Weighted Model Counting in FO^2 with Cardinality Constraints and Counting Quantifiers: A Closed Form Formula

Sagar Malhotra[1,2], **Luciano Serafini**[1]

[1]Fondazione Bruno Kessler, Italy
[2]University of Trento, Italy

Abstract

Weighted First-Order Model Counting (WFOMC) computes the weighted sum of the models of a first-order logic theory on a given finite domain. First-Order Logic theories that admit polynomial-time WFOMC w.r.t domain cardinality are called domain liftable. We introduce the concept of *lifted interpretations* as a tool for formulating closed-forms for WFOMC. Using lifted interpretations, we reconstruct the closed-form formula for polynomial-time FOMC in the universally quantified fragment of FO^2, earlier proposed by Beame et al. We then expand this closed-form to incorporate cardinality constraints, existential quantifiers, and counting quantifiers (a.k.a C^2) without losing domain-liftability. Finally, we show that the obtained closed-form motivates a natural definition of a family of weight functions strictly larger than symmetric weight functions.

Introduction

First-Order Logic (FOL) allows specifying structural knowledge with formulas containing variables ranging over all the domain elements. Probabilistic inference in domains described in FOL requires grounding (aka instantiation) of all the individual variables with all the occurrences of the domain elements. This grounding leads to an exponential blow-up of the complexity of the model description and hence the probabilistic inference.

Lifted inference (Poole 2003; Braz, Amir, and Roth 2005) aims at resolving this problem by exploiting symmetries inherent to the FOL structures. In recent years, *Weighted First-Order Model Counting* has emerged as a useful formulation for probabilistic inference in statistical relational learning frameworks (Getoor and Taskar 2007; Raedt et al. 2016). Formally, WFOMC (Chavira and Darwiche 2008) refers to the task of calculating the weighted sum of the models of a formula Φ over a domain of a given finite size $\text{WFOMC}(\Phi, w, n) = \sum_{\omega \models \Phi} w(\omega)$, where n is the cardinality of the domain and w is a *weight function* that associates a real number to each interpretation ω. FOL theories Φ and weight functions w which admit polynomial-time WFOMC w.r.t the domain cardinality are called *domain-liftable* (Van den Broeck et al. 2011). In the past decade, multiple extensions of FO^2 (the fragment of FOL with two variables) have been proven to be domain-liftable (Kazemi et al. 2016; Kuusisto and Lutz 2018; Kuzelka 2020b).

In this paper, instead of relying on an algorithmic approach to WFOMC, as in (Van den Broeck et al. 2011), our objective is to find a closed-form for WFOMC in FO^2 that can be easily extended to larger classes of first-order formulas. To this aim, we introduce the novel notion of *lifted interpretation*. Lifted interpretations allow us to reconstruct the closed-form formula for First-Order Model Counting (FOMC) in the universally quantified fragment of FO^2 as proposed in (Beame et al. 2015) and to extend it to larger classes of FO formulas. We see the following key benefits of the presented formulation:

1. *The formula is easily extended to FO^2 with existential quantifiers, cardinality constraints and, counting quantifiers, without losing domain-liftability*. A cardinality constraint on an interpretation is a constraint on the number of elements for which a certain predicate holds. Counting quantifiers admit expressions of the form $\exists^{\geq m} x \Phi(x)$ expressing that there exist at least m elements that satisfy $\Phi(x)$. Previous works have relied on Lagrange interpolation and Discrete Fourier Transform (Kuzelka 2020b) for evaluating cardinality constraints. In this work, we deal with cardinality constraints in a completely combinatorial fashion.
2. *We provide a complete and uniform treatment of WFOMC in the two-variable fragment*. Multiple extensions of FO^2 have been proven to be domain liftable (Kuzelka 2020b; Kuusisto and Lutz 2018; Van den Broeck, Meert, and Darwiche 2014). Most of these works rely extensively on a variety of logic-based algorithmic techniques. In this paper, we provide a uniform and self-contained combinatorial treatment for all these extensions.
3. *The formula computes WFOMC for a class of weight functions strictly larger than symmetric weight functions.* This extended class of weight functions allows modelling the recently introduced count distributions (Kuzelka 2020a).

Most of the paper focuses on First-Order Model Counting (FOMC) i.e. counting the number of models of a formula Φ over a finite domain of size n denoted by $\text{FOMC}(\Phi, n)$. We then show how WFOMC can be obtained by multiplying each term of the resulting formula for FOMC with the corresponding weight function. This allows us to separate the treatment of the counting part from the weighting part. The paper is therefore structured as follows: The next section describes the related work in the literature on WFOMC. We then present our formulation of the closed-form formula for FOMC given in (Beame et al. 2015) for the universally quantified fragment of FO^2. We then extend this formula to incorporate cardinality constraints, existential quantification, and counting quantifiers, dedicating one section to each of them respectively. The last part of the paper extends the formula for FOMC to WFOMC for the case of symmetric weight functions and for a larger class of weight functions that allow modeling count distributions (Kuzelka 2020a).

Related work

WFOMC was initially defined in (Van den Broeck et al. 2011). The paper provides an algorithm for Symmetric-WFOMC over universally quantified theories based on *knowledge compilation* techniques. The notion of a *domain lifted theory* i.e. a first-order theory for which WFOMC can be computed in polynomial time w.r.t domain cardinality was first formalized in (Van den Broeck 2011). The same paper shows that a theory composed of a set of universally quantified clauses containing at most two variables is domain liftable. (Van den Broeck, Meert, and Darwiche 2014) extends this procedure to theories in full FO^2 (i.e. where existential quantification is allowed) by introducing a skolemization procedure for WFOMC. These results are theoretically analysed in (Beame et al. 2015), which provides a closed-form formula for WFOMC in the universally quantified fragment of FO^2. (Kuusisto and Lutz 2018) extends the domain liftability results to FO^2 with a functionality axiom, and for sentences in *uniform one-dimensional fragment* U_1 (Kieronski and Kuusisto 2015). It also proposes a closed-form formula for WFOMC in FO^2 with functionality constraints. (Kuzelka 2020b) recently proposed a uniform treatment of WFOMC for FO^2 with cardinality constraints and counting quantifiers, proving these theories to be domain-liftable. With respect to the state-of-the-art approaches to WFOMC, we propose an approach that provides a closed-form for WFOMC with cardinality constraints and counting quantifiers from which the PTIME data complexity is immediately evident. Moreover, (Kuzelka 2020b) relies on a sequence of reductions for proving domain liftability of counting quantifiers in the two variable fragment, on the other hand, our approach relies on a single reduction and exploits the principle of inclusion-exclusion to provide a closed-form formula for WFOMC. Finally, (Kuzelka 2020a) introduces Complex Markov Logic Networks, which use complex-valued weights and allow for full expressivity over a class of distributions called *count distributions*. We show in the last section of the paper that our formalization is complete w.r.t. this class of distributions without using complex-valued weight functions.

FOMC for Universal Formulas

Let $\mathcal{L}$ be a first-order function free language with equality. A *pure universal formula* in $\mathcal{L}$ is a formula of the form $\forall x_1 \dots \forall x_m.\Phi(x_1, \dots, x_m)$, where $X = \{x_1, \dots, x_m\}$ is a set of m distinct variables occurring in $\Phi(x_1, \dots, x_m)$, and $\Phi(x_1, \dots, x_m)$ is a quantifier free formula that does not contain any constant symbol. We use the compact notation $\Phi(\boldsymbol{x})$ for $\Phi(x_1, \dots, x_m)$, where $\boldsymbol{x} = (x_1, \dots, x_m)$. Notice that we distinguish between the m-tuple of variables $\boldsymbol{x}$ and the *set* of variables denoted by X. We use C to denote the set of domain constants. For every $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_m)$, m-tuple of constants or variables, $\Phi(\boldsymbol{\sigma})$ denotes the result of uniform substitution of x_i with σ_i in $\Phi(\boldsymbol{x})$. If $\Sigma \subseteq X \cup C$ is the set of constants or variables of $\mathcal{L}$ and $\forall \boldsymbol{x}\Phi(\boldsymbol{x})$ a pure universal formula then:

$$\Phi(\Sigma) = \bigwedge_{\boldsymbol{\sigma} \in \Sigma^m} \Phi(\boldsymbol{\sigma}) \tag{1}$$

$\Phi(\Sigma)$ is a very convenient notion, for instance, grounding of a pure universal formula $\forall \boldsymbol{x}.\Phi(\boldsymbol{x})$ over a set of domain constants C, can be simply denoted as $\Phi(C)$. Furthermore, $\Phi(X)$ and $\Phi(\boldsymbol{x})$ have the following useful relationship:

Lemma 1. *For any arbitrary pure universal formula $\forall \boldsymbol{x}\Phi(\boldsymbol{x})$, the following equivalence holds:*

$$\forall \boldsymbol{x}\Phi(\boldsymbol{x}) \leftrightarrow \forall \boldsymbol{x}\Phi(X) \tag{2}$$

Example 1. *Let $\Phi(x, y) = A(x) \wedge R(x, y) \wedge x \neq y \rightarrow A(y)$, then $\Phi(X = \{x, y\})$ is the following formula*

$$\begin{aligned} &(A(x) \wedge R(x,x) \wedge x \neq x \rightarrow A(x)) \\ \wedge\, &(A(x) \wedge R(x,y) \wedge x \neq y \rightarrow A(y)) \\ \wedge\, &(A(y) \wedge R(y,x) \wedge y \neq x \rightarrow A(x)) \\ \wedge\, &(A(y) \wedge R(y,y) \wedge y \neq y \rightarrow A(y)) \end{aligned} \tag{3}$$

Due to Lemma 1, we can assume that in any grounding of $\forall x \forall y.\Phi(X = \{x, y\})$, two distinct variables x and y, are always grounded to different domain elements. This is because the cases in which x and y are grounded to the same domain element are taken into account by the conjuncts $\Phi(x, x)$ and $\Phi(y, y)$ in $\Phi(X)$. See, for instance, the first and the last conjunct of (3).

