Combining theories: the Ackerman and Guarded Fragments*

Carlos Areces[1,2] and Pascal Fontaine[1,3]

[1]INRIA Nancy-Grand Est, Nancy, France
[2]FaMAF, Universidad Nacional de Córdoba, Córdoba, Argentina
[3]Université de Nancy, Loria, Nancy, France
Carlos.Areces@gmail.com
Pascal.Fontaine@loria.fr

Abstract. Combination of decision procedures is at the heart of Satisfiability Modulo Theories (SMT) solvers. It provides ways to compose decision procedures for expressive languages which mix symbols from various decidable theories. Typical combinations include (linear) arithmetic, uninterpreted symbols, arrays operators, etc. In [7] we showed that any first-order theory from the Bernays-Schönfinkel-Ramsey fragment, the two variable fragment, or the monadic fragment can be combined with virtually any other decidable theory. Here, we complete the picture by considering the Ackermann fragment, and several guarded fragments. All theories in these fragments can be combined with other decidable (combinations of) theories, with only minor restrictions. In particular, it is not required for these other theories to be stably-infinite.

1 Introduction

Devising satisfiability decision procedures for the combination of logical theories has been a very active research subject during the last fifteen years. It is the theoretical background on which Satisfiability Modulo Theories (SMT) solvers are built. For instance, the set of literals

$$L = \{a \leq b, b \leq a + f(a), P(h(a) - h(b)), \neg P(0), f(a) = 0\}$$

can be shown to be unsatisfiable by an SMT solver, implementing the Nelson-Oppen framework [16] combining a decision procedure for the theory of uninterpreted symbols and a decision procedure for linear arithmetic. SMT solvers (see [2] for a thorough presentation of the techniques behind SMT solvers) are now widely used, notably for model-checking and formal verification.

Initial combination results (e.g., [16, 17]) imposed strong conditions to ensure decidability of the satisfiability problem for the combined theories, such as requiring the theories to be *stably infinite*, i.e., requiring any satisfiable set of literals within the theories to have an infinite model. Many theories, and specially, many theories interesting for formal verification of hardware and software,

* This work is partly supported by the ANR project DECERT.

are expressive enough to restrict the size of a model to be finite and, hence, are not stably infinite. In other words, stable infiniteness is a sufficient condition for theory combination, but it is too restrictive.

In recent years, much of the research in the area has focused on finding more relaxed conditions for combination that would ensure decidability of the combined theories. Tinelli and Zarba introduced in [19] the notion of *shiny theories* (see Definition 4) and proved that the disjoint combination of one shiny theory with an arbitrary (that is, not necessarily stably infinite) decidable theory is decidable. In [7] we considered the Bernays-Schönfinkel-Ramsey fragment, the two variables fragment, and the monadic fragment. These fragments include non stably infinite theories. We introduced the notion of *gentleness* (see Definition 5) and proved that the disjoint combination of one gentle theory with an arbitrary decidable theory (modulo a minor restriction[1]) is also decidable. All theories in the considered fragments are gentle.

In this article we first investigate the combination of *guarded fragments* of first-order logic with other decidable fragments. Guarded fragments, originally introduced in [1] as first-order counterparts of modal languages, are very expressive. In contrast to other well-known decidable classes, the guarded fragments impose no restriction on the number of variables, alternations of quantifiers, or symbol arity. Instead, quantification is restricted to occur only in guarded form. Relational properties such as symmetry of a relation (written as $\forall xy\,.\,R(x,y) \rightarrow R(y,x)$) can readily be expressed with these fragments, as well as various graph properties such as $\forall xy\,.\,R(x,y) \rightarrow \exists z\,.\,R(y,z)$ stating that every node with an incoming edge has an outgoing edge, or constraints such as $\forall yz\,.\,R(y,y,z) \rightarrow \bot$ which forbids certain kinds of tuples to appear in a relation.

In this article we will show that the guarded fragment [1], the loosely guarded fragment [20] and the packed guarded fragment [15] are shiny, and hence, they can be combined in a decidable way, with an arbitrary decidable theory. This can be seen as further explanation of the good computational behavior of many modal logics [21, 9].

To complete the picture of combination of theories from decidable first-order fragments, we also consider the well-known Ackermann fragment, i.e. formulas of the form $\exists^*\forall x\exists^*\varphi$, where φ is a function- and quantifier-free first-order formula. In this paper we will show that this fragment is gentle and, thus, easily combinable with arbitrary theories (with a minor restriction).

After introducing notations and definitions in Section 2, we discuss combination of decision procedures in the disjoint case in Section 3. Section 4 introduces the guarded fragments we will consider. The status of constants and equality in these fragments is sometimes unclear in the literature; since these are of foremost importance in our context, they will be handled with special care. Section 5 considers the Ackermann fragment. The proofs we present are straightforward but, to our knowledge, this is the first time that these fragments have been explored in the framework of combined theories.

[1] This theory should fall in one of the three cases of Theorem 3. These cases are such that unsuitable theories would be very particular.

2 Notations and Basic Definitions

A first-order language is a tuple $\mathcal{L} = \langle \mathcal{V}, \mathcal{F}, \mathcal{P} \rangle$ such that $\mathcal{V}$ is an enumerable set of variables, while $\mathcal{F}$ and $\mathcal{P}$ are sets of function and predicate symbols. Every function and predicate symbol has an arity. Nullary predicates symbols are called proposition symbols, and nullary function symbols are called constant symbols. A first-order language is called relational if it only contains function symbols of arity zero. A relational formula is a formula in a relational language.

Terms and formulas over the language $\mathcal{L}$ are defined in the usual way. An atomic formula is either an equality statement ($t = t'$) where t and t' are terms, or a predicate symbol applied to the right number of terms. Formulas are built from atomic formulas, Boolean connectives ($\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$), and quantifiers ($\forall$, $\exists$). A literal is an atomic formula or the negation of an atomic formula. The set of free variables $\text{Free}(\varphi)$ in a formula φ is defined as usual. A formula with no free variables is closed. A theory is a set of closed formulas. Two theories are disjoint if no predicate or function symbol appears in both theories; the theories can however share constants.

