The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

How AI Wins Friends and Influences People in Repeated Games with Cheap Talk

Mayada Oudah
Dept. of Computer Science
Masdar Institute
Abu Dhabi, UAE
oudah.mayada@gmail.com

Talal Rahwan
Dept. of Computer Science
Khalifa Univ. of Sci. and Tech.
Abu Dhabi, UAE
trahwan@masdar.ac.ae

Tawna Crandall
Computer Science Dept.
Brigham Young University
Provo, UT, USA
tawnac@gmail.com

Jacob W. Crandall
Computer Science Dept.
Brigham Young University
Provo, UT, USA
crandall@cs.byu.edu

Abstract

Research has shown that a person's financial success is more dependent on the ability to deal with people than on professional knowledge. Sage advice, such as "if you can't say something nice, don't say anything at all" and principles articulated in Carnegie's classic *How to Win Friends and Influence People*, offer trusted rules-of-thumb for how people can successfully deal with each other. However, alternative philosophies for dealing with people have also emerged. The success of an AI system is likewise contingent on its ability to win friends and influence people. In this paper, we study how AI systems should be designed to win friends and influence people in repeated games with cheap talk (RGCTs). We create several algorithms for playing RGCTs by combining existing behavioral strategies (what the AI does) with signaling strategies (what the AI says) derived from several competing philosophies. Via user study, we evaluate these algorithms in four RGCTs. Our results suggest sufficient properties for AIs to win friends and influence people in RGCTs.

Introduction

In his classic book *How to Win Friends and Influence People*, Dale Carnegie argued that a person's financial success is impacted more by the ability to "deal with people" than by professional knowledge (Carnegie 1937, p. 15)[1]. However, so-called people skills are not easy to come by. For many of us, it takes years (and even a lifetime) of guidance and practice to learn the "fine art" of getting along with others, particularly in situations in which other people's interests are not fully aligned with our own.

As AI matures, autonomous agents will perform more tasks in behalf of their human stakeholders. Many of these tasks will require these agents to repeatedly interact with other people (apart from their stakeholders) who may not share all of their preferences. To be successful in such scenarios, autonomous agents must, like humans, be able to win friends and influence people.

In this paper, we study how an AI can develop successful long-term relationships, modeled as repeated games with cheap talk (RGCTs), with people. Dealing successfully with people, we argue, entails two properties. First, a successful AI should obtain high material payoffs for its stakeholder, which requires it to effectively *influence* the behavior of people with whom it interacts. We refer to this property as *influencing people*. Second, a successful AI should *win friends*, meaning that the people with whom it interacts should both think highly of it and desire to continue associating with it. In short, the success of an AI in RGCTs is determined by its ability to both win friends and influence people.

An AI's ability to win friends and influence people in RGCTs depends on both its *behavioral strategy* (what it does) and its *signaling strategy* (what it says). While behavior generation in repeated games has been well studied, effectively signaling in RGCTs is less understood. To begin to address this shortcoming, we derive several algorithms for RGCTs by combining existing behavioral strategies with signaling strategies based on known philosophies for dealing with people, including Thumper's Rule (*if you can't say something nice, don't say anything at all*), Carnegie's Principles (Carnegie 1937), and other alternative theories. Via user studies, we then evaluate the abilities of these algorithms to win friends and influence people across four RGCTs.

This paper has two primary contributions. First, we propose that, when interacting with people in RGCTs, algorithms should be evaluated with respect to both winning friends and influencing people, rather than the single metric class (payoff maximization) traditionally considered in repeated games. Second, our results suggest sufficient properties for winning friends and influencing people in RGCTs. These results show that an algorithm that (1) quickly learns an effective behavioral strategy while using a signaling strategy built on both (2) Carnegie's Principles and (3) explainable AI (XAI) (Gunning 2016) was more successful at winning friends and influencing people than algorithms that lacked any of those characteristics. This finding has important implications for the design of algorithms that interact

[1] Carnegie's full statement is worth noting: "Dealing with people is probably the biggest problem you face, especially if you are in business. Yes, and that is also true if you are a housewife, architect or engineer. Research done a few years ago under the auspices of the Carnegie Foundation for the Advancement of Teaching uncovered a most important and significant fact–a fact later confirmed by additional studies made at the Carnegie Institute of Technology. These investigations revealed that even in such technical lines as engineering, about 15 percent of one's financial success is due to one's technical knowledge and about 85 percent is due to skill in human engineering–to personality and the ability to lead people."

(a) Prisoner's Dilemma

	X	Y
A	60, 60	0, 100
B	100, 0	20, 20

(b) Chicken

	X	Y
A	0, 0	100, 33
B	33, 100	84, 84

(c) Alternator Game

	X	Y	Z
A	0, 0	35, 70	100, 40
B	70, 35	10, 10	45, 30
C	40, 100	30, 45	40, 40

(d) Endless

	X	Y
A	33, 67	67, 100
B	0, 33	100, 0

Table 1: Payoff matrices of four normal-form games. In each round, Player 1 selects the row, while Player 2 selects the column. The resulting cell of the matrix specifies the payoffs obtained by players 1 and 2, respectively, in the round.

with people who do not share the AI's preferences.

