

Grounding Strategic Conversation: Using negotiation dialogues to predict trades in a win-lose game

Anaïs Cadilhac
IRIT
Univ. Toulouse, France
cadilhac@irit.fr

Nicholas Asher
IRIT, CNRS
Toulouse, France
asher@irit.fr

Farah Benamara
IRIT
Univ. Toulouse, France
benamara@irit.fr

Alex Lascarides
School of Informatics
Univ. Edinburgh, UK
alex@inf.ed.ac.uk

Abstract

This paper describes a method that predicts which trades players execute during a win-lose game. Our method uses data collected from chat negotiations of the game *The Settlers of Catan* and exploits the conversation to construct dynamically a partial model of each player’s preferences. This in turn yields equilibrium trading moves via principles from game theory. We compare our method against four baselines and show that tracking how preferences evolve through the dialogue and reasoning about equilibrium moves are both crucial to success.

1 Introduction

Rational agents act so as to maximise their *expected utilities*—an optimal trade off between what they *prefer* and what they *believe* they can achieve (Savage, 1954). Solving a game problem involves finding *equilibrium strategies*: an optimal action for each player that maximises his expected utility, assuming that the other players perform their specified action (Shoham and Leyton-Brown, 2009). Calculating equilibria thus requires knowledge of the other players’ preferences but almost all *bargaining* games occur under the handicap of imperfect information about this (Osborne and Rubinstein, 1994). Players therefore try to extract their opponents’ preferences from what they say, likewise revealing their own preferences in their own utterances. These elicited preferences guide an agent’s decisions, like choosing to make such and such a bargain with such and such a person. Tracking preferences through

dialogue is thus crucial for analyzing the agents’ strategic reasoning in real game scenarios.

In this paper, we design a model that maps what people say in a win-lose game into a prediction of exactly which players, if any, trade with each other, and exactly what resources they exchange. We use both statistics and logic: we use a corpus of negotiation dialogues to learn classifiers that map each utterance to its speech act and to other acts pertinent to bargaining; and we develop a symbolic algorithm that, from the classifiers’ output, dynamically constructs a model of each player’s preferences as the conversation proceeds (for instance, the preference to receive a certain resource, or to accept a certain trade). This preference model uses CP-nets (Boutilier et al., 2004), a representation of preferences for which algorithms for computing equilibrium strategies exist. We adapt those algorithms to predict the trades executed in the game.

The algorithm for constructing CP-nets uses *only* the output of our classifiers, which in turn rely entirely on shallow features in the raw text and robust parsers. Together they provide an end to end model, from raw text to a prediction of which trade, if any, occurred. We evaluate the various components of this (pipeline) algorithm separately, as well as the end to end model.

Our study exploits a corpus of negotiation dialogues from an online version of the win lose game *The Settlers of Catan*. Sections 2 and 3 describe the corpus and its annotation. Section 4 introduces our method for constructing the agents’ preferences from the dialogues. We use this in Section 5 to predict whether a trade is executed as a result of the

players' negotiations, and if so we predict who took part in the trade, and what they exchanged. Our method shows promising results, beating baselines that don't adequately track or reason about preferences. We compare our model to related work in Section 6 and point to future work in Section 7.

2 The game

The Settlers of Catan (www.catan.com) is a win-lose game that involves negotiations over restricted resources. Each player (three or more) acquires resources (of 5 types: ore, wood, wheat, clay, sheep), which they use in different combinations to build roads, settlements and cities, which in turn earns them points towards winning. The first player to 10 points wins. Players acquire resources in several ways, in particular through agreed trades with other players. Some methods (e.g., robbing) are hidden from view, so players lack complete information about their opponents' resources.

Our corpus contains conversations of humans playing an online version of *Settlers* (Afantenos et al., 2012). Players must converse in a chat interface to carry out trades. Each game contains several dozen self-contained bargaining dialogues. Our experiments use 10 *Settlers* games, consisting of more than 2000 individual dialogue turns (see Section 3).

Table 1 is a sample dialogue from the corpus. The sentences in the corpus have a relatively simple syntax, though many also exhibit long distance dependencies. However, these conversations are pragmatically complex. They exhibit complex anaphoric dependencies (e.g., utterance ID 4 in Table 1). Other pragmatic inferences, which are dependent on reasoning about intentions, speech acts and discourse structure, are also ubiquitous. For example, the question *Have you got any ore?* implies an offer for the speaker to receive ore in exchange for something from someone unspecified, and its response *I've got wheat* not only implies a willingness to exchange wheat for something, but as a response to the question it also implies a refusal to give any ore.

More generally, a dialogue turn in our corpus can express an offer, a counteroffer, an acceptance or rejection of an offer, or a commentary on the above or on moves in the game. All except the last provide clues about preferences: e.g., which players

a speaker wants to execute a trade with; or what resources to exchange. For instance, the utterance *Anybody have any sheep for wheat?* conveys several preferences. First, it conveys the speaker's preference to trade with someone unspecified. Other informative but underspecified preferences include: the speaker's preference to acquire some sheep over alternatives; and in a context where she receives sheep, a preference to give away some of her wheat over the alternatives. Crucially, it does not convey a preference to give away wheat in a context where she receives nothing or something other than sheep.

In line with a non-cooperative bargaining game, the preferences and offers that a speaker reveals are less specific than an executable trade requires, where the trading partners and the type of resources offered and received must all be defined. Such general dialogue moves are essentially information seeking—evidence that humans playing *Settlers* have imperfect information about their opponents' preferences. In fact, many offers to trade result in no trade being agreed to and executed. While observed negotiation failure would be puzzling in a bargaining game with perfect information (Osborne and Rubinstein, 1994), it occurs relatively frequently in *Settlers*.

