

# Towards Practical First-Order Model Counting

Ananth K. Kidambi ✉

Indian Institute of Technology Bombay, Mumbai, India

Guramrit Singh ✉

Indian Institute of Technology Bombay, Mumbai, India

Paulius Dilkas ✉ 🏠 

University of Toronto, Toronto, Canada

Vector Institute, Toronto, Canada

Kuldeep S. Meel ✉ 🏠 

University of Toronto, Toronto, Canada

---

## Abstract

First-order model counting (FOMC) is the problem of counting the number of models of a sentence in first-order logic. Since lifted inference techniques rely on reductions to variants of FOMC, the design of scalable methods for FOMC has attracted attention from both theoreticians and practitioners over the past decade. Recently, a new approach based on first-order knowledge compilation was proposed. This approach, called CRANE, instead of simply providing the final count, generates definitions of (possibly recursive) functions that can be evaluated with different arguments to compute the model count for any domain size. However, this approach is not fully automated, as it requires manual evaluation of the constructed functions. The primary contribution of this work is a fully automated compilation algorithm, called GANTRY, which transforms the function definitions into C++ code equipped with arbitrary-precision arithmetic. These additions allow the new FOMC algorithm to scale to domain sizes over 500,000 times larger than the current state of the art, as demonstrated through experimental results.

**2012 ACM Subject Classification** Theory of computation → Automated reasoning; Theory of computation → Logic and verification; Mathematics of computing → Combinatorics

**Keywords and phrases** first-order model counting, knowledge compilation, lifted inference

**Funding** *Ananth K. Kidambi*: TODO: funding

*Guramrit Singh*: TODO: funding

*Paulius Dilkas*: TODO: funding

*Kuldeep S. Meel*: TODO: funding

**Acknowledgements** Part of the research was conducted while all authors were at the National University of Singapore.

## 1 Introduction

*First-order model counting* (FOMC) is the task of determining the number of models for a sentence in first-order logic over a specified domain. The weighted variant, WFOMC, computes the total weight of these models, linking logical reasoning with probabilistic frameworks [32]. It builds upon earlier efforts in weighted model counting for propositional logic [3] and broader attempts to bridge logic and probability [15, 17, 20]. WFOMC is central to *lifted inference*, which enhances the efficiency of probabilistic calculations by exploiting symmetries [12]. Lifted inference continues to advance, with applications extending to constraint satisfaction problems [24] and probabilistic answer set programming [1]. Moreover, WFOMC has proven effective at reasoning over probabilistic databases [7] and probabilistic logic programs [18]. FOMC algorithms have also facilitated breakthroughs in discovering integer sequences [22] and developing recurrence relations for these sequences [5]. Recently,

these algorithms have been extended to perform sampling tasks [33].

The complexity of FOMC is generally measured by *data complexity*, with a formula classified as *liftable* if it can be solved in polynomial time relative to the domain size [10]. While all formulas with up to two variables are known to be liftable [29, 31], Beame et al. [2] demonstrated that liftability does not extend to all formulas, identifying an unliftable formula with three variables. Recent work has further extended the liftable fragment with additional axioms [23, 28] and counting quantifiers [13], expanding our understanding of liftability.

FOMC algorithms are diverse, with approaches ranging from *first-order knowledge compilation* (FOKC) to local search [16], Monte Carlo sampling [6], and anytime approximation [26]. Among these, FOKC-based algorithms are particularly prominent, transforming formulas into structured representations such as circuits or graphs. Notable examples include FORCLIFT [32] and its successor CRANE [5]. Another important algorithm, FASTWFOMC [27], uses cell enumeration as its foundation.

The CRANE algorithm marked a significant step forward, expanding the range of formulas handled by FOMC algorithms. However, it had notable limitations: it required manual evaluation of function definitions to compute model counts and introduced recursive functions without proper base cases, making it more complex to use. To address these shortcomings, we present GANTRY, a fully automated FOMC algorithm that overcomes the constraints of its predecessor. GANTRY can handle domain sizes over 500,000 times larger than previous algorithms and simplifies the user experience by automatically handling base cases and compiling function definitions into efficient C++ programs.

In Section 2, we cover some preliminaries, and in Section 3, we detail all our technical contributions. Finally, in Section 4, we present our experimental results, demonstrating GANTRY’s performance compared to other FOMC algorithms, and, in Section 5, we conclude the paper by discussing promising avenues for future work.

## 2 Preliminaries

In Section 2.1, we summarise the basic principles of first-order logic. Then, in Section 2.2, we formally define (W)FOMC and discuss the distinctions between three variations of first-order logic used for FOMC. Finally, in Section 2.3, we introduce the terminology used to describe the output of the original CRANE algorithm, i.e., functions and equations that define them.

We use  $\mathbb{N}_0$  to represent the set of non-negative integers. In both algebra and logic, we write  $S\sigma$  to denote the application of a *substitution*  $\sigma$  to an expression  $S$ , where  $\sigma = [x_1 \mapsto y_1, x_2 \mapsto y_2, \dots, x_n \mapsto y_n]$  signifies the replacement of all instances of  $x_i$  with  $y_i$  for all  $i = 1, \dots, n$ .

### 2.1 First-Order Logic

In this section, we will review the basic concepts of first-order logic as they are used in FOKC algorithms. We begin by introducing the format used internally by FORCLIFT and its descendants. Afterwards, we provide a high-level description of how an arbitrary sentence in first-order logic is transformed into this internal format.

A *term* can be either a variable or a constant. An *atom* can be either  $P(t_1, \dots, t_m)$  (i.e.,  $P(\mathbf{t})$ ) for some predicate  $P$  and terms  $t_1, \dots, t_m$  or  $x = y$  for some terms  $x$  and  $y$ . The *arity* of a predicate is the number of arguments it takes, i.e.,  $m$  in the case of the predicate  $P$  mentioned above. We write  $P/m$  to denote a predicate along with its arity. A *literal* can be either an atom (i.e., a *positive* literal) or its negation (i.e., a *negative* literal). An atom is *ground* if it contains no variables, i.e., only constants. A *clause* is of the form

Logic	Sorts	Constants	Variables	Quantifiers	Additional atoms
FO	one or more	✓	unlimited	$\forall, \exists$	$x = y$
$C^2$	one	✗	two	$\forall, \exists, \exists^{=k}, \exists^{\leq k}, \exists^{\geq k}$	—
$UFO^2 + CC$	one	✗	two	$\forall$	$ P  = m$

■ **Table 1** A comparison of the three logics used in FOMC. The 2<sup>nd</sup>–5<sup>th</sup> columns refer to: the number of sorts, support for constants, the maximum number of variables, and supported quantifiers, respectively. The last column lists supported atoms in addition to those of the form  $P(\mathbf{t})$  for a predicate  $P/n$  and an  $n$ -tuple of terms  $\mathbf{t}$ . Here:  $k$  and  $m$  are non-negative integers, with the latter depending on the domain size,  $P$  represents a predicate, and  $x$  and  $y$  are terms.

