Affordable Affective Avatars. Persuasion, Emotions and Language(s)

Boldur E. Bărbat Radu Cretulscu

Department of Computer Science and Automatic Control
"Lucian Blaga" University of Sibiu, 2400 Sibiu, Romania
Email: bbarbat@rdslink.ro, kretzulescu@isjsibiu.ro

Τό υφεστηκός δεί τέλος επιλογίζεσθαι
(We must take into account the goal sought)
Epicurus

Abstract. *P*ersuasive applications become vital. The rationale is: from *effectiveness* to *persuasion*, to *credibility*, to *emotivity*, to bio-mimetic *avatars*, to *human-like expression* of affective states. Economical restraints imply *affordable* characters, i.e. simulating emotions using traditional equipment. After summarizing the approach and the architecture described in previous papers – and the main software-engineering mechanisms – the paper focuses on perceptual interface affordability, based on implementing context-relevant affective features through usual devices and simple sensors. Emphasis is put on the rationale for blending speech and non-verbal communication in expressing emotions. The design-space dimensions are proposed accordingly. Among the conclusions: a) basic affective computing can be achieved using common configurations; b) due to their undemanding implementation, the agent-oriented mechanisms employed can be used as test-bed; c) the proportions of the language blend depend strongly on the *intended use*. Future work: a) affective computing functionality to existing persuasive avatars; b) other common devices (e.g., mobile phone).

1 Introduction

In the age of vigorous expansion of diverse web-centred applications, since *persuasion* is essential for *effectiveness* [1] in most fields involved (from e-commerce to e-therapy), captology (*C*omputers *A*s *P*ersuasive *T*echnology [13]) becomes vital too. However, to be *convincing* (in any sense), an entity must be *credible* (in both connotations: *trustworthy* and *expert*). Still, from ancient Greek rhetoric to modern advertising, credibility and *emotion* have been inextricably linked [3]. "The emotionless character is lifeless. It is a machine" [10]. Moreover, the failure of "computers to recognize, express, and have emotions severely limits their ability to act intelligently and interact naturally with us" [24]. Condensing the inference chain, the need for persuasive agents able to reveal emotions – and to control them likewise to humans – is entailed by direct implications: from *effectiveness* to *persuasion*, to *credibility*, to *emotivity*, to *human-like expression* of emotional states (including plausible progression in time); all that requires *multimodal interfaces*.

On the other hand, since current economical restrictions in our countries imply *affordable* artificial characters, emotions must be simulated through agent-based software using only conventional equipment. Thus, the main target is to bring closer such divergent requirements focusing on refining the emotional *reaction* (through pathematic agents) rather than on catching more emotional *stimuli* (through advanced sensors). For short, less hardware.

Since related work was outlined recently in [6,7] and, using a more general outlook – the role of agents in anthropocentric systems – in [5], to impair redundancy, the rest of the paper is organized as follows: Section 2 summarises its rationale, evolution, and approach. Section 3 emphasises the rationale for properly blending speech and non-verbal communication in expressing avatars affective states. Accordingly, the main aspects of pertinent design-space dimensions are proposed in Section 4. On this ground, Section 5 describes the current implementation. Conclusions and intentions are closing the paper.

2 Rationale, Evolution, and Approach

Rationale. In fact, the previous inference chain outlines the rationale; thus, here are only some additional links, regarding *affective computing* (AC) and *multimodal interfaces*, respectively.

The results of the first loop of the Delphi survey about the future of human-computer interaction set off by the IFAC Technical Committee on Human-Machine-Systems in 2002, gives 2010 as the likely year for the “state of the art” in *affective engineering* (technologies for recognizing emotions and for expressing them). Since users as well as developers are not quite ready for this recent challenge, raising their awareness is helpful. The capacity of most effective anthropomorphic agents to adapt their behaviour to those they encounter and to personalize their own conduct, speech, and gesture, can send emotional messages that affects the user's subconscious as well as conscious experience (crucial for e-commerce [15]). From a long-range perspective, it seems to be necessary to investigate “affective computing”, since “Much scientific evidence has come to light that emotions go far beyond their traditional stereotypical affects of irrationality and illogical behaviour in humans to be the foundation of rational thinking, perception, learning and other intelligent cognitive functions […]. Understanding emotional generation and developing a reasonable empirical model may hold the key to unlocking the potential of affective reasoning and decision-making in machines” [25]. In such a complex portrait of theories, it is crucial “to understand affect from a computational perspective” [30] (presenting Emotion-Based Control “as an alternative approach to the study of affective phenomena”). From an anthropocentric viewpoint, the raison d'être of AC is to provide a vehicle for “digital rhetoric”, vital in the Internet age.

