Approche non supervisée pour la gestion stochastique du dialogue

Fabrice Lefèvre and Renato de Mori

LIA - Université d'Avignon
{fabrice.lefevre, renato.demori}@univ-avignon.fr

ABSTRACT

Following recent studies in stochastic dialog management, an unsupervised approach is proposed to reduce the cost and complexity of a probabilistic POMDP-based dialog manager setup. A first decoding step derives semantic basic constituents from user utterances. Then the isolated units along with some relevant context features (previous system actions, previous user utterances...) are combined in vectors representing the current dialog situation. A clustering step is performed with these vectors and each partition of the derived space represents a particular dialog state. Any new utterance can be classified according to these automatic states and the system belief can be updated before the POMDP-based dialog manager can take a decision on the best next action to perform.

The approach is applied to the challenging French MEDIA task (tourist information and hotel booking). The MEDIA training corpus is semantically rich (over 80 basic concepts) and its 10k utterances are segmentally annotated in terms of basic concepts. Some insights on the method effectiveness are obtained by an analysis of the POMDP model training convergence using user simulation.

Keywords: spoken dialog systems, speech understanding, unsupervised clustering, POMDP

1. Introduction

Dans les systèmes de dialogue oral, le processus de décision pour la gestion du dialogue nécessite d'obtenir une interprétation des interactions de l'utilisateur. Les interprétations automatiques représentent les objets du domaine au travers, par exemple, d'instances de cadres sémantiques. Les modules de compréhension automatique de la Parole et de gestion du dialogue construisent progressivement les instances de cadres qui représentent la croyance du système dans les intentions de l'utilisateur avec l'objectif de satisfaire sa requête dès lors qu'elle sera suffisamment spécifiée pour permettre une action appropriée du système.

Un cadre sémantique est un modèle pour représenter les entités sémantiques et leurs propriétés. Trouver les valeurs des propriétés revient à remplir les champs du cadre. Pour atteindre cet objectif, le système dispose d'actions . Le choix de l'action à réaliser pour un tour particulier du dialogue fait partie de la stratégie du dialogue qui peut être établie suivant différentes approches.

Récemment, il a été proposé de considérer les cadres instanciés et les actions comme appartenant à un espace de représentation principal (*master space*) et d'introduire un espace réduit (*summary space*) dans lequel sont estimées les actions simplifiées à exécuter. Il est alors possible de concevoir des stratégies de dialogue avec un critère d'optimalité dans le cadre des processus de décision markoviens partiellement observables (POMDP).

Dans ce contexte, à un moment particulier, le système dispose d'une probabilité pour chaque état possible du dialogue. C'est son vecteur de croyance (*belief*). L'action à exécuter dans ce tour est décidée à partir d'un critère permettant d'obtenir une récompense (*reward*) maximale sur l'ensemble du dialogue. Et pas seulement la meilleure récompense immédiate, il y a donc planification. L'approche est utilisable uniquement si le nombre d'états possibles du dialogue est petit. En pratique, un état du dialogue doit être considéré pour toute combinaison possible des variables logiques utilisées dans le modèle de connaissance du système. Aussi pour pouvoir appliquer l'approche POMDP, les états du dialogue de l'espace principal doivent être regroupés dans des états plus généraux, formant l'espace réduit. Nous proposons de réaliser ce regroupement automatiquement et sans supervision afin d'en limiter le coût.

Le corpus MEDIA, utilisé dans nos travaux, contient des dialogues assez complexes, incluant des dizaines de tours de parole chacun et plus de 80 constituants sémantiques. Le corpus a été annoté sémantiquement. Une hypothèse raisonnable est que la progression du dialogue est inscrite dans l'espace principal des états, tandis que les actions liées aux propriétés et modalités du remplissage des champs d'information, comme les corrections ou les demandes de confirmation appartiennent à l'espace réduit. La conjoncture analysée et évaluée dans ce papier est que les actions réduites peuvent être déterminées exclusivement à partir de l'analyse des instances des constituants de base, sans s'attacher aux détails de la composition sémantique globale. Ainsi les états de l'espace réduit sont obtenus automatiquement en considérant que le regroupement des états dans l'espace principal est uniquement influencé par les constituants sémantiques de base. De cette façon les annotations peuvent être directement exploitées afin d'obtenir des modèles robustes.

