

# MODÉLISATION STOCHASTIQUE DU DIALOGUE PAR STRUCTURES SÉMANTIQUES

Florian Pinault

Affiliation (339, chemin des Meinajaries - Agroparc BP 1228 - 84911 AVIGNON Cedex 9, France)

Tél. : +33 04 90 84 35 04 - Fax : +33 04 90 84 35 01

Courriel : florian.pinault@univ-avignon.fr

## ABSTRACT

In Human-Machine interaction, mixed-initiative spoken dialogue systems are under investigation to allow users to speak freely to the machine. Nevertheless, natural language dialogue systems often lack the required robustness to ensure user satisfaction. A solution may lie in using a rich semantic representation and statistically modelling the dialogue sequence. This paper presents an overview of a Human-Machine dialogue system, an introduction to the statistical model of *Partially Observable Markov Decision Process* (POMDP) and a hierarchical semantic frame representation model. Then, an application of these two ideas is presented, using *summary spaces* instead of full spaces to make the computation more tractable. Some experimental results are presented, showing that this approach indeed increases system robustness. More investigations are needed to ensure practical implementations in full spaces will exhibit the same robustness.

## 1. INTRODUCTION

Dans le domaine de l'interaction homme-machine, les systèmes de dialogue à initiative mixte sont actuellement étudiés, afin de permettre aux utilisateurs de parler librement avec la machine. Cependant, les systèmes de dialogue en langue naturelle manquent souvent de la robustesse nécessaire pour assurer la satisfaction de l'utilisateur. Une solution consiste à utiliser une représentation sémantique riche du dialogue, ainsi qu'une modélisation statistique du cours du dialogue.

Dans cet article, nous présentons dans la section 1 une vue d'ensemble d'un système de dialogue générique. La section 2 donne une introduction sur les modèles statistiques utilisés : processus de décision de Markov totalement ou partiellement observables (*Partially Observable Markov Decision Process* POMDP). Dans la section 3 est détaillé un modèle de représentation structurée en graphe de *frames* sémantiques. L'application de ces idées est décrite dans la section 4, introduisant en particulier la notion d'espaces résumés dans lesquels est projeté le dialogue afin de rendre le problème d'optimisation calculable. Des résultats expérimentaux sont présentés dans la section 5.

### 1.1. Système de dialogue générique

Grâce aux dernières avancées dans le domaine de la reconnaissance vocale et à l'augmentation de la puissance de calcul des appareils mobiles, le dialogue homme-machine devient une technologie attendue avec impatience par le public.

Un système de dialogue peut être décrit de la façon suivante :

Considéré comme un système organique, le système perçoit son environnement à travers des *capteurs/stimuli*, modélise le monde à travers sa représentation de l'*état de dialogue* et génère une *action* vers l'extérieur.

Les capteurs/stimuli de la machine sont de deux types :

- capteurs/stimuli utilisateur : ce sont ceux qui perçoivent les actions de l'utilisateur telles que le signal acoustique du microphone ou l'entrée textuelle du clavier.
- capteurs internes : ils n'ont pas de liens directs avec l'utilisateur, c'est le cas par exemple de l'accès à une base de données (DB : DataBase).

Chaque capteur constitue un modèle à part entière appliqué par les algorithmes du système à ses stimuli d'entrée. Différents types de modèles peuvent être utilisés dans ces capteurs parmi lesquels on trouve les perceptrons et les réseaux de neurones, les classificateurs, des modélisations probabilistes apprises sur corpus ou par simulation, des règles logiques, des motifs (*patterns*). On trouve aussi parmi ces modèles, les algorithmes de reconnaissance vocale, de reconnaissance de concepts, de mots-clés, d'élaboration d'un contexte sémantique de dialogue. L'objectif de ces capteurs est de traiter l'information brute reçue par le système et d'élaborer un état de dialogue permettant une réponse efficace.

La modélisation du dialogue par le système représente l'idée que le système se fait de l'état du dialogue (*S : state of dialogue* – state of mind)). L'état de dialogue, élaboré par la succession de modèles ci-dessus, regroupe toute l'information utile pour choisir une réponse appropriée.

Le choix de l'action générée est alors déterminé par la stratégie de dialogue.

