CPT-L: an Efficient Model for Relational Stochastic Processes

Ingo Thon INGO.THON@CS.KULEUVEN.BE
Niels Landwehr NIELS.LANDWEHR@CS.KULEUVEN.BE
Luc De Raedt LUC.DERAEDT@CS.KULEUVEN.BE
Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, 3001 Heverlee, Belgium

Abstract

Agents that learn and act in real-world environments have to cope with both complex state descriptions and non-deterministic transition behavior of the world. Standard statistical relational learning techniques can capture this complexity, but are often inefficient. We present a simple probabilistic model for such environments based on CP-Logic. Efficiency is maintained by restriction to a fully observable setting and the use of efficient inference algorithms based on binary decision diagrams.

1. Introduction

Artificial intelligence aims at developing agents that learn and act in complex environments. Realistic environments typically feature a variable number of objects, relations amongst them, and non-deterministic transition behavior. Examples for such environments are massively multiplayer online role-playing games, where an agent or user has to make decisions based on his role in a social network and current relationships with other users. Standard probabilistic sequence models provide efficient inference and learning techniques, but typically cannot fully capture the relational complexity. On the other hand, statistical relational learning techniques are often too inefficient. In this paper, we present a simple model that occupies an intermediate position in this expressiveness/efficiency trade-off. More specifically, we contribute a novel representation, called CPT-L (for **CPT**ime-**L**ogic), that essentially defines a probability distribution over sequences of interpretations. Interpretations are relational state descriptions that are typically used in planning and many other applications of artificial intelligence. CPT-L can be considered a variation of CP-logic (Vennekens et al., 2006), a recent expressive logic for modeling causality. By focusing on the sequential aspect and deliberately avoiding the complications that arise when dealing with hidden variables, CPT-L is more restricted, but also more efficient to use than its predecessor and alternative formalisms within the artificial intelligence and statistical relational learning literature.

This is clear when positioning CPT-L w.r.t. the few existing approaches that can probabilistically model sequences of relational state descriptions. First, standard SRL-approaches (Getoor & Taskar, 2007) can be used in this setting by explicitly modeling time. However, such models are often intractable for complex sequential real-world domains. Second, relational STRIPS-based techniques (Zettlemoyer et al., 2005) are able to probabilistically model relational sequences. However, they are restricted by the fact that only one rule can "fire" at a particular point in time and thus only one aspect of the world can be changed. Finally, a third class of approaches (e.g. relational simple-transition models (Fern, 2005)) is concerned with modeling domains where the process generating the data is hidden. This is a significantly harder setting than the one discussed in this paper. The key contributions of our work are the introduction of 1) the CPT-L model for representing probability distributions over sequences of interpretations, 2) efficient algorithms for inference in CPT-L and 3) a simple but efficient Expectation-Maximization algorithm (EM) algorithms for parameter learning from fully observable example sequences.

2. CPT-L

A relational interpretation I is a set of ground facts $a_1, ..., a_N$ representing the objects and relations between them in the current state. A *relational stochastic process* defines a distribution $P(I_1, ..., I_T)$ over sequences of interpretations of length T, and thereby completely characterizes the transition behavior of the world.
The semantics of CPT-L is based on CP-logic, a probabilistic first-order logic that defines probability distributions over interpretations (Vennekens et al., 2006). CP-logic has a strong focus on causality and constructive processes: an (logical) 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

Appearing in the 6^{th} *International Workshop on Mining and Learning with Graphs*, Helsinki, Finland, 2008.

I_t to I_{t+1} (Markov assumption), are identical for all time steps (stationarity), and all causes and outcomes are observable. Models in CPT-L are also called CP-theories, and are defined as follows:

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

$$r = \underbrace{(h_1 : p_1) \vee \ldots \vee (h_n : p_n)}_{head(r)} \leftarrow \underbrace{b_1, \ldots, b_m}_{body(r)}$$

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

We shall also assume all variables appearing in the head of the rule also appear in its body. The intuition behind a rule is that whenever the (grounded) body of the rule holds in the current state I_t, one of the (grounded) heads will hold in the next state I_{t+1}. In this way, the rule models a (probabilistic) causal process as the condition specified in the body causes one (probabilistically chosen) atoms in the head to become true in the next time step. One of the main features of CPT-theories is that they are easily extended to include *background knowledge*, which can be any logic program (cf. (Bratko, 1990)). In the presence of background knowledge, we say that a ground rule is applicable in an interpretation I_t if its body $b_1\theta, \ldots, b_m\theta$ can be logically derived from I_t and the logic program B.

