, plus il sera facile pour

l'agent machine d'élever son niveau de connaissance et d'augmenter la probabilité de reconnaître le but et la croyance de l'agent humain.

## II.3.2 Modélisation de la stratégie du dialogue

### II.3.2.1 Degrés d'initiatives

Le degré d'initiatives est une manière de contrôler l'interaction entre l'agent humain et l'agent machine (système). Il y a trois manières, pour l'agent machine, de prendre l'initiative dans un dialogue : soit le système prend le contrôle total du dialogue (système-initiative), soit le système laisse l'utilisateur parler de ce qu'il veut (utilisateur-initiative), soit le système emploie une stratégie *initiative-mixte* (Chu-Carroll, MIMIC system, 2000) [CHU00]. Dans le cas, où le système prend le contrôle total du dialogue, il est plus facile d'orienter l'utilisateur de manière à ce qu'il reste dans le contexte du dialogue et qu'il réponde directement à la question, la charge du gestionnaire du dialogue, dans ce cas, est moins lourde, puisque le système contrôle les entrées et les sorties des données. Par conséquent, le contexte du dialogue est mieux maîtrisé, par exemple, le système sait qu'il est en train de demander une date de réservation ou une confirmation de date ; il saura ainsi entreprendre les actions appropriées par rapport au but. Quant à la stratégie utilisateur-initiative, le système laisse le choix à l'utilisateur de transmettre sa demande et le système essaie d'interpréter ses messages. Cette stratégie laisse plus de libertés d'expression à l'utilisateur, quoique ce dernier, ne puisse dire n'importe quoi au système. La troisième stratégie réside dans le mélange entre les deux stratégies précédentes. L'initiative du dialogue peut être prise, soit par l'utilisateur, soit par le système, par exemple, au début d'un dialogue, le système peut commencer le dialogue en adressant à l'utilisateur une phrase d'invitation « *comment puis-je vous aider ?* » Ensuite, le système laisse l'utilisateur parler, s'il y a un doute, le système reprend le contrôle. Ce mode de fonctionnement semble plus naturel et plus souple que les deux précédents modes, car il laisse vraiment la place à la communication mutuelle (écouter/parler), mais ce mode est aussi plus difficile à gérer pour le système car il faut tenir compte de plusieurs paramètres (taux d'erreurs de reconnaissance, synchroniser les tours de parole, détection de parole croisée<sup>10</sup>).

---

<sup>10</sup> Les sons de la parole entrent dans le système mais ne font pas partie du dialogue. Par exemple L'utilisateur parle à une autre personne pendant que son micro est allumé.

### II.3.2.2 Stratégie de confirmation

Selon le résultat du module de reconnaissance, la confirmation d'une réponse d'utilisateur peut être explicite (i.e. « *Avez-vous dit deux personnes?* ») ou implicite (i.e. *combien de nuits resterez-vous à deux?*)

Choisir une stratégie de confirmation en fonction du résultat de la reconnaissance est une façon élégante et intelligente de contrôler le déroulement du dialogue, mais il faut tenir compte aussi du comportement des utilisateurs et de leurs intentions. Une stratégie efficace répondant le plus possible à des situations dialogiques et satisfaisant le plus grand nombre d'utilisateurs, n'est certainement pas évidente à modéliser. Dans certains cas, les utilisateurs préfèrent une réponse explicite, car cela leur donne l'impression que le système les comprend, mais pour d'autres, cette stratégie de confirmation peut être irritante et ennuyeuse, car la conversation est longue, répétitive et moins naturelle.

## II.4 Apprentissage machine

Il est important de connaître l'objet que l'on veut faire apprendre à la machine, avant de lui fournir une méthode d'apprentissage. On veut dans ce cas, faire apprendre à la machine une stratégie optimale (efficace) pour qu'elle puisse réaliser les actes de dialogue souple et naturel. Un des champs de recherche en intelligence artificielle qui a fait l'objet de nombreuses publications depuis quelques années est l'apprentissage machine. Faire apprendre à une machine à imiter le comportement cognitif humain consiste, avant tout, à chercher à comprendre le processus d'apprentissage humain, à étudier les capacités d'acquérir la connaissance, le raisonnement et la généralisation de l'être humain. C'est une activité cognitive. Le choix du mode de représentation de connaissance (en l'occurrence, c'est la stratégie) a un impact direct sur la mise en œuvre de ce processus d'apprentissage.

Les récents travaux de recherche suggèrent que nous utilisons le formalisme du processus de décision de Markov (PDM) comme représentation de dialogue et comme base pour la modélisation des problèmes d'optimisation de stratégie du dialogue (Biermann et Long, 1996; Levin et al., 2000; Walker et al., 2000; Pietquin, 2004; Cuayahuilt et Renals, 2006). Ce processus est décrit en termes d'états (ou situations), d'actions et de stratégies. Les états représentent l'historique et le contexte du discours, les actions sont les actes de dialogue et les stratégies sont des décisions quant aux choix de l'action à exécuter pour passer d'un état à un autre.

Nous décrivons dans cette section, d'abord, la modélisation d'une stratégie de dialogue pour ensuite présenter les principes d'apprentissage machine lesquels permettent de faire apprendre à la machine, une stratégie optimale.

#### II.4.1 Principes d'apprentissage machine

On distingue trois types d'apprentissage machine: supervisé, non-supervisé et par renforcement.

L'apprentissage supervisé consiste à entraîner la machine pour faire apprendre les nouvelles connaissances avec des prototypes d'exemples d'observation. Une fois entraînée, la machine sera capable de faire de classifier en discriminant entre les classes connues à *priori*. Elle sera également capable de classifier les nouveaux objets qu'elle n'a jamais vus auparavant. L'apprentissage non-supervisé ne reçoit pas de prototypes d'exemples. La machine doit apprendre à détecter les corrélations entre les exemples de données pour se construire des groupes non connus à l'avance.

Quant à l'apprentissage par renforcement, il est une forme particulière de l'apprentissage supervisé, mais au lieu d'avoir les prototypes d'exemples (les exemplaires) en entrée, la machine reçoit des indications approximatives sur sa réponse. La machine apprend par des essais-erreurs en interaction avec l'environnement pour trouver le classement adéquat des données, par exemple, un joueur de tennis sait qu'il gagne un point, s'il a fait un bon coup et en perd un sur un coup fautif, mais personne ne lui dit comment il faut faire pour gagner un point. C'est à lui de réajuster son coup pour parvenir à gagner le match.

Les expériences de conception des systèmes de dialogue avec des connaissances préalables accélèrent la tâche d'apprentissage de la machine (Singh et al., 2002; Pietquin, 2004), par exemple, pour apprendre une stratégie optimale du dialogue, nous pouvons nous demander quel type d'apprentissage sera convenable dans ce domaine ? Supervisé, non-supervisé ou renforcement ? Schatzmann et al. (2006) ont remarqué que, dans les corpus de dialogue humain-machine, on peut avoir des indications quant à la réussite ou à l'échec du dialogue, mais on ne peut pas savoir laquelle des actions exécutées par la machine est la plus efficace (ou la plus optimale) à un stade donné du dialogue. Autrement dit, on ne peut pas savoir quelle est la stratégie optimale étant donné qu'on n'a pas d'exemples idéals du comportement du gestionnaire

de dialogues. Voilà pourquoi, l'apprentissage supervisé n'est pas approprié pour résoudre les problèmes d'optimisation des stratégies de dialogue (Walker, 1993). Cela dit, l'apprentissage supervisé peut trouver les stratégies les plus employées, mais certainement pas celles qui sont optimales. Quant à l'apprentissage par renforcement, il peut être employé par interaction directe avec l'utilisateur réel pour entraîner le système à trouver les stratégies optimales. Les corpus de dialogue réel sont nombreux en dialogue humain-machine, ils constituent des exemples par excellence pour faire apprendre une stratégie de dialogue à la machine. Mais, interagir en direct et en temps réel entre l'utilisateur réel et la machine avec des grands volumes de données demande beaucoup de temps et nécessite des ressources côté utilisateur réel. L'astuce dans ce cas, consiste à créer un utilisateur qui simule l'utilisateur réel pour interagir avec la machine. Cependant, on peut apprendre au système une stratégie optimale, sans l'utilisateur simulé, à partir d'un corpus minimal de données construits à la main. Mais, cette approche demande des connaissances de l'expert du dialogue. Nous illustrerons d'un exemple de cette approche utilisée dans le système NJFun, ensuite nous présenterons l'approche avec l'utilisateur simulé.

#### II.4.2 Apprentissage par renforcement sans utilisateur simulé

Le système NjFun, construit en 2002 [SIN02], est un système qui informe les utilisateurs sur les différentes activités au New Jersey. Ce système fonctionne en temps réel et est basé sur la plateforme du système de dialogue de Levin, Pieraccini, Eckert, Fabrizio et Narayanan qui a été construite en 1999 [SIN02]. NjFun est composé de cinq modules : reconnaissance, compréhension, gestion du dialogue, accès aux bases de données, synthèse de la parole.

Le module de reconnaissance utilise l'engin de reconnaissance Watson basé sur un modèle stochastique de langue et un modèle de compréhension. L'engin de reconnaissance Watson est entraîné à partir des exemples des dialogues des utilisateurs. Le module de gestion du dialogue utilise l'approche d'AR pour entraîner une stratégie optimale, le module d'accès aux bases de données permet de rechercher les informations depuis la page Web nj.online qui contient les informations relatives aux activités proposées par la ville telles que les visites aux zoos, aux parcs d'attraction, aux musées, etc. Le module de gestion du dialogue emploie le langage DMD<sup>11</sup> pour

---

<sup>11</sup> DMD (Dialogue manager Developer) est un langage utilisé pour développer les actions de dialogue aussi bien que de construire les stratégies de dialogue.



gérer le dialogue, enfin, le module de synthèse de la parole utilise la méthode de concaténation des diphtonges pour transformer la réponse issue du module de gestion du dialogue en signal acoustique. Il est important de remarquer que le module de génération de langue dans ce système n'est pas séparé des autres modules, car la génération du texte est incorporée dans la réponse générée par le module de gestion du dialogue.

L'apprentissage par renforcement utilise la méthode Monte Carlo pour calculer la fonction de coût. La particularité du choix de l'approche employée dans ce système réside dans le fait que l'apprentissage de la stratégie ne se fait que sur les cas difficiles. Dans ce rôle, l'expert de conception du dialogue est important, car c'est à lui de choisir des cas typiques pour faire l'apprentissage du système. D'autres systèmes (Levin et al., 2000; Scheffler, 2002; Pietquin, 2004) préfèrent employer un utilisateur simulé au lieu de choisir manuellement les cas typiques. L'utilisateur simulé a pour but d'entraîner plus de corpus de données sans intervention humaine, mais le choix de NJFun est justifié par le fait qu'il travaille avec un modèle déterminé et un corpus minimal de données pour générer une stratégie optimale.

Pour calculer la fonction de récompenses, le système emploie la technique valeur-binaire<sup>12</sup>, une simple technique qui prend la valeur 1, si le dialogue est satisfaisant, sinon -1. Cette technique permet aussi d'évaluer la performance du système en termes de succès de la tâche, de la qualité du dialogue, de l'efficacité et de l'utilisabilité.

