

Wizard of Oz Experiments for Companions

Jay Bradley

Oli Mival

David Benyon

Edinburgh Napier University
Centre for Interaction Design
School of Computing, Edinburgh,
EH10 5DT, +44 (0)131 455 2438

j.bradley@napier.ac.uk

o.mival@napier.ac.uk

d.benyon@napier.ac.uk

Abstract

Wizard of Oz experiments allow designers and developers to see the reactions of people as they interact with to-be-developed technologies. At the Centre for Interaction Design at Edinburgh Napier University we are developing a Wizard of Oz system to inform and further the design and development of Companion based technologies. Companions are intelligent, persistent, personalised, multimodal, natural language interfaces to the Internet and resources such as photo or music collections. They have the potential of turning our current human-machine interactions into human-machine relationships. In particular, a Companion prototype for reminiscing about a photo collection, called PhotoPal, is being used in our experiments. Several Wizard of Oz experiments have been run to assess people's reactions and thoughts about using a Companion interface. The feedback from these experiments has informed both the design direction and choice of development technologies going forward. The Wizard of Oz system has also been put to use in a classroom of young pupils and to aid adults make more productive use of the Internet for learning. Further experiments to investigate the appropriateness of Companion dialogue are planned.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – *evaluation/methodology, natural language, prototyping.*

General Terms

Design, Experimentation, Human Factors, Languages.

Keywords

Wizard of Oz, Companions, Interaction Design.

1. Introduction

Wizard of Oz experiments [6] refer to situations of requirements generation or evaluation where a human takes the place of a technology. Sometimes this is a technology that is yet to be developed, sometimes it is a technology that is not mature enough to reach a desired level of performance. Sometimes the technology is faked without the participant's

knowledge, but Wizard of Oz experiments can also be used as a participative design method where participants are well aware that the technology is faked [9].

Companions are envisaged as a new form of interaction; intelligent, personalised, persistent multimodal interfaces to the Internet [14]. Companions aim to go beyond task-oriented dialogue systems and should be capable of understanding natural human speech and through building and learning from a body of interactions should learn about their 'owners' and their resources. Benyon and Mival characterise companions as 'turning interactions into relationships' [2]. Companions will use automatic speech recognition to turn a person's speech into text. This text is then passed through a natural language understanding module and the output from there is passed onto a dialogue manager. The dialogue manager is, in its simplest form, a mapping from the semantics of an utterance to the semantics the Companion 'intends' to convey in response. The output from the dialogue manager is pushed through a natural language generation process producing text that is then passed through a text-to-speech module to be spoken. This last process, text-to-speech, is the only mature technology in the stack and hence the only module not replaced by our wizard. Looking further ahead we envisage more modalities such as gesture, facial expression and touch being involved in the interactions; other technologies that are yet to mature.

PhotoPal is a prototype Companion. PhotoPal discusses a person's digital photograph collection and ascertains information about the photos such as the names and ages of people in the photos, their relationships, the location of the photo and any anecdotes associated with it. We envisage that PhotoPal should support people reminiscing about periods of their life through their photo collection, possibly when there are limited opportunities to reminisce with other people. PhotoPal should also be utilitarian in helping to people to search, store, share and style their photos. In this paper we describe our use of Wizard of Oz techniques to understand, develop and evaluate PhotoPal.

In the next section we present the design challenge we face with this new type of multimodal, personalised, relationship-building interface. A description of the technology we are using in the Wizard of Oz system is given in Section 3. Section 4 discusses the application of our Wizard of Oz experiments so far. Finally, in Section 5 we outline our future plans for using the Wizard of Oz to measure changes in appropriateness of dialogue against changes in wizard behaviour.

2. The Design Challenge

The design challenge presented by companions is novel, interesting and significant. Companions are an example of something we could term 'semantic technologies'. Companions

© The Author 2009.

Published by the British Computer Society

deal with meanings rather than the syntactic interactions of mouse clicks and menu selections. We are just seeing the start of semantic technologies, with semantic tagging systems as found on 'Web 2.0' Internet sites such as Flickr and Del.icio.us. We are seeing the first attempts at interacting with these through novel interfaces such as tag clouds. Tim Berners Lee has suggested that the 'semantic web', based on tags will be the next development of Internet technology [3] and we envisage that Companions will be part of the new interaction with it.