From another broader viewpoint, *perceptual user interfaces* make human-machine-interaction (HMI)[1] “more natural and compelling by taking advantage of the ways in which people naturally interact which each other and with the world – both verbally

---

[1] We do not use the term “HCI” only to avoid the debate regarding the name – or architecture – of future devices, facing the “cellular-palmtop-convergence”.

and nonverbally" [29]. Moreover, it is crucial to intertwine the ways to *communicate* emotion. Accepting that communication (and primarily its phatic function, as well as its persuasion role) is eased through (the illusion of) lifelike features, an effective interface cannot survive only with text and image; instead it has to be anthropocentric (mainly, based on perception and multimodal; this is the specific rationale for Section 3).

Evolution. As regards the chronological evolution of this undertaking, from an agent-oriented (AO) architectural framework for adaptive medical applications, to captological agents showing traces of emotivity, it was summarized in [3]. Afterwards, the architecture for pathematic agents has been extended to exhibit different ethical attitudes (controlled by the physician in charge) and to assign them personality. This feature was modelled in two steps: a) A functional psychological type, the target of [4]. b) An identifiable personality (e.g., benevolent nurse, authoritative physician), aimed at in [6]. (Since for e-commerce a "Machiavellian" agent has been proposed [18], why not a "Hippocratic" one, for therapeutic persuasion?) Next, the focus was on believable agents, exploring a key source of credibility: human-like expression of emotional states [7]. All the phases mentioned, were illustrated by a test-bed application where an avatar tried to persuade a smoker to quit this harmful habit – except the last one [9], addressing paediatric purposes. The application was implemented in line with Fogg's ideas regarding the design space for persuasive technologies [13] and presented in its original structure, when captology was firstly applied to medical informatics [2].

Approach. For the quit-smoking application the approach – exposed in depth in [6] – was based on "micro-continuity" (pragmatically, a stepwise proceeding is safer and cheaper) at both *conceptual* and *implementation* levels. Despite being workable – a) all application versions endorsed the generic architecture they originated from; b) designing them according to the proposed guidelines was easy and unambiguous; c) although lacking logistic support to become commercial, the applications were relevant for potential users – the approach has been changed [9]. The reason: the leitmotiv question – natural for any applied research – "How many persons have been persuaded by the virtual therapist to quit smoking?" remained unanswered. The approach seemed to offer a solution searching for a problem. To be able to pass from a test-bed to an application having a measurable degree of end-user acceptance, the approach was amended, towards a strikingly modular, simple, fine-grain, and fast adjustable architectural framework, embracing a large palette of prospective uses. The research test-bed is not anymore an instance of a generic architecture, but a toolkit containing many small modules. Thus, due to small module size, testing is far more efficient, whereas compactness can lead to avatars carried out as hand-held technology or embedded in "ambiental intelligence".

Such an approach is suiting an undertaking aiming at affordability: the avatar is understood – and implemented – in its narrow meaning (a simple intelligent agent embodied by a bio-mimetic character), not in the broad one (a pseudo-being extending the genuine human).

3 Orality. Pros and Cons

The difficult problem of finding the best means to interact with avatars became recently more acute, due to the stage of a unique positive-feedback loop: WWW and avatars boost simultaneously the *requirements* as well as the *possibilities* (including the technological infrastructure) for intense communication. New, vigorously expanding interaction modalities – from ubiquitous computing to virtual reality – involve new kinds of communication means (from eudemonic interfaces to avatars). Thus, it is time for revisiting the effort towards intense dominance of oral communication in HMI. Since the scrutiny yields suggestions somehow against common opinion, the pros and cons of using spoken language are given in detail, following a former (more general) investigation [8, 11] but focusing on the emotional aspects.

