THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Revealing Resources in Strategic Contexts

Citation for published version:
Perret, J, Afantenos, S, Asher, N & Lascarides, A 2014, Revealing Resources in Strategic Contexts. in Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL). pp. 135-144.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)

Revealing Resources in Strategic Contexts

Jérémy Perret
IRIT,
Univ. Toulouse, France

Stergos Afantenos
IRIT,
Univ. Toulouse, France

Nicholas Asher
IRIT, CNRS,
France

{perret,stergos.afantenos,asher}@irit.fr

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

Abstract

Identifying an optimal game strategy often involves estimating the strategies of other agents, which in turn depends on hidden parts of the game state. In this paper we focus on the win-lose game *The Settlers of Catan* (or *Settlers*), in which players negotiate over limited resources. More precisely, our goal is to map each player's utterances in such negotiations to a model of which resources they currently possess, or don't possess. Our approach comprises three subtasks: (a) identify whether a given utterance (dialogue turn) reveals possession of a resource, or not; (b) determine the type of resource; and (c) determine the exact interval representing the quantity involved. This information can be exploited by a *Settlers* playing agent to identify his optimal strategy for winning.

1 Introduction

When resources are limited, there is a fine line between agents cooperating and competing with one another for those resources, especially in a win-lose game. The goal of every rational agent is to maximize his *expected utilities* by finding *equilibrium strategies*: that is, an action sequence for each player that is optimal in that no player would unilaterally deviate from his action sequence, assuming that all the other players perform the actions specified for them (Yoam Sholam and Kevin Leyton-Brown, 2009). Calculating equilibrium strategies thus involves reasoning about what's optimal for the other players, which in turn depends on which resources they possess and which resources they need. However, almost every kind of bargaining game occurs in a context of imperfect information (Osborne and Rubinstein, 1994), where the opponent's current resources are hidden or non-observable.

Indeed, imperfect information often results from deliberate obfuscation: if an opponent can accurately identify your resources then they can exploit it for their own strategic advantage. For instance, in *The Settlers of Catan* (or *Settlers*), our chosen domain of investigation here, Guhe and Lascarides (2014) develop a *Settlers* playing agent where game simulations show that making the agent omniscient about his opponents' resources enables him to achieve more successful negotiations (i.e., a significantly higher proportion of his trade offers are accepted) and a significantly higher win rate than his non-omniscient counterparts. So it is rational for players to balance achieving their desired trades with revealing as little as possible about their own resources, while at the same time attempting to elicit information about their opponents' resources.

In negotiations using natural language dialogue, eliciting information about an opponent's resources is often realized as a question; the opponent, on realizing the question's purpose, often avoids revealing their resources in their response. They use various communicative strategies to achieve this effect, such as making a counteroffer, being vague, or simply changing the subject.

In this paper, we are interested in determining how players can extract information about an opponent's resources from what they say during negotiation dialogues. In order to study how people commit, or don't commit, to the resources they have (or don't have), we have used a corpus of negotiation dialogues that take place during the win-lose game *The Settlers of Catan* in order to learn a statistical model that maps the utterances of the players to their commitments concerning the kind and number of resources they possess. In section 2 we describe our corpus in detail, as well as the phenomena that we are trying to capture. In section 3 we describe the annotation procedure that we have followed in order to obtain training and testing datasets. Section 4 describes the experiments we have performed and the results we have obtained. Section 5 describes the related work and conclusions and future work are in section 6.

2 The Corpus

Our model is trained on an existing corpus (see Afantenos et al. (2012)) of humans playing an online version of the game *The Settlers of Catan* (or *Settlers*, Teuber (1995); `www.catan.com`). *Settlers* is a win-lose game board game for 2 to 4 players. Each player acquires resources (ore, wood, wheat, clay, sheep) and uses them to build roads, settlements and cities. This earns Victory Points (VPs); the first player with 10 VPs wins. Players can acquire resources via the dice roll that starts each turn and through trading with other players—so players converse to negotiate trades. A player's decisions about what resources to trade depends on what he wants to build; e.g., a road requires 1 clay and 1 wood. Players can also lose resources: a player who rolls a 7 can rob from another player and any player with more than 7 resources must discard half of them. What's robbed or discarded is hidden, so players lack complete information about their opponents' resources. Consequently, agents can, and frequently do, engage in 'futile' negotiations that result in no trade (i.e., they miscalculate the equilibria).