$\forall x_1 \in \Delta_1. \forall x_2 \in \Delta_2 \dots \forall x_n \in \Delta_n. \phi(x_1, x_2, \dots, x_n)$ , where  $\phi$  is a disjunction of literals that only contain variables  $x_1, \dots, x_n$  (and any constants). We say that a clause is a (*positive*) *unit clause* if there is only one literal with a predicate, and it is a positive literal. Finally, a *formula* is a conjunction of clauses. Throughout the paper, we will use set-theoretic notation, interpreting a formula as a set of clauses and a clause as a set of literals.

► **Remark.** Conforming with previous work [32], the definition of a clause includes universal quantifiers for all variables within. While it is possible to rewrite the entire formula with all quantifiers at the front [8], the format we describe has proven itself convenient to work with.

There are two crucial preprocessing steps that transform an arbitrary sentence in first-order logic into the form used internally: Skolemization and rewriting the sentence as a conjunction of clauses. We describe the former in more detail. *Skolemization* [31] is a procedure that transforms a formula with existential quantifiers into a formula without existential quantifiers *with the same WFOMC*. (Note that it is different from the standard Skolemization that introduces function symbols [9].)

► **Example 1.** Skolemization transforms

$$\forall x \in \Gamma. \exists y \in \Delta. P(x, y) \quad (1)$$

into

$$\begin{aligned} & (\forall x \in \Gamma. Z(x)) \wedge \\ & (\forall x \in \Gamma. \forall y \in \Delta. Z(x) \vee \neg P(x, y)) \wedge \\ & (\forall x \in \Gamma. S(x) \vee Z(x)) \wedge \\ & (\forall x \in \Gamma. \forall y \in \Delta. S(x) \vee \neg P(x, y)). \end{aligned} \quad (2)$$

We will see how, with suitable weights on the new predicates  $S/1$  and  $Z/1$ , the WFOMC of Formula (2) is equal to the FOMC of Formula (1).

## 2.2 FOMC Algorithms and Their Logics

In Table 1, we outline the differences among three first-order logics commonly used in FOMC: FO,  $C^2$ , and  $UFO^2 + CC$ . First, FO is the input format for FORCLIFT\* and its extensions CRANE† and GANTRY. Second,  $C^2$  is often used in the literature on FASTWFOMC and

\* <https://github.com/UCLA-StarAI/Forclift>

† <https://doi.org/10.5281/zenodo.8004077>

related methods [13, 14]. Finally,  $\text{UFO}^2 + \text{CC}$  is the input format supported by the most recent implementation of  $\text{FastWFOMC}^\ddagger$ . The notation we use to refer to each logic is standard in the case of  $\text{C}^2$  and  $\text{UFO}^2 + \text{CC}$  [25] and redefined to be more specific in the case of  $\text{FO}$ . All three logics are function-free, and domains are always assumed to be finite. As usual, we presuppose the *unique name assumption*, which states that two constants are equal if and only if they are the same constant [19].

In  $\text{FO}$ , each term is assigned to a *sort*, and each predicate  $P/n$  is assigned to a sequence of  $n$  sorts. Each sort has its corresponding domain. These assignments to sorts are typically left implicit and can be reconstructed from the quantifiers. For example,  $\forall x, y \in \Delta. P(x, y)$  implies that variables  $x$  and  $y$  have the same sort. On the other hand,  $\forall x \in \Delta. \forall y \in \Gamma. P(x, y)$  implies that  $x$  and  $y$  have different sorts, and it would be improper to write, for example,  $\forall x \in \Delta. \forall y \in \Gamma. P(x, y) \vee x = y$ .  $\text{FO}$  is also the only logic to support constants, formulas with more than two variables, and the equality predicate. While we do not explicitly refer to sorts in subsequent sections of this paper, the many-sorted nature of  $\text{FO}$  is paramount to the algorithms presented therein.

► **Remark.** In the case of  $\text{FORCLIFT}$  and its extensions, support for a formula as valid input does not imply that the algorithm can compile the formula into a circuit or graph suitable for lifted model counting. However, it is known that  $\text{FORCLIFT}$  compilation is guaranteed to succeed on any  $\text{FO}$  formula without constants and with at most two variables [29, 31].

Compared to  $\text{FO}$ ,  $\text{C}^2$  and  $\text{UFO}^2 + \text{CC}$  lack support for constants, the equality predicate, multiple domains, and formulas with more than two variables. The advantage that  $\text{C}^2$  brings over  $\text{FO}$  is the inclusion of *counting quantifiers*. That is, alongside  $\forall$  and  $\exists$ ,  $\text{C}^2$  supports  $\exists^{=k}$ ,  $\exists^{\leq k}$ , and  $\exists^{\geq k}$  for any positive integer  $k$ . For example,  $\exists^{=1}x. \phi(x)$  means that there exists *exactly one*  $x$  such that  $\phi(x)$ , and  $\exists^{\leq 2}x. \phi(x)$  means that there exist *at most two* such  $x$ .  $\text{UFO}^2 + \text{CC}$ , on the other hand, does not support any existential quantifiers but instead incorporates (*equality*) *cardinality constraints*. For example,  $|P| = 3$  constrains all models to have *precisely three positive literals with the predicate  $P$* .

► **Definition 2 (Model).** Let  $\phi$  be a formula in  $\text{FO}$ . For each predicate  $P/n$  in  $\phi$ , let  $(\Delta_i^P)_{i=1}^n$  be a list of the corresponding domains. Let  $\sigma$  be a map from the domains of  $\phi$  to their interpretations as sets such that the sets are pairwise disjoint, and the constants in  $\phi$  are included in the corresponding domains. A structure of  $\phi$  is a set  $M$  of ground literals defined by adding to  $M$  either  $P(\mathbf{t})$  or  $\neg P(\mathbf{t})$  for every predicate  $P/n$  in  $\phi$  and  $n$ -tuple  $\mathbf{t} \in \prod_{i=1}^n \sigma(\Delta_i^P)$ . A structure is a model if it satisfies  $\phi$ .

► **Remark.** The distinctness of domains is important in two ways. First, in terms of expressiveness, a clause such as  $\forall x \in \Delta. P(x, x)$  is valid if predicate  $P$  is defined over two copies of the same domain and invalid otherwise. Second, having more distinct domains makes the problem more decomposable for the  $\text{FOKC}$  algorithm. With distinct domains, the algorithm can make assumptions or deductions about, e.g., the first domain of predicate  $P$  without worrying how (or if) they apply to the second domain.

While this work focuses on  $\text{FOMC}$ , we still define the weighted variant of the problem as  $\text{Skolemization}$  relies on weights even for unweighted  $\text{FOMC}$ .

► **Definition 3 (WFOMC instance).** A  $\text{WFOMC}$  instance comprises: a formula  $\phi$  in  $\text{FO}$ , two (rational) weights  $w^+(P)$  and  $w^-(P)$  assigned to each predicate  $P$  in  $\phi$ , and  $\sigma$  as described in Definition 2. Unless specified otherwise, we assume all weights to be equal to 1.

