It's All in the Game: Towards an Affect Sensitive and Context Aware Game Companion

Ginevra Castellano
Dept. of Computer Science
School of EECS
Queen Mary Univ. of London
United Kingdom
ginevra@dcs.qmul.ac.uk

Iolanda Leite, André Pereira,
Carlos Martinho, Ana Paiva
INESC-ID
Instituto Superior Técnico
Porto Salvo, Portugal

Peter W. McOwan
Dept. of Computer Science
School of EECS
Queen Mary Univ. of London
United Kingdom

Abstract

Robot companions must be able to display social, affective behaviour. As a prerequisite for companionship, the ability to sustain long-term interactions with users requires companions to be endowed with affect recognition abilities. This paper explores application-dependent user states in a naturalistic scenario where an iCat robot plays chess with children. In this scenario, the role of context is investigated for the modelling of user states related both to the task and the social interaction with the robot. Results show that contextual features related to the game and the iCat's behaviour are successful in helping to discriminate among the identified states. In particular, state and evolution of the game and display of facial expressions by the iCat proved to be the most significant: when the user is winning and improving in the game her feeling is more likely to be positive and when the iCat displays a facial expression during the game the user's level of engagement with the iCat is higher. These findings will provide the foundation for a rigorous design of an affect recognition system for a game companion.

1. Introduction

New applications are emerging where robots will play an important role as companions. As the average age of the population in many countries is increasing elderly care is likely to become more problematic. Robot companions may be able to assist users by providing additional functionalities. Companions could also be employed as personal assistants and represent a valuable tool for therapy and rehabilitation purposes, for example by encouraging interactions between people affected by social and cognitive disabilities (e.g., people with autism [20]). They could also be used in edutainment and entertainment applications, for example, design of socially intelligent and interactive toys.

To adopt the role of a companion, a robot must be able to display a social, affective behaviour [2] [5]. One of the most relevant prerequisites of companionship is the ability to sustain long-term interactions. To do so, it is important that robot companions contain a module for affect recognition that analyses both multimodal behavioural cues of the user and information about the context in which the interaction takes place. While researchers have been increasingly investigating affect recognition [24], the design of such a module for use in a human-robot interaction framework has not been extensively addressed yet.

A robot companion's recognition abilities must go beyond the detection of prototypical emotions and be sensitive to application-dependent affective states and cues. They must be trained using naturalistic, spontaneous expressions and integrate contextual information [3]. While some efforts in this direction have been reported in the literature (e.g., see [8] and [7]), work on affect recognition abilities for a specific interaction scenario that take into account contextual information are still not numerous.

In this paper we investigate the importance of contextual features for modelling user states in a specific interaction scenario where an iCat robot plays chess with children. We identify the user's states that the robot must be sensitive to in this particular scenario: the valence of the feeling experienced by the user and her engagement with the iCat. The modelling of these states and their evolution over time will allow the robot to behave in a more social and emphatic manner, which is important to sustain long-term interactions with children. Contextual features related to the game and the iCat's behaviour are also selected as a source of information that could be integrated with behavioural cues for the prediction of the user's states. We show how some of these features are significantly more effective than others in discriminating among the selected user states. In particular, state and evolution of the game and display of facial expressions by the iCat seem to play a major role in the discrimination of user states: when the user is winning and improving in the game her feeling is more likely to be positive and when the iCat displays a facial expression during the game the user's level of engagement with the iCat increases. We intend to exploit these results to carry out a rigorous design of an affect recognition system for a game companion.

The paper is organised as follows. In the next Section a brief literature review on systems designed to recognise application-dependent, spontaneous affective states in human-agent and human-robot interaction is presented. Afterwards, we present our interaction scenario, the user's states that we intend to model and the contextual features that can be extracted during the interaction between users and robot. Finally the data collection procedure is described and a discussion concerning the experimental results and the conclusion are reported.

2. Related work

While still not numerous, some examples of systems designed to recognise application-dependent, spontaneous affective states in human-agent and human-robot interaction were reported in the literature.

Kulic and Croft [10] proposed an HMM-based system capable of estimating valence and arousal elicited by viewing robot motions using physiological data such as heart rate, skin conductance and corrugator muscle activity. Kapoor et al. [7] designed a system that can automatically predict frustration of students interacting with a learning companion by using multimodal non-verbal cues including facial expressions, head movement, posture, skin conductance and mouse pressure data. Peters et al. [18] modelled user level of interest and engagement using eye gaze and head direction information during an interaction with a virtual agent displaying shared attention behaviour.

