Modal Markov Logic for Multiple Agents

Tivadar Papai
papai@cs.rochester.edu
Department of Computer Science
University of Rochester
Rochester, NY 14627

Henry Kautz
kautz@cs.rochester.edu
Department of Computer Science
University of Rochester
Rochester, NY 14627

Abstract

Modal Markov Logic for a single agent has previously been proposed as an extension to propositional Markov logic. While the framework allowed reasoning under the principle of maximum entropy for various modal logics, it is not feasible to apply its counting based inference to reason about the beliefs and knowledge of multiple agents due to magnitude of the numbers involved. We propose a modal extension of propositional Markov logic that avoids this problem by coarsening the state space. The problem stems from the fact that in the single-agent setting, the state space is only doubly exponential in the number of propositions in the domain, but the state space can potentially become infinite in the multi-agent setting. In addition, the proposed framework adds only the overhead of deciding satisfiability for the chosen modal logic on the top of the complexity of exact inference in propositional Markov logic. The proposed framework allows one to find a distribution that matches probabilities of formulas obtained from training data (or provided by an expert). Finally, we show how one can compute lower and upper bounds on probabilities of arbitrary formulas.

1 Introduction

The goal of this work is to provide a framework for probabilistic reasoning for modal logic in the multi-agent setting. More specifically, if we are given a set of modal formulas and their probabilities, the goal is to infer the probability of an arbitrary propositional modal logic formula. This problem formulation is different from the concept of using probabilistic Kripke structures (Shirazi and Amir 2007; 2008) where the probabilistic model was assumed to be already learned, and probabilistic queries were answered using this model.

(Papai, Kautz, and Stefankovic 2013) recently proposed a *single agent* modal extension to propositional Markov logic. In their extension the state space consisted of epistemic situations that describe both the state of the real world and the beliefs of an agent. Despite the fact that the state space was doubly exponential in the number of propositions, they provided a single exponential time algorithm for exact inference building on the idea of partitioning the state space into equivalence classes with respect to the formulas in the knowledge base, and reducing the computation of the partition function (hence inference as well) to computing the size of these partitions. The straightforward extension of this approach does not work for the multi-agent case. Therefore, we define a probability distribution *not* over epistemic situations, but, instead over equivalence classes of epistemic situations. This way the state space becomes less-fine grained, but probabilistic inference becomes more tractable. One of the main benefits of our proposed framework is that we are no longer restricted to reason about the beliefs of a single agent. We can use *any* propositional modal logic instead of propositional logic, with the extra computational cost of deciding satisfiability in our choice of modal logic.

The paper is organized as follows. First, we provide the background for the modal logics we are interested in and for Markov logic based probabilistic reasoning. Next, we describe the computational challenges of dealing with individual epistemic situations in a counting based inference algorithm in Sec. 2.3. In Sec. 3 we show how to define a probability distribution over sets of equivalent epistemic situations rather than over individual epistemic situations using a knowledge base of weighted modal logic formulas. We show how inference can be performed when we fix the set of query formulas in advance, and how approximate inference with lower and upper bounds on probabilities of formulas can be performed for a formula not part of the previously fixed set of queries. In Sec. 5 we show the connection between the distribution defined by the framework in (Papai, Kautz, and Stefankovic 2013) and ours. We point out the issue with weight learning in the fine grained state space of epistemic situations and how this issue is avoided in our framework in Sec. 6. We conclude the paper with summarizing our results in Sec. 7.

2 Background

2.1 Propositional Markov Logic

While *Markov Logic* (Domingos and Lowd 2009) is a first-order probabilistic logic, we are concerned here with its propositional subset. A propositional *Markov logic network* consists of a knowledge base $KB = \{(w_i, F_i) | i = 1, \dots, m\}$, where $w_i \in \mathbb{R}$ and F_i is a propositional formula over a fixed set of propositions $\Omega = \{p_1, \dots, p_{|\Omega|}\}$, and defines a probability distribution over truth assignments X to