A CPT-theory defines a distribution over possible successor states, $P(I_{t+1} \mid I_t)$, in the following way. Let $\mathbf{R}_t = \{r_1, ..., r_k\}$ denote the set of all ground rules applicable in the current state I_t. Each ground rule applicable in I_t will cause one of its head elements to become true in I_{t+1}. More formally, a *selection* σ is a mapping from rules r_i to indices j_i denoting that head element $h_{ij_i} \in head(r_i)$ is selected. In the stochastic process to be defined, I_{t+1} is a possible successor for the state I_t if and only if there is a selection σ such that $I_{t+1} = \{h_{1\sigma(1)}, ..., h_{k\sigma(k)}\}$. We say that σ *yields* I_{t+1} from I_t, denoted $I_t \xrightarrow{\sigma} I_{t+1}$, and define

$$P(I_{t+1}|I_t) = \sum_{\sigma: I_t \xrightarrow{\sigma} I_{t+1}} P(\sigma) = \sum_{\sigma: I_t \xrightarrow{\sigma} I_{t+1}} \Big(\prod_{(r_i, j_i) \in \sigma} p_{j_i} \Big) \quad (1)$$

where p_{j_i} is the probability associated with head element h_{ij_i} in r_i. As for propositional Markov processes, the probability of a sequence $I_1, ..., I_T$ given an initial state I_0 is defined by

$$P(I_1, ..., I_T) = P(I_1) \prod_{t=0}^{T} P(I_{t+1} \mid I_t). \quad (2)$$

Intuitively, it is clear that this defines a distribution over all sequences of interpretations of length T much as in the propositional case. More formally, the Kolmogorov extension theorem can be used to prove the following theorem:

Theorem 1 (Semantics of a CPT theory)**.** *Given an initial state I_0, a CPT-theory defines a discrete-time stochastic process, and therefore for $T \in \mathbb{N}$ a distribution $P(I_1, ..., I_T)$ over sequences of interpretations of length T.*

1. Initialize $f := true$
2. Compute applicable ground rules
   $\mathbf{R}_t = \{r\theta | body(r\theta) \text{ is true in } I_t\}$
3. $\forall r(r = (p_1 : h_1, ..., p_n : h_n) \leftarrow b_1, ..., b_m)$ in $\mathbf{R}_t$ do:
   (a) $f := f \wedge (r.h_1 \vee ... \vee r.h_n)$,
   (b) $f := f \wedge (\neg r.h_i \vee \neg r.h_j)$ for all $i \neq j$
4. For all facts $l \in I_{t+1}$
   (a) Initialize $g := false$
   (b) for all $r \in \mathbf{R}_t$ with $p : l \in head(r)$ do $g := g \vee r.l$
   (c) $f := f \wedge g$

Figure 1. Algorithm to convert the inference problem into a formula f. Concatenations $r.h$ of a rule r and head element h denote a propositional variable indicating that h was selected in r.[2]

3. Inference and Parameter Estimation

As for other probabilistic models, we can now ask several questions about the introduced CPT-L model:

- **Sampling**: how to sample sequences of interpretations $I_1, ..., I_T$ from a given CPT-theory $\mathcal{T}$ and initial interpretation I_0?
- **Inference**: given a CPT-theory $\mathcal{T}$ and a sequence of interpretations $I_1, ..., I_T$, what is $P(I_1, ..., I_T \mid \mathcal{T})$?
- **Parameter Estimation**: given the structure of a CPT-theory $\mathcal{T}$ and a set of sequences of interpretations, what are the maximum-likelihood parameters of $\mathcal{T}$?
- **Prediction**: Let $\mathcal{T}$ be a CPT-theory, $I_1, ..., I_t$ a sequence of interpretations, and F a first-order formula that constitutes a certain property of interest. What is the probability that F holds at time $t + d$, $P(I_{t+d} \models_B F \mid \mathcal{T}, I_1, ..., I_t)$?

Sampling from a CPT-theory $\mathcal{T}$ given an initial interpretation I_0 is straightforward due to the causal semantics employed in CP-logic. For $t \geq 0$, I_{t+1} can be constructed from I_t by finding all groundings $r\theta$ of rules $r \in \mathcal{T}$, and sampling for each $r\theta$ a head element to be added to I_{t+1}.

Inference: Because of the Markov assumption (Equation 2), the crucial task for solving the inference problem is to compute $P(I_{t+1} \mid I_t)$ for given I_{t+1} and I_t. According to Equation 1, this involves summing the probabilities of all selections yielding I_{t+1} from I_t. However, the number of possible selections σ is exponential in the number of ground rules $|\mathbf{R}_t|$ applicable in I_t, so a naive generate-and-test approach is infeasible. Instead, we present an efficient approach for computing $P(I_{t+1} \mid I_t)$ without explicitly enumerating all selections yielding I_{t+1}, which is strongly related to the inference technique discussed in (De Raedt et al., 2007) and also somewhat related to techniques discussed in (Chavira et al., 2006). The problem is first converted to a DNF formula over boolean variables such