3 Annotation

We have a multi-layered dialogue annotation scheme that includes: (1) a pre-annotation that segments the dialogue into turns which are further segmented into Elementary Discourse Units (EDUs) with the author of each turn automatically given; (2) a characterization of each EDU in terms of basic speech acts (assertion, question, request) as well as dialogue acts that are specific to bargaining (offers, counteroffers, etc.); and (3) associated information about the givable and/or receivable resources that EDUs express.

Two annotators received training on 77 dialogues, totaling 699 EDUs. They then both annotated the remaining dialogues independently (2741 EDUs and 511 dialogues in total). Kappas for inter-annotator agreement are given below.

3.1 Dialogue act annotation (Kappa=0.79)

Each turn logs what a player enters in the chat window and also aspects of the game state at the time:

ID	Dialogue Act	Text	Speaker	Addressee	Resource
1	Offer	i need clay, anyl have?	Rainbow	All	Receivable (clay, ?)
2	Refusal	Nope, sorry	inca	Rainbow	
3	Refusal	Not at the moment, unfortunately.	ariachiba	Rainbow	
4	Refusal	need mine sorry	Kittles	Rainbow	Not givable (Anaphoric, ?) Anaphora Link:(mine , clay)
5	Offer	no one has ore to giv?	Rainbow	All	Receivable (ore, ?)
6	Accept	oh yeah me	Kittles	Rainbow	
7	Counteroffer	ore for wheat again?	Kittles	Rainbow	Givable (ore, ?) Receivable (wheat, ?)
8	Accept	ya	Rainbow	Kittles	
9	Accept	ok	Kittles	Rainbow	

Table 1: Example of an annotated negotiation dialogue.

his resources, the state of the game board and a time stamp. The pre-annotation divides each turn into EDUs. The annotators then have to specify the *dialogue act* of each EDU: *Offer*, *Counteroffer*, *Accept* or *Refusal* (of an offer addressed to the emitter), and *Other*. *Other* labels units that either comment on strategic moves in the game or are not directly pertinent to bargaining. Annotators also specify the *addressee* of the EDU and its *surface type*: *Question*, *Request* or *Assertion*.

3.2 Resource type annotation ($\text{Kappa}=0.80$)

Annotators also specify for each EDU and its dialogue act an associated *feature structure*, which captures (partial) information that the EDU expresses about the type and quantity of resources that are of the following four attributes: *Givable*, *Not Givable*, *Receivable* or *Not Receivable*. These attributes can take Boolean combinations of resources as values via two operators AND and OR, that respectively stand for conjunction (the agent expresses two preferences and he prefers to achieve one of them if he cannot have both, such as *I need clay and wood*) and disjunction (free choice) of preferences (e.g., *I can give you clay or wood*). We allow attributes to have unknown values: the annotation tool inserts a ? in these cases. We also insist that the annotators resolve anaphoric dependencies when specifying values to attributes, as shown in EDU (4) in Table 1.

4 Dialogue act and resource prediction

Predicting the executed trades from the dialogues starts with three sub-tasks: automatically identifying each EDU’s dialogue act; detecting the EDU’s resources; and specifying the attributes of those resources (i.e., *Givable*, *Receivable*, etc.).

4.1 Identifying dialogue acts

As is well established, one EDU’s dialogue act depends on previous dialogue acts (Stolcke et al., 2000). In our corpus, *Accept* or *Reject* frequently follow *Offer* and *Counteroffer*. Since labeling is sequential, we use Conditional Random Fields (CRFs) to learn dialogue acts. CRFs have been shown to yield better results in dialogue act classification on online chat than HMM-SVN and Naive Bayes (Kim et al., 2012).

We use three types of features: lexical, syntactic and semantic. And we exploit them as unigrams and bigrams: unigrams associate the value of the feature with the current output class (level 0); bigrams take account of the value of the feature associated with a combination of the current output class and previous output class (level -1). 6 features were used exclusively as unigrams: the EDU’s position in the dialogue, its first and last words, its subject lemma, a boolean feature to indicate if the current speaker is the one that initiates the dialogue and the position of the speaker’s first turn in the dialogue.

We have 15 unigram and bigram features (at levels 0 and -1), as well as templates that combine feature values for the two levels. These include 14 boolean features that indicate if the EDU contains: bargaining verbs (e.g. *trade*, *offer*), references to another player (e.g. *you*), resource tokens as encoded in a task dedicated lexicon (e.g. *wheat*, *clay*), quantifiers (e.g. *one*, *none*), anaphoric pronouns, occurrences of “for” prepositional phrases (e.g. *wheat for clay*), acceptance words (e.g. *OK*), negation words, emoticons, opinion words (from (Benamara et al., 2011)), words of politeness, exclamation marks, questions, and finally whether the EDU’s speaker has talked previously in the dialogue.

The last feature gives the EDU speaker lemma. In addition, 3 unigram and bigram booleans indicate whether the current EDU contains the most frequent tokens, couple of tokens and syntactic patterns in our corpus. Finally, we use 2 composed bigram features that encode whether the EDU contains an acceptance or refusal word, given that the previous EDU is a question.

To assign sequential tags of dialogue acts within a negotiation dialogue, we use the CRF++ tool (crfpp.googlecode.com). Our data consists of 2741 EDUs in 511 dialogues. Each EDU is associated with a dialogue act resulting in 410 *Offer*, 197 *Counteroffer*, 179 *Accept*, 398 *Refusal* and 1557 *Other*. We use 10-fold cross-validation to evaluate our model, computing precision, recall and F-score for each class and global accuracy from the total number of true positives, false positives, false negatives and true negatives obtained by summing over all fold decisions. The results (in percent) are given in Table 2 (MaF is the average of F-scores of all the classes). Our model significantly outperforms the frequency-based baseline (MaF=14.5; Accuracy=56.8), with the best F-score achieved for *Other*. The least good results are for the two least frequent classes in our data. In addition to the frequency problem, the lower score for *Counteroffer* is mainly due to the model confusing it with *Offer*. Errors in the *Accept* class were often due to misspelling or to chat style conversation; e.g., *kk*, *yup*.

