

---

# Lifted Marginal MAP Inference\*

---

Vishal Sharma<sup>1</sup>, Noman Ahmed Sheikh<sup>2</sup>, Happy Mittal<sup>1</sup>, Vibhav Gogate<sup>3</sup> and Parag Singla<sup>1</sup>

<sup>1</sup>IIT Delhi, {vishal.sharma, happy.mittal, parags}@cse.iitd.ac.in

<sup>2</sup>IIT Delhi, nomanahmedsheikh@outlook.com

<sup>3</sup>UT Dallas, vgogate@hlt.utdallas.edu

## Abstract

Lifted inference reduces the complexity of inference in relational probabilistic models by identifying groups of constants (or atoms) which behave symmetric to each other. A number of techniques have been proposed in the literature for lifting marginal as well MAP inference. We present the first application of lifting rules for marginal-MAP (MMAP), an important inference problem in models having latent (random) variables. Our main contribution is two fold: (1) we define a new equivalence class of (logical) variables, called Single Occurrence for MAX (SOM), and show that solution lies at extreme with respect to the SOM variables, i.e., predicate groundings differing only in the instantiation of the SOM variables take the same truth value (2) we define a sub-class *SOM-R* (SOM Reduce) and exploit properties of extreme assignments to show that MMAP inference can be performed by reducing the domain of SOM-R variables to a single constant. We refer to our lifting technique as the *SOM-R* rule for lifted MMAP. Combined with existing rules such as decomposer and binomial, this results in a powerful framework for lifted MMAP. Experiments on three benchmark domains show significant gains in both time and memory compared to ground inference as well as lifted approaches not using SOM-R.

## 1 INTRODUCTION

Several real world applications such as those in NLP, vision and biology need to handle non-i.i.d. data as well as represent uncertainty. Relational Probabilistic mod-

els (Getoor and Taskar 2007) such as Markov logic networks (Domingos and Lowd 2009) combine the power of relational representations with statistical models to achieve this objective. The naïve approach to inference in these domains grounds the relational network into a propositional one and then applies existing inference techniques. This can often result in sub-optimal performance for a large number of applications since inference is performed oblivious to the underlying network structure.

Lifted inference (Kimmig, Mihalkova, and Getoor 2015) overcomes this shortcoming by collectively reasoning about groups of constants (atoms) which are identical to each other. Starting with the work of Poole (Poole 2003), a number of lifting techniques which lift propositional inference to the first-order level have been proposed in literature. For instance, for marginal inference, exact algorithms such as variable elimination and AND/OR search and approximate algorithms such as belief propagation and MCMC sampling have been lifted to the first-order level (cf. (de Salvo Braz, Amir, and Roth 2005; Gogate and Domingos 2011; G. Van den Broeck et al. 2011; Kersting, Ahmadi, and Natarajan 2009; Singla and Domingos 2008; Niepert 2012; Venugopal and Gogate 2012)). More recently, there has been increasing interest in lifting MAP inference (both exact and approximate) (Sarkhel et al. 2014; Mittal et al. 2014; Mladenov, Kersting, and Globerson 2014). Some recent work has looked at the problem of approximate lifting i.e., combining together those constants (atoms) which are similar but not necessarily identical (Van den Broeck and Darwiche 2013; Singla, Nath, and Domingos 2014; Sarkhel, Singla, and Gogate 2015).

Despite a large body of work on lifted inference, to the best of our knowledge, there is no work on lifted algorithms for solving marginal maximum-a-posteriori (MMAP) queries. MMAP inference is ubiquitous in real-world domains, especially those having latent (random) variables. It is well known that in many real-world do-

---

\*Paper accepted in UAI-18 (Sharma et al. 2018).

mains, the use of latent (random) variables significantly improves the prediction accuracy (Maaten, Welling, and Saul 2011). Moreover, the problem also shows up in the context of SRL domains in tasks such as plan and activity recognition (Singla and Mooney 2011). Therefore, efficient lifted methods for solving the MMAP problem are quite desirable.

MMAP inference is much harder than marginal (sum) and MAP (max) inference because sum and max operators do not commute. In particular, latent (random) variables need to be marginalized out before MAP assignment can be computed over the query (random) variables and as a result MMAP is NP-hard even on tree graphical models (Park 2002). Popular approaches for solving MMAP include variational algorithms (Liu and Ihler 2013), AND/OR search (Marinescu, Dechter, and Ihler 2014) and parity solvers (Xue et al. 2016).

In this paper, we propose the first ever lifting algorithm for MMAP by extending the class of lifting rules (Jha et al. 2010; Gogate and Domingos 2011; Mittal et al. 2014). As our first contribution, we define a new equivalence class of (logical) variables called *Single Occurrence for MAX (SOM)*. We show that the MMAP solution lies at extreme with respect to the SOM variables, i.e., predicate groundings which differ only in the instantiation of the SOM variables take the same truth (true/false) value in the MMAP assignment. The proof is fairly involved due to the presence of both MAX and SUM operations in MMAP, and involves a series of problem transformations followed by exploiting the convexity of the resulting function.

As our second contribution, we define a sub-class of SOM, referred to as *SOM-R (SOM Reduce)*. Using the properties of extreme assignments, we show that the MMAP solution can be computed by reducing the domain of SOM-R variables to a single constant. We refer to this as *SOM-R rule for lifted MMAP*. SOM-R rule is often applicable when none of the other rules are, and can result in significant savings since inference complexity is exponential in the domain size in the worst case.

Finally, we show how to combine SOM-R rule along with other lifting rules e.g., binomial and decomposer, resulting in a powerful algorithmic framework for lifted MMAP inference. Our experiments on three different benchmark domains clearly demonstrate that our lifting technique can result in orders of magnitude savings in both time and memory compared to ground inference as well as vanilla lifting (not using the SOM-R rule).

## 2 BACKGROUND

**First-Order Logic:** The language of first-order logic (Russell and Norvig 2010) consists of *constant*,

*variable*, *predicate*, and *function* symbols. A *term* is a variable, constant or is obtained by application of a function to a tuple of terms. Variables in first-order logic are often referred to as *logical variables*. We will simply refer to them as variables, henceforth. A *predicate* defines a relation over the set of its arguments. An *atom* is obtained by applying a predicate symbol to the corresponding arguments. A *ground atom* is an atom having no variables in it. Formulas are obtained by combining predicates using a set operators:  $\wedge$  (and),  $\vee$  (or) and  $\neg$  (not). Variables in a formula can be universally or existentially quantified using the operators  $\forall$  and  $\exists$ , respectively. A first-order theory (knowledge base) is a set of formulas. We will restrict our attention to function free finite first-order logic with Herbrand interpretation (Russell and Norvig 2010) and universally quantified variables. In the process of (partially) grounding a theory, we replace all (some) of the universally quantified variables with the possible constants in the domain. In the following, we will use capital letters (e.g.,  $X$ ,  $Y$  etc.) to denote logical variables and small case letters to denote constants. We will use  $\Delta_X = \{x_1, x_2, \dots, x_m\}$  denotes the domain of variable  $X$ .

**Markov Logic:** A Markov logic network (Domingos and Lowd 2009) (MLN)  $M$  is defined as a set of pairs  $\{f_i, w_i\}_{i=1}^n$  where  $f_i$  is a formula in first-order logic and  $w_i$  is the weight of  $f_i$ . We will use  $F(M)$  to denote the set of all the formulas in MLN. Let  $\mathcal{X}$  denote the set of all the logical variables appearing in MLN. An MLN can be seen as a template for constructing ground Markov networks. Given the domain  $\Delta_X$  for every variable  $X \in \mathcal{X}$ , the ground network constructed by MLN has a node for every ground atom and a feature for every ground formula. Let  $\mathcal{T}$  denote the set of all the predicates appearing in  $M$ . We will use  $\mathcal{T}_g$  to denote all the ground atoms corresponding to the set  $\mathcal{T}$  and  $t$  to denote an assignment, i.e. a vector of true/false values, to  $\mathcal{T}_g$ . The distribution specified by an MLN is given as:

$$P(\mathcal{T}_g = t) = \frac{1}{Z} e^{\sum_{i=1}^n \sum_{j=1}^{m_i} w_i f_{ij}(t)} \quad (1)$$

where  $m_i$  denotes the number of groundings of the  $i^{th}$  formula.  $f_{ij}$  represents the feature corresponding to the  $j^{th}$  grounding of the  $i^{th}$  formula. The feature is on if the corresponding formula is satisfied under the assignment  $t$  off otherwise.  $Z$  is the normalization constant. Equivalently, in the potential function representation, the distribution can be written as:

$$P(t) = \frac{1}{Z} \prod_{i=1}^n \prod_{j=1}^{m_i} \phi_{ij}(t) \quad (2)$$

where there is a potential  $\phi_{ij}$  for each  $f_{ij}$  such that  $\phi_{ij}(t) = e^{w_i f_{ij}(t)}$ .

