Advances in Modal Logic, Volume 3

F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyaschev, eds

CENTER FOR THE STUDY
OF LANGUAGE
AND INFORMATION

Contents

1

Modal logics with weak forms of recursion: PSPACE specimens

S. Demri

ABSTRACT. We analyze the computational complexity of extensions of the multimodal version of the standard modal logic K by finite addition of axiom schemes that can be read as the production rules of a formal grammar. By using proof-theoretical means, we show that every right linear grammar logic has a satisfiability problem in deterministic exponential time and we exhibit countably infinite classes of right linear grammar logics that contain weak forms of recursion for which the satisfiability problem can be solved in polynomial space.

1 Introduction

In order to explain the algorithmic properties of many modal logics, a possible approach consists in studying very expressive decidable logical theories in which can be easily embedded the modal logics. The modal μ-calculus (see e.g. (Kozen 1983)), the guarded fixed point logic μLGF (Grädel and Walukiewicz 1999) and the monadic second-order theory of two successors (Rabin 1969) are good representatives of such theories. Weaker logical theories such as the Propositional Dynamic Logic (see e.g. (Pratt 1979, Fischer and Ladner 1979)) and the guarded fragment (Andreka et al. 1998) are also serious candidates although strictly less expressive, respectively. Moreover, since both the μ-calculus and the guarded fixed point logic μLGF with bounded arity have an **EXPTIME**-complete satisfiability problem (Grädel and Walukiewicz 1999), those logical formalisms inherit the algorithmic properties of weak extensions of the modal logic K (by adding the universal modal connec-

Advances in Modal Logic, Volume 3
F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyaschev, eds.

tive for example) that are already of the same complexity level. By contrast, many standard modal logics are known to be in **PSPACE** (Ladner 1977, Halpern and Moses 1992) and this algorithmic property is not reflected by the analysis of most identified fragments of second order logic. As far as we know, a rare exception can be found in (Marx 1997) (see also (Lutz et al. 1999, Marx et al. 2000b)), where a **PSPACE** first-order fragment is defined capturing multimodal logics with K, T and B modal connectives plus inclusion, to quote a few examples of logics. This situation is all the more surprising since many modal logics with weak forms of recursion (K4, S4, ...) are also in **PSPACE**. So, what is the adequate first-order/second-order fragment that is responsible for the algorithmic behavior of **PSPACE** modal logics? This question could become meaningful for optimizing the efficiency of mechanical reasoning for such logics and thus avoiding a blind translation into a rich decidable logical theory, possibly algorithmically more expensive in the worst-case. Although we have no answer for this question, this motivates the developments made in this paper. We study a countably infinite class of multimodal logics that can be embedded uniformly into first-order logic with the relational translation. The target fragment is not known to belong to identified decidable fragments but we exhibit modal logics with weak forms of recursion that are in **PSPACE**.

In order to understand the **PSPACE** modal logics with weak forms of recursion, a class of multimodal logics that are worth investigating are the grammar logics defined in (Fariñas del Cerro and Penttonen 1988) that are closely related to formal grammars. With each production rule $i_1 \cdot \ldots \cdot i_n \rightarrow j_1 \cdot \ldots \cdot j_{n'}$ in the grammar is associated a reduction principle $[i_1] \ldots [i_n]\mathrm{p} \Rightarrow [j_1] \ldots [j_{n'}]\mathrm{p}$ (see e.g. (van Benthem 1976)) which is a particular form of Sahlqvist formula (Sahlqvist 1975). Observe that the logical view on grammar presented in (Kracht 2000) differs from the approach in (Fariñas del Cerro and Penttonen 1988). In the present paper, we mainly study the extensions of the multimodal logic K_m with m independent K modal connectives by finite addition of axiom schemes of the above form such that the associated finite set of production rules forms a right linear formal grammar. The right linear grammar logics contain a weak form of recursion although different from the one in the logics defined in (Halpern and Reif 1983) (e.g., we do not assume any determinism). For instance, consider the multimodal logic $\mathcal{L} = K_5 + [1]\mathrm{p} \Rightarrow [3][2]\mathrm{p}, [2]\mathrm{p} \Rightarrow [4][1]\mathrm{p}, [2]\mathrm{p} \Rightarrow [5]\mathrm{p}$. Each modal connective of $\mathcal{L}$ corresponds to a PDL modal connective (see e.g. (Demri 2000)). For example, [2] in $\mathcal{L}$ corresponds to the PDL modal connective $[(\mathsf{c}_4; \mathsf{c}_3)^*; (\mathsf{c}_2 \cup (\mathsf{c}_4; \mathsf{c}_1) \cup \mathsf{c}_5)]$. Though the PDL equivalent of [2] contains the star operator, in the paper we show that $\mathcal{L}$-satisfiability is in **PSPACE** (Theorem 8.1(1)). By contrast, the bimodal logic $K_2 + [1]\mathrm{p} \Rightarrow [1][2]\mathrm{p}$ is **EXPTIME**-complete (see Theorem 7.5). More generally, we wonder which grammar logics are in **PSPACE**. In the paper, by proof-theoretical means we characterize a class of right linear grammar logics that are in **PSPACE**.

A standard way to find **PSPACE** upper bound for modal logics consists in designing sound and complete tableaux-like calculi (see e.g. (Kripke 1963, Ladner 1977, Halpern and Moses 1992, Basin et al. 1997, Massacci 1998, Baader and Sattler 2000, Marx et al. 2000a)) augmented with adequate strategies, most of the time depth-first visit of the proof tree with a controlled amount of contractions. One can however distinguish the works that establish **PSPACE** upper bounds but not necessarily the tightest ones (see (Ladner 1977, Halpern and Moses 1992)) from the works that improve the space function by reducing the exponents of the polynomials (see e.g. (Huldelmaier 1993, Hudelmaier 1996)). This is also sometime a matter of natural chronology as determining the decidability status of a logic may precede its computational complexity characterization. The present work belongs rather to the first category since we wish to establish **PSPACE** complexity upper bounds by proof-theoretical means in a uniform way. Although we know that improvements are possible in many cases, we rather concentrate on the gain of generality and uniformity. Moreover, we want to refine the borderline between **EXPTIME**-hard right linear logics and **PSPACE** right linear logics in order to partially answer to the following question inspired from (Vardi 1997, Grädel 1999): why so many (multi)modal logics with weak forms of recursion are in **PSPACE**?

For any right linear grammar logic, we shall define an additive sequent calculi that is proved to be sound and complete. In the spirit of (Ohnishi and Matsumoto 1957), the calculi use neither labels nor a generalized form of sequents. Other kinds of sequent-style calculi for these logics already exist in the literature, see e.g. (Kracht 1996, Szałas 1996, Basin et al. 1998, Baldoni 1998) and we believe that the present formulation of the calculi is quite adequate to find complexity upper bounds mainly because of our treatment of contraction. Then, we show that given a right linear grammar $\mathcal{G}$ and a modal formula ϕ, deciding whether the formula is satisfiable in the extension of K_m with axiom schemes from $\mathcal{G}$ can be done in deterministic exponential-time in the size of $\mathcal{G}$ and ϕ. We refer to this problem as the *general satisfiability problem* for right linear grammar logics. The complexity upper bound is established by using a standard loop checking method dual to the one in (Pratt 1979). An extension is also presented for the global logical consequence problems. This improves upper bounds from (Baldoni 1998, Baldoni et al. 1998).

We also easily show that the general satisfiability problem for right linear grammar logics is **EXPTIME**-hard by exhibiting a decidable countably infinite class of **EXPTIME**-hard right linear grammar logics. Further classes of **EXPTIME**-hard grammar logics can be found in the companion paper (Demri 2000).

In the second part of the paper, we propose a characterization of **PSPACE** decision procedures from the sequent calculi that allows us to show uniformly that all the right linear grammar logics from some identified countably infinite classes of logics are in **PSPACE**. We have indeed found decidable sufficient syntactic properties of the right linear grammars that guarantee that the generated logics from the grammars are in **PSPACE**. All the complexity upper bounds established in the paper are obtained by analyzing proofs in the sequent calculi and thus this follows the proof-theoretic alternative described in Section 8 in (van Benthem 2000) to explain the algorithmic behavior of modal logics.

For instance, we are able to show that given a bimodal extension of K_2 obtained from K_2 by additing axiom schemes from either a left linear or a right linear grammar, deciding whether the satisfiability problem of the logic is in **PSPACE** can be done in linear-time in the size of the grammar. Although the right linear grammars generate the same class of languages as the left linear grammars, this correspondence is not relevant at the level of grammar logics.

2 Logics

Given the set $\mathrm{PRP} = \{\mathrm{p}_i : i \in \mathbb{N}\}$ of propositional variables, the set FORM of modal formulae is defined as the smallest set such that $\mathrm{PRP} \subseteq$ FORM and, if $\phi, \psi \in$ FORM, then $\phi \wedge \psi \in$ FORM, $\neg\phi \in$ FORM and for $i \geq 1$, $[i]\phi \in$ FORM. For $m \geq 1$, we write $\mathtt{L}_m$ to denote the restriction of the modal language to the modal connectives in $\{[i] : i \in \{1, \ldots, m\}\}$. Standard abbreviations include $\vee$, $\Rightarrow$, $\langle i \rangle$. A **necessity formula** is a formula of the form $[i]\phi$ for some $i \geq 1$. The set $\mathrm{sub}(\phi)$ of **subformulae** of the formula ϕ is defined in the standard way. The **modal depth** of an occurrence of a formula ψ in ϕ is the number of occurrences of modal connectives that dominate ψ in ϕ. We write $\mathrm{md}(\phi)$ to denote the maximal modal depth of the subformulae of ϕ. An occurrence of the subformula ψ in ϕ is **positive** [resp. **negative**] $\overset{\text{def}}{\Leftrightarrow}$ it is in the scope of an even [resp. odd] number of negations. The **possibility weight** [resp. **necessity weight**] of a formula ϕ, denoted $\mathrm{pw}(\phi)$ [resp. $\mathrm{nw}(\phi)$], is the number of occurrences of subformulae of the form $[i]\psi$ with negative [resp. positive] polarity. The notation Γ, ϕ, where Γ is a finite multi-set of formulae and ϕ is a formula, designates a multi-set which is the union of Γ with the singleton multi-set containing only ϕ. Let f be a map $f : \mathrm{FORM} \rightarrow \mathbb{N}$. If $\Gamma = \phi_1, \ldots, \phi_n$ is a finite multi-set of formulae, by $f^+(\Gamma)$ we mean the natural number $f(\phi_1) + \ldots + f(\phi_n)$. Similarly, for $i \in \{1, \ldots, m\}$, for any multi-set $\Gamma = \phi_1, \ldots, \phi_n$ of formulae, we write $[i]\Gamma$ [resp. $\neg\Gamma$] to denote $[i]\phi_1, \ldots, [i]\phi_n$ [resp. $\neg\phi_1, \ldots, \neg\phi_n$]. We also write $Set(\Gamma)$ to denote the set of formulae occurring in Γ and $\phi \in \Gamma$ for $\phi \in Set(\Gamma)$.

For any $\mathtt{L}_m$-formula ϕ, we write $\mathrm{r}(\phi)$ to denote the **rank** of ϕ; that is, the number of occurrences of members of $\mathrm{PRP} \cup \{\neg, \wedge\} \cup \{[i] : 1 \leq i \leq m\}$. For example $\mathrm{r}(\mathrm{p} \wedge (\mathrm{q} \wedge \neg \mathrm{p})) = 6$. Under reasonable hypothesis, the length of an $\mathtt{L}_m$-formula ϕ, noted $|\phi|$, is in $\mathcal{O}(\mathrm{r}(\phi) \times (log\ \mathrm{r}(\phi) + log\ m))$. As usual in complexity theory, the extra logarithmic factor is due to the fact that we need an index of size $log\ \mathrm{r}(\phi)$ for the different propositional variables.