---

<sup>‡</sup> <https://github.com/jan-toth/FastWFOMC.jl>

158 ► **Definition 4** (WFOMC [32]). *Given a WFOMC instance  $(\phi, w^+, w^-, \sigma)$  as in Definition 3,*  
 159 *the (symmetric) weighted first-order model count (WFOMC) of  $\phi$  is*

$$160 \quad \sum_{M \models \phi} \prod_{P(\mathbf{t}) \in M} w^+(P) \prod_{\neg P(\mathbf{t}) \in M} w^-(P), \quad (3)$$

161 *where the sum is over all models of  $\phi$ .*

162 ► **Example 5** (Counting functions). To define predicate  $P$  as a function from a domain  $\Delta$  to  
 163 itself, in  $\mathcal{C}^2$  one would write  $\forall x \in \Delta. \exists^=1 y \in \Delta. P(x, y)$ . In  $\text{UFO}^2 + \text{CC}$ , the same could be  
 164 written as

$$165 \quad (\forall x, y \in \Delta. S(x) \vee \neg P(x, y)) \wedge (|P| = |\Delta|), \quad (4)$$

166 where  $w^-(S) = -1$ . Although Formula (4) has more models compared to its counterpart in  
 167  $\mathcal{C}^2$ , the negative weight  $w^-(S) = -1$  makes some of the terms in Equation (3) cancel out.

168 Equivalently, in FO we would write

$$169 \quad (\forall x \in \Gamma. \exists y \in \Delta. P(x, y)) \wedge (\forall x \in \Gamma. \forall y, z \in \Delta. P(x, y) \wedge P(x, z) \Rightarrow y = z). \quad (5)$$

170 The first clause asserts that each  $x$  must have at least one corresponding  $y$ , while the second  
 171 statement adds the condition that if  $x$  is mapped to both  $y$  and  $z$ , then  $y$  must equal  $z$ . It is  
 172 important to note that Formula (5) is written with two domains instead of just one. However,  
 173 we can still determine the correct number of functions by assuming that the sizes of  $\Gamma$  and  
 174  $\Delta$  are equal. This formulation, as observed by Dilkas and Belle [5], can prove beneficial in  
 175 enabling FOKC algorithms to find efficient solutions.

## 176 2.3 Algebra

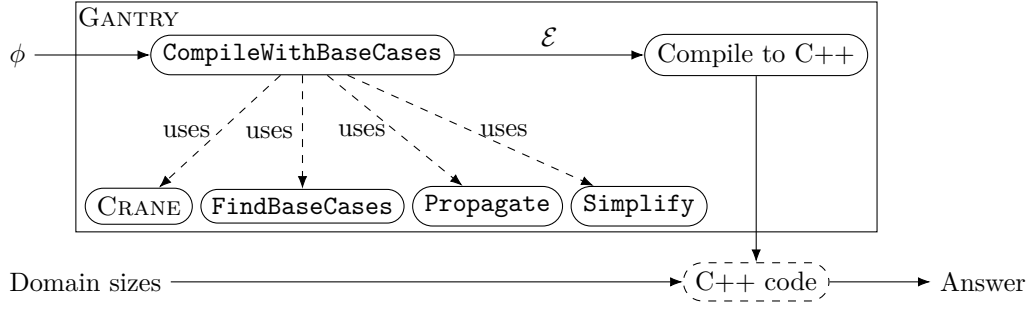
177 We write **expr** for an arbitrary algebraic expression. In the context of algebra, a *constant*  
 178 is a non-negative integer. Likewise, a *variable* can either be a parameter of a function or a  
 179 variable introduced through summation, such as  $i$  in the expression  $\sum_{i=1}^n \mathbf{expr}$ . A *function*  
 180 *call* is  $f(x_1, \dots, x_n)$  (or  $f(\mathbf{x})$  for short), where  $f$  is an  $n$ -ary function, and each  $x_i$  is an  
 181 algebraic expression consisting of variables and constants. A (function) *signature* is function  
 182 call that contains only variables. An *equation* is  $f(\mathbf{x}) = \mathbf{expr}$ , where  $f(\mathbf{x})$  is a signature.

183 ► **Definition 6** (Base case). *Let  $f(\mathbf{x})$  be a function call where each  $x_i$  is either a constant*  
 184 *or a variable. Then function call  $f(\mathbf{y})$  is a base case of  $f(\mathbf{x})$  if  $f(\mathbf{y}) = f(\mathbf{x})\sigma$ , where  $\sigma$  is a*  
 185 *substitution that replaces one or more  $x_i$  with a constant.*

186 ► **Example 7.** In equation  $f(m, n) = f(m - 1, n) + n f(m - 1, n - 1)$ , the only constant is  
 187 1, and the variables are  $m$  and  $n$ . The equation contains three function calls: one on the  
 188 left-hand side, and two on the right-hand side. The function call on the left-hand side is a  
 189 signature. Function calls such as  $f(4, n)$ ,  $f(m, 0)$ , and  $f(8, 1)$  are all considered base cases of  
 190  $f(m, n)$  (only some of which are useful).

## 191 3 Technical Contributions

192 Figure 1 provides an overview of GANTRY’s workflow. Section 3.1 describes the main  
 193 algorithm for completing the definitions of recursive functions with a sufficient set of base  
 194 cases. Sections 3.2 and 3.3 describe subsidiary algorithms for constructing a set of base cases  
 195 and their corresponding logical formulas. Section 3.4 explains the post-processing techniques  
 196 for ensuring accurate model counting. Additionally, Section 3.5 explains the process of  
 197 compiling function definitions into C++ code, greatly expanding upon the range of formulas  
 198 that could previously be handled by similar approaches [11].



■ **Figure 1** The outline of using GANTRY to compute the model count of a formula  $\phi$ . First, the formula is compiled into a set of equations, which are then used to create a C++ program. This program can be executed with different command line arguments to calculate the model count of  $\phi$  for different domain sizes. To accomplish this, the `CompileWithBaseCases` function makes use of the FOKC algorithm of CRANE, algebraic simplification techniques (denoted as `Simplify`), and two other auxiliary procedures.

■ **Algorithm 1** `CompileWithBaseCases( $\phi$ )`

---

**Input:** formula  $\phi$   
**Output:** set  $\mathcal{E}$  of equations

```

1   $(\mathcal{E}, \mathcal{F}, \mathcal{D}) \leftarrow \text{CRANE}(\phi);$ 
2   $\mathcal{E} \leftarrow \text{Simplify}(\mathcal{E});$ 
3  foreach base case  $f(\mathbf{x}) \in \text{FindBaseCases}(\mathcal{E})$  do
4     $\psi \leftarrow \mathcal{F}(f);$ 
5    foreach index  $i$  such that  $x_i \in \mathbb{N}_0$  do  $\psi \leftarrow \text{Propagate}(\psi, \mathcal{D}(f, i), x_i);$ 
6     $\mathcal{E} \leftarrow \mathcal{E} \cup \text{CompileWithBaseCases}(\psi);$ 

```

---

### 199 3.1 Completing the Definitions of Functions

200 Before describing the main contribution of this work, let us review the essential aspects of  
 201 FOKC as realised by CRANE. The input formula is compiled into: set  $\mathcal{E}$  of equations, map  $\mathcal{F}$   
 202 from function names to formulas, and map  $\mathcal{D}$  from function names and argument indices to  
 203 domains.  $\mathcal{E}$  can contain any number of functions, one of which (denoted by  $f$ ) represents the  
 204 solution to the FOMC problem. To compute the FOMC for particular domain sizes,  $f$  must  
 205 be evaluated with those domain sizes as arguments.  $\mathcal{D}$  records this correspondence between  
 206 function arguments and domains.

