G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

11 Multimodal Temporal Patterns for the Analysis of User's Involvement in Affective Interaction with Virtual Agents

Alessia AGLIATI, Fabrizia MANTOVANI, Olivia REALDON, Linda CONFALONIERI, Antonietta VESCOVO

Abstract. Given the growing interest in developing embodied virtual agents with multimodal communication and emotional expression abilities, the issue of user's involvement is a relevant topic to take into account in determining how to assess and interpret the quality of user-agent affective interaction. Main goal of this paper is the definition of a methodology for the analysis of user-agent interaction synchrony considered as an index of user's involvement. The proposed approach is based on recent advances in communication psychology, which on the one hand show the importance of considering the hidden temporal organization underlying communicative interaction and on the other hand provide a specific methodology for the structural analysis of the interactive flow (analysis of intra- and inter-individual multimodal behavioral patterns through Theme software). From a theoretical point of view, the crucial assumption is that the more synchronic interactions are, the more pleasant and fulfilling they are experienced, and consequently more related to positively valenced emotional states. Our main objective is to tune and to test this methodology (typically used in analyzing human-human communication exchanges) within user-agent interaction, in order to detect interactive temporal patterns of actions in affective interactions with virtual agents. This approach has been developed within the European project MYSELF, where we are a preliminary evaluation study of an interactive pedagogical agent by combining self-report measures, physiological measures and multimodal behavioral patterns approach is being carried out.

Contents

11.1 Introduction……..…228
11.2 Embodied Virtual Agents Devised to Participate in Communicative Interactions with Users...…....229
11.3 Bridging Multimodality and Interaction Synchrony within Affective Interactions ……………………………………………...…..................230
11.4 A Methodology for Evaluating the Quality of User-agent Interaction........231
11.5 User-Linda Agent Interaction Analysis Through Theme: a Preliminary Study...…………....... 236
11.6 Conclusions………………………………………………..……...…....238
11.7 References…………………………………………………………..…....238

emerging.communication

Studies on New Technology and Practices in Communication

Support us – Ask your institution/library/department to order our books

To order a book contact by fax or by e-mail the offices below:

IOS Press, Inc.
4502 Rachael Manor drive
Fairfax, VA 22032
U.S.A.
Tel.: +1 703 323 5600
Fax: +1 703 323 3668
sales@iospress.com

IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
The Netherlands
Tel: +31 20 688 3355
Fax: +31 20 620 3419
info@iospress.nl

IOS Press,
c/o Ohmsha, Ltd.
3-1 Kanda Nishiki-cho
Chiyoda-ku
Tokyo 101
Japan
Fax: +81 3 3233 2426
(Books only)

IOS Press/
Gazelle Book Services Ltd
White Cross Mills
Hightown
Lancaster LA1 4XS
United Kingdom
Tel.: +44 1524 68765
Fax: +44 1524 63232
sales@gazellebooks.co.uk

From Communication to Presence
Cognition, Emotions and Culture towards the Ultimate Communicative Experience
Festschrift in honor of Luigi Anolli
Volume 9 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: G. Riva, M.T. Anguera, B.K. Wiederhold, F. Mantovani
September 2006, 323 pp., hardcover
ISBN: Pending NEW
Price: US$134 / €107 / £72

Global Data Management
Volume 8 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: R. Baldoni, G. Cortese, F. Davide and A. Melpignano
July 2006, 376 pp., hardcover
ISBN: 1-58603-629-7 NEW
Price: US$175 / €140 / £95

The Hidden Structure of Interaction
From Neurons to Culture Patterns
Volume 7 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: L. Anolli, G. Riva, S. Duncan Jr. and M.S. Magnusson
May 2005, 304 pp., hardcover
ISBN: 1-58603-509-6
Price: US$126 / €105 / £73

Ambient Intelligence
The Evolution of Technology, Communication and Cognition Towards the Future of Human-Computer Interaction
Volume 6 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: G. Riva, F. Vatalaro, F. Davide and M. Alcañiz
January 2005, 316 pp., hardcover
ISBN: 1-58603-490-1
Price: US$126 / €105 / £74

Being There
Concepts, Effects and Measurements of User Presence in Synthetic Environments
Volume 5 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: G. Riva, F. Davide and W.A. IJsselsteijn
2003, 344 pp., hardcover
ISBN: 1-58603-301-8
Price: US$116 / €105 / £74

E-Business and E-Challenges
Volume 4 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: V. Milutinovic and F. Patricelli
2002, 360 pp., hardcover
ISBN: 1-58603-276-3
Price: US$116 / €105 / £74

Say not to Say: New Perspectives on Miscommunication
Volume 3 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: L. Anolli, R. Ciceri and G. Riva
2001, 288 pp., hardcover
ISBN: 1-58603-215-1
Price: US$116 / €105 / £74

Towards CyberPsychology
Mind, Cognition and Society in the Internet Age
Volume 2 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: G. Riva and C. Galimberti
2001, 326 pp., hardcover
ISBN: 1-58603-197-x
Price: US$116 / €105 / £74

Communications Through Virtual Technologies
Identity, Community and Technology in the Communication Age
Volume 1 Emerging Communication: Studies on New Technologies and Practices in Communication
Edited by: G. Riva and F. Davide
2001, 292 pp., hardcover
ISBN: 1-58603-162-7
Price: US$116 / €105 / £74

***IOS* Press**
http://www.booksonline.iospress.nl

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

11.1 Introduction

The challenge of developing virtual agents meant to be engaged as social interactive conversational partners within communicative exchanges with humans is currently being addressed with considerable empirical and theoretical efforts in the HCI field [1, 2, 3].

