In this chapter we first examine the requirements for social game-play robots in three game scenario types: robot play companions, robots and digitised games, robots and augmented reality. We consider issues relating to affect recognition, affective modelling in the robot, and robot expressive behaviour. We then discuss work in each of the three scenario types and how it has attempted to meet the requirements advanced. Finally the chapter considers key research issues for the future.
Chapter 17
Games Robots Play: Once More, with Feeling

Ruth Aylett

Abstract In this chapter we first examine the requirements for social game-play robots in three game scenario types: robot play companions, robots and digitised games, robots and augmented reality. We consider issues relating to affect recognition, affective modelling in the robot, and robot expressive behaviour. We then discuss work in each of the three scenario types and how it has attempted to meet the requirements advanced. Finally the chapter considers key research issues for the future.

Requirements for Social Game-Playing Robots

Discussion of robots in the context of digital gaming may seem a little paradoxical. After all, the major characteristic of robots is that they are part of the real physical world and not the virtual digital world. So how can they be involved in digital games? Indeed, RoboCup football is the most obvious intersection between robots and games. However this is a specialist area and as it involves robot-robot collaboration rather than human-robot interaction, it is not the topic of this chapter.

Three areas of application come to mind. The first and earliest extends the definition of game into play, and involves the creation of robot play-companions for children. The second draws on new collaborative display technologies such as multi-touch surfaces. Here, a robot acts in the real world in the roles a human might otherwise occupy, from fellow-player in digitally-supported board games [33], to intelligent tutor in serious games with educational purposes [34]. Finally, if the purely digital is moved into the real world via augmented reality approaches, then one can consider a robot actor that becomes part of an overall game experience containing both real and virtual elements.

All of these roles put the robot firmly into the relatively new research domain of social robotics [13]. A social robot can be defined as: “a physical entity embodied in a complex, dynamic, and social environment sufficiently empowered to behave in a manner conducive to its own goals and those of its community” [9]. Clearly
affect is a fundamental human characteristic in social situations, with a particular impact on non-verbal behaviour. In the human case, the three areas just mentioned could involve for example a human player gloating about the success of their move in a collaborative board game, a child ‘telling off’ a robot doll for bad behaviour, a human player in an augmented reality scenario expressing encouragement to a robot actor in a shared scenario.

It is considering these concrete scenario types that allows us to derive some generic requirements for game-playing robots. These relate to capturing user affective state, modelling the state of the user on the robot side, and robot expressive behaviour.

All three scenario types suggest that a social game-playing robot requires information about the affective state of the humans with whom they are interacting. Without this information it is very hard to respond appropriately, impacting the success of the robot’s action-selection mechanism. Detecting the affective state of a human interaction partner can be viewed as a difficult problem in social signal processing (as discussed in part II of this book), and remains work in progress.

The current state-of-the-art [42] is based on multi-modal approaches that try to fuse data from facial expression, voice, and sometimes, physiological data such as skin conductivity. However levels of accuracy on a small set of affective states (four or five) remain well below the 95% needed and this raises problems for social game-playing robot applications for which specific solutions must then be engineered.

There are theoretical problems as well as signal processing problems in assessing the affective state of a user. The term state may itself be one of them, given that human interaction partners move through a dynamic affective process in which affect is typically not very strong and more likely to be irritation than anger, boredom than fear, and amusement rather than outright happiness. A second well-known problem is that mapping between the behaviour captured by the robot’s sensors and the affective state of the user is not straightforward. A smile is a relatively easy expression to capture with modern sensors, but smiles are one of the most ambiguous facial expressions [21], able to express embarrassment, sorrow, anger and social welcome or agreement, among other things, as well as happiness.

Fortunately, a game-playing robot has a big advantage over a robot in less structured social situations. This is because the context for interaction is a game, with known rules and moves. As we will see later in the chess-playing companion robot, this helps a great deal because it becomes possible to infer affective state with a high likelihood of accuracy. In general, a combination of inference and sensor processing seems a promising approach, especially given the evidence for the human use of expectations in processing interaction data [38].

