Modeling The User Attitude Towards An ECA

Addolorata Cavalluzzi, Fiorella de Rosis, Irene Mazzotta and Nicole Novielli

Department of Informatics, University of Bari
Via Orabona 4, 70126 Bari, Italy
http://www.di.uniba.it/intint/
+39 080 5443282
{derosis, cavalluzzi, mazzotta, novielli}@di.uniba.it

ABSTRACT

We discuss the results of a Wizard of Oz study which was aimed at investigating which forms of 'social relationship' are established by users with ECAs and how this relationship can be measured with linguistic analysis. We describe how we built a model of the user attitude by learning a dynamic bayesian network from the corpus of data collected with the study.

Keywords

Embodied Animated Agents, empathy, natural language analysis, Bayesian networks.

1. INTRODUCTION

Embodied Conversational Agents (ECAs) are seen as a new metaphor of human-computer interaction which should give the users the illusion of cooperating with a human partner rather than just using a tool. The more these agents succeed in achieving this goal, the more their users are expected to show some sign of 'social relationship' with them: ECAs should be equipped to notice these signs and to respond appropriately. Although a number of evaluation studies have been produced, which describe how users see ECAs and how their vision is influenced by variations in the agent characteristics, the exact nature of the relationship between users and ECAs is still unclear. The Stanford group formulated, in the famous *media equation,* the hypothesis that social science theories may be applied in this domain [17]: recently, the need to specify the applicability conditions of this hypothesis and its rationale was advanced by several authors. Some studies proved that human interaction with technology is not exactly the same as the human-human one, and that humans tend to automatically adapt their dialog style when they are aware of interacting with a tool [19],[7],[5]. This finding brought to organize Wizard of Oz studies to investigate the nature of interaction with technology, either in natural language [6] or with artificial agents: the first corpora of dialogs collected with these studies contributed to elucidate how the user behaviour changes according to the interaction condition and the application domain.

We worked in the last four years at an ECA which is aimed at promoting appropriate eating habits. To design this system, we integrated knowledge from psychological theories about health promotion with analysis of a corpus of human-human dialogs in which the 'client' had serious smoking, drinking or eating problems. In the first prototype of our system, the ECA tried to emulate the behaviour of the 'human therapist', the underlying hypothesis being that the human-ECA relationship should aim at mirroring the human-human one [8]. To test whether this hypothesis was reasonable, we subsequently designed and prototyped a tool to perform Wizard of Oz studies with our ECAs in different conditions: the idea was to employ this tool as an iterative design method for our health promotion dialogs. In this paper, we discuss the results of a study which was aimed at investigating, in particular, which form of relationship is established with ECAs by users, how this relationship can be measured from linguistic analysis of the user moves and how a dynamic bayesian network model of the user attitude can be built from the collected corpus of data.

2. BACKGROUND

Empathy is a quite fuzzy concept: it implies listening skill and emotional intelligence, with the ability to identify with and understand another's situation, feelings and motives. It therefore requires some kind of cognitive evaluation of the interlocutor's situation and may occur even in absence of any expression of emotion by the 'empathizing interlocutor' [21]. Vaknin attributes to this concept a meaning which goes beyond pure emotion transmission, by claiming that: "The empathor empathizes not only with the empathee's emotions but also with his physical state and other parameters of existence" [24]. By accepting this definition, empathy may be defined as the process of entering into a warm social relationship with someone else, of being in a way involved in her goals and feelings: a concept closely related to friendship. The need, for an ECA, to show empathy towards the user has been broadly investigated. Cassel and Bickmore worked at endowing REA with the ability to apply some of the strategies which are employed by humans to facilitate trust and collaboration: increase intimacy and common ground over the course of the conversation, decrease interpersonal distance, use non explicit ways of achieving conversational goals and display expertise. These abilities were implemented by means of variations in the agent's language, the main of them being: (i) to introduce small talk to facilitate intimacy and build common ground; (ii) to induce emotional contagion by verbal and nonverbal affect expression and (iii) to increase credibility by means of expert's jargon [3]. Although an increase in the overall effectiveness of interaction induced by an empathic attitude of the agent could be proved by the evaluation studies performed in that project, much less clear was whether and how the users really felt (and showed) empathy for the ECA and whether feeling empathy