La première section de ce papier présente le corpus MEDIA. Ensuite la section 3 rappelle les principes du décodage des constituants sémantiques de base. Puis une présentation rapide des POMDP appliqués au problème de la gestion du dialogue est donnée section 4, suivie par la description de notre proposition pour la construction d'états du dialogue automatique et non-supervisée dans la section 5. La dernière section présente les expériences sur le corpus MEDIA.

2. Corpus MEDIA

Le corpus MEDIA [1] a été obtenu par le biais du protocole du Magicien d'Oz (WOZ) simulant un serveur téléphonique d'information touristique et de réservation d'hôtels. Huit catégories de scénarii de complexités variables ont été définies. Le corpus contient 1257 dialogues, de 250 locuteurs pour un total d'environ 70 heures de parole. Les données sont intégralement transcrites et annotées manuellement.

Un segment sémantique est composé par un triplet : modalité (+, -, ? et ~ pour *optionnel* comme dans "avec une chambre si possible"), nom de l'attribut et valeur normalisée (issue d'un dictionnaire ou d'une transformation régulière). La modalité est attribué à chaque segment, ce qui lève l'ambiguïté de phrases comme "pas à Paris à Nancy". Seuls les tours de parole de l'utilisateur sont annotés.

Le dictionnaire sémantique contient 83 attributs de base et 19 spécifieurs. Les relations et compositions entre segments ne sont pas disponibles. Les combinaisons attribut-spécifieur (1121 attributs potentiels) permettent de représenter les relations entre les segments principaux. Par exemple, l'attribut `payment-montant-entier-chambre` est obtenu par combinaison de l'attribut `payement-montant-entier` avec le spécifieur `chambre`. Cette annotation est comparable à un *shallow parsing* et la représentation hiérarchique de la requête peut être retrouvée à partir de sa représentation séquentielle [1]. Un total de 144 attributs distincts apparaît dans le corpus d'apprentissage (avec 2,2k valeurs normalisées distinctes).

3. Décodage conceptuel et constituants sémantiques

La compréhension de la parole est une tâche très complexe, y compris en comparaison avec la compréhension de la langue naturelle car, dans la parole, le sens est mélangé avec les autres informations comme l'identité du locuteur ou encore l'environnement sonore. Les phrases à traiter sont difficiles à analyser et les messages oraux souvent agrammaticaux. La connaissance mise en oeuvre est imparfaite et la transcription des messages oraux en terme d'hypothèses de mots est obtenue par un système automatique qui commet des erreurs. Deux types de constituants sont utilisés dans nos travaux : les actes de dialogue et les concepts.

Actes de dialogue
Un étiquetage semi-manuel des actes de dialogue (équivalent ici aux actions de la terminologie des POMDP) des tours de parole du WOZ a été réalisé à partir de règles. Les règles ont été corrigées et adaptées manuellement. Les interactions du WOZ sont complexes et ne peuvent pas toujours facilement être catégorisées par un unique acte de dialogue. Par exemple, les requêtes à la base de données sont très souvent annoncées par le WOZ conjointement avec une demande de confirmation : "Vous voulez un hôtel à Avignon en juin. Un instant, je recherche votre information.". Lorsque ces combinaisons apparaissent souvent, un acte composite est utilisé dans l'annotation. Sur 15 actes de dialogue retenus, 10 sont élémentaires et 5 composites.

