

Effective Interpolation and Preservation in Guarded Logics^{*}

Michael Benedikt

University of Oxford
michael.benedikt@cs.ox.ac.uk

Balder ten Cate

LogicBlox and UC Santa Cruz
btencate@ucsc.edu

Michael Vanden Boom

University of Oxford
michael.vandenboom@cs.ox.ac.uk

Abstract

Desirable properties of a logic include decidability, and a model theory that inherits properties of first-order logic, such as interpolation and preservation theorems. It is known that the Guarded Fragment (GF) of first-order logic is decidable and satisfies some preservation properties from first-order model theory; however, it fails to have Craig interpolation. The Guarded Negation Fragment (GNF), a recently-defined extension, is known to be decidable and to have Craig interpolation. Here we give the first results on effective interpolation for extensions of GF. We provide an interpolation procedure for GNF whose complexity matches the doubly exponential upper bound for satisfiability of GNF. We show that the same construction gives not only Craig interpolation, but Lyndon interpolation and Relativized interpolation, which can be used to provide effective proofs of some preservation theorems. We provide upper bounds on the size of GNF interpolants for both GNF and GF input, and complement this with matching lower bounds.

Categories and Subject Descriptors Theory of computation [Logic]

1. Introduction

Basic modal logic is known to have many attractive properties, both computationally and in its model theory. Given the limited expressiveness of basic modal logic, logicians have searched for more powerful languages that extend it while preserving its positive features.

Andréka et. al. ([2], 4.7) list “meta-properties” of modal logic that one would desire in an extension, including the following.

- The logic should have decidable satisfiability and validity problems.
- The logic should satisfy *interpolation theorems*, like the *Craig Interpolation Property* (CIP): if φ and ψ are in the logic, and φ entails ψ , then there should be a formula θ in the logic such that φ entails θ and θ entails ψ , and θ mentions only relations common to both φ and ψ .

^{*}Benedikt and Vanden Boom are supported by EPSRC grant EP/H017690/1 (The Engineering and Physical Sciences Research Council, UK), and ten Cate is supported by NSF grant IIS-1217869.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CSL-LICS 2014, July 14–18, 2014, Vienna, Austria.
Copyright © 2014 ACM 978-1-4503-2886-9...\$15.00.
<http://dx.doi.org/10.1145/2603088.2603108>

- The logic should satisfy *preservation theorems* such as the Łoś-Tarski theorem: if a sentence in the logic is preserved under superstructures, it should be equivalent to an existential sentence in the logic.
- The logic should have the *finite model property*: a satisfiable sentence should have a finite model.

Interpolation, in particular, is a desirable property since it implies the well-known *Beth Definability Property* (BDP), as well as its extension, the *Projective Beth Definability Property* (PBDP), stating that implicit specifications can be converted to explicit ones. This conversion from implicit to explicit is important in knowledge representation [24] and databases [20]. Interpolation has many applications, both in simplifying definitions and in verification [18].

Andréka et. al. proposed the *Guarded Fragment* (GF) as a candidate for an attractive logic. They showed that it is decidable, that it has the finite model property, and that it satisfies Łoś-Tarski. Unfortunately Hoogland, Marx, and Otto [14] showed that GF does not satisfy CIP. Later work [7] showed that PBDP fails for GF.

The *Guarded Negation Fragment* (GNF) is an extension of GF introduced by Bárány, ten Cate and Segoufin in [5]. It also extends the *Unary Negation Fragment* (UNF) introduced by ten Cate and Segoufin in [23]. Both UNF and GNF were proven decidable and shown to have the finite model property [5, 23]. They were also shown to satisfy many of the properties mentioned earlier: in [23], UNF was shown to have CIP and PBDP; in [7], GNF was shown to have CIP, PBDP, and several preservation properties including the analog of Łoś-Tarski. The combination of interpolation and the finite model property for these logics means that interpolation holds when validity is considered over finite models only. This is attractive for applications of these logics in databases, such as rewriting queries using views [20], since in databases typically only finite models are considered.

In this work we provide further evidence that GNF and UNF are attractive extensions of modal logic. We show that these logics satisfy the Lyndon interpolation property (LIP) and its generalization, the Relativized interpolation property (RIP), and moreover present the first constructive procedures for GNF and UNF interpolation, including CIP, LIP, and RIP. Roughly speaking, the LIP says that if φ entails ψ , we can find θ as in CIP, but with a relation symbol occurring positively in θ iff it occurs positively in both φ and ψ . The RIP says that the quantification pattern in θ must match the quantification pattern in φ and ψ .

We proceed by extending the *mosaic method* that has been applied in providing decidability for algebras of relations [21], modal logics [19], and temporal logics [17]. We present a variant of the method that gives a sound and complete decision procedure for the validity problem for GNF, but also allows interpolants to be read off from a proof of an entailment.

We show how this algorithm gives interpolants that witness the RIP and hence the LIP and CIP for GNF and UNF. Since the methods of [7, 23] were model-theoretic, this gives the first constructive proof of interpolation for GNF and UNF. We deduce

bounds on the size of interpolants for GNF, and also on the size of GNF interpolants for GF formulas. We use the RIP to prove that GNF and GF admit effective preservation theorems, which state that formulas satisfying certain semantic properties can be converted to specific forms. To the best of our knowledge, these are the first effective preservation theorems, even for GF.

We complement the upper bounds with lower bounds for the size of GNF interpolants for GF and UNF.

Due to space limitations, most proofs are deferred to the full version of this paper.

2. Preliminaries

Notation and conventions. We use $\mathbf{x}, \mathbf{y}, \dots$ to denote vectors of variables. For a formula ϕ , we write $\phi(\mathbf{x})$ to indicate that the free variables in ϕ are contained in $\mathbf{x}$, and write $\text{free}(\phi)$ for the actual free variables. The formula ϕ may also use constants $\text{con}(\phi)$. We use $\alpha, \beta, \dots$ to denote atomic formulas, and if we write $\alpha(\mathbf{x})$ then we assume that $\text{free}(\alpha) = \mathbf{x}$.

Classical query languages. We will make use of some basic notions of database theory – in particular the following “classical” query classes: conjunctive queries (CQs), unions of conjunctive queries (UCQs), first-order logic formulas (FO), existential FO, and positive existential FO. Abiteboul, Hull, and Vianu [1] is a good reference for these languages.

The basics of GNF and GF. The Guarded Negation Fragment (GNF) is built up inductively according to the grammar:

$$\phi ::= R\mathbf{t} \mid \exists \mathbf{x}.\phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid \alpha(\mathbf{x}) \wedge \neg\phi(\mathbf{x})$$

where R is either a relation symbol or the equality relation $x = y$, α is an atomic relation (including equality), and $\mathbf{t}$ is a tuple over variables and constants. Notice that any use of negation must occur conjoined with an atomic relation that contains all the free variables of the negated formula. Such an atomic relation is a *guard* of the formula.

The purpose of allowing equalities as guards is to ensure that every formula with at most one free variable can always be guarded. Thus GNF includes the Unary Negation Fragment (UNF), which is built up as above, but allowing negation only on formulas with at most one free variable.

A GNF formula is in *GN-normal form* if in its syntax tree, no disjunction is directly below an existential quantifier or a conjunction, and no existential quantifier is directly below a conjunction sign. Such GN-normal form formulas ϕ can be generated by the following recursive definition:

$$\begin{aligned} \phi &::= \bigvee_i \exists \mathbf{x}.\bigwedge_j \psi_{ij} \\ \psi &::= R\mathbf{t} \mid \alpha(\mathbf{x}) \wedge \neg\phi(\mathbf{x}) \end{aligned}$$

where R is either a relation symbol or the equality relation, α is an atomic relation (including equality), and $\mathbf{t}$ is a tuple over variables and constants. Thus, GN-normal form formulas can be viewed as formulas built up from atomic formulas using UCQs and guarded negation. Every GNF formula can be brought into GN-normal form, at the cost of an exponential increase in length and linear increase in the number of variables.

GNF should be compared to the *Guarded Fragment* (GF), typically defined via the grammar:

$$\phi ::= R\mathbf{t} \mid \exists \mathbf{x}.\alpha(\mathbf{x}\mathbf{y}) \wedge \phi(\mathbf{x}\mathbf{y}) \mid \phi \vee \phi \mid \phi \wedge \phi \mid \neg\phi(\mathbf{x})$$

where R is either a relation symbol or the equality relation, α is an atomic relation (including equality), and $\mathbf{t}$ is a tuple over variables and constants. Here it is the quantification that is guarded, rather than negation. As in GNF, we allow equality guards by default.

It is easy to see that every union of conjunctive queries is expressible in GNF. It is only slightly more difficult to verify that every GF sentence can be expressed in GNF [5].

Turning to fragments of first-order logic that are common in database theory, one can show that common classes of integrity constraints, such as inclusion dependencies (also known as referential constraints) are expressible in GNF, as well as many of the common dependencies used in data integration and exchange. Going further, the relational translations of many of the common description logic languages used in the semantic web (e.g. *ALC* and *ALCHIO* [4]) are known to be expressible in GF, and hence in GNF.

Not only is GNF an expressive fragment of FO, but it was shown to be decidable and to have the finite model property.

Theorem 1 ([5]). *A GNF sentence is satisfiable over all models iff it is satisfiable over finite models. Satisfiability can be tested effectively (and is 2EXPTIME-complete).*

Size of formulas. A formula ϕ given as input (to interpolation, preservation, etc.) is assumed to be in the standard tree representation of a formula, and the size of this representation is the number of symbols in ϕ .

For the output, however, we will usually represent the formula and measure its size as a node-labelled DAG (directed acyclic graph). The nodes represent formulas, and the edge relation connects a formula to its subformulas. We will also use DAGs with multiple roots to represent a set of formulas.

We often abuse notation by identifying a formula with a representation; given a representation of a formula ϕ , $|\phi|$ represents the size of the representation: the number of symbols if it is a formula, and the number of edges and labels if the representation is a DAG. Similarly, for a set S of formulas represented as a DAG, we write $|S|$ to denote the total number of nodes and edges in the DAG.

The *width* of ϕ , denoted $\text{width}(\phi)$, is the maximum number of free variables of any subformula of ϕ . For GF sentences, the width is bounded by the maximum arity of any relation.

Interpolation. For two formulas φ_L and φ_R , we write $\varphi_L \models \varphi_R$ if every model of φ_L is a model of φ_R , and in this case say that $\varphi_L \models \varphi_R$ is a *validity*. An *interpolant* for a validity is a formula θ such that $\varphi_L \models \theta \models \varphi_R$ and θ mentions only relations present in both φ_L and φ_R . The Craig interpolation theorem [8] states that if φ_L and φ_R are in FO, then there is an FO interpolant θ .