Ω as follows:

$$\Pr(X = x) = \frac{1}{Z}\exp(\sum_i w_i f_i(x)) \ , \qquad (1)$$

where $f_i(x) = 1$ if F_i is true under x, otherwise $f_i(x) = 0$, and where

$$Z = \sum_{x \in \mathcal{X}} \exp(\sum_i w_i f_i(x)) \ , \qquad (2)$$

is the partition function, and $\mathcal{X}$ denotes the set of all possible truth assignments to Ω, (*i.e*, $|\mathcal{X}| = 2^{|\Omega|}$). Note that (1) defines an exponential family of probability distributions (see *e.g.* (Wainwright and Jordan 2008; Koller and Friedman 2009; Murphy 2012)). Exponential families have the property that for a given set of f_i they describe the maximum entropy distribution that satisfies the set of consistent constraints $\mathbb{E}[f_i] = c_i$. *Consistent* here means that there exists a probability distribution that satisfies all the constraints simultaneously. We can interpret c_i as the probability of the propositional formula being satisfied under a randomly chosen truth assignment x, hence (1) defines the maximum entropy distribution over the state space of truth assignments to the propositions with the constraints $\mathbb{E}[f_i] = c_i$. The probability of an arbitrary propositional formula F over Ω is defined to be the probability of F being true under a randomly chosen truth assignment X, *i.e.*:

$$\Pr(F) = \sum_{x \in \mathcal{X} : F \text{ is satisfied under } x} \Pr(X = x) = \mathbb{E}[f_i] \ . \qquad (3)$$

Note that applying the principle of maximum entropy (Jaynes 1979) is an appealing choice, since it does not use any more information than what is given by the constraints of the form $\Pr(F_i) = c_i$, however, the defined distribution can become sensitive to the choice of state space, *e.g.*, when we are dealing with first-order logic formulas, or in our case when we reason with modal logic formulas (for further reading on the topic see, *e.g.*, (Halpern and Koller 1995; Jain, Barthels, and Beetz 2010)).

2.2 Modal Logics K45, KD45, and S5

Modal logics *K45*, *KD45* and *S5* (Chellas 1980) extend propositional or first-order logic by adding a non-truth-functional sentential operators; we will again only discuss the propositional case here. We use the symbol B_i to represent the i-th modal operator in the language. Where α is a well formed sentence, then $B_i\alpha$ is a well formed sentence.

Different modal operators for concepts such as belief, knowledge, desire, obligation, *etc.*, can be specified by the *axiom schemas* that they satisfy. In this paper, we will only discuss in detail modal logics *K45*, *KD45* and *S5*. The properties of this logic are summarized as the following axioms and rules (Fagin et al. 1995):

R1. From ϕ and $\phi \supset \psi$ infer ψ (Modus ponens)

R2. From ψ infer $B_i\psi$ (Knowledge Generalization)

A1. All tautologies of propositional calculus

A2. $(B_i\phi \wedge B_i(\phi \supset \psi)) \supset B_i\psi$ (Distribution Axiom)

A3. $B_i\phi \supset \phi$ (Knowledge Axiom)

A4. $B_i\phi \supset B_iB_i\phi$ (Positive Introspection Axiom)

A5. $\neg B_i\phi \supset B_i\neg B_i\phi$ (Negative Introspection Axiom)

A6. $\neg B_i false$ (Consistency Axiom)

We get *K45* if we take R1, R2, A1, A2, A4, and A5. Besides the axioms of *K45*, *KD45* contains A6 and *S5* contains A3. *S5* is generally used to represent knowledge, and *KD45* beliefs. *K45* is similar to *KD45*; however, it allows believing in contradicting statements.

A Kripke structure over a set of propositions Ω is a tuple $\mathcal{M} = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_A)$ where $S \neq \emptyset$ is the set of states, $\pi : S \to \mathcal{X}$, where $\mathcal{X}$ is the set of truth assignments over Ω and $\mathcal{K}_i \subseteq S \times S$ describes the relation corresponding to modal operator B_i. If $s \in S$ then for a propositional formula F, we have $\mathcal{M}, s \models F$ if F is satisfied under $\pi(s)$. For a formula BF, we have $\mathcal{M}, s \models B_iF$ iff $\forall (s, r) \in \mathcal{K}_i$: $\mathcal{M}, r \models F$. Moreover, $\mathcal{M}, s \models F_1 \wedge F_2$ iff $\mathcal{M}, s \models F_1$ and $\mathcal{M}, s \models F_2$, and $\mathcal{M}, s \models \neg F$ iff $\mathcal{M}, s \not\models F$.