### 1.2. Un modèle simple : motifs + règles

Un système primitif illustrant ce paradigme est le suivant :

On se place dans le cadre d'une interaction homme-machine munie d'un module de reconnaissance vocale (ASR : automatic speech recognition). L'utilisateur appelle le standard téléphonique et cherche à joindre un correspondant. La machine transmet l'appel au bon numéro ou prend un message.

Analysons l'échange suivant :

- SYST : Allô ? ( $a_{t-1}$ )

- USER : Est-ce que Dupont est disponible ? ( $o_t$ )
- SYST : transfère l'appel à Dupont. ( $a_t$ )

**Capteurs :** Un capteur est constitué d'un modèle de compréhension ( $m_i$ ) appliqué au texte transcrit à partir du signal acoustique de l'utilisateur. Un exemple simple de modèle est d'utiliser un motif représentant un modèle booléen :

	Nom	Type	Définition
Ex :	$m_1$	Booléen	"Je voudrais parler à X"
	$m_2$	Booléen	"X est disponible ?"
	$m_3$	Booléen	"X"
	$m_4$	Booléen	"un message pour X"

Si, à la date  $t$ , on observe  $o = \langle \dots \rangle$ , alors les motifs  $m_2$  et  $m_3$  sont instanciés. C'est-à-dire que les modèles  $m_2$  et  $m_3$  fournissent la valeur VRAI et  $m_1$  et  $m_4$  la valeur FAUX.

**État de dialogue :** À l'intérieur de la machine sont agrégés les stimuli analysés par les différents modèles afin de synthétiser une pensée (état de dialogue). La forme la plus appropriée de cet état de dialogue est celle qui permet de choisir une action.

Ex :  $M$  = premier motif instancié.

Dans le dialogue de standard téléphonique illustrant cet article, à la date  $t$ ,  $M = (m_2)$

Une définition plus complexe de l'état de dialogue est la suivante : l'état de dialogue  $s_t$  représente l'objectif exprimé par l'utilisateur, ainsi qu'une partie de l'historique du dialogue à la date  $t$ . Il inclut toute l'information disponible dans notre modèle pour que la machine prenne la décision la plus appropriée. À travers un ensemble de capteurs, l'état de dialogue est défini par une liste d'entrées (*slot*) remplie par des valeurs (*value*). Le tableau 1 montre un exemple d'état obtenu lorsque l'utilisateur a exprimé son objectif concernant la personne à contacter (*slot* Nom) mais n'a pas précisé le moyen de communication souhaité (*slot* Média).

Slot	Value	Valeurs possibles
Nom	Dupont	noms de l'annuaire, vide
Média	vide	téléphone fixe, boîte vocale, vide

**Table 1:** Un exemple d'état de dialogue de type *slot-value*

**Stratégie de dialogue :** La génération de l'action de la machine  $a$  s'effectue selon des règles définies par une fonction stratégie :  $a = \pi(M)$ .

Une stratégie peut-être définie par exemple par :

$$\pi(m_1) = \pi(m_2) = \pi(m_3) = a^{(1)}$$

$$\pi(m_4) = a^{(2)}$$

avec

$a^{(1)}$  = transfère l'appel à la personne mentionnée dans le motif.

$a^{(2)}$  = transfère l'appel à la boîte vocale de la personne mentionnée dans le motif.

### 1.3. Gestionnaires de dialogue

Les systèmes de **dialogue homme-machine** développés actuellement par les industries reposent essentiellement sur les standards Voice XML défini par le W3C ([www.w3.org/TR/voicexml20/](http://www.w3.org/TR/voicexml20/)). Le paradigme retenu est celui de machine à états finis utilisant des modèles de langage à grammaire non contextuelle réalisée par des experts. Ceci permet d'assurer au concepteur de garder le contrôle du déroulement du dialogue.

Un inconvénient de ces systèmes est leur rigidité, car ils ne permettent pas à l'utilisateur de s'exprimer naturellement puisque le cours du dialogue est préétabli. C'est pourquoi la recherche actuelle se tourne vers des modèles fondés sur une modélisation statistique de l'état du dialogue.

