Journal of Symbolic Computation 36 (2003) 163–192

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

Hyperresolution for guarded formulae

Lilia Georgieva[a], Ullrich Hustadt[b], Renate A. Schmidt[a,*]

[a] *Department of Computer Science, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[b] *Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, UK*

Received 1 October 2000; accepted 3 November 2001

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 of the guarded fragment which can be decided by hyperresolution. In particular, we prove decidability of hyperresolution with or without splitting for the fragment $GF1^-$ and point out several ways of extending this fragment without losing 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 Andréka et al. (1998, 1995) 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, the finite model property, and preservation under submodels. They show that the guarded fragment (GF) shares, indeed, all these properties with the basic modal logic *K*. Various extensions of the GF have been proposed and analysed with respect to these properties. The most well-known extension is the loosely GF, introduced in Andréka et al. (1998), and shown decidable in Ganzinger and de Nivelle (1999) and Grädel (1999b). Decidability has also been shown for the guarded fixpoint logic (Grädel, 1999a) and a monadic fragment of GF^2 with transitive guards (Ganzinger et al., 1999). The decision procedures for the GF and the various extensions exploit

* Corresponding author. Tel.: +44-(0)-161-275-6163; fax: +44-(0)-161-6236.
E-mail addresses: georgiel@cs.man.ac.uk (L. Georgieva), U.Hustadt@csc.liv.ac.uk (U. Hustadt), schmidt@cs.man.ac.uk (R.A. Schmidt).

0014-5793/03/$ - see front matter
doi:10.1016/S0747-7171(03)00034-8

different approaches: the finite model property, ordered resolution, alternating automata, or embeddings into monadic second-order logic. This is an interesting contrast to approaches in the literature on decidable modal logics and description logics, where tableaux-based decision procedures are predominantly used for testing satisfiability (see for example Donini et al., 1996; Goré, 1999).

In Lutz et al. (1999), Lutz, Sattler and Tobies investigate whether tableaux-based decision procedures exist for subclasses of the GF. They introduce a subclass of the GF, in particular, of the fragment GF1 which was introduced in Andréka et al. (1995). This subclass is called $GF1^-$, and is obtained 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. This grouping condition is important for termination in the tableaux procedure of Lutz et al.

In this paper we continue this line of investigation. However we exploit the close correspondence between tableaux-based decision procedures and hyperresolution combined with splitting which has been demonstrated for extended modal logics (de Nivelle et al., 2000; Hustadt and Schmidt, 2000b) and for description logics (Hustadt and Schmidt, 1999, 2000a). By using a structure preserving transformation of guarded formulae into clausal form we can recast the method of Lutz et al. in the framework of hyperresolution. The motivation for this shift in perspective is our interest in the applicability of resolution and hyperresolution methods as decision procedures.

Generally, hyperresolution is not a decision procedure for the entire GF. A simple example is provided by the guarded formula $p(y) \wedge \forall x(p(x) \rightarrow \exists z(p(z) \wedge \top))$ with clausal form $\{p(a), \neg p(x) \vee p(f(x))\}$, for which hyperresolution does not terminate. In contrast, in de Nivelle et al. (2000) and Hustadt and Schmidt (2000b) it is proved that hyperresolution with splitting is a decision procedure for a first-order encoding of the extended modal logic $K_{(m)}(\cap, \cup, \smile)$. Semantically $K_{(m)}(\cap, \cup, \smile)$ is defined by the class of frames in which the accessibility relations are closed under intersection, union and converse. In this paper we focus on the question as to whether the results for $K_{(m)}(\cap, \cup, \smile)$ can be extended to a generalized first-order logic fragment, for example, to the class $GF1^-$, and possibly to extensions of this class. Because the method of proving termination used in de Nivelle et al. (2000) and Hustadt and Schmidt (2000b) does not generalize to $GF1^-$, we investigate a different argument which is adapted from Lutz et al. (1999). This argument takes into consideration the form of the derived clauses and crucially depends on the grouping restriction in the definition of $GF1^-$. In the setting of hyperresolution it is immediately clear that the termination result can be extended to a larger class of guarded formulae than the class $GF1^-$ identified in Lutz et al. Thus, we obtain more general results than those previously known.

A problem closely related to the satisfiability problem is the problem of generating (counter-)models. It is well-known that hyperresolution can be employed with dual purpose, namely, as a reasoning method and a Herbrand model builder (Fermüller et al., 2001). Therefore, another topic in this paper is the use of hyperresolution as a procedure for automatically constructing Herbrand models for $GF1^-$ and the considered extension. The paper also considers how the method relates to other inference methods such as hypertableaux (Baumgartner et al., 1996), and how the work fits in with previous work on using hyperresolution as a decision procedure for first-order classes (Fermüller et al., 2001; Leitsch, 1993).

The structure of the paper is as follows. Some preliminary definitions are given in Section 2. Section 3 defines the fragment $GF1^-$ and describes the clausal normal form into which $GF1^-$ formulae are conveniently translated. The hyperresolution calculus is described in Section 4 and decidability of $GF1^-$ is shown. The topic of Section 5 is model building. Section 6 presents results on the computational properties of the decision procedure and the size of generated models. Section 7 analyzes and characterizes the precise relationship between the hyperresolution calculus and the semantic tableaux method of Lutz et al. Generalizations of the results for $GF1^-$ to a larger class of formulae are sketched in Section 8. The final section summarizes the contributions of this paper and concludes with some thoughts on further work.

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, p, q, r, atoms by A, B, literals by L, clauses by C, formulae by $\varphi, \phi, \psi, \vartheta, \alpha, \beta$ and sets of clauses by N.

An over-line indicates a sequence, for example, $\overline{x}$ denotes a finite sequence of variables and $\overline{s}$ denotes a finite sequence of terms. If $\overline{s} = (s_1, \ldots, s_n)$ then $\overline{f(\overline{s})}$ denotes a sequence of terms of the form $f_k(s_1, \ldots, s_n)$. 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 $\overline{t}$. By definition, $\overline{s} = \overline{t}$ iff $\overline{s} \subseteq \overline{t}$ and $\overline{t} \subseteq \overline{s}$. The union of the terms in $\overline{s}$ and $\overline{t}$ is denoted by $\overline{s} \cup \overline{t}$.

For any sequence $\overline{s}$ of terms (or formula ϕ) by var($\overline{s}$) (or var(ϕ)) 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 variables occurring in ϕ are all and only those in $\overline{x}$, regardless of the order they appear in ϕ and duplication of variables is possible.

An *expression* is a term, an atom, a literal or a clause. An expression is called *functional* if it contains a constant or a function symbol, and *non-functional*, otherwise. The set of all free variables occurring in an expression E, or in a set of expressions N, is denoted as $\mathcal{V}(E)$ or $\mathcal{V}(N)$. An expression is called *ground* if it contains no variables. For sets of expressions $|N|$ denotes the *cardinality* of the set N.

Clauses are disjunctions of literals, i.e. $C = L_1 \vee L_2 \vee \cdots \vee L_n$, they can also be regarded as multisets. As usual the symbols $\vee$ and $\neg$ denote disjunction and negation, respectively. A *positive* (resp., *negative)* clause contains only positive (resp., negative) literals. A clause is called *non-positive* if it contains at least one negative literal. A clause which consists of only one literal is called a *unit clause*. The empty clause is denoted by $\bot$. A *split component* of a clause $C \vee D$ is a subclause C such that C and D do not have any variables in common, i.e. are *variable disjoint*. A clause which cannot be split further is called a *maximally split* clause. Two formulae or clauses are said to be *variants* of each other if they are equal modulo variable renaming. Variant clauses are assumed to be equal.

A clause C is called *range-restricted* iff every variable occurring in the positive literals of C occurs also in the negative literals of C.

The *(term) depth* $\mathrm{dp}(t)$ of a term t is inductively defined as follows: (i) if t is a variable or a constant then $\mathrm{dp}(t) = 1$, and (ii) if $t = f(t_1, \ldots, t_n)$, then $\mathrm{dp}(t) = 1 + \max(\{\mathrm{dp}(t_i) \mid 1 \leq i \leq n\})$. The *term depth* $\mathrm{dp}(L)$ of a literal is defined to be the maximal depth of its argument terms and the *term depth* $\mathrm{dp}(C)$ of a clause is defined as the maximal term depth of the literals occurring in C.

We assume that a fixed finite signature Σ, i.e. a countable set of predicate symbols, a countable set of function symbols and a countable set of variables is given. With each predicate and each function symbol we associate a natural number n called the arity of the symbol. Given a set of clauses N, the Herbrand universe $\mathcal{HU}(\Sigma_N)$ over the signature Σ_N of N is the set of all ground terms built from the function and constant symbols in Σ_N. If there are no constants in the signature, a special constant symbol is introduced so that the $\mathcal{HU}(\Sigma_N)$ is not empty.

A first-order *interpretation* for a signature Σ_N is a structure $\mathcal{M} = \langle M, \cdot^I \rangle$, where M is a non-empty set and $\cdot^I$ is an *interpretation function* defined over the predicate symbols, the function symbols and the constant symbols. As usual $\cdot^I$ assigns an n-ary relation over M to an n-ary predicate symbol, an n-ary function from $M^n \to M$ to n-ary function symbols, and an element of M to constant symbols.

An assignment g for $\mathcal{M}$ is a mapping from the set of variables into M. Given an assignment g if x is a variable and $m \in M$, then $g^x_m(x) = m$ and $g^x_m(y) = g(y)$ for any variable y different from x. Analogously, if $x_1, \ldots, x_n$ are variables and $m_1, \ldots, m_n$ are elements of M, then $g^{[x_1,\ldots,x_n]}_{[m_1,\ldots,m_n]}(x_i) = m_i$ for every i, $1 \leq i \leq n$ and $g^{[x_1,\ldots,x_n]}_{[m_1,\ldots,m_n]}(y) = g(y)$ for any $y \notin \{x_1, \ldots, x_n\}$. Given an interpretation $\mathcal{M}$ and an assignment g for $\mathcal{M}$, the interpretation function can be extended to all terms by $x_i^I = g(x_i)$ and $f(t_1, \ldots, t_n)^I = f^I(t_1^I, \ldots, t_n^I)$. The satisfiability relation $\models$ is defined as

$$
\begin{array}{lll}
\mathcal{M}, g \models \top & & \\
\mathcal{M}, g \not\models \bot & & \\
\mathcal{M}, g \models P(t_1, \ldots, t_n) & \text{iff} & (t_1^I, \ldots, t_n^I) \in P^I \\
\mathcal{M}, g \models \neg\varphi & \text{iff} & \mathcal{M}, g \not\models \varphi \\
\mathcal{M}, g \models \varphi_1 \wedge \varphi_2 & \text{iff} & \mathcal{M}, g \models \varphi_1 \text{ and } \mathcal{M}, g \models \varphi_2 \\
\mathcal{M}, g \models \exists x_i \varphi & \text{iff} & \mathcal{M}, g^{x_i}_m \models \varphi \text{ for some } m \in M.
\end{array}
$$

If there exists an interpretation $\mathcal{M}$ and an assignment g such that $\mathcal{M}, g \models \varphi$, for a formula φ over Σ_N, then φ is *satisfiable* and $\mathcal{M}$ *satisfies* φ.

A *Herbrand interpretation* H is a set of ground atoms. By definition a ground atom A is *true* in the interpretation H if $A \in H$ and it is *false* in the interpretation H if $A \notin H$, $\top$ is true in all interpretations and $\bot$ is false in all interpretations. The truth value of the binary logical connectives $\vee$ and $\wedge$ is defined as follows: a conjunction of two ground atoms A and B is true in the interpretation H iff both A and B are true in H and respectively, a disjunction of ground atoms is true in H iff at least one of A or B is true in the interpretation. The truth value of a formula depends on the truth value assigned to its atomic subformulae. A clause C is true in an interpretation H iff for all ground substitutions σ there is a literal L in $C\sigma$ which is true in H. If an expression is true in an interpretation H then H is referred to as a *Herbrand model* of the expression.

3. The fragment GF1$^-$

In the language of GF1$^-$ 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 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.

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

Note how the role of $\overline{x}$ and $\overline{y}$ may be interchanged in the guard.

Grouping is a global condition of predicate symbols, so that all occurrences of a predicate symbol in a guard position must satisfy the given grouping of the predicate symbol. 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.

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

The grouping of the predicate symbols r_1 and r_2 is $(2, 2)$ and $(2, 1)$, respectively, while for the remaining predicate symbols the grouping is immaterial. Free variables in GF1$^-$ formulae are implicitly existentially quantified. Note that making the existential quantification of x and y explicit in the second example would result in a formula which is not in GF1$^-$ (it would still be a guarded formula though). Other examples of guarded formulae which do not belong to GF1$^-$ are $\forall xy(q(x, x, y) \rightarrow \bot))$ and $\forall xy(p(x, y) \rightarrow \exists zp(y, z))$.