Definition 1 (Lifted interpretation)**.** *A* lifted interpretation *τ of a pure universal formula $\forall \boldsymbol{x}\Phi(\boldsymbol{x})$ is a function that assigns to each atom of $\Phi(X)$ either 0 or 1 (0 means false and, 1 means true) and assigns 1 to $x_i = x_i$ and 0 to $x_i = x_j$ if $i \neq j$.*

Lifted interpretations are different from FOL interpretations as they assign truth values to the atoms that contain free variables. Instead, lifted interpretations are similar to m-types (Kuusisto and Lutz 2018)(we will later formalize this similarity), where m is the number of variables in the language $\mathcal{L}$. The *truth value* of a pure universal formula Φ under the lifted interpretation τ denoted by $\tau(\forall \boldsymbol{x}\Phi(\boldsymbol{x}))$, can be computed by applying the classical semantics for propositional connectives to the evaluations of the atoms in $\Phi(X)$. With abuse of notation, we sometimes write also $\tau(\Phi(X))$ instead of $\tau(\forall \boldsymbol{x}\Phi(\boldsymbol{x}))$.

Example 2. *The following is an example of a lifted interpretation for the formula* (3) *of Example 1:*

	$A(x)$	$R(x,x)$	$A(y)$	$R(y,y)$	$R(x,y)$	$R(y,x)$
τ	0	1	1	1	0	1
	τ_x		τ_y		τ_{xy}	

We omit the truth assignments of equality atoms since they are fixed for all lifted interpretations. We have that $\tau((3)) = 0$.

As highlighted in the previous example, any lifted interpretation τ can be decomposed into a set of partial lifted interpretations τ_Y where $Y \subseteq X$. Notice that τ_Y assigns truth value to all the atoms that contain *all* the variables Y. For instance, in Example 2, $\tau_{\{x,y\}}$ (denoted as τ_{xy} in the example) assigns to atoms $R(x,y)$ and $R(y,x)$ only and not to the atoms $R(x,x)$ and $A(x)$. In general, we will use the simpler notation τ_{xy} to denote the partial lifted interpretation $\tau_{\{x,y\}}$.

Beame et al. in (Beame et al. 2015) provide a mathematical formula for computing $\text{FOMC}(\Phi, n)$, where Φ is a pure universal formula in FO^2, i.e., sentences of the form $\forall xy.\Phi(x,y)$. In the following, we reconstruct this result using the notion of lifted interpretations. As it will be clearer later, using lifted interpretations allow us to seamlessly extend the result to larger extensions of FO^2 formulas.

Let $\forall xy.\Phi(x,y)$ be a pure universal formula. Let u be the number of atoms whose truth values are assigned by τ_x i.e., first-order atoms containing only the variable x. Let $P_1(x), \ldots, P_u(x)$ be an ordering of these atoms[1]. There are 2^u possible partial lifted interpretations τ_x which could assign truth values to these atoms. We assume that each such partial lifted interpretation τ_x is indexed by an integer i, where $0 \leq i \leq 2^u - 1$. Hence, the i-th partial lifted interpretation τ_x is defined as $\tau_x(P_j(x)) = bin(i)_j$, where $bin(i)_j$ represents the value of the j^{th} bit in the binary encoding of i, for all $1 \leq j \leq u$. We use $i(x)$ to denote the conjunction of a maximally consistent set of literals (atoms and negated atoms) containing only the variable x which are satisfied by the i-th τ_x. For instance, in Example 1, $A(x)$ and $R(x,x)$ are the atoms assigned by τ_x. Assuming the order of atoms to be $(A(x), R(x,x))$, we have that $\tau_x = 1$ implies that $A(x)$ is interpreted to be false and $R(x,x)$ is interpreted to be true. Also, $i(x)$ denotes $\neg A(x) \wedge R(x,x)$ if $i = 1$ and $\neg A(x) \wedge \neg R(x,x)$ if $i = 0$. We use similar notation for atoms assigned by $\tau_y(y)$ i.e. atoms containing only the variable y. Furthermore, we use $i(c)$ to denote the conjunction of ground atoms containing only one constant c. In Example 1, $i(c)$ denotes $\neg A(c) \wedge R(c,c)$ if $i = 1$ and $\neg A(c) \wedge \neg R(c,c)$ if $i = 0$. Clearly, $i(x)$ exactly corresponds to 1-types. Furthermore, given an interpretation ω if $\omega \models i(c)$ then we say that c is of 1-type i.

For a pure universal formula Φ and $0 \leq i \leq j \leq 2^u - 1$, let n_{ij} be the number of lifted interpretations τ that satisfy Φ such that $\tau_x = i$ and $\tau_y = j$. Formally:

$$n_{ij} = |\{\tau \mid \tau \models \Phi(\{x,y\}) \wedge i(x) \wedge j(y)\}|$$

[1]The list includes atoms of the form $P(x,x)$ for binary predicate P.

Example 3 (Example 1 cont'd)**.** *The set of atoms containing only x or only y in the formula* (3) *are $\{A(x), R(x,x)\}$ and $\{A(y), R(y,y)\}$ respectively. In this case $u = 2$. The partial lifted interpretations τ_x and τ_y corresponding to the lifted interpretation τ of Example 2 are: $\tau_x = 1$ and $\tau_y = 3$. n_{13} is the number of lifted interpretations satisfying* (3) *and agreeing with $\tau_x = 1$ and $\tau_y = 3$. In this case $n_{13} = 2$. The other cases are as follows:*

n_{00}	n_{01}	n_{02}	n_{03}	n_{11}	n_{12}	n_{13}	n_{22}	n_{23}	n_{33}
4	4	2	2	4	2	2	4	4	4

Theorem 1 (Beame et al.)**.** *For any pure universal formula $\forall xy.\Phi(x,y)$*

$$\text{FOMC}(\forall xy.\Phi(x,y), n) = \sum_{\sum \boldsymbol{k}=n} \binom{n}{\boldsymbol{k}} \prod_{0 \leq i \leq j \leq 2^u - 1} n_{ij}^{\boldsymbol{k}(i,j)} \tag{4}$$

where $\boldsymbol{k} = (k_0, ..., k_{2^u-1})$ is a 2^u-tuple of non-negative integers, $\binom{n}{\boldsymbol{k}}$ is the multinomial coefficient and

$$\boldsymbol{k}(i,j) = \begin{cases} \frac{k_i(k_i-1)}{2} & \text{if } i = j \\ k_i k_j & \text{otherwise} \end{cases}$$

We provide the proof for Theorem 1 along with some additional Lemmas in the appendix. Intuitively, k_i represents the number of constants c of 1-type i. Hence, for a given $\boldsymbol{k}$, we have $\binom{n}{\boldsymbol{k}}$ possible interpretations. Furthermore, given a pair of constants c and d such that c is of 1-type i and d is of 1-type j, the number of truth assignments to binary predicates containing both c and d is given by n_{ij} independently of all other constants. Finally, the exponent $\boldsymbol{k}(i,j)$ accounts for all possible pair-wise choices of constants given a $\boldsymbol{k}$ vector.

Notice that the method in (Beame et al. 2015) requires additional $n + 1$ calls to a counting oracle for dealing with equality. Lifted interpretations on the other hand allow us to fix the truth values of the equality atoms, by assuming (w.l.o.g.) that different variables are assigned to distinct objects in $\forall \boldsymbol{x}.\Phi(X)$. The equality atoms then contribute to the model count only through n_{ij}, hence, allowing us to deal with equality in constant time w.r.t domain cardinality.