## 2. MODÉLISATION STATISTIQUE

### 2.1. Modèle de processus de décision de Markov (MDP)

Une introduction de référence sur les **processus de décision de Markov** (Markov Decision Process, MDP) est présentée dans [Sut98].

Afin de se placer dans le cadre théorique des MDPs pour la gestion du dialogue homme-machine, on définit en plus des états de dialogue  $s_t$  et des actions  $a_t$ , la chaîne de Markov de transition d'état conditionnée par les  $a_t$  et une fonction de récompense  $R$  de la façon suivante.

L'action  $a_t$  est déterminée par la machine, c'est la réponse du système à la date  $t$ .

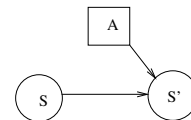
**Récompense** La récompense immédiate reçue au temps  $t$  est  $r_t$  elle est définie grâce à la fonction de récompense immédiate (*reward*)  $R$  par  $r_t = R(s_t, a_t, s_{t-1})$ . On fixe généralement  $R = -1$  sauf pour les états de dialogues terminant le dialogue par un succès ou un échec :

- $\forall s, R(s, a, s') = +20$  lorsque  $a$  et  $s'$  correspondent.
- $\forall s, R(s, a, s') = -20$  lorsque  $a$  et  $s'$  ne correspondent pas.

**Hypothèse de Markov** L'hypothèse principale du modèle MDP est de supposer la propriété de Markov suivante sur les distributions conjointes des  $s_t$  et  $a_t$  :

$$P(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, \dots) = P(s_{t+1}|s_t, a_t) \quad (1)$$

Ce modèle est représenté par le réseau Bayésien suivant (voir [Jen07] pour une introduction aux réseaux Bayésiens) :



**Stratégie optimale** Une **stratégie** est une fonction  $\pi$  qui à chaque état associe une action :  $\pi(s_t) = a_t$ .

La **valeur  $V^\pi$  d'une stratégie** est définie récursivement par :

$$V^\pi(s) = E(r_t + \gamma \cdot V^\pi(s_{t+1}) | s_t = s, a_t = \pi(s)) \quad (2)$$

où  $\gamma$  est le facteur d'escompte (*discount factor*).

$Q$  est l'espérance de la valeur d'une action  $a$  dans un état  $s$  :

$$Q^\pi(s, a) = E(r_t + \gamma \cdot V^\pi(s_{t+1}) | s_t = s, a_t = a) \quad (3)$$

On peut également définir matriciellement  $Q$  par :

$$Q^\pi = (I - \gamma T \Pi^\pi)^{-1} R \quad (4)$$

avec  $T$  et  $\Pi^\pi$  matrices de transition et de stratégie définies de façon appropriée. Voir par exemple, [Man04] pour une description matricielle des MDP.

Une **stratégie optimale** est celle qui maximise  $V$  :

$$V^{\pi^*}(s) = \arg \max_{\pi} (V^\pi(s)) \quad (5)$$

La solution optimale dépend fortement de la fonction récompense choisie, de l'ensemble des actions possibles et du modèle de transition d'états. Dans le cas du standard téléphonique, on obtient un système qui pose des questions tant que l'état de dialogue n'est pas suffisamment rempli pour transférer l'appel.

**Limitations** L'hypothèse de Markov est très forte et n'est généralement pas vérifiée. En effet, elle implique que l'état du dialogue reste le même lorsque l'utilisateur répète plusieurs fois sa requête. Or, il est clair que l'action optimale doit évoluer vers une solution de repli en cas de dialogue ne fonctionnant pas bien ("Je ne vous comprend pas, je vous passe un opérateur", ou plus simplement "au revoir" et raccrocher). Une solution consiste à ajouter un *slot* "patience" dans l'état de dialogue qui compte le nombre de répétition de l'utilisateur pour donner au système l'information nécessaire au choix de l'action de rechange.

De plus, le module de reconnaissance vocale (*Automatic Speech Recognition* ASR) ne fournit pas la transcription exacte de ce qu'a dit l'utilisateur mais un ensemble d'hypothèses probabilisées (treillis ou *lattice*). Les systèmes de dialogue MDP peuvent être très sensibles aux erreurs de l'ASR. La section suivante présente un modèle statistique permettant de prendre en compte cette incertitude.

