Abstract

Model counting and weighted model counting are key problems in artificial intelligence. Marginal inference can be reduced to model counting in many statistical-relational systems, such as Markov Logic. One common approach used by model counters is splitting a theory into disjoint subtheories, performing model counting on the subtheories, and then caching the result. If an identical subtheory is encountered again in the search, the cached result is used, greatly reducing runtime. In this work we introduce a way to cache symmetric subtheories compactly, which could potentially decrease required cache size, increase cache hits, and decrease runtime of solving.

Introduction

Given a clausal propositional theory, the problem of model counting is the problem of determining the number of assignments to the variables of the theory that satisfy it. In the weighted variety of the problem, weights are assigned to the clauses (or predicates) of the theory, the weight of a possible world is the sum of the weights of the satisfied clauses (or weights of the literals assigned to true), and the weighted model counting problem is to find the sum of the weights of the satisfying possible worlds. Model counting and its variants are #P-complete problems, as it is the counting version of the SAT problem.

Model counting and its weighted variants are important problems in artificial intelligence, as many important problems we want to solve can be reduced to model counting. In probabilistic graphical models, such as Markov networks, the problem of exact marginal inference can be reduced to weighted model counting. Model counting can therefore also be used to solve marginal inference for other representations of graphical models, such as statistical-relational learning systems like Markov Logic (Richardson and Domingos 2006).

It is these systems in particular from which we draw inspiration. Although model counting, as stated, is a problem over propositional theories, it is often the case that the problems people typically want to solve are more easily expressed in a higher level language, such as first order logic. The input to the model counter is created through the process of propositionalizing (or grounding) the first order representation of the problem.

In propositional theories, it is possible to detect and leverage symmetries between variables to reduce solving and counting time. However, the number of propositional symmetries may be exponential, so using them effectively typically involves heuristically choosing which symmetries to account for. In first order theories, however, it is possible to compactly represent many symmetries and exploit them. This is the common thread that runs through many lifted inference algorithms for both marginal and MPE inference, such as e.g. (Mittal et al. 2014), (Singla and Domingos 2008) and many more.

In this paper, we will focus on symmetries over terms, that is, constants in the theory. Specifically, we will exploit the class of term equivalent symmetries, which were introduced in various forms in earlier works such as (Poole 2003; Bui, Huynh, and Riedel 2013). Two terms are term equivalent if they may be swapped without altering the theory. These symmetries can be efficiently detected in evidence (Kopp, Singla, and Kautz 2015), a quality that makes them especially well-suited to the type of problems typically formulated in systems like Markov Logic.

The main contributions of this paper are threefold: first, we introduce a canonical form, a way of representing many symmetric theories compactly, second we give a procedure for putting a theory in canonical form, and lastly, we describe a method for using the canonical form to exploit symmetries in model counting.

Related Work

Our work has connections to research in both the machine learning and constraint-satisfaction research communities. Developments include lifted versions of variable elimination (Poole 2003; de Salvo Braz, Amir, and Roth 2005), belief propagation (Singla and Domingos 2008; Singla, Nath, and Domingos 2014), and DPLL (Gogate and Domingos 2011). The approach of defining symmetries using group theory and detecting them by graph isomorphism is shared by Bui et al.’s work on lifted variational inference (Bui, Huynh, and Riedel 2013). Niepert gives a group-theoretic formalism of symmetries at the level of constants (Niepert 2012; 2013), applying them to MCMC methods. Kopp et. al de-
fine the class of symmetries we exploit in this work (Kopp, Singla, and Kautz 2015). Bui notes that symmetry groups can be defined on the basis of unobserved constants in the domain, while the symmetries we exploit can be explicitly found in the evidence. Two lines of work in SRL make use of problem transformations. First-order knowledge compilation (Van den Broeck 2013; den Broeck, Meert, and Darwiche 2014) transforms a relational problem into a form for which MPE, marginal, and MAP inference is tractable. Recent work on MAP inference in Markov Logic has identified special cases where a relational formula can be transformed by replacing a quantified formula with a single grounding of the formula (Mittal et al. 2014).

