You, by proxy: Advances in virtual teachers

Andrew Marriott

Department of Computing Curtin University of Technology

This paper examines your future as a teacher and the many faces that you may wear. The paper first briefly outlines the relevant background areas of educational technology, embodied conversational agents and dialogue management. It then discusses some of the researcher's recent projects in software based mentoring and in embodied conversational agents, and uses these to propose various realistic future teaching and learning scenarios. It is hoped that this will foster discussion and/or persuade other teachers to move beyond the comfort zone of existing IT usage in their teaching.

Keywords: mentoring, embodied conversational agents, tutorial dialogue systems, dialogue management

Introduction

There are many different ways for someone to learn - see EMTECH (2004) for various paradigms on teaching and learning. Table 1 (adapted from Reinhardt (1995) who paraphrased from Means (1993)) indicates how the philosophy underpinning education has evolved to meet the new challenges, and especially the opportunities provided by the technology of ubiquitous cheap computing. However, whilst discussing the benefits of computer based instruction Kearsley, Hunter, & Seidel (1984) indicate that "while technology can be a tremendous multiplier on good ideas, it does not, in itself, produce them". The technology is there to help, not to solve.

Changing educational paradigms				
Old model New model		Technology implications		
Classroom lectures	Individual exploration	Networked PCs with access to information		
Passive absorption	Apprenticeship	Requires skills development and simulations		
Individual work	Team learning	Benefits from collaborative tools and Email		
Omniscient teacher	Teacher as guide	Relies on access to experts over network		
Stable content	Fast changing content	Requires networks and publishing tools		
Homogeneity	Diversity	Requires a variety of access tools / methods		

Table 1 Changing educational paradigms (adapted from Reinhardt, 1995))

Tutorial Dialogue Systems (TDS) (Graesser, VanLehn, Rose, Jordan, & Harter, 2001; Rose & Aleven, 2002) encourage aspects of the "New Model" outlined in Table 1 such as individual and non-linear learning (Beeman et al., 1987), apprenticeship (Brown, Collins, & Duguid, 1989) and guidance from experts (Mayer, 1998). For an ageless learner, ubiquitous cheap computing can be an enabling technology but it must be coupled with appropriate content and paradigms for learning..

One of the most important outcomes of a good educational system is "the cultivation of individual initiative in Students" (Schank & Cleary, 1994). In this respect, any developed educational system must encourage exploration by the user, provide rich information resources, and also a guiding hand to assist learning. The guiding hand can be either from peers or from an expert. The pedagogies and "andragogies" (adult learning theories (Knowles, 1984)) of the system must enable a learning style that is effective and motivating for the learner (Brickell, 1993).

One mechanism that can help in motivation is by using embodied conversational agents (ECA): sound, graphics and knowledge can convey ideas faster than technical documents alone. Koda, & Maes (1996) in their article "Agents with Faces: The effects of Personification of Agents" investigated the most "favourable" interface through the use of a poker game and 4 computer players, all with different personas. Two important results for ECAs were found:

- Faces made the game more engaging.
- The human face rated 1st in intelligence and comfort, and was 2nd to the dog cartoon in likeability.

People liked faces in interfaces!

Virtual or mythical Talking Heads (TH) have "existed" for a long time (Senior, 1985). These THs act as mediators or communicators of a message and facilitators of outcomes. They may be virtual such as in the people simulator *The Sims* by Maxis (Stepnik, 2000). They can be as lifelike as in the Famous-3d website http://www.famous3d.com/web/gallery/prerecorded/sorryjohnny or as abstract but well meaning as Microsoft's Wizard Paperclip (http://www.iarchitect.com/metaphor.htm). Their effectiveness in communicating does not normally correlate to their physical appearance. Pelachaud (1996) however, has suggested that integrating such non-verbal behaviour as facial gestures, expressions and emotions, with expressive speech, would increase the realism and hence the effectiveness of an ECA.