An $\mathtt{L}_m$**-frame** is a structure $\mathcal{F} = \langle W, R_1, \ldots, R_m \rangle$ such that W is a nonempty set and for $i \in \{1, \ldots, m\}$, R_i is a binary relation on W. An $\mathtt{L}_m$**-model** is a structure $\mathcal{M} = \langle W, R_1, \ldots, R_m, V \rangle$ such that $\langle W, R_1, \ldots, R_m \rangle$ is an $\mathtt{L}_m$-frame and V is a valuation $V : \mathrm{PRP} \rightarrow \mathcal{P}(W)$. The standard definition of the satisfiability relation $\models$ is omitted here (see e.g. (Blackburn et al. 2001)). An $\mathtt{L}_m$-formula ϕ is said to be **true** in the $\mathtt{L}_m$-model $\mathcal{M}$ (written $\mathcal{M} \models \phi$) $\stackrel{\text{def}}{\Leftrightarrow}$ for all $x \in W$, $\mathcal{M}, x \models \phi$. An $\mathtt{L}_m$-formula ϕ is said to be **true** in the $\mathtt{L}_m$-frame $\mathcal{F}$ (written $\mathcal{F} \models \phi$) $\stackrel{\text{def}}{\Leftrightarrow}$ ϕ is true in all the $\mathtt{L}_m$-models based on $\mathcal{F}$.

In this paper, a **modal logic** $\mathcal{L}$ is understood as a pair $\langle \mathtt{L}_m, \mathcal{S} \rangle$ where $\mathtt{L}_m$ is a modal language with m modal connectives and $\mathcal{S}$ is a nonempty class of $\mathtt{L}_m$-frames. The class $\mathcal{S}$ is usually defined in terms of properties that the relations in the frames of $\mathcal{S}$ are supposed to satisfy. An $\mathtt{L}_m$-formula is said to be $\mathcal{L}$**-satisfiable** $\stackrel{\text{def}}{\Leftrightarrow}$ there is an $\mathtt{L}_m$-model based on some $\mathcal{F} \in \mathcal{S}$ and $x \in W$ such that $\mathcal{M}, x \models \phi$. An $\mathtt{L}_m$-formula is said to be $\mathcal{L}$**-valid** $\stackrel{\text{def}}{\Leftrightarrow}$ for all the $\mathcal{L}$-models $\mathcal{M}$ based on some frame in $\mathcal{S}$, ϕ is true in $\mathcal{M}$. $\mathcal{L}$-satisfiability and $\mathcal{L}$-validity can be easily extended to finite sets of formulae understood as conjunctions.

3 Grammar Logics

For any alphabet Σ (finite set of symbols), we write Σ^* [resp. Σ^+] to denote the set of [resp. nonempty] finite strings built over elements of Σ. ϵ denotes the empty string and $u_1 \cdot u_2$ denotes the concatenation of two strings. For any finite string u, we write $|u|$ to denote its length. For any $u \in \Sigma^*$, we write u^k to denote the string composed of k copies of u. By convention, $u^0 = \epsilon$.

A **(formal) grammar** $\mathcal{G}$ is a quadruple $\mathcal{G} = \langle N, \Sigma, P, S \rangle$ such that

N and Σ are disjoint finite sets of **nonterminal symbols** and **terminal symbols**, respectively (in the paper we allow Σ empty). P is a finite set of **production rules**, each production rule is of the form $u \rightarrow v$ such that $u \in (N \cup \Sigma)^* N (N \cup \Sigma)^*$ and $v \in (N \cup \Sigma)^*$. Finally, $S \in N$ is a special symbol called the **start symbol** (see e.g. (Hopcroft and Ullman 1979)). For the grammar $\mathcal{G}$, the size of $\mathcal{G}$, denoted $|\mathcal{G}|$, is

$$|\mathcal{G}| \stackrel{\text{def}}{=} (card(N) + card(\Sigma) + \Sigma_{u \rightarrow v \in P}(|u \cdot v| + 1)) \times log(card(N) + card(\Sigma))$$

Let $\Rightarrow_{\mathcal{G}}$ be the direct derivation relation defined as the subset of $(N \cup \Sigma)^* \times (N \cup \Sigma)^*$ such that $u \Rightarrow_{\mathcal{G}} v \stackrel{\text{def}}{\Leftrightarrow}$ there is a production rule $u' \rightarrow v' \in P$ such that $u = u_1 \cdot u' \cdot u_2$, $v = u_1 \cdot v' \cdot u_2$, $u_1, u_2 \in (N \cup \Sigma)^*$. Let $\Rightarrow_{\mathcal{G}}^*$ be the reflexive and transitive closure of $\Rightarrow_{\mathcal{G}}$. For $i \in (N \cup \Sigma)$, we write $\mathtt{L}_i(\mathcal{G})$ to denote the set of strings $\{u \in \Sigma^* : i \Rightarrow_{\mathcal{G}}^* u\}$. For instance, for $i \in \Sigma$, $\mathtt{L}_i(\mathcal{G}) = \{i\}$. A grammar $\mathcal{G}$ is said to be **strongly finite** [resp. **finite**] $\stackrel{\text{def}}{\Leftrightarrow}$ for $i \in N$, $\{u \in (N \cup \Sigma)^* : i \Rightarrow_{\mathcal{G}}^* u\}$ is finite [resp. $\mathtt{L}_i(\mathcal{G})$ is finite]. It is possible that for some $i \in N$, $\mathtt{L}_i(\mathcal{G})$ is empty although $\mathcal{G}$ is not strongly finite.

In the rest of the paper we assume that each grammar $\langle N, \Sigma, P, S\rangle$ satisfies $N = \{1, \ldots, k\}$ for some $k \geq 1$, $\Sigma = \{k+1, \ldots, m\}$ for some $k \leq m$ (we allow Σ to be empty) and $S = 1$.

Let $\mathcal{G} = \langle N, \Sigma, P, S\rangle$ be a grammar and $\rightsquigarrow$ be the binary relation in $N \times (N \cup \Sigma)^*$ such that $i \rightsquigarrow u \stackrel{\text{def}}{\Leftrightarrow}$ either there is $j \rightarrow u \in P$ such that $i \Rightarrow_{\mathcal{G}}^* j$ or $u \in N$ and $i \Rightarrow_{\mathcal{G}}^* u$. If $\mathcal{G}$ is right [resp. left] linear, then by using the technique for eliminating the unit production rules (Hopcroft and Ullman 1979), one can compute in polynomial-time in $|\mathcal{G}|$ a right [resp. left] linear grammar $\mathcal{G}' = \langle N', \Sigma', P', S'\rangle$ such that $N = N'$, $\Sigma = \Sigma'$, $S = S'$, $P \subseteq P'$ and for $\langle i, u\rangle \in (N \cup \Sigma)^*$, $i \rightsquigarrow u$ in $\mathcal{G}$ iff $i \rightarrow u \in P'$. For $i \in N$, we have $\{u \in (N \cup \Sigma)^* : i \Rightarrow_{\mathcal{G}}^* u\} = \{u \in (N \cup \Sigma)^* : i \Rightarrow_{\mathcal{G}'}^* u\}$. In the case when $\mathcal{G}$ is either left linear or right linear (called **regular** in the sequel), $\rightsquigarrow$ can be computed in polynomial-time in $|\mathcal{G}|$ and for all $i \in N$, $\sum_{i \rightsquigarrow u} |u| \leq |\mathcal{G}|$. Although $\mathcal{G}$ and $\mathcal{G}'$ generate the same language ($\mathtt{L}_1(\mathcal{G}) = \mathtt{L}_1(\mathcal{G}')$), in the sequel we do not assume that the grammars are necessarily of the form of $\mathcal{G}'$. Indeed, grammars generating the same language, may engender different grammar logics. The binary relation $\rightsquigarrow$ is used in Section 4 to define sequent calculi.

Let $\mathcal{G}$ be a grammar and $\mathcal{S}$ be a class of $\mathtt{L}_m$-frames. We write $\mathcal{S}^{\mathcal{G}}$ to denote the subset of $\mathcal{S}$ such that for any $\mathcal{F} = \langle W, R_1, \ldots, R_m\rangle \in \mathcal{S}$, $\mathcal{F} \in \mathcal{S}^{\mathcal{G}} \stackrel{\text{def}}{\Leftrightarrow}$ for any production rule $i_1 \ldots i_{k'} \rightarrow j_1 \ldots j_{k''}$ in $\mathcal{G}$, $R_{j_1} \circ \ldots \circ R_{j_{k''}} \subseteq R_{i_1} \circ \ldots \circ R_{i_{k'}}$. For the logic $\mathcal{L}_m = \langle \mathtt{L}_m, \mathcal{S}_m\rangle$ where $\mathcal{S}_m$ is the class of all the $\mathtt{L}_m$-frames, we write $\mathcal{L}_m^{\mathcal{G}}$ to denote the logic $\langle \mathtt{L}_m, \mathcal{S}_m^{\mathcal{G}}\rangle$. $\mathcal{L}_m^{\mathcal{G}}$ is said to be a **grammar logic** (Fariñas del Cerro and Penttonen 1988).

For any string $u = i_1 \cdot \ldots \cdot i_n$ in $\{1, \ldots, m\}^*$, we write R_u to denote $R_{i_1} \circ \ldots \circ R_{i_n}$. When $u = \epsilon$, $R_u \stackrel{\text{def}}{=} \{\langle x, x\rangle : x \in W\}$. Moreover, we write $[u]\phi$ to denote the $\mathtt{L}_m$-formula $[i_1] \ldots [i_n]\phi$ where $u = i_1 \cdot \ldots \cdot i_n$. If $u = \epsilon$, then $[u]\phi$ is simply ϕ.

Theorem 3.1 *Let $\mathcal{G} = \langle N, \Sigma, P, S\rangle$. For $u, v \in (N \cup \Sigma)^*$, (I) $u \Rightarrow^*_{\mathcal{G}} v$ iff (II) $[u]\mathrm{p} \Rightarrow [v]\mathrm{p}$ is $\mathcal{L}^{\mathcal{G}}_m$-valid iff (III) for all $\mathcal{L}^{\mathcal{G}}_m$-models $R_v \subseteq R_u$.*

The equivalence between (II) and (III) is a classical correspondence result in modal logic theory (see e.g. (van Benthem 1984)). (I) implies (II) can be proved by induction on the length of the derivation whereas (II) implies (I) can be shown by using part of the proof of Theorem 3 in (Chagrov and Shehtman 1994). In order to study the grammar logic $\mathcal{L}^{\mathcal{G}}_m$, what is essential is the value of the set P of production rules whereas once P is fixed, the value of the start symbol S and the distribution of the terminal and nonterminal symbols are immaterial for $\mathcal{L}^{\mathcal{G}}_m$-satisfiability. Hence, semi-Thue rewriting systems are also appropriate to define grammar logics.

The **general satisfiability problem** GSP(REG) [resp. GSP(LIN), GSP(RLIN_f), GSP(RLIN)] for regular grammar [resp. linear grammar, finite right linear grammar, right linear grammar] logics is defined as follows:

- Inputs: a regular [resp. right linear, finite right linear, linear] grammar $\mathcal{G}$ and an $\mathtt{L}_m$-formula ϕ;
- Question: Is ϕ $\mathcal{L}^{\mathcal{G}}_m$-satisfiable?

The above general satisfiability problems can be viewed as syntactic variants of satisfiability problems for fragments of the well-known description logic $\mathcal{ALC}$ augmented with role value maps (see details in (Demri 2000)).

It is known that the multimodal logic K_m, $m \geq 1$, has a **PSPACE**-complete satisfiability problem (see e.g. (Halpern and Moses 1992)). Adding a regular set of modal axioms preserves the **PSPACE** complexity lower bound.

