

Un assistant vocal personnalisable

Tatiana Ekeinhor-Komi ^{1,3} Hajar Falih ² Christine Chardenon ¹
Romain Laroche ² Fabrice Lefevre ³

(1) Orange Labs, 2 Avenue Pierre Marzin, 22300 Lannion

(2) Orange Labs, 38-40 Rue du Général Leclerc, 92130 Issy les Moulineaux

(3) LIA-CERI, Université d'Avignon, France

prenom.nomsanstiret@orange.com, fabrice.lefevre@univ-avignon.fr

Résumé. Nous proposons la démonstration d'un assistant personnel basé sur une architecture distribuée. Un portail vocal relie l'utilisateur à des applications. Celles-ci sont installées par l'utilisateur qui compose de ce fait son propre assistant personnel selon ses besoins.

Abstract. We introduce a personal assistant based on a distributed architecture. A portal connects user to applications. Applications are installed by a user who compose his own assistant according to his needs.

Mots-clés : Système de dialogue, applications du traitement automatique du langage naturel, assistant personnel.

Keywords: Dialogue system, natural language processing applications, personal assistant.

Présentation du démonstrateur

Un assistant vocal est un système de dialogue qui converse avec son utilisateur en langage naturel, afin de répondre aux divers besoins de celui-ci. Avec Google Now et Siri, les assistants vocaux se sont démocratisés dans la vie de tous les jours. Leur utilité croît avec les maisons intelligentes, objets connectés, etc. Plus que jamais, on attend de ces systèmes qu'ils soient capables de gérer des dialogues portant sur des sujets divers et variés.

Dans le but de concevoir des systèmes de dialogue multi-domaines (Hsu *et al.*, 2002; Lee *et al.*, 2012; Planells *et al.*, 2013), l'approche dominante a été d'étendre les domaines d'un système existant (Gašić *et al.*, 2013). Ceci fonctionne bien pour une extension d'un voire deux domaines. Mais cela ne permet pas de gérer la dynamique de dizaines d'applications et encore moins de réaliser l'extension automatiquement. C'est pourquoi, le modèle du démonstrateur s'inspire plutôt du modèle distribué de (Lin *et al.*, 1999) car celui-ci permet d'ajouter ou de supprimer un domaine sans perturber le fonctionnement de l'existant. Ce modèle considère le système comme un ensemble constitué d'un module central et de sous-modules correspondant aux domaines possibles du système de dialogue. Le module central assure la liaison entre l'utilisateur et les sous-modules. Dans le cas du démonstrateur, les domaines correspondent à des applications de dialogue. C'est la modularité du système qui permet la composition des applications sur demande de l'utilisateur. De plus cette personnalisation permet à l'utilisateur de n'installer que des applications qui lui seront utiles. Nous proposons la démonstration d'un assistant personnel basé sur une architecture distribuée. Un portail vocal relie l'utilisateur à des applications. Celles-ci sont installées par l'utilisateur qui compose de ce fait son propre assistant personnel selon ses besoins.

Le démonstrateur est une mise en œuvre préliminaire du modèle distribué, où chaque domaine est une application de dialogue indépendante, appelée à partir d'une application centrale, dénommée *Portail*. Le démonstrateur est donc constitué du *Portail* et d'une application spéciale : le *Store*. Il sert à installer, mettre à jour ou désinstaller des applications. A titre d'illustration trois applications sont actuellement proposées sur le *Store* : *MeteoOrange* est une application de météo qui se connecte au site de météo d'Orange ; *NewsAFP* est une application d'actualités catégorisée en plusieurs thèmes (politique, sport, affaires), il se connecte au site de l'AFP (Agence France-Presse) ; *ShoppingList* est une application qui permet de gérer sa liste de course. A chaque nouvelle mise à jour (installation, désinstallation), le portail doit intégrer les données sémantiques fournies par l'application. Ces données permettent au portail d'interpréter une requête et de sélectionner

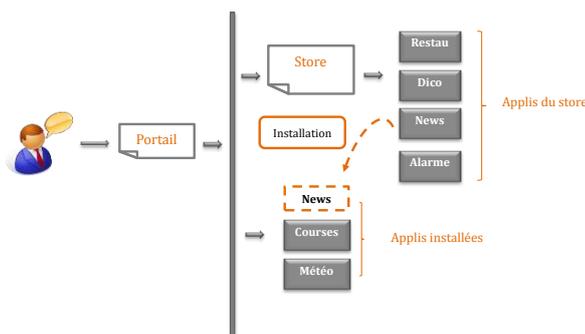


FIGURE 1 – Architecture générale

l'application correspondante. L'architecture est présentée dans la figure 1. Les applications ont été développées avec l'outil interne d'Orange : Disserto (Laroche *et al.*, 2010). Une interface web a été développée pour réaliser la démonstration. L'API de Google est utilisée pour la reconnaissance vocale, la transcription textuelle et la synthèse vocale. Le tout est déployé sur un serveur d'application.

Exemple de dialogue démontrant l'installation à la volée d'une application (comme illustré dans la figure 1)

U : Quels sont les actualités politiques du jour ?

S : Désolé, nous ne pouvons donner suite à votre demande. Voulez-vous rechercher sur le Store une application à installer ?

U : Oui.

S : Voulez-vous faire une recherche par nom ou par description ?

U : Par nom.

S : Donnez le nom de l'application à installer

U : Je veux installer NewsAFP.

S : Confirmez-vous l'installation de NewsAFP ?

U : Oui.

S : L'application NewsAFP est désormais installée sur votre portail.

U : Quelles sont les actualités politiques ?

S : [Résultats]

Ce sujet embrasse les problématiques de conception de systèmes de dialogue multi-domaines par l'intégration de systèmes existants d'une part mais aussi de la gestion de dialogues complexes d'autre part. L'étape suivante de notre travail sera de permettre la transition cohérente entre les différentes applications au cours de l'interaction ce qui supposera d'améliorer les algorithmes de fusion et de désambiguïsation qui sont implantés pour le moment sous une forme simple dans un modèle à base de règles.

Références

- GAŠIĆ M., BRESLIN C., HENDERSON M., KIM D., SZUMMER M., THOMSON B., TSIAKOULIS P. & YOUNG S. (2013). Pomdp-based dialogue manager adaptation to extended domains.
- HSU W.-T., WANG H.-M. & LIN Y.-C. (2002). The design of a multi-domain chinese dialogue system. In *International Symposium on Chinese Spoken Language Processing*.
- LAROCHE R., BRETIER P. & PUTOIS G. (2010). Enhanced monitoring tools and online dialogue optimisation merged into a new spoken dialogue system design experience. In *Proceedings of Interspeech*, Chiba (Japan).
- LEE S. R. D. L. I., LEE S. H. G. G. & KIM M. K. K. (2012). A hierarchical domain model-based multi-domain selection framework for multi-domain dialog systems.
- LIN B.-S., WANG H.-M. & LEE L.-S. (1999). A distributed architecture for cooperative spoken dialogue agents with coherent dialogue state and history. In *Asru*, volume 99, p.4 : Citeseer.
- PLANELLIS J., HURTADO L.-F., SEGARRA E. & SANCHIS E. (2013). A multi-domain dialog system to integrate heterogeneous spoken dialog systems.