

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 This research was funded in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2024-05956, and the Digital Research Alliance of Canada (alliancecan.ca).

Acknowledgements The first two authors contributed equally. 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 [28]. It builds upon earlier efforts in weighted model counting for propositional logic [4] and broader attempts to bridge logic and probability [12, 14, 17]. WFOMC is central to *lifted inference*, which enhances the efficiency of probabilistic calculations by exploiting symmetries [10]. Lifted inference continues to advance, with applications extending to constraint satisfaction problems [22] and probabilistic answer set programming [1]. Moreover, WFOMC has proven effective at reasoning over probabilistic databases [8] and probabilistic logic programs [16]. FOMC algorithms have also facilitated breakthroughs in discovering integer sequences [19] and developing recurrence relations for these sequences [6]. Recently, these algorithms have been extended to perform sampling tasks [29].

The complexity of FOMC is generally measured by *data complexity*, with a sentence classified as *liftable* if it admits a polynomial-time solution relative to the domain size [9]. While all sentences with up to two variables are known to be lifttable [25, 27], Beame et al. [3] demonstrated that liftability does not extend to all sentences, identifying an unlifttable sentence with three variables. Recent work has extended the lifttable fragment with additional axioms [20, 24] and counting quantifiers [11], expanding our understanding of liftability.

FOMC algorithms are diverse, with approaches ranging from *first-order knowledge compilation* (FOKC) to cell enumeration [23], local search [13], and Monte Carlo sampling [7]. (See the technical appendix for a more detailed comparison of the state-of-the-art FOMC algorithms.) FOKC-based algorithms are particularly prominent, transforming sentences into structured representations such as circuits or graphs. Even when multiple algorithms can solve the same instance, FOKC algorithms are known to find polynomial-time solutions where the polynomial has a lower degree than other approaches [6]. The recently developed ability of a FOKC algorithm to formulate solutions in terms of recursive functions [6] is also noteworthy since the only other proposed alternative is to guess recursive relations [2]. Notable examples of FOKC algorithms include FORCLIFT [28] and its extension CRANE [6].

The CRANE algorithm marked a significant step forward, expanding the range of sentences 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 difficult 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 those of previous algorithms and simplifies the user experience by automatically managing base cases and compiling function definitions into efficient C++ programs.

The paper is structured as follows. Section 2 covers the necessary preliminaries, including notation, terminology, and background information about (W)FOMC and FOKC. Section 3 describes all aspects of GANTRY:

- its overall structure and interactions with CRANE,
- how we identify a sufficient set of base cases for a recursive function,
- how we construct the sentence that describes each base case, and
- how we compile function definitions into a C++ program with caching mechanisms that ensure efficiency.

Then, Section 4 presents our experimental results, demonstrating GANTRY’s performance compared to other FOMC algorithms. Finally, Section 5 concludes the paper by discussing promising avenues for future work.

2 Preliminaries

We begin this section by describing some notation that we will use throughout the paper. Then, in Sections 2.1 and 2.2, we introduce the basic terminology of first-order logic and formally define (W)FOMC. Section 2.3 outlines the principles of FOKC, particularly in the context of CRANE. Finally, in Section 2.4, we introduce the algebraic terminology used to describe the output of CRANE, i.e., functions and equations that define them.

Notation

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* σ to an expression S , where $\sigma = [x_1 \mapsto y_1, x_2 \mapsto$

89 $y_2, \dots, x_n \mapsto y_n]$ signifies the replacement of all instances of x_i with y_i for all $i = 1, \dots, n$.
 90 Additionally, for any variable n and $a, b \in \mathbb{N}_0$, let $[a \leq n \leq b] := \begin{cases} 1 & \text{if } a \leq n \leq b \\ 0 & \text{otherwise} \end{cases}$.

91 2.1 First-Order Logic

92 This section reviews the basic concepts of first-order logic used in FOKC algorithms. We focus
 93 specifically on the format used internally by FORCLIFT and its descendants. See Section 2.3.1
 94 for a brief overview of how GANTRY transforms an arbitrary sentence in first-order logic into
 95 this internal format.

96 A *term* can be either a variable or a constant. An *atom* can be either $P(t_1, \dots, t_m)$ (i.e.,
 97 $P(\mathbf{t})$) for some predicate P and terms t_1, \dots, t_m , or $x = y$ for some terms x and y . The
 98 *arity* of a predicate is the number of arguments it takes, i.e., m in the case of the predicate
 99 P mentioned above. We write P/m to denote a predicate along with its arity. A *literal*
 100 can be either an atom (i.e., a *positive* literal) or its negation (i.e., a *negative* literal). An
 101 atom containing no variables, only constants, is called *ground*. A *clause* is of the form
 102 $\forall x_1 \in \Delta_1. \forall x_2 \in \Delta_2 \dots \forall x_n \in \Delta_n. \phi(x_1, x_2, \dots, x_n)$, where ϕ is a disjunction of literals that
 103 contain only the variables x_1, \dots, x_n (and any constants). We say that a clause is a (*positive*)
 104 *unit clause* if there is only one literal with a predicate, and it is a positive literal. Finally, a
 105 *sentence* is a conjunction of clauses. Throughout the paper, we will use set-theoretic notation,
 106 interpreting a sentence as a set of clauses and a clause as a set of literals.

107 ► **Remark.** Conforming to previous work [28], the definition of a clause includes universal
 108 quantifiers for all its variables. While it is possible to rewrite the entire sentence with all
 109 quantifiers at the front, the format we describe has proven convenient for practical use.

110 2.2 First-Order Model Counting

111 In this section, we will formally define FOMC and its weighted variant. Although this
 112 work focuses on FOMC, computing the FOMC using GANTRY requires using WFOMC
 113 for sentences with existential quantifiers. For such sentences, preprocessing (described in
 114 Section 2.3.1) introduces predicates with non-unary weights that must be accounted for to
 115 compute the correct model count.