Theorem 3.2 *Let $\mathcal{G}$ be either a regular grammar or a context-free grammar with a nonempty set of terminal symbols. Then, $\mathcal{L}^{\mathcal{G}}_m$-satisfiability is* **PSPACE**-*hard.*

A natural proof consists in reducing satisfiability for either the modal logic K or the modal logic T into $\mathcal{L}^{\mathcal{G}}_m$-satisfiability. The only difficulty in the proof is to show that for any binary relation R there is an $\mathcal{L}^{\mathcal{G}}_m$-frame with $R_m = R$. Additionally, GSP(LIN) is undecidable. This can

$$\Gamma, \phi \vdash \Delta, \phi \text{ (initial sequents)} \quad \frac{\Gamma \vdash \Delta, \phi}{\Gamma, \neg\phi \vdash \Delta} \ (\neg \vdash) \quad \frac{\Gamma, \phi \vdash \Delta}{\Gamma \vdash \Delta, \neg\phi} \ (\vdash \neg)$$

$$\frac{\Gamma, \phi_1, \phi_2 \vdash \Delta}{\Gamma, \phi_1 \wedge \phi_2 \vdash \Delta} \ (\wedge \vdash) \quad \frac{\Gamma \vdash \Delta, \phi_1 \quad \Gamma \vdash \Delta, \phi_2}{\Gamma \vdash \Delta, \phi_1 \wedge \phi_2} \ (\vdash \wedge)$$

FIGURE 1 Initial sequents and standard rules for propositional connectives

be proved by reducing the problem of empty intersection between linear languages into GSP(LIN) (see e.g. (Rozenberg and Salomaa 1994)) by using either prefixed tableaux calculi (Baldoni et al. 1998) or the equational characterization of context-free languages (Demri 2000).

4 Sequent calculi

4.1 Definitions

Let $\mathcal{G}$ be a right linear grammar. The basic syntactic objects in the calculi are **sequents**. A **sequent** is an expression of the form $\Gamma \vdash \Delta$ where Γ and Δ are finite multi-sets of formulae, Γ is the **antecedent** and Δ the **succedent**. We write SEQ_m to denote the class of all sequents built over the modal language L_m. The additive sequent calculus $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ for the logic $\mathcal{L}_m^{\mathcal{G}}$ contains the rules from Figure 1 for the propositional fragment of $\mathcal{L}_m^{\mathcal{G}}$. The rules are read upwards and other rules depending of $\mathcal{G}$ are presented below. The left-hand side introduction rule $[i]$ is defined below only if $i \Rightarrow_{\mathcal{G}}^* \epsilon$:

$$\frac{\Gamma, [i]\phi, \phi \vdash \Delta}{\Gamma, [i]\phi \vdash \Delta} \ ([i] \vdash)$$

The right-hand side introduction rule for $[i]$ is defined as follows. For $i \in \{1, \ldots, k\}$, let $\rightsquigarrow (i) = \{u_{i,1}, \ldots, u_{i,l_i}\}$ where given a binary relation R on U and $x \in U$, $R(x) \stackrel{\text{def}}{=} \{y \in U : \langle x, y\rangle \in R\}$. For $j \in \{1, \ldots, m\}$ and for $i \in \{1, \ldots, k\}$, let $start_j^i = \{\langle i, k'\rangle : u_{i,k'} = j \cdot v_{i,k'}\}$. Let $i \in \{k+1, \ldots, m\}$. The $(\vdash [i])$ rule is defined as follows:

$$\frac{\bigcup_{\langle 1,k'\rangle \in start_i^1} [v_{1,k'}]\Gamma_1, \ldots, \bigcup_{\langle k,k'\rangle \in start_i^k} [v_{k,k'}]\Gamma_k, \Gamma_i \vdash \phi}{\Gamma', [1]\Gamma_1, \ldots, [k]\Gamma_k, [i]\Gamma_i \vdash [i]\phi, \Delta} \ (\vdash [i])$$

Moreover, we assume that in Γ', there is no formula of the form $[j]\psi$ for some $j \in \{1, \ldots, k\} \cup \{i\}$. It remains to define for $i \in \{1, \ldots, k\}$, the $(\vdash [i])$ rule:

$$\frac{\Gamma'_1, \ldots, \Gamma'_k \vdash \phi}{\Gamma', [1]\Gamma_1, \ldots, [k]\Gamma_k \vdash [i]\phi, \Delta} \ (\vdash [i])$$

where for $j \in \{1, \ldots, k\}$, if $j \rightsquigarrow i$, then $\Gamma'_j \stackrel{\text{def}}{=} \Gamma_j$, otherwise $\Gamma'_j \stackrel{\text{def}}{=} \emptyset$

(empty multi-set). Moreover, we assume that in Γ', there is no formula of the form $[j]\psi$ for some $j \in \{1, \ldots, k\}$.

For $i \in \{1, \ldots, m\}$ we abbreviate the $(\vdash [i])$ rule by

$$\frac{\Gamma^{[i]} \vdash \phi}{\Gamma \vdash [i]\phi, \Delta}\ (\vdash [i])$$

where the appropriate definition of $\Gamma^{[i]}$ from Γ is immediate from the above cases.

The $([i] \vdash)$ rules and the $(\vdash [i])$ rules are defined from the strings u satisfying $i \rightsquigarrow u$ and not only from the ones satisfying $i \rightarrow u \in P$. This is the price we may have to pay in the $(\vdash [i])$ rules since we want to introduce a $[i]$-formula at the right-hand side and to take into account the grammatical properties of the logic simultaneously.

An implicit contraction is operated in the $([i] \vdash)$-rule (if $i \Rightarrow^*_{\mathcal{G}} \epsilon$) whereas implicit contractions can be found also in the applications of the $(\vdash [i])$-rule but this depends on the structure of $\mathcal{G}$ (see e.g. Section 8 for calculi with no implicit contractions). In order to get the **PSPACE** upper bounds, our main task is to control contraction and this requires a careful analysis.

Example 4.1 Let $\mathcal{G}^{rl} = \langle \{1,2\}, \{3,4\}, \{1 \rightarrow 3{\cdot}3{\cdot}1, 1 \rightarrow 2, 1 \rightarrow 4{\cdot}2, 2 \rightarrow 4 \cdot 4 \cdot 4 \cdot 2\}, 1\rangle$ be a right linear grammar. The $(\vdash [4])$ rule is defined as follows:

$$\frac{[2]\Gamma_1, [4][4][2](\Gamma_1, \Gamma_2), \Gamma_4 \vdash \phi}{\Gamma', [1]\Gamma_1, [2]\Gamma_2, [4]\Gamma_4 \vdash [4]\phi, \Delta}\ (\vdash [4])$$

Observe the duplication of Γ_1.

As is usual, a **proof** Π in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$ is a tree whose nodes are labelled by sequents satisfying the following conditions: the topmost sequents of Π are initial sequents and every sequent of Π, except the lowest one is an upper sequent of an inference whose lower sequent is also in Π. A sequent $\Gamma \vdash \Delta$ is **provable** in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m \overset{\text{def}}{\Leftrightarrow}$ there is a proof where the lowest sequent is $\Gamma \vdash \Delta$. A formula ϕ is **provable in** $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m \overset{\text{def}}{\Leftrightarrow}$ the sequent $\emptyset \vdash \phi$ (also noted $\vdash \phi$) is provable in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$. A sequent $\Gamma \vdash \Delta$ is **consistent** $\overset{\text{def}}{\Leftrightarrow}$ $\Gamma \vdash \Delta$ is not provable in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$.

A sequent $\Gamma \vdash \Delta$ is **downward saturated** $\overset{\text{def}}{\Leftrightarrow}$ for $\phi \in \Gamma$ and for $\psi \in \Delta$:

- $\phi = \phi_1 \wedge \phi_2$ implies $\phi_1, \phi_2 \in \Gamma$; $\phi = [i]\phi_1$ and $i \Rightarrow^*_{\mathcal{G}} \epsilon$ imply $\phi_1 \in \Gamma$;
- $\phi = \neg\phi_1$ implies $\phi_1 \in \Delta$; $\psi = \neg\psi_1$ implies $\psi_1 \in \Gamma$;
- $\psi = \psi_1 \wedge \psi_2$ implies either $\psi_1 \in \Delta$ or $\psi_2 \in \Delta$.

A **derivation** is defined as a proof except that the topmost sequents are not necessarily initial sequents. The derivations are supposed to grow upwards. If Π is a derivation of $\Gamma \vdash \Delta$ we write $\sigma = \Gamma_0 \vdash \Delta_0 \prec (r_1)\Gamma_1 \vdash \Delta_1 \prec (r_2) \ldots \prec (r_{n-1})\Gamma_n \vdash \Delta_n$ to denote the fact that there is an initial segment σ of a branch in Π from the root $\Gamma_0 \vdash \Delta_0 = \Gamma \vdash \Delta$ such that for $i \in \{0, \ldots, n-1\}$, $\Gamma_{i+1} \vdash \Delta_{i+1}$ is one of the premisses of the inference of the rule (r_{i+1}) with conclusion $\Gamma_i \vdash \Delta_i$. We omit to write the r_i's when they are of no use. We write $Ant(\sigma)$ [resp. $Suc(\sigma)$] to denote the antecedent set $\bigcup_{0 \leq i \leq n} Set(\Gamma_i)$ [resp. the succedent set $\bigcup_{0 \leq i \leq n} Set(\Delta_i)$]. We write $last(\sigma)$ [resp. $first(\sigma)$] to denote $\Gamma_n \vdash \Delta_n$ [resp. $\Gamma_0 \vdash \Delta_0$]. The sequence σ is said to be **local** $\overset{\text{def}}{\Leftrightarrow}$ no r_i is the right-hand side introduction rule $(\vdash [j])$ for some $j \in \{1, \ldots, m\}$. The sequence σ is said to be **consistent** $\overset{\text{def}}{\Leftrightarrow}$ all the $\Gamma_i \vdash \Delta_i$ are consistent. The sequence σ is said to be **maximal** $\overset{\text{def}}{\Leftrightarrow}$ σ is local and $Ant(\sigma) \vdash Suc(\sigma)$ is downward saturated. The maximal and consistent sequences play the role of downward saturated sets in the standard terminology for tableaux (see e.g. (Goré 1999) for further details and historical notes). This complication is due to the fact that we consider multi-sets instead of sets in the sequents. The reward is that we can more easily control contraction and this shall be helpful to get **PSPACE** complexity upper bounds.

4.2 Properties

Let ϕ be a formula. The closure of ϕ with respect to $\mathcal{G}$ is the smallest set $\mathrm{cl}_{\mathcal{G}}(\phi)$ of formulae such that $\mathrm{cl}_{\mathcal{G}}(\phi)$ is closed under subformulae, $\mathrm{sub}(\phi) \subseteq \mathrm{cl}_{\mathcal{G}}(\phi)$ and if $i \rightsquigarrow u$ and $[i]\psi \in \mathrm{cl}_{\mathcal{G}}(\phi)$, then $[u]\psi \in \mathrm{cl}_{\mathcal{G}}(\phi)$. One can prove that $card(\mathrm{cl}_{\mathcal{G}}(\phi))$ is bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$. We write $\mathrm{SEQ}(\phi)$ to denote the set of sequents $\Gamma \vdash \Delta$ such that $\bigcup_{\psi \in \Gamma, \Delta} \mathrm{cl}_{\mathcal{G}}(\psi) \subseteq \mathrm{cl}_{\mathcal{G}}(\phi)$.

Lemma 4.2 *Let $\Gamma \vdash \Delta$ be a sequent. Then, every formula occurring in a derivation of $\Gamma \vdash \Delta$ belongs to $\bigcup_{\psi \in \Gamma, \Delta} \mathrm{cl}_{\mathcal{G}}(\psi)$.*

The (easy) proof is by induction on the depth of the proof tree. Following for instance the terminology from (Goré 1999), $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ has therefore the analytical superformula property and obviously $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ does not have necessarily the subformula property.

Lemma 4.3 *$\Gamma \vdash \Delta$ is a provable sequent in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ iff $\Gamma \vdash \Delta$ has a proof in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ such that all the initial sequents are of the form $\Gamma', \mathrm{p} \vdash \mathrm{p}, \Delta'$ where p is a propositional variable.*

The proof of Lemma 4.3 is standard. It is sufficient to show in the induction step that for every initial sequent $\Gamma', \phi \vdash \phi, \Delta'$ with $\mathrm{r}(\phi) \geq 2$, there is a proof of $\Gamma', \phi \vdash \phi, \Delta'$ such that all the initial sequents are of the form $\Gamma', \psi \vdash \psi, \Delta'$ with $\mathrm{r}(\psi) < \mathrm{r}(\phi)$.