**Marginal MAP (MMAP):** Let the set of all predicates  $\mathcal{T}$  be divided into two disjoint subsets  $\mathcal{Q}$  and  $\mathcal{S}$ , referred to as MAX and SUM predicates, respectively. Let  $q$  (resp.  $s$ ) denote an assignment to all the groundings of the predicates in  $\mathcal{Q}$  (resp.  $\mathcal{S}$ ). Note that  $\mathcal{T} = \mathcal{Q} \cup \mathcal{S}$ , and given assignment  $t$  to  $\mathcal{T}$ ,  $t = q \cup s$ . Then, the *marginal-MAP (MMAP)* problem for MLNs can be defined as:

$$\arg \max_q \sum_s \prod_{i=1}^n \prod_{j=1}^{m_i} \phi_{ij}(q, s) = \arg \max_q W_M(q) \quad (3)$$

$$\text{where, } W_M(q) = \sum_s \prod_{i=1}^n \prod_{j=1}^{m_i} \phi_{ij}(q, s)$$

$W_M(q)$  is referred to as the MMAP objective function for the MLN  $M$ , and its solution  $q^* = \arg \max_q W_M(q)$  is referred as the MMAP solution. Note that we can get rid of  $Z$  in equation 3, since we are only interested in finding the maximizing assignment and  $Z$  is a constant.

**Preliminaries:** We will assume that our MLN is in Normal Form (Mittal et al. 2014) i.e., (a) no constants appear in any of the formulae (b) if  $X$  and  $Y$  appear at the same predicate position in one or more formulae, then  $\Delta_X = \Delta_Y$ . Any MLN can be converted into normal form by a series of mechanical operations. We will also assume that formulas are standardized apart i.e., we rename the variables such that the sets of variables appearing in two different formulae are disjoint with each other. We define an equivalence relation  $\sim$  over the set of variables such that  $X \sim Y$  if (a)  $X$  and  $Y$  appear at the same predicate position OR (b)  $\exists Z$  such that  $X \sim Z$  and  $Y \sim Z$ . We will use  $\tilde{X}$  to denote the equivalence class corresponding to variable  $X$ . Variables in the same equivalence class must have the same domain due to the normal form assumption. We will use  $\Delta_{\tilde{X}}$  to refer to the domain of the variables belonging to  $\tilde{X}$ .

Finally, though our exposition in this work is in terms of MLNs, our ideas can easily be generalized to other representations such as weighted parfactors (de Salvo Braz, Amir, and Roth 2005) and probabilistic knowledge bases (Gogate and Domingos 2011).

### 3 SINGLE OCCURRENCE FOR MMAP

#### 3.1 Motivation

In this work, we are interested in lifting the marginal-MAP (MMAP) problem. Since MMAP is a problem harder than both marginal and MAP inference, a natural question to examine would be if existing lifting techniques for MAP and marginal inference can be extended to the case of MMAP. Or further still, if additional rules can be discovered for lifting the MMAP problem. Whereas many of the existing rules such as decomposer

and binomial<sup>1</sup> (Jha et al. 2010; Mittal et al. 2015) extend in a straightforward manner for MMAP, unfortunately the SO rule (Mittal et al. 2014), which is a powerful rule for MAP inference, is not directly applicable.

In response, we propose a new rule, referred to as Single Occurrence for MAX Reduce (SOM-R), which is applicable for MMAP inference. We first define a variable equivalence class, referred to as SOM, which requires that (1) no two variables in the class appear in the same formula (2) at least one of the variables in the class appears in a MAX predicate. We further define a sub-class of SOM, referred to as SOM-R, which imposes a third condition (3) either all the SUM predicates in the theory contain a SOM variable or none of them does. Our SOM-R rule states that domain of SOM-R variables can be reduced to a single constant for MMAP inference. Consider the following example MLN, henceforth referred to as  $M_1$ :

$$\begin{aligned} w_1 &: Frnds(X, Y) \wedge Parent(Z, X) \Rightarrow Knows(Z, Y) \\ w_2 &: Knows(U, V) \\ SUM &: Parent \quad MAX : Frnds, Knows \end{aligned}$$

The equivalence classes in this example are given by  $\{X\}$ ,  $\{Y, V\}$  and  $\{Z, U\}$ . It is easy to see that each of these equivalence classes satisfy the three conditions above and hence, SOM-R rule can be applied over them. This makes the MMAP inference problem independent of the size of the domain and hence, it can be solved in  $O(1)$  time. Ground inference has to deal with  $O(m^2)$  number of ground atoms resulting in  $O(\exp(cm^2))$  complexity in the worst case<sup>2</sup>, where  $c$  is a constant. Further, in the absence of the SOM-R rule, none of the existing lifting rules apply and one has to resort to partial grounding again resulting in worst case exponential complexity.

We note that conditions for identifying SOM and SOM-R specifically make use of the structure of the MMAP problem. Whereas condition 1 is same as Mittal et al.'s SO condition, condition 2 requires the variables in the SOM class to belong to a MAX predicate. Condition 3 (for SOM-R) further refines the SOM conditions so that domain reduction can be applied.

We prove the correctness of our result in two phases. First, we show that SOM equivalence class implies that MMAP solution lies at extreme, meaning that predicate groundings differing only in the instantiation of the SOM class take the same truth value. Second, for the sub-class SOM-R, we further show that domain can be reduced to a single constant for MMAP. Here, we rely on the properties of extreme assignments.

<sup>1</sup>applicable when the binomial predicate belongs to MAX

<sup>2</sup>Inference complexity is exponential in the number of ground atoms. Here, we assume  $|\Delta_X| = |\Delta_Y| = |\Delta_Z| = m$

Our proof strategy makes use of a series of problem transformations followed by using the convexity of the resulting function. These algebraic manipulations are essential to prove the correctness of our result, and are some of the important contributions of our paper. Next, we describe each step in detail. The proofs of theorems (and lemmas) marked with (\*) are in the supplement.

### 3.2 SOM implies Extreme Solution

We introduce some important definitions. We will assume that we are given an MLN  $M$ . Further, we are interested in solving an MMAP problem over  $M$  where the set of MAX predicates is given by  $\mathcal{Q}$ .

**Definition 1.** (Single Occurrence for MAX) We say that a variable equivalence class  $\tilde{X}$  is Single Occurrence for MAX (SOM) if (a)  $\forall i, f_i \in F(M)$ , there is at most one variable from the set  $\tilde{X}$  occurring in  $f_i$  (b) there exists a variable  $X \in \tilde{X}$  and a predicate  $P \in \mathcal{Q}$ , such that  $X$  appears in  $P$ .

Next, we define the notion of an extreme assignment.

**Definition 2.** (Extreme Assignment) Let  $\tilde{X}$  be a variable equivalence class. An assignment  $q$  to MAX predicates  $\mathcal{Q}$  lies at extreme (with respect to  $\tilde{X}$ ), if  $\forall P \in \mathcal{Q}$ , all the groundings of  $P$  with the same instantiation to variables  $\mathcal{X} - \tilde{X}$ , take the same value in  $q$ .

In  $M_1$ , an extreme assignment with respect to variable equivalence class  $\{Y, V\}$  will assign the same truth value to the ground atoms  $Knows(z, y_1)$  and  $Knows(z, y_2)$ ,  $\forall z \in \Delta_Z$  and  $\forall y_1, y_2 \in \Delta_Y$ . We next define the notion of an MLN variablized with respect to a variable equivalence class.

**Definition 3.** (Variablized MLN) Let  $\tilde{X}$  be an equivalence class. Let  $M_{\tilde{X}}$  be the MLN obtained by instantiating (grounding) the variables in the set  $\mathcal{X} - \tilde{X}$ . We say that  $M_{\tilde{X}}$  is variablized (only) with respect to the set  $\tilde{X}$ .

For instance in  $M_1$ , variablizing with respect to the equivalence class  $\{Y, V\}$  results in MLN with formulas similar to:

$$\begin{aligned} w_1 &: Frnds(x, Y) \wedge Parent(z, x) \Rightarrow Knows(z, Y) \\ w_2 &: Knows(u, V) \end{aligned}$$

where  $x, z$  and  $u$  are constants belonging to respective domains.  $Frnds(x, Y)$ ,  $Knows(z, Y)$  and  $Knows(u, V)$  can be treated as unary predicates over the equivalence class  $\{Y, V\}$  since  $x, z$  and  $u$  are constants. Similarly,  $Parent(z, x)$  can be treated as a propositional predicate.

It is important to note that,  $M_{\tilde{X}}$  represents the same distribution as  $M$ . Further,  $M_{\tilde{X}}$  can be converted back into normal form by introducing a new predicate for every

combination of constants appearing in a predicate. We now define one of the main theorems of this paper.

**Theorem 1.** Let  $M$  be an MLN and let  $\tilde{X}$  be a SOM equivalence class. Then, an MMAP solution for  $M$  lies at extreme with respect to  $\tilde{X}$ .