Examples are also reported of systems where recognition of application-dependent, spontaneous affective states and cues is achieved by exploiting higher level information such as the context where an affective state originates. Kapoor and Picard [8], for example, proposed an approach for the detection of interest in a learning environment by combining non-verbal cues and information about the learner's task (e.g., level of difficulty and state of the game). COSMO is a pedagogical agent that provides feedback to a learner's problem-solving activities by generating contextually appropriate expressive behaviours [13]. Another example within the field of pedagogical agents is Prime Climb, an educational game created by Conati et al. [4]. In this scenario an intelligent pedagogical agent helps users to succeed in a math game. Inspired by the OCC model of cognitive appraisal [17], they developed an affective model that aims to predict multiple emotions experienced by users. The combination of the latter with a model of the user's knowledge was investigated in order to determine how the agent could intervene so as to maximise the trade-off between student learning and engagement. Martinho et al. [14] proposed a framework to infer affective states of children that collaborate in a virtual theatre to create stories based on their behaviour and experienced situations. These situations are interpreted using the OCC model to allow the system to make inferences about possible affective states of the users.

3. Towards an affect sensitive and context aware game companion

The aim of this research is to build an affect sensitive companion that plays board games such as chess. In order to design a game companion capable of sustaining long-term interactions, the recognition of user states is an issue of primary importance. For this reason, the selection of user states to be modelled needs to take into account the specific interaction scenario.

Behavioural and contextual features used to train an affect recognition system must also be selected depending on the application. In our specific interaction scenario, contextual information is likely to improve the understanding of behavioural cues displayed by the user and was therefore identified as an important focus of investigation for the design of a context aware affect recognition system.

3.1. Scenario

The scenario involves a social robot, the iCat [23], that plays chess with children using an electronic chessboard (see Figure 1). The iCat acts as a game companion, helping children to improve their chess skills [11]. While playing with the iCat, children receive feedback from their moves on the chessboard through the robot's facial expressions, which are generated by an affective system influenced by the state of the game. The iCat's affective system is self-oriented, which means that when the user makes a good move, the iCat displays a sad facial expression, and when the user makes a bad move, it expresses positive reactions. The robot's affective expressions are determined by the

Figure 1. A user playing chess with the iCat.

emotivector [15], an anticipatory system that generates an affective signal resulting from the mismatch between the expected and the sensed values of the sensor to which it is coupled to. In this case, the emotivector is coupled to the values of the chess evaluation function that determines the moves played by the iCat. After each move played by the user, the chess evaluation function returns a new value, updated according to the current state of the game. The emotivector system captures this value and, by using the history of evaluation values, an expected value is computed (applying the moving averages prediction algorithm [6]). Based on the mismatch between the expected and the actual sensed value (the last value received from the evaluation function), the system generates one out of nine affective signals for that perception [12]. For instance, after three moves in the chess game, if the iCat has already captured an opponent's piece, it might be expecting to keep the advantage in the game (i.e., expecting a reward) after the user's next move. Therefore, if the user makes a move that is even worse than the one the iCat was expecting (e.g., by putting her queen in a very dangerous position), the generated affective signal will be a "stronger reward", which means "this state of the game is better than what I was expecting". Figure 2 shows the emotivector model: the first row displays the possible outcomes when reward (R) is expected, the second row shows the possible outcomes when the state is expected to remain the same and the third row contains the possible outcomes when punishment (P) is expected (e.g., when, after the user's move, the state of the game is worse than expected). Depending on whether the sensed value is higher, within, or lower than a confidence interval (computed based on the history of mismatches between the expected and the sensed values), the affective signal will belong to the first column (more reward), second (as expected) or third (more punishment).

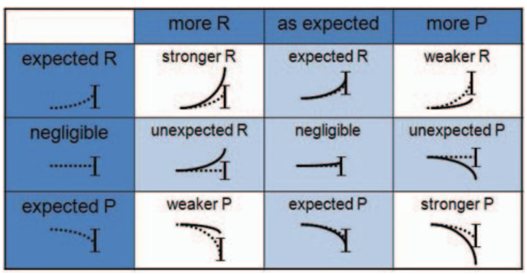

Figure 2. Emotivector model.