Lyndon proved a stronger interpolation property by considering the polarity of relations in the formulas. Given a formula ϕ , we define $\text{occ}(\phi)$ to be the set describing the *polarity of occurrences of atomic relations* (not including the equality relation). A relation occurs positively (respectively, negatively) in a formula if it is within the scope of an even (respectively, odd) number of negations. For every atomic relation S in ϕ (where S is not the equality relation), $(S, +) \in \text{occ}(\phi)$ (respectively, $(S, -) \in \text{occ}(\phi)$) iff S occurs positively (respectively, negatively) in ϕ . Lyndon’s interpolation theorem [15] states that the interpolant θ can be taken so that $\text{occ}(\theta) \subseteq \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$.

A more recent relativized interpolation theorem of Otto [22] states that if φ_L and φ_R both use only quantification *relativized to a set of unary predicates* $\mathbb{U}$, then the same is true of the interpolant. A relativized quantifier for $\mathbb{U}$ is $\exists \mathbf{x}. U_1 x_1 \wedge \dots \wedge U_k x_k \wedge \psi(\mathbf{x}\mathbf{y})$ or $\forall \mathbf{x}. U_1 x_1 \wedge \dots \wedge U_k x_k \rightarrow \psi(\mathbf{x}\mathbf{y})$ where $\mathbf{x} = x_1 \dots x_k$ and $U_i \in \mathbb{U}$ for $i \in \{1, \dots, k\}$. As shown in [22], the relativized interpolation theorem gives an alternative means of showing preservation theorems. We explain and expand upon this in Section 7.

Although there are proofs of Craig interpolation for FO that construct interpolants efficiently given proofs that witness the validity (e.g. [10]), these do not give bounds on the size of interpolants relative to the size of φ_L and φ_R . In fact, it has been shown that for

FO validities, there is no computable bound on the sizes of the interpolants relative to the sizes of the original validities [11].

We say that a logic has the Craig Interpolation Property (CIP) if given φ_L and φ_R in the logic, there is an interpolant that is again in the logic. Similarly we can talk about the Lyndon Interpolation Property (LIP) and Relativized Interpolation Property (RIP). Hoogland et. al [14] showed that the CIP fails for GF. Marx [16] argued that one of the applications of CIP, the Projective Beth Definability Property (PBDP) holds for GF as well as for an extension called the “Packed Fragment”. He conjectured that the proof could be made effective using the “mosaic method”. The argument in [16] was flawed, and indeed neither the guarded fragment nor the packed fragment has PBDP [7]. Nevertheless, here we use the mosaic technique suggested by Marx to show effective CIP, LIP, RIP, and PBDP for GNF (see Theorem 9).

3. The Mosaic Method

In its simplest form, a mosaic is just a finite set of formulas. When it satisfies certain coherence properties, it can be viewed as a collection of formulas describing part of a model. For instance, in the mosaic system for GF described in Section 4.6 of [2], a mosaic can describe the complete subformula type of some guarded set of elements in a model.

For modal and guarded logics, it turns out that the existence of a (finite) set of mosaics satisfying certain coherence properties is enough to ensure the existence of a full model. This allows mosaics to be used to decide satisfiability.

We adapt this method to check whether $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, and if so, extract an interpolant for $\varphi_L \models \varphi_R$ from the mosaics. The purpose of this section is to describe the general framework that we are using. We utilize this framework in later sections with different versions of mosaics depending, for instance, on whether the input formulas are in GF versus GNF, or whether we are aiming for Lyndon interpolation versus Relativized interpolation. As a result, some of the terms in this section are only defined informally; the formal definitions appear in later sections since they depend on the particular mosaic variant that we are using.

We start by fixing some sentences φ_L and φ_R over signatures σ_L and σ_R , respectively. The first step is to define a finite set Γ of formulas that are relevant to checking satisfiability of $\varphi_L \wedge \neg\varphi_R$. This is usually some extension of the subformula closure of φ_L and φ_R . To be suitable for computing an interpolant for $\varphi_L \models \varphi_R$, we also annotate these formulas with a *provenance* $X \in \{L, R\}$ indicating that the formulas are related to φ_X . This parallels other constructive interpolation methods, e.g. [10]. Given a provenance X , we write $\bar{X}$ for the *dual provenance* (i.e. $\bar{L} = R$ and $\bar{R} = L$). Throughout the paper, we use X, Y, Z to range over provenances L and R .

Given some finite set of *parameters* c , we define $\Gamma(c)$ as the set of formulas from Γ with parameters from c substituted for any free variables. Parameters are just a special type of variable. We write $\mathbf{a}, \mathbf{b}, \dots$ to denote vectors of parameters, and given a formula ϕ , we write $\phi(\mathbf{x}\mathbf{b})$ to indicate that the free variables in ϕ are contained in $\mathbf{x}$ and the parameters in ϕ are contained in $\mathbf{b}$. We write $\text{free}(\phi)$ (respectively, $\text{par}(\phi)$) for the actual free variables (respectively, parameters). For atomic α , if we write $\alpha(\mathbf{x}\mathbf{b})$ then we assume that $\text{free}(\alpha) = \mathbf{x}$ and $\text{par}(\alpha) = \mathbf{b}$.

A *mosaic* τ over some finite set of parameters c (of size bounded by the width of the original sentences) is then defined to be a subset of formulas from $\Gamma(c)$. Each mosaic is annotated with L or R , and an X -mosaic has additional restrictions on the $\bar{X}$ -formulas appearing in it. This helps prove properties about the interpolants that we later extract from these mosaics.

Mosaics can be *linked* together. For instance, if a formula $Y : \exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b})$ appears in an X -mosaic τ , then this can be linked to some Y -mosaic τ' using parameters $\mathbf{a}\mathbf{b}$ and containing the formula $Y : \psi(\mathbf{a}\mathbf{b})$ (for some new parameters $\mathbf{a}$). In addition, the linked mosaic τ' must contain the maximum amount of information from τ about formulas using only shared parameters $\mathbf{b}$ (subject to any constraints on formulas in τ') and cannot contain any new information about formulas using only shared parameters. We say an existential requirement $Y : \exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b})$ in τ is *fulfilled* if there is a formula $Y : \psi(\mathbf{a}\mathbf{b})$ in τ , or there is some τ' such that τ and τ' are linked via $Y : \exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b})$.

In order for a mosaic to correspond to a part of a model it must satisfy certain coherence properties. Within a single mosaic, this means that it must be *internally consistent*. For instance, we may want to require that if it contains a conjunction, then it contains both conjuncts. The exact definition will depend on the particular version of mosaics that we are considering. A set of mosaics is coherent if every mosaic is internally consistent and it is *saturated*: every existential requirement is fulfilled by a mosaic in the set.

We then fix some finite set $\mathcal{M}$ of mosaics that suffices to prove the following key *mosaic theorem*:

$\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is a saturated set of internally consistent mosaics from $\mathcal{M}$ containing a mosaic with φ_L and $\neg\varphi_R$.

In one direction, the idea is to extract a saturated set of internally consistent mosaics from a model. In the other (harder) direction, we must show that a saturated set of internally consistent mosaics from $\mathcal{M}$ can be used to construct a model of $\varphi_L \wedge \neg\varphi_R$.

This gives us a procedure for deciding whether or not $\varphi_L \wedge \neg\varphi_R$ is satisfiable. Consider the following *mosaic elimination procedure* over $\mathcal{M}$.

- At stage $i = 1$, every internally inconsistent mosaic is eliminated.
- At stage $i > 1$, a mosaic τ is eliminated if it has an existential requirement that is not fulfilled in τ , and every linked τ' fulfilling it was already eliminated.

This process continues until a fixpoint is reached. The mosaic theorem can then be used to show that $\varphi_L \wedge \neg\varphi_R$ is satisfiable iff there is a mosaic containing φ_L and $\neg\varphi_R$ that is never eliminated.

For interpolation, we are interested in the case when $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, since this means $\varphi_L \models \varphi_R$.

It turns out that we can construct a *mosaic interpolant* θ_X^c for each X -mosaic τ eliminated using the procedure above. These mosaic interpolants satisfy the property that $\tau_X \models \theta_X^c$ and $\theta_X^c \models \neg\tau_{\bar{X}}$, where τ_Y denotes the conjunction of all of the Y -formulas from $\Gamma(c)$ in τ and the negation of all of the Y -formulas from $\Gamma(c)$ that satisfy the constraints on formulas in τ but are not in τ .

We start by constructing these interpolants for mosaics eliminated because of internal inconsistencies. The interpolants for such mosaics are simple: atomic formulas or the negation of atomic formulas. Proceeding by induction on the stage at which the mosaic is eliminated, we then build up more complicated interpolants for the other mosaics. Intuitively, each mosaic interpolant describes the reason the mosaic was eliminated.

Since the sequence of elimination steps for the mosaics provide a proof that $\varphi_L \wedge \neg\varphi_R$ is unsatisfiable, it is not surprising that the desired interpolant for $\varphi_L \models \varphi_R$ can then be constructed from these mosaic interpolants.

It is helpful to compare this to other constructive interpolation approaches. In [10], an interpolant is constructed from a proof tree for $\varphi_L \wedge \neg\varphi_R$ in a tableau proof system for FO. The construction starts at the leaves (where very simple interpolants suffice), and builds up to an interpolant that holds for the original validity at the root of the proof tree. Because of the “one-step” nature of the proof

rules, there are straightforward interpolation construction rules at each stage.

Using the mosaic method, the definition of the interpolant at the inductive stages is more complicated. Unlike a traditional proof system, however, the mosaics allow precise control over the parameters appearing at each stage. This is advantageous for constructing an interpolant in GNF, since we need to be able to guard certain parameters in order to remain in GNF. Indeed, this precise control over the parameters was one of the motivations for building GNF interpolants from mosaics rather than from a more traditional proof system.

4. GF Mosaics

We start by describing the mosaic method for validity testing and interpolation construction when the input consists of GF sentences φ_L and φ_R over signatures σ_L and σ_R , respectively. The purpose of this section is to give a more gentle introduction to mosaics, without the complications that come from handling GNF, relativized quantifiers, or equality.

We assume the sentences φ_L and $\neg\varphi_R$ are in GF and are in negation normal form (negations are pushed to the inside) and do not use equality. These sentences may use constants, and we let $e := \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$.

Closure. For GF satisfiability, we restrict formulas in mosaics to the subformula closure of φ_L and $\neg\varphi_R$, with a slight extension to handle atoms. The *closure* $\text{cl}(\varphi)$ of a GF formula φ is the smallest set C of formulas containing φ such that:

- if $\psi_1 \wedge \psi_2 \in C$ or $\psi_1 \vee \psi_2 \in C$, then $\psi_1, \psi_2 \in C$;
- if $\exists x.\alpha(\mathbf{x}\mathbf{y}) \wedge \psi(\mathbf{x}\mathbf{y}) \in C$, then $\alpha(\mathbf{x}\mathbf{y}), \psi(\mathbf{x}\mathbf{y}) \in C$;
- if $\forall x.\alpha(\mathbf{x}\mathbf{y}) \rightarrow \psi(\mathbf{x}\mathbf{y}) \in C$, then $\neg\alpha(\mathbf{x}\mathbf{y}), \psi(\mathbf{x}\mathbf{y}) \in C$;
- if $\alpha(\mathbf{x}\mathbf{y}) \in C$, then $\exists x.\alpha(\mathbf{x}\mathbf{y}) \in C$.

