



HAL
open science

AgentSlang : une Plateforme pour Systèmes Interactifs Distribués Multi-Incarnation

Sébastien Adam, Sandratra Rasendrasoa, Maël Bouabdelli, Alexandre Pauchet, Julien Saunier

► **To cite this version:**

Sébastien Adam, Sandratra Rasendrasoa, Maël Bouabdelli, Alexandre Pauchet, Julien Saunier. AgentSlang : une Plateforme pour Systèmes Interactifs Distribués Multi-Incarnation. WACAI 2020 - Workshop sur les Affects, Compagnons artificiels et Interactions, Jun 2020, Saint Pierre d'Oléron, France. 10.5220/0004907305960603 . hal-02933465v2

HAL Id: hal-02933465

<https://inria.hal.science/hal-02933465v2>

Submitted on 25 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AgentSlang : une Plateforme pour Systèmes Interactifs Distribués Multi-Incarnation

Sandratra Rasendrasoa² Maël Bouabdelli¹
Alexandre Pauchet¹ Julien Saunier¹ Sebastien Adam²

1-prénom.nom@insa-rouen.fr

2-prénom.nom@univ-rouen.fr

Normandie Université, INSA Rouen-Normandie, Université de Rouen-Normandie, LITIS

ABSTRACT

Les systèmes interactifs incarnés, virtuels ou robotiques, sont de plus en plus riches en terme de capacités dialogiques et affectives, ainsi que de comportements multimodaux. Cependant, cette richesse rend difficile la conception de nouveaux systèmes. Dans ce cadre, une approche par composants permet d'une part de réutiliser ceux-ci dans différents contextes et d'autre part, de répartir les traitements inhérents à l'interaction. Dans cet article, une démonstration des nouvelles fonctionnalités de la plateforme AgentSlang est présentée. Les améliorations apportées sont principalement de deux ordres : (1) l'intégration de mécanismes de dialogue fondés sur des modèles issus de l'apprentissage automatique, et (2) le support des agents robotiques. Ces ajouts permettent d'améliorer le potentiel de prototypage rapide d'AgentSlang, notamment pour étudier un modèle d'interaction au travers de plusieurs types d'incarnations.

KEYWORDS

Intelligence artificielle, interaction humain-agent, interaction humain-robot, système d'interaction distribué

1 INTRODUCTION

L'interaction entre humain et compagnon artificiel nécessite de nombreuses fonctionnalités : reconnaissance vocale, traitement du langage, gestion du dialogue, génération de réponses, etc. Chacune de ces tâches est difficile mais peut être traitée séparément, ce qui implique de développer et réutiliser des briques logicielles. Afin de pouvoir réutiliser ces différentes briques fonctionnelles, il est nécessaire de faire abstraction au maximum de leurs architectures logicielles. Une approche par composants répond à ces problématiques. Par ailleurs, cette approche permet également de répartir aisément la puissance de calcul, si besoin entre plusieurs processeurs. L'objectif de ces travaux est ainsi de proposer un système interactif capable de traiter de bout en bout une interaction humain-compagnon incarné, quelle que soit cette incarnation (robot ou agent virtuel).

2 LA PLATEFORME AGENTSLANG

AgentSlang¹ [10] est une plateforme permettant de développer rapidement des Systèmes Interactifs Distribués [10]. Conçue dans le respect d'une approche orientée composant, la structure modulaire d'AgentSlang permet d'assurer l'intégration des différents services internes au système, via des composants réutilisables dédiés à chaque fonctionnalité du système.

D'autres solutions ont été proposées pour développer des systèmes interactifs. En particulier, ICT Virtual Human Toolkit (VHT)² [5] et GRETA³ [8] sont deux systèmes interactifs dédiés à la conception d'agents conversationnels animés. VHT propose une collection de composants pour la majorité des tâches nécessaires aux systèmes interactifs, telles que la reconnaissance et la synthèse vocale, la gestion du dialogue et des comportements non verbaux. GRETA propose une architecture développée en Java permettant d'interagir avec un personnage virtuel du même nom. [8]. D'une part, VHT est la plateforme la plus aboutie, mais son code fermé rend son utilisation contraignante et ne permet pas de modifier ses composants. D'autre part, GRETA est plus orientée sur l'étude des agents virtuels, en particulier dans leur composante d'animation.

