A 'Companion' ECA with Planning and Activity Modelling (Short Paper)

Marc Cavazza, Cameron Smith, Daniel Charlton, Li Zhang
University of Teesside
School of Computing
Middlesbrough, United Kingdom
+44 1642 218121
{m.o.cavazza, c.g.smith, d.charlton, l.zhang}@tees.ac.uk

Markku Turunen, Jaakko Hakulinen
University of Tampere
Department of Computer Sciences
Tampere, Finland
+358 3 3551 8559
{Markku.Turunen, Jaakko.Hakulinen}@cs.uta.fi

ABSTRACT
In this paper, we describe the development of an Embodied Conversational Agent (ECA) implementing the concept of a companion, i.e. an agent supporting the persistent representation of user activities and dialogue-based communication with the user. This first experiment implements a Health and Fitness companion aimed at promoting a healthier lifestyle. The system operates by generating an 'ideal' plan of daily activities from background knowledge and dialogue interaction with the user. This plan then becomes an activity model, which will later be instantiated by reports from the user and analysed by the agent from the perspective of initial objectives. At various stages of the day, the plan can still be adapted through further dialogue. The agent is embodied using a wireless rabbit (Nabaztag™) device situated in the user's home. After describing the planning component, based on Hierarchical Task Networks (HTN) and the spoken dialogue system, we present a working example from the system illustrating its behaviour through various phases of user activity generation, updating and re-planning.

Categories and Subject Descriptors
H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems.

General Terms
Algorithms, Human Factors.

Keywords
Embodied Conversational Agents, Planning, Human-Computer Dialogue, Assistive Systems.

1. INTRODUCTION
The successful development of ECA opens the way for many new applications. Alongside training, education and entertainment applications, virtual advisors [4] and personal assistants [2] of all kinds have attracted considerable interest in recent years. A new paradigm for virtual assistants has emerged in the form of companions [21], defined by Forbus and Hinrichs [7] as being able to interact with users over sustained periods of time, while also possessing robust reasoning abilities.

Cite as: A 'Companion' ECA with Planning and Activity Modelling (Short Paper), Cavazza, M., Smith, C., Charlton, D., Zhang, L., Turunen, M., Hakulinen, J., *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. XXX-XXX.

Figure 1: The Nabaztag™ device

In this paper, we describe the development of a physically embodied Health and Fitness Companion (HFC), which aims at promoting healthier lifestyle for a typical user as office worker. It is a central feature of this application to operate in an anytime, persistent fashion both in terms of knowledge use and in terms of dialogue sessions. The user can decide to interact with the HFC to request specific advice but, in the long term, its main mode of operation should be to embed such advice inside more open conversation whose topics will be dictated by the context in which they take place (time of the day, user expected or intended activities).

The HFC is embodied using the Nabaztag™ device (Figure 1), a commercial wireless rabbit character [18] already recognised as one of the most successful ubiquitous computing devices in terms of consumer adoption and potential for applications.

2. RELATION TO PREVIOUS WORK
This work relates to previous research in several ways, both in terms of similar applications and through its underlying technical choices in planning and dialogue.

Several groups have described assistive systems for daily life or office work, although not all of them as ECA. The *Autominder* system [12] [14] is an autonomous mobile robot that can 'live' in the home of an older individual, and provide him or her with reminders about daily plans". The CALO project aims at developing a personal assistant helping an office worker to deal with information and task overload [13] [2]. The POLLy system