contributed to their overall evaluation. Finding a circumstantiated answer to this question is crucial for designing an ECA which is aware of the user attitude and is able to react appropriately. If we assume that the user-agent relationship is symmetrical, we may hypothesize that users display their empathic attitude towards the agent with the same forms of expression which are employed by ECAs to this aim: in particular, attempts to increase intimacy and decrease interpersonal distance, attempts to establish a common ground and use of affective language. Humorous acts may also be taken as an offer of sympathy, as indirect indices of attempt to manifest an empathic relationship with the agent: "When the participants are in the mood for jokes, joke telling occurs naturally and there is some meta-level cooperation" [18].

Though not being synonyms, friendship and empathy are closely related concepts. Friendship may involve varying types and degrees of companionship, intimacy, affection and mutual assistance. It is influenced, again, by interpersonal attraction but also by rewards, which should outweigh costs such as irritation or disappointment. In advice giving dialogs, rewards are affected by the subject's expectation (information and, maybe, also fun). Therefore, even if (as in our study) subjects are pre-informed that the ECA with which they are going to interact is still in a prototypical stage, their involvement is probably affected by the degree of satisfaction in the information received and by how pleasant they find interacting with it. The subjects' overall evaluation of the ECA and the dialog will probably depend on their personality, their interest for the dialog topic, their previous level of information on that topic and others.

3. OUR STUDY

As we said, our study was aimed at studying whether some form of social relationship might occur in interacting with an ECA. As we wanted to apply measuring methods that went beyond subjective ratings of the agent characteristics, we employed an experimental setting which was based on a Wizard of Oz tool. This tool enables us to perform experiments in various conditions, by varying the physical aspect of the agent, its expressivity, the dialog moves, the evaluation questionnaire and other factors. Data of various kinds may be collected: subjects may be asked to evaluate the individual agent moves as well as its overall behaviour. On the other side, the resulting corpus of human-agent dialogs may be employed to perform more analytical studies of the subjects attitude by means of a linguistic analysis of their moves. The head-only embodied agents we employ in our experiments are built with a commercial software (Haptek, see website): their voice may be rendered with a text-to-speech (TTS) synthesizer in Italian or in English. This flexibility enables us to diversify the dialog content, that is the 'moves' the agent can pronounce and to employ a gallery of characters with a more or less realistic voice and more or less emphasized facial expressions. In the study described in this paper, we manipulated these parameters in a controlled way, by setting the study conditions at every step according to the particular hypothesis we wanted to test in that step. Our application domain was that of health promotion (in particular, suggestions about diet), in which we already got a considerable experience with the evaluation of character's monologs [1].

3.1 Method

To insure uniformity of experimental conditions throughout the whole study, we established some rules the wizard was requested to follow. After every subject move, the wizard selected her next move so as to respect a well defined dialog plan and to insure, at the same time, internal coherence in every dialog. This was achieved by a careful preliminary training of the wizard and by employing the same wizard with all subjects. We employed an head-only character with a rather realistic and pleasant aspect (figure 1) and with two kinds of voices: a mechanical and not much natural one (produced with the Microsoft TTS in Italian) and a much more natural one (produced with Loquendo: see website). During the dialog, the subject could evaluate every single agent move by clicking on one of the icons at the right side of the window which indicate, respectively, whether the expression was considered as 'nice', 'unclear' or 'bad'. At the end of the experiment, a final questionnaire was displayed on the video, to collect an evaluation of several features of the message and of the agent, each with a Likert scale from 1 to 6. Items in this questionnaire measured how much *credible*, *plausible*, *clear*, *useful* and *persuasive* was the message and how much *sincere*, *likable*, *natural*, *intelligent* and *competent* was the agent.