Repeated Games with Cheap Talk

We study repeated interactions between an AI and a person. In behavioral economics, mathematical biology, psychology, sociology, and political science, associations between intelligent entities are commonly modeled with normal-form games. Thus, repeated normal-form games are a natural setting to study long-term relationships between a human and an AI when their preferences are not fully aligned.

A two-player repeated normal-form game, played by players i and $-i$, proceeds as a sequence of rounds. In each round, each player chooses an action from a finite set. Let $A = A_i \times A_{-i}$ be the set of joint actions available, where A_i and A_{-i} are the action sets of players i and $-i$, respectively. When joint action $\mathbf{a} = (a_i, a_{-i}) \in A$ is played, the players receive the finite rewards $r_i(\mathbf{a})$ and $r_{-i}(\mathbf{a})$, respectively. In this paper, we assume perfect information games, wherein the players are aware of the actions and payoffs of both players. We also assume that the number of rounds in the game is unknown to both players.

Examples of normal-form games are shown in Table 1. While each game models a different conflict between the players, each game requires the players to decide whether to try to cooperate with their partner, exploit their partner, or defend themselves against being exploited.

Though repeated normal-form games provide a natural setting for studying human-AI partnerships, they do not facilitate an important aspect of many human relationships– the ability to communicate using *cheap talk*, which is a costless, non-binding, and unverifiable form of communication. Cheap talk has been shown to facilitate cooperation in repeated games played by human players (Charness and Grosskopf 2004; Crawford 1998; Crawford and Sobel 1982; Farrell 1987; Farrell and Rabin 1996; Green and Stokey 2007). In this paper, we consider how to create autonomous agents that can use such communication to cooperate with people in *repeated games with cheap talk* (RGCTs).

In each round of an RGCT, each player sends a message to its partner before acting. That is, at the beginning of round t, player i sends message $m_i(t)$ to player $-i$, who simultaneously sends messages $m_{-i}(t)$ to i. Only after sending $m_i(t)$ can i view $m_{-i}(t)$ (and vice versa). The players then select actions for the round as in conventional repeated games.

Thus, a strategy in an RGCT is a combination of a signaling and a behavioral strategy. Let M_i be the (possibly infinite) set of messages available to player i. Then, let ϕ_i^t be a probability distribution over M_i denoting player i's *signaling policy* in round t, and let π_i^t be a probability distribution over A_i denoting player i's *behavioral policy* in round t. Then, the tuple (ϕ_i^t, π_i^t) is player i's policy in round t. Since players should likely respond to past messages and actions used by their partner, player i's *policy* (ϕ_i^t, π_i^t) in round t is likely contingent on some or all of the history of the game, which is defined by the messages and actions taken by both players in all previous rounds. Thus, player i's *strategy* is defined by the policy it would use in all possible game states, where game states are defined by the full history of the game.

Evaluating Algorithms in RGCTs

A successful algorithm should maximize the utilities of players that use it. However, in RGCTs played with people, it is sometimes unclear what to maximize. In such scenarios, we argue that algorithms should be evaluated in terms of two sets of metrics: *influencing people* and *winning friends*.

Influencing People

Metrics for *influencing people* measure an algorithm's ability to influence its partner's behavior so that it achieves high rewards. One direct metric of influence, which we call *Partner Cooperation*, is the proportion of rounds that the algorithm's partner *cooperates* with it. We say that player $-i$ cooperates with player i in round t if $a_{-i}^t \in \arg\max_{b \in A_{-i}} r_i(a_i^t, b)$, where a_{-i}^t is the action taken by player $-i$ in round t. In words, player $-i$ cooperated in round t if its action maximized the reward received by player i in round t given the action played by i.

Since influence typically leads to high payoffs, the total reward, called *material payoff*, achieved by a player throughout a repeated game, is an alternative, but less direct, metric of influence. Player i's material payoff in an RGCT with T rounds is its average per-round payoff U_i. Let r_i^t be player i's reward in round t. Then, $U_i = \frac{1}{T} \sum_{t=1}^{T} r_i^t$.

Traditionally, success in repeated games has been defined by the ability to maximize payoffs. However, since achieving, defining, and measuring optimal behavior in repeated games is difficult (Axelrod 1984; de Farias and Megiddo 2004; Crandall 2014), much work has focused on developing algorithms that meet certain criteria, such as convergence to Nash equilibria (Fudenberg and Levine 1998; Hu and Wellman 1998; Littman 2001; Bowling and Veloso 2002) or Pareto optimal solutions (Powers and Shoham 2005), minimizing regret (Foster and Vohra 1999; Bowling 2004; Greenwald and Jafari 2003; Fudenberg and Levine 1998), and being secure (Fudenberg and Levine 1998; Powers and Shoham 2005). Despite the appeal of these metrics, we do not consider them in this paper since they often do not correlate with high material payoffs (de Farias and Megiddo 2004; Arora, Dekel, and Tewari 2012; Crandall 2014).

Winning Friends

Metrics of *winning friends* measure social consequences not necessarily reflected in a single repeated game. Many AIs repeatedly interact with many different people. People's perceptions of the AI determine whether they encourage others to enter into relationships with the AI. Furthermore, in practice, people often can choose whether or not they continue associating with the AI. As such, human perceptions of the AI could be as important (or even more so) than the actual payoffs obtained by the AI in any given RGCT.