Dialogue act	Precision	Recall	F-score
<i>Other</i>	87.4	93.1	90.1
<i>Offer</i>	80.0	81.0	80.5
<i>Counterof.</i>	64.8	53.3	58.5
<i>Accept</i>	65.1	53.1	58.5
<i>Refusal</i>	81.7	73.9	77.6
Macro-averaged F-score (MaF)			73.0
Accuracy			83.0

Table 2: Results for dialogue act classification.

4.2 Finding resource text spans

Since the resource vocabulary in *The Settlers of Catan* is a closed set composed of words denoting specific resources (e.g., *clay*, *wood*) and their synonyms (*brick*), we use a simple rule to detect them: a noun phrase (NP) is a resource text span if and

only if it contains a lemma from our resource lexicon. A closed set resource vocabulary is common to many different types of negotiation dialogues. We used the Stanford parser (Klein and Manning, 2003) to obtain the NPs: there are 4361 NPs, where (by the gold standard annotations) 21% are resources and 79% are not. We obtain an F-score of 96.9% and accuracy of 97.9%, clearly beating both the frequency and random baselines for this task.

4.3 Recognizing the type of resources

Recall that each resource within an EDU can be the value of four types of attributes: *Givable*, *Receivable*, *Not Givable* or *Not Receivable* (cf. Section 3.2). We predict these attributes using CRFs with the following features. 8 features are used as unigram at the current and the previous EDU level: the speaker, the EDU’s subject, the dialogue act, and (if present) the lemma of a bargaining verb, and 4 boolean features indicate if the EDU contains an opinion word, a reference to another speaker, if the resource comes after a “for” and if it contains a refusal word. These features also serve as bigrams at the current EDU level. Additionally, we have a set of unigram and bigram boolean features that indicate if the current EDU contains the most frequent verbs in the corpus. And finally, we use a feature that encodes the combination subject/bargaining verb in the current EDU.

We used CRF++ to implement our classifier. Our corpus data consists of 1077 Resources, split into 510 *Receivable*, 432 *Givable*, 116 *Not Givable* and 19 *Not Receivable*. We use again 10-fold cross-validation to evaluate our model and compute the results by summing over all fold decisions. We present them (in percent) in Table 3. They beat the frequency-based baseline (MaF=16.1; Accuracy=47.4), although performance on the *Not Receivable* class is poor probably due to its low frequency in the data.

Ambiguities make this task challenging. For instance, *anyone wheat for clay?* can mean that the speaker wants to receive wheat and give clay or the opposite, and resolving which meaning is intended involves reasoning not only with the previous and/or the following EDU, but also sometimes EDUs with long distance attachments, which are not supported by our classifier and require a full discourse parser.

Res. type	Precision	Recall	F-score
<i>Receivable</i>	66.8	71.4	69.0
<i>Givable</i>	62.6	59.7	61.1
<i>Not Giv.</i>	88.1	89.7	88.9
<i>Not Rec.</i>	0	0	0
Macro-averaged F-score (MaF)			54.8
Accuracy			67.4

Table 3: Results for resource type classification.

5 Predicting Players’ Strategic Actions

We aim to capture the evolution of commitments to certain preferences as the dialogue proceeds so as to predict the agents’ bargaining behavior. In other words, we wish to predict which of the 61 possible trade actions is executed at the end of each dialogue. The possible trades vary over which partner the player whose turn it is trades with (3 options in a 4 player game), the resources exchanged (assuming each partner gives one type of resource and receives another type yields $5 \times 4 = 20$ possibilities), or there is no trade; i.e., $(3 \times 20) + 1 = 61$ possible actions in the hypothesis space (we predict the types of resources that are exchanged, but not their quantity).

We predict the executed action by identifying the equilibrium trade entailed by the model of the players’ preferences, which in turn we construct dynamically from the output of the classifiers in Section 4. We use the attributes of resources in the EDUs (*Givable*, etc.) to identify the preference that a speaker conveys in the EDU, and we use the dialogue acts (*Offer*, *Accept*, etc.) to update a model of the preferences expressed so far in the dialogue with this new preference (see Section 5.2). Our model of preferences consists of a set of partial CP-nets, one for each player (see Section 5.1 for details). The resulting CP-nets are then used to infer the executed trading action (if any) automatically, via well-understood principles from game theory for identifying rational behavior (Bonzon, 2007).

5.1 CP-Nets

Following Cadilhac et al. (2011), we use CP-nets (Boutilier et al., 2004) to model preferences and their dependencies. CP-nets are compatible with the kind of partial information about preferences that utterances reveal, and inference with CP-nets is com-

putationally efficient.

Just as Bayesian nets are a graphical model that exploits probabilistic conditional independence to provide a compact representation of a joint probability distribution (Pearl, 1988), CP-nets are a graphical model that exploits *conditional preferential independence* to provide a compact representation of the preference order over all outcomes. The CP-net structures the decision maker’s preferences under a *ceteris paribus* assumption: outcomes are compared, other things being equal.