The (*guarded*) *quantifier depth* $\mathrm{gqd}(\varphi)$ of a formula φ in GF1$^-$ is defined as follows: (i) $\mathrm{gqd}(\top) = \mathrm{gqd}(\bot) = \mathrm{gqd}(P(\overline{x})) = 0$, (ii) if ϕ and ψ are GF1$^-$ formulae then $\mathrm{gqd}(\neg\phi) = \mathrm{gqd}(\phi)$ and $\mathrm{gqd}(\phi \wedge \psi) = \mathrm{gqd}(\phi \vee \psi) = \max(\mathrm{gqd}(\phi), \mathrm{gqd}(\psi))$, and (iv) if $\overline{x}$ is a non-empty sequence of variables and $\exists\overline{x}\phi$ and $\forall\,\overline{x}\phi$ are GF1$^-$ formulae, then $\mathrm{gqd}(\exists\overline{x}\phi) = \mathrm{gqd}(\forall\,\overline{x}\phi) = 1 + \mathrm{gqd}(\phi)$, independent of the number of variables in $\overline{x}$.

We assume that all formulae are in negation normal form, i.e. negation is pushed inwards to occur only in front of predicate symbols. Furthermore, we assume that occurrences of $\top$ and $\bot$ are respectively replaced by appropriate tautologous and contradictory formulae. This assumption is not crucial; it is made to simplify the clausal class associated with GF1$^-$ and consequently the proofs are slightly more elegant. Short reflection 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. Baaz et al., 1994; Hustadt and Schmidt, 2000b; Plaisted and Greenbaum, 1986). The fundamental idea is the replacement of particular subformulae by atoms with new predicate symbols. This renaming preserves satisfiability and unsatisfiability.

To present the particular form of renaming we use, we need to define the notion of position of a (sub)formula within a formula. A *position* is a word over the natural numbers. The set $\mathrm{Pos}(\varphi)$ of positions of a given formula φ is inductively defined as follows: (i) the empty word ϵ is in $\mathrm{Pos}(\varphi)$, (ii) for $1 \le i \le n$, $i\lambda \in \mathrm{Pos}(\varphi)$ if $\varphi = \varphi_1 \star \cdots \star \varphi_n$ and $\lambda \in \mathrm{Pos}(\varphi_i)$ where $\star$ is a first-order operator. If $\lambda \in \mathrm{Pos}(\varphi)$, then $\varphi|_\epsilon = \varphi$ and $\varphi|_{i\lambda} = \varphi_i|_\lambda$ where $\varphi = \varphi_1 \star \cdots \star \varphi_n$.

The renaming associates with each element λ of $\mathrm{Pos}(\varphi)$ a predicate symbol Q_λ and a literal $Q_\lambda(x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are the free variables of $\varphi|_\lambda$, and the symbol Q_λ does not occur in φ. Two symbols Q_λ and $Q_{\lambda'}$ are equal only if $\varphi|_\lambda$ and $\varphi|_{\lambda'}$ are variant formulae. The *definition of* Q_λ is the formula:

$$\mathrm{Def}_\lambda(\varphi) = \forall x_1 \ldots x_n (Q_\lambda(x_1, \ldots, x_n) \to \varphi|_\lambda).$$

The *definitional form* $\mathrm{Def}_\Lambda(\varphi)$ of φ is inductively defined by:

$$\mathrm{Def}_\emptyset(\varphi) = \varphi \qquad \text{and}$$
$$\mathrm{Def}_{\Lambda\cup\{\lambda\}}(\varphi) = \mathrm{Def}_\Lambda(\varphi[Q_\lambda(x_1, \ldots, x_n) \mapsto \lambda]) \wedge \mathrm{Def}_\lambda(\varphi),$$

where λ is maximal in $\Lambda \cup \{\lambda\}$. (Here, $\varphi[A \mapsto \lambda]$ denotes the formula obtained by replacing the subformula at position λ in φ with A.)

Theorem 3.1 (e.g. Plaisted and Greenbaum, 1986). *Let φ be a first-order formula. For any $\Lambda \subseteq \mathrm{Pos}(\varphi)$, $\mathrm{Def}_\Lambda(\varphi)$ can be computed in polynomial time, and φ is satisfiable iff $\mathrm{Def}_\Lambda(\varphi)$ is satisfiable.*

Corollary 3.1. *For any given GF1$^-$ formula φ and any $\Lambda \subseteq \mathrm{Pos}(\varphi)$, $\mathrm{Def}_\Lambda(\varphi)$ can be computed in polynomial time, and φ is satisfiable iff $\mathrm{Def}_\Lambda(\varphi)$ is satisfiable.*

We denote the result of the transformation of a first-order formula φ to clausal form by $\mathrm{Cls}(\varphi)$. We assume that in this transformation the free variables of φ are treated as existentially quantified and are replaced by distinct Skolem constants.

The use of structural transformation prior to the conversion to clausal form has two major advantages.

1. If a first-order formula is translated directly to its clausal form $\mathrm{Cls}(\varphi)$, the size of $\mathrm{Cls}(\varphi)$ can be exponential in the size of φ. If Λ is the set of all positions of (non-atomic) subformulae of φ, then the size of $\mathrm{Cls}(\mathrm{Def}_\Lambda(\varphi))$ is linear in the size of φ.
2. The application of structural transformation considerably simplifies the form of clauses that are obtained from φ.

In the case of GF1$^-$, we require that Λ contains the positions of all non-atomic subformulae of the formula under consideration 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 are referred to as *definitional*

$Q_\varphi(\overline{a})$	if φ is the input formula
$\neg Q_\varphi(\overline{x}) \vee \neg P(\overline{x})$	if $\varphi = \neg P(\overline{x})$
$\neg Q_\varphi(\overline{x}) \vee \neg \mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$	if $\varphi = \forall \overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y}))$
$\left.\begin{array}{l} \neg Q_\varphi(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})}) \\ \neg Q_\varphi(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})}) \end{array}\right\}$	if $\varphi = \exists \overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \wedge \phi(\overline{y}))$
$\left.\begin{array}{l} \neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) \\ \neg Q_\varphi(\overline{z}) \vee Q_\psi(\overline{y}) \end{array}\right\}$	if $\varphi = \phi(\overline{x}) \wedge \psi(\overline{y}) \quad \overline{z} = \overline{x} \cup \overline{y}$
$\neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) \vee Q_\psi(\overline{y})$	if $\varphi = \phi(\overline{x}) \vee \psi(\overline{y}) \quad \overline{z} = \overline{x} \cup \overline{y}$

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

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. To simplify our presentation, we assume that if φ is an atomic formula $P(\overline{x})$, then the symbol Q_φ stands for the predicate symbol P. By definition, we let $\mathcal{G}(\overline{x}, \overline{y})$ represent 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 $f_k(x_1, \ldots, x_n)$, where the arguments of each of the f_k are exactly the elements of $\overline{x}$.

Theorem 3.2. *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 linear time, and (iii) φ is satisfiable iff N is satisfiable.*

4. Hyperresolution for GF1$^-$

To decide GF1$^-$ we use a calculus $\mathsf{R}^{\mathsf{hyp}}$ 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:

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

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 $\neg A_{n+1} \vee \cdots \vee \neg A_{2n} \vee D$ is referred to as the *negative* premise and all the other premises in the resolution rule are referred to as *positive* premises. Hence, the positive premises of a hyperresolution inference step have to be positive clauses.

Factors are generated by the positive factoring rule:

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

where C is a positive clause and 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 to avoid applications of splitting to clauses containing duplicate literals.

The splitting rule is similar to disjunction elimination in semantic tableaux.

$$\frac{N \cup \{C_1 \vee C_2\}}{N \cup \{C_1\} \mid N \cup \{C_2\}}$$

where C_1 and C_2 are variable disjoint. That is, if the clause set N contains a clause C which can be split into variable disjoint clauses C_1 and C_2. The original clause becomes redundant and the resolution refutation is continued independently on $N \cup \{C_1\}$ or $N \cup \{C_2\}$.

A *derivation* from a set of clauses N is a finitely branching tree T with root N. The tree is expanded either by adding a successor node N'' to one of the leaf nodes N' of the tree such that $N'' = N' \cup \{C\}$ where C is either a factor or a hyperresolvent of clauses in N' or by adding two successor nodes $N' \cup \{C\}$ and $N' \cup \{D\}$ to a leaf node $N' \cup \{C \vee D\}$ where C and D are variable disjoint. A derivation T is a *refutation* if for every path $N = N_0, N_1, \ldots$ the clause set $\bigcup_j N_j$ contains the empty clause. A derivation T from N is called *fair* if for any path $N = N_0, N_1, \ldots$ in the tree T, with *limit* $N_\infty = \bigcup_j \bigcap_{k \geq j} N_k$, it is the case that each clause C that can be deduced from non-redundant premises in N_∞ is contained in some set N_j.

The calculus is compatible with a general notion of redundancy (Bachmair and Ganzinger, 1994, 2001).

Theorem 4.1 (Bachmair and Ganzinger, 1994; Bachmair et al., 1993). *Let T be a fair $\mathsf{R}^{\mathsf{hyp}}$ derivation from N. If $N, N_1, \ldots$ is a path with limit N_∞, then N_∞ is saturated up to redundancy. Furthermore, N is satisfiable if and only if there exists a path in T with limit N_∞ such that N_∞ is satisfiable.*

Theorem 4.2 (Bachmair and Ganzinger, 1994; Bachmair et al., 1993). *Let T be a fair $\mathsf{R}^{\mathsf{hyp}}$ derivation from N. N is unsatisfiable if and only if for every path $N = N_0, N_1, \ldots$ the clause set $\bigcup_j N_j$ contains the empty clause.*

We restrict our attention to derivations generated by strategies such that the positive premises of any hyperresolution step are positive ground unit clauses. For $\mathrm{GF1}^-$ this can be achieved by performing suitable splitting and factoring inferences before hyperresolution inferences. Furthermore, we assume that no inference step is performed twice with the same premises. Since we are able to prove termination of any such derivation for sets of $\mathrm{GF1}^-$ clauses, any such strategy is fair.

For the classes of clause sets we consider in this paper the positive premises are always ground, in particular, because we use splitting, the positive premises are always 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. (1999), and are defined next. The intuition underlying these notions is that the uni-nodes represent the vertices and bi-nodes the edges in a bidirectional tree. Uni-node clauses can be viewed as local constraints and bi-node clauses as transitional constraints.

A (multi-)set $\{t_1, \ldots, t_n\}$ (or sequence $\bar{t} = (t_1, \ldots, t_n)$) 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 (multi-)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 it is a unit clause L and the set of the arguments of L is a uni-node (bi-node). The empty clause $\perp$ is a special type of uni-node without direct successors.

The sets $\{a, a, b\}$, $\{h_Q(a, b)\}$, $\{h_Q(a, b), g_Q(a, b)\}$, $\{h_Q(a, b), g_Q(a, b), f_Q(a, b)\}$ 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 an existentially quantified subformula and f_Q, 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 $\{a, b\}$. The set $\{a, f_Q(a, b)\}$ is neither a uni-node nor a bi-node.

Lemma 4.1. *Suppose a finite signature is given.*

1. *The cardinality of any uni-node and any bi-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.*

Proof. In order to prove property 1 we use the definition of a uni-node (above) and the assumption that the signature is finite. Let X be a uni-node. Then by definition either each element in X is a constant or each element has the form $t_i = f_Q(\overline{s})$, for a fixed $\overline{s}$ and a fixed predicate symbol Q. By assumption the signature is finite, thus there are only finitely many distinct elements in the uni-node X which are constants. The number of distinct terms in X which have the form $t_i = f_Q(\overline{s})$ is bounded by the number of function symbols associated with the predicate symbol Q. The cardinality of a bi-node is finite, too, because it is the disjoint union of two uni-nodes.