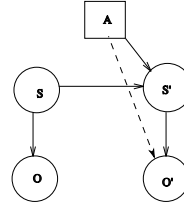
## 2.2. Modèle de MDP partiellement observables (POMDP)

Les POMDP (Partially Observable MDP) ont donné lieu à une abondante littérature. Une introduction est présentée dans [Son71]. Diverses méthodes ont été proposées pour leur résolution (c.à.d. trouver la stratégie  $\pi$  optimale ou l'approximer)[Bon02].

**Modèle statistique** Dans le cadre des MDP, la valeur de l'état de dialogue courant  $s_t$  est supposée connue par la machine. Ce qui est loin d'être le cas lors d'un dialogue où les problèmes d'incompréhension sont légion. L'utilisation de POMDP vise à modéliser l'incertitude sur l'état de dialogue. Différencier l'observation  $o_t$  de l'état réel du dialogue  $s_t$  inconnu permet de modéliser l'interaction de façon plus réaliste et de gagner en robustesse.

La chaîne de Markov (observable) du MDP est remplacée dans un POMDP par une chaîne de Markov cachée (*Hidden Markov Model* HMM), également conditionnée par l'action du système  $a_t$ .

Le réseau bayésien correspondant est présenté figure 1.



**Figure 1:** Graphe Bayésien d'un POMDP. La dépendance directe entre  $a$  et  $o'$  (en pointillés) est parfois omise.

Ce qui signifie que la loi de  $o_t$  sachant l'historique complet du dialogue ne dépend que de l'état courant et de la dernière action du système ; et que celle de  $s_t$  ne dépend que de l'état précédent et de l'action. La dépendance entre  $a$  et  $o$  est généralement omise car on inclut dans  $s'$  l'information de  $a$  pertinente pour  $o'$ .

$$P(o_t | s_t, s_{t-1}, o_{t-1}, a_{t-1}, \dots) = P(o_t | a_t, s_t) \quad (6)$$

$$P(s_{t+1} | s_t, a_t, s_{t-1}, o_{t-1}, \dots) = P(s_{t+1} | s_t, a_t) \quad (7)$$

**Vecteur de croyance  $b_t$**  Dans un POMDP, une stratégie n'est plus une fonction de l'état  $a_t = \pi(s_t)$  qui ne peut être connue par le système. En revanche, on définit  $\pi$  sur un vecteur de croyance  $b(s)$  (*belief*).

Le **vecteur de croyance**, défini par l'équation 8, est une distribution sur l'espace des états qui représente la probabilité d'être effectivement dans un état donné sachant tout l'historique observé du dialogue.

$$b_t(s) = P(s_t | o_t, a_t, o_{t-1}, \dots) \quad (8)$$

Il peut être mis à jour à chaque étape en fonction de la valeur observée  $o_t$  :  $a_t = \pi(b_t)$  selon l'équation 9 :

$$b_{t+1}(s') = \alpha P(o' | s', a) \sum_s P(s' | s, a) b(s) \quad (9)$$

où  $\alpha$  est une constante de normalisation.

Le vecteur  $b_t$ , en tant que statistique de l'historique observé ( $o_t, a_t, o_{t-1}, \dots$ ), est **suffisant** pour l'état courant de dialogue  $s_t$ .

D'autres représentations (*Predictive State Representations* PSR) ont été proposées (voir par exemple [Sin04] ou [Jae98]) pour généraliser la notion de vecteur de croyance.

L'écriture matricielle des POMDP s'enrichit d'une matrice d'observation  $O$  :

$$Q^\pi = (I - \gamma O \cdot T \Pi^\pi)^{-1} R \quad (10)$$

Une présentation des stratégies des POMDP vues comme des contrôleurs à états finis est exposée dans [Far08]

**Limitations** Si les POMDPs permettent effectivement de tenir compte de l'incertitude due aux erreurs de compréhension, leur principal inconvénient est de ne pouvoir être résolus lorsque le nombre d'états est trop grand.

L'état de dialogue doit être défini de façon à fournir un maximum d'information avec un nombre restreint d'états

possibles. De même, le nombre d'observations et d'actions doit être aussi petit que possible. Le passage par une représentation sémantique structurée permet de telles définitions.

