Some challenges for interface design
from an Intelligent User Interfaces point of view

Oliviero Stock
ITC-irst
Centro per la Ricerca Scientifica e Tecnologica
38050 Povo, Trento, Italy
stock@itc.it

Introduction

In this paper I will start with some synthetic comments originated from presentations given at two sessions, Emotions and Design and Memory and Narration. After having discussed some aspects I find relevant for the aforementioned themes, I will focus on two particular themes I believe symbolize what a novel class of human oriented, intelligent interfaces can be.

1. Discussion on Emotions and Design

In recent times Andrew Ortony, author of one of the most influential works in the field of cognition and emotions [A. Ortony, G. L. Clore and A. Collins, 1988] and Don Norman, probably the most influential author in interaction design, have taken a position similar among themselves regarding the relations between emotions and design. In synthesis, the designer has a motivational focus and a design focus (the latter being concerned with concepts such as utility and emotions). The user's emotion processing is of different levels: visceral, dealing with the appearance of the artefact; behavioural, concerned with functionality and usability of the artefact; reflective, concerned with the appeal, the image and the brand. They have elaborated their views along those lines.
Some comments of mine are summarized here below.
Probably to their view it is useful to add a temporal dimension. In fact emotions may be future-oriented, for instance as it occurs when you choose a product. Product selection is clearly influenced by immediate emotions, but in part one can talk of an expectation of direct emotional experience in the future. There are then short-term oriented emotions, typical of the moment of the direct use of the artefact. The key concepts here are usability, attention, engagement. Finally there is the recall of the emotional experience. In this case the timing is after the use of the artefact. Also this one is of vital importance because, for instance it leads to the establishment of fidelity toward a brand.

A second line of comments concerns the relation between computing and emotions, in particular the area of emotion processing modelling. There is a growing interest in Artificial Intelligence and emotions. As we are getting some insight about emotion processing, we try to model some aspects into intelligent interfaces. Areas are: a) detection of the user's emotional state, abducting it from some perceptive indicator; b) automatic dynamic emotion induction, as in the case of emotion provoking presentations: the system tries to obtain a certain emotional state of the user in a given context; c) experiencing emotions or –leaving aside philosophical questions- displaying emotion indicators, relevant if we are to build robots that exhibit a human-like behaviour.

A final comment regards personalization and emotion flow. In the future user modelling will encompass also the modelling of the personality and mood of the particular user and the dynamic emotion flow. It is certainly something scaring if privacy and ethical issues are not taken into account appropriately. On the other side the technology will offer that and there will be situations in which it will be useful.

In conclusion I believe there is a large potential for artefacts that are able to process emotions. We are only at the beginning of our understanding how emotions and rationality are to work together in a complex dynamic interface, but no doubt we shall have incremental advancements in this field.

2. Discussion on Memory and Narration

In his presentation, GiorgioDe Michelis explains that memory is narrative. Therefore support is needed for making memory accessible and "navigable"; tools are needed for updating and transforming it so that it is a fundamental instrument for the community. Ontologies constitute then a basic component for the realization of any support system of this kind. He elaborated his view making reference also to projects initiated under the umbrella of the I-cubed program of the European Fifth FP.

Some comments follow.

In the first place I would like to emphasize that time is the scarcest resource we have, and preserving and appreciating memory requires that we take into account factors centred on the potential beneficiaries resources: attention and contextualization, so that the relevant memory is linked to the specific situation and task the user is involved in.

Support for narration and accessing memory is something where intelligent information presentation can contribute substantially (see below). I mean systems that can provide automated narration tailored to the specific individual, taking into account his tastes and interests and that may take into account the context, the specific task and what has happened and has been reported up to then. Automated personalized report generation is one case that has some of the characteristics we are talking about. Some first experiments have been conducted in that direction.

Another aspect is concerned with the effectiveness of narration. If systems are to produce reports and narrate past events automatically, we want that the quality of the output is such that the audience is attracted and hooked up. Both initial attention and continuous engagement of the audience must be goals of the narrating system. We would like that the system is able to induce emotions in the audience. It must be able to build anticipation and then surprise, build tension and release tension. As Brenda Laurel emphasized, interfaces must have many of the characters of theatre [Laurel, 1991].

