

# How Was Your Day? A Companion ECA

Marc Cavazza  
School of Computing  
Teesside University  
Middlesbrough TS1 3BA  
[m.o.cavazza@tees.ac.uk](mailto:m.o.cavazza@tees.ac.uk)

Raul Santos de la Camara  
Telefonica I+D  
C / Emilio Vargas 6  
28043 Madrid  
[e.rsai@tid.es](mailto:e.rsai@tid.es)

Markku Turunen  
University of Tampere  
Kanslerinrinne 1  
FI-33014  
[mturunen@cs.uta.fi](mailto:mturunen@cs.uta.fi)

## ABSTRACT

We demonstrate a “Companion” ECA, which is able to provide advice and support to the user, taking into account emotions expressed by her through dialogue. The integration of all required multimodal I/O components is based on interaction strategies defining the shape of dialogue, on the ECA’s response times, and on the underlying affective strategy. The system supports free conversation on an everyday life scenario in which the user comments her day at the office.

## Categories and Subject Descriptors

H.5.2 [User Interfaces] (D.2.2, H.1.2, I.3.6)

## General Terms

Experimentation, Human Factors.

## Keywords

Embodied Conversational Agents, Affective Interfaces, Human-Computer Dialogue.

## 1. INTRODUCTION

Embodied Conversational Agents have often been associated to personal assistants, helping the user to access information by means of task-oriented dialogue. Following recent developments in affective dialogue systems [1], a new breed of ECA can be envisioned, whose primary purpose is indeed to hold conversations with the users, and the relationship they can establish. These ECA, which can play the role of Companions, still differ from chatterbots, in that they require an appropriate understanding of user utterances to hold conversations in which they can reassure, comfort or advise the user.

We are demonstrating the prototype of such an ECA with which the user can engage in free conversation, albeit on a limited set of topics. The scenario for the demonstration consists for the user to discuss her day at work, which we have named the “How Was Your Day?” application (henceforth HWYD). This constitutes an everyday life domain which nevertheless carries affective content. The system supports conversation rather than dialogue, breaking several conventions of human-computer dialogue: it allows lengthy user utterances for both the user and the agent, multiple dialogue turns from the user, as well as user interruptions during the ECA turns. Furthermore, it is strongly

based on user initiative, with the ECA behaving as a sympathetic, yet active, listener, responding by providing in advice or encouragement.

## 2. SYSTEM OVERVIEW AND ARCHITECTURE

The demonstration presents itself as an on-screen ECA with which the user engages in dialogue. The agent will acknowledge, occasionally asking questions about events, and provide advice or reassurance once it has gathered sufficient information. The system regularly supports dialogues of over 15 minutes in duration.

An overview of the system interface is presented in Figure 1: in addition to the ECA, this interface features a detailed trace allowing various forms of “glass box” demonstrations in which system behavior (e.g. user interruptions, barge-in) can be analysed in real time.

The HWYD Companion prototype integrates multiple software components using a blackboard-like architecture, the *Inamode* framework developed at Telefonica I+D, in which message transmission is carried out using XML formatted messages over plain text sockets. The prototype uses Nuance™ dictation system for speech recognition, with emotional speech recognition carried out in parallel using the *EmoVoice* system [2]. The ASR transcript is also analysed by an emotional text processing module and its output is fused with that of *EmoVoice*. The transcript is processed through an Information Extraction system which recognizes key events such as company restructuring, individual promotion, redundancies, meetings, changes in relationship between colleagues (20 in the current version, 40+ in the target demonstration). Event recognized are attached an emotional category corresponding to the utterance which introduced them. Overall dialogue is under the control of a rule-based Dialogue Model, which decides when to ask the user for further information about an event, and when to actually respond to the user giving comfort, warnings or advice.

These are decided and generated using a plan-based utterance generator which can produce long utterances based on an appraisal of the user’s situation. The ECA is developed around the Haptex™ toolkit, and is controlled using an FML-like language [3]. Text-to-Speech (TTS) is based on an emotional extension of Loquendo™’s TTS system, developed as part of this project.

**Cite as:** How Was Your Day? a Companion ECA, Marc Cavazza, Raul Santos de la Camara, Markku Turunen and the COMPANIONS Consortium *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1629-1630 Copyright © 2010, International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). All rights reserved.