It has long been man's dream to be able to interact or converse with something man made: G.B. Shaw's "Pygmalion", the android in Fritz Lang's "Metropolis", Robbie the Robot in "Forbidden Planet", HAL in Clarke's "2001". A seminal argument in this debate is seen in the collection ``Can a machine think?" by Alan Turing (Turing, 1950). He argued that if a judge could not decide if a hidden contestant was human or not simply based upon their ability to carry on a question / answer conversation, then for all intents and purposes, the contestant was "human" regardless. The computer either imitates a human or is "intelligent"

Arguably, one of the first successful conversational agents to imitate humans was Eliza, developed by Joseph Weizenbaum in 1966. It was renown more for its "illusion of understanding" than for its Artificial Intelligence techniques. It did not understand what was typed into it as conversation but its abilities were convincing. It did not use formal Natural Language Processing techniques but classified the input text into patterns that it could use in its answers. For example, user enters "I like fish". Eliza replies with "Tell me more about why you like fish". Eliza has recognised "I" and replaced it with "you", has seen the action and the object of the sentence and used that as well.

Chatterbots are pattern matching conversational agents constructed to simulate conversation and/or provide useful information (Mauldin, 1994). Mauldin has continued his foray into chatterbot design recently with the construction of a new breed of chatterbots called verbots (see **Figure 1**). These verbots are not only chatterbots, but are "embodied" as Talking Heads with the use of 2½ D graphics and synthesised voices and include virtual personalities.

Marriott (1999) details research on a software based mentoring system – Mentor - that uses similar pattern matching techniques for natural language processing and for producing relevant information rich responses to help students in their university units. Although that study did not use a Talking Head, the system has been integrated with an MPEG-4 compliant Talking Head system (see **Figure 2**) – this could be used as a Virtual Lecturer.



Figure 1: Verbot: A visual chatterbot (http://www.verbots.com)

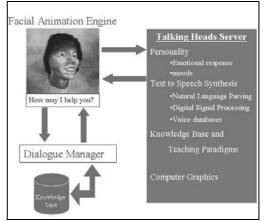


Figure 2: The functionality of a talking head dialogue system

Mentoring by proxy

The Mentor System (Marriott, 2002) supports various educational paradigms and is effective in helping users to learn (Marriott & Shortland-Jones, 2003) (Marriott, 2004b). The architecture (see Figure 3.)) is client server based, and is currently about 70,000 lines of Java code in 170 different classes spread over about 50 packages organised in a hierarchical fashion. The user interface to Mentor (see Figure 4) is HTML based and so can show multi- and hyper- media information as well as text. The three main regions of the interface cater for asking questions, a history of these requests, and the answer to the request. Typically, as seen in Figure 4, the request is for help on some assessable item. The system is also pro-active in that it prompts the user with questions such as "Have you started your assignment yet?" The system typically guides a user through a Directed Learning Path planned by the lecturer.

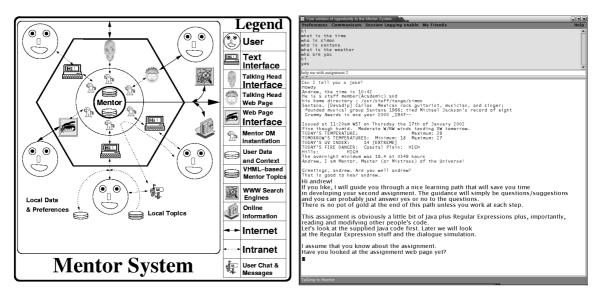


Figure 3: Schematic of mentor system

Figure 4: Mentor system user interface

In each study, users were asked via a 5 point Likert scale: "*Do you think it was beneficial or helpful to use the Mentor System*?" (1=> not at all, 5=>very beneficial). **Table 2** shows the summarised results of this question for each study (see Marriott (2004b) for a further discussion). Assuming a value of 3 or above is a positive indication of benefit, it can be seen that the 2^{nd} and 3^{rd} studies were marginally beneficial (given the standard deviation value and the '% > 3' figures).