### 3. REPRÉSENTATION SÉMANTIQUE STRUCTURÉE

Cette section présente le module de compréhension permettant de transformer un enregistrement acoustique de dialogue en transcription textuelle, puis de l'annoter en concepts, puis en *frame* au contenu sémantique plus élaboré. Cette annotation en *frame* est finalement utilisée pour définir des suites d'états  $s_t$ , observations  $o_t$  et actions  $a_t$ .

Une fois appliquées sur un corpus, ces annotations permettent également d'estimer la dynamique du système : les probabilités de transition  $T = P(s'|s, a)$  et d'observation  $Z = P(o'|s', a)$ .

#### 3.1. Corpus MEDIA

La campagne d'évaluation EVALDA [BM05] a fourni un corpus de dialogues de demandes d'informations touristiques. Le corpus MEDIA est constitué de 1257 dialogues (18831 tours de parole utilisateur) enregistrés par la technique du magicien d'Oz (l'utilisateur interagissant avec un humain simulant la machine). Le tableau 2 présente un extrait du corpus :

woz :	...souhaitez vous écouter le guide
spk :	Non
woz :	Quelle information voulez-vous obtenir
spk :	Je voudrais réserver un hôtel
woz :	Donnez le nom de la ville où vous souhaitez aller
spk :	Je voudrais aller à Évreux
woz :	À Dreux à quelle période souhaitez vous y séjourner
spk :	Euh c'est pas Dreux c'est Évreux
woz :	à Évreux à quelle période souhaitez vous y séjourner

**Table 2:** Extrait du corpus MEDIA (transcriptions texte exactes).

Les dialogues ont été transcrits en texte de deux façons. Les transcriptions manuelles sont considérées comme "exactes" tandis que les transcriptions ASR (Automatic Speech Recognizer) sont des observations bruitées.

#### 3.2. Annotation en concepts

Un module de compréhension (Speech Language Understanding SLU) a été appliqué aux transcriptions pour générer les annotations en concepts.

Le dictionnaire sémantique contient 83 concepts. Une étiquette *mode* pouvant prendre trois valeurs différentes (positif/négatif/optionnel) est ajoutée aux concepts. L'annotation contient également la valeur prise par le concept.

Un extrait d'annotation en concept est présenté table 3 pour la phrase utilisateur "euh c' est pas Dreux c' est Évreux"

Le module de compréhension dans lequel est réalisé l'an-

transcription	concept	mode	valeur
euh c'est	null		
pas Dreux	localisation-ville	-	Dreux
c'est	null		
Évreux	localisation-ville	-	Évreux

**Table 3:** Extrait du corpus MEDIA : annotation en concepts

notation en concepts utilise un réseau Bayésien dynamique (DBN) (voir [Meu09] pour une présentation plus précise du modèle)

#### 3.3. Annotation en frames

Lors de l'annotation en *frames*, l'ensemble des concepts présents dans le tour de dialogue est structuré en *frames* sémantiques.

texte	Je voudrais euh un hôtel plutôt euh vers Nice euh oui un hôtel à Nice
concepts	vouloir, null, hôtel, null, Nice, null, accept, hôtel, Nice
frames	F1 : HOTEL(ville = 'Nice' , nom = ' ' ) F2 : VOULOIR(objet = F1)

**Table 4:** Exemple simplifié d'annotation en frame

Pour réaliser l'annotation du corpus en *frame* à partir des transcriptions texte et des annotations en concepts, on utilise un réseau Bayésien dynamique (DBN) ainsi que des règles expertes définies manuellement (voir [Meu08] pour le processus détaillé).

## 4. DÉFINITION DU MODÈLE POMDP

#### 4.1. Définition des états de dialogue

À la date  $t$ , un état de dialogue est défini d'après l'historique du dialogue  $h_t$ , de manière déterministe :  $s_t = M(h_t)$ . L'application de la fonction  $M$  implique une très forte réduction de la quantité d'information, éliminant toute information présente dans  $h_t$  qui soit inutile pour la sélection de l'action.