Ce système a su montrer par des techniques simples en AR (Q-Learning, MDP) que l'on peut optimiser les stratégies de dialogue dont les tâches sont bien décrites (42 stades, 2 actions possibles sur chaque stade). Mais, surtout, selon les auteurs Singh et al.(2002) [SIN02] ces résultats ont démontré que les applications d'AR permettent aux concepteurs d'optimiser empiriquement des stratégies de dialogue en recherchant à travers des espaces d'états plus larges que s'ils avaient été exploités avec des méthodes plus classiques. Néanmoins, plusieurs défis restent à résoudre afin de rendre le système plus robuste en termes de stratégie dialogique, comme par exemple, de trouver des méthodes pour réduire l'espace d'état à une taille gérable.

#### II.4.3 Apprentissage par renforcement avec un utilisateur simulé

---

<sup>12</sup> *Binary completion* en anglais.

Le projet TALK<sup>13</sup> (2006), utilise l'approche AR avec un utilisateur simulé. L'idée d'avoir un utilisateur simulé de dialogue pour identifier les stratégies de dialogue a été proposée par plusieurs chercheurs comme Levin et al. (1998, 2000) [LEV98], [LEV00], Litman et al. (2000) [LIT00], Walker (2000) [WAT00], Scheffler et Young (2001,2002) [SCH00],[SCH02], Pietquin et Renals, (2002) [PIE02], Pietquin et Dutoit (2006) [PIE06], Schatzmann et al. (2005,2006) [SCH05], [SCH06].

Levin et al. (1998) [LEV98] ont prouvé qu'un système de dialogue peut converger vers une stratégie raisonnable, même si on l'apprend avec une connaissance vide au départ. Levin, Pieraccini et Eckert (2000) [LEV00] ont fait des simulations au niveau de l'intention plutôt qu'au niveau acoustique. Mais leur simulation donne souvent des réponses irréalistes (i.e. « *Quelle est votre destination?* » « *New York* » ; « *Quelle est votre ville de départ?* » « *New York* »). Scheffler et Young (2000) ont trouvé que l'utilisateur simulé peut aider à trouver les stratégies de dialogue que les experts et développeurs ne voient pas. Litman et al. (2000) [LIT00] ont conclu que l'approche d'apprentissage par renforcement (AR) permet, en recherchant dans un espace d'états, d'optimiser empiriquement une stratégie de dialogue de façon beaucoup plus efficace (en recherchant dans un espace d'états) qu'en employant les méthodes traditionnelles (concevoir et tester empiriquement plusieurs versions de stratégies avant d'en choisir une pour l'implanter dans le système). En 2001, les travaux de Lin et Lee [LIN01] ont démontré que le simulateur est très utile dans l'analyse et la conception des systèmes de dialogue. En 2002, Levin et al. [LEV02] ont continué leurs travaux dans cette optique et ont aussi démontré qu'on peut apprendre un comportement relativement complexe des utilisateurs en choisissant quelques critères simples, par exemple, choisir une bonne représentation de l'espace d'état et un bon modèle d'utilisateur simulé. Leurs études sont comparables aux recherches heuristiques conçues par plusieurs autres groupes de chercheurs. Pietquin et Renals, (2002) [PIE02], dans leurs essais avec le réseau bayésien, ont démontré que, après plusieurs tests de simulation, le système reste stationnaire et semble converger vers une stratégie optimale. Chung (2004) [CHU04] a démontré que le simulateur faciliterait le développement des nouvelles applications des systèmes de dialogue, car il y a des millions de dialogues artificiels disponibles à développer et à tester. Les travaux de Schatzmann et al. (2005,2006) [SCH05], [SCH06] ont avancé que la recherche future dans le domaine du dialogue devraient inclure l'apprentissage machine basé sur l'utilisateur simulé. Ils ont démontré

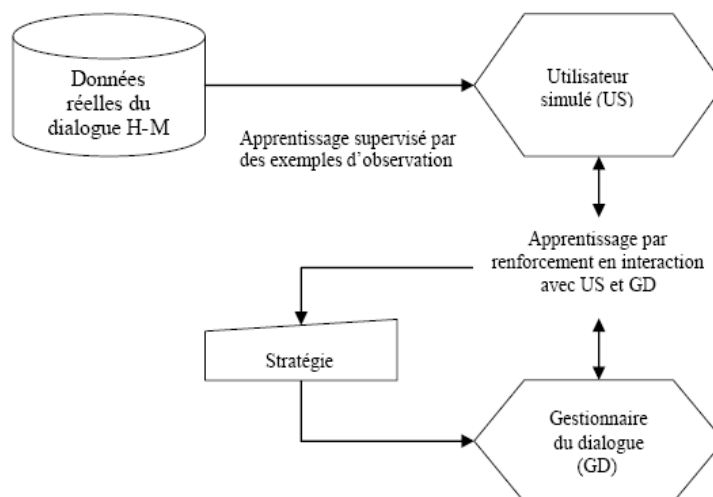
---

<sup>13</sup> Projet TALK. Disponible sur : (<http://www.talk-project.org/>) [consulté le 1 oct 2006].

que le choix des modèles d'utilisateurs a un impact significatif sur la stratégie apprise, ce qui veut dire que si on développe des modèles d'utilisateurs simulés plus proches des utilisateurs réels, alors cela rendra la simulation et l'environnement simulé plus réaliste. Par conséquent, cela fera augmenter la chance de développer des stratégies optimales. Cependant, en 2006, Pietquin et Dutoit [PIE06] ont affirmé que la simulation de dialogue et l'AR ne modélise pas entièrement la conception de dialogue, mais une première stratégie acceptable pourrait être obtenue avec la méthode bayésienne.

#### II.4.4 Architecture d'AR avec un utilisateur simulé

L'architecture générale de l'apprentissage avec un simulateur est illustrée dans la figure 2.2. D'abord, il faut apprendre au simulateur des stratégies optimales. Cet apprentissage se fait par la technique d'apprentissage supervisé en utilisant des exemples d'observation à partir des données réelles provenant du dialogue humain-machine. Ensuite, le simulateur interagit en direct avec le système qui apprend à réajuster sa stratégie en fonction des réponses et des remarques du simulateur. Cette forme d'entraînement permet de générer un grand nombre de dialogues et donc de créer des scénarios de stratégie qui ne se présentent pas dans le dialogue humain-machine. Ce qui est un avantage énorme pour l'exploration des différentes possibilités stratégiques.



**Figure 2.2** L'utilisateur simulé apprend des stratégies optimales du dialogue à la machine (d'après Schatzmann et al., 2006)

#### II.4.5 Méthodes d'AR

Rappelons que la stratégie optimale du dialogue peut être vue comme un problème d'optimisation. En AR, Il y a trois principales méthodes d'AR qui permettent de trouver les stratégies optimales : la méthode dynamique de programmation (DP), méthode Monte Carlo (MC) et la méthode de différence temporelle (DT) [SUT98].

La méthode DP est à la base de toutes les méthodes d'AR, elle est efficace et est la plus simple à implanter. Elle sert de valeur-itération et de description complète à tous les stades (connaissance complète de l'environnement) pour optimiser la stratégie dialogique qui est représentée par les associations stades-actions dans un processus de décision de Markov (PDM). Cependant, cette méthode est très coûteuse en termes d'opérations computationnelles. La méthode MC, quant à elle, est basée sur l'expérience et elle privilégie trois paramètres : les stades, les actions et les fonctions de récompense. La combinaison de ces deux méthodes DP et MC fait naître la troisième méthode appelée différence temporelle (DT). Dans cette méthode, les deux algorithmes les plus connus sont Q-Learning et Sarsa (Watkin; 1989, 1992) [WAT89], [WAT92]. L'algorithme Q-Learning permet de trouver l'action optimale en réalisant les expériences, sans avoir besoin de connaître tout le domaine (i.e. cet algorithme n'a pas besoin de connaissances prédéterminées). La différence entre Q-Learning et Sarsa se situe dans le mécanisme d'apprentissage par renforcement : *'Q-learning agents acquire optimal behaviors paying no attention to risk (e.g., danger of receiving large negative rewards), while Sarsa agents cannot acquire optimal behaviors but acquire behaviors that avoid risk.'* (Takadama et H. Fujita, 2005) [TAK05].

Il existe d'autres algorithmes qui combinent les différentes méthodes DP et DT comme l'algorithme 'Eligibility Traces' qui est basé sur un mécanisme d'affectation temporaire de valeur<sup>14</sup> [SUT98], [KAE96]. Puisque le problème avec l'AR est d'affecter la bonne valeur à la fonction de récompense pendant le déroulement du dialogue, alors cette valeur ne peut être connue qu'à la fin du dialogue (lorsque le PDM est au stade final). Donc, l'astuce de cet algorithme consiste à donner une valeur arbitraire en fonction de l'état actuel du dialogue et à le réajuster à la fin du dialogue.

---

<sup>14</sup> Terme traduit en anglais *temporal credit assignment*

Un autre problème d'AR est l'écart entre *l'exploitation* et *l'exploration* des différentes possibilités de stratégies. L'agent apprenant doit *exploiter* les stratégies apprises, afin d'obtenir la valeur de récompense, mais il doit aussi *explorer* des stratégies nouvelles afin de découvrir des meilleures actions. Pour être efficace dans son apprentissage, l'agent apprenant doit choisir entre essayer différentes actions et distinguer celles qui semblent être les plus efficaces. Les méthodes de base pour cela sont  *$\varepsilon$ -greedy* et *softmax* [SUT98], [KAE96], [PIE04]. La méthode  *$\varepsilon$ -greedy* consiste à choisir une meilleure action en fonction de la probabilité  $\varepsilon$ , sinon choisir une action aléatoire. La méthode *softmax*, quant à elle, est basée sur la méthode de Boltzmann, qui choisit l'action en basant sur sa *Q-valeur*.

## II.5 La plateforme informatique

La conception d'une application de dialogue humain-machine demande des experts en dialogue, en informatique aussi en acoustique. Ces applications demandent des développements manuels comme les systèmes NJFun, Clarissa, Hotel-Demo.

Par exemple, le système NJFun utilise le langage DMD pour la gestion du dialogue. A partir d'un minimum de règles de gestion des interactions de dialogue conçues à la main, le système s'entraîne par la méthode d'apprentissage par renforcement à trouver des stratégies optimales, pour ensuite implanter dans le système la stratégie choisie.