An interpretation $\mathcal{I}$ for a first-order language $\mathcal{L}$ provides a non empty domain D, a total function $\mathcal{I}[f] : D^r \rightarrow D$ of appropriate arity for every function symbol f, a predicate $\mathcal{I}[p] : D^r \rightarrow \{\top, \bot\}$ of appropriate arity for every predicate symbol p, and an element $\mathcal{I}[x] \in D$ for every variable x. By extension, an interpretation defines a value in D for every term, and a truth value for every formula. The cardinality of an interpretation is the cardinality of its domain. The notation $\mathcal{I}_{x_1/d_1,\ldots,x_n/d_n}$ for $x_1, \ldots, x_n$ different variables stands for the interpretation that agrees with $\mathcal{I}$, except that it associates $d_i \in D$ to the variable x_i, $1 \leq i \leq n$. Given an interpretation $\mathcal{I}$ on domain D, an extension $\mathcal{I}'$ of $\mathcal{I}$ is an interpretation on a domain including D such that $\mathcal{I}'$ restricted to the domain D is exactly $\mathcal{I}$.

A model of a formula (or a theory) is an interpretation in which the formula (resp., every formula in the theory) evaluates to true. A formula or theory is satisfiable if it has a model, and it is unsatisfiable otherwise. A formula G is $\mathcal{T}$-satisfiable if it is satisfiable in the theory $\mathcal{T}$, that is, if $\mathcal{T} \cup \{G\}$ is satisfiable. A $\mathcal{T}$-model of G is a model of $\mathcal{T} \cup \{G\}$. A formula G is $\mathcal{T}$-unsatisfiable if it has no $\mathcal{T}$-models. A decidable theory $\mathcal{T}$ is a theory such that the $\mathcal{T}$-satisfiability problem for sets of literals in the language of $\mathcal{T}$ is decidable.

The bold notation $\boldsymbol{x}$ denotes a tuple, and stands for a sequence of variables or constants (or both) depending on the context. For instance, in $\forall \boldsymbol{x}\, \varphi$, formula φ is quantified universally over all variables in $\boldsymbol{x}$. Expressions such as $p(\boldsymbol{x})$, $p(\boldsymbol{y}, c)$, $p(\boldsymbol{z}, \boldsymbol{d})$ and $\mathcal{I}_{\boldsymbol{x}/\boldsymbol{d}}$, where p is a predicate and $\mathcal{I}$ an interpretation, may be used, with straightforward meaning. When used with set operators, tuples behave like the set of the elements in the tuple, whereas $|\boldsymbol{x}|$ gives the length of the tuple.

3 Combination of Theories

To study the satisfiability of a set of literals like

$$L = \{a \leq b, b \leq a + f(a), P(h(a) - h(b)), \neg P(0), f(a) = 0\}$$

that mixes symbols from the integer linear arithmetic theory $\mathcal{T}_1$ and the theory of uninterpreted symbols $\mathcal{T}_2$, one uses a combination framework to design a decision procedure for the joint language from the simple component decision procedures for one theory only. To divide the above satisfiability problem into problems for the component decision procedures, a *separation* is first built by introducing fresh uninterpreted constants[2], to produce an equisatisfiable problem:

$$L_1 = \{a \leq b,\, b \leq a + v_1,\, v_1 = 0,\, v_2 = v_3 - v_4,\, v_5 = 0\}$$
$$L_2 = \{P(v_2),\, \neg P(v_5),\, v_1 = f(a),\, v_3 = h(a),\, v_4 = h(b)\}.$$

The set L_1 only contains arithmetic symbols and uninterpreted constants. The symbols in L_2 are all uninterpreted. The decision procedure for linear arithmetic and the one for uninterpreted symbols can thus handle the sets L_1 and L_2 respectively. However, although L is unsatisfiable in $\mathcal{T}_1 \cup \mathcal{T}_2$, L_1 is $\mathcal{T}_1$-satisfiable, and L_2 is $\mathcal{T}_2$-satisfiable; it is not sufficient for the decision procedures for $\mathcal{T}_1$ and $\mathcal{T}_2$ to only examine the satisfiability of their part of the separation. Indeed, the decision procedures also have to "agree" on the symbols that are shared, namely the uninterpreted constants in the set $S = \{a, b, v_1, v_2, v_3, v_4, v_5\}$. In order to make sure that both decision procedures will interpret those shared symbols coherently, the notion of arrangement is useful:

Definition 1. *An arrangement $\mathcal{A}$ for a set of constant symbols S is a maximal satisfiable set of equalities and inequalities $a = b$ or $a \neq b$, with $a, b \in S$.*

That is, an arrangement $\mathcal{A}$ for S cannot be extended with any equality or inequality over S and remain consistent.

The following theorem (other formulations can be found in [18, 19, 8]) then states the completeness of the combination of decision procedures:

Theorem 1. *Assume $\mathcal{T}_1$ and $\mathcal{T}_2$ are theories over the disjoint languages $\mathcal{L}_1$ and $\mathcal{L}_2$, and L_i ($i = 1, 2$) is a set of literals in $\mathcal{L}_i$ augmented by a finite set of fresh constant symbols S. Then $L_1 \cup L_2$ is $\mathcal{T}_1 \cup \mathcal{T}_2$-satisfiable if and only if there exists an arrangement $\mathcal{A}$ of S, a cardinality k, and two models $\mathcal{M}_1$ and $\mathcal{M}_2$ of cardinality k, such that $\mathcal{M}_1$ is a $\mathcal{T}_1$-model of $\mathcal{A} \cup L_1$ and $\mathcal{M}_2$ is a $\mathcal{T}_2$-model of $\mathcal{A} \cup L_2$.*

Intuitively, if a set of literals is satisfiable in the combination of theories, a model of this set defines in a straightforward way an arrangement and two models with the same cardinality for the two parts of the separation. The converse is also true: from models of the two parts of the separation (augmented with the arrangement), it is possible to build a unique model for both parts, since both models agree on the cardinality, and on the interpretation of the shared constants in S (thanks to the arrangement). The cardinality condition is essential to be able to map elements in both domains together into a unique domain.