Table 2:	Was	it	beneficial	or	helpful
----------	-----	----	------------	----	---------

Study	#	Mean	Med'n	std dev	% <u>></u> 3
S2-2001	38	2.58	3.00	0.919	55
S1-2002	31	3.29	3.00	1.160	75
S2-2002	69	3.39	3.00	0.826	87

Table 3: How effective was it

Study	#	Mean	Med'n	std dev	% <u>></u> 3
S2-2001	35	2.23	2.00	0.877	40
S1-2002	31	3.13	3.00	0.885	77
S2-2002	74	3.54	4.00	0.879	88

Users were also asked: "*How effective or useful was it for this purpose*?" Table 3 shows the results: assuming a value of 3 or above is a positive indication, it can be seen that the 2nd and 3rd studies were marginally effective. The "Mean" value of the rankings shown in Table 2 and Table 3, whilst not being impressive for a human teacher, are seen as being very positive quantitative results for a software based mentoring system. Issues raised from the qualitative responses can be found in Marriott (2004a).

The Mentor system was an effective proxy for the lecturer, available 24 hours a day. It required a different approach to teaching, it raised new issues in learning.

However, since the students mainly accessed the system in noisy computer laboratories, it was not appropriate to add a Talking Head to the system. Also, in the university's learning environment, it was felt that a Talking Head would not have added any extra value.

Distance education by proxy

A distance education environment often has a different set of learning obstacles. The student is isolated with little peer interaction, may have limited access to learning materials such as libraries, and may at best have only email or telephone contact with the Lecturer in Charge of the unit. Problems or issues that arise in assimilating the learning material may not be able to be readily addressed, or the student may not be able to explore new material or even cover the existing material due to a lack of guidance from experienced tutors.

A Virtual Lecturer (VL) or Virtual Tutor, who may resemble the Lecturer in Charge in face and sound, could reduce some of these problems. The system could be supplied on a CD along with the normal teaching materials. On booting up the CD, the VL would introduce itself to the student, install itself on the user's system and start helping the student with the prepared material. Most usefully, the VL could answer questions from the student *at any time*, and suggest new avenues of investigation, similar to the Mentor system. This help could be based upon proven educational paradigms and use hyper and multimedia material in answering student requests. If an Internet connection was also available to the system, Web sites could be mined for extra dynamic information. New material could also be obtained from the Lecturer's own teaching Web site. For those not familiar with search engines, the VL could provide a familiar front end that also reduced the number of irrelevant sites.

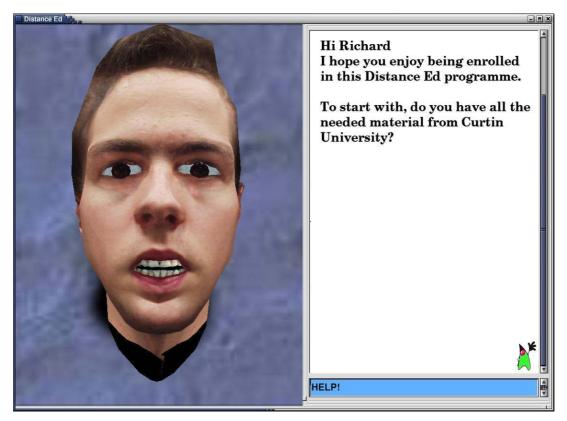


Figure 5: A distance education talking head connected to mentor system

Marriott (2002) indicates how a VL was created using an MPEG-4 TH and the Mentor system. This system has **not** been used or evaluated in distance education, but there is no reason, other than time and preparation ,why it could not be used. Disciplines could use appropriate famous icons, such as Einstein

for physics, Marie Curie for Chemistry, as teachers. Distance education students could have access to multiple THs each with their own expertise, features and personalities (for example, see the many faces of the Virtual Human Assistant in Frederic Pohl's "Beyond the Blue Event Horizon" (Pohl, 1980)).