207 Algorithm 1 presents our overall approach for compiling a formula into equations that  
 208 include the necessary base cases. To begin, we use the FOKC algorithm of the original  
 209 CRANE to compile the formula into the three components:  $\mathcal{E}$ ,  $\mathcal{F}$ , and  $\mathcal{D}$ . After some algebraic  
 210 simplification,  $\mathcal{E}$  is passed to the `FindBaseCases` procedure (see Section 3.2). For each base  
 211 case  $f(\mathbf{x})$ , we retrieve the logical formula  $\mathcal{F}(f)$  associated with the function name  $f$  and  
 212 simplify it using the `Propagate` procedure (explained in detail in Section 3.3). We do this by  
 213 iterating over all indices of  $\mathbf{x}$ , where  $x_i$  is a constant, and using `Propagate` to simplify  $\psi$  by  
 214 assuming that domain  $\mathcal{D}(f, i)$  has size  $x_i$ . Finally, on line 6, `CompileWithBaseCases` recurses  
 215 on these simplified formulas and adds the resulting base case equations to  $\mathcal{E}$ . Example 8  
 216 below provides more detail.

217 ► **Remark.** Although `CompileWithBaseCases` starts with a call to CRANE, the proposed  
 218 algorithm is not just a post-processing step for FOKC because Algorithm 1 is recursive and

219 can issue more calls to CRANE on various derived formulas.

220 ► **Example 8** (Counting bijections). Consider the following formula (previously examined by  
221 Dilkas and Belle [5]) that defines predicate  $P$  as a bijection between two sets  $\Gamma$  and  $\Delta$ :

$$\begin{aligned} & (\forall x \in \Gamma. \exists y \in \Delta. P(x, y)) \wedge \\ & (\forall y \in \Delta. \exists x \in \Gamma. P(x, y)) \wedge \\ 222 & (\forall x \in \Gamma. \forall y, z \in \Delta. P(x, y) \wedge P(x, z) \Rightarrow y = z) \wedge \\ & (\forall x, z \in \Gamma. \forall y \in \Delta. P(x, y) \wedge P(z, y) \Rightarrow x = z). \end{aligned}$$

223 We specifically examine the first solution returned by GANTRY for this formula.

224 After line 2, we have

$$\begin{aligned} 225 \quad \mathcal{E} &= \left\{ \begin{aligned} f(m, n) &= \sum_{l=0}^n \binom{n}{l} (-1)^{n-l} g(l, m), \\ g(l, m) &= g(l-1, m) + mg(l-1, m-1) \end{aligned} \right\}; \\ 226 \quad \mathcal{D} &= \{ (f, 1) \mapsto \Gamma, (f, 2) \mapsto \Delta, (g, 1) \mapsto \Delta^\top, (g, 2) \mapsto \Gamma \}, \end{aligned}$$

227 where  $\Delta^\top$  is a new domain. (We omit the definition of  $\mathcal{F}$  as the formulas can get a bit  
228 verbose.) Then **FindBaseCases** identifies two base cases:  $g(0, m)$  and  $g(l, 0)$ . In both cases,  
229 **CompileWithBaseCases** recurses on the formula  $\mathcal{F}(g)$  simplified by assuming that one of the  
230 domains is empty. In the first case, we recurse on the formula  $\forall x \in \Gamma. S(x) \vee \neg S(x)$ , where  
231  $S$  is a predicate introduced by Skolemization with weights  $w^+(S) = 1$  and  $w^-(S) = -1$ .  
232 Hence, we obtain the base case  $g(0, m) = 0^m$ . In the case of  $g(l, 0)$ , **Propagate**( $\psi, \Gamma, 0$ )  
233 returns an empty formula, resulting in  $g(l, 0) = 1$ .

234 It is worth noting that these base cases overlap when  $l = m = 0$  but remain consistent  
235 since  $0^0 = 1$ . Generally, let  $\phi$  be a formula with two domains  $\Gamma$  and  $\Delta$ , and let  $n, m \in \mathbb{N}_0$ .  
236 Then the FOMC of **Propagate**( $\phi, \Delta, n$ ) assuming  $|\Gamma| = m$  is the same as the FOMC of  
237 **Propagate**( $\phi, \Gamma, m$ ) assuming  $|\Delta| = n$ .

238 Finally, the main responsibility of the **Simplify** procedure is to handle the algebraic  
239 pattern  $\sum_{m=0}^n [a \leq m \leq b] f(m)$ . Here:  $n$  is a variable,  $a, b \in \mathbb{N}_0$  are constants, and  $f$   
240 is an expression that may depend on  $m$ . Additionally,  $[a \leq m \leq b] = \begin{cases} 1 & \text{if } a \leq m \leq b \\ 0 & \text{otherwise} \end{cases}$ .  
241 **Simplify** transforms this pattern into  $f(a) + f(a+1) + \dots + f(\min\{n, b\})$ . For instance,  
242 in the case of Example 8, **Simplify** transforms  $g(l, m) = \sum_{k=0}^m [0 \leq k \leq 1] \binom{m}{k} g(l-1, m-k)$   
243 into  $g(l, m) = g(l-1, m) + mg(l-1, m-1)$ .

## 244 3.2 Identifying a Sufficient Set of Base Cases

245 Algorithm 2 summarises the implementation of **FindBaseCases**. It considers two types of  
246 arguments when a function  $f$  calls itself recursively: constants and arguments of the form  
247  $x_i - c_i$ . Here,  $c_i$  is a constant, and  $x_i$  is the  $i$ -th argument of the signature of  $f$ . When the  
248 argument is a constant  $c_i$ , a base case with  $c_i$  is added. In the second case, a base case is  
249 added for each constant from 0 up to (but not including)  $c_i$ .

250 ► **Example 9**. Consider the recursive function  $g$  from Example 8. **FindBaseCases** iterates  
251 over two function calls:  $g(l-1, m)$  and  $g(l-1, m-1)$ . The former produces the base case  
252  $g(0, m)$ , while the latter produces both  $g(0, m)$  and  $g(l, 0)$ .



---

**Algorithm 2** FindBaseCases( $\mathcal{E}$ )

---

**Input:** set  $\mathcal{E}$  of equations  
**Output:** set  $\mathcal{B}$  of base cases  
1  $\mathcal{B} \leftarrow \emptyset$ ;  
2 **foreach** function call  $f(\mathbf{y})$  on the right-hand side of an equation in  $\mathcal{E}$  **do**  
3      $\mathbf{x} \leftarrow$  the parameters of  $f$  in its definition;  
4     **foreach**  $y_i \in \mathbf{y}$  **do**  
5         **if**  $y_i \in \mathbb{N}_0$  **then**  $\mathcal{B} \leftarrow \mathcal{B} \cup \{f(\mathbf{x})[x_i \mapsto y_i]\}$ ;  
6         **else if**  $y_i = x_i - c_i$  for some  $c_i \in \mathbb{N}_0$  **then**  
7             **for**  $j \leftarrow 0$  **to**  $c_i - 1$  **do**  $\mathcal{B} \leftarrow \mathcal{B} \cup \{f(\mathbf{x})[x_i \mapsto j]\}$ ;

---

253 It can be shown that the base cases identified by **FindBaseCases** are sufficient for the  
254 algorithm to terminate.<sup>4</sup> For the remainder of this section, let  $\mathcal{E}$  denote the equations  
255 returned by **CompileWithBaseCases**.

