

# Simulation du processus de croyance mutuelle de la compréhension dans le dialogue (*grounding process*) à l'aide des réseaux bayésiens

Stéphane Rossignol, Olivier Pietquin, Michel Ianotto

Équipe IMS

Supélec – Campus de Metz, 2, rue Édouard Belin ; 57070 Metz, France

stephane.rossignol@supelec.fr

http://ims.metz.supelec.fr/

## ABSTRACT

User simulation has become an important trend of research in the field of spoken dialogue systems because collecting and annotating real interactions with users is often expensive and time consuming. Yet, such data are generally required for designing and assessing efficient dialogue systems. The general problem of user simulation is thus to produce as many as necessary natural, various and consistent interactions from as few data as possible. In this paper, we propose a user simulation method based on Bayesian Networks (BN) that is able to produce consistent interactions in terms of user goal and dialogue history but also to simulate the grounding process that often appears in human-human interactions. The BN is trained on a database of 1234 human-machine dialogues in the TownInfo domain (a tourist information application). Experiments with a state-of-the-art dialogue system (REALL-DUDE/DIPPER/OAA) have been realized and promising results are presented.

**Keywords:** Bayesian network, spoken dialogue systems, grounding process

## 1. Introduction

Les systèmes de dialogue oral sont à présent très répandus et sont utilisés dans un grand nombre de domaines (de la réservation de vols à l'aide en ligne à la réparation). Mettre en place de telles interfaces basées sur la parole est habituellement un processus itératif qui comprend plusieurs cycles de prototypage, de tests et de validation. Les tests et les validations requièrent des interactions entre le système courant et des utilisateurs humains, ce qui rend ces phases chères et très gourmandes en temps. Pour cette raison, la simulation d'utilisateurs est devenue, depuis une dizaine d'années, un très important domaine de recherche. La simulation d'utilisateurs peut dès lors être utilisée pour tester les performances d'un système et pour optimiser celui-ci [4, 7]. Il peut s'agir par exemple d'optimiser la politique du gestionnaire de dialogue (GD) à l'aide de méthodes d'apprentissage par renforcement. La simulation d'utilisateurs ne doit pas être confondue avec la modélisation d'utilisateurs. Dans un système de dialogue, cette dernière est utilisée en général dans des buts internes, comme par exemple pour représenter la connaissance qu'a l'utilisateur de l'état d'avancement du dialogue, ou pour inférer le but de l'utilisateur [2, 5], ou encore pour simuler le comportement du module de compréhens-

sion [8] de la parole. Par contre, l'objectif de la simulation d'utilisateurs est la création d'un grand nombre d'interactions simulées avec un système de dialogue, et de ce fait l'utilisateur peut être considéré comme étant extérieur au système.

L'utilisateur simulé (US) présenté dans cet article est basé sur les réseaux bayésiens (RB). Ce modèle a été choisi pour plusieurs raisons. Premièrement, les RB sont des modèles génératifs et peuvent donc être utilisés aussi bien pour inférer que pour générer des données, ce qui est bien sûr requis pour la simulation. Deuxièmement, il s'agit d'un paradigme statistique, qui peut donc générer une grande variété de dialogues cohérents statistiquement. Troisièmement, les paramètres des RB peuvent être soit fixés par des experts, soit entraînés à partir d'une base de données. Étant donné que la collecte de données se révèle souvent un processus difficile, l'apport d'un expert peut être très utile. Pour finir, il existe, pour effectuer l'inférence dans les RB et l'entraînement des RB, un grand nombre d'outils libres.

Cet article se fonde sur des travaux précédents [7] où les RB sont utilisés pour simuler des utilisateurs, mais il met l'accent sur deux contributions nouvelles. Tout d'abord, le modèle a été modifié pour générer des problèmes de différence de croyance mutuelle de la compréhension dans le dialogue (*grounding process*). Ce processus de *grounding* est considéré ici comme étant le processus utilisé par deux locuteurs d'une même conversation pour s'assurer qu'ils partagent la même connaissance de l'état d'avancement du dialogue. En pratique, ceci veut dire que l'utilisateur simulé réagira automatiquement si un problème de transmission de l'information est détecté, en donnant de nouvelle information correcte au système. Le but principal de ce travail est d'entraîner des politiques de GD qui prennent en compte de tels problèmes de *grounding* [6]. Ensuite, le modèle est entraîné avec des données réelles concernant des interactions homme-machine ; puis il est testé en combinaison avec un système de dialogue de l'état de l'art (l'environnement REALL-DUDE/DIPPER/OAA [3, 1]).