Each of the nine affective signals is then mapped onto a different facial expression of the iCat. For example, in the presence of a "stronger reward", the iCat displays a facial expression of excitement. On the other hand, in case of a "stronger punishment", the robot shows a scared expression. For the generation of facial expressions, animations provided by the iCat software library were used (see [1] for an investigation on how users perceive the iCat's emotional expressions).

3.2. User states

Given this specific interaction scenario, an affect sensitive companion should be able to capture user states that are both related to the game and the social interaction with the iCat. The *valence of the feeling* experienced by the user was chosen to measure the degree to which the user's affect is positive or negative [21]. This categorisation of affect appears to be adequate for the purpose of describing the overall feeling that the user is experiencing throughout the game. On the other hand, the user's *engagement with the iCat* was chosen to describe the level of social interaction established between them. We believe that the correct interpretation of the user's level of engagement with the iCat is a factor of primary importance for the establishment of a long-term interaction. Engagement has been defined as "the value that a participant in an interaction attributes to the goal of being together with the other participant(s) and continuing the interaction" [19] and "the process by which two (or more) participants establish, maintain and end their perceived connection" [22]. In human-robot interaction, engagement has been related to the user establishing eye contact with the robot at a comfortable distance [16]. We regard engagement with the iCat as being characterised by an affective and attention component (see [18] for a similar view on engagement in human-agent interaction). We consider the user as engaged with the iCat if she is willing to interact and maintain the interaction with it. This relates, for example, to the amount of time the user looks at the iCat and the presence of an active component in the behaviour displayed by the user, regardless of whether the iCat is doing something or not and whether the iCat's behaviour or the current state of the game induce a positive or a negative feeling in the user.

3.3. Contextual features

The choice of using contextual features for the modelling of the user's states relies on the assumption that, when the user is playing with the iCat, the experienced affective states are related to the events happening in the game or to the behaviour and expressions exhibited by the robot. The contextual features selected for investigation in this interaction scenario are listed as follows.

Game state: is a value that represents the condition of advantage/disadvantage of the user in the game. This value is computed by the same chess evaluation function that the iCat uses to plan its own moves. The more the value of the game state is positive, the more the user is in a condition of advantage with respect to the iCat, the more it is negative, the more the iCat is winning.

Game evolution: is given by the difference between the current and the previous value of the game state. A positive value for game evolution indicates that the user is improving in the game, while a negative value means that the user's condition is getting worse with respect to the previous move.

Captured pieces: indicates if, in the last move played by the user and the iCat, a piece was captured either by the user or the iCat.

User sensations: are calculated with the same mechanism used by the iCat to generate its affective reactions, but from the user's perspective, i.e., taking into account the user's game state. This feature attempts to predict the user's sensations during the game. As an example, consider the game description used to illustrate the iCat's affective reactions in Section 3.1, but from perspective of the user: after three moves in the game the user has lost one piece, so she might expect the iCat to continue holding the advantage (i.e., expecting a "punishment"). If the user plays a terrible move, she might be experiencing something closer to a "stronger punishment" sensation.

iCat facial expressions: as described earlier, after each move played by the user, the iCat displays a facial expression, generated by the emotivector system.

3.4. Data collection

An experiment was performed at a primary school where once a week children have two hours of chess lessons. 5 eight year old children (3 male and 2 female) took part in the experiment. Each participant played two different exercises, one with low and one with medium difficulty. By using different levels of difficulty in each exercise we expected to elicit different types of affective states and behaviours in the children. The exercises consisted of middle game chess positions chosen by a chess instructor who was familiar with each student's chess skills. The children were told to play and try to win against the iCat. The robot begins the interaction by inviting the user to play. The user is always the first to play a move. After each move played by the user the iCat asks her to play its move, as its embodiment does not allow it to do so itself. Each interaction session ends when the user completes both exercises (i.e., by winning, loosing or drawing).

For each exercise, a log file containing the contextual features was created in real-time during the interaction. After each user move, a new entry was written in the log, containing the time since the beginning of the interaction (for synchronisation purposes) and the current value of the contextual features: the game state from the user's perspective, the game evolution, whether there were pieces captured either as a consequence of the current user move or the upcoming iCat move, the user sensations, and the upcoming iCat facial expression.