We will prove the above theorem by defining a series of problem transformations. In the following, we will work with MLN  $M$  and  $\tilde{X}$  as a SOM variable equivalence class. We will use  $\mathcal{Q}$  and  $\mathcal{S}$  to denote set of MAX and SUM predicates, respectively.  $q$  and  $s$  will denote the assignments to respective predicate groundings (see Background (section 2)).

#### 3.2.1 Problem Transformation (PT) 1

**Objective PT1:** Convert MMAP objective into a form which only has unary and propositional predicates.

**Lemma 1.** Let  $M_{\tilde{X}}$  denote the MLN variablized with respect to SOM equivalence class  $\tilde{X}$ . Then,  $M_{\tilde{X}}$  contains only unary and propositional predicates. Further, the MMAP objective can be written as:

$$\arg \max_q W_M(q) = \arg \max_q W_{M_{\tilde{X}}}(q)$$

The proof that  $M_{\tilde{X}}$  only has unary and propositional predicates follows immediately from the definition of  $M_{\tilde{X}}$  (defn. 3) and the fact that  $\tilde{X}$  is SOM. Further, since  $M$  and  $M_{\tilde{X}}$  define the same distribution, we have the equivalence of the MMAP objectives. Since,  $M_{\tilde{X}}$  only has unary and propositional predicates, we will split the assignment  $q$  to groundings of  $\mathcal{Q}$  into  $(q_u, q_p)$  where  $q_u$  and  $q_p$  denote the assignments to groundings of unary and propositional predicates, respectively. Similarly, for assignment  $s$  to groundings of  $\mathcal{S}$ , we split  $s$  as  $(s_u, s_p)$ .

#### 3.2.2 Problem Transformation 2

**Objective PT2:** In the MMAP objective, get rid of propositional MAX predicates.

**Lemma 2.\*** Consider the MMAP problem over  $M_{\tilde{X}}$ . Let  $q_p$  be some assignment to propositional MAX predicates. Let  $M'_{\tilde{X}}$  be an MLN obtained by substituting the truth value in  $q_p$  for propositional predicates. Then, if  $M'_{\tilde{X}}$  has a solution at extreme for all possible assignments of the form  $q_p$  then,  $M_{\tilde{X}}$  also has a solution at extreme.

Therefore, in order to prove the extrema property for  $M_{\tilde{X}}$ , it is sufficient to prove it for a generic MLN  $M'_{\tilde{X}}$ , i.e., without making any assumptions on the form of  $q_p$ .

For ease of notation, we will drop the prime in  $M'_{\tilde{X}}$  and simply refer to it as  $M_{\tilde{X}}$ . Therefore, we need to show that the solution to the following problem lies at extreme:

$$\arg \max_{q_u} W_{M_{\tilde{X}}}(q_u)$$

where the propositional MAX predicates have been gotten rid of in  $M_{\tilde{X}}$ .

### 3.2.3 Problem Transformation 3

**Objective PT3:** In the MMAP objective, get rid of unary SUM predicates using inversion elimination (de Salvo Braz, Amir, and Roth 2005).

First, we note that the MMAP objective:

$$W_{M_{\tilde{X}}}(q_u) = \sum_{s_p, s_u} \prod_{i=1}^n \prod_{j=1}^{m_i} \phi_{ij}(q_u, s_p, s_u)$$

can be equivalently written as:

$$W_{M_{\tilde{X}}}(q_u) = \sum_{s_p, s_u} \prod_{i=1}^n \prod_{j=1}^m \phi'_{ij}(q_u, s_p, s_u)$$

where  $m = |\Delta_{\tilde{X}}|$ .  $\phi'_{ij}(q_u, s_p, s_u) = \phi_{ij}(q_u, s_p, s_u)$  if  $f_i$  contains a variable from  $\tilde{X}$ , else  $\phi'_{ij}(q_u, s_p, s_u) = \phi_{ij}(q_u, s_p, s_u)^{\frac{1}{m}}$  otherwise. It is easy to see this equivalence since the only variables in the theory are from the class  $\tilde{X}$ . When  $f_i$  contains a variable from  $\tilde{X}$ , it has exactly  $m_i = m$  groundings since  $\tilde{X}$  is SOM. On the other hand, if  $f_i$  does not contain a variable from  $\tilde{X}$ , it only contains propositional predicates. Then we raise it to power  $\frac{1}{m}$ , and then multiply  $m$  times in the latter expression to get an equivalent form.

Next, we use inversion elimination (de Salvo Braz, Amir, and Roth 2005) to get rid of unary SUM predicates.

**Lemma 3.** *MMAP problem over  $M_{\tilde{X}}$  can be written as:*

$$\arg \max_{q_u} W_{M_{\tilde{X}}}(q_u) = \arg \max_{q_u} \sum_{s_p} \prod_{j=1}^m \Theta_j(q_u, s_p)$$

where  $\Theta_j$  is a function of unary MAX and propositional SUM predicates groundings  $q_u$  and  $s_p$ , respectively.

*Proof.* We can write the MMAP objective  $W_{M_{\tilde{X}}}(q_u)$  as:

$$\begin{aligned} &= \sum_{s_p, s_u} \prod_{i=1}^n \prod_{j=1}^m \phi'_{ij}(q_u, s_p, s_u) \\ &= \sum_{s_p, s_u} \prod_{j=1}^m \prod_{i=1}^n \phi'_{ij}(q_u, s_p, s_u) \\ &= \sum_{s_p, s_u} \prod_{j=1}^m \Phi_j(q_u, s_p, s_u) \\ &= \sum_{s_p} \sum_{s_{u_1}, s_{u_2}, \dots, s_{u_m}} \prod_{j=1}^m \Phi_j(q_u, s_p, s_{u_j}) \end{aligned} \quad (4)$$

(apply inversion elimination)

$$\begin{aligned} &= \sum_{s_p} \prod_{j=1}^m \sum_{s_{u_j}} \Phi_j(q_u, s_p, s_{u_j}) \\ &= \sum_{s_p} \prod_{j=1}^m \Theta_j(q_u, s_p) \end{aligned}$$

**Proof Explanation:** Second equality is obtained by interchanging the two products. Third equality is obtained by defining  $\prod_i \phi'_{ij}(q_u, s_p, s_u) = \Phi_j(q_u, s_p, s_u)$ . In fourth equality, we have made explicit the dependence of  $\Phi_j$  on  $s_{u_j}$  i.e. the groundings corresponding to the  $j^{\text{th}}$  constant.

**Inversion Elimination** (de Salvo Braz, Amir, and Roth 2006): Since  $\Phi_j$  only depends on  $s_{u_j}$  (among  $s_u$ ) groundings, we can use inversion elimination to invert the sum over  $s_{u_j}$  and product over  $j$  in the fifth equality.

**Final Expression:** We define  $\Theta_j(q_u, s_p) = \sum_{s_{u_j}} \Phi_j(q_u, s_p, s_{u_j})$ .

Note that, at this point, we have only propositional SUM and unary MAX predicates in the transformed MMAP objective.

### 3.2.4 Problem Transformation 4

**Objective PT4:** Exploit symmetry of the potential functions in the MMAP objective.

We rename  $q_u$  to  $q$  and  $s_p$  to  $s$  for ease of notation in Lemma 3. The MMAP objective can be written as:

$$W_{M_{\tilde{X}}}(q) = \sum_s \prod_{j=1}^m \Theta_j(q_j, s) \quad (5)$$

Here,  $q = (q_1, q_2, \dots, q_m)$  and  $q_j$  represents the assignment to the unary MAX predicate groundings corresponding to constant  $j$ . In the expression above, we have made explicit the dependence of  $\Theta_j$  on  $q_j$ . We make the following two observations.

1) Due to the normal form assumption, all the groundings of a first-order logic formula behave identical to each other (up to renaming of constants). Hence, the resulting potential function  $\Theta_j$ 's are also identical to each other.

2) If there are  $r$  unary MAX predicates in  $M_{\tilde{X}}$ , then each  $q_j$  can take  $R = 2^r$  possible values<sup>3</sup>.

Therefore, the value of the product  $\prod_{j=1}^m \Theta_j(q, s)$  in the RHS of Equation 5 depends only on the number of different types of values  $q_j$ 's take in  $q$  (and not on which  $q_j$  takes which value). Let  $\{v_1, v_2, \dots, v_R\}$  denote the set of  $R$  different values that  $q_j$ 's can take. Given a value  $v_l$ ,

<sup>3</sup>since there are  $r$  predicate groundings for each  $j$  and each is Boolean valued

let  $N_l$  denote the number of times  $v_l$  appears in  $q$ . Next, we state the following lemma.

**Lemma 4.** *The MMAP problem can be written as:*

$$\arg \max_q W_{M_{\tilde{X}}}(q) = \arg \max_{N_1, N_2, \dots, N_R} \sum_s \prod_{l=1}^R f_l(s)^{N_l}$$

subject to the constraints that  $\forall l, N_l \geq 0, N_l \in \mathbb{Z}$  and  $\sum_l N_l = m$ . Here,  $f_l(s) = \Theta_j(v_l, s)$ .

