Hyperresolution for Guarded Formulae

Lilia Georgieva[1], Ullrich Hustadt[2] and Renate A. Schmidt[1]

[1] Department of Computer Science, University of Manchester
Manchester M13 9PL, United Kingdom, {georgiel,schmidt}@cs.man.ac.uk
[2] Centre for Agent Research and Development, Manchester Metropolitan University
Manchester M1 5GD, United Kingdom, U.Hustadt@doc.mmu.ac.uk

Abstract. This paper investigates the use of hyperresolution as a decision procedure and model builder for guarded formulae. In general hyperresolution is not a decision procedure for the entire guarded fragment. However we show that there are natural fragments which can be decided by hyperresolution. In particular, we prove decidability of hyperresolution with or without splitting for the fragment $\mathrm{GF1}^-$ and point out several ways of extending this fragment without loosing decidability. As hyperresolution is closely related to various tableaux methods the present work is also relevant for tableaux methods. We compare our approach to hypertableaux, and mention the relationship to other clausal classes which are decidable by hyperresolution.

1 Introduction

In [1, 2] Andréka, van Benthem and Neméti investigate whether there exist natural fragments of first-order logic extending the modal fragment which corresponds to basic modal logic (via the relational translation) sharing some or all of the properties of modal logics, including decidability, Craig interpolation, bisimulation invariance, Beth definability, finite model property, and preservation under submodels. They show that the *guarded fragment* (GF) shares all these properties with basic modal logic. Decidability was also shown for various extensions of the guarded fragment, like the loosely guarded fragment [12, 15], guarded fixpoint logic [16], or monadic GF^2 with transitive guards [13]. The various decision procedures exploit the finite model property, use ordered resolution, alternating automata, or embeddings into monadic second-order logic. This is an interesting contrast to the literature on decidable modal logics and description logics, where tableaux-based decision procedures are predominant for testing satisfiability (see for example [10, 14]).

In [22] Lutz, Sattler and Tobies investigate whether tableaux-based decision procedures exist for subclasses of the guarded fragment. In this paper we continue the line of investigation making use, however, of the close correspondence between tableaux-based decision procedure for modal logics and hyperresolution combined with splitting on an encoding of modal formulae in clausal logic, as previously demonstrated in [9, 20], and in [18, 19] for description logics. By using a structure preserving transformation of guarded formulae into clausal form we are able to recast the method of Lutz et al. in a first-order setting using in particular hyperresolution with splitting. In this setting it is immediately clear that the termination result can be extended to a larger class of guarded formulae than the class $\mathrm{GF1}^-$ identified in Lutz et al. Thus one of our aims is to study the extent to which hyperresolution can be used as a decision procedures for reasoning about guarded formulae. Generally hyperresolution is not a decision procedure for the entire guarded fragment. A simple example is provided by the guarded formula $p(y) \land \forall x(p(x) \to \exists z(p(z) \land \top))$ with clausal form $\{p(a), \neg p(x) \lor p(f(x))\}$. The method of proving termination used in [9, 20] in the case of modal logics does not generalise to GF1$^-$. We investigate a different argument adapted from Lutz et al. which takes into consideration the form of the derived clauses. The obtained results are more general than those previously known. Another aim is to study how the method relates to other inference methods such as hypertableaux [8], and how the work fits into the bigger picture of hyperresolution as a decision procedure [11, 21].

The structure of the paper is as follows. We give preliminary definitions in Section 2. Section 3 contains a definition of GF1$^-$ and the schematic form of the corresponding clausal class. Decidability of GF1$^-$ by hyperresolution with splitting is shown in Section 4. In Section 5 we extend the decidability results to clausal classes outside the fragment, in Section 6 we describe related calculi and give an account of other clausal classes, solvable by hyperresolution. The final section is the Conclusion.

2 Preliminaries

First-order variables are denoted by x, y, z, terms are denoted by s, t, u, constants by a, b, functions by f, g, h, predicate symbols by P, Q, G, atoms by A, literals by L, clauses by C, formulae by ϕ, φ and ψ, and sets of clauses by N. An over-line indicates a sequence, for example, $\overline{x}$ will denote a finite sequence of variables and $\overline{s}$ will denote a finite sequence of terms. If $\overline{s} = (s_1, \ldots, s_n)$ then $\overline{f(\overline{s})}$ will denote a sequence of terms of the form $f_k(s_1, \ldots, s_n)$.

For any sequence $\overline{s}$ of terms (or formula ϕ) by $\mathrm{var}(\overline{s})$ (or $\mathrm{var}(\phi)$) we denote the set of variables that occur freely in $\overline{s}$ (or ϕ). We also write $\phi(\overline{x})$ to indicate that the free variable occurring in ϕ are $\overline{x}$. Let X be a set of variables. A variable sequence over X is defined to be a finite sequence $\overline{x} \in X^n$. If $\overline{s}$ and $\overline{t}$ are sequences of terms then $\overline{s} \subseteq \overline{t}$ means that every term in $\overline{s}$ also occurs in $\overline{t}$.

3 The Fragment GF1$^-$

The fragment GF1$^-$ is obtained from the fragment GF1 which was introduced in [2] by restricting the way the variables may occur in guards. A formula φ belongs to GF1 if any quantified subformula ψ of φ has the form $\exists\overline{y}(G(\overline{x}, \overline{y}) \wedge \phi(\overline{y}))$ or $\forall\overline{y}(G(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y}))$. In formulae of GF1$^-$ the atoms $G(\overline{x}, \overline{y})$ in guard positions need to satisfy an additional grouping condition.