For each different modal logic, Kripke structures with different properties are associated. Reflexive, symmetric, and transitive relations (equivalence relations) are associated with modal operators that satisfy *S5*. Euclidean, serial, and transitive relations are associated with *KD45*. While Euclidean, and transitive relations are associated with *K45*. For a more detailed description of Kripke structures see, *e.g.*, (Chellas 1980) or (Fagin et al. 1995).

A Kripke structure with a distinguished state (generally denoting the real world) is called a pointed Kripke structure or *epistemic situation*, hence an epistemic situation $\sigma = (s, S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_A)$ where $s \in S$. We call two epistemic situations σ_1 and σ_2 equivalent if for every formula F we have $\sigma_1 \models F$ if and only if $\sigma_2 \models F$. We call two epistemic situations σ_1 and σ_2 equivalent with respect to a set of formulas $\mathcal{F}$ if for every formula $F \in \mathcal{F}$ we have $\sigma_1 \models F$ if and only if $\sigma_2 \models F$.

Using this definition of equivalence, we can partition situations into equivalence classes.

In the single-agent case we can enumerate all the non-equivalent epistemic situations for *K45*, *KD45* and *S5*, *i.e.*, we can select a member from each equivalence class by storing the worlds the agent considers possible and a distinguished real world state (Fagin et al. 1995). Hence, in the single-agent case all the non-equivalent epistemic situations could consitute the state space Σ, where $|\Sigma| \approx 2^{2^{|\Omega|}} 2^{|\Omega|}$ for all the above mentioned modal logics. However, in the multi-agent case, there are infinitely many non-equivalent epistemic situations, since two agents can make statements about each other's beliefs about each other's belief *ad infinitum*. We can still create equivalence classes with respect to a bounded depth of modal formulas and consider them in the limit when the bound on the depth approaches infinity. The depth of a formula is defined as follows:

- $d(\phi) = 0$ if ϕ is a primitive proposition,
- $d(\neg\phi) = d(\phi)$,
- $d(\phi \wedge \psi) = max(d(\phi), d(\psi))$,
- $d(\phi \vee \psi) = max(d(\phi), d(\psi))$,

- $d(B_i\phi) = d(\phi) + 1$,

In the multi-agent case, if we consider only formulas of depth D the number of non-equivalent situations with respect to these formulas would be in the range of $2^{|\Omega|}\left(2^{2^{|\Omega|}\left(2^{2^{|\Omega|}(\dots)^A}\right)^A}\right)^A$, where A is the number of modal operators (agents), and we altogether have $D+1$ levels of exponentiations nested into each other. If Σ_D^{K45}, Σ_D^{KD45}, and Σ_D^{S5} denotes the sets of non-equivalent epistemic situations *w.r.t.* formulas of depth at most D for modal logics *K45*, *KD45*, and *S5*, respectively, then it can be shown that:

$$|\Sigma_D^{K45}| = 2^{|\Omega|}(2^{|\Sigma_D^{K45}|})^A$$
$$|\Sigma_D^{KD45}| = 2^{|\Omega|}(2^{|\Sigma_D^{K45}|}-1)^A$$
$$|\Sigma_D^{K45}| = 2^{|\Omega|}(2^{|\Sigma_D^{K45}-1|})^A$$

where A is the number of agents. The main challenge in the multi-agent setting is to deal with a state space that grows with this rate.

In the multi-agent setting deciding whether an arbitrary formula is satisfiable, *i.e.*, if there exists an epistemic situation where the formula holds is a *PSPACE*-complete problem (Halpern and Moses 1992; Fagin et al. 1995).

2.3 Defining Distribution over Epistemic Situations

Here we review how a probability distribution was defined in (Papai, Kautz, and Stefankovic 2013), and describe the problem that arise in generalizing it to the multi-agent case.