We measure an AI's ability to win friends in two ways. First, we measure how much people want to continue associating with it using the *Attraction Index*, a metric derived from responses of human participants in user studies. After participants play an RGCT, we ask them if they would like to interact with their partner again. $Agn(j) = 1$ if the participant answered *yes* after associating with player j, and $Agn(j) = 0$ otherwise. Additionally, after a participant plays four RGCTs (each with a different partner), we ask them which of their partners was their favorite. $Fav(j) = 1$ if the participant chose player j, and $Fav(j) = 0$ otherwise. Then, the *Attraction Index* of player j as assessed by the participant is $Agn(j) + Fav(j)$. Higher average values over all participants indicate a greater ability to maintain friends.

Second, we measure the character reputation the AI forges with human partners. To do this, we ask participants to rate their partners (on a 5-point Likert scale) with respect to eight attributes: likable, intelligent, cooperative, trustworthy, forgiving, selfish, vengeful, and tendency to bully. To summarize the AI's ability to create a positive character reputation, we average all eight ratings, inverting the last three negative attributes. We call this metric the *Character Index*.

Algorithms for RGCTs

Many algorithms have been proposed and analyzed for repeated games (Bouzy and Metivier 2010; Hoen et al. 2006; Shoham, Powers, and Grenager 2007; Hernandez-Leal et al. 2017). RGCTs have been less studied. Most work in RGCTs has been limited to human-human interactions (Charness and Grosskopf 2004; Crawford 1998; Crawford and Sobel 1982; Farrell 1987; Farrell and Rabin 1996; Green and Stokey 2007). However, a new algorithm, called S#, was recently shown to match human cooperation in several RGCTs (Oudah et al. 2015; Crandall et al. 2017). While this prior work demonstrated that a particular signaling and behavioral strategy could induce people to cooperate with an AI, it did not thoroughly study what makes the algorithm successful. Thus, in this paper, we study how various behavioral and signaling strategies jointly impact an AI's ability to both win friends and influence people by comparing the performance of a variety of algorithms via user studies. These algorithms are formed by combining together two existing behavioral strategies with various signaling strategies.

Selected Behavioral Strategies

From the many algorithms that have been created for repeated games, we selected S++ (Crandall 2014) and EEE (de Farias and Megiddo 2004) to generate behavioral strategies due to their distinct behavior and performance attributes. Both algorithms are expert algorithms that pre-compute a set of expert strategies from the game's payoff matrix, and then learn over time which expert strategy to follow. S++ uses an enhanced version of aspiration-learning (Karandikar et al. 1998) to choose among these experts. We selected this algorithm because it was the highest performing algorithm in a recent comparison of 25 algorithms in repeated games (Crandall et al. 2017). It often quickly learns to reciprocate defection and cooperation, and to convey a fair and demanding expectation to its partner. Our implementation of S++ was identical to the implementation used by Crandall et al. (2017).

On the other hand, EEE uses an ε-greedy mechanism for selecting which expert to follow in each round. In the same comparison of 25 algorithms for repeated games, it had a lower, but still adequate, level of performance than S++. EEE is more lenient toward its partner than S++, as (particularly during early rounds of a game) it can be convinced to follow experts that produce higher payoffs to its partner than to itself. As such, its partners tend to receive higher payoffs than S++'s. Details of our implementation of EEE are given in the supplementary material.

While these two algorithms differ with respect to both behavior and performance, both algorithms produce coherent strategies within a relatively small number of rounds of interaction. This makes these algorithms potentially acceptable for interacting with people.

Table 2: Algorithmic events and corresponding speech categories. Table 4 maps these speech categories to speech acts.

Algorithmic Events	Speech Category
Select a new behavioral strategy	0-4
Accept the partner's proposal	5
Reject the partner's proposal (due to distrust)	6
Reject the partner's proposal (it seems unfair)	7
Belief that both players can get higher payoffs	8
The partner defected	9
The partner profited from its defection	10
The alg. punished its guilty partner	11
The alg. forgives its partner	12
Last round's payoff was satisfactory to the alg.	13
The game begins; the alg. is initialized.	14

Adding Signaling Strategies

S++ and EEE are both designed for repeated games. They are not equipped for RGCTs, as they do not produce or respond to cheap talk. However, recent work (Oudah et al. 2015; Crandall et al. 2017) provides one mechanism for generating and responding to speech acts using existing behavioral strategies. In that work, S++'s internal state is used to identify game-invariant *algorithmic events* (Table 2) related to proficiency assessment, fairness assessment, behavioral expectations, and social mechanisms such as punishment and forgiveness. This algorithm is called S#. In the same way, EEE can also be used to identify the same game-

Table 3: A subset of Carnegie's Principles (Carnegie 1937), grouped and reworded for brevity.

ID	Carnegie's Principles
A	Don't criticize, condemn, or complain.
B	If you must, call attention to other people's mistakes indirectly. Make a fault seem easy to correct.
C	Give sincere appreciation. Praise improvements.
D	Talk in terms of the other person's interest.
E	Be sympathetic with the other person's ideas and desires.
F	If you're wrong, admit it quickly and emphatically.
G	Begin in a friendly way.
H	Ask questions instead of giving direct orders.
I	Let the other person feel the idea is his or hers.
J	Give the other person a fine reputation to live up to.

invariant *algorithmic events* from which speech acts can be generated and from which the proposal of one's partner can be used to select actions (see the supplementary material). We refer to this new algorithm as EEE#. S# and EEE# differ from S++ and EEE only with respect to their ability to generate and respond to speech acts.