As well as trying to capture the affect of interaction partners, there are strong arguments for a game-playing robot itself attempting to impact the user’s affective state. This may be for dramatic or pedagogical reasons, but it may also be a way of simplifying the sensor-processing problem just discussed. If a robot performs a game-action that is intended to please a human player, and the human then smiles, this makes happiness the most likely interpretation of the smile.
Both inferring user affect from situation and changing user affect via actions require a robot that has a model of affect \([1]\) such that it is able to carry out some aspects of Theory of Mind (ToM) processing \([7]\) about the situation of the human interaction partner. This need not be an explicit model – it could be embedded in a set of behaviour rules or stimulus-action pairs. It could be as specific as a rule that takes the metric of goodness of a chess move and maps that onto the likely happiness of the person that made the move. However there are advantages in an explicit model where the game-playing robot needs to use natural language outputs as part of its action repertoire and where the robot has a more flexible set of roles and functions rather than specialising in one thing, like playing chess. Cognitive appraisal models \([31]\) are useful in this context, especially because running in one direction they can generate robot expressive behaviour, while running in the other, they offer a simulation view of ToM affect issues (e.g. ‘What does my model say I would feel if I did that?’) \([7, 30]\).

There are a number of reasons for wanting the robot to be able to generate contextually appropriate robot expressive behaviour. One is that communicating the robot’s affective state is one way of impacting that of its interaction partner. Thus robot expressive behaviour can be perceived by human players as a commentary on the game-actions of the robot, but also as comment on the game-actions of human players. However there is also a more generic reason for producing expressive behaviour.

Each of the three scenario types above supports a set of game-based social roles both for the robot and its human interaction partners, and each role has an associated set of expectations about behaviour and goals. A robot can signal adherence to its expected role by creating transparency about its current goals and by advertising its coming behaviours before it executes them. One could do this verbally, but in smooth human-human interaction, this is far from the only or even the dominant modality. Taking an intentional stance \([6]\) means that in human-human interaction the participants are continuously monitoring each other’s goals and updating their expectations about behaviour, largely through the use of non-verbal behaviours such as gaze, facial expression, posture, gesture, non-verbal vocalisations \([29]\).

While the generation of expressive behaviour has been extensively researched for graphical characters \([32]\), and attempts have been made to create standard interfaces supporting expressive mark-up languages \([19, 40]\), social robots raise a set of new issues, primarily due to the very variable physical embodiments they may involve. Indeed, issues relating to embodiment are fundamental throughout HRI research. For a reasonably humanoid robot like the Nao, it may be feasible to apply similar methods as for graphical characters \([22]\). However robots need not be humanoid – they could be more animal in form, as in the case of the Kismet \([3]\) or Paro \([16]\) robots, or very machine-like in design. There is no absolute requirement that they have facial features (the Nao has almost none) or indeed anything like a face; limbs with which to gesture; or an articulated body supporting posture. If issues of naturalism are moot with graphical humanoid characters, they are much more so in general with robots.
While we have argued that games help with some of the problems of affective processing in social robots, there is at least one area in which they may make these problems more difficult. Games provide a rare example of a social situation in which participants may be entitled to deceive fellow players about their goals and likely behaviour [7]. The card game Poker is an obvious example, but there are many others: board games such as Risk where secret alliances between players may be formed, and Mafia/Werewolf, where the point is to dissimulate if you are a Mafia member/Werewolf [7]. In these cases, generating robot expressive behaviour requires a more indirect mapping between what the affective model says and what the robot expresses. ToM processing is needed for the robot to assess the likely impact of its choice of deceiving behaviour. Assessing the expressive behaviour of human players also becomes a much more difficult problem.

In the rest of this chapter we consider recent work in the three game types just outlined, focusing on the interplay between embodiment and expressive behaviour, and to its relationship to empathic engagement between robot and human interaction partners. We conclude with a look at the most significant continuing research issues.

Robots as Play Companions

The introduction of robots as play companions has come both through the evolution of actual toys in the direction of robot functionality and through the development of research prototypes.