256 ► **Theorem 10** (Termination). *Let  $f$  be an  $n$ -ary function in  $\mathcal{E}$  and  $\mathbf{x} \in \mathbb{N}_0^n$ . Then the*  
257 *evaluation of  $f(\mathbf{x})$  terminates.*

258 We prove Theorem 10 using double induction. First, we apply induction to the number  
259 of functions in  $\mathcal{E}$ . Then, we use induction on the arity of the ‘last’ function in  $\mathcal{E}$  according to  
260 some topological ordering. We begin with a few observations that stem from previous [5, 32]  
261 and this work. Let  $\mathcal{E}$  represent the equations returned by **CompileWithBaseCases**.

262 ► **Observation 11.** *For each function  $f$ , there is precisely one equation  $e \in \mathcal{E}$  with  $f(\mathbf{x})$  on*  
263 *the left-hand side where all  $x_i$ ’s are variables (i.e.,  $e$  is not a base case). We refer to  $e$  as the*  
264 *definition of  $f$ .*

265 ► **Observation 12.** *There is a topological ordering of all functions  $(f_i)_i$  in  $\mathcal{E}$  such that*  
266 *equations in  $\mathcal{E}$  with  $f_i$  on the left-hand side do not contain function calls to  $f_j$  with  $j > i$ .*  
267 *This condition prevents mutual recursion and other cyclic scenarios.*

268 ► **Observation 13.** *For every equation  $(f(\mathbf{x}) = \text{expr}) \in \mathcal{E}$ , the evaluation of  $\text{expr}$  terminates*  
269 *when provided with the values of all relevant function calls.*

270 ► **Corollary 14.** *If  $f$  is a non-recursive function with no function calls on the right-hand*  
271 *side of its definition, then the evaluation of any function call  $f(\mathbf{x})$  terminates.*

272 ► **Observation 15.** *For any equation  $f(\mathbf{x}) = \text{expr}$ , if  $\mathbf{x}$  contains only constants, then  $\text{expr}$*   
273 *cannot include any function calls to  $f$ .*

274 Additionally, we introduce an assumption about the structure of recursion.

275 ► **Assumption 16.** *For every equation  $(f(\mathbf{x}) = \text{expr}) \in \mathcal{E}$ , every recursive function call*  
276  *$f(\mathbf{y}) \in \text{expr}$  satisfies the following:*

- 277 ■ *Each  $y_i$  is either  $x_i - c_i$  or  $c_i$  for some constant  $c_i$ .*
- 278 ■ *There exists  $i$  such that  $y_i = x_i - c_i$  for some  $c_i > 0$ .*

---

<sup>4</sup> Note that characterising the fine-grained complexity of the solutions found by GANTRY or other FOMC algorithms is an emerging area of research. These questions have been partially addressed in previous work [5, 25] and are orthogonal to the goals of this section.



Finally, we assume a particular order of evaluation for function calls using the equations in  $\mathcal{E}$ . Specifically, we assume that base cases are considered before the recursive definition. The exact order in which base cases are considered is immaterial.

► **Assumption 17.** *When multiple equations in  $\mathcal{E}$  match a function call  $f(\mathbf{x})$ , preference is given to an equation with the most constants on its left-hand side.*

With the observations and assumptions mentioned above, we are ready to prove Theorem 10. For readability, we divide the proof into several lemmas of increasing generality.

► **Lemma 18.** *Assume that  $\mathcal{E}$  consists of just one unary function  $f$ . Then the evaluation of a function call  $f(x)$  terminates for any  $x \in \mathbb{N}_0$ .*

**Proof.** If  $f(x)$  is captured by a base case, then its evaluation terminates by Corollary 14 and Observation 15. If  $f$  is not recursive, the evaluation of  $f(x)$  terminates by Corollary 14.

Otherwise, let  $f(y)$  be an arbitrary function call on the right-hand side of the definition of  $f(x)$ . If  $y$  is a constant, then there is a base case for  $f(y)$ . Otherwise, let  $y = x - c$  for some  $c > 0$ . Then there exists  $k \in \mathbb{N}_0$  such that  $0 \leq x - kc \leq c - 1$ . So, after  $k$  iterations, the sequence of function calls  $f(x), f(x - c), f(x - 2c), \dots$  will be captured by the base case  $f(x \bmod c)$ . ◀

► **Lemma 19.** *Generalising Lemma 18, let  $\mathcal{E}$  be a set of equations for one  $n$ -ary function  $f$  for some  $n \geq 1$ . Then the evaluation of  $f(\mathbf{x})$  terminates for any  $\mathbf{x} \in \mathbb{N}_0^n$ .*

**Proof.** If  $f$  is non-recursive, the evaluation of  $f(\mathbf{x})$  terminates by previous arguments. We proceed by induction on  $n$ , with the base case of  $n = 1$  handled by Lemma 18. Assume that  $n > 1$ . Any base case of  $f$  can be seen as a function of arity  $n - 1$ , since one of the parameters is fixed. Thus, the evaluation of any base case terminates by the inductive hypothesis. It remains to show that the evaluation of the recursive equation for  $f$  terminates, but that follows from Observation 13. ◀

**Proof of Theorem 10.** We proceed by induction on the number of functions  $n$ . The base case of  $n = 1$  is handled by Lemma 19. Let  $(f_i)_{i=1}^n$  be some topological ordering of these  $n > 1$  functions. If  $f = f_j$  for  $j < n$ , then the evaluation of  $f(\mathbf{x})$  terminates by the inductive hypothesis since  $f_j$  cannot call  $f_n$  by Observation 12. Using the inductive hypothesis that all function calls to  $f_j$  (with  $j < n$ ) terminate, the proof proceeds similarly to the Proof of Lemma 19. ◀

### 3.3 Propagating Domain Size Assumptions

Algorithm 3, called **Propagate**, modifies the formula  $\phi$  based on the assumption that  $|\Delta| = n$ . When  $n = 0$ , some clauses become vacuously satisfied and can be removed. When  $n > 0$ , partial grounding is performed by replacing all variables quantified over  $\Delta$  with constants. (None of the formulas examined in this work had  $n > 1$ .) Algorithm 3 handles these two cases separately. For a literal or a clause  $C$ , the set of corresponding domains is denoted as  $\text{Doms}(C)$ .

In the case of  $n = 0$ , there are three types of clauses to consider:

1. those that do not mention  $\Delta$ ,
2. those in which every literal contains variables quantified over  $\Delta$ , and
3. those that have some literals with variables quantified over  $\Delta$  and some without.