116 ► **Definition 1** (Structure, model). Let ϕ be a sentence. For each predicate P/n in ϕ , let
 117 $(\Delta_i^P)_{i=1}^n$ be a list of the corresponding domains. Let σ be a map from the domains of ϕ
 118 to their interpretations as finite sets, ensuring the sets are pairwise disjoint and contain
 119 the corresponding constants from ϕ . A *structure* of ϕ is a set M of ground literals defined
 120 by adding to M either $P(\mathbf{t})$ or $\neg P(\mathbf{t})$ for every predicate P/n in ϕ and every n -tuple
 121 $\mathbf{t} \in \prod_{i=1}^n \sigma(\Delta_i^P)$. A *structure* is a *model* if it makes ϕ valid.

122 ► **Example 2** (Counting bijections). Let us consider the following sentence (previously
 123 examined by Dilkas and Belle [6]) that defines predicate P as a bijection between two
 124 domains Γ and Δ :

$$\begin{aligned}
 & (\forall x \in \Gamma. \exists y \in \Delta. P(x, y)) \wedge \\
 & (\forall y \in \Delta. \exists x \in \Gamma. P(x, y)) \wedge \\
 & (\forall x \in \Gamma. \forall y, z \in \Delta. P(x, y) \wedge P(x, z) \Rightarrow y = z) \wedge \\
 & (\forall x, z \in \Gamma. \forall y \in \Delta. P(x, y) \wedge P(z, y) \Rightarrow x = z).
 \end{aligned} \tag{1}$$

Let σ be defined as $\sigma(\Gamma) := \{1, 2\}$ and $\sigma(\Delta) := \{a, b\}$. Sentence (1) has two models:

$$\{P(1, a), P(2, b), \neg P(1, b), \neg P(2, a)\} \quad \text{and} \quad \{P(1, b), P(2, a), \neg P(1, a), \neg P(2, b)\}.$$

► **Remark.** The distinctness of domains is significant in two respects. First, in terms of expressiveness, a clause such as $\forall x \in \Delta. P(x, x)$ is valid if predicate P operates over two copies of the same domain and invalid otherwise. Second, having more distinct domains makes the problem more decomposable for the 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.

► **Definition 3** (WFOMC instance). A WFOMC instance *comprises*:

- a sentence ϕ ,
 - two (rational) weights $w^+(P)$ and $w^-(P)$ assigned to each predicate P in ϕ , and
 - σ as described in Definition 1.
- Unless specified otherwise, we assume all weights are equal to 1.

► **Definition 4** (WFOMC [28]). Given a WFOMC instance (ϕ, w^+, w^-, σ) as in Definition 3, the (symmetric) weighted first-order model count (WFOMC) of ϕ is

$$\sum_{M \models \phi} \prod_{P(\mathbf{t}) \in M} w^+(P) \prod_{\neg P(\mathbf{t}) \in M} w^-(P), \quad (2)$$

where the sum is over all models of ϕ .

2.3 Crane and First-Order Knowledge Compilation

As our work builds on CRANE, in this section, we will briefly outline the steps CRANE goes through to compile a sentence into a set of function definitions. We divide the inner workings of the algorithm into two stages: preprocessing and compilation.

2.3.1 Preprocessing

This stage transforms an arbitrary sentence into the format described in Section 2.1, primarily by eliminating existential quantifiers. For example, the first conjunct of sentence (1), i.e.,

$$\forall x \in \Gamma. \exists y \in \Delta. P(x, y) \quad (3)$$

is transformed into

$$\begin{aligned} & (\forall x \in \Gamma. Z(x)) \wedge \\ & (\forall x \in \Gamma. \forall y \in \Delta. Z(x) \vee \neg P(x, y)) \wedge \\ & (\forall x \in \Gamma. S(x) \vee Z(x)) \wedge \\ & (\forall x \in \Gamma. \forall y \in \Delta. S(x) \vee \neg P(x, y)), \end{aligned} \quad (4)$$

where $Z/1$ and $S/1$ are two new predicates with $w^-(S) = -1$. One can verify that the WFOMC of sentences (3) and (4) is the same.

2.3.2 Compilation

After preprocessing, CRANE compiles the sentence into the triple $(\mathcal{E}, \mathcal{F}, \mathcal{D})$, where \mathcal{E} is the set of equations, and \mathcal{F} and \mathcal{D} are auxiliary maps. \mathcal{F} maps function names to sentences. \mathcal{D} maps function names and argument indices to domains. \mathcal{E} can contain any number of functions, one of which (which we will always denote as f) represents the solution to the FOMC problem. Computing the FOMC for specific domain sizes involves invoking f with those sizes as inputs. \mathcal{D} records this correspondence between function arguments and domains.

► **Example 5.** CRANE compiles sentence (1) for bijection counting into

$$\mathcal{E} = \left\{ \begin{array}{l} f(m, n) = \sum_{l=0}^n \binom{n}{l} (-1)^{n-l} g(l, m), \\ g(l, m) = \sum_{k=0}^m [0 \leq k \leq 1] \binom{m}{k} g(l-1, m-k) \end{array} \right\};$$

$$\mathcal{D} = \{ (f, 1) \mapsto \Gamma, (f, 2) \mapsto \Delta, (g, 1) \mapsto \Delta^\top, (g, 2) \mapsto \Gamma \},$$

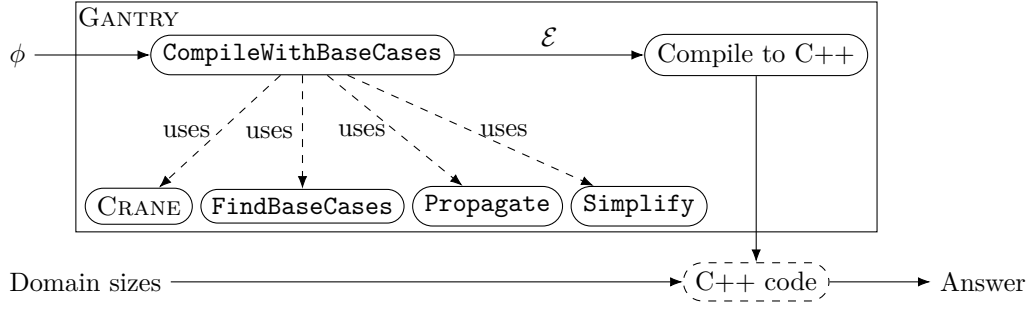
where Δ^\top is a newly introduced domain. (We omit the definition of \mathcal{F} as the sentences can become quite verbose.) To compute the number of bijections between two sets of cardinality 3, one would evaluate $f(3, 3)$; however, the definition of g is incomplete: g is a recursive function presented without any base cases. \mathcal{D} encodes that in $f(m, n)$, m and n represent $|\Gamma|$ and $|\Delta|$, respectively. Similarly, in $g(l, m)$, l represents $|\Delta^\top|$, and m represents $|\Gamma|$.

