

# Merging Cognition and Affect in HCI

(Guest Editor's Introduction)

*Fiorella de Rosis*

Intelligent Interfaces

Department of Informatics, University of Bari

derosis@di.uniba.it

E' importante considerare gli aspetti affettivi del ragionamento umano, nella costruzione di agenti intelligenti? O il crescente numero di iniziative in questo campo non è altro che una moda, destinata ad essere superata in favore di altri slogan?

Noi crediamo che una moda non sia, ma che si tratti di una esigenza reale. L'accettabilità di un sistema è sempre più legata alla sua capacità di interagire con l'utente in modo 'naturale'. E comportarsi in modo naturale richiede la capacità di riprodurre forme di ragionamento più complesse di quello puramente logico e modelli di decisione non puramente 'razionale'. Anzi, richiede (come propongono Lisetti e Gmytrasiewicz in questo Issue) di rivedere il concetto stesso di 'rationality' by including, in this term, the role that emotions and attitudes<sup>1</sup> play during the human decision-making process. Un comportamento naturale richiede inoltre la capacità di superare, nella comunicazione, l'ipotesi restrittiva che i messaggi scambiati siano sempre sinceri, comprensibili e 'neutrali' nella forma. Occorre invece affrontare situazioni che sono comuni nella human social life, come le forme più o meno sottili di inganno, il dubbio, l'ironia, lo humor o la manifestazione di empatia, che sono creati da esigenze di intrattenimento e di politeness. L'intreccio tra forme di ragionamento e forme di comunicazione 'naturali' è ovvio: simulare e riconoscere dubbio, inganno o empatia non implica soltanto sviluppare metodi di generazione e interpretazione 'superficiale' di frasi, ma anche riflettere su quali forme di ragionamento (su di sé e sull'Interlocutore) soggiacciono a tali forme di comunicazione. Ad esempio, la conoscenza dello stato emotivo (di chi parla e di chi ascolta) o della predisposizione a particolari stati emotivi, dovuta a temperamento o a tratti di personalità, influenzano notevolmente la scelta delle strategie di persuasione (Sillince and Minors, 1991).

---

<sup>1</sup> La definizione di concetti come "attitude, personality and emotion is still rather controversial. A remind of the theories behind these concepts may be found in Hudlicka (this issue) and also in ....

Analogamente, alcune forme di humor possono essere generate mediante una manipolazione di variabili che influenzano lo stato emotivo dell'Interlocutore (Elliott, 1997).

L'esigenza di naturalezza nell'interazione con le tecnologie comporta quindi problemi nuovi, e forse sollecita anche l'elaborazione di metodi nuovi, per l'IA, se è vero che le applicazioni complesse hanno storicamente lanciato nuove sfide al mondo della teoria (ed è proprio questo, l'interesse di una Rivista come AAI). Proviamo allora a chiederci da dove sorgono, i problemi legati alla realizzazione di una interazione naturale. Se è lecito cercare fattori comuni in ricerche così diverse fra loro, come quelle evocate in questa Introduzione e di cui i lavori inclusi in questo Special Issue sono esempi significativi, l'*ambiguità* e la *combinazione di diversi fattori* sembrano essere aspetti rilevanti per molti dei problemi di metodo affrontati.

Prendiamo il caso del *dubbio* ('a particular attitude towards a proposition in a dialog, which expresses the Interlocutor's difficulty in accepting the proposition), che è il tema del lavoro di Sandra Carberry e colleghi (this issue). Alla base del riconoscimento e della generazione di espressioni di dubbio, nel dialogo fra agenti umani e artificiali, c'è la capacità di rappresentare la struttura delle loro menti (belief di primo e second'ordine, con le loro strength e i loro legami), il contesto in cui la proposition considerata è stata prodotta e la 'credibilità' dei fatti comunicati in relazione alla 'affidabilità delle fonti informative' citate. Sono, questi, elementi simili a quelli che condizionano, insieme ad altri (come la 'plausibilità' dei dati e il 'rischio di essere scoperti'), la possibilità di generare e riconoscere comunicazioni ingannevoli, che siano meno banali dell'aperta menzogna (Carofiglio et al, 2001; de Rosis et al, submitted). Anche se è possibile definire delle regole che combinano questi elementi per stabilire condizioni di riconoscimento e generazione del dubbio only from the analysis of propositions (e il lavoro citato, ammirevolmente, lo fa, e con risultati interessanti), la sola elaborazione del linguaggio naturale produce tuttavia (a giudizio degli stessi autori) risultati ancora non pienamente soddisfacenti. Di qui l'idea di considerare, in futuro, aspetti non verbali della comunicazione (prosody and gesture) e il modo in cui comunicazione verbale e non verbale si combinano. In prospettiva, le variabili da considerare crescono, dunque; il ruolo di ciascuna di esse non è definibile in modo chiaro ed univoco, e i problemi legati alla loro combinazione si fanno più pesanti.