*Proof.* Proof follows from the fact that  $\Theta_j$ 's are symmetric to each other and that the  $q_j$ 's take a total of  $m$  possible (non-unique) assignments since  $\Delta_{\tilde{X}} = m$ .

We say that an assignment  $N_1, N_2, \dots, N_R$  subject to the constraints:  $\forall l, N_l \geq 0$  and  $\sum_l N_l = m$  is at *extreme* if  $\exists l$  such that  $N_l = m$ . Note that for  $R \geq 2$ , extreme assignment also implies that  $\exists l, N_l = 0$ . We have the following lemma.

**Lemma 5.\*** *The solution to the MMAP formulation  $\arg \max_q W_{M_{\tilde{X}}}(q)$  lies at extreme iff solution to its equivalent formulation:*

$$\arg \max_{N_1, N_2, \dots, N_R} \sum_s \prod_{l=1}^R f_l(s)^{N_l}$$

subject to the constraints  $\forall l, N_l \geq 0, N_l \in \mathbb{Z}$  and  $\sum_l N_l = m$  lies at extreme.

### 3.2.5 Proving Extreme

**Lemma 6.** *Consider the optimization problem:*

$$\arg \max_{N_1, N_2, \dots, N_R} \sum_s g(s) \times \prod_{l=1}^R f_l(s)^{N_l}$$

subject to the constraints  $N_l \geq 0, \sum_l N_l = m$ .  $g(s)$  is an arbitrary real-valued function independent of  $l$ . The solution of this optimization problem lies at extreme.

*Proof.* Note that it suffices to prove this theorem assuming  $N_l$ 's are real-valued. If the solution is at extreme with real-valued  $N_l$ 's, it must also be at extreme when  $N_l$ 's are further constrained to be integer valued. We will use induction on  $R$  to prove the result. Consider base case of  $R = 2$ , the function becomes  $\arg \max_{N_1} \sum_s f_1(s)^{N_1} f_2(s)^{m-N_1} \times g(s)$ . This function is convex and has its maximum value at  $N_1 = m$  or  $N_1 = 0$  (see supplement for a proof).

Assuming that the induction hypothesis holds for  $R = k$ . We need to show for the case when  $R = k + 1$ . We will prove it by contradiction. Assume that the solution to this problem does not lie at extreme. Then, in this solution, it must be the case that  $N_l \neq 0, \forall l$ . If

not, we can then reduce the problem to a  $k$  sized one and apply our induction hypothesis to get an extreme solution. Also, clearly  $N_l < m, \forall l$ . Let  $N_{k+1}$  has the optimal value of  $N_{k+1}^*$  in this solution. Then, substituting the optimal value of this component in the expression, we can get the optimal value for  $(N_1, N_2, \dots, N_k)$  by solving  $\arg \max_{N_1, N_2, \dots, N_k} \sum_s g'(s) \times \prod_{l=1}^k f_l(s)^{N_l}$ , subject to  $\sum_{l=1}^k N_l = m - N_{k+1}^*$ . Here,  $g'(s) = g(s) \times f_{k+1}(s)^{N_{k+1}^*}$ . Using the induction hypothesis, the solution for this must be at extreme, i.e.  $\exists l, N_l = 0$  since  $k \geq 2$ . This is a contradiction.

**Corollary 1.** *The solution to the optimization problem*

$$\arg \max_{N_1, N_2, \dots, N_R} \sum_s \prod_{l=1}^R f_l(s)^{N_l}$$

subject to the constraints  $\forall l, N_l \geq 0, N_l \in \mathbb{Z}$  and  $\sum_l N_l = m$  lies at extreme.

**Theorem 1 (Proof):** Corollary 1 combined with Lemma 5, Lemma 4, Lemma 3, Lemma 2 and Lemma 1 proves the theorem.

### 3.3 SOM-R Rule for lifted MMAP

We will first define the SOM-R (SOM Reduce) equivalence class which is a sub-class of SOM. Following our notation, we will use  $\mathcal{Q}$  and  $\mathcal{S}$  to denote the set of MAX and SUM predicates, respectively in the MMAP problem.

**Definition 4.** We say that an equivalence class of variables  $\tilde{X}$  is SOM-R if (a)  $\tilde{X}$  is SOM (b)  $\forall P \in \mathcal{S}, P$  contains a variable from  $\tilde{X}$  OR  $\forall P \in \mathcal{S}, P$  does not have a variable from  $\tilde{X}$ .

Note that if  $|\mathcal{S}| = 1$ , then any SOM equivalence class is also necessarily SOM-R. Next, we exploit the properties of extreme assignments to show that domain of SOM-R variables can be reduced to a single constant for MMAP inference. We start with the definition of a reduced MLN.

**Definition 5.** (Reduced MLN) Let  $\{(f_i, w_i)\}_{i=1}^n$  denote the set of (weighted) formulas in  $M$ . Let  $\tilde{X}$  be a SOM-R equivalence class with  $|\Delta_{\tilde{X}}| = m$ . We construct a reduced MLN  $M^r$  by considering the following 2 cases:

CASE 1:  $\forall P \in \mathcal{S}, P$  contains a variable from  $\tilde{X}$

- $\forall f_i \in F(M)$  containing a variable  $X \in \tilde{X}$ , add  $(f_i, w_i)$  to  $M^r$ .
- $\forall f_i \in F(M)$  not containing a variable  $X \in \tilde{X}$ , add  $(f_i, \frac{1}{m} \times w_i)$  to  $M^r$ .

CASE 2:  $\forall P \in \mathcal{S}, P$  does not contain a variable from  $\tilde{X}$

- $\forall f_i \in F(M)$  containing a variable  $X \in \tilde{X}$ , add  $(f_i, w_i \times m)$  to  $M^r$ .

- $\forall f_i \in F(M)$  not containing a variable  $X \in \tilde{X}$ , add  $(f_i, w_i)$  to  $M^r$ .

In each case, we reduce the domain of  $\tilde{X}$  to a single constant in  $M^r$ .

We are ready to state our SOM-R rule for lifted MMAP.

**Theorem 2.** (SOM-R Rule for MMAP) *Let  $\tilde{X}$  be a SOM-R equivalence class. Let  $M^r$  be the reduced MLN in which domain of  $\tilde{X}$  has been reduced to single constant. Then, MMAP problem can be equivalently solved over  $M^r$ .*

*Proof.* Let  $\mathcal{Q}$  denote the set of MAX predicates in the problem. We prove the above theorem in two parts. In Lemma 7 below, we show that for every extreme assignment (with respect to  $\tilde{X}$ )  $q$  to groundings of  $\mathcal{Q}$  in  $M$ , there is a corresponding extreme assignment  $q^r$  in  $M^r$  (and vice-versa). In Lemma 8, we show that given two extreme assignments,  $q$  and  $q^r$  for the respective MLNs, the MMAP value at  $q$  (in  $M_{\tilde{X}}$ ) is a monotonically increasing function of the MMAP value at  $q^r$  (in  $M_{\tilde{X}}^r$ ). These two facts combined with the fact that MMAP solution to the original problem is at extreme (using Theorem 1) prove the desired result. Next we prove each result in turn.

**Lemma 7.** *Let  $\mathbf{q}$  (resp.  $\mathbf{q}^r$ ) denote the sets of extreme assignments to the groundings of  $\mathcal{Q}$  in  $M$  (resp.  $M^r$ ). There exists a one to one mapping between  $\mathbf{q}$  and  $\mathbf{q}^r$ .*

*Proof.* Instead of directly working with  $M$  and  $M^r$ , we will instead prove this lemma for the corresponding variablized MLNs  $M_{\tilde{X}}$  and  $M_{\tilde{X}}^r$ . This can be done since the process of variablization preserves the distribution as well as the set of extreme assignments. Let  $q$  denote an extreme assignment to MAX predicates in  $M_{\tilde{X}}$ . We will construct a corresponding assignment  $q^r$  for MAX predicate in  $M_{\tilde{X}}^r$ . Since  $\tilde{X}$  is SOM-R,  $M_{\tilde{X}}$  has only unary and propositional predicates, whereas  $M_{\tilde{X}}^r$  is full ground since the domain of  $\tilde{X}$  is reduced to a single constant.

First, let us consider a propositional MAX predicate  $P$  in  $M_{\tilde{X}}$ . Since  $P$  is ground both in  $M$  and  $M^r$ , we can assign the value of  $P$  in  $q^r$  to be same as  $q$ . Next, let us consider a unary predicate  $P$ . Let the assignments to the  $m$  groundings of  $P$  in  $q$  be given by the set  $\{q_{P_j}\}$  where  $1 \leq j \leq m$ . Since  $q$  is extreme, each element in the set  $\{q_{P_j}\}$  takes the same truth value. We can simply assign this value to the ground appearance of  $P$  in  $M_{\tilde{X}}^r$ . Hence, we get a mapping from  $q$  to  $q^r$ . It is easy to see that we can get a reverse mapping from  $q^r$  to  $q$  in a similar manner. Hence, proved.