Compilation is performed primarily by applying (*compilation*) rules to sentences. CRANE has two modes depending on the selection process for compilation rules when multiple alternatives are available. The first option is to use greedy search: there is a list of rules, and the first applicable rule is the one that gets used, disregarding all the others. The second option is to use a combination of greedy and *breadth-first search* (BFS). In this approach, we classify each compilation rule as greedy or non-greedy. Greedy rules are applied as soon as possible at any stage of the compilation process, while BFS is executed over all applicable non-greedy rules, identifying the solution that necessitates the smallest number of non-greedy rules.

2.4 Algebra

In this paper, we use both logical and algebraic constructs. While the rest of Section 2 focused on the former, this section describes the latter. We write **expr** for an arbitrary algebraic expression. In the context of algebra, a *constant* is a non-negative integer. Likewise, a *variable* can either be a parameter of a function or a variable introduced through summation, such as i in the expression $\sum_{i=1}^n \mathbf{expr}$. A *function 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 algebraic expression consisting of variables and constants. A (function) *signature* is a function call that contains only variables. Given two function calls, $f(\mathbf{x})$ and $f(\mathbf{y})$, we say that $f(\mathbf{y})$ *matches* $f(\mathbf{x})$ if $x_i = y_i$ whenever $x_i, y_i \in \mathbb{N}_0$. An *equation* is $f(\mathbf{x}) = \mathbf{expr}$, where $f(\mathbf{x})$ is a function call.

► **Definition 6** (Base case). *Let $f(\mathbf{x})$ be a function call where each x_i is either a constant or a variable. Then the function call $f(\mathbf{y})$ is a base case of $f(\mathbf{x})$ if $f(\mathbf{y}) = f(\mathbf{x})\sigma$, where σ is a substitution that replaces one or more variable x_i 's with a constant while leaving constants unchanged.*



■ **Figure 1** Using GANTRY to compute the model count of a sentence ϕ . First, GANTRY compiles ϕ into a set of equations, which form the basis for creating a C++ program. Executing the program with different command line arguments calculates the model count of ϕ for various domain sizes. To accomplish this, the `CompileWithBaseCases` procedure employs CRANE, algebraic simplification techniques (denoted as `Simplify`), and two other auxiliary procedures.

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

198 3 Technical Contributions

199 Figure 1 provides an overview of GANTRY’s workflow. We will briefly describe and motivate
 200 each procedure before going into more detail in the corresponding subsection.

201 `CompileWithBaseCases` (see Section 3.1), the core procedure of GANTRY, is responsible
 202 for completing the function definitions produced by CRANE with the necessary base cases.
 203 To do so, it may recursively call itself (and CRANE) on other sentences. We prove that the
 204 number of such recursive calls is upper bound by the number of domains.

205 Section 3.1 also describes the `Simplify` procedure for algebraic simplification. It is crucial
 206 for simplifying, e.g., a sum of n terms, only some of which are non-zero. More generally, the
 207 equations returned by CRANE often benefit from easily detectable algebraic simplifications
 208 such as $0 \cdot \text{anything} = 0$ and $\text{anything}^0 = 1$.

209 `FindBaseCases` (described in Section 3.2) inspects a set of equations to identify a sufficient
 210 set of base cases for a given set of equations. We prove that the returned set of base cases
 211 ensures that the evaluation of the resulting function definitions will never get stuck in an
 212 infinite loop.

213 Section 3.3 introduces the `Propagate` procedure, which takes a sentence ϕ , a domain
 214 Δ , and $n \in \mathbb{N}_0$. It returns ϕ transformed under the assumption that $|\Delta| = n$, with n
 215 new constants added and all variables quantified over Δ eliminated. For example, when
 216 computing a base case such as $f(0, y)$, `Propagate` will significantly simplify ϕ with the
 217 assumption that the domain associated with the first parameter of f (i.e., $\mathcal{D}(f, 1)$) is empty.
 218 `CompileWithBaseCases(Propagate(ϕ , $\mathcal{D}(f, 1)$, 0))` will then return the equations for the
 219 base case $f(0, y)$.

220 Section 3.4 describes a new *smoothing* procedure that ensures `Propagate` preserves
 221 the correct model count. Smoothing is a well-known technique in knowledge compilation
 222 algorithms for propositional model counting [5]. Although FOMC algorithms have used
 223 smoothing before [28], our setting requires a novel approach.