L'élaboration de la fonction  $M$  est réalisé par une succession de capteurs (*features*) permettant d'analyser les perceptions brutes de la machine. En effet, l'historique du dialogue contient toute l'information du dialogue passé jusqu'au temps  $t$  : les annotations en frames de l'utilisateur, les actions de la machine, et les états précédents. Le contenu de l'historique constitue les *features*  $F_i$  de premier niveau. Des *features* additionnels  $F$  de niveaux supérieurs sont définis à partir des premiers  $F_i$  :

- Fonction prédéfinie :  $F = \text{card}\{F_i\}$ ,
- Combinaison linéaire :  $F = \sum_i \theta_i.F_i$ ,
- Maximisation inverse :  $F = \arg \max_i (F_i)$ ,
- Clause logique, par exemple :  $F = (F_1 \vee F_2) \wedge F_3$ .

Les *features* peuvent être de nature diverse : entiers, booléens, réels, vecteurs, ensemble de *frames*...

L'ensemble des valeurs de  $h_t$  est virtuellement de taille infinie car  $h_t$  contient des graphes de *frames* de taille non fixée. Le tableau 5 donne quelques exemples de *features*. Dans le corpus MEDIA utilisé,  $h_t$  contient en moyenne 49 frames ( $\sigma \simeq 31$ ). Par exemple la *frame* Person peut apparaître jusqu'à 34 fois dans certains dialogues.

<i>features</i>	Définition
$F_{Query,u}$	$\text{card}\{\text{frame } Query \in u_t\}$
$F_{Hotel,u}^{\neq \emptyset}$	$\text{card}\{\text{non-empty frame } Hotel \in u_t\}$
$F_{Query}$	$\theta_1 F_{Query,u} + \theta_2 F_{Request,u} + \dots + \theta_7 F_{Number,u}$ ( $\theta_1 = +5, \theta_2 = -3, \dots, \theta_7 = +1$ )
$F_{RecentS3}$	$(\tilde{s}_{t-1} = 3) \vee (\tilde{s}_{t-2} = 3)$
...	

**Table 5:** Exemples de *features*.  $u_t$  (resp.  $w_t$ ) est l'ensemble des *frames* dans le  $t$ -ème tour de parole utilisateur (resp. système).

Un ensemble de *features* secondaires est ainsi défini jusqu'à un *feature* particulier  $F = \tilde{s}_t$  appelé état résumé.  $\tilde{s}_t$  est à valeur dans  $\{1, \dots, 18\}$ . L'optimisation et l'évaluation du modèle POMDP seront effectuées sur l'espace des états résumés (c.à.d. en considérant des suites de  $\tilde{s}_t$ ).

#### 4.2. Définition des observations

Dans notre contexte, nous considérons que les observations sont des versions bruitées des états. Les bruits considérés sont l'ASR (module de reconnaissance vocale) et le SLU (module d'annotation en concepts et en *frames*). La définition de *features* d'observation et d'observations résumées  $\tilde{o}_t$  peut être réalisée de la même façon que pour  $s$ .

Deux niveaux de bruit ont été considérés :

- ASR + SLU : le signal acoustique automatiquement transcrit (ASR) puis automatiquement annoté (SLU).
- SLU : le signal acoustique transcrit exactement (humain) puis automatiquement annoté (SLU).

Par contraste, l'état est défini à partir des données exactes :

- $s$  : Le signal acoustique transcrit exactement (humain) puis annoté exactement (humain).

#### 4.3. Définition des actions

Les actions de la machine sont définies également de manière semi-automatique en utilisant des règles expertes appliquées aux *frames* du dialogue. Le protocole du Magicien d'Oz (*Wizard of Oz* : woz) utilisé lors de la création du corpus donne lieu à des phrases stéréotypées pour les tours de parole de la machine. Les actions du magicien d'Oz sont donc définies avec beaucoup plus de précision que les états et les observations. On a donc défini les actions du woz directement à partir des mots de la transcription exacte, sans utiliser de *frame*.

On se place également pour les actions dans le cadre d'un espace d'actions résumées. On considère par exemple l'action résumée "demander une information" (*request*) au lieu de l'action complète "demander le nom de la ville" (*request(city - name)*).