A rule is **invertible** $\overset{\text{def}}{\Leftrightarrow}$ for every inference of the rule, the conclusion has a proof iff the premises have proofs.

Lemma 4.4 *The rules* $(\wedge \vdash)$*,* $(\vdash \wedge)$*,* $(\neg \vdash)$*,* $(\vdash \neg)$ *and* $([i] \vdash)$ *if* $i \Rightarrow^*_{\mathcal{G}} \epsilon$ *are invertible.*

The proof of Lemma 4.4 uses Lemma 4.3 and is not difficult to show. For instance, invertibility of $([i] \vdash)$ if $i \Rightarrow^*_{\mathcal{G}} \epsilon$, is immediate.

Lemma 4.5 *Let* $\Gamma \vdash \Delta$ *be a consistent sequent. Then, there is a maximal and consistent sequence* $\sigma = \Gamma_0 \vdash \Delta_0 \prec \Gamma_1 \vdash \Delta_1 \prec \ldots \prec \Gamma_n \vdash \Delta_n$ *with* $\Gamma_0 \vdash \Delta_0 = \Gamma \vdash \Delta$.

Proof. Since $\Gamma \vdash \Delta$ is consistent, we know that no proof of $\Gamma \vdash \Delta$ exists and there is no propositional variable occurring in both Γ and Δ. Let $\Gamma_0 \vdash \Delta_0 = \Gamma \vdash \Delta$. If $\Gamma_0 \vdash \Delta_0$ is downward saturated, then we are done. Now suppose that $\sigma_i = \Gamma_0 \vdash \Delta_0 \prec \Gamma_1 \vdash \Delta_1 \prec \ldots \prec \Gamma_i \vdash \Delta_i$ is a local sequence, $Ant(\sigma_i) \vdash Suc(\sigma_i)$ is not downward saturated and each $\Gamma_j \vdash \Delta_j$, $0 \leq j \leq i$, is consistent. For instance, suppose that $\phi_1 \wedge \phi_2 \in Ant(\sigma_i)$ and $\phi_1, \phi_2 \notin Ant(\sigma_i)$. Hence, $\phi_1 \wedge \phi_2 \in \Gamma_i$ since otherwise $\phi_1, \phi_2 \in \Gamma_{i'}$ for some $i' < i$. Apply the $(\wedge \vdash)$-rule to an occurrence of $\phi_1 \wedge \phi_2$ in $\Gamma_i \vdash \Delta_i$ and let $\Gamma_{i+1} \vdash \Delta_{i+1}$ be $(\Gamma_i \setminus \{\phi_1 \wedge \phi_2\}), \phi_1, \phi_2 \vdash \Delta_i$. Since the $(\wedge \vdash)$-rule is invertible (see Lemma 4.4), $\Gamma_{i+1} \vdash \Delta_{i+1}$ is also consistent. If $\Gamma_0 \vdash \Delta_0 = \Gamma \vdash \Delta$ and $Ant(\sigma_i) \vdash Suc(\sigma_i)$ is not downward saturated for some other reason, we use a similar reasoning with obvious adaptations (we may have to choose between two branches). Lemma 4.2 guarantees termination and the length of σ can be bounded by $card(\mathrm{cl}_{\mathcal{G}}(\Gamma, \Delta))$. $\dashv$

4.3 Completeness

Following (Rautenberg 1983) (see also (Goré 1999)), we introduce the central notion of model graph. We prove completeness using the well-known technique due to Schütte (see e.g. (Schütte 1967, Takeuti 1975)).

Definition 4.6 A **model graph** for some sequent $\Gamma \vdash \Delta$ is an $\mathtt{L}_m$-frame of the form $\langle W, R_{1+m}, \ldots, R_{2\times m}\rangle$ such that W is a countable set of maximal and consistent sequences such that

1. for $\sigma \in W$, $Ant(\sigma) \cup Suc(\sigma) \subseteq \bigcup_{\psi \in \Gamma, \Delta} \mathrm{cl}_{\mathcal{G}}(\psi)$;
2. there is σ_0 such that $Set(\Gamma) \subseteq Ant(\sigma_0)$ and $Set(\Delta) \subseteq Suc(\sigma_0)$;
3. for $\sigma \in W$, if $[i]\phi \in Suc(\sigma)$ for some $i \in \{1, \ldots, m\}$, then there is $\sigma' \in W$ such that $\sigma R_{i+m} \sigma'$ and $\phi \in Suc(\sigma')$;

4. for $\sigma, \sigma' \in W$, if $\sigma R_{i+m} \sigma'$ and $j \rightsquigarrow i \cdot u$ for some $\langle i, j\rangle \in \{1, \ldots, m\} \times \{1, \ldots, k\}$, and $[j]\phi \in Ant(\sigma)$, then $[u]\phi \in Ant(\sigma')$;
5. for $\sigma, \sigma' \in W$, if $\sigma R_{i+m} \sigma'$ for some $i \in \{k+1, \ldots, m\}$, and $[i]\phi \in Ant(\sigma)$, then $\phi \in Ant(\sigma')$.

A direct consequence of Definition 4.6(4) is that Definition 4.6(5) holds true even for $i \in \{1, \ldots, k\}$. The cornerstone of the completeness proof is the following result.

Theorem 4.7 *If there is a model graph for $\Gamma \vdash \Delta$, then $Set(\Gamma) \cup Set(\neg\Delta)$ is $\mathcal{L}^{\mathcal{G}}_m$-satisfiable.*

Proof. (sketch) For $i \in \{1, \ldots, k\}$, there is $l_i \geq 0$ such that $i \rightarrow u_{i,1}, \ldots, i \rightarrow u_{i,l_i}$ are the only production rules in P having i as left-hand side. Let W be a countable non-empty set and $R_{m+1}, \ldots, R_{2\times m}$ be binary relations on W. Let $f : \mathcal{P}(W \times W)^m \rightarrow \mathcal{P}(W \times W)^m$ be the map such that

$$f(R_1, \ldots, R_m) \stackrel{\text{def}}{=}$$

$$\langle (R_{1+m} \cup \bigcup_{1 \leq j \leq l_1} R_{u^\star_{1,j}}), \ldots, (R_{k+m} \cup \bigcup_{1 \leq j \leq l_k} R_{u^\star_{k,j}}), R_{k+1+m}, \ldots, R_{2\times m}\rangle$$

where $u^\star_{i,j}$ is a string obtained from $u_{i,j}$ by replacing $i \in \{k+1, \ldots, m\}$ by $i+m$. Let $\leq$ be the binary relation on $\mathcal{P}(W \times W)^m$ defined as follows: $\langle R_1, \ldots, R_m\rangle \leq \langle R'_1, \ldots, R'_m\rangle \stackrel{\text{def}}{\Leftrightarrow}$ for all $i \in \{1, \ldots, m\}$, $R_i \subseteq R'_i$. The structure $\langle \mathcal{P}(W \times W)^m, \leq\rangle$ is a complete lattice and f is continuous and order-preserving. By Kleene's Theorem, the least fixed point of f exists and is equal to

$$\mu(f) = \bigcup_{i \in \mathbb{N}} f^i(\emptyset, \ldots, \emptyset) \stackrel{\text{def}}{=} \langle \mathcal{R}_1, \ldots, \mathcal{R}_m\rangle.$$

By construction, $\langle W, \mathcal{R}_1, \ldots, \mathcal{R}_m\rangle$ is an $\mathcal{L}^{\mathcal{G}}_m$-frame such that for $i \in \{1, \ldots, m\}$, $R_{i+m} \subseteq \mathcal{R}_i$ and $\mathcal{R}_i \subseteq (R_{m+1} \cup \ldots \cup R_{2\times m})^*$.

Let $\langle W, R_{1+m}, \ldots, R_{2\times m}\rangle$ be a model graph for $\Gamma \vdash \Delta$. We define an $\mathcal{L}^{\mathcal{G}}_m$-model $\mathcal{M} = \langle W, \mathcal{R}_1, \ldots, \mathcal{R}_m, V\rangle$ as follows: $\langle \mathcal{R}_1, \ldots, \mathcal{R}_m\rangle$ is the least fixed point of f defined with R_{1+m}, …, $R_{2\times m}$ and for any propositional variable p, $V(\mathrm{p}) \stackrel{\text{def}}{=} \{\sigma \in W : \mathrm{p} \in Ant(\sigma)\}$.

By induction on formulae we can show that for all $\sigma \in W$, for $\phi \in Ant(\sigma) \cup Suc(\sigma)$, if $\phi \in Ant(\sigma)$, then $\mathcal{M}, \sigma \models \phi$ otherwise $\mathcal{M}, \sigma \not\models \phi$. Moreover, there is $\sigma_0 \in W$ such that $Set(\Gamma) \subseteq Ant(\sigma_0)$ and $Set(\Delta) \subseteq Suc(\sigma_0)$. Consequently, $\mathcal{M}, \sigma_0 \models Set(\Gamma) \cup Set(\neg\Delta)$. $\dashv$

Theorem 4.8 *For any sequent $\Gamma \vdash \Delta$, $(\bigwedge_{\phi \in \Gamma} \phi) \Rightarrow (\bigvee_{\phi \in \Delta} \phi)$ is $\mathcal{L}^{\mathcal{G}}_m$-valid iff the sequent $\Gamma \vdash \Delta$ is provable in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$.*

Proof. The soundness proof is standard by using an induction on the depth of the proof tree. In order to prove completeness, we assume that $(\bigwedge_{\phi\in\Gamma}\phi) \Rightarrow (\bigvee_{\phi\in\Delta}\phi)$ is $\mathcal{L}^{\mathcal{G}}_m$-valid and suppose that $\Gamma \vdash \Delta$ is consistent. The first step is to create a maximal and consistent sequence σ_0 starting with $\Gamma \vdash \Delta$ (see Lemma 4.5). Since σ_0 is consistent, $last(\sigma_0)$ is consistent. We use this fact to construct an $\mathtt{L}_m$-frame whose infinite limit will be a graph model. This is a meta-level construction for which we can visit all derivations for $last(\sigma_0)$, choosing nodes at will, since all such derivations cannot be completed as proofs. We use the successor relations R_i, $1 \leq i \leq m$ while building this frame. By Theorem 4.7, $(\bigwedge_{\phi\in\Gamma}\phi) \wedge (\bigwedge_{\phi\in\Delta}\neg\phi)$ is $\mathcal{L}^{\mathcal{G}}_m$-satisfiable which will lead to a contradiction.

Let us show how to build the model graph. If no $[i]\phi$ occurs in the succedent part of $last(\sigma_0)$, then the structure $\langle\{\sigma_0\}, \emptyset, \ldots, \emptyset\rangle$ is a model graph for $\Gamma \vdash \Delta$. Otherwise, for $i \in \{1, \ldots, m\}$, let $\psi^1_i, \ldots, \psi^{s_i}_i$ be all the formulae such that $[i]\psi^j_i$ occurs in the succedent part of $last(\sigma_0)$. Let Γ' be the antecedent part of $last(\sigma_0)$. Since $last(\sigma_0)$ is consistent, for $i \in \{1, \ldots, m\}$, for $j \in \{1, \ldots, s_i\}$, ${\Gamma'}^{[i]} \vdash \psi^j_i$ is also consistent. For $i \in \{1, \ldots, m\}$, for $j \in \{1, \ldots, s_i\}$, create a maximal and consistent sequence $\sigma_{i,j}$ starting with ${\Gamma'}^{[i]} \vdash \psi^j_i$ and put $\sigma_0 R_i \sigma_{i,j}$. The sequences $\sigma_{i,j}$ are said to be of level 1. $\sum_{i=1}^m s_i$ is bounded by $|\mathcal{G}| \times (\sum_{\psi\in\Gamma} \mathrm{r}(\psi) + \sum_{\psi\in\Delta} \mathrm{r}(\psi))$ and this shall hold true at any level. Continue to create the nodes of further levels in a similar way. Either this procedure can go forever (but the infinite limit frame is a model graph for $\Gamma \vdash \Delta$) or the procedure stops after a finite amount of time (the number of level is finite as well as the resulting model graph). $\dashv$

5 GSP(RLIN) is EXPTIME-complete

No single rule $(\vdash [i])$ for some $i \in \{1, \ldots, m\}$ in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$ captures all the properties of the relation R_i in the $\mathcal{L}^{\mathcal{G}}_m$-frames unlike the combination of all the introduction rules for necessity formulae. As in the *single steps calculi* in (Massacci 1994) (see also (Goré 1999, Massacci 2000)), the closure property of the $\mathcal{L}^{\mathcal{G}}_m$-frames (see the proof of Theorem 4.7) is encoded step by step and this is the key point to characterize the complexity of GSP(RLIN). That is why, we can improve the complexity upper bound of GSP(RLIN) from (Baldoni 1998, Baldoni et al. 1998).