Communicative exchanges are the framework within which interpersonal relationships are generated and fostered, hence granting interaction sequences with predictability and regularity [4]. At the same time, the shaping of interpersonal relationships is due to emotions that, as situated practices occurring through repeated interactions [5], give sense, direct and modify the relationship patterns. As such, emotions grow up, are redefined and modified within relationships, while relationships provide the context for the generation of emotions and their management. Within this perspective, emotions, being eventually rooted in interactions, should not be seen as involuntary intra-personal events, but, rather, as resulting from personal choices and options, that become manifest within the communication flow. Within the spatial and temporal boundaries of interactive communicative exchanges, emotions, more often than thoughts and intentions, tend to be expressed, mainly by non verbal extra-linguistic systems (like facial expressions, gestures, posture, etc.). As widely agreed in literature, emotion expression and management is mainly performed via non verbal signaling systems, that, in the relationship-building process tend to be synchronically shaped and tuned in order to attain the desired goal [6, 7].

Within this theoretical framework, the evaluation of affective interactions with virtual agents begs the relevant issue regarding how and at which degree multimodal agent's and user's expressions reciprocally match in a synchronic way.

Basically, the crucial assumption is that, the more synchronic interactions are, the more pleasant and fulfilling they are experienced, and consequently more related to positively valenced emotional states. Therefore it is reasonable to suppose that, on the one hand, positively valenced emotions contribute in shaping and maintaining functionally adaptive relationships, and on the other hand, complex synchronic interactions, featured by coordinated interpersonal timing, are related to more adaptive relationships, at least in terms of attachment security [8, 9].

Given these premises, this contribution mainly aims at introducing a pioneering evaluation method of human-agent interaction quality, by detecting interactive multimodal temporal patterns of actions occurring within user-agent affective interactions. To reach this goal, recent contributions in developing multimodal embodied conversational agents will be introduced. Then, starting from the premise about multimodality expression abilities as an essential requisite for both interactants in order to eventually converge in their communication styles, theoretical perspectives connecting interaction synchrony with positively valenced interpersonal relationships will be discussed, presenting tentative researches conducted on this theme within user-agents interactions.

After that, a methodology for detecting the degree in quality (as resulting from different signaling systems) and quantity (i.e., frequencies) of co-occurring patterns

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

of action, considered as valid indexes of interaction synchrony, will be proposed. Finally, a preliminary study aimed at evaluating user-agent affective interactions by employing this methodology will be briefly presented.

11.2 Embodied Virtual Agents Devised to Participate in Communicative Interactions with Users

When conceived as communicative partners, embodied virtual agents need to be equipped with multimodal expressive abilities. As, among others, Anolli [4] pointed out, besides language, there are several other communicational devices to show interactants' own communicative design, like the paralinguistic (or supra-segmental), the face and gestures system, the gaze, the proxemic and the haptic, as well as the chronemic. Each of these communicative systems bears its contribution and participates in defining the meaning of a communicative act in an autonomous way. Therefore, meaning is not connected with a unique and exclusive signaling system, but comes out of the network of semantic and pragmatic connections between different signaling systems (multimodal configurations). Such semantic and pragmatic connections are organized, coordinated and made convergent, in communicative exchanges, according to the principle of semantic synchrony, which enables interlocutors to make explicit the meaning of their own communicative intentions in a unitary and coherent way [10].

Along this theoretical line, research trends on Embodied Conversational Agents (ECAs) [11], focusing specifically on communicative and conversational agents, concern the development of humanoid software agents that use speech, gaze, gesture, intonation and other signalling systems in the way humans use them in the communication process [12, 13].

The assumption of this perspective can be considered as one of the most relevant turning points in the development of embodied virtual agents. Indeed, in the early nineties, the challenge in this domain of study was focused basically on agents' appearance. More specifically, establishing and defining the agents' visual physical appearance in terms of anthropomorphism and human-likeness basically represented the main rationale in agents' design and development [14-16].

From the late nineties, studies on agents' embodiment and animation have started taking into account the agent being featured and acting as an interlocutor and a conversational partner, engaged and participating in a communicative interaction [12], animated in order to use different signalling systems [1, 17, 18].

Considerable efforts are being spent on this issue, with interesting and convergent findings from the several research groups addressing it. De Rosis and colleagues proposed a research project aimed at implementing Greta, a 3D Embodied agent that can be animated in real-time, able to coherently communicate complex information through the combination and the tight synchronization of verbal and nonverbal signals [19]. Consistent with the issue of multimodal animation, Cosi and colleagues proposed Lucia, an Italian talking head based on a modified Cohen-Massaro's Model [20]. Likewise, within the MIT group [13, 21], virtual agents named LAURA (a physical trainer) and REA (a salesperson) were developed, with the aim of supplying them with a wide range of conversational affordances, including the synchronic use

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

of hand gestures, body posture, eye gaze, and facial expressions, to act as believable partners in communicative interactions.

11.3 Bridging Multimodality and Interaction Synchrony within Affective Interactions

11.3.1 Theoretical Perspectives

Although agents' multimodal expression abilities are recognized as a basic requisite in developing virtual agents, the issue connected with their synchronizing ability and the role such ability plays in affective interactions and in establishing and maintaining positively valenced relationships still need to be addressed in order to provide the development of agents' architecture with a robust theoretical backbone and to gather a valid frame for their evaluation.