Example 4 (Example 1 continued)**.** *Consider a domain of 3 elements (i.e., $n = 3$). Each term of the summation* (4) *is of the form*

$$\binom{3}{k_0, k_1, k_2, k_3} \prod_{i=0}^{3} n_{ii}^{\frac{k_i(k_i-1)}{2}} \prod_{\substack{i<j \\ i=0}}^{3} n_{ij}^{k_i k_j}$$

which is the number of models with k_0 elements for which $A(x)$ and $R(x,x)$ are both false; k_1 elements for which $A(x)$ is false and $R(x,x)$ true, k_2 elements for which $A(x)$ is true and $R(x,x)$ is false and k_3 elements for which $A(x)$ and $R(x,x)$ are both true. For instance: $\binom{3}{2,0,0,1} n_{00}^1 n_{03}^2 = \binom{3}{2,0,0,1} 4^1 \cdot 2^2 = 3 \cdot 16 = 48$ is the number of models in which 2 elements are such that $A(x)$ and $R(x,x)$ are false and 1 element such that $A(x)$ and $R(x,x)$ are both true.

FOMC for Cardinality Constraints

Cardinality constraints are arithmetic expressions that impose restrictions on the number of times a certain predicate is interpreted to be true. A simple example of a cardinality constraint is $|A| = m$, for some unary predicate A and positive integer m. This cardinality constraint is satisfied by any interpretation in which $A(c)$ is interpreted to be true for exactly m distinct constants c in the domain C. A more complex example of a cardinality constraint could be: $|A| + |B| \leq |C|$, where A, B and C are some predicates in the language.

For every interpretation ω of the language $\mathcal{L}$ on a finite domain C, we define $A^\omega = \{c \in C \mid \omega \models A(c)\}$ if A is unary, and $A^\omega = \{(c,d) \in C \times C \mid \omega \models A(c,d)\}$ if A is binary. ω satisfies a cardinality constraint ρ, in symbols $\omega \models \rho$, if the arithmetic expression, obtained by replacing $|A|$ with $|A^\omega|$ for every predicate A in ρ, is satisfied.

If a cardinality constraint involves only unary predicates, then we can exploit Theorem 1 considering only a subset of $\boldsymbol{k}$'s. The multinomial coefficient $\binom{n}{\boldsymbol{k}}$ counts the models that contain exactly k_i elements of 1-type i, the cardinality of the unary predicates in these models are fully determined by $\boldsymbol{k}$.

To deal with cardinality constraints involving binary predicates, we have to expand the formula (4) by including also the assignments to binary predicates. This implies extending the $\boldsymbol{k}$ vector in order to consider assignments to atoms that contain both variables x and y. Let $R_0(x,y), R_1(x,y), \ldots, R_b(x,y)$ be an enumeration of all the atoms in $\Phi(X)$ that contain both variables x and y. Notice that the order of variables leads to different atoms, for instance in Example 1, we have two binary atoms $R_1(x,y) = R(x,y)$ and $R_2(x,y) = R(y,x)$.

For every $0 \leq v \leq 2^b - 1$ let v denote the v^{th} partial lifted interpretation τ_{xy}, such that τ_{xy} assigns $bin(v)_j$ to the j-th binary atom $R_j(x,y)$ for every $1 \leq j \leq b$. As for the unary case, $v(x,y)$ represents the conjunction of all the literals that are satisfied by v. For instance, in example 1, $v(x,y)$ denotes $\neg R(x,y) \wedge R(y,x)$ if $v = 1$ and $R(x,y) \wedge \neg R(y,x)$ if $v = 3$. Clearly, the set of 2-types in the language of the formula Φ correspond to $i(x) \wedge j(x) \wedge v(x,y)$ .We define n_{ijv} as follows:

$$n_{ijv} = |\{\tau \mid \tau \models \Phi(\{x,y\}) \wedge i(x) \wedge j(y) \wedge v(x,y)\}|$$

Notice that $n_{ij} = \sum_{v=0}^{2^b-1} n_{ijv}$ and that n_{ijv} is either 0 or 1.

Example 5. *For instance n_{13} introduced in Example 3 expands to $n_{130} + n_{131} + n_{132} + n_{133}$ where n_{13v} corresponds to the following assignments:*

$A(x)$	$R(x,x)$	$A(y)$	$R(y,y)$	$R(x,y)$	$R(y,x)$	v	n_{13v}
0	1	1	1	0	0	0	$n_{130} = 1$
				0	1	1	$n_{131} = 0$
				1	0	2	$n_{132} = 1$
				1	1	3	$n_{133} = 0$
$\tau_x = 1$		$\tau_y = 3$		$\tau_{xy} = v$			

By replacing n_{ij} in equation (4) with its expansion $\sum_{v=0}^{2^b-1} n_{ijv}$ we obtain that $\text{FOMC}(\forall xy.\Phi(x,y), n)$ is equal to

$$\sum_{\sum \boldsymbol{k} = n} \binom{n}{\boldsymbol{k}} \prod_{0 \leq i \leq j \leq 2^u - 1} \left(\sum_{0 \leq v \leq 2^b - 1} n_{ijv} \right)^{\boldsymbol{k}(i,j)}$$
$$= \sum_{\boldsymbol{k},\boldsymbol{h}} \binom{n}{\boldsymbol{k}} \prod_{0 \leq i \leq j \leq 2^u - 1} \binom{\boldsymbol{k}(i,j)}{\boldsymbol{h}^{ij}} \prod_{0 \leq v \leq 2^b - 1} n_{ijv}^{h_v^{ij}} \quad (5)$$

where, for every $0 \leq i \leq j \leq 2^u - 1$, $\boldsymbol{h}^{ij}$ is a vector of 2^b integers that sum up to $\boldsymbol{k}(i,j)$. To simplify the notation we define the function $F(\boldsymbol{k}, \boldsymbol{h}, \Phi)$ where Φ is a pure universal formula as follows

$$F(\boldsymbol{k}, \boldsymbol{h}, \Phi) = \binom{n}{\boldsymbol{k}} \prod_{0 \leq i \leq j \leq 2^u - 1} \binom{\boldsymbol{k}(i,j)}{\boldsymbol{h}^{ij}} \prod_{0 \leq v \leq 2^b - 1} n_{ijv}^{h_v^{ij}}$$

where h_v^{ij} is the v-th element of the vector $\boldsymbol{h}^{ij}$, which represents the number of pairs of constants of distinct elements that satisfy the 2-type $i(x) \wedge j(y) \wedge v(x,y)$. We will now show that the $(\boldsymbol{k}, \boldsymbol{h})$ vectors contain all the necessary information for determining the cardinality of the binary predicates.

For every $\mathcal{L}$-interpretation ω on the finite domain C, we define $(\boldsymbol{k}, \boldsymbol{h})^\omega = (\boldsymbol{k}^\omega, \boldsymbol{h}^\omega)$ with $\boldsymbol{k}^\omega = \langle k_0^\omega, \ldots, k_{2^u-1}^\omega \rangle$ such that k_i^ω is the number of constants $c \in C$ such that $\omega \models i(c)$. $\boldsymbol{h}^\omega$ is equal to $\{(\boldsymbol{h}^{ij})^\omega\}_{0 \leq i \leq j \leq 2^b - 1}$, where $(\boldsymbol{h}^{ij})^\omega = \langle (h_0^{ij})^\omega, \ldots (h_{2^b-1}^{ij})^\omega \rangle$ such that $(h_v^{ij})^\omega$ is the number of pairs (c,d) with $c \neq d$ such that $\omega \models i(c) \wedge j(d) \wedge v(c,d)$ if $i < j$. When $i = j$, $(h_v^{ii})^\omega$ is equal to the count of the unordered pairs (c,d) (i.e. only one of the (c,d) and (d,c) is counted) for which $\omega \models i(c) \wedge i(d) \wedge v(c,d)$.

Lemma 2. *For every predicate P, and interpretations ω_1 and ω_2, $(\boldsymbol{k}, \boldsymbol{h})^{\omega_1} = (\boldsymbol{k}, \boldsymbol{h})^{\omega_2}$ implies $|P^{\omega_1}| = |P^{\omega_2}|$.*

Proof. Let $(\boldsymbol{k}, \boldsymbol{h})$ be a vector such that $(\boldsymbol{k}, \boldsymbol{h}) = (\boldsymbol{k}, \boldsymbol{h})^\omega$. The Lemma is true iff $(\boldsymbol{k}, \boldsymbol{h})$ uniquely determines the cardinality of P^ω. If P^ω is a unary predicate whose atom is indexed by s in the ordering of the unary atoms, then the cardinality of P^ω can be given as $\sum_{i=0}^{2^u-1} bin(i)_s \cdot k_i$. Similarly, if P is binary then in order to count P^ω, we need to take into account both $\boldsymbol{k}$ and $\boldsymbol{h}$. Let $P(x,x)$ be the atom indexed s i.e. P_s, let $P(x,y)$ be the atom indexed l i.e. P_l and let $P(y,x)$ be the atom indexed r i.e. P_r, then the cardinality of P if P is binary is given as $\sum_{i=0}^{2^u-1} bin(i)_s \cdot k_i + \sum_{i \leq j} \sum_{v=0}^{2^b-1} (bin(v)_l + bin(v)_r) \cdot h_v^{ij}$. □

Example 6. *Consider formula* (3) *with the additional conjunct $|A| = 2$ and $|R| = 2$. The constraint $|A| = 2$ implies that we have to consider $\boldsymbol{k}$ such that $k_2 + k_3 = 2$. $|R| = 2$ constraint translates to only considering $(\boldsymbol{k}, \boldsymbol{h})$ with $k_1 + k_3 + \sum_{i \leq j}(h_1^{ij} + h_2^{ij} + 2h_3^{ij}) = 2$.*

For a given $(\boldsymbol{k}, \boldsymbol{h})$, we use the notation $\boldsymbol{k}(P)$ to denote cardinality of P if P is unary and $(\boldsymbol{k}, \boldsymbol{h})(P)$ if P is binary. Using Lemma 2, we can conclude that $\text{FOMC}(\Phi \wedge \rho, n)$ where Φ is a pure universal formula with 2 variables can be computed by considering only the $(\boldsymbol{k}, \boldsymbol{h})$'s that satisfy ρ, i.e., those $(\boldsymbol{k},\boldsymbol{h})'s$ where ρ evaluates to true, when $|P|$ is substituted with $(\boldsymbol{k},\boldsymbol{h})(P)$ when P is binary, and $\boldsymbol{k}(P)$ when P is unary.