Papai et al. extended propositional Markov logic by a modal operator corresponding to one of modal logics *K45*, *KD45*, or *S5*. A probability distribution over epistemic situations was defined first and the probability of a formula is defined as the sum of the probabilities of the epistemic situations where the formula holds. This idea can be generalized to any set of epistemic situations Σ. Given a non-empty set of epistemic situations Σ over a fixed Ω propositions, we can define the probability of $\sigma \in \Sigma$ as:

$$\Pr(\sigma) = \frac{1}{Z}\exp(\sum_i w_i \mathbb{I}[\sigma \models F_i]) \ , \qquad (4)$$

where $\mathbb{I}$ is the indicator function and the partition function Z is defined as:

$$Z = \sum_{\sigma\in\Sigma}\exp(\sum_i w_i \mathbb{I}[\sigma \models F_i]) \ . \qquad (5)$$

The probability of a formula ϕ (modal or non-modal) is defined as:

$$\Pr(\phi) = \sum_{\sigma\in\Sigma:\sigma\models\phi}\Pr(\sigma) \ . \qquad (6)$$

If the knowledge base consists of the weighted formulas $\{(w_1, F_1), \dots, (w_n, F_n)\}$, let $\mathcal{T}$ be the set of length n Boolean vectors, and for $t \in \mathcal{T}$ let $\Phi(t)$ denote the conjunction where the ith term is F_i if $t_i =$ true, otherwise it is $\neg F_i$.

Inference can be reduced to computing the partition function in (5), which can be rewritten as:

$$Z = \sum_{t\in\mathcal{T}} N(\Phi(t))\exp(w(t)) \ , \qquad (7)$$

where $N(\phi)$ denotes the number of epistemic situations where ϕ holds, *i.e.*, $N(\phi) = |\{\sigma \in \Sigma | \sigma \models \phi\}|$.

The probability of any query formula F can be written as:

$$\Pr(F) = \frac{1}{Z}\sum_{t\in\mathcal{T}} N(\Phi(t)\wedge F)\exp(w(t)) \ . \qquad (8)$$

The benefit of the single-agent case is that there are finitely many non-equivalent epistemic situations, hence Σ could be chosen to be an enumeration of all non-equivalent epistemic situations *w.r.t* to our choice of modal logic.

Papai et al. showed how $N(F)$ could be computed in $2^{O(|F|+|\Omega|)}$ time using basic counting rules and the inclusion-exclusion principle.

In the multi-agent setting, as it was pointed out in the previous section, even if we are dealing with formulas with bounded depth D, the state space gets too large to describe the quantities such as $N(\Phi(t))$ with sufficient precision unless we use more than $2^{2^{\dots|\Omega|}}$ space, where we have D nestings of exponentiations. To avoid dealing with such large numbers, we could try to restrict the state space to contain a subset of all possible epistemic situations that does not grow that fast as we increase D. However, this way we would lose the flexibility that an agent can have arbitrary beliefs about another agent's beliefs and so on. Moreover, if the state space grows with D, we are likely to face the issue of having always 0 probability for the formulas in the knowledge base so long as the weights are finite in the knowledge base (we discuss this in Sec. 6). Therefore, we follow a different approach and define probability distribution over *sets* of epistemic situations.

3 Defining Distribution over Partitions of Epistemic Situations

To avoid the computational issues presented in the previous section, and to have a way of assigning probabilities different from 0 and 1 for arbitrary formulas in the infinite state space of non-equivalent epistemic situations in the multi-agent case, we define probability distributions over the set of *equivalence classes of epistemic situations* rather than over the set of individual epistemic situations. Given a knowledge base $KB = \{(w_1, F_1), \dots, (w_n, F_n)\}$, $t \in \mathcal{T}$ partitions a set of non-equivalent situations Σ into at most 2^n disjoint partitions. Using the previous definition of $\Phi(t)$, a consistent $\Phi(t)$ creates a partition Π_t defined as follows:

$$\Pi_t \triangleq \{\sigma \in \Sigma | \sigma \models \Phi(t)\} \ . \qquad (9)$$

We can define a distribution over partitions

$$\Pr(\Pi_t) = (1/Z)c(\Phi(t))\exp(\sum_i w_i \mathbb{I}[t_i]) \qquad (10)$$
$$= (1/Z)c(\Phi(t))\exp(w(t)) \ ,$$

where $c(F)$ is a function that returns 1 if and only if F is a modally consistent formula, and 0 otherwise[1]. The partition function Z is defined as:

$$Z = \sum_{t \in \mathcal{T}} c(\Phi(t)) \exp(w(t)) \ . \tag{11}$$

In the single-agent the decision problem associated with the computation of $c(F)$ is *NP*-complete and for multi-agents it is *PSPACE*-complete for modal logics *K45*, *KD45*, and *S5* (Fagin et al. 1995).