What means "language(s)"? In principle, any system of signs, socially codified, used for communication; here, only those used in HMI. The undertones are as usual: the singular brings to mind "tongue" (orality), whereas the plural suggests all the others. The problem core is neither to give up the intent to use spoken language (since its convenience is obvious) nor to eliminate completely other languages (since their occurrence is inexorable), but to find the right blend, in particular to debate upon the overwhelmingly belief that orality is a panacea and, consequently, it should take the lion share in future (affective) interfaces. Moreover, the users are facing two paradoxes: a) it is hard to believe that just the (unsympathetic) technological advance enables a better fulfilment of their aspirations; b) some of the (seemingly new) languages they are afraid of, are older and, hence, "more natural" than the verbal one (for instance, the body language can be used again in telecommunications, after its century long lockout through telephone and radio). Indeed, the tacit premise that, apart from orality – as pinnacle of "naturalness" – most other communication means are artificial, hence potentially harmful, should be amended. Thus, the rest of this section is shaped as two batteries of arguments. The first covers the (more) stable reasons for (or against) orality, stemming from the persistent anthropocentric requirements; the next lays out contextual technological factors perceived rather as acting as triggers (or inhibitors) for the former group.

Human-centred arguments.

a) Since spoken language is the cornerstone of inter-human communication, it is expected to play a crucial role in communication between humans and *any* entity or machine.

b) "What makes humans different from all other animals? […] our ability to recognize and to produce speech. Speech is so much part of being human that people with IQ scores as low as 50 or brains as small as 400 grams fully comprehend and speak at a competent level (Slobin, 1979)" [20].

c) In the context of complex and/or vital interaction – where the user is too "hands-busy, eyes-busy" to resort to conventional interfaces – voice command becomes essential (for many handicapped persons it is the *only* way out; games follow soon).

d) "Speech has proved useful for store-and-forward messages and for alerts in busy environments […] and shows promise for telephone-based services" [27].

e) „Who are you? Your voice alone can be used to verify your personal identity – unobtrusively and invisibly. "It's me!" embodies the expectation that the sound of one's voice is sufficient for the hearer to recognize the speaker" [19].

Now some (less recognized) general objections:

a) The gap between the *atemporality* of text and image, and the *protensity* (*evolution* i*n time*), of sound[2] provides new worth to “verba volant, scripta manent”. Communication has to be lasting and stable until it was assessed. Robustness implies either more attention, or less speed, or repetitions.

b) „Spoken language is effective for human-human interaction but often has severe limitations when applied to human-computer-interaction. Speech is slow for presenting information, is transient and therefore difficult to review or edit, and interferes significantly with other cognitive tasks” [27].

The following four reasons are just *some* instances of b).

c) Besides the background noise interfering with speech, occurs also the noise brought in by the speaker through unarticulated sounds like „ummm” or „uh”.

d) Inter-individual differences (including those caused by dialects, bad health, emotion, etc.) can reduce communication effectiveness, sometimes beyond an acceptable threshold.

e) “The part of the human brain that transiently holds chunks of information and solves problems also supports speaking and listening. Therefore, working on tough problems is best done in quiet environments – without speaking or listening to someone. However, because physical activity is handled in another part of the brain, problem solving is compatible with routine physical activities like walking[3] and driving. In short, humans speak and walk easily but find it difficult to speak and think at the same time” [27]. In computer jargon: sharing the same memory shrinks drastically the possibility of parallel processing.

f) This hindrance has repercussions also on others, destroying the (physical and psychical) microclimate necessary for effective work. (Imagine a research laboratory with a few machines with vocal input!)

g) As the rest of nature, humans are *multimodal* they use a blend of concurrent communication means based on at least two of the main channels: visual, auditory, and haptic.

The following five reasons are again *some* instances of g).

h) “A primary goal of speaking is to be understood. […] humans use "hyperarticulate" speech (such as increased pauses and elongated words) when encountering people with comprehension difficulties” [20]. However, this “Lombard effect” [23] – is often not enough: we have to resort to other languages – those given by nature to *all* animals, mainly gestures. “During multimodal communication, we speak, shift eye gaze, gesture, and move in a powerful flow of communication” [23].

i) Feeling that we have a weaker voice than other animals, we use it rather at short distance, whereas gestures are the oldest “long distance” communication means.

j) The “Perceptual Bandwidth” – jargon name of the rich palette of sensorial experiences accessible to man [26] – generates amazing synergy: the reaction time to a combined stimulus is noticeable under the minimal value for separate stimuli.

k) Humans are very proficient in discriminating messages received simultaneously. (The “cocktail party effect” is a notorious, albeit unimodal, example.)