Clarissa, un navigateur vocal expérimental dont le procédé est basé sur la plateforme commerciale des outils de Nuance (structure XML, SICstus Prolog<sup>15</sup>, SVM<sup>16</sup>), est actuellement déployé sur la station spatiale internationale (International Space Station-ISS). Et ce représente probablement le premier système de dialogue oral dans l'espace. C'est un logiciel qui permet à des astronautes à bord de l'ISS d'interagir avec de diverses procédures de vol spatial à l'aide des communications verbales plutôt qu'avec les méthodes conventionnelles (i.e. écran, souris et

---

<sup>15</sup> Sicstus Prolog. Disponible sur : <http://www.sics.se/sicstus/> [consulté le 20-avril-2006]

<sup>16</sup> SVM : support vectors machines. Disponible sur : <http://svmlight.joachims.org/> [consulté le 20-avril-2006]

clavier), lesquelles ont été précédemment employées sur l'ISS (International Space Station) pour lire des procédures (la plupart du temps écrites dans le format pdf). Ces méthodes conventionnelles requièrent des manipulations manuelles des astronautes. Elles limitent leurs actions pour d'autres tâches (Rayner et al., 2005) [RAY05]. La gestion du dialogue dans Clarissa est basée sur une architecture de TRIPS [ALL00] et de TrindiKit [LAR00]. Les langages utilisés dans Clarissa sont XML, SICSus Prolog et les outils de Nuance.

Hotel-Demo utilise les outils de Nuance pour développer une version de démonstration de dialogue en utilisant les outils de Nuance comme VBuilder, OSR, VoiceXML.

## **II.6 Conclusion**

Nous avons parcouru l'état de l'art sur quatre aspects entourant le développement d'un SDOHM : les théories du dialogue humain-humain, les modèles de dialogue humain-machine, l'apprentissage et la plate-forme informatique. Nous nous sommes concentrés sur les points importants du contrôle des interactions de dialogue et de sa stratégie, sur l'apprentissage par renforcement avec et sans utilisateur simulé. Une stratégie de dialogue est dite optimale si les actions exécutées satisfont le but recherché par l'agent humain. Faire apprendre une stratégie de dialogue à une machine c'est lui donner la possibilité de contrôler des interactions de dialogue. Ce rôle est assuré par le gestionnaire du dialogue qui fait partie des tâches du module de gestion du dialogue.

L'approche AR avec un utilisateur simulé semble prometteuse dans la recherche de la stratégie optimale de dialogue. Les avantages de cette méthode sont nombreux : 1) il y a des millions de dialogues artificiels disponibles qui sont prêts à utiliser pour développer et tester de nouvelles applications ; 2) l'utilisateur simulé facilite le développement des nouvelles applications (Chung, 2004) [CHU04] et 3) l'apprentissage par renforcement et l'utilisateur simulé permet d'explorer et d'exploiter les stratégies nouvelles que l'approche traditionnelle manuelle ne pourrait pas la voir.

## CHAPITRE III

### NOTRE APPROCHE MÉTHODOLOGIQUE

#### III.1 Introduction

Dans ce chapitre, nous proposons une approche d'apprentissage par renforcement à l'aide d'un utilisateur simulé pour optimiser automatiquement les stratégies de dialogue, pour une application dans le domaine de la réservation des chambres d'hôtel. Une version sans apprentissage du système Hotel-Demo a été développée par la société Nuance en utilisant les approches classiques (concevoir manuellement plusieurs stratégies de dialogue, tester, optimiser et choisir une stratégie optimale pour implantation). Le problème d'une telle conception est qu'il devient lourd et coûteux d'y apporter des modifications. De plus, il est difficile de trouver une stratégie optimale répondant aux différents comportements des utilisateurs. Notre travail de recherche consiste à créer un environnement de simulation qui incorpore une architecture d'apprentissage permettant d'automatiser la tâche de conception d'une stratégie optimale de dialogue. Ainsi, cet environnement permettra de réduire le temps et le coût de développement.

Pour rendre plus claire et évidente l'explication de notre approche, nous limitons les tâches et but de notre SDOHM. Concernant le cas de la réservation des chambres d'hôtels dans l'application Hotel-Demo de Nuance, nous limitons les variables d'états et d'actions. Ainsi, l'utilisateur n'aura besoin de fournir que trois informations : la date de réservation (date d'entrée, date de sortie), le nombre de personnes que concerne cette réservation (adulte ou enfant) et le type de chambres (vue sur mer, fumeur, non fumeur). De plus, nous formalisons notre système de dialogue par sa description en termes des états, un ensemble d'actions et une stratégie dialogue qui cadrent dans le processus de décision de Markov. Nous définissons, par la suite, une fonction objective qui sert de mesure de performance pour déterminer la stratégie optimale. Ceci est une

condition nécessaire pour l'apprentissage automatique. Cette fonction modélise la perception et la satisfaction de l'utilisateur.

### III.2 Exemple du système Hotel-Demo de Nuance

Considérons l'exemple d'un dialogue oral entre un agent humain<sup>17</sup> (U : utilisateur) et un agent machine (S : Système) où le but du système est d'enregistrer une demande de réservation des chambres d'hôtel. Les informations requises sont le nombre de personnes (adultes ou enfants), les dates de réservation (début, fin), le type de chambre (vue sur mer, fumeur, non fumeur, etc.).

Ce système de dialogue comprend plusieurs composants comme le XML<sup>18</sup> grammars, l'OSR (OpenSpeech Recognizer), le VoiceXML et le Vbuilder. Certains de ces composants sont la propriété de Nuance comme l'OSR et le Vbuilder. La version de démonstration du système Hotel-Demo a aussi pour objectif de démontrer que les outils de Nuance sont suffisamment robustes pour permettre aux concepteurs et aux experts en dialogue de construire des systèmes de dialogue performants et efficaces. Ce système de dialogue emploie les stratégies suivantes (voir tableau 3.1).

1- **Initiative-mixte**. Le système ou l'utilisateur peut prendre l'initiative du déroulement du dialogue, par exemple, pendant l'ouverture du dialogue par le système (1), l'utilisateur peut interrompre le système en disant « *I'm ready* » (2). À n'importe quel moment du dialogue, l'utilisateur peut provoquer le retour à l'étape précédente de la conversation à l'aide de la commande « *go back* »(10).

2- **Information entrante multiple**. L'utilisateur peut donner plus d'une réponse au système « *I'd like to reserve one room for this week-end* » (4). Il y a deux propositions données par l'utilisateur dans cet énoncé dont le système doit extraire le *sens* : (a) vouloir réserver une chambre et (b) donner les dates de réservation. Le système doit d'abord extraire l'information

---

<sup>17</sup> Les termes agent humain, humain ou utilisateur ont la même signification. Les termes agent machine, machine ou système sont sémantiquement identiques.

<sup>18</sup> Le langage XML (eXtensible Markup Language), développé par la W3C. Disponible sur : <http://w3.org> [consulté le 10 sept 2004]



relative à la réservation d'une chambre et traduire le mot *week-end* en deux attributs *date de début* et *date de fin* de réservation, ainsi il peut déduire le nombre de jours réservés.

3- **Vérification explicite et implicite.** Lorsqu'il y a ambiguïté dans l'énoncé (taux d'erreurs de reconnaissance élevés, formes anaphores, ellipses, agrammaticales, etc.), le système utilise la vérification explicite (5) pour s'assurer d'avoir bien capté l'information. Autrement, le système utilise la vérification implicite c'est-à-dire inclure la réponse de l'utilisateur dans la question suivante (la vérification implicite n'est pas illustrée dans cet exemple).

1	<b>S: Welcome to Hotel demo... you can use..... Are you ready?</b>
2	U: I'm ready
3	<b>S: When would you like to reserve?</b>
4	U: I'd like to reserve one room for this week-end
5	<b>S: Ok, you'd like to reserve one room, check in Friday, check out Sunday, right?</b>
6	U: Yes
7	<b>S: How many people would you like to reserve?</b>
8	U: Ten adults
9	<b>S: Sorry, I didn't get that. How many people would you like to reserve?</b>
10	U: Go back
11	<b>S: How many people?</b>
12	U: Three adults and one child
13	<b>S: Ok, you'd like to reserve one room, check in Friday and check out Sunday for three adults and one child, right?</b>
14	U: Yes
15	<b>S: Your reservation number is 123459. Thank you for calling Hotel Demo.</b>

**Tableau 3.1** Exemple d'un dialogue extrait du système Hotel-Demo développé par Nuance (2006)

4- **Méthode 'form-filling/slot-filling'** (remplissage des formes ou variables). Le système emploie cette méthode pour transférer l'information de l'utilisateur vers le système de dialogue. Ce transfert d'information peut être vu comme une extraction sémantique des phrases et être représentée sous forme de concept. Les concepts sont représentés comme un ensemble de paires d'attributs-valeurs, par exemple le concept *réservation* contient les attributs tels que *nombre de chambres*, *les dates* et *le nombre de personnes*. Chaque attribut a une valeur correspondante sauf pour l'attribut *date* qui se réfère au concept *date* (attributs *date de début* et *date de fin*).

5- **Historisation/récapitulation des données.** Lorsque les informations sont complètement fournies, en supposant que le domaine de réservation ne nécessite que ces trois concepts (nombre de chambres, dates et nombre de personnes), le système récapitule

l'information, vérifie explicitement et termine le dialogue en donnant à l'utilisateur un numéro de réservation (13 à 15).

### III.3 Cheminement méthodologique

La figure 2.2 (chapitre II) montre l'architecture générale de notre système d'apprentissage avec un utilisateur simulé qui remplace l'agent humain pour interagir avec l'agent machine en assistant dans son apprentissage. Nous décrivons la démarche à suivre pour concrétiser cette approche, afin de pouvoir améliorer le système Hotel-Demo actuel.

1. À partir des transcriptions brutes (sans annotation sémantique) de dialogue réel entre l'utilisateur et le système Hotel-Demo, créer un corpus de données avec annotations sémantiques.
2. Créer un utilisateur simulé et entraîner ce dernier avec le corpus construit en 1. La méthode d'apprentissage supervisée est employée à cette étape. Une fois apprise, elle permettra à l'utilisateur simulé d'interagir en direct avec le système pour lui apprendre une stratégie optimale.
3. Construire les états et les actions du système selon le formalisme de PDM pour la tâche de réservation de chambres d'hôtel (nombre d'actions, nombre d'états, de stratégies).
4. Développer un système de dialogue vide, c'est-à-dire sans aucune stratégie de dialogue.
5. Implanter l'algorithme d'apprentissage Q-Learning dans ce système. Cet algorithme est basé sur la description dans Cuayahuilt et al., (2005) [CUA05].
6. Spécifier les valeurs de la fonction de récompense. Ces valeurs permettent au système de réajuster son action, afin de trouver la stratégie optimale.
7. Spécifier les paramètres de la fonction objective qui est cruciale pour l'évaluation du système final (pas pour évaluer l'utilisateur simulé).
8. Connecter l'utilisateur simulé au système pour apprendre la stratégie optimale.

9. Implanter cette stratégie dans le système Hotel-Demo actuel.

### III.4 Dialogue humain-humain

Nous basons sur la théorie de communication humain-humain, des actes de langages et de la croyance pour interpréter l'intention de l'utilisateur. Ce choix permet au système de modéliser sa stratégie de réponse en fonction du but de l'application et pour résoudre l'ambiguïté dans l'énoncé due à un mauvais résultat de reconnaissance ou une mauvaise interprétation.

### III.5 Dialogue humain-machine

Basé sur les approches proposées dans la littérature, nous avons choisi une représentation markovienne pour représenter les actes de dialogue. Cette représentation nous permet de modéliser les actes de dialogue complexes. Elle facilitera l'apprentissage de stratégie optimale par la suite.

