HCSNet Priority Area Workshop on Effective Interactive Interfaces

NICTA @ Australian Technology Park
Sydney
27-28 September 2006

Abstracts

Welcome

Welcome to the HCSNet Priority Area Workshop on Effective Interactive Interfaces. The aim of this workshop is to bring together researchers from disciplines as diverse as speech processing, linguistics, psychology, human-computer interaction, graphics, music and audio-visual communication to shed new light on how we can improve human-machine interfaces. Questions to be considered include: What makes interaction "effective", in the sense of enhancing user task performance and task efficiency? What interaction strategies can be deployed within human-machine interfaces, particularly dialogue and Multimodal interfaces, to make them more effective? What aspects of human-human interaction lead to successful collaborative task performance, and how can these be leveraged in the design of interactive interfaces? What techniques can be used to make interaction more robust? Do "natural" styles of interaction lead to more effective interfaces? What role do emotions and embodiedness play?

The workshop will include short presentations by all participants describing their specific interests in order to identify shared interests, understand different perspectives, and take first steps towards developing future collaborations. Panels and break-out sessions will provide opportunities for group discussion and planning, and there will be special emphasis on defining a series of multi-disciplinary topic themes that have the potential to lead to specific projects.

For their help with organisation of various aspects of the workshop, our thanks to James Ballantine, Fang Chen, Rosemary Eliott, Michelle Jablonski, Barbara Munday and Darlene Williams. We are also grateful to NICTA for the use of conference facilities at the Australian Technology Park.

We thank all our speakers and delegates for their participation in the Workshop. We hope you enjoy the papers and discussions and look forward to seeing you at HCSNet events in the future.

Lawence Cavedon, NICTA and RMIT University
Robert Dale, Macquarie University
Convenors

About

HCSNet

http://www.hcsnet.edu.au

The ARC Research Network in Human Communication Science – HCSNet - was awarded five years' funding by the Australian Research Council in late 2004. The aim of HCSNet is to promote and facilitate interdisciplinary research in human communication science by connecting leading researchers in language, speech and sonics.

Priority Research Areas in HCSNet are:

- Human and Machine Speech
- Effective Interfaces
- Next-Generation Search Technology
- Human Communication Disorders
- Perception and Action

By generating an explosion of new approaches and knowledge, the network aims to build Australia's reputation as a leader in communication science and technology via advances in areas as diverse as automatic speech recognition, distress call monitoring, hearing prostheses, web interfaces, and data retrieval and data mining systems.

Getting involved in HCSNet is easy: visit www.hcsnet.edu.au to sign up as a member of the network. You'll be added to our online profile database, and automatically receive our weekly electronic newsletter, *HCSNet Update,* which will keep you informed of HCSNet activities, including the annual SummerFest, and events in the range of HCSNet disciplines. Australian-based HCSNet members can apply for funding under our various programs.

Workshop Information

Schedule
Wednesday, 27 September
Lunch: 12.30pm - 1.30pm
Program commences: 1.30pm
Afternoon tea: 3.00pm - 3.30pm
Program concludes: 5.30pm

Thursday, 28 September
Program commences: 9.00am
Morning Tea: 10.30am - 11.00am
Lunch: 12.30pm - 1.30pm
Afternoon Tea: 3.00pm - 3.30pm
Program concludes: 5.00pm

Registration
Please note that all workshop participants will be required to wear security badging (NICTA and HCSNet logos) on both days of the workshop.

Workshop Dinner
HCSNet will host a Workshop Dinner at 7:00pm on Wednesday, 27 September at Almustafa Lebanese Restaurant, 23 Glebe Point Road, Glebe.

Taxis
Taxis can be booked through NICTA – please advise Barbara Munday, Research Administrator, if you wish to make a booking.

