

## A Social Robot as a Card Game Player

Filipa Correia,<sup>1</sup> Patrícia Alves-Oliveira,<sup>2</sup> Tiago Ribeiro,<sup>1</sup> Francisco S. Melo,<sup>1</sup> Ana Paiva<sup>1</sup>

<sup>1</sup>INESC-ID, Lisbon, Portugal and Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>2</sup>Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal; CIS-IUL, Lisbon, Portugal; and INESC-ID, Lisbon, Portugal

### Abstract

This paper describes a social robotic game player that is able to successfully play a team card game called *Sueca*. The question we will address in this paper is: how can we build a social robot player that is able to balance its ability to play the card game with natural and social behaviours towards its partner and its opponents. The first challenge we faced concerned the development of a competent artificial player for a hidden information game, whose time constraint is the average human decision time. To accomplish this requirement, the Perfect Information Monte-Carlo (PIMC) algorithm was used. Further, we have performed an analysis of this algorithm's possible parametrizations for games trees that cannot be fully explored in a reasonable amount of time with a MinMax search. Additionally, given the nature of the *Sueca* game, such robotic player must master the social interactions both as a partner and as an opponent. To do that, an emotional agent framework (FAtiMA) was used to build the emotional and social behaviours of the robot. At each moment, the robot not only plays competitively but also appraises the situation and responds emotionally in a natural manner. To test the approach, we conducted a user study and compared the levels of trust participants attributed to the robots and to human partners. Results have shown that the robot team exhibited a winning rate of 60%. Concerning the social aspects, the results also showed that human players increased their trust in the robot as their game partners (similar to the way to the trust levels change towards human partners).

As interactive entertainment expands, computer games are progressively moving from the virtual world back to the physical world. Augmented reality games, haptic interfaces in gaming, touch tables, etc, are some of the types of interactivity placing human players in physically situated entertainment experiences. In parallel with this move into the physical world, artificial partners and opponents can also be created to exist in such physical world. To do that, the area of entertainment robots offers challenging opportunities as it explores the role of a robot as a game player. In general, social robots can contribute with new and broad ways of creating socially engaging interactions with humans in entertainment contexts. The challenges of these human-robot interactions may vary from game to game. Some games, when played in the physical world not only hold complex social

behaviours, but they also hinder performance aspects related with the competitive nature of the game itself.

Furthermore, in many types of games, current advances in Artificial Intelligence (AI) over the past few years has shown that strong and powerful algorithms combined with significant amounts of data, are able to defeat human world champions of these games (see for example, the game of Go). These results raise our expectations and people are starting to consider such artificial agents as fierce competitors. Yet, when we consider multi-player games, where the social environment becomes more relevant, and when the games are played in the physical world, how will people perceive a social robotic player compared to human standards? Will people be willing to trust a social robot to be his partner in a team game?

To address these questions we created a social autonomous robotic partner for the *Sueca* card game, with a twofold goal of both playing competitively and interacting socially with the other players. The development of such robotic game player introduces some challenging aspects, in particular finding the best balance between social responses and computations related with the game, in order for the socially intelligent agent to produce natural and human-like behaviours.

Another important challenge of creating an intelligent agent in this social context is the time constraint on the computation of a hidden information card game. State-of-the-art approaches, for instance PIMC, promise good results on the *Sueca* domain according to the game properties. However, the full computation of multiple perfect information games is not time-efficient and will hinder natural interaction in a game with human players. Therefore, this paper also explores how the algorithm's parametrizations affect the game results in order to choose the best performance-time configuration.

Finally, by using an expressive robot that is able to express emotions, provide spoken feedback, and respond socially, the game experience can be created balancing these social and game competencies. In this case, we built the social competencies by using an emotional agent framework (FAtiMA) which allows for emotional appraisal to occur and fire social and emotional behaviours. At each moment, the robot not only plays competitively, but also appraises the situation and responds emotionally to the game situations.

