DipTools: Experimental Data Visualization Tool for the DipGame Testbed (Demonstration)

Angela Fabregues, David López-Paz and Carles Sierra
Artificial Intelligence Research Institute (IIIA-CSIC)
Campus Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain
{fabregues, sierra}@iiia.csic.es, dlopez@dipgame.org

ABSTRACT

DipGame is a testbed for negotiation. It permits to test negotiation algorithms, even if enriched with argumentation, trust or reputation techniques. It is very appropriate to run experiments that mix humans and agents. In this demonstration we introduce a tool to visualise data obtained from DipGame experiments.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems

General Terms

Experimentation

Keywords

application, visualisation tool, testbed, diplomacy game

1. INTRODUCTION

Diplomacy is a rather popular game. It is very adequate for MAS research because negotiation is key to win. In the game, players represent seven European Great Powers that decide alliances, select whom to ask for help, argue with other players, get information about other players immediate objectives, or find out what the others know. From the point of view of AI research, Diplomacy is a multiagent system environment where competitive self interested agents need to cooperate to obtain better outcomes. This is done through negotiation. Players can be incarnated by software agents and compete either with other agents or with humans. During every phase of a game,[1] software agents exchange proposals and observe how their counterparts (software or human) behave. Thus they can build a model of the other agents' beliefs, desires and intentions. This model is key to decide whom to trust and whom to betray and when. The game is therefore very appropriate to experiment argumentation, negotiation, trust or reputation models.

[1]A game is composed of a sequence of phases, where negotiation and movements happen.

Cite as: DipTools: Experimental Data Visualization Tool for the DipGame Testbed (Demonstration), A. Fabregues, D. López-Paz and C. Sierra, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Yolum, Tumer, Stone and Sonenberg (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1315-1316.

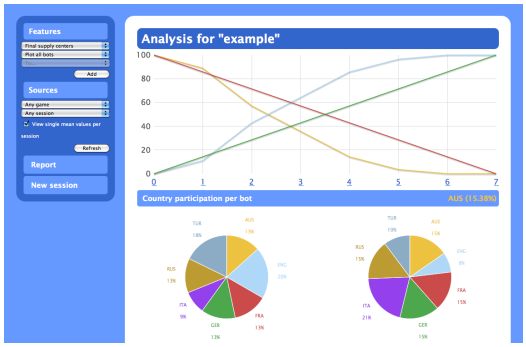

Figure 1: Screenshot of the tool

In order to facilitate that MAS researchers experiment with this game we created DipGame [1]. It is both a website for humans to play the game and a testbed to run experiments. As argued in [2, 4], Diplomacy is a flexible and rich domain for a multiagent systems testbed.

The testbed is in production and available to everyone at `http://www.dipgame.org`. What we introduce in this demonstration is *DipTools*, a visualisation tool that enriches the testbed with support for experimental data analysis, see Figure 1.

Probably the most popular visualisation tool used by AI researchers for their experiments is Gnuplot (`http://www.gnuplot.info`). It is a useful tool but the generation of the data files in the appropriate format and the selection of its settings are quite tedious when you are interested in the analysis of several variables. Often, researchers complain about the lack of tools similar to GapMinder (`http://www.gapminder.org/`) to represent their results. It is a web-based visualisation tool that is very flexible —it allows for several variables to be represented, and interactive —charts can be created aggregating variables dynamically. Concretely, the most important experimental analysis in MAS research is the *relationships* among agents. Instead of just comparing an agent against another, we would like to compare the relationships among sets of them. This kind of analysis is not possible to be done with visualisation tools like GapMinder. Diptools aims at bringing to the DipGame testbed users, and to the MAS community in general, the possibility of using an experiment visualisation tool that is interactive, flexible, and web based. Moreover, it eases the analysis not only of individual agent behaviours but also of relationships between agents.

We describe the visualisation tool in section 2 and provide an example in section 3.

2. DIPTOOLS

An experiment is defined as a set of sessions each one containing a set of games. Sessions are used in DipTools to allow the experimenter to group together the data from games ran using the same settings, it is usually useful to compare results obtained from different settings. Several experiments can be stored but only one can be visualised at any time.

There are three families of charts: (i) for a single game, (ii) for a game session and (iii) for the whole experiment. The chart of a single game represents on the x-axis the phases of the game. On the y-axis it permits to display a numerical variable. For example, the amount of deals reached by an agent.

Given a game session, the tool allows to plot variable values over the games of the sessions. This chart can be used to check whether the performance of a bot was similar or not in all session games. We can plot, for instance, the degree of interaction with other agents or the ranking of the bot at the end of each game.