All interaction sessions were recorded with three video cameras: one capturing the frontal view (e.g., the face of the children), one the side view and one the iCat. The videos recorded with the frontal camera were annotated in terms of user states and contextual features by three annotators. The annotators only had access to the frontal videos displaying the children's behaviour, in order to minimise the factors influencing their decision. 72 video segments containing children displaying different types of behaviour were selected from different phases of the game, starting from the 10 collected videos (two for each participant), using ANVIL, a free video annotation tool [9]. We observed that children display affective behaviour during the whole game and with a higher frequency after the iCat generates an affective reaction. Each video segment had a duration of approximately 7 seconds. Before starting the annotations, the annotators agreed on the meaning of each label to describe the user's state in the videos. As for the valence of the feeling, annotators could choose one out of three options: "positive", "negative" and "cannot say". They were not requested to differentiate between feeling towards the game and feeling towards the iCat. To describe the engagement with the iCat, annotators could choose among "engaged with the iCat", "not engaged with the iCat" and "cannot say". Note that in this experiment the annotators were not requested to code the intensity of the user's state. Each annotator associated labels with each video segment working separately. The results of the three annotation processes were then compared for each video segment: a label was selected to describe the state of the user in a video segment when it was chosen by two or three of the annotators. In case each of the annotators chose a different label, the video segment was labelled as "cannot say". From the annotation process, we randomly selected 15 video segments labelled as "positive", 15 as "negative", 15 as "engaged with the iCat" and 15 as "not engaged with the iCat". Each group of videos contains 3 samples for each participant.

After the annotation of the user's states, the extracted contextual features were added to the annotation of each video. To associate the appropriate log entry to each video, we selected the one that occurred within or right before the video segment.

4. Experimental results

Statistical analysis was performed in order to study if and how the selected contextual features co-occur with the user's states identified in our interaction scenario (i.e., *feeling* and *engagement with the iCat*). Note that this experiment is not intended to be a comprehensive investigation of the relationship between contextual features and user states, but rather it aims to uncover insights to inform choices made in the design of an affect recognition system for a game companion.

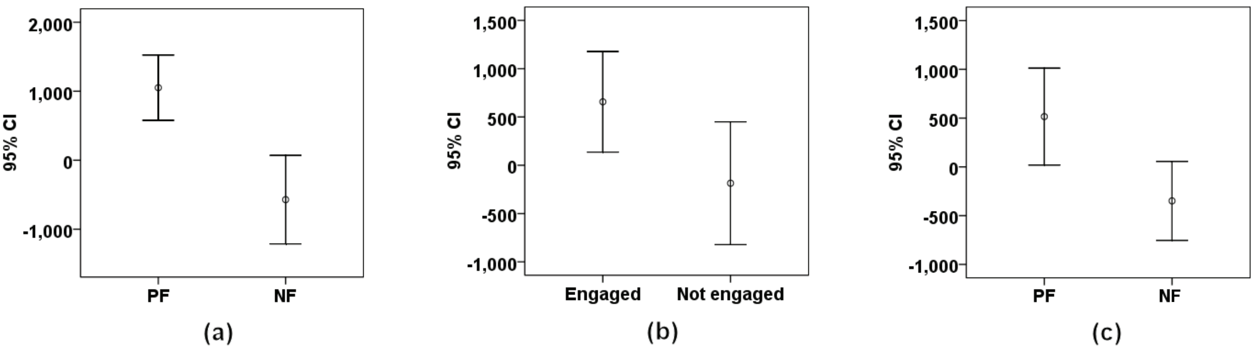

Figure 3. Error bar chart for game state and feeling (a), game state and engagement (b) and game evolution and feeling (c). The confidence interval (CI) is 95 %. The graph shows that when the feeling is positive (PF) the mean of the game state and the game evolution (see the square in the middle of the error bar on the Y axes) is higher than when the feeling is negative (NF) and when the user is engaged with the iCat the mean of the game state is higher than when the user is not engaged with the iCat.

4.1. Game state

In order to investigate whether there is a significant difference between the means of the game state in correspondence with the two conditions for valence of feeling (*positive* and *negative*) and engagement with the iCat (*engaged with the iCat* and *not engaged with the iCat*), two repeated measures t tests (N = 15) were performed for the game state (the dependent variable) with the valence of the feeling and the engagement with the iCat as the independent variables (two levels) respectively in the first and in the second test. It was predicted that higher values of the game state would be observed for *positive* and *engaged with the iCat* samples and that the game state would better differentiate *positive* from *negative* than *engaged with the iCat* from *not engaged with the iCat*.