Next, we state the relationship between the MMAP values obtained by the extreme assignments in  $M$  and  $M^r$ .

**Lemma 8.** \* *Let  $M$  be an MLN and  $M^r$  be the re-*

*duced MLN with respect to the SOM-R equivalence class  $\tilde{X}$ . Let  $q$  and  $q^r$  denote two corresponding extreme assignments in  $M$  and  $M^r$ , respectively. Then,  $\exists$  a monotonically increasing function  $g$  such that  $W_M(q) = g(W_{M^r}(q^r))$ .*

The proof of Lemma 8 exploits inversion elimination and symmetry of potential functions over a variablized MLN similar to their use in Section 3.2. These combined with Lemma 7 become our key insights for reducing the complexity of MMAP inference significantly compared to existing methods (see supplement for details).

**Corollary 2.** *SOM-R rule for MMAP problem subsumes SO rule for MAP problem given by Mittal et al. (2014).*

The corollary follows from the fact that MAP is a special case of MMAP when all the predicates are MAX.

## 4 ALGORITHMIC FRAMEWORK

SOM-R rule can be combined with existing lifted inference rules such as lifted decomposition and conditioning (Jha et al. 2010; Gogate and Domingos 2011) (with minor modifications) to yield a powerful algorithm for solving MMAP (see Algorithm 1). The algorithm takes as input an MLN  $M$ , the set of MAX predicates  $\mathcal{Q}$ , SUM predicates  $\mathcal{S}$  and a ground MMAP solver  $gSol$ . It has six steps. In the first step, the algorithm checks to see if the MLN, along with  $\mathcal{Q}$  and  $\mathcal{S}$  can be partitioned into disjoint MLNs that do not share any ground atoms. If this condition is satisfied, then the MMAP solution can be constructed by solving each component independently and simply concatenating the individual solutions. In the next three steps, we apply the decomposer (Jha et al. 2010), SOM-R (this work) and binomial rules (Jha et al. 2010; Gogate and Domingos 2011) in order. The former two reduce the domain of all logical variables in the equivalence class to a constant and thus yield exponential reductions in complexity. Therefore, they are applied before the binomial rule which creates  $O(m)$  ( $|\Delta_{\tilde{X}}| = m$ ) smaller sub-problems. In the algorithm,  $M^d$  refers to an MLN obtained from  $M$  by setting the domain of  $\tilde{X}$  to a single constant and we assume that  $|\Delta_{\tilde{X}}| = m$ . Similarly,  $M^r$  refers to the MLN obtained from  $M$  by applying the SOM-R rule (see Definition 5).

The binomial rule (steps 4a and 4b) efficiently conditions on the unary predicates and can be applied over the SUM as well as MAX predicates. However, care must be taken to ensure that all MAX predicates are instantiated before the SUM predicates. Therefore, the binomial rule is applied over the SUM predicates only when the MLN has no MAX predicates (Step 4b). In the algorithm,  $M_k$  refers to the MLN obtained from  $M$  by setting exactly  $k$  groundings of  $P$  to true and the remaining to false.

If none of the lifting rules are applicable and the MLN has only ground atom, we return the solution returned by the propositional solver  $gSol$ . Otherwise, if not all predicates are ground, we resort to partial grounding, namely we heuristically ground a logical variable and recurse on the corresponding MLN  $M'$ .

Finally, note that the algorithm returns the exponentiated weight of the MMAP assignment. The assignment can be recovered by tracing the recursion backwards.

**Heuristics:** (a) Binomial: In case of multiple possible binomial applications, we pick the one which results in the application of other lifting rules (in the priority order described above) using a one step look ahead. In case of a tie, we pick the one with maximum domain size.

(b) Partial Grounding: We pick the equivalence class which results in further application of lifting rules (in the priority order) using a one step look ahead. In case of a tie, we pick the one which has smallest domain size.

## 5 EXPERIMENTS

The goal of our experiments is two fold. First, we would like to examine the efficacy of lifting for MMAP. Second, we would like to analyze the contribution of SOM-R rule in lifting. Towards this end, we compare the following three algorithms: (1) Ground: ground inference with no lifting whatsoever (2) Lifted-Basic: lifted inference without use of the SOM-R rule <sup>4</sup> (3) Lifted-SOM-R: using all our lifting rules including SOM-R. For ground inference, we use a publicly available <sup>5</sup> base (exact) solver built on top of And/Or search developed by Marinescu et al. (2014).

We experiment with three benchmark MLNs: (1) Student (Sarkhel et al. 2014) (2) IMDB (Mittal et al. 2016) (3) Friends & Smokers (FS) (Domingos and Lowd 2009). All the datasets are described in the lower part of Figure 1 along with the MAP predicates used in each case; the remaining predicates are treated as marginal predicates. Weights of the formulas were manually set.

We compare the performance of the three algorithms on two different metrics: (a) time taken for inference (b) memory used. We used a time-out of 30 minutes for each run. Memory was measured in terms of the number of formulas in the ground network in each case. We do not compare the solution quality since all the algorithms are guaranteed to produce MMAP assignments with same

<sup>4</sup>We use the rules described in Algorithm 1. For Lifted-Basic, too many applications of the binomial rule led to blow up. So, we restricted the algorithm to a single binomial application and before any partial grounding. Lifted-SOM-R had no such issues.

<sup>5</sup><https://github.com/radum2275/merlin>

---

### Algorithm 1 Lifted MMAP

---

**Input:** MLN  $M, \mathcal{Q}, \mathcal{S}, gSol$

**Output:** MMAP value

**Begin:**

```
//1. Disjoint Sub-Theories
if  $M$  can be partitioned into disjoint MLNs  $M_1, \dots, M_t$  that
share no atoms then
    return  $\prod_{i=1}^t$  liftedMMAP( $M_i, \mathcal{Q}_i, \mathcal{S}_i$ )
//2. Decomposer
if there exists a decomposer  $\tilde{X}$  in  $M$  then
    return [liftedMMAP( $M^d, \mathcal{Q}, \mathcal{S}, gSol$ )] $m$ ;
//3. SOM-R (see Defn. 5)
if there exists a SOM-R class  $\tilde{X}$  in  $M$  then
    return liftedMMAP( $M^r, \mathcal{Q}, \mathcal{S}, gSol$ );
//4a. Binomial over MAX
if there exists a unary predicate  $P \in \mathcal{Q}$  then
    return  $\max_k$  liftedMMAP( $M_k, \mathcal{Q} - \{P\}, \mathcal{S}, gSol$ );
//4b. Binomial over SUM
if  $\mathcal{Q} = \emptyset$  and there exists a unary predicate  $P \in \mathcal{S}$  then
    return  $\sum_{k=0}^m \binom{m}{k}$  liftedMMAP( $M_k, \mathcal{Q}, \mathcal{S} - \{P\}, gSol$ );
//5. Check if fully Ground
if  $M$  is fully Ground then
    return apply( $M^l, \mathcal{Q}, \mathcal{S}, gSol$ );
else
    //6. Partial Grounding
     $M' =$  Heuristically ground an equivalence class  $\tilde{X}$  in  $M$ 
    return liftedMMAP( $M', \mathcal{Q}, \mathcal{S}, gSol$ );
```

**End.**

---

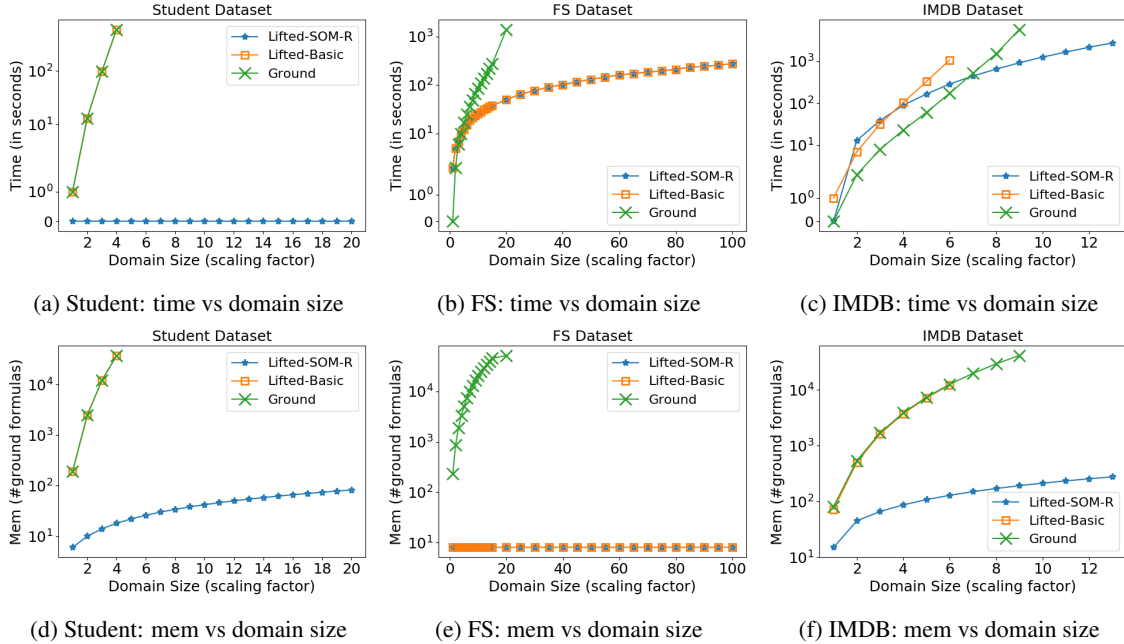
(optimal) probability. All the experiments were run on a 2.20 GHz Xeon(R) E5-2660 v2 server with 10 cores and 62 GB RAM.