Problemi simili si ritrovano nel riconoscimento di altre attitude, come gli stati di 'cognitive overload', from speech variables such as sentence fragments and articulation rate (Berthold and Jameson, 1999). Infine, livelli di complessità confrontabili si incontrano anche nel mapping between emotional states, the factors that are responsible for activating them (their causes), the signals that are employed in expressing them and the changes in the agents' behavior that are

produced by the new emotional situation (effects). Il triggering e l'espressione delle emozioni sono caratterizzate dalle stesse condizioni di *aspecificità e multifattorialità* che rendono complesso lo studio epidemiologico delle patologie prevalenti nelle società industrializzate. L'attivazione di una emozione è, in genere, il risultato della combinazione di più fattori: eventi che occorrono nel contesto considerato, ma anche variabili individuali, legati cioè alle caratteristiche del singolo agente (come la sua personalità e il suo temperamento) e sociali, e cioè testimoni delle relazioni che intercorrono fra l'agente e il suo o i suoi interlocutori. Una sola causa non basta, spesso, a produrre l'insorgere di una emozione: è la loro combinazione che lo fa. E, d'altra parte, il contesto gioca un ruolo essenziale anche nell'espressione di emozioni provate: è noto ad esempio che la reazione emotiva di un agente e la sua decisione di manifestare o no l'emozione provata sono influenzate, di nuovo, dal contesto culturale, etnico, sociale, in cui l'agente si muove o da cui proviene. *Esempi* (De Carolis et al, 2001). Ed è anche ben noto che un segnale facciale può ammettere interpretazioni diverse, fortemente dipendenti dal contesto. Ad esempio, nella 'componential appraisal theory' di Kaiser et al, facial expressions convey several meanings at the same time: smiles and frowns may be employed as 'speech regulation signals', 'speech-related signals', 'means for signalling relationship', 'indicators for cognitive processes' and so on. In addition, prototypical expressions do very rarely occur in spontaneous interactions: many depend on the situational context and the interindividual variability is very high. (Kaiser et al, ?). Similar considerations may be found in LaFrance (2002), who shows that there are different types of smiles ('miserable', 'angry', 'skeptical' etc), each expressing a mixed emotion state (La France, 2002).

Lo stesso problema di ambiguità nell'interpretazione è sollevato, again, nel lavoro di Carberry e colleghi (this issue), per sentence del tipo "At my age?", o ""To me?" ed altre che possono essere interpretate sia come manifestazioni di dubbio che come semplici richieste di chiarimento.

Deriva, da tutto ciò, che un sistema che si ponga l'obiettivo di una interazione naturale si trova ad affrontare l'annoso problema del sensing, della rappresentazione e dell'interpretazione del *contesto*. Secondo l'analisi di alcuni psicologi cognitivi ben noti, come Elliott<sup>2</sup>, lo spettro delle

---

<sup>2</sup> Elliott (1993) analizza le variabili che influenzano l'intensità delle emozioni in Affective Reasoner classificandole in tre grandi categorie:

- *simulation-event variables*: those whose value changes independently of situation interpretation mechanisms (goal realization or blockage, appealingness, certainty, sense of reality, temporal proximity, surprisingness,...)
- *stable disposition variables*: those involved in an agent's interpretation of situations, which tend to be constant and help to determine the agent's personality (appraisal bias, importance to agent of achieving goal or not having it blocked or of not having standards violated,...);
- *mood-relevant variables*, that contribute to the agent's mood state (arousal, valence bias, anxiety-invincibility etc)

variabili di contesto è davvero ampio, e l'importanza di ciascuna di esse non è ben definita, malgrado gli studi su questo tema risalgano ormai a diversi anni.

