

Traitement du contenu verbal et analyse des sentiments dans les systèmes d'interactions humain-agent

Mots clés :

- **Directeur de thèse** : Catherine PELACHAUD
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire Traitement et Communication de l'Information
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

Secteur: Traitement Automatique du Langage Naturel, Dialogue Homme-Machine Le domaine du sentiment analysis et de l'opinion mining est un domaine en plein essor avec l'arrivée en masse de données textuelles sur le web comportant des expressions d'opinions par les citoyens (critiques de films, débats sur les commentaires de forums, tweets) (El-Bèze et al. 2010)). Les recherches en traitement automatique des langues se mobilisent sur le développement de méthodes de détection d'opinion dans les textes en s'appuyant sur ces nouvelles ressources. La diversité des données et des applications industrielles faisant appel à ces méthodes multiplient les défis scientifiques à relever avec, notamment, la prise en compte des différents contextes d'énonciation (e.g., contexte social et politique, personnalité du locuteur) et la définition du phénomène d'opinion à analyser en fonction du contexte applicatif. Ces méthodes d'analyse des sentiments dans les textes s'étendent également depuis peu à l'oral en passant par l'analyse des transcriptions automatiques issues de systèmes de reconnaissance automatique de la parole pour des problématiques d'indexation d'émissions radiophoniques ou de centres d'appels (Clavel et al., 2013), et peuvent être ainsi corrélées aux méthodes d'analyse acoustique/prosodique des émotions (Clavel et al., 2010). Autre domaine scientifique en plein essor, celui des agents conversationnels animés (ACA) fait intervenir des personnages virtuels interagissant avec l'humain. Les ACA peuvent prendre un rôle d'assistant comme les agents conversationnels présents sur les sites de vente (Suignard, 2010), de tuteur dans le cadre des Serious Games (Chollet et al. 2012) ou encore de partenaire dans le cadre des jeux vidéos. Le défi scientifique majeur pour ce domaine est l'intégration, au sein de l'ACA, de la composante affective de l'interaction. Il s'agit d'une part de prendre en compte les comportements affectifs et des attitudes sociales de l'humain et d'autre part de les générer de façon pertinente. Nous proposons pour cette thèse de travailler sur la détection des opinions et des sentiments dans un contexte d'interaction multimodale de l'humain avec un agent conversationnel animé, sujet jusqu'à maintenant peu étudié par la "communauté agent". En effet, d'un côté, les ACA réagissent à des contenus émotionnels essentiellement non verbaux (Schröder et al., 2011) et de l'autre côté, les ACA "assistant" réagissent à partir des contenus verbaux informatif (Suignard, 2010) sans prendre en compte les opinions ou les sentiments exprimés par l'utilisateur. Des premières études ont été réalisées sur la reconnaissance de l'affect dans le langage dans un contexte d'interaction avec un agent (Osherenko et al., 2009) mais celles-ci restent envisagées indépendamment de la stratégie de dialogue. Les développements de la thèse s'intégreront dans la plateforme GRETA qui repose sur l'architecture SAIBA, une architecture globale unifiée développée par la "communauté agent" pour la génération de comportements multimodaux (Niewiadomski et al., 2011). Greta permet de communiquer avec l'humain en générant chez l'agent une large palette de comportements expressifs verbaux et non verbaux (Bevacqua et al., 2012). Elle peut simultanément montrer des expressions faciales, des gestes, des regards et des mouvements de têtes. Cette plateforme a notamment été intégrée dans le cadre du projet SEMAINE avec le développement d'une architecture temps-réel d'interaction humain-agent (Schröder et al., 2011) qui inclut des analyses acoustiques et vidéos, un système de gestion du dialogue et, du côté de la synthèse, le système Text To Speech OpenMary et l'agent virtuel de la plateforme GRETA. A l'instar de ce projet, la détection d'opinions et de sentiments envisagée dans la thèse interviendra en entrée des modèles d'interactions multi-modaux de la plateforme. La stratégie de dialogue multimodale associée à ces entrées relatives au contenu verbal devra être définie et intégrée dans la plateforme GRETA.

