Travian Meta Alliances ⋆

Ingo Thon[1], Niels Landwehr[2], and Luc De Raedt[1]

[1]Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, 3001 Heverlee, Belgium
[2]Department of Computer Science, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany
ingo.thon@cs.kuleuven.be, landwehr@cs.uni-potsdam.de, luc.deraedt@cs.kuleuven.be

Abstract.

1 CPT-L

Let us first introduce some terminology. A logical *atom* is an expression of the form $p(t_1, \ldots, t_n)$ where p/n is a *predicate symbol* and the t_i are *terms*. Terms are built up from constants, variables, and functor symbols. Constants are denoted in lower case (such as a), variables in upper case (such as Z), and functors by f/k where k is the arity of functor f. The set of all atoms is called a *language* $\mathcal{L}$. *Ground* expressions do not contain variables. Ground atoms will be called *facts*. A substitution θ is a mapping from variables to terms, and $b\theta$ is the atom obtained from b by replacing variables with terms according to θ. As an example, consider the substitution $\theta = \{Z/a\}$ that replaces variable Z with a, as in $b\theta = p(a)$ for $b = p(Z)$.

Complex world states can now be described in terms of *interpretations*. An interpretation I is a set of ground facts $\{a_1, \ldots, a_N\}$. These ground facts can represent objects in the current world state, their properties, and any relationship between objects. As an example, consider the representation of the state of a multiplayer game in terms of an interpretation as depicted in Figure 1. The semantics of CPT-L is based on CP-logic, a probabilistic first-order logic that defines probability distributions over interpretations [1]. CP-logic has a strong focus on causality and constructive processes: an interpretation is incrementally constructed by a process that adds facts which are probabilistic *outcomes* of other already given facts (the *causes*). CPT-L combines the semantics of CP-logic with that of (first-order) Markov processes. Causal influences only stretch from I_t to I_{t+1} (Markov assumption), are identical for all time steps (stationarity), and all causes and outcomes are observable.

Definition 1. *A* **CPT-theory/model** *is a set of rules of the form*

$$r = (h_{1,1} \wedge \ldots \wedge h_{1,k_1} : p_1) \vee \ldots \vee (h_{n,1} \wedge \ldots \wedge h_{1,k_n} : p_n) \longleftarrow b_1, \ldots, b_m$$

where the $h_{i,j}$ are logical atoms, $p_i \in [0,1]$ are probabilities s.t. $\sum_{i=1}^{n} p_i = 1$, and the b_l are literals (i.e., atoms or their negation).

⋆ This is an excerpt of a paper accepted for publication in the Machine Learning Journal Special issue on Mining and Learning with Graphs and Relations.

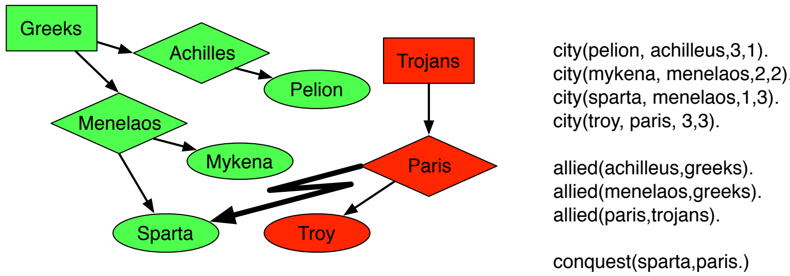

Fig. 1: Example for the state of a multiplayer game represented as a graph structure and, equivalently, as a logical interpretation. The rectangles in graphical representation refer to alliances, diamonds to players, and ellipsis to cities. The last two arguments of city in the logical representation refer to the location of the city.

A conjunction $h_{i,1} \wedge \ldots \wedge h_{i,k_i}$ in $head(r)$ will also be called a *head element*, and its probability p_i will be denoted by $P(h_{i,1} \wedge \ldots \wedge h_{i,k_i} \mid r)$. The meaning of a rule is that whenever $b_1\theta, \ldots, b_m\theta$ holds for a substitution θ in the current state I_t, exactly one head element $h_{i,1}\theta \wedge \ldots \wedge h_{i,k_i}\theta$ is chosen from $head(r)$ and all its conjuncts $h_{i,j}\theta$ are added to the next state I_{t+1}.

Example 1. Consider the following CPT-theory for the *travian world* domain:

$$conq(P, C) : 0.039 \vee nil : 0.961 \longleftarrow conq(P, C'), city(C', \_, \_, \_, P'), city(C, \_, \_, \_, P')$$

The rule encodes that a player is likely to conquer a city of a player he or she already attacked in the previous time step.