Graphical artefacts such as the Tamagotchi [8] of the later 1990s established that human-like or even sophisticated embodiment is not required to evoke attachment to a play character. The Tamagotchi was a small and very low-resolution graphical artefact on a small plastic key-fob, loosely based on a chicken. Its behaviour appeared autonomous; the human was responsible for feeding it virtual food and giving it caring behaviour through frequent interactions to which it would respond. If neglected or ‘old’ it would ‘die’. Its small behaviour repertoire was wholly expressive of its inner state since there was no functional content, no ‘task’ that a Tamagotchi had to carry out. These behaviours were enough to create a caring, empathic relationship on the human side and there were anecdotal accounts of children mourning when their Tamagotchi ‘died’, raising ethical issues that had not been addressed in the design [37].

This was followed in 1999 by the Sony Aibo [14], marketed as a robot pet/adult toy and embodied as a mechanical-looking dog. A design criterion was ‘lifelikeness’: this could have been trivially addressed by making the appearance more naturalistic, for example by providing a furry exterior. However instead this was interpreted as a behavioural requirement, that of ‘maximising the complexity of responses and movements of the robot [14]. This did have embodiment implications. Complex movement requires a higher number of degrees of freedom so that different body parts can be moved in different combinations. Complex responses also require sufficient sensor input. In addition, designers saw the need for multiple motivations
for movement, and mechanisms for producing non-repetitive behaviour so that the
Aibo did not become predictable like a normal machine.

Again, almost all Aibo behaviour was expressive, driven by four ‘instincts’: affection, investigation, exercise, and appetite and expressing the six Ekman ‘basic emotions’: joy, sadness, anger, fear, surprise and disgust. Emotions allowed different
to be produced in the same external situation because they reflected the
impact of earlier interactions and thus altered the context. Thus a happy Aibo would
produce a paw extension behaviour when its sensors detected a hand in front of it,
while an angry Aibo would not. This underlines the close relationship between an
architectural model of emotion in a robot and the generation of a rich behavioural
repertoire.

The Aibo was a relatively successful consumer product by the standard of social
robots, though never on the scale of Sony’s other range of devices, which may
explain why it was eventually discontinued. The scale and type of engagement
with Aibos was researched by a number of groups, and in [15], analysis of more
than 2000 forum postings by Aibo owners suggested a high degree of affective
engagement.

Other robot toys have followed, with the Pleo, a small and relatively cheap
dinosaur, having some commercial success. However analysis of Pleo in the home
brought out quite sharply the current limitations of robots as free-play toys. After
the novelty effect had worn off, the Pleo’s limitations led to diminishing use over
[11]. These limitations included straightforward issues like short battery life.
The Pleo could not act as a toy when it was being recharged because its battery had
to be extracted from its belly and put into a charger. The Aibo had a much more
sophisticated autonomous charging system in which the robot would move itself
into the charger, but this reflected its much higher cost. Even then, it too could not
act as a plaything while being charged.

While this is an issue of embodiment rather than robot game-behaviour, it
also seems likely that such robot toys still have too little flexibility for a child’s
open-ended play as well as too little functionality to be able to drive the play
interaction themselves for more than a limited period. These limitations motivated
the development of the iCat Chess Companion developed in the EU project LIREC
between 2008 and 2012.

The iCat [4], designed as a research tool by the Philips Eindhoven laboratory, was
what is known as an interface robot, designed to rest on a desk and communicate
rather than to move around. It took the form of a small yellow plastic cat with
programmable head movement, eyes, mouth and eyebrows. The cat had a schematic
body but no legs (see Fig. 17.1). It was capable of flexible facial expressions, glance
and head movement as well as speech output. Only a limited number were produced
and it was never commercially available. The design is not naturalistic and raises no
expectations of real cat behaviour.

An iCat was developed into a chess opponent for children in a Portuguese chess
club. An advantage of this application is that the role of a chess opponent is well
understood and also quite limited. The game is turn-based. Functionally, such a
robot needs to be able to play chess, an easy requirement to meet given many
mature chess engines are already available. To avoid the more difficult physical functionality of moving the pieces, the child player was asked to move the piece for the iCat.