Given a knowledge base, probabilities of $\Phi(t)$ can be inferred using any inference algorithm applicable for Markov logic networks with the extra additional step of rejecting the modally inconsistent states (or samples in case of relying on sampling based algorithms such as, *e.g.*, MC-SAT (Poon and Domingos 2006)) , thus yielding an exponential running time algorithm for computing the partition function Z exactly in the multi-agent case for modal logics *K45*, *KD45*, and *S5*.

Using (10) we can assign probabilities to situations if we assume that each situation in a given equivalence class contributes with equal weight to the probability of the partition. Hence, for $\sigma \in \Pi_t$ we have:

$$\Pr(\sigma) = \frac{1}{N(\Phi(t)} \Pr(\Pi_t) \ . \tag{12}$$

(10) defines a probability distribution over Π_t, and hence over formulas $\Phi(t)$, however, it does not directly define the probability of an arbitrary query formula Q. If we follow the definition (6) from Sec. 2.3, we get:

$$\Pr(Q) = (1/Z) \sum_{t \in \mathcal{T}} c(\Phi(t)) \frac{N(\Phi(t) \wedge Q)}{N(\Phi(t))} \exp(w(t)) \tag{13}$$

This way we would still need to compute the ratios $\frac{N(\Phi(t) \wedge Q)}{N(\Phi(t))}$; hence we are not much better off compared to the approach that assigns probabilities directly to epistemic situations. We propose two workarounds to avoid dealing with this computational challenge.

In the first proposed solution, we assume that we know in advance the query formulas we are interested in; therefore we can create partitions using the query formulas in addition to the formulas in the knowledge base. Let $F_1, \ldots, F_n$ be again the formulas in the knowledge base, and let $Q_1, \ldots, Q_m$ be the query formulas of interest. We can then choose $\mathcal{T}$ to be truth assignments of length $n + m$ and $\Phi(t)$ to be a conjunction composed of positive and negated forms of $F_1, \ldots, F_n, Q_1, \ldots, Q_m$. This approach is sensitive to the choice of Q, which is not surprising, since in general maximum entropy models can be sensitive to the choice of state space (see, *e.g.*, (Grove, Halpern, and Koller 1994)). The same would occur if we were to define probability distribution over epistemic situations directly and would change the state space from non-equivalent epistemic situations with respect to formulas of depth D (Σ_D) to non-equivalent epistemic situations with respect to formulas of depth $D + 1$ (Σ_{D+1}). In fact, it can always be shown that it is enough to consider a minimal set of formulas such that the truth values of these formulas always determine the truth values of the formulas both in the knowledge base and in the set of possible queries. *E.g.*, if $B_1(p \wedge q)$, $B_1 q$ are formulas in the knowledge base and $B_1 p$ is a query formula, then it is enough to partition Σ according to truth assignments to $B_1 p$ and $B_1 q$. We discuss in more details how to partition the state space in the next section.

The second approach provides bounds on the probability of arbitrary query formulas. The key idea is that $\frac{N(\Phi(t) \wedge Q)}{N(\Phi(t))} = 0$ if $\Phi(t) \wedge Q$ is modally inconsistent. Consequently,

$$\sum_{t \in \mathcal{T}} \mathbb{I}[\Phi(t) \wedge Q \text{ is consistent}] \Pr(\Phi(t)) \geq \Pr(Q) \tag{14}$$

$$1 - \sum_{t \in \mathcal{T}} \mathbb{I}[\Phi(t) \wedge \neg Q \text{ is consistent}] \Pr(\Phi(t)) \leq \Pr(Q) \tag{15}$$

Therefore, if we want to answer a query Q not being among the previously fixed set of query formulas, we can still get an estimate for the probability of Q. The exact probability would depend on the choice of Σ we are dealing with, but the exact probability would be within the bounds provided by (14). If we, *e.g.*, fix the state space to be the partitions of Σ_D created by truth assignments to $F_1, \ldots, F_n, Q_1, \ldots, Q_m$ then after changing the state space to the partitions of Σ_{D+1} we would assign a different probability to Q, but it would still be within the bounds in (14). The statement would hold even in the limit $D \to \infty$.