In the proof of 2 we use the definition of a uni-node and property 1. Assume $\mathcal{G}(\overline{s}, \overline{t})$ is a uni-node. Then by definition of a uni-node, either all the s_i in $\overline{s}$ and all t_j in $\overline{t}$ are constants, or there exists a sequence of ground terms $\overline{u}$ such that all terms s_i, t_j have the form $f_Q(\overline{u})$, where f_Q is a function symbol associated with predicate symbol Q. Since the number of constants as well as the number of function symbols f_Q associated with predicate symbol Q is finite, and since $\overline{s}$ is given, the number of uni-nodes $\mathcal{G}(\overline{s}, \overline{t})$ is finitely bounded. Assume $\mathcal{G}(\overline{s}, \overline{t})$ is a bi-node, and $\overline{t}$ is a direct successor of $\overline{s}$. By definition any direct successor of a uni-node sequence (set) of terms $\overline{s}$ is of the form $\{f_Q^1(\overline{u}), \ldots, f_Q^n(\overline{u})\}$, for some n, where $f_Q^1, \ldots, f_Q^n$ are function symbols, associated with a predicate symbol Q and $\overline{u} = \overline{s}$, and $\overline{u}$ is non-empty. So by the same argument as before, the number of direct successors of $\overline{s}$ is finitely bounded, since $\overline{s} = \overline{u}$ is given. In the case where $\overline{s}$ is a direct successor of $\overline{t}$, each s_i in $\overline{s}$ has the form $f_Q(\overline{u})$ with $\overline{t} = \overline{u}$. Since the number of possibilities of forming $\overline{t}$ given $\overline{u}$ is finitely bounded, the result follows. This proves property 2.

Property 3 follows from properties 1 and 2 and by the definitions of a uni-node and a bi-node. □

It is worth noting that without a restriction on the depth of terms, the number of uni-nodes is not finitely bounded for a finite signature. Thus, an important result for the proof of termination is Lemma 4.4 which proves the existence of a term depth bound for inferred clauses.

What are the properties of inferred clauses and 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 steps are uni-nodes. The exceptions are inferences with definitional clauses of the form $\neg Q_\psi(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$, which produce bi-node conclusions. As factoring is applied only to positive clauses, and positive clauses in any $\mathsf{R}^{\mathsf{hyp}}$ derivation for $\mathrm{GF1}^-$ clauses (more generally, for any range restricted set of clauses) are always ground, factoring has the effect of eliminating duplicate literals in ground clauses. For this reason no special consideration is given to factoring inference steps in subsequent proofs.

Lemma 4.2. *In any $\mathsf{R}^{\mathsf{hyp}}$ derivation from N:*

1. *The negative premise of any inference step is a definitional clause while at least one of the positive premises is a uni-node.*
2. *All derived clauses are either empty or positive ground clauses which can be split into positive ground unit clauses of the form: $Q_\psi(\overline{s})$ 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.*
3. *If $Q_\psi(\overline{s})$ or $\mathcal{G}(\overline{s})$ are uni-nodes occurring in the derivation, then all terms in $\overline{s}$ have the same depth. If $\mathcal{G}(\overline{s}, \overline{t})$ is a bi-node occurring in the derivation and $\overline{t}$ is a direct successor of $\overline{s}$, then all terms in $\overline{s}$ have the same depth d and all terms in $\overline{t}$ have the same depth $d + 1$.*

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. Since all arguments of $Q_\varphi(\overline{a})$ are constants, they have the same depth. The inductive hypothesis is that properties (1)–(3) hold for the premises and conclusions of the first n inference steps in any path of the derivation.

In the inductive step we consider inference step $n + 1$. Consider the resolution steps where the negative premise is a definitional clause introduced for an existentially quantified $\mathrm{GF1}^-$ formula. Assume that the positive premise $Q_\psi(\overline{s})$ is a uni-node. There are two possibilities. (i) The negative premise is a clause $\neg Q_\psi(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})})$. The argument set of the conclusion $Q_\phi(\overline{f(\overline{s})})$ is a sequence 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_ψ. This means that the conclusion of this resolution step is a uni-node. Furthermore, if all terms in $\overline{s}$ have the same depth, then also all the terms in $\overline{f(\overline{s})}$ have the same depth. (ii) The negative premise is a clause $\neg Q_\psi(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$. By definition, any of the function symbols f_i in $\overline{f(\overline{s})}$ is associated with Q_ψ. Consequently, $\overline{f(\overline{s})}$ is a direct successor of $\overline{s}$. Therefore the conclusion $\mathcal{G}(\overline{s}, \overline{f(\overline{s})})$ is a bi-node. Obviously, since all terms in $\overline{s}$ have the same depth d, all the terms in $\overline{f(\overline{s})}$ have depth $d + 1$.

Consider also hyperresolution inference steps involving a definitional clause $\neg Q_\psi(\overline{x}) \vee \neg \mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$ corresponding to universally quantified GF1$^-$ formula. The two positive premises have the form $Q_\psi(\overline{s})$ and $\mathcal{G}(\overline{s}, \overline{t})$. By assumption $Q_\psi(\overline{s})$ 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. By the inductive hypothesis, all the argument terms of the uni-node $\mathcal{G}(\overline{s}, \overline{t})$ have the same depth. This implies that also all the terms of $Q_\phi(\overline{t})$ have the same depth. 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. The grouping restriction ensures that the sequences of variables $\overline{x}$ and $\overline{y}$ have the same length as the sequences of terms $\overline{s}$ and $\overline{t}$, respectively. This implies that all the variables in $\overline{x}$ are instantiated with terms from $\overline{s}$ only. Similarly for $\overline{y}$ and $\overline{t}$, i.e. there is no variable in $\overline{x}$ (resp. $\overline{y}$) which can be instantiated with a term from $\overline{t}$ (resp. $\overline{s}$). Hence, the conclusion has the form $Q_\phi(\overline{t})$ where $\overline{t}$ is a ground uni-node. Since $\overline{t}$ is a direct successor of $\overline{s}$, or vice versa, it follows by the inductive hypothesis that all terms in $\overline{t}$ have the same depth.

The remaining inference possibilities are resolution steps between a positive premise $Q_\psi(\overline{s})$ and negative premises of the form $\neg Q_\psi(\overline{x}) \vee \neg P(\overline{x})$, $\neg Q_\psi(\overline{x}) \vee Q_\phi(\overline{y})$ where $\overline{y} \subseteq \overline{x}$, or $\neg Q_\psi(\overline{x}) \vee Q_\phi(\overline{y}) \vee Q_\vartheta(\overline{z})$ where $\overline{y}, \overline{z} \subseteq \overline{x}$. Obviously, properties 1–3 hold for such inferences. □

Lemma 4.3. *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 property follows immediately from the case analysis in the proof of Lemma 4.2. Inspecting all resolution inference 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 property is a direct consequence of Lemma 4.2 and property 1. □

The main technical lemma is the following:

Lemma 4.4. *Let* φ *be a* GF1$^-$ *formula and let* N *be the corresponding clause set. The term depth of any clause in a derivation from* N *is bounded by one plus the guarded quantifier depth of* φ*.*

Proof. If Q_ψ is a predicate symbol introduced by the structural transformation of φ for a subformula at position λ in φ, then we can define the guarded quantifier depth $\mathrm{gqd}(Q_\psi)$ of Q_ψ as $\mathrm{gqd}(\varphi|_\lambda)$, i.e. $\mathrm{gqd}(Q_\psi) = \mathrm{gqd}(\psi)$. The guarded quantifier depth of any other predicate symbol is defined to be zero.

We define a complexity measure $\nu(C)$ for uni-nodes and show for any $\mathsf{R}^{\mathsf{hyp}}$ derivation from N that

(i) the term depth of a positive unit clause $P(\overline{s})$ in the derivation does not exceed $\mathrm{gqd}(\varphi) - \mathrm{gqd}(P) + 1$, and

(ii) the complexity of the conclusion of an arbitrary inference step is always smaller than the complexity of one of the uni-node premises of the inference step.

The particular complexity measure we use, allows us to establish the required upper bound on the term depth of clauses in the derivation.

We define a partial ordering $\succ_d$ on predicate symbols by $S_1 \succ_d S_2$, 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 $A = (\neg)P(t_1, \ldots, t_n)$ be a uni-node. Then $\nu(A) = (\text{gqd}(P), \text{dp}(A), P)$, that is, the complexity measure of $\nu(A)$ of A is given by the ordered tuple consisting of the guarded quantifier depth of the predicate symbol of A, the term depth of A, and the predicate symbol of A. For the empty clause we define $\nu(\bot) = (0, 0, tt)$ where tt is a new symbol which is smallest with respect to $\succ_d$. We compare complexity measures by the lexicographic combination $\succ$ of $>_\mathbb{N}$, $>_\mathbb{N}$, and $\succ_d$.

To prove (ii) it would actually be sufficient if the complexity measure of a uni-node A would consist of the predicate symbol of A alone. However, the additional components of $\nu(A)$ as defined above allow further interesting observations about $\mathsf{R}^{\mathsf{hyp}}$-derivations from N.

The proof proceeds by induction over the number of inference steps in the derivation from N. In the base case, as yet no inference steps have been performed and the only positive uni-node in N is the unit clause $Q_\varphi(\overline{a})$ of term depth 1. Since $\text{gqd}(Q_\varphi) = \text{gqd}(\varphi)$, $Q_\varphi(\overline{a})$ satisfies the upper bound of $\text{gqd}(\varphi) - \text{gqd}(Q_\varphi) + 1 = 1$ on the term depth of positive uni-nodes in the derivation.

The inductive hypothesis is that properties (i) and (ii) hold for the first n inference steps of the derivation. We consider inference step $n + 1$. If this inference step is the derivation of the empty clause from two positive uni-nodes $Q_\psi(\overline{s})$ and $Q_{\neg\psi}(\overline{s})$, and the definitional clause $\neg Q_{\neg\psi}(\overline{x}) \vee \neg Q_\psi(\overline{x})$, then ψ is an atom. Since $\text{dp}(Q_{\neg\psi}(\overline{s})) \geq 1$, we have $\nu(Q_{\neg\psi}(\overline{s})) = (0, \text{dp}(Q_{\neg\psi}(\overline{s})), Q_{\neg\psi}) \succ (0, 0, tt) = \nu(\bot)$. So, the inference step satisfies property (ii).

Consider the derivation of $Q_\phi(\overline{t})$ from $Q_\psi(\overline{s})$ and $\neg Q_\psi(\overline{z}) \vee Q_\phi(\overline{x})$ where $\overline{x} \subseteq \overline{z}$ and $\overline{t} \subseteq \overline{s}$. The negative premise of this inference step is a definitional clause introduced for a conjunction, that is, $\psi = \phi \wedge \vartheta$. Therefore, $\text{gqd}(Q_\psi) = \text{gqd}(Q_\phi)$. Let $\text{gqd}(Q_\psi)$ be d_q. Let the term depth of $Q_\psi(\overline{s})$ be d_t. By Lemma 4.2 all terms in $\overline{s}$ have the same depth. Since $\overline{t} \subseteq \overline{s}$ it follows that the term depth of $Q_\phi(\overline{t})$ is also d_t. Thus, if property (i) holds for $Q_\psi(\overline{s})$ then it also holds for $Q_\phi(\overline{t})$. Furthermore, it follows from $Q_\psi \succ_d Q_\phi$ that $\nu(Q_\psi(\overline{s})) = (d_q, d_t, Q_\psi) \succ (d_q, d_t, Q_\phi) = \nu(Q_\phi(\overline{t}))$. The case involving an inference step with a definitional clause of the form $\neg Q_\psi(\overline{z}) \vee Q_\phi(\overline{x}) \vee Q_\vartheta(\overline{y})$ introduced for a disjunctive subformula of φ is similar to the previous case.

Next we consider the derivation of $Q_\phi(\overline{f(\overline{s})})$ from $Q_\psi(\overline{s})$ and $\neg Q_\psi(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})})$. The negative premise of this inference step is a definitional clause introduced for an existentially quantified subformula $\exists \overline{x}(A \wedge \phi(\overline{y}))$ of φ. By the definition of the guarded quantifier depth of predicate symbols, $\text{gqd}(Q_\psi) = \text{gqd}(Q_\phi) + 1$. On the other hand, $\text{dp}(Q_\phi(\overline{f(\overline{s})})) = \text{dp}(Q_\psi(\overline{s})) + 1$. If $Q_\psi(\overline{s})$ satisfies property (i), that is, $\text{dp}(Q_\psi(\overline{s})) \leq \text{gqd}(\varphi) - \text{gqd}(Q_\psi) + 1$, then also

$$\begin{aligned} \text{dp}(Q_\phi(\overline{f(\overline{s})})) &= \text{dp}(Q_\psi(\overline{s})) + 1 \\ &\leq \text{gqd}(\varphi) - \text{gqd}(Q_\psi) + 2 = \text{gqd}(\varphi) - \text{gqd}(Q_\phi) + 1. \end{aligned}$$

Thus, property (i) also holds for $Q_\phi(\overline{f(\overline{s})})$. Concerning property (ii) we observe that $\mathrm{gqd}(Q_\psi) > \mathrm{gqd}(Q_\phi)$ implies

$$\begin{aligned}\nu(Q_\psi(\overline{s})) &= (\mathrm{gqd}(Q_\psi), \mathrm{dp}(Q_\psi(\overline{s})), Q_\psi) = (\mathrm{gqd}(Q_\phi) + 1, \mathrm{dp}(Q_\psi(\overline{s})), Q_\psi)\\ &\succ (\mathrm{gqd}(Q_\phi), \mathrm{dp}(Q_\psi(\overline{s})) + 1, Q_\phi) = (\mathrm{gqd}(Q_\phi), \mathrm{dp}(Q_\phi(\overline{f(\overline{s})})), Q_\phi)\\ &= \nu(Q_\phi(\overline{f(\overline{s})})).\end{aligned}$$

Therefore, property (ii) holds for this inference step.