Il challenge di chi lavora nel settore dell'affective computing è quindi innanzi tutto quello della *costruzione di modelli* validi, che combinino questi fattori in modi appropriati; e poiché i passi del processo di riconoscimento, modellizzazione e reazione sono diversi, e si combinano in modo non necessariamente predefinito, occorre studiare *le architetture* che meglio rispondono alle esigenze di questo settore. Infine, nessuno ha la presunzione di aver ancora definito il modello 'ideale': al massimo, si ha la convinzione di aver prodotto una proposta nuova, che affronta e risolve alcuni dei problemi aperti, ma che va sottoposta al vaglio di una valutazione seria, in forme diverse. Per questa ragione, si tende oggi a costruire *tool* o *testbed* domain-independent che consentano ad esperti di dominio di realizzare i modelli attraverso il tipico processo di 'costruzione per approssimazioni successive' sperimentato da secoli in fisica. E i casi di studio orientati a specifiche applicazioni vengono presentati come puramente esemplificativi.

Questo Special Issue presenta lavori che contribuiscono allo studio degli aspetti citati. In particolare, per quanto riguarda lo sviluppo di *modelli* per la simulazione di emozioni, si evincono, in essi, esempi di due *tendenze* in atto da alcuni anni, fra loro alternative, per la soluzione dei problemi di aspecificità e multifattorialità prima citati:

- a. Il metodo che, seguendo il seminal work di Ortony (1988), si basa su *funzioni matematiche* (tipicamente, una funzione log-esponenziale) per combinare i diversi fattori causali o predisponenti in misure di intensità delle emozioni o di mood. Gli esempi classici in questo campo sono Affective Reasoner (Elliott), Em (Reilly) e Cathexis (Velasquez). Il sistema adottato da Prendinger e collaboratori in SCREAM (this issue) segue questo filone. Qui, ognuna delle variabili (individuali e di contesto) assume un valore in un range, e l'intensità dello stato affettivo è calcolato attraverso una funzione del tipo di quelle menzionate.
- b. Il metodo che propone gli Hidden Markov Model come soluzione al problema. In questo caso, aspecificità e multifattorialità vengono risolte tentando di individuare le relazioni esistenti fra i vari fattori e rappresentandole in un grafo di relazioni, in cui l'ambiguità è definita in termini di incertezza e manipolata. Il primo passo in quest'ambito was quello compiuto qualche anno fa da Ball e Breese (2000) nella modellizzazione di aspetti 'shallow' delle situazioni emotive (modi di espressione). Il lavoro di Conati in questo Special Issue è uno sviluppo notevole in questa direzione. L'applicazione considerata è quella dell'interazione con gli educational game. I modelli

sono dinamici, in quanto tengono conto del fatto che the student's learning and emotional state evolves gradually over time, so that the values in consecutive time instants are not independent of each other. Il metodo adottato è quindi quello dei Dynamic Belief Networks (*ref*). Questa soluzione è simile a quella proposta in de Rosis et al (in press) per rappresentare il mixing delle emozioni e il loro time decay. Nel lavoro di Lisetti e Gmytr, il modello considera, in particolare, la transizione fra stati emotivi diversi. The Affective Knowledge Representation schema is a probabilistic frame-based representation in which the large number of factors describing causes, characteristics and consequences of emotions are described in a hierarchy of frames and the way that transitions among emotional states occurs is represented, again, in Bayesian Networks.

La differenza principale fra le due soluzioni è nel fatto che, nella prima, l'effetto prodotto dalla combinazione di più variabili in una funzione matematica è abbastanza prevedibile (il logaritmo di una somma di esponenziali tende a isolare i fattori più forti trascurando l'effetto di quelli più deboli). L'effetto prodotto dalla concatenazione di queste funzioni per definire il mood, la .... Non è tuttavia prevedibile in modo altrettanto chiaro.