---

[2] In any context; essential for *feeling* (in the case of *music*) or for *understanding* (in the case of *speech*).

[3] Remember Aristotle’s “peripatetic” school. (The authors’ comment.)

l) Perceptions are processed holistically, as demonstrated by the McGurk Effect: “syllables "sound" different if the lips suggest a different sound. Humans tend to integrate voices and faces when they are combined. People "hear lips and see voices"” [20]

m) Because of the tendencies of humans to personify technology, trust specialists, and form relationships [15] the danger of over-expectations is huge. “The basic problem is that if the system sounds too much like a human, the user can reasonably expect it to understand like a human” [12]. “Individuals behave toward and make attributions about voice systems using the same rules and heuristics they would normally apply to other humans” [20].

n) Those “tongue-independent” objections, supplemented by the Babel-Tower hurdles. Just one example how multimodal interfaces can diminish them: an already internationally widespread gesture is the simple place holder for “quotation marks”…

Technological factors. Despite remarkable achievements (*VoiceXML*, voice biometrics, etc.), it is more relevant to see the technological limitations, because “Human-human relationships are rarely a good model for designing effective user interfaces. Spoken language is effective for human-human interaction but often has severe limitations when applied to human-computer interaction” [27]. The (obvious) cause: the patterns and the dynamics of HMI relationships are too different from those of inter-human ones.

a) “The technology continues to improve but does not yet come close to approximating the speech-recognition capacity of an adult human, or even a teenager” [17]. This general assertion splits into several types of failure, three of them being illustrated below.

b) What computers are recognizing are rather *words* (sound waves) than *speech* (genuine understanding). The systems “are likely to respond to fluently spoken dialogue only if it falls within their specialized task domains” [12].

c) “Spoken dialogue systems involve recognition-based technology that is probabilistic in nature and therefore subject to misinterpretations. The Achilles’ heel limiting widespread commercialisation of this technology is the rate of errors and lack of graceful error handling [22].

d) Even worse: “current estimates indicate a 20% - 50% decrease in recognition rates when speech is delivered under the following conditions: during natural spontaneous interaction; by diverse speakers (such as those with accents); or in a natural field environment” [22]. Unfortunately, thus, the value of reason “pro” c) is severely reduced, since in industrial environments interaction is spontaneous, speakers are very different, and the noise level is high.

e) Many of the recognition difficulties lie not in technological failures, but in the very nature of speech. Although made up by phonemes, speech is based also on intonation, pauses, timbre, that can alter its meaning. “The syntactic aspects of prosody, such as rising tone for questions, are important for a system’s recognition and generation of sentences” [27].

4 Design-space Dimensions for Affective Avatars

The design space for captological agents described in [2] has been enhanced with the dimensions/aspects employed in [6]. Those regarding emotivity and ethics [3,7] are implemented, the other three are only modelled; all of them are mirrored in embryonic form in an anthropomorphic avatar (James) and in a zoomorphic one (Teddy), both presented in Figure 1. Here the design space is extended and refined in line with the paper's target: a) the dimensions relating to emotivity are divided into three categories: *emulating* emotion, *stimulating* it, and *expressing* it; b) all aspects can be implemented with affordable configurations and mechanisms.

Figure 1a. James (calm)

Figure 1b. James (upset)

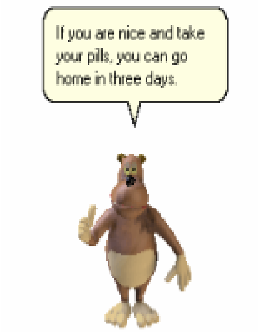

Figure 1c. Teddy (calm)

Figure 1d. Teddy (upset)

Fig 1. Avatars in two emotional states

In this context, Frijda's four phases of "emotional involvement" [14] get different weights: *Perception* (of events relevant to the individual's concerns) maintains its usual importance, but focus is rather on *accidental factors* (e.g., an instant triggering event: if the child starts playing during the dialogue with the e-therapist), than on *taxonomical elements* (e.g., number of pills taken). *Reaction* (based on inner goal repercussion) has enhanced significance, since it is vital to test the emotional behaviour. *Reasoning* (evaluating implications of changes, by inferring new knowledge from new facts) becomes somewhat irrelevant. *Action* (based on the former evaluations and according to resource availability), although not essential per se, is used as an "exteriorisation amplifier".