Players in the corpus described in Afantenos et al. (2012) must chat in an online interface in order to negotiate trades, and each move in the chat interface is automatically aligned with the current game state—so one can compare what an utterance reveals about possessed resources with what the speaker actually possesses, and so identify examples of obfuscation (e.g., see Table 1). The corpus consists of 59 games, and each game contains dozens of individual negotiation dialogues, each dialogue consisting of anywhere from 1 to over 30 dialogue turns. In our experiments, we have used 7 games consisting of more than 2000 dialogue turns (see Section 3).

Table 1 contains an excerpt from one of the dialogues. In turn 157 the player "gotwood4sheep" asks if anyone has any wood, implying that he wants to negotiate an exchange of resources where he receives wood. Player "ljaybrad123" is the first to reply, negatively, implicating that he has no wood. Turn 158 is thus annotated with the information that the player "ljaybrad123" is revealing that he has 0 wood.[1] In turn 159 player "gotwood4sheep" persists in his attempt to negotiate, referring directly to player "tomas.kostan" and making a more specific trade offer, of ore in exchange for wood. He has thus revealed that he possesses at least one ore. The player "tomas.kostan" acknowledges that he has wood (so this turn is annotated with the information that "tomas.kostan" has at least one wood) but that this resource is important to him. "tomas.kostan" then proposes 2 ore in exchange for 1 wood (again, this turn is annotated with the information that "tomas.kostan" possesses at least one wood). "gotwood4sheep" in turn 162 explicitly says that he has only one ore and not two, so this turn is annotated with the information that player "gotwood4sheep" has exactly 1 ore. In the end the negotiation fails since for "tomas.kostan" a wood is currently worth more to him than what "gotwood4sheep" is currently offering.

Note that revealed resources depend not only on the content of the individual utterance but also on its semantic connection to the discourse context. For example, the dialogue turn 158 (*no*) reveals nothing about resources on its own; it is the fact that it is connected to the question 157 with a QAP (Question-Answer-Pair) relation that commits "ljaybrad123" to having 0 wood. Similarly, 160 is an Acknowledgment to 159 and so reveals that "tomas.kostan" possesses at least one wood.

3 Annotations

The corpus has been annotated with information at multiple levels, including dialogue boundaries, turns within dialogues, speech acts (offers, counteroffers, refusals, etc.), as well as discourse relations following SDRT (Asher and Lascarides, 2003). Full details are in Afantenos et al. (2012); here we provide in Tables 2 and 4 statistics of only

[1] In this paper, we simplify our task by ignoring the fact that players can lie. As matter of fact, manual analysis of the corpus logs show that players rarely lie concerning their resources, preferring instead to conceal relevant information by avoiding giving a direct answer.

Dialogue turn	Player	Utterance
157	gotwood4sheep	anyone got wood?
158	ljaybrad123	no
159	gotwood4sheep	ore for a wood, tomas?
160	tomas.kostan	yes but i need mine
161	gotwood4sheep	ore more?
162	tomas.kostan	2 ore for a wood?
163	gotwood4sheep	i don't have 2, sorry, just the one
164	gotwood4sheep	early doors, early offers :)
165	tomas.kostan	then i cannot make you a deal
166	tomas.kostan	sry
167	gotwood4sheep	ah dommage :(

Table 1: Excerpt from a dialogue

Number of speech turns in dialogue	Dialogue count
1-5	112
6-10	63
11-15	23
16-20	13
21 and more	23

Table 2: Dialogue statistics

the relevant annotations that we used to train our models. Our model mostly exploits the QAP and Q-Elab relations to infer revealed resources; see Section 4 for details, including the performance of our trained model for identifying discourse relations.

We manually annotated each utterance with its corresponding revealed resource. Two of this paper's authors were involved in this annotation effort. After a thorough examination of the dialogues in an initial game, they settled on the format of the annotations and the guide for performing the annotation task. The annotation format is as follows. Each speech turn corresponding to a revealed resource is annotated with a pair: a resource name, and the quantity interval which the player reveals, representing the lower and upper bound of the resource. For example, in dialogue turn 158 of table 1 player "ljaybrad123" declares that he has no wood, so this dialogue turn is annotated as **`(wood, [0,0])`**. In dialogue turn 159 player "gotwood4sheep" reveals he has at least one ore, so this turn is annotated as **`(ore, [1,+∞])`**. Revelations of multiple resources are associated with multiple pairs.