Résumé du projet de recherche (Langue 2)

La thèse portera sur le développement conjoint de méthodes de détection des opinions et des sentiments et de stratégies de dialogue humain-agent. Les méthodes envisagées sont des méthodes hybrides mêlant apprentissage statistique et règles expertes. Pour les stratégies de dialogue, le doctorant pourra s'appuyer sur les travaux réalisés dans le cadre du moteur de dialogue DISCO (Rich et al., 2012) et du moteur développé dans le projet Semaine (Schröder et al., 2011). Les méthodes développées pourront également s'appuyer sur des analyses de corpus humain-humain ou de type Magicien d'Oz (McKeown et al., 2012) et un protocole d'évaluation de ces méthodes devra être mis en place. En particulier, pour répondre à cet objectif, la thèse devra aborder les problématiques suivantes: - la définition des types d'opinions et de sentiments pertinents à considérer en entrée du moteur de dialogue. Il s'agira d'aller au-delà de la distinction classique entre opinions positives et opinions négatives, peu pertinente dans ce contexte, en s'appuyant sur les modèles issus de la psycholinguistique (Martin and White, 2007); - l'identification des marqueurs lexicaux, syntaxiques, sémantiques et dialogiques des opinions et des sentiments; - la prise en compte du contexte d'énonciation: les règles implémentées pourront intégrer différentes fenêtres d'analyse: la phrase, le tour de parole et les tours de paroles antérieurs; - la prise en compte des problématiques temps-réel de l'interaction: des stratégies de dialogues seront définies en fonction des différentes fenêtres d'analyse afin de proposer des stratégies d'interactions à différents niveaux de réactivité. Par exemple, certains mots-clés pourront être utilisés comme déclencheurs de backchannel en temps réels et la planification des comportements de l'agent pourra être ajustée au fur et à mesure de l'avancement de l'interaction.

Informations complémentaires (Langue 1)

Ces travaux de thèse interviennent en complémentarité des travaux réalisés sur les interactions non verbales dans le cadre du projet européen FP7 TARDIS prenant comme application les Serious games dans le cas d'un entraînement à l'entretien d'embauche (<http://tardis.lip6.fr/presentation>) et des travaux réalisés sur le traitement des signaux sociaux dans le cadre du réseau d'excellence SSPNET (<http://sspnet.eu/>) Une collaboration avec Candy Sidner, professeur au département Computer Science du Worcester Polytechnic Institute et experte en modèles computationnels d'interactions verbales et non verbales et à l'origine du moteur de dialogue DISCO (Richet et al. 2012) sera également mise en place.

Informations complémentaires (Langue 2)

Références: E. Bevacqua, E. de Sevin, S.J. Hyniewska, C. Pelachaud (2012), A listener model: Introducing personality traits, *Journal on Multimodal User Interfaces*, special issue Interacting ECAs, Elisabeth André, Marc Cavazza and Catherine Pelachaud (Guest Editors), 6:27-38, 2012. M. Chollet, M. Ochs and C. Pelachaud (2012), Interpersonal stance recognition using non-verbal signals on several time windows, *Workshop Affect, Compagnon Artificiel, Interaction*, Grenoble, November 2012, pp. 19-26 C. Clavel and G. Richard (2010). Reconnaissance acoustique des émotions, *Systèmes d'interactions émotionnelles*, C. Pelachaud, chapitre 5, 2010 C. Clavel, G. Adda, F. Cailliau, M. Garnier-Rizet, A. Cavet, G. Chapuis, S. Courcinous, C. Danesi, A-L. Daquo, M. Deldossi, S. Guillemin-Lanne, M. Seizou, P. Suignard (2013). Spontaneous Speech and Opinion Detection: Mining Call-centre Transcripts. In *Language Resources and Evaluation*, avril 2013. M. El-Bèze, A. Jackiewicz, S. Hunston, Opinions, sentiments et jugements d'évaluation, *Revue TAL 2010*, Volume 51 Numéro 3. J.R. Martin, P.R.R. White (2007) *Language of Evaluation: Appraisal in English*, Palgrave Macmillan, Novembre 2007 G. McKeown, M. Valstar, R. Cowie, R., M. Pantic, M. Schroder (2012) The SEMAINE Database: Annotated Multimodal Records of Emotionally Colored Conversations between a Person and a Limited Agent, *IEEE Transactions on Affective Computing*, Volume: 3, Issue: 1, Page(s): 5- 17, Jan.-March 2012 R. Niewiadomski, S. Hyniewska, C. Pelachaud (2011), Constraint-Based Model for Synthesis of Multimodal Sequential Expressions of Emotions, *IEEE Transactions of Affective Computing*, vol. 2, no. 3, 134-146, Juillet 2011. A. Osherenko, E. Andre, T. Vogt (2009), Affect sensing in speech: Studying fusion of linguistic and acoustic features, *International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009 C. Rich, C. L. Sidner (2012), Using Collaborative Discourse Theory to Partially Automate Dialogue Tree Authoring. *IWA 2012*: 327-340 M. Schröder, E. Bevacqua, R. Cowie, F. Eyben, H. Gunes, D. Heylen, M.ter Maat, G. McKeown, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M.I Valstar, and M. Wöllmer (2011), Building Autonomous Sensitive Artificial Listeners, *IEEE Transactions of Affective Computing*, pp. 134-146, Octobre 2011. P. Suignard, (2010) *NaviQuest: un outil pour naviguer dans une base de questions posées à un Agent Conversationnel*, WACA, Octobre 2010