More formally, let V be a finite set of variables whose combination of values determine all outcomes O . Then a *preference relation* \succeq over O is a reflexive and transitive binary relation with strict preference \succ defined as: $o \succeq o'$ and $o' \not\succeq o$. Indifference, written $o \sim o'$, means $o \succeq o'$ and $o' \succeq o$. Definition 1 defines *conditional preference independence* and Definition 2 defines *CP-nets*: the graphical component \mathcal{G} of a CP-net specifies for each variable $X \in V$ its *parent variables* $Pa(X)$ that affect the agent’s preferences over the values of X , such that X is conditionally preferentially independent of $V \setminus (\{X\} \cup Pa(X))$ given $Pa(X)$.

Definition 1 Let V be a set of variables, each variable X_i with a domain $D(X_i)$. Let $\{X, Y, Z\}$ be a partition of V . X is **conditionally preferentially independent** of Y given Z if and only if $\forall z \in D(Z)$, $\forall x_1, x_2 \in D(X)$ and $\forall y_1, y_2 \in D(Y)$, $x_1 y_1 z \succeq x_2 y_1 z$ iff $x_1 y_2 z \succeq x_2 y_2 z$.

Definition 2 $\mathcal{N}_V = \langle \mathcal{G}, \mathcal{T} \rangle$ is a **CP-net** on variables V , where \mathcal{G} is a directed graph over V , and \mathcal{T} is a set of Conditional Preference Tables (CPTs). That is, $\mathcal{T} = \{\text{CPT}(X_j) : X_j \in V\}$, where $\text{CPT}(X_j)$ specifies for each combination p of values of the parent variables $Pa(X_j)$ either $p : x_j \succ \bar{x}_j$, $p : \bar{x}_j \succ x_j$ or $p : x_j \sim \bar{x}_j$ where the $\bar{}$ symbol sets the variable to false.

We discuss below how a CP-net predicts rational action, but first we describe how CP-nets are constructed from the dialogues. In the *Settlers* corpus, preferences involve a quadruplet $(o, a, \langle r, q \rangle)$ where: o is the preference owner, a is the addressee, r is the resource and q is its quantity. So each variable in the CP-nets we construct is such a quadruplet, and for each variable the possible values are *Givable* (*Giv*), *Not Givable* (*Giv*), *Receiv-*

able (Rcv) and Not Receivable (\overline{Rcv}).

For example, the utterance *Anyone want to give me a wheat for a clay?* expresses two preferences: one for receiving wheat, represented by the variable $P_w = (A, All, <wheat, 1>)$; and given this preference, another for giving clay, represented by $P_c = (A, All, <clay, 1>)$ (where A is the name of the speaker). The corresponding CP-Net is Figure 1.

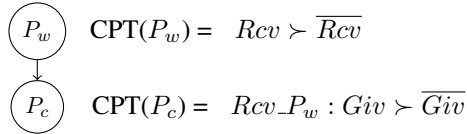


Figure 1: An example CP-net

5.2 Modeling players' preferences

As stated above, we first automatically acquire a CP-net from each EDU by using the EDU's dialogue act and the attributes (*Givable*, etc.) of its resources. We then apply the rules presented in (Cadilhac et al., 2011) to dynamically construct a preference model of the dialogue overall: this uses an equivalence between their coherence relations and our dialogue acts. Our CP-nets reasoning model handles uncertain information and noise because it use as input only the outputs of the statistical models described in Section 4, and these prior models handle uncertain information and noise. The symbolic rules for constructing CP-nets have complete coverage over any possible combination of classes that are output by the statistical models, and so they are robust. We give our rules below where π_i stands for EDU ID i .

Offers. Because an *Offer* may specify or refine an existing preference or offer, we must model how the preferences expressed in an EDU that's an *Offer* updates the prior declared preferences. So, while our annotations treat *Offer* as a property of EDUs, we treat them here as *binary relations*: $Offer(\pi_1, \pi_2)$, where the second term, π_2 , is the actual EDU whose dialogue act is *Offer* and π_1 is the set of EDUs occurring between π_2 and the last EDU uttered by the same speaker. *Offers* then have a similar effect on the CP-net as the coherence relation *Elaboration* presented in (Cadilhac et al., 2011). That is, to automatically update the CP-net constructed so far with a current EDU that's an *Offer*, the two step rule for $Offer(\pi_1, \pi_2)$ is:

1. to update the speaker's CP-net according to the preferences expressed in π_1 , and

2. if π_2 expresses preferences, to enrich the CP-net with these new preferences so that each variable in π_2 depends on each variable in π_1 .

Counteroffers. They specify or modify the terms of a previous *Offer* or *Counteroffer*. Their purpose is to give new information to refine the negotiation. Like *Offers* they must also receive a contextually dependent interpretation. The rule is quite similar to that for *Offer*; however, *Counteroffer* can modify or *correct* elements in a previously introduced offer. So for $Counteroffer(\pi_1, \pi_2)$, the rule is :

1. to *partially* update the speaker's CP-net according to the preferences expressed in π_1 which do not have the same resource type (*Givable*, *Receivable*) than the ones in π_2 .
2. same as step 2 *Offer* rule.

Accepts and Refusals. As they are answers to *Offers* and *Counteroffers*, they behave like question answer pairs (*QAPs*) presented in (Cadilhac et al., 2011). Because we are not doing full discourse parsing, we once again approximate its effects by making *Accepts* and *Refusals* respond to the set of EDUs between the current EDU and the speaker's last turn.

Accepts are positive responses to *Offers* or *Counteroffers* and are *de facto* similar to $QAP(\pi_1, \pi_2)$ where π_2 is *Yes*. Thus, the rule is, as for *Offer*, to update and enrich the CP-net.