As far as the valence of feeling is concerned, results show that there is a significant difference between the means of the game state: when the feeling is positive the values of the game state are higher than when the feeling is negative [$t(14) = 3.61$; $p < 0.01$]. Regarding the engagement with the iCat, the values of the game state are significantly higher when the children are engaged with the iCat than when they are not [$t(14) = 2.43$; $p < 0.05$], as predicted, given that it is expected that the situation of the game would affect the user's willingness to interact with the iCat. Figure 3a and Figure 3b show the error bar graphs that illustrate the confidence intervals for each sample mean. Note that game state discriminates better between *positive* and *negative* than between *engaged with the iCat* and *not engaged with the iCat* samples: this is also in line with our hypothesis, as it was expected that the game state would have greater effects on the valence of the feeling than the engagement with the iCat.

Summary of results: when the user is winning, her feeling tends to be positive and her level of engagement with the iCat is higher.

4.2. Game evolution

To investigate whether there is a significant difference between the means of the game evolution in correspondence with the two conditions for valence of feeling and engagement with the iCat, two repeated measures t tests (N = 15) were performed for the game evolution (the dependent variable) with the valence of the feeling and the engagement with the iCat as the independent variables (two levels) respectively in the first and in the second test. It was predicted that higher values of the game evolution would be observed for the samples annotated as *positive* and *engaged with the iCat* and that the game evolution would better differentiate *positive* from *negative* than *engaged with the iCat* from *not engaged with the iCat* samples.

In terms of the valence of feeling, when the feeling is positive the values of the game evolution are significantly higher than when the feeling is negative [$t(14) = 2.62$; $p < 0.01$]. Figure 3c shows the error bar graph that illustrates the confidence intervals for each sample mean. As far as the engagement with the iCat is concerned, no significant difference was found between the means of the game evolution. First of all, even though this result does not confirm our hypothesis, it was expected that the game evolution would have greater effects on the valence of the feeling than the level of engagement with the iCat. Furthermore, improvement in the game does not necessarily mean that the user is winning: the user can improve from one move to another while at the same time being in an overall condition of disadvantage. On the other hand, while it was found that both game state and game evolution significantly affect the valence of the feeling ($p < 0.01$), the computed significance was different ($p = 0.002$ for the game state and $p = 0.010$ for the game evolution).

Summary of results: when the user is improving her condition in the game, her feeling tends to be more positive than negative.

4.3. Captured pieces

In order to establish whether there is an association between the pieces captured during the game and the feeling and engagement with the iCat, two chi-square tests were carried out by considering only the video samples where the event of capturing pieces was present. Feeling (two levels, N = 10) and engagement with the iCat (two levels, N = 8) were considered as the influencing variables respectively in the first and the second test and the pieces captured during the game (two levels: *captured by the user*, *captured by the iCat*) as the dependent variable in both tests. It was expected that the captured pieces would influence the feeling and that pieces *captured by the user* would be associated with a positive feeling, while pieces *captured by the iCat* with a negative feeling.

Results show that there is no significant association between engagement with the iCat and the event of capturing pieces, but a significant association was found between the latter and the feeling [$\chi^2 = 4.29$, $df = 1$, $p < 0.05$]: as expected, when the feeling is positive most of the times the user has captured a piece (60 %), while when it is negative it is more likely that a piece has been captured by the iCat (100 %).

Summary of results: when the user captures a piece during the game her feeling is more likely to be positive, while when a piece is captured by the iCat the user's feeling is more likely to be negative.

4.4. User sensations

To study the association between user sensations and feeling and engagement with the iCat two chi-square tests were performed. Feeling and engagement with the iCat were regarded as the influencing variables (two levels, N = 30) for the first and second test respectively and the user's sensations (eight levels) as the dependent variable in both tests. It was expected that an association would be found between the valence of the feeling and specific categories of the sensations: *expected P*, *expected R*, *stronger P*, *stronger R*, *unexpected P*, *unexpected R*, *negligible* and *none*. No samples in our corpus of videos included instances of the categories *weaker P* and *weaker R*, so that the latter were not considered in the analysis.