Concepts
Les concepts sont liés au sens littéral de la phrase. Les techniques stochastiques pour le décodage conceptuel sont maintenant très répandues comme alternative efficace aux méthodes à base de règles. Elles sont séquentielles : i.e. elle trouve la séquence de concepts qui à la plus grande probabilité d'avoir été générée par la séquence de mots considérée. Le formalisme des réseaux bayesiens dynamiques (DBN) a été utilisé pour permettre le développement de systèmes de compréhension probabilistes multi-niveaux [2]. Dans ce travail, nous utilisons un décodage conceptuel conforme à [2] : une première étape de décodage avec les DBN produit une séquence d'attributs/valeurs puis 2 étapes de classification avec les CRF enrichissent cette séquence avec les modalités et les spécifieurs. Le taux d'erreur en concepts est d'environ 26% à partir des transcriptions manuelles.

4. Gestion du dialogue et POMDP

Nous rappelons brièvement les fondements de l'utilisation des processus de décision markovien partiellement observables pour la gestion du dialogue (plus de détails dans [3, 4, 5]). Le principe est que, lors de chaque tour de parole, le module de gestion met à jour sa croyance dans la situation du dialogue qui est représentée par une distribution de probabilité sur les états possibles du dialogue. Le choix de l'action suivante dépend alors de ce vecteur de croyance plutôt que d'un état en particulier. Les probabilités pour le système d'être dans un état au prochain échange dépendra alors de la croyance actuelle, de l'action choisie et de la réaction de l'utilisateur à cette action.

Pour l'état i lors de l'échange k, la probabilité $p_k(s_i)$ d'être dans l'état s_i est :

$$p_k(s_i) = f(a_{k-1}, b_{k-1}) \tag{1}$$

La probabilité $p_k(s_i)$ est une fonction de la croyance b_{k-1} et de l'action a_{k-1} au tour $k-1$.

L'action du système la plus adaptée est sélectionnée en suivant une stratégie optimale. L'apprentissage par renforcement [6] offre une famille de méthodes pour sélectionner les actions optimales dans un environnement défini par un processus de décision markovien (MDP). Une discussion plus avancée sur l'utilisation des MDP pour les gestionnaire de dialogue est disponible dans [3]. La stratégie de dialogue π sélectionne une action système, basée sur l'état actuel du dialogue :

$$a_k = \pi(s_k) \tag{2}$$

Les transitions entre états sont caractérisées par une distribution de probabilités :

$$p(s_k|s_{k-1}, a_{k-1}) \tag{3}$$

La combinaison de l'action choisie et de l'état au tour k détermine la récompense immédiate :

$$r_k = R(s_{k-1}, a_{k-1}, s_k) \tag{4}$$

où R est une fonction tabulée établie manuellement. L'objectif est de trouver une stratégie qui maximise la récompense totale future :

$$R_k = \sum_{\tau=k+1}^{K} r_\tau \tag{5}$$

Toutes les méthodes d'apprentissage par renforcement nécessite de déterminer une fonction de la valeur considérant l'état courant :

$$V_\pi(s) = E_\pi(r_k|s_k = s) \tag{6}$$

Cette fonction peut être décomposée :

$$Q^\pi(s, a) = E_\pi(r_k|s_k = s, a_{k-1} = a) \tag{7}$$

Classifieur	Appr.	Test (#3524)		
	(#14282)	REF	SLU	ASR+SLU
WER	0,0	0,0	0,0	33,5
CER	0,0	0,0	21,3	43,4
1-NN	0,0	8,4	26,3	30,1
20-NN	16,0	16,2	30,0	33,4
SVM	3,9	3,9	23,5	29,2

Tab. 1: Taux d'erreur de classification sur les ensembles d'apprentissage et de test pour 3 classifieurs (1-PPV, 20-PPV et SVM). Conditions contrastives : à partir des annotations de référence (*REF*), décodées des transcriptions manuelles (*SLU*) ou depuis les hypothèses de reconnaissance (*ASR+SLU*).

La stratégie optimale pour l'état s est alors déterminée comme suit :

$$\pi'(s) = \arg\max_a Q^\pi(s,a) \quad (8)$$

Les méthodes permettant de déterminer la stratégie optimale sont proposées dans [6]. Malheureusement, les MDP ne sont pas utilisables en pratique car ils requièrent une connaissance complète de l'état du système, ce qui est irréaliste. De même les buts de l'utilisateur ne peuvent pas être directement observés.