Given parameters c , $\text{cl}_L(c)$ (respectively, $\text{cl}_R(c)$) consists of formulas from $\text{cl}(\varphi_L)$ (respectively, $\text{cl}(\neg\varphi_R)$) with free variables replaced by parameters from c or constants from e and labelled with provenance L (respectively, R).

Observe that the formulas in $\text{cl}_L(c) \cup \text{cl}_R(c)$ respect the polarity of relations in the original formulas. This will be important for extracting Lyndon-style interpolants later, where the polarity of relations needs to be respected.

Guards. We say α is an L-atom (respectively, R-atom) if $(\alpha, +) \in \text{occ}(\varphi_L)$ (respectively, $(\alpha, -) \in \text{occ}(\varphi_R)$). Usually, any atom is a possible guard in a GF or GNF formula. For the purposes of interpolation, however, we need *atoms in the common signature* (atoms for short). We say α is an L-catom (respectively, R-catom) if $(\alpha, +) \in \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$ (respectively, $(\alpha, -) \in \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$).

Given a collection τ of formulas from $\text{cl}_L(c) \cup \text{cl}_R(c)$, we say $b \subseteq c$ is *X-guarded* in τ (respectively, *X-cguarded* in τ) if b is empty, or if there is some $X : \exists x.\alpha(\mathbf{x}b) \in \tau$ where α is an X-atom (respectively, X-catom). We allow x to be empty, in which case the X-guard or X-cguard for b is actually of the form $X : \alpha(b)$.

GF mosaics. A GF *X-mosaic* τ for $\varphi_L \wedge \neg\varphi_R$ using parameters c is a subset of $\Gamma(c) := \text{cl}_L(c) \cup \text{cl}_R(c)$ such that c is X-guarded in τ , and if $\tilde{X} : \psi \in \tau$, then $\text{par}(\psi)$ is X-cguarded in τ . We write $\tau(c)$ to emphasize that τ uses parameters c .

Each X-mosaic provides information about an X-guarded set of parameters. Unlike the mosaics in [2], we place further restrictions on the formulas in the mosaics; namely, we restrict $\tilde{X}$ -formulas in an X-mosaic. Roughly speaking, the $\tilde{X}$ -formulas are the “dangerous” formulas that could potentially introduce a negation when we try to extract GNF interpolants from these mosaics later, so we require that any such formulas have an X-cguard in the mosaic (a guard that could be used in the interpolant).

Let τ be a GF *X-mosaic* over parameters c . Assume $Y : \eta \in \tau$ is an *existential requirement* i.e. η is of the form $\exists x.\alpha(\mathbf{x}b) \wedge \psi(\mathbf{x}b)$. Then τ is *linked* to τ' via $Y : \eta$ (written $\tau \rightarrow_{Y:\eta} \tau'$) if

- τ' is a GF *Y-mosaic* over parameters $\mathbf{a}b$ with $\mathbf{a} \cap c = \emptyset$;
- $Y : \alpha(\mathbf{x}b)[\mathbf{a}/x] \in \tau'$ and $Y : \psi(\mathbf{x}b)[\mathbf{a}/x] \in \tau'$;
- for all $Y : \psi'(\mathbf{b}) \in \tau$, $Y : \psi'(\mathbf{b}) \in \tau'$;
- for all $\tilde{Y} : \psi'(\mathbf{b}) \in \tau$ such that $\text{par}(\psi')$ is Y-cguarded in τ' , $\tilde{Y} : \psi'(\mathbf{b}) \in \tau'$;
- for all $Z : \psi'(\mathbf{b}) \in \tau'$, $Z : \psi'(\mathbf{b}) \in \tau$.

These conditions ensure that the maximum amount of information about the shared parameters b is passed from τ to τ' (subject to the restrictions on $\tilde{Y}$ formulas in the Y-mosaic τ'), and no new information about b is added in τ' .

Fix some set P of parameters of size $2 \cdot \text{width}(\varphi_L \wedge \neg\varphi_R)$, and let $\mathcal{M}$ be the set of GF mosaics over parameters P . The following bound on the number of mosaics is easy to show.

Proposition 2. *There is a polynomial function p (independent of φ_L and φ_R) such that if $|\varphi_L| + |\varphi_R| = n$, then there are at most $2^{2^{p(n)}}$ mosaics in $\mathcal{M}$. Moreover, for a fixed bound on the arity of relations, there are at most $2^{p(n)}$ mosaics in $\mathcal{M}$.*

Coherence. We say a GF *X-mosaic* τ is *internally consistent* if there is no atomic $\alpha \in \sigma_L \cup \sigma_R$ such that $Y : \alpha, Z : \neg\alpha \in \tau$, and τ satisfies the following *downward closure* properties:

- if $Y : \alpha(\mathbf{a}b) \in \tau$, then $Y : \exists x.\alpha(\mathbf{x}b) \in \tau$;
- if $Y : \exists x.\alpha(\mathbf{x}ab) \in \tau$, then $Y : \exists xy.\alpha(\mathbf{x}yb) \in \tau$;
- if $Y : \psi_1 \wedge \psi_2 \in \tau$, then $Y : \psi_1, Y : \psi_2 \in \tau$;
- if $Y : \psi_1 \vee \psi_2 \in \tau$, then $Y : \psi_1 \in \tau$ or $Y : \psi_2 \in \tau$;
- if $Y : \forall x.\alpha(\mathbf{x}b) \rightarrow \psi(\mathbf{x}b) \in \tau$ and there is $\mathbf{a} \subseteq c \cup e$ such that $Z : \alpha(\mathbf{x}b)[\mathbf{a}/x] \in \tau$ for some $Z \in \{L, R\}$, then $Y : \psi(\mathbf{x}b)[\mathbf{a}/x] \in \tau$.

A set of GF mosaics is said to be *saturated* if for every mosaic $\tau(c)$ in the set and for every existential requirement $Y : \eta \in \tau$ for η of the form $\exists x.\alpha(\mathbf{x}b) \wedge \psi(\mathbf{x}b)$, either (i) $Y : \alpha(\mathbf{x}b)[\mathbf{a}/x]$ and $Y : \psi(\mathbf{x}b)[\mathbf{a}/x]$ are in τ for some $\mathbf{a} \subseteq c \cup e$ (so η is fulfilled in τ) or (ii) τ is linked to some mosaic τ' in the set via $Y : \eta$ (so η is fulfilled in τ'). Finding a saturated set of internally consistent mosaics containing the original sentences provides a sound and complete method for GF satisfiability testing.

Theorem 3. *$\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is a saturated set of internally consistent mosaics from $\mathcal{M}$ that contains a mosaic τ such that $L : \varphi_L, R : \neg\varphi_R \in \tau$.*

We can now use the mosaic elimination procedure described in the previous section, and Theorem 3 ensures that $\varphi_L \wedge \neg\varphi_R$ is satisfiable iff there is some mosaic containing $L : \varphi_L$ and $R : \neg\varphi_R$ that is not eliminated using this process.

Constructive interpolation. Now assume $\varphi_L \models \varphi_R$, and consider the mosaic elimination procedure for the GF mosaics $\mathcal{M}$ for $\varphi_L \wedge \neg\varphi_R$. We write $\mathcal{N}'$ for the set of mosaics for $\varphi_L \wedge \neg\varphi_R$ that are eliminated using this procedure, and $\mathcal{N}_i$ for the set of mosaics that have been removed by stage i . By Theorem 3, we know that every mosaic containing $L : \varphi_L$ and $R : \neg\varphi_R$ must be in $\mathcal{N}'$.

We first show how to construct a mosaic interpolant θ_X^τ for each X-mosaic $\tau(c)$ in $\mathcal{N}'$. Roughly speaking, θ_X^τ is an interpolant (in the usual sense) for $\tau_X \models \neg\tau_{\bar{X}}$, where

$$\tau_X := \bigwedge_{Y:\psi \in \tau} \psi \quad \wedge \quad \bigwedge_{Y:\psi \in \bar{\tau}} \neg\psi$$

and $\bar{\tau}$ is the set consisting of formulas $Z : \psi \in \text{cl}_Y(c) \setminus \tau$ such that $Z = X$, or $Z = \tilde{X}$ and $\text{par}(\psi)$ is X-cguarded in τ . In other words, $\bar{\tau}$ contains the formulas from the closure that satisfy the

requirements for being in τ (in terms of X -cguardedness), but are not in τ . Intuitively, θ_X^τ describes the reason τ was eliminated, i.e. it expresses why $\tau_X \wedge \tau_{\tilde{X}}$ is not satisfiable.

We have the following Lyndon-style mosaic interpolant properties.

Lemma 4. *For each GF X -mosaic $\tau(c) \in \mathcal{N}'$ (the set of eliminated mosaics), we can construct a DAG representation of a formula θ_X^τ such that*

- $\tau_X \models \theta_X^\tau$ and $\theta_X^\tau \models \neg\tau_{\tilde{X}}$;
- $\text{occ}(\theta_X^\tau) \subseteq \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$ if $X = L$,
 $\text{occ}(\neg\theta_X^\tau) \subseteq \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$ if $X = R$;
- $\text{con}(\theta_X^\tau) \subseteq \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$;
- $\text{par}(\theta_X^\tau) \subseteq c$;
- θ_X^τ is in GNF, even when parameters are viewed as free variables.

Moreover, if $n = |\varphi_L| + |\varphi_R|$ then there is a DAG representation of $\Theta_i := \{\theta_X^\tau : \tau \text{ is a } Z\text{-mosaic in } \mathcal{N}_i\}$ such that $|\Theta_i| \leq i \cdot 2^{p'(n)}$ ($|\Theta_i| \leq i \cdot 2^{p'(n)}$ for the fixed arity case), where p' is some polynomial function independent of φ_L and φ_R .

Proof. We first introduce some additional notation. We use $\top$ (respectively, $\perp$) as an abbreviation for the GNF sentence $\exists x.x = x$ (respectively, $\exists x.(x = x \wedge \neg(x = x))$). During the interpolant construction, we will need to include guards in order to remain in GNF, so for any tuple $\mathbf{b}$ of parameters that are X -cguarded in τ , we define a formula $\text{gdd}_X^\tau(\mathbf{b})$. If $\mathbf{b} = \emptyset$, then set $\text{gdd}_X^\tau(\mathbf{b}) := \top$. Otherwise, there is some formula $X : \exists x.\beta(\mathbf{x}\mathbf{b}) \in \tau$ for some X -catom β , and we set $\text{gdd}_X^\tau(\mathbf{b}) := \exists x.\beta(\mathbf{x}\mathbf{b})$ (note that $\mathbf{x}$ can be empty, in which case this is just an atomic formula). We will also write $\text{gdd}_X^\tau(\mathbf{b}) \wedge \neg\psi$ to indicate some formula $\exists x.(\beta(\mathbf{x}\mathbf{b}) \wedge \neg\psi)$.