---

**Algorithm 3**  $\text{Propagate}(\phi, \Delta, n)$ 


---

**Input:** formula  $\phi$ , domain  $\Delta$ ,  $n \in \mathbb{N}_0$   
**Output:** formula  $\phi'$

```

1  $\phi' \leftarrow \emptyset$ ;
2 if  $n = 0$  then
3   foreach clause  $C \in \phi$  do
4     if  $\Delta \notin \text{Doms}(C)$  then  $\phi' \leftarrow \phi' \cup \{C\}$ ;
5     else
6        $C' \leftarrow \{l \in C \mid \Delta \notin \text{Doms}(l)\}$ ;
7       if  $C' \neq \emptyset$  then
8          $l \leftarrow$  an arbitrary literal in  $C'$ ;
9          $\phi' \leftarrow \phi' \cup \{C' \cup \{\neg l\}\}$ ;
10  else
11     $D \leftarrow$  a set of  $n$  new constants in  $\Delta$ ;
12    foreach clause  $C \in \phi$  do
13       $(x_i)_{i=1}^m \leftarrow$  the variables in  $C$  with domain  $\Delta$ ;
14      if  $m = 0$  then  $\phi' \leftarrow \phi' \cup \{C\}$ ;
15      else  $\phi' \leftarrow \phi' \cup \{C[x_1 \mapsto c_1, \dots, x_m \mapsto c_m] \mid (c_i)_{i=1}^m \in D^m\}$ ;

```

---

320 Clauses of Type 1 are transferred to the new formula  $\phi'$  without any changes. For clauses of  
 321 Type 2,  $C'$  is empty, so these clauses are filtered out. As for clauses of Type 3, a new kind of  
 322 smoothing is performed, which will be explained in Section 3.4.

323 In the case of  $n > 0$ ,  $n$  new constants are introduced. Let  $C$  be an arbitrary clause in  $\phi$ ,  
 324 and let  $m \in \mathbb{N}_0$  be the number of variables in  $C$  quantified over  $\Delta$ . If  $m = 0$ ,  $C$  is added  
 325 directly to  $\phi'$ . Otherwise, a clause is added to  $\phi'$  for every possible combination of replacing  
 326 the  $m$  variables in  $C$  with the  $n$  new constants.

327 ► **Example 20.** Let  $C \equiv \forall x \in \Gamma. \forall y, z \in \Delta. \neg P(x, y) \vee \neg P(x, z) \vee y = z$ . Then  $\text{Doms}(C) =$   
 328  $\text{Doms}(\neg P(x, y)) = \text{Doms}(\neg P(x, z)) = \{\Gamma, \Delta\}$ , and  $\text{Doms}(y = z) = \{\Delta\}$ . A call to  
 329  $\text{Propagate}(\{C\}, \Delta, 3)$  would result in the following formula with nine clauses:

$$\begin{aligned}
 & (\forall x \in \Gamma. \neg P(x, c_1) \vee \neg P(x, c_1) \vee c_1 = c_1) \wedge \\
 & (\forall x \in \Gamma. \neg P(x, c_1) \vee \neg P(x, c_2) \vee c_1 = c_2) \wedge \\
 & \quad \vdots \\
 & (\forall x \in \Gamma. \neg P(x, c_3) \vee \neg P(x, c_3) \vee c_3 = c_3).
 \end{aligned}$$

334 Here,  $c_1$ ,  $c_2$ , and  $c_3$  are the new constants.

### 335 3.4 Smoothing the Base Cases

336 *Smoothing* modifies a circuit to reintroduce eliminated atoms, ensuring the correct model  
 337 count [4, 32]. In this section, we describe a similar process performed on lines 7–9 of  
 338 Algorithm 3. Line 7 checks if smoothing is necessary, and lines 8 and 9 execute it. If the  
 339 condition on line 7 is not satisfied, the clause is not smoothed but omitted.

340 Suppose  $\text{Propagate}$  is called with arguments  $(\phi, \Delta, 0)$ , i.e., we are simplifying the formula  
 341  $\phi$  by assuming that the domain  $\Delta$  is empty. Informally, if there is a predicate  $P$  in  $\phi$  unrelated

342 to  $\Delta$ , smoothing preserves all occurrences of  $P$  even if all clauses with  $P$  become vacuously  
 343 satisfied.

344 ► **Example 21.** Let  $\phi$  be

$$345 \quad (\forall x \in \Delta. \forall y, z \in \Gamma. Q(x) \vee P(y, z)) \wedge \quad (6)$$

$$346 \quad (\forall y, z \in \Gamma'. P(y, z)), \quad (7)$$

347 where  $\Gamma' \subseteq \Gamma$  is a domain introduced by a compilation rule. It should be noted that  $P$ , as a  
 348 relation, is a subset of  $\Gamma \times \Gamma$ .

349 Now, let us reason manually about the model count of  $\phi$  when  $\Delta = \emptyset$ . Predicate  $Q$  can  
 350 only take one value,  $Q = \emptyset$ . The value of  $P$  is fixed over  $\Gamma' \times \Gamma'$  by Clause (7), but it can vary  
 351 freely over  $(\Gamma \times \Gamma) \setminus (\Gamma' \times \Gamma')$  since Clause (6) is vacuously satisfied by all structures. Therefore,  
 352 the correct FOMC should be  $2^{|\Gamma|^2 - |\Gamma'|^2}$ . However, without line 9, **Propagate** would simplify  
 353  $\phi$  to  $\forall y, z \in \Gamma'. P(y, z)$ . In this case,  $P$  is a subset of  $\Gamma' \times \Gamma'$ . This simplified formula has  
 354 only one model:  $\{P(y, z) \mid y, z \in \Gamma'\}$ . By including line 9, **Propagate** transforms  $\phi$  to

$$355 \quad (\forall y, z \in \Gamma. P(y, z) \vee \neg P(y, z)) \wedge (\forall y, z \in \Gamma'. P(y, z)),$$

356 which retains the correct model count.

357 It is worth mentioning that the choice of  $l$  on line 8 of Algorithm 3 is inconsequential  
 358 because any choice achieves the same goal: constructing a tautological clause that retains  
 359 the literals in  $C'$ .

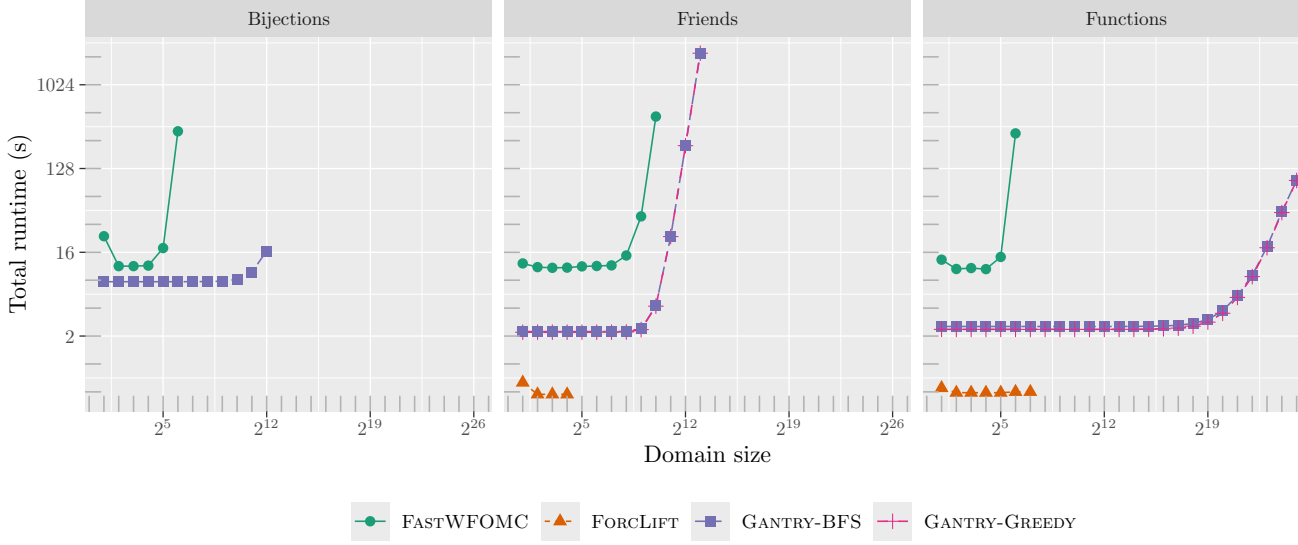
### 360 3.5 Generating C++ Code

361 In this section, we will describe the final step of GANTRY as outlined in Figure 1. This step  
 362 involves translating the set of equations  $\mathcal{E}$  into C++ code. The resulting C++ program  
 363 can then be compiled and executed with different command-line arguments to compute the  
 364 model count of the formula for various domain sizes.