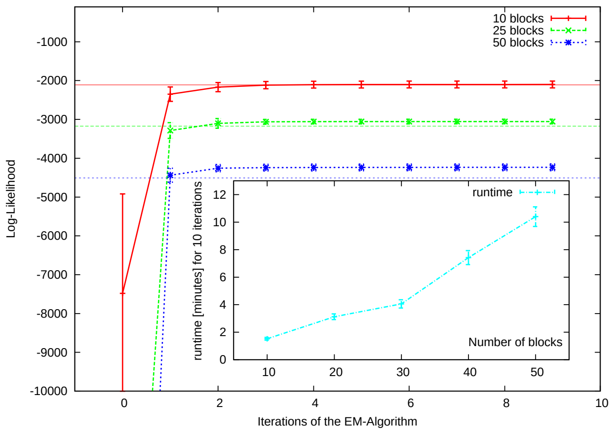

Figure 2. Large graph: per-sequence log-likelihood on training data as a function of the EM iteration together with the log-likelihood for the gold standard model. Small graph: Running time of EM as a function of the number of blocks in the world model.

that assignments to variables correspond to selections, and satisfying assignments to selections yielding I_{t+1}. The formula is then compactly represented as a binary decision diagram (BDD), and $P(I_{t+1} \mid I_t)$ efficiently computed from the BDD using dynamic programming. Although finding satisfying assignments for DNF formulae is a hard problem in general, the key advantage of this approach is that existing, highly optimized BDD software packages can be used. The conversion of a given inference problem to a DNF formula f is briefly sketched by the pseudocode given in Figure 1.

Learning can be realized by an expectation-maximization approach, where the hidden information is the head element used in the application of a grounded rule. Sufficient statistics for the maximization step are the expected number of times a head element h has been selected in rule r, and the key algorithmic challenge is to compute these expectations efficiently. This can be realized using the same BDD structure as for inference, with a dynamic programming algorithm related to the forward-backward procedure used in hidden Markov models.

4. Experimental Evaluation

The proposed CPT-L model has been evaluated in a stochastic version of the well-known *blocks world* domain. The domain was chosen because it is truly relational and also serves as a popular artificial world model in agent-based approaches such as planning and reinforcement learning. Furthermore, it is an example for a domain in which multiple aspects of the world can change concurrently — for instance, a block can be moved from A to B while at the same time a stack collapses, spilling all of its blocks on the floor. In an experiment, we explore the convergence behavior of the EM algorithm for CPT-L. The world model together is implemented by a (gold-standard) CPT-theory $\mathcal{T}$, and a training set of 20 sequences of length 50 each is sampled from $\mathcal{T}$. From this data, the parameters are re-learned using EM. Figure 2, large graph, shows the convergence behavior of the algorithm on the training data for different numbers of blocks in domain, averaged over 15 runs. It shows rapid and reliable convergence. Figure 2, small graph, shows the running time of EM as a function of the number of blocks. The scaling behavior is roughly linear, indicating that the model scales well to reasonably large domains. Absolute running times are also low, with about 1 minute for an EM iteration in a world with 50 blocks.This is in contrast to other, more expressive modeling techniques which typically scale badly to domains with many objects. The difference between the log likelihood on an independent test set of the gold-standard model and the learned model, were by four orders of magnitudes smaller than the difference to a random model. Manual inspection of the learned model also shows that parameter values are on average very close to those in the gold-standard model.

5. Conclusions and Future Work

We have introduced CPT-L, a probabilistic model for sequences of relational state descriptions. In contrast to other approaches that address this setting, the focus in CPT-L is on computational efficiency rather than maximal expressivity. The main interesting directions for future work is to further evaluate representation power and scaling behavior of the model in challenging real-world domains.

References

Bratko, I. (1990). *Prolog programming for artificial intelligence*. Addison-Wesley. 2nd Edition.

Chavira, M., Darwiche, A., & Jaeger, M. (2006). Compiling relational bayesian networks for exact inference. *Int. J. Approx. Reasoning*, *42*, 4–20.

De Raedt, L., Kimmig, A., & Toivonen, H. (2007). Problog: A probabilistic Prolog and its application in link discovery. *ICJAI-07* (pp. 2462–2467).

Fern, A. (2005). A simple-transition model for relational sequences. *IJCAI-05* (pp. 696–701).

Getoor, L., & Taskar, B. (Eds.). (2007). *Statistical relational learning*. MIT press.

Vennekens, J., Denecker, M., & Bruynooghe, M. (2006). Representing causal information about a probabilistic process. *Logics In Artificial Intelligence* (pp. 452–464).

Zettlemoyer, L. S., Pasula, H., & Kaelbling, L. P. (2005). Learning planning rules in noisy stochastic worlds. *AAAI-05* (pp. 911–918).