The literature surrounding the use of symmetries in constraint satisfaction and model counting is quite extensive. Here we give just a few examples of key developments in the field. Symmetry detection has been based either on graph isomorphism on propositional theories as in the original work by by Crawford et. al. (Crawford et al. 1996); by interchangeability of row and/or columns in CSPs specified in matrix form (Meseguer and Torras 2001); by checking for other special cases of geometric symmetries (Sellmann and Hentenryck 2005), or by determining that domain elements for a variable are exchangeable (Audemard, Benhamou, and Hentenryck 2006). Researchers have suggested symmetry-aware modifications to backtracking CSP solvers for variable selection, branch pruning, and no-good learning (Meseguer and Torras 2001; Flener et al. 2009). A recent survey of symmetry breaking for CSP (Walsh 2012) described alternatives to the lex-leader formulation of SBPs, including one based on Gray codes.

Symmetries over terms

In this section we provide the background necessary to understand the types of symmetries we will exploit in the subsequent sections. Symmetry-breaking for satisfiability testing, introduced by Crawford et. al. (Crawford et al. 1996), is based on concepts from group theory. A permutation σ is a mapping from a set L to itself. A permutation group is a set of permutations that is closed under composition and contains the identity and a unique inverse for every element. A literal is an atom or its negation. A clause is a disjunction over literals. A CNF theory T is a set (conjunction) of clauses. Let L be the set of literals of T. We consider only permutations that respect negation, that is σ(¬l) = ¬σ(l) (l ∈ L). The action of a permutation on a theory, written σ(T), is the CNF formula created by applying σ to each literal in T. We say σ is a symmetry of T if it results in the same theory i.e. σ(T) = T.

A model M is a truth assignment to the atoms of a theory. The action of σ on M, written σ(M), is the model where σ(M)(P) = M(σ(P)). The key property of σ being a symmetry of T is that M |= T iff σ(M) |= T. The orbit of a model M under a symmetry group Σ is the set of models that can be obtained by applying any of the symmetries in Σ. A symmetry group divides the space of models into disjoint sets, where the models in an orbit either all satisfy or all do not satisfy the theory.

In this work, we will consider symmetries over the terms (constants) of a first order relational theory. A relational theory is a tuple T = (F, E), where F is a set of first-order formulas and E is a set of evidence. We restrict the formulas in F to be built from predicates, variables, quantifiers, and logical connectives, but no constants or function symbols. E is a set of ground literals; that is, literals built from predicates and constant symbols. Universal and existential quantification is over the set of the theory’s constants D (i.e. the constants that appear in its evidence). Optionally, the formulas or predicates may be weighted.

In (Kopp, Singla, and Kautz 2015), two classes of symmetries over terms of relational theories were formalized, and methods to exploit the symmetries for MPE inference were given. The first was the set of term symmetries. A term symmetry is a permutation σ of the terms in the theory such that σ(T) = T. Note that a permutation of terms over a ground theory induces a permutation of atoms of the theory. The second class of symmetries over terms is a special case of the former called term equivalent symmetries. A term equivalent symmetry is a partitioning of the terms in the theory such that if two terms C1 and C2 appear in the same term equivalent class, they can be permuted without changing the theory. In other words, they define equivalence classes of terms. Both classes of symmetries can be detected over the evidence of a theory.

Consider a relational theory that models allocation of computational resources to distributed computing tasks. The domain might have a type for CPUs in a cluster, cores on a CPU, and computational tasks. It may have predicates that describe properties of the computational resources, as well as the requirements of the tasks. We might want to determine the best (MPE) allocation of resources to tasks, or we might use the theory to describe a stochastic scheduling algorithm, and be interested in the (marginal) probability of a task being mapped to a particular core. In this domain, all of the cores that belong to the same CPU belong to the same term equivalent symmetry group. Furthermore, if two CPUs have the same number and types of cores, then there are term symmetries that permute the CPUs while also permuting their respective cores.

The methods for exploiting these symmetries involved adding clauses to the theory that restricted the search space while guaranteeing that at least one assignment in each orbit was left in the space. For model counting, this technique cannot be applied straightforwardly, since reducing the search space will reduce the number of models. In this work, we leverage the class of term equivalent symmetries in a way that does not directly modify the search space, and is therefore suitable for model counting.