Per i modelli basati sui belief network, invece, soltanto la simulazione in condizioni diverse permette di 'calibrare' i parametri introdotti. D'altra parte, la rappresentazione della mente degli agenti mediante una estensione del modello belief-desire-intentions di Rao e Georgeff ( ) che includa anche personalità, emozioni e mood presenta diversi vantaggi. Innanzitutto, it enables driving the *consistent behavior* of agents from a model of their internal cognitive state (the beliefs, desires and intentions that activate appraisal of emotions, regulate the decision of whether to show or to hide them and finally drive externalized actions (Allbeck and Badler, 2001). Vale, in questo caso, la distinzione fra 'surface models' ("aimed at achieving external behaviors and display of traits which convey the impression of believable personalities") e 'deep generative cognitive models', "in which ... emotions and personality... are integrated in the respective architectures as necessary components for the achievement of rationality and social competences" (Petta and Trappl, 1997). Occorre notare che i cognitive model hanno il vantaggio di offrire maggiori garanzie di costruzione di comportamenti consistenti: e la consistenza è un fattore determinante della believability (Reilly?). Se, infatti, è un sistema di belief, desire e goal a guidare sia l'attivazione degli stati emotivi che la loro esternazione e l'espressione di attitudini come dubbio, deception, suspicion eccetera, sarà la consistenza di questo insieme a dare qualche garanzia di consistenza nel comportamento. In addition, questo secondo tipo di modelli ha il vantaggio di poter essere usato sia in senso generativo che in senso riconoscitivo, mediante propagazione di evidenze (rispettivamente) sulle cause scatenanti o sui

segnali osservati. E' quindi possibile ipotizzare, ad esempio, qual'è lo stato emotivo dell'Interlocutore mediante una propagazione di evidenze sugli eventi occorsi in un insieme di belief del second'ordine (l'immagine che l'Agente ha dello stato mentale dell'Interlocutore stesso) e anche tentare di interpretare le *ragioni* di un suo comportamento emotivo: *“Mi sembra arrabbiato. Perché? Forse ha interpretato negativamente la mia affermazione? E' possibile, visto che è un permaloso e che non sa che io tendo a scherzare!”*. Piuttosto che intrattenersi in discussioni di principio sul metodo più adeguato, sarebbe forse venuto il momento di porre a confronto i due metodi su un unico scenario, per verificare differenze e vantaggi / svantaggi di ognuno di essi.

L'applicazione prevalente dei metodi di affective computing è nell'interazione fra singolo utente e sistema, in quei settori (education, army, embodied animated agents) in cui l'interazione naturale è una condizione essenziale per l'accettabilità del sistema. Uno spunto (applicativo e di metodo) nuovo è quello proposto da Hudlicka nel suo testbed MAMID (this issue), in cui equipes of agents with various combinations of profiles and affective states interact, and the different responses that may be activated in these interactions are simulated. The interest of MAMID is not only in enabling the simulation of multiple interactions, but also in proposing a parametrized cognitive architecture model that may be applied, in principle, to any domain. This system includes several modules, motivated by cognitive theories about attention, memory, decision making that are described in the paper. Knowledge structure and processing steps are fine-grained and a number of parameters have to be defined, in building a specific application.

Per ottenere la possibilità di adattare l'aspetto esterno dell'Agente Artificiale al contesto in cui questo dovrà mostrarsi, si tende a separare la sua componente riflessiva (la 'mente') da quella comunicativa, più immediatamente visibile per l'Utente (il 'body'). Questo è da sempre vero, nei sistemi di generazione del linguaggio naturale che sono basati, in prevalenza, su architetture pipeline in cui la parte 'strategica' e di microplanning sono separate da quella, finale e 'tattica', di generazione superficiale (Reiter and Dale, 2000). Soluzioni più raffinate, anche se nella stessa direzione, vengono proposte ora per gli Agenti Embodied. Qui, il collegamento fra componente strategica e tattica è spesso realizzato attraverso la marcatura del messaggio che il body dell'Agente dovrà realizzare, sottoforma di file XML. La definizione dei tag è strettamente legata alle capacità espressive dell'Agente. In alcune proposte (come quella di Prendergast, this issue), i tag corrispondono direttamente ai segnali 'non verbali' che il body associa ad ogni parte del discorso. Per garantire una maggiore indipendenza del body dalla mind, è possibile però anche definire tag che indichino i meaning (cioè le funzioni comunicative) piuttosto che i signal.

Ad esempio: il tipo di performativo ('order') piuttosto che l'espressione facciale che gli corrisponde ('....'). Associando poi ad ogni Agente una tabella meaning-signal, è possibile realizzare il messaggio che l'Agente stesso deve trasmettere in modo adattato al contesto, al tipo di tecnologia disponibile (pc vs handheld etc) all'ambito culturale ecc (de Rosis et al, in press).