Data counts	
Number of games	7
Speech turns	2460
Relation count by type	
Question-answer pair	687
Comment	443
Continuation	250
Acknowledgement	230
Result	182
Q-Elab	161
Elaboration	150
Contrast	140
Explanation	79
Clarification question	52
Narration	43
Alternation	42
Correction	41
Parallel	40
Conditional	32
Background	19

Table 3: Annotation of discourse relations

To test the consistency and difficulty of the task, both annotators independently annotated a single game after settling on the above format and instructions for annotation. Over 422 speech turns, the resulting kappa coefficient of inter-annotator agreement is **0.94**, enough to validate our annotation method. The remaining 6 games were then

Speech turns	2201
Dialogues	263
Word count	9121
Turns revealing resources	452 (21% of turns)

Table 4: Dataset overview

annotated, for which statistics can be found in tables 4 and 2. Most dialogues appear to be short, frequently consisting of comments on the game status, which do not call for answers. Trade negotiations are usually longer, with player emitting offers and counteroffers, sometimes competitively. Revelations of resources are present in 21% of dialogue turns.

4 Experiments and Results

4.1 Formulating the problem

As mentioned earlier, our goal is to predict whether a given turn reveals that its emitter possess a resource, and if so the type of the resource and its quantity in the form of an interval. Although players could potentially reveal having a specific number of resources (e.g., line 163 in table-1), in most cases the players reveal either having zero resources (interval $[0, 0]$) or having at least one (interval $[1, \infty]$), and in few occasions, players reveal that they have more than one (interval $[2, \infty]$) or exactly two resources ($[2, 2]$). In most of the cases, a revelation of having zero resources is manifested through the player rejecting a trade offer by stating that they don't have the resource desired by their opponent.

Using a single classifier to predict from an NL string the revelation of a particular type of resource, or no revelation of any resource, would involve classifying each utterance into 6 classes: one for each of the 5 types of resources, and one for revealing that no resources are possessed. But such a model would fail to take full advantage of the following facts. First, the NL strings that reveal a resource are relatively invariant, save for the particular resource type; in other words, the ways in which people talk about their possession of clay is the same as their talk about possessing wood, save for the words "clay" vs. "wood". Secondly, it is easy to specify the properties of a revelation (both the type of resource and quantity) when we know a given utterance exhibits a revelation. Given these observations, we decided to divide the prediction process into two subtasks:

- Determine if a given speech turn reveals a resource or not;
- For those utterances that do reveal a possessed resource, determine the type of resource and its associated quantity interval.

4.2 Features

Our goal was to learn a function

$$f : \mathcal{X} \mapsto \{0, 1\}$$

where every $\mathbf{x} \in \mathcal{X}$ corresponds to a vector representing a dialogue turn and $\{0, 1\}$ represents the fact that there is a revelation concerning an underspecified resource from the part of the dialogue act emitter.

The features that we have extracted for every dialogue turn can be summarized in the following categories:

- Contextual features: positioning of the turn in the dialogue;
- Lexical features: single words present in the utterance;
- Pattern-related features: recurring speech structures associated with revealed resources;
- Relational features: discourse relationships with other turns.

These features are listed more extensively in Table 5. Non-relational features are extracted directly from the underlying text. In order to compute the relational features—essentially whether a pair of dialogue turns are linked with a *Question-answer pair* (QAP) or a *Question-Elaboration* (Q-Elab) discourse relation—we used the results of a separate classifier for the prediction of discourse relations. This classifier was trained on 7 games consisting of 2460 dialogue turns. We used a MaxEnt model, as in the case of predicting revealed resources (see below for more details). We selected, for this classifier, a subset of the feature set used for the task of predicting revealed resources. More specifically, we used only the *Contextual* and *Lexical* features shown in Table 5. Although the model we have used was a general one, capable of predicting the full set of discourse relations listed in Table 3, for this series of experiments we were only interested in the QAP and Q-Elab relations. Results for these relations are shown in Table 6.