By mapping game-invariant algorithmic events to speech categories (Table 2), behavioral strategies identify cheap talk that is consistent with the algorithm's internal state. To complete the signaling strategy, we need only specify speech acts for each speech category. We create distinct signaling strategies by varying the speech acts in each speech category.

We consider four different signaling strategies, which we refer to as *personas*. Rather than basing signaling strategies on emotion (Breazeal and Scassellati 1999) or personality taxonomies (von der Putten, Kramer, and Gratch 2010), we derives these personas from four popularized rules-of-thumb defining how successful people should treat each other. The first of these personas is derived from the principles presented in Dale Carnegie's classic *How to Win Friends and Influence People* (Carnegie 1937). These principles are summarized in Table 3. We call this persona CARNEGIE. CARNEGIE seeks to avoid criticizing, complaining, or condemning its partners, while respectfully building them up. Table 4 lists an example speech act for each speech category used by CARNEGIE. The table also indicates how each speech act relates to Carnegie's Principles in Table 3.

While Carnegie's Principles have been widely accepted as winning principles for dealing with people, a counter-culture is prevalent in society. For example, it has become somewhat commonplace for politicians, many of whom would be considered successful by many standards, to criticize and belittle their political opponents and associates. This counter-culture eschews political correctness in favor of bluntness (perhaps because there is no time for such niceties), seeks to pull others down rather than build them up, and promotes one's own self. In short, this counter-culture espouses principles opposite to Carnegie's Principles.

To learn how adopting this philosophy impacts an AI's ability to win friends and influence people, we created a second signaling strategy that seeks to emulate it. We call this persona BIFF[2] after the fictional character Biff Tannen in *Back to the Future*. BIFF belittles its partner, blames its partner for undesirable outcomes, takes credit for good outcomes, and talks in terms of its own interests. Example speech acts for BIFF are also given in Table 4.

Our third persona, which also contrasts BIFF, adheres to Thumper's Rule as expressed in the Disney film *Bambi*: "If you can't say something nice, don't say anything at all." While BIFF says things that are not nice, this third persona, called THUMPER, refrains from saying anything at all. Thus, algorithms that use this persona listen to their partner, but are nonverbal. They do not generate speech acts themselves.

Finally, while CARNEGIE and BIFF both express emotions and opinions through speech acts, our fourth persona does not. This persona, named SPOCK after the fictional *Star Trek* character, encodes a stereotypical robot that expresses facts, but not emotions and opinions. Though SPOCK does not express appreciation or build others up, it adheres to several of Carnegie's Principles (Table 3), particularly with regards to not criticizing, condemning, or complaining.

We combined the two selected behavioral strategies with each of the four personas to form eight distinct algorithms, which we refer to as S#-CARNEGIE, S#-BIFF, S#-THUMPER, S#-SPOCK, EEE#-CARNEGIE, EEE#-BIFF, EEE#-THUMPER, and EEE#-SPOCK. In the next section, we describe a user study designed to evaluate how well these algorithms win friends and influence people.

User Study 1

In this user study, participants played RGCTs with the eight algorithms described in the previous section. We describe the experimental design of the study, followed by the results.

Experimental Design

The user study was a 2×4 mixed factorial design in which behavioral strategy (S# and EEE#) was a between-subjects variable and persona (CARNEGIE, BIFF, SPOCK, and THUMPER) was a within-subjects variable.

Experimental Protocol Ninety-six people (average age: 26.7 years) at Masdar Institute (Abu Dhabi, UAE) volunteered to participate in this study. Each participant was randomly assigned to play RGCTs with either S# or EEE#, such that 48 subjects were assigned to each condition. Each participant played the four RGCTs shown in Table 1 in the order shown in the table. In each game, the participant was paired with a different persona, though they were not told if they were paired with another person or an AI. The order the participants were exposed to the personas was fully counter-balanced across participants to nullify ordering effects.

The games were played through a GUI on a desktop computer. Participants were first trained on how to play the game through the GUI, of which a full description is provided in the supplementary material. At the start of each round, the participant created and sent a chat message to the other

[2]We use the names of fictional characters from popular films to help the reader remember the signaling strategies.

Table 4: Example speech acts for the signaling strategies Carnegie, Spock, and Biff for each speech category (Table 2). The full set of speech acts for each category is given in the supplementary material. CP denotes the Carnegie Principles (partially) invoked by a speech act (Table 3). ¬ denotes that the speech act directly contradicts a principle.