Algorithm 1 `CompileWithBaseCases(ϕ)

---`**Input:** sentence ϕ **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    $\mathcal{E} \leftarrow \mathcal{E} \cup \text{CompileWithBaseCases}(\psi);$ 

```

224 With the help of other procedures outlined above, `CompileWithBaseCases` returns a set
 225 of equations that fully cover the base cases of all recursive functions. While these equations
 226 can be intriguing and valuable in their own right, users of FOMC algorithms typically expect
 227 a numerical answer. Thus, Section 3.5 describes how GANTRY compiles these equations into
 228 a C++ program that one can execute with different command-line arguments to compute
 229 the model count for various combinations of domain sizes.

230 3.1 Completing the Definitions of Functions

231 Algorithm 1 presents our overall approach for compiling a sentence into equations that
 232 include the necessary base cases. First, we use `CRANE` to compile the sentence into three
 233 components: \mathcal{E} , \mathcal{F} , and \mathcal{D} (as described in Section 2.3.2). After some algebraic simplifications
 234 (described below), the algorithm passes \mathcal{E} to the `FindBaseCases` procedure (see Section 3.2).
 235 For each base case $f(\mathbf{x})$, we retrieve the sentence $\mathcal{F}(f)$ associated with the function name f
 236 and simplify it using the `Propagate` procedure (explained in detail in Section 3.3). We do
 237 this by iterating over all indices of \mathbf{x} , where x_i is a constant, and using `Propagate` to simplify
 238 ψ by assuming that domain $\mathcal{D}(f, i)$ has size x_i . Finally, on line 6, `CompileWithBaseCases`
 239 recurses on these simplified sentences and adds the resulting base case equations to \mathcal{E} .

240 Simplify

241 The main responsibility of the `Simplify` procedure is to handle the algebraic pattern
 242 $\sum_{m=0}^n [a \leq m \leq b] f(m)$. Here, n is a variable, a and b are constants, and f is an expression
 243 that may depend on m . `Simplify` transforms this pattern into $f(a) + f(a+1) + \dots +$
 244 $f(\min\{n, b\})$.

245 ► **Example 8.** We return to the bijection-counting problem from Example 2 and its initial
 246 solution described in Example 5. `Simplify` transforms

$$247 \quad g(l, m) = \sum_{k=0}^m [0 \leq k \leq 1] \binom{m}{k} g(l-1, m-k)$$

248 into

$$249 \quad g(l, m) = g(l-1, m) + mg(l-1, m-1).$$

250 Then `FindBaseCases` identifies two base cases: $g(0, m)$ and $g(l, 0)$. In both cases,
 251 `CompileWithBaseCases` recurses on the sentence $\mathcal{F}(g)$ simplified by assuming that one
 252 of the domains is empty. In the first case, we recurse on the sentence $\forall x \in \Gamma. S(x) \vee \neg S(x)$,

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 RHS 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 where S is a predicate introduced by preprocessing with weights $w^+(S) = 1$ and $w^-(S) = -1$.
254 Hence, we obtain the base case $g(0, m) = 0^m$. In the case of $g(l, 0)$, **Propagate**($\psi, \Gamma, 0$)
255 returns an empty sentence, resulting in $g(l, 0) = 1$. While these base cases overlap when
256 $l = m = 0$, they remain consistent since $0^0 = 1$.

257 Generally, let ϕ be a sentence with two domains Γ and Δ , and let $n, m \in \mathbb{N}_0$. Then the
258 FOMC of **Propagate**(ϕ, Δ, n), assuming $|\Gamma| = m$, is the same as the FOMC of **Propagate**(ϕ ,
259 Γ, m), assuming $|\Delta| = n$.

260 Although **CompileWithBaseCases** starts with a call to **CRANE**, the proposed algorithm
261 is not merely a post-processing step for FOKC because Algorithm 1 is recursive and can
262 issue additional calls to **CRANE** on various derived sentences. We conclude this section by
263 bounding the recursion depth of the **CompileWithBaseCases** procedure, thereby also proving
264 that the algorithm terminates.

265 ► **Theorem 9.** *Let ϕ be a sentence with n domains. The maximum recursion depth of*
266 ***CompileWithBaseCases**(ϕ) is then n .*

267 The proof of this theorem relies on two observations regarding the algorithms presented
268 in Sections 3.2 and 3.3.

269 ► **Observation 10.** *Each base case returned by **FindBaseCases** contains at least one constant*
270 *(in line with Definition 6).*

271 ► **Observation 11.** *For any sentence ϕ , domain Δ , and $n \in \mathbb{N}_0$, **Propagate**(ϕ, Δ, n) returns*
272 *a sentence with no variables quantified over Δ .*

273 **Proof.** We proceed by induction on n . If $n = 0$, then ϕ is essentially a propositional
274 formula, and **CRANE** compiles it into an equation of the form $f = \text{expr}$ with no ‘function
275 calls’. Suppose that—for all sentences with at most n domains—**CompileWithBaseCases**
276 terminates with a recursion depth of at most n . Let ϕ be a sentence with $n + 1$ domains. By
277 Observation 10, each base case on line 3 of Algorithm 1 has at least one constant. Therefore,
278 by Observation 11, after line 5, the sentence ψ has at most n domains. Thus, line 6 terminates
279 with a recursion depth of at most n by the inductive hypothesis, completing the inductive
280 proof. ◀

3.2 Identifying a Sufficient Set of Base Cases

281 Algorithm 2 summarises the implementation of **FindBaseCases**. It considers two types of
282 arguments when a function f calls itself recursively: constants and arguments of the form
283

284 $x_i - c_i$. When the argument is a constant c_i , line 5 adds a base case with c_i to the set of all
 285 base cases \mathcal{B} . In the second case, line 7 adds a base case to \mathcal{B} for each constant from 0 to
 286 (but not including) c_i .

287 ► **Example 12.** Consider the recursive function g from Example 5. `FindBaseCases` iterates
 288 over two function calls: $g(l-1, m)$ and $g(l-1, m-1)$. The former produces the base case
 289 $g(0, m)$, while the latter produces both $g(0, m)$ and $g(l, 0)$.