4 Choosing the Partitions

When we choose $\mathcal{T}$ to be the truth assignments to $F_1, \ldots, F_n, Q_1, \ldots, Q_m$, we might create many inconsistent truth assignments. Consider, *e.g.*, if in the single-agent setting we have formulas $p_1 \wedge Bp_1, p_2 \wedge \neg Bp_1, p_3 \wedge Bp_1, p_4 \wedge \neg Bp_1, \ldots, p_{|\Omega|-1} \wedge Bp_1, p_{|\Omega|} \wedge \neg Bp_1$ in the knowledge base. Any truth assignment that assigns true to a formula at an odd position and true to an even position is modally inconsistent, hence using any sampling algorithm for Markov logic most of our samples would be rejected for being modally inconsistent. Let us define *modal atom* to be formulas that start with a modal operator. As a solution to this particular case we should not choose $\mathcal{T}$ to be all the possible truth assignments to $F_1, \ldots, F_n, Q_1, \ldots, Q_m$; instead, we choose the truth assignments to be all the propositions together with those *modal atoms* from $F_1, \ldots, F_n, Q_1, \ldots, Q_m$ that are not part of a modal atom within the same formula. *E.g.*, from $B_2 p \vee B_1 B_2 p$ we would extract two modal atoms, while from $B_1 B_2 p$ only one. Although the size of $\mathcal{T}$ may increase, in domains where different formulas in the knowledge base, build on information about the same basic beliefs of agents,

[1]We placed $c(F)$ outside of the exponential to make clear its role as a filter. Alternatively, $c(F)$ could be added to the summation in the exponent, by having it return 0 for a modally consistant theory and $-\infty$ for an inconsistent one.

we will no longer create more than necessary modally inconsistent truth assignments. More importantly, the set of queries we can answer can grow exponentially with the size of $\mathcal{A}$.

5 Connection between the Two Approaches

When using the state space of truth assignments to $F_1, \dots, F_n, Q_1, \dots, Q_m$ or to truth assignments to modal atoms and propositions $\mathcal{A}$, one might ask what happens if we keep adding query formulas (modal atoms).

We show the connection between the two approaches for defining maximum entropy distributions over epistemic situations and sets of epistemic situations. Then, we prove that in the limit of adding all possible non-equivalent modal atoms, we end up defining the same distribution as we would get if we considered all non-equivalent epistemic situations to be Σ and defined the probability distribution using (4).

We first show that for a given knowledge base and $\mathcal{A}$ (the set of propositions and modal atoms), how one can find a Σ such that both approaches will assign the same probabilities to every query formulas.

Theorem 1. *For a given knowledge base and $\mathcal{A}$ it is always possible to find a set of situations Σ s.t. every query formula will be assigned the same probability by the different approaches when using the same knowledge base.*

Proof. We can construct Σ as follows. For every consistent truth assignment t there must exist a situation σ_t, where the propositional formula $\Phi(t)$ corresponding to t (a conjunction of propositional and modal literals) is satisfied, but any other $\Phi(t')$ is not, where t' is a consistent truth assignment different from t. Since every consistent truth assignment t can decide whether a formula is satisfied or not both in the knowledge base and among the queries, the same subset of formulas in the knowledge base and the queries are satisfied in σ_t as satisfied under truth assignment t. Therefore, if

$$\Sigma = \{\sigma_t | t \in \mathcal{T} \text{ and } \Phi(t) \text{ is consistent}\} ,$$

then we have

$$\Pr(\Phi(t)) = \Pr(\sigma_t) ,$$

where on the left hand side of the equation the distribution is defined over partitions created by truth assignments while on the right hand side over situations. Consequently, every query formula will be assigned the same probability using the two different state spaces. □

We only prove here the converse of the statement for the single-agent case when Σ contains all the non-equivalent epistemic situations (for modal logics *K45*, *KD45*, or *S5*). For a given set of propositional and modal atoms $\mathcal{A}$ let $\mathcal{S}(\mathcal{A})$ be the family of equivalence classes of situations that the formulas $\Phi(t)$ – corresponding to truth assignments t to $\mathcal{A}$ – would create, *i.e.*, $\mathcal{S}(\mathcal{A})$ contains all the partitions Π_t where t ranges over all the possible truth assignments to $\mathcal{A}$. Since it is enough to consider only depth one formulas over a finite Ω, $\mathcal{A}$ is always finite in the single-agent case. (In the multi-agent case we can make a similar statement when we only consider formulas of bounded depth.).