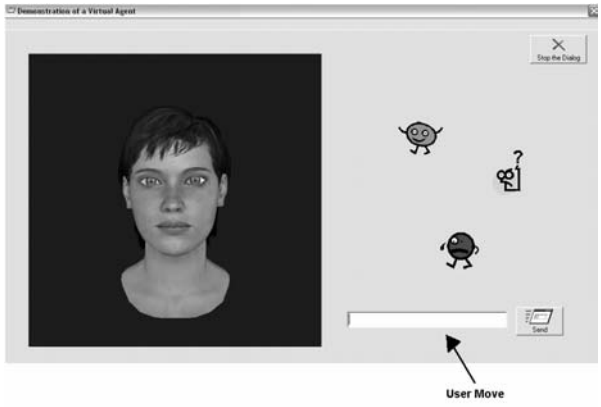

Figure 1: The character employed in our Woz studies

Dialogs were stored in a log at the end of interaction with every subject, to be analysed also from a 'qualitative' and deeper viewpoint. We defined, first of all, two measures of the subject attitude during the dialog:

- *Level of involvement*: a function of the average number of subject moves in a dialog and of their average length, and
- *Degree of initiative*: a function of the ratio between questions raised by the subject and overall moves.

These measures were integrated with a set of 'signs of social relationship' that we drew from a linguistic analysis of the subject moves. These signs enabled us to evaluate the *degree and kind of social relationship* of the subject with the agent and to assess the relation between overall evaluation of the agent and the dialog (with the final questionnaire), level of involvement, degree of initiative and forms of expression of social relationship.

3.2 Main Results

We performed 6 tests, with 5 subjects in each of them (Table 1). These tests were considered as steps of an 'iterative design' of our ECA: therefore, in designing every step we considered the results of the previous ones to find out the main limits of the ECA and revise its behaviour. After the first three tests, we could stabilize the agent moves and behaviour and we recruited subjects with a different cultural background, to evaluate the possible role played by this factor.

A pre-test questionnaire enabled us to verify that the six groups of subjects were comparable in their level of knowledge, habits and interest for healthy eating, and in the importance given to it. They belonged to the same age group (23 to 26) and were equi-distributed in gender. The length of the dialogs (in n of adjoint pairs[1]) ranged from 9 to 60 and increased only slightly with the number of overall moves among which the wizard could choose her answers (22.4 for T1&T2, vs 25.5 in T3-T6). The average length of moves for every subject ranged from 29 to 95 characters.

Table 1: tests performed

Test ID	Ag move available	Subject background	Agent behaviour
T1	53	Degree in humanities	'cold' style; Microsoft TTS
T2	53	Degree in humanities	'warm' style; Microsoft TTS
T3	84	Degree in humanities	intermediate style; 'social' agent moves added; Microsoft TTS
T4	84	Student in CS	as in T3; Loquendo TTS
T5	84	PhD Student in CS	as in T3; Loquendo TTS
T6	84	PhD student in CS	as in T3; Loquendo TTS

To analyse which factors influenced overall evaluation of the dialog, we computed a *message rating* and an *agent rating* as averages of the scores assigned by every subject to the various features (of the message and of the agent respectively) in the final questionnaire. A multiple regression analysis showed that the message rating was associated positively with the ratings in the initial questionnaire and the percentage of subject moves tagged as 'social' (see next Section). On the contrary, it was correlated negatively with the dialog duration (n. of moves), the average length (in characters) of subject moves and the percentage of questions in a dialog. This showed that the subjects' evaluation of the message was not associated positively (as we expected) with their degree of involvement and of initiative in the dialog. The percentage of social moves was associated positively with the degree of involvement, while it was correlated negatively with their level of initiative. The subjects' background was the factor which mostly influenced their behaviour: computer scientists made shorter dialogs with shorter moves, a larger proportion of questions and a lower proportion of social moves than subjects with a background in humanities. More detailed data on this quantitative analysis of our corpus may be found in [9].

Overall, our corpus included 708 subject moves, that we asked three raters to label manually, to identify those of them which showed some sign of the subject attempt to establish a social relationship with the ECA. The following are the language features that we considered as signs of this kind. For each of them, we provide an example of adjoint pair which was extracted from the logs of our experiments and translated from Italian: some pairs belong to several classes.

a. Friendly self-introduction

The first move of the ECA is to briefly introduce herself by describing her name and role. The subjects sometimes answer by briefly introducing themselfs as well [2]:

Oz: Hi. My name is Valentina. I'm here to suggest you how to improve your diet.

S: *Hi, my name is Isa and I'm curious to get some information about education to healthy eating*

b. Familiar expressions

Some subjects employ a familiar language in their moves:

Oz: Are you attracted by sweets?