However though being able to play chess is the starting point for a chess opponent, as a child’s leisure activity it also has a social dimension, with comment about the moves and how each player is doing. Thus it is a tractable environment in which to explore the impact of robot expressive behaviour and the development of an empathic engagement, in which the child player is willing to assign to the robot an ability to understand how they, the child, feel.

The iCat first used the game state reported by the chess engine to adopt a sad or happy expression depending on whether its game position was strong or weak. A study [24] showed that this allowed the child player to more accurately assess the game state than in a version without the happy and sad expressions. Next, an affect detection system was developed so as to assess the engagement of the child player so that the iCat could respond empathically, that is, with actions taking the child’s affect into account.

The difficulty of generalised affect detection is not the only reason for developing a scenario-dependant system, as was done in this case. The affective states of interest are themselves scenario dependant [5], concerned both with the content of the interaction (playing chess in this case) and with other interaction parameters such as whether the user is standing or sitting, moving or stationary, one or many, facing or not the social robot. Many affect systems have been developed using actors; in this case the aim was to use in-the-wild data of children actually playing chess.
A person-independent Bayesian learning system was developed using a video corpus of children playing chess, but knowledge from the task and the social context was also included and improved the recognition rates [5]. Smiles, gaze at the robot, game state and game evolution were used to generate probability values for the child’s positive, neutral and negative valence at each turn. The iCat would generate an expected next move for the user, and then compare the actual move to this expectation. In the style of cognitive appraisal, this comparison was used to generate an iCat emotion and an accompanying expressive behaviour for it. There were nine of these depending on whether the next move was better or worse than expected [5]: Excited, Confirm, Happy, Arrogant, Think, Shocked, Apologise, Angry, Scared. Some of these – Happy, Angry, Scared – correspond to straightforward emotions, and others – Confirm, Think, Apologise – do not. This underlines the point that affect for a social robot is frequently based around quite complex affective states. Evaluation showed that the child players noticed the empathic elements of the iCat interaction, and had a more positive attitude to it as a result [23]. An interesting aspect was that though the iCat had no speech recognition capabilities, the child players did not remark upon this and it did not seem to impact the interaction.

**Robots and Digitised Games**

The second area involving robots and digital games identified above generalises the iCat example into a robot player of digital board games. This exploits the growing availability and popularity of large multi-touch surfaces that allow a board game to be ported into a digital form. A characteristic of such games is that they are often multi-player so that social interaction is correspondingly more substantial and more complex than in a two-player game like chess.

Given that full natural language interaction using speech is well beyond the state of the art and one cannot expect a social robot to sustain a conversation, this may seem likely to worsen the problems already discussed. However one should bear in mind that in multiple player games, human-human interaction becomes more important. With two human players and a robot, human-human interaction occupies a third of the possible interactions space. With three human players and a robot it is half of the possible interactions space. Thus the robot is no longer wholly responsible for the social experience.

Moreover multi-person conversation in a game context is often fairly unstructured, with stereotypical speech actions like announcing whose turn it is, and repartee, rather than sustained conversation, around the game state.

Early work on a social robot for a multi-player game took place in the LIREC project [27] in which an EMYS robot head (see Fig. 17.2) was incorporated into a multi-touch surface version of a game called Risk [33]. This is a popular board game for three to six players in which players run armies and try to conquer territories. Successful play involves the creation of informal alliances and joint attacks as well as deception and betrayal.
Believable verbal and non-verbal behaviours were seen as key requirements and a robotic embodiment as a way of supporting non-verbal behaviours. As in this article, the ability to model affective state was specified as a requirement, as well as the ability to simulate social roles commonly found in board games. These roles included Helper, a role already discussed for the iCat chess companion, but also Dominator, in which the agent tries to influence a human player to take a specific action. A role not so far discussed was Exhibitionist – behaviour intended to grab attention. Finally, a role specific to a multi-player game was Negotiator – mediating between two other players.