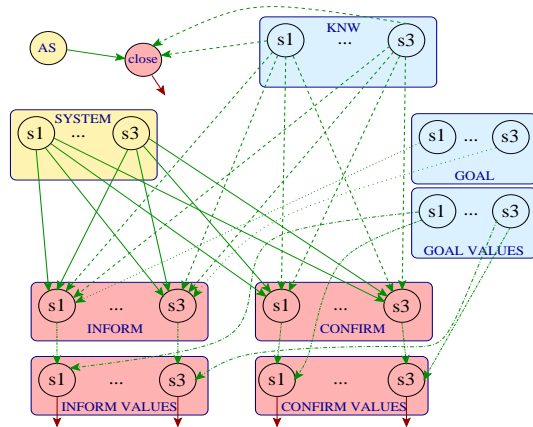
La tâche considérée se place dans le domaine du TownInfo. Elle concerne l'aide aux touristes. La tâche consiste à fournir des informations concernant des restaurants dans une ville donnée. Ceci peut être considéré comme étant une tâche de remplissage d'un formulaire à trous, où les trois différents attributs requérant d'être informés sont : le type de nourriture,

la gamme de prix et l'emplacement dans la ville. Ces attributs peuvent respectivement prendre 3, 3 et 5 valeurs.

## 2. Description du modèle

### 2.1. Les réseaux bayésiens

Le cœur du réseau bayésien utilisé ici est décrit dans [7]. Les lecteurs sont invités à se référer à ce travail pour des compléments théoriques. À la base, ce RB s'inspire de l'idée que la réponse de l'utilisateur à une action du système est influencée par le type du message reçu (AS), par le but de l'utilisateur (GOAL) et par sa connaissance (KNW) concernant l'état d'avancement du dialogue. La réponse de l'utilisateur peut être de différents types : INFORM, CONFIRM, et une action spéciale consistant à vouloir clore le dialogue. La figure 1 montre en détails la structure de l'US. Pour des raisons de portabilité et pour être à même d'intégrer cet US dans le système de dialogue visé, il a été implanté en C++, en utilisant la librairie Smile (<http://genie.sis.pitt.edu/about.html#smile>). Smile est une librairie de classes C++ qui mettent en place des modèles de graphes probabilistes et de théorie de la décision, comme les RB, etc.



**Fig. 1:** Utilisateur simulé à base de réseaux bayésiens

Le nœud AS

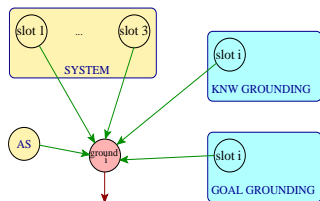


Fig. 2: La composante de *grounding*

donner l’information concernant cet attribut.

### 3. Expérimentations

#### 3.1. Interaction avec REALL-DUDE/DIPPER/OAA

L’utilisateur simulé a été interfacé avec le système de dialogue oral inclus dans l’environnement REALL-DUDE/DIPPER/OAA (voir [3] et [1]). Cet environnement, originellement, a pour but l’entraînement de politiques de gestion de dialogue via des méthodes d’apprentissage par renforcement. Jusqu’à présent, la politique utilisée pour nos expérimentations a été entraînée indépendamment de l’US présenté ici et est utilisée seulement pour des tests. L’US accompagne chaque hypothèse qu’il envoie au GD d’un score de confiance simulé, qui serait donné, entre autres, par le reconnaiseur de parole (ASR). Si cette confiance est trop petite, le GD peut décider de demander à l’utilisateur de confirmer.

Dans cette section, sont présentées des statistiques calculées sur des dialogues obtenus en interfaçant l’US présenté dans cet article et l’environnement REALL-DUDE/DIPPER/OAA. Comme la tâche comprend trois attributs et comme l’US est configuré pour ne pas informer le GD au sujet de plus d’un attribut par tour, le nombre minimum de tours nécessaires pour atteindre la fin du dialogue est quatre : un par attribut, plus le tour “Close”. Il faut noter qu’un tour est défini ici comme étant un couple  $\langle \text{sys act} \rangle / \langle \text{user act} \rangle$  (excepté pour l’action “Close”, puisque c’est le GD qui a la possibilité de stopper les dialogues).

En considérant seulement les dialogues où un problème de *grounding* est survenu, et pas la totalité des dialogues, les résultats de la table 1 sont obtenus. Ils se révèlent prometteurs. Le taux de dialogues plus long que 5 tours diminue de 30.6 %.