Refusals are instead negative responses and behave like $QAP(\pi_1, \pi_2)$ where π_2 is *No*. For $Refusal(\pi_1, \pi_2)$, there is no update of the preferences expressed in π_1 . Instead, we enrich the CP-net with the *Non Givable* and *Non Receivable* information obtained from the negation of the preferences expressed in the previous *Offer* or *Counteroffer*. We then enrich the CP-net based on any new preferences expressed in π_2 . If there is a conflict between the value of a variable to be updated and the current value in the CP-net, we apply the *Correction* rule: all occurrences of the old value are replaced by the new value in π_2 .

Other. This category pertains to content that does not directly relate to trading in the game, and so we choose to ignore resources expressed in the EDUs with this dialogue act.

At the end of the negotiation dialogue, to predict exactly what trade is executed (if any), the method

checks if there are complete and reciprocal preferences expressed in the CP-nets that respectively represent the declared preferences of two agents A and B . This is done in two steps. First, we use the logic of CP-nets to determine each agent’s best outcome $bestO_A$ and $bestO_B$ from their respective CP-nets (we’ll discuss how shortly). Secondly, we compare these best outcomes: if they correspond to the same trade, we predict that this trade was executed; if not, we predict no trade is executed. Specifically, $bestO_A$ (resp. $bestO_B$) corresponds to a preference for receiving a resource r_1 from an agent B (or from all the agents indifferently) and for giving a resource r_2 to this (or these) agent(s). We predict that A gives B r_2 and B gives A r_1 if and only if: $bestO_A = Rcv(A, B, r_1) \wedge Giv(A, B, r_2)$ and $bestO_B = Rcv(B, A, r_2) \wedge Giv(B, A, r_1)$.

The first step—computing each agent’s best outcome from his CP-net—can be found in linear time using the *forward sweep algorithm* (Boutilier et al., 2004): sweep through the CP-net’s graph from top to bottom, instantiating each variable with its preferred value, given the values that are (already) assigned to its parents. This algorithm is *sound* with respect to the semantics of CP-nets.

Example. We apply this method for constructing CP-nets and determining the executed trade to the negotiation dialogue presented in Table 1.

π_1 The EDU is an *Offer*, so Rainbow’s CP-net is updated according to π_1 ’s content.

$$CPT(R, All, \langle clay, ? \rangle) = Rcv \succ \overline{Rcv}$$

π_2 It’s a *Refusal*, so we update inca’s CP-net with the negation of the preferences expressed in Rainbow’s offer.

$$CPT(I, R, \langle clay, ? \rangle) = \overline{Giv} \succ Giv$$

π_3 Idem for ariachiba.

$$CPT(A, R, \langle clay, ? \rangle) = \overline{Giv} \succ Giv$$

π_4 Idem for Kittles where the preferences expressed in this EDU are redundant with the negation of the preferences in Rainbow’s offer.

$$CPT(K, R, \langle clay, ? \rangle) = \overline{Giv} \succ Giv$$

π_5 It’s an *Offer*, so Rainbow’s CP-net is first updated according to previous EDUs (π_2 to π_4 until his last speaking), then according to the content of π_5 .

$$CPT(R, All, \langle clay, ? \rangle) = Rcv \succ \overline{Rcv} \quad (\text{inactive})$$

$$CPT(R, I, \langle clay, ? \rangle) = \overline{Rcv} \succ Rcv$$

$$CPT(R, A, \langle clay, ? \rangle) = \overline{Rcv} \succ Rcv$$

$$CPT(R, K, \langle clay, ? \rangle) = \overline{Rcv} \succ Rcv$$

$$CPT(R, All, \langle ore, ? \rangle) = \overline{Rcv}(R, I, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, A, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, K, \langle clay, ? \rangle): Rcv \succ \overline{Rcv}$$

The introduction of the preference to receive ore conflicts with the prior one for receiving clay. So the method adds to the associated CPT the label “inactive” to indicate that this is older and should be ignored if the preference about ore is satisfied.

π_6 The EDU is an *Accept*, so Kittles’s CP-net is updated according to previous EDUs (only π_5).¹

$$CPT(K, R, \langle ore, ? \rangle) = \overline{Giv}(K, R, \langle clay, ? \rangle): Giv \succ \overline{Giv}$$

π_7 The EDU is a *Counteroffer*. Since she is the last speaker, her CP-net gets updated only according to the content of the current EDU, to obtain:

$$CPT(K, R, \langle ore, ? \rangle) = \overline{Giv}(K, R, \langle clay, ? \rangle): Giv \succ \overline{Giv}$$

$$CPT(K, R, \langle wheat, ? \rangle) = \overline{Giv}(K, R, \langle clay, ? \rangle) \wedge Giv(K, R, \langle ore, ? \rangle): Rcv \succ \overline{Rcv}$$

π_8 The EDU is an *Accept*, so Rainbow’s CP-net is updated according to previous EDUs (π_6 and π_7):

$$CPT(R, K, \langle ore, ? \rangle) = \overline{Rcv}(R, I, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, A, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, K, \langle clay, ? \rangle): Rcv \succ \overline{Rcv}$$

$$CPT(R, K, \langle wheat, ? \rangle) = \overline{Rcv}(R, I, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, A, \langle clay, ? \rangle) \wedge \overline{Rcv}(R, K, \langle clay, ? \rangle) \wedge Rcv(R, K, \langle ore, ? \rangle): Giv \succ \overline{Giv}$$

π_9 It’s an *Accept* with nothing new to update.

At the end of the dialogue, these agents’ CP-nets (correctly) predict that Kittles gave ore to Rainbow in exchange for wheat.