Formally, every n-ary predicate symbol P is associated with a unique pair (i, j) of positive integers such that $0 < i, j$, and $i + j = n$, which is called the *grouping* of the predicate symbol. Often we will write $P^{(i,j)}$ to make P's grouping explicit.

The set of formulae in GF1$^-$ is defined to be the smallest set satisfying the following conditions: (i) $\top$ and $\bot$ are GF1$^-$ formulae, (ii) if P is an n-ary predicate symbol and $\overline{x}$ is a sequence of n variables, then $P(\overline{x})$ is a GF1$^-$ formula, (iii) if ϕ and ψ are GF1$^-$ formulae then so are $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, and (iv) if $\phi(\overline{y})$ is a GF1$^-$ formula, $G^{(i,j)}$ is a predicate symbol with grouping (i, j), and $\overline{x}, \overline{y}$ are non-empty variable sequences of length i and j with no variables in common, then the following formulae are GF1$^-$ formulae:

$$\exists\overline{y}(G^{(i,j)}(\overline{x}, \overline{y}) \wedge \phi(\overline{y})) \qquad \forall\overline{y}(G^{(i,j)}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y}))$$
$$\exists\overline{x}(G^{(i,j)}(\overline{x}, \overline{y}) \wedge \phi(\overline{x})) \qquad \forall\overline{x}(G^{(i,j)}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{x})).$$

$$
\begin{array}{ll}
Q_\varphi(\overline{a}) & \text{if } \varphi \text{ is the input formula} \\
\neg Q_\varphi(\overline{x}) \vee \neg P(\overline{x}) & \text{if } \varphi = \neg P(\overline{x}) \\
\neg Q_\varphi(\overline{x}) \vee \neg \mathcal{G}(\overline{x},\overline{y}) \vee Q_\phi(\overline{y}) & \text{if } \varphi = \forall \overline{y}(\mathcal{G}(\overline{x},\overline{y}) \rightarrow \phi(\overline{y})) \\
\neg Q_\varphi(\overline{x}) \vee \mathcal{G}(\overline{x},\overline{f(\overline{x})}) & \multirow{2}{*}{\text{if } \varphi = \exists \overline{y}(\mathcal{G}(\overline{x},\overline{y}) \wedge \phi(\overline{y}))} \\
\neg Q_\varphi(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})}) & \\
\neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) & \multirow{2}{*}{\text{if } \varphi = \phi(\overline{x}) \wedge \psi(\overline{y}) \quad \overline{z} = \overline{x} \cup \overline{y}} \\
\neg Q_\varphi(\overline{z}) \vee Q_\psi(\overline{y}) & \\
\neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) \vee Q_\psi(\overline{y}) & \text{if } \varphi = \phi(\overline{x}) \vee \psi(\overline{y}) \quad \overline{z} = \overline{x} \cup \overline{y}
\end{array}
$$

Fig. 1. Schematic clausal form for GF1$^-$ formulae.

Note how the role of $\overline{x}$ and $\overline{y}$ may be interchanged in the guard. Occurrences of atoms in guard positions such that (iv) is satisfied are said to *satisfy the grouping restriction.*

Examples of GF1$^-$ formulae are the following:

$$
q(x) \wedge \exists x, y(r_1(z,z,x,y) \wedge p(x,y)), \quad p(x,y) \wedge p(y,x),
$$
$$
\forall xy(r_2(x,y,z) \rightarrow (p(x,y) \wedge \exists z(r_2(x,y,z) \wedge q(z)))).
$$

The grouping of the predicate symbols r_1 and r_2 is $(2,2)$ and $(2,1)$, while for the remaining predicate symbols the grouping is immaterial. The second formula is implicitly existentially quantified. If it were to be explicitly quantified it would no longer belong to GF1$^-$ although it would be a guarded formula. Other examples of guarded formulae which do not belong to GF1$^-$ are $\forall xy(q(x,x,y) \rightarrow \bot))$ and $\forall xy(p(x,y) \rightarrow \exists z p(y,z))$.

We will assume that all formulae are in negation normal form. Short reflexion will convince the reader that the transformation to negation normal form does not take us outside the boundaries of GF1$^-$. The transformation of GF1$^-$ formulae into clausal form makes use of structural transformation, also known as definitional form transformation or renaming (cf. e.g. [3, 20]). In particular, we require the introduction of new symbols for non-atomic subformula of the original formula with the exception of implications and conjunctions immediately below quantifiers. The transformation maps GF1$^-$ formulae to guarded formulae in a certain form, which, when clausified, render clauses satisfying the schematic presentation of Fig. 1. The non-positive clauses[1] will be referred to as *definitional clauses.* The symbol Q_φ is a new symbol introduced for the subformula φ indicated in the index. Thus Q_φ can be thought of as the name for the subformula φ. By definition, $\mathcal{G}(\overline{x},\overline{y})$ represents either $G(\overline{x},\overline{y})$ or $G(\overline{y},\overline{x})$. $\overline{a}$ stands for a sequence of constants, and $\overline{f(\overline{x})}$ is a sequence of terms $f_k(x_1,\dots,x_n)$, where the arguments of each of the f_k are are exactly the elements of $\overline{x}$.

Theorem 1. *Suppose φ is any GF1$^-$ formula. Let N be the set of clauses obtained from φ by negation normal form transformation, the above renaming and clausification. Then, (i) any clause in N has one of the forms in Fig. 1, (ii) the conversion of φ to N can be performed in polynomial time, and (iii) φ is satisfiable iff N is satisfiable.*

[1] A *non-positive clause* contains a negative literal, whereas a *positive clause* contains only positive literals.