290 In the rest of this section, we will show that the base cases identified by `FindBaseCases`
 291 are sufficient for the algorithm to terminate.¹ Let \mathcal{E} denote the equations returned by
 292 `CompileWithBaseCases`.

293 ► **Theorem 13.** *Given an n -ary function f in \mathcal{E} and $\mathbf{x} \in \mathbb{N}_0^n$, the evaluation of $f(\mathbf{x})$
 294 terminates.*

295 We prove Theorem 13 using double induction. First, we apply induction to the number
 296 of functions in \mathcal{E} . Then, we use induction on the arity of the ‘last’ function in \mathcal{E} according to
 297 a topological ordering. Before proving Theorem 13, we make a few observations about this
 298 and previous [6, 28] work.

299 ► **Observation 14.** *For each function f , there is precisely one equation $e \in \mathcal{E}$ with $f(\mathbf{x})$
 300 on the LHS where all x_i ’s are variables (i.e., e is not a base case). We refer to e as the
 301 definition of f .*

302 ► **Observation 15.** *There is a topological ordering $(f_i)_i$ of all functions in \mathcal{E} such that
 303 equations in \mathcal{E} with f_i on the LHS do not contain function calls to f_j with $j > i$.*

304 Observation 15 prevents mutual recursion and other cyclic scenarios.

305 ► **Observation 16.** *For each equation $(f(\mathbf{x}) = \text{expr}) \in \mathcal{E}$, the evaluation of expr terminates
 306 when provided with the values of all relevant function calls.*

307 ► **Corollary 17.** *If f is a non-recursive function with no function calls on the RHS of its
 308 definition, then the evaluation of any function call $f(\mathbf{x})$ terminates.*

309 ► **Observation 18.** *For each equation $(f(\mathbf{x}) = \text{expr}) \in \mathcal{E}$, if \mathbf{x} contains only constants, then
 310 expr cannot include any function calls to f .*

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

312 ► **Assumption 19.** *For each equation $(f(\mathbf{x}) = \text{expr}) \in \mathcal{E}$, every recursive function call
 313 $f(\mathbf{y}) \in \text{expr}$ satisfies the following:*

- 314 ■ *each y_i is either $x_i - c_i$ or c_i for some constant c_i and*
- 315 ■ *there exists i such that $y_i = x_i - c_i$ for some $c_i > 0$.*

316 Finally, we assume a particular order of evaluation for function calls using the equations
 317 in \mathcal{E} ; specifically, base cases precede the recursive definition.

318 ► **Assumption 20.** *When multiple equations in \mathcal{E} match a function call $f(\mathbf{x})$, preference is
 319 given to the equation with the most constants on its LHS.*

¹ 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 [6, 21] and are unrelated to the goals of this section.

320 With the observations and assumptions mentioned above, we are ready to prove Theo-
 321 rem 13. For readability, we divide the proof into several lemmas of increasing generality.

322 ► **Lemma 21.** *Assume that \mathcal{E} consists of just one unary function called f . Then the evaluation*
 323 *of $f(x)$ terminates for any $x \in \mathbb{N}_0$.*

324 **Proof.** If $f(x)$ matches a base case, then its evaluation terminates by Corollary 17 and Ob-
 325 servation 18. If f is not recursive, the evaluation of $f(x)$ also terminates by Corollary 17.

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

330 ► **Lemma 22.** *Generalising Lemma 21, let \mathcal{E} be a set of equations for one n -ary function f*
 331 *for some $n \geq 1$. The evaluation of $f(\mathbf{x})$ terminates for any $\mathbf{x} \in \mathbb{N}_0^n$.*

332 **Proof.** If f is non-recursive, the evaluation of $f(\mathbf{x})$ terminates by previous arguments. We
 333 proceed by induction on n , with the base case of $n = 1$ handled by Lemma 21. Assume that
 334 $n > 1$. Any base case of f can be seen as a function of arity $n - 1$ since one of the parameters
 335 is constant. Thus, the evaluation of any base case terminates by the inductive hypothesis.
 336 It remains to show that the evaluation of the recursive equation for f terminates, but this
 337 follows from Observation 16. ◀

338 **Proof of Theorem 13.** We proceed by induction on the number of functions n . The base
 339 case of $n = 1$ is handled by Lemma 22. Let $(f_i)_{i=1}^n$ be some topological ordering of these
 340 $n > 1$ functions. If $f = f_j$ for $j < n$, then the evaluation of $f(\mathbf{x})$ terminates by the inductive
 341 hypothesis since f_j cannot call f_n by Observation 15. Using the inductive hypothesis that
 342 all function calls to f_j (with $j < n$) terminate, the proof proceeds similarly to the proof of
 343 Lemma 22. ◀

344 **3.3 Propagating Domain Size Assumptions**

345 Algorithm 3, called **Propagate**, modifies the sentence ϕ based on the assumption that $|\Delta| = n$.
 346 When $n = 0$, some clauses become vacuously satisfied and can be removed. When $n > 0$,
 347 partial grounding replaces all variables with domain Δ with constants. (None of the sentences
 348 examined in this work had $n > 1$.) Algorithm 3 handles these two cases separately. For a
 349 literal or clause C , we write the set of corresponding domains as $\text{Doms}(C)$. In the case of
 350 $n = 0$, there are three types of clauses to consider:

- 351 1. those that do not mention Δ ,
- 352 2. those in which every literal contains variables quantified over Δ and
- 353 3. those with some literals containing such variables and some without.

354 We transfer clauses of Type 1 to the new sentence ϕ' without any changes. For clauses of
 355 Type 2, C' is empty, so these clauses are filtered out. As for clauses of Type 3, lines 7–9
 356 perform a new kind of smoothing, the explanation of which we defer to Section 3.4.

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

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

Input: sentence ϕ , domain Δ , $n \in \mathbb{N}_0$
Output: sentence ϕ'