Ainsi un système de dialogue doit baser sa stratégie sur des données incomplètes et les POMDP sont utilisés. Ils sont équivalents aux MDP avec un espace d'état continu : l'état réel du dialogue est remplacé par un vecteur de croyances (une probabilité par état). Des présentations de l'utilisation des POMDP pour la gestion du dialogue sont disponibles dans [4, 5, 7].

Comme les états du dialogue ne sont pas directement connus du système, des observations sont prises en compte et la relation entre les observations et les états sous-jacents est modélisée par une distribution de probabilités :

$$p(o_k|a_{k-1}, s_k) \quad (9)$$

Les probabilités d'observation, combinées aux probabilités de transition d'états, permettent de mettre à jour l'état de croyance (fonction f ci-dessus). L'algorithme de l'itération de la valeur calcule de façon itérative la meilleure valeur pour V, en considérant les actions choisies et fournit ainsi une solution optimale. La difficulté provient du coût de calcul très important de l'algorithme optimal. Des variantes sous-optimales existent, mais elles sont aussi limitées à quelques dizaines d'observations, états et actions. Des exemples d'approches permettant de réduire le nombre d'états ont été présentées dans [8, 7]. La construction automatique des états du dialogue est aussi un moyen de contrôler leur quantité.

5. Regroupement et classification des états du dialogue

Un état du dialogue peut être défini à partir de l'interprétation sémantique de la requête utilisateur, en prenant en compte quelques informations pragmatiques comme le contexte courant du dialogue et les informations déjà réunies lors des interactions précédentes. L'objectif principal de l'interprétation sémantique est de catégoriser complètement chaque échange de l'utilisateur. Aussi les relations entre les unités de base doivent être déterminées de sorte à obtenir le sens global de la phrase.

Il est communément admis que l'interprétation sémantique dans un dialogue homme-machine est le résultat d'une inférence logique. Des modèles logiques du premier ordre (et même probabilistes comme les MLN) peuvent être mis en oeuvre afin de proposer une interprétation structurée basée sur la première étape de décodage conceptuel. Cette approche présente des inconvénients : 1. soit des experts sont nécessaires pour obtenir un ensemble de règles conséquent (l'implémentation initiale des règles est coûteuse et sujette à erreurs, généralement une adaptation des règles devra être réalisée en ligne), 2. soit une annotation de haut niveau (de type FrameNet ou Propbank) doit être réalisée ce qui est aussi coûteux et complexe.

Nous proposons d'obtenir l'interprétation des requêtes de l'utilisateur de manière automatique. Pour cela, un ensemble de paramètres est défini basé sur les concepts de base et quelques autres informations contextuelles (les concepts du tour de parole précédent, les actions précédentes de l'utilisateur et du système). A partir de cette représentation, des regroupements de vecteurs de requête sont construits de façon non-supervisée.

En fait, une légère supervision est introduite dans la façon dont les paramètres sont sélectionnés et combinés. Nous établissons quelques règles pour décider des paramètres :
- la composition doit être fondée uniquement sur les unités conceptuelles et les actes de dialogue ;
- dans la mesure où l'information des constituants de base peut être sur-spécifiée et redondante pour les besoins de gestion de dialogue, les concepts sont projetés sur un sous-ensemble qui regroupe les informations comparables (e.g. `location-ville` et `location-proche` sont groupés dans un unique concept `location`), et aussi les valeurs ne sont pas prises en compte (sauf celles des concepts dialogiques qui peuvent conduire à des interprétation différentes, e.g. `réponse`, `commande-tache`...) ;
- les informations dépendantes du contexte sont nécessaires : le vecteur de requête courant prend en compte le vecteur précédent (en cas de modalité négative, si un concept était présent il est annulé).