4 Hyperresolution for GF1$^-$

To decide GF1$^-$ we use a calculus, denoted $\mathsf{R}^{\mathsf{hyp}}$, of positive hyperresolution combined with splitting.

Positive hyperresolution resolves positive clauses with a non-positive clause always producing a positive conclusion or the empty clause. More precisely, a hyperresolvent is derived according to the following rule:

$$\frac{C_1 \vee A_1 \quad \ldots \quad C_n \vee A_n \quad \neg A_{n+1} \vee \ldots \vee \neg A_{2n} \vee D}{(C_1 \vee \ldots \vee C_n \vee D)\sigma}$$

where (i) σ is the most general unifier such that $A_i\sigma = A_{n+i}\sigma$ for any i, $1 \leq i \leq n$, and (ii) $C_i \vee A_i$ and D are positive clauses, for any i, $1 \leq i \leq n$. The premise $\neg A_{n+1} \vee \ldots \neg A_{2n} \vee D$ is referred to as the *negative* premise and all the other premises in the resolution rule are referred to as *positive* premises. Consequently, the positive premise of an inference step has to be a positive clause.

Factors are generated by the rule:

$$\frac{C \vee A_1 \vee A_2}{(C \vee A_1)\sigma}$$

where σ is the most general unifier of A_1 and A_2. Factoring is not required for the completeness of our decision procedure, but it helps avoiding applications of splitting to clauses containing duplicate literals.

The splitting rule originates from semantic tableaux. If the clause set N contains a clause C it can be split into clauses C_1 and C_2, provided that C_1 and C_2 are variable disjoint. The original clause becomes redundant and the resolution refutation is performed independently on $N \cup \{C_1\}$ and $N \cup \{C_2\}$.

The order in which these rules are applied in $\mathsf{R}^{\mathsf{hyp}}$ is (i) factoring, (ii) splitting and (iii) hyperresolution. As usual we make a minimal assumption that at any stage in the derivation the clause set contains no duplicate clauses. For soundness and refutational completeness of $\mathsf{R}^{\mathsf{hyp}}$ see [23] or [4, 5].

For the classes of clause sets we consider in the present paper the positive premises will always be ground, in particular, because we use splitting, the positive premises will always be ground *unit* clauses, and the conclusions are always positive ground clauses. Crucial for termination is that the unit clauses are always either *uni-node* or *bi-node*. These are notions inspired by Lutz et al. [22].

Let N denote the clauses obtained from a given GF1$^-$ formula. We will show that any derivation by $\mathsf{R}^{\mathsf{hyp}}$ from N is finite. For this we define some basic concepts.

A function symbol g is said to be *associated with* a predicate symbol Q iff N contains a definitional clause of the form $\neg Q(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$ in which g occurs.

A set $\{t_1 \ldots t_n\}$ (or sequence $\overline{t} = (t_1, \ldots, t_n)$) of ground terms is called a *uni-node* iff either each t_i, $1 \leq i \leq n$, is a constant, or there exists a predicate symbol Q and a sequence of ground terms $\overline{s}$, such that each t_i, $1 \leq i \leq n$, has the form $f_Q(\overline{s})$, where f_Q is a function symbol associated with Q. A uni-node X_2 is called a *direct successor* of a uni-node X_1 iff there is a predicate symbol Q such that for each element t of X_2 there is a function symbol f_Q, associated with Q, and $t = f_Q(\overline{s})$, where $\overline{s}$ is a sequence of exactly the elements of X_1. A set (or sequence) of ground terms is called a *bi-node* iff it can be presented as a union $X_1 \cup X_2$ of two non-empty disjoint uni-nodes X_1 and

X_2 such that X_2 is a direct successor of X_1. A ground literal is a *uni-node* (*bi-node*) iff the set of its arguments is a uni-node (bi-node). A clause is a *uni-node* (*bi-node*) iff the set of the arguments of all literals in it is a uni-node (bi-node). The empty clause $\perp$ is a special type of uni-node without direct successors.

The sets $\{a, a, b\}$, $\{h_Q(a, b)\}$, $\{h_Q(a, b), g_Q(a, b)\}$ are examples of uni-nodes, while examples of bi-nodes are $\{a, b, h_Q(a, b)\}$ and $\{a, b, g_Q(a, b), h_Q(b, a)\}$. Here, Q is assumed to be a symbol introduced for a existentially quantified subformula and g_Q and h_Q are function symbols associated with the same predicate symbol Q. Observe that both $\{g_Q(a, b)\}$ $\{h_Q(b, a)\}$ are direct successors of $\{a, b\}$.

Lemma 1. *Suppose a finite signature is given.*

1. *The cardinality of any uni-node is finitely bounded.*
2. *For any given uni-node $\overline{s}$, the number of uni-nodes, and bi-nodes, of the form $\mathcal{G}(\overline{s}, \overline{t})$ is finitely bounded.*
3. *Every uni-node clause has a bounded number of direct successors, which are uni-nodes.*

Now let us take a closer look at the properties of inferences in $\mathsf{R}^{\mathsf{hyp}}$. We note that except for one positive ground unit clause N contains only definitional clauses which are non-positive and non-ground. The negative premise of a resolution inference step is always a definitional clause in N, and maximally split conclusions of most resolution inference step are uni-nodes. The exceptions are inferences with $\neg Q_\exists(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$ which produce bi-node conclusions. Formally:

Lemma 2. *In any $\mathsf{R}^{\mathsf{hyp}}$ derivation from N all derived clauses are either empty or positive ground clauses which can be split into positive ground unit clauses of the form: $Q_\varphi(\overline{s})$ or $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$, where $\overline{s}$ is a uni-node. That is, maximally split conclusions are either uni-nodes or bi-nodes.*

Proof. The proof is by induction over an arbitrary derivation. In the first step of the derivation there is only one possible positive premise, namely the ground unit clause $Q_\varphi(\overline{a})$, which is a uni-node. In the induction step we assume that the positive premise $Q_\varphi(\overline{s})$ of the derivation step under consideration is a uni-node, while premises $G(\overline{s}, \overline{t})$ are either uni-nodes or bi-nodes.