5.3 Evaluation and results

We compare our model against four baselines. Since none of these baselines support reasoning about equilibrium moves, they all rely on the presence of an *Accept* act to predict there was a trade, and its absence to predict there wasn’t. The baselines differ, however, in how they identify the trading partners and resources in an executed trade. The **first baseline** predicts a trade according to the first *Offer* and the last person to *Accept*, and if the *Offer* doesn’t specify one of the resources then it is chosen randomly (similar random choices complete all partial predictions in all the models we consider here): e.g., for Table 1 this would predict that Kittles gave clay to Rainbow (which is incorrect) in exchange for

¹Due to lack of space, in the following CP-nets, we do not copy the inactive CPTs and CPTs about *Not Givable* or *Not Receivable* resources.

something that's chosen randomly (which will probably be incorrect). The **second baseline** uses the last *Offer* and the last person to *Accept*: e.g., for Table 1 this predicts that Kittles gave ore to Rainbow (correct) for something random (probably incorrect). The **third baseline** uses the last *Offer* or *Counteroffer*, whichever is latest, and the last person to *Accept*: e.g., for Table 1 this correctly predicts that Kittles gave ore to Rainbow in exchange for wheat. And the **fourth baseline**, uses *default unification* between the prior *Offers* or *Counteroffers* and the current one to resolve any of the current offer's elided parts and to replace specific values in prior offers with conflicting specific values in the current offer (Ehlen and Johnston, 2013). One then takes the executed trade to be the result of this unification process at the point where the last *Accept* occurs. This makes the same predictions as the third baseline for Table 1, but outperforms it in the corpus example (1) by predicting the correct and *complete* trade (i.e., Rainbow gave Kittles sheep for wheat, rather than for something random):

- (1) Rainbow: i need clay ore or wheat
 Kittles: i got wheat
 Rainbow: i cn giv sheep
 Kittles: ok

We performed the evaluation on the data presented in Sections 3 and 4: 254 dialogues in total since we ignore dialogues that contain only *Others*. 90 of these dialogues end with a trade being executed and 2 of them end with 2 trades. A random baseline would give 1.6% accuracy (given the 61 possible trading actions) and a frequency baseline (always choose no trade) gives 64.1% accuracy.

Table 4 presents the accuracy figures for all the models when calculated from the gold standard labels rather than the classifiers' predicted labels from Section 4, so that we can compare the models in isolation of the classifiers' errors. McNemar's test shows that our model significantly outperforms all the baselines ($p < 0.05$). A predicted trade counts as correct only if it specifies the right participants and the correct type of resources offered and received (we ignore their quantity). True Positives (TP) are thus examples where the model correctly predicts not only that a trade happened, but also the correct partners and resources; Wrong Positives (WP), on

the other hand, constitute a correct prediction that there was a trade but errors on the partners and/or resources involved (so WPs undermine accuracy). True Negatives (TN) are examples where the model correctly predicts there was no trade (so TPs and TNs contribute to accuracy). False Positives (FP) and False Negatives (FN) are respectively incorrect predictions that there was a trade, or that there was no trade.

While Table 4 does not reflect this, the first three baselines tend to predict *incomplete* information about the trade even when what they do predict is correct: that is, they predict the correct addressee and the owner but resort to random choice for a resource that's missing from the *Offer* or *Counteroffer* that predicts which trade occurred. For the first baseline 34 examples are like this; for the second and third baselines it's 32. In contrast, this problem occurs only once with the fourth baseline, and all the trades predicted by our method are complete, making random choice unnecessary. Moreover, the first three baselines often make incorrect predictions about the addressee or resources exchanged because in contrast to our model and the fourth baseline, they don't track how potential trades *evolve* through a *sequence* of offers and counteroffers.

Even though the fourth baseline, which uses *default unification* to track the content of the current offer, is smart and gives good results, it has statistically significant lower accuracy than our model. One major problem with the fourth baseline is that, in contrast to our model, it does not track each player's *attitude* towards the current offer. Instead, like all our baselines, it relies on the presence of an *Accept* act to predict that there's a trade.² But several corpus examples are like (2), in which a trade is executed but there's no *Accept* act, thus yielding a False Negative (FN) for all four baselines:

- (2) Joel: anyone have sheep or wheat
 Cardlinger: neither :(
 Joel: will give clay or ore
 Euan: not just now
 Jon: got a wheat for a clay
 (*Joel gives clay to Jon and receives wheat*)

²We tried a baseline that doesn't rely on the presence of an *Accept* act, but rather predicts a trade whenever *default unification* yields a complete offer. It performed worse than the fourth baseline.

So overall, our analysis shows that using CP-nets significantly outperforms all baselines that don't model how preferences evolve in the dialogue, and error analysis yields evidence that our model outperforms the fourth baseline because our model supports *reasoning* about player preferences, rational behavior and equilibrium strategies.

1st baseline: first Offer/last Accept					
TP	FP	FN	TN	WP	Accuracy
24	14	30	150	38	68.0
2nd baseline: last Offer/last Accept					
TP	FP	FN	TN	WP	Accuracy
29	6	32	158	31	73.0
3rd baseline: last (Counter)Offer/last Accept					
TP	FP	FN	TN	WP	Accuracy
39	4	23	160	30	77.7
4th baseline: default unification					
TP	FP	FN	TN	WP	Accuracy
64	4	23	160	5	87.5
Our method					
TP	FP	FN	TN	WP	Accuracy
75	4	15	160	2	91.8

Table 4: Results for trade prediction. TP, FP, FN, TN and WP are the True and False Positives, False and True Negatives and Wrong Positives.

Table 5 presents the results for the end to end evaluation, where trade predictions are made from the classifiers' output from Section 4 rather than the gold standard labels. As expected, performance decreases due to the classifiers' errors, mainly on the type of resources (*Givable*, etc.). But our method still significantly outperforms all the baselines with an accuracy of 73.4% when the baselines obtain values between 60.9% and 68.4%.