Workshop Participants

Stephen Barrass	The University of Canberra	Stephen.Barrass@canberra.edu.au
Michael Bui	The University of Sydney	mbui3089@mail.usyd.edu.au
Densil Cabrera	The University of Sydney	densil@arch.usyd.edu.au
Lawrence Cavedon	NICTA & RMIT University	lawrence.cavedon@nicta.com.au
Fang Chen	NICTA, Sydney	Fang.Chen@nicta.com.au
Kelvin Cheng	The University of Sydney	kcheng@it.usyd.edu.au
Eric Choi	NICTA, Sydney	eric.choi@nicta.com.au
Robert Dale	Macquarie University	rdale@ics.mq.edu.au
Sam Ferguson	The University of Sydney	sferguson@ihug.com.au
Roland Goecke	Australian National University	roland.goecke@anu.edu.au
Edward Ivanovic	The University of Melbourne	edwardi@csse.unimelb.edu.au
Andrew Johnston	University of Technology, Sydney	aj@it.uts.edu.au
Judy Kay	The University of Sydney	judy@it.usyd.edu.au
Trent Lewis	Flinders University	Trent.Lewis@flinders.edu.au
Martin Luerssen	Flinders University	martin.luerssen@flinders.edu.au
Andrew Lum	The University of Sydney	alum@it.usyd.edu.au
Gregor McEwan	NICTA, Sydney	gregor.mcewan@nicta.com.au
Matthew McGill	The University of New South Wales	mmcgill@cse.unsw.edu.au
William Niu	The University of Sydney	niu@it.usyd.edu.au
Garth Paine	The University of Western Sydney	ga.paine@uws.edu.au
Cecile L. Paris	CSIRO ICT Centre	Cecile.Paris@csiro.au
David Powers	Flinders University	David.Powers@flinders.edu.au
Claude Sammut	The University of New South Wales	claude@cse.unsw.edu.au
Claudia Schremmer	CSIRO ICT Centre	Claudia.Schremmer@csiro.au
Rolf Schwitter	Macquarie University	schwitt@ics.mq.edu.au
Ian Stevenson	University of Western Sydney	i.stevenson@uws.edu.au
Cara Stitzlein	CSIRO ICT Centre	Cara.Stitzlein@csiro.au
Ronnie Taib	NICTA, Sydney	Ronnie.Taib@nicta.com.au
Dian Tjondronegoro	Queensland University of Technology	dian@qut.edu.au
Aiman Turani	The University of Sydney	aimant@ee.usyd.edu.au
Rudi Vernik	DSTO	Rudi.Vernik@unisa.edu.au
Anja Wessels	CSIRO ICT Centre	Anja.Wessels@csiro.au

Abstracts

Stephen Barrass
The University of Canberra

My research interests are in interaction design, haptic-audio narrative, creativity in collaborative virtual environments, data sonification, information aesthetics, and auditory scene analysis.

I was involved in the development of a voice operated home environment for a quadraplegic that was in continuous usage in Bega from 1990-1995. A 286 PC with 250 word recognition was used to control the bed, television, phone, lights and do word processing. When this custom built system broke down the user could not find a replacement and went back to having to call a person every time he had to change the telly or move the bedhead. Now we are about ten doublings in technology sophistication later and the voice operated environment has not become any closer to the practical reality we believed we were on the verge of back then. I am hoping to find out why it hasn't happened, and if it ever will, or why it never will, with the help of the others at this workshop.

Michael Bui
The University of Sydney

Pertinent research interests include:

- The process of decision making relies on having a significant amount of data that needs to be analysed in order to make an informed decision.
- Decision making can be viewed as being equivalent to path planning as both involve going from one state to another.
- Therefore, it may be possible to apply path planning techniques to support decision making(depending on the data). However, such data can be high dimensional and thus making it difficult to apply traditional path planning techniques.
- The Self-Organizing Map algorithm(in this case I use the Geodesic Self-Organizing Map) may be used to perform dimensionality reduction, to map n-dimensional data to 2D or 3D, which would then allow the application of path planning techniques, such as distance transformations, which have typically been employed in the area of robotics.

Densil Cabrera
The University of Sydney

My research interests are in acoustics and audio, including aspects of auditory display, spatial audio, psychoacoustics and room acoustics. Some research is concerned with interfaces, such as applying psychoacoustical models to auditory graphing, design of auditory alerts for air traffic control consoles, effectiveness of simple spatial audio interfaces, and other aspects of auditory perception and auditory display. I also work with audio and acoustic systems, including the assessment of speech intelligibility and system quality more generally. I run an acoustics laboratory which has substantial equipment and facilities for precise measurement and analysis in these areas. Thoughts on the workshop theme: (i) there is great potential for exploiting perceptual and cognitive models in designing and optimizing auditory interfaces; (ii) many situations have big limitations with regard to acoustic and audio systems, posing interesting design challenges.

Fang Chen
NICTA, Sydney

Cognitive Load Measurement through Multimodal Behavior Patterns

Multimodal interfaces expand the communication channel between the system and the user allowing users to express themselves more naturally and interact with complex information with more freedom of expression. One of the many cited advantages of multimodal interfaces is their ability to facilitate effortful complex tasks over unimodal interfaces. These strategies often result in changes to the way multimodal constructions are planned and executed.

Cognitive load refers to the amount of mental effort imposed by a particular task and has been associated to the limited capacity of working memory. I will start with an overview of the state of the art in cognitive load measurement. Recent research has shown that users' multimodal constructions exhibit significant changes as they self-manage their cognitive load when faced with tasks of increasing complexity. Our research focuses on extending the accepted benefits of multimodal interaction by using it to detect fluctuations in cognitive load will be stressed. The primary advantage of this approach is that cognitive load can be determined implicitly by monitoring variations of specific multimodal features during day to day tasks. Such unobtrusive measures may help determine users' cognitive load in real time and adapt information content selection and presentation (multimodal output generation) accordingly, in order to ensure optimal user performance.