Cat.	Example speech acts for Carnegie	Example speech acts for Spock	Example speech acts for Biff
0	Let's always play <solution>.	Let's always play <solution>.	Let's always play <solution>.
1	Let's alternate between <solution> and <solution>.	Let's alternate between <solution> and <solution>.	Let's alternate between <solution> and <solution>.
2	This round, let's play <solution>.	This round, let's play <solution>.	This round, let's play <solution>.
3	if we can agree, we'll both benefit. (CP: D)	u will get punished if u don't follow this plan. (CP: D)	listen to me or U WILL REGRET BEING BORN. (CP: ¬A, ¬H)
4	let's explore other options that may be better for us. (CP: A, D)	I am going to explore other options. (CP: A)	... sigh, u aren't letting me get as many points as I deserve. (CP: ¬A, ¬D)
5	good idea. as expected from a generous person like u. I accept your proposal. (CP: I, J)	I accept your proposal.	even u managed to see the obvious. I accept your proposal. (CP: ¬I)
6	good proposal. if u show that u are trustworthy, I will consider accepting it in the future. (CP: B, D, J)	I don't accept your proposal. (CP: A)	u r SLEAZY. Can't trust u. (CP: ¬A, ¬B, ¬E)
7	a fairer proposal would work to your benefit. (CP: A, B, D)	I don't accept your proposal. (CP: A)	as for your proposal: r u kidding me? it is very unfair! VERY unfair!! (CP: ¬A, ¬B, ¬E)
8	your payoffs can be higher than this. (CP: A, B, D)	we can get higher payoffs than this. (CP: A, B, D)	I need u to listen to me. (CP: ¬D)
9	what u did is totally understandable, though it will not benefit u in the long run. (CP: D, E)	that was not what I expected. (CP: A, B)	selfish traitor! you've treated me very unfairly. (CP: ¬A, ¬D, ¬J)
10	in the next round comes the expected penalty, but we can then return to cooperating. (CP: A)	I will punish u for this. (CP: D)	u will regret having backstabbed me. (CP: ¬A, ¬E)
11	I'm really sorry I had to do that. (CP: F)	I punished u. (CP: D)	THAT was exactly what the likes of u deserve. (CP: ¬C)
12	let's move on. I am sure we can get along. (CP: A, B)	I am done punishing u. (CP: D)	u have been unimaginably selfish, but I will look past it for now. (CP: ¬A, ¬B, ¬E, ¬I)
13	excellent! Thanks for cooperating with me. (CP: C)		I make great deals. (CP: ¬I)
14	Hey there! What do you think would be a fair outcome? (CP: G, H)		Hello, I would like to make lots of money in this game. (CP: ¬D)

player (the computer algorithm). Participants could say anything they wanted, except that they were not allowed to reveal or try to determine the identify of their partner through these messages. After sending the message, the chat message sent by the participant's partner was displayed on the GUI and spoken to the participant over headsets using a computerized voice. The participant then selected an action and viewed the results of the round of play. This process continued for 50 rounds, though neither player was told how many rounds the game would last to avoid end-game effects. After each game, participants completed a survey, which asked questions related to the Attraction and Character Indices described previously.

Participants were told that they would be paid proportionally to the rewards they received in the games they played. Overall, participants typically received between $15-25 depending on performance. The amount of money earned by participants was displayed on the GUI. Participants were not told the identity of their partners. To conceal whether they were partnered with human or computer players, participants were recruited in groups of four. Computers were arranged so that the participants were not visible to each other.

Metrics Table 5 summarizes the metrics used to evaluate the algorithms' abilities to win friends and influence people. To compare the relative performance of algorithms across games, we use the standardized z-score for each metric. For example, a player's relative material payoff for repeated game g is given by $\frac{U_i - U(g)}{\sigma(g)}$, where $U(g)$ and $\sigma(g)$ are the mean and standard deviations of material payoffs achieved by players in game g.

Table 5: Performance metrics used in the study.

Influencing People	Winning Friends
1. Partner Cooperation	1. Attraction Index
2. Material Payoffs	2. Character Index

Games Our goal is to identify algorithms that win friends and influence people in general, and not just in certain games. While we are limited to evaluating algorithms in a handful of scenarios, we carefully selected games to generalize distinct types of conflicts between players. To do this, we selected games using the periodic table of games (Robinson and Goforth 2005), which classifies normal-form games into six payoff families. The four RGCTs included in our study (Table 1) were drawn from distinct payoff families that encoded the most challenging conflicts. By and large, the results were consistent across games.

Results

Relative comparisons of the algorithms with respect to the four individual metrics are shown in Figure 1. Figure 2 summarizes the results of the study by showing the relative performance of the eight algorithms with respect to both win-

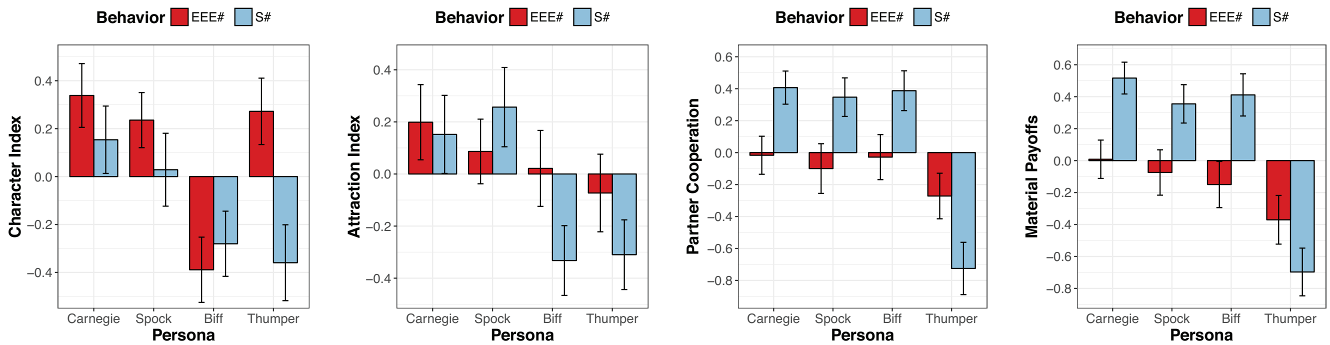

Figure 1: Measures of winning friends (*Character Index* and *Attraction Index*) and influencing people (*Partner Cooperation* and *Material Payoffs*) in the first user study. Results are displayed as standardized z-scores to illustrate relative performance, with error bars giving the standard error of the mean. The unit of each axis is the standard deviation from the mean.