We proceed by induction on i , the stage at which τ was eliminated. We define the interpolants and then sketch the proof of correctness in the inductive case.

Base case. Consider an X -mosaic $\tau(c) \in \mathcal{N}_1$. Then τ has an internal inconsistency, and we consider the following cases.

- Assume one of these conditions holds for $Y \in \{L, R\}$:
 - $Y : \alpha(\mathbf{b}), Y : \neg\alpha(\mathbf{b}) \in \tau$;
 - $Y : \alpha(\mathbf{a}\mathbf{b}) \in \tau$, but $Y : \exists x.\alpha(\mathbf{x}\mathbf{b}) \notin \tau$;
 - $Y : \exists x.\alpha(\mathbf{x}\mathbf{a}\mathbf{b}) \in \tau$, but $Y : \exists x\mathbf{y}.\alpha(\mathbf{x}\mathbf{y}\mathbf{b}) \notin \tau$;
 - $Y : \psi_1 \wedge \psi_2 \in \tau$, but $Y : \psi_1 \notin \tau$ or $Y : \psi_2 \notin \tau$;
 - $Y : \psi_1 \vee \psi_2 \in \tau$, but $Y : \psi_1 \notin \tau$ and $Y : \psi_2 \notin \tau$;
 - $Y : \alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}], Y : \forall x.\alpha(\mathbf{x}\mathbf{b}) \rightarrow \psi(\mathbf{x}\mathbf{b}) \in \tau$, but $Y : \psi(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}] \notin \tau$.

If $Y = X$, then $\theta_X^\tau := \perp$. If $Y = \tilde{X}$, then $\theta_X^\tau := \top$.

- If $X : \alpha(\mathbf{b}), \tilde{X} : \neg\alpha(\mathbf{b}) \in \tau$, then $\theta_X^\tau := \alpha(\mathbf{b})$.
- If $X : \neg\alpha(\mathbf{b}), \tilde{X} : \alpha(\mathbf{b}) \in \tau$, then $\theta_X^\tau := \text{gdd}_X^\tau(\mathbf{b}) \wedge \neg\alpha(\mathbf{b})$.
- If $X : \alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}], \tilde{X} : \forall x.\alpha(\mathbf{x}\mathbf{b}) \rightarrow \psi(\mathbf{x}\mathbf{b}) \in \tau$, but $\tilde{X} : \psi(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}] \notin \tau$, then $\theta_X^\tau := \alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}]$.
- If $\tilde{X} : \alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}], X : \forall x.\alpha(\mathbf{x}\mathbf{b}) \rightarrow \psi(\mathbf{x}\mathbf{b}) \in \tau$, but $X : \psi(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}] \notin \tau$, then $\theta_X^\tau := \text{gdd}_X^\tau(\mathbf{a}'\mathbf{b}) \wedge \neg\alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}]$, where $\mathbf{a}'\mathbf{b} = \text{par}(\alpha(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}])$.

Inductive case. Now consider an X -mosaic $\tau(c) \in \mathcal{N}_i \setminus \mathcal{N}_{i-1}$. This means there is some $Y : \eta = Y : \exists x.\alpha(\mathbf{x}\mathbf{b}) \wedge \psi(\mathbf{x}\mathbf{b}) \in \tau$ that is not fulfilled in τ , and if $\tau \rightarrow_{Y:\eta} \tau'$ then $\tau' \in \mathcal{N}_{i-1}$.

Assume $\text{par}(\psi) = \mathbf{b} \subseteq c$, and fix some $\mathbf{a} \in P \setminus c$.

We introduce some additional notation. For $Z \in \{L, R\}$, we define $\tau|_Z^\mathbf{b}$ to be the Z -mosaic that results from taking the Z -formulas in τ that only use parameters from $\mathbf{b}$, and then unioning with the $\tilde{Z}$ -formulas in τ whose parameters are contained in $\mathbf{b}$ and are Z -cguarded in τ . For mosaics v, v' and $Z \in \{L, R\}$, we write $v'(\mathbf{a}\mathbf{b}) \supseteq_Z v$ if v' uses only parameters from $\mathbf{a}\mathbf{b}$ and is obtained

from v by adding only Z -formulas (so the $\tilde{Z}$ -formulas in v and v' are identical).

If $Y = X$, then define θ_X^τ to be

$$\bigvee_{\substack{v'(\mathbf{a}\mathbf{b}) \supseteq_X \tau|_X^\mathbf{b} \\ \text{s.t. } \tau \rightarrow_{X:\eta} v'}} \exists x. \left(\bigwedge_{\substack{\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v' \\ \text{s.t. } \tau \rightarrow_{X:\eta} \tau'}} \theta_X^{\tau'}[\mathbf{x}/\mathbf{a}] \right).$$

Informally, this interpolant expresses that there is a way to fulfil the existential requirement by adding some X -formulas to $\tau|_X^\mathbf{b}$, but no matter what $\tilde{X}$ -formulas we add, the resulting linked mosaic τ' has already been eliminated.

If $Y = \tilde{X}$, then define θ_X^τ to be

$$\text{gdd}_X^\tau(\mathbf{b}) \wedge \neg \left(\bigvee_{\substack{v'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} \tau|_{\tilde{X}}^\mathbf{b} \\ \text{s.t. } \tau \rightarrow_{\tilde{X}:\eta} v'}} \exists x. \left(\bigwedge_{\substack{\tau'(\mathbf{a}\mathbf{b}) \supseteq_X v' \\ \text{s.t. } \tau \rightarrow_{\tilde{X}:\eta} \tau'}} \theta_X^{\tau'}[\mathbf{x}/\mathbf{a}] \right) \right).$$

In both cases, notice that the parameters $\mathbf{a}\mathbf{b}$ in the inductively defined interpolants $\theta_X^{\tau'}$ are guarded by the atom $\alpha(\mathbf{a}\mathbf{b})$ appearing in the existential requirement $Y : \exists x.\alpha(\mathbf{x}\mathbf{b}) \wedge \psi(\mathbf{x}\mathbf{b})$. We cannot necessarily guard the quantification in the interpolant using $\alpha(\mathbf{x}\mathbf{b})$ since α may not be in the common signature. This is why the resulting interpolant might not be in GF, even when the input sentences are in GF. However, in case $Y = \tilde{X}$, we can always guard $\mathbf{b}$ in the common signature using $\text{gdd}_X^\tau(\mathbf{b})$. This guard is guaranteed to exist because the existential requirement is a $\tilde{X}$ -formula using parameters $\mathbf{b}$ in an X -mosaic. This ensures that the interpolant is in GNF. The occurrence of constants and the polarity of relations follows from the inductive hypothesis and the fact that the polarity of occurrences of relations are preserved when moving to mosaics.

We now sketch part of the proof of correctness, showing that $\tau_X \models \theta_X^\tau$ and $\theta_X^\tau \models \neg\tau_{\tilde{X}}$ when $Y = X$.

Assume there is a model $\mathfrak{M}$ for τ_X . Since $X : \eta \in \tau$, this means that $\mathfrak{M} \models \exists x.\alpha(\mathbf{x}\mathbf{b}) \wedge \psi(\mathbf{x}\mathbf{b})$, so there are elements $\mathbf{a}' \in M$ and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\mathbf{a}^{\mathfrak{M}'} := \mathbf{a}'$ such that $\mathfrak{M}' \models \alpha(\mathbf{a}\mathbf{b}) \wedge \psi(\mathbf{a}\mathbf{b})$.

Take v' to be the union of $\tau|_X^\mathbf{b}$ and the set of formulas $X : \psi'$ in $\text{cl}_X(\mathbf{a}\mathbf{b})$ such that $\mathfrak{M}' \models \psi'$. It can be checked that $\tau \rightarrow_{X:\eta} v'$. By construction, $\mathfrak{M}' \models v'_X$. Moreover, for any choice of $\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v'$, $\mathfrak{M}' \models \tau'_X$ (since the X -formulas are identical in v' and τ'). By the inductive hypothesis, this means that $\mathfrak{M}' \models \theta_X^{\tau'}$. Hence, $\mathfrak{M}' \models \bigwedge_{\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v'} \theta_X^{\tau'}$ and $\mathfrak{M} \models \theta_X^\tau$.

Now assume that there is a model $\mathfrak{M}$ of θ_X^τ . Then there is some $v'(\mathbf{a}\mathbf{b}) \supseteq_X \tau|_X^\mathbf{b}$ with $\tau \rightarrow_{X:\eta} v'$, elements $\mathbf{a}' \in M$, and an expansion $\mathfrak{M}'$ of $\mathfrak{M}$ with the interpretation $\mathbf{a}^{\mathfrak{M}'} := \mathbf{a}'$, such that $\mathfrak{M}' \models \bigwedge_{\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v'} \theta_X^{\tau'}$.

Note that for all $\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v'$ such that $\tau \rightarrow_{X:\eta} \tau'$, $\mathfrak{M}' \models \theta_X^{\tau'}$ and consequently, the inductive hypothesis implies $\mathfrak{M}' \models \neg\tau_{\tilde{X}}'$. In particular, consider $\tau' := v' \cup S'$ where S' is the set of formulas $\tilde{X} : \psi'$ in $\text{cl}_{\tilde{X}}(\mathbf{a}\mathbf{b})$ such that $\mathfrak{M}' \models \psi'$, $\text{par}(\psi') \cap \mathbf{a} \neq \emptyset$, and $\tilde{X} : \psi'$ is X -cguarded in v' as required in an X -mosaic. It is clear that $\tau'(\mathbf{a}\mathbf{b}) \supseteq_{\tilde{X}} v'$ and $\tau \rightarrow_{X:\eta} \tau'$, so $\mathfrak{M}' \models \neg\tau_{\tilde{X}}'$ as described above.