Corollary 1 (of Theorem 1). *For every pure universal formula Φ and cardinality constraint ρ,* $\text{FOMC}(\Phi \wedge \rho, n) = \sum_{\boldsymbol{k},\boldsymbol{h}\models\rho} F(\boldsymbol{k},\boldsymbol{h},\Phi)$

FOMC for Existential Quantifiers

In this section, we provide a proof for model counting in the presence of existential quantifiers. The key difference in our approach w.r.t (Beame et al. 2015) is that we make explicit use of the principle of inclusion-exclusion and we will later generalize the same approach to counting quantifiers. We will first provide a corollary of the principle of inclusion-exclusion.

Corollary 2 ((Wilf 2006) section 4.2). *Let Ω be a set of objects and let $\mathcal{S} = \{S_1,\dots,S_m\}$ be a set of subsets of Ω. For every $\mathcal{Q} \subseteq \mathcal{S}$, let $N(\supseteq \mathcal{Q})$ be the count of objects in Ω that belong to all the subsets $S_i \in \mathcal{Q}$, i.e., $N(\supseteq \mathcal{Q}) = \left|\{\bigcap_{S_i\in Q} S_i\}\right|$. For every $0 \leq l \leq m$, let $s_l = \sum_{|\mathcal{Q}|=l} N(\supseteq \mathcal{Q})$ and let e_0 be count of objects that do not belong to any of the S_i in $\mathcal{S}$, then*

$$e_0 = \sum_{l=0}^{m}(-1)^l s_l \tag{6}$$

Any arbitrary formula in FO^2 can be reduced to an equisatisfiable reduction called Scott's Normal Form (SNF) (Scott 1962). Moreover, SNF preserves FOMC as well as WFOMC if all the new predicates and there negation are assigned a unit weight (Kuusisto and Lutz 2018). A formula in SNF has the following form:

$$\forall x\forall y.\Phi(x,y) \wedge \bigwedge_{i=1}^{q} \forall x\exists y.\Psi_i(x,y) \tag{7}$$

where $\Phi(x,y)$ and $\Psi_i(x,y)$ are quantifier-free formulae.

Theorem 2. *For an FO^2 formula in Scott's Normal Form as given in* (7)*, let $\Phi' = \forall xy.(\Phi(x,y) \wedge \bigwedge_{i=1}^{q} P_i(x) \rightarrow \neg\Psi_i(x,y))$ where P_i's are fresh unary predicates, then:*

$$\text{FOMC}((7),n) = \sum_{\boldsymbol{k},\boldsymbol{h}}(-1)^{\sum_i \boldsymbol{k}(P_i)} F(\boldsymbol{k},\boldsymbol{h},\Phi') \tag{8}$$

Proof. Let Ω be the set of models of $\forall xy.\Phi(x,y)$ over the language of Φ and $\{\Psi_i\}$ (i.e., the language of Φ' excluding the the predicates P_i) and on a domain C consisting of n elements. Let $\mathcal{S} = \{\Omega_{ci}\}_{c\in C,\ 1\leq i\leq q}$ be the set of subsets of Ω where Ω_{ci} is the set of $\omega \in \Omega$, such that $\omega \models \forall y.\neg\Psi_i(c,y)$. Notice that for every model ω of (7) $\omega \not\models \forall y \neg\Psi_i(c,y)$, for any pair of i and c i.e. ω is not in any Ω_{ci}. Also, for every ω, if $\omega \in \Omega_{ci}$ then $\omega \not\models$ (7). Hence, $\omega \models$ (7) if and only if $\omega \notin \Omega_{ci}$ for all c and i. Therefore, the count of models of (7) is equal to the count of models in Ω which do not belong to any Ω_{ci}.

If we are able to compute s_l (as introduced in Corollary 2), then we could use Corollary 2 for computing cardinality of all the models which do not belong to any Ω_{ci} and hence $\text{FOMC}((7),n)$.

For every $0 \leq l \leq n\cdot q$, let us define

$$\Phi'_l = \Phi' \wedge \sum_{i=1}^{q}|P_i| = l \tag{9}$$

We will now show that s_l is exactly given by $\text{FOMC}((9),n)$.

Every model of Φ'_l is an extension of an $\omega \in \Omega$ that belongs to at least l elements in $\mathcal{S}$. In fact, for every model ω of $\forall xy.\Phi(x,y)$ i.e $\omega \in \Omega$, if $\mathcal{Q}'$ is the set of elements of $\mathcal{S}$ that contain ω, then ω can be extended into a model of Φ'_l in $\binom{|Q'|}{l}$ ways. Each such model can be obtained by choosing l elements in Q' and interpreting $P_i(c)$ to be true in the extended model, for each of the l chosen elements $\Omega_{ci} \in Q'$. On the other hand, recall that $s_l = \sum_{|\mathcal{Q}|=l} N(\supseteq Q)$. Hence, for any $\omega \in \Omega$ if $\mathcal{Q}'$ is the set of elements of $\mathcal{S}$ that contain ω, then there are $\binom{|\mathcal{Q}'|}{l}$ distinct subsets $\mathcal{Q} \subseteq \mathcal{Q}'$ such that $|\mathcal{Q}| = l$. Hence, we have that ω contributes $\binom{|\mathcal{Q}'|}{l}$ times to s_l. Therefore, we can conclude that

$$s_l = \text{FOMC}(\Phi'_l,n) = \sum_{|\mathcal{Q}|=l} N(\supseteq Q)$$

and by the principle of inclusion-exclusion as given in Corollary 2, we have that :

$$\begin{aligned}
\text{FOMC}((7),n) = e_0 &= \sum_{l=0}^{n\cdot q}(-1)^l s_l \\
&= \sum_{l=0}^{n\cdot q}(-1)^l \text{FOMC}(\Phi'_l,n) \\
&= \sum_{l=0}^{n\cdot q}(-1)^l \sum_{\boldsymbol{k},\boldsymbol{h}\models\sum_i|P_i|=l} F(\boldsymbol{k},\boldsymbol{h},\Phi') \\
&= \sum_{\boldsymbol{k},\boldsymbol{h}}(-1)^{\sum_i \boldsymbol{k}(P_i)} F(\boldsymbol{k},\boldsymbol{h},\Phi')
\end{aligned}$$

□

FOMC for Counting Quantifiers

Counting quantifiers are expressions of the form $\exists x^{\geq m} y.\Psi$, $\exists^{\leq m} y.\Psi$, and $\exists^{=m} y.\Psi$. The extension of FO^2 with such quantifiers is denoted by C^2 (Rosen, Graedel, and Otto 1997). In this section, we show how FOMC in C^2 can be performed by exploiting the formula for FOMC in FO^2 with cardinality constraints. We assume that the counting quantifier $\exists^{\leq m} y.\Psi$ is expanded to $\bigvee_{k=0}^{m} \exists^{=k} y.\Psi$, and the quantifiers $\exists^{\geq m} y.\Psi$ are first transformed to $\neg(\exists^{\leq m-1} y.\Psi)$ and then expanded. We are therefore left with quantifiers of the form $\exists^{=m} y.\Psi$. Hence, any C^2 formula can be transformed into a formula of the form $\Phi_0 \wedge \bigwedge_{k=1}^{q} \forall x.(A_k(x) \leftrightarrow \exists^{=m_k} y.\Psi_k)$ that preserves FOMC, where[2] Φ_0 is a pure universal formula obtained by replacing every occurrence of the sub-formula $\exists^{=m_k} y.\Psi_k$ with $A_k(x)$, where A_k is a fresh predicate. W.l.o.g, we can assume that Ψ_k is the atomic

[2]We assume that Φ_0 contains no existential quantifiers as they can be transformed as described in Theorem 2.

formula $R_k(x,y)$. We will now present a closed-form for FOMC of $\Phi_0 \wedge \bigwedge_k \forall x.(A_k(x) \leftrightarrow \exists^{=m_k} y.R_k(x,y))$. For the sake of notational convenience, we use $\Phi_{i..j}$ to denote $\bigwedge_{i\le s\le j} \Phi_s$ for any set of formulas $\{\Phi_s\}$.