The argument for inference steps with negative premises of the form $\neg Q_\psi(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$ is similar to the previous case. Let the positive premise be $Q_\psi(\overline{s})$ with term depth d_t. Note that $\mathcal{G}$ is not a predicate symbol introduced during the structure transformation of φ. So $\mathrm{gqd}(\mathcal{G}) = 0$. Furthermore, the definitional clause under consideration has been introduced for an existentially quantified subformula of φ. Thus, $\mathrm{gqd}(Q_\psi) \geq 1$. If $Q_\psi(\overline{s})$ satisfies property (i), that is, $\mathrm{dp}(Q_\psi(\overline{s})) \leq \mathrm{gqd}(\varphi) - \mathrm{gqd}(Q_\psi) + 1$, then

$$\begin{aligned}\mathrm{dp}(\mathcal{G}(\overline{s}, \overline{f(\overline{s})})) &= \mathrm{dp}(Q_\psi(\overline{s})) + 1 \leq \mathrm{gqd}(\varphi) - \mathrm{gqd}(Q_\psi) + 2\\ &\leq \mathrm{gqd}(\varphi) + 1 = \mathrm{gqd}(\varphi) - \mathrm{gqd}(\mathcal{G}) + 1.\end{aligned}$$

Obviously, $\mathrm{gqd}(Q_\psi) > \mathrm{gqd}(\mathcal{G}) = 0$. It follows that

$$\begin{aligned}\nu(Q_\psi(\overline{s})) &= (\mathrm{gqd}(Q_\psi), \mathrm{dp}(Q_\psi(\overline{s})), Q_\psi)\\ &\succ (0, \mathrm{dp}(Q_\psi(\overline{s})) + 1, \mathcal{G}) = (\mathrm{gqd}(\mathcal{G}), \mathrm{dp}(\mathcal{G}(\overline{s}, \overline{f(\overline{s})})), \mathcal{G}).\end{aligned}$$

Therefore, property (ii) holds for this inference step.

Finally, consider the derivation of $Q_\phi(\overline{t})$ from unit clauses $Q_\psi(\overline{s})$, $\mathcal{G}(\overline{s}, \overline{t})$, and the definitional clause $\neg Q_\psi(\overline{x}) \vee \neg\mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$. It follows from Lemma 4.2(3) that $\mathrm{dp}(\mathcal{G}(\overline{s}, \overline{t})) \leq \mathrm{dp}(Q_\psi(\overline{s})) + 1$. So, $\mathrm{dp}Q_\phi(\overline{t}) \leq \mathrm{dp}(Q_\psi(\overline{s})) + 1$. Since the definitional clause under consideration has been introduced for a universally quantified subformula of φ we have that $\mathrm{gqd}(Q_\psi) = \mathrm{gqd}(Q_\phi) + 1$. If $Q_\psi(\overline{s})$ satisfies property (i), then

$$\mathrm{dp}(Q_\phi(\overline{t})) \leq \mathrm{dp}(Q_\psi(\overline{s})) + 1 \leq \mathrm{gqd}(\varphi) - \mathrm{gqd}(Q_\psi) + 2 = \mathrm{gqd}(\varphi) - \mathrm{gqd}(Q_\phi) + 1.$$

Concerning the complexity of $\nu(Q_\phi(\overline{t}))$ and $\nu(Q_\psi(\overline{s}))$ we obtain

$$\begin{aligned}\nu(Q_\psi(\overline{s})) &= (\mathrm{gqd}(Q_\psi), \mathrm{dp}(Q_\psi(\overline{s})), Q_\psi) = (\mathrm{gqd}(Q_\phi) + 1, \mathrm{dp}(Q_\psi(\overline{s})), Q_\psi)\\ &\succ (\mathrm{gqd}(Q_\phi), \mathrm{dp}(Q_\phi(\overline{t})), Q_\phi) = \nu(Q_\phi(\overline{t})).\end{aligned}$$

This concludes the proof of properties (i) and (ii) for all inference steps and clauses in an arbitrary derivation from N. □

The considerations in the proof allow the following additional observations:

- If $\mathcal{G}$ is a predicate symbol of a guard atom in φ, and $\{Q_{\psi_1}, \ldots, Q_{\psi_n}\}$ is the set of all predicate symbols such that N contains a clause of the form $\neg Q_{\psi_i}(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$ or $\neg Q_{\psi_i}(\overline{x}) \vee \neg\mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$, then the term depth of any bi-node $\mathcal{G}(\overline{s}, \overline{t})$ occurring in a derivation from N is bounded from above by $\max(\{\mathrm{gqd}(\varphi) - \mathrm{gqd}(Q_{\psi_i}) + 1 \mid 1 \leq i \leq n\})$.
- If a uni-node $Q_\psi(\overline{s})$ with term depth d_t is used as a premise in an inference step with a uni-node conclusion $Q_\phi(\overline{t})$ with term depth greater than d_t, then $\mathrm{gqd}(Q_\psi) > \mathrm{gqd}(Q_\phi)$.

It is interesting to compare Lemma 4.4 with the corresponding results for the modal logic $K_{(m)}(\cap, \cup, \smile)$ in de Nivelle et al. (2000) and Hustadt and Schmidt (2000b). Theorem 7.3 in de Nivelle et al. (2000) states that the depth of any clause derived from the translation of a $K_{(m)}(\cap, \cup, \smile)$ formula ϑ in negation normal form using resolution with maximal selection (or hyperresolution) is bounded by the number of diamond (existential) subformulae in ϑ. Because the clausal form of the particular translation of $K_{(m)}(\cap, \cup, \smile)$ formulae are instances of GF1$^-$ clauses, one might expect, in analogy, that the term depth of any clause in a derivation from $\mathrm{Cls}(\mathrm{Def}_\Lambda(\varphi))$ for a GF1$^-$ formula φ in negation normal form is bounded by the number of existentially quantified subformulae in φ. The following example shows that this bound is too tight.

$$\varphi = R(x, x) \wedge \forall y(R(x, y) \to \forall z(R(y, z) \to \exists u(R(z, u) \wedge P(u)))).$$

The corresponding clause set N contains the following clauses:

(1) $Q_0(a)$
(2) $\neg Q_0(x) \vee R(x, x)$
(3) $\neg Q_0(x) \vee \neg R(x, y) \vee Q_1(y)$
(4) $\neg Q_1(x) \vee \neg R(x, y) \vee Q_2(y)$
(5) $\neg Q_2(x) \vee R(x, f(x))$
(6) $\neg Q_2(x) \vee P(f(x))$.

We obtain the following derivation by $\mathsf{R}^{\mathsf{hyp}}$ from N.

[(1), (2)]	(7)	$R(a, a)$
[(1), (7), (3)]	(8)	$Q_1(a)$
[(8), (7), (4)]	(9)	$Q_2(a)$
[(9), (5)]	(10)	$R(a, f(a))$
[(9), (6)]	(11)	$P(f(a))$
[(1), (10), (3)]	(12)	$Q_1(f(a))$
[(8), (10), (4)]	(13)	$Q_2(f(a))$
[(13), (5)]	(14)	$R(f(a), f(f(a)))$
[(13), (6)]	(15)	$P(f(f(a)))$
[(12), (14), (4)]	(16)	$Q_2(f(f(a)))$
[(16), (5)]	(17)	$R(f(f(a)), f(f(f(a))))$
[(16), (6)]	(18)	$P(f(f(f(a))))$.

Here, [(1), (7), (3)] denotes that the negative premise (3) is resolved with the two positive premises (1) and (7).

The guarded quantifier depth of φ is 3. By Lemma 4.4 the term depth of clauses in any derivation from N is bounded by $\mathrm{gqd}(\varphi) + 1 = 4$. This is obviously the case. However, φ contains only one existentially quantified subformula. A tighter bound on the term depth of derived clauses based solely on the number of existentially quantified subformulae of a GF1$^-$ formula is not possible.

The example formula φ also shows that GF1$^-$ allows the formulation of a form of ‘local reflexivity’ which means it shares some properties with the fragment of first-logic corresponding to the propositional modal logic KT, which is characterized by the class of reflexive frames. In fact, Lemma 4.4 describes one of these properties, namely, that the term depth of derived clauses is linear in the number of universal and existential quantifiers in the input formula.

Lemma 4.5. *Let φ be a formula in* GF1$^-$ *and let N be the corresponding clause set. The number of clauses derivable from N is finitely bounded.*

Proof. By Lemma 4.2(2) all derived clauses are ground clauses. By Lemma 4.4 there is an upper bound on the term depth of these derived clauses. Since there are only boundedly many ground clauses up to a given term depth, the derivation must eventually terminate. □

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

Theorem 4.3. *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.*
2. *If T is a fair derivation from N then (i) If $N(= N_0), N_1, \ldots$ is a path with limit N_∞, N_∞ is saturated up to redundancy. (ii) φ is satisfiable if and only if there exists a path in T with limit N_∞ such that N_∞ is satisfiable. (iii) φ is unsatisfiable if and only if for every path $N(= N_0), N_1, \ldots$ the clause set $\bigcup_j N_j$ contains the empty clause.*

Proof. This is a consequence of Lemmas 4.1–4.5, Corollary 3.1 and Theorems 3.2, 4.1 and 4.2. □

The decision procedure we have presented looks very similar to the decision procedures based on refinements of resolution using maximal selection of negative literals for expressive modal logics and description logics, which are described in de Nivelle et al. (2000) and Hustadt and Schmidt (1999, 2000a,b). The main difference is the way in which we prove termination. In the proofs of de Nivelle et al. (2000), 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 an ordering 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. The proofs in this section extend to the generalizations of GF1$^-$ discussed in the Section 8.

5. Model building for GF1$^-$

It is well-known that hyperresolution, like tableaux methods, can be used to construct models for satisfiable formulae (Fermüller et al., 2001). 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 is given by the set of ground unit clauses in an open branch of the derivation tree.

Theorem 5.1. *Assume that φ is a formula in* GF1$^-$. *Let N be the clausal form of* $\mathrm{Def}_\Lambda(\varphi)$, *and let N_∞ denote the saturation of N by* $\mathsf{R}^{\mathsf{hyp}}$. *Let H be the set of positive ground unit clauses in N_∞. If N_∞ does not contain the empty clause, then H is a model of N_∞ and N.*

Proof. In order to prove that H is a model of N_∞ we have to show that every ground instance of a clause in N_∞ is true in H.

The maximally split conclusions of the resolution derivation leading to N_∞ are positive ground unit clauses by Lemma 4.2, and are true in H, because by definition H contains the positive ground unit clauses in N_∞. The remaining clauses in N_∞ are the definitional clauses which were already present in N. We consider a ground instance $C\sigma$ of such a definitional clause. $C\sigma$ has the form $\neg A_1\sigma \vee \cdots \vee \neg A_n\sigma \vee B_1\sigma \vee \cdots \vee B_m\sigma$, with $n > 0$ and $m \geq 0$.

Case 1. Assume that there exists an i, $1 \leq i \leq n$, such that $A_i\sigma \notin H$. Then $\neg A_i\sigma$ is true in H and, therefore, $C\sigma$ is true in H.

Case 2. Assume that $A_i\sigma \in H$ for all i, $1 \leq i \leq n$. We have to show that there exists a j, $1 \leq j \leq m$, such that $B_j\sigma \in H$. Since $A_i\sigma \in H$ we have that $A_i\sigma \in N_\infty$ for every i, $1 \leq i \leq n$. Thus a hyperresolution inference of C with positive premises $A_1\sigma, \ldots, A_n\sigma$ is possible. Since each clause in N is range restricted, the conclusion $B_1\sigma \vee \cdots \vee B_m\sigma$ of the inference step is ground. Due to the application of splitting one of the $B_j\sigma$, $1 \leq j \leq m$, has been added to the clausal set. As we do not use any form of redundancy elimination, a clause that is once generated is never deleted. (It is straightforward to see that, even if we allow subsumption or other forms of redundancy elimination, $B_j\sigma$ still persists.) So, $B_j\sigma$ is an element of N_∞ and, therefore, also an element of H. Thus $C\sigma$ is true in H.

Therefore, we have proved that H is a model of N and N_∞. □

Corollary 5.1. *A finite model for every satisfiable formula in* GF1$^-$ *can be constructed on the basis of* $\mathsf{R}^{\mathsf{hyp}}$.