[2] Traditionally, combination schemes use variables for this role. Since variables will be used in quantifiers in the following sections, for consistency and clarity we will rather use uninterpreted constants here.

Relying on the above theorem, an algorithm implementing a satisfiability decision procedure for the combination of two disjoint decidable theories $\mathcal{T}_1$ and $\mathcal{T}_2$ could be as follows:

1. Build a separation (L_1, L_2) for the set of literals L which mix symbols from $\mathcal{T}_1$ and $\mathcal{T}_2$. L_1 contains symbols from $\mathcal{T}_1$ only and symbols from a finite set of fresh constant symbols S, and likewise for L_2;
2. Guess an arrangement $\mathcal{A}$ for the set of constants shared between L_1 and L_2;
3. If $\mathcal{A} \cup L_i$ is $\mathcal{T}_i$-satisfiable for $i = 1, 2$, then, if there exists a (finite or infinite) cardinality k such that $\mathcal{A} \cup L_i$ has a model of cardinality k for $i = 1, 2$, then L is $\mathcal{T}_1 \cup \mathcal{T}_2$-satisfiable. Otherwise, $\mathcal{A} \cup L$ is $\mathcal{T}_1 \cup \mathcal{T}_2$-unsatisfiable.

If we want to ensure that the above algorithm is indeed a decision procedure, two issues needs to be solved. First, as we presented it above the algorithm is non-deterministic but this is not fundamental. Since the number of arrangements for a fixed finite set of constants is finite (although large), the non-deterministic choice can be turned into a loop over this set. The second issue is, however, essential. It involves being able to compare the cardinalities of the models for both parts of the arrangement. To handle this problem, combination of decision procedures and SMT solvers usually consider only stably infinite theories:

Definition 2. *A theory $\mathcal{T}$ is said to be stably infinite when every $\mathcal{T}$-satisfiable set of literals has a model with cardinality $\aleph_0$.*

By definition, when dealing with stably infinite theories, if both parts of the separation are satisfiable in their corresponding theory, then both have an infinite model of cardinality $\aleph_0$.

Consider again the above example. As both the theory for uninterpreted symbols and the theory of integer linear arithmetic are stably infinite, the set of literals L in our example is $\mathcal{T}_1 \cup \mathcal{T}_2$-satisfiable if and only if there exists an arrangement $\mathcal{A}$ of the seven variables in S such that $\mathcal{A} \cup L_i$ is $\mathcal{T}_i$-satisfiable for $i = 1$ and $i = 2$. No such arrangements exist. Indeed, consider an arrangement $\mathcal{A}$ such that $\mathcal{A} \cup L_1$ is $\mathcal{T}_1$-satisfiable and $\mathcal{A} \cup L_2$ is $\mathcal{T}_2$-satisfiable. Such an arrangement contains $a = b$, otherwise $\mathcal{A} \cup L_1$ would not be $\mathcal{T}_1$-satisfiable. It also contains $v_3 = v_4$ since $\mathcal{A} \cup L_2$ is $\mathcal{T}_2$-satisfiable, and as a consequence, $v_2 = v_5$ should also be in $\mathcal{A}$. But if $\mathcal{A}$ contains $v_2 = v_5$, $\mathcal{A} \cup L_2$ is not $\mathcal{T}_2$-satisfiable.

Considering stably infinite theories only is one way to fulfill the cardinality requirement for disjoint combination. It is, however, very restrictive. While some very useful theories are stably infinite, many are not. For instance, there exist theories that only have finite models. Many first-order decidable classes allow to write formulas that constrain the cardinality of the models. Consider, for example, the Ackermann theory $\varphi = \{\forall x \,.\, p(c) \rightarrow (x = a \vee x = b)\}$ that requires the cardinality of the model to be at most two whenever $p(c)$ is true. While $\varphi \cup \{\neg p(c)\}$ does have infinite models, $\varphi \cup \{p(c)\}$ only has finite models.

Of course, there are other ways to ensure that the cardinality requirement is fulfilled. They allow to build decision procedures for union of theories that are

not all stably infinite. To examine the cardinality requirements in Theorem 1, the notion of spectrum[3] is convenient

Definition 3. *The* spectrum *of a theory $\mathcal{T}$ is the set of cardinalities k such that $\mathcal{T}$ is satisfiable in a model of cardinality k.*

Theorem 1 now becomes: $L_1 \cup L_2$ is $\mathcal{T}_1 \cup \mathcal{T}_2$-satisfiable if and only if there exists an arrangement $\mathcal{A}$ of S, such that the spectra of $\mathcal{T}_1 \cup \mathcal{A} \cup L_1$ and $\mathcal{T}_2 \cup \mathcal{A} \cup L_2$ have a non-empty intersection. The intersection of spectra is the crucial difficulty for combination frameworks. Fortunately, the spectrum for many theories (as we will see for guarded fragments and the Ackermann class) is such that the computation of the intersection with another spectrum is easy.

Some theories (e.g., the empty theory, the theory of partial orders, the theory of total orders) have spectral properties that allow combination with any other decidable disjoint theory; these are called *shiny* theories [19].

Definition 4. *A decidable theory $\mathcal{T}$ is* shiny *if, for every $\mathcal{T}$-satisfiable set of literals L, there is a finite computable number k such that the spectrum of $\mathcal{T} \cup L$ is the set of cardinalities greater than or equal to k.*

In the following sections, we show that the guarded, the loosely guarded, and the packed guarded fragments are all shiny. They can thus be combined with any disjoint theory:

Theorem 2. *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two disjoint decidable theories sharing only constants. If $\mathcal{T}_1$ is shiny then $\mathcal{T}_1 \cup \mathcal{T}_2$ is decidable.*

In [7], we show that theories in the Bernays-Schönfinkel-Ramsey class, the two variables fragment, and the monadic fragment have interesting spectral properties, though weaker than shininess. Every theory $\mathcal{T}$ in these classes is gentle:

Definition 5. *A theory $\mathcal{T}$ is* gentle *if, for every $\mathcal{T}$-satisfiable set of literals L, the spectrum of $\mathcal{T} \cup L$ can be computed and is either*

- *a finite set of finite cardinalities*
- *the union of a finite set of finite cardinalities and all the (finite and infinite) cardinalities greater than a computable finite cardinality; it is thus co-finite.*

The definition of shininess and gentleness are quite similar; considering only sufficiently large cardinalities, both notions express the same property. Notice that a shiny theory is also gentle. Furthermore, the union of disjoint gentle theories is also a gentle theory [7]. Some widely used theories are not gentle, but in practical cases they can be combined with gentle theories [7]:

Theorem 3. *Given a gentle theory $\mathcal{T}$ and another disjoint theory $\mathcal{T}'$, the $\mathcal{T} \cup \mathcal{T}'$-satisfiability problem for sets of literals written in the union of their language is decidable if one of the following cases holds:*

[3] The spectrum of a theory is usually defined as the set of the *finite* cardinalities of its models. We here slightly extend the definition for convenience.

