Live Generation of Interactive Non-Verbal Behaviours (Extended Abstract)

Ken Prepin
LTCI/TSI, Telecom-ParisTech/CNRS,
37-39 rue Dareau,
75014, Paris, France
ken.prepin@telecom-paristech.fr

Catherine Pelachaud
LTCI/TSI, Telecom-ParisTech/CNRS,
37-39 rue Dareau,
75014, Paris, France
catherine.pelachaud@telecom-paristech.fr

ABSTRACT

Psychology, robotic and virtual agents communities commonly claim that to enable natural interaction to take place within dyad of agents, the dyad must be the siege of dynamical coupling: to give to their partners a feeling of natural interaction, interactants be human, robotic or virtual, must be able to make the dynamic of their behaviour emerge both from their own internal states and from their partner's behaviours. However, most virtual agents engines for interactions model communication as a step by step phenomenon, where pre-scripted signals and corresponding feedbacks alternate. We propose here an agent architecture which generates non-verbal behaviours in live, influenced by both the internal state of the agent and the continuously incoming reaction of its partner. This architecture enables an agent facing either another agent or a human, to emphasise shared behaviours (called *Snowball effect*), to decrease un-shared behaviours as well as to align dynamically with its partner's behaviour.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems

General Terms

Theory

Keywords

Human-robot/agent interaction, Peer to peer coordination, Emergent behavior, Modeling the dynamics of MAS, Agent commitments

1. MODEL PRINCIPLES

During a dyadic interaction, partners' behaviours are influenced by both their internal state and the continuously incoming reactions of their partner. When a behaviour is triggered (a smile, a head-nod) how its dynamics will develop through time is not defined ahead of time: decay, emphasis or complete change of the behaviour will depend on the course of the interaction, influenced by the live partner's reactions.

In virtual agent systems, such a dynamical coupling capability is still lacking even though it is necessary for the occurrence of natural interactions [9]. Implementing coupling capabilities require to deal with dynamic adaptation of agent animation in real-time [3, 4, 6].

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

The model of live generation of interactive non-verbal behaviours that we propose, is implemented in a Neural Network (NN) simulator (Leto/Prometheus (developed in the ETIS lab. by Gaussier et al. [2])) interfaced with a virtual agent engine (the listener agent Greta and its backchannels engine proposed by Bevacqua and Pelachaud [1] and implemented in the SEMAINE platform for Sensitive Artificial Listners (SAL) [8]).

We propose here an agent architecture which generates non-verbal behaviours (head movements and multimodal sequential expressions) *on the fly*, influenced by both the internal state of the agent and the continuously incoming reaction of its partner. The resulting behaviour of a dyad of agents having such an architecture is a *snowball effect* on shared behaviours (when coupling occurs), a decay of not-shared behaviours (when coupling is disrupted), the ability for the two agents involved in the interaction to evaluate their partners engagement by detecting *snowball effects*.

Our model relies on the three properties of every natural communication described bellow:

P1 - Interaction feedbacks modify the course of actions *on the fly*. If the feedbacks from interaction partner are not fast enough regarding the action length, coupling and synchrony cannot occur between partners [6]. In our architecture, the agent which performs the action can have feedbacks concerning this action while s/he is performing the action: the action is commonly built by agent's intentions and partner's feedbacks.

P2 - Perception-Action mapping. There is a natural/structural tendency to imitate the other and to better perceive the other when s/he imitates you.
This mapping has two components: first a default mirror mapping; second a mapping between different actions (cf. *backchannels* [3]).

P3 - Action are built as interpolated sequences of basic elements. Niewiadomski et al. [5] indicates that to be able to convey subtle emotional states, Multimodal Sequential Expressions are needed.

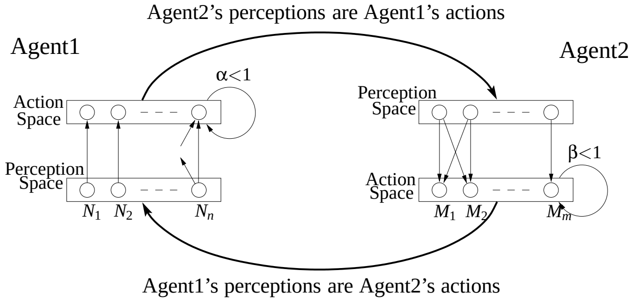

Figure 1: Scheme of the interactive loop within the dyad.

A *Snowball Effect* on shared actions and a decrease of non-shared actions result from P1 and P2 (see Fig.1 for a principle scheme of a dyad of agents implementing P1 and P2).

2. SNOWBALL EFFECT

We assume here and model the fact that during dialogue, sequences of emotional signals (P3) are induced by the coupling and the mutual reinforcement occurring between agents (P1 and P2). Let us consider the example of polite vs friendly smile: if you smile politely to somebody who smiles back at you but more friendly, without interruption, your smile could evolve in a friendly smile as well (see Fig.2).