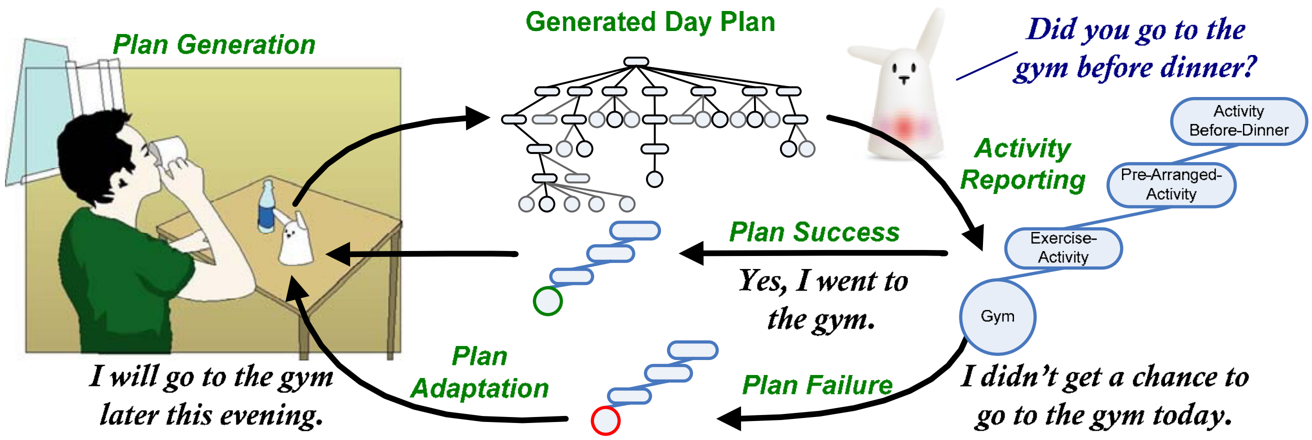

Figure 2: The various phases of interaction with the HFC: Plan generation, activity reporting, plan adaptation or replanning.

[10] has been developed to research politeness in the context of task-based interactions (more specifically, cooking). In this system, dialogue would however take place over plan execution.

Several dialogue systems have used plans as underlying knowledge models, in particular for the representation of joint user-system tasks since the original TRAINS project [5]. Similar approaches to decompositional planning as task representations or baseline plans have been described, for instance, in TRIPS' "straw plans" [6] or WITAS' "recipes" [9].

3. SYSTEM OVERVIEW

For this first prototype, we have devised an interaction scenario which assumes that the HFC is located at the user's home, and consequently the user will only interact with it during specific phases of his working day: in the morning before leaving for work and in the evening just after returning from work but before any further leisure activities. This in turn determines various phases for relating dialogue to planning (Figure 2), and for the nature of dialogue itself:

- plan generation: to plan the day's activities ahead (e.g. in the morning, sometimes leaving certain options open for later in the day).

- activity reporting: to report on activities which took place during the day to instantiate *a posteriori* the task model. This type of dialogue depends mostly on the user but has to be primed by relevant questions from the Nabaztag™.

- plan adaptation: to adapt a portion of the plan or re-plan an entire phase of the day before it takes place, depending on changing user conditions rather than on the outcome of previous phases

In line with the philosophy of a companion agent, we want to depart from task-related dialogue sessions during which the user would be systematically asked for required parameters, with the system leading dialogue and acknowledging all user input. More natural and asynchronous communication can be based on the fact that the agent possesses background knowledge on the user's preferences and her activities. For instance, when elaborating a plan for the user's daily activities, the system will only enquire about specific situations (e.g. the weather conditions or the user's mood) or the user's preferences.

4. SYSTEM ARCHITECTURE

The first HFC prototype, as presented in Figure 2, is implemented with a generic agent-based architecture designed for adaptive spoken dialogue systems [19]. It has been used in several spoken dialogue systems, including a multilingual spoken dialogue system [20]. In the HFC, this architecture is extended to support interaction with virtual and physical Companions.

Our architecture is based on distributed but coordinated components, shared system knowledge and a general system-level adaptation mechanism. The system architecture is distributed so that different managers and agents can run on different computers and platforms. It is similar to certain central components found in other speech architectures, such as the HUB in the Communicator architecture [16], and the Facilitator in the Open Agent Architecture [11].