**Results:** For each of the graphs in Figure 1, we plot time (memory) on y-axis (log-scale) and domain size on x-axis. Time is measured in seconds. Since we are primarily concerned about the scaling behavior, we use the number of ground formulae as a proxy for the actual memory usage. Domain size is measured as a function of a scaling factor, which is the number by which (all of) the starting domain sizes are multiplied. We refer to domain descriptions (Figure 1) for the starting sizes.

Figures 1a and 1d compare the performance of the three algorithms on the Student dataset. None of the lifting rules apply for Lifted-Basic. Hence, its performance is identical to Ground. For Lifted-SOM-R, all the variables (except teacher(T)) can be reduced to a single constant, resulting in significant reduction in the size of the ground theory. Lifted-SOM-R is orders of magnitude better than Ground and Lifted-Basic for both time and memory.

Figures 1b and 1e compare the three algorithms on the FS dataset. Here, Lifted-Basic performs identical to Lifted-SOM-R. This is because binomial rule applies in the beginning on Smokes, following which theory decomposes. We never need to apply SOM-R rule on this domain. Both Lifted-SOM-R and Lifted-Basic perform





Student (Sarkhel et al. 2014)	IMDB (Mittal et al. 2016)
Teaches(T, C) $\wedge$ Takes(S, C) $\Rightarrow$ JobOffer(S, M)	WorksWith(P1,P2) $\Rightarrow$ Act(P1); WorksWith(P1,P2) $\Rightarrow$ Dir(P2);
<b>MAP Predicate:</b> Takes(S, C), JobOffer(S, M)	Dir(P1) $\wedge$ Act(P2) $\wedge$ Mov(M,P1) $\wedge$ Mov(M,P2) $\Rightarrow$ WorksWith(P2,P1);
<b>size:</b> teachr(T):2,course(C):3,comp(M):4,stud(S):6	Dir(P1) $\wedge$ Act(P2) $\wedge$ Mov(M,P2) $\wedge$ WorksWith(P2,P1) $\Rightarrow$ Mov(M,P1);
<b>FS (Domingos and Lowd 2009)</b>	Dir(P1) $\wedge$ Act(P2) $\wedge$ Mov(M,P1) $\wedge$ WorksWith(P2,P1) $\Rightarrow$ Mov(M,P2);
Smokes(P) $\Rightarrow$ Cancer(P);	Dir(P1) $\wedge$ Act(P2) $\Rightarrow$ WorksWith(P2,P1);
Smokes(P1) $\wedge$ Friend(P1, P2) $\Rightarrow$ Smokes(P2);	<b>MAP Predicates:</b> Act(P), Dir(P), Mov(M,P)
<b>MAP Predicates:</b> Smokes(P), Cancer(P)	<b>size:</b> person(P):3, movie(M):2
<b>size:</b> person(P):5	

Figure 1: Results and rules of Student, FS and IMDB datasets. "size" gives initial domain sizes for each case.

significantly better than Ground on this domain (in both time and memory).

IMDB dataset (Figures 1c and 1f) presents a particularly interesting case of interspersed application of rules. For Lifted-SOM-R, SOM-R rule applies on movie(M) variables, simplifying the theory following which binomial rule can be applied on Mov, Dir and Act predicates. Theory decomposes after these binomial applications. For Lifted-Basic, though binomial rule can be applied on Dir, Act the movie variables still remain, eventually requiring for partial grounding. Surprisingly, Ground does slightly better than both the lifted approaches for smaller domains for time. This is due to the overhead of solving multiple sub-problems in binomial without much gain since domains are quite small. Lifted-SOM-R has a much better scaling behavior for larger domains. It also needs significantly less memory compared to both other approaches.

In none of the above cases, Lifted-SOM-R has to ever partially ground the theory making a very strong case

for using Lifted-SOM-R for MMAP inference in many practical applications. Overall, our experiments clearly demonstrate the utility of SOM-R in the scenarios where other lifting rules fail to scale.

## 6 CONCLUSION

We present the first lifting technique for MMAP. Our main contribution is the SOM-R rule, which states that the domain of a class of equivalence variables, referred to as SOM-R, can be reduced to a single constant for the purpose of MMAP inference. We prove the correctness of our rule through a series of problem transformations followed by the properties of what we refer to as extreme assignments. Our experiments clearly demonstrate the efficacy of our approach on benchmark domains. Directions for future work include coming up with additional lifting rules, approximate lifting and lifting in presence of constraints (Mittal et al. 2015), all in the context of MMAP, and experimenting with a wider set of domains.

## **Acknowledgements**

Happy Mittal is supported by the TCS Research Scholars Program. Vibhav Gogate and Parag Singla are supported by the DARPA Explainable Artificial Intelligence (XAI) Program with number N66001-17-2-4032. Parag Singla is supported by IBM Shared University Research Award and the Visvesvaraya Young Faculty Research Fellowship by the Govt. of India. Vibhav Gogate is supported by the National Science Foundation grants IIS-1652835 and IIS-1528037. Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views or official policies, either expressed or implied, of the funding agencies.

## References

- [de Salvo Braz, Amir, and Roth 2005] de Salvo Braz, R.; Amir, E.; and Roth, D. 2005. Lifted first-order probabilistic inference. In *Proc. of IJCAI-05*, 1319–1325.
- [de Salvo Braz, Amir, and Roth 2006] de Salvo Braz, R.; Amir, E.; and Roth, D. 2006. MPE and partial inversion in lifted probabilistic variable elimination. In *Proc. of AAAI-06*, 1123–1130.
- [Domingos and Lowd 2009] 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.
- [G. Van den Broeck et al. 2011] G. Van den Broeck; Taghipour, N.; Meert, W.; Davis, J.; and Raedt, L. D. 2011. Lifted probabilistic inference by first-order knowledge compilation. In *Proc. of IJCAI-11*.
- [Getoor and Taskar 2007] Getoor, L., and Taskar, B., eds. 2007. *Introduction to Statistical Relational Learning*. MIT Press.
- [Gogate and Domingos 2011] Gogate, V., and Domingos, P. 2011. Probabilistic theorem proving. In *Proc. of UAI-11*, 256–265.
- [Jha et al. 2010] Jha, A. K.; Gogate, V.; Meliou, A.; and Suciu, D. 2010. Lifted inference seen from the other side : The tractable features. In *Proc. of NIPS-10*, 973–981.
- [Kersting, Ahmadi, and Natarajan 2009] Kersting, K.; Ahmadi, B.; and Natarajan, S. 2009. Counting belief propagation. In *Proc. of UAI-09*.
- [Kimmig, Mihalkova, and Getoor 2015] Kimmig, A.; Mihalkova, L.; and Getoor, L. 2015. Lifted graphical models: a survey. *Machine Learning* 99(1):1–45.
- [Liu and Ihler 2013] Liu, Q., and Ihler, A. T. 2013. Variational algorithms for marginal map. *Journal of Machine Learning Research* 14(1):3165–3200.
- [Maaten, Welling, and Saul 2011] Maaten, L.; Welling, M.; and Saul, L. K. 2011. Hidden-unit conditional random fields. In *International Conference on Artificial Intelligence and Statistics*, 479–488.
- [Marinescu, Dechter, and Ihler 2014] Marinescu, R.; Dechter, R.; and Ihler, A. T. 2014. And/or search for marginal map. In *Proc. of UAI-14*, 563–572.
- [Mittal et al. 2014] Mittal, H.; Goyal, P.; Gogate, V.; and Singla, P. 2014. New rules for domain independent lifted MAP inference. In *Proc. of NIPS-14*, 649–657.
- [Mittal et al. 2015] Mittal, H.; Mahajan, A.; Gogate, V.; and Singla, P. 2015. Lifted inference rules with constraints. In *Advances in Neural Information Processing Systems*, 3519–3527.
- [Mittal et al. 2016] Mittal, H.; Singh, S. S.; Gogate, V.; and Singla, P. 2016. Fine grained weight learning in markov logic networks. In *Proc. of IJCAI-16 Wkshp. on Statistical Relational AI*.
- [Mladenov, Kersting, and Globerson 2014] Mladenov, M.; Kersting, K.; and Globerson, A. 2014. Efficient lifting of map lp relaxations using k-locality. In *Proc. of AISTATS-14*, 623–632.
- [Niepert 2012] Niepert, M. 2012. Markov chains on orbits of permutation groups. In *In Proc. UAI-14, CA, USA, August 14-18, 2012*, 624–633.
- [Park 2002] Park, J. 2002. MAP Complexity Results and Approximation Methods. In *Proc. of UAI-02*.
- [Poole 2003] Poole, D. 2003. First-order probabilistic inference. In *Proc. of IJCAI-03*, 985–991.
- [Russell and Norvig 2010] Russell, S. J., and Norvig, P. 2010. *Artificial Intelligence - A Modern Approach*. Pearson Education.
- [Sarkhel et al. 2014] Sarkhel, S.; Venugopal, D.; Singla, P.; and Gogate, V. 2014. Lifted MAP inference for Markov logic networks. In *Proc. of AISTATS-14*, 895–903.
- [Sarkhel, Singla, and Gogate 2015] Sarkhel, S.; Singla, P.; and Gogate, V. G. 2015. Fast lifted map inference via partitioning. In *Advances in Neural Information Processing Systems*, 3240–3248.
- [Singla and Domingos 2008] Singla, P., and Domingos, P. 2008. Lifted first-order belief propagation. In *Proc. of AAAI-08*.
- [Singla and Mooney 2011] Singla, P., and Mooney, R. 2011. Abductive Markov logic for plan recognition. In *Proc. of AAAI-11*, 1069–1075.
- [Singla, Nath, and Domingos 2014] Singla, P.; Nath, A.; and Domingos, P. 2014. Approximate lifted belief propagation. In *Proc. of AAAI-14*, 2497–2504.
- [Van den Broeck and Darwiche 2013] Van den Broeck, G., and Darwiche, A. 2013. On the complexity