Within human face-to-face interactions, people naturally tend to synchronize their rhythms of speech and body movements: many biological processes are known to be cyclical or rhythmical in nature and to be driven by the central nervous system, and there is evidence to suggest that human interactions have rhythms as well [22]. Such a rhythm has to do, and develops over time, within communicators' reciprocal matching and convergence in the way they shape their verbal and non verbal communication. In other words, people tend to converge their communication style (encompassing vocal features and non verbal behavior) towards the style used by the listener, even exhibiting exactly the same behaviors (*mirroring phenomenon*). From this perspective, rhythm and synchrony in the communicative interaction can be seen as powerful indexes of participation, agreement and closeness between interactants, showing that one is "with" his/her interlocutor in his/her attention and expectancies [23]. Moreover, according to the Accommodation Theory [24, 25], synchrony reflects *reciprocity,* since it meets the underlying psychological need to evoke listeners' social approval, to attain communication efficacy and to maintain positive social identities. Consistent with this point of view, results in recent researches exploring the relationship between interaction synchrony and initial attraction in first-met couples, showed that a higher degree in interaction synchrony was more strongly connected with participants' interest in their partner in the first minute than in last minute of their interaction [26].

These findings support the idea that interaction synchrony as resulting in complex co-occurring multimodal patterns of action (thus showing a more robust convergence in communication styles) can be viewed as valid and accurate indexes of perceived closeness, reciprocity and more positively valenced relationships. As such, interaction synchrony can be considered as a way of shaping interactions, which varies along the communicative flow as a function of reciprocal participation and involvement, hence placing a suitable frame for sharing positive affect.

Indeed, in turn, emotions are viewed as a communicatively structured, convention-based and rule-governed process [6, 7, 27], that are originated in, and give sense to one's patterns of relationships. As such, they can be thought of as embodied practises through which people build and organize interactions in their life worlds. According to Goodwin [5], emotions are situated practices within specific sequential positions

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. ***Festschrift in honor of Luigi Anolli***
IOS Press, Amsterdam, 2006, (c) All rights reserved – http://www.emergingcommunication.com

in interaction, whose relevant unit for analysis is not the individual, but the sequential organization of action in interaction.

11.3.2 Tentative applications to human-artificial agents interactions

Starting from these theoretical perspectives, it is likely to assume that fulfilling and satisfying user-agent relationships are related to high degree in interaction synchrony between user and agent.

A tentative application of these concepts to human-computer interaction has recently been proposed by Bailenson and Nick [28]. In their study, assuming the theoretical paradigm of the *chameleon effect* [29], participants interacted with a virtual agent in an immersive virtual reality environment. The agent either mimicked the participant's head movements at 4-s delay or utilized pre-recorded movements of another participant as it verbally presented an argument. Mimicking agents were perceived as more persuasive and received more positive trait ratings that nonmimickers, despite participant's inability to explicitly detect the mimicry.

Moreover, research trends on *humanoid robots* [30, 31, 32] seem to follow the same pathway, highlighting how (as well as in human-human interactions) mutual coordination between human's and robot's body movements is a valuable key aspect to warrant compelling smoothness and naturalness in the communication exchange. Empirical results by Sakamoto and colleagues [31, 33] showed that such cooperative body movements as eye-contact and synchronized arm movements mutually related to a more positive user's subjective evaluation of the robot.

In sum, the effectiveness in synthonically and synchronically using bodily expressions and movements, as well as other communicative signalling systems, represents the added value in producing compelling human-like bonding between humans and artificial agents, be they Embodied Conversational Agents or robots.

11.4 A Methodology for Evaluating the Quality of User-agent Interaction

This methodology roots in the same theoretical premises but is aimed to demonstrate that user's tendency to synchronize and in case even mimic the agent is per se a signal of involvement in the interaction with him/her.

Consequently, the proposed method to assess the user's involvement in interaction with a virtual agent, may provide fruitful advances in this field by measuring interaction synchrony. In particular, a structural analysis of interaction aimed at hidden temporal organization detection should be integrated with a detailed qualitative analysis of visible behaviors. The methodology below describe meets this requirement, since it is observational in nature and aims at interactive multimodal patterns detecting, using Theme software [34, 35].

11.4.1 The Theme Method

11.4.1.1 General Description

Theme software is a tool for the detection and analysis of patterns in time-based data. It detects relationships that human observers typically overlook, and that are not

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006, *– http://www.emergingcommunication.com*

commonly found by means of traditional statistical methods. Theme uses a unique pattern detection algorithm that has especially been designed for behavioral research. It gives insight into the structure of behavior in time, and gives measures for its complexity and organization.

The main assumption of this method is that behaviors are made of multimodal temporal patterns (*T-patterns*), which are sequences of behaviors regularly recurring within a certain temporal period [36]. Given the link between synchrony and positive affect in interactions, it is likely to expect that detected temporal patterns indicate the user involvement in the ongoing interaction with the agent.

