

<b>Domain</b>	Interaction
<b>Domain - extra</b>	Dialogue et apprentissage
<b>Year</b>	2013
<b>Starting</b>	Septembre 2013
<b>Status</b>	Open
<b>Subject</b>	Dialogue et apprentissage
<b>Thesis advisor</b>	Sophie Rosset (LIMSI, TLP)
<b>Co-advisors</b>	co-encadrant : Gabriel Illouz (LIMSI, ILES)
<b>Laboratory</b>	LIMSI
<b>Collaborations</b>	

### ***Abstract***

Ce projet s'inscrit dans le thème des systèmes de dialogue apprenant de l'utilisateur et à l'utilisateur. Si d'une part les systèmes de dialogue en langue naturelle (écrite ou parlée) et d'autre part les systèmes apprenant à base d'apprentissage actif ont fait l'objet de très nombreux travaux, peu ont encore tenté de réunir ces deux champs d'activités. Le cadre applicatif dépendra des projets en cours mais essentiellement différents cadres tels que l'interfaçage avec un domaine formel (analyses statistiques), l'apprentissage de comportements sociaux, la réalisation de tâches de renseignements, etc. serviront de base d'expériences.

### ***Contexte***

Ce projet s'inscrit dans le thème des systèmes de dialogue apprenant la réalisation d'une tâche par l'interaction avec un utilisateur expert. Il s'agit de concevoir un système de dialogue capable d'apprendre à réaliser une tâche et que dans ce contexte il apprenne à interagir pour augmenter ses connaissances. Au-delà de l'apprentissage d'une tâche, il s'agit d'apprendre à modéliser la prise de décision quant au comportement optimal à adopter pour obtenir les informations manquantes à la réalisation d'une tâche.

Si les systèmes de dialogue en langue naturelle (écrite ou parlée) et les systèmes à base d'apprentissage actif ont fait l'objet de très nombreux travaux, peu ont tenté de réunir ces deux aspects. Dans ce cadre, de nombreuses questions se posent et notamment quelles sont les limites de l'apprentissage par l'interaction ? Quel amorçage minimal est nécessaire ? Pour restreindre le champ d'étude, les cadres applicatifs porteront sur l'acquisition de compétences statistiques en liaison avec un langage formel (R par exemple) ou encore des compétences systèmes comme la manipulation de fichiers via des langages de script.

## **Objectifs**

L'objectif de cette thèse est de proposer une approche permettant d'apprendre à modéliser le comportement d'un système de dialogue pour apprendre à réaliser une tâche.

On partira de formes de raisonnement par induction et par analogie minimales de manière à étudier les limites de ces approches pour qu'un système apprenne à apprendre une tâche et un domaine par l'interaction. Il conviendra également de proposer un modèle pour discriminer les informations pertinentes à extraire des énoncés en langue naturelle produits dans un cadre interactif. Cette discrimination résultera d'un apprentissage validé par l'interaction avec l'utilisateur expert.

L'analyse des énoncés dans un cadre interactif pose le problème de la variation linguistique et surtout des phénomènes de références (anaphores et déictiques). L'apprentissage doit intégrer la résolution de ses phénomènes et sera aussi validé par l'interaction avec l'utilisateur expert.

La thèse permettra d'apporter des éléments de réponses à des questions plus générales telles que :

- quelles sont les limites de l'induction et de l'analogie pour apprendre ?
- l' « apprentissage tutoré » (avec un utilisateur expert) a-t-il des limites ?
- le savoir et la compétence sont-ils transmissibles via un média numérique ?
- jusqu'à quel point un tel système apprenant pourra à son tour être utile à un utilisateur non expert ?

## **Programme de travail**

Année 1 :

A partir d'un premier prototype de dialogue implémentant un modèle à base d'induction et un à base d'analogie pour l'apprentissage, il conviendra d'analyser et évaluer précisément l'état d'évolution des systèmes afin de proposer des adaptations et évolutions des modèles :

- état de l'art (dialogue, apprentissage actif et modèle d'apprentissage à base d'induction et d'analogie)
- analyse des cas d'échecs et de succès des systèmes existants et propositions des premières méthodologies d'évaluation en tenant compte de l'évolution des systèmes
- production d'un premier article

Année 2 :

- utilisation de l'interaction pour détecter et traiter les cas d'échecs (éventuellement par le biais de score de confiance sur les décisions et acquis précédents)
- généralisation de l'apprentissage
- identification de l'expert le plus approprié pour expliquer à l'utilisateur ce que le système sait ne pas savoir faire (extension du contexte)
- publications dans les domaines concernés (TAL, EIAH, Dialogue et pragmatique, IA)

Année 3 :

- validation de l'approche sur un autre langage formel de même type (commandes shell, etc.)
- rédaction de la thèse
- stage doctoral à l'étranger si état d'avancement de rédaction de la thèse au bout de 6 mois satisfaisant
- article de revue de synthèse des travaux

## ***Informations supplémentaires***

### Références

K. Jokiinen and M. McTear. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers. 2009.

F. Pinault. Apprentissage par renforcement pour la généralisation des approches automatiques dans la conception des systèmes de dialogue oral. Thèse de doctorat en informatique. Université d'Avignon et des Pays du Vaucluse. 2011.

James Allen, Nathanael Chambers, George Ferguson, Lucian Galescu, Hyuckchul Jung, Mary Swift, William Taysom. PLOW: A Collaborative Task Learning Agent. AAI 2007.

J Caelen, L Villaseñor. Dialogue homme-machine et apprentissage, In Apprentissage par l'interaction, 1997.

Alexis Bondu. Apprentissage actif par modèles locaux. Thèse d'informatique. Université d'Angers. 2008.

### ***Prérequis***

- connaissance en TAL, apprentissage et IA

### ***Détails***