I would also like to emphasize the importance of multimodal presentation [see below]. Narration may include various modality of expression appropriately used and coordinated. In particular the use of images and even the automatic production of moving images coordinated with a given text are something that we are beginning to produce in form of prototypes [Zancanaro, Stock and Alfaro, 2003].

3. Intelligent interfaces and information presentation

What we want are interfaces that understand us, that are non intrusive, natural and powerful, that adapt to us, that help us focus our attention and memorize, and that are pleasant and entertaining.
Natural language as a means of communication is an obvious aspiration.
Natural language processing has been a focus of research for many years; it produced many ideas and potentiality in the area of interaction, especially with the development of the field of computational dialogue. The last decade has seen a transformation of the field, due mainly to two factors: a) availability of a large quantity of linguistic data, and dramatic increase of computer power and memory that allows to process them; b) introduction of short term competition in the field, basically imported from the speech research tradition. These factors have been at the basis of a more engineering oriented development, as opposed to an ambition of understanding cognitive processes, and specifically to the prevailing emphasis on statistical methods, rather then knowledge-based methods. Speech technology, a culturally different area, has in the meanwhile produced notable results, and speech recognition can be realistically be integrated in many interfaces. Yet we speak of limited dialogue capabilities, currently appropriate only for certain applications.
The Natural Language Processing community has contributed also to the emergence of intelligent user interfaces.
Here I would like to focus briefly on Intelligent Multimodal Information Presentation [see Stock and Zancanaro, in press].

At the root of the theme of Intelligent Information Presentation we can consider several scientific areas, but at least three are fundamental. Probably the first to be mentioned is Natural Language Generation, the branch of natural language processing that deals with the automatic production of texts. The field normally is described as investigating communicative goals, the dynamic choice of what to say, the planning of the overall rhetorical structure of the text (called sometime strategic planning), the actual realization of sentences on the basis of grammar and lexicon (sometimes called tactical planning), and so on. With a similar objective but with different means, the field of Adaptive Hypermedia combines hypertext (hypermedia) and user modeling. Adaptive Hypermedia systems build a model of the goals, preferences and knowledge of the individual user and use this throughout the interaction for adaptation of the hypermedia to the needs of the user. By keeping a model of some aspects of the user's characteristics, the system can adapt to and aid the user in navigating and filtering information that best suits his or her goals. A third important field is computer graphics; it has experienced a fundamental passage toward the end of the Eighties, when it was understood that graphics production should start from internal representations and communicative goals in a way similar to language production. This passage has led to the possibility of developing multimodal systems, that in output would consider the available modalities, possibly the context and the user characteristics, and operate so that the message is allocated and realized in a coordinated way on several media.
Intelligent Interactive Information Presentation has gone further along that line: it relates to the ability of a computer system to

automatically produce multimodal information presentations, taking into account the specifics about the user, such as needs, interests and knowledge, and engaging in a collaborative interaction that helps the retrieval of relevant information and its understanding on the part of the user. It may include dimensions such as entertainment and education, opening important connections to areas that were not related to the world of human-computer interaction, such as for instance broadcasting or cinematography. This vision has led to novel concrete aggregations. This is evident in a number of projects, in Europe as well as in America or in Japan, where the teams have included very diverse expertise.

Having briefly discussed Intelligent Interactive Information Presentation, I would like to get into some more specific elements of intelligence I think will be important for interface design: persuasion and humour.

4. Automating persuasive communication

Future intelligent interfaces will have contextual goals to pursue. They may aim at inducing the user - or in general the audience - to perform some actions in the real world. They will have to take into account the "social environment", exploit the situational context, and value emotional aspects in communication.

Some foreseeable scenarios of this kind are: dynamic advertisement, preventive medicine, social action. In all these scenarios rational reasoning is not enough. For intention adoption often what really matters is not only the content but the overall impact of the communication.

We want to provide the interface with the capability of reasoning on the effectiveness of the message, as well as on the high-level goals and content [Guerini, Stock and Zancanaro, 2003]. According to Perelman [Perelman and Olbrechts-Tyteca, 1969], persuasion is a skill that human beings use in order to make their partners perform certain actions or collaborate in various activities. Argumentation has often been considered as addressing similar points.