Category	Description
Contextual	Speaker initiated the dialogue
Contextual	First utterance of the speaker in the dialogue
Contextual	Position in dialogue
Lexical	Contains resource name
Lexical	Ends with exclamation mark
Lexical	Ends with interrogation mark
Lexical	Contains possessive pronouns
Lexical	Contains modal modifiers
Lexical	Contains question words
Lexical	Contains a player's name
Lexical	Contains emoticons
Lexical	First and last words
Pattern-related	Contains a possession structure, such as *I have (no) X*
Pattern-related	Contains a query structure, such as *I need X*
Pattern-related	Contains *X for Y*
Relational	Is predicted as question wrt another speech turn
Relational	Is predicted as answer wrt another speech turn

Table 5: Feature set description

Question-answer pair		
Precision	Recall	F1 score
83.8	86.8	85.3
Q-Elab		
Precision	Recall	F1 score
53.3	57.9	55.5

Table 6: Results for the relation prediction task.

4.3 Statistical model

For our classifier, we used a regularized maximum entropy (MaxEnt, for short) model (Berger et al., 1996). In MaxEnt, the parameters of an exponential model of the following form are estimated:

$$P(b|t) = \frac{1}{Z(c)} \exp\left(\sum_{i=1}^{m} w_i f_i(t,c)\right)$$

where t represents the current dialogue turn and c the outcome (i.e., revelation of a resource or not). Each dialogue turn t is encoded as a vector of m indicator features f_i (see table 5 for more details). There is one weight/parameter w_i for each feature f_i that predicts its classification behavior. Finally, $Z(c)$ is a normalization factor over the different class labels (in this case just two, whether we have a revelation of a resource or not), which guarantees that the model outputs probabilities.

In MaxEnt, the values for the different parameters $\hat{w}$ are obtained by maximizing the log-likelihood of the training data T with respect to the model (Berger et al., 1996):

$$\hat{w} = \underset{w}{\operatorname{argmax}} \sum_{i}^{T} \log P(c^{(i)}|t^{(i)})$$

Various algorithms have been proposed for performing parameter estimation (see (Malouf, 2002) for a comparison). Here, we used the Limited Memory Variable Metric Algorithm implemented in the MegaM package.[2] We used the default regularization prior that is used in MegaM.

4.4 Predicting the type and quantity of revealed resource

From our observations, the majority of utterances revealing resources fall into one the following two categories:

- Self-contained: resource and quantity can be deduced from the utterance alone, such as *I have no ore*;

[2]Available from `http://www.cs.utah.edu/~hal/megam/`.

Type	Keywords
Negation	no, not, don't
Second-person	you, someone, anyone
Possession	got, have, give, spare, offer
Query	want, need, get
For	for

Table 7: Markers used in type prediction

- Contextual: some information is deduced from another utterance. Both usually form a question-answer pair, such as *Do you have any wheat ? – Yes.*

We created five marker categories, described in Table 7, from the most frequent words appearing in revealing utterances. We designed a rule-based model using these markers; their combination allows us to pinpoint where the resource the player reveals is mentioned. For example, in the utterance *anyone has sheep for ore?*, the second-person marker *anyone* and the possession marker *has* indicate that the first mentioned resource is the one wanted by the player, which he doesn't reveal as possessing. Moreover, the presence of a *for* marker indicates that the players offers a resource. Hence, the resource following the marker, *ore*, is possessed by the player.

Such a rule system allows us to analyze a single utterance. However, in the case of a QAP, we often fail to retrieve data from the answer utterance alone. A second pass is thus performed on the question utterance, giving us enough context to deduce revealed resources. For example, in the QAP *anyone have wood ? – none, sorry*, in the second utterance, the negation marker *none* implies the absence of an unknown resource. The processing of the first utterance reveals that *wood* is requested by another player. We conclude that the answering players possess no wood.

We first tested our rule model on reference data, knowing exactly (from the annotations) which speech turns contain revealed resources, and which discourse relations link them. We then used the model on predicted data (discourse relations as well as dialogue turns representing revealed resources), effectively creating a full end-to-end system.