In this talk, I will describe an experiment designed to identify the relationships between combined speech and manual gesture input structures and users' cognitive load. The two input modalities are very familiar to users and psychologically closely interrelated, both in terms of planning and execution. Assessing a user's cognitive load implicitly through their multimodal behaviour requires identifying a number of indices that reliably reflect

fluctuations. Our hypothesis is that variations in redundant and complementary multimodal constructions can reflect cognitive load changes experienced by the user. The feasibility of using rates of redundant constructions or even complementary constructions in multimodal input as an index of cognitive load is supported by the results of our study. I will illustrate multimodal patterns that may be monitored to detect cognitive load variations based on symptomatic behavioural features. I will conclude with a discussion on the enormous impact such methods may bring to the design of human computer interaction systems, but highlight the current limitations of the pattern acquisition methodology. Directions for future work will also be addressed.

Kelvin Cheng
The University of Sydney

Large displays have been shown to provide higher productivity gain and user satisfaction compared to traditional monitor displays for the personal computers. Yet in most cases, the user interaction still consist of the computer mouse - a device that was only designed for the PC and only performs moderately well when scaled to large displays. Various novel technologies have been introduced to make large interactive interfaces more effective and direct, less intrusive and less cumbersome. Touch screen is a good example, but it is not scalable. The current trend is to use our own hand as the input device so we can point at the screen directly using the index finger. The goal of many on-going researches is to investigate the use of computer vision and the pointing gesture to interact effectively with large displays.

In psychology, the pointing gesture belongs to the class of "deictics" gestures, which serves two main functions: to indicate a direction or to pinpoint a certain object. In collaboration with psychologists, computer scientists can determine the most natural way for human to point at objects from a distance. They can also determine the kind of gestures that is more natural to use, and allow us to interact with computers (HCI) similar to interacting with humans (HHI).

Julien Epps
NICTA, Sydney

Three of Australia's publicly funded research organisations (CSIRO, DSTO and NICTA) have recently joined forces to lead a new national research program, known as the HxI Initiative, in ICT-Augmented Human Interactivity. The HxI Initiative focuses on the "x" factor of Human Interactivity, where the effective application of ICT itself is used to augment human cognitive and social abilities, often seen as the limiting factor in highly networked organisations and societies. The term 'HxI' was coined to describe these trends towards ubiquity and human experience. The "x" in the term 'HxI' represents research from a number of disciplines (e.g. cognitive science, computer and information science, psychology, sociology, and interaction design) that collectively improve the ability of humans to interact with information, with each other, and with their environments through the effective application of ICT.

In particular, the Braccetto project, which is the first and foundational project of the initiative, undertakes research into how the effective application of ICT in the area of Mixed Presence Groupware can assist teams of co-workers to collaborate more effectively across a distance. Braccetto is undertaking research into the principles underlying effective, intense distributed collaboration and is implementing the results as new capabilities for supporting teams involved in creative activities such as collaborative design, planning, analysis, and decision-making. These new approaches will be evaluated against various underlying teamwork mechanics, focusing specifically on increasing the effectiveness of distributed teams through enhanced awareness.

Roland Goecke
Australian National University

My research interests can largely be summarised as being in signal processing for Human-Computer Interaction (HCI), with a particular focus on video and audio processing and related areas. Specific research interests currently include face and facial feature tracking by active appearance models, automatic detection of humans in video and separation from the background, gesture recognition, facial expression recognition, multimodal affective computing, infrared imaging for health-care applications, and audio-video speech processing. As one can see, multimodal HCI is a general theme in my research and with that comes an interest in multimodal signal fusion, i.e. how can one best integrate data from multiple sensors in multiple modalities.

From a technology point of view, we find ourselves at a point in time where hardware prices have dropped to such an extent that we find computer technology more and more embedded in everyday items such as mobile phones, cars, digital cameras etc. With such a widespread use of technology the issue arises of how humans can interact most effectively with this technology. In the past, it was a case of the human having to adapt to the particular input and output formats of the computer system. Nowadays, we are at a transition point towards more human-like interfaces, such as automatic speech recognition, gesture recognition and facial expression recognition. In my opinion, this does not mean that traditional means of HCI (keyboards, computer mice, screens) will become obsolete, but that we will see more differentiated and application-specific interfaces that will thus be more effective. The interaction between disciplines such as computer science, psychology, linguistics and speech science, cognitive science and artificial intelligence will play an important role in developing such effective interfaces. Similar to interfaces becoming more application-specific, I also foresee that collaboration and interaction between disciplines will be more specific towards particular interfaces.