Nous assumons que notre système est orienté but et nous trouvons importance d'assurer un bon échange à chaque étape du dialogue. Cet échange est mesurable par le nombre de questions implicites (i.e. « *When do you want to reserve for 2 adults?* »), de questions explicites (i.e. « *Do you want to reserve for 2 adults?* ») et de questions répétitives (i.e. « *Sorry, can you repeat please?* ») du système, ainsi que par le nombre de questions sans réponse. En effet, cette quantification représente la performance du système. Elle représente la fonction de performance (ou fonction de coût) qui est basée sur la qualité d'échange et non pas sur la longueur du dialogue comme l'envisage English et Heeman (2005) [ENG05].

#### III.5.1 Degré d'initiatives

La meilleure façon de choisir un degré d'initiatives qui soit plus près de la communication s'établissant naturellement entre les humains est de combiner dynamiquement tous ces degrés d'initiatives en fonction de taux de reconnaissance. Si le taux de reconnaissance est faible (trop d'erreurs de reconnaissance), alors la stratégie système-initiative s'applique. Si le taux de reconnaissance est très élevé et le module de compréhension performant (extraction sémantique les énoncés) alors on peut utiliser la stratégie utilisateur-initiative. Dans certains cas, on peut commencer avec une stratégie mixte-initiative et revenir à la stratégie système-initiative ou utilisateur-initiative, en fonction des paramètres de reconnaissance.

Nous proposons une technique de paramétrage pour choisir le degré d'initiatives comme suit :

soit un ensemble de variables de seuil de tolérance  $\Delta$  basé sur le taux d'erreurs de reconnaissance issu du module de reconnaissance de la parole et sur le taux d'erreurs d'interprétations sémantiques issu du module de compréhension de la parole. Le système se basera sur cette valeur, selon les contraintes suivantes pour choisir quel degré d'initiatives à employer :

$\Delta_1$  est la somme minimum acceptable de  $TE_r$  et  $TE_c$  où  $TE_r$  désigne le taux d'erreurs de reconnaissance et  $TE_c$  désigne taux d'erreurs d'interprétation sémantique,

$\Delta_2$  est la somme maximum acceptable de  $TE_r$  et  $TE_c$  avec  $\Delta_1 < \Delta_2$ .

Lors d'un acte de dialogue (lorsque l'utilisateur a donné une réponse), on obtient une valeur courante  $X = TE_r + TE_c$ .

- Si  $X < \Delta_1$  la performance du système est bonne, on emploie le mode user-initiative.
- Si  $\Delta_1 < X < \Delta_2$  la performance du système est correcte, on emploie le mode mixte-initiative.
- Si  $X > \Delta_2$  la performance est mauvaise, le système prend le contrôle du dialogue (système-initiative).

### III.5.2 Stratégie de confirmation

Nous proposons une stratégie de confirmation dynamique qui tient compte des valeurs de  $\Delta_1$  et de  $\Delta_2$  définies dans la section degré d'initiatives.

Lors d'un acte de dialogue (lorsque l'utilisateur a donné une réponse), on a une valeur courante  $X = TE_r + TE_c$ .

- Si  $X < \Delta_1$  la performance du système est haute, on n'emploie pas la stratégie de confirmation, mais le système passe à la question suivante.
- Si  $\Delta_1 < X < \Delta_2$  la performance du système est correcte, on emploie la stratégie de confirmation implicite.
- Si  $X > \Delta_2$  la performance est mauvaise, le système emploie la stratégie de confirmation explicite pour s'assurer la qualité de réponse.

### III.5.3 Dialogue comme processus de décision de Markov

L'approche AR demande une représentation du dialogue en termes d'actions, d'états et de stratégies, lesquels sont des propriétés de Markov. L'apprentissage par renforcement qui cadre avec les propriétés markovienne, s'intitule le processus de décision de Markov. Nous décrivons toutes les variables des propriétés de Markov servant à décrire notre dialogue dans le domaine de la réservation des chambres d'hôtel. Nous décrivons également des méthodes d'apprentissage et des paramètres des fonctions de mesure de performance qui sont indispensables pour l'apprentissage.

### III.5.3.1 Les actions

Les actions du système de dialogue représentent les possibilités des réponses et des questions que le système peut fournir à l'utilisateur. Nous définirons huit types d'actions du système :

- 1- Ouverture de dialogue ou invité. (i.e. *"How can I help you? are you ready?"*).
- 2- Question. Cette action correspond à la stratégie système-initiative. L'utilisateur répond à une question spécifique pour remplir une variable par exemple la question « *when do you want to leave?* » est pour connaître la date de départ.
- 3- Clarifier une information par une vérification
  - a. Implicite. Cette action est appelée lorsque la reconnaissance est haute (i.e. « *what kind of rooms do you want for 2 days?* »)
  - b. Explicite. Cette action est appelée lorsque la reconnaissance est moyenne (i.e. « *Do you want to leave Saturday?* »)
  - c. Erreur. Cette action est appelée lorsque la reconnaissance est imprécise (i.e. « *sorry, I didn't get that..Can you repeat that?* » )
- 4- Aider. Cette action permet d'informer l'utilisateur sur la façon dont il doit fournir les renseignements. (i.e. « *if you want to leave tomorrow just say tomorrow.* »)
- 5- Confirmation totale. Cette action est appelée lorsque toutes les variables sont remplies (i.e. « *Ok, you'd like to reserve one room, check in Friday and check out Sunday for three adults and one child, right?* »)
- 6- Fermeture du dialogue (i.e. « *Thank you for calling Hotel Demo. Good bye.* »)

### III.5.3.2 Les états

Puisque notre système est défini comme étant orienté vers un but, l'enchaînement d'un état à un autre est conçu comme une sorte d'automate déterministe<sup>19</sup>. Lorsque le but est atteint, on dit que son état est terminal ; autrement, son état est non terminal.

Les états de notre système de dialogue sont censés être représentatifs de toutes les connaissances du système (les ressources internes et externes qui interagissent avec le système). Dans notre exemple, les ressources internes sont les entrées/sorties du module gestion de dialogue et la ressource externe est l'utilisateur (simulé ou non). Les états du système incluent quatre attributs : les deux dates de réservation (date d'entrée, date de sortie), le nombre de personnes pour qui les chambres sont réservées (adulte ou enfant) et le type de chambre (vue sur mer, fumeur, non fumeur).

Pour contrôler toutes les actions du dialogue, notre système de dialogue représente les états de dialogue sous forme de vecteurs opérationnels. Ces vecteurs sont des valeurs numériques portant le statut de chaque variable. On peut voir ces valeurs numériques comme une sorte de 'photo' prise pendant une séquence de dialogue. Le tableau 3.2 illustre les différentes valeurs pouvant être affectées à un état. Un simple calcul montre que pour 8 actions avec 4 variables d'états, on obtient  $(8^4) = 4096$  possibilités. De plus, la variable d'état **Attribut** est la variable ayant le plus de valeur (4 pour **Attribut**, 2 pour **valeur**, 3 pour **Erreur** et 3 pour **Répétition**). Donc, le nombre maximum de possibilités de stratégie est environ un peu moins de  $(4096)^4 = 281474976710656$ . Cela devient vite inadmissible pour un système qui ne peut gérer ces innombrables possibilités. Il faut optimiser manuellement les stratégies « évidentes » pour faciliter l'apprentissage machine, par exemple, au début du dialogue, il ne peut pas y avoir une action de vérification implicite, ni de confirmation totale ou de fermeture de dialogue à moins que le taux de reconnaissance soit très mauvais, mais même dans cette condition, le système ne prend pas l'initiative de mettre fin à un dialogue. Ce travail de prétraitement permet de réduire énormément l'espace d'états de dialogue, par conséquent, le système évite d'introduire inutilement les cas anodins dans l'apprentissage.

---

<sup>19</sup> Finite state Machine. Disponible sur : [http://en.wikipedia.org/wiki/Finite\\_state\\_machine](http://en.wikipedia.org/wiki/Finite_state_machine) [consulté le 15 jan 2004]

Variables d'état	Abréviation	Valeurs	Explications
Attribut	A	0, 1, 2, 3	Indique l'attribut courant sur lequel le système travaille (0 = date d'entrée, 1 date de sortie, 2 = nombre de personne, 3 = type de chambre)
Valeur	V	0,1	0 = valeur non obtenue, 1 = valeur obtenue
Erreurs	E	0, 1, 2	Cette valeur permet au système de choisir l'action de clarification. 0 = mauvais, le système choisira l'action vérification d'erreur. De plus si le nombre de répétition = 2, l'action aide sera choisie à la place de l'action vérification d'erreur. 1 = moyen, action vérification explicite sera choisie 2 = correct, action vérification implicite sera choisie
Répétition	R	0, 1, 2	0 = aucune répétition, première fois, 1 = une fois, action vérification explicite sera choisie, 2 = deux fois, action aide sera choisie.

**Tableau 3.2** Les paramètres des états de dialogue et leurs valeurs.

### III.5.3.3 Les stratégies

La stratégie détermine quelle action doit être choisie pour passer à l'état suivant. Le but de stratégie optimale consiste à choisir le couplage adéquat entre action et état en tenant compte de degré d'initiatives et de la stratégie de confirmation choisies. Ce choix s'effectuera par l'apprentissage par renforcement en connectant le système avec l'utilisateur simulé.

## III.6 Apprentissage

Notre approche d'AR est basée sur une architecture décrite dans (Levin et al., 2000; Schatzmann et al. 2006). Elle comporte deux étapes. D'abord, l'utilisateur simulé est entraîné avec les corpus de dialogue issus de Hotel-Demo. Ensuite, le système, en interaction directe avec l'utilisateur simulé, apprend à réajuster sa stratégie par rapport aux réponses et aux remarques de l'utilisateur simulé. Ce qui n'est pas possible avec un utilisateur réel, en raison du grand volume de données à explorer et à exploiter pour trouver les stratégies optimales. Nous employons l'algorithme Q-Learning pour l'apprentissage par renforcement en raison de sa simplicité et de son efficacité à trouver une action optimale, sans avoir besoin de connaissances *à priori*.

Les motivations pour adopter cette architecture résident dans le fait que l'approche d'AR a été appliquée en intelligence artificielle (IA) pour résoudre de nombreux problèmes nécessitant un agent d'apprentissage qui interagit avec son environnement dans le but d'augmenter la performance (problèmes de contrôle, de recherche d'opérations, de jeux, etc.) (Sutton et al., 1998;

Singh et al., 2002) [SUT98], [SIN02]. Mais, l'avantage principal de l'approche d'AR réside dans son potentiel pour calculer une stratégie optimale à l'intérieur d'un large espace d'états et sa capacité d'apprendre en direct (*online*). Cependant, cette approche présente aussi des défis considérables, car il faut choisir adéquatement les corpus de données pour entraîner le simulateur et ensuite, choisir les bons paramètres d'apprentissage en tenant compte du comportement des utilisateurs, du taux d'erreurs de reconnaissance, de la tâche de l'application, etc. De plus, il faut définir une fonction de calcul de coût permettant de trouver le chemin optimal parmi de multiples possibilités de solutions, enfin il faut choisir judicieusement la fonction objective permettant d'évaluer la qualité de la stratégie.