Proof. Let φ be a satisfiable formula in GF1$^-$ with free variables $x_1, \ldots, x_k$. Let N be the clausal form of $\mathrm{Def}_\Lambda(\varphi)$, let N_∞ denote the saturation of N by $\mathsf{R}^{\mathsf{hyp}}$, and let H be the Herbrand model of N_∞. Furthermore, let $a_1, \ldots, a_k$ be the Skolem constants introduced for the free variables of φ in the transformation of φ to clausal form. We construct an interpretation $\mathcal{M}$ as follows. The domain M of $\mathcal{M}$ contains all ground terms in N_∞. There are only finitely many ground terms in N_∞, thus the domain M is finite. Note that φ contains no constant or function symbols. The interpretation P^I of a predicate symbol of arity n is defined by $(t_1, \ldots, t_n) \in P^I$ iff $P(t_1, \ldots, t_n) \in N_\infty$ for all $t_1, \ldots, t_n$ in M.

Next we prove that there exists an assignment g such that $\mathcal{M}, g \models \varphi$. The proof is by induction over the structure of φ starting with its atomic subformulae. We show that if ψ is a subformula of φ with free variables $x_1, \ldots, x_n$ and there exist terms $t_1, \ldots, t_n$ such that $Q_\psi(t_1, \ldots, t_n)$ is in N_∞, then $\mathcal{M}, g \models \psi$ for any assignment g with $g(x_i) = t_i$, for every i, $1 \leq i \leq n$. Then, since $Q_\varphi(a_1, \ldots, a_k)$ is in N_∞, we know that $\mathcal{M}, g \models \varphi$ for the assignment g with $g(x_i) = a_i$, for every k, $1 \leq i \leq k$.

Base case. Consider an atomic subformula $\psi = P(x_1, \ldots, x_n)$. Assume that there is a ground unit clause $P(t_1, \ldots, t_n)$ in N_∞, where $t_1, \ldots, t_n$ are ground terms. By construction

$t_1, \ldots, t_n$ are also elements of M and by definition of P^I, $(t_1, \ldots, t_n) \in P^I$. Thus, $\mathcal{M}, g \models \psi$ for any g with $g(x_i) = t_i$, $1 \le i \le n$.

Suppose ψ has the form $\neg\alpha$ with free variables $x_1, \ldots, x_n$. Since φ is in negation normal form, α is an atomic formula $P(x_1, \ldots, x_n)$. We assume that there are terms $t_1, \ldots, t_n$ such that $Q_\psi(t_1, \ldots, t_n) \in N_\infty$. As a consequence, we have that $\neg Q_\psi(x_1, \ldots, x_n) \vee \neg Q_\alpha(x_1, \ldots, x_n) \in N$ and $Q_\psi(t_1, \ldots, t_n) \in N_\infty$. Then $Q_\alpha(t_1, \ldots, t_n)$ is not an element of N_∞, since otherwise we would be able to deduce the empty clause. So, by definition of P^I, $(t_1, \ldots, t_n) \notin P^I$. Let g be any assignment with $g(x_i) = t_i$ for all i, $1 \le i \le n$. Then, $\mathcal{M}, g \not\models \alpha$ and therefore $\mathcal{M}, g \models \psi$.

Inductive hypothesis. If ω is a strict subformula of some subformula ψ of φ with free variables $x_1, \ldots, x_n$ and there exist terms $t_1, \ldots, t_n$ such that $Q_\omega(t_1, \ldots, t_n)$ is in N_∞, then $\mathcal{M}, g \models \omega$ for any assignment g with $g(x_i) = t_i$, for every i, $1 \le i \le n$.

Inductive step. In the inductive step we look at a subformula ψ of φ with free variables $x_1, \ldots, x_n$. We assume there exist terms $t_1, \ldots, t_n$ such that $Q_\psi(t_1, \ldots, t_n)$ is in N_∞. We want to show that $\mathcal{M}, g \models \psi$ for any assignment g with $g(x_i) = t_i$, for every i, $1 \le i \le n$.

Case 1. Suppose ψ has the form $\alpha_1 \wedge \alpha_2$. Then

$$\begin{aligned}
&\neg Q_\psi(z_1, \ldots, z_n) \vee Q_{\alpha_1}(x_1, \ldots, x_m) \in N \\
&\neg Q_\psi(z_1, \ldots, z_n) \vee Q_{\alpha_2}(y_1, \ldots, y_k) \in N \\
&\ Q_\psi(t_1, \ldots, t_n) \in N_\infty
\end{aligned}$$

where $\{z_1, \ldots, z_n\} = \{x_1, \ldots, x_m\} \cup \{y_1, \ldots, y_k\}$. We assume without loss of generality that $z_1, \ldots, z_n$, $x_1, \ldots, x_m$, and $y_1, \ldots, y_k$ are the free variables of ψ, α_1, and α_2, respectively. Let σ be the substitution $\{z_1/t_1, \ldots, z_n/t_n\}$. Then also $Q_{\alpha_1}(x_1, \ldots, x_m)\sigma$ and $Q_{\alpha_2}(y_1, \ldots, y_k)\sigma$ are in N_∞, since N_∞ is saturated under $\mathsf{R}^{\mathsf{hyp}}$. Let g_1 be any assignment with $g_1(x_i) = x_i\sigma$ for every i, $1 \le i \le m$ and let g_2 be any assignment with $g_2(y_j) = y_j\sigma$ for every j, $1 \le j \le k$. By the inductive hypothesis, $\mathcal{M}, g_1 \models \alpha_1$ and $\mathcal{M}, g_2 \models \alpha_2$. Now, let g be any assignment with $g(z_i) = z_i\sigma = t_i$, for every i, $1 \le i \le n$. Note that g coincides with g_1 and g_2 on the free variables of α_1 and α_2, respectively. Thus, $\mathcal{M}, g \models \alpha_1 \wedge \alpha$ and $\mathcal{M}, g \models \psi$.

Case 2. The case that ψ has the form $\beta_1 \vee \beta_2$ is analogous to the previous case.

Case 3. Consider a universally quantified subformula $\psi = \forall\ \overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y}))$. We assume there are terms $t_1, \ldots, t_n$ such that $Q_\psi(t_1, \ldots, t_n)$ is in N_∞. N and N_∞ also contains

$$\neg Q_\varphi(x_1, \ldots, x_n) \vee \neg\mathcal{G}(x_1, \ldots, x_n, y_1, \ldots, y_m) \vee Q_\phi(y_1, \ldots, y_m).$$

Let $s_1, \ldots, s_m$ be arbitrary elements of M. First, assume that the ground unit clause $\mathcal{G}(t_1, \ldots, t_n, s_1, \ldots, s_m)$ is in N_∞. Then, we can derive $Q_\phi(s_1, \ldots, s_m)$. By the inductive hypothesis, any assignment h with $h(y_j) = s_j$ for every j, $1 \le j \le m$, $\mathcal{M}, h \models \phi$ holds. In addition, for $g'' = h^{[x_1,\ldots,x_n]}_{[t_1,\ldots,t_n]}$, we have $\mathcal{M}, g'' \models \mathcal{G}(\overline{x}, \overline{y})$ as well as $\mathcal{M}, g'' \models \phi$. Again, $\mathcal{M}, g'' \models \mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y})$. Second, assume that $\mathcal{G}(t_1, \ldots, t_n, s_1, \ldots, s_m)$ is not in N_∞. Then $\mathcal{M}, g'' \not\models \mathcal{G}(x_1, \ldots, x_n, y_1, \ldots, y_m)$. So, $\mathcal{M}, g'' \models \mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y})$. Taking both cases together we see that for any assignment g with $g(x_i) = t_i$ for every i, $1 \le i \le n$, $\mathcal{M}, g \models \mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y})$ and, therefore, $\mathcal{M}, g \models \psi$.

Case 4. Consider an existentially quantified subformula $\psi = \exists \overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \wedge \phi(\overline{y}))$. We assume that there are terms $t_1, \ldots, t_n$ such that $Q_\psi(t_1, \ldots, t_n) \in N_\infty$. N and N_∞ contain also the definitional clauses for Q_ψ, that is

$$\neg Q_\psi(x_1, \ldots, x_n) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})}) \in N \quad \text{and}$$
$$\neg Q_\psi(x_1, \ldots, x_n) \vee Q_\phi(\overline{f(\overline{x})}) \in N.$$

Since N_∞ is saturated under $\mathsf{R}^{\mathsf{hyp}}$, also

$$Q_\phi(f_1(t_1, \ldots, t_n), \ldots, f_m(t_1, \ldots, t_n)) \quad \text{and}$$
$$\mathcal{G}(t_1, \ldots, t_n, f_1(t_1, \ldots, t_n), \ldots, f_m(t_1, \ldots, t_n))$$

are in N_∞. Let g be any assignment with $g(x_i) = t_i$ for every i, $1 \leq i \leq t_i$. Furthermore, let g' be $g^{[y_1,\ldots,y_m]}_{[f_1(t_1,\ldots,t_n),\ldots,f_m(t_1,\ldots,t_n)]}$. By the inductive hypothesis, $\mathcal{M}, g' \models Q_\phi(\overline{y})$ and $\mathcal{M}, g' \models \mathcal{G}(\overline{x}, \overline{y})$. Thus, $\mathcal{M}, g \models \exists \overline{y} \mathcal{G}(\overline{x}, \overline{y}) \wedge Q_\phi(\overline{y})$ and, therefore, $\mathcal{M}, g \models \psi$. □

6. Some upper bounds

By considering the ground constraints in $\mathsf{R}^{\mathsf{hyp}}$ derivation trees it is possible to estimate the maximal computational space requirements for derivations in $\mathsf{R}^{\mathsf{hyp}}$ and to determine the maximal size of Herbrand models.

Lemma 6.1. *Let φ be a* GF1$^-$ *formula and N the corresponding set of clauses. Let Σ be the signature of N, let s be the size of Σ, let a be the maximum of (i) the maximal arity of function symbols in N and (ii) the maximal arity of predicate symbols in N. The space requirements of uni-node or bi-node clauses up to term depth d over Σ is of the order of magnitude $a^d s^{da^d}$.*

Proof. In a similar way as in de Nivelle and de Rijke (2003) we calculate the number of significant symbols (i.e. all symbols other than brackets and ',') of each uni-node depending on its term depth. Let a_1 and a_2 be the maximal arity of any of the function symbols and any of the predicate symbols, respectively. Then by definition $a = \max(a_1, a_2)$. By assumption the maximal number of significant symbols of a uni-node clause of term depth 1 is $1 + a_2 \leq 1 + a$. The maximal number of significant symbols of a uni-node clause of a term depth 2 is $1 + a_2(1 + a_1) \leq 1 + a + a^2$. The maximal size of a uni-node clause of term depth 3 is $1 + a_2(1 + a_1(1 + a_1)) \leq 1 + a(1 + a + a^2) \leq 1 + a + a^2 + a^3$. Thus the maximal size of a uni-node of term depth d is smaller than $1 + \cdots + a^d = (a^{d+1} - 1)/(a - 1) \leq d \cdot a^d$, when $a > 1$. Then the number of uni-nodes of depth d can be estimated by $s^{d \cdot a^d}$, where s stands for the total number of function symbols plus the total number of constant symbols plus the total number of predicate symbols. Then the space requirements for uni-nodes up to term depth d over a finite signature of size s is bounded by $s^{d \cdot a^d} \cdot a^d$.

The size of the bi-node clauses depends on the number of free variables of each subformula. Let m stand for the maximal number of free variables of each quantified subformula of φ. Then each bi-node clause of term-depth d consists of m uni-nodes of term-depth $d - 1$ and $a_2 - m$ uni-nodes of term-depth d. Hence the maximal number of significant symbols of a bi-node clause over a signature of a bounded size s is bounded by $s^{d \cdot a^d} + s^{(d-1) \cdot a^{d-1}}$, which in turn is bounded by $2s^{d \cdot a^d}$. □

Because the maximal term depth in derivations is linear in the size n of φ (Lemma 4.4), the space requirements of uni-node or bi-node clauses over Σ is of the order of magnitude $a^n s^{na^n}$.

Lemma 6.2. *Let φ be a* GF1$^-$ *formula and N the corresponding set of clauses. Let s and a be as in the previous lemma. Then the maximal size of the model of φ, constructed by the conclusions of the* $\mathsf{R}^{\mathsf{hyp}}$ *derivation, is of the order of magnitude $a^n s^{na^n}$, where n is the length of φ.*

Proof. The model of any satisfiable GF1$^-$ formula φ is constructed from the conclusions in the $\mathsf{R}^{\mathsf{hyp}}$ derivation. We estimate the size of the model depending on the maximal number of ground unit clauses which could be in it. The result follows then by Lemmas 4.4 and 6.1. □

Therefore:

Theorem 6.1. *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 size of the model is at most double exponential in the length of φ.*

In Georgieva et al. (2001) we consider complexity issues in more depth. More specifically, we describe a polynomial space decision procedure of optimal worst-case space and time complexity for GF1$^-$. We also consider the problem of minimal Herbrand model generation for GF1$^-$, discuss various approaches to this problem and compare their space complexities.