A requirement also present in the iCat chess companion and clearly very significant for any social robot in or out of a game was that of being able to recognise the human players, with an associated ‘greet’ behaviour, and to remember interactions with them over time [33].

Again, this is in principle a more difficult requirement in a multi-player game than in a single-player one, where the robot has no choice of player with whom to interact. However in a turn-based board game, players normally play in a fixed order and sit in a stable configuration round the board, so sophisticated facial recognition is not needed. The game context comes to the rescue once more. In addition, a microphone array can determine which player is speaking, though in a multi-player game overlapping utterances and interruptions are very common.

This allows a gaze behaviour to be developed so that the game-playing robot can ‘look’ at the player with whom they are interacting. As the iCat chess companion showed, this is a very powerful social behaviour and is probably one of the factors supporting acceptable social interaction without speech recognition. In a board game, players also often look at active areas of the board, and on a multi-touch surface it is relatively easy for a robot to determine which these are given the use of touch to make a move.
Fortunately it is a great deal easier to generate speech utterances for a social game-playing robot than it would be to recognise the speech of the other players. The game itself can be used to determine when the robot should speak, and in the work described here, a *relevance* value was used to select significant game events and prevent the robot tediously commenting on everything. Highly relevant events included game attacks on the robot’s forces, or conquests by a player the robot disliked. Like and dislike were established via alliances and attacks. A further source of utterances was the die roll, where ‘good’ or ‘bad’ values often produce human utterances about luck.

This work has been extended into a serious game environment in the EMOTE project [10], in which the sustainable energy game Enercities [17] was converted from a single-user web-based system to a multi-player touch-table based game [34]. The robot is a torso-only version of the Aldebaran Nao (Fig. 17.3).

Where the robot player of Risk was simple a player in a similar role to the human players, in Enercities, the robot player is expected in addition to play a tutorial role. This underlines that a Robot Digital game player may sometimes be more than just a player, and its model of the game may need to be more sophisticated than telling it which is the best move in relation to the game rules. In the Enercities case, there may be cases where the robot should play a non-optimal move from the point of view of game score if this produces good teaching for the human players. The social roles involved are now different from those of the peer-player robot of Risk: for example an Exhibitionist role becomes wholly inappropriate, while the Negotiator role is folded into a pedagogical approach in which the trade-off between possible player-choices are one of the learning points. Careful analysis of social roles is needed as much as game rule competence for a successful digital game-playing robot.
Robots and Augmented Reality Game Experiences

The vision of Pervasive Gaming is to take digital games out of the computer-based virtual world and into the real physical world using an augmented reality approach [28]. This has so far largely involved the use of handheld devices, but the incorporation of actor-robots seems an interesting and logical step to contemplate. Most initial work in this field has focused on theatre [12, 18, 26], and there are good reasons why this should be so.

A theatrical stage is a much more controlled environment than a pervasive game, which could be room-based but often takes place in the normal physical world. The stage is a privileged environment to which extra resources can be added, for example to give robot actors other information than that directly available to their sensors. In addition, theatrical drama is usually pre-scripted, and while this does not remove the need for real-time responsiveness, interacting with human actors following a script is much less difficult than interacting with improvising pervasive game players. In particular, the problem of speech-based natural language interaction is finessed by knowing in advance what utterances fellow actors will make.

However it is also true that a theatrical performance is required to engage a large audience of spectators, where a pervasive game involves a much smaller number of participants. Participants must themselves take timely actions and are therefore more able to cope with robot actor errors, as long as they are not too frequent and substantial. As we saw in the case of the iCat and digital board game robots, as long as relevant robot utterances are combined with appropriate expressive behavior, especially glance, participants may overlook the absence of natural language interaction. However, spectators are likely to notice robot actor errors very quickly and these are more likely to impact their enjoyment. For this reason, much theatrical robot work involves a greater or lesser degree of tele-operation, especially where a real-world performance is involved [41].

One approach that is argued can help humans to accommodate robot actor errors is clearly signaling the behavior it is about to carry out and then expressing a reaction to its success or failure, much as animators do with graphical characters [18]. Robot actors make a good testbed for exploring this approach since it is common in human drama too, and is an example of how integrating robots into game environments may have a useful spin-off for more general social interaction.