S: *I'm crazy for them.*

or introduce dialectal expressions or proverbs:

Oz: I know: somebody may think that eating, and maybe living, in a messy way is pleasant, and maybe they are right. But, in the long term, negative effects may occur.

S: *Somebody says that one day as a lion is better than a hundred days as a sheep.*

or argue informally about the suggestion received

Oz: There seems to be a common agreement on the idea that limiting the amount of fat, in particular the 'saturated' one, is a fundamental rule of healthy dieting.

S: *But this takes away the pleasure of eating!*

c. Personal information

Providing personal information even when not requested may be seen as a sign of intimacy, as in the following examples:

Oz: Do you like sweets? Do you ever stop in front of the display window of a beautiful bakery?

S: *Very much! I'm greedy!*

d. Humor and irony

As we said in Section 2, answering with humorous forms to the agent's questions or suggestions is a sign of 'offer of sympathy'; for example:

Oz: I know we risk to enter into private issues. But did you ever try to ask yourself which are the reasons of your eating habits?

S: *Unbridled life, with light aversion towards healthy food.*

[1] An *adjoint pair* is a couple of adjacent wizard-subject moves in the dialog.

[2] Oz stays for 'Wizard', S for 'Subject'

e. Personal questions about the agent

These may be seen as signs of attempts, by the subject, to induce the agent to reciprocate manifestations of intimacy and decrease interpersonal distance:

Oz: What did you eat at lunch?

S: *Meet-stuffed peppers. How about you?*

f. Benevolent or polemic comments

These may be seen as signs of involvement or disappointment; for instance:

Oz: I'm sorry, I'm not much expert in this domain.

S2: *OK: but try to get more informed, right?*

g. Requests to carry on interaction

If, when the agent tries to close the dialog, the subject asks to carry it on, this may be seen as a sign of engagement:

Oz: Goodbye. It was really pleasant to interact with you. Come back when you wish.

S1: *But I would like to chat a bit more with you.*

3.3 Discussion

A comment about the role of the subject's background: quantitative and qualitative analysis of the dialogs showed that subjects with a background in computer science differed considerably, in their attitude towards the ECA, from those with a background in humanities. Their high level of initiative in the dialog corresponded to an attitude of 'trying to challenge the system' rather than being really interested in getting information; they seemed to be less involved in conversating with the ECA and more 'cold' in their answers and comments than subjects with a background in humanities. Although the corpus we collected so far is not large enough to enable us to make any strong conclusion, we would tend to believe that ECAs are not the ideal form of interaction for computer scientists, at least in their present stage of development: they promise a really natural interaction that they are not really able, so far, to keep.

To be effective in its advice-giving purpose, the agent should be ableto recognize the attempts to establish a social relationship displayed by the user, and react appropriately. When interaction with the agent is via a keyboard, the only information source available is text: for this reason, we tried to assess whether and and how a predictive user model may be built with this kind of information.

4. LABELLING THE CORPUS

We defined a markup language for 'signs of social relationship' and extracted 237 moves from the corpus of WOZ dialogs, to ask three independent ratersto annotate them. We considered the label of a move as 'agreed' when at least two raters gave it the same value: agreement rates were therefore computed with a principle of 'majority voting'. Their values for the various signs of social relationship are shown in Table 2. There was a high agreement among the raters for 'Friendly self-introduction', 'Talks about self', 'Questions about the agent', 'Irony', 'Favourable comments' and 'Friendly farewell', and a very low agreement on 'Familiar style'. Favourable comments were more frequent but less agreed than negative ones. This might be due to a imprecise definition of the mentioned categories in the markup language but also to an objective difficulty of recognizing these signs from language features only rather than from a combination of multimedia signs.

'Percentage agreement' indices provide an immediately interpretable measure of the quality of a labelling method and of the difficulty of recognizing the expression of some aspect of the mental state in the language employed. However, they suffer of beingindependent of the agreement rate one would expect by chance: 'chance-corrected measures' like kappa [10] respond to this need. As the value of kappa depends on how skew is the distribution of the considered variable, given an agreement rate the value of kappa depends on the frequency of the sign of interest. For instance: humor and irony were very unfrequent phenomena in our corpus; therefore, their kappa was low even if the agreement rate was not bad. A similar cosideration can be applied to 'comments'.