### III.6.1 Paramètres pour l'apprentissage par renforcement

#### III.6.1.1 Propriété de Markov

Lorsque l'état du monde (objet, concept, etc.) peut être décrit de manière compacte, formelle et dans un ordre de séquence successive captant toute (ou presque toute) l'information, alors on dit que cet état représente des propriétés markoviennes ou simplement un état de Markov. Les propriétés de Markov sont formellement décrites dans une équation mathématique (équ.1) :

$$Pr \{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t, r_t, s_{t-1}, a_{t-1}, \dots, r_1, s_0, a_0\}, \quad (1)$$

La probabilité de passer à l'état  $s_{t+1}$  et de recevoir une valeur de récompense  $r_{t+1}$  dépend de l'action  $a_t$ , de  $r_t$  et de l'état  $s_t$  qui lui-même dépend de l'état à l'instance  $t-1$  et ainsi de suite jusqu'à l'état initial  $s_0$  et l'action initiale  $a_0$  l'état. Ce qui permet de généraliser comme ceci :

$$Pr \{s_{t+1} = s', r_{t+1} = r \mid s_t, a_t\}. \quad (2)$$

L'idée principale de propriété de Markov réside dans les paramètres de décision qui influencent le choix d'une action au lieu d'une autre action. Autrement dit, l'état ou la valeur de récompense à l'instance  $t+1$  dépend seulement de l'action et de l'état à l'instance  $t$  (instance présent) (equ. 2)

#### III.6.1.2 Valeur de la fonction de récompense



On détermine les valeurs pour la fonction de récompense qui servent d'indicateurs de réponses (feed-back) dans l'apprentissage par renforcement. De plus, la fonction de récompense permet de remplacer la méthode d'évaluation classique proposée par PARADISE, car cette méthode (demande) exige des commentaires de la part des utilisateurs réels.

Nous pouvons attribuer les valeurs suivantes pour la fonction de récompense telle qu'indiquée dans le tableau 3.3

Chaque attribut complété	10
Chaque attribut confirmé	10
Chaque question explicite	5
Chaque question implicite	5
Chaque question erreur, aide ou répétitive	-5
Abandon	-10
Maximum score possible sans erreur pour remplir les 4 attributs en une séquence de dialogue (une confirmation par attribut, plus un échange d'invité au début, une confirmation totale et une fermeture	$(25 \times 4) + 3 = 103$

**Tableau 3.3** Valeurs définies pour la fonction de récompense.

### III.6.1.3 Fonction objective

Il existe plusieurs méthodes pour mesurer cette performance. Par exemple, Levin et al. (2000, 2002) définissent une fonction objective  $C = \sum C_i$  où  $C$  représente la somme des mesures de performances telles que le nombre d'interactions, d'erreurs et d'attributs non complétés, etc. Levin et al. illustrent leur proposition avec un exemple du système de dialogue qui consiste à obtenir le jour et le mois en un nombre minimum d'interactions. Puisque, nous avons une stratégie de dialogue basée sur la qualité d'échange et non pas sur le nombre de tours d'interactions, alors nous choisissons les paramètres d'évaluation suivants pour la fonction objective  $C_i = W_{imp}(N_{imp}) + W_{exp}(N_{exp}) + W_{rep}(N_{rep}) + W_{inc}(N_{inc})$ .

La variable  $C_i$  représente le coût total des paramètres de récompense. Les variables  $W_{imp}$ ,  $W_{exp}$ ,  $W_{rep}$ ,  $W_{inc}(N_{inc})$  représentent les poids associés aux valeurs  $N$  déterminant le rapport de force parmi ces coûts. Les variables  $N_{imp}$ ,  $N_{exp}$ ,  $N_{rep}$  et  $N_{inc}$  représentent respectivement le nombre de questions implicites, explicites, répétitives et de variables incomplètes représentant les questions sans réponses. Les rapports  $N_{imp}/N_{exp}$ ,  $N_{imp}/N_{rep}$ ,  $N_{exp}/N_{rep}$  donnent une mesure d'estimation de la performance du système.

### III.7 Plateforme informatique

Étant donné la grande diversité des outils existant sur le marché et la pluridisciplinaire du domaine de recherche, nous choisissons une plateforme flexible et adaptative pour qu'elle puisse nous offrir des modules flexibles et adaptatifs dans d'autres domaines. Nous aurons le choix entre les outils de Nuance (pour les modules de reconnaissance et synthèse du signal de parole et de gestion de dialogue) et la plateforme SATIM pour l'analyse textuelle.

### III.8 Conclusion

Nous avons proposé un cadre de conceptualisation de stratégie de dialogue pour implanter dans un SDOHM. Nous avons aussi proposé une démarche méthodologique pour réaliser et implanter l'approche d'apprentissage par renforcement à l'aide d'un utilisateur simulé dans la machine. Nous avons donné des détails sur des paramètres de MDP (actions, états, stratégies) et des variables d'apprentissage (fonction objective, valeur de récompense). Une fois apprise, nous implanterons cette stratégie dans la machine.

Pour concrétiser cette approche, nous établissons une liste de tâches à réaliser pour démontrer notre démarche. Cette liste est conçue en fonction des étapes de cheminement méthodologique et en fonction de la plate-forme informatique choisie. Pour l'instance, nous utilisons la plate-forme de Nuance et de ses outils (V-Builder 3.0, OSR, ASR, TTS, etc.) :

- Décrire les logiciels utilisés (les modules de Nuances et quelques exemples de code)
- Donner des pseudo-codes des modules importants
- Montrer des diagrammes sur interfaces des différents modules du système.
- Comment les modules de Nuances sont utilisés/adaptés dans cette réalisation
- Comment le corpus de données (Hotel-Demo) est construit (sa taille, de quels exemples est-il extrait ?, etc.) ?
- Analyser de sensibilité sur les rapports des valeurs de N de la fonction objective. Par exemple, si on varie un peu les méthodes, est-ce que ces rapports changent ? Et dans quelles proportions?, ..
- Donner un tableau des résultats des rapports quand on utilise le système construit avec le corpus de données ci-dessus. Si le corpus est plus grand ou il s'étend sur plusieurs domaines, le système produit-il les mêmes résultats ou différents (meilleurs, pire ou inchangé)? Et pourquoi ?
- Comment comparer les performances du système proposé avec d'autres systèmes ?

## CHAPITRE IV

### ÉVALUATION

#### IV.1 Méthode d'évaluation

Walker et al. (1997) [WAL97] ont proposé la méthode d'évaluation PARADISE qui a été largement acceptée par la communauté scientifique pour évaluer les systèmes de dialogue humain-machine. Cette méthode a été appliquée dans l'évaluation du système INSPIRE [MÖL04] (*'smart home système'*). Cependant, elle a fait l'objet de plusieurs critiques par un grand nombre de chercheurs, notamment Williams et Young (2004) [WIL04]<sup>20</sup>. Le point nébuleux de cette méthode est qu'il est difficile de définir de manière objective et constante le niveau de satisfaction des usagers quant à la tâche à accomplir. Ainsi, le lien entre la satisfaction des usagers et le temps de complétion d'une tâche n'a pas été mis en évidence. Même si les usagers ne manifestent pas leur mécontentement, il ne serait pas acceptable qu'un système prenne plus que le temps nécessaire pour répondre à la demande des usagers. De plus, il n'est pas démontré dans le modèle de Walker que le succès de la tâche dépend de la stratégie de gestion des erreurs [SKA03], c'est-à-dire que la stratégie de gestion des erreurs ne soit pas considérée comme pertinente et primordiale pour accomplir la tâche dans un système du dialogue humain-machine. Ce qui est totalement inacceptable dans cette méthode d'évaluation, car on sait qu'une meilleure gestion des erreurs permettrait d'augmenter la chance de compléter la tâche et donc de donner satisfaction aux usagers. Cependant, il est vrai qu'il est difficile d'évaluer la satisfaction des usagers, donc difficile d'évaluer un système de dialogue, étant donné que le dialogue est subjectif et souvent réalisé dans des contextes particuliers (domaine complexe, comportement des usagers, tâches complexes, etc.). Néanmoins, sachant que le corpus devant être utilisé dans cette

---

<sup>20</sup>Young Researchers' Roundtable on Spoken Dialog Systems. Disponible sur : <http://www.cs.cmu.edu/~dod/YRR/proceedings.pdf> [consulté le 17 avril 2006]

recherche provient des enregistrements de conversations réelles au cours des tests de demande de réservation de chambre d'hôtel réalisé avec la version Hôtel-Démo actuelle, nous avons relevé quelques points importants pour évaluer les résultats de la recherche.

**Au niveau de satisfaction globale du dialogue.** La pertinence de la stratégie employée dans des situations différentes se mesure par la satisfaction des agents humains (AH) ainsi que par la réponse donnée par le système (réponse courte, précise et désambiguïsée). La rhétorique du dialogue (l'art de bien parler) joue aussi un rôle non négligeable dans le succès d'un système de dialogue. Une stratégie est dite optimale lorsqu'elle permet à l'AH de rencontrer son but (l'AH confirmera à la fin du dialogue si le système a compris son but ou non). Nous distinguons 4 situations dans lesquelles le système doit faire face au dialogue oral :

**Situation a :** l'utilisateur ne connaît pas le système, mais coopère avec le système (il suit les directives du système et fait des contre-indications),

**Situation b :** l'utilisateur ne connaît pas le système et ne coopère pas avec le système (il donne des réponses erronées, hors contexte, change des buts, revient sur la confirmation, pose des questions),

**Situation c :** l'utilisateur connaît le système et il coopère à 100%, mais anticipe les questions posées par le système (réponse à multiples concepts « *Nous sommes 4 et voulons 2 lits Kings* »),

**Situation d :** l'utilisateur connaît le système et il coopère à 100% sans contrainte (dialogue idéal),

**Au niveau de stratégie de confirmation.** La gestion des ambiguïtés est effectuée par des vérifications explicites et/ou implicites. Par exemple, si le taux d'erreurs de reconnaissance est élevé, le système utilise la vérification explicite, dans le cas contraire, le système choisit l'option de vérification implicite. Dans ce cas, le rôle de l'évaluateur est de relever la pertinence des choix des questions du système. Les questions explicites doivent être posées de façon convenable pour ne pas alourdir le système.

**Au niveau de l'interprétation des énoncés.** L'interprétation du type d'acte de langage : forme illocutoire, anaphore (répétition sémantique) et ellipse (les références à un ou plusieurs niveaux historiques). Cette mesure est particulièrement nécessaire dans les situations b et c ci-haut, où l'utilisateur cherche à faire échouer le système.

Nous cherchons à évaluer la capacité à retrouver les liens entre les propositions et à maintenir un état épistémologique du système, c'est-à-dire à demeurer rationnel et cohérent dans le domaine d'application, tout en gardant une souplesse du dialogue pour éviter que l'agent humain sorte du contexte de discours. Cet exemple illustre un dialogue contenant des anaphores et ellipses : (U : Utilisateur, S : Système).