Theorem 2. *If $\mathcal{A}$ contains all the possible propositional and modal atoms w.r.t. to modal logic K45(, KD45, or S5) then the members of Σ_{K45} (Σ_{KD45}, Σ_{S5}) are contained in $\mathcal{S}(\mathcal{A})$ as singletons.*

Proof (sketch). Since any consistent truth assignment t to $\mathcal{A}$ is capable of deciding whether any modal formula F is true or not, the number of consistent truth assignments to $\mathcal{A}$ must be equal to $|\Sigma_{K45}|$, $|\Sigma_{KD45}|$ or $|\Sigma_{S5}|$ depending on the underlying modal logic used, and since these numbers are the upper bounds on the number of non-equivalent situations every epistemic situation in a singleton of $\mathcal{S}(\mathcal{A})$ must be equivalent to a $\sigma \in \Sigma_{K45}$ ($\sigma \in \Sigma_{KD45}$ or $\sigma \in \Sigma_{S5}$). □

Note that instead of enumerating all the possible modal atoms, it suffices for $\mathcal{A}$ to contain all the modal atoms of the form $B\neg(l_1 \wedge \dots \wedge l_Q)$, where every l_i is either p_i or $\neg p_i$ where $\Omega = \{p_1, \dots, p_Q\}$. Intuitively, this way every consistent truth assignment would correspond to the situation where the real world state is described by the truth assignment to the propositions and a world with label $l_1 \wedge \dots l_Q$ would be in the set of possible worlds if $B\neg(l_1 \wedge \dots l_Q)$ is set to false, since $\neg B\neg(l_1 \wedge \dots l_Q)$ is only true if the world with label $l_1 \wedge \dots l_Q$ is among the possible worlds.

Also, we could generalize $\mathcal{A}$ to contain arbitrary formulas instead of propositional and modal atoms. Even in this setting Theorem 2 would hold if we enumerate all the (countable infinite) formulas.

The analogue could be stated for the multi-agent case for using the state space Σ_D and then taking the limit $D \to \infty$.

6 Weight Learning

The task of weight learning is that given a set of propositional modal formulas $F_1, \dots, F_n$ with their probabilities $c_1, \dots, c_n$, find a weighted set of formulas which if we use as our knowledge base then the distribution defined by (13) will satisfy the constraints $\Pr(F_i) = c_i$. Generally, we require the constraints to be consistent, hence there must exist a distribution satisfying the constraints, which is guaranteed when we are working with training data. If the constraints come from experts, we can use, *e.g.*, the framework proposed in (Papai, Ghosh, and Kautz 2012).

One more justification for defining a probability distribution over sets of epistemic situations rather than epistemic situations is that in the latter approach even simple formulas such as, *e.g.*, $B_i p$ will always have probability 0 in the limit $D \to \infty$ (assuming finite weights). For modal logic *K45* it can be shown that even if we only have 2 agents,

$$N(B_i p) = 2^{|\Omega|} 2^{2^{|\Omega|-1} 2^{|\Sigma_{D-2}|}} 2^{|\Sigma_{D-1}|} ,$$

while

$$|\Sigma_D| = 2^{|\Omega|} (2^{|\Sigma_{D-1}|})^2 ,$$

hence

$$N(B_i p)/|\Sigma_D| = 2^{2^{|\Omega|-1}|\Sigma_{D-2}| - |\Sigma_{D-1}|} .$$

Since

$$|\Sigma_{D-1}| = 2^{|\Omega|} \left(2^{|\Sigma_{D-2}|}\right)^2 ,$$

we have

$$\lim_{D\to\infty} N(B_i p)/|\Sigma_D| = 0 \ ,$$

therefore, there is no finite weight for which $\Pr(B_i p) > 0$ could be achieved. Even if we fix D, we still have the challange of dealing with large numbers during the weight learning procedure.