To evaluate such robotic game player, we have considered two central aspects for assessing when playing in the card game: (1) competitiveness between team opponents and (2) cooperation between team partners. Performance can be measured by the number of games won, and simulations made with the developed algorithm to test it against other artificial players. But, most importantly, to assess how humans perceive the robot as a game player we conducted a user study with 60 participants and measured the human players' trust levels in their partners (before and after playing). Concerning the competence of the robot, the results show that the robot is able to play competitively with human players, achieving a winning rate of 60%. We also compared the trust level on the robotic partner with the trust level on human partners. The results show that human players significantly increased their trust in the robot as their game partners (in a similar way to the trust levels they perceive to have towards human partners).

## Background

PIMC is a search algorithm suitable for partially observable environments. It uses a Monte-Carlo methodology and deals with the imperfect information by the determinization technique. In other words, the hidden information is sampled several times and the best move is computed by solving exactly or heuristically in each perfect information game.

The main two disadvantages of this approach were pointed by Frank and Basin in 1998 (Frank and Basin 1998), and were called strategy fusion and non-locality. The first one refers to decisions that would only make sense for certain sampled distributions and when applied turned out to be poor decisions. The non-locality results from the fact that the value of a game tree node only considers its children values, however, in an imperfect information game, some guesses might be done using values of the non-local sub-tree.

Nonetheless, this algorithm already counts with some successful implementations as the Ginsberg's Intelligent Bridge player (GIB) (Ginsberg 2001), Skat (Buro et al. 2009), or Hearts (Sturtevant 2008). Their remarkable results seem to reveal the effectiveness of determinization. However, there were still difficulties in understanding the strong results of this algorithm. As such, Long et al. (Long et al. 2010) not only have analysed the previously mentioned problems of PIMC search, but the authors have also shown how three different properties of a game can influence the success of PIMC. The first property is *leaf correlation*, which refers to how likely it is to affect a the payoff of a player in the neighbourhood of a leaf. When the probability of all siblings having the same payoff values is higher, the correlation value increases. Secondly, *bias* indicates the chance of a player being preferred over another. Finally, the last game characteristic that has been pointed is *disambiguation factor*, that denotes how rapidly the hidden information is revealed. The authors analysed these properties for the game of Skat, indicating a considerably good performance of PIMC, due to its values of *leaf correlation*, *bias*, and *disambiguation factor* (Long et al. 2010). This analysis strongly led us to consider PIMC in our first implementation for the *Sueca* domain, and

to expect a reasonable performance, considering the obvious similarities between these two trick-taking card games.

## Related Work

Robots are being developed and introduced in social environments with humans as tools for assistive (Feil-Seifer and Mataric 2005), educational (Castellano et al. 2013) and even entertainment purposes (Pereira, Prada, and Paiva 2012). In the scope of this work, we will focus on entertaining activities - namely playing the card game of *Sueca* - with an artificial robotic companion.

The gaming experience with an artificial player is an engaging experience due to the agent's sociability and embodiment (Behrooz, Rich, and Sidner 2014), either virtual or physical. Therefore, game playing scenarios became more popular to analyse the influence of physically embodied agents on different aspects that might be related with the game experience or with the perception of the artificial player. For instance, physical embodiment can provide a more immersive user experience, an improved game feedback and a more believable social interaction (Pereira et al. 2008). Additionally, other studies revealed that empathic behaviour can positively affect how children perceive a robot in a game playing scenario (Leite et al. 2012).

Henceforth, different playful scenarios have been developed, in which robots and humans interact and play together. Sometimes humans can behave as opponents towards each other and towards the robot, like in the Risk game scenario (Pereira, Prada, and Paiva 2012). In this game, the robot behaves in the game as a socially aware agent that can play and socially interact with the human players. Its social behaviour is driven by its ability to have memory of the game actions of other players, be emotionally aware and consider social roles in the game (e.g., as Dominant or Exhibitionist towards the other players). In a more recent scenario, a robot played the Mastermind game with elderly people as a mean to provide social support in a form of entertainment (Johnson et al. 2016). Although participants have indeed recognised the robot to have an entertaining associated value, the hypothesis that participants would enjoy more a robot displaying behavioural patterns associated with the game progress when compared to randomly displayed behaviours was not supported.