```

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  $\phi' \leftarrow \phi' \cup \{C[x_1 \mapsto c_1, \dots, x_m \mapsto c_m] \mid (c_i)_{i=1}^m \in D^m\}$ ;

```

361 ► **Example 23.** Let $C := \forall x \in \Gamma. \forall y, z \in \Delta. \neg P(x, y) \vee \neg P(x, z) \vee y = z$. Then $\text{Doms}(C) =$
 362 $\text{Doms}(\neg P(x, y)) = \text{Doms}(\neg P(x, z)) = \{\Gamma, \Delta\}$, and $\text{Doms}(y = z) = \{\Delta\}$. A call to
 363 $\text{Propagate}(\{C\}, \Delta, 3)$ would result in the following sentence with nine clauses:

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

368 Here, c_1 , c_2 , and c_3 are the new constants.

3.4 Smoothing the Base Cases

370 Smoothing modifies a circuit to reintroduce eliminated atoms, ensuring the correct model
 371 count [5, 28]. This section describes a similar process performed on lines 7–9 of Algorithm 3.
 372 Line 7 checks if smoothing is necessary, and lines 8 and 9 execute it. If the condition on
 373 line 7 is not satisfied, the clause is not smoothed but omitted.

374 Suppose $\text{CompileWithBaseCases}$ calls Propagate with arguments $(\phi, \Delta, 0)$, i.e., we are
 375 simplifying the sentence ϕ by assuming that the domain Δ is empty. Informally, if there is a
 376 predicate P in ϕ unrelated to Δ , smoothing preserves all occurrences of P , even if all clauses
 377 with P become vacuously satisfied.

378 ► **Example 24.** Let ϕ be

$$(\forall x \in \Delta. \forall y, z \in \Gamma. Q(x) \vee P(y, z)) \wedge \tag{5}$$

$$(\forall y, z \in \Gamma'. P(y, z)), \tag{6}$$

■ **Algorithm 4** A sketch of the C++ program for the equations in Example 5, particularly highlighting the recursive definition of the function g .