Since $\mathfrak{M}' \models \neg\tau_{\tilde{X}}'$, there is some conjunct χ in $\tau_{\tilde{X}}'$ such that $\mathfrak{M}' \models \neg\chi$. Consider some conjunct $\chi(\mathbf{a}\mathbf{b})$ in $\tau_{\tilde{X}}'$ that actually uses some parameters from $\mathbf{a}$. Then by choice of S' , $\mathfrak{M}' \models \chi(\mathbf{a}\mathbf{b})$,

so this formula cannot witness the fact that $\mathfrak{M}' \models \neg\tau'_{\tilde{X}}$. This means there must be some conjunct $\chi(\mathbf{b})$ in $\tau'_{\tilde{X}}$ that only uses parameters from $\mathbf{b}$ such that $\mathfrak{M}' \models \neg\chi(\mathbf{b})$. If $\tilde{X} : \chi(\mathbf{b})$ is in τ' , then $\tilde{X} : \chi(\mathbf{b}) \in \tau$ by the definition of $\tau \rightarrow_{X:\eta} \tau'$. Likewise, if $\tilde{X} : \chi(\mathbf{b}) \notin \tau'$, then it must be the case that $\text{par}(\chi)$ is X -cguarded in τ' , so $\tilde{X} : \chi(\mathbf{b}) \notin \tau$ (otherwise it would contradict the definition of $\tau \rightarrow_{X:\eta} \tau'$). In either case, this means $\chi(\mathbf{b})$ must be a conjunct in $\tau_{\tilde{X}}$. Hence, $\mathfrak{M} \models \neg\tau_{\tilde{X}}$. $\square$

We can prove the following constructive Lyndon interpolation result using the mosaic interpolants from Lemma 4.

Theorem 5. *Let φ_L and φ_R be GF sentences without equality over signatures σ_L and σ_R , respectively. If $\varphi_L \models \varphi_R$ and $|\varphi_L| + |\varphi_R| = n$, then we can construct a DAG representation of a GNF interpolant θ such that*

- $\varphi_L \models \theta$ and $\theta \models \varphi_R$;
- $\text{occ}(\theta) \subseteq \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$;
- $\text{con}(\theta) \subseteq \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$;
- the DAG representation of θ is of size at most $2^{2^{p(n)}}$ for some polynomial function p independent of φ_L and φ_R (and this can be improved to a size of at most $2^{p(n)}$ when the bound on the arity of relations is fixed).

For brevity, in this theorem and throughout the paper, we give only bounds on the output size, not the running time of the algorithms. However the proofs will show that the worst-case running time is bounded by a polynomial in the output size. In particular, the running time of the interpolation algorithm above is doubly-exponential in the input (but only singly-exponential for bounded arity GF sentences without equality).

Proof of Theorem 5. Let

$$\theta := \bigvee_{\tau(\emptyset) \supseteq_L \{L : \varphi_L\}} \bigwedge_{\substack{\tau'(\emptyset) \supseteq_R \tau \\ \text{s.t. } \tau' \in \mathcal{N}'}} \theta_L^{\tau'}$$

where $\mathcal{N}'$ is the set of eliminated mosaics and we write $v'(\emptyset) \supseteq_Z v$ if v' is over parameters $\emptyset$ and is obtained from v by adding only Z -formulas.

We prove that $\varphi_L \models \theta \models \varphi_R$.

Assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \varphi_L$. Let $\tau(\emptyset) \supseteq_L \{L : \varphi_L\}$ be the set of sentences from $\text{cl}_L(\emptyset)$ that are true in $\mathfrak{M}$. For any $\tau'(\emptyset) \supseteq_R \tau$ such that $\tau' \in \mathcal{N}'$, $\mathfrak{M} \models \tau'_L$. By Lemma 4, this means that $\mathfrak{M} \models \theta_L^{\tau'}$, so $\mathfrak{M} \models \theta$.

Now assume there is a structure $\mathfrak{M}$ such that $\mathfrak{M} \models \theta$. Then there is some $\tau(\emptyset) \supseteq_L \{L : \varphi_L\}$ such that $\mathfrak{M} \models \theta_L^{\tau'}$ for all $\tau'(\emptyset) \supseteq_R \tau$ such that $\tau' \in \mathcal{N}'$. Let S' be the set of formulas $R : \psi$ in $\text{cl}_R(\emptyset)$ such that $\mathfrak{M} \models \psi$. Consider $\tau' := \tau \cup S' \cup \{R : \neg\varphi_R\} \supseteq_R \tau$. By completeness of the mosaic system (Theorem 3), every mosaic containing $L : \varphi_L$ and $R : \neg\varphi_R$ must be in $\mathcal{N}'$. Hence, $\mathfrak{M} \models \theta_L^{\tau'}$. By Lemma 4 this implies that $\mathfrak{M} \models \neg\tau'_R$. But every formula $R : \psi$ in S' was chosen such that $\mathfrak{M} \models \psi$. The only way that $\mathfrak{M} \models \neg\tau'_R$ is if $\mathfrak{M} \models \varphi_R$.

The other properties in Theorem 5 follow from Lemma 4. $\square$

Although GF does not have CIP, it was shown in [14] that GF does admit the following weaker form of interpolation.

Theorem 6 ([14]). *For GF sentences $\varphi_L \models \varphi_R$ there is a GF sentence θ such that $\varphi_L \models \theta$ and $\theta \models \varphi_R$ and θ only uses relations in the common signature or relations occurring as a guard for some quantification in φ_L or φ_R .*

We remark that our mosaic interpolation method can be modified to produce GF interpolants like this, with the same DAG-size bounds as in Theorem 5, by modifying the inductive case of Lemma 4 to include the guard that is present already in the existential requirement.

5. GNF Mosaics

We now aim to describe mosaics that can handle φ_L and φ_R in GNF. For now, we assume φ_L and φ_R are GNF sentences, over signatures σ_L and σ_R , that are in GN-normal form, use ordinary quantifiers rather than relativized quantifiers, and do not use equality. Let $e := \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$.

Closure. It turns out that just taking the subformula closure is not sufficient to determine satisfiability of $\varphi_L \wedge \neg\varphi_R$ from a set of coherent mosaics. Instead, the mosaics must contain additional information about CQ-shaped subformulas (subformulas of the form $\exists \mathbf{x}. \bigwedge_i \psi_i(\mathbf{x}\mathbf{y})$).

Our solution is to add specializations to the closure. Informally, each specialization describes a way in which the original CQ-shaped formula could be satisfied in some tree-like structure.

Consider a GN-normal form formula $\phi(\mathbf{y})$ that is of the form $\exists \mathbf{x}. \bigwedge_i \psi_i(\mathbf{x}\mathbf{y})$. A *specialization* of ϕ is a formula ϕ' obtained from ϕ by the following operations:

- select a subset z of $\mathbf{x}$ (call variables from $\mathbf{y}z$ the *inside variables* and variables from $\mathbf{x} \setminus z$ the *outside variables*);
- select a partition $\mathbf{x}_1, \dots, \mathbf{x}_k$ of the outside variables, with the property that for every ψ_j , either ψ_j has no outside variables or all of its outside variables are contained in some partition element $\mathbf{x}_i$;
- let χ_0 be the conjunction of the ψ_i using only inside variables, and let χ_j be the conjunction of the ψ_i using outside variables and satisfying $\text{par}(\psi_i) \subseteq \mathbf{x}_j\mathbf{y}z$;
- set $\phi'(\mathbf{y}z)$ to be $\chi_0(\mathbf{y}z) \wedge \bigwedge_{j \in \{1, \dots, k\}} \exists \mathbf{x}_j. \chi_j(\mathbf{x}_j\mathbf{y}z)$.

We are now ready to define the closure. Each formula is labelled with a provenance $X \in \{L, R\}$ and a *polarity* $p \in \{+, -\}$. The *closure* $\text{cl}(X^p : \varphi)$ is the smallest set C of formulas containing $X^p : \varphi$ and such that:

- if $X^+ : \alpha(\mathbf{y}) \wedge \neg\psi(\mathbf{y}) \in C$, then $X^+ : \alpha(\mathbf{y}), X^- : \psi(\mathbf{y}) \in C$;
- if $X^- : \alpha(\mathbf{y}) \wedge \neg\psi(\mathbf{y}) \in C$, then $X^- : \alpha(\mathbf{y}), X^+ : \psi(\mathbf{y}) \in C$;
- if $X^p : \bigwedge_i \psi_i \in C$ or $X^p : \bigvee_i \psi_i \in C$, then $X^p : \psi_i \in C$ for all i ;
- if $X^p : \exists \mathbf{x}. \psi(\mathbf{x}\mathbf{y}) \in C$, then $X^p : \phi' \in C$ for all specializations ϕ' of $\exists \mathbf{x}. \psi(\mathbf{x}\mathbf{y})$;
- if $X^+ : \alpha(\mathbf{x}\mathbf{y}) \in C$, then $X^+ : \exists \mathbf{x}. \alpha(\mathbf{x}\mathbf{y}) \in C$.

Given parameters c , we let $\text{cl}_L(c)$ (respectively, $\text{cl}_R(c)$) consist of formulas from $\text{cl}(L^+ : \varphi_L)$ (respectively, $\text{cl}(R^- : \varphi_R)$) with free variables replaced by parameters from c or constants from e .

Roughly speaking, if $X^p : \psi \in \text{cl}_L(c) \cup \text{cl}_R(c)$, then the provenance $X \in \{L, R\}$ indicates whether this formula came from a subformula of φ_L or $\neg\varphi_R$, and the polarity $p \in \{+, -\}$ indicates whether this subformula occurred positively or negatively (i.e., whether it occurred within the scope of an even or odd number of negations). We write $p\psi$ to mean ψ (respectively, $\neg\psi$) if p is $+$ (respectively, $-$).

As in GF mosaics, we remark that the formulas in $\text{cl}_L(c) \cup \text{cl}_R(c)$ respect the polarity of relations in the original formulas.

Guards. Let τ be a collection of formulas from $\text{cl}_L(c) \cup \text{cl}_R(c)$. We say $\mathbf{b} \subseteq c$ is X -guarded in τ (respectively, X -cguarded in τ) if $\mathbf{b}$ is empty, or if there is some $X^+ : \exists \mathbf{x}. \alpha(\mathbf{x}\mathbf{b}) \in \tau$ where α is an X -atom (respectively, X -catom). (The definition of X -atom and X -catom is the same as in Section 4.)

GNF mosaics. Let $\mathbf{c}$ be a tuple of parameters such that $|\mathbf{c}| \leq \text{width}(\varphi_L \wedge \neg\varphi_R)$. A GNF X -mosaic τ for $\varphi_L \wedge \neg\varphi_R$ using parameters $\mathbf{c}$ is a subset of $\Gamma(\mathbf{c}) := \text{cl}_L(\mathbf{c}) \cup \text{cl}_R(\mathbf{c})$ such that if $\tilde{X}^p : \psi \in \tau$, then $\text{par}(\psi)$ is X -cguarded in τ .

Like the GF mosaics, the $\tilde{X}$ -formulas are restricted in X -mosaics. Unlike a GF mosaic, this restriction only applies to positive $\tilde{X}$ -formulas.