7. Semantic tableaux versus $\mathsf{R}^{\mathsf{hyp}}$

Next, we investigate the relationship between resolution and tableaux proof systems for GF1$^-$. We describe a tableaux proof system for GF1$^-$, which is abstracted from Lutz et al. (1999), and show that $\mathsf{R}^{\mathsf{hyp}}$ polynomially simulates the tableaux proof system for GF1$^-$, and vice versa.

Given two proof systems $\mathcal{A}$ and $\mathcal{B}$, the system $\mathcal{A}$ *polynomially simulates* the proof system $\mathcal{B}$ if there is a function g, computable in polynomial time, that maps proofs in $\mathcal{B}$ for any given formula φ to proofs in $\mathcal{A}$ for φ. A system $\mathcal{A}$ *polynomially simulates derivations* (as well as proofs) of a system $\mathcal{B}$ if there is a function g, computable in polynomial time, such that for any formula φ, g maps derivations from φ in $\mathcal{B}$ to derivations in $\mathcal{A}$ from φ (de Nivelle et al., 2000).

For a GF1$^-$ formula φ in negation normal form with free variables $\overline{x} = x_1, \ldots, x_n$ let $\varphi\{\overline{x}/\overline{a}\}$, where $\overline{a} = a_1, \ldots, a_n$, denote the formula obtained from φ by replacing all occurrences of x_i by a_i for every i, $1 \leq i \leq n$. A derivation for φ in the tableaux method of Lutz et al. (1999) is a finitely branching tree T with root $\{\varphi\{\overline{x}/\overline{a}\}\}$. In the following we write X, φ instead of $X \cup \{\varphi\}$. The tree is expanded by adding one or two successor nodes, consisting of sets of formulae, to one of the clash-free leaf nodes of the tree according to the tableaux rules described below. A leaf node contains a *clash* iff it contains the formula $\perp$, otherwise it is *clash-free*. A leaf node is *complete* iff no successor nodes can be added to it by one of the tableaux rules. The derivation terminates if either all leaf nodes contain a clash or there is a complete leaf node.

Derivation of falsum: $$\frac{X, \phi, \neg\phi}{X, \phi, \neg\phi, \bot}$$

Conjunction: $$\frac{X, \phi \wedge \psi}{X, \phi \wedge \psi, \phi, \psi}$$

provided that $\{\phi, \psi\} \not\subseteq X$.

Disjunction: $$\frac{X, \phi \vee \psi}{X, \phi \vee \psi, \phi \mid X, \phi \vee \psi, \psi}$$

provided that $\{\phi, \psi\} \cap X = \emptyset$.

Existential quantification: $$\frac{X, \exists\overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \wedge \phi(\overline{y}))}{X, \exists\overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \wedge \phi(\overline{y})), \mathcal{G}(\overline{a}, \overline{b}), \phi(\overline{y}/\overline{b})}$$

provided $\overline{b}$ is a sequence of fresh constants and there are no constants $\overline{c}$ such that $\{\mathcal{G}(\overline{a}, \overline{c}), \phi(\overline{y}/\overline{c})\} \subseteq X$.

Universal quantification: $$\frac{X, \forall\overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \rightarrow \phi(\overline{y})), \mathcal{G}(\overline{a}, \overline{b})}{X, \forall\overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \rightarrow \phi(\overline{y})), \mathcal{G}(\overline{a}, \overline{b}), \phi(\overline{y}/\overline{b})}$$

provided that $\phi(\overline{y}/\overline{b}) \notin X$.

Theorem 7.1 (Lutz et al., 1999). *A formula φ in* GF1$^-$ *is satisfiable iff the rules can be used to construct a tableaux which contains a branch $\mathcal{B}$ such that the endpoint of $\mathcal{B}$ is a complete and clash-free set of formulae.*

Theorem 7.2 (Lutz et al., 1999). *For a signature of bounded arity the tableaux algorithm can be implemented to run in polynomial space.*

Before proving the simulation results formally, we illustrate the idea by an example showing the tableaux and resolution derivations for the GF1$^-$ formula

$$\varphi = \forall x(r(x, y, z) \rightarrow p(x)) \wedge \exists x(r(x, y, z) \wedge \neg p(x)).$$

Tableaux derivation for φ:

$$\begin{aligned}
X_1 &= \{\forall x(r(x, a, b) \rightarrow p(x)) \wedge \exists x(r(x, a, b) \wedge \neg p(x))\} \\
X_2 &= X_1 \cup \{\forall x(r(x, a, b) \rightarrow p(x)), \exists x(r(x, a, b) \wedge \neg p(x))\} \\
X_3 &= X_2 \cup \{r(c, a, b), \neg p(c)\} \\
X_4 &= X_3 \cup \{p(c)\} \\
X_5 &= X_4 \cup \{\bot\}.
\end{aligned}$$

The endpoint of the branch contains a clash. Since no alternative tableau can be constructed for X_1, the original formula is unsatisfiable.

The corresponding resolution derivation starts from the clausal set N, obtained from φ after a renaming of each non-atomic subformula with the exception of implications and conjunctions immediately below quantifiers.

$$\begin{aligned}\mathrm{Def}_\Lambda(\varphi) = {} & Q_\wedge(y,z)\wedge\\ & \forall y,z(Q_\wedge(y,z) \rightarrow (Q_\forall(y,z) \wedge Q_\exists(y,z)))\wedge\\ & \forall y,z(Q_\forall(y,z) \rightarrow \forall x(r(x,y,z) \rightarrow p(x)))\wedge\\ & \forall y,z(Q_\exists(y,z) \rightarrow \exists x(r(x,y,z) \wedge \neg p(x)))\\ N = \{ & Q_\wedge(a,b),\\ & \neg Q_\wedge(y,z) \vee Q_\forall(y,z),\\ & \neg Q_\wedge(y,z) \vee Q_\exists(y,z),\\ & \neg Q_\forall(y,z) \vee \neg r(x,y,z) \vee p(x),\\ & \neg Q_\exists(y,z) \vee r(f(y,z),y,z),\\ & \neg Q_\exists(y,z) \vee \neg p(f(y,z))\}.\end{aligned}$$

Resolution derivation for φ:

$$\begin{aligned}N_1 &= N\\ N_2 &= N_1 \cup \{Q_\forall(a,b)\}\\ N_3 &= N_2 \cup \{Q_\exists(a,b)\}\\ N_4 &= N_3 \cup \{r(f(a,b),a,b)\}\\ N_5 &= N_4 \cup \{p(f(a,b))\}\\ N_6 &= N_5 \cup \{\bot\}.\end{aligned}$$

The clause set N_6 contains the empty clause. Since the branch on which N_6 occurs is the only one in our derivation, the formula φ is unsatisfiable.

The correspondence between the tableaux derivation and the resolution derivation is straightforward. Let g be the function that maps the constant c in the tableaux derivation to the term $f(a,b)$ in the resolution derivation. All other terms in the tableaux derivation are mapped to themselves. Furthermore, g maps subformulae of φ to predicate symbols in N such that $g(P(\overline{x})) = P$ if $P(\overline{x})$ is atomic, and $g(\psi) = Q_\psi$ otherwise. Then

$$\begin{aligned}&Q_\wedge(a,b) = g(\varphi)(x,y)\{x/a, y/b\}\\ &Q_\exists(a,b) = g(\exists x(r(x,y,z) \wedge \neg p(x)))(y,z)\{y/a, z/b\}\\ &Q_\forall(a,b) = g(\forall x(r(x,y,z) \rightarrow p(x)))(y,z)\{y/a, z/b\}\\ &r(f(a,b),a,b) = g(r(x,y,z))(x,y,z)\{x/g(c), y/a, z/b\}\\ &p(f(a,b)) = g(p(x))(x)\{x/g(c)\}.\end{aligned}$$

For every formula ϑ in the tableaux derivation there is a ground unit clause C generated in the $\mathsf{R}^{\mathsf{hyp}}$ derivation such that $g(\vartheta)(\overline{x})\delta = C$, where $\overline{x}$ are the free variables of C and δ is a suitable substitution.

Extending the simulation results of de Nivelle et al. (2000) and Hustadt and Schmidt (2000a) we prove that:

Theorem 7.3. *There is a polynomial simulation of the tableaux system of Lutz et al. (1999) for* GF1$^-$ *by* $\mathsf{R}^{\mathsf{hyp}}$.

Proof. We show that $\mathsf{R}^{\mathsf{hyp}}$ simulates the tableaux derivation stepwise. Let φ be a formula in GF1$^-$ and let $(X_1, \ldots, X_n)$ be a branch in the tableaux derivation starting from φ.

Then there exists a branch $(N_1, \ldots, N_k)$ in the $\mathsf{R}^{\mathsf{hyp}}$ derivation for some $2n \geq k \geq n$ starting from N and a function g such that for every formula $\vartheta(\overline{x})\gamma$ in X_n where $\vartheta(\overline{x})$ is a subformula of φ and γ is a substitution which maps the free variables $\overline{x}$ of ϑ to constants there exists a ground unit clause $g(\vartheta)(\overline{x})\delta$ in N_k where $\delta(x_i) = g(\gamma(x_i))$ for every x_i in $\overline{x}$, with $g(\vartheta(\overline{x})\gamma) = g(\vartheta)(\overline{x})\delta$.

The proof is by induction on n, which stands for the length of the branch in the tableaux derivation.

Base case. If $n = 1$ then the tableaux consists of the single node $X_1 = \{\varphi\gamma\}$ with $\gamma = \{\overline{x}/\overline{a}\}$. We assume without loss of generality that in the clausal form transformation, we have used the same constant symbols $\overline{a}$ to instantiate the free variables in $\mathrm{Def}_\Lambda(\varphi)$. Thus, the function g maps these constant symbols to themselves. The clause set $N = \mathrm{Cls}(\mathrm{Def}_\Lambda(\varphi))$ contains one ground clause, namely $Q_\varphi(\overline{a})$. We let $k = 1$ and $N_1 = N$, and g maps φ to Q_φ.

Inductive step. Suppose that the result holds for a derivation of length n, that is, if $(X_1, \ldots, X_n)$ is a branch in the tableaux derivation from φ, then there exists a branch $(N_1, \ldots, N_k)$ in the $\mathsf{R}^{\mathsf{hyp}}$ derivation for some $2n \geq k \geq n$ from N and a function g such that for every formula $\vartheta(\overline{x})\gamma$ in X_n, where $\vartheta(\overline{x})$ is a subformula of φ and γ is a substitution which maps the free variables $\overline{x}$ of ϑ to constants there exists a ground unit clause $g(\vartheta)(\overline{x})\delta$ in N_k, where $\delta(x_i) = g(\gamma(x_i))$ for every x_i in $\overline{x}$, with $g(\vartheta(\overline{x})\gamma) = g(\vartheta)(\overline{x})\delta$.

We show that the claim holds also for derivations of length $n + 1$. The proof is by case analysis of the tableaux rule applied to the endpoint X_n of the branch.

Case 1. Suppose the conjunction rule is applied to the formula $\vartheta(\overline{z})\gamma = \phi\gamma_1 \wedge \psi\gamma_2$ in X_n where γ_1 and γ_2 map the free variables of ϕ and ψ to constants. The branch is extended by the successor node $X_{n+1} = X_n \cup \{\phi\gamma_1, \psi\gamma_2\}$. By the inductive hypothesis there is a branch in the $\mathsf{R}^{\mathsf{hyp}}$ derivation with endpoint N_k and a function g such that there exists a ground unit clause $g(\vartheta)(\overline{z})\delta$ with $g(\vartheta)(\overline{z})\gamma = g(\vartheta)(\overline{z})\delta$, where $\delta(z_i) = g(\gamma(z_i))$ for every z_i in $\overline{z}$. Since ϑ is a non-atomic formula, N_k also contains the definitional clauses $\neg Q_\vartheta(\overline{z}) \vee Q_\phi(\overline{x})$ and $\neg Q_\vartheta(\overline{z}) \vee Q_\psi(\overline{y})$. Then the conjunction rule is simulated by two hyperresolution steps between the ground clause $Q_\vartheta(\overline{z})\delta$ and these two clauses, producing the ground resolvents $Q_\phi(\overline{x})\delta$ and $Q_\psi(\overline{y})\delta$. Next, extend g so that it maps ϕ to Q_ϕ and ψ to Q_ψ.