Dès que les paramètres ont été sélectionnés, un algorithme classique de regroupement est appliqué sur les vecteurs de l'ensemble d'apprentissage [1]. Du fait de la nature des vecteurs, les approches permettant le traitement de données éparses et (principalement) binaires sont mises en oeuvre (e.g. [9]). Dans la mesure où ils doivent être utilisés dans le contexte des POMDP, le nombre de regroupements ne doit pas dépasser quelques dizaines. Cette étape fournit un étiquetage de référence pour les données d'apprentissage, ensuite chaque nouvelle requête de l'utilisateur pourra être classée selon la référence.

Comme rappelé dans la section 4, deux ensembles de probabilités sont nécessaires pour définir les modèles POMDP : $p(s_k|s_{k-1}, a_{k-1})$ et $p(o_k|a_{k-1}, s_k)$ ($\simeq p(o_k|s_k)$). Les états obtenus automatiquement sont utilisés pour apprendre les transitions entre états à partir de l'ensemble d'apprentissage. Les probabilités d'observation sont apprises par confrontation entre le résultat du regroupement automatique et de la classification. Le manque de données est une difficulté classique lors de la mise au point des modèles POMDP. Afin d'obtenir les estimations pour toutes les distributions de probabilité, les modèles de langage factorisés (FLM) ont été utilisés en combinaison avec la technique du repli parallèle généralisé (GPB) [10].

[1] L'outil WEKA http ://www.cs.waikato.ac.nz/ml/weka a été utilisé pour nos expériences.

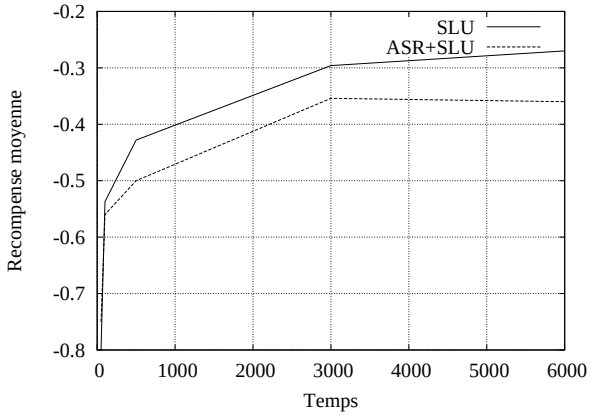

Fig. 1: Convergence de la récompense moyenne lors de l'apprentissage des modèles POMDP en fonction du temps d'apprentissage (en secondes).

6. Expériences

Afin d'évaluer la méthode proposée, un gestionnaire à base de POMDP est créé pour la tâche MEDIA. L'ensemble d'apprentissage comprend 14k requêtes, l'ensemble de test 3k. Le regroupement est réalisé simultanément sur les 2 ensembles de sorte à obtenir aussi une référence pour l'ensemble de test.

Afin d'évaluer la robustesse de l'approche, les données de test sont classées selon 3 variantes : 1. les annotations sémantiques manuelles sont utilisées pour construire les vecteurs de requête, 2. les vecteurs sont obtenus à partir des séquences de concept décodées à partir de la transcription manuelle des requêtes, 3. les séquences de concept proviennent des hypothèses de transcription automatique. Deux classifieurs sont comparés : un classifieur à base d'exemples [9] (une version des k-plus proches voisins, bien adaptée aux données éparses) et un classifieur à noyau SVM [11].

Le Tableau 1 donne les résultats d'identification pour les 2 classifieurs. Le taux d'erreur en mots est 33,5% [12], les taux d'erreur en concepts sont 21,3% sur les données parfaitement transcrites et 43,4% sur la reconnaissance (seulement 2 modalités sont considérées, et aucun spécifieur). Les classes étant automatiques, les performance sont naturellement bonnes quand les vecteurs de test sont obtenus dans les mêmes conditions que les vecteurs d'apprentissage (colonne "Test/REF" dans le Tableau 1). Toutefois, une baisse importante de performance est observée lorsque la classification est réalisée sur le décodage en concepts (entre 15 et 20%) et qui est encore aggravée avec la reconnaissance (entre 3 et 6%).

