From Thalamus to Skene: High-level behaviour planning and managing for mixed-reality characters

Tiago Ribeiro, André Pereira, Eugenio di Tullio, Patrícia Alves-Oliveira, and Ana Paiva

INESC-ID & Instituto Superior Técnico, Universidade de Lisboa

1 Main Research Themes

This paper presents work developed by the authors during the last three years on interactive scenarios featuring expressive robots that embody intelligent virtual agents (IVAs). Our work has focused on understanding how a robotic character's behaviour can be modelled and expressed in a body-independent form while providing users with continuous interactions. It has contributed to both the LIREC[1] and current EMOTE[2] EU FP7 Projects.

1.1 LIREC - Thalamus and EMYS-Risk

During the LIREC project, we developed a socially interactive scenario in which the EMYS robot plays the Risk boardgame with three human players (Pereira et al., 2014). The Risk game was implemented as a multimedia application on a touch-table. EMYS takes into account the game state, history of interactions and perception of the environment in order to be perceived as a socially present artificial opponent. Its behaviour is managed by our own EmysToolkit software and an initial version of Thalamus, which handled the scheduling and synchronization of BML speech and face actions. The system was inspired by SAIBA (Kopp et al., 2006), using expressive utterances to model the intention of the robot. These were composed of dialogue acts annotated with animation instructions, and triggered from EMYS's artificial intelligence.

1.2 BML and Perception

While acting in a real-world setting, the behaviour of a character depends on external sources (perception) in order to pose correctly towards the users. We ran a case-study with two robots that interacted by using BML while perceiving the external environment (Ribeiro et al., 2012). The study showed how the imperfection of the robot's sensors makes it risky to schedule BML events and expect regular consistency.

[1] `www.lirec.eu`

[2] `www.emote-project.eu`

1.3 EMOTE - Thalamus as a Censys system

The EMOTE project aims at developing empathic robotic tutors that can interact with school children through multimedia applications (scenarios). Our requirements for the IVA architecture are now for it to be independent of the type of robot used and to allow to reuse components in different scenarios. We also felt the need for a flexible way of structuring the mind and the body of the agent, especially because of the different types of physical devices that our agent is composed of (robot, external sensors, touch-device).

With this in mind, we created the Censys model of agents, which proposes that there is no need to explicitly define a Mind or a Body in an agent. These components can actually be built out of several interacting processes, which exchange information (Ribeiro et al., 2013). Thalamus was therefore restructured in order to become a modular component-integration framework aimed at IVAs (Ribeiro et al., 2014).

1.4 EMOTE - Skene

Skene is a semi-autonomous behaviour planner that is being developed in the EMOTE project. It is designed to be reused in different applications and embodiments, and in SAIBA fits as a behaviour planner. As we described previously, our interaction environment consists of both robots and other devices (e.g., multi-touch table (MTT)). Skene is the component in which all of them meet. Its input is a high-level behaviour description language (Skene Utterances) and perception information (such as target locations). Its output consists both of scheduled BML and non-BML actions (like sounds or application commands).

Skene Utterances essentially drive some state machines and behaviour generation mechanisms. Such functionality was inspired by the expressive utterances previously used on EMYS but now include instructions for Gazing, Pointing, Waving, Animating, Sound, Head-Nodding and Application instructions. The following is an example of a Skene Utterance:

<GAZE(*student*)> <WAVE(*throughMap*)>

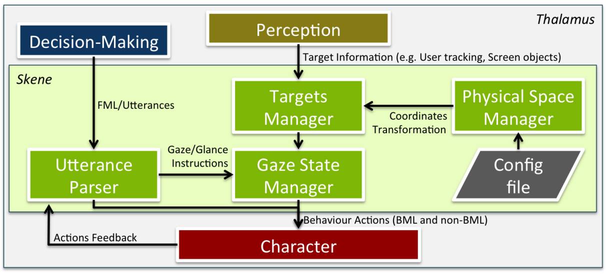

Figure 1: Skene's components.

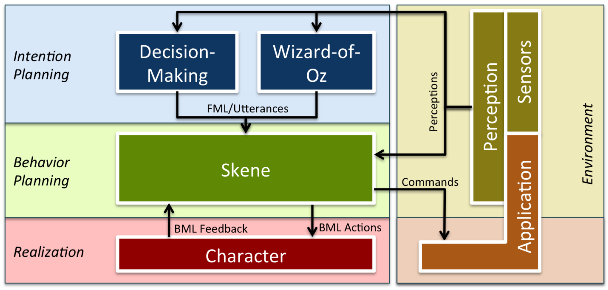

Figure 2: A general structure of our SAIBA-based IVAs.

`Every map has a` **<POINT(*scale*)>** `scale that` **<HEADNOD(*1*)>** `can help you. Some maps represent` **<ANIMATE(*metaphWorld*)>** `a country, while others represent smaller areas like this one` **<ANIMATE(*metaphDichotomicRight*)>.**

The language was developed alongside with non-technical partners from psychology working in EMOTE who provided us with results of human-human interaction analysis to inform on the design of the agent's behaviour (Alves-Oliveira et al., 2014). It is being designed to be simple, human-readable and to provide a link between the psychologists' annotation process and the character's behaviour design.

Skene includes some semi-automated behaviour, such as a gaze-state machine that can keep the character following specific targets while providing it with additional behaviour (e.g gaze-aversion). This gaze-state machine is being developed based on annotations from human-human interactions combined with existing literature (Andrist et al., 2014).