4.1 Speech Input and Output

The Communication Manager handles all input and output management. It includes devices and engines that provide interfaces to technology components. Most importantly, in the HFC it includes components to control Loquendo™ ASR and TTS components and the physical agent interface. The system uses recognition grammars in "Speech Recognition Grammar Specification" (W3C) format that are dynamically selected by the Modality Manager according to the current dialogue state. Dynamic grammar generation also takes place in certain situations.

In the first prototype natural language understanding is based on the concept-spotting approach, using heavily "Semantic Interpretation for Speech Recognition (SISR) Version 1.0" (W3C) format information. Semantic information provided by the SISR tags is combined with the dialogue state to construct predicates compatible with the planning domain.

Natural language generation is implemented with a concept-based approach, mostly using templates. The main starting point is predicate-form task descriptions formed by the cognitive model. Further details and contextual information are retrieved from dialogue history, the user model, and potentially other sources. Finally, SSML (Speech Synthesis Markup Language) 1.0 tags are used for controlling the Loquendo™ synthesizer.

4.2 The Physical Agent Interface

For a physical agent interface, the jNabServer software was created to handle communication with the Nabaztag™. The Nabaztag™ device can handle various forms of interaction, from voice to touch (button press), and from RFID 'sniffing' to ear movements. It can respond by moving its ears, by displaying or changing the color of its four LED lights. It can also play sounds which can be music, synthesised speech or other voices.

4.3 Dialogue Management

The Dialogue Manager takes care of conversational strategies and communicates with the planner that generates the user activity model. Together, they use hierarchical task decomposition and a dialogue stack similar to CMU Agenda [15] and RavenClaw [3] systems. The dialogue manager maintains a dialogue history tree and communicates facts and user preferences to the planner at the various stages of plan elaboration and task instantiation (Figure 2). The planner (implemented in Allegro Common Lisp) is connected to the software architecture via the Cognitive Model Manager. The integration between the Planner and the dialogue system is based on a mapping between the dialogue lexicon semantics and the Planning domain, as presented in detail in Section 5.

5. ACTIVITY MODEL PLANNING

In order to fulfil his role as an assistant, the system generates a global plan corresponding to an ideal course of action for the user's daily activities. The system uses planning techniques to generate a reasoned top-down decomposition of user activities, implicitly ordered to follow the rhythm of a normal day itself. The central idea of our approach is that this plan in turn becomes a task model representing potential user activities which will be instantiated by user reports.

Establishing the plan consists in generating user activities in a way which maximizes energy expenditure and minimises food intake, within the boundaries of normal activities. There is an implicit agreement that the user will actually follow the plan for the 'standard' part of her activities, and for those actions explicitly discussed with the Nabaztag™.

5.1 Plan Generation

We use Hierarchical Task Network Planning with a total-order forward decomposition algorithm [8], which has been specifically extended to incorporate semantic knowledge in the decomposition process. That is to say, when there are multiple applicable methods, selection of the most appropriate method is based on a heuristic approach that uses semantic categorisation.

To illustrate this, we look at the high level task of travelling to work from home ('Medium-Distance-Travel Home Work', which is part of the Plan-Day domain). The task can be decomposed into eight different options depending on how the user will travel to work. In terms of the AND/OR tree, this involves a root node holding the 'Medium-Distance Travel Home Work' task with an OR-branch node holding five task nodes (see Figure 3).