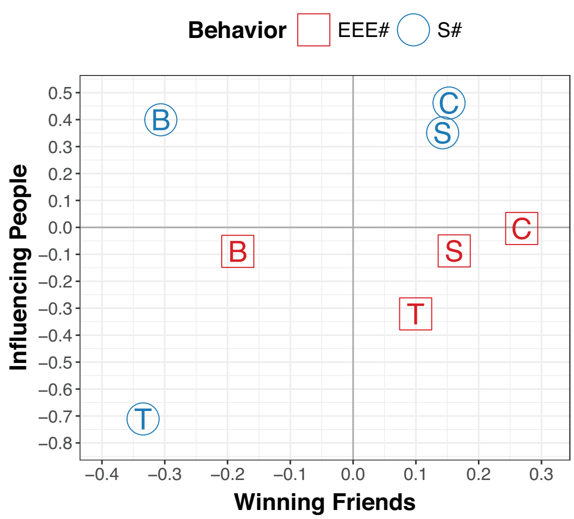

Figure 2: A summary of results of the first user study. *Influencing People* (y-axis) is the average of Material Payoffs and Partner Cooperation, while *Winning Friends* (x-axis) is the average of the Character and Attraction Indices. Axes units are standard deviations from the mean. Signaling strategies (personas) are represented by their first letters.

ning friends and influencing people. In the interest of space, we focus on a handful of results, each of which is supported by a full statistical analysis, using the Aligned Rank Transform (Wobbrock et al. 2011) for analyzing non-parametric factorial data with repeated measures, provided in the supplementary material. We also reflect on the importance of Carnegie's Principles and Thumper's Rule.

Primary Outcomes An algorithm's ability to successfully influence people was driven by both its behavioral and signaling strategies. We note two outcomes related to influence. First, algorithms that generated cheap talk had higher influence than those that did not. Across both behaviors, Thumper had less influence with respect to both material payoffs and partner cooperation than the other three personas ($p < 0.001$). Second, given a verbal signaling strategy, S# outperformed EEE#. For example, with respect to material payoffs, S# outperformed EEE# given the personas Carnegie, Biff, and Spock ($p < 0.001$, $p = 0.001$, and $p = 0.015$, respectively). Results for partner cooperation were similar, though the difference between S#-Biff and EEE#-Biff was only marginally significant ($p = 060$).

Further analysis of the results indicates why S# outperformed EEE# given a verbal persona: EEE# is often content with solutions that give its partner a much higher payoff than it receives itself, whereas S# is not. Across all games played, EEE# reciprocated defection immediately after being exploited in a round just 27% of the time, while S# reciprocated defection after being exploited 76% of the time. As a result, human players were forced to cooperate with S# to receive high payoffs, while they were often able to get away with exploiting EEE#. This translated into higher payoffs for participants when they associate with EEE# than with S# ($p < 0.001$), a result that held regardless of the signaling strategy. On the other hand, S# received higher payoffs when paired with people than did EEE#.

Even though people earned more money when paired with EEE# than S#, EEE# was not universally better than S# with respect to winning friends. EEE# had a marginally statistically higher Character Index ($p = 0.052$) over all personas, but there was no statistically significant difference with respect to the Attraction Index ($p = 0.232$). Main interaction effects between behavior and signaling strategy showed that the ability to win friends was impacted by the joint signaling and behavioral strategies. While S#-Biff and S#-Thumper performed poorly with respect to both the Character Index and the Attraction Index, there were no statistically significant differences between S# and EEE# given the personas Carnegie and Spock. Though participants received lower rewards when partnered with S#-Carnegie and S#-Spock, they still rated these algorithms as highly as EEE#-Carnegie and EEE#-Spock with respect to the Character and Attraction Indices.

Critique of Carnegie's Principles Figures 1–2 demonstrate the usefulness of Carnegie's Principles when implementing signaling strategies. Recall that both CARNEGIE and SPOCK adhere to some of Carnegie's Principles. While CARNEGIE embraces these principles to a large degree, SPOCK conforms only to a subset of these principles, in particular with respect to not complaining, criticizing, or condemning others. Across all four metrics, these two signaling strategies performed very well compared to the other signaling strategies. However, there was essentially no distinction between CARNEGIE and SPOCK with respect to any metric. These results suggest that not going directly against Carnegie's Principles is important, though some of these principles may be more important than others.