Table 2: Agreement among raters

Signs of social relationship	**Agreement rate**	**Frequency**	**Kappa**
Friendly self-introduction	.98	4 %	.87
Familiar style	.33	59 %	.16
Talks about self	.73	33 %	.64
Questions about agent	.70	31 %	.56
Humor and irony	.84	5 %	.36
Favourable comments	.82	5%	.42
Neutral comments	.68	81 %	
Negative comments	.86	12 %	
Friendly farewell	.93	7 %	.65

We committed the role of recognizing the social attitude of the user towards the ECA to a combination of a very simple parser with a dynamic user model. Criteria applied in parsing are described in Table 3: they combined knowledge about the sign semantic with analysis of word salience in the corpus: a word was considered as 'salient' for a category if it applied more often in the category than in other parts of the corpus [13]. The predictive capacity of he parser for every sign was evaluated from confusion matrices in terms of sensitivity (true positives over all the positive cases), specificity (true negatives over all the negative cases) and proportion of correctly classified cases (% of CCC). Table 3 shows that the specificity was high for all signs while the sensitivity was low for some of them (negative comments and familiar style in particular).

5. LEARNING A MODEL OF THE USER'S EMPATHIC ATTITUDE

The function we assign to our user model is to enable the agent to infer how the social attitude of the user evolves during the dialog in relation to the dialog history. Social attitude is a 'hidden' variable in the model, while observable ones are the user's stable

characteristics, the context in which the move was entered and its linguistic features (parser results). As the relationships among these variables are uncertain, we represent this model with a dynamic bayesian network.

Table 3: Recognition criteria and predictive capacity of the parser

Signs	**Criteria**	**Sensitivity**	**Specificity**	**% of CCC**
Friendly self-introduction	Expressions of greetings ('ciao', 'hello',..) or of self-presentation ('my name is…')	0.91	0.98	0.97
Familiar style	Agent name ('Valentina'), interjections ('!', 'Hurrah',…), friendly lexicon ('papa', 'mummy', 'greedy', 'chat', 'my passion', 'dear', …), dialectal expressions ('cute', 'espressino',…), diminutive or espressive forms ('little sweet', 'fatty', …)	0.36	0.96	0.79
Talks about self	Personal pronouns ('I', 'my', 'to me',..), auxiliary verbs ('I have', 'I am',…), expressions of knowledge ('I know', I believe',…), of attitude ('I try', 'I think', 'I tend to', 'I care of',…), domain verbs ('I eat', 'I drink',.. all at the first person	0.81	0.79	0.80
Question about the agent	Similar to the previous one, but at the second person	0.80	0.92	0.91
Positive comments	Expressions of agreement ('OK', 'right', 'good', 'true',…), of attitude ('I agree', 'I trust',…), of opinion about the agent ('That's kind of you',…)	0.56	0.94	0.92
Negative comments	Objections ('no', 'but',…) negative evaluations about the agent ('you are rude', 'you don't know', 'you don't understand') or about the message received ('this is too much','too little', …)	0.16	0.98	0.93
Friendly farewell	2. Expressions of farewell ('bye', 'see you soon', …), of thanking and wishes ('thanks', …)	0.83	0.97	0.96

We applied the K2 Bayesian network learning algorithm [4] which is employed in Bayesware (http://www.bayesware.com) to the corpus of 712 dialog moves. This algorithm reduces the space of possible BN structures by receiving as a binding an ordering of the variables from which the network will be built. It is appropriate in our case (and, we claim, more in general in learning user models from datasets) because it enables distinguishing 'trigger variables' (at the top level of the network) from those describing the resulting behavior of the user (leaf nodes). In our case, variables introduced belonged to the following categories (see Table 4):

a. *Stable and known subject characteristics*: background (in humanities or in computer science) and gender;
b. *Dynamically evolving, unknown subject characteristics*: attitude towards the ECA;
c. *Single move, easily recognizable characteristics*: context (type of previous agent's move) and subject's move length;
d. *'hidden' subject move characteristics*: these are the signs of social relationship described in Section 3.2;
e. *'observable' linguistic features of the subject move*: we developed a very simple parsing method to recognize only six categories of expressions according to short word sequences (one, two or three words).