Edward Ivanovic
The University of Melbourne

My research interests are in instant messaging style dialogue systems. One obvious application of this type of dialogue is in user interfaces. With the number of options in typical applications such as word processors increasing, current user interfaces have resorted to using devices such as multi-tabbed dialogue boxes in an effort to logically group hundreds of options. This becomes difficult for users to navigate and find the desired functionality.

Dialogue interfaces would allow a user to tell the application what to do using natural language. In the case of ambiguous requests, applications would have the opportunity to ask clarification questions. Accomplishing this requires some degree of semantic representation to match a function with the many ways in which it can be described. This may eventually encompass a wide range of user feedback combined with reinforcement learning where systems learn what feature the user is referring to with various utterances.

Andrew Johnston
University of Technology, Sydney

In our research we have been developing interactive software "toys" for musicians which are intended to encourage musical exploration. Musical exploration in this sense is a deliberately broad term which might include:

- improvisation with the toys;
- composition for "traditional" instrument augmented by the toys; and/or
- exploration of characteristics of the musician's sound/technique.

A technique we have been using to facilitate user engagement is the use of physical models (ie. mass-spring-damper systems) to mediate between the live music produced by the musician-user and the audio-visual response of the software. This seems to result in amore intuitively understandable interaction paradigm that musicians respond well to.

This research is in progress, but our intention is to use evaluation techniques from HCI to inform an iterative approach to interaction design. We would be very interested to discuss our work with those researching the psychology of music, as we feel that research into perceived relationships between sound and physical forces will be relevant to our work.

Judy Kay
The University of Sydney

Here is a list of the things that our group is doing that seem relevant:
Tabletop interaction: social interfaces, metadata markup and pervasive file system interaction (Trent Apted, Anthony Collins, Judy Kay, Glen Whitaker)
Learner modelling for reflection: largescale modelling of learning in the context of an HCI course based on blended elearning; learning to think like a programmer with design and analysis supported with reflective activities (Judy Kay, Lichao Li, Andrew Lum)
Keep-in-touch: intergenerational communication with appliances that support natural asynchronous voice messages plus images with highly flexible interaction options from magic mirrors, magic wands, touch.... (Patrick Burns, Veasna Hoy, Judy Kay, Bob Kummerfeld, Geoff Langdale, Glen Tregoning)
MyPlace: personalised delivery of information about places, people, sensors and things driven by active user modelling (Mark Assad, David Carmichael, Judy Kay, Bob Kummerfeld, William Niu)
Never mind the quality, just feel the width: exploiting vast amounts of learning data to support teaching and learning, making use of visualisations, data mining of activity, interactions and temporal patterns (Judy Kay, Irena Koprinksa, Andrew Lum, Nicolas Maisonneuve, Peter Reimann, Adam Ullman, Kalina Yacef, Osmar Zaine).

Andrew Lum
University of Sydney

This work is concerned with the ways light-weight ontologies can support scrutability for large user models and the user modelling process. It explores the role that light-weight ontologies can play, and how they can be exploited, for the purpose of creating and maintaining large, scrutable user models consisting of hundreds of components. We address problems in four key areas: ontology creation, metadata annotation, creation and maintenance of large user models, and user model visualisation, with a goal to provide a simple and adaptable approach that maintains scrutability. Each of these key areas presents a number of challenges that we address.

Our solution is the development of a toolkit, LOSUM, which consists of a number of tools to support the user modelling process. It incorporates light-weight ontologies to fulfill a number of roles: aiding in metadata creation, providing structure for large user model visualisation, and as a means to reason across granularities in the user model. In conjunction with this, LOSUM also features a novel visualisation tool, SIV, which performs a dual role of ontology and user model visualisation, supporting the process of ontology creation, metadata annotation, and user model visualisation.

Gregor McEwan and Saul Greenberg
NICTA, Sydney

Informal Awareness and Casual Interaction with Community Bar

Various studies of white collar work sites report that a large portion of peoples' time is spent in unplanned, casual interactions with co-workers [6][9]. These interactions are stimulated by physical proximity: people acquire informal awareness of each other, such as knowledge about presence, activity, and availability, which leads to opportunities for people to engage in light-weight casual interactions. Casual interactions are unplanned, brief, frequent, and usually engage small groups of people familiar with one another. While seemingly mundane, these casual interactions prove important as they make the transition to tightly-coupled collaboration easier. However, the same studies also found that these interactions are severely affected by physical separation. This means that distributed communities of co-workers miss out on these interaction opportunities. In response, are a myriad of tools providing mechanisms for displaying informal awareness information that lead to casual interactions, e.g. Instant Messengers [7], chat rooms / MUDS [3], and video-based media spaces [1].