– $\mathcal{T}'$ *is gentle;*
– $\mathcal{T}'$ *is a decidable finitely axiomatized first-order theory;*
– $\mathcal{T}'$ *is a decidable theory that only admits a fixed finite (possibly empty) known set of finite cardinalities for its models, and possibly infinite models.*

In the next sections we will prove that guarded fragments are shiny, and that the Ackermann theories are gentle.

4 The Guarded Fragments

The guarded fragment (GF) was originally introduced in [1] as a suitable counterpart and generalization of modal logics. To make this article self contained, let us start with a brief recapitulation of modal logics (see [3, 4] for further details). Consider the language defined as

$$\mathcal{BML} := p_i \mid \neg\varphi \mid \varphi \vee \psi \mid \Diamond\varphi,$$

where p_i is a propositional symbol and $\varphi, \psi \in \mathcal{BML}$. Syntactically, the language $\mathcal{BML}$ (the basic modal language) is a slight extension of propositional logic (we have only added the unary operator $\Diamond$). Semantically, on the other hand, the change is radical. We interpret formulas of $\mathcal{BML}$ on first-order relational models $M = \langle D, \mathcal{I}\rangle$ over a signature with only one binary relational symbol R and uncountably many propositional symbols $\{p_1, p_2, \ldots\}$. Given such a model M and an element a in the domain, semantics is defined as follows:

$$\begin{aligned}
M[p](a) = \top &\text{ iff } \mathcal{I}[p](a)\\
M[\neg\varphi](a) = \top &\text{ iff } M[\varphi](a) = \bot\\
M[\varphi \vee \psi](a) = \top &\text{ iff } M[\varphi](a) = \top \text{ or } M[\psi](a) = \top\\
M[\Diamond\varphi](a) = \top &\text{ iff for some } b \in D,\ \mathcal{I}[R](a, b) \text{ and } M[\varphi](b) = \top.
\end{aligned}$$

These semantic conditions should tip us off on the close connections between modal and first-order languages. Indeed, it is simple to define an equivalence preserving translation from the former to the latter. Define recursively the translation Tr_x for x a first-order variable as:

$$\begin{aligned}
\mathsf{Tr}_x(p) &= P(x)\\
\mathsf{Tr}_x(\neg\varphi) &= \neg\mathsf{Tr}_x(\varphi)\\
\mathsf{Tr}_x(\varphi \vee \psi) &= \mathsf{Tr}_x(\varphi) \vee \mathsf{Tr}_x(\psi)\\
\mathsf{Tr}_x(\Diamond\varphi) &= \exists y.R(x, y) \wedge \mathsf{Tr}_y(\varphi),
\end{aligned}$$

where y is a new variable, not yet used in the translation. A simple induction shows that for any formula $\varphi \in \mathcal{BML}$, any model M (in the proper signature) and any element a in the domain of M, $M[\varphi](a) = \top$ iff $M[\mathsf{Tr}_x(\varphi)](a) = \top$. In other words, $\mathcal{BML}$ can be seen as nothing else than a fragment of first-order logic in disguise. But $\mathcal{BML}$ is only *one* among many modal logics. Other modal operators such as the inverse modality, the universal modality, the difference modality, etc. can be defined (see [4] for details). Most of them can be translated

into first-order logic preserving equivalence. A natural question is then, whether it is possible to define a fragment of first-order logic that can be the range of these translations, and that will preserve the common modal aspects of all these logical languages. The answer to this question was the guarded fragment GF.

Definition 6. *A formula* γ guards *another formula* φ *if every free variable of* φ *also occurs free in* γ *(i.e.,* $\mathrm{Free}(\varphi) \subseteq \mathrm{Free}(\gamma)$*).*

Definition 7. *A formula in the* guarded fragment GF *of first-order logic is a relational formula such that all quantified sub-formulas are of the form* $\forall \boldsymbol{x} \,.\, \gamma \to \psi$ *or* $\exists \boldsymbol{x} \,.\, \gamma \wedge \psi$ *where*

- γ *is an atom, but not an equality,*
- ψ *is guarded by* γ*,*
- $\boldsymbol{x}$ *is a tuple of variables in* $\mathrm{Free}(\gamma)$*,*

The atom γ *is called the* guard*.*

Formulas in the fragment might contain an arbitrary number of variables (i.e., GF is not contained in any finite variable fragment of first-order logic). Similarly, formulas in GF might contain an arbitrary number of quantifier alternations, and hence they cannot be defined in terms of prenex normal form prefixes. Also, the arity of relational symbols is not bounded. Moreover, many natural properties expressible in first-order logic fall in GF. Some examples, besides those we mentioned in Section 1, are:

$\forall x \,.\, R(x,x)$	reflexivity
$\forall x \,.\, \neg R(x,x)$	irreflexivity
$\exists v_1 \,.\, (R(a,v_1) \wedge \exists v_2 \,.\, (R(v_1,v_2) \wedge R(v_2,b)))$	there is a path of length 3 between a and b

On the other hand, some simple formulas, such as transitivity $\forall xyz \,.\, (R(x,y) \wedge R(y,z)) \to R(x,z)$, are neither in GF nor equivalent to any formula of GF (i.e., transitivity is not expressible in GF).