Emulating emotion. Basically, the avatar must be driven by intentional features – for both actions and attitudes (otherwise the illusion is broken). Likewise, adaptive attitudes are needed, since real-time interaction involves unpredictability (when the interaction is not pre-scripted, not every situation can he foreseen – at least, at the avatar design stage). Adaptivity is vital for the avatar to interact with humans in a compelling manner (most of all when the interactant is a child).

Whilst Frijda's appraisal theory offers the framework, the specific course of action remains based on the categories of emotional reactions derived from the OCC cognitive theory of emotions [21]. In line with this point of view, emotions are *induced by actions* rather than by *consequence of events.*

Personality is conveyed in everything from an avatar's movement to its choice of message. It "is really a holistic package of consistent attributes all working together to contribute to an overall impression for the user" [15]. It is reached in two stages: a) psychological type: based on Jung's theory, the functional attitudinal type is that of the practical brains [5]; b) a distinct, identifiable personality: Hippocrates, Merlin, Donald, grandmother … [6].

Persuasion time horizon: now seconds for avatar "emotions", next longer spans for "moods" and "feelings" [7]).

Stimulating emotion. A result of the required context-sensitivity: emotional *stimuli* (at architectonic level) have to be treated as *exceptions* (at implementation level). Hence, almost all affective responsiveness must be included in exception handlers, with two valuable "human-like" side effects: asynchronous and non-deterministic emotion emergence [3,7].

Obviously, emotion is triggered by patient behaviour. To remain affordable, user behaviour is perceived through usual devices (e.g., mouse) and simple sensors: in Figure 1, James was upset by a smoke detector whereas Teddy, by a motion detector.

Expressing emotion. The main *languages* are: a) *Animation*. Deeply rooted in anthropogenesis, the ability to connect motion to life was a question of survival [15]; hence, animated avatars are more credible. b) *Gesture*. "Animation is movement; gestures are movements that contain information" [16]. Body language can and shall be added to verbal messages. c) *Spoken language*. The avatars use synthesized speech (Text-To-Speech Engine) and text in cartoon word balloons.

Protensity. Expressing emotion is linked not only to its duration, but also to the staying power of its impact – in terms of perseverance, persistence, endurance, etc. If emotion is expected to start rather abruptly, its decay is slowed down (the rate depending on many factors); however, if emotion just vanishes or is switched off, it is visibly faked. Similarly, its start moment is uncertain since the first two phases of "emotional involvement" (perception and reaction) are fuzzy [7].

5 Current Implementation

Since test-bed applications are rather *grown* (from a design-space kernel) than *built* and implementation issues (mechanisms, scenario-chunk instances, development tools) have been detailed in [3,7], here is summarised only the current stage [9] of the mechanisms implemented previously.

- *Exception handling.* To achieve responsiveness to stimuli (since triggers are *exogenous*) and *consistency* of communication (since they are *asynchronous*), the system is interrupt-driven.

- *Clone-based polymorphism*. The cloning process (moments, pace, clone dissimilarity, etc.) is managed by the emotion exception handler.

- *(Dynamic) thread priorities*. Since agents are both *autonomous* and *lasting*, a simple way to "humanise" their emotions is to emulate the "dormant" human behaviour features, emerging only when impelled. As each clone is supposed to represent a (slightly) different mental state of the initial, "emotionally normal", avatar ("clone#0"), all clones are threads coexisting from the beginning, but with lowest priority (or suspended). The only significant hindrance to priority-driven emotion control is the restricted range of available priorities: by Windows 2000/XP only five (when the implementation is kept within a single process). That is why: a) the composed avatar has only four clones; b) "dynamic" is still in parentheses.