A profound difference between this type of scenario and the ones so far considered is that the robot must navigate. Robot toys may move but they do not have to navigate, and they could, equally, be moved by the child. Robot board-game players should not move away from the board. Robot actors must move according to the logic of the dramatic situation, whether to specific locations, or in order to interact with specific human players. It is possible to design a drama or indeed a game around the known limitations of the robot, but the experience a 2014 performance in Pittsburgh involving the HERB robot revealed that precisely because the robot had wheels and could move, the audience expected it to do so in performance [41].
A novel approach to dealing with robot movement constraints explored *migration* — a transfer of the robot ‘character’ into a graphical embodiment and back [20]. This was applied to a treasure hunt game in which an Emys robot (as in Fig. 17.4) was transferred in the form of a small animated version to a handheld device the accompanied the player as they collected clues [2]. This has not been tried in theatre applications, where it seems inappropriate for a static audience, but could be a promising approach for pervasive games.

This work demonstrated that although a robot has greater physical presence than a graphical character [36], a common appearance and voice are strong enough cues for users to assume a continuity of personality between very different embodiments. A study in which one version of the migration allowed the character to refer to memories across embodiments while another version had the character forget interactions from earlier embodiments, showed that this had no effect at all on the user perception of it being ‘the same’ character [2].

**Key Research Issues**

In this article we have looked at three possible ways in which robots can be involved in digital games: as a robot toy, as a board game play companion, and as an actor in pervasive games. Research work has already been carried out in the first two scenario types, though much of it is still some way from real-world application. The third scenario-type is still very much in the area of future research.
A common theme in all three is the difficulty in knowing the state of the human interaction partners involved. The level of difficulty is dependent on the exact application, with the iCat chess companion a good example of an application in which it is less difficult and a pervasive game one in which it is more difficult. Social signal recognition is related to basic configuration issues — whether the user is sitting, moving around, one or many; to the specific signals to be detected — easier for some facial expressions such as smiles, harder for many others; and above all to the social context. This context encompasses the overall social framework, which includes the rules and gameplay in the case of games, the specific interaction taking place in its game context, and the social roles involved.

A continuing research challenge is finding applications in which these factors allow successful incorporation of the game-playing robot without requiring functionality too far away from the current state-of-the-art.

As with other robotic applications, running an application over the long-term rather than in short lab-based experiments is also a challenge. This is partly related to obvious constraints like battery life and the need for recharging, as discussed earlier. Running complex robot software continuously over periods of weeks is a further technical challenge. However there are more interesting social issues. Robots have a pronounced novelty effect because most people rarely interact with them and have often never come across a real-world example. This produces an atypical social interaction over the short-term. Continuing social interaction beyond the novelty effect is a research issue that is only slowly being confronted.

Even a relatively successful game-playing robot such as the iCat chess companion was less successful as time passed and the novelty effect faded. Child players looked at it less, smiled at it less, and commented upon its behaviour less as time went on [23]. It is likely that this was related to the range and variety of its social behaviour and that over time it became too predictable. Thus developing the social repertoire of a game-playing robot requires careful thought if it is to succeed over the longer term.

This is a difficult design problem but also a difficult evaluation problem. Iterating and reiterating a design that may take several weeks to evaluate requires a substantial commitment of research resource as well as access to a large number of willing participants over long periods.

We finish by reiterating the generic requirements discussed above. Expressive behaviour underpinned by affective processing is at least as crucial to functional and engaging game-playing robots as an ability to play good moves.

Acknowledgements Some of the work discussed in this paper was partially funded by the EU FP7 ICT-215554 project LIREC (Living with Robots and Interactive Companions) and the EU FP7 ICT-317923 project EMOTE. The author is solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein. The author wishes to acknowledge the partners of both projects and especially the GAIPS group at INESC-ID.

References
AUTHOR QUERY

AQ1. Refs. [25, 35, 39] are not cited in text. Please provide in-text citation or should we delete from the reference list.