Theorem 3. *Let Φ be the follwoing C^2 formula :*

$$\Phi_0 \wedge \bigwedge_{k=1}^{q} \forall x.(A_k(x) \leftrightarrow \exists^{=m_k} y.R_k(x,y))$$

where Φ_0 is a pure universal formula in FO^2. Let us define the following formulas for each k, where $1 \le k \le q$:

$\Phi_1^k = \bigwedge_{i=1}^{m_k} \forall x \exists y.A_k(x) \vee B_k(x) \to f_{ki}(x,y)$

$\Phi_2^k = \bigwedge_{1\le i<j\le m_k} \forall x \forall y.f_{ki}(x,y) \to \neg f_{kj}(x,y)$

$\Phi_3^k = \bigwedge_{i=1}^{m_k} \forall x \forall y.f_{ki}(x,y) \to R_k(x,y)$

$\Phi_4^k = \forall x.B_k(x) \to \neg A_k(x)$

$\Phi_5^k = \forall x \forall y.M_k(x,y) \leftrightarrow ((A_k(x) \vee B_k(x)) \wedge R_k(x,y))$

$\Phi_6^k = |A_k| + |B_k| = |f_{k1}| = \cdots = |f_{km_k}| = \frac{|M_k|}{m_k}$

where[3] *B_k, f_{ki} and M_k are fresh predicates. Then* $\text{FOMC}(\Phi,n)$ *is given as:*

$$\sum_{(\boldsymbol{k},\boldsymbol{h})\models \bigwedge_k \Phi_6^k} \frac{(-1)^{\sum_k \boldsymbol{k}(B_k)+\sum_{k,i}\boldsymbol{k}(P_{ki})} F(\boldsymbol{k},\boldsymbol{h},\Phi')}{\prod_k m_k!^{\boldsymbol{k}(A_k)}}$$

where Φ' is obtained by replacing each Φ_1^k with $\bigwedge_{i=1}^{m_k} \forall x \forall y.P_{ki}(x) \to \neg(A_k(x) \vee B_k(x) \to f_{ki}(x,y))$ in $\Phi_0 \wedge \bigwedge_k \Phi_{1..5}^k$ and P_{ki} are fresh unary predicates.

Lemma 3. *If $\omega \models \Phi_0 \wedge \bigwedge_{k=1}^q \Phi_{1..6}^k$ then every $c \in A_k^\omega \cup B_k^\omega$ has exactly m_k R_k-successors i.e., $\omega \models \exists^{=m_k} y.R(c,y)$.*

Proof. If $c \in A_k^\omega \cup B_k^\omega$, then by Φ_1^k, c has an f_{ki}-successor for every $1 \le i \le m_k$. Φ_2^k implies that c has distinct f_{ki} and f_{kj} successor for any choice of i and j. Φ_3^k implies that any f_{ki}-successor of c is also an R_k-successor. Hence, c has at least m_k R_k-successors.

Axiom Φ_5^k implies that c has exactly as many R_k-successors as M_k-successors. Hence, c has at-least m_k M_k-successors. Furthermore, by Φ_4^k we have that A_k^ω and B_k^ω are disjoint. Hence, using Φ_6^k, we can conclude that c has exactly m_k M_k-successors. Finally, using Φ_5^k we can conclude that c has exactly m_k R_k-successors. □

Proof (of Theorem 3). First notice that every model ω of Φ can be extended to $\prod_k m_k!^{A_k^\omega}$ models of $\Phi_0 \wedge \bigwedge_k \Phi_{1..6}^k$ by interpreting B_k in the empty set, f_{ki} in the set of pairs $\langle c,d\rangle$ for $c \in A_k^\omega$ and d being the i-th R_k-successor of c (for some ordering of the R_k-successors) and M_k according to the definition given in Φ_k^5.

Let Ω the set of models of $\Phi_0 \wedge \bigwedge_{k=1}^q \Phi_{1..6}^k$ restricted to the language of Φ, M_k and f_{ki} (i.e., the language of $\Phi_0 \wedge \bigwedge_{k=1}^q \Phi_{1..6}^k$ excluding the predicates B_k) and on a domain C consisting of n elements.

Notice that Ω contains also the models that are not extensions of some model of Φ. Therefore, in the first part of the proof we count the number of extensions of models of Φ in Ω and successively we will take care of the over-counting due to the multiple interpretations of f_{ki}'s.

Let $\mathcal{S} = \{\Omega_{ck}\}$ be the the set of subsets of Ω such that if $\omega \in \Omega_{ck}$ then $\omega \models \neg A_k(c) \wedge \exists^{=m_k} y.R_k(c,y)$. Due to Lemma 3, if $\omega \in \Omega$ then $\omega \models \bigwedge_k \forall x.A_k(x) \to \exists^{=m_k} y.R_k(x,y)$. Hence, in order to count the models of Φ in Ω we only need to count the number of models in Ω that satisfy $\bigwedge_k \forall x \exists^{=m_k} y.R_k(x,y) \to A_k(x)$, equivalently, the number of models that belong to none of the Ω_{ck}. Hence, if we are able to evaluate s_l (as introduced in Corollary 2) then we can use Corollary 2 to count the set of models in Ω that satisfy Φ.

Let $\omega \in \Omega$. Let us define Φ_l for $l \ge 0$ as follows:

$$\Phi_l = \Phi_0 \wedge \bigwedge_k \Phi_{1..6}^k \wedge \left(\sum_k |B_k| = l\right) \tag{10}$$

Firstly, let $\mathcal{Q}'$ be the set of elements in $\mathcal{S}$ that contain ω. By Lemma 3, ω can be extended in $\binom{|\mathcal{Q}'|}{l}$ models of Φ_l. Each such extension can be achieved by choosing l elements in $\mathcal{Q}'$, and interpreting $B_k(c)$ to be true in the extended model iff Ω_{ck} is a part of the l chosen elements. On the other hand, recall that $s_l = \sum_{|\mathcal{Q}|=l} N(\supseteq \mathcal{Q})$. Every ω that is contained in all the elements of $\mathcal{Q}'$, contributes $\binom{|\mathcal{Q}'|}{l}$ to s_l. Hence, $s_l = \text{FOMC}(\Phi_l, n)$. Using inclusion-exclusion principle (corollary 2), we have that the number of models which do not belong to any of the Ω_{ck} are:

$$\sum_l (-1)^l s_l = \sum_l (-1)^l \text{FOMC}(\Phi_l, n) \tag{11}$$

Hence, we have the count of models of Φ in Ω. But notice that this is the count of the models of Φ in the language of $\Phi_0 \wedge \bigwedge_k \Phi_{1..6}^k$ excluding B_k, where there are the additional predicates $\{f_{ki}\}$. Since every interpretation with $|A_k^\omega| = r_k$ can be extended in $m_k!^{r_k}$ models of Φ due to the permutations of $\{f_{ki}\}_{i=1}^{m_k}$, to obtain FOMC on the language of Φ we have to take into account this over-counting[4]. This can be obtained by introducing a cardinality constraint $|A_k| = r_k$ for every A_k and dividing by $m_k!^{r_k}$ for each k and $r_1...r_q$ values. Giving the following expression for $\text{FOMC}(\Phi,n)$:

$$\sum_{l,r_k} (-1)^l \frac{\text{FOMC}(\Phi_l \wedge \bigwedge_k |A_k| = r_k, n)}{\prod_k m_k!^{r_k}} \tag{12}$$

Also notice that Φ_1^k contains m_k existential quantifiers, to eliminate them we use the result of Theorem 2. We introduce m_k new unary predicates $P_{k1},\dots,P_{km_k}$ for each k, and replace each Φ_1^k with $\bigwedge_i \forall x \forall y.P_{ki}(x) \to \neg(A_k(x) \vee B_k(x) \to f_{ki}(x,y))$.