AgentSlang propose ainsi une alternative open-source et indépendante de l'incarnation à ces deux plateformes. Une liste non-exhaustive de ses fonctionnalités est présentée ci-après :

- Afin de traiter les entrées utilisateurs, plusieurs modules sont disponibles : (1) un module de transcription vocale automatique via l'API Google Speech ; (2) des composants permettant le traitement du langage naturel via Senna [2], un système versatile permettant notamment l'annotation grammaticale du discours et la reconnaissance d'entités nommées.

- Pour la gestion du dialogue, la plateforme propose des composants réalisés à l'aide de Syn!bad [9], un langage d'expressions rationnelles basé sur des synonymes. Le modèle de dialogue proposé suit une approche déterministe : une correspondance est faite entre l'entrée utilisateur et un modèle réalisé via Syn!bad ; une réponse est ensuite générée en fonction du motif identifié.

- La plateforme permet aussi de traiter l'aspect émotionnel de l'interaction, via un composant de détection de la valence émotionnelle du texte [10].

- Il est enfin possible de générer des comportements et les réponses à une requête utilisateur, par l'intermédiaire d'un agent virtuel, via des composants : (1) qui traitent la synthèse vocale en employant par exemple le logiciel Cereproc Voice⁴, et (2) qui génèrent et transmettent un fichier BML (Behaviour Markup Language) vers un outil de représentation du comportement compatible avec ce format tel que MARC [3] ou GRETA [8].

3 NOUVEAUX COMPOSANTS

De nouveaux composants ont été intégrés à la plateforme de façon à (1) intégrer une nouvelle approche concernant les mécanismes de dialogue, (2) assurer l'interfaçage avec des agents robotiques.

²<https://vhtoolkit.ict.usc.edu/>

³<https://github.com/isir/greta>

⁴<https://www.cereproc.com/fr/products/voices>

¹<https://github.com/AgentSlang/agentslang.github.io>

3.1 Mécanismes de dialogue

Trois fonctionnalités ont été intégrées à la plateforme pour enrichir la gestion du dialogue : (1) Un composant pour la détection automatique de l'interlocuteur dans un contexte multi-partie a été proposé par Usman Malik et al. [6]. (2) Un composant d'annotation automatique des actes dialogiques (AD) a également été ajouté [7] pour faciliter l'interprétation des entrées utilisateurs. (3) Un composant fondé sur le système DrQA [1] et le chatbot proposé par [4] a été intégré pour traiter deux cas d'usage : le dialogue ouvert et un système de question réponse à partir de transcriptions vocales de l'utilisateur.

Ces nouveaux composants ont l'avantage d'avoir une implémentation dans le même langage de programmation, ce qui a motivé notre choix de les intégrer à la plateforme. Ils facilitent l'extraction d'informations pertinentes et la génération du dialogue à partir des entrées utilisateurs, et permettent aux utilisateurs d'interagir sur des sujets plus divers avec les agents interfacés à la plateforme.

3.2 Gestion des robots

Les composants AgentSlang ont été initialement conçus pour de l'interaction humain-agent virtuel. Des composants nécessaires à l'intégration de la communication avec des robots ont été développés. Du fait que ROS⁵ est le middleware le plus utilisé en robotique, des noeuds ROS ont été développés pour gérer la communication avec AgentSlang.

Les communications entre composants AgentSlang s'effectuent via la bibliothèque de messagerie asynchrone ZeroMQ⁶. Cette bibliothèque haut niveau a été comparée avec d'autres solutions et présente deux avantages : elle est plus efficace, et elle supporte plusieurs types de sockets [10]. De plus, cette bibliothèque dispose d'implémentations dans la majorité des langages de programmation, ce qui permet de communiquer entre différents programmes tout en faisant abstraction de leur langage. Pour les motifs évoqués plus haut mais aussi pour des raisons d'uniformité, cette API est utilisée pour les communications entre AgentSlang et ROS, à travers un mécanisme de publication/abonnement. Afin de ne pas disséminer la localisation des composants entre noeuds ROS et composants AgentSlang, ainsi que de permettre une réutilisation dans le cas d'incarnations par des agents virtuels, l'approche de conception centre les tâches de traitement dialogique génériques au sein d'AgentSlang. Cette architecture est présentée sur la figure 1. Les applications agent virtuel ou robot disposent d'un centre décisionnel commun et seule la partie émission et réception des informations varie selon l'incarnation.