Our work includes development of an awareness and interaction tool. Our design perspective was to ground development in social science theory. In particular, we were motivated by the Locales Framework [4], a comprehensive theoretical group interaction framework, as well as the Focus/Nimbus model of awareness [8]. We derived and combined principles from these theories and applied them to the design of Community Bar (CB), a groupware tool that supplies ad hoc groups with rich awareness information leading to casual interaction. CB also leverages and extends two previously introduced design ideas. First, media items [5] are used as groupware building blocks to offer rich multimedia awareness and interaction capabilities. Second, these items are embedded within the sidebar metaphor [2], where people see awareness information at the screen's periphery, and can selectively drill down to more information and interaction. This work made two contributions: firstly, the derivation of theory-based design principles for informal awareness and casual interaction systems, and secondly, the design and implementation of a system, CB, that follows these principles.

Bly, S.A., Harrison, S.R., and Irwin S. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment, in *Comm. ACM*, 3, 1, (1993), 28-47.
Cadiz, JJ, Venolia, G.D., Jancke, G., and Gupta, A. Designing and deploying an information awareness interface. *Proc ACM CSCW* (2002), 314-323.
Curtis, P., Nichols, D. A. MUDs Grow Up: Social Virtual Reality in the Real World. *Proc 39th IEEE COMPCON* (1994), 193-200.
Fitzpatrick, G. *The Locales Framework: Understanding and Designing for Wicked Problems*. Kluwer Academic Publishers, (2003).
Greenberg, S. and Rounding, M. The Notification Collage: Posting Information to Public and Personal Displays. *Proc ACM CHI*, (2001), 515-521.
Kraut, R., Egidio, C., Galegher, J. Patterns of Contact and Communication in Scientific Research Collaboration. In *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work.* Lawrence Erlbaum Associates Publishers, 1990, 149-181.

Nardi, B.A., Whittaker, S., and Bradner, E. Interaction and Outeraction: Instant Messaging in Action, *Proc ACM CSCW* (2000), 79-89.
Rodden, T. Populating the Application: A Model of Awareness for Cooperative Applications. *Proc. ACM CHI,* 1996, 88-96.
Whittaker, S., Frolich, D., and Daly-Jones, O. Informal workplace communication: What is it like and how might we support it? *Proc ACM CSCW,* (1994).131-138.

Matthew McGill and Claude Sammut
The University of New South Wales

A scripting language for multi-modal user interfaces

After developing several multi-modal user interfaces for applications such as a museum tour guide, portable conversational assistant and intelligent entertainment centre, we created a scripting language that would allow the construction of modular scripts that could be reused in different applications. This feature was originally envisaged as allowing the development of script libraries that would enable new multimodal applications to be scripted more easily by allowing common interactions to be imported from a library. As a result the language FrameScript was created.

FrameScript is a multi-paradigm language for scripting of multimodal interfaces. Included in the language are rule-based processing, frame representations and simple functional evaluation. FrameScript has been developed for use in the development of multimodal interaction managers in multimodal applications. It does this by extending the language used in ProBot for scripting conversational agents to allow it to respond to any events that a system can detect or generate and can be represented as a frame. This allows scripts to be written that respond to not just spoken inputs but also to clicks from a mouse or touch screen, or even recognized gestures. It also allows scripts to be written to provide a system with a level of proactivity as the scripts can initiate interactions with users in response to system events, such as the arrival of a new email or the detection of a change in environment.

By accepting input in the form of frames FrameScript can provide interaction management in a multimodal system whose messages between components can be represented as frames. This allows it to be used with both EMMA and Mica multimodal architectures.

William Niu
The University of Sydney

My research interests include ontological reasoning, ontology learning, user modelling, and ubiquitous computing. As more and more services are available in ubiquitous computing environment, it gets harder for a user to effectively and efficiently find out the more relevant ones in different contexts. One way to address this problem is to model the user and his or her context. A context is any information that helps identify the user's situation. Extraction of contextual information is not always straightforward. This lead us to use light-weight ontologies to mine implicit contextual information. In order to minimise the laborious work in constructing ontologies, we propose to learn ontologies from different document sources as well as borrow parts of existing ontologies. The test-bed we have been using is our newly built School of IT building. The document sources being used include the building manual, the staff directory, and the course timetable.

Garth Paine
The University of Western Sydney

The Thummer(tm) Mapping Project - ThuMP

This paper presents the Thummer Mapping Project (ThuMP), an industry partnership project between ThumMotion P/L and The University of Western Sydney (UWS). A crucial step in the development of new musical interfaces is the design of controllable and sometimes externally observable relationships between the performer's physical gestures and the parameters that dictate the generation of the instrument's sound, a process called control mapping.

The ThuMP project is engaging in the development of a new electronic musical interface/instrument based on a re-evaluation of the performer's relationship with the performance interface. It sought to go back to examine musical interfaces that are broadly agreed to be successful and have persisted for a long time; acoustic instruments, namely, string and wind instruments and because of the nature of the Thummer. interface, the piano accordion and concertina. The ThuMP project posits that approaching the challenge of musical interface design from the musician's perspective might enable a detailed understanding of the subtle mechanisms of feedback and control that allow the support virtuosic technique.