Theorem 5.1 *GSP(RLIN) is in* **EXPTIME**.

Proof. We use a technique that is dual to the method in (Pratt 1979) that shows that PDL satisfiability is in **EXPTIME**. By Theorem 4.8, ϕ is $\mathcal{L}^{\mathcal{G}}_m$-satisfiable iff $\neg\phi$ is not provable in $\mathrm{G}\mathcal{L}^{\mathcal{G}}_m$ and therefore we concentrate on validity instead of satisfiability since our procedure is deterministic. Actually, we use a variant of $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$, namely $\mathrm{SETG}\mathcal{L}_m^{\mathcal{G}}$, where the sequents are pairs of finite sets (instead of finite multi-sets). A contraction is explicitly operated for the rules $(\vdash \wedge)$, $(\wedge \vdash)$, $(\neg \vdash)$ and $(\vdash \neg)$. Following the developments of Section 4, one can show that $X \vdash Y$ is provable in $\mathrm{SETG}\mathcal{L}_m^{\mathcal{G}}$ iff the formula $(\bigwedge_{\phi \in X} \phi) \Rightarrow (\bigvee_{\phi \in Y} \phi)$ is $\mathcal{L}_m^{\mathcal{G}}$-valid. Let ϕ be a formula for which we want to know whether ϕ is $\mathcal{L}_m^{\mathcal{G}}$-valid. The rules of the proof system $\mathrm{SETG}\mathcal{L}_m^{\mathcal{G}}$ can be computed in polynomial-time in $|\mathcal{G}|$ since the binary relation $\rightsquigarrow$ can be computed in polynomial-time in $|\mathcal{G}|$ and all the rules are of size in $\mathcal{O}(|\mathcal{G}|)$. Their applicability can be checked in polynomial-time in $|\mathcal{G}|$ and in the size of the premiss and conclusion sequents. The cost of the computation of $\mathrm{SETG}\mathcal{L}_m^{\mathcal{G}}$ is relevant here since $\mathcal{G}$ is part of the inputs.

The cardinality of the set $\mathrm{cl}_{\mathcal{G}}(\phi)$ is at most $|\mathcal{G}| \times \mathrm{r}(\phi)$. Let $\mathrm{SETSEQ}(\phi)$ be the finite set of sequents $X \vdash Y$ for $X, Y \subseteq \mathrm{cl}_{\mathcal{G}}(\phi)$. $\mathrm{SETSEQ}(\phi)$ is obviously a subset of the countably infinite set $\mathrm{SEQ}(\phi)$ where only sets can occur in the sequents of $\mathrm{SETSEQ}(\phi)$. The cardinality of $\mathrm{SETSEQ}(\phi)$ is bounded by $2^{|\mathcal{G}| \times \mathrm{r}(\phi)+1}$ and the size of each $X \vdash Y$ is at most $2 \times (|\mathcal{G}| \times |\phi|)^2$. Obviously, $\vdash \phi$ belongs to $\mathrm{SETSEQ}(\phi)$. We define a sequence of sets $Z_1 \subseteq Z_2 \subseteq Z_3 \subseteq \ldots$ included in $\mathrm{SETSEQ}(\phi)$. Z_1 is defined as the set of sequents $X \vdash Y$ from $\mathrm{SETSEQ}(\phi)$ such that $X \cap Y \neq \emptyset$. Now suppose that Z_i is defined and let us define Z_{i+1}. For $X \vdash Y \in \mathrm{SETSEQ}(\phi)$, $X \vdash Y \in Z_{i+1} \stackrel{\text{def}}{\Leftrightarrow}$ either (C1) $X \vdash Y \in Z_i$ or (C2) there are $X_1 \vdash Y_1, X_l \vdash Y_l \in Z_i$ such that

$$\frac{X_1 \vdash Y_1 \quad \ldots \quad X_l \vdash Y_l}{X \vdash Y}\ (r)$$

is a correct inference of the rule (r) in $\mathrm{SETG}\mathcal{L}_m^{\mathcal{G}}$. The index l takes the value either 1 or 2 according to the form of the rule (r). Since $card(X) + card(Y) \leq 2 \times |\mathcal{G}| \times \mathrm{r}(\phi)$, each formula in $X \vdash Y$ can be principal in only one rule, checking the condition (C2) can be done in exponential time in $|\mathcal{G}| + |\phi|$. If $\vdash \phi \in Z_{i+1}$, we stop and return 'yes' ϕ is $\mathcal{L}_m^{\mathcal{G}}$-valid. Otherwise, we continue the construction. Since $Z_i \subseteq Z_{i+1}$ and $\mathrm{SETSEQ}(\phi)$ has at most $2^{|\mathcal{G}| \times \mathrm{r}(\phi)+1}$ elements, this construction terminates after at most exponentially many stages. Computing Z_{i+1} can be done in deterministic exponential time in $|\mathcal{G}| + |\phi|$. Hence, the whole construction can be done in deterministic exponential time in $|\mathcal{G}| + |\phi|$. Whenever $Z_i = Z_{i+1}$ and $\vdash \phi \notin Z_i$, we return 'no', ϕ is not $\mathcal{L}_m^{\mathcal{G}}$-valid. The procedure can be shown to be correct. $\dashv$

The procedure in the proof of Theorem 5.1 is more suited for proving theoretical results than for being used in applications. It can be viewed as the addition of a highly inefficient loop checking to the calculi $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$

(see e.g. (Ladner 1977, Fitting 1983, Cerrito and Cialdea Mayer 1997, Heuerding 1998) for related matters). Besides the **EXPTIME** upper bound is sharp enough.

Theorem 5.2 *For $m \geq 2$, there is a decidable countably infinite set of right linear grammars such that $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is* **EXPTIME***-hard.*

Proof. (sketch) Let $\mathcal{G}_i$, $i \geq 1$, be the right linear grammar

$$\langle \{1\}, \{2, \ldots, m\}, \{1 \rightarrow \epsilon, 1 \rightarrow 2^i \cdot 1\}, 1\rangle.$$

For any two different prime numbers n, n', we have $\mathtt{L}_1(\mathcal{G}_n) \neq \mathtt{L}_1(\mathcal{G}_{n'})$. By Theorem 3.1, this guarantees that we have defined a countably infinite set of essentially different right linear grammar logics. Let $\mathtt{L}(\Box)$ be the standard modal language for the modal logic K. Let K-GSAT be the set of standard modal formulae ϕ such that there is a Kripke model $\mathcal{M} = \langle W, R, V\rangle$ satisfying $\mathcal{M} \models \phi$. The global satisfiability problem K-GSAT is known to be **EXPTIME**-hard (Chen and Lin 1994) (see also (Hemaspaandra 1996)). Let ϕ be a formula of $\mathtt{L}(\Box)$. One can show that ϕ belongs to K-GSAT iff $\bigwedge_{0 \leq \alpha \leq i-1} [2^\alpha \cdot 1]\phi'$ is $\mathcal{L}_m^{\mathcal{G}_i}$-satisfiable where ϕ' is obtained from ϕ by replacing every occurrence of $\Box$ by [2]. $\dashv$

Theorem 5.3 *GSP(RLIN) is* **EXPTIME***-complete.*

There is also a natural log-space transformation from GSP(RLIN) into satisfiability for the μ-calculus with multiple fixed points (see e.g. (Park 1981, Streett and Emerson 1989, Grädel et al. 2000). By way of example, the translation $t([2]\phi)$ of $[2]\phi$ for the right linear grammar logic $\mathcal{L}$ from Section 1 is

$$\nu X_2(X_1, X_2, X_3).\langle [1]X_3 \wedge [3]X_2, [2]X_3 \wedge [4]X_1 \wedge [5]X_3, t(\phi)\rangle$$

where X_3 is used as a renaming variable. The principle behind this example can be generalized to any modal connective of a right linear grammar logic. However, by using a polynomial-time transformation into satisfiability for PDL with finite automata one can show that GSP(REG), the extension of GSP(RLIN) with left linear grammars, is **EXPTIME**-complete (Demri 2000).

The general global logical consequence problem for right linear grammar logics GGLC(RLIN) takes as inputs a right linear grammar $\mathcal{G}$ and two $\mathtt{L}_m$-formulae ϕ, ψ and checks whether for all $\mathcal{L}_m^{\mathcal{G}}$-models $\mathcal{M}$, $\mathcal{M} \models \phi$ implies $\mathcal{M} \models \psi$. For any right linear grammar $\mathcal{G}$, GLC($\mathcal{L}_m^{\mathcal{G}}$) denotes the problem obtained from GGLC(RLIN) by fixing the grammar to $\mathcal{G}$. The calculus G$\mathcal{L}_m^{\mathcal{G}}$ can be extended in order to deal with GGLC(RLIN). The sequents are of the form $\Gamma \vdash_\phi \Delta$ for the $\mathtt{L}_m$-formula ϕ. $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ is extended by writing $\vdash_\phi$ instead of $\vdash$. However, one rule is added:

$$\frac{\Gamma, \phi \vdash_\phi \Delta}{\Gamma \vdash_\phi \Delta}\ glc_\phi$$

glc_ϕ is an obvious adaptation of existing rules for capturing global logical consequence (see e.g. (Fitting 1983, Heuerding 1998, Massacci 2000)). One can show that for any sequent $\Gamma \vdash_\phi \Delta$, the formula $\langle \phi, (\bigwedge_{\psi \in \Gamma} \psi) \Rightarrow (\bigvee_{\psi \in \Delta} \psi)\rangle \in \mathrm{GLC}(\mathcal{L}_m^{\mathcal{G}})$ iff the sequent $\Gamma \vdash_\phi \Delta$ is derivable in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}} + glc_\phi$. Similarly, one can show that GGLC(RLIN) is in **EXPTIME** by adapting the proof of Theorem 5.1.

6 A sufficient condition for PSPACE decision procedures

The completeness proof of Theorem 4.8 and the proof of Lemma 4.5 induce a depth-first systematic procedure to determine whether a sequent $\Gamma \vdash \Delta$ is provable in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ or not. In the proof of Lemma 4.5, in order to obtain a maximal sequence σ starting from a given sequent in a derivation $\vdash \phi$, essential backtracking points are introduced when the rule $(\vdash \wedge)$ needs to be applied. Moreover, the length of such a sequence σ is bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$. By Lemma 4.2 and since $card(\mathrm{cl}_{\mathcal{G}}(\phi))$ is bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$, the number of backtracking points is bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$. Similarly, in the proof of Theorem 4.8, other backtracking points are introduced when the rule $(\vdash [i])$ for some $i \in \{1, \ldots, m\}$ needs to be applied. The number of such backtracking points is also bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$. So one can use a bit-string of length $|\mathcal{G}| \times \mathrm{r}(\phi)$ to remember which choices have been already tried. Consequently, in order to show that a given right linear grammar logic $\mathcal{L}_m^{\mathcal{G}}$ has a polynomial space satisfiability problem (or equivalently a polynomial space validity problem), it is sufficient to show that there is a polynomial $p(.)$ such that in the depth-first systematic procedure for proving $\vdash \phi$, the $(\vdash [i])$ rules are applied on a branch at most $p(|\phi|)$ times (see also the proof of Lemma 6.1). This means that in the proof of Theorem 4.8 the number of levels is finite and is bounded by a polynomial in the size of the input sequent $\Gamma \vdash \Delta$.