Overall, the mentioned research studies prove the broad application of entertainment robotics among different age groups, at the same time they constitute insightful considerations for the design and development of any social entertainment robot. We aim to extend this line of research, by developing a scenario in which a robot plays either as an opponent and as a partner in the traditional card game of *Sueca*. This card game is a well-known game in Portugal and (young) adults and elderly play usually the game in social contexts of interaction, providing joyful moments together.

## The *Sueca* card game

*Sueca* is a Portuguese trick-taking card game played by four players divided into two teams. It uses a French deck with the cards 2-6, Q, J, K, 7 and A valuing 0, 2, 3, 4, 10 and

11 points respectively, and summing 120 points across all suits. The rules are quite similar to any other trick-taking card game: if possible, players need to follow the suit of the first played card in the trick (lead suit); and the team players win the trick points if one of their cards has the highest value belonging to the trump suit or the lead suit. At the end, a score between 61-90 points accumulates one victory for that team, a score between 91-119 points accumulates two victories for that team, and a score of 120 points accumulates four victories for that team. When a player does not follow a suit lead suit while holding it in his hand, the game ends and the opponent team accumulates four victories. These victory points determine the winning team in a session of  $n$  games.

*Sueca* is a non-deterministic game, since it includes what is called the element of chance by the cards being randomly dealt at the beginning. Additionally, as the cards of each player are hidden from the other players, it is considered an imperfect information game.

The general winning strategy is to play the highest card of a non-trump suit when other players probably still have that suit and, therefore, cannot cut the trick with a trump. This situation typically occurs in the first tricks for each different lead suit. The trump cards are usually saved for cutting tricks.

### The Social Robotic player

Our social robotic player is responsible for prescribing natural and human-like behaviours based on both social responses and computations related with the game. The ultimate socially intelligent decision maker includes two modules: the *Game Module* – that holds our proposed task AI and is responsible for choosing, at each play, among the set of possible moves, while also updating the agent’s beliefs from the game events–; and the *Social Module* – that holds the “social AI” of the robot, and is responsible for selecting utterances of verbal and nonverbal behaviours according to both game events and the beliefs from the *Game Module*.

### The Game Module

During the development of the proposed AI for the *Sueca* card game, we went through three main stages: (1) creating a benchmark player based on rules for further evaluation; (2) applying PIMC algorithm in our domain; (3) enhancing the previous approach considering our requirements.

### The Rule-based player

The Rule-based player roughly replicates the general game-play strategy of non-professional human players. Its rule-based procedure starts by identifying the highest cards of each allowed suit for the current play; it then assesses whether or not to play each card by checking (i) if it is the highest unplayed card of that suit; and (ii) if there are less than 5 cards from that same suit in its hand, except for the trump suit. The condition (i) ensures no other player has a higher card of the same suit, while condition (ii) helps to measure the probability of another player cutting the trick. If both conditions are true, it plays one of its highest cards, otherwise it plays a random zero-value card.

We evaluated the performance of the rule-based player by running two experiments involving rule-based players and random<sup>1</sup> players. The first experiment consisted in 1,000 independent games involving a team of *one* rule-based player and *one* random player against a team of *two* random players. The second experiment consisted in 1,000 independent games involving a team of *two* rule-based players against a team of *two* random players. In the first experiment, the team with the rule-based player attained a winning rate of 56.6%, while in the second simulation, the team with two rule-based players attained a winning rate of 61.1%.

These results illustrate two important aspects. First of all, they highlight the cooperative nature of the game (by comparing results of the first and second experiments), which also matches the practical experience of human players: a good player matched with a bad teammate may not be able to properly leverage its proficiency during play. A second important aspect is that the rules of the game will often impose severe restrictions on the cards that a player may play. In those circumstances, the random player will essentially be as good as the optimal player, since no decision is actually taking place. This means that the card distribution *among the two teams* has a significant impact in the outcome of the game—which explains the surprisingly high winning rate of the random players.

Finally, it is worth noting that, as will soon as it will become apparent in Section “Evaluation and Discussion”, the performance of the rule-based agent against our player is similar to the performance of human players against our player, suggesting that the rule-based player is able to properly capture the fundamental game-play of regular human players and, as such, provides an adequate baseline for comparison.