2 Experimental Evaluation[1]

Chat Room Domain This domain is concerned with the analysis of user interaction in chat rooms. We have monitored a number of IRC chat rooms in real time, and recorded who was sending messages to whom using the PieSpy utility [2]. This results in dynamically changing graphs of user interaction, representing the social network structure among chat room participants, cf. Figure 2 (left). We learn these dynamics using separate models for different chat rooms. The resulting set of models can be used to visualize commonalities and differences in the behavior displayed in different chat rooms, thereby characterizing the underlying user communities.

2.1 Experiments in the Chat Room Domain

For our experiments in the chat room domain, we have selected the following 7 well-frequented IRC chat rooms: $[football \mid iphone \mid computer \mid poker \mid math \mid politics]@irc.efnet.net$, and *travian@irc.travian.org*. Each chat room was monitored

[1] The implementation, models and data will be made available soon at http://www.ingothon.de/

using the PieSpy utility [2], generating a sequence graphs (c.f. Figure 2 (left)). For each chat room, we selected the first 100 observations in a sequence.

We have hand-coded a simple CPT-theory $\mathcal{T}$ for this domain, which makes use of a number of graph-theoretic properties defined in the background knowledge, such as graph centrality, node degree, closeness, betweenness, and co-citation. As an example rule, consider

$$communicates(P1, P2) : 0.1 \vee nil : 0.9 \longleftarrow cocitation(P1, P2, CC)$$

encoding that two chat participants start talking to each other if there is a third participant with whom they have both talked before. For each chat room we learn the probabilistic parameters of a CPT-theory, resulting in 7 CPT-theories $\mathcal{T}_1, \ldots, \mathcal{T}_7$ with the same rule structure but different parameters. Learning took about 10 seconds per theory $\mathcal{T}_i$. The learned CPT-theories can be seen as a probabilistic representation of the typical interaction behavior among members of that chat room. We evaluated the likelihood $P(S_i \mid \mathcal{T}_j)$ of each sequence S_i under the learned CPT-theory $\mathcal{T}_j$. This gives an indication as to how well the behavior in chat room i is explained by the model learned for chat room j.

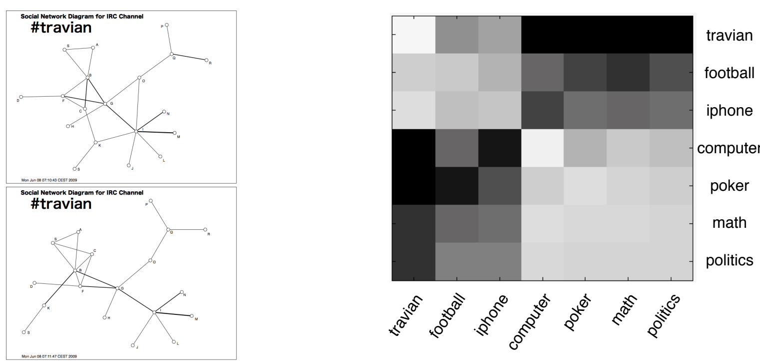

Fig. 2: (Left) User interaction graphs from the Chat Room Domain. Shown are two subsequent different time points during the observation sequence recorded for the *irc.travian.org* chat room. (Right) Plot of the likelihood $P(S_i \mid \mathcal{T}_j)$ of a sequence S_i (corresponding to chat room i) under the CPT-theory $\mathcal{T}_j$ (learned on chat room j). Rows correspond to models $\mathcal{T}_j$ and columns to sequences S_i. Lighter colors indicate higher likelihoods.

The result of this experiment is visualized in Figure 2. We can distinguish different clusters of chat rooms, or, equivalently, user communities. For instance, chat rooms that are concerned with recreational topics such as *travian* and *football* (as well as *iphone*) are clearly distinguishable from chat rooms concerned with more “serious” topics such as *math* and *politics*. Manual inspection of the learned rule parameters showed that in the “serious” chat domains the likelihood of a communication between two players mostly depends on the betweenness and degrees of the nodes involved, while in the “recreational” chats shared cocitations are more important.

References

1. Vennekens, J., Denecker, M., Bruynooghe, M.: Representing Causal Information About a Probabilistic Process. In: Logics In Artificial Intelligence. Volume 4160 of Lecture Notes in Computer Science. (2006) 452–464
2. Mutton, P.: Inferring and visualizing social networks on Internet Relay Chat. In: Proceedings of the 8th International Conference on Information Visualization (IV-2004). (2004) 35–43