This observation is not really surprising, but all the point now is to refine the above statement in order to be able to prove the polynomial space upper bounds for countably many right linear grammar logics.

Let $\langle S, \ll\rangle$ be a well-founded set, $meas : \mathrm{SEQ}_m \rightarrow S$ be a map and $p_{card}(.)$, $p_{length}(.)$ be polynomials such that for any $\mathtt{L}_m$-formula ϕ,

(C3) for $\Gamma_1 \vdash \Delta_1, \ldots, \Gamma_n \vdash \Delta_n \in \mathrm{SEQ}(\phi)$, $meas(\Gamma_n \vdash \Delta_n) \ll \ldots \ll$

$meas(\Gamma_1 \vdash \Delta_1) \ll meas(\vdash \phi)$ implies $n < p_{card}(|\phi|)$;

(C4) for any inference

$$\frac{\Gamma_1 \vdash \Delta_1 \quad \dots \quad \Gamma_l \vdash \Delta_l}{\Gamma \vdash \Delta} \ (r)$$

(a) with (r) different from $(\vdash [i])$ for all $i \in \{1, \dots, m\}$ ($l \in \{1, 2\}$ according to the form of the rule (r));

(b) $\Gamma_1 \vdash \Delta_1, \dots, \Gamma_l \vdash \Delta_l, \Gamma \vdash \Delta \in \mathrm{SEQ}(\phi)$;

we have for $i \in \{1, l\}$, either $meas(\Gamma_i \vdash \Delta_i) \ll meas(\Gamma \vdash \Delta)$ or $meas(\Gamma_i \vdash \Delta_i) = meas(\Gamma \vdash \Delta)$;

(C5) for any sequence of sequents in $\mathrm{SEQ}(\phi)$,

$$\sigma_0 \prec (\vdash [i_1])\Gamma_1 \vdash \Delta_1 \prec \sigma_1 \prec (\vdash [i_2]) \dots \sigma_{n-1} \prec (\vdash [i_n])\Gamma_n \vdash \Delta_n$$

such that for $i \in \{0, \dots, n-1\}$, σ_i is maximal and $n > p_{length}(|\phi|)$, we have $meas(\Gamma_n \vdash \Delta_n) \ll meas(first(\sigma_0))$.

The condition (C3) holds true when $\{meas(\Gamma \vdash \Delta) : \Gamma \vdash \Delta \in \mathrm{SEQ}(\phi)\}$ is finite since $\langle S, \ll \rangle$ is a well-founded set. For example, this is the case with $meas(\Gamma \vdash \Delta) = \mathrm{md}(\bigwedge_{\psi \in \Gamma, \Delta} \psi)$.

Lemma 6.1 *Let $\mathcal{G}$ be a right linear grammar. If there exist a well-founded set $\langle S, \ll \rangle$, a map meas $: \mathrm{SEQ}_m \to S$ and polynomials $p_{card}(.)$, $p_{length}(.)$ satisfying the conditions (C3), (C4) and (C5), for any $\mathtt{L}_m$-formula ϕ, then $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is in* **PSPACE**.

Proof. (sketch) The proof is by an easy verification by using the depth-first systematic procedure induced by the proofs of Lemma 4.5 and Theorem 4.8 in order to check whether $\vdash \phi$ is provable in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$. $\dashv$

The above method is not new (see e.g. (Ladner 1977, Fitting 1983, Massacci 1998)) but it allows us to use a uniform depth-first procedure from the proof of Theorem 4.8 and from the proof of Lemma 4.5.

Let $\mathcal{G}^{rl}$ be the right linear grammar defined in Example 4.1. Because of the duplication of Γ_1 in the application of the $(\vdash [4])$-rule, none of the obvious measures work to prove the polynomial space upper bound. Obviously, $\mathcal{G}^{rl}$ is finite since $\mathtt{L}_1(\mathcal{G}^{rl})$ and $\mathtt{L}_2(\mathcal{G}^{rl})$ are empty.

7 Finiteness implies PSPACE

This section is devoted to show that for any finite right linear grammar $\mathcal{G}$, the $\mathcal{L}_m^{\mathcal{G}}$-satisfiability problem is in **PSPACE** (forthcoming Theorem 7.4 is even a bit more general). We need to introduce a few definitions. Let $\mathcal{G}$ be a finite right linear grammar. The set N of non terminal symbols can be partionned into three sets N_1, N_2 and N_3 such that for $i \in N$,

- $i \in N_1 \overset{\text{def}}{\Leftrightarrow} \mathtt{L}_i(\mathcal{G}) = \emptyset$;

- $i \in N_2 \overset{\text{def}}{\Leftrightarrow} i \notin N_1$ and for some $j \in N_1$ and $u \in \Sigma^*$, $i \Rightarrow^*_{\mathcal{G}} u \cdot j$;
- $i \in N_3 \overset{\text{def}}{\Leftrightarrow} i \notin N_1 \cup N_2$.

N_1, N_2 and N_3 can be computed in polynomial-time in $|\mathcal{G}|$. A $N_2 \rightarrow N_1$**-derivation** is a sequence of the form

$$i_0 \Rightarrow_{\mathcal{G}} u_1 \cdot i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_{n-1} \cdot i_{n-1} \Rightarrow_{\mathcal{G}} u_1 \ldots u_n \cdot i_n$$

such that $n \geq 1$, $\{i_0, \ldots, i_{n-1}\} \subseteq N_2$ and $i_n \in N_1$. The main properties of the partition $\{N_1, N_2, N_3\}$ are the following.

Lemma 7.1 *Let $\mathcal{G}$ be a finite right linear grammar and $\{N_1, N_2, N_3\}$ be the partition on N defined above. Then,*

(I) if $i \Rightarrow^*_{\mathcal{G}} u$ with $i \in N_1$, then $u \in \Sigma^* \cdot N_1$;
(II) if $i_0 \Rightarrow_{\mathcal{G}} u_1 \cdot i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n \cdot i_n$ is a $N_2 \rightarrow N_1$-derivation, then
  1. *for $j \in \{0, \ldots, n-2\}$, $i_j \neq i_{j+1}$;*
  2. *$n \leq card(N_2) - 1$; $|u_1 \cdot \ldots \cdot u_n| \leq |\mathcal{G}|$;*

(III) if $i \Rightarrow^*_{\mathcal{G}} u$ with $i \in N_3$, then $u \in (\Sigma^* \cdot N_3) \cup \Sigma^*$;
(IV) if $i \Rightarrow^*_{\mathcal{G}} u$ with $i \in N_2 \cup N_3$ and $u \in \Sigma^*$, then $|u| \leq |\mathcal{G}|$.

For $i \in N \cup \Sigma$, we define $w_{\mathcal{G}}(i)$, the **weight** of i in $\mathcal{G}$, as follows:

- for $i \in N_1 \cup \Sigma$, $w_{\mathcal{G}}(i) \overset{\text{def}}{=} 1$;
- for $i \in N_3$, $w_{\mathcal{G}}(i) \overset{\text{def}}{=} max\{1 + |u| : u \in \mathtt{L}_i(\mathcal{G})\}$;
- for $i \in N_2$,

$$w_{\mathcal{G}}(i) \overset{\text{def}}{=} max(\{1 + |u| : u \in \mathtt{L}_i(\mathcal{G})\} \cup \{1 + |u_1 \cdot \ldots \cdot u_n| :$$
$$i \Rightarrow_{\mathcal{G}} u_1 i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n i_n \ N_2 \rightarrow N_1\text{-derivation}\}).$$

Observe that for $i \in N \cup \Sigma$, $w_{\mathcal{G}}(i) \leq |\mathcal{G}|$. Lemma 7.2 contains the main properties of the map $w_{\mathcal{G}}(.)$.

Lemma 7.2 *Let $\mathcal{G}$ be a finite right linear grammar. Then,*

(I) if $i \rightsquigarrow u \cdot i'$ with $i \in N_3$ and $u \in \Sigma^+$, then $i' \in N_3$ and $|u| - 1 + w_{\mathcal{G}}(i') < w_{\mathcal{G}}(i)$;
(II) if $i \rightsquigarrow u$ with $i \in N_3$ and $u \in \Sigma^+$, then $|u| < w_{\mathcal{G}}(i)$;
(III) if $i \rightsquigarrow u$ with $i \in N_2$ and $u \in \Sigma^+$, then $|u| < max\{1 + |v| : v \in \mathtt{L}_i(\mathcal{G})\} \leq w_{\mathcal{G}}(i)$;
(IV) if $i \rightsquigarrow u \cdot i'$ with $i \in N_2$, $i' \in N_3$ and $u \in \Sigma^+$, then $|u| - 1 + w_{\mathcal{G}}(i') < w_{\mathcal{G}}(i)$;
(V) if $i \rightsquigarrow u \cdot i'$ with $i, i' \in N_2$ and $u \in \Sigma^+$, then $|u| - 1 + w_{\mathcal{G}}(i') < w_{\mathcal{G}}(i)$.

Proof. By way of example, we show (V). Assume $i \rightsquigarrow u \cdot i'$ for some $i, i' \in N_2$ and $u \in \Sigma^+$. It is obvious that

$$max\{|v| + 1 : v \in \mathtt{L}_{i'}(\mathcal{G})\} + |u| - 1 < max\{|v| + 1 : v \in \mathtt{L}_i(\mathcal{G})\}.$$

It remains to show that

$$max\{1 + |u_1 \cdot \ldots \cdot u_n| : i' \Rightarrow_{\mathcal{G}} u_1 i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n i_n\} + |u| - 1$$
$$< max\{1 + |u_1 \cdot \ldots \cdot u_n| : i \Rightarrow_{\mathcal{G}} u_1 i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n i_n\}.$$

Since $i \rightsquigarrow u \cdot i'$, there are non terminal symbols $i_0, \ldots, i_\alpha$, $\alpha \geq 0$, such that $i_0 = i$ and $i_0 \Rightarrow_{\mathcal{G}} i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} i_\alpha \Rightarrow_{\mathcal{G}} u \cdot i'$. In the case all the non terminal symbols $i_0, \ldots, i_\alpha$ belongs to N_2 we can easily conclude the proof since

$$max\{1 + |u_1 \cdot \ldots \cdot u_n| : i \Rightarrow_{\mathcal{G}} u_1 i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n i_n\}$$
$$\geq |u| + max\{1 + |u_1 \cdot \ldots \cdot u_n| : i' \Rightarrow_{\mathcal{G}} u_1 i_1 \Rightarrow_{\mathcal{G}} \ldots \Rightarrow_{\mathcal{G}} u_1 \ldots u_n i_n\}.$$

Suppose $i_j \in N_3$ for some $j \in \{1, \ldots, \alpha\}$. If $i_j \in N_3$, then there is some $j' \in N_1$ such that $i_j \Rightarrow^*_{\mathcal{G}} v \cdot j'$, which leads to a contradiction by definition of the partition $\{N_1, N_2, N_3\}$. Suppose $i_j \in N_1$ for some $j \in \{1, \ldots, \alpha\}$. Hence $i_j \Rightarrow^*_{\mathcal{G}} u \cdot i'$ and $u \cdot i' \notin \Sigma^* \cdot N_1$ which is in contradiction with Lemma 7.1(I). $\dashv$

