

Learning interaction between conflicting human agents and their assistants

Boris Galitsky

Knowledge-Trail Inc 9 Charles Str., Natick MA 01760
bgalitsky@knowledge-trail.com

Boris Kovalerchuk

Dept. of Computer Science, Central Washington University,
Ellensburg, WA, 98926
borisk@cwu.edu

Abstract

We build the generic methodology based on machine learning and reasoning to detect the patterns of interaction between conflicting agents, including humans and their assistants.

Learning behavior via communicative actions

Over last decade, substantial efforts have gone into designing intelligent artificial assistants. These assistants will help human agents in a variety of everyday tasks, as well as in tasks involving assistants as well as human agents in a conflict. Therefore such assistant should be capable of reasoning about handling and resolving conflicts, learning from previous experience.

It has been well understood that to be truly useful, the assistants needs to be personalized, be trusted and become partners, and be able to learn from the human agent as well as from its own mistakes. In addition to that, to be involved in a domain where some form of interaction while conflict resolution is required, assistants need to be enabled with machine learning framework to categorize conflicts, predict their outcomes, and suggest resolution strategies. Moreover, a desired machine learning framework should be reusable from domain to domain, follow the general laws of inter-human conflicts, and be computationally tractable.

In our earlier studies of learning behavior of conflicting human agents we discovered that scenarios of inter-human conflicts can be categorized and the outcomes of these conflicts can be predicted, based on the structure of communicative actions of these agents, abstracting from particular domain knowledge. In the current study we explore how the graph-based mechanism of learning communicative actions can be incorporated into intelligent

assistants designed for a number of distinct domains involving a conflict dialogue.

One of the main problems to be solved by assistants while dealing with inter-human conflict is how to reuse the previous experience with similar agents. We deploy a machine learning technique for handling scenarios of interaction between conflicting human agents. The developed scenario representation and comparative analysis techniques are first applied (and then evaluated) to the classification of customer complaints. Then the proposed technique is employed in such problems as predicting an outcome of international conflicts, assessment of an attitude of a security clearance candidate, mining emails for suspicious emotional profiles and revealing suspicious behaviour of cell phone users.

Successful use of the proposed methodology in rather distinct domains shows its adequacy for mining human attitude-related data in a wide range of applications. The result of this study is that as long as intelligent assistants are programmed with the common mechanisms of interactions between human agents, they can learn conflict resolution in a wide variety of domains.

Evaluation in various domains

To demonstrate that the proposed representation language of labeled graphs is adequate to represent scenarios of interactions between human agents in various domains, we performed the evaluation of coding to graph / decoding from graph and evaluate distortion of communicative action - related information. We conducted the evaluation with respect to the criteria on how the suggested model based on communicative actions can represent real-world scenarios.

We started the evaluation from textual complaints which were downloaded from the public website PlanetFeedback.com in 2005. For the purpose of this evaluation, each complaint was manually coded as a sequence of communicative actions, being assigned with a particular status. The usability and adequacy of our

formalism was evaluated on the basis of a team of individuals divided into three classes: *complainants*, *company representatives*, and *judges*.

Complainants had a task to read a textual complaint and draw a graph so that another team member (a company representative) could comprehend it (and briefly sketch the plot as a text). A third team member (judge) then compared the original complaint and the one written by the company representative as perceived from the form. The result of this comparison was the judgment on whether the scenario structure has been dramatically distorted in respect to the validity of a given complaint.

It must be noted that less than 15% of complaints were hard to capture by means of communicative actions. We also observed that about a third of complaints lost important details and could not be adequately restored (although they might still be properly related to a class). Nevertheless, one can see that the proposed representation mechanism is adequate for representing so complex and ambiguous structures as textual complaints in most cases.

Conducting the evaluation of adequateness in other domains, we split the members of evaluation team into *reporters*, *assessors* and *judges*. Reporters represented scenarios as graphs, and assessors decoded the perceived structure of communicative actions back into text. Finally, the judges compared the original description (be it text or other media in the case of wireless interaction) with the respective originals (Table 1).

Domain/dataset	Number of scenarios	Percentage of scenarios which were successfully represented as graphs by experts	Percentage of scenarios which were (at least partially) reconstructed from the graph	Percentage of scenarios which were properly represented and reconstructed	Percentage of properly related to a class (being adequately represented), 2 classes
Complaints-Bank 1	20	85	75	65	72
Complaints-Bank 2	20	80	75	60	75
Complaints-Bank 3	20	95	85	75	78
Conflict between communities of agents	2	50	50	50	No eval
Emotional interactions (revealed communicative actions plus emotions)	12	75	67	58	60
Induced conflict interactions while security clearance assessment	4	No eval	No eval	No eval	No eval
Wireless interaction under possible suspicious behavior	38	84	74	55	61
Average	18.7	78.2	71	60.5	69.2

Table 1: Evaluation of the adequacy of representation language

For the banks, one can track deviation of one dataset versus another, which is 10-15% of the third set versus the first two sets. This is due to the lower variability of scenarios, which makes it easier to represent and reconstruct it (classification accuracy is comparable).

Recognition for banking complaints is almost as accurate as coding via graph (representation), but not the reconstruction of the structure of interactions between complainants and their opponents.

For an average number of almost 19 scenarios per dataset, almost 80% can be somehow represented via labeled graphs, about 70% reconstructed from graph without major loss of the conflict structure, and 60% both correct representation and reconstruction. The classification accuracy of relating to one out of two classes is close to the reconstruction accuracy.

Conclusions

We explored the role of communicative actions in representing various kinds of conflicts in multiagent systems and observed that proper formalization of communicative actions is essential to judge on conflicts. A machine learning approach to relate a formalized conflict scenario to a class is proposed, which takes into account structures of communicative actions represented via labeled graphs. For machine learning, the scenarios are represented as a sequence of communicative actions attached to agents; these actions are grouped by subjects.

The structure of graphs, as well as the number and structure of classes depend on a domain, but the criteria of sequences of communicative actions have been shown

useful to express commonalities between scenarios. Hence domain-independent communicative actions' representation via labeled graphs, once developed, can be reused from one conflict domain to another. At the same time, having the common representation language, scenarios from one domain are dissimilar to the ones from another domain, so only the knowledge about communicative actions structure is common between these domains. In each domain, graph structures are different, so we cannot export experience from domain to domain.

Based on the Speech Act theory, we designed the set of attributes for communicative actions and showed how the procedure of relating a complaint to a class can be implemented as Nearest Neighbor learning machinery. We believe that suggested approach is

appropriate for deployment in decision support settings in the mentioned domains.