Persuasion is a wider concept, in our view: argumentation can be regarded as a resource for persuasion, while negotiation puts the accent on interactivity in argumentation.

In the first place it is a "superset" of argumentation: while argumentation is concerned with the goal of making the receiver believe a certain proposition, persuasion is concerned with the goal of making the receiver perform a certain action. The link relies on the fact that, apart coercion, the only way to make someone doing something is to change his beliefs.

"It is impossible to directly modify the Goals […] of an Autonomous Cognitive Agent. In order to Influence him (to modify his goals) another Agent is obliged to modify the former's beliefs supporting those Goals" [Castelfranchi, 1996].

In this prospect argumentation is a resource for persuasion.

The statement that there is more than argumentation in persuasion refers as well to the fact that persuasion is concerned also with a-rational elements. Examples are inducing emotions as a factor for obtaining a given result, or the use of specific language for threatening or promising. They all can be regarded as resources for inducing the receiver to act in a desired way.

Natural argumentation comes closer to persuasion, as it is also concerned, for example, with the problem of the *adequacy* - effectiveness - of the message. Even in professional setting, such as juridical argumentation, we know extra-rational elements can play a major role.

Persuasion mechanisms have to include the four following aspects:

1. The cognitive state of the participants (beliefs and goals of both the user and the interface)
2. Their social relations (social power, shared goals etc.)
3. Their emotional state (both the emotional state of the user and the one expressed by the system)
4. The context in which the interaction takes place.

The beliefs and goals of both the user and the system about the domain of the interaction: they are pre-requisite for a persuasive interaction to take place, since persuasion is a communication leading to belief adoption, with the overall goal of inducing an action of the user by modifying his pre-existent goals.
For instance in a museum guide system aimed at persuading visitors to see some exhibits (a theme we have explored within a large project devoted to cultural heritage appreciation, called PEACH, Stock, Zancanaro and Not, in press), we can instantiate all these elements. Social relations exist between the visitor and the system (the system playing the role of a competent guide) and between the visitor and other relevant persons such as experts, parents and so on. Emotional elements can enhance or lower the effectiveness of the message. In current work we focus on the role of the emotional state of the receiver (how it affects strategy selection) and on the emotion the system has to convey (express) to maximize the effectiveness of the message. Persuasion strategies can make use of contextual elements, e.g.. making reference to a painting the visitor has seen previously ("this painting is by the same author of …") can enhance the probability the user stops in front of the current painting.

5. Automatically generated humour

As for humor, I think we share the view that without it our species cannot survive. In future human-machine interaction, humans will demand a naturalness and effectiveness that requires the incorporation of models of possibly all human cognitive capabilities, including the handling of humor.
Computer-human interaction needs to evolve beyond usability and productivity. There is a wide perception in the field that the future is in themes such as entertainment, fun, emotions, aesthetic pleasure, motivation, attention, engagement and so on. Humour is an essential element in communication: it is strictly related to the themes mentioned above, and probably humans cannot survive without it. While it is generally considered merely a way to induce amusement, humour provides an important way to influence the mental state of people to improve their activity. Even though humour is a very complex capability to reproduce, it is realistic to model some types of humour production and to aim at implementing this capability in computational systems.

Humour is a powerful generator of emotions. As such it has an impact on people's psychological state, directs their attention, influences the processes of memorization and of decision-making, and creates desires and emotions. Actually, emotions are an extraordinary instrument of motivation and persuasion because those who are capable of transmitting and evoking them, have the power to influence other people's opinions and behaviour. Humour, therefore, allows a conscious and constructive use of the affective states generated by it. The affective induction through verbal language is particularly interesting and humour is one of the most effective ways of achieving it: the purposeful use of humorous techniques enables us to induce positive emotions and mood and to exploit their cognitive and behavioural effects. For example, the persuasive effect of humour and emotions is well known and widely employed in advertising. Advertisements have to be both short and meaningful, and are able to convey information and emotions at the same time.
Humour acts not only upon emotions, but also on human beliefs. A joke plays on the beliefs and expectations of the hearer. By infringing on them, it causes surprise and then hilarity. Jesting at beliefs and opinions, humour induces irony and accustoms people not to take themselves too seriously. Sometimes simple wit can sweep away a negative outlook that place limits on people desires and abilities, and makes people overcome self-concern and pessimism that often prevents them from pursuing more ambitious goals and objectives.
Finally, humour encourages creativity. The change of perspective caused by humorous situations induces new ways of interpreting the same event. By stripping away clichés and commonplaces, and stressing their inconsistency, people can be more open to new ideas and points of view. Creativity redraws the space of possibilities and delivers unexpected solutions to problems. Actually, creative stimuli constitute one of the most effective impulses for human activity.