### PIMC in the *Sueca* domain

The decision of using the PIMC algorithm in the *Sueca* domain relates to our main requirements of time-efficient online computation and also to the previous results in the Skat domain (Long et al. 2010), which is a similar trick-taking card game. Its pseudo-code is described in Algorithm 1.

---

#### Algorithm 1 PIMC search pseudo-code.

---

```

1: procedure PIMC(InfoSet  $I$ , int  $N$ )
2:   for all  $m \in \text{Moves}(I)$  do
3:      $val[m] = 0$ 
4:   for all  $i \in \{1..N\}$  do
5:      $x = \text{Sample}(I)$ 
6:     for all  $m \in \text{Moves}(I)$  do
7:        $val[m] += \text{PerfectInfoValue}(x, m)$ 
8:   return  $\underset{m}{\text{argmax}}\{val[m]\}$ 

```

---

We also illustrate in Figure 1 how PIMC can be applied in this card game. The example shows an 8<sup>th</sup> trick of the game, where each player still holds two cards in their hands. The first step is to sample card distributions or configurations for other players’ unknown hands. Then, it calculates

<sup>1</sup>Random players only played allowed moves

the reward of playing each card in its own hand for every sampled distribution. These two steps are repeated  $N$  times. The chosen card to play is, therefore, the one with the maximum accumulated reward for all the sampled distributions. This algorithm contains two important steps referred in Figure 1 as *Sample* and *Search*, which are discussed below in greater detail.

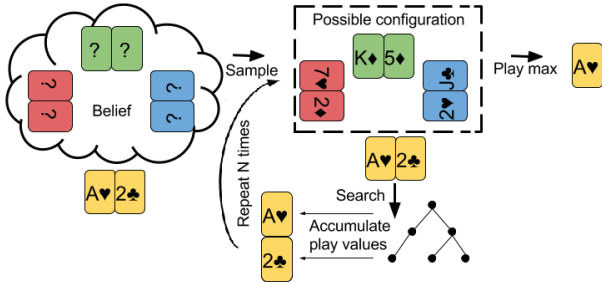


Figure 1: Illustration of the PIMC procedure in the 8<sup>th</sup> trick.

**Sample** In order to sample possible card distributions for the other three players with their real conditions, the artificial player must store all relevant information to update its beliefs on other players’ game states. Therefore, the information set summarises all visible information during a game, and also inferred information based on certain events. The player must keep an instance of the information set per game and update it when necessary. In this case, it stores the known hand of the player itself and a deck with all the cards whose owner is unknown. As a result, each time another player plays a card, it should be removed from that deck.

This data allows to sample distributions with more accurate information, considering the closer they are to the real world conditions, the more accurate the returning value of each search will be. Additionally, the information set keeps track of suits per player and, when a player does not follow the lead-suit of a trick, it removes that suit from the player possible suits. By possessing this information, sampling possible distributions gets even closer to the real world. However, it increases the complexity of the sampling process. The current sampling method builds a Constraint Satisfaction Problem (CSP) where the variables are the unplayed cards; the corresponding domains are the set of players that still have that suit; and the constraints are the number of times a player can be assigned to a card.

**Search** As mentioned above, PIMC has to calculate the reward of playing each possible card, for each sampled world. Since a sampled distribution assigns the remaining cards to players, every game can be handled as a perfect information game and the MinMax algorithm can perform this search.

However, exploring full game trees is a hard computational task, in light of our time constraint. This is particularly critical in the early plays of the game, since the number of computed game trees is the number of possible moves times the number of sampled distributions, and the size of

each game tree depends on the tricks that are left to finish the game. We set a time constraint of two 2 seconds for deliberation for each of the ten tricks. This time constraint was based on the observation of human *Sueca* players during a pilot study, where they took an average time of 2 seconds playing each card during a game.

To respect the time constraint, the MinMax search performed depth cuts on the game trees, which could only be fully explored from the 5<sup>th</sup> trick on. However, considering important game decisions might occur during the first tricks, this naive solution may compromise the results. To circumvent such difficulty, we propose a different approach to compute the perfect information games in the PIMC algorithm, presented in the continuation.