Case 2. Suppose the disjunction rule is applied to the formula $\vartheta(\overline{z})\gamma = \phi\gamma_1 \vee \psi\gamma_2$. An application of the disjunction rule leads to two successor nodes one of which is chosen to extend the branch under consideration. Without loss of generality, let $X_{n+1} = X_n \cup \{\phi\gamma_1\}$. This case is analogous to the previous one. The disjunction rule is simulated by one hyperresolution step followed by an application of the splitting rule. First, we derive the ground clause $(Q_\phi(\overline{x}) \vee Q_\psi(\overline{y}))\delta$. Second, we replace this clause by $Q_\phi(\overline{x})\delta$ using the splitting rule. We extend g to map ϕ to Q_ϕ.

Case 3. Suppose the existential quantification rule is applied to the formula $\vartheta(\overline{x})\gamma = \exists\overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \wedge \phi(\overline{y}))\gamma$ with $\overline{x} = x_1, \ldots, x_m$ and $\overline{y} = y_1, \ldots, y_n$. Let $\overline{b}$ be a sequence of n fresh constants and let $\gamma'(x_i) = \gamma(x_i)$ for every i, $1 \leq i \leq m$, and $\gamma'(y_j) = b_j$ for every j, $1 \leq j \leq n$. Then X_{n+1} is equal to $X_n \cup \{\phi(\overline{y})\gamma', \mathcal{G}(\overline{x}, \overline{y})\gamma'\}$. By the inductive hypothesis there is a ground clause $g(\vartheta)(\overline{x})\delta$ in our clause set corresponding to $\vartheta(\overline{x})\gamma$. The clause set contains the definitional clauses $\neg Q_\vartheta(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})})$ and $\neg Q_\vartheta(\overline{x}) \vee Q_\phi(\overline{f(\overline{x})})$. With two hyperresolution inference steps we derive ground clauses $\mathcal{G}(\overline{x}, \overline{f(\overline{x})})\delta$ and $Q_\phi(\overline{f(\overline{x})})\delta$. We extend g to map $\mathcal{G}$ to itself and ϕ to Q_ϕ. Note that the b_i are fresh constants, that is, $g(b_i)$ is not yet defined. However, for each constant b_i there is a corresponding Skolem term $f_i(\overline{x})\delta$ in both clauses we have derived. So, we define for every i, $1 \leq i \leq n$, $g(b_i) = f_i(\overline{x})\delta$. It is straightforward to see that this definition yields the desired effect.

Case 4. Suppose the universal quantification rule is applied to the two formulae $\vartheta(\overline{x})\gamma_1 = \forall\overline{y}(\mathcal{G}(\overline{x}, \overline{y}) \rightarrow \phi(\overline{y}))\gamma_1$ and $\mathcal{G}(\overline{x}, \overline{y})\gamma_2$ where $\gamma_1(x_i) = \gamma_2(x_i)$ for every x_i in $\overline{x}$. Then $X_{n+1} = X_n \cup \{\phi(\overline{y})\gamma_2\}$. Again, by the inductive hypothesis, we have ground clauses $g(\vartheta)(\overline{y})\delta_1$ and $g(\mathcal{G})(\overline{x}, \overline{y})\delta_2$ in N_k. By the construction of g, we have $g(\vartheta) = Q_\vartheta$ and $g(\mathcal{G}) = \mathcal{G}$. The clause set N_k also contains the clause $\neg Q_\varphi(\overline{x}) \vee \neg\mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{y})$. With a single hyperresolution inference step we derive $Q_\phi(\overline{y})\delta_2$. We extend g to map ϕ to Q_ϕ.

Case 5. An application of the derivation of the 'falsum rule' to X_n containing formulae $\phi(\overline{x})\gamma$ and $\neg\phi(\overline{y})\gamma'$ leads to $X_{n+1} = X_n \cup \{\perp\}$. Note that $\phi(\overline{x})\gamma = \phi(\overline{y})\gamma'$. By the inductive hypothesis we have already derived ground clauses $g(\phi)(\overline{x})\delta$ and $g(\neg\phi)(\overline{y})\delta'$ corresponding to $\phi(\overline{x})\gamma$ and $\neg\phi(\overline{y})\gamma'$, respectively. Note that $\overline{x}\delta = \overline{y}\delta'$. Since φ is in negation normal form, $\phi(\overline{x})$ has to be an atomic formula $P(\overline{x})$. Thus, $g(\phi)(\overline{x})\delta = P(\overline{x})\delta$. The clause set under consideration contains a definitional clause $\neg Q_{\neg\phi}(\overline{y}) \vee \neg P(\overline{y})$. A single hyperresolution step with this clause, $g(\neg\phi)(\overline{y})\delta' = Q_{\neg\phi}(\overline{y})\delta'$ and $g(\phi)(\overline{x})\delta = P(\overline{x})\delta$ leads to the derivation of the empty clause.

Thus we have proved that each application of a tableaux rule can be simulated by one or two inference steps of $\mathsf{R}^{\mathsf{hyp}}$. Therefore, every tableaux derivation for GF1$^-$ can be polynomially mapped to a derivation by $\mathsf{R}^{\mathsf{hyp}}$. □

Similarly, the $\mathsf{R}^{\mathsf{hyp}}$ rules can be identified and reformulated as tableaux rules, using the inverse of the mapping g, cf. Fig. 2.

Theorem 7.4. *There is a polynomial simulation of* $\mathsf{R}^{\mathsf{hyp}}$ *for* GF1$^-$ *by a moderate extension of the tableaux system of* Lutz et al. *(1999).*

Proof. It is necessary to add a simplification rule to the tableaux calculus which simulates positive factoring. □

8. Generalization

From the analysis in the previous sections, particularly the investigation of the behaviour of $\mathsf{R}^{\mathsf{hyp}}$ on GF1$^-$ clauses in Section 4, it is not difficult to observe that the results can be strengthened to cover a larger class than GF1$^-$, as long as the inferred clauses have the same syntactic structure as before, i.e. are uni-nodes and bi-nodes, and the grouping restriction is preserved. In this section we mention some ways of extending GF1$^-$ and its corresponding clausal class without losing the termination property of hyperresolution.

According to the definition of GF1$^-$, the quantified variables in the formulae must be exactly the free variables of non-guard formulae. Hyperresolution is a decision procedure

Resolution inferences $\rightsquigarrow$ Tableaux inferences

$$\frac{\begin{array}{l} Q_P(\overline{s}) \\ P(\overline{s}) \\ \neg Q_P(\overline{x}) \vee \neg P(\overline{x}) \end{array}}{\bot} \quad \rightsquigarrow \quad \frac{P(\overline{a}),\ \neg P(\overline{a})}{\bot}$$

$$\begin{array}{l} \dfrac{\begin{array}{l} Q_\varphi(\overline{s}) \\ \neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) \\ \neg Q_\varphi(\overline{z}) \vee Q_\psi(\overline{y}) \end{array}}{Q_\varphi(\overline{t}),\quad Q_\varphi(\overline{u})} \\ \text{where } \overline{s} = \overline{t} \cup \overline{u} \end{array} \quad \rightsquigarrow \quad \frac{\phi(\bar{b}) \wedge \psi(\bar{c})}{\phi(\bar{b}),\ \psi(\bar{c})}$$

$$\begin{array}{l} \dfrac{\begin{array}{l} Q_\varphi(\overline{s}) \\ \neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{x}) \vee Q_\psi(\overline{y}) \end{array}}{Q_\phi(\overline{t}) \mid Q_\psi(\overline{u})} \\ \text{where } \overline{s} = \overline{t} \cup \overline{u} \end{array} \quad \rightsquigarrow \quad \frac{\phi(\bar{b}) \vee \psi(\bar{c})}{\phi(\bar{b}) \mid \psi(\bar{c})}$$

$$\frac{\begin{array}{l} Q_\varphi(\overline{s}) \\ \neg Q_\varphi(\overline{z}) \vee \mathcal{G}(\overline{z}, \overline{f(\overline{z})}) \\ \neg Q_\varphi(\overline{z}) \vee Q_\phi(\overline{f(\overline{z})}) \end{array}}{\mathcal{G}(\overline{s}, \overline{f(\overline{s})}),\quad Q_\phi(\overline{f(\overline{s})})} \quad \rightsquigarrow \quad \frac{\exists \overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \wedge \phi(\overline{y}))}{\mathcal{G}(\overline{a}, \overline{b}),\ \phi(\overline{b})}$$

$$\frac{\begin{array}{l} Q_\varphi(\overline{s}) \\ \mathcal{G}(\overline{s}, \overline{t}) \\ \neg Q_\varphi(\overline{z}) \vee \neg \mathcal{G}(\overline{z}, \overline{y}) \vee Q_\phi(\overline{y}) \end{array}}{Q_\phi(\overline{t})} \quad \rightsquigarrow \quad \frac{\mathcal{G}(\overline{a}, \overline{b}),\ \forall \overline{y}(\mathcal{G}(\overline{a}, \overline{y}) \rightarrow \phi(\overline{z}))}{\phi(\overline{b})}$$

Fig. 2. Simulation of hyperresolution by tableaux.

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.

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

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

$$\begin{array}{lll} \neg Q_\forall(\overline{x}) \vee \neg \mathcal{G}(\overline{x}, \overline{y}) \vee Q_\phi(\overline{z}) & & \text{where } \overline{z} \subseteq \overline{y} \\ \neg Q_\exists(\overline{x}) \vee \mathcal{G}(\overline{x}, \overline{f(\overline{x})}) & \neg Q_\exists(\overline{x}) \vee Q_\phi(\overline{t}) & \text{where } \overline{t} \subseteq \overline{f(\overline{x})}. \end{array}$$

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).

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

where $\overline{x}_i \subseteq \overline{x}, \overline{y}_i \subseteq \overline{y}$ $(1 \leq i \leq n)$, $\overline{x} \cap \overline{y} = \emptyset$ and $\overline{z} \subseteq \overline{y}$. An essential condition is that each of the atoms $\mathcal{G}_i(\ldots)$, where $0 \leq i \leq n$, satisfy the grouping restriction (as suggested by the notation) and the clause includes at least one guard $\neg\mathcal{G}_0(\overline{x}, \overline{y})$. This ensures that the conclusion is a ground clause. On the first-order level, this means we can allow formulae of the form:

$$\forall\, \overline{y}((\mathcal{G}_0(\overline{x}, \overline{y}) \wedge (\neg)\mathcal{G}_1(\overline{x}_1, \overline{y}_1) \wedge \cdots \wedge (\neg)\mathcal{G}_n(\overline{x}_n, \overline{y}_n)) \rightarrow \phi(\overline{z})),$$

where $\overline{x}_i \subseteq \overline{x}, \overline{y}_i \subseteq \overline{y}$ $(1 \leq i \leq n)$, $\overline{x} \cap \overline{y} = \emptyset$ and $\overline{z} \subseteq \overline{y}$. Note that due to the restrictions of the positions of quantified and free variables in GF1$^-$ the equivalent formula does not belong to the fragment, although it is a guarded formula.

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

Disjunctions in the guard expression are permitted provided none of the atoms are negated:

$$\forall\, \overline{y}((\mathcal{G}_1(\overline{x}_1, \overline{y}_1) \vee \cdots \vee \mathcal{G}_n(\overline{x}_n, \overline{y}_n)) \rightarrow \phi(\overline{z})),$$

where $\overline{x} = \overline{x}_1 \cup \cdots \cup \overline{x}_n$, $\overline{y} = \overline{y}_1 \cup \cdots \cup \overline{y}_n$, $\overline{x} \cap \overline{y} = \emptyset$, and $\overline{z} \subseteq \overline{y}_1 \cap \cdots \cap \overline{y}_n$. The corresponding clause set includes clauses of the following form:

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

Such formulae fall outside the GF and the loosely GF.

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 $\mathcal{G}_i(\cdots)$ are required to satisfy the grouping restriction. Clausification produces clauses of the form:

$$\neg Q_\exists(\overline{x}) \vee (\neg)\mathcal{G}_{i_1}(\overline{x}_{i_1}, \overline{f_{i_1}(\overline{x})}) \vee \cdots \vee (\neg)\mathcal{G}_{i_m}(\overline{x}_{i_m}, \overline{f_{i_m}(\overline{x})})$$
$$\neg Q_\exists(\overline{x}) \vee Q_\phi(\overline{g(\overline{x})}),$$

where $1 \leq i_j \leq n$ for each $1 \leq j \leq m$, and $\overline{x} = \overline{x}_1 \cup \cdots \cup \overline{x}_n$.