It would be tempting to assume that the student could actually "talk" to the VL. Unfortunately, although Automatic Speech Recognition systems are becoming more accurate, they are not yet acceptable (but see Kadous & Sammut, 2002).

Elderly learning by proxy

To understand the problems of life long learning for the Elderly, we must understand the needs of users (Bergman & Johnson, 1995). As people age, their abilities change. This process includes a decline over time in their physical and cognitive abilities as well as their physical abilities declining at different rates relative to each other. When compared to their younger counterparts, the individual variability of sensory, physical and cognitive functionality of people increases with age and different techniques are required to deal with the widely appearing problems with cognition, e.g. dementia and memory dysfunction.

The elderly can be divided roughly into three groups (Gregor, Newell, & Zajicek, 2002):

- **Fit older people**: Who do not consider themselves, nor appear disabled but whose needs have changed with age.
- **Frail older people**: Who have suffered reduction in many of their functionalities, one of which may be considered a disability. For example, they may have limited mobility.
- **Disabled people who have aged**: Whose disabilities have affected their aging process making them dependent on other faculties that may also be declining.

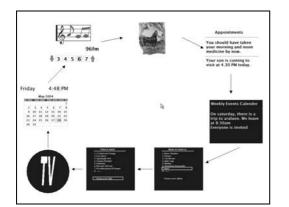
Most computer applications demand the use of memory, sight and strategy building; all faculties that decrease with age. Most are not designed to meet the requirements of the first time user, who is expected to know the required concepts in advance (EJEISA, 1996).

Experiential user interfaces seek ways to support new forms of communication. Experiential thought is when we perceive and react to events efficiently and effortlessly to make decisions without reflection (Lindh, 1997). In a sense, disruptions of the user's conscious world are limited by relying on the subconscious. There is a benefit from incorporating the experiential paradigm into an application to assist the elderly. This would reduce the demands on cognition and learning, and on fluid memory (memory to solve problems that have no solutions derivable from formal training or cultural practices).

Figure 6 shows the channel concept of the Pandora System (Holic, 2004). This system has been built using an experiential paradigm to assist the elderly who have mobility and/or memory problems. This system is a computer with a TV card in it and looks and behaves like a TV to the elderly user. It is assumed that they will be familiar with a TV and will not feel threatened by it. The system uses the dialogue manager of the Mentor system and is controlled by a 5 button TV remote control to reduce complexity: the user can change channels as normal. Unlike a TV, extra channels are also available to the user: channel 11 may be the radio, channel 12 a calendar, channel 13 an appointments diary, channel 14 a list of available CDs or music, etc. All these can be chosen using the simple remote control.

The system's main purpose was to remind user's of appointments or events such as taking medicine, as well as providing a simple encapsulated service for the delivery of TV and radio programs, music from CDs, movies from DVD's or via broadband, and normal or spoken word books from the Internet. Access to all these features is via the "channel" concept.

Other hyper-media learning materials such as online museums and art galleries are available although access to these is not currently implemented. These materials could be data mined and then filtered by the system – a proxy teacher who reduces the noise of irrelevant sites - to present the user with a menu of images or gallery spaces that could be accessed by the user as extra channels. It may be necessary for the primary carer of the elderly person (such as a son or daughter) to tell the Mentor system various preferences such as artist name or time period – "*Mum likes pictures by Vincent*." Similarly, a TV recording facility could be used to time shift Open University programs or documentaries.



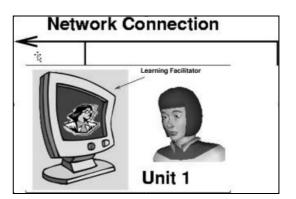
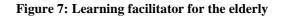


Figure 6: Information environment for the elderly

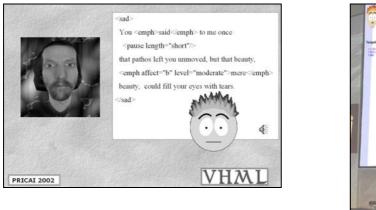


Similar to distance education, a TV TH could be used as a learning facilitator (see Figure 7). Along with the previously mentioned cognition problems however, other issues may produce obstacles to the effective learning for the elderly. For example, Age Care experts have indicated that using a realistic Talking Head as an interface may be detrimental in anchoring an elderly person into reality if they have the onset of dementia. It is quite common for some elderly people to talk to their TV and to blur the boundary between reality and TV.