However, with Companions we want to go further than simple key-word tags. We want to associate objects from a domain of application with whole conversations in natural language that a person had with a Companion. These conversations will be highly domain specific, at least to start with, but will grow over time. Already we have effective spoken natural language interactions in domains characterised by structured tasks such as buying cinema tickets and train tickets. What we do not have is ways of joining up these natural language interactions, learning about individuals or engaging in less structured activities.

The dialogue capabilities of companions will need to embrace a whole new set of concepts if relationships are to be formed. Persuasion is one of them and pro-activity another. The dialogues will need politeness and humour. They will also need explanation, rationale, discussion, disagreement and argumentation.

Interaction design will need to understand and develop a new set of techniques that will enable people to work at this level. And interaction design must do this as the inter-networked world becomes increasingly complex. New methodologies and new attitudes to design will be needed. Designing for relationships is very different than designing for function. Interaction design has always embraced the importance of form and as well as function and now it is taking on board emotional design too [10]. Companions demand a further step to deal with the characteristics described in Section 4 and to design for relationships.

Companionship is about an accessible, pleasing relationship with an interactive source in which there has been placed a social and emotional investment. There is a level of trust, compatibility and familiarity within this relationship that results in a feeling of security, content and general wellbeing. Companions are designed to enact these relationships, evoking an emotional investment, an attribution of personality and the provision of social roles. Relationships are built up, evolve and are maintained over time.

The social impact of changing interactions into relationships is significant. We already have stories of people having road accidents because they were paying too much attention to, and becoming too concerned about their Tamagotchi, the virtual pet. We know that the 'persona effect' [8] can have impact on interaction and how significant the 'media equation' ('media equals real life') is for people [12]. There are important social and ethical issues involved if we draw people into having relationships with devices, or with a computationally enabled ambient environment. There are issues concerned with exactly what these relationships might be like and what happens if they go wrong.

Companions bring about a significant change in the relationships between people and technologies. They introduce new moral and ethical issues to our discipline and radically alter interaction design. Although the term 'Companion' is deliberately chosen to evoke an anthropomorphic response in

the reader, Companions may be embodied as an intelligent building or product. Ambient Intelligence faces the same issues as Companions, and designers of all interfaces of the future will have to come to terms with designing for a level of engagement with technology that changes interactions into relationships.

Companions are a development of agents. Agents appear in the literature as software agents, interface agents or embodied conversational agents (ECA). ECAs have typically been more concerned with behaviours [11]. Interface agents have focused on dealing with some specific aspects of HCI. Some early thoughts on interacting with interface agents did highlight speech as a key element [10].

Bickmore and Picard argue that maintaining relationships involves managing expectations, attitudes and intentions [4]. They emphasise that relationships are long-term built up over time through many interactions. Relationships are fundamentally social and emotional, persistent and personalised. Citing Kelley they say that relationships demonstrate interdependence between two parties – a change in one results in a change to the other. Relationships demonstrate unique patterns of interaction for a particular dyad, a sense of 'reliable alliance'.

It is these characteristics of relationships as rich and extended forms of affective and social interaction that we are trying to tease apart. Benyon and Mival describe the characteristics of companions in terms of utility, form, personality, emotion, social aspects and trust [2]. Utility concerns how useful the Companion is; from 'useful uselessness' (for example, pets) to a virtual butler. Form concerns the embodiment of the Companion; as an on-screen avatar, a physical, digitally-enabled object or as a background sense of presence. The Companion will need some form of personality if people are to form relationships with it. It should be capable of behaving emotionally and of recognising emotions. It must be consistent and trustworthy and have appropriate social attitudes.

3. The Wizard of Oz Architecture

To meet this design challenge we have developed a Wizard of Oz architecture along with protocols for using the system to investigate the design space and experience of digital companions. We hypothesise that by iteratively building and testing Wizard of Oz mock technologies with people, we can inform and direct the interaction design to produce a highly suitable interaction. The term we have coined for this is 'Designing by Wizard of Oz' [5]. Our Wizard of Oz system is replacing several immature technologies and supporting the interaction design and look and feel of the eventually developed Companion interactions and as such is a reasonable development undertaking in itself. This contradicts the usual 'quick and dirty' application of Wizard of Oz experiments early on in the design of technologies. We intend to keep our Wizard of Oz system, develop it in line with user feedback and our observations and ultimately use it to develop suitable interfaces for Companion led applications.