We define the maps pw^a : FORM $\rightarrow \mathbb{N}$ and pw^s : FORM $\rightarrow \mathbb{N}$ as follows ("a" stands for antecedent and "s" for succedent). Let ϕ be an $\mathtt{L}_m$-formula. Let $[i]\psi$ be the occurrence of a formula occurring negatively [resp. positively] in ϕ. To be precise, one should define the notion of occurrence (as a finite sequence of natural numbers for instance). For the sake of simplicity, this is omitted here. In order to define $\mathrm{pw}^a(\phi)$ [resp. $\mathrm{pw}^s(\phi)$], we define an auxiliary value $\mathrm{pw}^a(\phi, [i]\psi)$. Let $[i_1]\psi_1, \ldots, [i_n]\psi_n$, be the positive [resp. negative] occurrences of necessity formulae of ϕ such that the very occurrence of $[i]\psi$ is a subformula of each $[i_j]\phi_{i_j}$ (if any). In the case when some element of N_1 is in $\{i_1, \ldots, i_n\}$, $\mathrm{pw}^a(\phi, [i]\psi) \stackrel{\text{def}}{=} 0$ [resp. $\mathrm{pw}^s(\phi, [i]\psi) \stackrel{\text{def}}{=} 0$], otherwise $\mathrm{pw}^a(\phi, [i]\psi) \stackrel{\text{def}}{=} 1 + \sum_{j=1}^{n} w_{\mathcal{G}}(i_j)$ [resp. $\mathrm{pw}^s(\phi, [i]\psi) \stackrel{\text{def}}{=} 1 + \sum_{j=1}^{n} w_{\mathcal{G}}(i_j)$]. We are now in a position to define $\mathrm{pw}^a(\phi)$ [resp. $\mathrm{pw}^s(\phi)$]. $\mathrm{pw}^a(\phi) \stackrel{\text{def}}{=} \sum \mathrm{pw}^a(\phi, [i]\psi)$ [resp. $\mathrm{pw}^s(\phi) \stackrel{\text{def}}{=} \sum \mathrm{pw}^s(\phi, [i]\psi)$] where the sum is on the set of negative [resp. positive] occurrences of necessity formula in ϕ. For each $\psi \in \mathrm{cl}_{\mathcal{G}}(\phi)$, the number of negative occurrences of necessity formula is bounded by $|\mathcal{G}| \times \mathrm{r}(\phi)$ as well as the number of positive occurrences of necessity formula. So for $\psi \in \mathrm{cl}_{\mathcal{G}}(\phi)$,

$$max(\mathrm{pw}^a(\psi), \mathrm{pw}^s(\psi)) \leq |\mathcal{G}|^2 \times \mathrm{r}(\phi)^2 \times (|\mathcal{G}| + 1).$$

Theorem 7.3 *Let $\mathcal{G}$ be a finite right linear grammar logic. Then, $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is* **PSPACE**-*complete.*

Deciding whether $\mathcal{G}$ is a finite right linear grammar can be done in polynomial time in $|\mathcal{G}|$.

Proof. (sketch) Let us define S, $meas$, p_{card} and p_{length}. We shall write $\mathrm{md}_N(\Gamma \vdash \Delta)$ to denote the maximal nesting of modal connectives $[i]$ with $i \in N$ in $\bigwedge_{\psi \in \Gamma, \Delta} \psi$. Observe that for $\Gamma \vdash \Delta \in \mathrm{SEQ}(\phi)$, $\mathrm{md}_N(\Gamma \vdash \Delta) \leq \mathrm{r}(\phi)$. We write $\mathrm{pw}'(\Gamma \vdash \Delta)$ to denote

$$\mathrm{pw}'(\Gamma \vdash \Delta) \stackrel{\text{def}}{=} max(\{\mathrm{pw}^a(\psi) : \psi \in \Gamma\} \cup \{\mathrm{pw}^s(\psi) : \psi \in \Delta\}).$$

- $\langle S, \ll \rangle \stackrel{\text{def}}{=} \langle \mathbb{N}^2, < \rangle$ where $<$ is the standard lexicographical ordering on $\mathbb{N}^2$ extending the standard $<$ on $\mathbb{N}$;
- $meas(\Gamma \vdash \Delta) \stackrel{\text{def}}{=} \langle \mathrm{md}_N(\Gamma \vdash \Delta), \mathrm{pw}'(\Gamma \vdash \Delta) \rangle$;
- $p_{card}(x) = 1 + |\mathcal{G}|^2 \times (|\mathcal{G}| + 1) \times x^3$; $p_{length}(x) = 1$.

Let ϕ be an $\mathtt{L}$-formula. Condition (C3) is satisfied because

$$card(\{meas(\Gamma \vdash \Delta) : \Gamma \vdash \Delta \in \mathrm{SEQ}(\phi)\}) \leq |\mathcal{G}|^2 \times (|\mathcal{G}| + 1) \times |\phi|^3.$$

The condition (C4) is obviously satisfied. In order to check the condition (C5), by way of example consider an inference of a $(\vdash [i])$ rule for some $i \in \Sigma$ (see notations in Section 4.1). Since for $j \in N$ and $u \in (N \cup \Sigma)^*$, $j \rightsquigarrow u$ implies $u \in \Sigma^* \cup \Sigma^* \cdot N$, for $j \in \{1, \ldots, k\}$, $\mathrm{md}_N([j]\Gamma_j) \geq \mathrm{md}_N(\bigcup_{\langle j,k' \rangle \in start_i^j} [v_{j,k'}]\Gamma_j)$. Additionnally, $\mathrm{md}_N([i]\phi) = \mathrm{md}_N(\phi)$, which guarantees that $\mathrm{md}_N(.)$ does not strictly increase (when reading the rules upwards).

In order to show that $\mathrm{pw}'(.)$ strictly decreases, it is sufficient to see that $\mathrm{pw}^s([i]\phi) = 1 + \mathrm{pw}^s(\phi)$ and to check that

- for $j \in N_1$, $max(\{\mathrm{pw}^a(\psi) : \psi \in [j]\Gamma_j\}) = max(\{\mathrm{pw}^a(\psi) : \psi \in \bigcup_{\langle j,k' \rangle \in start_i^j} [v_{j,k'}]\Gamma_j\}) = 0$ by Lemma 7.1(I);
- for $j \in N_2$ [resp. $j \in N_3$],
$$max(\{\mathrm{pw}^a(\psi) : \psi \in [j]\Gamma_j\}) \geq$$
$$max(\{\mathrm{pw}^a(\psi) : \psi \in \bigcup_{\langle j,k' \rangle \in start_i^j} [v_{j,k'}]\Gamma_j\})$$
by Lemma 7.2(III,IV,V) [resp. by Lemma 7.2(I,II)].

$\dashv$

Theorem 7.4 below is slightly stronger than Theorem 7.3.

Theorem 7.4 *GSP(RLIN$_f$) is in* **PSPACE**.

The proof of Theorem 7.4 follows the line of the proof of Theorem 7.3 by observing that for any sequent $\Gamma \vdash \Delta \in \mathrm{SEQ}(\phi)$, $\mathrm{md}_N(\Gamma \vdash \Delta) \leq \mathrm{r}(\phi)$ and $\mathrm{pw}'(\Gamma \vdash \Delta) \leq |\mathcal{G}|^2 \times (|\mathcal{G}|+1) \times |\phi|^2$. Moreover, from the space analysis in Section 6, we can decide whether $\vdash \phi$ is provable in $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ in polynomial space in $|\mathcal{G}| + |\phi|$.

It seems difficult to extend Theorem 7.4 to a larger class of finite context-free grammars. Indeed, for the finite left linear grammar $\mathcal{G}^{ll} = \langle \{1\}, \{2\}, \{1 \rightarrow 1 \cdot 2\}, 1 \rangle$, $\mathcal{L}_2^{\mathcal{G}^{ll}}$-satisfiability is already **EXPTIME**-hard (see Theorem 7.5(1) below). By contrast, for any strongly finite context-free grammar $\mathcal{G}$, $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is in **PSPACE**. This can be shown by translation into PDL without Kleene star. By analogy to Theorem 7.4, it is open whether the general satisfiability problem for strongly finite context-free grammar logics is in **PSPACE**.

By using Theorem 7.3 and (Demri 2000), one can characterize the complexity of all the bimodal regular grammar logics.

Theorem 7.5 *Let $\mathcal{L}_2^{\mathcal{G}}$ be a bimodal regular grammar logic, that is $N \cup \Sigma = \{1, 2\}$ ($m = 2$) and $S = 1$.*

1. *If $1 \rightarrow 1 \cdot 2^i$ is a production rule of $\mathcal{G}$ for some $i \geq 1$, then, $\mathcal{L}_2^{\mathcal{G}}$-satisfiability is* **EXPTIME***-complete.*
2. *If $\mathcal{G}$ is right linear and finite [resp. infinite], then $\mathcal{L}_2^{\mathcal{G}}$-satisfiability is* **PSPACE***-complete [resp.* **EXPTIME***-complete].*

Theorem 7.5 exhausts all the possibilities of bimodal regular grammar logics. As a corollary, given a bimodal regular grammar logic $\mathcal{L}_2^{\mathcal{G}}$, deciding whether $\mathcal{L}_2^{\mathcal{G}}$ is **PSPACE**-complete can be done in linear time in $|\mathcal{G}|$. **EXPTIME**-hardness is roughly due to the fact that the concerned bimodal logics contain a modal connective [2] and another one [1] that is a variant of its reflexive and transitive closure (noted [2*]), see also (Spaan 1993, Sattler 1996, Castilho et al. 1999) for logics with a similar attribute. For instance, if $\mathcal{M}, x \models \langle 1 \rangle \phi \wedge [1]\psi$ for some $\mathcal{L}_2^{\mathcal{G}^{ll}}$-model $\mathcal{M}$, then there is $y \in R_1(x)$ such that $\mathcal{M}, y \models \phi$ and for all $z \in R_2^*(y)$, $\mathcal{M}, z \models \psi$, in symbols $\mathcal{M}, y \models \phi \wedge [2^*]\psi$. These are typically the formulae of that form that are responsible for the **EXPTIME**-hardness of PDL (Fischer and Ladner 1979, Spaan 1993). However, not every variant [1] of [2*] leads to **EXPTIME**-hardness. Consider the right linear grammar $\mathcal{G} = \langle \{1\}, \{2\}, \{1 \rightarrow 2 \cdot 1\}, 1 \rangle$ that is closely related to $\mathcal{G}^{ll}$. Although [2] [resp. [1]] can be viewed as the PDL modal connective $[\mathtt{c}_2]$ [resp. $[\mathtt{c}_2^*; \mathtt{c}_1]$], $\mathcal{L}_2^{\mathcal{G}}$ cannot isolate the modal connective $[\mathtt{c}_2^*]$ (or equivalently $[2^*]$) because in $[\mathtt{c}_2^*; \mathtt{c}_1]$, there is always a last step that is a $\mathtt{c}_1$ transition. Indeed, we have shown (see Theorem 7.3) that $\mathcal{L}_2^{\mathcal{G}}$-satisfiability is in **PSPACE**. By Theorem 3.2 we conclude that $\mathcal{L}_2^{\mathcal{G}}$-satisfiability is **PSPACE**-complete.

8 Infinity and PSPACE

Theorem 7.3 roughly states that finiteness for right linear grammars implies an **PSPACE** complexity upper bound for the corresponding grammar logics. By contrast, we can show that infinity does not imply **EXPTIME**-hardness. The class of right linear grammar logics introduced below contains for $m \geq 2$, countably infinite **PSPACE** logics $\mathcal{L}_m^{\mathcal{G}}$ with $N \cup \Sigma = \{1, \ldots, m\}$.

Theorem 8.1 *Let $\mathcal{G} = \langle N, \Sigma, P, S\rangle$ be a right linear grammar such that for $i \in N$, (1) $i \leadsto u$ implies $|u| \geq 1$ and (2) $i \leadsto j \cdot u$ and $i \leadsto j \cdot u'$ for some $j \in \Sigma$ imply $u = u'$. Then, $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is* **PSPACE***-complete.*

If $\mathcal{G}$ is viewed as a finite automaton $\mathcal{A}$, the assumption (2) in Theorem 8.1 can be interpreted as a requirement on the determinism of $\mathcal{A}$.