A canonical form for symmetric theories

We introduce the notion of a canonical form for ground first-order clausal theories under a set of symmetries, and define a particular canonical form for such theories under term equivalent symmetries.

Let T be a ground, first-order clausal theory, and let Σ be a set of symmetries over terms, literals, or clauses in T. Let there be an ordering on all of the possible theories. We
say that another theory \( T' \) is symmetric to \( T \) under term partition \( Z \) if there is a permutation \( \sigma \) that respects \( \Sigma \) such that \( \sigma(T) = T' \). A theory \( T \) is in canonical form if for every \( T' \) that is symmetric to \( T \) under \( \Sigma \), \( T \) comes before \( T' \) in the ordering.

First we will precisely define an ordering on ground theories. There are many possible orderings, here we give just one.

Put an ordering on the constants of the theory: \( C_1, \ldots, C_n \). Put an ordering on the predicates of the theory and their negations: \( P_1, \neg P_1, \ldots, P_k, \neg P_k \). This induces an ordering on the literals of the theory: a literal \( L \) comes before \( L' \) in the ordering if the predicate of \( L \) comes before that of \( L' \) in the ordering of predicates. If they are the same, then \( L \) comes before \( L' \) if, when considering the arguments in order, the first differing argument of \( L \) comes before that of \( L' \) in the ordering of constants. When considering the literals of a clause in turn, we always do so according to this order.

Now we define an ordering on clauses, which is actually a hierarchy of orderings. A level of the hierarchy is only considered if the comparison criteria of the above levels are equal for the clauses to be compared. The highest ordering is the ordering of weights: A clause \( C \) comes before \( C' \) in the ordering if the weight of \( C \) is less than the weight of \( C' \) i.e. \( w(C) < w(C') \). If the theory does not have weighted clauses, this level is omitted. The next ordering is the ordering on clause lengths. A clause \( C \) comes before \( C' \) in the ordering if \( w(C) = w(C') \) and \( C \) has fewer literals than \( C' \), i.e. \( l(C) < l(C') \). The final ordering is the ordering on literals. A clause \( C \) comes before \( C' \) in the ordering if \( w(C) = w(C'), l(C) = l(C') \), and the first differing literal of \( C \) comes before that of \( C' \) in the ordering of literals.

Finally we define the ordering on theories. Given two theories, we consider each clause of the theories in the order defined above. A theory \( T \) comes before \( T' \) in the ordering if the first differing clause of \( T \) comes before that of \( T' \) in the ordering of clauses.

Now that we have an ordering on the theories, we define a canonical form for term equivalent symmetries. Let \( T \) be a ground, first-order clausal theory, and \( Z \) be a term equivalent partition of the constants in \( T \). We say that \( T \) is in canonical form if for every \( T' \) that is symmetric to \( T \) under \( Z \), \( T \) comes before \( T' \) in the ordering of theories.

A procedure for canonicalization

Now that we have a precisely defined canonical form for theories under term equivalent partitions, we give a procedure for converting a theory \( T \) to canonical form under term equivalent partition \( Z \). First, for each clause \( C \) in \( T \), we reorder the literals of \( C \) so that they appear in the ordering of literals defined in the previous section. If the length of the clause is \( l(C) \), then this takes time \( O(l(C) \log(l(C))) \). Let \( l \) be the length of the longest clause, and \(|T|\) be the number of clauses in the theory. To complete this step for each clause in the theory, it takes time \( O(|T| \cdot l \log(l)) \). Next, we reorder the clauses of \( T \) according to the ordering of clauses defined in the previous section. This takes time \( O(|T| \log(|T|)) \).