4th baseline: default unification					
TP	FP	FN	TN	WP	Accuracy
23	12	37	152	32	68.4
Our method					
TP	FP	FN	TN	WP	Accuracy
34	10	43	154	15	73.4

Table 5: Results for the end to end trade prediction.

6 Related Work

6.1 Dialogue act modeling

Most work on dialogue act modeling focuses on spoken dialogue (Stolcke et al., 2000; Fernández et al., 2005; Keizer et al., 2002). But live chats introduce specific complications (Kim et al., 2012): ill-formed data, abbreviations and acronyms, emotional indicators and entanglement (especially for multi-party chat). Among related work in this emerging field, Joty et al. (2011) use unsupervised learning to model dialogue acts in Twitter, Ivanovic (2008) and Kim et al. (2010) analyze one-to-one online chat in a customer service domain, and Wu et al. (2002) and Kim et al. (2012) predict dialogue acts in a multi-party setting. We used a similar classifier to predict dialogue acts as the one reported in (Kim et al., 2012) and evaluation yields similar results.

This paper proposes an approach to dialogue act identification in online chat that aims to predict strategic actions like bargaining. Compared to (Sidner, 1994) and DAMSL (Core and Allen, 1997), our domain level annotation is much more detailed: we not only predict moves like *Accept* but also features like the *Givable* and *Receivable* resources. Our general speech act typology of EDUs lacks intentional descriptions of speech acts, however. This reflects a conscious choice to specify the semantics of each act purely by the public commitments made to offer or to receive goods.

6.2 Preference extraction

While preference extraction from non-linguistic actions is well studied (Chen and Pu, 2004; Fürnkranz and Hüllermeier, 2011), their extraction from spontaneous conversation has received little attention. To our knowledge, the only existing work is (Asher et al., 2010; Cadilhac et al., 2011; Cadilhac et al., 2012) which we build on. Cadilhac et al. (2011) compute CP-nets from *coherence relations*, found in the annotation of the Verbmobil corpus (Baldrige and Lascarides, 2005). Here we adapt their algorithm from coherence relations to *unary* dialogue acts. Further, while they assume that preferences are given, here we apply versions of the NLP techniques from Cadilhac et al (2012) to *estimate* the preferences of EDUs automatically. And we go further than any of these works by using the elicited pref-

ferences to infer the domain-level actions that result from information exchanged in the conversation.

In this respect, our work relates to models for *grounding language*, where semantic parsing techniques are used to automatically map linguistic instructions to domain-level actions (Artzi and Zettlemoyer, 2013; Kim and Mooney, 2013). Our domain of application is more challenging, however: to our knowledge, this is the first attempt to map *non-cooperative* dialogues into predictions about domain-level actions. We can tackle these strategic scenarios because we exploit a logic of preferences as part of our model, yielding inferences about rational action even when agents’ preferences conflict.

Compared to previous work, our task is new. Our aim is not to predict what dialogue act to perform next, but what non verbal action should be performed, mapping dialogue acts to non verbal actions. The difference between our work and other work on grounding is that we are grounding non-cooperative dialogue rather than instructions in a cooperative setting. There is no prior work of which we’re aware that maps a non-cooperative dialogue into a prediction about which joint non-verbal action the agents will do as a result of what they’ve learned about their opponent through conversation. Furthermore, both the CP-net and the fourth baseline, whose accuracy is quite high (making it a hard baseline to beat), use the dialogue history as they incrementally build up the preference model.

6.3 Predicting strategic actions

Modeling player behavior in real-time strategy games is a growing research area in AI. These models can be used to identify common strategic states, discover new strategies as they emerge or predict an opponents future actions and so help players to optimize their choices. For example, Schadd et al. (2007) develop a hierarchical opponent model in the game *Spring*, Dereszynski et al. (2011) reason about strategic behavior in *StarCraft* using hidden Markov models and Amato and Shani (2010) use reinforcement learning to acquire a policy for switching among high-level strategies in *Civilization IV*.

In comparison, we propose a novel approach for predicting strategic action based on the *symbolically* formalized preferences that each agent commits to in spontaneous conversation. Our approach thus deals

with imperfect information by exploiting the agents’ declared preferences. By predicting what bargain (if any) will take place, we are able to verify the correctness of our preference descriptions. Our task is a subtask of learning a strategy over an entire game space, but our approach yields good predictive results on relatively little data—an advantage of exploiting CP-nets and the symbolic rules that guide their evolution from observable evidence.

7 Conclusion

We have proposed a linguistic approach to strategy prediction in spontaneous conversation, exploiting dialogue acts to build a partial model of the agents’ declared preferences. Our method tracks how preferences evolve during the dialogue, which we use to infer their bargaining behavior, i.e. what resources, if any, are exchanged, and by whom.

We based our study on a corpus collected using an online version of *The Settlers of Catan*. Negotiations in this game mirror complex real life negotiations and provide a fruitful arena to study strategic conversation. Evaluation shows that our approach provides more accurate and complete information about trades than baselines that don’t track how an offer evolves through the dialogue, and we also argued that game-theoretic reasoning about rational behavior has advantages over relying on the presence or absence of an *Accept* act to make predictions.

Our approach, however, does not exploit discourse structure, which is needed to properly handle long distance dependencies of offers on prior material. We will exploit this in future work to improve our results. We also plan to investigate other aspects of strategic reasoning on a larger dataset.

We have proposed a method that relies on a typology of dialogue acts that is domain sensitive. However, in other work we have shown how to adapt our algorithms to several domains (Cadilhac et al., 2012). In future work, we plan to link our preference extraction algorithms to an automatically acquired discourse structure for a given text. This will provide a domain independent means for extracting preferences from dialogue.