### The Hybrid player

In order for the algorithm to respect the time constraint without bounding the depth of the search, we contribute a *hybrid player* that computes the perfect information game with a MinMax search only from a certain trick on; up to that trick, each player node in the tree search selects a card according to a predefined play strategy. The trick from which we chose to compute full perfect information games is the 5<sup>th</sup> trick, because it is the highest that allows an average decision time of 2 seconds for a reasonable  $N$  ( $N = 50$ ).

The predefined play strategy is a stochastic version of the rule-based strategy adopted in the baseline players. It consists of a set of rules that eliminate potentially poor plays from the set of available plays, and then randomises between the remaining possibilities (usually a very small set).

Summarising, we introduce another loop into the PIMC algorithm (around line 7 of Algorithm 1), which causes the PIMC algorithm to calculate  $M$  times, instead of only once, the reward of playing each possible card for the  $N$  sampled worlds. In other words, this approach computes a total amount of  $N \times M$  roll-outs or perfect information games. The performance effects of the  $N$  and  $M$  parameters are further discussed.

### Parametrizations of Hybrid player

The Hybrid player that uses the PIMC algorithm can be configured by the  $N$  and  $M$  parameters, which correspond to the number of sampled distributions, and the number of computed perfect information games for each card in each sampled distribution, respectively.

The results presented in Table 1 correspond to simulations of 1,000 independent games, where a team with one Hybrid player and one Rule-based player played against a two Rule-based players’ team, for each different parametrization. The chosen parametrizations denote all possible combinations for  $N$  and  $M$  within  $\{1, 5, 10\}$ . These values were set considering a maximum number of 100 computed perfect information games (given by  $N \times M$ ), which already exceeds our time constraints.

By analysing the table above, we can generally conclude that the higher the  $M$  and  $N$  parameters are, the higher the mean points and the winning rate are, although suggesting that the observed growth may eventually stabilise. Comparing directly parametrizations of  $N = 1$  with  $M = 1$  for

Table 1: Mean points, standard deviations and winning rates obtained by the team with *one* Hybrid player and *one* Rule-based player against to *two* Rule-based players in 1,000 games for each parametrization

	N = 1	N = 5	N = 10
M = 1	58,8 ± 26,8 47,3%	61,2 ± 26,6 52,4%	61,4 ± 26,2 54,2%
M = 5	59,4 ± 26,5 50,3%	62,8 ± 25,8 55,8%	62,3 ± 25,6 54,6%
M = 10	61,4 ± 25,7 52,9%	63,1 ± 25,5 56%	63,2 ± 25,9 57%

the same number of computed perfect information games, one can conclude that sampling more is better than compute more games in the same sampled distribution. However, the bottom right square of 4 parametrizations does not lead to a similar conclusion. Due to the non-determinism of our Hybrid player, computations within the same sampled distribution require less sampled worlds in order to achieve a similar performance. What could easily have turned into redundant computations led indeed to a reduction of the execution time of all perfect information games. Naturally, a better sampling process requires more execution time, which in our case with the CSP, caused an impractical time consuming as the  $N$  parameter increases.

Additionally, the decision mean times of our Hybrid player were measured for each trick and the red-coloured configurations present the mean time of at least one trick higher than 2 seconds. The orange-coloured configurations have mean times lower than 2 seconds, although with large standard deviations presenting a high probability of taking more than 2 seconds in the decision procedure of some tricks. Consequently, we excluded the red-colored parametrizations of the table ( $\{M = 5, N = 10\}$ ,  $\{M = 10, N = 5\}$ , and  $\{M = 10, N = 10\}$ ), and selected the parametrization  $\{M = 5, N = 5\}$  as the best trade-off between performance and execution time.

Furthermore, it is important to note that this result was achieved by a team consisting of two players with a significant difference in skill, which will generally perform below the ability of the most skilled player. For that reason, we ran a second set of 1,000 games, now involving a team of two Hybrid players against a team of two Rule-based players. We used the chosen configuration of  $\{M = 5, N = 5\}$  for both Hybrid players. As expected, in this second set of experiments, the Hybrid-player-team attained a significantly larger winning rate of 63.7% ( $66.031 \pm 24.168$  points).