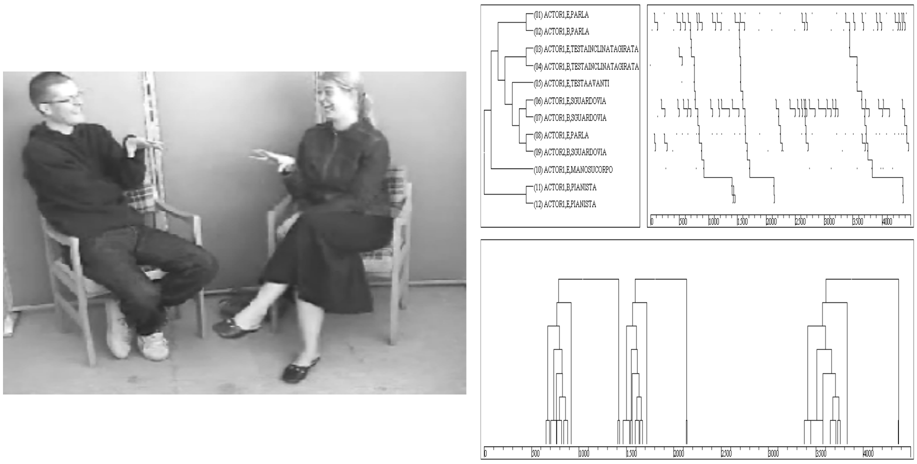


Figure 1. *This* figure includes a picture and a T-pattern of a dyadic interaction between two Icelandic participants. The picture on the left presents the "pianist" gesture displayed only by Icelandic subjects. The T-patterns show the different use of turnover displayed by Italians and Icelanders. The T-pattern on the right refers to an Icelandic interaction between friends, it includes twelve events ((1) *subjectB, e, turn*; (2) *subjectA, b, turn*; (3) subjectA, b, tilt head; (4) subjectA, e, tilt head; (5) subjectA, b, head forward; (6) subjectA, e, look away; (7) subjectB, b, look away; (8) subjectA, e, turn; (9) subjectB, b, look away; (10) subjectA, e, hand on body; (11) subjectA, b, pianist; (12) subjectA, e, pianist), and it recurs three times in the same observation period.

Theme application to studies on emotion and communication is turning out to be heuristically useful and powerful. Applicative opportunities encompass several kinds of face-to-face interactions: from conversation in cross-cultural perspective [22, 34] to adult-children emotion communication [37], to friendship [38] and courtship interactions [39], to clinical married couples exchanges [40] and siblings. Results, in such studies, give evidence to the idea that a higher degree in interaction synchrony is related to the perception of closeness and bonding between interactants.

Examples of T-patterns are illustrated in Figure 1. T-pattern on the left was detected from a dyadic interaction between Italian friends, whilst T-pattern on the right was detected from a dyadic interaction between Icelandic friends. Results drawn from this study suggested a relationship between interaction synchrony and culture [41, 42].

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

11.4.1.2 The Theme Procedure

The Theme procedure entails two main steps: first, coding individual's behaviors following a grid; second, detecting multimodal temporal patterns.

11.4.1.2.1 Coding Behavior

So far, non verbal behavior only has been coded and analyzed using Theme. However, considerable efforts are being spent in analyzing behavioral patterns related to verbal productions by means of Theme, since the quality of interaction is closely linked also to speech production [40, 43].

To code participants' facial and body movements, a grid has to be developed. Behavior units, divided in categories, have to be mapped into the grid. Facial expressions, gaze and head orientation, trunk and shoulders, arms and hands, self-contact, turn-taking are usually included in the grid and coded. Behaviors included in the grid are identified by using specific action units (AU) defined by Facial Action Coding System [44] for upper face, lower face, gaze and head orientation. Other behavior units for hands, arms, trunk and shoulders movements and self-contact are usually taken from grids used in previous studies on nonverbal behavior [39, 41, 45].

Category	**Behavior units**
Hands	rotation, hand contact, crossed hands, hand on face, hand on body, hand on chair, hand on hair, palm inward, palm forward, palm down, palm upward, palm back, palm outside, fist, bag-hand, mirror-hands, rubbing hands, ax-hand, hitchhiking, ring, hands away, ball-hands
Fingers	pointing, counting, negation, pianist
Arms	arms upward, crossed arms, arm forward, open arms, dangling arms, arms back, bend arms
Head	head forward, head down, head back, tilt head l/r, tilt head up/down, negation by head
Trunk	trunk forward, tilt trunk l/r, trunk back, swinging
Face	AU 6 + AU 12 + AU 26; AU 6 + AU 12 + AU 27
Gaze	avert gaze
Shoulders	shoulders shrug
Turn	Speak

Table 1. Example of grid. Fifty behavior in nine categories.

11.4.1.2.2 Detecting Multimodal Temporal Patterns

The occurrences of all event-types within the observation period herein taken into account constitute the so-called *T-data* (or T-dataset) which is the input to the T-pattern detection algorithm. T-pattern algorithm has been implemented in a specialized software package, called Theme, [34, 35] (http://www.patternvision.com and http://www.noldus.com). A T-pattern is essentially a combination of such event-types occurring in the same order with the time distances between consecutive pattern components remaining *relatively* invariant (that is, the time difference between A and B will be $x \pm y$). To be detected a T-pattern should recur at least three times (minimal occurrence) within the observation period.

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*

A schematic representation of a T-pattern is shown in Figure 2. Especially, the letters in the upper line correspond to specific event-type appearing in proportion to the time of their occurrence. Thus, the line 1 is the visual representation of the temporal sequence of movements shown by a given subject during his/her interaction with the other interactant. Within the upper line, a sequence of four event-types – *A, B, C*, and *D* – recurs but it is masked by the occurrence of two other event-types, *W* and *K*. Even a highly expert human observer by visually inspecting the data string would most unlikely be able to detect any kind of pattern. On the contrary, a T-pattern analysis allows to identify the repeated pattern *A, B, C*, and *D* because of its consistent temporal structure. The T-pattern detection algorithm enables an analyst to discover repeated temporal patterns, even when various other event-types occur between the components of the pattern itself, masking its detection at a naked eye.

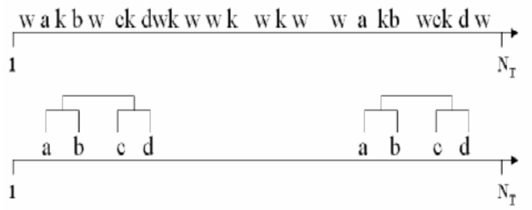


Figure 2. Example of repeated T-pattern

The pattern in Figure 2 illustrates how a larger pattern ((AB) (CD)) is detected as a combination of the two simpler patterns (AB) and (CD). Even in moderate data sets, the number of potential T-patterns is very high. For instance, when the potential number of event-types is 100, the number of potential patterns involving up to 10 event-types is many orders of magnitude greater than 100^{10} if all possible time windows are also considered. Even for supercomputers, it becomes an impossible task to search for each possible temporal pattern separately.