1	<i>U: I'd like to reserve one room with view of beach, a beautiful view please [forme anaphore]</i>
2	<b>S: How many people would you like to reserve?</b>
3	<i>U: Hum... how much for 2 adults?</i>
4	<b>S: 100\$ per night and per adult</b>
5	<i>U: Ok, 2 [forme ellipse mais ambiguë, sous-entendu 2 nuits ou 2 adultes ?]</i>
6	<b>S: Ok, would you like to reserve for 2 adults with view of the beach?</b>
7	<i>U: Yes</i>

**Tableau 4.1** Exemple de dialogue contenant des anaphores et ellipses

## IV.2 Conclusion

Nous avons proposé la méthode d'évaluation PARADISE pour tester la performance de Hotel-Demo. Par cette évaluation, nous voulions vérifier notre hypothèse et notre méthodologie concernant le modèle de dialogue humain-machine, en particulier la stratégie de dialogue (degré initiatives et la stratégie de confirmation), l'apprentissage machine (la stratégie apprise est-elle optimale ?) et enfin concernant la plate-forme informatique, nous n'avons pas de mesures spécifiques pour vérifier si les programmes développés dans cette plate-forme sont flexibles et adaptés avec d'autres bibliothèques de fonctions. Cela est attribuable à la difficulté de mettre en place une telle mesure pour une plate-forme informatique. Cependant, nous pouvons facilement tester la flexibilité de la plate-forme en observant comment les programmes sont structurés (code source libre, programmation modulaire, création des librairies de fonction, d'objets, la facilité de modifier un code, etc).

## CHAPITRE V

### DISCUSSION

#### V.1 Introduction

Le but ultime de notre travail est de réaliser des outils d'apprentissage qui permettront aux développeurs novices de réaliser des systèmes de dialogue performants en termes de stratégies de dialogue et ainsi rendre davantage de robustesse et de naturel aux futurs systèmes. L'approche que nous avons choisie est basée sur la méthode stochastique. Nous pensons qu'il sera intéressant de comparer notre méthode avec l'approche bayésienne (Pietquin, 2004) pour savoir si les deux méthodes pourraient se fondre. Nous sommes en cours de réflexion sur ces points. Cependant, l'apprentissage machine pourrait être efficace si et seulement si nous possédions des données cohérentes.

#### V.2 Conclusion préliminaire

Durant notre recherche, nous avons constamment remis en question nos hypothèses et notre réflexion : l'hypothèse de départ est-elle juste ? Le processus de décision de Markov donne-il des résultats plus valides que d'autres méthodes, par exemple ceux du réseau Bayésien ou ceux de méthodes traditionnelles (conception manuelle)? Si oui, dans quels cas? La stratégie choisie (degrés d'initiatives et stratégie de confirmation) permet-elle de réaliser des dialogues souples et naturels?

Selon nous, un système de dialogue humain-machine est considéré comme efficace, robuste et performant lorsqu'il répond à ces critères :

- 1- Le système est capable de maintenir un échange naturel et souple du dialogue.

- 2- À la fin de la conversation ou du dialogue, les deux parties arrivent à un consensus (consentement) ou un compromis général qui satisfasse les deux parties.
- 3- L'abandon du dialogue n'est pas dû au système (mauvaise réception du signal de la parole, temps de réponse trop long, mauvaise interprétation, erronée, incompréhension, mauvaise génération de la parole, réponse du système inappropriée).
- 4- Il n'est pas nécessaire d'avoir la même capacité cognitive que l'homme pour pouvoir entretenir un dialogue avec lui (notre hypothèse de départ).
- 5- Approfondir l'apprentissage machine, car c'est la seule façon pour un système d'élever sa connaissance et sa performance.
- 6- Choisir une plate-forme informatique flexible et adaptative pour pouvoir gérer des changements et des évolutions de programmes.
- 7- Travailler sur les modèles de données, c'est-à-dire les annotations sémantiques doivent être plus précises, car les corpus de données ont un impact sur l'apprentissage machine surtout lorsque l'apprentissage est basée sur les exemples d'observation.

La puissance des ordinateurs n'est pas un facteur déterminant pour un système de dialogue humain-machine robuste, ni le module de reconnaissance et de synthèse de la parole. Ces modules ne font que traduire les représentations symboliques des structures de surface de la parole en une structure de surface du texte et vis-versa. Quant au module de l'interprétation au sens de compréhension de la parole, il est important de bien comprendre (interpréter) les énoncés en entrée du système pour traiter correctement les données, mais cela n'est pas une fonction majeure du SDOHM. Concernant la gestion dialogique par l'emploi d'une stratégie optimale, cette gestion contribue partiellement à formuler une réponse intelligente, elle guide et oriente le dialogue dans un sens ou dans un autre sans se perdre dans le dialogue. Le module de génération de la parole s'occupe, à partir des règles grammaticales et de la stratégie de réponse prises en aval par la gestion de stratégie dialogique, de construire des réponses sous forme de propositions (phrases, énoncés), autrement dit, le module de génération construit une structure de surface supposant de représenter la connaissance – le message parlé – Mais la génération de la parole est en quelque sorte commanditée par la gestion de stratégie dialogique. Il n'y a pas de réponse

intelligente sans une gestion intelligente. Donc, la puissance des ordinateurs ne permet seulement de rendre le dialogue plus agréable à entendre (voix plus naturelle, moins hachée, temps de réponse plus rapide), mais le cœur du problème reste toujours de savoir comment générer une sortie vocale cohérente avec le contexte du discours, autrement dit, comment générer une connaissance pertinente dans une situation donnée? Nous émettons l'hypothèse que l'outil adéquat pour générer la connaissance se trouve en partie dans la stratégie dialogique, car sans la stratégie dialogique, le module de génération ne peut donner une réponse intelligente. Par exemple, si une personne possède toutes ces capacités physiques et mentales, mais qui ne sait pas communiquer, elle ne peut pas établir un dialogue avec les autres. Tout comme un SDOHM, sans une stratégie de dialogue bien contrôlé, le dialogue ne peut pas se dérouler de façon souple et naturelle.

Nous parlons de la façon dont un système de dialogue est construit en regard de la science cognitive. Ce système, doit-il ressembler à notre comportement de communication entre humains? ou alors, en tenant compte des limites d'intelligence de la machine et de sa performance en calcul, devons-nous nous inspirer de cette théorie (de la communication humain-humain) pour trouver une autre façon d'établir une communication entre l'humain et la machine? (l'oiseau versus l'avion, la marche versus la voiture, etc.) Tel est notre point de réflexion pour de futures directions de recherche.

### V.3 Plan d'avancement

En suivant les étapes définies dans la section cheminement méthodologique, nous avons établi approximativement un plan de thèse tel que :

- oct.-06 : état de l'art, développement expérience 1 (système sans apprentissage)
- déc.-06 : conception du système avec apprentissage
- jan-07 : analyser les résultats (système avec apprentissage) + écriture des chapitres
- mars-07 : conclusions + 2 articles acceptés pour AIAI2007 (Cinquième Conférence Internationale sur l'Automatisation Industrielle 2007)
- avril-07 : dépôt
- juillet-07 : défens



## BIBLIOGRAPHIE

- [ABE99] A. Abella and A. Gorin. 1999. Construct Algebra: Analytical Dialog Management. In *Proc. ACL*, Washington D.C.
- [ALL00] J. Allen, D. Byron, M. Dzikovska, G. Ferguson, L. Galescu and A. Stent. 2000. An Architecture for a Generic Dialogue Shell. *Natural Language Engineering*. Vol. 6. Cambridge University Press.
- [ALL01] J. Allen, D. Byron, M. Dzikovska, G. Ferguson, L. Galescu and A. Stent. 2001. Towards conversational human-computer interaction. *AI Magazine*, 22(4):27-38.
- [ALL76] J. Allwood. 1976. *Linguistic Communication as Action and Cooperation*. Gothenburg Monographs in Linguistics 2. Göteborg University, Department of Linguistics, Sweden.
- [ALL79] J. Allen. 1979. A plan-based approach to speech act recognition. Technical Report 131/79, University of Toronto. PhD thesis.
- [ALL96] J. Allen, B.W. Miller, E.K. Ringger and T. Sikorski. 1996. A Robust System for Natural Spoken Dialogue. *Proc. of ACL '96*.
- [AMA93] R. Amalberti, C. Valot. 1993. Le Magicien d'Oz. CERMA, journée du PRC Rhône-Alpes.
- [ANT01] J.Y. Antoine, J.Goulian. 2001. Word order variations and spoken man-machine dialogue in French : a corpus analysis on the ATIS domain. *Corpus Linguistics'2001*, UCREL, Lancaster.
- [ANT01b] J.Y. Antoine, J.Goulian. 2001. Étude des phénomènes d'extraction en français parlé sur deux corpus de dialogue oral finalisé. *TAL*, 42:1-20.
- [ARA97] M. Araki, T.Watanabe and S. Doshita. 1997. Evaluating dialogue strategies for recovering from misunderstandings. In *Proc. IJCAI Workshop on Collaboration Cooperation and Conflict in Dialogue Systems*, pages 13-18.
- [ASA05] H. Asai, T. Koshizen, M. Watanabe, H. Tsujino and K. Aihara. 2005. Cognitive User Modeling Computed by a Proposed Dialogue Strategy Based on an Inductive Game Theory. *Studies in Computational Intelligence (SCI)*, 7, 325-351. Springer-Verlag Berlin Heidelberg.
- [AUS62] J.L. Austin. 1962. *How to Do Things with Words*. Harvard University Press, Cambridge, Massachusetts.
- [BAL99] B. Balentine, D.P. Morgan. 1999. How to build a *Speech Recognition Application*, Enterprise Intergration Group., CA.
- [BIL91] E. Bilange. 1991. An approach to oral dialogue modelling. In *Proc. of Second Venaco Workshop on the structure of multimodal dialogue*, Acquafredda di Maratea, pp. 1-12.
- [BLA05] N. Blaylock and J. Allen. 2005. Generating artificial corpora for plan recognition. In Liliana Ardissono, Paul Brna, and Antonija Mitrovic, editors, *User Modeling 2005*, in *Lecture Notes in Artificial Intelligence*. Springer, Edinburgh, 3538:179-188
- [BOE88] B. Boehm. 1988. A Spiral Model of Software Development and Enhancement. *IEEE Computer*, 21(5).
- [BOH03] D. Bohus and A. Rudnicky. 2003. RavenClaw: Dialog Management Using Hierarchical Task Decomposition and an Expectation Agenda. In *Proc. of Eurospeech*, Geneva, Switzerland, pp. 597-600.
- [CAR83] S.K. Card, T.P. Moran and A. Newell. 1983. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, 1983.
- [CHA92] D. J. Chalmers. 1992. Subsymbolic Computation and the Chinese Room. In *The Symbolic and Connectionist Paradigms: closing the gap*, ed: John Dinsmore, Lawrence Erlbaum Associates, pp. 25-48.
- [CHO56] N. Chomsky. 1956. Three Models for Description of Languages. *IRE. Transaction on Information Theory*, pp. 113-124.
- [CHO65] N. Chomsky. 1965. *Aspect of the Theory of Syntax*. MIT Press, Cambridge, MA, 1965.
- [CHU00a] J. Chu-Carroll. 2000. MIMIC: An Adaptive Mixed Initiative Spoken Dialogue System for Information Queries. *ANLP*, pp. 97-104
- [CHU00b] J. Chu-Carroll and S. Carberry. 2000. Conflict resolution in collaborative planning dialogs. *Int. J. Hum.-Comput. Stud.* 53(6): 969-1015.