Guarded fragments have been defined and redefined repeatedly, looking for the largest fragment of first-order logic with a nice 'modal' behavior. The original definition of [1] contained the restriction on equality atoms not appearing in guards we introduced above. This restriction was later removed (even though the exact status of equality in the different definitions of guarded fragments is sometimes unclear), but it is crucial for the results we will present.

Suppose we eliminate this restriction. Then equality atoms could occur as a guard in one of the following shapes (let's consider only universal quantification):

1. $\forall x \,.\, x = x \to \psi(x)$
2. $\forall x \,.\, x = y \to \psi(x,y)$
3. $\forall xy \,.\, x = y \to \psi(x,y)$
4. $\forall x \,.\, x = c \to \psi(x)$

Cases 2 and 4 can be rewritten as $\psi(y, y)$ and $\psi(c)$, respectively, eliminating the quantifier and resulting in a formula in GF. Cases 1 and 3 rewrite to $\forall x \,.\, \psi(x)$ and $\forall x \,.\, \psi(x, x)$, respectively. The resulting formulas in the scope of the quantifier contain at most one free variable, but this variable is not guarded. Without the restriction on the use of equality in guards GF would include formulas such as $\forall x \,.\, x = a_1 \vee \ldots x = a_n$ that restricts the domain to a finite cardinality smaller or equal to n. These improper guarded formulas would invalidate the good properties necessary for combining GF theories (see Corollary 1 below).

Many good properties of GF are shown in [1]. In particular, the authors prove that its satisfiability problem is decidable (it is actually 2ExpTime-complete, and only ExpTime-complete if the number of variables is bounded by any finite number k, see [12]), and that the fragment has the finite-model property (i.e., every satisfiable formula is satisfied in a finite model).

Different variations of GF were introduced, gradually relaxing the conditions imposed on the guard to obtain larger fragments. We present the loosely guarded fragment introduced in [20], and the packed guarded fragment introduced in [15].

Definition 8. *A formula in the* loosely guarded fragment LGF *of first-order logic is a relational formula such that all quantified sub-formulas are of the form* $\forall \boldsymbol{x} \,.\, \gamma \rightarrow \psi$ *or* $\exists \boldsymbol{x} \,.\, \gamma \wedge \psi$ *where*

- $\gamma = \alpha_1 \wedge \cdots \wedge \alpha_m$ *is an equality-free conjunction of atoms,*
- ψ *is guarded by* γ*,*
- *for every variable* y *in* $\boldsymbol{x}$ *and every variable* $z \in \mathrm{Free}(\gamma)$ *with* $y \neq z$*, there is at least one atom* α_j *that contains both* y *and* z

The conjunction of atoms γ *is called the* guard*.*

Notice that GF is a proper subset of LGF. The loosely guarded fragment is decidable [20] and has the finite model property [14]. Its satisfiability problem is 2ExpTime-complete [12].

Definition 9. *A formula in the* packed guarded fragment PGF *of first-order logic is a relational formula such that all quantified sub-formulas are of the form* $\forall \boldsymbol{x} \,.\, \gamma \rightarrow \psi$ *or* $\exists \boldsymbol{x} \,.\, \gamma \wedge \psi$ *where*

- $\gamma = \alpha_1 \wedge \cdots \wedge \alpha_m$ *is an equality-free conjunction of atoms and existentially-quantified atomic formulas,*
- ψ *is guarded by* γ*,*
- *for every variables* $y, z \in \mathrm{Free}(\gamma)$ *there is at least one conjunct* α_j *such that* $\{y, z\} \subseteq \mathrm{Free}(\alpha_j)$

The conjunction γ *is called the* guard*.*

Although LGF is not a subset of PGF, PGF is (strictly) more expressive than LGF: any LGF formula can be rewritten to a logically equivalent PGF formula (see [10]). The packed guarded fragment is also known as the clique-guarded fragment. Both definitions are equivalent [10]. The packed guarded fragment is

decidable and has the finite model property [14]. The satisfiability problem for PGF is 2ExpTime-complete [13].

The status of constants in guarded fragments has sometimes been vague. Constants are crucial for our goal, as they will be used to link the combined theories. Notice that in our definitions, all guarded fragments allow constants. The following theorem, adapted from [12], shows that constants can always be added to guarded fragments without interfering with decidability, the finite model property or complexity.

Theorem 4. *Adding constants to the languages for* GF, LGF *and* PGF, *preserves decidability, the finite model property, and complexity.*

Proof. Assume φ is a formula in GF, LGF or PGF with constants from a finite set C. Let $\boldsymbol{c}$ be a sequence containing all constants in C. Let G be the set of all predicates occurring in guards (remember that guards are equality free, so G does not include equality). For every n-ary predicate $p \in G$, let p' be a fresh $(n+|\boldsymbol{c}|)$-ary predicate. The formula φ' is built from φ by replacing every occurrence $p(\boldsymbol{x})$, for every $p \in G$ and every sequence of variables and constants $\boldsymbol{x}$ by $p'(\boldsymbol{x}, \boldsymbol{c})$. Let Z be a fresh $|\boldsymbol{c}|$-ary predicate. The formula $\psi = \exists \boldsymbol{c}\,(Z(\boldsymbol{c}) \wedge \varphi')$ — where the constants $\boldsymbol{c}$ in φ are variables in ψ — is equisatisfiable to φ. From a model of φ it is possible to build a model on the same domain for ψ, and conversely, thus the finite model property (and consequently, decidability) is preserved. ψ is constant-free, and it is properly guarded (in the same fragment GF, LGF or PGF than φ). Replacing constants by variables may involve a polynomial growth of the formula. This does not affect the 2ExpTime-complete complexity. □

4.1 The Spectra of Guarded Fragments

The following theorem states that an interpretation of a formula in the guarded fragments GF, LGF or PGF, can always be extended by new elements without changing the truth value of the considered formula. Intuitively, it suffices for those new elements to be "disconnected" from the other elements, that is, those new elements make every guard false. This, together with the finite model property, will directly imply that these fragments are shiny.