To deal with this issue, simple patterns are detected first – that is, identifying relationships between two event-types such as the (*AB*) or (*CD*) relationship in Figure 2 – while more complex patterns are detected as patterns of patterns (a so called "bottom-up" search strategy). The simpler patterns (*AB*) and (*CD*) are detected first and, then, the larger pattern (*(A B)(C D)*) as a combination of these. The new larger pattern may then become a part of even more complex patterns as it combines with other simple or complex patterns.

A further phase of the T-pattern detection deals with completeness competition between all the detected patterns. In this phase, those patterns that are less complete versions of one or more alternate patterns are deleted. As a matter of fact, during the detection process, a pattern P_X = (*ABCDE*) may be partially detected as, for instance, (*ACDE*) or (*BDE*) or (*ABCE*). Since components of P_X are missing, these three patterns constitute less complete description of the underlying patterning, and consequently they are eliminated. This completeness competition ensures that only the most complete patterns survive and constitute the outcome of the detection process [47].

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

11.4.1.3 The Theme Output

Using Theme, three kinds of data should be obtained, indicating interaction synchrony. In particular, the first two kinds of data, *pattern number and complexity* outputs, reveal the hidden temporal organization of interaction. The third output (Event-Types in Raw Data) has to do with the type of behavior exhibited by each actor during their interaction.

Number of interactive patterns. In each observed interaction, it is possible to detect N. different interactive patterns including both actors' behaviors: interactive patterns imply that both the actors jointly contribute to their combination. Number of patterns is used as an indicator of interaction synchrony: higher number of patterns implies a higher number of co-occurring sequences of behaviors within the observational period. Thus, higher number of patterns means a higher interaction synchrony.

Pattern complexity. It refers to the number of event-types included in each pattern. Each observation period is defined by the average of patterns complexity usually exploited as a further indicator of interaction synchrony, since it indicates that synchrony occurs on different verbal and non-verbal systems of communication.

Event types in raw data. It refers to the proportional occurrences (%) of a given event-type recorded in the observation period. Each of these data are represented in Figure 3.

A further kind of data should be obtained using Theme, by means of a qualitative approach, allowing to measure *mirroring phenomenon* [24]. Within each interactive pattern it is possible to identify identical behaviors simultaneously exhibited by both the actors. For instance, Icelandic friends showed in the right picture of figure 1 simultaneously exhibit the hand-gesture called "*pianist*". The number of these behaviors integrate quantitative data on interaction synchrony.

11.4.2 Using Theme in Human-computer Interaction Analysis

Given the methodological framework above advanced, and considering that a higher degree of interpersonal interaction's synchrony tightly links to the experienced quality of relationship in interacting with virtual agents as well as with humans, some advantages should be pointed in using Theme for evaluating the quality of human-agent interaction.

The first benefit is its quantitative and qualitative integrated approach to the analysis of interactions. Whilst the quantitative approach is aimed at detecting the hidden temporal structure of interaction, the qualitative one makes possible the analysis of visible behaviors and gives a measure of mirroring. In such way, user's involvement in interaction with an agent may be evaluated throughout a structural measure (interactive patterns) that reveals his tendency to synchronize to the agent's movements. A qualitative measure (mirroring) of user-agent interaction may be provided as well, indicating the user's tendency to mimicry agent's movements.

Quantitative and qualitative measures of emotional involvement may be fruitfully integrated with subjective measures (self-report questionnaires) and physiological measurement (changes in subject's heart rate, skin temperature, skin conductance, respiration rate, etc.) provided by other existing approaches in evaluating the quality of human-computer interaction, [48, 49, 50].

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

A further strength of this method is its lack of intrusion for users, avoiding risks of users' discomfort and inhibition related to survey-devices application. A video-camera and a webcam for virtual agent's and users' movements and facial expressions recording are only required.
Moreover, the collecting data session is simultaneous with the user-agent interaction: it makes possible the replication of the procedure on a significant number of subjects (forty at least), allowing proper statistical analysis. Obviously, observed behaviors are coded and analyzed post-session from videotapes.