- [CHU04] G. Chung. 2004. Developing a Flexible Spoken Dialog System Using Simulation. In *ACL*, pp. 63-70, Barcelona, Spain.
- [CHU98] J. Chu-Carroll and M.K Brown. 1998. An evidential model for tracking initiative in collaborative dialogue interactions. *User Modeling and User-Adapted Interaction*, 8(3-4):215-253.
- [CHU99] J. Chu-Carroll. 1999. Form-Based Reasoning for Mixed-Initiative Dialogue Management in Information-Query Systems. In *Proc. of Eurospeech*, Budapest, Hungary, pp. 1519-1522.
- [CUA05] H. Cuayahuitl, S. Renals, O. Lemon and H. Shimodaira. 2005. Human-Computer Dialogue Simulation Using Hidden Markov Models. in *Proc. of IEEE ASRU Workshop*, Cancun, Mexico.
- [CUA06] H. Cuayahuitl, S. Renals, O. Lemon and H. Shimodaira. 2006. Reinforcement Learning of Dialogue Strategies With Hierarchical Abstract Machines. In *Proc. of IEEE/ACL Workshop on Spoken Language Technology*. Palm Beach, Aruba.
- [ECK97] W. Eckert, E. Levin and R. Pieraccini. 1997. User Modeling for Spoken Dialogue System Evaluation. in *Proc. of IEEE ASRU Workshop*, Santa Barbara, Cal., USA.
- [ELZ95] S. Elzer. 1995. The role of user preferences and problem-solving knowledge in plan recognition for expert consultation systems. In *Working Notes of the IJCAI-95 Workshop on The Next Generation of Plan Recognition Systems*, pages 37-41, Montreal, Canada.
- [ENG05] M. English and P. Heeman. 2005. Learning mixed initiative dialog strategies by using reinforcement learning on both conversants. In *Proc. of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pp. 1011-1018, Vancouver, Canada.
- [FAB04a] G. Di Fabbrizio and C. Lewis. 2004. Florence: a Dialogue Manager Framework for Spoken Dialogue Systems. In *Proc. of the 8th International Conference on Spoken Language Processing (ICSLP 2004)*, Jeju Island, Korea.
- [FAB04b] G. Di Fabbrizio and C. Lewis. 2004. An XPath-based Discourse Analysis Module for Spoken Dialogue Systems. *The 9th International World Wide Web Conference*, New York.
- [FIL68] Ch. J. Fillmore. 1968. *The case for case*. Universals in Linguistic Theory, pages 1-90.
- [FOU04] Y. Fouquet. 2004. *Modélisation des attentes en dialogue oral*. Thèse de doctorat d'informatique, Université Grenoble I.
- [GAI93] B.R. Gaines and M.L.G. Shaw. 1993. Eliciting Knowledge and Transferring It Effectively to a Knowledge-Based System. *IEEE Trans. Knowl. Data Eng.* 5(1): 4-14.
- [GEO05] K. Georgila, J. Henderson and O. Lemon. 2005. Learning user simulations for information state update dialogue systems. In *Eurospeech*, Lisbon, Portugal.
- [GOD00] D. Goddeau and J. Pineau. 2000. Fast Reinforcement Learning of Dialog Strategies. *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP2000)*.
- [GOL05] C. Golanski, J. Caelen. 2005. *Test d'utilisabilité de l'agent conversationnel « Angela »*, WACA, Grenoble, France.
- [GOR02] G. Gorrell, I. Lewin and M. Rayner. 2002. Adding Intelligent Help to Mixed Initiative Spoken Dialogue Systems. In *Proc. of ICSLP*.
- [GRI89] H.P. Grice. 1989. *Study in the way of words* Cambridge. Harvard UP.
- [HOC04] M. Rayner and B.A. Hockey. 2004. Side Effect Free Dialogue Management in a Voice Enabled Procedure Browser. In *Proc. of INTERSPEECH*, Jeju Island, South Korea.
- [HOC05] B.A. Hockey and M. Rayner. Proceedings of the AAAI Workshop on Spoken Language Understanding, Pittsburgh, PA. 2005.
- [HOC88] J.M. Hoc. 1988. *Cognitive Psychology of Planning*. Academic Press London.
- [JAR05] I. Jars. 2005. Contribution des Sciences Sociales dans le domaine de l'Intelligence Artificielle Distribuée : ALONE, un modèle hybride d'agent apprenant. Thèse de Doctorat, Université Claude Bernard - Lyon.
- [JAR05a] I. Jars, N. Kabachi et M. Lamure. 2005. Proposal for a Vygotsky's theory based approach for learning in MAS. *The AAAI-04 Workshop on Agent Organizations: Theory and Practice*, San Jose, California.
- [JOH96] B.E. John and D.E. Kieras. 1996. Using GOMS for user interface design and evaluation: Which technique?. *ACM Transactions on Computer-Human Interaction*, 3, pp. 287-319.
- [KAE96] L.P. Kaelbling, M.L. Littman and A.W. Moore. 1996. Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research*, 4:237-285.
- [LAR00] S. Larsson and D. Traum. 2000. Information State and Dialogue Management in the TRINDI Dialogue Move Engine Toolkit. In *Natural Language Engineering*, 1:1-17.
- [LAV92] N. Lavrac and I. Mozetic. 1992. Second generation knowledge acquisition methods and their application to medicine. In Keravnou, E. (Ed.) *Deep Models for Medical Knowledge Engineering*, pages 177-198. Elsevier.

- [LEV00] E. Levin, R. Pieraccini, W. Eckert. 2000. A Stochastic Model of Human Machine Interaction for Learning Dialog Strategies. In *Proc. of the IEEE ICASSP*, Istanbul, Turkey, pp. 1883-1886.
- [LEV81] S. Levinson. 1981. Some pre-observations on the modelling of dialogue. *Discourse Processes*, 4(1):93-116.
- [LEV83] S. Levinson. 1983. *Pragmatics*. Cambridge University Press.
- [LEV97] E. Levin, R. Pieraccini and W. Eckert. 1997. A Stochastic Model of Computer-Human Interaction for Learning Dialog Strategies. In *Proc. of Eurospeech*, Rhodes, Greece, pp. 1883-1886.
- [LEV97b] E. Levin, R. Pieraccini and W. Eckert. 1997. Learning Dialogue Strategies within the Markov Decision Process Framework. In *Proc. of ASRU'97*, Santa Barbara, California.
- [LIN01] B. Lin and L. Lee. 2001. Computer-Aided Analysis and Design for Spoken Dialogue Systems Based on Quantitative Simulations. In *Proc. of the IEEE, Transactions on Speech and Audio Processing*, 9(5):534-548.
- [LIT00] D.J. Litman, S. Pan. 2000. Predicting and Adapting to Poor Speech Recognition in a Spoken Dialogue System. In *Proc. of 7th National Conference on AI (AAAI-2000)*, Austin, Texas.
- [LIT84] D.J. Litman, J. Allen. 1984. A Plan Recognition Model for Clarification Subdialogues. In *Proc. of Coling84*, pp. 302-311.
- [LIT85] D.J. Litman. 1985. Plan Recognition and Discourse Analysis : an Integrated Approach for Understanding Dialogues. Thèse de l'Université de Rochester.
- [LOP03] R. Lopez-Cozar, A. De la Torre, J.C. Segura and A.J. Rubio. 2003. Assessment of Dialogue Systems by Means of a New Simulation Technique. In *Speech Communication*, 40:387-407.
- [LUZ95] D. Luzzati. 1995. *Le dialogue verbal homme-machine*, Masson, Paris.
- [MEU05] J.G. Meunier. 2005. Module 1: La cognition. Note du cours de Doctorat en Informatique Cognitif, Département de Philosophie, UQAM. [page web]. [www.unites.uqam.ca/philosophie/cours/DIC81004/textes/notes\\_cours1-06.pdf](http://www.unites.uqam.ca/philosophie/cours/DIC81004/textes/notes_cours1-06.pdf); [consulté le 10-10-2005].
- [MEU06] J.G. Meunier. 2006. Introduire à la philosophie du langage. Partie 1 : L'action langagière. Note du cours de philosophie PHI-1004. Département de Philosophie, UQAM.
- [MIN00] W. Minker and S. Bennacef. 2000. *Parole et dialogue homme-machine*. Paris, Edition d'Eyrolles, 212.p.
- [MIN02] W. Minker. 2002. Overview on recent activities in speech understanding and dialogue systems evaluation. In *ICSLP-2002*, pp. 337-340.
- [MIN97] W. Minker, S. Bennacef and J.L. Gauvain. 1997. A Stochastic Case Frame for Natural Language Understanding. *Proc. of ICSLP'97*.
- [MIT97] T. Mitchel. 1997. *Machine Learning*. Mc Graw Hill.
- [MOE89] J. Moeschler. 1989. *Modélisation du Dialogue : Représentation de l'Inférence Argumentative*. Hermès, Paris.
- [MÖL04] S. Möller, J. Krebber, A. Raake and al. 2004. INSPIRE: Evaluation of a Smart-Home System for Infotainment Management and Device Control. In *Proc. 4th Int. Conf. on Language Resources and Evaluation (LREC 2004)*, P-Lisbon, 5:1603-1606.
- [MYE91] R. B. Myerson. 1991. *Game Theory: Analysis of Conflict*. Harvard University Press.
- [MYE98] B.A. Myers. 1998. A Brief History of Human Computer Interaction Technology. *ACM interactions*. Vol. 5, no. 2, March, 1998. pp. 44-54.
- [NÉE99] F. Néel, and W. Minker. 1999. *Computational Models of Speech Pattern Processing*. Springer Verlag, Berlin/Heidelberg, Germany.
- [NER93] P. Nerzic. 1993. Erreurs et échecs dans le dialogue oral homme-machine, détection et réparation. Thèse de l'Université de Rennes I.
- [NGU04] H. Nguyen et J. Caelen. 2004. *Multi-session Management in Spoken Dialogue System*. C. Lemaître, C.A. Reyes, J.A. Gonzalez éditeurs, *Advances in Artificial Intelligence*, IBERAMIA 2004, Springer ed.
- [NGU05] H. Nguyen. 2005. Dialogue homme-machine : modélisation de multisession. Thèse de doctorat. Université Joseph Fournier - Grenoble I.
- [NGU06] H.P. Nguyen, C. Dan. 2006. A Basic Mathematical Framework for Conceptual Graphs. *IEEE Transactions on Knowledge and Data Engineering*, 18(2):261-271.
- [NGU07] M.Q. Nguyen, Philip H.P. Nguyen, T.H Nguyen, D. O'Shaughnessy. 2007. A Proposed AI Planning Approach for Staff Management in a Service Center. To appear In *Proc. Of Fifth International Conference on Industrial Automation*. ETS, Montreal. Paper accepted.
- [NGU07] M.Q. Nguyen, Philip H.P. Nguyen, T.H Nguyen, J.G. Meunier, D. O'Shaughnessy. 2007. Apprentissage par renforcement à l'aide d'un utilisateur simulé pour optimiser automatiquement les stratégies de dialogue. À paraître dans Cinquième Conférence Internationale sur l'Automatisation Industrielle. ETS, Montréal. Papier accepté.