The papers in this Issue all derive from contributions to a Workshop on "Attitudes, Personality and Emotions in User-Adapted Interaction", that was organized in Sonthofen, in the scope of the User Modeling 01 Conference (electronic Proceedings of this Workshop are available at the url: <http://aos2.di.uniba.it:8080/ws-um01.html>). This Special Issue could not include all contributions that were presented at the Workshop, and the papers are different from the versions published in the informal Workshop Proceedings: these improvements reflect the fastly moving situation in this domain and the lively debate in this area. To give an idea of this debate and of differences in the viewpoints of people involved, the papers in this Issue are concluded with a Panel Discussion. Again, this is not a transcript of the discussion that was done during the Workshop, but is the result of a subsequent email debate among the participants. I wish to personally acknowledge all the participants for accepting to contribute to what I believe results to be an interesting discussion. Although I'm convinced that this is a good-value Special Issue (and hope that the readers of AAI will enjoy with it), I also believe that this is only a step forward in a process that still requires a lot of more work to produce a computer that interacts in a really 'natural' way with its user: that is, the kind of computer that is beautiful, warm, has a sense of humor, is able to distinguish good from evil, appreciates strawberries with cream, is able to feel love and to induce the User to fall in love for it (him? her?) etc, that Turing was dreaming of not so many years ago.

### **Main References**

J.M. Allbeck and N.J. Badler (2001). Consistent communication with control. In C. Pelachaud and I. Poggi (Eds): *Proceedings of the Workshop on "Representing, annotating and evaluating non-verbal and verbal communicative acts to achieve Contextual Embodied Agents"*. Autonomous Agents 2001.

J. Ball and J. Breese (2000). Emotion and personality in a conversational agent. In S. Prevost, J. Cassell, J. Sullivan and E. Churchill (Eds). *Embodied Conversational Characters*. MIT Press.

A. Berthold and A. Jameson (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (Ed), *Proceedings of the 7<sup>th</sup> International Conference of User Modeling*. Springer.

- J. G. Carbonell (1980). Towards a process model of human personality traits. *Artificial Intelligence*, 15.
- V. Carofiglio, F. de Rosis and C. Castelfranchi (2001). Ascribing and weighting beliefs in deceptive information exchanges. In M.Bauer, P.J.Gmytrasiewicz and J Vassileva (Eds), *Proceedings of the 8<sup>th</sup> International Conference of User Modeling*. Springer.
- F. de Rosis, C. Pelachaud, I. Poggi, V.Carofiglio and B. De Carolis (in press). From Greta's mind to her face: modeling the dynamics of affective states in a Conversational Embodied Agent. Submitted to the *International Journal of Human-Computer Studies*.
- P. Ekman: Facial expressions. In Dalglish, T., & Power, M. (1999). *Handbook of Cognition and Emotion*. New York: John Wiley & Sons Ltd.
- C. Elliott and A. Ortony (1992). Point of view: modeling the emotions of others. *Proceedings of the 14<sup>th</sup> Annual Conference of the Cognitive Science Society*.
- C. Elliott (1993). Variables influencing the intensity of simulated affective states. AAAI Technical Report for the Spring Symposium on "Reasoning about Mental States: formal theories and applications". Stanford.
- C. Elliott (1994). Using the Affective Reasoner to support social simulations. *Proceedings of IJCAI '93*.
- C. Elliott (1999). Why boys like motorcycles: using emotion theory to find structure in humorous stories. *Proceedings of EBAA '99*.
- M. La France (2002). What's in a robot's smile? The many meanings of positive facial expressions. In R Aylett and L Canamero (Eds), *Animating expressive characters for social interactions*. Symposium of the AISB'02 Convention.
- A. Ortony (1988). Subjective importance and computational models of emotions. In V. Hamilton, G.H. Bauer and N.H. Frjida (Eds). *Cognitive perspectives on emotion and motivation*. Kluwer.
- J. Pearl (2000). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufman Publishers.
- P. Petta and R.Trappl (1999). Personalities for synthetic actors: current issues and some perspectives. In R. Trappl and P. Petta (Eds). *Creating personalities for synthetic actors*.

A.S. Rao and M.P. Georgeff (1991). Modeling rational agents within a BDI architecture. In J. Allen, R. Fikes and R. Sandewall (Eds). *Proceedings of the 2<sup>nd</sup> International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufman.

J.A.A. Sillince and R.H. Minors (1991). What makes a strong argument? Emotions, highly-placed values and role-playing. *Communication and Control*. 24, 3-4-

M.A. Walker, J.E. Cahn and S.J Wittaker (1997). Improvising linguistic style: social and affective bases for agent personality. *Proceedings of Autonomous Agents'97*.