Now consider some of the most interesting inference possibilities. Consider the resolution steps, where the positive premise is a definitional clause, introduced for an existentially quantified $\mathrm{GF1}^-$ formula. Assume that the positive premise $Q_\varphi(\overline{s})$ is a uni-node. There are two possibilities. (i) The negative premise is a clause $\neg Q_\varphi(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})})$. Because of the grouping restrictions, the argument set of the conclusion $Q_\phi(\overline{f(\overline{s})})$ is a sequence $\overline{f(\overline{s})}$ of terms $f_k(s_1, \ldots s_n)$, where $s_1, \ldots s_n$ are exactly the elements of $\overline{s}$ and f_k is associated with Q_φ. Hence the conclusion of this resolution step is a uni-node. (ii) The negative premise is a clause $\neg Q_\varphi(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$. Due to the grouping restrictions, the argument set of the conclusion $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$ is such that the arguments of the elements $\overline{f(\overline{s})}$ are exactly the elements of $\overline{s}$ and each of the functional symbols in $\overline{f(\overline{s})}$ is associated with Q_φ. By definition, this means that $\overline{f(\overline{s})}$ is a direct successor of $\overline{s}$. Therefore the conclusion $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$ is a bi-node.

Consider also hyperresolution involving a definitional clause $\neg Q_\varphi(\overline{x}) \vee \neg\mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$, corresponding to universally quantified $\mathrm{GF1}^-$ formula and positive premises

$Q_\varphi(\overline{s})$ and $\mathcal{G}(\overline{s}, \overline{t})$. By assumption $Q_\varphi(\overline{s})$ is a uni-node. $\mathcal{G}(\overline{s}, \overline{t})$ is either a uni-node or a bi-node. In the first case, since the argument set of the conclusion $Q_\phi(\overline{t})$ is a subset of the argument set of the positive premise $\mathcal{G}(\overline{s}, \overline{t})$ the conclusion is also a uni-node. In the second case, the argument set of $\mathcal{G}(\overline{s}, \overline{t})$ consists of two distinct uni-nodes $\overline{s}$ and $\overline{t}$, such that one of them is a direct successor of the other. Hence, the argument set of the conclusion $Q_\phi(\overline{t})$ consists of ground terms $\overline{t}$, such that $\overline{t}$ is a uni-node. □

Lemma 3. *In any* $\mathsf{R}^{\mathsf{hyp}}$ *derivation from* N*:*

1. *Every ground clause which is a bi-node is an instantiated guard atom of the form* $\mathcal{G}(\overline{s}, \overline{t})$*, where* $\overline{s}$ *and* $\overline{t}$ *are uni-nodes.*
2. *If* C *and* D *are uni-nodes, such that* D *is a direct successor of* C*, then* D *is derived from* C *and a bi-node.*

Proof. The first statement follows immediately from Lemma 2. Inspecting all resolution steps we observe that every ground clause which is a bi-node is an instantiated guard atom of the form $\mathcal{G}(\overline{s}, \overline{t})$, where $\overline{s}$ and $\overline{t}$ are uni-nodes and $\overline{s}$ is a direct successor of $\overline{t}$ or vice versa. The second statement is a direct consequence of Lemma 2 and statement 1. □

In order to prove termination we define a dependency relation $\succ_d$ and a complexity measure μ on predicate symbols. The *dependency relation* $\succ_d$ is defined on predicate symbols S_1, S_2 and $S_1 \succ_d S_2$, if there is a definitional clause $\neg Q_\varphi(\overline{x}) \vee C$, such that $S_1 = Q_\varphi$ and S_2 occurs in C. Let tt be a new symbol which is smallest with respect to $\succ_d$. Define $>_D$ to be an ordering on predicate symbols, which is compatible with the transitive closure of $\succ_d$. The relation $\succ_d$ forms a directed acyclic graph, in which the inner nodes are labelled with definitional symbols and the leaves are labelled with predicate symbols of the original formula. The graph forms a tree which is isomorphic to the subformula tree of the original formula. Thus it is not difficult to see that, because the number of different predicate symbols in N is finite, the ordering $>_D$ is well-founded.

For the input clause set N corresponding to the class of formulae in GF1$^-$ (cf. Fig. 1), the following relationships hold: $Q_\varphi >_D P$ if $\varphi = \neg P(\overline{x})$, $Q_\varphi >_D G$ and $Q_\varphi >_D Q_\phi$ if $\varphi = \exists \overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \wedge \phi(\overline{y}))$, and so on.

The complexity measure which we define next will be ordered by $>_D$. The *complexity measure* μ of a uni-node clause is defined by:

$$\mu(C) = \begin{cases} Q & \text{if } C = (\neg)Q(\overline{s}) \\ tt & \text{if } C \text{ is the empty clause.} \end{cases}$$

Lemma 4. *The measure of any derived uni-node is always smaller (with respect to* $<_D$*) than the measure of the positive uni-node premise with an introduced predicate symbol.*

The main technical lemma is the following.