La théorie des actes de parole a été développée en linguistique [Aus62] [Sea69] et appliquée pour établir la taxonomie DAMSL [Cor97] dont s'inspirent les définitions utilisées dans notre modélisation. Suivant [Wil07], on fait l'hypothèse que les actions résumées peuvent être correctement déterminées par les informations de l'état résumé.

#### 4.4. Estimations des lois de probabilités du modèle

Les probabilités conditionnelles de transition  $T = P(s'|s, a)$  et d'observation  $Z = P(o|s, a)$  du modèle POMDP sont estimées à l'aide d'un modèle de bigrammes (avec repli) sur les données d'entraînement du corpus MEDIA. L'implémentation a été effectuée par l'outil SRILM [Bil03].

### 5. EXPÉRIENCES

#### 5.1. Optimisation

Les stratégies sous-optimales approchant la solution optimale du POMDP ont été obtenues en utilisant *pomdp-solve* [Meu99] en limitant l'exploration à un horizon fini (2 à 10).

#### 5.2. Évaluation de la stratégie

Pour évaluer une stratégie  $\pi$ , solution optimale d'un POMDP, on simule 1000 dialogues obtenus en suivant la stratégie  $\pi$  avec la même dynamique ( $T$  et  $Z$ ) qui a servi à optimiser  $\pi$ .

L'évaluation objective des systèmes de dialogue (et des systèmes de TALN (Traitement automatique de la Langue Naturelle)) est loin d'être parfaite. En effet, lors des simulations de dialogues, on simule un utilisateur par un tirage aléatoire de  $\tilde{s}$  et  $\tilde{o}$  suivant les lois  $T$  et  $Z$ . Ainsi, les performances du modèle sont évaluées en réutilisant les paramètres de ce même modèle (ici  $Z$  et  $T$ ). La conception d'utilisateurs simulés plus complexes [Sch07] permet d'évaluer plus précisément les stratégies obtenues, même si le principe d'auto-évaluation reste le même.

Une solution pourrait venir des industriels, ceux-ci disposant de systèmes en situation interagissant avec des utilisateurs réels.

#### 5.3. Résultats

Afin d'estimer l'impact de l'erreur faite par le module de reconnaissance en concepts (SLU) par rapport à l'erreur totale (ASR+SLU), on compare les deux systèmes de dialogue SLU et ASR+SLU.

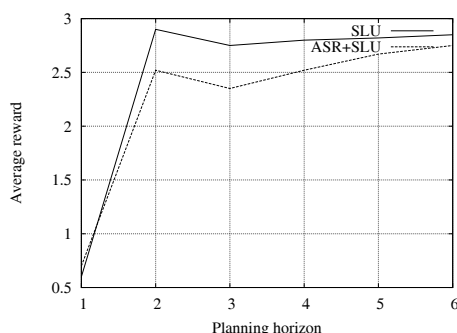
La table 6 montre les récompenses obtenues en moyenne par ces différents systèmes. La table 7 détaille la récompense moyenne selon le temps d'apprentissage.

	Moyenne	Écart-type
Récompense ASR+SLU	2.71	4.48
Récompense ref+SLU	2.88	4.49

**Table 6:** Récompenses obtenues (horizon = 6)

On constate que la différence entre les systèmes SLU et ASR+SLU est faible, ce qui montre la robustesse du système relativement aux erreurs de l'ASR. Cette robustesse peut être en grande partie attribuée à l'utilisation d'espaces résumés. En effet, dans certains cas, l'observation approximant l'état ainsi que l'état exact sont projetés sur un même état résumé. Alors, l'erreur ASR n'entraîne aucun bruit dans le modèle résumé.

On constate que les performances du système n'augmentent que très légèrement après un horizon de 2, ce qui



**Table 7:** Récompense moyenne obtenue par la stratégie optimisée à un horizon de 1 à 6

signifie que la planification à plus long terme a peu d'impact sur la récompense obtenue. Ceci pourrait être dû à une définition trop grossière des états de dialogue résumés qui restreint la recherche de la stratégie optimale aux stratégies reposant sur l'information contenue dans l'état résumé, oubliant une partie de l'information disponible dans l'historique de dialogue.