In this context, computational humour deserves particular attention because of its potential to change computers into an extraordinary creative and motivational tool for human activity. Machines equipped with humorous capabilities become able to play an active role in inducing users' emotions and beliefs, and in providing a motivational support.

Deep modeling of humor in all of its facets is not something for the near future; the phenomena are too complex, humor is one of the most sophisticated forms of human intelligence. It is AI-complete: the problem of modeling it is as difficult to solve as the most difficult Artificial Intelligence problems. But some steps can be followed to achieve results [Stock, 2003]. And when something is realized we can note that humor has the methodological advantage (unlike, say, computer art) of leading to more directly falsifiable theories: the resulting humorous artifacts can be tested on human subjects in a rather straightforward manner.
Basically, in order to be successfully humorous, a computational system should be able to: recognize situations appropriate for humor; choose a suitable kind of humor for the situation; generate an appropriately humorous output; and, if there is some form of interaction or control, evaluate the feedback.

And indeed society needs humor, not just for entertainment. In the current business world, humor is considered to be so important that companies may hire 'humor consultants'. Humour can be used "to criticize without alienating, to defuse tension or anxiety, to introduce new ideas, to bond teams, ease relationships and elicit cooperation" [Binsted, 1996].

So, looking at computational humor from an application-oriented point of view, one assumption is that in future human-machine interaction, humans will demand a naturalness and effectiveness that requires the incorporation of models of possibly all human cognitive capabilities, including the handling of humor.

There are many *practical* settings where computational humor will add value. Among them there are: business world applications (such as advertisement, e-commerce, etc…), general computer-mediated communication and human-computer interaction [Morkes, Kernal and Nass, 1999], increase in the friendliness of natural language interfaces, educational and edutainment systems.

Not necessarily applications need to emphasize interactivity. For instance there are important prospects for humor in automatic information presentation. In the Web age presentations will become more and more flexible and personalized and they will require humor contributions for electronic commerce developments (e.g. product promotion, getting selective attention, help in memorizing names etc.) more or less as it happened in the world of broadcasted advertisement.

We are concerned with systems that automatically produce humorous output (rather than systems that appreciate humor). Some of the fundamental competencies are within the range of the state of the art of natural language processing.

In one form or in another humor is most often based on some form of incongruity. In verbal humor this means that at some level different interpretations of material must be possible (and some not detected before the culmination of the humorous process) or various pieces of material must cause perception of specific forms of opposition. Natural language processing research has often dealt with ambiguity in language. A common view is that ambiguity is an obstacle for deep comprehension. Most current text processing systems attempt to reduce the number of possible interpretations of the sentences, and a failure to do so is seen as a weakness of the system. The potential for ambiguity, however, can be seen as a positive feature of natural language. Metaphors, idioms, poetic language and humor use the multiple senses of texts to suggest connections between concepts that cannot, or should not, be stated explicitly. Fluent users of natural language are able to both use and interpret ambiguities inherent in the language and verbal humor is one of the most regular uses of linguistic ambiguity.

So far, very limited effort has been put on building computational humour prototypes. The few existing ones are concerned with rather simple tasks, normally in limited domains. Probably the most important attempt to create a computational humour prototype is the work of Binsted [1996]. Recently the first European project devoted to computational humour, HAHAcronym [Stock and Strapparava, 2003] part of the Future Emerging Technologies section of the Fifth European Framework Program was completed.