Research presentation by proxy

Figure 8 shows a single slide from an entire Powerpoint presentation for a workshop on Embodied Conversational Agents held in Tokyo in 2002. The authors could not be there so a photo realistic ECA of one of the authors was prepared, along with Spiky Boy (Beard & Reid, 2002) – the 2 ½ D cartoon character in Figure 8 – and these two ECAs presented the research paper by proxy. Note that both ECA presenters were displayed via videos within Powerpoint.

In Figure 9, Spiky Boy is a "live" presentation aide (Beard, 2004). That is, he is generated in real time as a Powerpoint add on not a video. The application was able to assist the presenter during the seminar by giving information that otherwise may have been forgotten and gone unmentioned.



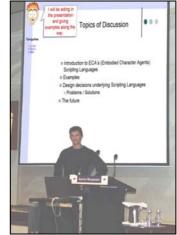


Figure 8: PRICAI'02 workshop Powerpoint presentation by proxy (Tokyo, Japan)

Figure 9: Presentation aide for AAMAS2003 workshop (adapted from Beard (2004))

Note that current technology severely restricts the interaction between aides such as Spiky Boy and either the presenter or the audience. Current interaction is via the keyboard only. As previously mentioned, Automatic Speech Recognition systems are still quite limited and although Keyword Recognition systems

- the ability to recognise a single utterance such as "yes" or "no", or a simple phrase such as "turn left" - are more robust, this would not be useful for carrying on a conversation with Spiky Boy.

You by proxy

Figure 10 shows a Spiky Boy Business Card (Beard, 2004). This Business Card CD contains an ECA system that on being inserted into a PC will bring up a Spiky Boy enabled Web page. Beard (2004) has produced several of these: a demonstrator for a product, a seminar presenter, and an eRésumé system that has Spiky Boy extolling the virtues of his creator. The creator, by proxy!

The ECA system on the Business Card CD could also contain a Dialogue Manager such as the Mentor system to answer questions about the person or company that the card represents. For example, a business card for the International Office of Curtin University could contain knowledge useful to overseas students about course structures, fees, application procedures, housing, cost of living, etc. A Virtual Curtin Representative could give the student a guided tour of the university and answer questions about the various Schools and facilities.

Figure 12 shows some of the non photo realistic avatars that companies such as Eptamedia Srl. make available for clients. These typically represent the client in a Web site and may be tailored to conform to the client's corporate logo. For example, the Western Australian company "Hospital Benefit Fund" would probably use an avatar on their web pages that resembles their Teddy Bear corporate image. Photo realistic avatars simply require front and side photographs of the appropriate person for mapping onto the 3D geometric model (see examples in Figure 5, Figure 8 and Figure 11).





Figure 11: "You, by Proxy" calling card

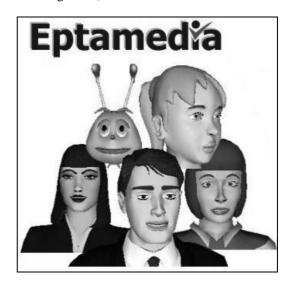


Figure 12: MPEG-4 avatars from Eptamedia

Finally, Figure 11 shows a futuristic example of the business card. 'You, by proxy' is a high density Flash Memory stick combined with a wireless transmitter such as BlueTooth or WiFi, that can broadcast a proxy of yourself that others might use or interrogate. It may be that you actively make this proxy available to others by specifically broadcasting, or it may simply be available to anyone who queries the stick. Public Key Encryption (Zimmermann, 1995) can be used to allow levels of information access to various people. General information about you may be available to the casual observer, whereas specific information may be made available by you to specific audiences.