Next, we iterate over the arguments of the predicates of the theory in order, building a permutation \( \sigma \). For each argument \( C_i \) we encounter we do the following: If \( C_i \) has not been encountered before, we find the first constant \( C_i' \) in the ordering that is term equivalent to \( C_i \) under \( Z \), add \( C_i \rightarrow C_i' \) to \( \sigma \), and delete \( C_i' \) from \( Z \). If \( C_i \) was encountered before (or equivalently, if \( C_i \) appears on the left hand side of \( \sigma \)), it is skipped. Note that because we delete \( C_i' \) from \( Z \), no constant will appear on the right hand side of \( \sigma \) more than once. Likewise, because we skip constants that have already been considered, no constant will appear on the left hand side of \( \sigma \) more than once. The procedure ends when either all of the arguments have been iterated over, or when every constant in the theory appears on both sides of \( \sigma \).

If we store the term equivalent partitions in sorted arrays, we can use a simple pointer to make the lookup and deletion of \( C_i' \) a constant time operation. Thus, if the number of arguments appearing in the theory is \( a \), then this step takes time \( O(a) \).

Finally, to return the canonical form, we simply apply the permutation we built to the input theory, i.e. \( T' = \sigma(T) \). This takes time proportional to the number of terms in the theory, so \( O(a) \). Applying the permutation can actually be done in-place while computing it.

**Theorem 1.** The theory \( T' \) found by the above procedure is the canonical form of \( T \) under \( Z \).

**Proof.** This proof takes two parts. First we will prove that \( T' \) is equivalent to \( T \) under \( Z \), then we will prove that \( T' \) is the first such equivalent theory. By construction, the permutation \( \sigma \) produced by the above procedure only permutes a term to a term that appears in the same term equivalent partition defined by \( Z \). Furthermore, \( \sigma \) is one-to-one and onto. Therefore applying \( \sigma \) yields a term equivalent theory under \( Z \).

Next we show that \( T' \) is the first such theory in the ordering. Suppose not. Suppose there was a theory \( T_0 \) that was term equivalent with \( T \) and appeared before \( T' \) in the ordering. By symmetry, \( T_0 \) is equivalent to \( T' \) under \( Z \). Furthermore, by symmetry, the clauses of \( T' \) and \( T_0 \) have the same weights and predicate signatures. By the definition of the ordering, the first pair of terms that differ between \( T_0 \) and \( T' \), \( C_0 \) and \( C' \) respectively, must be such that \( C_0 \) comes before \( C' \) in the ordering. However, if that were true, then we must have chosen a term from a partition of \( Z \) that was not least in the ordering when building \( T' \). This is not how the procedure is defined. Contradiction. Therefore \( T' \) is the first theory in the ordering that is equivalent to \( T \) under \( Z \).

Therefore \( T' \) is the canonical form of \( T \).

Next we give a worked example. We apply the given procedure to two symmetric theories, and find that they have the same canonical form.

**Example 1** Suppose that after detecting term equivalent symmetries from evidence, we found the term equivalent
partition to be $Z = \{\{A, B, C\}, \{X, Y, Z\}\}$.

$P(A) \lor Q(Y)$  &  $P(B) \lor Q(X)$

$Q(Z)$ & $Q(Y)$

$P(C) \lor Q(X) \lor P(Z)$ & $P(A) \lor Q(Z) \lor P(Y)$

Let the ordering on terms and predicates be alphabetical. We first rewrite the theory according to the induced ordering.

$Q(Z)$ & $Q(Y)$

$P(A) \lor Q(Y)$ & $P(B) \lor Q(X)$

$P(C) \lor P(Z) \lor Q(X)$ & $P(A) \lor P(Y) \lor Q(Z)$

We will perform the permutation in place. We iterate over the terms in the theory. On the left, we have $Z \rightarrow X$, and on the right, $Y \rightarrow X$. We apply this partial symmetry to the first term.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(B) \lor Q(X)$

$P(C) \lor P(Z) \lor Q(X)$ & $P(A) \lor P(Y) \lor Q(Z)$

We continue to the next term. On the left, it is already the least term, so we have $A \rightarrow A$. On the right, $B \rightarrow A$.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(A) \lor Q(X)$

$P(C) \lor P(Z) \lor Q(X)$ & $P(A) \lor P(Y) \lor Q(Z)$

On the left, $Y$ is already the least term, so we have $Y \rightarrow Y$. On the right, $X \rightarrow Y$.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(A) \lor Q(X)$