## The Social Module

The social module is able react upon a received game event by prescribing adequate interaction that use both verbal and non-verbal behaviours. Those behaviours were based on human players, since we conducted previously a user centred study to analyse how people behave during a *Sueca* card game. The collected behaviours were then categorised according to the corresponding game event, the game state and other relevant features of the *Sueca* game. For instance, the

analysed participants usually used an encouraging tone towards their partners and a competitive tone towards their opponents. As a result, an important consideration was that the social agent should differentiate utterances towards opponents and towards partners.

Moreover, we used Fearnot Affective Mind Architecture (FAtiMA) as the emotional agent’s architecture (Dias, Mascarenhas, and Paiva 2011), providing our social robot the ability to appraise each new event according to its goals in the game (e.g. winning), and therefore produce an adequate emotion. Each triggered emotion in our social robot is expressed through the physical posture of the embodiment, and also to index the subcategory of some utterances (e.g. gloating, resentment, happy, or pity when a player does a move). In other words, this emotional agent’s architecture allowed us to balance social and competitive aspects of our social robot player.

## Evaluation and Discussion

In order to evaluate the proposed social robot that plays the *Sueca* card game, we conducted a user study where the agent integrates the best achieved parametrization (see the previous subsection). Besides measuring the performance of the robot team as a competition factor, we also measured the level in which humans trust the social robot as their partner in *Sueca*. Henceforth, we have set up the environment in which the robot interacts with human players and analysed their perceived trust levels (see Figure 2).

We hypothesise that the trust levels that human players will report towards their robotic partner will be similar to the trust levels that human players report on their human partners. The motivation behind this hypothesis is related with the fact that the measure of human-robot trust (Schaefer 2013) was reported by its authors as a complex concept dependent upon many factors, but most importantly, upon robot-related factor as its performance. Our social robotic player not only uses behaviours that were inspired by human fashion as its performance was evaluated as good for simulations against a rule-based agent.

**Procedures and Methodology** The game playing experiment involved three humans playing *Sueca* with the autonomous robot EMotive headY System (EMYS) over a multi-touch table using physical cards, as Figure 2 suggests. Each session lasted about an hour and included the time participants took to answered to two sets of questionnaires: one before and one after playing with EMYS (*i.e.*, pre- and post-questionnaires). Moreover, the study followed the Ethical norms of conduct in which participants signed the consent form and assented to participate in the study.

Firstly, each participant selected his team player in a draw and according to each partner - having a human or robot partner -, participants answered to the pre-questionnaire called *Human-Robot Trust Questionnaire* (Schaefer 2013) (a modified version of the questionnaire was used for the participants with a human partner), aimed at measuring the expectation of trust towards their partner in the game. The *Human-Robot Trust Questionnaire* is a validated questionnaire that measures trust perception in HRI scenarios, con-

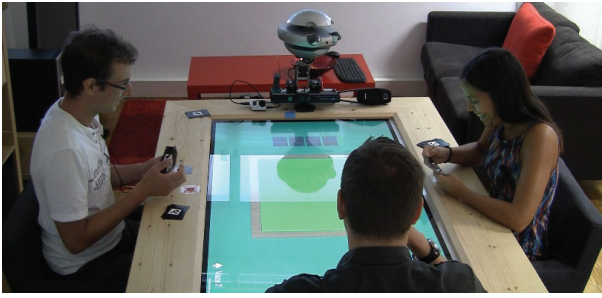


Figure 2: Example of a card game session with EMYS.

sidering the theoretical complexity of the construct.

Then, a researcher explained the game rules with a standard deck, and played some tricks until all participants agreed upon the game dynamics. This stage was performed to disambiguate the possible changes in the card game rules that sometimes occur, and to have a standardised version played during this study. After reviewing the *Sueca* game, participants moved to the touch table and started a session of five games with or against EMYS, considering their partner assigned in the initial draw.

When the five games ended, the participants answered to the post-questionnaires of *Human-Robot Trust Questionnaire* (Schaefer 2013), aimed at measuring the trust levels towards their partner after the game play experience. Moreover, participants also filled in a questionnaire related with demographic questions.