Acknowledgments

This work is supported by ERC grant 269427 STAC.

References

- Stergos Afantenos, Nicholas Asher, Farah Benamara, Anaïs Cadilhac, Cédric Dégremont, Pascal Denis, Markus Guhe, Simon Keizer, Alex Lascarides, Oliver Lemon, Philippe Muller, Soumya Paul, Verena Rieser, and Laure Vieu. 2012. Developing a corpus of strategic conversation in the settlers of catan. In *Proceedings of the 1st Workshop on Games and NLP (GAMNLP-12)*.
- Christopher Amato and Guy Shani. 2010. High-level reinforcement learning in strategy games. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'10)*, pages 75–82.
- Yoav Artzi and Luke Zettlemoyer. 2013. Weakly supervised learning of semantic parsers for mapping instructions to actions. *Transactions of the Association for Computational Linguistics*, 1:49–62.
- Nicholas Asher, Elise Bonzon, and Alex Lascarides. 2010. Extracting and modelling preferences from dialogue. In *IPMU*, pages 542–553.
- Jason Baldridge and Alex Lascarides. 2005. Annotating discourse structures for robust semantic interpretation. In *Proceedings of the 6th IWCS*.
- Farah Benamara, Baptiste Chardon, Yannick Mathieu, and Vladimir Popescu. 2011. Towards context-based subjectivity analysis. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 1180–1188, Chiang Mai, Thailand.
- Elise Bonzon. 2007. *Modélisation des interactions entre agents rationnels : les jeux booléens*. PhD thesis, Université Paul Sabatier, Toulouse.
- Craig Boutilier, Craig Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. 2004. Cp-nets: A tool for representing and reasoning with conditional *ceteris paribus* preference statements. *Journal of Artificial Intelligence Research*, 21:135–191.
- Anaïs Cadilhac, Nicholas Asher, Farah Benamara, and Alex Lascarides. 2011. Commitments to preferences in dialogue. In *Proceedings of SIGDIAL*, pages 204–215. ACL.
- Anaïs Cadilhac, Nicholas Asher, Farah Benamara, Vladimir Popescu, and Mohamadou Seck. 2012. Preference extraction from negotiation dialogues. In *European Conference on Artificial Intelligence (ECAI)*, pages 211–216. IOS Press.
- Li Chen and Pearl Pu. 2004. Survey of preference elicitation methods. Technical report.
- Mark G. Core and James F. Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *Working Notes of the AAI Fall Symposium on Communicative Action in Humans and Machines*.
- Ethan W. Dereszynski, Jesse Hostetler, Alan Fern, Thomas G. Dietterich, Thao-Trang Hoang, and Mark Udarbe. 2011. Learning probabilistic behavior models in real-time strategy games. In *AIIDE*.
- Patrick Ehlen and Michael Johnston. 2013. A multi-modal dialogue interface for mobile local search. In *IUI Companion*, pages 63–64.
- Raquel Fernández, Jonathan Ginzburg, and Shalom Lappin. 2005. Using machine learning for non-sentential utterance classification. In *Proceedings of the 6th SIGDial Workshop on Discourse and Dialogue*, pages 77–86.
- Johannes Fürnkranz and Eyke Hüllermeier, editors. 2011. *Preference Learning*. Springer.
- Edward Ivanovic. 2008. Automatic instant messaging dialogue using statistical models and dialogue acts. In *Masters thesis, The University of Melbourne*.
- Shafiq R. Joty, Giuseppe Carenini, and Chin-Yew Lin. 2011. Unsupervised modeling of dialog acts in asynchronous conversations. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 1807–1813.
- Simon Keizer, Rieks op den Akker, and Anton Nijholt. 2002. Dialogue act recognition with bayesian networks for dutch dialogues. In *Proceedings of the 3rd SIGDial Workshop on Discourse and Dialogue*, pages 88–94. Association for Computational Linguistics.
- Joohyun Kim and Raymond J. Mooney. 2013. Adapting discriminative reranking to grounded language learning. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL-2013)*, Sofia, Bulgaria.
- Su Nam Kim, Lawrence Cavedon, and Timothy Baldwin. 2010. Classifying dialogue acts in 1-to-1 live chats. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 862–871.
- Su Nam Kim, Lawrence Cavedon, and Timothy Baldwin. 2012. Classifying dialogue acts in multi-party live chats. In *26th Pacific Asia Conference on Language, Information and Computation*, pages 463–472.
- Dan Klein and Christopher D. Manning. 2003. Accurate unlexicalized parsing. In *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pages 423–430.
- Martin Osborne and Ariel Rubinstein. 1994. *A Course in Game Theory*. MIT Press.
- Judea Pearl. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Leonard Savage. 1954. *The Foundations of Statistics*. John Wiley.

- Frederik Schadd, Sander Bakkes, and Pieter Spronck. 2007. Opponent modeling in real-time strategy games. In *Games and Simulation GAMEON*, pages 61–68.
- Yoav Shoham and Kevin Leyton-Brown. 2009. *Multia-gent Systems: Algorithmic, Game-Theoretic and Logical Foundations*. Cambridge University Press.
- Candace Sidner. 1994. An artificial discourse language for collaborative negotiation. In *AAAI*, volume 1, pages 814–819. MIT Press, Cambridge.
- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol V. Ess-dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. In *Computational Linguistics*, pages 26:339–373.
- Tianhao Wu, Faisal M. Khan, Todd A. Fisher, Lori A. Shuler, and William M. Pottenger. 2002. Posting act tagging using transformation-based learning. In *Foundations of Data Mining and knowledge Discovery*, pages 319–331. Springer.