365 Each equation in  $\mathcal{E}$  is compiled into a C++ function, along with a separate cache for  
 366 memoisation. Let us consider an arbitrary equation  $e = (f(\mathbf{x}) = \mathbf{expr}) \in \mathcal{E}$ , and let  $\mathbf{c} \in \mathbb{N}_0^n$   
 367 represent the arguments of the corresponding C++ function. The implementation of  $e$   
 368 consists of three parts. First, we check if  $\mathbf{c}$  is already present in the cache of  $e$ . If it is,  
 369 we simply return the cached value. Second, for each base case  $f(\mathbf{y})$  of  $f(\mathbf{x})$  (as defined in  
 370 Definition 6), we check if  $\mathbf{c}$  *matches*  $\mathbf{y}$ , i.e.,  $c_i = y_i$  whenever  $y_i \in \mathbb{N}_0$ . If this condition is  
 371 satisfied,  $\mathbf{c}$  is redirected to the C++ function that corresponds to the definition of the base  
 372 case  $f(\mathbf{y})$ . Finally, if none of the above cases apply, we evaluate  $\mathbf{c}$  based on the expression  
 373  $\mathbf{expr}$ , store the result in the cache, and return it.

## 374 4 Experimental Evaluation

375 Our empirical evaluation sought to compare the runtime performance of GANTRY with the  
 376 current state of the art, namely FASTWFOMC and FORCLIFT. It is worth remarking that  
 377 FORCLIFT does not support arbitrary precision, and returns error for cases that requires  
 378 arbitrary precision reasoning. Our experiments involve two versions of GANTRY: GANTRY-  
 379 GREEDY and GANTRY-BFS. Like its predecessor, GANTRY has two modes for applying  
 380 compilation rules to formulas: one that uses a greedy search algorithm similar to FORCLIFT  
 381 and another that combines greedy and breadth-first search.



**Figure 2** The runtime of the algorithms as a function of the domain size. Note that both axes are on a logarithmic scale.

The experiments were conducted using an Intel Skylake 2.4 GHz CPU with 188 GiB of memory and CentOS 7. C++ programs were compiled using the Intel C++ Compiler 2020u4. FASTWFOMC ran on Julia 1.10.4, while the other algorithms were executed on the Java Virtual Machine 1.8.0\_201.

## 4.1 Benchmarks

We compare these algorithms using three benchmarks from previous studies. The first benchmark is the function-counting problem from Example 5, previously examined by Dilkas and Belle [5]. The second benchmark is a variant of the well-known ‘Friends and Smokers’ Markov logic network [21, 30]. In  $C^2$ , FO, and  $UFO^2 + CC$ , this problem can be formulated as

$$(\forall x, y \in \Delta. S(x) \wedge F(x, y) \Rightarrow S(y)) \wedge (\forall x \in \Delta. S(x) \Rightarrow C(x))$$

or, equivalently, in conjunctive normal form as

$$(\forall x, y \in \Delta. S(y) \vee \neg S(x) \vee \neg F(x, y)) \wedge (\forall x \in \Delta. C(x) \vee \neg S(x)).$$

Finally, we include the bijection-counting problem previously utilised by Dilkas and Belle [5]. Its formulation in FO is described in Example 8. The equivalent formula in  $C^2$  is

$$(\forall x \in \Delta. \exists^{=1} y \in \Delta. P(x, y)) \wedge (\forall y \in \Delta. \exists^{=1} x \in \Delta. P(x, y)).$$

Similarly, in  $UFO^2 + CC$  the same formula can be written as

$$(\forall x, y \in \Delta. R(x) \vee \neg P(x, y)) \wedge (\forall x, y \in \Delta. S(x) \vee \neg P(y, x)) \wedge (|P| = |\Delta|),$$

where  $w^-(R) = w^-(S) = -1$ .

The three benchmark families cover a wide range of possibilities. The ‘friends’ benchmark stands out as it uses multiple predicates and can be expressed in FO using just two variables

without cardinality constraints or counting quantifiers. The ‘functions’ benchmark, on the other hand, can still be handled by all the algorithms, but it requires cardinality constraints, counting quantifiers, or more than two variables. Lastly, the ‘bijections’ benchmark is an example of a formula that FASTWFOMC can handle but FORCLIFT cannot.

For evaluation purposes, we ran each algorithm on each benchmark using domains of sizes  $2^1, 2^2, 2^3$ , and so on, until an algorithm failed to handle a domain size due to timeout, out of memory error, or out of precision errors. While we separately measured compilation and inference time, we primarily focus on total runtime, dominated by the latter.

## 4.2 Results

Figure 2 presents a summary of the experimental results. Only FASTWFOMC and GANTRY-BFS could handle the bijection-counting problem. For this benchmark, the largest domain sizes these algorithms could accommodate were 64 and 4096, respectively. On the other two benchmarks, FORCLIFT had the lowest runtime. However, due to its finite precision, it only scaled up to domain sizes of 16 and 128 for ‘friends’ and ‘functions’, respectively. FASTWFOMC outperformed FORCLIFT in the case of ‘friends’, but not ‘functions’, as it could handle domains of size 1024 and 64, respectively. Furthermore, both GANTRY-BFS and GANTRY-GREEDY performed similarly on both benchmarks. Similarly to the ‘bijections’ benchmark, GANTRY significantly outperformed the other two algorithms, scaling up to domains of size 8192 and 67,108,864, respectively.

Another aspect of the experimental results that deserves separate discussion is compilation. Both Julia and Scala use just-in-time (JIT) compilation, which means that FASTWFOMC and FORCLIFT take longer to run on the smallest domain size, where most JIT compilation occurs. In the case of GANTRY, it is only run once per benchmark, so the JIT compilation time is included in its overall runtime across all domain sizes. Additionally, while FORCLIFT’s compilation is generally faster than that of GANTRY, neither significantly affects overall runtime. Specifically, FORCLIFT compilation typically takes around 0.5s, while GANTRY compilation takes around 2.3s.

Based on our experiments, which algorithm should be used in practice? If the formula can be handled by FORCLIFT and the domain sizes are reasonably small, FORCLIFT is likely the fastest algorithm. In other situations, GANTRY is expected to be significantly more efficient than FASTWFOMC regardless of domain size, provided both algorithms can handle the formula.