Cecile L. Paris and Natalie Colineau
CSIRO ICT Centre

Tailoring as contributing to effectiveness in information seeking tasks

In our work, we look at providing users with the information they need, in a manner that is understandable and useful and to them. A major aspect of our approach is to tailor the information to users, in particular to a user's task and context. We see tailoring as contributing to the effectiveness of an interface. With the increasing amount of data and information available, finding what one is looking for is not an easy task. The idea of being able to provide to users information tailored to their needs (or preferences) is thus an attractive proposition. This has led to a substantial body of work in User Modelling, Adaptive and Recommender Systems, Intelligent User Interfaces (IUI) and Intelligent Multi-Media Presentation Systems. It is assumed that providing targeted information will result in a more effective communication.

While there has been a number of evaluations to test whether the customisation of information makes a difference to the user in terms of whether they prefer it to non-tailored information, or, in cases where the information provided was meant to influence behaviour, whether it led to more behavioural change than non-tailored information, the impact of tailoring on the efficiency of information seeking has not been tested. This specific question is the one we addressed in our work. We wanted to learn about the effectiveness of the tailoring on users' information task performance, and, in particular, whether or not having tailored information helps them finding the information they need. More generally, we look at identifying what in a user model (or in the context) can affect the system to make a more effective, useful and natural interface - with information and information spaces, in our case.

This issue arises not only in the context of textual interfaces but also multimodal and spoken interfaces, when a user interfaces with information. A number of elements can play a role, including tasks, the environment, emotions, etc. If they are to be exploited for a more effective interaction, issues of modelling and acquiring them also arise. A number of disciplines can contribute to this problem (as illustrated by the fact that the User Modelling community includes researchers from, at least, HCI, Language Technology, Psychology, Communication Science, Artificial Intelligence, Cognitive Science, Computer Science, Linguistics and Speech Science, Sociology, and other related fields.

David Powers, Darius Pfitzner, Kenneth Treharne, Martin Luerssen
Flinders University

Matching and Monitoring Channel Load

User interfaces are about transferring information. Even simple interfaces are multimodal and involve two-way communication and multiple dimensions. Miller's well known Magical Number Seven survey (1956) of cognitive limitations has spawned considerable exploration of the information capacity of individual dimensions as well as combinations of dimensions and/or modalities. However, user interface design has tended to neglect the cognitive capacity of the user and certain multimodal and 3D interfaces have been shown to reduce user performance rather than increase it.

The Flinders Artificial Intelligence Laboratory is focusing on this area from a number of perspectives: Human Factors study of utility and combination of static and dynamic visualization attributes for information retrieval; Human Factors study of choice and ranking of keywords for document description and search; EEG analysis of cognitive load during skill acquisition; Thinking Head study of role of emotional and expressional extensions of language in Human Computer Interaction.

David Powers, Trent Lewis and Richard Leibbrandt
Flinders University

Audio-Visual Speech Recognition and Synthesis

Our research on multimodal communication employs both acoustic and non-acoustic cues, using training to different speakers and noise sources, separation and identification of relevant and irrelevant sources, tuning of recognition systems to identified source factors, conditioning to different lighting, signal and noise conditions, use of syntactic patterns and fusion of multimodal information.

There is a synergy between our work with EEG in which muscular artefact is considered undesirable but forms a major part of the signal, and AVSR in which EMG signal correlates with visemic features. Similar techniques can be used to separate and identify multiple sources by combining ICA with source identification. We regard prosody, emotional content, source identity and characteristics as information that should be analysed in tandem with speech recognition/synthesis, each conditioned by the others. New fusion techniques are being applied across all kinds of information whether acoustic, visual, EMG, phonetic, prosodic, syntactic, semantic, emotive or identificatory.

Claudia Schremmer
CSIRO ICT Centre

Intense Distributed Collaboration for Creative Team Members in the Digital Media Production Industry: High-Quality Videoconferencing within a Personal Workspace

Collaboration over a distance has been a major research driver in telecommunication and high-bandwidth application development. If collaboration across remote locations is applied to teams of people rather than groups (a team is a purposeful group of people; in a team, people have specific roles, tasks, and objectives to achieve), the thrust on the mediating technology is generally higher (Daft et al, 1986). In the business project Virtual Media Office within the CSIRO Networking Technologies Lab, we investigate remote collaboration to teams of intensely interacting creative animation artists in the Digital Media Production industry. Based on field observations of how the digital animation artists interact with each other in the current co-located setting in Sydney's inner-city headquarters (Schremmer at al., 2006a), we have determined a first development project to create a new, user-status aware control interface to CSIRO's current implementation of the Virtual Tearoom high-quality videoconferencing system for use within the confined place of an artist's personal workspace. A video link within a personal workplace requires a higher degree of ambience of the communications technology; see (Schremmer, 2006b) for a more detailed discussion.