Table 4: Variables associated with nodes in figure 2

Variable category	**Variable name**	**Label**
Stable user characteristics	Background	Back
	Gender	Gend
Context	Type of last Agent move	Ctext
	Type of user move	Mtype
	User move length	Mleng
Dynamic variable	User attitude towards the agent	Frsh
Signs of social attitude	Familiar style	Fstyl
	Friendly self-introduction	Fsint
	Talks about self	Perin
	Questions about agent	Qagt
	Irony	Irony
	Friendly farewell	F-Fw
	Comments	Comm
Results of parsing	Cues of familiar style	P-fst
	Cues of friendly self introduction	P-Fsin
	Cues of talks about self	P-pin
	Cues of questionsto the agent	P-qag
	Cues of friendly farewell	P-Ffw
	Cues of comments	P-comm

We tested several kinds of models, by setting a search order according to the intended use of the model: stable and known subject characteristics were taken as network roots, linguistic features as leaves and other variables as intermediate nodes. We initially learned automatically both the structure and the parameters of the network, and subsequently refined interactively the 'best' model found in terms of MLL (maximum logarithmic likelihood) by introducing a few links between some nodes, to avoid problems in evidence propagation due to d-separation properties or to too long paths between observable nodes and the main node to monitor (friendship). The resulting structure is shown in figure 2: in this figure, node labels correspond to the variables described in Table 4.

Some comments on the static part of the model. Conditional probability tables associated with the leaf nodes represents the results of validity analysis of the parser. Familiar style is employed especially when their attitude towards the agent is 'friendly' and in humorous moves: these prevail in subjects with a background in humanities. Very long moves are more likely in subjects with this background. Users may omit farewell and self-introduction when their social attitude is not much friendly: in this case, they may ask questions especially after receiving a suggestion or a negative answer by the agent, or even immediately after it introduced itself. User comments frequently come after agent comments especially when the user attitude is friendly, ... and so on.