Les paramètres des POMDP sont appris sur le corpus MEDIA entier. La convergence de l'algorithme de l'itération de la valeur est rapportée dans les courbes de la Figure 1. Une version "par grille" (*grid-based* ou *point-based*) de l'algorithme a été utilisée [13]. Plusieurs valeurs ont été testées pour le nombre de vecteurs de croyance utilisés pour représenter la fonction de valeur. 500 vecteurs apparaît un compromis raisonnable entre performance et temps d'apprentissage. Les récompenses sont moyennées sur les récompenses immédiates accumulées durant 10k dialogues simulés (en utilisant les distributions de probabilité des POMDP comme modèle d'utilisateur). La nature non-supervisée de la définition des états de dialogue ne permet pas d'affecter des récompenses à des états particuliers. Pour contourner ce problème, les états "début" et "fin" du dialogue sont ajoutés. Les dialogue simulés partent obligatoirement de l'état début et sont interrompus lorsque que l'état fin est rencontré ou après un nombre maximal d'échanges (200).

Deux ensembles de paramètres sont testés pour la définition des POMDP. Ils diffèrent par la distribution des observations. Dans le premier cas, les observations sont construites depuis les séquences de concepts décodées à partir des transcriptions manuelles (23,5% d'erreur de classification), dans le second à partir des transcriptions automatiques (29,2%). Les deux courbes de la Figure 1 montrent que, en ce qui concerne la convergence de l'apprentissage, la technique proposée semble faiblement sensible à l'augmentation des erreurs de decodage.

7. Conclusion

Dans cet article, une approche pour la définition non-supervisée des états du dialogue dans un système de dialogue homme-machine a été introduite et étudiée. Cette méthode a permis le développement d'un gestionnaire de dialogue à base de POMDP sur la tâche MEDIA.

Les expériences préliminaires montrent la faisabilité de l'approche. Une évaluation avec des utilisateurs réels devra être menée afin de valider en pratique la méthode.

Références

[1] H. Bonneau-Maynard et al, "Results of the French EVALDA-MEDIA evaluation campaign for literal understanding", LREC, 2006.

[2] F. Lefèvre, "Dynamic Bayesian Networks and Discriminative Classifiers for Multi-stage Semantic Interpretation", IEEE ICASSP, 2007.

[3] E. Levin, R. Pieraccini and W. Eckert, "Using Markov Decision Processes For Learning Dialogue Strategies", IEEE ICASSP, 1998.

[4] N. Roy, J. Pineau and S. Thrun, "Spoken dialog management using probabilistic reasoning", ACL, 2000.

[5] S. Young et al, "The Hidden Information State Approach to Dialog Management", IEEE ICASSP, 2007.

[6] R. Sutton and A. Barto, "Reinforcement Learning : An Introduction", In Adaptive Computation and Machine Learning", MIT Press, 1998.

[7] J. Williams and S. Young, "Partially observable Markov decision processes for spoken dialog systems", In Computer Speech and Language, 21 :393-422, 2007.

[8] C. Boutilier and D. Poole, "Computing optimal policies for partially observable decision processes using compact representations", AAAI, 1996.

[9] D. Aha and D. Kibler, "Instance-based learning algorithms", In Machine Learning, 6 :37-66, 1991.

[10] J. Bilmes and K. Kirchhoff, "Factored language models and generalized parallel backoff", HLT-NAACL, 2003.

[11] S. Keerthi et al., "Improvements to Platt's SMO algorithm for SVM classifier design", In Neural Computation, 13(3) :637-649, 2001.

[12] C. Servan et al, "Conceptual decoding from word lattices : application to the spoken dialogue corpus MEDIA", ICSLP, 2006.

[13] M. Spaan and N. Vlassis, "Perseus : randomized point-based value iteration for POMDPs", Technical Report, Universiteit van Amsterdam, 2004.