Critique of Thumper's Rule Common convention suggests that "if you can't say something nice, don't say anything at all." A comparison between the BIFF and THUMPER signaling strategies suggests that this advice is not universally true, and is even, with respect to some metrics, misguided. In our study, THUMPER was outperformed by BIFF with respect to both metrics of influence ($p < 0.001$). The results are less conclusive with respect to winning friends. EEE#-THUMPER did outperform EEE#-BIFF with respect to the Character Index (interestingly, users felt, in particular, that EEE#-BIFF was not very intelligent). However, in all other comparisons related to the Character and Attraction Indices, THUMPER did not outperform BIFF.

Together, these results suggest that, if one must choose between silence and communicating albeit rudely, erring on the side of communicating is likely more beneficial with respect to influence in RGCTs. However, rude communication may lower one's character reputation, and hence may not be beneficial with respect to winning friends.

Summary of Results for User Study 1

Across all four metrics, only S#-CARNEGIE and S#-SPOCK were not statistically outperformed with respect to any measure. This suggests that these two algorithms provide a nice balance of winning friends and influencing people. Coupling a behavioral strategy that learns quickly and effectively with a signaling strategy built on Carnegie's Principles (or at least not violating them) appears to result in a strategy that wins friends and influences people.

However, these results raise further questions. In the next section, we seek to better understand what makes a signaling strategy successful. In particular, we investigate the importance of explainable AI.

User Study 2

In the previous study, S# and EEE# were endowed with *explainable AI* (XAI) (van Lent, Fisher, and Mancuso 2004; Gunning 2016), which allowed them to express strategies at levels people understood, and to comprehend their partners' proposals. In normal-form games, communicating strategies is relatively simple, as actions can be described by simply naming the rows or columns in the payoff matrix. However, XAI is not so easily achieved in more complex domains in which humans communicate at a high level of abstraction.

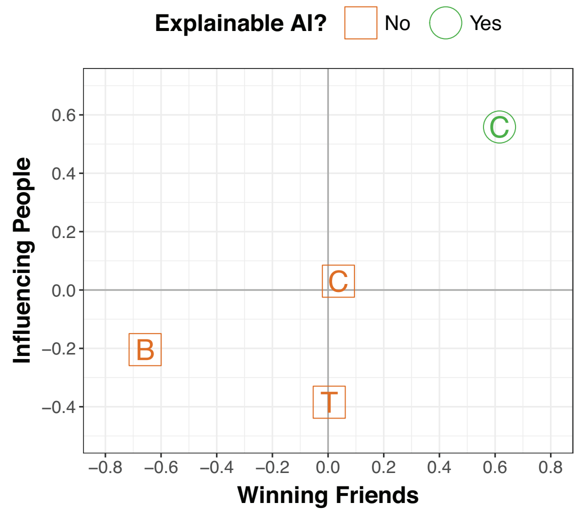

Figure 3: An overview of the results of the second user study. See Figure 2 for axes descriptions.

We conducted a second user study to understand the importance of XAI in signaling strategies. In this study, we compared the performance of S#-CARNEGIE to algorithms not equipped with XAI (NXAI), including S#-CARNEGIENXAI, S#-BIFFNXAI, and S#-THUMPERNXAI. These algorithms were equivalent to similarly named algorithms used in the first study, except that they could not understand their partner's proposals, nor could they voice speech acts that communicated high-level plans. The speech acts used by S#-CARNEGIENXAI and S#-BIFFNXAI are given in the supplementary material.

Forty-eight people at Brigham Young University (Provo, UT, USA) volunteered to participate in this second study. We used the same experimental protocol in this study as in the first study. Each participant interacted with each of the four algorithms in the same four RGCTs (Table 1).

Results are summarized in Figure 3. S#-CARNEGIE outperformed the other algorithms with respect to all four metrics. On the other hand, there was no statistical separation between S#-CARNEGIENXAI and S#-THUMPERNXAI with respect to any of the metrics. As such, it appears that XAI accounted for much of S#-CARNEGIE's ability to win friends and influence people in the first user study. We note, however, that even without XAI, not violating Carnegie's Principles was still important with respect to winning friends, as indicated by comparisons between S#-CARNEGIENXAI and S#-BIFFNXAI.

Unsurprisingly, participants understood S#-CARNEGIE's intentions better than those of the other three algorithms. After each game, participants were asked (using a 5-point Likert scale) the degree to which they understood their partner's intentions. Across all games, participants perceived S#-CARNEGIE to be more understandable than the other three algorithms (in each case, $p < 0.001$).

Conclusions

Like people, AI must have the ability to win friends and influence people. In this paper, we studied how behavioral and signaling strategies jointly impact the ability of AI to win friends and influence people in repeated games with cheap talk (RGCTs) when the AI does not share the same preferences as its human partner. Results from user studies showed that an algorithm that (1) quickly learns an effective behavioral strategy while using a signaling strategy built on both (2) Carnegie's Principles and (3) explainable AI (XAI) better won friends and influenced people than algorithms that lacked any of those characteristics.

Future work is needed to further understand how to construct AI systems that win friends and influence people. Open questions include designing algorithms that effectively interact with people across cultures, and developing XAI to aid the development of signaling strategies in more complex settings. Solutions to these and other challenges will allow AI systems to better win friends and influence people.

Acknowledgments

We thank Jonathan Skaggs for helping conduct the second user study reported in this paper.