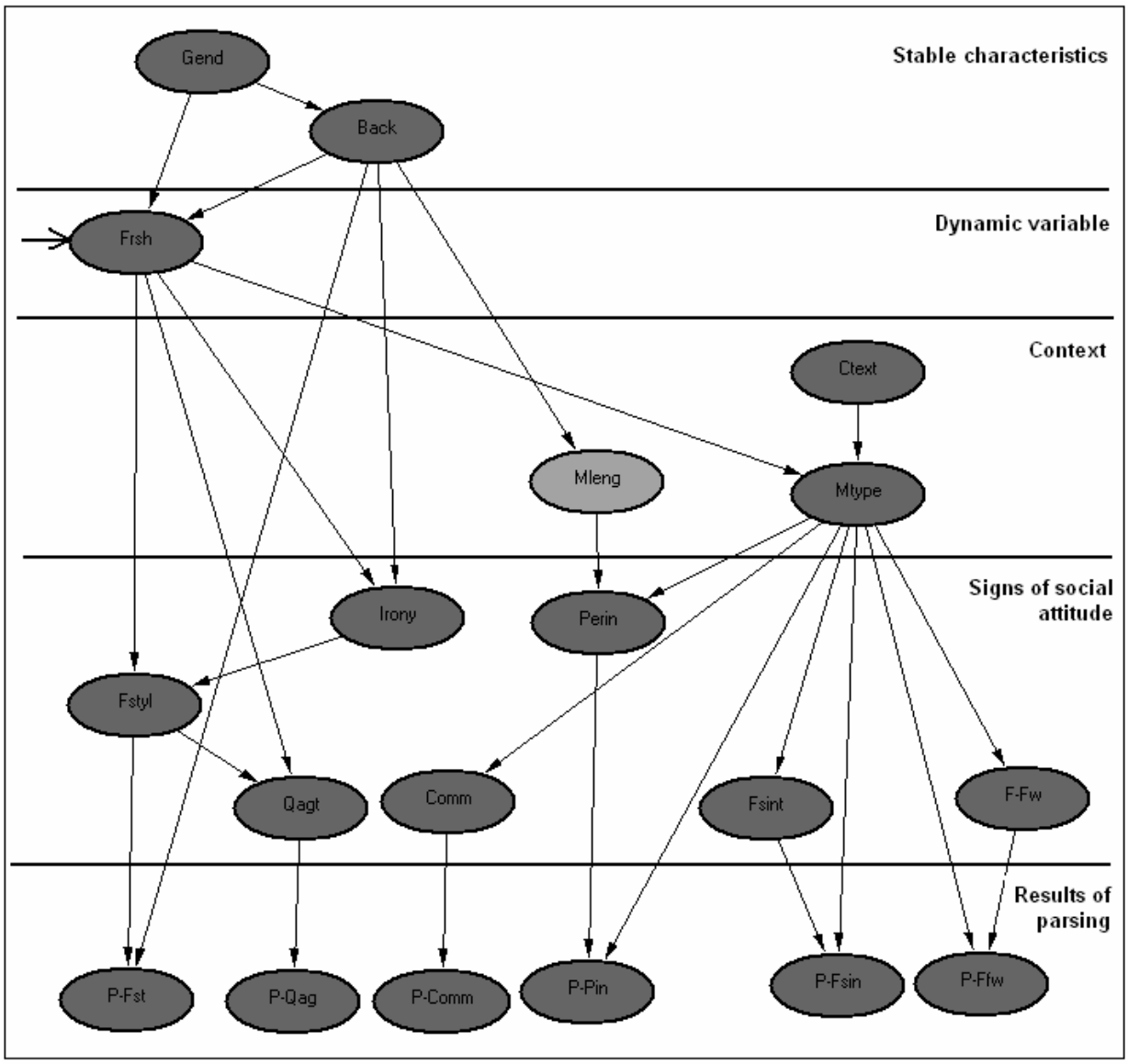

Figure 2: Outline of the user model

In the dynamic belief network, every time slice corresponds to a user move, the prior of the user's gender and background do not change from time to time, and the only temporal link between two consecutive time slices is assigned to the social attitude (Frship). The model develops as follows:

- *at the beginning of interaction*, the DBN is initialized by propagating evidence about stable user characteristics; the probability of friendship takes a prior value which corresponds to this category of users;
- *after every user move*, the length of the user move (Mleng) is calculated, some of its linguistic features are analysed with the parser and these results are introduced and propagated in the network together with evidence about the agent's move (Ctext);
- *the new probability of the 'friendship' node* is read and contributes to planning the next agent move.

Combining results of parsing with knowledge about the subject, the context and the attitude of the user in the previous phases of the dialog enables us to reduce the difficulty in the recognition of the user attitude by means of linguistic analysis which was acknowledged by several researchers (for instance, [14], [2], [11]). We did, so far, only an 'internal' evaluation of the predictive value of the dynamic model on a small subset of the dialogs. Models induced from a sample of data tend to be overtrained, that is to expect future events to be like the events onwhich they were trained [16]. We therefore plan to extend testing to an external dataset, after collecting a new corpus of dialogs from subjects with similar characteristics of those included in the training set (young people with different backgrounds).

6. CONCLUSIONS

We learned a lot from our experience of iterative prototyping of affective health promotion dialogs. We initiated our studies with the belief that a key requirement of dialog simulation was the recognition of the emotional state of the users. This is true when the user problems are serious and therefore produce a strong emotional state (as in the case of natural dialogs with a therapist about drinking and smoking). On the contrary, when the user problems are less serious, different kinds of emotions emerge in interaction: rather than strong 'individual' emotions like fear, joy, anxiety, relief etc, softer 'social' emotions like sympathy or antipathy, tenderness, contempt, sense of belonging occur [22]. To increase the effectiveness of advice-giving, the ability to recognize the degree of involvement of the user and to manifest reciprocity of social relationship seems to be more important than displaying realistic expressions of emotions in the agent's face. This opens complex problems, like recognising and responding to humorous acts [23], formulating moves in a 'familiar' style, adding the ability to talk about 'self' and so on: and this, as everybody knows, is a typical category of 'open problems' in ECA's design and implementation.

7. ACKNOWLEDGEMENTS

This work was financed, in part, by HUMAINE, the European Human-Machine Interaction Network on Emotion (EC Contract 507422). We thank Berardina De Carolis and Giuseppe Clarizio for cooperating in the implementation of the WoZ tool, Loquendo for providing us their software in the scope of a scientific cooperation and Haptek Inc. for assisting us in the application of their software.