By considering a representation of the Environment, Skene is not bounded to a specific physical set-up. The Physical Space Manager (PSM) keeps a model of the environment surrounding the character. Pointing, gazing or waving all require to know the targets' physical position in relation to the character. The PSM is configured with the resolution, dimensions and position of the MTT relative to the robot. The application informs it about the screen coordinates of relevant GUI and game elements; these coordinates can then be converted to angles to be used by expressive behaviours like Gazing or Pointing.

2 Current Architectures and Standards

The systems we described have followed on SAIBA, with some refinements as we illustrate in Figure 2. We have used both Wizard-of-Oz and autonomous modules to generate and feed Utterances to Skene. These Utterances drive the selection and automated generation of behaviour, both regarding the character's expressiveness and its actions in some virtual environment. It also considers the perception from both the virtual and the real environment in order to keep the behaviour consistent with the surroundings. Our main addition to the SAIBA model is the consideration of the interaction environment as part of the whole architecture. Our architecture is being used in two EMOTE scenarios, in a collaboration with EPFL[3], and on some MSc projects at IST.

3 Future Architectures and Standards for IVAs

Our vision is that a successful IVA architecture will allow us to take components from one IVA, and reuse them in another IVA and application. That requires a standardization of interfaces through which the components communicate and the definition of roles/functions of some components. We have been using Thalamus as the backbone for our robotic IVAs, which by following the Censys model, gives us flexibility to separate components and reuse them across applications. Although neither Censys nor Thalamus are architectures, inspiration from Censys might bring forward some of the flexibility we envision to other SAIBA-based systems and the future of IVAs. We also believe that the representation and connection with the surrounding environment must be considered by such an architecture, as the future will bring mixed-reality IVAs that interact with the users through both through some physical form and multimedia applications.

4 Suggestions for Discussion

The following list contains the topics we would like to discuss during this workshop:

- PML and FML syntax;
- How robots and environments fit into SAIBA;
- BML on non-anthropomorphic characters;
- Designing and developing scenarios with non-technical partners (e.g. from psychology and arts).

[3] École polytechnique fédérale de Lausanne

References

P. Alves-Oliveira, S. Janarthanam, A. M. Candeias, A. Deshmukh, T. Ribeiro, H. Hastie, A. Paiva, and R. Aylett. 2014. Towards Dialogue Dimensions for a Robotic Tutor in Collaborative Learning Scenarios. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*.

S. Andrist, X. Z. Tan, M. Gleicher, and B. Mutlu. 2014. Conversational Gaze Aversion for Human-like Robots. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 25–32.

S. Kopp, B. Krenn, and S. Marsella. 2006. Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent Virtual Agents*, pages 205–217.

A. Pereira, R. Prada, and A. Paiva. 2014. Improving social presence in human-agent interaction. In *Proceedings of the 32nd ACM Conference on Human Factors in Computing Systems*, pages 1449–1458. ACM.

T. Ribeiro, M. Vala, and A. Paiva. 2012. Thalamus: Closing the mind-body loop in interactive embodied characters. In *12nd International Conference on Intelligent Virtual Agents*, pages 189–195.

T. Ribeiro, M. Vala, and A. Paiva. 2013. Censys: A Model for Distributed Embodied Cognition. In *13th International Conference on Intelligent Virtual Agents*, pages 58–67.

T. Ribeiro, E. D. Tullio, L. J. Corrigan, A. Jones, F. Papadopoulos, R. Aylett, G. Castellano, and A. Paiva. 2014. Developing Interactive Embodied Characters using the Thalamus Framework: A Collaborative Approach. In *14th International Conference on Intelligent Virtual Agents*.

Biographical Sketch

Tiago Ribeiro is an eclectic researcher seeking harmony between arts and technology. He has been collaborating internationally on research projects like LIREC and EMOTE, and also with CMU, focusing especially on animation of expressive robots. He is currently in an early stage of obtaining his PhD, in which he pursues artist-oriented intelligent robot animation.

André Pereira is a multidisciplinary researcher focused in the design, implementation and evaluation of social robots. His research is mainly focused in studying how to create socially present board game opponents.

Eugenio Di Tullio obtained his MSc in Computer Science at Universit degli Studi di Bari Aldo Moro, University of Bari (Italy) in 2012. Since then he worked at INESC-ID on the Machinima Project and joined the Emote project in January 2014.

Patrícia Alves-Oliveira is graduated in Clinical and Health Psychology by Instituto Superior de Psicologia Aplicada and Universidade de Aveiro and has done work in the area of human sexuality and emotion with an evolutionary point of view. She is currently a research assistant in INESC-ID at Instituto Superior Técnico, Technical University of Lisbon, working in human-robot interaction. Her interests regard the improvement of human-robot interaction, enabling a balanced future co-existence.

Ana Paiva is a research group leader of GAIPS at INESC-ID and an Associate Professor at Instituto Superior Técnico, Technical University of Lisbon. Her research is focused on the affective elements in the interactions between users and computers with concrete applications in robotics, education and entertainment.

Acknowledgments

This work was partially supported by the European Commission (EC) and was funded by the EU FP7 ICT-317923 project EMOTE and partially supported by national funds through FCT - Fundação para a Ciência e a Tecnologia, under the project PEst-OE/EEI/LA0021/2013. The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.