Baseline (accuracy : 82.1)			
	Precision	Recall	F1 score
H_+	54.7	73.7	.628
H_-	92.5	84.2	.882
Our method (accuracy : 89.2)			
	Precision	Recall	F1 score
H_+	75.2	70.6	.728
H_-	95.2	94.0	.933

Table 8: Results for the task of deterring whether a turn reveals a resource. H_+ represents the hypothesis that the dialogue turn does reveal a resource, while H_- the hypothesis that it doesn't.

4.5 Results

The classifier was trained using 10-fold cross-validation. For every training round, we partition the data by dialogues. With the speech turns belonging to 90% of them, we form the model training set. The turns from the remaining 10% are used as test data. We compared our method to a baseline, which does not involve machine learning. This naive model predicts revealed resource whenever a resource is mentioned by name in the utterance.

After performing ten rounds of cross-validation on the training data, we achieve a F1 score of **0.72** for the positive hypothesis *"This speech turn reveals a resource"*. The opposite class ("There is no revealed resource in this turn") has an F1 score of 0.93, achieving thus a global accuracy of 89.2%. Detailed results for our model and baseline are shown in Table 8.

Results for the prediction of resource type quantity interval are shown in table 9. As we can see, prediction of the type of resource that a player's dialogue turn reveals has an accuracy of 77% on the manually annotated instances, which falls down to 61.5% when using the results of the first classifier as input. Interval prediction on the other hand has has an accuracy of 79.9% when using manually annotated results which falls down to 65.7% when using the results of the first classifier as input. Note as well that we have implemented a baseline for both systems. Concerning resource type, the baseline randomly attributes a resource to utterances labeled as revealing one. The baseline for interval prediction assigns the most frequent interval. Results are also shown in table 9.

In table 10 we report results on the pipeline combining the three tasks. The accuracy of 57.1% does not include the instances that have been classified as not revealing any resources by the first classifier. When we evaluate both classes the accuracy goes up to 86.3%.

Accuracy	on manual annotations	on the output of the first classifier
	Baseline	
Resource type	0.165	0.146
Interval	0.559	0.328
	Our method	
Resource type	0.770	0.615
Interval	0.799	0.657

Table 9: Baseline and evaluation of predicting resource type and interval.

	Accuracy
On all instances	0.863
Only on instances classified as revealing a resource	0.571

Table 10: Results of the pipeline, that is prediction of the exact triplets **`(resource, [lower bound, upper bound])`**.

4.6 Discussion

The first step of our prediction process, locating turns revealing resources, yields very encouraging results (see Table 8): we are able to retrieve such turns with an F1 score of over 0.72, while they represent only 21% of all speech turns. On the other hand our system does not perform very well on the detection of resource type as well as the associated interval. This is to be expected: since we have split our system in three parts, there is error propagation in the pipeline. On the other hand jointly predicting the triplets is not a viable solution either, since this would lead to a great number of classes (six as we have mentioned above, multiplied by all the possible values for lower and upper bounds). We would like though to note that we greatly outperform both baselines for each of the last two tasks.

One way to improve the quality of our prediction would be to add more relational features. As context plays a critical part in determining the meaning of an utterance, features associated to its relational neighbors should be taken into account. This is true for the prediction of whether a dialogue turn reveals a resource as well as for the prediction of its type.

Accuracy for this last task is not very satisfying. The main reasons for this, which can serve as the basis for future improvements, include:

- Ambiguous *for* patterns. The utterance ***X*** *for* ***Y*** can be interpreted two ways : either as a revealing possession of **X** or **Y**. This is ambiguous even for the players themselves since often they pose a clarification question. Observation shows that the latter (possession of **Y**) is more frequent. The rule model implements this behavior as default when encountering such a pattern. In actual dialogues, this ambiguity is resolved by a follow-up question (*Which one are you offering ?*) or by the game context (dice rolls and resource distribution) which we haven't access to.

- Long-distance resource anaphora. On most trade negotiations, the resource being traded isn't mentioned by name at every point of the discussion, but rather referred to implicitly. When this carries over several speech turns, it becomes increasingly difficult to determine the traded resource (solving the anaphora) from a later utterance. Incorporating anaphora resolution could definitively improve our results.

- Uncommon idioms. Some utterances, such as *I'm oreless*, or *I just discarded all of my sheep*, employ rare vocabulary (with respect to the corpus) to describe resource possession. Incorporating more lexical information is necessary.