- [OSH00] D. O'shaughnessy. 2000. *Speech Communication: Human and Machine*. Addison-Wesley Publishing Co., Reading, MA, 1987; 2nd edition, 536 pages, IEEE Press, 568 pages.
- [OST89] N.D.M. Ostler. 1989. LOQUI: How Flexible can a formal prototype be?. In *M.M.Taylor, F. Neel & D.G. Bouwhuis (Eds.), The Structure of Multimodal Dialogue*, North-Holland:Netherlands.
- [PIE02] O. Pietquin and S. Renals. 2002. ASR System Modeling for Automatic Evaluation and Optimization of Dialogue Systems. In *Proc. of the IEEE ICASSP*, Orlando, USA, pp. 46-49.
- [PIE04] O. Pietquin. 2004. A Framework for Unsupervised Learning of Dialogue Strategies. Presses Universitaires de Louvain, *SIMILAR Collection*, ISBN 2-930344-63-6.
- [PIE05] O. Pietquin and R. Beaufort. 2005. Comparing ASR Modeling Methods for Spoken Dialogue Simulation and Optimal Strategy Learning. *Proc. of Interspeech/Eurospeech*, Lisbon, Portugal.
- [PIE06] O. Pietquin and T. Dutoit. 2006. A Probabilistic Framework for Dialog Simulation and Optimal Strategy Learning. *IEEE Transactions on Audio, Speech and Language Processing*, 14(2):589-599.
- [PIER01] R. Pieraccini, S. Caskey, K. Dayanidhi, B. Carpenter, B. and M. Phillips. 2001. ETUDE, A Recursive Dialogue Manager With Embedded User Interface Patterns. In *Proc. of IEEE ASRU Workshop*, Madonna di Campiglio, Italy.
- [PRI90] P. Price. 1990. Evaluation of spoken language systems: The ATIS domain. In *Proc. 1990 DARPA Speech and Natural Language Workshop*, Morgan Kaufmann.
- [RAM89a] L.A. Ramshaw. 1989. A metaplan model for problem-solving discourse. In *Proc. of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*, pp. 35-42, Manchester, England.
- [RAM89b] L.A. Ramshaw. 1989. Pragmatic Knowledge for Resolving Ill-Formedness. PhD thesis, University of Delaware, Newark, Delaware.
- [RAM91] L.A. Ramshaw. 1991. A three-level model for plan exploration. In *Proc. of the 29th ACL*, pp. 39-46, Berkeley, CA.
- [RAY05] M. Rayner, B. Hockey, J.M. Renders, N. Chatzichrisafis and K. Farrell. 2005. Spoken Language Processing in the Clarissa Procedure Browser. *International Computer Science Institute*, California.
- [REB02] J. Rebecca. 2002. How can a dialogue system compensate for speech recognition deficiencies?. Department of Linguistics, Goteborgs University. Available at : [http://www.speech.kth.se/~rolf/gslt\\_papers/RebeccaJonson.pdf](http://www.speech.kth.se/~rolf/gslt_papers/RebeccaJonson.pdf)
- [RIC01] C. Rich, C.L. Sidner and N.B. Lesh. 2001. COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction. *Artificial Intelligence Magazine*, 22(4):15-25.
- [RIC98] C. Rich and C.L. Sidner. 1998. COLLAGEN:A collaboration manager for software interface agents. *User Modeling and User-Adapted Interaction*, 8(3-4):315-350.
- [ROL05] C. Rolland. 2005. L'ingénierie des méthodes : une visite guidée. e-TI, Vol.01, <http://www.revue-eti.net/document.php?id=726>.
- [ROU00] J. Rouillard. 2000. Hyperdialogue sur Internet – Le système HALPIN. thèse de l'Université Joseph Fourier. CLIPS-IMAG.
- [ROU01] E. Roulet, L. Filliettaz, A. Grobet. 2001. Un modèle et un instrument d'analyse de l'organisation du discours. Berne, Peter Lang.
- [ROU81] E. Roulet. 1981. Echanges, interventions et actes de langage dans la structure de la conversation. *Etudes de linguistique appliquée*, 44:5-39.
- [RUB98] J. Rumbaugh, I. Jacobson, G. Booch. 1998. *The Unified Modeling Language Reference Manual*. Addison-Wesley.
- [RUS03] S. Russell, P. Norvig. 2003. *Artificial Intelligence: A modern approach*, Pearson Education, 2003, 2nd edition.
- [SCH01] K. Scheffler, S. Young. 2001. Corpus-Based Simulation for Automatic Strategy Learning and Evaluation. In *Workshop on Adaptation in Dialogue Systems (NAACL)*, Pittsburgh, Pennsylvania, USA.
- [SCH02] K. Scheffler, S. Young. 2002. Automatic learning of dialogue strategy using dialogue simulation and reinforcement learning. In *Proc. Human Language Technology*, San Diego, 12-18.
- [SCH05] J. Schatzmann, K. Geogila, S. Young. 2005. Quantitative Evaluation of User Simulation Techniques for Spoken Dialogue Systems. In *Proc. of Workshop on Discourse and Dialogue*, Lisbon, Portugal.
- [SCH06] J. Schatzmann, K. Weilhammer, M. Stuttle, S. Young. 2006. A Survey of Statistical User Simulation Techniques for Reinforcement-Learning of Dialogue Management Strategies. *Knowledge Engineering Review*.
- [SCH99] K. Scheffler and S. Young. 1999. *Simulation of Human-Machine Dialogues* Cambridge, U.K.: Engineering Dept., Cambridge University, Tech. Rep. CUED/F-INFENG/TR 355, 1999.
- [SEA69] J.R. Searle. 1969. *Speech Acts*. Cambridge : Cambridge University Press.

- [SEA80] J.R. Searle. 1980. Minds, brains and programs. *Behavioral and Brain Sciences*, 3:417-424.
- [SEA85] J.R. Searle and D. Vanderveken. 1985. *Foundations of Illocutionary Logic*. Cambridge : Cambridge University Press.
- [SIN02] S. Singh, D. Litman, M. Kearns and M. Walker. 2002. Optimizing Dialogue Management with Reinforcement Learning: Experiments with the NJFun System. *Journal of Artificial Intelligence Research*, 16:105–133.
- [SIN99] S. Singh, M. Kearns, D. Litman and M. Walker. 1999. Reinforcement Learning for Spoken Dialogue Systems. *Proc. of NIPS'99*, Denver, USA.
- [SOU00] B. Souvignier, A.Kellner, B. Rueber, H. Schramm and F. Seide. 2000. The thoughtful elephant - strategies for spoken dialog systems. In *IEEE Transactions on Speech and Audio Processing*, 8(1):51-62.
- [STA06a] The Stanford University, 'Stanford Encyclopedia of Philosophy: SHRDLU', [page web]. <http://hci.stanford.edu/~winograd/shrdlu/>; [consulté 29-2-2006].
- [STA06b] The Stanford University, 'Stanford Encyclopedia of Philosophy.', [page web]. <http://plato.stanford.edu/entries/chinese-room/>; [consulté 29-2-2006]
- [STE99] A. Stent, J. Dowding, J.M. Gawron, E.O. Bratt and R. Moore. 1999. The CommandTalk spoken dialogue system. In *Proc. of the 37th Annual Meeting of the Association for Computational Linguistics*, pp. 183-190.
- [SUT98] R. Sutton and A. Barto. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- [TAK05] k. Takadama and H. Fujita. 2005. Toward Guidelines for Modeling Learning Agents in Multiagent-Based Simulation: Implications from Q-Learning and Sarsa Agents. In *Multi-Agent and Multi-Agent-Based Simulation*, Springer Berlin / Heidelberg, ISBN 978-3-540-25262-7, 3415:159-172.
- [TAL99] M. Tallis, J. Kim, J. and Y. Gil. 1999. User studies of knowledge acquisition tools: Methodology and lessons learned. In *Proc. of KAW*.
- [THO04] C. Thompson, M. Goker, P. Langley. 2004. A Personalized System for Conversational Recommendations. In *Journal of Artificial Intelligence Research*, 21:1-36.
- [TUR50] A.M. Turing. 1950. Computing Machinery and Intelligence. *Mind* 49: 433-460. Site Internet <http://www.turing.org.uk/turing/> consulté mai-2005.
- [VAN91] D. Vanderveken. 1991. Meaning and Speech Acts. *Formal Semantics of Success and Satisfaction*, vol. 2, Cambridge: CUP.
- [VIL99] L. Villasenor-Pineda. 1999. Contribution à l'apprentissage dans le dialogue homme-machine. Thèse de doctorat, Université Joseph-Fourier - Grenoble I.
- [WAL00] M.A. Walker. 2000. An Application of Reinforcement Learning to Dialogue Strategy Selection in a Spoken Dialogue System for Email. In *Journal of Artificial Intelligence Research*, 12:387-416.
- [WAL97] M.A. Walker, D.J. Litman, C.A. Kamm and A. Abella. 1997. PARADISE: A Framework for Evaluating Spoken Dialogue Agents. In *Proc. of the 35th Annual Meeting of the Association for Computational Linguistics (ACL-97)*, pp. 271-280, Madrid, Spain.
- [WAL98] M. Walker, J. Fromer and S. Narayanan. 1998. Learning optimal dialogue strategies: A case study of a spoken dialogue agent for email. In *Proc. of ACL/COLING 98*.
- [WAT89] C. Watkins. 1989. Learning from delayed rewards. PhD Thesis, University of Cambridge, England.
- [WAT92] C. Watkins and P. Dayan. 1992. Technical note: Q-learning. *Machine Learning*, 8:279-292.
- [WIL04] J. D. Williams and S. Young. 2004. Characterizing Task-Oriented Dialog using a simulated ASR Channel. *Proc. of the ICSLP*.
- [WIN72] T. Winograd. 1972. *Understanding Natural Language*. Edinburgh, Edinburgh University Press.
- [XUE04] A. Xuereb, J. Caelen. 2004. Un modèle d'interprétation pragmatique en dialogue homme-machine basé sur la SDRT. Actes de TALN'04, *XIème Conférence sur le Traitement Automatique du Langage Naturel*, ISBN 2-9518235-5-5, pp. 505-514, Fès, France.
- [YEH05] H. Ye, S. Young. 2005. Improving Speech Recognition Performance of Beginners in Spoken Conversational Interaction for Language Learning. *Proc. of Interspeech05*, Lisbon, Portugal.
- [ZUE00] V.W. Zue, J.R. Glass. 2000. Conversational Interfaces: Advances and Challenges. In *Proc. of the IEEE. Special Issue on Spoken Language Processing*, 88(8).