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

Theorem 8.1. *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.*
2. *If T is a fair derivation from N then* (i) *If $N(= N_0), N_1, \ldots$ is a path with limit N_∞, N_∞ is saturated up to redundancy.* (ii) *φ is satisfiable if and only if there exists a path in T with limit N_∞ such that N_∞ is satisfiable.* (iii) *φ is unsatisfiable if and only if for every path $N(= N_0), N_1, \ldots$ the clause set $\bigcup_j N_j$ contains the empty clause.*

Proof. Termination follows from Theorem 4.3, since all derived clauses from the formulae in the extensions of GF1$^-$ by hyperresolution with splitting are either uni-nodes or bi-nodes. □

Theorem 8.2. *Let φ be a satisfiable formula in the above extension. A finite model for φ can be constructed on the basis of* $\mathsf{R}^{\mathsf{hyp}}$.

Similarly, as in the previous section (and de Nivelle et al., 2000), 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 8.3. *Hyperresolution and factoring 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 Robinson (1965) if factoring includes positive and negative factoring. Otherwise, soundness and completeness follows from Bachmair and Ganzinger (1994, 2001). Termination follows from Theorems 4.3 and 8.1, since all derived clauses by hyperresolution without splitting are formed from uni-node and bi-node literals appearing in the corresponding $\mathsf{R}^{\mathsf{hyp}}$ derivation tree. □

9. Related work

Related Calculi. Apart from the semantic tableaux calculus of Lutz et al. (1999), whose relationship to $\mathsf{R}^{\mathsf{hyp}}$ was considered in Section 7, there are other inference calculi closely related to $\mathsf{R}^{\mathsf{hyp}}$. These include resolution with maximal selection of negative literals, hypertableaux and its descendants. These connections are useful since, not only do they present new perspectives, they also allow the interchange of search pruning mechanisms between the different inference systems, and, more practically, make available a larger array of provers for automating reasoning about problems formulated in GF1$^-$.

We already mentioned resolution with maximal selection of negative literals (Bachmair and Ganzinger, 2001) which has been used in a translation-based approach to modal logic and description logic reasoning (de Nivelle et al., 2000; Hustadt and Schmidt, 1999, 2000a,b). Resolution with maximal selection of negative literals can be viewed as hyperresolution with positive factoring (Bachmair and Ganzinger, 2001), and thus amounts to the same as $\mathsf{R}^{\mathsf{hyp}}$ (with or without splitting).

Hypertableaux was introduced by Baumgartner et al. (1996). 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) $C = \neg B_1 \vee \cdots \vee \neg B_n \vee A_1 \vee \cdots A_m$ 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) $B_1\sigma \wedge \cdots \wedge B_n\sigma$, 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 leaves 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 always the identity substitution. 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. Therefore, hypertableaux also provides a decision procedure and model building procedure for $\mathrm{GF1}^-$ and the considered extension. (So do the descendants of hypertableaux (Baumgartner, 1998, 2000) for that matter.) For practical considerations the link between $\mathsf{R}^{\mathsf{hyp}}$ and hypertableaux allows us to transfer several improvements of hypertableaux discussed in Baumgartner et al. (1996). These include factorization and level cut. Factorization has the effect that different branches 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 (Weidenbach, 2001) or backjumping used in tableaux methods (Hustadt and Schmidt, 1998). (On the side we remark that hyperresolution with splitting avoids the 'memory management' problem of hyperresolution highlighted in Baumgartner et al., 1996.)

A resolution based decision procedure for the full GF without equality is presented by de Nivelle and de Rijke (2003). Their method uses ordered resolution with a non-liftable ordering that is incomplete in general, but complete for the GF. To deal with the loosely GF without equality a combination of this method with a non-trivial modification of hyperresolution is used. Ganzinger and de Nivelle (1999) describe a decision procedure for the guarded and loosely GF with equality based on ordered paramodulation with selection.

Related Clausal Classes. We have already referred to the related clausal classes associated with modal and description logics. A related class is the encoding in clausal form of the extended multi-modal logic $K_{(m)}(\cap, \cup, \smile)$ (de Nivelle et al., 2000; Hustadt and Schmidt, 2000b) and the corresponding description logic $\mathcal{ALB}_D$ (Hustadt and Schmidt, 2000a). This class is subsumed by the clausal class of Section 3.

Other clausal classes decidable by hyperresolution are investigated in Fermüller et al. (2001) and Leitsch (1993) and include the classes $\mathcal{PVD}$ and $\mathcal{KPOD}$. A set of clauses N belong to $\mathcal{PVD}$ (positive variable dominated) if for every clause C in N, the following conditions hold: (i) The variables in the positive part of C are a subset of the variables of the negative part of C. (ii) The maximal term-depth of each variable in the positive part of C is smaller or equal to the maximal term-depth of the same variable in the positive part of C.

A set of clauses N belongs to $\mathcal{KPOD}$ (Krom positive occurrence dominated) if: (i) All clauses C in N are Krom, i.e. $|C| \leq 2$. (ii) For every variable x contained in the positive part of a clause C, the number of occurrences of x in the positive part of C is smaller than the number of occurrences of x in the negative part of C.

Obviously, the sets of clauses we obtain from $\mathrm{GF1}^-$ formulae in general do not satisfy condition (ii) of the definition of $\mathcal{PVD}$ nor do they satisfy condition (i) of the definition of $\mathcal{KPOD}$. For $\mathcal{PVD}$ the syntactic restrictions on the class imply that during a derivation by hyperresolution the depth of the conclusions does not increase (Fermüller et al., 2001; Leitsch, 1993). This is unlike the case for GF1$^-$. For $\mathcal{KPOD}$ the term depth of conclusions may increase during a derivation (Leitsch, 1993). 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 GF1$^-$ formulae.

Termination for $\mathcal{PVD}$ and $\mathcal{KPOD}$ 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 $\{A\sigma \mid \sigma \in \sigma_0, \mu(A\sigma) \leq k\}$ 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 formalized in this framework.

10. Conclusion and further work

The presented work is a continuation of ideas and techniques developed in Hustadt and Schmidt (1999, 2000a,b) for extended propositional modal logics, making use of concepts introduced in Lutz et al. (1999). We have considered the use of hyperresolution as a decision procedure for guarded formulae in GF1$^-$ as well as extensions of this fragment. We have also considered the use of hyperresolution for automatically building models and analysed the close relationship to tableaux approaches. The latter can be exploited to extract a tableaux system for the extension of GF1$^-$ discussed in Section 8. An advantage of using hyperresolution is in the availability of a number of theorem provers which can be used without adaptation as decision procedures for GF1$^-$ and the considered extension (for example, FDPLL, OTTER, PROTEIN, SPASS, and Vampire).

Currently we are looking into defining an abstract atom complexity measure in analogy to Leitsch (1993) which would generalize the specific complexity measures and orderings used in the termination proofs presented in this paper and in de Nivelle et al. (2000) and Hustadt and Schmidt (1999, 2000a,b). We are also attempting to define a larger solvable class which would accommodate more formulae outside the GF and the loosely GF. Further, it would be of interest whether it is possible to extend the approach to the entire GF, possibly by using blocking conditions in the spirit of Ganzinger et al. (1997).

Acknowledgements

We thank the reviewers for helpful comments. The work was partially supported by research grants GR/M36700 and GR/R92035 from the UK Engineering and Physical Sciences Research Council.

References

Andréka, H., Németi, I., van Benthem, J., 1998. Modal languages and bounded fragments of predicate logic. J. Philos. Logic 27 (3), 217–274.

Andréka, H., van Benthem, J., Németi, I., 1995. Back and forth between modal logic and classical logic. Bull. IGPL 3 (5), 685–720.
Baaz, M., Fermüller, C., Leitsch, A., 1994. A non-elementary speed-up in proof length by structural clause form transformation. In: Proc. LICS' 94, IEEE Computer Society Press, pp. 213–219.
Bachmair, L., Ganzinger, H., 1994. Rewrite-based equational theorem proving with selection and simplification. J. Logic Comput. 4 (3), 217–247.
Bachmair, L., Ganzinger, H., 2001. Resolution theorem proving. In: Robinson, A., Voronkov, A. (Eds.), Handbook of Automated Reasoning, vol. I. Elsevier, pp. 19–99 (Chapter 2).
Bachmair, L., Ganzinger, H., Waldmann, U., 1993. Superposition with simplification as a decision procedure for the monadic class with equality. In: Proc. KGC' 93, LNCS, vol. 713. Springer, pp. 83–96.
Baumgartner, P., 1998. Hyper tableau: the next generation. In: Proc. TABLEAUX' 98, LNAI, vol. 1397. Springer, pp. 60–76.
Baumgartner, P., 2000. FDPLL: A first-order Davis-Putnam-Logeman-Loveland procedure. In: Automated Deduction—CADE-17, LNAI, vol. 1831. Springer, pp. 200–219.
Baumgartner, P., Furbach, U., Niemelä, I., 1996. Hyper tableaux. In: Proc. JELIA' 96, LNAI, vol. 1126. Springer, pp. 1–17.
de Nivelle, H., de Rijke, M., 2003. Deciding the guarded fragments by resolution. J. Symbolic Comput. 35 (1), 21–58.
de Nivelle, H., Schmidt, R.A., Hustadt, U., 2000. Resolution-based methods for modal logics. Logic J. IGPL 8 (3), 265–292.
Donini, F.M., Lenzerini, M., Nardi, D., Schaerf, A., 1996. Reasoning in description logics. In: Brewka, G. (Ed.), Principles in Knowledge Representation. CSLI Publications, Stanford, pp. 191–236.
Fermüller, C., Leitsch, A., Hustadt, U., Tammet, T., 2001. Resolution decision procedures. In: Robinson, A., Voronkov, A. (Eds.), Handbook of Automated Reasoning. vol. II, Elsevier, pp. 1791–1849 (Chapter 25).
Ganzinger, H., de Nivelle, H., 1999. A superposition decision procedure for the guarded fragment with equality. In: Proc. LICS' 99. IEEE Computer Society Press, pp. 295–303.
Ganzinger, H., Meyer, C., Veanes, M., 1999. The two-variable guarded fragment with transitive relations. In: Proc. LICS' 99. IEEE Computer Society, pp. 24–34.
Ganzinger, H., Meyer, C., Weidenbach, C., 1997. Soft typing for ordered resolution. In: Automated Deduction—CADE-17, LNAI, vol. 1249. Springer, pp. 321–335.
Georgieva, L., Hustadt, U., Schmidt, R.A., 2001. Computational space efficiency and minimal model generation for guarded formulae. In: Proc. LPAR' 2001, LNAI, vol. 2250. Springer, pp. 85–99.
Goré, R., 1999. Tableau methods for modal and temporal logics. In: D' Agostino, M., Gabbay, D., Hähnle, R., Posegga, J. (Eds.), Handbook of Tableau Methods. Kluwer, pp. 297–396.
Grädel, E., 1999a. Decision procedures for guarded logics. In: Automated Deduction—CADE-16, LNAI, vol. 1632. Springer, pp. 31–51.
Grädel, E., 1999b. On the restraining power of guards. J. Symbolic Logic 64, 1719–1742.
Hustadt, U., Schmidt, R.A., 1998. Simplification and backjumping in modal tableau. In: Proc. TABLEAUX' 98, LNAI, vol. 1397. Springer, pp. 187–201.
Hustadt, U., Schmidt, R.A., 1999. On the relation of resolution and tableaux proof systems for description logics. In: Proc. IJCAI' 99, Morgan Kaufmann, pp. 110–115.
Hustadt, U., Schmidt, R.A., 2000a. Issues of decidability for description logics in the framework of resolution. In: Automated Deduction in Classical and Non-Classical Logics, LNAI, vol. 1761. Springer, pp. 191–205.
Hustadt, U., Schmidt, R.A., 2000b. Using resolution for testing modal satisfiability and building models. In: Gent, I.P., van Maaren, H., Walsh, T. (Eds.), SAT 2000: Highlights of Satisfiability Research in the Year 2000, IOS Press, Amsterdam.

Leitsch, A., 1993. Deciding clause classes by semantic clash resolution. Fund. Inform. 18, 163–182.

Lutz, C., Sattler, U., Tobies, S., 1999. A suggestion of an *n*-ary description logic. In: Proc. DL' 99, CEUR Electronic Workshop Proceedings, vol. 22. Linköping University, pp. 81–85.

Plaisted, D.A., Greenbaum, S., 1986. A structure-preserving clause form translation. J. Symbolic Comput. 2, 293–304.

Robinson, J.A., 1965. Automatic deduction with hyper-resolution. Int. J. Comput. Math. 1 (3), 227–234.

Weidenbach, C., 2001. Combining superposition, sorts and splitting. In: Robinson, A., Voronkov, A. (Eds.), Handbook of Automated Reasoning, vol. II, Elsevier, pp. 1965–2013 (Chapter 27).