## 6. CONCLUSION

Dans cet article, l'incertitude de la machine sur l'état réel du dialogue est modélisée statistiquement par un POMDP. De plus, l'utilisation de structures sémantiques riches pour modéliser l'état de dialogue présentée dans cet article permet de définir des états de dialogue correspondant mieux à la réalité que de simples motifs ou un modèle de *slot - value*. La complexité calculatoire induite peut être atténuée par l'utilisation d'états résumés macroscopiques.

Les simulations de dialogues ont permis d'évaluer les stratégies optimisées sur les POMDP, ce qui a montré que les erreurs de compréhension (en mots et concepts) ont peu d'impact sur les performances du système en terme de récompense (*reward*). Un prototype est en cours de construction et des tests avec des utilisateurs réels seront bientôt disponibles permettant une évaluation approfondie de l'approche proposée.

## RÉFÉRENCES

[Aus62] Austin, J. (1962), "How to do thing with words", *Oxford Clarendon Press*.

[Bil03] Bilmes, J. and Kirchhoff, K. (2003), "Factored language models and generalized parallel backoff", *Proceedings of HLT/NACCL*, pp. 4–6.

[BM05] Bonneau-Maynard, H., Rosset, S., Ayache, C., Kuhn, A. and Mostefa, D. (2005), "Semantic Annotation of the French Media Dialog Corpus", *Ninth European Conference on Speech Communication and Technology*.

[Bon02] Bonet, B. (2002), "An  $\epsilon$ -optimal grid-based algorithm for partially observable Markov decision processes", *Proceedings of the Nineteenth International Conference on Machine Learning (ICML)*.

[Cor97] Core, M., Allen, J. and Traum, D. (1997), "Coding Dialogues with the DAMSL Annotation Scheme", *Working Notes :AAAI Fall Symposium*

*on Communicative Action in Humans and Machines*, pp. 28–35.

[Far08] Fard, M.M., Pineau, J. and Sun, P. (2008), "A variance analysis for pomdp policy evaluation", *AAAI Press*.

[Jae98] Jaeger, H. (1998), "Discrete-time, discrete-valued observable operator models : A tutorial", Tech. rep.

[Jen07] Jensen, F. and Nielsen, T. (2007), *Bayesian Networks and Decision Graphs*, Springer.

[Man04] Mannor, S., Simester, D., Sun, P. and Tsitsiklis, J. (2004), "Bias and variance in value function estimation", *ACM International Conference Proceeding Series ACM Press New York, NY, USA*.

[Meu99] Meuleau, N., Kim, K., Kaelbling, L. and Cassandra, A. (1999), "Solving POMDPs by searching the space of finite policies", *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, pp. 417–426.

[Meu08] Meurs, M.J., Lefèvre, F. and Mori, R.D. (2008), "A bayesian approach to semantic composition for spoken language interpretation", in *ISCA Interspeech*, Brisbane, Australia.

[Meu09] Meurs, M., Lefèvre, F. and De Mori, R. (2009), "Spoken language interpretation : On the use of dynamic bayesian networks for semantic composition", in *ICASSP*.

[Sch07] Schatzmann, J., Thomson, B., Weilhammer, K., Ye, H. and Young, S. (2007), "Agenda-based user simulation for bootstrapping a POMDP dialogue system", *ACL*.

[Sea69] Searle, J. (1969), *Speech Acts : An Essay in the Philosophy of Language*, Cambridge University Press.

[Sin04] Singh, S., James, M.R. and Rudary, M.R. (2004), "Predictive state representations : a new theory for modeling dynamical systems", in *AUAI '04 : Proceedings of the 20th conference on Uncertainty in artificial intelligence*, Arlington, Virginia, United States : AUAI Press, pp. 512–519.

[Son71] Sondik, E. (1971), "The optimal control of partially observable decision processes", *Ph. D. thesis, Stanford University, Stanford, California, USA*.

[Sut98] Sutton, R. and Barto, A. (1998), *Reinforcement Learning : An Introduction*, Cambridge, MA : MIT Press.

[Wil07] Williams and Young (2007), "Scaling pomdps for spoken dialog management", *IEEE Audio, Speech and Language Processing*.