5 Related Work

Work on dialogue has traditionally focused on spoken dialogue and especially on the modeling of spoken dialogue acts (Stolcke et al., 2000; Bangalore et al., 2006; Fernández et al., 2005; Keizer et al., 2002). Recently a growing interest has emerged in working with written dialogues which can take the form either of a synchronous communication (two or multiparty live chats) or asynchronous communication (fora, email exchanges, etc). (Joty et al., 2013) are focused on the detection and labeling of topics within asynchronous discussions, more specifically email exchanges and blogs, using unsupervised methods. (Tavafi et al., 2013) are focused on the supervised learning of dialogue acts in a broad range of domains including both synchronous and asynchronous communication. They use a multi-class SVM approach as well as two structured prediction approaches (SVM-HMM and CRFs). (Wu et al., 2002) are interested in the prediction of dialogue acts in a multi-party setting. (Joty et al., 2011) focus on the modeling of dialogue acts in asynchronous discussions (emails and fora) using unsupervised approaches. Finally, (Kim et al., 2012) are interested in the classification of dialogue acts in multi-party live chats, using a naive Bayes classifier.

Revealing a resource can be viewed as a commitment by a player that she possesses a specific resource. Public commitments have been extensively studied from a theoretical point of view in linguistics (Asher and Lascarides, 2008a; Asher and Lascarides, 2008b; Lascarides and Asher, 2009) or elsewhere (Prakken, 2005; Bentahar et al., 2005; Chaib-draa et al., 2006; Prakken, 2006; El-Menshawy et al., 2010). As far as automatic detection of public commitments is concerned, in either synchronous or asynchronous conversation, to the best of our knowledge this is the first work to explore this issue. The closest work to our own is that of (Cadilhac et al., 2013) who use the live chats from the game of *The Settlers of Catan* as well. It is concerned with the detection of dialogue acts, the detection of the resources that are givable and receivable, as well as the predictions of players' strategic actions via the use of CP-nets.

6 Conclusions

Developing a strategy in any kind of game requires reasoning about the opponents' strategies. In a win-lose game, such as the board game *Settlers* on which our experiments were based, a crucial ingredient in reasoning about everyone's strategies, including one's own, is beliefs about what resources each player possesses. Information about their resources can be inferred from observable non-verbal actions, such as the dice roll that starts each turn. Here, we provided a model for inferring information about possessed resources from verbal actions in a non-cooperative setting, where players have an incentive to conceal such information.

Our model divided the task into a three subtasks: (a) first, identify whether a dialogue turn reveals the speaker to possess a specific resource, or not; and, if so (b) identify the type of that resource, and (c) its quantity. We addressed task (a) using a statistical model of logistic regression, achieving overall accuracy of 89.2% with an F-score for the positive class (the turn reveals possession information) of 72.8% (in spite of this class comprising only 21% of the data). Our prediction of the type resource possessed and its quantity was achieved through a symbolic model, since the number of classes (5 resources, unlimited quantity intervals) makes training on the available data too sparse. While there is clearly room for improvement (61.5% accuracy on resource type; 65.7% accuracy on their quantity), our models beat a random baseline for estimating the resource type and the frequency baseline for predicting its quantity.

As we mentioned earlier, game simulations using an existing *Settlers* agent from Guhe and Lascarides (2014) show that the agent benefits if all the players' resources are made observable to him. But that's not the realistic scenario for this game, and the *Settlers* agent from (Guhe and Lascarides, 2014) for whom resources aren't made observable doesn't use the negotiation dialogues as any evidence at all about possessed resources. Instead, the agent relies only on dice rolls, build actions and robbing to update his beliefs, and so by ignoring conversation the agent can miss crucial evidence for who has what. In future work, we plan to enhance the belief model of the *Settlers* agent from Guhe and Lascarides (2014) to exploit our (noisy) model for mapping conversation to the players' resources, and evaluate whether this richer source of evidence for inferring the hidden aspects of the game state improves the agent's performance, both for successfully negotiating and winning the overall game.