The main goal of HAHAcronym was the realization of an acronym ironic re-analyzer and generator as a proof of concept in a focalized but non restricted context. In the first case the system makes fun of existing acronyms, in the second case, starting from concepts provided by the user, it produces new acronyms, constrained to be words of the given language. And, of course, they have to be funny.
HAHAcronym is based on various resources for natural language processing, adapted for humour. Many components are present but simplified with respect to more complex scenarios and some general tools have been developed for the humorous context. A funda-mental tool is an incongruity detector/generator: in practice there is a need to detect semantic mismatches between expected sentence meaning and other readings, along some specific dimension (i.e. in our case the acronym and its context).
Testing the humorous quality of texts or other verbal expressions is not an easy task. There are some relevant studies though, such as [Ruch, 1996]. For HAHAcronym an evaluation was set with a group of 30 American university students. They had to evaluate the system production, along a scale of five levels of amusement (from *very_funny* to *not_funny*). The outcome was very encouraging. The system performance with humorous strategies and the one without such strategies (maintaining only syntactic correctness) were totally different. None of the "humorous" re-analyses proposed to the students were rejected as completely non-humorous. Almost 70 % were rated funny enough (without humorous strategies the figure was less than 8%). In the case of generation of new acronyms results were positive in 52% of the cases.
The results of the HAHAcronym project have been positive and a neat prototype resulted, aimed at a very specific task, but operating without restrictions of domain. It turns out that it can be even useful *per se*, but for us it is the starting point for getting more serious about humour. In the future we intend to work on sentences, and the idea is to eventually study how to introduce forms of dynamically produced humour in the conversation of an artificial embodied character.

Conclusions

Starting from the discussion of design and emotions and memory and narration, I have argumented in favour of a view that emphasizes intelligent communication capabilities in future interfaces. In particular I think the emotion dimension and the personality dimension must become part of what we will realize in individual-oriented and context-aware systems. I have introduced the field of Intelligent Information Presentation and then moved to two particular themes on which I have been working with my colleagues lately: persuasion and humor. In both cases some initial prototypes show that expectations of concrete practical results for interface design are not totally ridiculous.

Bibliography

K. Binsted Machine humour: An implemented model of puns. Ph.D. Thesis, University of Edinburgh, 1996

C. Castelfranchi Reasons: Beliefs structure and goal dynamics. Mathware & Soft Computing, 3(2):233--247, 1996

A. Ortony, G. L. Clore and A. Collins The Cognitive Structure of Emotions. Cambridge University Press, 1988.

M. Guerini, O. Stock and M. Zancanaro 'Toward Intelligent Persuasive Interfaces' Proceedings of the IJCAI 2003 Workshop on Natural Argumentation, Acapulco, 2003.

B. Laurel Computer as Theater Addison Wesley, Reading, Mass., 1991

J. Morkes, H. Kernal, and C. Nass. 'Effects of humor in task-oriented human-computer interaction and computer-mediated communication: a direct test of SRCT theory'. Human-Computer Interaction, 14:395-435, 1999.

C. Perelman and L. Olbrechts-Tyteca The new Rhetoric: a treatise on Argumentation, Notre Dame Press, 1969.

W. Ruch Special issue: Measurement approaches to the sense of humor. Humor: International Journal of Humor Research, 9, 1996.

O. Stock 'Password Swordfish: Verbal Humour in the Interface' Humor: International Journal of Humor Research. Vol 16, n.3. pp. 281-295. Mouton de Gruyter, 2003.

O. Stock, and C. Strapparava 'Getting Serious with the Development of Computational Humor' Proceedings of IJCAI 03, the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, 2003.

O. Stock, M. Zancanaro (eds.) Intelligent Multimodal Information Presentation. Kluwer Academic Publishers, Dordrecht, in press.

O. Stock, M.Zancanaro, E.Not 'Intelligent Interactive Information Presentation for Cultural Tourism' in O. Stock, M. Zancanaro (eds.) Intelligent Multimodal Information Presentation. Kluwer Academic Publishers, Dordrecht, in press.

M. Zancanaro, O. Stock and I. Alfaro, 'Cinematic Techniques for Mobile Presentations' in L. Chittaro (ed.) Human-Computer Interaction with Mobile Devices and Services. Proceedings of 5th International Symposium, Mobile HCI 2003, 384-389, Springer Verlag LNCS, Berlin, 2003.