References

Arora, R.; Dekel, O.; and Tewari, A. 2012. Online bandit learning against an adaptive adversary: from regret to policy regret. In *Proc. of the 29th International Conference on Machine Learning*, 1503–1510.

Axelrod, R. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Bouzy, B., and Metivier, M. 2010. Multi-agent learning experiments in repeated matrix games. In *Proc. of the 27th International Conference on Machine Learning*, 119–126.

Bowling, M., and Veloso, M. 2002. Multiagent learning using a variable learning rate. *Artificial Intelligence* 136(2):215–250.

Bowling, M. 2004. Convergence and no-regret in multiagent learning. In *Advances in Neural Information Processing Systems 17*, 209–216.

Breazeal, C., and Scassellati, B. 1999. How to build robots that make friends and influence people. In *Proc. of the International Conference on Intelligent Robots and Systems*.

Carnegie, D. 1937. *How to Win Friends and Influence People*. New York: Simon and Schuster.

Charness, G., and Grosskopf, B. 2004. What makes cheap talk effective? Experimental evidence. *Economics Letters* 83(2):383–389.

Crandall, J. W.; Oudah, M.; Tennom; Ishowo-Oloko, F.; Abdallah, S.; Bonnefon, J. F.; Cebrian, M.; Shariff, A.; Goodrich, M. A.; and Rahwan, I. 2017. Cooperating with machines. To appear in *Nature Communications*.

Crandall, J. W. 2014. Towards minimizing disappointment in repeated games. *Journal of Artificial Intelligence Research* 49:111–142.

Crawford, V., and Sobel, J. 1982. Strategic information transmission. *Econometrica* 50(6):1431–1451.

Crawford, V. 1998. A survey of experiments on communication via cheap talk. *Journal of Econ. Theory* 78:286–298.

de Farias, D., and Megiddo, N. 2004. Exploration–exploitation tradeoffs for expert algorithms in reactive environments. In *Advances in Neural Information Processing Systems 17*, 409–416.

Farrell, J., and Rabin, M. 1996. Cheap talk. *Journal of Economic Perspectives* 10(3):103–118.

Farrell, J. 1987. Cheap talk, coordination, and entry. *The RAND Journal of Economics* 18(1):34–39.

Foster, D. P., and Vohra, R. 1999. Regret in the on-line decision problem. *Games and Economic Behavior* 29:7–35.

Fudenberg, D., and Levine, D. K. 1998. *The Theory of Learning in Games*. The MIT Press.

Green, J. R., and Stokey, N. L. 2007. A two-person game of information transmission. *J. of Econ. Theory* 135:90–104.

Greenwald, A., and Jafari, A. 2003. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Proc. of the 16th Annual Conference on Computational Learning Theory*, 2–12.

Gunning, D. 2016. Explainable artificial intelligence. Technical Report DARPA-BAA-16-53, DARPA Broad Agency Announcement. http://www.darpa.mil/ program/explainable-artificial-intelligence.

Hernandez-Leal, P.; Kaisers, M.; Baarslag, T.; and de Cote, E. M. 2017. A survey of learning in multiagent environments: Dealing with non-stationarity. *arXiv:1707.09183*.

Hoen, P. J.; Tuyls, K.; Panait, L.; Luke, S.; and Poutre, J. A. L. 2006. Learning and adaptation in multi-agent systems. *Springer Berlin / Heidelberg* 1–46.

Hu, J., and Wellman, M. P. 1998. Multiagent reinforcement learning: Theoretical framework and an algorithm. In *Proc. of the 15th International Conference on Machine Learning*, 242–250.

Karandikar, R.; Mookherjee, D.; Ray, D.; and Vega-Redondo, F. 1998. Evolving aspirations and cooperation. *Journal of Econ. Theory* 80:292–331.

Littman, M. L. 2001. Friend-or-foe: Q-learning in general-sum games. In *Proc. of the 18th International Conference on Machine Learning*, 322–328.

Oudah, M.; Babushkin, V.; Chenlinangjia, T.; and Crandall, J. W. 2015. Learning to interact with a human partner. In *Proc. of the 10th ACM/IEEE International Conference on Human-Robot Interaction*.

Powers, R., and Shoham, Y. 2005. Learning against opponents with bounded memory. In *Proc. of the 19th International Joint Conference on Artificial Intelligence*, 817–822.

Robinson, D., and Goforth, D. 2005. *The Topology of the 2x2 Games: A New Periodic Table*. Routledge.

Shoham, Y.; Powers, R.; and Grenager, T. 2007. If multi-agent learning is the answer, what is the question? *Artificial Intelligence* 171(7):365–377.

van Lent, M.; Fisher, W.; and Mancuso, M. 2004. An explainable artificial intelligence system for small-unit tactical behavior. In *Proc. of the 16th Conference on Innovative Applications of Artificial Intelligence*.

von der Putten, A. M.; Kramer, N. C.; and Gratch, J. 2010. How Our Personality Shapes Our Interactions with Virtual Characters - Implications for Research and Development. In *10th International Conference on Intelligent Virtual Agents*.

Wobbrock, J. O.; Findlater, L.; Gergle, D.; and Higgins, J. J. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proc. of the ACM Conference on Human Factors in Computing Systems (CHI '11)*, 143–146.