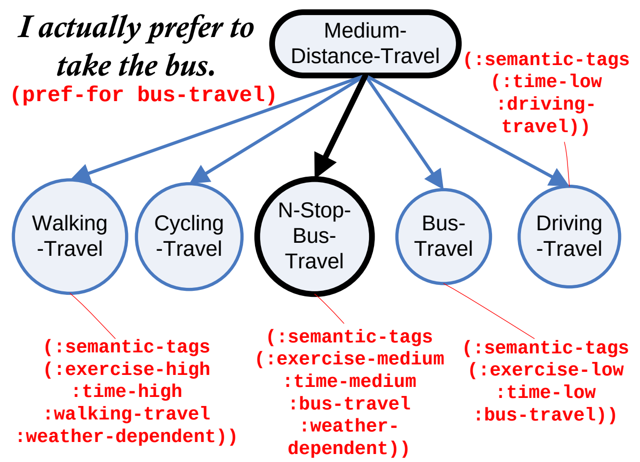

Figure 3: Plan Generation with Semantic Knowledge

The various options outline various ways of getting to work including walking or taking the bus. (The 'n Stop' option for bus indicates getting off a couple of stops early, so as to get more exercise.) The definition of each in the domain includes semantic tags which are used to contrast the differing properties of each option. These semantic descriptions correspond to domain knowledge which should be activated from dialogue.

The initial state of the planner (in Figure 3) contains a recent preference for bus travel generated from dialogue with the user. This preference ensures that Medium-Distance-Travel-n-Stop-Bus scores higher than the other Medium-Distance-Travel options and thus this task is selected to decompose further.

5.2 Activity Reporting

Once a plan has been generated it becomes a task model for user activities and rests on the assumption that the user will generally follow the plan, with however potential for departing from it. It is thus necessary to update the task from the user herself at different stages, for instance when the user returns from work, following the cycle of interaction described on Figure 2. This is done by traversing the AND/OR graph defining the plan and marking task nodes as completed or failed based on the information available (although strictly speaking, this is not a case of the plan "failing" as it is only used as a resource).

5.3 Plan Adaptation

Plan adaptation consists in surface modifications to the planned activities [17]: from a task decomposition perspective, adaptation can be formalised as only involving the lower levels of task decomposition. After the plan has been generated the user may wish to change some aspect of it without generating a whole new plan. This is accomplished by the user rejecting a current task which results in the planner being re-activated and backtracking to the nearest overarching OR branch and generating a new sub-plan from the remaining nodes.

6. CONCLUSIONS AND FURTHER WORK

We have described a first implementation of a 'Companion' ECA generating and analysing user activities so as to influence his/her behaviour. We have adapted the level of plan elaboration to several factors, amongst which the constraints of interacting only when the user is at home as well as a desire to allow more flexible interaction and to avoid the type of complex negotiation and acknowledgement seen in related dialogue systems.

However, a natural extension of the system is to support some phases of real-time dialogue-based Mixed-Initiative Planning [1], in which the user would take a greater interest in the details of his daily activities. There is probably a balance to be found between user control and the burden of interaction and negotiation.

7. ACKNOWLEDGMENTS

This work was funded in part by the Companions project (www.companions-project.org) sponsored by the European Commission as part of the Information Society Technologies (IST) programme under EC grant number IST-FP6-034434.

Nabaztag™ is a trademark of Violet™, who is thanked for authorizing the development by some of the authors of the local web server "jNabServer" used in these experiments.

8. REFERENCES

[1] Allen, J., Chambers, N., Ferguson, G., Galescu, L., Jung, H., Swift, M., and Taysom, W., 2007. PLOW: A Collaborative Task Learning Agent. Proceedings of the Twenty-Second Conference on Artificial Intelligence (AAAI-07). Vancouver, Canada, pp. 22-26.

[2] Berry, P.M., Albright, C., Bowring, E., Conley, K., Nitz, K., Pearce, J.P, Peintner, B., Saadati, S., Tambe, M., Uribe, T.E., Yorke-Smith, N., 2006, Conflict negotiation among personal calendar agents. AAMAS 2006: 1467-1468

[3] Bohus, D., Rudnicky, A, 2003. RavenClaw:. Dialog Management Using Hierarchical Task Decomposition and an Expectation Agenda. In Proceedings of Eurospeech 2003: 597-600.

[4] Cavalluzzi, A., Carofiglio, V., de Rosis, F., 2004. Affective Advice Giving Dialogs. In: Elisabeth André, Laila Dybkjær, Wolfgang Minker, Paul Heisterkamp (Eds.): Affective Dialogue Systems, Tutorial and Research Workshop, ADS 2004, Kloster Irsee, Germany, June 14-16, 2004, Lecture Notes in Computer Science n. 3068, Springer, pp. 77-88.