[3] If Φ_0 is obtained after a transformation as described in Theorem 2, then we can add the term $\sum_g k(P_g)$ to the exponent of (-1), for the set of unary predicates $\{P_g\}$ introduced to deal with existential quantifiers. Also, any cardinality constraint on predicates of Φ_0 can be easily conjuncted and incorporated into $\wedge_k \Phi_6^k$

[4] Notice that M_k leads to no additional models of Φ as interpretations of M_k are uniquely determined by A_k and R_k by Φ_5^k

By Theorem 2 we have that $\text{FOMC}(\Phi, n)$ is equal to:

$$\sum_{(\boldsymbol{k},\boldsymbol{h})\models\bigwedge_k \Phi_6^k} \frac{(-1)^{\sum_k \boldsymbol{k}(B_k)+\sum_{k,i}\boldsymbol{k}(P_{ki})} F(\boldsymbol{k},\boldsymbol{h},\Phi')}{\prod_k m_k!^{\boldsymbol{k}(A_k)}}$$

where Φ' is the pure universal formula $\Phi_0 \wedge \bigwedge_{k=1}^q \Phi_{2..5}^k \wedge \bigwedge_{i,k} P_{ki}(x) \to \neg(A_k(x) \vee B_k(x) \to f_{ki}(x,y))$. □

Weighted First-Order Model Counting

All the FOMC formulas introduced so far can be easily extended to weighted model counting by simply defining a positive real-valued weight function $w(\boldsymbol{k},\boldsymbol{h})$ and adding it as a multiplicative factor to $F(\boldsymbol{k},\boldsymbol{h},\Phi)$ in all FOMC formulas. The case of Symmetric-WFOMC can be obtained by defining $w(\boldsymbol{k},\boldsymbol{h})$ as follows:

$$w(\boldsymbol{k},\boldsymbol{h}) = \prod_{P\in\mathcal{L}} w(P)^{(\boldsymbol{k},\boldsymbol{h})(P)} \cdot \bar{w}(P)^{(\boldsymbol{k},\boldsymbol{h})(\neg P)}$$

where $w(P)$ and $\bar{w}(P)$ associate positive real values to predicate P and it's negation respectively. But symmetric-weight functions are clearly not the most general class of weight functions. (Kuzelka 2020a) introduced a strictly more expressive class of weight functions which also preserves domain liftability. These weight functions can express count distributions, which are defined as follows:

Definition 2 (Count distribution (Kuzelka 2020a))**.** *Let* $\Phi = \{\alpha_i, w_i\}_{i=1}^m$ *be a Markov Logic Network defining a probability distribution* $p_{\Phi,\Omega}$ *over a set of possible worlds (we call them assignments) of a formula* Ω*. The count distribution of* Φ *is the distribution over* m*-dimensional vectors of non-negative integers* $\boldsymbol{n}$ *given by*

$$q_\Phi(\Omega,\boldsymbol{n}) = \sum_{\omega\models\Omega,\ \boldsymbol{n}=\boldsymbol{N}(\Phi,\omega)} p_{\Phi,\Omega}(\omega) \tag{13}$$

where $\boldsymbol{N}(\Phi,\omega) = (n_1,\dots,n_m)$*, and* n_i *is the number of grounding of* α_i *that are true in* ω*.*

(Kuzelka 2020a) shows that count distributions can be modelled by Markov Logic Networks with complex weights. In the following, we prove that if each α_i is in FO^2, count distributions can be expressed by a $w(\boldsymbol{k},\boldsymbol{h})$.

Theorem 4. *Every count distribution over a set of possible worlds of a formula* Ω *definable in* FO^2 *can be modelled with a weight function on* $(\boldsymbol{k},\boldsymbol{h})$*, by introducing* m *new predicates* P_i *and adding the axioms* $P_i(x) \leftrightarrow \alpha_i(x)$ *and* $P_j(x,y) \leftrightarrow \alpha_j(x,y)$*, if* α_i *and* α_j *has one and two free variables respectively, and by defining:*

$$q_\Phi(\Omega,\boldsymbol{n}) = \frac{1}{Z} \sum_{(\boldsymbol{k},\boldsymbol{h})(P_i)=n_i} w(\boldsymbol{k},\boldsymbol{h}) \cdot F(\boldsymbol{k},\boldsymbol{h},\Omega) \tag{14}$$

where $Z = \text{WFOMC}(\Omega, w, n)$ *is the partition function.*

Sketch. The proof is a simple consequence of the fact that all the models agreeing with a count statistic $\boldsymbol{n}$ can be counted using cardinality constraints which agree with $\boldsymbol{n}$. Any such cardinality constraint correspond to a specific set of $(\boldsymbol{k},\boldsymbol{h})$ vectors. Hence, we can express arbitrary probability distributions over count statistics by picking real valued weights for $(\boldsymbol{k},\boldsymbol{h})$ vector. We defer the full proof to appendix. □

Example 7. *In the example proposed in (Kuzelka 2020a), they model the distribution of a sequence of 4 coin tosses such that the probability of getting odd number of heads is zero, and the probability of getting even number of heads is uniformly distributed. In order to model this distribution, we introduce a predicate* $H(x)$ *over a domain of* 4 *elements, we also define* Ω *as* $\top$*. This means that every model of this theory is a model of* Ω*. Notice that this distribution cannot be expressed using symmetric weights, as symmetric weights can only express binomial distribution for this language. But we can define weight function on* $(\boldsymbol{k},\boldsymbol{h})$ *vector. In this case* $\boldsymbol{k} = (k_0, k_1)$ *such that* $k_0 + k_1 = 4$*. Since there are no binary predicates we can ignore* $\boldsymbol{h}$*. Intuitively,* k_0 *is the number of tosses which are not heads and* k_1 *is the number of tosses which are heads. If we define the weight function as* $w(k_0,k_1) = 1 + (-1)^{k_1}$ *by applying* (14) *we obtain the following probabilities:*

$$q(\Omega,(4,0)) = \frac{\binom{4}{4}\cdot(1+1)}{16} = \frac{1}{8}$$
$$q(\Omega,(3,1)) = \frac{\binom{4}{3}\cdot(1-1)}{16} = 0$$
$$q(\Omega,(2,2)) = \frac{\binom{4}{2}\cdot(1+1)}{16} = \frac{3}{4}$$
$$q(\Omega,(1,3)) = \frac{\binom{4}{1}\cdot(1-1)}{16} = 0$$
$$q(\Omega,(0,4)) = \frac{\binom{4}{0}\cdot(1+1)}{16} = \frac{1}{8}$$

which coincides with the distribution obtained by (Kuzelka 2020a). Notice, that such a distribution cannot be expressed through symmetric weight functions and obligates the use of a strictly more expressive class of weight functions. In this example we are able to obtain this using real valued weight functions, whereas (Kuzelka 2020a) relies on complex valued weight functions.

We are able to capture count distributions without loosing domain liftability. Furthermore, we do not introduce complex or even negative weights, making the relation between weight functions and probability rather intuitive.

Conclusion

In this paper, we have presented a closed-form formula for FOMC of universally quantified formulas in FO^2 that can be computed in polynomial time w.r.t. domain cardinality. From this, we are able to derive a closed-form expression for FOMC in FO^2 formulas in Scott's Normal Form, extended with cardinality constraints and counting quantifiers. These extended formulas are also computable in polynomial time, and therefore they constitute lifted inference algorithms for C^2. All the formulas are extended to cope with weighted model counting in a simple way, admitting a larger class of weight functions than symmetric weight functions. All the results have been obtained using combinatorial principles, providing a uniform treatment to all these fragments.

References

Beame, P.; Van den Broeck, G.; Gribkoff, E.; and Suciu, D. 2015. Symmetric Weighted First-Order Model Counting. In *Proceedings of the 34th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS '15, 313–328. New York, NY, USA: Association for Computing Machinery. ISBN 9781450327572.

Braz, R. D. S.; Amir, E.; and Roth, D. 2005. Lifted First-Order Probabilistic Inference. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, 1319–1325. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Chavira, M.; and Darwiche, A. 2008. On Probabilistic Inference by Weighted Model Counting. *Artif. Intell.*, 172(6–7): 772–799.

Getoor, L.; and Taskar, B. 2007. *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*. The MIT Press. ISBN 0262072882.

Kazemi, S. M.; Kimmig, A.; Van den Broeck, G.; and Poole, D. 2016. New Liftable Classes for First-Order Probabilistic Inference. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, 3125–3133. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781510838819.

Kieronski, E.; and Kuusisto, A. 2015. Uniform one-dimensional fragments with one equivalence relation. In *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Kuusisto, A.; and Lutz, C. 2018. Weighted Model Counting beyond Two-Variable Logic. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '18, 619–628. New York, NY, USA: Association for Computing Machinery. ISBN 9781450355834.

Kuzelka, O. 2020a. Complex Markov Logic Networks: Expressivity and Liftability. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124. PMLR.

Kuzelka, O. 2020b. Weighted First-Order Model Counting in the Two-Variable Fragment With Counting Quantifiers. arXiv:2007.05619.

Poole, D. 2003. First-Order Probabilistic Inference. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, IJCAI'03, 985–991. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Raedt, D.; Kersting; Natarajan; and Poole. 2016. *Statistical Relational Artificial Intelligence: Logic, Probability, and Computation*. Springer.

Rosen, E.; Graedel, E.; and Otto, M. 1997. Two-variable logic with counting is decidable. In *Logic in Computer Science, Symposium on*, 306. Los Alamitos, CA, USA: IEEE Computer Society.

Scott, D. S. 1962. A decision method for validity of sentences in two variables. *Journal of Symbolic Logic*, 27: 377.