Theorem 5. *Given any interpretation M on domain D for a formula φ in* GF, LGF *or* PGF, *then for every $D' \supset D$ there is an extension M' of M on domain D' such that $M'[\varphi] = M[\varphi]$.*

Proof. Given an interpretation M on domain D for a formula φ in GF, LGF or PGF, the interpretation M' on D' is defined as follows:

- for every constant a, $M'[a] = M[a]$;
- for every variable $x \in \text{Free}(\varphi)$, $M'[x] = M[x]$;
- for every n-ary predicate p, and for $a_i \in D'$ $(1 \leq i \leq n)$
  - $M'[p](a_1, \ldots a_n) = M[p](a_1, \ldots a_n)$ if $a_i \in D$ for all i $(1 \leq i \leq n)$;
  - $M'[p](a_1, \ldots a_n) = \bot$ otherwise.

To be able to handle PGF as the two other fragments in the following, first consider an existentially-quantified atomic formula $\gamma = \exists \boldsymbol{x} \,.\, p(\boldsymbol{y})$. Notice that (1) for any interpretation M'' defined as M' is above, but assigning at least one free variable of γ to an element in $D' \setminus D$, $M''[\gamma] = \bot$ (2) for any interpretation M' defined as above, $M'[\gamma] = M[\gamma]$. The first point is direct. To prove the second, notice that if $M[\gamma] = \top$, then $M_{\boldsymbol{x}/\boldsymbol{d}}[p(\boldsymbol{y})] = \top$ for some tuple $\boldsymbol{d}$ of elements in D. Then $M'_{\boldsymbol{x}/\boldsymbol{d}}[p(\boldsymbol{y})]$ is also true and as a consequence, $M'[\gamma] = \top$. If $M[\gamma] = \bot$, notice that, for any tuple $\boldsymbol{d}$ of elements in D', $M'_{\boldsymbol{x}/\boldsymbol{d}}[p(\boldsymbol{y})] = M_{\boldsymbol{x}/\boldsymbol{d}}[p(\boldsymbol{y})] = \bot$ if all arguments of p are assigned to elements in D, and $M'_{\boldsymbol{x}/\boldsymbol{d}}[p(\boldsymbol{y})] = \bot$ if one argument of p is in $D' \setminus D$. As a consequence $M'[\gamma] = \bot$.

Theorem 5 is proved by showing by structural induction that $M'[\varphi] = M[\varphi]$, for M' defined from M as above. It is trivial if φ is atomic, a negation, or a Boolean combination of several formulas. The only remaining cases are the quantified constructions.

Let $\varphi = \forall x_1 \ldots x_n \,.\, \gamma \to \psi$ (where γ is the guard) belong to GF, LGF or PGF. For simplicity and without loss of generality, assume that $x_i \in \text{Free}(\gamma \to \psi)$ for every $i \in \{1, \ldots n\}$. Consider an interpretation M on domain D for φ, $D' \supset D$, and M' as defined above. For $d_1, \ldots d_n \in D'$ one of the two cases hold:

- if $d_i \in D' \setminus D$ for some $i \in \{1, \ldots n\}$, then $M'_{x_1/d_1, \ldots x_n/d_n}[\gamma] = \bot$, and hence $M'_{x_1/d_1, \ldots x_n/d_n}[\gamma \to \psi] = \top$. Indeed, since γ is a guard, x_i appears free in γ. Since the guard is either (GF) an atom, (LGF) a conjunction of atoms, (PGF) or a conjunction of atoms and existentially quantified atoms, the atom having x_i as an argument is interpreted as false, and so is the whole guard.
- if $d_i \in D$ for all $i \in \{1, \ldots n\}$, then $M'_{x_1/d_1, \ldots x_n/d_n}$ and $M_{x_1/d_1, \ldots x_n/d_n}$ agree on $(\gamma \to \psi)$, i.e., $M'_{x_1/d_1, \ldots x_n/d_n}[\gamma \to \psi] = M_{x_1/d_1, \ldots x_n/d_n}[\gamma \to \psi]$. Indeed, by the inductive hypothesis, $M'_{x_1/d_1, \ldots x_n/d_n}[\psi] = M_{x_1/d_1, \ldots x_n/d_n}[\psi]$, for all $d_1, \ldots d_n \in D$. Furthermore, for all $d_1, \ldots d_n \in D$ then $M'_{x_1/d_1, \ldots x_n/d_n}[\gamma] = M_{x_1/d_1, \ldots x_n/d_n}[\gamma]$. This is trivial for GF and LGF thanks to the inductive hypothesis, since guards are Boolean combinations of atoms. This is also true for PGF, given the previous remarks on existentially-quantified atomic formulas. Hence, $M'_{x_1/d_1, \ldots x_n/d_n}[\gamma \to \psi] = M_{x_1/d_1, \ldots x_n/d_n}[\gamma \to \psi]$.

It follows that $M'[\varphi] = M[\varphi]$. The existential case is handled similarly. □

Corollary 1. *Any theory in* GF, LGF, *or* PGF *is shiny.*

Proof. Assume $\mathcal{T}$ is a theory in GF, LGF, or PGF. For any set of literals L in the language of $\mathcal{T}$, $\mathcal{T} \cup L$ is also a theory in GF, LGF, or PGF. Thanks to the finite model property of GF, LGF, and PGF, if $\mathcal{T} \cup L$ is satisfiable, it has a finite model. It is thus possible to compute the minimum cardinality of $\mathcal{T} \cup L$. Furthermore, thanks to the previous theorem, its spectrum is an unbounded interval. □

5 The Ackermann Class

The Ackermann class (with equality) is the set of formulas of the form

$$\exists \, . \, z_1 \cdots \exists z_n \, . \, \forall x \, . \, \exists y_1 \, . \, \cdots \exists y_m \, . \, \varphi(x, y_1, \ldots, y_m, z_1, \ldots, z_n),$$

where $\varphi(x, y_1, \ldots, y_m, z_1, \ldots, z_n)$ is quantifier-free and function-free. Checking the satisfiability of formulas of the above form can be reduced (using Skolemization) to checking the satisfiability of formulas without existental quantifiers of the form $\psi = \forall x \, . \, \varphi(x, f_1(x), \ldots, f_m(x))$, where $\varphi(x, y_1, \ldots, y_m)$ is quantifier-free and function-free.