[5] Ferguson, G., Allen, J. and Miller, B., 1996. TRAINS-95: Towards a mixed-initiative planning assistant, in Brian Drabble, ed., Proc. Third Conference on Artificial Intelligence Planning Systems (AIPS-96), pages 70–77, Edinburgh, Scotland, 1996.

[6] Ferguson, G., and. Allen, J. F., 1998. TRIPS: An Integrated Intelligent Problem-Solving Assistant, Proc. of the National Conference on Artificial Intelligence (AAAI-98), Madison, WI, pages 567–573.

[7] Forbus, K. and Hinrichs, T., 2004. Companion Cognitive Systems: A step towards human-level AI. To appear in Proceedings of the 2004 AAAI Fall Symposium on Achieving Human-level Intelligence through Integrated Systems and Research.

[8] Ghallab, M.; Nau, D.; and Traverso, P., 2004. Automated Planning: Theory and Practice. Morgan Kaufmann.

[9] Gruenstein, A., and Cavedon, L., 2004. Using an activity model to address issues in task-oriented dialogue interaction over extended periods, AAAI Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation.

[10] Gupta, S., Walker, M.A. and Romano, D.M., 2007. How Rude are You?: Evaluating Politeness and Affect in Interaction. Second International Conference on Affective Computing and Intelligent Interaction, ACII 2007, Lisbon, Portugal, pp. 203-217.

[11] Martin, D. L., Cheyer, A. J., Moran, D. B. The Open Agent Architecture: A framework for building distributed software systems. Applied Artificial Intelligence: An International Journal. Volume 13, Number 1-2, January-March 1999: 91-128, 1999.

[12] McCarthy, C.E. and Pollack, M. A, 2002. Plan-Based Personalized Cognitive Orthotic. Proceedings of the AIPS 2002 Conference.

[13] Myers, K. L. and Yorke-Smith, N., 2007. Proactivity in an Intentionally Helpful Personal Assistive Agent, in Proceedings of the AAAI Spring Symposium on Intentions in Intelligent Systems, AAAI Press, March 2007.

[14] Pollack, M. E., 2002. Planning technology for intelligent cognitive orthotics. In Proceedings of the 6th International Conference on Automated Planning and Scheduling.

[15] Rudnicky, A., and Xu,W, 1999. An agenda-based dialog management architecture for spoken language systems. In Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop, I–337.

[16] Seneff, S., Hurley, E., Lau, R., Pao, C., Schmid, P., Zue, V. Galaxy-II: a Reference Architecture for Conversational System Development. In Proceedings of ICSLP98, 931-934, 1998.

[17] Tate, A., Levine, J., Jarvis, P., and Dalton, J., 2000, Using AI Planning Technology for Army Small Unit Operations, Proceedings of the Fifth International Confer-ence on Artificial Intelligence Planning Systems (AIPS 2000), May 2000.

[18] Tomitsch, M., Kappel, K., Lehner, A. & Grechenig, T., 2007. Towards A Taxonomy For Ambient Information Systems. Pervasive 2007 Workshop on the Issues of Designing and Evaluating Ambient Information Systems.

[19] Turunen, M., Hakulinen, J., Räihä, K.-J., Salonen, E.-P., Kainulainen, A., Prusi, P, 2005. An architecture and applications for speech-based accessibility systems. IBM Systems Journal, Vol. 44, No. 3 (pp. 485-504).

[20] Turunen, M., Salonen, E.-P., Hartikainen, M., Hakulinen, J, 2004. Robust and Adaptive Architecture for Multilingual Spoken Dialogue Systems. In Proceedings of ICSLP 2004.

[21] Wilks, Y., 2006. Artificial Companions as a new kind of interface to the future Internet, Oxford Internet Institute, Research Report No. 13.