and approximation of binary evidence in lifted inference. In *Advances in Neural Information Processing Systems*, 2868–2876.

[Venugopal and Gogate 2012] Venugopal, D., and Gogate, V. 2012. On lifting the Gibbs sampling algorithm. In *Proc. of NIPS-12*, 1664–1672.

[Xue et al. 2016] Xue, Y.; Ermon, S.; Gomes, C. P.; Selman, B.; et al. 2016. Solving marginal map problems with np oracles and parity constraints. In *In Proc. of NIPS-16*, 1127–1135.

## Supplementary Material

### Lemmas

We will start by proving Lemma 0, which will be used in the proof of Lemma 2 .

**Lemma 0.** *Given a function  $f(x, y)$ , let  $(x^*, y^*)$  be a maximizing assignment, i.e.,  $(x^*, y^*) = \arg \max_{x, y} f(x, y)$ . Then,  $\forall y'$  s.t.  $y' = \arg \max_y f(x^*, y)$ ,  $(x^*, y')$  is also a maximizing assignment.*

*Proof.* We can write the following (in)equality:

$$f(x^*, y^*) \leq \max_y f(x^*, y) = f(x^*, y')$$

But since  $(x^*, y^*)$  was the maximizing assignment for  $f(x, y)$ , it must be the case that  $f(x^*, y^*) = f(x^*, y')$ . Hence,  $(x^*, y')$  must also be a maximizing assignment. Hence, proved.

**Lemma 2.** *Consider the MMAP problem over  $M_{\bar{X}}$ . Let  $q_p$  be an assignment to the propositional MAP predicates. Let  $M'_{\bar{X}}$  be an MLN obtained by substituting the truth value in  $q_p$  for propositional predicates. Then, if  $M'_{\bar{X}}$  has a solution at extreme for all possible assignments of the form  $q_p$  then,  $M_{\bar{X}}$  also has a solution at extreme.*

*Proof.* The MMAP problem can be written as:

$$\arg \max_{q_p, q_u} \sum_{s_p, s_u} W_{M_{\bar{X}}}(q_p, q_u, s_p, s_u) \quad (6)$$

here,  $q_p, q_u$  denote an assignment to the propositional and unary MAX predicate groundings in  $M_{\bar{X}}$ , respectively. Similarly,  $s_p, s_u$  denote an assignment to the propositional and unary SUM predicate groundings in  $M_{\bar{X}}$ , respectively. Let  $q_p^*$  denote an optimal assignment to the propositional MAX predicates. Then, using Lemma 0, we can get the MMAP assignment  $q_u^*$  as a solution to the following problem:

$$\arg \max_{q_u} \sum_{s_p, s_u} W_{M'_{\bar{X}}}(q_u, s_u, s_p) \quad (7)$$

where  $M'_{\bar{X}}$  is obtained by substituting the truth assignment  $q_p^*$  in  $M_{\bar{X}}$ . Since,  $M'_{\bar{X}}$  has a solution at extreme  $\forall q_p$ , it must also be at extreme when  $q_p = q_p^*$ . Hence,  $q_u^*$  must be at extreme. Hence, proved.

**Lemma 5.** *The solution to the MMAP formulation  $\arg \max_q W_{M_{\bar{X}}}(q)$  lies at extreme iff solution to its equivalent formulation:*

$$\arg \max_{N_1, N_2, \dots, N_R} \sum_s \prod_{l=1}^R f_l(s)^{N_l} \quad (8)$$

subject to the constraints  $\forall l, N_l \geq 0, N_l \in \mathbb{Z}$  and  $\sum_l N_l = m$  lies at extreme.

*Proof.* If  $\arg \max_{N_1, N_2, \dots, N_R} \sum_s \prod_{l=1}^R f_l(s)^{N_l}$  lies at extreme then  $\exists l$  such that  $N_l = m$  and  $N_{l'} = 0, \forall l' \neq l$ . Let  $v_l$  denote the value taken by the groundings of the unary MAX predicates corresponding to index  $l$ . Since  $N_l = m$ , it must be the case that all the groundings get the identical value  $v_l$ . Hence, the solution to  $\arg \max_q W_{M_{\bar{X}}}(q)$  lies at extreme. Similar proof strategy holds for the other way around as well.

**Lemma. (Induction Base Case in Proof of Lemma 6):** *Let  $f_1(s), f_2(s)$  and  $g(s)$  denote real-valued functions of a vector valued input  $s$ <sup>6</sup>. Further, let each of  $f_1(s), f_2(s), g(s)$  be non-negative. Then, for  $N \in \mathbb{R}$ , we define a function  $h(N) = \sum_s f_1(s)^N f_2(s)^{m-N} \times g(s)$  where the domain of  $h$  is further restricted to be in the interval  $[0, m]$ , i.e.,  $0 \leq N \leq m$ . The maxima of the function  $h(N)$  lies at  $N = 0$  or  $N = m$ .*

*Proof.* First derivative of  $h(N)$  with respect to  $N$  is:

$$\frac{dh}{dN} = \sum_s \left( f_1(s)^N f_2(s)^{m-N} g(s) \times [\log(f_1(s)) - \log(f_2(s))] \right)$$

Second derivative of  $h(N)$  with respect to  $N$  is given as:

$$\frac{d^2h}{dN^2} = \sum_s \left( f_1(s)^N f_2(s)^{m-N} g(s) \times [\log(f_1(s)) - \log(f_2(s))]^2 \right) \geq 0$$

The inequality follows from the fact that each of  $f_1, f_2, g$  is non-negative. Hence, the second derivative of  $h(N)$  is non-negative which means the function is convex. Therefore, the maximum value of this function must lie at the end points of its domain, i.e, either at  $N = 0$  or at  $N = m$ .

**Lemma 8.** *Let  $M$  be an MLN and  $M^r$  be the reduced MLN with respect to the SOM-R equivalence class  $\bar{X}$ . Let  $q$  and  $q^r$  denote two corresponding extreme assignments in  $M$  and  $M^r$ , respectively. Then,  $\exists$  a monotonically increasing function  $g$  such that  $W_M(q) = g(W_{M^r}(q^r))$ .*

*Proof.* First, we note that if we multiply the weight  $w_i$  of a formula  $f_i$  in an MLN by a factor  $k$ , then, the corresponding potential  $\phi_{ij}$  (i.e., potential corresponding to the  $j^{\text{th}}$  grounding of the  $i^{\text{th}}$  formula) gets raised to the

<sup>6</sup>Recall that  $s$  was an assignment to all the propositional SUM predicates in our original Lemma.

power  $k$ . If  $w_i$  gets replaced by  $k \times w_i$ , then, correspondingly,  $\phi_{ij}$  gets replaced by  $(\phi_{ij})^k \forall j$ . We will use this fact in the following proof.