Let τ be a GNF X -mosaic over parameters $\mathbf{c}$. Assume $Y^+ : \eta \in \tau$ for η an *existential requirement* of the form $\exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b})$. Then τ is *linked* to τ' via $Y^+ : \eta$ (written $\tau \rightarrow_{Y^+ : \eta} \tau'$) if

- τ' is a GNF Y -mosaic over $\mathbf{a}\mathbf{b}$ with $\mathbf{a} \cap \mathbf{c} = \emptyset$;
- $Y^+ : \psi(\mathbf{a}\mathbf{b}) \in \tau'$;
- for all $Y^p : \psi'(\mathbf{b}) \in \tau$, $Y^p : \psi'(\mathbf{b}) \in \tau'$;
- for all $\tilde{Y}^p : \psi'(\mathbf{b}) \in \tau$ such that either $\text{par}(\psi')$ is Y -cguarded in τ' or $p = -$, $\tilde{Y}^p : \psi'(\mathbf{b}) \in \tau'$;
- for all $Z^p : \psi'(\mathbf{b}) \in \tau'$, $Z^p : \psi'(\mathbf{b}) \in \tau$.

As in the GF mosaics, the maximum amount of information about shared parameters is passed through a link, and no new information about shared parameters is added in τ' .

Fix some set P of parameters of size $2 \cdot \text{width}(\varphi_L \wedge \neg\varphi_R)$, and let $\mathcal{M}$ be the set of GNF mosaics over parameters P . The following result is proven by a routine calculation.

Proposition 7. *There is a polynomial function p (independent of φ_L and φ_R) such that if $|\varphi_L| + |\varphi_R| = n$, then there are at most $2^{2^{p(n)}}$ mosaics in $\mathcal{M}$. This bound holds even if φ_L and φ_R are GNF sentences but are not in GN-normal form.*

Coherence. A GNF X -mosaic $\tau(\mathbf{c})$ is *downward closed* if the following properties are satisfied:

- if $Y^+ : \alpha(\mathbf{b}) \wedge \neg\psi(\mathbf{b}) \in \tau$, then $Y^+ : \alpha(\mathbf{b}) \in \tau$ and $Y^- : \psi(\mathbf{b}) \in \tau$;
- if $Y^- : \alpha(\mathbf{b}) \wedge \neg\psi(\mathbf{b}) \in \tau$ and $Z^+ : \alpha(\mathbf{b}) \in \tau$, then $Y^+ : \psi(\mathbf{b}) \in \tau$;
- if $Y^+ : \bigwedge_i \psi_i \in \tau$ (respectively, $Y^- : \bigvee_i \psi_i \in \tau$), then $Y^+ : \psi_i \in \tau$ (respectively, $Y^- : \psi_i \in \tau$) for all i ;
- if $Y^+ : \bigvee_i \psi_i \in \tau$ (respectively, $Y^- : \bigwedge_i \psi_i \in \tau$), then $Y^+ : \psi_i \in \tau$ (respectively, $Y^- : \psi_i \in \tau$) for some i ;
- if $Y^- : \exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b}) \in \tau$, then for all specializations $\phi'(\mathbf{y}\mathbf{z})$ of $\exists \mathbf{x}.\psi(\mathbf{x}\mathbf{y})$ and for all $\mathbf{c}' \subseteq \mathbf{c} \cup \mathbf{e}$ with $|\mathbf{c}'| = |\mathbf{z}|$, $Y^- : \phi'(\mathbf{b}\mathbf{z})[\mathbf{c}'/\mathbf{z}] \in \tau$.

The last rule in this definition is a generalization of the rule for universal requirements in the GF mosaics. It helps ensure that when we construct a model from a saturated set of internally consistent mosaics, a CQ-shaped subformula that is asserted to be false in one mosaic, does not become true in the constructed model.

As before, a GNF X -mosaic τ is *internally consistent* if τ is downward closed and there is no atomic formula α and no $Y, Z \in \{L, R\}$ such that $Y^+ : \alpha \in \tau$ and $Z^- : \alpha \in \tau$.

A set of GNF mosaics is said to be *saturated* if for every mosaic $\tau(\mathbf{c})$ in the set and for every existential requirement $Y^+ : \eta \in \tau$ for η of the form $\exists \mathbf{x}.\psi(\mathbf{x}\mathbf{b})$, either (i) $Y^+ : \psi(\mathbf{x}\mathbf{b})[\mathbf{a}/\mathbf{x}]$ is in τ for some $\mathbf{a} \subseteq \mathbf{c} \cup \mathbf{e}$ (so η is fulfilled in τ) or (ii) τ is linked to some mosaic τ' in the set via $Y^+ : \eta$ (so η is fulfilled in τ').

These requirements lead to the following key theorem.

Theorem 8. *$\varphi_L \wedge \neg\varphi_R$ is satisfiable if and only if there is a saturated set of internally consistent GNF mosaics from $\mathcal{M}$ that contains a mosaic τ such that $L^+ : \varphi_L, R^- : \varphi_R \in \tau$.*

Using the elimination procedure described earlier, this gives an alternative proof of the 2EXPTIME upper bound for satisfiability of GNF (as stated in Theorem 1).

Extensions of GNF mosaics. We can consider various extensions of these mosaics, to handle input with free variables, relativized quantifiers, and equality. We briefly summarize some of these extensions.

Relativization. For input that uses $\mathbb{U}$ -relativized quantifiers, a *relativized X -mosaic* $\tau(\mathbf{c})$ is a collection of $\mathbb{U}$ -relativized GNF sentences with parameters from $\mathbf{c}$. Each relativized mosaic also has a set of distinguished parameters $\mathbf{d} \subseteq \mathbf{c}$ (which we call the relativized parameters in the mosaic). These parameters $\mathbf{d}$ in τ will be removed from θ_X^+ using existential quantification when constructing the mosaic interpolants. Hence, we add further restrictions on the $\tilde{X}$ -formulas that use parameters from $\mathbf{d}$ to ensure that we can relativize any quantification over $\mathbf{d}$. In particular, if $d \in \mathbf{d}$ appears in a $\tilde{X}$ -formula in an X -mosaic τ then there must be some relativizer $X^+ : Ud \in \tau$ such that $U \in \mathbb{U}$ is an X -catom.

Equality. We have ignored the presence of equality in the mosaic construction, but it can be accommodated by normalizing the use of equality in formulas, and then putting additional “equality consistency requirements” on the mosaics.

Constructive interpolation. Using these extensions of the GNF mosaics, we can lift the restrictions placed on φ_L and φ_R at the beginning of this section, and derive the following constructive interpolation results, subsuming Theorem 5. This proves constructive RIP for GNF, and is a key contribution.

Theorem 9. *Let φ_L and φ_R be GNF (respectively, equality-free UNF) formulas over signatures σ_L and σ_R , respectively. If $\varphi_L \models \varphi_R$ and $|\varphi_L| + |\varphi_R| = n$, then we can construct a DAG representation of a GNF (respectively, UNF) interpolant θ such that*

- $\varphi_L \models \theta$ and $\theta \models \varphi_R$;
- $\text{occ}(\theta) \subseteq \text{occ}(\varphi_L) \cap \text{occ}(\varphi_R)$;
- $\text{free}(\theta) \subseteq \text{free}(\varphi_L) \cap \text{free}(\varphi_R)$;
- $\text{con}(\theta) \subseteq \text{con}(\varphi_L) \cup \text{con}(\varphi_R)$;
- if φ_L and φ_R are $\mathbb{U}$ -relativized, for $\mathbb{U}$ a distinguished set of unary relations from $\sigma_L \cup \sigma_R$, then θ is $\mathbb{U}$ -relativized (when treating $\top$ and $\perp$ as atomic formulas);
- the DAG representation of θ is of size at most $2^{2^{p(n)}}$ for some polynomial function p independent of φ_L and φ_R (and this can be improved to a size of at most $2^{p(n)}$ when φ_L and φ_R are in GF without equality and the bound on the arity of relations is fixed).

Note that using this interpolation method, we can also construct an *FO* interpolant that can be represented by a doubly-exponential sized formula – e.g. by applying the theorem above, and then converting efficiently from a DAG representation to a formula [3].

Beth definability. As a corollary of Theorem 9, we get doubly exponential bounds on the DAG size of explicit GNF definitions coming from Projective Beth Definability. The argument is standard (see [13]).

6. Lower Bounds

We have matching lower bounds for the interpolation results stated in Theorem 9.

Theorem 10. *There is a polynomial function p and a family of UNF sentences $\chi_n \models \chi'_n$ without equality such that $|\chi_n| + |\chi'_n| \leq p(n)$, and there is no GNF interpolant θ_n for $\chi_n \models \chi'_n$ with size at most $2^{2^{n-1}}$, even when θ_n is represented via a DAG.*

The sentences χ_n and χ'_n describe certain *indexed AND/OR trees*, finite trees where the leaves are labelled with 0 or 1, and other nodes are labelled with AND or OR. Moreover, each node is indexed by a tree that describes the depth of the node (up to depth 2^{2^n}).

Formally, an indexed AND/OR tree is a structure over signature ρ_n consisting of the following predicates: 1. Unary predicates IndValOne , $\text{IndDepth}_0, \dots, \text{IndDepth}_n$, and a ternary connective IndChild describe the index of one node. Informally, $\text{IndDepth}_0, \dots, \text{IndDepth}_n$ and IndChild describe the structure of a binary tree of depth n (with IndDepth_0 labelling the root), while IndValOne distinguishes the leaves of this tree with value one. Each such tree can be assigned a value, when viewing the leaves as 2^n bits of a binary number. 2. Unary predicates InputOne and IsOr along with binary predicate AOChild describe the AND/OR structure, with IsOr representing an OR node and InputOne distinguishing the inputs on the leaves that have value one.

The signatures of χ_n and χ'_n extend ρ_n using two copies of a predicate AOValOne for the intermediate value of the circuit computation at each node. Hence ρ_n will be the common signature for the interpolant.

Roughly speaking, the formula χ_n asserts

there is an indexed AND/OR tree of depth 2^{2^n} , and the value (using unary relation AOValOne) of any AND/OR tree in the structure is 1.

Likewise, χ'_n asserts

the value (using unary relation $\text{AOValOne}'$) of any AND/OR tree in the structure is 1.

We have $\chi_n \models \chi'_n$, and there is a polynomial function p such that for all n , $|\chi_n| + |\chi'_n| \leq p(n)$.

GN bisimulation game. In order to prove that there is no “small” GNF interpolant for $\chi_n \models \chi'_n$, we use a variation of GN bisimulation [5].

A position in the k -width version of the GN bisimulation game between $\mathfrak{A}$ and $\mathfrak{B}$ (relative to some signature σ) is a *partial rigid homomorphism* f from $\mathfrak{A}$ to $\mathfrak{B}$ or vice versa with $|\text{dom}(f)| \leq k$. This means that f is a partial homomorphism and for any guarded tuple $c \in \text{dom}(f)$, $f \upharpoonright_c$ (the restriction of f to c) is a partial isomorphism. We say the *active structure* is the structure containing $\text{dom}(f)$.