Finally, given the overall experiment, the tool allows to chart the average of a selected variable over all the games of each session. This option is used in the example provided in section 3 and illustrated in Figure 2. It is a quick way to visualise the overall performance of our agents.

There are many useful variables that can be displayed and that are related to a player (e.g. the number of successful movements[2]) or to the interaction of two players (e.g. the number of attacks between them). The experimenter just needs to select the observable variables and the involved agents (one or two). An observable variable can be complex as, for instance, the number of times that *simpleBot* has attacked Germany or the number of attacks that *simpleBot* has performed. The tool allows the experimenter to easily define such observable variables, as well as chart several of them at the same time.

In addition to point chart displays, DipTools provides pie charts that are ideal to represent exclusive variable values as, for example, what percentage of victories were obtained by a particular agent depending on what Great Power it was representing. The tool also provides text reports where the data is provided in tabular form.

3. EXAMPLE

To perform an experiment a user should download all resources from `http://www.dipgame.org` and implement a number of agents. In this example we assume that two agents have been implemented, one of them capable of negotiating [3]. We assume that an experiment is performed with 8 sessions where the games in each session had 0, 1, ..., or 7 instances of the negotiating agent and the rest of the players were instances of the non negotiating agent, e.g. session 4 has 4 instances of the negotiating agent and 3 instances of the non negotiating agent. 100 games are performed in each session. After running the experiment, the 800 games, we load the log files containing the results of the experiment into DipTools.

With DipTools we can then choose the variables we are interested in to produce charts and reports. For example, in Figure 2 we can see a chart on the overall experiment where the percentage of games won by every agent is represented. Note that the number of players of each agent type competing in each session is different. We can say that the negotiating agent performs better than the non negotiating one because its percentage of victories is larger than the expected results in case all agents were equal.

This paper is completed with a video demostration available at `http://www.dipgame.org/media/AAMAS2011demo`.

[2]Sometimes the players do not succeed in performing their movements because of collisions with the movements of other players.

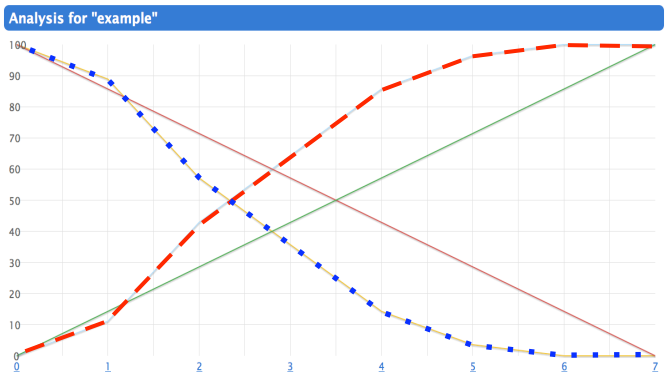

Figure 2: Percentage of games won per session. The dashed line represents the percentage of victories of negotiating agents and the doted line the percentage of victories of non negotiating agents. The continuous lines (increasing and decreasing) represent the expected percentage of the negotiating and non-negotiating agents in case they all were equal. This particular graphic shows that the negotiating agents perform better in the experiment.

4. ACKNOWLEDGMENTS

Research supported by the Agreement Technologies CONSOLIDER project under contract CSD2007-0022 and INGENIO 2010, by the Agreement Technologies COST Action, IC0801, and by the Generalitat de Catalunya under the grant 2009-SGR-1434. This work is supported under project CBIT, which is funded by Spain's Ministry of Science and Innovation under grant number TIN2010-16306.

5. REFERENCES

[1] A. Fabregues, D. Navarro, A. Serrano, and C. Sierra. Dipgame: a testbed for multiagent systems (extended abstract). In *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, 2010.

[2] A. Fabregues and C. Sierra. Diplomacy game: the test bed. *PerAda Magazine, towards persuasive adaptation*, 2009. `http://www.perada-magazine.eu/view.php?source=1761-2009-08-03`.

[3] A. Fabregues and C. Sierra. An agent architecture for simultaneous bilateral negotiations. In *Proceedings of the 13è Congrés Internacional de l'Associació Catalana d'Intel·ligència Artificial (CCIA 2010)*, pages 29–38, Espluga de Francolí, Tarragona, 2010.

[4] S. Kraus, D. Lehmann, and E. Ephrati. An automated diplomacy player. In D. Levy and D. Beal, editors, *Heuristic Programming in Artificial Intelligence: The 1st Computer Olympia*, pages 134–153. Ellis Horwood Limited, 1989.