As in the case of Lemma 7, we will instead work with the variabilized MLNs  $M_{\tilde{X}}$  and  $M_{\tilde{X}}^r$ , respectively. Let  $q = (q_p, q_u)$  be the MMAP assignment for  $\mathcal{Q}$  in  $M_{\tilde{X}}$  and similarly  $q = (q_p^r, q_u^r)$  be the MMAP assignment for  $\mathcal{Q}$  in  $M_{\tilde{X}}^r$ .

For MLN  $M_{\tilde{X}}$ , the MMAP objective  $W_M$  at  $(q_p, q_u)$  can be written as  $W_{M_{\tilde{X}}}(q_p, q_u) =$ :

$$\sum_{s_p, s_u} \left( \prod_{i=1}^r \prod_{j=1}^m \phi_{ij}(q_p, q_u, s_p, s_u) \prod_{k=1}^t \phi_k(q_p, s_p) \right) \quad (9)$$

where  $\phi_{ij}$  are potentials over formulas containing some  $X \in \tilde{X}$  and  $\phi_k$  are potentials over formulas which do not contain any  $X \in \tilde{X}$ . In particular, note that we have separated out the formulas which involve a variable from the class  $\tilde{X}$  from those which don't.  $r$  denotes the count of the formulas of the first type and  $t$  denotes the count of the formulas of the second type. We will use this form in the following proof.

Let the reduced domain of  $\tilde{X}$  in  $M^r$  is given by  $\{x_1\}$ , i.e., the only constant which remains in the domain is corresponding to index  $j = 1$ . Next we prove the above lemma for the two cases considered in Definition 5:

**CASE 1:**  $\forall P \in \mathcal{S}, P$  contains a variable from  $\tilde{X}$

In this case  $M_{\tilde{X}}$  and  $M_{\tilde{X}}^r$  will not contain any propositional SUM predicate i.e.  $s_p = \emptyset$ .

In this case, while constructing  $M_{\tilde{X}}^r$ , for formulas not containing some  $X \in \tilde{X}$  we divided the weight by  $m$ . This combined with the result stated in the beginning of this proof, the MMAP objective for  $M_{\tilde{X}}^r$  can be written as:

$$\begin{aligned} W_{M_{\tilde{X}}^r}(q_p, q_u) &= \sum_{s_u} \left( \prod_{i=1}^r \phi_{i1}(q_p, q_{u_1}, s_{u_1}) \prod_{k=1}^t \phi_k(q_p)^{\frac{1}{m}} \right) \\ &= \left( \sum_{s_{u_1}} \prod_{i=1}^r \phi_{i1}(q_p, q_{u_1}, s_{u_1}) \right) \prod_{k=1}^t \phi_k(q_p)^{\frac{1}{m}} \end{aligned}$$

Next for MLN  $M_{\tilde{X}}$  we have,  $W_{M_{\tilde{X}}}(q_p, q_u) =$

$$\begin{aligned} &\sum_{s_u} \left( \prod_{i=1}^r \prod_{j=1}^m \phi_{ij}(q_p, q_u, s_u) \prod_{k=1}^t \phi_k(q_p) \right) \\ &= \left( \sum_{s_u} \prod_{j=1}^m \prod_{i=1}^r \phi_{ij}(q_p, q_u, s_u) \right) \prod_{k=1}^t \phi_k(q_p) \end{aligned}$$

$$\begin{aligned} &= \left( \sum_{s_{u_j}} \prod_{j=1}^m \prod_{i=1}^r \phi_{ij}(q_p, q_{u_j}, s_{u_j}) \right) \prod_{k=1}^t \phi_k(q_p) \\ &= \left( \prod_{j=1}^m \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_j}, s_{u_j}) \right) \prod_{k=1}^t \phi_k(q_p) \\ &= \left( \prod_{j=1}^m \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_1}, s_{u_j}) \right) \prod_{k=1}^t \phi_k(q_p) \\ &= \left( \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_1}, s_{u_j}) \right)^m \prod_{k=1}^t \phi_k(q_p) \\ &= \left( \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_1}, s_{u_j}) \right)^m \prod_{k=1}^t \left( \phi_k(q_p)^{\frac{1}{m}} \right)^m \\ &= \left( \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_1}, s_{u_j}) \right)^m \left( \prod_{k=1}^t \phi_k(q_p)^{\frac{1}{m}} \right)^m \\ &= \left( \sum_{s_{u_j}} \prod_{i=1}^r \phi_{ij}(q_p, q_{u_1}, s_{u_j}) \prod_{k=1}^t \phi_k(q_p)^{\frac{1}{m}} \right)^m \\ &= \left( W_{M_{\tilde{X}}^r}(q_p, q_u) \right)^m \end{aligned}$$

First equality comes by removing  $s_p$  from Equation 9. In second equality we switch the order of two products. In third equality we have made explicit the dependence of  $\phi_{ij}$  on  $q_{u_j}$  and  $s_{u_j}$  i.e. groundings corresponding to  $j^{th}$  constant. In fourth equality we use inversion elimination (de Salvo Braz, Amir, and Roth 2005) to invert the sum over  $s_{u_j}$  and product over  $j$ . Next, since  $\tilde{X}$  is SOM-R, from Theorem 1 we know  $q_u$  lies at extreme i.e.  $\forall j, q_{u_j} = q_{u_1}$ , so we replace all  $q_{u_j}$  by  $q_{u_1}$  in fifth equality. Next, after summing out  $s_{u_j}$  all  $\phi_{ij}$  will behave identically<sup>7</sup>, so we reduce  $\prod_{j=1}^m$  to exponent  $m$ . In the next steps we do basic algebraic manipulations to show  $W_{M_{\tilde{X}}}(q) = g(W_{M_{\tilde{X}}^r}(q^r))$  where  $g$  is function defined as  $g(x) = x^m$ , i.e.,  $g$  is monotonically increasing. Hence, proved.

**CASE 2:**  $\forall P \in \mathcal{S}, P$  doesn't contain a variable from  $\tilde{X}$   
In this case  $M_{\tilde{X}}$  will not contain any unary SUM predicate i.e.  $s_u = \emptyset$ .

In this case for the reduced MLN  $M_{\tilde{X}}^r$  we multiply the weight of formulas containing some  $X \in \tilde{X}$  by  $m$  and domain of  $\tilde{X}$  is reduced to a single constant. Combining this fact along with the result shown in the beginning of this proof, MAP objective for  $M_{\tilde{X}}^r$  is given by:

$$W_{M_{\tilde{X}}^r}(q_p, q_u) = \sum_{s_p} \left( \prod_{i=1}^r \phi_{i1}(q_p, q_{u_1}, s_p)^m \prod_{k=1}^t \phi_k(q_p, s_p) \right)$$

<sup>7</sup> $\phi_{ij}$ 's are identical to each other up to renaming of the index  $j$ , due to the normal form assumption.

Next for MLN  $M_{\tilde{X}}$  we have,  $W_{M_{\tilde{X}}}(q_p, q_u) =$

$$\begin{aligned}
& \sum_{s_p} \left( \prod_{i=1}^r \prod_{j=1}^m \phi_{ij}(q_p, q_u, s_p) \prod_{k=1}^t \phi_k(q_p, s_p) \right) \\
&= \sum_{s_p} \left( \prod_{i=1}^r \prod_{j=1}^m \phi_{ij}(q_p, q_{u_j}, s_p) \prod_{k=1}^t \phi_k(q_p, s_p) \right) \\
&= \sum_{s_p} \left( \prod_{i=1}^r \prod_{j=1}^m \phi_{ij}(q_p, q_{u_1}, s_p) \prod_{k=1}^t \phi_k(q_p, s_p) \right) \\
&= \sum_{s_p} \left( \prod_{i=1}^r (\phi_{i1}(q_p, q_{u_1}, s_p))^m \prod_{k=1}^t \phi_k(q_p, s_p) \right) \\
&= W_{M_{\tilde{X}}^r}(q_p, q_u)
\end{aligned}$$

First equality comes by removing  $s_u$  from Equation 9. In second equality we have made explicit the dependence of  $\phi_{ij}$  on  $q_{u_j}$  i.e. groundings corresponding to  $j^{th}$  constant. Next, since  $\tilde{X}$  is SOM-R, from Theorem 1 we know  $q_u$  lies at extreme i.e.  $\forall j, q_{u_j} = q_{u_1}$ , so we replace all  $q_{u_j}$  by  $q_{u_1}$  in third equality. Last equality comes from the fact that  $\phi_{ij}$ 's are identical to each other up to re-naming of the index  $j$  as argued earlier. Hence we can write  $\prod_j \phi_{ij}$  as  $(\phi_{i1})^m$ . Hence, in this case, we have  $W_{M_{\tilde{X}}}(q) = W_{M_{\tilde{X}}^r}(q^r)$  implying that the function  $g$  is identity (and hence, monotonically increasing).

From proofs of Case 1 and Case 2 we conclude that  $\exists$  a monotonically increasing function  $g$  such that  $W_M(q) = g(W_{M^r}(q^r))$ .