```

1 initialise  $\text{Cache}_{g(0,m)}$ ,  $\text{Cache}_{g(l,0)}$ ,  $\text{Cache}_g$ , and  $\text{Cache}_f$ ;
2 Function  $g_{0,m}(m)$ : ...
3 Function  $g_{l,0}(l)$ : ...
4 Function  $g(l, m)$ :
5   if  $(l, m) \in \text{Cache}_g$  then return  $\text{Cache}_g(l, m)$ ;
6   if  $l = 0$  then return  $g_{0,m}(m)$ ;
7   if  $m = 0$  then return  $g_{l,0}(l)$ ;
8    $r \leftarrow g(l-1, m) + mg(l-1, m-1)$ ;
9    $\text{Cache}_g(l, m) \leftarrow r$ ;
10  return  $r$ ;
11 Function  $f(m, n)$ : ...
12 Function Main:
13    $(m, n) \leftarrow \text{ParseCommandLineArguments}()$ ;
14   return  $f(m, n)$ ;

```

where $\Gamma' \subseteq \Gamma$ is a domain introduced by a compilation rule. Note that P , as a relation, is a subset of $\Gamma \times \Gamma$.

Now, let us reason manually about the model count of ϕ when $\Delta = \emptyset$. Predicate Q can only take one value, $Q = \emptyset$. The value of P is fixed over $\Gamma' \times \Gamma'$ by clause (6), but it can vary freely over $(\Gamma \times \Gamma) \setminus (\Gamma' \times \Gamma')$ since clause (5) is vacuously satisfied by all structures. Therefore, the correct FOMC should be $2^{|\Gamma|^2 - |\Gamma'|^2}$. However, without line 9, **Propagate** would simplify ϕ to $\forall y, z \in \Gamma'. P(y, z)$. In this case, P is a subset of $\Gamma' \times \Gamma'$. This simplified sentence has only one model: $\{P(y, z) \mid y, z \in \Gamma'\}$. By including line 9, **Propagate** transforms ϕ to

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

which retains the correct model count.

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

3.5 Generating C++ Code

In this section, we will describe the final step of GANTRY as outlined in Figure 1, i.e., translating the set of equations \mathcal{E} into C++ code. Recall that this step is crucial for the usability of the algorithm; otherwise, function definitions would remain purely mathematical, with no convenient way to compute the model count for particular domain sizes. Once a C++ program is produced, it can be executed with different command-line arguments to determine the model count of the sentence for various domain sizes.

See Algorithm 4 for the typical structure of a generated C++ program. Each equation in \mathcal{E} turns into a C++ function with a separate cache for memoisation. Hence, Algorithm 4 has a function and a cache for $f(\cdot, \cdot)$, $g(\cdot, \cdot)$, $g(\cdot, 0)$, and $g(0, \cdot)$. The implementation of an equation consists of three parts. First (on line 5), we check if the arguments already exist in the corresponding cache. If they do, we return the cached value. Second (on lines 6 and 7), we check if the arguments match any of the base cases (as defined in Section 2.4). If so,

we redirect the arguments to the C++ function for that base case. Finally, if none of the above cases applies, we evaluate the arguments based on the expression on the RHS of the equation, store the result in the cache, and return it.

4 Experimental Evaluation

Our empirical evaluation sought to compare the runtime performance of GANTRY with the current state of the art, namely FASTWFOMC² [21, 23] and FORCLIFT³. Our experiments involved two versions of GANTRY: GANTRY-GREEDY and GANTRY-BFS. Like its predecessor (see Section 2.3.2), GANTRY has two modes for applying compilation rules to sentences: one that uses a greedy search algorithm similar to FORCLIFT and another that combines greedy and BFS.

The experiments used an Intel Skylake 2.4 GHz CPU with 188 GiB of memory and CentOS 7. We used the Intel C++ Compiler 2020u4 for C++ programs, Julia 1.10.4 for FASTWFOMC, and the Java Virtual Machine 1.8.0_201 for FORCLIFT and GANTRY. Although implemented in different languages, GANTRY and FASTWFOMC use the same GNU Multiple Precision Arithmetic Library for arbitrary-precision arithmetic.

We ran each algorithm on each benchmark using domains of size $2^1, 2^2, 2^3$, and so on until an algorithm failed to handle a domain size due to a timeout (of one hour) or out-of-memory or out-of-precision errors. While we separately measured compilation and inference time, we primarily focus on total runtime, dominated by the latter. We verified the accuracy of the numerical answers using the corresponding integer sequences in the On-Line Encyclopedia of Integer Sequences [15].

4.1 Benchmarks

We compare these algorithms using three benchmarks from previous work. The first benchmark is the bijection-counting problem from Example 2. The second benchmark is a variant of the well-known *Friends & Smokers* Markov logic network [18, 26], which takes the form of

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

In this sentence, we have three predicates, S , F , and C , that denote smoking, friendship, and cancer, respectively. The first clause states that friends of smokers are also smokers, and the second clause asserts that smoking causes cancer. Common additions to this sentence include making the friendship relation symmetric and assigning probabilities to each clause. Finally, we include the function-counting problem [6]

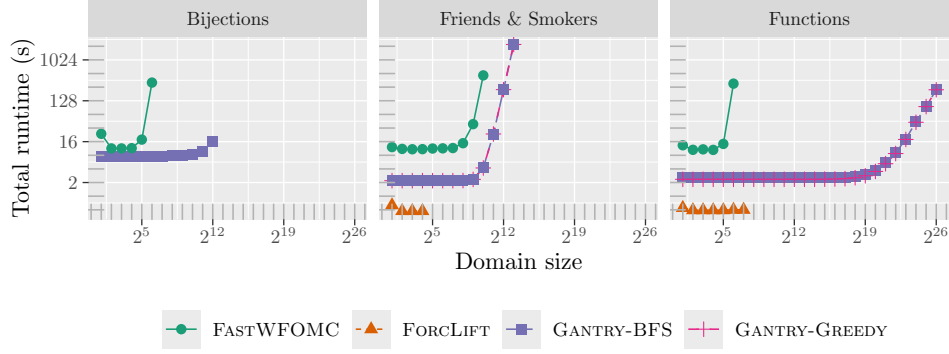
$$(\forall x \in \Gamma. \exists y \in \Delta. P(x, y)) \wedge (\forall x \in \Gamma. \forall y, z \in \Delta. P(x, y) \wedge P(x, z) \Rightarrow y = z) \quad (7)$$

as our third benchmark. Here, predicate P represents a function from Γ to Δ . The first clause asserts that each x must have at least one corresponding y , while the second clause ensures the uniqueness of such a y .

► **Remark.** We formulate the *Bijections* and *Functions* benchmarks using two domains, Γ and Δ , as this formulation is known to help FOKC algorithms find efficient solutions [6]. To compare GANTRY and FORCLIFT with FASTWFOMC, which has no support for multiple domains, we set $|\Gamma| = |\Delta|$.

² <https://github.com/jan-toth/FastWFOMC.jl>

³ <https://github.com/UCLA-StarAI/Forclift>



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

The three benchmarks cover a wide range of possibilities. The *Friends & Smokers* benchmark uses multiple predicates and only two variables without employing any constructs specific to only one of the (W)FOMC algorithms. The *Functions* benchmark, on the other hand, can still be handled by all the algorithms, but its formulation relies on algorithm-specific constructs. Namely, for FORCLIFT and GANTRY, the sentence is written using three variables. Since FASTWFOMC supports at most two variables, we rewrite the sentence (7) with two variables and a *cardinality constraint* (see the technical appendix for details about translating the benchmarks to other first-order logics). Lastly, the *Bijections* benchmark is an example of a sentence that FASTWFOMC can handle but FORCLIFT cannot.

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, since it can only handle model counts smaller than 2^{31} , it only scales up to domain sizes of 16 and 128 for *Friends & Smokers* and *Functions*, respectively. FASTWFOMC outperformed FORCLIFT in the case of *Friends & Smokers*, 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.