References

Stergos Afantenos, Nicholas Asher, Farah Benamara, Anaïs Cadilhac, Cédric Degremont, Pascal Denis, Markus Guhe, Simon Keizer, Alex Lascarides, Oliver Lemon, Philippe Muller, Soumya Paul, Vladimir Popescu, Verena Rieser, and Laure Vieu. 2012. Modelling strategic conversation: model, annotation design and corpus. In *Workshop on the Semantics and Pragmatics of Dialogue, Paris, France*. Université Paris 7, septembre.

Nicholas Asher and Alex Lascarides. 2003. *Logics of*

Conversation. Studies in Natural Language Processing. Cambridge University Press, Cambridge, UK.

Nicholas Asher and Alex Lascarides. 2008a. Commitments, beliefs and intentions in dialogue. In *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue (Londial)*, pages 35–42, London.

Nicholas Asher and Alex Lascarides. 2008b. Making the right commitments in dialogue. In *Workshop on Implicatures*. University of Michigan Linguistics and Philosophy.

Srinivas Bangalore, Giuseppe Di Fabbrizio, and Amanda Stent. 2006. Learning the structure of task-driven human-human dialogs. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 201–208. Association for Computational Linguistics.

Jamal Bentahar, Bernard Moulin, and Brahim Chaib-draa. 2005. Specifying and implementing a persuasion dialogue game using commitments and arguments. In Iyad Rahwan, Pavlos Moratis, and Chris Reed, editors, *Argumentation in Multi-Agent Systems*, volume 3366 of *Lecture Notes in Computer Science*, pages 130–148. Springer Berlin Heidelberg.

A. Berger, S. Della Pietra, and V. Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1):39–71.

Anais Cadilhac, Nicholas Asher, Farah Benamara, and Alex Lascarides. 2013. Grounding strategic conversation: Using negotiation dialogues to predict trades in a win-lose game. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 357–368, Seattle, Washington, USA, October. Association for Computational Linguistics.

Brahim Chaib-draa, Marc-Andr Labrie, Mathieu Bergeron, and Philippe Pasquier. 2006. Diagal: An agent communication language based on dialogue games and sustained by social commitments. *Autonomous Agents and Multi-Agent Systems*, 13(1):61–95.

Mohamed El-Menshawy, Jamal Bentahar, and Rachida Dssouli. 2010. Modeling and verifying business interactions via commitments and dialogue actions. In Piotr Jdrzejowicz, NgocThanh Nguyen, RobertJ. Howlet, and LakhmiC. Jain, editors, *Agent and Multi-Agent Systems: Technologies and Applications*, volume 6071 of *Lecture Notes in Computer Science*, pages 11–21. Springer Berlin Heidelberg.

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.

Markus Guhe and Alex Lascarides. 2014. Game strategies in the settlers of catan. In *Proceedings of the IEEE Conference in Computational Intelligence in Games (CIG)*, Dortmund.

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.

Shafiq Joty, Giuseppe Carenini, and Raymond Ng. 2013. Topic segmentation and labeling in asynchronous conversations. *Journal of AI Research (JAIR)*, 47:521–573.

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, Philadelphia, July. Association for Computational Linguistics.

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.

Alex Lascarides and Nicholas Asher. 2009. Agreement, disputes and commitments in dialogue. *Journal of Semantics*, 26(2):109–158.

Robert Malouf. 2002. A comparison of algorithms for maximum entropy parameter estimation. In *Proceedings of the Sixth Workshop on Natural Language Learning*, pages 49–55, Taipei, Taiwan.

M. J. Osborne and A. Rubinstein. 1994. *A Course in Game Theory*. MIT Press.

Henry Prakken. 2005. Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation*, 15(6):1009–1040.

Henry Prakken. 2006. Formal systems for persuasion dialogue. *The Knowledge Engineering Review*, 21:163–188, 6.

Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–373.

Maryam Tavafi, Yashar Mehdad, Shafiq Joty, Giuseppe Carenini, and Raymond Ng. 2013. Dialogue act recognition in synchronous and asynchronous conversations. In *Proceedings of the SIGDIAL 2013 Conference*, pages 117–121, Metz, France, August. Association for Computational Linguistics.

K. Teuber. 1995. *Die Siedler von Catan: Regelheft.* Kosmos Verlag, Stuttgart, Germany.

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.

Yoam Sholam and Kevin Leyton-Brown. 2009. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.