Results show a significant association between the valence of the feeling and the user sensations [$\chi^2 = 18.33$, $df = 7$, $p < 0.05$]. Figure 4 shows how the frequencies of the different categories of the user's sensations are linked to the feeling. The feeling is more likely to be positive rather than negative for conditions of reward (20 % vs. 6.7 % for *expected R*, 13.3 % vs. 0 % for *stronger R* and 13.3 % vs. 6.7 % for *unexpected R*) rather than punishment, which are associated more with a negative feeling rather than with a positive feeling (46.7 % vs. 0 % for *stronger P* and 20 % vs. 0 % for *unexpected P*). No difference between positive and negative feeling was found for expected punishment (*expected P*, 20 % vs. 20 %). Furthermore, when the user is predicted as expecting the state of the game not to change from one move to another (*negligible*) the feeling is more likely to be positive (13.3 % vs. 0%) and the same occurs in correspondence with particular circumstances of the game where no expectation is involved, such as the very beginning of the game before the user plays her first move, when the iCat gives up and when the user wins, loses or draws (*none*, 20 % vs. 0 %). No significant association was found between the engagement with the iCat and the user's sensations: this suggests that the level of engagement of the user is not necessarily related to any specific condition of expectation in the game.

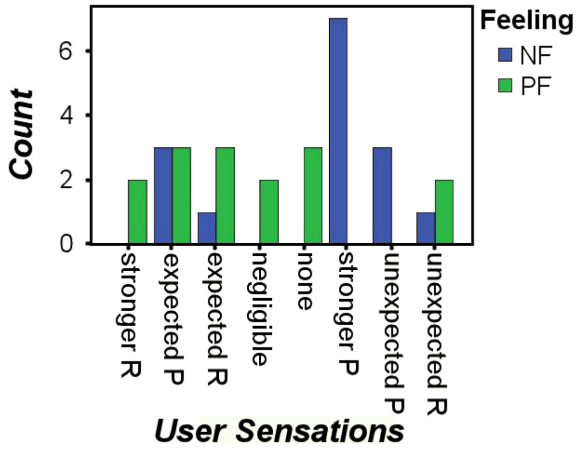

Figure 4. Clustered bar chart for user sensations and feeling (PF = positive feeling; NF = negative feeling). The graph shows that most of the times the feeling is positive rather than negative in case of expected reward (*expected R*, 3 counts vs. 1), stronger reward (*stronger R*, 2 counts vs. 0) and unexpected reward (*unexpected R*, 2 counts vs. 1), while the feeling is more likely to be negative rather than positive in correspondence with stronger punishment (*stronger P*, 7 counts vs. 0) and unexpected punishment (*unexpected P*, 3 counts vs. 0).

Summary of results: in correspondence with conditions of reward most of the times the user's feeling is positive, while in correspondence with conditions of punishment the feeling is more likely to be negative.

4.5. iCat facial expressions

Two chi-square tests were carried out to investigate whether the generation of a facial expression by the iCat would influence the feeling or the level of engagement of the user. In these tests, the presence of a facial expression generated by the iCat was used as the influencing variable (two levels, N = 30) and the feeling and the engagement with the iCat (two levels) as the dependent variables for the first and second test respectively. It was hypothesised that the generation of a facial expression by the iCat would increase the level of engagement.

Results confirm the hypothesis: when the iCat generates a facial expression it is more likely that the user is engaged with it (86.7 %), while when no facial expression is displayed, most of the times the user is not engaged with the iCat (53.3 %) [$\chi^2 = 5.4$, $df = 1$, $p < 0.05$]. No significant association between the presence of a facial expression and the feeling was found. Two chi-square tests were also carried out to investigate whether the generation of a specific facial expression by the iCat would influence the feeling or the level of engagement, but no significant association was found.

Summary of results: when the iCat displays a facial expression during the game, the level of engagement of the user towards the iCat increases.

5. Conclusion

This paper addresses affect recognition as a prerequisite for a robot companion to sustain long-term interactions with users. In a naturalistic scenario where a robot plays chess with children, user states that are important for a robot to recognise are identified and the role of context is explored.