One might notice that the runtime of FASTWFOMC and FORCLIFT is slightly higher for the smallest domain size. This peculiarity is the consequence of *just-in-time* (JIT) compilation. As GANTRY is only run once per benchmark, we include the JIT compilation time in its overall runtime across all domain sizes. Additionally, while the compilation time of FORCLIFT is generally lower compared to GANTRY, neither significantly affects the overall runtime. Specifically, FORCLIFT compilation typically takes around 0.5s, while GANTRY compilation takes around 2.3s.

Based on our experiments, which algorithm should one use in practice? If FORCLIFT can handle the sentence and the domain sizes are reasonably small, it is likely the fastest algorithm. In other situations, GANTRY will likely be significantly more efficient than FASTWFOMC regardless of domain size, provided both algorithms can handle the sentence.

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 sentences than FORCLIFT. The ability to efficiently handle large domain sizes is particularly crucial in the weighted setting, as illustrated by the *Friends & Smokers* example, where the model captures complex social networks with probabilistic relationships. Without this scalability, these models would have limited practical value.

Future directions for research include conducting a comprehensive experimental comparison of FOMC algorithms to better understand their comparative performance across various sentences. The capabilities of GANTRY could also be characterised theoretically, for example, by proving completeness for logic fragments liftable by other algorithms [11, 20, 24]. 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 sentences.

References

- 1 Damiano Azzolini and Fabrizio Riguzzi. Lifted inference for statistical statements in probabilistic answer set programming. *Int. J. Approx. Reason.*, 163:109040, 2023. doi:10.1016/J.IJAR.2023.109040.
- 2 Jáchym Barvínek, Timothy van Bremen, Yuyi Wang, Filip Zelezný, and Ondřej Kuželka. Automatic conjecturing of P-recursions using lifted inference. In *ILP*, volume 13191 of *Lecture Notes in Computer Science*, pages 17–25. Springer, 2021. doi:10.1007/978-3-030-97454-1_2.
- 3 Paul Beame, Guy Van den Broeck, Eric Gribkoff, and Dan Suciu. Symmetric weighted first-order model counting. In *PODS*, pages 313–328. ACM, 2015. doi:10.1145/2745754.2745760.
- 4 Mark Chavira and Adnan Darwiche. On probabilistic inference by weighted model counting. *Artif. Intell.*, 172(6-7):772–799, 2008. doi:10.1016/J.ARTINT.2007.11.002.
- 5 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. doi:10.3166/JANCL.11.11-34.
- 6 Paulius Dilkas and Vaishak Belle. Synthesising recursive functions for first-order model counting: Challenges, progress, and conjectures. In *KR*, pages 198–207, 2023. doi:10.24963/KR.2023/20.
- 7 Vibhav Gogate and Pedro M. Domingos. Probabilistic theorem proving. *Commun. ACM*, 59(7):107–115, 2016. doi:10.1145/2936726.
- 8 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. URL: <http://sites.computer.org/debull/A14sept/p6.pdf>.
- 9 Manfred Jaeger and Guy Van den Broeck. Liftability of probabilistic inference: Upper and lower bounds. In *StarAI@UAI*, 2012. URL: <https://starai.cs.kuleuven.be/2012/accepted/jaeger.pdf>.
- 10 Kristian Kersting. Lifted probabilistic inference. In *ECAI*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 33–38. IOS Press, 2012. doi:10.3233/978-1-61499-098-7-33.
- 11 Ondřej Kuželka. Weighted first-order model counting in the two-variable fragment with counting quantifiers. *J. Artif. Intell. Res.*, 70:1281–1307, 2021. doi:10.1613/JAIR.1.12320.
- 12 Nils J. Nilsson. Probabilistic logic. *Artif. Intell.*, 28(1):71–87, 1986. doi:10.1016/0004-3702(86)90031-7.

- 525 13 Feng Niu, Christopher Ré, AnHai Doan, and Jude W. Shavlik. Tuffy: Scaling up statistical
526 inference in Markov logic networks using an RDBMS. *Proc. VLDB Endow.*, 4(6):373–384,
527 2011. doi:10.14778/1978665.1978669.
- 528 14 Vilém Novák, Irina Perfilieva, and Jiri Mockor. *Mathematical principles of fuzzy logic*, volume
529 517. Springer Science & Business Media, 2012.
- 530 15 OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2025. Published
531 electronically at <http://oeis.org>.
- 532 16 Fabrizio Riguzzi, Elena Bellodi, Riccardo Zese, Giuseppe Cota, and Evelina Lamma. A
533 survey of lifted inference approaches for probabilistic logic programming under the distribution
534 semantics. *Int. J. Approx. Reason.*, 80:313–333, 2017. doi:10.1016/J.IJAR.2016.10.002.
- 535 17 Dragan Z. Šaletić. Graded logics. *Interdisciplinary Description of Complex Systems: INDECS*,
536 22(3):276–295, 2024.
- 537 18 Parag Singla and Pedro M. Domingos. Lifted first-order belief propagation. In *AAAI*, pages
538 1094–1099. AAAI Press, 2008. URL: [http://www.aaai.org/Library/AAAI/2008/aaai08-173.](http://www.aaai.org/Library/AAAI/2008/aaai08-173.php)
539 php.
- 540 19 Martin Svatos, Peter Jung, Jan Tóth, Yuyi Wang, and Ondřej Kuželka. On discovering
541 interesting combinatorial integer sequences. In *IJCAI*, pages 3338–3346. ijcai.org, 2023.
542 doi:10.24963/IJCAI.2023/372.
- 543 20 Jan Tóth and Ondřej Kuželka. Lifted inference with linear order axiom. In *AAAI*, pages
544 12295–12304. AAAI Press, 2023. doi:10.1609/AAAI.V37I10.26449.
- 545 21 Jan Tóth and Ondřej Kuželka. Complexity of weighted first-order model counting in the
546 two-variable fragment with counting quantifiers: A bound to beat. In *KR*, 2024. doi:
547 10.24963/KR.2024/64.
- 548 22 Pietro Totis, Jesse Davis, Luc De Raedt, and Angelika Kimmig. Lifted reasoning for combina-
549 torial counting. *J. Artif. Intell. Res.*, 76:1–58, 2023. doi:10.1613/JAIR.1.14062.
- 550 23 Timothy van Bremen and Ondřej Kuželka. Faster lifting for two-variable logic using cell
551 graphs. In *UAI*, volume 161 of *Proceedings of Machine Learning Research*, pages 1393–1402.
552 AUAI Press, 2021. URL: <https://proceedings.mlr.press/v161/bremen21a.html>.
- 553 24 Timothy van Bremen and Ondřej Kuželka. Lifted inference with tree axioms. *Artif. Intell.*,
554 324:103997, 2023. doi:10.1016/J.ARTINT.2023.103997.
- 555 25 Guy Van den Broeck. On the completeness of first-order knowledge compilation for lifted
556 probabilistic inference. In *NIPS*, pages 1386–1394, 2011. URL: [https://proceedings.neurips.](https://proceedings.neurips.cc/paper/2011/hash/846c260d715e5b854ffad5f70a516c88-Abstract.html)
557 cc/paper/2011/hash/846c260d715e5b854ffad5f70a516c88-Abstract.html.
- 558 26 Guy Van den Broeck, Arthur Choi, and Adnan Darwiche. Lifted relax, compensate and then
559 recover: From approximate to exact lifted probabilistic inference. In *UAI*, pages 131–141.
560 AUAI Press, 2012. URL: [https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&](https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=2349&proceeding_id=28)
561 smnu=2&article_id=2349&proceeding_id=28.
- 562 27 Guy Van den Broeck, Wannes Meert, and Adnan Darwiche. Skolemization for weighted
563 first-order model counting. In *KR*. AAAI Press, 2014. URL: [http://www.aaai.org/ocs/](http://www.aaai.org/ocs/index.php/KR/KR14/paper/view/8012)
564 index.php/KR/KR14/paper/view/8012.
- 565 28 Guy Van den Broeck, Nima Taghipour, Wannes Meert, Jesse Davis, and Luc De Raedt. Lifted
566 probabilistic inference by first-order knowledge compilation. In *IJCAI*, pages 2178–2185.
567 IJCAI/AAAI, 2011. doi:10.5591/978-1-57735-516-8/IJCAI11-363.
- 568 29 Yuanhong Wang, Juhua Pu, Yuyi Wang, and Ondřej Kuželka. Lifted algorithms for symmetric
569 weighted first-order model sampling. *Artif. Intell.*, 331:104114, 2024. doi:10.1016/J.ARTINT.
570 2024.104114.