$P(C) \lor P(Z) \lor Q(X)$ & $P(A) \lor P(Y) \lor Q(Z)$

Next term, $C \rightarrow B$ on the left and $A \rightarrow B$ on the right.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(A) \lor Q(Y)$

$P(B) \lor P(Z) \lor Q(X)$ & $P(B) \lor P(Y) \lor Q(Z)$

Next term, we apply the $Z \rightarrow X$ from before on the left, and the $Y \rightarrow X$ from before on the right.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(A) \lor Q(Y)$

$P(B) \lor P(X) \lor Q(X)$ & $P(B) \lor P(X) \lor Q(Z)$

Last term, $X \rightarrow Z$ on the left, and $Z \rightarrow Z$ on the right.

$Q(X)$ & $Q(X)$

$P(A) \lor Q(Y)$ & $P(A) \lor Q(Y)$

$P(B) \lor P(X) \lor Q(Z)$ & $P(B) \lor P(X) \lor Q(Z)$

We see that the two different theories have the same canonical form. By inspection, we see that they are equivalent under the set of term symmetries.

### Applying canonical form in model counting

Exact model counters such as RelSAT (Bayardo Jr. and Schrag 1997) and Cachet (Sang et al. 2004) leverage a caching procedure to count models more efficiently. Consider an arbitrary place in the search space, consisting of a partial assignment to the variables and a simplified version of theory. In a systematic search, the algorithm will count the number of models in this simplified theory, then backtrack to explore other parts of the search space. It is possible that, at a different point in the search space, the simplified theory will be identical to the simplified theory we previously counted the models of. Thus, these algorithms, upon counting the models in a simplified theory, will cache the simplified theory and the number (or weight) of models it has in a lookup table. Then, if it is encountered again, the value can be reused without performing the counting again.

This technique is even more effective when used in conjunction with theory decomposition. These systematic search algorithms will, at a given point in the search space, detect if the simplified theory can be split into two or more disjoint subtheories. A theory $T$ can be partitioned in such a way if it contains two or more disjoint subsets of it’s clauses $T_1, \ldots, T_k$ such that none of the $T_i$ share any variables. Since they share no variables, their models can be counted independently. In a caching algorithm, it is the disjoint subtheories and their counts that are stored in the lookup table.

One problem with this method is knowing which subtheories to cache. If we cache every subtheory, our cache will take up too much memory. If too few, we may do redundant computation. Furthermore, algorithms that are not aware of symmetries will perform model counting on subtheories that are symmetrical, despite the fact that the result will be the same. We propose a method of leveraging canonical form to reduce the number of symmetrical subtheories that are recomputed. Furthermore, the technique will reduce the memory needed to cache many subtheories.

With the foundation provided in the previous section, the technique is actually very simple. We detect the term equivalent partition once before we begin the search. Each time a theory is split into disjoint subtheories, we convert the subtheories into canonical form with respect to the partition. If the canonical form of the subtheory does not appear in the cache we compute the number of models and cache the result with the canonical form as the key. If it does appear in the cache, that means the original subtheory is identical to or term equivalent to a subtheory that we saw previously. We save space in the cache because each cached theory is actually representative of every theory symmetric to it under the term equivalent partition. Furthermore, we reduce redundant computations because when we count the number of models in a subtheory, we reuse the value for every theory that is symmetric to it under the term equivalent partition, rather than just subtheories that are exactly identical to it. It is hypothesized that leveraging this technique will result in significant speedups in high-symmetry domains.
Conclusion and future work

We have provided a canonical form for ground first order CNF theories under a term equivalent partitioning, as well as an efficient procedure for converting a theory to canonical form. We then described a simple technique to increase the solving efficiency of exact model counters by leveraging this canonical form.

Short term future work includes implementing this system by modifying an existing exact model counter, and comparing the technique with the stock version, as well as other algorithms that leverage these types of symmetries. Longer term goals include creating a canonical form and efficient procedure that canonicalizes theories under a set of term symmetries, a class of symmetries of which term equivalent symmetries are a special case.

References


2014, Québec City, Québec, Canada., 2497–2504. AAAI Press.