During the workshop, we'd like to discuss the concept of ambience in technology-mediated communication, tools to improve the perception of ambience (e.g., user-status information, design of user control interfaces, auditive feedback, visual .blurring. of a video connection), and methods to evaluate the effectiveness of ambient perception.
The Virtual Media Office project is aligned with the .HxI Initiative. (abstract submitted by Rudi Vernik, Belinda Kellar, Julien Epps, Claudia Schremmer) and .Evaluation Criteria from a Human Factor Psychological Perspective. (abstract submitted by Anja Wessels and Cara Stitzlein).

Rolf Schwitter
Macquarie University

My main research interest centres around controlled natural languages, in particular controlled natural language design and interface design for controlled natural languages. Controlled natural languages are well-defined subsets of natural languages that have been restricted to its grammar and its lexicon. Grammatical restrictions result in less complex and less ambiguous sentences, while lexical restrictions reduce the size of the lexicon and the meaning of words for a particular application domain. Using a machine-processable controlled natural language, texts become easier to read and understand for humans, and are easier to process for machines (in particular for reasoning services). However, it is an open research question, what kind of interface techniques best support authors writing texts in controlled natural language.

Over the last couple of years we developed an intelligent text editor for writing texts in a machine-oriented controlled natural language. The writing process is guided by text- and menu-based predictive interface techniques. This has the advantage that the user does not need to remember the rules of the language. Using this approach we can guarantee that the resulting text has the same formal properties as the underlying knowledge representation language. I am especially interested in getting feedback on our interface approach to write text in controlled natural language and in hearing about other effective interface strategies for knowledge acquistion, in particular in the context of the Semantic Web. Our ultimate goal is to find an interface strategy for controlled natural languages which supports lay persons in an unobtrusive way. In summary: what can HCI research offer here?

Ian Stevenson
The University of Western Sydney

Effective Interactive Interfaces for Musical Performance
Work on designs for new musical interfaces (Paine and Stevenson, 2005, Stevenson, 2005, Hewitt and Stevenson, 2003) has identified specific design criteria centring on evaluating models of musical communication in performance and understanding existing modes of physical performance interaction. The next phase of this research involves developing methods of evaluation to measure the effectiveness of these performance interfaces. The research proposes measures of accuracy, precision and ease of use based on control data extracted from the interfaces in performance conditions. A new measure of effectiveness based on a comparison of control data and analysed audio and video will be evaluated.

This presentation hopes to elicit ideas and discussion on evaluating physical interfaces based on analysis of cross-modal data sets.

HEWITT, D. & STEVENSON, I. (2003) EMIC - Extended Mic-Stand Interface Controller. *NIME03.* Montreal, McGill University.
PAINE, G. & STEVENSON, I. (2005) The Thummer Mapping Project – ThuMP. *Australasian Computer Music Conference.* Brisbane, ACMA.
STEVENSON, I. (2005) Design Issues for New Performance Systems. *Australasian Computer Music Conference.* Brisbane, Queensland University of Technology.

Ronnie Taib
NICTA, Sydney

I am interested in how and why people interact multimodally, depending on many factors such as the application, physical context and current cognitive load. In particular, I work on speech and gesture interaction (hands, head) and explore the semantic and temporal relationships that exist between them in user inputs. I also have interest on how multimodal output generation can be used to optimise the information acquisition by the user, especially when related to the cognitive load. I am involved in the design and deployment of user studies focusing on those aspects, in particular involving real-life scenarios and subjects from industry partners.

Some important issues:

- Evaluating the effectiveness of systems is an open issue in my view, as it is classically measured on error rates and task completion times. A user-centric approach is required, as affective and cognitive issues may impact the long term-performance of the systems. Some advances have been made when using educational methods;
- A related issue is about the design of test tasks that can induce predictable levels of cognitive load (low, medium, high) so that the performance of multimodal systems can be assessed during user studies. This field would benefit from inputs from the psychology domain;
- User customisation is probably on of the major issues faced by multimodal systems, since more modalities allow for a multitude of user profiles to be drawn. Introducing context awareness further expands it.

Dian Tjondronegoro
Queensland University of Technology

Research area: Multi-modal and Intelligent Interfaces, specifically:

- Personalized and context-aware delivery of video "on-the-move" contents with total users' control.
- Harmonization of multimedia contents (i.e. by optimizing the design, capture, editing, and visualization of video, images, text and sound) to make interaction appealing, immersive, intuitive and enjoyable.
- User-centred development of innovative applications that exploit emerging technologies, such as mobile devices and wireless Internet technologies. Refer to Figure below for examples of such applications.
- Using web-based multimedia platforms to enhance teaching/learning experience.
- Innovative user input methods for virtual immersive environments.