Lemma 5. *Let* φ *be a formula in GF1*$^-$ *and let* N *be the corresponding clause set. The number of uni-nodes derivable from* N *is finitely bounded.*

Proof. The proof uses an inductive argument over the ordering $>_D$ on uni-nodes in a derivation tree. We show by induction on the measure of predicate symbols that in any derivation from N there are only finitely many clauses C associated with complexity measure Q_ϕ for some subformula ϕ of φ.

Base case. There is only one uni-node in N with measure Q_φ. The predicate symbol Q_φ has no positive occurrences in any of the definitional clauses in N. So, no new clauses with predicate symbol Q_φ are derivable from N. Therefore all derivable uni-nodes have measure strictly smaller than the measure of Q_φ. Hence, there are only finitely many uni-nodes with measure Q_φ.

Induction hypothesis. There are only finitely many derivable uni-nodes with measure greater than Q_ϕ, for some Q_ϕ. We prove that there are only finitely many uni-nodes with measure Q_ϕ.

Induction step. We consider all possible ways of deriving uni-nodes with measure Q_ϕ from the definitional clauses in the input clause set N.

Consider an inference step between a definitional clause $\neg Q_\rho(\overline{z}) \vee Q_\phi(\overline{x})$ introduced for a conjunction of GF1$^-$ formulae and a positive premise $Q_\rho(\overline{s})$. Since $Q_\rho >_D Q_\phi$, by induction hypothesis there are only finitely many uni-nodes with measure Q_ρ. Then there are only finitely many uni-nodes of the form $Q_\phi(\overline{t})$ since the set of arguments of $Q_\phi(\overline{t})$ is a subset of the argument set of $Q_\rho(\overline{s})$ because $\overline{x} \subseteq \overline{z}$. The case where an inference step with the definition of a disjunction is performed is analogous.

Next we consider the resolution steps involving definitional clauses resulting from existentially quantified formulae, $\neg Q_\rho(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})})$ and $\neg Q_\rho(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$, and a positive premise clause $Q_\rho(\overline{s})$. By the induction hypothesis there are finitely many uni-nodes of the form $Q_\rho(\overline{s})$. The conclusion of the first resolution step is a uni-node, which is a direct successor of the positive premise. By Lemma 1.3 for each uni-node $Q_\rho(\overline{s})$ there are only finitely many uni-nodes of the form $Q_\phi(\overline{f(\overline{s})})$. The conclusion $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$ of the second resolution step is a bi-node. By Lemma 1.2 there are only finitely many bi-nodes of the form $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$ for a given $\overline{s}$.

Further, we consider a resolution step, involving a definitional clause $\neg Q_\varphi(\overline{x}) \vee \neg\mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$ introduced for universally quantified formulae. By Lemma 1.2 for each $Q_{\rho)}(\overline{s})$ there are only finitely many bi-nodes of the form $\mathcal{G}(\overline{s}, \overline{t})$ and also only finitely many uni-nodes of the form $\mathcal{G}(\overline{s}, \overline{t})$. This proves there are only finitely many uni-nodes of the form $Q_\phi(\overline{t})$.

Finally, any inference step involving a definitional clause for a negated atomic subformula results in the empty clause. The complexity measure of the empty clause is *tt* and by definition there is only one clause with this measure. □

Now, we can state the first main theorem of this section.

Theorem 2. *Let φ be a GF1$^-$ formula and let N be the corresponding clause set. Then,*

1. *any* $\mathsf{R}^{\mathsf{hyp}}$ *derivation from N terminates, and*
2. *φ is unsatisfiable iff the* $\mathsf{R}^{\mathsf{hyp}}$ *saturation of N contains the empty clause.*

This theorem subsumes corresponding results for an encoding of the extended modal logic $K_{(m)}(\cap, \cup, \smile)$ in [9, 20].

In addition we can prove the following results.

Lemma 6. *There is no nesting of the same functional symbol in any clause in the derivation.*

Lemma 7. *The term depth of any clause in the derivation is bounded by the quantifier depth of the original formula.*

It is well-known that hyperresolution, like tableaux methods, can be used to construct models for satisfiable formulae [11]. Clearly, in the present application if $\mathsf{R}^{\mathsf{hyp}}$ terminates without having produced the empty clause then it takes no extra effort to construct a model. A model will be given by the set of ground unit clauses in an open branch of the derivation tree. We have:

Theorem 3. *Let φ be a satisfiable formula in GF1$^-$.*

1. *A finite model for φ can be constructed on the basis of $\mathsf{R}^{\mathsf{hyp}}$.*
2. *The maximum size of the model is exponential in the length of φ.*

Extending the simulation results of [9, 19] we get:

Theorem 4. *There is a polynomial bisimulation between $\mathsf{R}^{\mathsf{hyp}}$ and the tableaux system of Lutz et al. [22] for GF1$^-$.*

The described decision procedure looks very similar to the decision procedures based on negative selection for expressive modal logics and description logics, which are described in [9, 18–20]. The main difference is the way in which we prove termination. In the proofs of [9], for instance, an ordering is defined under which all conclusions of inference steps are smaller than every premise, while here this is only true for uni-node premises (with introduced predicate symbols). In the case of guarded formulae a complexity measure on all clauses would not work because predicate symbols can occur in guard and non-guard positions and consequently such an ordering would be cyclic. In addition, we cannot rely solely on the well-foundedness property of the ordering on the complexity measure, but also have to exploit the type of the conclusions, obtained in the derivation and by semantic analysis show that there are only finitely many. With the notions of uni-node and bi-node a finer characterisation can be given of what happens during inference. The proof here extends to generalisations of GF1$^-$ discussed in the next section.