In line with current *agile* development trends we take the view that the design and development of novel applications are not distinct activities. With interpreted scripting languages like ruby and interface building frameworks like Flex, along with several supporting libraries it is possible to create and alter experimental interfaces quickly. This supports the tight turnaround between design, development and testing of user experience. Barriers to sophisticated development frameworks and speed of development practices mean that development can

almost match design practices in terms of flexibility and freedom of expression, whilst also providing a somewhat usable application to use in testing. Molin also comments on how effective Wizard of Oz experiments can be for rapid iterative development [9].

In Wizard of Oz experiments, the wizard takes the place of the computer technology, which of course is not an easy thing to do. Salber and Coutaz argue that the task of the wizard in Wizard of Oz experiments is cognitively highly demanding [13]. The wizard has to respond as if he or she were some technology, and so must be consistent in content, style and pace of response. They suggest training for wizards is important along with using pre-stored answers. We also used pre-stored answers, but have subsequently moved away from this to a predictive text system.

The current Wizard of Oz system consists of two web-based interfaces and some server side components to join them together. Both interfaces are written using Adobe's Flex framework with the use of Papervision3D for 3D effects and the Actionscript Physics Engine to create the illusion of solidity for the photos. The photos are placed on a canvas that is masked by a frame. The canvas expands beyond the frame and can be moved around with simulated physics (bouncing off the limits of its movement, momentum and friction). The photos can similarly be *thrown* around the canvas. The user's interface (Figure 1) contains the Companion's representative avatar, the photo browsing and selecting area and a text box where the Companion's current and previous utterances are shown. The avatar is generated using the CrazyTalk Studio application, a package allowing one to quickly turn any image into an animated, lip-synching talking avatar. The user's and wizard's interface are almost identical to look at. The only difference is that the text box on the user's interface shows the wizard's utterances whereas the wizard's text box is for the wizard to type into. The wizard can hear the user's voice but the user only hears the wizard's utterances through the text-to-speech function of the CrazyTalk avatar. The photos and canvas area are kept in synchronisation so the actions of one can be seen on both interfaces. The synchronisation and streaming of the user's voice is handled by a Red5 server. A small ruby on rails application handles users and photo uploads.

The advantages of using web-based technologies are two fold: there are existing libraries and frameworks to allow reasonably quick development and deployment of interfaces and; the portability and accessibility of web applications should make it easier to run user evaluation sessions. Unfortunately the two proprietary technologies, Adobe's flash player and CrazyTalk browser plug-in, limit the portability of the interfaces and we are looking for more open solutions. Using a server side database it should be possible to automate collecting and collating of data about user interaction such as length of interaction, modalities used, frequency of speech and other modalities, et cetera. although we have not yet implemented this.

The first version of the user's interface (Figure 2) showed the avatar, its current utterance and a single picture. It used only HTML and JavaScript. A single file on a server facilitated sending text from the wizard to the user and moving forward and backward through the photos through pressing buttons on the wizard's interface. The first iteration of the wizard's interface contained a box for typing what was to be said and *next photo* and *previous photo* buttons. It was judged to be too difficult to talk via typing and operate the next/ previous buttons so a second wizard interface was built. The second

wizard's interface was massively more complex (Figure 3) than it was in our first version and is in our current version. This second version included canned text buttons that triggered frequently used utterances when clicked such as, 'Please tell me about this photo' and, 'Who are these people?' The interfaces have moved through several iterations and through each we have learned how to improve the design. We believe that performing the role of the wizard gives one a good insight into the challenges of designing for the Companions project. In the future we plan to hold participatory design workshops where participants get the chance to be a wizard to test the hypothesis that being a wizard can enlighten, inform and inspire better contributions from the participants than if they had not been wizards.