**Sample** A group of 60 participants participated in this study ( $M = 24,31$   $SD = 3,852$ ; 20 females, 39 males, 1 unknown). Out of the 60 subjects, 40 played the game with a human partner and 20 played with EMYS. These distributions aimed to collect a valid number of answers from EMYS' partners. Furthermore, the majority of the participants confirmed that their knowledge of the game was at a medium level.

**Competence** Out of the 20 performed sessions with human players, the robot's team won 12 and drawn 1 session of games, achieving a winning rate of 60%. The average difference of victory points between the robot's team and the other team was 1.5 with a standard deviation of 4.796.

**Trust Results and Discussion** To analyse if the trust levels changed after the experience of interaction in the *Sueca* card game a Mixed ANOVA statistical test was run on data. The results revealed a statistically significant difference ( $p = 0.03$ ) when comparing the pre- and post-levels of trust perceived by the participants towards their partners. This means that the trust levels increased after playing the card game when having both human and the robotic partners. Figure 3a shows the mean trust levels before and after playing the game, and we can see they increase from 70.5% to 75.0%, respectively.

This result shows that human players perceive an increase in their trust towards their partners after having played the game with them - either having played with a human or a

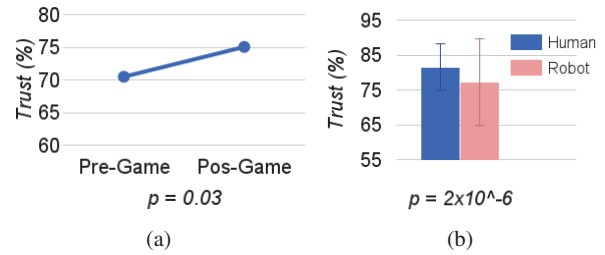


Figure 3: (a) Trust level comparison between the pre-game and post-game. (b) Trust levels towards the robotic and human partners after playing the game.

robotic partner. This result translates that the experience of having had interacted with a robotic partner in the game is similar with the experience of having interacted with a human partner. In turn, this tells us that both the social component and the game module developed for the robot are valid in the context of playing the game of *Sueca* in a team that considers both partners and opponents.

When analysing the trust levels of the pre- and post-questionnaires in relation with the type of the partner (human or robotic partner), the results did not reveal to be statistically significant ( $p = 0.65$ ), although the trust increased in both partner types. This means that, despite having a human or a robotic partner in the game, the trust levels remained similar and above average *i.e.*, above 50%.

This result enables us to corroborate our study hypothesis. By considering the results found, we can suggest that interacting with a human partner in the card game seems to show similar results in terms of trust when compared with playing the game with a robotic partner, taking into account the initial levels of trust. This means that the players had a similar experience of trust towards their partner in the case - being that partner a robot or a human.

Additionally, we run the statistical Welch Test to the data to analyse if the levels of trust were influenced by the type of game partner after playing the game. Henceforth, when taking into account only the trust levels perceived at the end of the interaction, results showed to be statistically significant between having a robot partner or a human partner ( $p < 0.01$ ). Participants' trust levels towards human partners was higher than the trust levels towards the robot as a partner (trust mean value of 81.538% and of 77.215% for human and robot partners, respectively - see Figure 3b).

By looking at Figure 3b we can see that the trust levels are indeed similar, however, trust acquired slightly higher values in human partners (81.538%) comparing with robotic partner (77.215%). This result is somehow expected as it shows that after playing the game the trust in human partners is slightly higher, denoting that humans tend to trust more in other humans when playing the game. Another reason that can explain this result relates with the fact that robots are a novel type of technology and therefore, unknown to

most people. The (possible and expectable) strangeness of interacting with a robot for the first time when playing the game can be a inhibitor factor for higher levels of trust. It is to note thought, that the difference between trust levels is not statistically significant, as mentioned earlier.

## Conclusion

In this paper, we have presented the design, development and test of a robotic game partner and opponent in a multi-player card game. Our results reveal that we have succeeded in this development and suggest that the behaviour of the robot, in conjunction with its ability to play the game well, led to a desirable balance of interaction.