5 Generalisation

In this section we mention different ways of extending GF1$^-$ and its corresponding clausal class without loosing decidability by hyperresolution.

According to the definition of GF1$-$, the quantified variables in the GF1$^-$ formulae must be exactly the free variables of non-guard formulae. Hyperresolution is a decision procedure for a more general fragment, defined so that the quantified sequences of variables in the non-guard formulae are a subset of the quantified variables.

$$\exists\overline{y}(\mathcal{G}(\overline{x},\overline{y}) \wedge \phi(\overline{z})) \qquad\qquad \forall\overline{y}(\mathcal{G}(\overline{x},\overline{y}) \rightarrow \phi(\overline{z}))$$

where $\overline{z} \subseteq \overline{y}$. The resulting clausal forms are

$$\neg Q_\forall(\overline{x}) \vee \neg\mathcal{G}(\overline{x},\overline{y}) \vee Q_\phi(\overline{z}) \qquad\qquad \text{where } \overline{z} \subseteq \overline{y}$$

$$\neg Q_\exists(\overline{x}) \vee \mathcal{G}(\overline{x},\overline{f(\overline{x})}) \qquad \neg Q_\exists(\overline{x}) \vee Q_\phi(\overline{t}) \qquad\qquad \text{where } \overline{t} \subseteq \overline{f(\overline{x})}$$

If $\overline{z}$ is the empty sequence then Q_ϕ is a propositional symbol. In general, this means that ϕ is a closed subformula, but due to restriction (iv) in the definition of GF1$^-$, namely, that the variable sequences $\overline{x}$ and $\overline{y}$ may not be empty, it follows that ϕ is a propositional formula.

The restriction in GF1$^-$ that a guard is a single atom can be relaxed. Certain complex guards which may include negation can be allowed. If we consider what happens in a hyperresolution inference step then it is not difficult to see that inferences with definitional clauses like the following produce uni-node and bi-node conclusions (after splitting).

$$\neg Q_\forall(\overline{x}) \vee \neg\mathcal{G}_0(\overline{x},\overline{y}) \vee (\neg)\mathcal{G}_1(\overline{x^{(1)}},\overline{y^{(1)}}) \vee \ldots \vee (\neg)\mathcal{G}_n(\overline{x^{(n)}},\overline{y^{(n)}}) \vee Q_\phi(\overline{z})$$

where $\overline{x^{(i)}} \subseteq \overline{x}$, $\overline{y^{(i)}} \subseteq \overline{y}$ $(1 \leq i \leq n)$ and $\overline{z} \subseteq \overline{y}$. An essential condition is that each of the atoms $G_0(\ldots)$, $G_i(\ldots)$ satisfies the grouping restriction (as suggested by the notation) and the clause includes at least one guard, $\neg G_0(\overline{x},\overline{y})$. This ensures that the conclusion is a ground clause. On the first-order level, this means we can allow formulae of the form:

$$\forall\overline{y}((\mathcal{G}_0(\overline{x},\overline{y}) \wedge (\neg)\mathcal{G}_1(\overline{x^{(1)}},\overline{y^{(1)}}) \wedge \ldots \wedge (\neg)\mathcal{G}_n(\overline{x^{(n)}},\overline{y^{(n)}})) \rightarrow \phi(\overline{z})).$$

Disjunctions in the guard expression are permitted provided none of the atoms is negated and $\overline{z} \subseteq \overline{y} \cap \overline{y^{(1)}} \cap \ldots \cap \overline{y^{(n)}}$.

$$\forall\overline{y}((\mathcal{G}_1(\overline{x^{(1)}},\overline{y^{(1)}}) \vee \ldots \vee \mathcal{G}_n(\overline{x^{(n)}},\overline{y^{(n)}})) \rightarrow \phi(\overline{z}))$$

The corresponding clause set includes clauses of this form, where $\overline{x} = \overline{x^{(1)}} \cup \ldots \cup \overline{x^{(n)}}$.

$$\neg Q_\forall(\overline{x}) \vee \neg\mathcal{G}_i(\overline{x^{(i)}},\overline{y^{(i)}}) \vee Q_\phi(\overline{z})$$

Such formulae fall outside the guarded and loosely guarded fragment.

As the introduced negative literal in a clause associated with an existentially quantified formula contains all the variables of the clause we can be much more permitting in this case:

$$\exists\overline{y}(F \wedge \phi(\overline{z})),$$

where F is any Boolean combination of atoms $\mathcal{G}_1(\overline{x^{(1)}},\overline{y^{(1)}}), \ldots, \mathcal{G}_n(\overline{x^{(n)}},\overline{y^{(n)}})$. Again, the $\mathcal{G}_i(\ldots)$ are required to satisfy the grouping restriction. Clausification produces clauses of the form:

$$\neg Q_\exists(\overline{x}) \vee (\neg)\mathcal{G}_{i_1}(\overline{x^{(i_1)}},\overline{f^{(i_1)}(\overline{x})}) \vee \ldots \vee (\neg)\mathcal{G}_{i_m}(\overline{x^{(i_m)}},\overline{f^{(i_m)}(\overline{x})})$$
$$\neg Q_\exists(\overline{x}) \vee Q_\phi(\overline{g(\overline{x})}).$$

Other generalisations are conceivable, but this is the subject of ongoing work. At this stage we have the following results.

Theorem 5. *Let φ be a formula in the above extension of GF1$^-$ and let N be the corresponding clause set. Then,*

1. *any* $\mathsf{R}^{\mathsf{hyp}}$ *derivation from N terminates, and*

2. φ is unsatisfiable iff the $\mathsf{R}^{\mathsf{hyp}}$ saturation of N contains the empty clause.