Theorem 6. *The class of formulas of the form $\psi = \forall x \, . \, \varphi(x, f_1(x), \ldots, f_m(x))$, where $\varphi(x, y_1, \ldots, y_m)$ is quantifier and function-free (constants are allowed) has the finite model property.*

The proof may be found for instance in [5]. The following theorem will allow to determine that the Ackermann theories are gentle. An equivalent property for the Ackermann fragment is discussed in [6].

Theorem 7. *Consider a formula $\psi = \forall x \, . \, \varphi(x, f_1(x), \ldots, f_m(x))$, where formula $\varphi(x, y_1, \ldots, y_m)$ is quantifier and function-free (constants are allowed). If ψ has a model of cardinality κ strictly greater than the number of constants in ψ, then it has models with any cardinality greater than κ.*

Proof. Consider a model $\mathcal{M}$ of ψ on domain D such that $|D|$ is greater than the number of constants in ψ. Then there exists an extension $\mathcal{M}'$ on any domain D' with $D \subset D'$ that is also a model of ψ.

Let $\Phi(x) = \varphi(x, f_1(x), \ldots, f_m(x))$ and $d \in D$ be an element of the domain, not assigned by $\mathcal{M}$ to a constant in the formula. Obviously $\mathcal{M}_{x/d}$ is a model of $\Phi(x)$. Consider $d' \in D' \setminus D$. For every n-ary predicate p, and n-uple $\mathbf{d}'$ of elements in $(D \setminus \{d\}) \cup \{d'\}$, let $\mathbf{d}$ be a n-uple of elements in D obtained from $\mathbf{d}'$ by changing d' by d whenever d' is an element of the tuple $\mathbf{d}'$, and set $\mathcal{M}'[p](\mathbf{d}') = \mathcal{M}[p](\mathbf{d})$. For every function f_i $(1 \leq i \leq m)$ let $\mathcal{M}'[f_i](d') = \mathcal{M}[f](d)$ if $\mathcal{M}[f](d) \neq d$, and let $\mathcal{M}'[f_i](d') = d'$ otherwise. Functions and predicates are only partially defined above, but they can be completed arbitrarily without any influence on the result. One can show by structural induction that $\mathcal{M}'_{x/d'}[\Phi(x)] = \mathcal{M}'_{x/d}[\Phi(x)]$.

Indeed, according to our definition of $\mathcal{M}'$,

- $\mathcal{M}'_{x/d'}[x] = d'$ whereas $\mathcal{M}_{x/d}[x] = d$,
- $\mathcal{M}'_{x/d'}[c] = \mathcal{M}_{x/d}[c]$ for every constant c in $\Phi(x)$,
- $\mathcal{M}'_{x/d'}[f(x)] = \mathcal{M}_{x/d}[f(x)]$ if $\mathcal{M}_{x/d}[f(x)] \neq d$,
- $\mathcal{M}'_{x/d'}[f(x)] = d'$ if $\mathcal{M}_{x/d}[f(x)] = d$.

Thus, for every atom $p(t_1, \ldots, t_n)$ (respectively, $t_1 = t_2$) in $\Phi(x)$, $\mathcal{M}'_{x/d'}$ and $\mathcal{M}'_{x/d}$ assign the same values to every t_i except that $\mathcal{M}'_{x/d'}$ assigns d' instead of

d for $\mathcal{M}'_{x/d}$. Finally, thanks to the way $\mathcal{M}'$ extends the assignment of predicates, $\mathcal{M}'_{x'/d'}[p(t_1,\ldots,t_n)] = \mathcal{M}_{x/d}[p(t_1,\ldots,t_n)]$ (respectively, $\mathcal{M}'_{x'/d'}[t_1 = t_2] = \mathcal{M}_{x/d}[t_1 = t_2]$). □

Corollary 2. *The spectrum of an Ackermann theory can be computed and expressed either as a finite set of natural numbers, or as the union of a finite set of natural numbers with the set of all the (finite or infinite) cardinalities greater than a natural. The Ackermann theories are gentle.*

Proof. Given $\psi = \forall x\,.\,\varphi(x, f_1(x),\ldots,f_m(x))$, where formula $\varphi(x, y_1,\ldots,y_m)$ is quantifier and function-free, it is possible to establish if ψ has a model of cardinality greater than the number n of constants in ψ. Indeed, formula $\psi' = \psi \wedge \bigwedge_{0\le i<j\le n} a_i \neq a_j$ (where the a_i's are fresh constants) is also in the decidable Ackermann class, and is satisfiable if and only if ψ has a model of cardinality greater than or equal to $n+1$. If ψ' is unsatisfiable, ψ has no model of cardinality greater than or equal to $n+1$. If ψ' is satisfiable, a decision procedure to get the smallest cardinality $m > n$ of the models of ψ can just be a simple test of the (finite) interpretations of increasing cardinality size starting from $n+1$; this procedure will indeed terminate, and ψ will accept models for every cardinality greater than or equal to m. It then only remains to check if ψ accepts models for the cardinalities between 1 and n, which can be done by considering the finitely many interpretations. □

6 Conclusions

The first frameworks to combine disjoint decidable theories were very restrictive: the combined theories were required to be stably infinite. Later results led to more liberal frameworks. In particular, it was proved in [19] that shiny theories are combinable with any other disjoint decidable theory. We have showed that any theory in the guarded fragment, in the loosely guarded fragment, or in the packed guarded fragment, is shiny.

Another well known decidable class with equality (the only relevant classes in our context) that was not yet proved to have good combining properties is the Ackermann class. We showed here that, although not shiny, Ackermann theories are gentle, and, as such, are combinable with non-stably infinite theories with minor requirements. Together with [7], this work then covers the major first-order decidable classes. Interestingly, *all of them are at least gentle.*

The Rabin and the Shelah classes are, respectively, extensions of the Löwenheim class (studied in [7]) and the Ackermann class (studied here), with one unary function. Both are decidable [5]. However, both have infinity axioms [5], and they also contain formulas restricting the cardinality of their models to a finite number. Hence, they are neither shiny nor gentle, and not even stably infinite. It is still an open problem whether these classes have spectral properties that allow liberal combinations. A solution to this problem would probably involve more complex combination frameworks than those discussed in this paper.