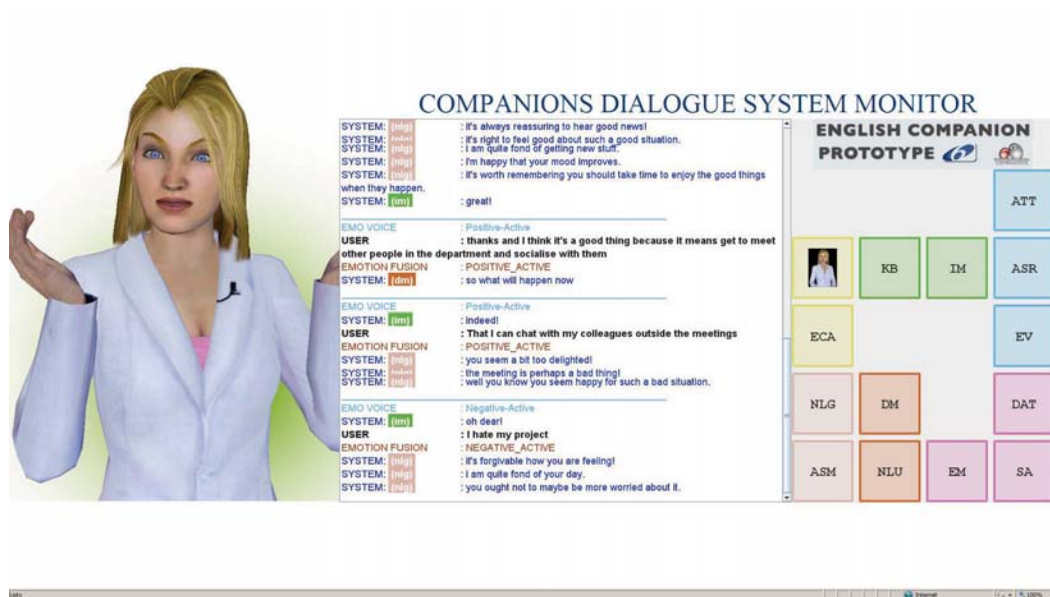


Figure 1: The HWYD Companion in operation: the central portion of the screen traces the dialogue, whilst the right-hand section displays which current software module is activated.

### 3. SYSTEM BEHAVIOUR

Interaction between the ECA and the user is based on the notion of “affective loops”, which control the timing of the agent response. The short loop corresponds to the “real-time” (< 500 ms) reaction of the ECA to the perceived user attitude and provides a backchannel for interaction. It consists in aligning its non verbal behavior (facial expressions and gestures) to the affective content of the latest user utterances, as well as acknowledging the user intervention with short expressions, on a random basis. User interruptions constitute a specific case in which ECA feedback has to be prompt and appropriate to be realistic. Although similar in essence to the short loop, the ECA response is under the control of a specific Interruption Manager. Typically an interruption will take place during a longer ECA utterance, often one attempting at providing comfort or advice. The IM interrupts instantaneously the ECA utterance, prompts the appropriate multimodal expression in a way similar to the short loop, and triggers the analysis of the interrupting utterance.

Finally, the main loop corresponds to the deliberative aspects of the ECA behaviour and is under the control of the Dialogue Model. It consists in constructing event representations from the user report, and once sufficient information has been gathered, in providing comfort or advice to the user. The actual content will depend on an appraisal of the user response to the events she reports and in under the control of the Affective Strategy Module (ASM). The ASM produces a plan to influence the user through several communication operators, each of which is translated into an ECA utterance, forming a short tirade (typically 4-5 utterances) providing advice to the user.

### 4. ADDITIONAL AUTHORS

This demo is presented on behalf of the whole COMPANIONS Consortium (<http://www.companions.org>).

### 5. ACKNOWLEDGMENTS

This work was funded by the Companions project sponsored by the European Commission as part of the Information Society Technologies (IST) programme under EC grant number IST-FP6-034434. The *EmoVoice* system has been used courtesy of the University of Augsburg (Multimedia Concepts and Applications).

### 6. REFERENCES

- [1] André, E., Dybkjær, L., Minker, W., and Heisterkamp, P. (Eds.): Affective Dialogue Systems, Tutorial and Research Workshop, ADS 2004, Kloster Irsee, Germany, June 14-16, 2004, Proceedings. Lecture Notes in Computer Science 3068 Springer 2004.
- [2] Vogt, T., André, E. and Bee, N. (2008). EmoVoice – A framework for online recognition of emotions from voice. In: Proceedings of Workshop on Perception and Interactive Technologies for Speech-Based Systems, Springer, Kloster Irsee, Germany, (June 2008).
- [3] Hernández, A., López, B., Pardo, D., Santos, R., Hernández, L., Relaño Gil, J. and Rodríguez, M.C. (2008) Modular definition of multimodal ECA communication acts to improve dialogue robustness and depth of intention. In: Heylen, D., Kopp, S., Marsella, S., Pelachaud, C., and Vilhjálmsson, H. (Eds.), AAMAS 2008 Workshop on Functional Markup Language.