Theorem 6. *Let φ be a satisfiable formula in the above extension.*

1. *A finite model for φ can be constructed on the basis of $\mathsf{R}^{\mathsf{hyp}}$.*
2. *The maximum size of the model is exponential in the length of φ.*

Similar as in de Nivelle et al. [9], macro inferences in $\mathsf{R}^{\mathsf{hyp}}$ (for N) can be identified and reformulated as tableaux inference rules, providing a sound and complete tableaux decision procedure for the extension.

Finally we note:

Theorem 7. *Hyperresolution without splitting is a sound, complete and terminating inference procedure for the clausal classes associated with GF1$^-$ and the considered extension.*

Proof. Soundness and completeness is proved in [23]. Termination follows from Theorems 2 and 5, since all derived clauses by hyperresolution without splitting are formed from uni-node and bi-node literals appearing in the corresponding derivation tree for $\mathsf{R}^{\mathsf{hyp}}$. □

6 Related Work

Related Calculi. Decidability of formulae in GF1$^-$ is shown by semantic tableaux in Lutz et al. [22]. As stated in Section 4 there is a polynomial bisimulation between $\mathsf{R}^{\mathsf{hyp}}$ and this tableaux system. The precise relationship is in fact a stepwise simulation, namely, each application of a tableaux rule can be simulated by a macro inference step of $\mathsf{R}^{\mathsf{hyp}}$.

Hyperresolution can be viewed as resolution with maximal selection of negative literals, see [5] and [9, 18–20]. This connection provides an avenue for refining $\mathsf{R}^{\mathsf{hyp}}$ with mechanisms for pruning the search space.

Another useful connection is the connection to *hypertableaux*, introduced by Baumgartner, Furbach and Niemelä [8]. Given a finite set N of input clauses and a selection function S, the hypertableaux procedure generates a literal tree and at each stage of the derivation every open branch is a partial representation of a potential model for N. Initially the hypertableaux consists of a single node marked open. In subsequent steps a hypertableaux is obtained from a literal tree T by attaching child nodes to the open branch selected by S in T. The child nodes are

$$A_1\sigma\pi,\ \ldots,\ A_m\sigma\pi,\ \neg B_1\sigma\pi,\ \ldots,\ \neg B_n\sigma\pi,$$

if (i) $C = \neg B_1 \vee \ldots \vee \neg B_n \vee A_1 \vee \ldots A_m$ is a clause from N, $0 \leq m, n$, (ii) σ is a most general substitution such that the minimal Herbrand model of the set of (universal closures of the) literals in the selected branch satisfies (the universal closure of) $B_1\sigma \wedge \ldots \wedge B_n\sigma$, and (iii) π is a substitution for $C\sigma$ such that the positive literals in $C\sigma\pi$ do not share variables. C is called the extending clause, and π is called a purifying substitution. The new branches with negative leafs are immediately marked 'closed'.

The close link between hypertableaux and hyperresolution with splitting is evident. A drawback of hypertableaux is the guessing of the purifying substitution.

For the clausal classes considered in the previous sections all hyperresolvents are ground, which implies that the purifying substitution is never needed. That is, for our application hypertableaux and hyperresolution with splitting are essentially the same. Consequently, the results for $\mathsf{R}^{\mathsf{hyp}}$ are also true for hypertableaux, and another alternative to $\mathsf{R}^{\mathsf{hyp}}$ is hypertableaux (and for that matter also the descendants of hypertableaux [6, 7]). For practical considerations this link allows us to transfer several improvements of hypertableaux discussed in [8] to $\mathsf{R}^{\mathsf{hyp}}$. These include factorisation and level cut. Factorisation has the effect that different branch represent disjoint partial models. This can be achieved in our case by modification of the splitting rule to: If the clause set N contains a ground clause $C_1 \vee C_2$ then the resolution refutation is performed independently on $N \cup \{C_1\}$ and $N \cup \{\neg C_1, C_2\}$. The level cut improvement corresponds to branch condensing (used in SPASS [24]) or backjumping (used in tableaux methods [17]). On the side we remark that hyperresolution with splitting avoids the 'memory management' problem of hyperresolution highlighted in [8].

Related Clausal Classes. We have already referred to the related clausal classes associated with modal and description logics. Other clausal classes decidable by hyperresolution are investigated in [11, 21] and include the classes $\mathcal{PVD}$ and $\mathcal{KPOD}$.

For $\mathcal{PVD}$ the syntactic restrictions on the class imply that during a derivation by hyperresolution the depth of the conclusions does not increase [11, 21], unlike in GF1−. For $\mathcal{KPOD}$ the term depth of conclusions increases during the derivation [21]. However, essential for $\mathcal{KPOD}$ is the restriction of clauses to Krom form ($|C| \leq 2$), which does not apply to clauses originating from the definitional form of $\mathrm{GF1}^-$ formulae.

Termination for these classes is shown in terms of an atom complexity measure μ, defined as a function from atoms to natural numbers with the following properties: (i) $\mu(A) \leq \mu(A\sigma)$ for all atoms A and all substitutions σ, (ii) for all natural numbers k and any finite signature Σ it is true that for all atoms A, the set $\{A\sigma | \sigma \in \sigma_0, \mu(A\sigma) \leq k\}$ is finite, where σ_0 is the set of all ground substitutions over Σ, (iii) μ is extended to literals by $\mu(A) = \mu(\neg A)$, and to clauses by $\mu(\{L_1, \ldots, L_n\}) = \max\{\mu(L_i) \mid 1 \leq i \leq n\}$. Our complexity measure does not have the second property. It is open whether decidability of the classes considered in this paper can be formalised in this framework.