We now show that if $F_1, \ldots, F_n$ are the formulas in the knowledge base and if there is a distribution over epistemic situations that satisfy the given constraints, then there must be a distribution over sets of epistemic situations where the constraints are satisfied. To prove this statement, we will use the fact that if there is a distribution that satisfies $\Pr(F_i) = c_i$ then there is a maximum entropy distribution coming from an exponential family of probability distributions that satisfies this distribution (see, *e.g.*, (Wainwright and Jordan 2008)). Suppose D is finite and we learned a distribution over epistemic situations with weights $w_1, \ldots, w_n$, then the knowledge base $\{(w(t), \Phi(t))|t \in \mathcal{T}$ and $\Phi(t)$ is consistent$\}$ defines the same distribution. Comparing (7) and (6) with (11) and (10), we can conclude every $\Phi(t)$ will have the same probability if we use a knowledge base $\{(w(t) \ln N(\Phi(t)))|t \in \mathcal{T}$ and $\Phi(t)$ is consistent$\}$ to define a probability distribution over the partitions created by $\Phi(t)$. This proves that a distribution over sets of epistemic situations exist that satisfies the $\Pr(F_i) = c_i$ constraints, hence one exists when we use a knowledge base with only formulas $F_1, \ldots, F_n$.

7 Conclusions

We showed how to extend propositional Markov logic with modal operators for multiple agents. In contrast to the previous approach (Papai, Kautz, and Stefankovic 2013), we defined a probability distribution over sets of epistemic situations rather than over individual epistemic situations. Each set contained an equivalence class of epistemic situations with respect to a set of selected formulas. We gave guidelines how to select these formulas. The main advantage of our approach compared to the one described in (Papai, Kautz, and Stefankovic 2013) is that we no longer have to deal with infeasibly large numbers. Moreover, the straightforward extension of the framework of Papai et. al. (2013) suffers from the issue of simple formula probabilities converging to 0 as we increase allowed maximum depth for modal formulas; however, this does not arise in our proposed framework. Computation of the partition function (and hence performing exact inference) can be accomplished in single exponential time for all of the modal logics we discussed, regardless of the number of agents or the number of nestings of modal operators we have. Finally, inference algorithms based on sampling and designed for the standard Markov logic framework can be applied in our framework by simply adding an extra rejection step.

8 Acknowledgments

This research was supported by grants from ARO (W991NF-08-1-0242), ONR (N00014-11-10417), NSF (IIS-1012017), DOD (N00014-12-C-0263), and a gift from Intel.

References

Chellas, B. F. 1980. *Modal logic: an introduction*. Cambridge University Press.

Domingos, P., and Lowd, D. 2009. *Markov Logic: An Interface Layer for Artificial Intelligence*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.

Fagin, R.; Halpern, J. Y.; Moses, Y.; and Vardi, M. Y. 1995. *Reasoning About Knowledge*. MIT Press.

Grove, A. J.; Halpern, J. Y.; and Koller, D. 1994. Random worlds and maximum entropy. *J. Artif. Intell. Res. (JAIR)* 2:33–88.

Halpern, J., and Koller, D. 1995. Representation dependence in probabilistic inference. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, 1852–1860.

Halpern, J. Y., and Moses, Y. 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence* 54(3):319–379.

Jain, D.; Barthels, A.; and Beetz, M. 2010. Adaptive Markov Logic Networks: Learning Statistical Relational Models with Dynamic Parameters. In *19th European Conference on Artificial Intelligence (ECAI)*, 937–942.

Jaynes, E. T. 1979. *Where do we stand on Maximum Entropy?* The MIT Press. 15–118.

Koller, D., and Friedman, N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Murphy, K. P. 2012. *Machine learning: a probabilistic perspective*.

Papai, T.; Ghosh, S.; and Kautz, H. A. 2012. Combining subjective probabilities and data in training markov logic networks. In Flach, P. A.; Bie, T. D.; and Cristianini, N., eds., *ECML/PKDD (1)*, volume 7523 of *Lecture Notes in Computer Science*, 90–105. Springer.

Papai, T.; Kautz, H.; and Stefankovic, D. 2013. Reasoning under the principle of maximum entropy for modal logics K45, KD45, and S5. In *Theoretical Aspects of Rationality and Knowledge (TARK 2013)*.

Poon, H., and Domingos, P. 2006. Sound and efficient inference with probabilistic and deterministic dependencies. In *AAAI*, 458–463. AAAI Press.

Shirazi, A., and Amir, E. 2007. Probabilistic modal logic. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence.*, 489–495. Elsevier.

Shirazi, A., and Amir, E. 2008. Factored models for probabilistic modal logic. In *Proceedings of the 23rd national conference on Artificial intelligence*, AAAI'08, 541–547. AAAI Press.

Wainwright, M. J., and Jordan, M. I. 2008. *Graphical Models, Exponential Families, and Variational Inference*. Hanover, MA, USA: Now Publishers Inc.