Van den Broeck, G. 2011. On the Completeness of First-Order Knowledge Compilation for Lifted Probabilistic Inference. In *Proceedings of the 24th International Conference on Neural Information Processing Systems*, NIPS'11, 1386–1394. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781618395993.

Van den Broeck, G.; Meert, W.; and Darwiche, A. 2014. Skolemization for Weighted First-Order Model Counting. In *Proceedings of the Fourteenth International Conference on Principles of Knowledge Representation and Reasoning*, KR'14, 111–120. AAAI Press. ISBN 1577356578.

Van den Broeck, G.; Taghipour, N.; Meert, W.; Davis, J.; and De Raedt, L. 2011. Lifted Probabilistic Inference by First-Order Knowledge Compilation. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Three*, IJCAI'11, 2178–2185. AAAI Press. ISBN 9781577355151.

Wilf, H. S. 2006. *Generatingfunctionology*. USA: A. K. Peters, Ltd. ISBN 1568812795.

APPENDIX

FOMC for Universal Formulas

Proof of Lemma 1. For any $\boldsymbol{x'} \in X^m$, we have that $\forall \boldsymbol{x} \Phi(\boldsymbol{x}) \rightarrow \forall \boldsymbol{x} \Phi(\boldsymbol{x'})$ is valid. Which implies that $\forall \boldsymbol{x} \Phi(\boldsymbol{x}) \rightarrow \bigwedge_{\boldsymbol{x'} \in X^m} \forall \boldsymbol{x} \Phi(\boldsymbol{x'})$ is also valid. Since $\forall$ and $\wedge$ commute, we have that $\forall \boldsymbol{x}.\Phi(\boldsymbol{x}) \rightarrow \forall \boldsymbol{x}.\Phi(X)$. The viceversa is obvious since $\Phi(\boldsymbol{x})$ is one of the conjuncts in $\Phi(X)$. ☐

Proposition 1. *For every pure universal formula $\Phi(\boldsymbol{x})$, every permutation π of X and every lifted interpretation τ for $\Phi(X)$, $\tau(\Phi(X)) = \tau_\pi(\Phi(X))$; where $\tau_\pi(P(x_i, x_j, \dots) = \tau(P(\pi(x_i), \pi(x_j), \dots)$, for every atom $P(x, y, \dots)$.*

Proof. If $\tau(\Phi(X)) = 0$ then $\tau(\Phi(\boldsymbol{x'})) = 0$ for some $\boldsymbol{x'} \in X^m$. This implies that $\tau_\pi(\Phi(\pi^{-1}(\boldsymbol{x'}))) = 0$, which implies that $\tau_\pi(\Phi(X)) = 0$. The proof of the opposite direction follows form the fact that $(\tau_\pi)_{\pi^{-1}} = \tau$. ☐

In order to prove Theorem 1, we first introduce the following notation and we also introduce Lemma 4 and Lemma 5. For any set of constants C and any 2^u-tuple $\boldsymbol{k} = (k_0, \dots, k_{2^u-1})$ such that $\sum \boldsymbol{k} = |C|$, let $\mathbb{C}_{\boldsymbol{k}}$ be any partition $(C_i)_{i=0}^{2^u-1}$ of C such that $|C_i| = k_i$. We define $\Phi(\mathbb{C}_{\boldsymbol{k}})$ as follows:

$$\Phi(\mathbb{C}_{\boldsymbol{k}}) = \Phi(C) \wedge \bigwedge_{i=0}^{2^u-1} \bigwedge_{c \in C_i} i(c) \tag{15}$$

Example 8. *Examples of $\mathbb{C}_{(1,0,2,0)}$, on $C = \{a, b, c\}$ are $\{\{a\}, \emptyset, \{b, c\}, \emptyset\}$ and $\{\{b\}, \emptyset, \{a, c\}, \emptyset\}$.*

$$\begin{aligned} \Phi(\{\{a\}, \emptyset, \{b, c\}, \emptyset\}) = \Phi(C) &\wedge \neg A(a) \wedge \neg R(a, a) \\ &\wedge A(b) \wedge \neg R(b, b) \\ &\wedge A(c) \wedge \neg R(c, c) \end{aligned}$$

Note there are $\binom{3}{1,0,2,0} = 3$ *such partitions, and all the* $\Phi(\mathbb{C}_{\boldsymbol{k}})$ *for such partitions will have the same model count. These observations have been formalized in lemma 4*

Lemma 4. $\mathrm{MC}(\Phi(C)) = \sum_{\boldsymbol{k}} \binom{n}{\boldsymbol{k}} \mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}}))$*, where* $\mathrm{MC}(\alpha)$ *denotes the model count of an arbitrary propositional formula* α.

Proof. Let $\mathbb{C}_{\boldsymbol{k}}$ and $\mathbb{C}'_{\boldsymbol{k}}$, be two partitions with the same $\boldsymbol{k}$. Notice that $\mathbb{C}'_{\boldsymbol{k}}$ can be obtained by applying some permutation on C from $\mathbb{C}_{\boldsymbol{k}}$. From Proposition 1 we have that

$$\mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}})) = \mathrm{MC}(\Phi(\mathbb{C}'_{\boldsymbol{k}}))$$

Furthermore notice that if $\mathbb{C}_{\boldsymbol{k}}$ is different from $\mathbb{C}'_{\boldsymbol{k}'}$ then $\Phi(\mathbb{C}_{\boldsymbol{k}})$ and $\Phi(\mathbb{C}'_{\boldsymbol{k}'})$ cannot be simultaneously satisfied. This implies that

$$\mathrm{MC}(\Phi(C)) = \sum_{\boldsymbol{k}} \sum_{\mathbb{C}_{\boldsymbol{k}}} \mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}}))$$

Since there are $\binom{n}{\boldsymbol{k}}$ partitions of C, of the form $\mathbb{C}_{\boldsymbol{k}}$, then

$$\mathrm{MC}(\Phi(C)) = \sum_{\boldsymbol{k}} \binom{n}{\boldsymbol{k}} \mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}}))$$

□

Lemma 5. *For any partition* $\mathbb{C}_{\boldsymbol{k}} = \{C_0, \dots, C_{2^u-1})$

$$\mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}})) = \prod_{\substack{c \neq d \\ c,d \in C}} n_{i_c i_d}$$

where for all $c, d \in C$, $0 \leq i_c, i_d \leq 2^u - 1$ *are the indices such that* $c \in C_{i_c}$ *and* $d \in C_{i_d}$.

Proof. $\Phi(\mathbb{C}_{\boldsymbol{k}})$ can be rewritten in

$$\bigwedge_{\substack{\{c,d\} \subseteq C \\ c \neq d}} \Phi^{i_c, i_d}(\{c, d\})$$

$\Phi^{i_c,i_d}(\{c,d\}))$ is obtained by replacing each atom $P_j(c)$ with $\top$ if $bin(i_c)_j = 1$ and $\bot$ otherwise and each atom $P_j(d)$ with $\top$ if $bin(i_d)_j = 1$ and $\bot$ otherwise. Notice that all the atoms of $\Phi^{i_c,i_d}(\{c,d\})$ contain both c and d. Furthermore notice that if $\{c,d\} \neq \{e,f\}$ then $\Phi^{i_c,i_d}(\{c,d\}))$ and $\Phi^{i_e,i_f}(\{e,f\}))$ do not contain common atoms. Finally we have that $\mathrm{MC}(\Phi^{i_c,i_d}(\{c,d\})) = n_{i_c i_d}$. Hence

$$\mathrm{MC}\left(\bigwedge_{\substack{c,d \in C \\ c \neq d}} \Phi^{i_c,i_d}(\{c,d\}) \right) = \prod_{\substack{c \neq d \\ c,d \in C}} n_{i_c i_d}$$

□

Finally, we provide the following proof for Theorem 1.

Proof of Theorem 1. Notice that $\mathrm{FOMC}(\Phi(\boldsymbol{x}), n) = \mathrm{MC}(\Phi(C))$ for a set of constants C with $|C| = n$. Therefore, by Lemma 4, to prove the theorem it is enough to show that for all $\boldsymbol{k}$, $\mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}})) = \prod_{0 \leq i \leq j \leq 2^u-1} n_{ij}^{\boldsymbol{k}(i,j)}$. By the Lemma 5 we have that $\mathrm{MC}(\Phi(\mathbb{C}_{\boldsymbol{k}})) = \prod_{c \neq d} n_{i_c i_d}$. Then:

$$\prod_{c \neq d} n_{i_c i_d} = \prod_i \prod_{\substack{c \neq d \\ c,d \in C_i}} n_{ii} \cdot \prod_{i<j} \prod_{\substack{c \in C_i \\ d \in C_j}} n_{ij}$$

$$= \prod_i n_{ii}^{\binom{k_i}{2}} \cdot \prod_{i<j} n_{ij}^{k_i k_j} = \prod_{0 \leq i \leq j < 2^u} n_{ij}^{\boldsymbol{k}(i,j)}$$

□

As a final remark for this section, notice that the computational cost of computing n_{ij} is constant with respect to the domain cardinality. We assume the cost of multiplication to be constant. Hence, the computational complexity of computing (4) depends on the domain only through the multinomial coefficients $\binom{n}{\boldsymbol{k}}$ and the multiplications involved in $\prod_{ij} n^{\boldsymbol{k}(i,j)}$. The computational cost of computing $\binom{n}{\boldsymbol{k}}$ is polynomial in n and the total number of $\binom{n}{\boldsymbol{k}}$ are $\binom{n+2^u-1}{2^u-1}$, which has $\left(\frac{e \cdot (n+2^u-1)}{2^u-1}\right)^{2^u-1}$ as an upper-bound. Also, the $\prod_{ij} n^{\boldsymbol{k}(i,j)}$ term has $O(n^2)$ multiplication operations. Hence, we can conclude that the (4) is computable in polynomial time with respect to the domain cardinality.