Starting in position f , one round of the game consists of the following: (i) Spoiler can restrict to some subset $c \subseteq \text{dom}(f)$, and then the game proceeds from position $f \upharpoonright_c$; or (ii) if $\text{dom}(f)$ is a guarded tuple c , then Spoiler can choose to switch structures, and the game proceeds from position $(f \upharpoonright_c)^{-1}$; or (iii) if $|\text{dom}(f)| < k$, Spoiler can select an element c such that c is in the active structure but $c \notin \text{dom}(f)$, Duplicator can choose d in the inactive structure such that $f[c \mapsto d]$ is a partial rigid homomorphism, and then the game proceeds from position $f[c \mapsto d]$ (Duplicator immediately loses if she is unable to choose a d such that $f[c \mapsto d]$ is a partial rigid homomorphism).

We write $\mathfrak{A}, a \xrightarrow{k, m}_{\text{gn}[\sigma]} \mathfrak{B}, b$ if Duplicator has a winning strategy in the k -width m -round guarded negation bisimulation game relative to signature σ starting from a partial rigid homomorphism $g : a \mapsto b$. A winning strategy for Duplicator implies agreement between $\mathfrak{A}$ and $\mathfrak{B}$ on certain GNF formulas.

Proposition 11. Assume $\mathfrak{A}, a \xrightarrow{k, m}_{\text{gn}[\sigma]} \mathfrak{B}, b$, i.e. Duplicator has a winning strategy in the k -width, m -round GN-bisimulation game between $\mathfrak{A}$ and $\mathfrak{B}$ (relative to signature σ) starting from position $a \mapsto b$. If $\varphi(x)$ is a DAG representation of a GNF formula over σ such that $|\varphi| \leq m$, $\text{width}(\varphi) \leq k$, and $\text{free}(\varphi) = x$, then $\mathfrak{A} \models \varphi(a)$ implies $\mathfrak{B} \models \varphi(b)$.

Counterexamples. The specific family of counterexamples that we use to prove Theorem 10 are an inductively defined family of indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ of increasing depth. The value

of $\mathfrak{A}_i^n$ (respectively, $\mathfrak{B}_i^n$) is 1 (respectively, 0), but the trees have some identical subtrees that make them hard to distinguish with “small” GNF sentences.

Roughly speaking, we show that for $i = 2^{2^n-1}$, Duplicator has a winning strategy in the i -round i -width GN bisimulation game relative to signature ρ_n between the indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ (starting from the empty homomorphism). By Proposition 11, this is enough to ensure that $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ cannot be distinguished by GNF sentences with a DAG representations of size at most i .

Supposing that there is a “small” (size at most 2^{2^n-1}) DAG representation of an interpolant θ_n for $\chi_n \models \chi'_n$ implies that $\mathfrak{A}_i^n \models \theta_n$ since $\mathfrak{A}_i^n \models \chi_n$ by construction. But $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ cannot be distinguished by such “small” formulas, so this means that $\mathfrak{B}_i^n \models \theta_n$, and hence $\mathfrak{B}_i^n \models \chi'_n$. This is impossible since $\mathfrak{B}_i^n$ is an indexed AND/OR tree with value 0, not value 1 as asserted by χ'_n .

We now describe in more detail this family of counterexamples $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ for all $i \leq 2^{2^n-1}$. For notational simplicity in the description below, we write $\mathfrak{A}_i$ and $\mathfrak{B}_i$ instead of $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$.

The basic structure of these AND/OR trees is defined recursively as follows, where we write a to denote a single a -labelled node, and $a(T_1, T_2)$ to denote a tree where the root is labelled a and the children are the roots of trees T_1 and T_2 :

- $\mathfrak{A}_0 := 1$ and $\mathfrak{B}_0 := 0$;
- for $i > 0$, $\mathfrak{A}_i := \text{AND}(\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1}), (\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1})))$ and $\mathfrak{B}_i := \text{AND}(\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{A}_{i-1}), (\text{OR}(\mathfrak{B}_{i-1}, \mathfrak{B}_{i-1})))$.

Evaluating $\mathfrak{A}_i$ (respectively, $\mathfrak{B}_i$) gives value 1 (respectively, 0).

A separate index tree (binary tree of depth n) is then attached to each node in the trees described above, where the index correctly describes the depth of this node in the AND/OR tree.

We view these indexed AND/OR trees as structures over the signature $\rho_n = \sigma_n \cap \sigma'_n$. In particular, this means that there are no internal computations using AOValOne or $\text{AOValOne}'$.

Let $h : b \mapsto a$ be a position (partial rigid homomorphism) in the k -width GN bisimulation game between some $\mathfrak{B}_i$ and $\mathfrak{A}_i$. We can assume that h is restricted to elements in the AND/OR tree (rather than the elements in the index trees). This is without loss of generality since h can always be chosen to preserve the depth of positions, and the local structure of the index trees is identical in $\mathfrak{A}_i$ and $\mathfrak{B}_i$.

As mentioned earlier, we want to show that Duplicator has a winning strategy in certain GN bisimulation games between $\mathfrak{B}_i$ and $\mathfrak{A}_i$. In order to prove this inductively, we need to introduce the notion of “safe” positions h , which represent good starting positions for Duplicator.

We first introduce some additional notation. We write $\mathfrak{B}_i(b)$ for the restriction of $\mathfrak{B}_i$ to the subtree rooted at b . For $i > 0$, let C_i (respectively, D_i) denote the set of nodes in $\mathfrak{A}_i$ (respectively, $\mathfrak{B}_i$) at depth 2 (these are the roots of the $\mathfrak{A}_{i-1}$ and $\mathfrak{B}_{i-1}$ subtrees of $\mathfrak{A}_i$ and $\mathfrak{B}_i$). We write $b \upharpoonright_b$ for the restriction of b to those elements appearing in $\mathfrak{B}_i(b)$.

We now define inductively what it means for h to be (i, k, m) -safe. We say a strategy in the m -round game is (i, k, m) -safe if it uses only (i, k, m') -safe positions where m' is the number of remaining moves in the game. We say $h : b \mapsto a$ is (i, k, m) -safe if it satisfies the following properties:

- (S1) for all $b \in b$, $\text{depth}(b) = \text{depth}(h(b))$;
- (S2) if $i > 0$, then there exists a mapping $f : D_i \rightarrow C_i$ such that for all $b, b' \in D_i$,
 - if b, b' are siblings, then $f(b), f(b')$ are siblings;
 - if $d \in b$ is the parent of b , then $h(d)$ is the parent of $f(b)$;
 - if $d \in b$ appears in $\mathfrak{B}_i(b)$, then $h(d)$ appears in $\mathfrak{A}_i(f(b))$;
 - $\mathfrak{B}_i(b), b \upharpoonright_b \xrightarrow{k, m}_{\text{gn}[\rho_n]} \mathfrak{A}_i(f(b)), h(b \upharpoonright_b)$ via a strategy for Duplicator that is $(i-1, k, m)$ -safe.

Likewise, we say a position $h : \mathbf{a} \mapsto \mathbf{b}$ is (i, k, m) -safe if the conditions (S1) and (S2) hold with $\mathfrak{B}_i, \mathbf{b}, D_i$ exchanged with $\mathfrak{A}_i, \mathbf{a}, C_i$, respectively.

Observe that the empty partial rigid homomorphism h from $\mathfrak{B}_i$ to $\mathfrak{A}_i$ is (i, k, m) -safe: (S1) vacuously holds, and (S2) holds since (if $i > 0$) we can choose f such that sibling relationships are preserved, and $\mathfrak{A}_{i-1}$ (respectively, $\mathfrak{B}_{i-1}$) subtrees in $\mathfrak{B}_i$ are mapped to $\mathfrak{A}_{i-1}$ (respectively, $\mathfrak{B}_{i-1}$) subtrees in $\mathfrak{A}_i$. Similarly for the empty partial rigid homomorphism from $\mathfrak{A}_i$ to $\mathfrak{B}_i$.

More importantly, we can show that Duplicator has a winning (i, k, m) -safe strategy when starting from an (i, k, m) -safe position.

Lemma 12. *For all $k, i \leq 2^{2^n-1}$, and $m \leq i$,*

- *if $\mathbf{b} \mapsto \mathbf{a}$ is (i, k, m) -safe then $\mathfrak{B}_i, \mathbf{b} \xrightarrow{\text{gn}[\rho_n]}^{k, m} \mathfrak{A}_i, \mathbf{a}$ via a strategy for Duplicator that is (i, k, m) -safe;*
- *if $\mathbf{a} \mapsto \mathbf{b}$ is (i, k, m) -safe then $\mathfrak{A}_i, \mathbf{a} \xrightarrow{\text{gn}[\rho_n]}^{k, m} \mathfrak{B}_i, \mathbf{b}$ via a strategy for Duplicator that is (i, k, m) -safe.*

This implies that for $i = 2^{2^n-1}$, Duplicator has a winning strategy in the i -round i -width GN bisimulation game relative to signature ρ_n between the indexed AND/OR trees $\mathfrak{A}_i^n$ and $\mathfrak{B}_i^n$ (starting from the empty homomorphism). As argued above, this can be combined with Proposition 11 to prove the lower bound stated in Theorem 10.

Guarded fragment. A variant of this construction inspired by [12] can be used to get a lower bound for the size of interpolants of GF validities.

Theorem 13. *There is a polynomial function p and a family of GF sentences $\chi_n \models \chi'_n$ with relations of unbounded arity (respectively, bounded arity) such that $|\chi_n| + |\chi'_n| \leq p(n)$, and there is no GNF interpolant θ_n for $\chi_n \models \chi'_n$ with size at most 2^{2^n} (respectively, size at most 2^n), even when θ_n is represented by a DAG.*

Beth definability. A similar argument can be used to prove doubly-exponential lower bounds on the DAG-size of GNF explicit definitions coming from Beth Definability.

7. Preservation Theorems

In this section, we consider some preservation theorems for GNF and GF that follow from the constructive interpolation results of Theorem 9. We remark that the results in this section hold when preservation and equivalence is considered over finite structures only, as well as over general structures.

We first consider the analog of the Łoś-Tarski theorem. A formula φ is *preserved under extensions* if $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{A} \models \varphi$ together imply $\mathfrak{B} \models \varphi$ (see, e.g., [9, 22]). It was shown in [7] that every GNF formula preserved under extensions has an equivalent *existential GNF formula*, which is a GNF formula where no quantifier is in the scope of a negation symbol. We can use relativized GNF interpolation to get a bound on the size of this existential formula. This contrasts with the result of Dawar et. al. [9] that for general first-order formulas there is no elementary bound for the Łoś-Tarski theorem.

Corollary 14 (Analog of Łoś-Tarski Theorem). *Let φ be a GNF formula over signature σ with $|\varphi| = n$. If φ is preserved under extensions, then we can construct a DAG representation of an equivalent existential GNF formula φ' such that $|\varphi'| \leq 2^{2^{p(n)}}$ for some polynomial function p independent of φ .*

Proof. We follow Otto [22]. Let $\mathbb{U} := \{U^1, U^2\}$ for distinct unary predicates U^1, U^2 not occurring in σ .