8. REFERENCES

[1] Berry, D., Butler, L. and de Rosis, F. *Evaluating GRETA. The importance of consistency of behaviour in a multimodal animated agent*, in press.

[2] Carberry, S., Lambert, L. and Schroeder, L. Towards recognising and conveying an attitude of doubt via natural language. *Applied Artificial Intelligence*, 2002.

[3] Cassell, J. and Bickmore, T. Negotiated collusion: modelling social language and its relationship effects in intelligent agents. *User Modelling and User-Adapted Interaction*, 13, 1-2, 2003.

[4] Cooper, G.F. and Herskovitz, E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* , 9, 1992.

[5] Coulston, R., Oviatt, S. and Darves, C. Amplitude Convergence in Children's Conversational Speech With Animated Personas. In *Proceedings of the 7th International Conference on Spoken Language Processing* (ICSLP2002). J. Hansen and B. Pellom (Eds). 2002.

[6] Dahlback, N., Joensson, A. and Ahrenberg L. Wizard of Oz Studies. Why And How. In *Proceedings of the Int Workshop on IUI*, 1993.

[7] Darves, C. and Oviatt, S. Adaptation of Users' Spoken Dialogue Patterns in a Conversational Interface. In *Proceedings of the 7th International Conference on Spoken Language Processing* (ICSLP2002). J. Hansen and B. Pellom (Eds). 2002.

[8] de Rosis, F., Pelachaud, C., Poggi, I., Carofiglio V., and De Carolis, B. From Greta's Mind to her Face: Modeling the Dynamics of Affective States in a Conversational Embodied Agent. *International Journal of Human-Computer Studies*. B.R.Gaines. Vol.59 (1-2) (2003).

[9] de Rosis, F., Cavalluzzi, A., Mazzotta, I. and Novielli, N.: Can embodied conversational agents induce empathy in users? *AISB Virtual Social Characters Symposium*, 2005.

[10] B.Di Eugenio and M. Glass. The Kappa statistics, a second look. *Computational linguistics*, 2004

[11] Gill, A. and Oberlander, J. Taking care of the linguistic features of extraversion. *24th Annual Conference of the Cognitive Sciene Society*, 2002.

[12] Haptek website: http://www.haptek.com

[13] C.M.Lee, S.S. Narayanan, and R. Pieraccini. Combining acoustic and language information for emotion recognition. *Proceedings of ICSLP*. 2002..

[14] Litman, D., Forbes, K. and Silliman, S.Towards emotion prediction in spoken tutoring dialogues. In *Proceedings of HLT/NAACL* 2003.

[15] Loquendo website: http://www.loquendo.com/

[16] C.D. Manning and H. Schutze. *Foundations of statistical natural language processing*. The MIT Press, 1999.

[17] Nass, C., Isbister K. and Lee, E-J. *Truth is beauty: Researching Embodied Conversational Agents*. In J. Cassell, J. Sullivan, S. Prevost and E Churchill: Embodied Conversational Agents. The MIT Press, 2000.

[18] Nijholt, A. Observations on humorous act construction. In *Proceedings of EMCSR* 2004.
http://wwwhome.cs.utwente.nl/~anijholt/artikelen/emcsr2004.pdf

[19] Oviatt, S. and Adams, B. *Designing and Evaluating Conversational Interfaces With Animated Characters*. In J Cassell, J Sullivan, S Prevost and E Churchill: Embodied Conversational Agents. The MIT Press, 2000.

[20] Picard, R. What does it mean for a computer to 'have' emotions? In R. Trappl, P. Petta and S. Pays (Eds): Emotions in humans and artefacts. *A Bradford Book*, MIT Press, 2002.

[21] Poggi, I. Emotions from mind to mind. In A Paiva (Ed): *Proceedings of the Workshop on "Empathic Agents"*. AAMAS 2004.

[22] Poggi, I. and Magno-Caldognetto, E. Il parlato emotivo. Aspetti cognitivi, linguistici e fonetici. In *Proceedings of the Conference "Il Parlato italiano"*. D'Auria, Naples, 2003

[23] Stock, O. and Strapparava, C. An experiment in automated humorous output production. In *Proceedings of Intelligent User Interfaces*, 2003.

[24] Vaknin, S. On Empathy.
http://samvak.tripod.com/empathy.html