- *Flexible temporal windows*. As regards emotion progression, to reflect its asymmetry, the rise is expressed by shifting fast ($\Delta t = 0.5s$) from clone#0 (Figure 1a, 1c), through clone#2, to clone#4 (Figure 1b, 1d), whereas emotion fall involves all clones in slower succession ($\Delta t = 2.5s$). Intensity itself is expressed in multi-modal way, altering some of the avatar parameters: voice (pace and pitch of utterance), dimensions, limbs position, speed of moving across the screen, message conveyed.

- *Synergistic interaction*. Seen as black box, the *FETCH* (*F*uzzy *E*motion *T*ime-*C*ontrol *H*andler) block is identified through an input constituted by exogenous and asynchronous events (and/or sequence of events) and an output consisting of temporary behaviour alteration. The problem is neither fuzziness, (since the start moment is not critical) nor emotion increase (since the rise is abrupt), but its lengthy fading away (with only four clones it is hard to attain the granularity required to return gradually to the initial discourse). Hence, the exception handler interrupts the discourse of clone#0, resumes clone#2 (who is silent), and then clone#4 (who utters an emotional message). Now, clone#3 comes in with a phrase marking the return to the first discourse, clone#2 is again silent, and clone #1 resumes the persuasion.

Other features: a) At the beginning, the physician can select a persuasion strategy for the upcoming session, (a "carrot" or a "stick" strategy, both with three strength levels). b) The patient has (very limited) options to personalize the interaction with the avatar. c) Corollary: to discriminate the access rights of physicians, nurses, parents, and children, according to profiles and levels (vital for telemedicine) – a logon procedure was added.

6 Conclusions and Intentions

Since the paper treated an approach and an instantiation this section is divided into factual conclusions (A) and broad-spectrum ones (B).

A1. Primitive characters showing emotional features are relevant in shrinking the gap between the *appearance* of artificial emotions and of human ones. Due to the undemanding implementation on usual configurations, the avatars (mainly the agent-oriented mechanisms employed) are useful as test-bed for any aspects of human-like avatar behaviour.

A2. As regards emotions, three temporal aspects must be considered in concert: a) Time range (seconds or minutes). b) Asymmetrical intensity variation (especially, gradual decay rate). c) Fuzziness (mainly, fuzzy interval between triggering events and emotion emergence).

A3. An effective way to express the relationships between emotions and time is to assign agent-oriented semantics to mechanisms used hitherto as technical means (mostly, to improve run-time efficiency). Relevant results can be achieved with usual configurations and extensive multithreading.

B1. Basic affective computing – at least, simple avatars simulating emotions – can be achieved using common interfaces and software.

B2. The proportions of the language blend depend strongly on the *intended use* and – in a gradually diminishing degree – on the *devices involved* because nature – humans included – was multimodal from the very beginning, and interfaces started recently to become too. Thus, humans can afford to remain *natural* and build avatars similarly.

B3. However, to improve performance radically, genuine perceptual interfaces are needed, since the main limitation is the "perceptual bottleneck".

The main *short-range intention* is to design and implement limited non-deterministic pathematic avatar behaviour (by refining the dynamic-priority mechanism and implementing variants of fuzzy temporal windows). Later, *middle-range targets* are:

- Adapting the mechanisms to longer time spans (minutes), to express agent "moods".

- Extending the approach to cover: a) other common devices (e.g., mobile phone); b) various aspects of persuasion strategies (starting with the "carrot and stick" one); c) the relationship between artificial emotions and deontological ethics in captological applications (first in medical informatics, afterwards in e-commerce).

References

1. Andersen K. *Persuasion Theory and Practice*. Allyn and Bacon, Boston, (1971)
2. Bărbat B., C.B Zamfirescu, G. Costache. Agent-Oriented Captology for Medical Informatics. In: *Medical Infobahn for Europe. Proceedings MIE & GMDS* Conferences (2000) 465-469
3. Bărbat B.E. Pathematic Agents for Medical Informatics. *Proceedings International NAISO Congress on Information Science Innovations (ISI)* 2001, 647-653
4. Bărbat B.E. Captological Agents in Therapy, Between Intentions and Behaviour. *Proceedings 10th World Congress on Medical Informatics (MEDINFO),* (2001), 571
5. Bărbat B.E. Agent-Oriented Captology for Anthropocentric Systems. In *Large Scale Systems: Theory and Applications* (F.G. Filip et al., eds.), Elsevier, IFAC Publications, (2001) 214-219
6. Bărbat B.E. Hippocratic Agents. Rationale and Design Space. *Proceedings MIE Special Topic Conference on Healthcare Telematics Support in Transition Countries*, EUROBIT, Timişoara, (2002) 92-98