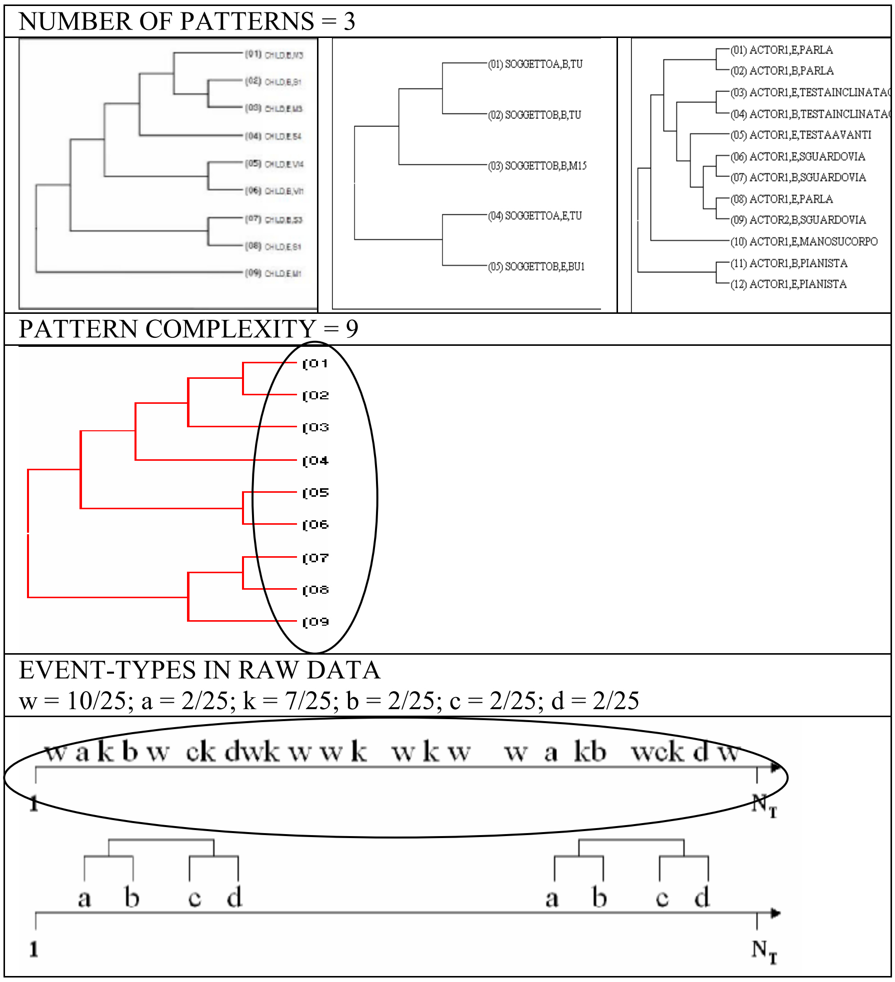


Figure 3. Pattern number, pattern complexity and event-types in raw data are shown with graphical example.(N Event-Types = 25)

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

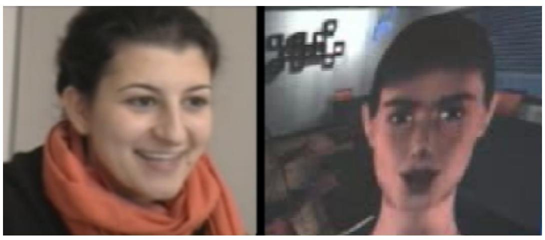
11.5 User-Linda Agent Interaction Analysis Through Theme: a Preliminary Study

A first application of the Theme method in human-computer interaction analysis has been developed within the European project MYSELF - "*Multimodal e*-learning System based on *Simulations*, Role-Playing, Automatic Coaching and Voice Recognition interaction for Affective Profiling" (http://www.myself-proj.it) [51, 52], where a preliminary evaluation study of an interactive pedagogical agent is being carrying out, by combining self-report questionnaires, physiological measures and analysis of multimodal behavioral temporal patterns (synchrony and accommodation). The study basically aims at investigating user-agent interaction quality, within a learning environment involving an animated pedagogical agent called 'Linda'. Agent's multimodal expressivity and treatment of users' expectations' on the forthcoming interaction were considered as main factors.

Taking into account the agent's multimodal expressivity as independent variable, participants were randomly assigned to two different conditions, operationalized as follows: *(a_1)* agent with high multimodal expressive animation ('Linda' agent with voice, blink, lip movements, smile, eyebrow raise, head movements); *(a_2)* agent with limited multimodal expressive animation ('Linda' agent with voice, blink and lip movements). The main hypothesis under investigation is that the occurrence and the complexity of interactive patterns will be greater in the condition of 'Linda' agent with high multimodal expressive animation. Both virtual agent's and user's behavior was observed, coded and analysed using Theme software.

Even though further specific data analyses have to be carried out, first results show that participants tend to enact more synchronized behavioral patterns when interacting with 'Linda' agent with high multimodal expressive animation, compared to the condition of 'Linda' agent with limited multimodal expressive animation (see Figure 4). Thus, the occurrence of a higher degree of interaction synchrony between user and agent along the whole interactive flow, indicates that users seem to naturally experience greater participation and involvement within the interaction and higher relational closeness to agents with high multimodal expressive animation.

This trend of results is consistently corroborated by self-report measures indicating that participants clearly reported more fulfilling and positively valenced experience in condition of interaction with highly multimodal expressive agent.



G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006,

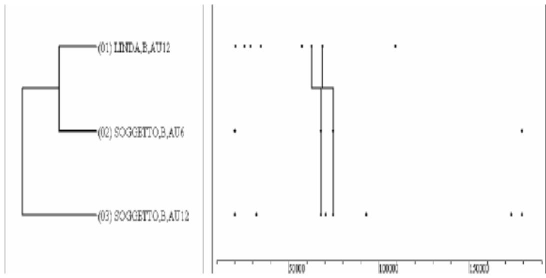


Figure 4. Example of synchronized behavioral patterns of a user interacting with 'Linda' agent with high multimodal expressive animation

11.6 Conclusions

Tuning and employing the Theme method to evaluate user-agent interaction does entail a substantial change in the perspective of the evaluation process itself: not only a rigorous device and a procedure for the detection of patterns of interaction is advanced, but also a radically different view on how to segment and define the phenomenon to be evaluated is proposed.

Indeed, the minimal unit of analysis is identified in interaction, hence entailing that the object to be observed and explored is not the outcome of contingent though frequent connections between two isolated individuals (user and agent), but is social in its nature, since the communicative process is intrinsically social in itself. Moreover, not only does communication require at least two interlocutors to be generated and performed, but also requires participation [53], as a necessary condition for both interactants to define the relationship between them.

Given that relationships-building processes via communication exchanges substantially affect emotions and the perceived valence of the quality of the interaction, the Theme method reveals itself as a fruitful option in evaluating human-agent interaction, by providing valid indexes of the affective quality of the interaction itself. Therefore, Theme represents the affordance for a valuable connection between cutting edge research trends in communication psychology and current technological advances.