Since φ is preserved under extensions, we have

$$\neg \exists y. (U^1 y \wedge \neg U^2 y) \wedge \bigwedge_{t \in \mathbf{x} \cup \mathbf{e}} U^1 t \wedge \varphi^1 \models \varphi^2$$

where $\text{free}(\varphi) = \mathbf{x}$, $\text{con}(\varphi) = \mathbf{e}$, and φ^i denotes the relativization of φ to U^i . For readability, here and in the similar arguments below, the equality guard for the negation of the first conjunct in the antecedent has been omitted.

Let N be the size of this validity, which is linear in n . By Theorem 9, the previous GNF validity admits a $\mathbb{U}$ -relativized GNF interpolant θ over signature $\sigma \cup \mathbb{U}$ with a DAG representation of size at most $2^{2^{p(N)}}$ for some polynomial function p independent of φ .

Since U^1 does not occur on both sides of the validity, U^1 cannot appear at all in θ . Since U^2 does not occur negatively on the left hand side of the original validity, U^2 cannot appear negatively in θ . This means that every existential quantifier in θ must occur positively (since every quantifier must be relativized by U^2 , and U^2 can only occur positively).

Replacing every occurrence of some $U^2 z$ in θ by $z = z$ results in a GNF formula φ' over σ that is equivalent to φ , and such that any quantifier occurs within the scope of an even number of negations. By pushing the negations to the atomic level (and potentially duplicating guards to cover any negated atoms), we have a doubly-exponential size DAG representation of an existential GNF formula. $\square$

Similarly, we can use the Lyndon interpolation result to prove a relationship between monotonicity and positivity. We say a formula φ is *monotone* if it is preserved when a structure is modified via adding tuples to the interpretation of some relation. We say a formula φ is *positive* if all relation symbols (except possibly equality) occur positively (within the scope of an even number of negations). A formula is *domain-independent* (see [1]) if it depends only on the interpretations of relation symbols.

Corollary 15 (Monotone=Positive). *Let φ be a GNF formula over signature σ with $|\varphi| = n$ such that φ is monotone. Then we can construct a DAG representation of an equivalent positive GNF formula φ' such that $|\varphi'| \leq 2^{2^{p(n)}}$ for some polynomial function p independent of φ . Furthermore, if φ is domain-independent, then φ' can be taken to be positive and in existential GNF.*

Proof. Since φ is monotone, we have

$$\bigwedge_{R \in \sigma} \neg \exists \mathbf{y}. (R^1 \mathbf{y} \wedge \neg R^2 \mathbf{y}) \wedge \varphi^1 \models \varphi^2 \quad (1)$$

where φ^i denotes φ with each $R \in \sigma$ replaced with R^i . Applying Theorem 9 yields an interpolant only using relations of the form R^2 for $R \in \sigma$. Moreover, these relations must occur positively. Replacing each R^2 with R , yields φ' .

The above argument is modified in the domain-independent case by using additional predicates for the domains. More precisely, we relativize the antecedent and consequent of the entailment (1) by fresh unary predicates U^1 and U^2 , respectively; and we add $\bigwedge_{R \in \sigma; i, j \in \{1, 2\}} \neg \exists \mathbf{x}. (\bigwedge_{x \in \mathbf{x}} U^i x \wedge R^i \mathbf{x} \wedge \neg \bigwedge_{x \in \mathbf{x}} U^j x)$ as conjunct to the antecedent. Applying Theorem 9 as before, and replacing each R^2 with R , as well as $U^2 z$ by $z = z$, yields the desired formula φ' . $\square$

Guarded fragment. We now turn to effective preservation for GF. Recall that Andr  ka et. al. [2] prove a Łoś-Tarski theorem for GF, but with no bound. Here we state a constructive version.

Theorem 16. *Let φ be a GF formula over signature σ with $|\varphi| = n$. If φ is preserved under extensions, then we can construct an equivalent existential GF formula φ' such that $|\varphi'| \leq 2^{2^{2^{2p(n)}}}$ for a polynomial function p independent of φ .*

Proof. We sketch the argument for the boolean case. The proof proceeds by applying Corollary 14 to obtain a DAG representation of some GNF existential χ' equivalent to φ of doubly-exponential size. We can convert χ' to χ'' that is a union of χ''_i , where each χ''_i is a “GNCQ” – a conjunctive query with atomic negation, where every negated atom is guarded. Further, each GNCQ χ''_i within χ'' is at most exponential in the size of χ' , and the overall size of χ'' (as a formula) is at most doubly exponential in the size of χ' . Let A_i be the set of queries in GF of the form $\exists \mathbf{y}. \bigwedge_j A_{ij}$, where $\bigwedge A_{ij}$ is obtained by identifying variables in χ''_i and then adding on at most $2^{|\varphi|}$ additional positive atoms. We claim that χ' is equivalent to $\bigvee_i \bigvee_{Q \in A_i} Q$. Since the set of indices i is doubly-exponential in the size of χ' and each A_i has size doubly-exponential in χ' , this gives the desired bound. Clearly each query in A_i implies χ''_i and hence implies φ . On the other hand, consider a model M satisfying φ . Given any model M there is another structure M^* agreeing with M on all GF sentences that has a guarded tree decomposition. Indeed, the guarded unravelling of M (see, e.g., [2]) gives such an M^* . M^* must satisfy φ and therefore must satisfy some χ''_i via a homomorphism h . Let C be the image of h and Q_C be the conjunctive query that describes the restriction of M^* to the bags containing C – such a query can be written in GF by constructing it inductively from the leaves of the tree up. The size of Q_C is at most the size of χ''_i times the maximal size of bags – the latter being a factor which is at most exponential in the size of the schema. Note that every element $c \in C$ corresponds to exactly one variable y_c of Q_C . We let Q'_C extend Q_C by adding, for every negated atom in χ''_i , the result of replacing each variable x_i by $y_{h(x_i)}$. Then Q_C also holds in M^* and $Q'_C \in \text{GF}$. Hence Q'_C holds in M and is in A_i , which completes the argument. $\square$

We get similar bounds for characterizing monotone domain-independent GF formulas.

Theorem 17. *Let φ be a GF formula over signature σ with $|\varphi| = n$. If φ is monotone and domain-independent then we can construct an equivalent formula φ' that is positive and is in existential GF, and such that $|\varphi'| \leq 2^{2^{2^{2p(n)}}}$ for a polynomial function p independent of φ .*

Proof. Again we focus on the case where φ is boolean.

By Corollary 15, φ is equivalent to a positive and existential GNF formula χ' whose DAG representation has doubly-exponential size. As above, we can convert this to a union $\chi'' = \bigvee_i \chi''_i$, with χ''_i in this case being a conjunctive query, but possibly including guarded inequalities. As before, this results in a doubly-exponential blow up in the overall size, but only a singly-exponential blow up in the size of each CQ.

Again letting A_i be the set of queries in GF of the form $\exists \mathbf{y}. \bigwedge_j A_{ij}$, where $\bigwedge A_{ij}$ is obtained by identifying variables in χ''_i and then adding on at most $2^{|\varphi|}$ additional positive atoms, we can show that χ' is equivalent to $\bigvee_i \bigvee_{Q \in A_i} Q$. $\square$

8. Conclusions

We have shown that GNF is a rich logic that admits interpolation and preservation theorems with elementary complexity, in sharp contrast to full first-order logic. In doing this, we developed the mosaic method for use in interpolation. In future work, we will look

at revisiting the applications of GNF interpolation to databases (as in [6]) in light of these results.

The bounds for interpolation are tight for GNF and UNF. We leave open the question of tightness of our bounds for preservation, as well as the tightness of interpolation for GF with restricted arity and equality. We also do not know the status of many other preservation theorems for GNF, such as preservation under surjective homomorphism, and the status of monotone=positive for non-domain independent GF formulas.

References

- [1] S. Abiteboul, R. Hull, and V. Vianu. *Foundations of Databases*. Addison-Wesley, 1995.
- [2] H. Andr  ka, I. N  meti, and J. van Benthem. Modal languages and bounded fragments of predicate logic. *J. Phil. Logic*, 27(3):217–274, 1998.
- [3] J. Avigad. Eliminating definitions and skolem functions in first-order logic. *ACM Trans. Comput. Log.*, 4(3), 2003.
- [4] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The description logic handbook*. Cambridge, 2003.
- [5] V. B  r  ny, B. ten Cate, and L. Segoufin. Guarded negation. In *ICALP*, 2011.
- [6] V. B  r  ny, M. Benedikt, and P. Bourhis. Access patterns and integrity constraints revisited. In *ICDT*, 2013.
- [7] V. B  r  ny, M. Benedikt, and B. ten Cate. Rewriting guarded negation queries. In *MFCS*, 2013.
- [8] W. Craig. Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *JSL*, 22(3):269–285, 1957.
- [9] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Model theory makes formulas large. In *ICALP*, 2007.
- [10] M. Fitting. *First-order logic and automated reasoning*. Springer, 1996.
- [11] H. Friedman. The complexity of explicit definitions. *Advances in Math.*, 20(1):18 – 29, 1976.
- [12] E. Gr  del. On the restraining power of guards. *JSL*, 64(4):1719–1742, 1999.
- [13] E. Hoogland. *Definability and interpolation: model-theoretic investigations*. PhD thesis, University of Amsterdam, 2000.
- [14] E. Hoogland, M. Marx, and M. Otto. Beth definability for the guarded fragment. In *LPAR*, 1999.
- [15] R. C. Lyndon. An interpolation theorem in the predicate calculus. *Pacific J. Math.*, 9:129–142, 1959. ISSN 0030-8730.
- [16] M. Marx. Queries determined by views: pack your views. In *PODS*, 2007.
- [17] M. Marx, S. Mikulas, and M. Reynolds. The mosaic method for temporal logics. In *TABLEAUX*, 2000.
- [18] K. L. McMillan. Applications of Craig Interpolation to model checking. In *CSL*, 2004.
- [19] S. Mikul  s. Taming first-order logic. *Journal of the IGPL*, 6(2):305–316, 1998.
- [20] A. Nash, L. Segoufin, and V. Vianu. Views and queries: Determinacy and rewriting. *ACM Trans. Database Syst.*, 35(3), 2010.
- [21] I. N  meti. *Free Algebras and Decidability in Algebraic Logic*. PhD thesis, Hungarian Academy of Sciences, Budapest, 1986.
- [22] M. Otto. An interpolation theorem. *BSL*, 6(4):447–462, 2000.
- [23] B. ten Cate and L. Segoufin. Unary negation. In *STACS*, 2011.
- [24] B. ten Cate, E. Franconi, and I. Seylan. Beth definability in expressive description logics. In *IJCAI*, 2011.