#### 3.2. Statistiques sur les dialogues

Dans cette section, sont présentées des statistiques calculées sur les dialogues obtenus en interfaçant deux versions de l’US présenté dans cet article et l’environnement REALL-DUDE/DIPPER/OAA. La première version de l’US est telle que les paramètres du RB ont des valeurs heuristiquement déterminées par un expert. La seconde version est telle que ces paramètres sont appris. La base de données utilisée pour l’entraînement contient 1234 dialogues. Elle comprend des dialogues beaucoup plus compliqués que ceux obtenus pour la tâche considérée ici. Douze attributs en tout sont présents. Les valeurs pour certains de ces attri-

Tab. 1: Nombre moyen de tours requis pour atteindre la fin des dialogues ; pourcentage de dialogues pour lesquels la fin est atteinte en plus de 5 tours

	moyenne	> 5 tours
sans composante <i>grounding</i>	5.254	29.22 %
avec composante <i>grounding</i>	5.069	20.29 %

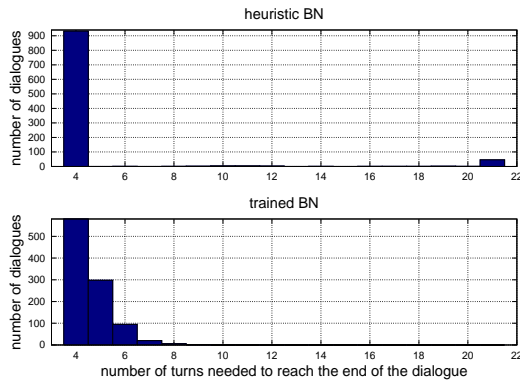
Tab. 2: Nombre moyen de tours requis pour atteindre la fin des dialogues ; nombre max de tours ; nombre min de tours ; pourcentage de dialogues pour lesquels la fin est atteinte en 4 tours ; pourcentage de dialogues pour lesquels la fin est atteinte en moins de 9 tours

	moyenne	max	min	4 tours	< 9 tours
h-RB	4.969	21	4	93.20 %	93.40 %
e-RB	4.577	9	4	58.00 %	99.90 %

buts sont requis par le GD, comme “type de nourriture”, etc., réclamant donc une action “Request” du système et une action “Inform” de l’utilisateur ; d’autres sont requis par l’utilisateur, comme “adresse”, etc. ; la “gamme de prix” peut être requise par les deux protagonistes. De plus, plus de dix actions différentes, du système et de l’utilisateur, sont considérées (voir [10] pour une liste exhaustive). Finalement, plus d’un attribut et plus d’une action peuvent être présentés pendant chaque tour, aussi bien considérant les messages du GD que ceux de l’utilisateur. Ainsi, la base de données a nécessité d’être labellée avec soin, dans l’objectif de fournir des données utilisables pour cet article. La base de données a été décrite avec plus de détails dans [9], où elle a été utilisée pour entraîner des stratégies du GD.

Dans la table 2, sont présentés les résultats obtenus en utilisant le RB heuristique (h-RB) et le RB entraîné (e-RB). Mille dialogues ont été simulés pour chaque RB. Sur la figure 3, est présenté l’histogramme du nombre de tours requis pour atteindre la fin des dialogues. Considérant le h-RB, on peut remarquer que la plupart des dialogues sont, en effet, complétés en quatre tours, comme attendu (93.4 %). Ceci est dû au fait qu’il a été spécialement mis en place pour obtenir des dialogues aussi courts que possible.

L’e-RB donne des dialogues plutôt plus longs que ceux donnés par le h-RB. Ceci ne peut pas être remarqué si l’on considère le nombre moyen de tours ; il faut considérer le pourcentage de dialogues qui ont nécessité exactement quatre tours pour atteindre leur fin : ce pourcentage passe de 93.2 % à 58.0 %. Cependant, de façon très prometteuse, le nombre moyen de tours et le pourcentage de dialogues qui ont nécessité moins de neuf tours pour atteindre leur fin sont plutôt meilleurs quand on utilise l’e-RB ; de plus il faut remarquer que les très longs dialogues (plus de neuf tours), indiquant une profonde incompréhension entre le GD et l’US, ont complètement disparus. Ceci indique que les dialogues obtenus avec l’e-RB sont plus naturels, au moins du point de vue du GD, que les



**Fig. 3:** Histogramme du nombre de tours requis pour atteindre la fin du dialogue – haut : le RB heuristique est utilisé – bas : le RB entraîné est utilisé

**Tab. 3:** Nombre de tours requis pour chaque attribut ; h-RB, plus longs dialogues conservés ; h-RB, plus longs dialogues éliminés ; e-RB ; et base de données

h-RB	h-RB sans longs dialogues	e-RB	base de données
1.6563	1.3986	1.5257	1.5661

dialogues obtenus avec le h-RB. La distribution des actions est en fait plus en accord avec les données, ce qui montre de nouveau que les dialogues simulés sont plus naturels avec l’e-RB.