11.7 References

[1] J. Cassell, J. Sullivan, S. Prevost and E. Churchill, Embodied Conversational Agents. The MIT Press, Cambridge, MA, 2000.
[2] J.N. Bailenson, J. Blascovich, A.C. Beall and J.M. Loomis, Interpersonal Distance in Immersive Virtual Environments. In Personality and Social Psychology Bulletin, 2003.
[3] L. Schilbach, A.M. Wohlschlaeger, N.C. Kraemer, A. Newen, N.J. Shah, G.R. Fink, and K. Vogeley, Being with virtual others: Neural correlates of social interaction. Neuropsychologia, (2006), 44, 718-730.
[4] L. Anolli, The Detection of the Hidden design of Meaning. In: L. Anolli, S. Duncan, M.S. Magnusson, G. Riva, Eds. The Hidden Structure of Interaction, 23-50. IOS Press, Amsterdam, 2005.
[5] M.H. Goodwin and C. Goodwin, Emotion within Situated Activity. In: Linguistic Anthropology: A Reader, Blackwell Press, Oxford. 239-253, 2000.

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. ***Festschrift in honor of Luigi Anolli***
IOS Press, Amsterdam, 2006, (c) All rights reserved – http://www.emergingcommunication.com

[6] L. Anolli, Psicologia della Cultura [Psychology of Culture]. Il Mulino, Bologna, 2004.
[7] A. Fridlund, Human Facial Expression. An Evolutionary View. Academic Press, San Diego, 1994.
[8] A.A. Hane, S. Feldstein and V.H. Dernetz, The Relation Between Coordinated Interpersonal Timing and Maternal Sensitivity in Four-Month-Old. Journal of Psycholinguistic Research, (2003), 32, 5.
[9] R.A. Isabella and J. Belsky, Interactional synchrony and the origins of infant-mother attachment: A replication study. Child Development, (1991), 62, 373-384.
[10] L. Anolli, Psicologia della Comunicazione [Communication Psychology]. Il Mulino, Bologna, 2006.
[11] C. Pelachaud and I. Poggi, Multimodal Embodied Agents. In: C. Pelachaud e I. Poggi, Eds. Multimodal Communication and Context in Embodied Agents. Proceedings of the Workshop W7 at the 5th International Conference on Autonomous Agents, Montreal, Canada 95-101, 2001.
[12] J. Cassell, More than just another pretty face: Embodied Conversational Agents. Communication of the ACM, (2000), 43, 4, 70-78.
[13] J. Cassell, T. Bickmore, L. Campbell, H. Viljailmsson and H. Yan, More than just a pretty face: Conversational protocols and the affordances of embodiment. Knowledge- Based System, (2001), 14, 55-64.
[14] J. Bates, The role of emotion in believable agents. Communications of the ACM, (1994), 37, 7, 122-125.
[15] T.J., Norman, Motivated Goal and Action Selection. In: Working Notes: AISB workshop, Models or Behaviors-Which Way Forward for Robotics?. University College London, Department of Computer Science research note 94/32, 1994.
[16] T.J. Norman, N.R. Jennings, P. Faratin and E.H. Mamdani, Designing and implementing a multi-agent architecture for business process management . In: J.P. Müller, M.J. Wooldridge, N.R. Jennings, Eds. Intelligent Agents III: Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Vol. 1193 of Lecture Notes in Artificial Intelligence, Springer Verlag, 261-275, 1997.
[17] D.N. Chorafas, Agent technology handbook. Mc Graw Hill, New York, 1997.
[18] S. Franklin, Autonomous Agents as Embodied AI. Cybernetics and Systems, special issue on Epistemological Issues in Embodied AI, (1997), 28, 499-520.
[19] F. De Rosis, C. Pelachaud, I. Poggi, V. Carofiglio and B. De Carolis, From Greta's mind to her face: modelling the dynamics of affective states in a conversational Embodied Agent. International Journal of Human-Computer Studies, (2003), 59, 81-118.
[20] P. Cosi, A. Fusaro and G. Tisato, Lucia: A new Italian talking head based on a modified Cohen-Massaro's Labial Coarticulation Model. Proceedings of EuropeSpeech 2003, Geneva, Switzerland, 3, 2269-2272, 2003.,
[21] T. Bickmore and R.W. Picard, Towards Caring Machines. Conference on Human Factors in Computing Systems, 2004.
[22] L. Anolli, S. Duncan, M.S. Magnusson and G. Riva, The Hidden Structure of Interaction. Amsterdam, IOS Press, 2005.
[23] A. Kendon, Movement coordination in social interaction: Some examples described. Acta Psychologica, (1970), 32, 1-25,
[24] H. Giles, N. Coupland and J. Coupland, Accommodation Theory: communication, context and consequence. In: H. Giles, N. Coupland & J. Coupland, Eds. Contexts of Accommodation, 1-68. Cambridge University Press, Cambridge, 1991.
[25] R.L. Street Jr. and D. Buller, Patients' characteristics affecting physician- patient nonverbal communication. Human Communication Research, (1988), 15, 60-90.
[26] K. Sakaguchi, G.K. Jonsson and T. Hasegawa, Initial interpersonal attraction and movement synchrony in mixed-sex dyads. In: L. Anolli, S. Duncan, M.S. Magnusson, G. Riva, Eds. The Hidden Structure of Interaction, IOS Press, Amsterdam, 107-122, 2005.
[27] J.M. Fernández-Dols, F. Sánchez, P. Carrera and M.A. Ruiz Belda, What are the spontaneous expressions of basic emotions? New views on old questions. Journal of Nonverbal Behavior, (1997), 21, 3, 163-177,
[28] J.N. Bailenson and Y. Nick, Automatic assimilation of nonverbal gestures in immersive virtual environment. Psychological Science, (2005), 16,10, 814-819.
[29] T.L. Chartrand and J.A. Bargh, The chameleon effect: The perception-behavior link and social interaction. Journal of Personality and Social Psychology, (1999), 76. 893-910.
[30] K. Hirai, M. Hirose, Y. Haikawa and T. Takenaka, The development of HONDA humanoid robot. Proceedings of the IEEE INt. Conf. on Robotics and Automation, 1321-1326, 1998.
[31] T. Kanda, H. Ishiguro, M. Imai and R. Nakatsu, Development and evaluation of an interactive humanoid robot "Robovie". IEEE International Conference on Robotics and Automation, 1848-1855, 2002.