Le support des incarnations robotiques et la possibilité d'interagir sur des sujets plus variés ont permis la réalisation d'une première application de nos travaux : un dialogue ouvert entre un robot et un interlocuteur humain via la plateforme.

4 CONCLUSION

Dans cet article de démonstration, nous présentons les évolutions récentes de la plateforme AgentSlang, qui enrichissent la gestion du dialogue grâce à des composants de détection automatique de l'interlocuteur, d'annotation des AD et de deux modèles de dialogue.

De plus, il est désormais possible d'effectuer aisément des interactions à l'aide de robots par réutilisation des composants décisionnels utilisés dans le cas d'une interaction humain-agent virtuel. Ces nouvelles fonctionnalités facilitent la réalisation d'expérimentations. Dans un premier temps, il serait intéressant de comparer les nouveaux modèles de langue avec ceux initialement inclus. Dans un second temps, l'impact de l'incarnation sur l'interaction humain-agent selon différents contextes pourra être mesuré.

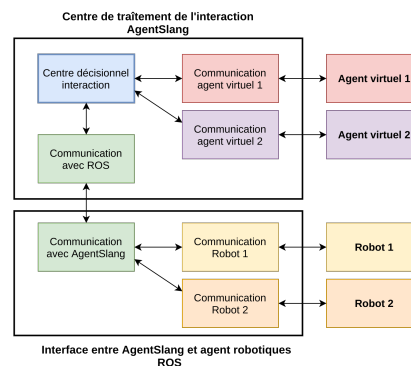


Figure 1: Schéma simplifié qui présente le support multi-incarnation de AgentSlang

REMERCIEMENTS

Ces recherches ont été financées par l'Union Européenne dans le cadre du projet INCA, convention FEDER numéro 19E00858.

REFERENCES

- [1] Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. 2017. Reading wikipedia to answer open-domain questions. (2017), 1870–1879.
- [2] Roman Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. 2011. Natural language processing (almost) from scratch. *Journal of machine learning research* 12, Aug (2011), 2493–2537.
- [3] Matthieu Courgeon, Jean-Claude Martin, and Christian Jacquemin. 2008. User's gestural exploration of different virtual agents' expressive profiles. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3*. International Foundation for Autonomous Agents and Multiagent Systems, 1237–1240.
- [4] Braden Hancock, Antoine Bordes, Pierre-Emmanuel Mazare, and Jason Weston. 2019. Learning from Dialogue after Deployment: Feed Yourself, Chatbot! (2019), 3667–3684.
- [5] Arno Hartholt, David Traum, Stacy C Marsella, Ari Shapiro, Giota Stratou, Anton Leuski, Louis-Philippe Morency, and Jonathan Gratch. 2013. All together now. In *International Workshop on Intelligent Virtual Agents*. Springer, 368–381.
- [6] Usman Malik, Mukesh Barange, Naser Ghannad, Julien Saunier, and Alexandre Pauchet. 2019. A Generic Machine Learning Based Approach for Addressee Detection In Multiparty Interaction. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. 119–126.
- [7] Usman Malik, Mukesh Barange, Julien Saunier, and Alexandre Pauchet. 2018. Performance Comparison of Machine Learning Models Trained on Manual vs ASR Transcriptions for Dialogue Act Annotation. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 1013–1017.
- [8] Radoslaw Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud. 2009. Greta: an interactive expressive ECA system. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. 1399–1400.
- [9] Ovidiu Șerban. 2014. Syn! Bad: A Synonym-based Regular Expression Extension For Knowledge Extraction Task. *Studia Universitatis Babeș-Bolyai, Informatica* 59, 1 (2014).
- [10] Ovidiu Șerban and Alexandre Pauchet. 2014. AgentSlang: A New Distributed Interactive System - Current Approaches and Performance. In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART, INSTICC, SciTePress*, 596–603. <https://doi.org/10.5220/0004907305960603>

⁵<https://www.ros.org/>

⁶<https://zeromq.org/>