La table 3 présente le nombre de tours requis par attribut, respectivement quand le h-RB est utilisé les plus longs dialogues étant conservés, quand le h-RB est utilisé les plus longs dialogues n’étant pas conservés (ils reflètent seulement la condition d’arrêt incluse dans le GD), quand l’e-RB est utilisé, et considérant la base de données. Clairement, l’e-RB donne des dialogues plus réalistes, en terme de nombre de tours requis pour chaque attribut.

#### 4. Conclusion et perspectives

Dans cet article, un modèle d’utilisateur simulé basé sur les réseaux bayésiens est proposé pour simuler des dialogues homme-machine (au niveau de l’intention), réalistes, et incluant des comportements de *grounding*. Notre but était de montrer l’intérêt de simuler le processus de *grounding*, qui survient souvent dans les dialogues humain-humain. Ceci est fait en comparant le nombre de tours requis pour atteindre la fin d’un dialogue quand différentes configurations du modèle proposé sont utilisées. Plusieurs perspectives sont de plus envisagées.

Premièrement, cet Utilisateur Simulé sera utilisé pour l’entraînement des politiques du GD, ce considérant le paradigme de l’apprentissage par renforcement. Deuxièmement, des résultats préliminaires concernant l’interaction de l’US présenté ici avec un GD indépendamment développé sont montrés. Il est prévu à court terme d’interagir avec le système de dialogue inclus dans l’environnement REALL-DUDE/DIPPER/OAA d’une manière beaucoup plus intensive et systématique. Ceci nous permettra de

comparer le nombre de tours obtenus avec les Utilisateurs Simulés basés sur les RB et le nombre de tours obtenus avec des utilisateurs humains, d’analyser le taux de complétion de la tâche correspondant, etc. Ceci nous permettra aussi d’entraîner la politique utilisée dans les POMDP implantés dans l’environnement REALL-DUDE/DIPPER/OAA. Troisièmement nous souhaiterions utiliser la capacité des RB d’apprendre en ligne leurs paramètres afin d’améliorer le caractère naturel des dialogues simulés, alors que des utilisateurs réels sont en train d’interagir avec le système ; et nous souhaiterions ré-entraîner des politiques du GD à partir de ces interactions réelles.

#### Références

- [1] J. Bos, E. Klein, O. Lemon, and T. Oka. DIPPER : Description and Formalisation of an Information-State Update Dialogue System Architecture. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, pages 115–124, 2003.
- [2] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse. The Lumiere Project : Bayesian User Modeling for Inferring the Goals and Needs of Software Users. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998.
- [3] O. Lemon, X. Liu, D. Shapiro, and C. Tollander. Hierarchical Reinforcement Learning of Dialogue Policies in a Development Environment for Dialogue Systems : REALL-DUDE. In *10th SemDial Workshop on the Semantics and Pragmatics of Dialogue ; BRANDIAL*, 2006.
- [4] E. Levin, R. Pieraccini, and W. Eckert. A Stochastic Model of Human-Machine Interaction for Learning Dialog Strategies. In *IEEE Transactions on Speech and Audio Processing*, volume 8, pages 11–23, 2000.
- [5] H. Meng, C. Wai, and R. Pieraccini. The Use of Belief Networks for Mixed-Initiative Dialog Modeling. In *Proceedings of the 8th International Conference on Spoken Language Processing (ICSLP)*, 2000.
- [6] O. Pietquin. Learning to Ground in Spoken Dialogue Systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 4, pages 165–168, 2007.
- [7] O. Pietquin and T. Dutoit. A Probabilistic Framework for Dialog Simulation and Optimal Strategy Learning. In *IEEE Transactions on Audio, Speech, and Language Processing*, volume 14, pages 589–599, 2006.
- [8] O. Pietquin and T. Dutoit. Dynamic Bayesian Networks for NLU Simulation with Applications to Dialog Optimal Strategy Learning. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2006.
- [9] J. D. Williams and S. Young. Partially Observable Markov Decision Processes for Spoken Dialogue Systems. In *Computer Speech and Language*, volume 21, pages 231–422, 2007.
- [10] S. Young. CUED Standard Dialogue Acts. Technical report, Cambridge University Engineering Dept, 2007.