Guarded fragments have been extended beyond PGF, even to include fragments of second order logic. The fixed point guarded fragment μGF, for example, was introduced in [11] extending GF with fixed point operators. But unlike GF, LGF, and PGF, μGF even though decidable, does not have the finite model property, and hence it is not shiny. We conjecture, though, that Theorem 5 can be extended to μGF proving it stably infinite.

Our motivation here was mainly to study the decidability of combinations of disjoint theories, without having a practical applications in mind. However, the guarded fragments are highly promising from the point of view of applications. Indeed, since they can easily express graph properties, we beleive implementations will trigger concrete applications. As a toy example of what can be handled by a combination with the guarded fragments, consider the conjunction of the following formulas[4]:

$$\forall x\, y\, .\, R(x,y) \rightarrow \forall z\, .\, \big(R(y,z) \wedge R(z,x)\big) \rightarrow (x = y \vee y = z \vee z = x)$$
$$R(a,b) \wedge R(b,c) \wedge R(a,c)$$
$$f(b) = f(a) + 1 \wedge f(c) = f(b) + 1$$

This set of formulas is unsatisfiable: the first formula enforces 3-edges loop to have at least one reflexive edge, the second states the existence of a 3-edge loop through a, b and c, and the last formula (using uninterpreted function f and some arithmetic) enforces a, b and c to be distinct, which leads to a contradiction. This formula can be dealt with a classical Nelson-Oppen combination framework since all theories are stably-infinite.

In [22], the authors show that it is possible to combine non-disjoint theories from various decidable classes, those theories sharing monadic predicates. This results in a very expressive language. A future direction for research will be to study if the guarded fragments can also be included in such a framework for combining *non-disjoint* theories.

Acknowledgment: we would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

1. H. Andréka, I. Németi, and J. van Benthem. Modal logics and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27(3):217–274, June 1998.
2. C. Barrett, R. Sebastiani, S. A. Seshia, and C. Tinelli. Satisfiability modulo theories. In A. Biere, M. J. H. Heule, H. van Maaren, and T. Walsh, editors, *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 26, pages 825–885. IOS Press, Feb. 2009.
3. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

[4] The first formula can also be written in the Bernays-Schönfinkel-Ramsey class, which contains only gentle theories. It is thus more convenient to consider this formula as in the LGF fragment, containing only shiny theory

4. P. Blackburn, F. Wolter, and J. van Benthem, editors. *Handbook of Modal Logics*. Elsevier, 2006.
5. E. Börger, E. Grädel, and Y. Gurevich. *The Classical Decision Problem*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1997.
6. B. Dreben and W. D. Goldfarb. *The Decision Problem: Solvable Classes of Quantificational Formulas*. Addison-Wesley, Reading, Massachusetts, 1979.
7. P. Fontaine. Combinations of theories for decidable fragments of first-order logic. In S. Ghilardi and R. Sebastiani, editors, *Frontiers of Combining Systems (FroCoS)*, volume 5749 of *LNCS*, pages 263–278. Springer Verlag, 2009.
8. P. Fontaine and E. P. Gribomont. Combining non-stably infinite, non-first order theories. In W. Ahrendt, P. Baumgartner, H. de Nivelle, S. Ranise, and C. Tinelli, editors, *Selected Papers from the Workshops on Disproving and the Second International Workshop on Pragmatics of Decision Procedures (PDPAR 2004)*, volume 125 of *ENTCS*, pages 37–51, July 2005.
9. E. Grädel. Why are modal logics so robustly decidable? In *Current Trends in Theoretical Computer Science. Entering the 21st Century*, pages 393–408. World Scientific, 2001.
10. E. Grädel. Guarded fixed point logics and the monadic theory of countable trees. *Theoretical Computer Science*, 288(1):129–152, 2002.
11. E. Grädel and I. Walukiewicz. Guarded fixed point logic. In *Logic In Computer Science (LICS)*, pages 45–54, Washington, USA, 1999. IEEE Computer Society.
12. E. Grädel. On the restraining power of guards. *Journal of Symbolic Logic*, 64:1719–1742, 1998.
13. E. Grädel. Decision procedures for guarded logics. In *Automated Deduction – CADE16. Proceedings of 16th International Conference on Automated Deduction, Trento, 1999*, volume 1632 of *LNCS*. Springer, 1999.
14. I. M. Hodkinson. Loosely guarded fragment of first-order logic has the finite model property. *Studia Logica*, 70(2):205–240, 2002.
15. M. Marx. Tolerance logic. *Journal of Logic, Language and Information*, 10(3):353–374, 2001.
16. G. Nelson and D. C. Oppen. Simplifications by cooperating decision procedures. *ACM Transactions on Programming Languages and Systems*, 1(2):245–257, Oct. 1979.
17. C. Tinelli and M. T. Harandi. A new correctness proof of the Nelson–Oppen combination procedure. In F. Baader and K. U. Schulz, editors, *Frontiers of Combining Systems (FroCoS)*, Applied Logic, pages 103–120. Kluwer, Mar. 1996.
18. C. Tinelli and C. Ringeissen. Unions of non-disjoint theories and combinations of satisfiability procedures. *Theoretical Computer Science*, 290(1):291–353, Jan. 2003.
19. C. Tinelli and C. G. Zarba. Combining non-stably infinite theories. *Journal of Automated Reasoning*, 34(3):209–238, 2005.
20. J. van Benthem. Dynamic bits and pieces. Technical Report LP-1997-01, ILLC, University of Amsterdam, Jan. 1997.
21. M. Vardi. Why is modal logic so robustly decidable? In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 31, pages 149–184. AMS, 1997.
22. T. Wies, R. Piskac, and V. Kuncak. Combining theories with shared set operations. In S. Ghilardi and R. Sebastiani, editors, *Frontiers of Combining Systems (FroCoS)*, volume 5749 of *LNCS*, pages 366–382. Springer Verlag, 2009.