Another related class is the encoding in clausal form of the extended multi-modal logic $K_{(m)}(\cap, \cup, \smile)$ and the corresponding description logic $\mathcal{ALB}_D$ [19, 20]. This class is subsumed by the clausal class of Section 3.

7 Conclusion and further work

We have considered the use of hyperresolution as a decision procedure for guarded formulae. Our contribution is in the different approach to showing decidability and in considering several extensions of the fragment. We have discussed related calculi and compared them to hyperresolution. The decision procedure is practical. Standard resolution provers can be used without adaptation (for example, FDPLL, OTTER, PROTEIN, SPASS, and VAMPIRE). Currently we are looking into defining an abstract atom complexity measure μ in analogy to Leitsch [21] which would generalise the specific complexity measures and orderings used in the termination proofs presented in this paper and in [9, 18–20]. We are also attempting to define a larger solvable class which would accommodate more formulae outside the guarded fragment. Further it would be of interest to extend the approach to the entire guarded fragment, by using proper blocking conditions in the context of resolution.

References

1. H. Andréka, I. Németi, and J. van Benthem. Modal languages and bounded fragments of predicate logic. *J. Philos. Logic*, 27(3):217–274, 1998.
2. H. Andréka, J. van Benthem, and I. Németi. Back and forth between modal logic and classical logic. *Bull. IGPL*, 3(5):685–720, 1995.
3. M. Baaz, C. Fermüller, and A. Leitsch. A non-elementary speed-up in proof length by structural clause form transformation. In *Proc. LICS'94*, pp. 213–219. IEEE Computer Society Press, 1994.
4. L. Bachmair and H. Ganzinger. Rewrite-based equational theorem proving with selection and simplification. *J. Logic Computat.*, 4(3):217–247, 1994.
5. L. Bachmair and H. Ganzinger. Resolution theorem proving. In J. A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*. Elsevier, 2000. To appear.
6. P. Baumgartner. Hyper tableau: The next generation. In *Proc. TABLEAUX'98*, vol. 1397 of *LNAI*, pp. 60–76. Springer, 1998.
7. P. Baumgartner. FDPLL: A first-order Davis-Putnam-Logeman-Loveland procedure. In *Automated Deduction—CADE-17*, LNAI. Springer, 2000. To appear.
8. P. Baumgartner, U. Furbach, and I. Niemelä. Hyper tableaux. In *European Workshop on Logic in AI (JELIA'96)*, vol. 1126 of *LNAI*, pp. 1–17. Springer, 1996.
9. H. de Nivelle, R. A. Schmidt, and U. Hustadt. Resolution-based methods for modal logics. *Logic J. IGPL*, 8(3):265–292, 2000.
10. F. M. Donini, M. Lenzerini, D. Nardi, and A. Schaerf. Reasoning in description logics. In G. Brewka, editor, *Principles in Knowledge Representation*, Studies in Logic, Language and Information, pp. 191–236. CSLI Publications, Stanford, 1996.
11. C. G. Fermüller, A. Leitsch, U. Hustadt, and T. Tammet. Resolution decision procedures. In *Handbook of Automated Reasoning*. Elsevier, 2000. To appear.
12. H. Ganzinger and H. de Nivelle. A superposition decision procedure for the guarded fragment with equality. In *Proc. LICS'99*, pp. 295–303. IEEE Computer Society Press, 1999.
13. H. Ganzinger, C. Meyer, and M. Veanes. The two-variable guarded fragment with transitive relations. In *Proc. LICS'99*, pp. 24–34. IEEE Computer Society, 1999.
14. R. Goré. Tableau methods for modal and temporal logics. In M. D'Agostino, D. Gabbay, R. Hähnle, and J. Posegga, editors, *Handbook of Tableau Methods*. Kluwer, 1999.
15. E. Grädel. On the restraining power of guards. Manuscript. Submitted to the *J. Symbolic Logic*, 1998.
16. E. Grädel. Decision procedures for guarded logics. In *Automated Deduction—CADE-16*, vol. 1632 of *LNAI*, pp. 31–51. Springer, 1999.
17. U. Hustadt and R. A. Schmidt. Simplification and backjumping in modal tableau. In *Proc. TABLEAUX'98*, vol. 1397 of *LNAI*, pp. 187–201. Springer, 1998.
18. U. Hustadt and R. A. Schmidt. On the relation of resolution and tableaux proof systems for description logics. In *Proc. IJCAI'99*, pp. 110–115. Morgan Kaufmann, 1999.
19. U. Hustadt and R. A. Schmidt. Issues of decidability for description logics in the framework of resolution. In *Automated Deduction in Classical and Non-Classical Logics*, vol. 1761 of *LNAI*, pp. 191–205. Springer, 2000.
20. U. Hustadt and R. A. Schmidt. Using resolution for testing modal satisfiability and building models. To appear in the *SAT 2000* Special Issue of *J. Automated Reasoning*, 2000.
21. A. Leitsch. Deciding clause classes by semantic clash resolution. *Fundamenta Informatica*, 18:163–182, 1993.
22. C. Lutz, U. Sattler, and S. Tobies. A suggestion of an n-ary description logic. In *Proc. DL'99*, pp. 81–85. Linköping University, 1999.
23. J. A. Robinson. Automatic deduction with hyper-resolution. *Internat. J. Computer Math.*, 1(3):227–234, 1965.
24. C. Weidenbach. SPASS: Combining superposition, sorts and splitting. In J. A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*. Elsevier, 2000. To appear.