The main goal of this work was to study if a social robot can be perceived as a trustable game partner in the entertaining activity of playing the card game of *Sueca*. However, studying trust is not a straightforward accomplishment due to the complex nature of this construct (Hancock et al. 2011), and according to the authors, the robot's factors as performance have an important role in the perception of trust. Therefore, we carefully considered the time constraint of playing in a social and natural context in the development of our social robotic player. Moreover, we conducted an analysis of possible parametrizations for the state-of-the-art algorithm for solving imperfect information card games (PIMC) in order to achieve the best compromise between execution time and performance. We believe the technique explored in our Hybrid Player can be beneficial when the constraints and requirements are similar to the ones we had – namely, computing optimal or nearly optimal values for each perfect information game tree within a human decision time.

From the results of the user study, it seems that yes - *we do seem to trust in a robot to be our game partner in a card game*. Also, the fact that the trust in the robot is similar to the trust in the human partner show the success of the social Human-Robot Interaction (HRI) interaction. Finally, we conclude that trust is an important construct in HRI entertainment scenarios.

## Acknowledgements

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT-UID/CEC/500 21/2013), through the project AMIGOS (PTDC/EEISII/7174/2014), and through the project LAW TRAIN (Ref. H2020-FCT-2014/ 653587). Filipa Correia acknowledges an FCT grant (Ref. SFRH/BD/118031/2016).

## References

Behrooz, M.; Rich, C.; and Sidner, C. 2014. On the Sociability of a Game-Playing Agent: A Software Framework and Empirical Study. *Intelligent Virtual Agents SE - 6* 8637:40–53.

Buro, M.; Long, J. R.; Furtak, T.; and Sturtevant, N. R. 2009. Improving state evaluation, inference, and search in trick-based card games. In *IJCAI*, 1407–1413.

Castellano, G.; Paiva, A.; Kappas, A.; Aylett, R.; Hastie, H.; Barendregt, W.; Nabais, F.; and Bull, S. 2013. Towards empathic virtual and robotic tutors. In *Artificial Intelligence in Education*, 733–736. Springer.

Dias, J. a.; Mascarenhas, S.; and Paiva, A. 2011. Fatima modular: Towards an agent architecture with a generic appraisal framework. *International Workshop on Standards for Emotion Modeling* 1–8.

Feil-Seifer, D., and Matarić, M. J. 2005. Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, 465–468. IEEE.

Frank, I., and Basin, D. 1998. Search in games with incomplete information: A case study using bridge card play. *Artificial Intelligence* 100(1-2):87–123.

Ginsberg, M. L. 2001. Gib: Imperfect information in a computationally challenging game. *Journal of Artificial Intelligence Research* 303–358.

Hancock, P. A.; Billings, D. R.; Schaefer, K. E.; Chen, J. Y.; De Visser, E. J.; and Parasuraman, R. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53(5):517–527.

Johnson, D. O.; Cuijpers, R. H.; Pollmann, K.; and van de Ven, A. A. 2016. Exploring the entertainment value of playing games with a humanoid robot. *International Journal of Social Robotics* 8(2):247–269.

Leite, I.; Castellano, G.; Pereira, A.; Martinho, C.; and Paiva, A. 2012. Modelling empathic behaviour in a robotic game companion for children: an ethnographic study in real-world settings. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 367–374. ACM.

Long, J. R.; Sturtevant, N. R.; Buro, M.; and Furtak, T. 2010. Understanding the success of perfect information monte carlo sampling in game tree search. In *AAAI*.

Pereira, A.; Martinho, C.; Leite, I.; and Paiva, A. 2008. icat, the chess player: the influence of embodiment in the enjoyment of a game. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3*, 1253–1256. International Foundation for Autonomous Agents and Multiagent Systems.

Pereira, A.; Prada, R.; and Paiva, A. 2012. Socially present board game opponents. In *Advances in Computer Entertainment*. Springer. 101–116.

Schaefer, K. E. 2013. *The perception and measurement of human-robot trust*. Ph.D. Dissertation, University of Central Florida Orlando, Florida.

Sturtevant, N. 2008. An analysis of UCT in multi-player games. *Computers and Games* 5131 LNCS:37–49.