7. Bărbat B.E. Emotions and Time in Captological Agents. *Proceedings 3rd International NAISO Symposium on Engineering of Intelligent Systems*, (abstract; paper on CD-rom), (2002)
8. Bărbat B.E. From Languages to Language and Back. Orality in HCI. *Proceedings COST-269 WG "Extended Human"* meeting, Montegrotto, *http://www.cost269.org/documents* (2002)
9. Bărbat B.E., D. Luca. Emotional Disneyland Characters for Paediatric Purposes. Accepted at *MIE2003 Medical Informatics Europe. The New Navigators: from Professionals to Patients*, St Malo, (2003)
10. Bates J. The Role of Emotion in Believable Agents. *Communications of the ACM,* special issue on Intelligent Agents, 37(7):122-125, (1994)
11. B Bărbat B.E., Crețulescu R.. Digital Ethics. Agents, Between Machiavelli and Hippocrates, The Good, the Bad and the Irrelevant, (COST 269) Helsinki, (2003)
12. Boyce S.J. Natural Spoken Dialogue Systems for Telephony Applications. *Communications of the ACM*, 43(9):29-34, (2000)
13. Fogg B.J. Persuasive Technologies. *Communications of the ACM*, 42(5):27-29, (1999)
14. Frijda N.H. *The Emotions*. Cambridge University Press, (1987)
15 Heckman C.E., J.O. Wobbrock. Put your Best Face Forward: Anthropomorphic Agents, e-Commerce Consumers, and the Law. *Proceedings 4th International Conference on Autonomous Agents*. Barcelona, (2000) 435-442
16. Kurtenbach G., E.A. Hulteen. Gestures in Human-Computer Communication. In *The Art of Human-Computer Interface Design* (B. Laurel, ed.), Addison-Wesley, Reading, MA (1990)
17. Lai J. Conversational Interfaces. *Communications of the ACM*, 43(9):24-27, (2000)
18. Litzinger T., J. Wise. E-Commerce: the Virtual Battlefield. *Proceedings 4th IFIP/IEEE International BASYS Conf*erence (2000), 227-234
19. Markowitz J.A. Voice Biometrics. *Communications of the ACM*, 43(9):66-74, (2000)
20. Nass C., L. Gong. Speech Interfaces from an Evolutionary Perspective. *Communications of the ACM*, 43(9):36-43, (2000)
21. Ortony A., G Clore, A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, (1988)
22. Oviatt S. Taming Recognition Errors with a Multimodal Interface. *Communications of the ACM*, 43(9):45-51, (2000)
23. Oviatt S., P. Cohen. Multimodal Interfaces That Process What Comes Naturally. *Communications of the ACM*, 43(3):45-53, (2000)
24. Picard R.W. *Does HAL Cry Digital Tears? Emotions and Computers. HAL's Legacy: 2001's Computer as Dream and Reality*. D.G. Stork (ed), MIT Press (1997)
25. Ray P., M. Toleman, D. Lukose. Could Emotions be the Key to Real Artificial Intelligence? *Proceedings ICSC Symposia on Intelligent Systems and Applications (ISA)*, CD-ROM, (2000)
26. Reeves B., C. Nass. Perceptual Bandwidth. *Communications of the ACM*, 43(3):65-70, (2000)
27. Shneiderman B. The Limits of Speech Recognition. *Communications of the ACM*, 43(9):63-65, (2000)
28. Slobin D. *Psycholinguistics*. 2nd Edition. Scott, Foreman &Co., Glenview, IL, (1979)
29. Turk M., G. Robertson. Perceptual User Interfaces. *Communications of the ACM*, 43(3):33-34, (2000)
30. Velásquez J.D. From Affect Programs to Higher Cognitive Emotions: An Emotion-Based Control Approach. *Proceedings 3rd Int*ernational *Conference on Autonomous Agents*. Seattle, WA, (1999)