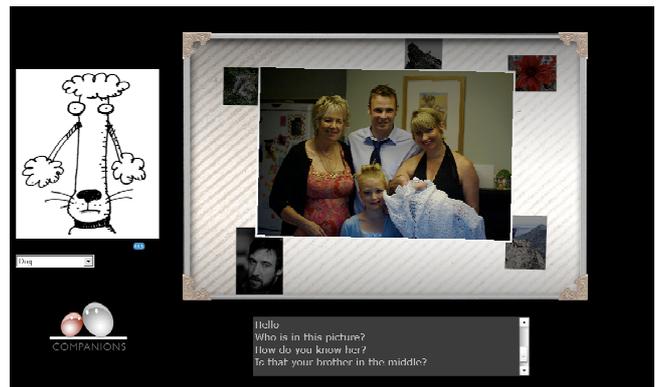


Figure 1: The current version of the user's interface.



Figure 2: The first version of the user's interface.

4. Application

The Wizard of Oz architecture has been used, so far, in two situations, with a third implementation pending. The methodology has been deployed imitating the Companions Project's PhotoPal prototype as well as in a classroom of school children who were learning about World War II from an avatar of Winston Churchill. It is also about to be used in a pilot study investigating the potential value of a Learning Companion for adult learners, intended to provide support for identifying, planning and achieving their own projects of learning.

The vast majority of work done has been investigating the notion of companionable interfaces to support the development of the PhotoPal Companion prototype. Many users have used the system informally and much testing has been done this way. Formally we have had twelve users of the Wizard of Oz system. Transcripts from these sessions were passed on to other members of the Companions project for use as *gold standard*

measures calculated automatically from the audio and transcriptions will inform the continuing design and development of Companion interfaces.

In summary, using Wizard of Oz experiments has enabled us to gain a greater insight into how people interact with digital companions. Our research has provided feedback to the development teams in the Companions Project and produced one example viable platform for delivering Companions technologies. It will allow us to perform appropriateness analysis on *gold standard* dialogue between humans and Companions that would otherwise not be possible.

6. Acknowledgements

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

Brian O'Keefe and Ingi Helgason undertook much of the empirical work. Thank you to Rebecca Eynon and Chris Davies from the Oxford Internet Institute for their input.

7. References

- [1] Benyon, D., Hansen, P. and Webb, N. (2008) *Evaluating Human-Computer Conversation in Companions*. Fourth International Workshop on Human-Computer Conversation, Bellagio, Italy.
- [2] Benyon, D. and Mival, O. (2008) *Landscaping Personification Technologies: from Interactions to Relationships*. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI 2008).
- [3] Berners-Lee, T., Hendler, J. and Lassila, O. (May, 2001) *The Semantic Web*. In Scientific America.
- [4] Bickmore T. and Picard R. (2005) *Establishing and maintaining long-term human-computer relationships*. ACM Transactions on Computer-Human Interaction (TOCHI), Volume 12 Issue 2.
- [5] Bradley, J., Mival, O. and Benyon, D. (2008). *A Novel Architecture for Designing by Wizard of Oz*. In: Proceedings of CREATE08.
- [6] Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). *Wizard of Oz studies - why and how*. In Proceedings of the 1st international Conference on User interfaces. IUI '93.
- [7] Eynon, R and Davies, C. (2009) *A Companion for learning in everyday life*. In, Y. Willks (ed) *Artificial Companions: technical, philosophical and social issues* Benjamins: Amsterdam.
- [8] Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., and Bhogal, R. S. (1997) *The persona effect: Affective impact of animated pedagogical agents*. In Proceedings of CHI 1997, 359-366.
- [9] Molin, L. (2004) *Wizard-of-Oz Prototyping for Cooperative Interaction Design of Graphical User Interfaces*. In proceedings of Nordichi 2004.
- [10] Norman, D. (2004) *Emotional Design*. Basic Books.
- [11] Pelachaud, C. (2005) *Brave new topics 2: affective multimodal human-computer interaction: Multimodal expressive embodied conversational agents*. In proceedings of the 13th annual ACM international conference on Multimedia. MULTIMEDIA '05 p.683 - 689
- [12] Reeves B. and Nass, C. (1996) *The Media Equation*. CSLI Publications; Stanford, CA.
- [13] Salber, D. and Coutaz, J. (1993) *Applying the Wizard of Oz Technique to the Study of Multimodal Systems*. In Proc. HCI'93, 219-230.
- [14] Wilks, Y. (2006) *Artificial Companions as a new kind of interface to the future Internet*. Oxford Internet Institute, Research Report 13, October 2006.