Similarly, you may enable access to Teaching and Learning material for your students. Lectures may become a once a week meeting for a 30 second wireless broadcast plus a 2 hour discussion session for problems. What face will you wear in the future?

Conclusions

In conclusion, future discussion should not be about the use of an ECA in Learning applications simply because of the ability of a face (either natural or virtual) to attract and hold the attention of a student. The discussion instead should be concerned with whether to use (or misuse) a virtual teacher to emulate a real teacher within an environment as close as possible to the real world or, on the contrary, to exploit the potentialities of the virtual teacher to create something completely new, something that will help the student to learn in new and interesting ways!

References

- Beard, S. (2004). *MetaFace: A Virtual Face Framework and Metaphor*. Unpublished PhD., Curtin University of Technology, Perth, Western Australia.
- Beard, S. & Reid, D. (2002, 16 July 2002). *MetaFace and VHML: A First Implementation of the Virtual Human Markup Language*. Paper presented at the AAMAS workshop on Embodied Conversational Agents let's specify and evaluate them!, Bologna, Italy.
- Beeman, W. O., Anderson, K. T., Bader, G., Larkin, J., McClard, A. P., McQuillan, P., & Shields, M. (1987, November 1987). *Hypertext and Pluralism: From Linear to Non-linear Thinking*. Paper presented at the ACM Hypertext '87 Conference.
- Bergman, E. & Johnson, E. (1995). Towards Accessible Human-Computer Inter-action. *Intellect*, Oxford, UK.
- Brickell, G. (1993). Navigation and learning style. *Australian Journal of Educational Technology*, 9(2), 103-114, http://www.ascilite.org.au/ajet/ajet9/brickell.html
- Brown, J. S., Collins, A. & Duguid, S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42. [verified 24 Oct 2004]
- http://www.exploratorium.edu/IFI/resources/museumeducation/situated.html
- EJEISA (1996). *Basic characteristics of users with special needs and their telematics requirements*. [29 March 2004, verified 24 Oct 2004] http://www.ejeisa.com/nectar/respect/6.1/
- EMTECH (2004). *Learning Theories*. [verified 24 Oct 2004] http://www.emtech.net/learning_theories.htm
- Graesser, A. C., VanLehn, K., Rose, C. P., Jordan, P. W. & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine*, 22, 39-52. [verified 24 Oct 2004] http://www.findarticles.com/p/articles/mi_m2483/is_4_22/ai_82129226
- Gregor, P., Newell, A. F. & Zajicek, M. (2002). Designing for dynamic diversity interfaces for older people. Paper presented at the Fifth International ACM conference on Assistive Technologies, Edinburgh, Scotland.
- Holic, A. (2004). *Assistive User Interfaces For the Elderly*. Unpublished Honours Thesis, Curtin University of Technology, Perth, Australia.
- Kadous, M. W. & Sammut, C. (2002). Mobile conversational characters. Paper presented at the HF2002: Virtual Conversational Characters: Applications, Methods, and Research Challenge, Melbourne, Australia. http://www.smartinternet.com.au/SITWEB/publication/files/112_\$\$\$_4057/P03_001.pdf
- Kearsley, G., Hunter, B. & Seidel, R. J. (1984). Two decades of computer based instruction projects: What have we learned? *Technological Horizons in Education Journal*, 10(4), 88-96.
- Knowles, M. (1984). Andragogy in Action. San Francisco: Jossey-Bass.