Important issues and potential points of interaction with other disciplines:

- While consuming mobile video, users should have total control over the play-back, non-linear navigation, and customizable personalized summaries [HCI]
- Multi-media can be used independently or synchronized effectively to attract human multi-senses and enrich cognitive tasks [cognitive and psychology science].
- To support intuitive services on wireless and mobile setting, we need to model and support users. behaviour, tasks and activities while interacting with the interface and collaborating with others [communication and sociology science]
- We can exploit gesture- or speech- recognition and other interaction methods, such as face recognition, to automate human inputs. For example, we can design an emotionally-aware living room that personalizes the music, lighting and wallpaper in the house [artificial intelligence]
- We can support enjoyable e-learning by incorporating virtual immersive environment (games-like) for university students. Conversational agents may impede non-linear and asynchronous teaching-learning environment, thus more emphasis should be placed on fully interactive and interruptible sequence of activities [Education]

Aiman Turani
The University of Sydney

Designing successful collaborative learning activities is a new focus of research within the e-learning community. The social dimension inside the traditional face-to-face collaborative learning is important and must be included in the online learning designs. There have been a number of projects that enable designing and implementing online activates, yet these projects are based on asynchronous collaborative technology. Designing asynchronous collaborative learning is applied on activates that take place over long period of time such as, group projects, PBL, etc but not for short term activities that are based on pedagogical techniques (Brainstorm, Debate, Buzz group, etc). However, generic synchronous tools are not always appropriate or sufficient for implementing the wide range of face-to-face pedagogical techniques in the online environment. These techniques might have complex internal structures associated with different roles. Therefore, there is an obvious need to structure and describe the synchronous collaborative learning in a formal way (collaboration scripting). This has been implemented in Beehive, a new web application framework for designing and supporting synchronous collaborative learning.

Anja Wessels, Cara Stitzlein
CSIRO ICT Centre

Evaluation Criteria from a Human Factor Psychological Perspective

Since any (computer-mediated) interaction is inherently social, the consideration of human factors is a prerequisite for the success of proper representation of the nature of interaction. Therefore, we would like to present a number of psychologically-derived quality criteria useful in the evaluation of "interaction effectiveness and efficiency". These criteria such as feedback, cognitive load, and motivation are borne from theory and observable in human-human interactions. For instance, feedback (as information) acts as a stimulus- improving attentional focus on relevant content. Within the interaction itself, it moderates motivational behaviour, realizable in outcomes of task performance. Feedback as an evaluative component implies that the computer-mediation (communication and coordination) provides sufficient channels for the sourcing, transmitting, and receiving of information necessary for successful interaction.

Apart from above mentioned quality criteria, these psychological insights shall be integrated in early developmental phases of new or modified ICT. A model, like the reference model used in Braccetto, enables a fit between several streams such as capability (i.e. technological work practice and content), work context (i.e. activities and artefacts), and teamwork mechanics (i.e. cognitive, motivational and conative processes with respect to the perceiving and acting bodies of the participants in the interaction). Using a model like this allows consideration of additional criteria from organizational and technical disciplines, providing a common language absent in some ICT agendas. In this way, evaluative frameworks include the human factors perspective.

Wayne Wobcke
The University of New South Wales

Our research focuses on an agent-based approach to dialogue management for speech interaction with a system of personal assistant applications, such as e-mail and calendar assistants. The dialogue model is implemented using a BDI agent architecture (JACK Intelligent Agents) and consists of a mixture of reusable domain-independent discourse level plans and task-specific domain plans. The basic decision of the dialogue manager is to select which plan to use in a given situation, based on the current user utterance, discourse history and salient items. We have also integrated the Alkemy symbolic machine learning system to enable the dialogue manager to adapt its responses to the user's preferences and context (e.g. the current location and device). The current version of the system runs on a PDA platform over a wireless network, using Dragon NaturallySpeaking for speech input.

Currently we are conducting a number of user studies to evaluate the prototype. Our experience so far has highlighted a number of issues regarding "effectiveness", including (i) the differences between task-focused usability evaluations and qualitative user satisfaction evaluations, including whether there is any meaning to correlations between task-completion rates and user satisfaction, (ii) the degree to which it is possible to evaluate a mobile application in a laboratory setting, (iii) the difficulty of evaluating specific components of a complex system compared to evaluating a whole system end-to-end, especially where we rely on black-box speech recognition software, and (iv) how to enable users to find the "right" language to interact with the system. Our approach has been to enable fairly "natural", though constrained, language (open vocabulary, extensive use of pronouns, but confined to e-mail and calendar tasks) with the aim that users will accommodate their speech input over time to the type of language accepted by the system; we also expect users will require some initial training. The work thus involves conversational agents, HCI, intelligent user interfaces and user modelling.