G. Riva, M.T. Anguera, B.K. Wiederhold and *F. Mantovani* (Eds.)
From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. *Festschrift in honor of Luigi Anolli*
IOS Press, Amsterdam, 2006, (c) All rights reserved – http://www.emergingcommunication.com

[32] T. Ono, M. Imai and H. Ishiguro, A model of embodied communication with gestures between humans and robots. Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society, 8, 732-737, 2001.
[33] D. Sakamoto, T. Kanda, T. Ono, M. Kamashima, M. Imai and I. Hiroshii, Cooperative embodied communication emerged by interactive humanoid robots. Human-Computer Studies, 2005, 62, 247-265.
[34] M.S. Magnusson, Modelling complex real-time behavioral streams as optimised sub-sets of mutually exclusive and nested T-patterns. In: Measuring Behavior 2000. Proceedings of the 3rd International Conference on Methods and Techniques in Behavioral Research (15-18) August 2000. Nijmegen, The Netherlands, 2000.
[35] M.S. Magnusson, Understanding social interaction: Discovering hidden structure with model and algorithms. In: L. Anolli, S. Duncan, M.S. Magnusson & G. Riva, Eds. The Hidden Structure of Interaction, IOS Press, Amsterdam, 3-22, 2005.
[36] I. Eibl-Eibesfeldt, Ethology: The biology of behavior. Rinehart & Winston, New York, 1970.
[37] A. Agliati, L. Anolli and G.K. Jonsson, Time patterns of emotional Communication in children. Measuring Behavior 2002. Proceedings of the 4th International Conference on Methods and Tecniques in Behavioral research. Noldus Infomation Technology, Amsterdam, 2002.
[38] G.K. Jonsson, Self-Esteem, friendship and verbal and non verbal interaction, 1996. On-line: http://www.hi.is/~msm/poster1996gkj.htm .
[39] K. Grammer, K. B. Kruck and M.S. Magnusson, The courtship dance: patterns of nonverbal synchronization in opposite- sex encounters. Journal of Nonverbal Behavior, (1998), 22, 3-29.
[40] L. Pullini, Communication and discommunication in married couples: interactive nonverbal behavior and temporal structure analysis. Unpublished doctoral dissertation, Catholic University, Milan, 2006.
[41] A. Agliati, A. Vescovo and L. Anolli, Conversation patterns in Icelandic and Italian people: Similarities and differences in rhythm and accommodation. In: L. Anolli, S. Duncan, M.S. Magnusson & G. Riva, Eds. The Hidden Structure of Interaction, IOS Press, Amsterdam, 210-223, 2005.
[42] A. Agliati, A. Vescovo and L. Anolli, A new methodological approach to nonverbal behavior analysis in cultural perspective. Behavior Research Method, (in press).
[43] A. Blanchet, M. Batt, A. Trognon and L. Mass, Language and behavior patterns in a therapeutic interaction sequence. In L. Anolli, S. Duncan, M.S. Magnusson, G. Riva, Eds. The Hidden Structure of Interaction, IOS Press, Amsterdam, 210-223, 2005.
[44] P. Ekman and W.V. Friesen, Manual for Facial Action Coding System. Consulting Psychology Press, Palo Alto, CA, 1978.
[45] D. Morris, Human and his gestures (L'uomo e i suoi gesti). Arnoldo Mondatori, Milano, 1990.
[46] W.C. McGrew, An ethological study of children's behavior. Academic Press, New York, 1972.
[47] A. Borrie, G.K. Jonsson and M.S. Magnusson, Temporal pattern analysis and its applicability in sport: an explanation and preliminary data. Journal of Sport Science, (2002), 20, 845-852.
[48] K. Hook, K. Isbister and J. Laaksolhati, Sensual evaluation instrument. Proceedings of CHI 2005 Conference. Workshop on "Evaluating Affective Interfaces", 2005.
[49] R.W. Picard and S.B. Daily, Evaluating affective interactions: Alternative to asking what users feel. Proceedings of CHI 2005 Conference. Workshop on "Evaluating Affective Interfaces", 2005.
[50] R.L. Mandryk, Evaluating Affective Environments Using Physiological Measures. Proceedings of CHI 2005 Conference. Workshop on "Evaluating Affective Interfaces", 2005.
[51] L. Anolli, F. Mantovani, M. Balestra, A. Agliati, O. Realdon, V. Zurloni, M. Mortillaro, A. Vescovo and L. Confalonieri, The Potential of Affective Computing in E-Learning: MYSELF project experience. Proceedings of INTERACT 2005 Conference - Workshop on "eLearning and Human-Computer Interaction: Exploring Design Synergies for more Effective Learning Experiences". Rome, 12-15 september 2005.
[52] L. Anolli, F. Mantovani, M. Mortillaro, A. Vescovo, A. Agliati, L. Confalonieri, O. Realdon, V. Zurloni and A. Sacchi, A Multimodal Database as a Background for Emotional Synthesis, Recognition and Training in E-Learning Systems. In: J. Tao, T. Tan, R. Picard, Eds. Affective Computing and Intelligent Interaction, Lecture Notes in Computer Science, Springer Verlag, Berlin, 567-573, 2005.
[53] A. Duranti, Linguistic anthropology. Cambridge: University Press, Cambridge, 1997.