FOMC for Cardinality Constraints

In the following, we provide some examples to better explain $(\boldsymbol{k}, \boldsymbol{h})$ vectors.

Example 9. *To count the models of* (3) *with the additional constraint that* A *is balanced i.e.,* $\frac{n}{2} \leq |A| \leq \frac{n+1}{2}$, *we have to consider only the terms where* $\boldsymbol{k}$ *is such* $\frac{n}{2} \leq \boldsymbol{k}(A) \leq \frac{n+1}{2}$. *Equivalently, we should consider only the* $\boldsymbol{k}$ *such that* $\frac{n}{2} \leq k_2 + k_3 \leq \frac{n+1}{2}$. *(Notice that* k_2 *is the number of elements that satisfy* $A(x)$ *and* $\neg R(x,x)$ *and* k_3 *is the number of elements that satisfy* $A(x)$ *and* $R(x,x)$*).*

Example 10. *A graphical representation of the pair* $\boldsymbol{k}, \boldsymbol{h}$ *for the formula* (3) *is provided in the following picture:*

	k_0	k_1	k_2	k_3
k_0	$h_0^{00}\ h_1^{00}$ $h_2^{00}\ h_3^{00}$	$h_0^{01}\ h_1^{01}$ $h_2^{01}\ h_3^{01}$	$h_0^{02}\ h_1^{02}$ $h_2^{02}\ h_3^{02}$	$h_0^{03}\ h_1^{03}$ $h_2^{03}\ h_3^{03}$
k_1		$h_0^{11}\ h_1^{11}$ $h_2^{11}\ h_3^{11}$	$h_0^{12}\ h_1^{12}$ $h_2^{12}\ h_3^{12}$	$h_0^{13}\ h_1^{13}$ $h_2^{13}\ h_3^{13}$
k_2			$h_0^{22}\ h_1^{22}$ $h_2^{22}\ h_3^{22}$	$h_0^{23}\ h_1^{23}$ $h_2^{23}\ h_3^{23}$
k_3				$h_0^{33}\ h_1^{33}$ $h_2^{33}\ h_3^{33}$

This configuration represent the models in which a set C *of* n *constants are partitioned in four sets* $C_0, \dots, C_3$, *each* C_i *containing* k_i *elements (hence* $\sum k_i = n$*). Furthermore, for each pair* C_i *and* C_j *the relation* $D^{ij} = C_i \times C_j$ *is partitioned in 4 sub relations* $D_0^{ij}, \dots, D_3^{ij}$ *where each* D_v^{ij} *contains* h_v^{ij} *pairs (hence* $\sum_v h_v^{ij} = \boldsymbol{k}(i,j)$*). For instance if the pair* $(c,d) \in D_2^{12}$ *it means that we are considering* *assignments that satisfy* $\neg A(c) \wedge R(c,c) \wedge A(d) \wedge \neg R(d,d) \wedge R(c,d) \wedge \neg R(d,c)$.

Weighted First Order Model Counting

All the FOMC formulas introduced in the paper can be extended to WFOMC by defining a real-valued function on $(\boldsymbol{k}, \boldsymbol{h})$ and adding it as a multiplicative factor, for instance if Φ is in SNF, as given in equation (7), then it's WFOMC can be defined as follows:

$$\text{WFOMC}(\Phi, n) = \sum_{\boldsymbol{k},\boldsymbol{h}} (-1)^{\sum_i \boldsymbol{k}(P_i)} w(\boldsymbol{k}, \boldsymbol{h}) F(\boldsymbol{k}, \boldsymbol{h}, \Phi')$$

where $w(\boldsymbol{k}, \boldsymbol{h})$ is a real-valued weight function and Φ' is the transformed formula as described in Theorem 2.

Symmetric Weight Functions

Theorem 5. *For all* Φ *in* C^2 *and for arbitrary cardinality constraint* ρ*, symmetric-WFOMC can be obtained from FOMC by defining the following weight function:*

$$w(\boldsymbol{k}, \boldsymbol{h}) = \prod_{P \in \mathcal{L}} w(P)^{(\boldsymbol{k},\boldsymbol{h})(P)} \cdot \bar{w}(P)^{(\boldsymbol{k},\boldsymbol{h})(\neg P)}$$

where $w(P)$ *and* $\bar{w}(P)$ *are real valued weights on predicate* P *and it's negation respectively.*

Proof. The proof is a consequence of the observation that $F(\boldsymbol{k}, \boldsymbol{h}, \Phi)$ is the number of models of Φ that contain $\boldsymbol{k}(P)$ elements that satisfy P if P is unary, and $(\boldsymbol{k}, \boldsymbol{h})(P)$ pairs of elements that satisfy P, if P is binary. □

Expressing Count Distributions

In the following we provide the proof for Theorem 4.

Proof of Theorem 4. Since Ω is a FO^2 formula, then we can compute FOMC as follows:

$$\text{FOMC}(\Omega, n) = \sum_{\boldsymbol{k},\boldsymbol{h}} F(\boldsymbol{k}, \boldsymbol{h}, \Omega)$$

Let us define $w(\boldsymbol{k}, \boldsymbol{h})$ for each $\boldsymbol{k}, \boldsymbol{h}$ as follows:

$$w(\boldsymbol{k}, \boldsymbol{h}) = \frac{1}{F(\boldsymbol{k}, \boldsymbol{h}, \Omega)} \sum_{\substack{\omega \models \Omega \\ N(\alpha_1,\omega)_1 = (\boldsymbol{k},\boldsymbol{h})(P_1) \\ \cdots \\ N(\alpha_m,\omega)_m = (\boldsymbol{k},\boldsymbol{h})(P_m)}} p_{\Phi,\Omega}(\omega)$$

This definition implies that the partition function Z is equal to 1. Indeed:

$$\begin{aligned}
Z &= \text{WFOMC}(\Omega, w, n) \\
&= \sum_{\boldsymbol{k},\boldsymbol{h}} w(\boldsymbol{k}, \boldsymbol{h}) \cdot F(\boldsymbol{k}, \boldsymbol{h}, \Omega) \\
&= \sum_{\boldsymbol{k},\boldsymbol{h}} \sum_{\substack{\omega \models \Omega \\ N(\alpha_1,\omega)_1 = (\boldsymbol{k},\boldsymbol{h})(P_1) \\ \cdots \\ N(\alpha_m,\omega)_m = (\boldsymbol{k},\boldsymbol{h})(P_m)}} p_{\Phi,\Omega}(\omega) \\
&= \sum_{\omega \models \Omega} \sum_{\substack{\boldsymbol{k},\boldsymbol{h} \\ N(\alpha_1,\omega)_1 = (\boldsymbol{k},\boldsymbol{h})(P_1) \\ \cdots \\ N(\alpha_m,\omega)_m = (\boldsymbol{k},\boldsymbol{h})(P_m)}} p_{\Phi,\Omega}(\omega) \\
&= \sum_{\omega \models \Omega} p_{\Phi,\Omega}(\omega) \\
&= 1
\end{aligned}$$

Hence,

$$\begin{aligned}
q_\Phi(\Omega, \boldsymbol{n}) &= \sum_{(\boldsymbol{k},\boldsymbol{h})(P_i) = n_i} F(\boldsymbol{k}, \boldsymbol{h}, \Omega) \cdot w(\boldsymbol{k}, \boldsymbol{h}) \\
&= \sum_{(\boldsymbol{k},\boldsymbol{h})(P_i) = n_i} \sum_{\substack{\omega \models \Omega \\ N(\alpha_1,\omega)_1 = (\boldsymbol{k},\boldsymbol{h})(P_1) \\ \cdots \\ N(\alpha_m,\omega)_m = (\boldsymbol{k},\boldsymbol{h})(P_m)}} p_{\Phi,\Omega}(\omega) \\
&= \sum_{\substack{\omega \models \Omega \\ N(\alpha_1,\omega)_1 = n_1 \\ \cdots \\ N(\alpha_m,\omega)_m = n_m}} p_{\Phi,\Omega}(\omega)
\end{aligned}$$

□