## 5 Conclusion and Future Work

In this work, we have presented a scalable automated FOKC-based approach to FOMC. Our algorithm involves completing the definitions of recursive functions and subsequently translating all function definitions into C++ code. Empirical results demonstrate that GANTRY can scale to larger domain sizes than FASTWFOMC while supporting a wider range of formulas than FORCLIFT. The ability to efficiently handle large domain sizes is particularly crucial in the weighted setting, as illustrated by the ‘friends’ example discussed in Section 4, where the model captures complex social networks with probabilistic relationships. Without this scalability, the practical usefulness of these models would be limited.

Future directions for research include conducting a comprehensive experimental comparison of FOMC algorithms to better understand their comparative performance across various formulas. The capabilities of GANTRY could also be characterised theoretically, e.g. by proving completeness for specific logic fragments like  $C^2$ . Additionally, the efficiency

of FOMC algorithms can be further analysed using fine-grained complexity, which would provide more detailed insights into the computational demands of different formulas.

## References

- 1 Damiano Azzolini and Fabrizio Riguzzi. Lifted inference for statistical statements in probabilistic answer set programming. *Int. J. Approx. Reason.*, 163:109040, 2023.
- 2 Paul Beame, Guy Van den Broeck, Eric Gribkoff, and Dan Suciu. Symmetric weighted first-order model counting. In *PODS*, pages 313–328. ACM, 2015.
- 3 Mark Chavira and Adnan Darwiche. On probabilistic inference by weighted model counting. *Artif. Intell.*, 172(6-7):772–799, 2008.
- 4 Adnan Darwiche. On the tractable counting of theory models and its application to truth maintenance and belief revision. *Journal of Applied Non-Classical Logics*, 11(1-2):11–34, 2001.
- 5 Paulius Dilkas and Vaishak Belle. Synthesising recursive functions for first-order model counting: Challenges, progress, and conjectures. In *KR*, pages 198–207, 2023.
- 6 Vibhav Gogate and Pedro M. Domingos. Probabilistic theorem proving. *Commun. ACM*, 59(7):107–115, 2016.
- 7 Eric Gribkoff, Dan Suciu, and Guy Van den Broeck. Lifted probabilistic inference: A guide for the database researcher. *IEEE Data Eng. Bull.*, 37(3):6–17, 2014.
- 8 Peter G. Hinman. *Fundamentals of mathematical logic*. CRC Press, 2018.
- 9 Wilfrid Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997.
- 10 Manfred Jaeger and Guy Van den Broeck. Liftability of probabilistic inference: Upper and lower bounds. In *StarAI@UAI*, 2012.
- 11 Seyed Mehran Kazemi and David Poole. Knowledge compilation for lifted probabilistic inference: Compiling to a low-level language. In *KR*, pages 561–564. AAAI Press, 2016.
- 12 Kristian Kersting. Lifted probabilistic inference. In *ECAI*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 33–38. IOS Press, 2012.
- 13 Ondrej Kuželka. Weighted first-order model counting in the two-variable fragment with counting quantifiers. *J. Artif. Intell. Res.*, 70:1281–1307, 2021.
- 14 Sagar Malhotra and Luciano Serafini. Weighted model counting in FO2 with cardinality constraints and counting quantifiers: A closed form formula. In *AAAI*, pages 5817–5824. AAAI Press, 2022.
- 15 Nils J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28(1):71–87, 1986.
- 16 Feng Niu, Christopher Ré, AnHai Doan, and Jude W. Shavlik. Tuffy: Scaling up statistical inference in Markov logic networks using an RDBMS. *Proc. VLDB Endow.*, 4(6):373–384, 2011.
- 17 Vilém Novák, Irina Perfilieva, and Jiri Mockor. *Mathematical principles of fuzzy logic*, volume 517. Springer Science & Business Media, 2012.
- 18 Fabrizio Riguzzi, Elena Bellodi, Riccardo Zese, Giuseppe Cota, and Evelina Lamma. A survey of lifted inference approaches for probabilistic logic programming under the distribution semantics. *Int. J. Approx. Reason.*, 80:313–333, 2017.
- 19 Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach (4th Edition)*. Pearson, 2020.
- 20 Dragan Z. Šaletić. Graded logics. *Interdisciplinary Description of Complex Systems: INDECS*, 22(3):276–295, 2024.
- 21 Parag Singla and Pedro M. Domingos. Lifted first-order belief propagation. In *AAAI*, pages 1094–1099. AAAI Press, 2008.
- 22 Martin Svatos, Peter Jung, Jan Tóth, Yuyi Wang, and Ondrej Kuželka. On discovering interesting combinatorial integer sequences. In *IJCAI*, pages 3338–3346. ijcai.org, 2023.
- 23 Jan Tóth and Ondrej Kuželka. Lifted inference with linear order axiom. In *AAAI*, pages 12295–12304. AAAI Press, 2023.

- 496 24 Pietro Totis, Jesse Davis, Luc De Raedt, and Angelika Kimmig. Lifted reasoning for combina-  
497 torial counting. *J. Artif. Intell. Res.*, 76:1–58, 2023.
- 498 25 Jan Tóth and Ondřej Kuželka. Complexity of weighted first-order model counting in the  
499 two-variable fragment with counting quantifiers: A bound to beat, 2024. URL: <https://arxiv.org/abs/2404.12905>, arXiv:2404.12905.
- 500 26 Timothy van Bremen and Ondrej Kuželka. Approximate weighted first-order model counting:  
501 Exploiting fast approximate model counters and symmetry. In *IJCAI*, pages 4252–4258.  
502 ijcai.org, 2020.
- 503 27 Timothy van Bremen and Ondrej Kuželka. Faster lifting for two-variable logic using cell  
504 graphs. In *UAI*, volume 161 of *Proceedings of Machine Learning Research*, pages 1393–1402.  
505 AUAI Press, 2021.
- 506 28 Timothy van Bremen and Ondrej Kuželka. Lifted inference with tree axioms. *Artif. Intell.*,  
507 324:103997, 2023.
- 508 29 Guy Van den Broeck. On the completeness of first-order knowledge compilation for lifted  
509 probabilistic inference. In *NIPS*, pages 1386–1394, 2011.
- 510 30 Guy Van den Broeck, Arthur Choi, and Adnan Darwiche. Lifted relax, compensate and then  
511 recover: From approximate to exact lifted probabilistic inference. In *UAI*, pages 131–141.  
512 AUAI Press, 2012.
- 513 31 Guy Van den Broeck, Wannes Meert, and Adnan Darwiche. Skolemization for weighted  
514 first-order model counting. In *KR*. AAAI Press, 2014.
- 515 32 Guy Van den Broeck, Nima Taghipour, Wannes Meert, Jesse Davis, and Luc De Raedt. Lifted  
516 probabilistic inference by first-order knowledge compilation. In *IJCAI*, pages 2178–2185.  
517 IJCAI/AAAI, 2011.
- 518 33 Yuanhong Wang, Juhua Pu, Yuyi Wang, and Ondrej Kuželka. Lifted algorithms for symmetric  
519 weighted first-order model sampling. *Artif. Intell.*, 331:104114, 2024.
- 520