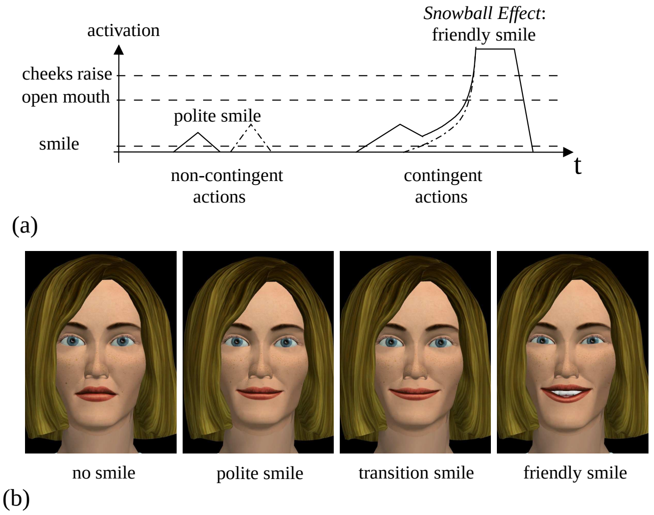

Figure 2: Snowball effect of shared actions. (a) Example of dynamics which would be obtained for smile depending on agents'action contingency. Solid line is for Agent1's activation of smile, dotted line is for Agent2's (b) Politeness smile to amusement smile transition on one of the two agents: frames are generated *on the fly*.

We model P3 enabling agents to recruit a sequence of basic facial signals *on the fly*, depending on the emphasize of their own actions induced by P1 and P2: in our example when facial signals corresponding to a *polite smile* are emphasized enough, new signals, corresponding to *friendly smile* are triggered.
The sequence of emotional signals is not pre-defined but is mainly induced by the mutual reinforcement between agents: the mutual reinforcement occurring when agents perform contingent actions, leads to a specific shared sequence of actions(see Fig.2(a)).

Fig.2(b) shows what the animation from polite to amusement smile looks like with our architecture for live generation of behaviour: the frames are computed *on the fly*, they are structurally building a single continuous behaviour.

3. CONCLUSION

The resulting behaviour of a dyad of agents having such an architecture (i.e. implementing of the three properties P1, interaction feedbacks modify the course of actions *on the fly*; P2, perception-action mapping; P3, action are built as sequences of basic elements) is a *snowball effect* on shared behaviours and a decrease of non-shared behaviours: perception directly influences action; actions last long enough to enable several perceptions to influence them.

This behaviour of the dyad has a direct impact on the interactants' sensitivity to their rapport: *snowball effect* occurs only if interactants share a common state and if they are both aware of each other (cross-perception). That makes the occurrence of the snowball effect a marker of the dyadic state and particularly of partners contingency [7]. To know if whether or not they have an effective interaction, agents can directly observe their own sequential expressions dynamics: novelty and complex sequences is equivalent to high coupling and exchange between partners.

At short term, our aim is to test this architecture between an agent and a human. Snowball effect would be enabled giving to interactants a cue of their level of contingency.

Demos of the *snowball effect* obtained with our architecture an be seen on:
`http://www.tsi.telecom-paristech.fr/mm/?p=778`.

Acknowledgements

This work has been partially financed by the European Project NoE SSPNet, the project SYNC from Institut Télécom and the European Project VERVE. Nothing could have been done without the Leto/Prometheus NN simulator, lend by the Philippe Gaussier's team (ETIS lab, Cergy-Pontoise, France).

4. REFERENCES

[1] E. Bevacqua, S. Hyniewska, and C. Pelachaud. Positive influence of smile backchannels in ecas. In *International Workshop on Interacting with ECAs as Virtual Characters (AAMAS 2010)*, Toronto, Canada, Oct. 2010.

[2] P. Gaussier and S. Zrehen. Avoiding the world model trap: An acting robot does not need to be so smart! *Journal of Robotics and Computer-Integrated Manufacturing*, 11(4):279–286, 1994.

[3] J. Gratch, A. Okhmatovskaia, F. Lamothe, S. Marsella, M. Morales, R. J. Werf, and L.-P. Morency. Virtual rapport. In *Proceedings of 6th International Conference on Intelligent Virtual Agents:IVA'06*, pages 14–27, Sept 2006.

[4] G. Jonsdottir, K. Thorisson, and E. Nivel. Learning smooth, human-like turntaking in realtime dialogue. In H. Prendinger, J. Lester, and M. Ishizuka, editors, *Intelligent Virtual Agents*, volume 5208 of *Lecture Notes in Computer Science*, pages 162–175. Springer Berlin / Heidelberg, 2008.

[5] R. Niewiadomski, S. Hyniewska, and C. Pelachaud. Constraint-based model for synthesis of multimodal sequential expressions of emotions. *IEEE Transactions of Affective Computing*, 2(99):1–13, 2011.

[6] K. Prepin and C. Pelachaud. Effect of time delays on agents' interaction dynamics. In Tumer, Yolum, Sonenberg, and Stone, editors, *The Tenth International Conference on Autonomous Agents and Multi Agents Systems, AAMAS2011*, pages 1–8. IFAAMAS, 2011.

[7] K. Prepin and C. Pelachaud. Shared understanding and synchrony emergence: Synchrony as an indice of the exchange of meaning between dialog partners. In J. Filipe, editor, *Third International Conference on Agents and Artificial Intelligence, ICAART2011*, pages 1–10. Springer, 2011.

[8] M. Schröder. The semaine api: Towards a standards-based framework for building emotion-oriented systems. *Advances in Human-Computer Interaction*, 2010:21, 2010.

[9] F. J. Varela. *Invitation aux sciences cognitives*. Editions du Seuil, Paris, 1995.