Koda, T. & Maes, P. (1996). Agents with faces: The effects of personification of agents. Paper presented at the HCI'96, 20-23 August. London, UK. [verified 24 Oct 2004] http://citeseer.ist.psu.edu/rd/7694471%2C181867%2C1%2C0.25%2CDownload/http://citeseer.ist.psu .edu/cache/papers/cs/2755/http:zSzzSztomoko.www.media.mit.eduzSzpeoplezSztomokozSzdocszSzH CI_final.doc.pdf/koda96agents.pdf

- Lindh, E. (1997). *Sensational interfaces: Function and passion*. Paper presented at the STIMDI Conference, September. Linkoping, Sweden.
- Marriott, A. (1999). A lifelong mentor system. In K. Martin, N. Stanley and N. Davison (Eds), Teaching in the Disciplines/ Learning in Context. Proceedings of the 8th Teaching Learning Forum, The University of Western Australia, February. Perth: UWA. http://lsn.curtin.edu.au/tlf/tlf1999/marriott.html
- Marriott, A. (2002). A Facial Animation case study for HCI: the VHML-based Mentor System. In I. Pandzic & R. Forchheimer (Eds.), *MPEG-4 Facial Animation - The standard, implementations and applications*. London: John Wiley.

Marriott, A. (2004a). *Mentor:* "lotsa progress in future". *Transforming Knowledge into Wisdom: Holistic Approaches to Teaching and Learning*. Proceedings HERDSA 2004, 4-7 July. Miri, Sarawak. [verified 24 Oct 2004] http://herdsa2004.curtin.edu.my/Contributions/NRPapers/P053-jt.pdf

- Marriott, A. (2004b, February 9th-10th). *Mentor : Really annoying. but quite helpful.* Paper presented at the Teaching and Learning Forum 2004, Perth, Western Australia.
- Marriott, A., & Shortland-Jones, B. (2003). The Mentor System. Paper presented at the "Tutorial Dialogue Systems: With a View Towards the Classroom" Workshop Proceedings of 11th International Conference on Artificial Intelligence in Education, 18 July 2003, Sydney, Australia.
- Mauldin, M. L. (1994). Chatterbots, Tinymuds, And The Turing Test: Entering The Loebner Prize Competion. Paper presented at the AAAI-94, Seattle, USA.
- Mayer, R. E. (1998). Cognitive, metacognitive and motivational aspects of problem solving. *Instructional Science*, 26(1,2), 49-63,
- Means, B., Blando, J., Olson, K., Middleton, T., Morocca, C., Remz, A., & Zorfass, J. (1993). Using Technology to Support Education Reform. Washington, USA: US Government Printing Office.
- Pelachaud, C., Badler, N. & Steedman, M. (1996). Generating Facial Expressions for Speech. *Cognitive Science*, 20(1).
- Pohl, F. (1980). Beyond the Blue Event Horizon. London: Gollancz SF.

Reinhardt, A. (Ed) (1995). New ways to learn. *BYTE*, March [verified 24 Oct 2004] http://www.byte.com/art/9503/sec7/art1.htm

Rose, C. P. & Aleven, V. (2002). Paper presented at the Workshop on Empirical Methods for Tutorial Dialogue Systems, 4 June, San Sebastion, Spain.

Schank, R. & Cleary, C. (1994). *Engines for Education*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. [verified 24 Oct 2004] http://www.engines4ed.org/hyperbook/

- Senior, M. (1985). The Illustrated Who's Who in Mythology. London: Guild Publishing.
- Stepnik, M. (2000). Reviews: The Sims. PC PowerPlay, April, 58-63. [verified 24 Oct 2004] http://www.pcpowerplay.com.au/review_content.cfm?id=57

Turing, A. M. (1950). Can a machine think? *Minds and Machines*, 433-460. Zimmermann, P. R. (1995). *The Official PGP User's Guide*. Cambridge, MA, USA: MIT Press.

Please cite as: Marriott, A. (2004). You, by proxy: Advances in virtual teachers. In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Eds), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 587-595). Perth, 5-8 December. http://www.ascilite.org.au/conferences/perth04/procs/marriott.html

Copyright © 2004 Andrew Marriott

The author assigns to ASCILITE and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The author also grants a non-exclusive licence to ASCILITE to publish this document on the ASCILITE web site (including any mirror or archival sites that may be developed) and in printed form within the ASCILITE 2004 Conference Proceedings. Any other usage is prohibited without the express permission of the author.