The main contribution of this paper consists of exploiting different levels of contextual information to model naturalistic user states that originate both from the task that the user is accomplishing (i.e., playing chess in this case) and the social interaction with the robot. Contextual information related to the game and the behaviour displayed by the iCat are analysed as possible indicators to be used to discriminate among the identified states. Results highlight a key role of game state, game evolution, event of capturing pieces and user sensations (as a result of the mismatch between their expectations and what really happened in the game) to discriminate between positive and negative feeling (i.e., the user's feeling tends to be more positive than negative when the user is winning or improving in the game, when a piece is captured by the user instead of by the iCat and in case of conditions of reward rather than punishment during the game). Game state and display of facial expressions by the iCat proved successful to predict the level of engagement with the iCat: higher levels of engagement with the iCat were found when the user is winning in the game and when the iCat displays facial expressions. These findings will provide the foundation for the design of an affect recognition system where context can be used with behavioural cues to predict the user's state towards the game and the robot.

6. Acknowledgements

This work was supported by the EU FP7 ICT-215554 project LIREC (LIving with Robots and intEractive Companions). The authors would like to thank Christopher Peters for providing useful comments.

References

[1] C. Bartneck, J. Reichenbach, and A. Breemen. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. In *Proceedings of the Design and Emotion Conference 2004*, pages 32–51, Ankara, Turkey, 2004.

[2] C. Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2):119–155, July 2003.

[3] G. Castellano, R. Aylett, A. Paiva, and P. W. McOwan. Affect recognition for interactive companions. In *Workshop on Affective Interaction in Natural Environments (AFFINE), ACM International Conference on Multimodal Interfaces (ICMI'08)*, Chania, Crete, Greece, 2008.

[4] C. Conati and H. Maclaren. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, 2009.

[5] K. Dautenhahn. Socially intelligent robots: Dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):679–704, 2007.

[6] E. Hannan, P. Krishnaiah, and M. Rao. *Handbook of Statistics 5: Time Series in the Time Domain*. Elsevier, 1985.

[7] A. Kapoor, W. Burleson, and R. W. Picard. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8):724–736, 2007.

[8] A. Kapoor and R. W. Picard. Multimodal affect recognition in learning environments. In *ACM International Conference on Multimedia*, pages 677–682, 2005.

[9] M. Kipp. Spatiotemporal coding in ANVIL. In E. L. R. A. (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may 2008.

[10] D. Kulic and E. A. Croft. Affective state estimation for human-robot interaction. *IEEE Transactions on Robotics*, 23(25):991–1000, 2007.

[11] I. Leite, C. Martinho, A. Pereira, and A. Paiva. iCat: An affective game buddy based on anticipatory mechanisms. In *AAMAS '08: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, pages 1229–1232, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems.

[12] I. Leite, A. Pereira, C. Martinho, and A. Paiva. Are emotional robots more fun to play with? In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*, pages 77–82, Aug. 2008.

[13] J. C. Lester, S. G. Towns, and P. J. FitzGerald. Achieving affective impact: Visual emotive communication in lifelike pedagogical agents. *The International Journal of Artificial Intelligence in Education*, 10:278–291, 1999.

[14] C. Martinho, I. Machado, and A. Paiva. A cognitive approach to affective user modeling. *Affective interactions: Towards a new generation of computer interfaces*, pages 64–75, 2000.

[15] C. Martinho and A. Paiva. Using anticipation to create believable behaviour. In *American Association for Artificial Intelligence Technical Conference*, pages 1–6, Boston, july 2006.

[16] M. P. Michalowski, S. Sabanovic, and R. Simmons. A spatial model of engagement for a social robot. In *Proceedings of the 9th IEEE International Workshop on Advanced Motion Control*, 2006.

[17] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, July 1988.

[18] C. Peters, S. Asteriadis, K. Karpouzis, and E. de Sevin. Towards a real-time gaze-based shared attention for a virtual agent. In *Workshop on Affective Interaction in Natural Environments (AFFINE), ACM International Conference on Multimodal Interfaces (ICMI'08)*, Chania, Crete, Greece, 2008.

[19] I. Poggi. *Mind, hands, face and body. A goal and belief view of multimodal communication*. Weidler, Berlin, 2007.

[20] B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn. Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction Studies*, 5(2):161–198, 2004.

[21] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 1980.

[22] C. L. Sidner, C. D. Kidd, C. H. Lee, and N. B. Lesh. Where to look: A study of human-robot engagement. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, pages 78–84, Funchal, Madeira, Portugal, 2004. ACM, New York, NY, USA.

[23] A. van Breemen, X. Yan, and B. Meerbeek. iCat: An animated user-interface robot with personality. In *AAMAS '05: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 143–144, New York, NY, USA, 2005. ACM.

[24] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), January 2009.