Proof. (sketch) Let us define S, $meas$, p_{card} and p_{length}:

- $\langle S, \ll\rangle \stackrel{\text{def}}{=} \langle \mathbb{N}, <\rangle$; $p_{card}(x) = x + 1$; $p_{length}(x) = 1$;
- for any sequent $\Gamma \vdash \Delta$, $meas(\Gamma \vdash \Delta) = \mathrm{pw}^+(\Gamma) + \mathrm{nw}^+(\Delta)$ (see Section 2 for the definitions of $\mathrm{pw}^+(.)$ and $\mathrm{nw}^+(.)$).

The assumptions of Lemma 6.1 can be shown to hold. $\dashv$

For any grammar $\mathcal{G}$ satisfying the assumptions of Theorem 8.1, the proof system $\mathrm{G}\mathcal{L}_m^{\mathcal{G}}$ involves no rule with implicit contraction. For $m \geq 5$, let $\mathcal{G} = \langle \{1, 2, 3\}, \{4, \ldots, m\}, P, 1\rangle$ be the infinite right linear grammar with P defined as the union of $\{1 \rightarrow 4^m\}$ with

$$\bigcup_{i \in \{5, \ldots, m\}} \{1 \rightarrow i^{(2^m)} \cdot 2, 2 \rightarrow (i \cdot (i+1) \cdot \ldots \cdot m)^i \cdot 3, 3 \rightarrow i^i \cdot (i+1)^{i+1} \cdot \ldots \cdot m^m 1\}$$

The grammar $\mathcal{G}$ is not finite but by Theorem 8.1, $\mathcal{L}_m^{\mathcal{G}}$-satisfiability is in **PSPACE**.

In Theorem 8.1, it is unlikely that one can significantly relax the condition $|u| \geq 1$ by allowing $|u| = 0$. By Theorem 7.5(2), $\mathcal{L}_2^{\mathcal{G}}$-satisfiability is **EXPTIME**-hard for the right linear grammar $\mathcal{G} = \langle \{1\}, \{2\}, \{1 \rightarrow \epsilon, 1 \rightarrow 2 \cdot 1\}, 1\rangle$.

9 An open problem

By a proof-theoretical analysis, we have designed polynomial space decision procedures in a uniform framework for countably infinite right linear grammar logics. Understanding the complexity/decidability status of context-free grammar logics by proof-theoretical means seems to be a challenge worth being attacked in order to further characterize the complexity of modal logics. More precisely, let $\mathcal{G}$ be a context-free grammar such that for $i \in N$, $\{u \in (\Sigma \cup N)^* : i \Rightarrow^*_{\mathcal{G}} u\}$ is a regular language. In (Demri 2000) it is shown that $\mathcal{L}^{\mathcal{G}}_m$-satisfiability can be polynomially reduced to PDL satisfiability by replacing any occurrence of $[i]$ by $[\pi_i]$ where π_i is a regular expression (program) generating precisely the language $\{u \in (\Sigma \cup N)^* : i \Rightarrow^*_{\mathcal{G}} u\}$. Hence, $\mathcal{L}^{\mathcal{G}}_m$-satisfiability is in **EXPTIME** and showing that $\mathcal{L}^{\mathcal{G}}_m$-satisfiability is in **PSPACE** is equivalent to showing that PDL satisfiability restricted to the program expressions $\pi_1, \ldots, \pi_m$ (if $\Sigma \cup N = \{1, \ldots, m\}$) is in **PSPACE**. Is there a natural class of **PSPACE** context-free grammar logics? K4, S4 should preferably fall into this hypothetical class.

Acknowledgments

I wish to express my gratitude to the anonymous referee for useful suggestions and remarks, to G. de Giacomo and E. Grädel for pointing me to works related to modal μ-calculus with simultaneous fixed points, to the participants of AiML-ICTL 2000 for their feedback on this work and to the AiML organizers for their kind invitation.

References

Andreka, H., I. Nemeti, and J. van Benthem. 1998. Modal languages and Bounded Fragments of Predicate Logic. *JPL* 27(3):217–274.

Baader, F., and U. Sattler. 2000. Tableau algorithms for Description Logics. In *TABLEAUX-9*, 1–18. LNAI 1847, Springer.

Baldoni, M. 1998. *Normal Multimodal Logics: Automated Deduction and Logic Programming.* Doctoral dissertation, Università degli Studi di Torino.

Baldoni, M., L. Giordano, and A. Martelli. 1998. A Tableau Calculus for Multimodal Logics and Some (Un)Decidability results. In *TABLEAUX-7*, 44–59. LNAI 1397, Springer.

Basin, D., S. Matthews, and L. Vigano. 1997. A new method for bounding the complexity of modal logics. In *Computational Logic and Proof Theory*, 89–102. LNCS 1289, Springer.

Basin, D., S. Matthews, and L. Viganò. 1998. Natural Deduction for Non-Classical Logics. *Studia Logica* 60(1):119–160.

Blackburn, P., M. de Rijke, and Y. Venema. 2001. *Modal Logic.* Cambridge University Press. to appear.

Castilho, M., O. Gasquet, and A. Herzig. 1999. Formalizing action and change in modal logic I: the frame problem. *JLC* 9(5):701–735.

Cerrito, S., and M. Cialdea Mayer. 1997. A Polynomial Translation of S4 into T and Contraction-free Tableaux for S4. *Logic Journal of the IGPL* 5(2):287–300.

Chagrov, A., and V. Shehtman. 1994. Algorithmic aspects of propositional tense logics. In *CSL-8*, 442–455. LNCS 933, Springer.

Chen, C., and I. Lin. 1994. The complexity of propositional modal theories and the complexity of consistency of propositional modal theories. In *LFCS-3*, 69–80. LNCS 813, Springer.

Demri, S. 2000. The complexity of regularity in grammar logics and related modal logics. Submitted.

Fariñas del Cerro, L., and M. Penttonen. 1988. Grammar logics. *Logique et Analyse* 123–134.

Fischer, M., and R. Ladner. 1979. Propositional Dynamic Logic of Regular Programs. *Journal of Computer and System Sciences* 18:194–211.

Fitting, M. 1983. *Proof methods for modal and intuitionistic logics.* D. Reidel Publishing Co.

Goré, R. 1999. Tableaux methods for modal and temporal logics. In *Handbook of Tableaux Methods*, ed. M. d'Agostino, D. Gabbay, R. Hähnle, and J. Posegga, 297–396. Kluwer.

Grädel, E. 1999. Why are modal logics so robustly decidable? *Bulletin of the EATCS* 68:90–103.

Grädel, E., C. Hirsch, and M. Otto. 2000. Back and forth between guarded and modal logics. In *LICS'2000*, 217–228.

Grädel, E., and I. Walukiewicz. 1999. Guarded Fixed Point Logic. In *LICS'99*, 45–54.

Halpern, J., and Y. Moses. 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *AI* 54:319–379.

Halpern, J., and H. Reif. 1983. The propositional dynamic logic of deterministic, well-structured programs. *TCS* 27:127–165.

Hemaspaandra, E. 1996. The price of Universality. *NDJFL* 37(2):173–203.

Heuerding, A. 1998. *Sequent Calculi for Proof Search in Some Modal Logics.* Doctoral dissertation, University of Bern.

Hopcroft, J., and J. Ullman. 1979. *Introduction to automata theory, languages, and computation.* Addison-Wesley Publishing Company, Reading, M.A.

Hudelmaier, J. 1996. Improved decision procedures for the modal logics K, T and S4. In *CSL'95*, 320–334. LNCS 1092, Springer.

Huldelmaier, J. 1993. An O(n log n)-space decision procedure for intuitionistic propositional logic. *JLC* 3(1):63–75.

Kozen, D. 1983. Results on the Propositional μ-calculus. *TCS* 27:333–354.

Kracht, M. 1996. Power and Weakness of the modal display calculus. In *Proof theory of modal logic*, ed. H. Wansing, 93–121. Kluwer.

Kracht, M. 2000. Logic and Syntax - A Personal Perspective. In *Advances in Modal Logics, Vol. 2*, ed. M. Zakharyaschev, K. Segerberg, M. de Rijke, and H. Wansing, 355–384. CSLI Publications.

Kripke, S. 1963. Semantical analysis of modal logic I: normal modal propositional calculi. *Zeitschrift für Mathematik Logik und Grundlagen der Mathematik* 9:67–96.

Ladner, R. 1977. The computational complexity of provability in systems of modal propositional logic. *SIAM Journal of Computing* 6(3):467–480.

Lutz, C., U. Sattler, and S. Tobies. 1999. A suggestion for an n-ary description logic. In *International Workshop on Description Logics, Sweden*, 81–85.

Marx, M. 1997. Complexity of Modal Logics of Relations. Technical report. ILLC.

Marx, M., S. Mikulas, and M. Reynolds. 2000a. The Mosaic Method for Temporal Logics. In *TABLEAUX-9*, 324–340. LNAI 1847, Springer.

Marx, M., S. Mikulas, and S. Schlobach. 2000b. Labelled Deduction for the guarded fragment. In *Labelled Deduction*, ed. D. Basin, M. D'Agostino, D. Gabbay, and S. Matthews, 195–214. Kluwer.

Massacci, F. 1994. Strongly analytic tableaux for normal modal logics. In *CADE-12*, 723–737. LNAI 814, Springer.

Massacci, F. 1998. *Efficient Approximate Deduction and an Application to Computer Security*. Doctoral dissertation, Università Degli Studi Di Roma "'La Sapienza".

Massacci, F. 2000. Single Steps Tableaux for Modal Logics. *JAR* 24(3):319–364.

Ohnishi, M., and K. Matsumoto. 1957. Gentzen Method in modal calculi. *Osaka Mathematical Journal* 9:113–130.

Park, D. 1981. Concurrency and Automata on infinite sequences. In *5th GI Conference on Theoretical Computer Science*, 167–183. LNCS 104, Springer.

Pratt, V. 1979. Models of program logics. In *Proceedings of the 20th IEEE Symposium on Foundations of Computer Science*, 115–122.

Rabin, M. 1969. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society* 41:1–35.

Rautenberg, W. 1983. Modal Tableau Calculi and Interpolation. *JPL* 12:403–423.

Rozenberg, G., and A. Salomaa. 1994. *Cornerstones of Undecidability*. Prentice Hall.

Sahlqvist, H. 1975. Completeness and correspondence in the first and second order semantics for modal logics. In *3rd Scandinavian Logic Symposium, 1973*, 110–143. North Holland.

Sattler, U. 1996. A concept language extended with different kinds of transitive roles. In *20. Deutsche Jahrestagung für Künstliche Intelligenz*. LNM 1137, Springer.

Schütte, K. 1967. *Proof theory*. Springer-Verlag.

Spaan, E. 1993. *Complexity of Modal Logics*. Doctoral dissertation, ILLC, Amsterdam University.

Streett, R., and A. Emerson. 1989. An automata theoretic decision procedure for the propositional μ-calculus. *Information and Computation* 81:249–264.

Szałas, A. 1996. On natural deduction in first-order fixpoint logics. *Fundamenta Informaticae* 26(1):81–94.

Takeuti, G. 1975. *Proof Theory*. North-Holland Publishing Company.

van Benthem, J. 1976. Modal Reduction principles. *JSL* 2:301–312.

van Benthem, J. 1984. Correspondence Theory. In *Handbook of Philosophical Logic, Volume II*, ed. D. Gabbay and F. Günthner, 167–247. Reidel, Dordrecht.

van Benthem, J. 2000. Modal Logic in Two Gestalts. In *Advances in Modal Logics, Vol. 2*, ed. M. Zakharyaschev, K. Segerberg, M. de Rijke, and H. Wansing, 91–118. CSLI Publications.

Vardi, M. 1997. Why is modal logic so robustly decidable? In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science 31, American Mathematical Society*, 149–183.

Stéphane Demri
Lab. Spécification et Vérification
ENS de Cachan & CNRS UMR 8643
61 Av. Pdt. Wilson
94235 Cachan Cedex, France
email: `demri@lsv.ens-cachan.fr`
URL: `http://www.lsv.ens-cachan.fr/~demri/`
On leave from Lab. LEIBNIZ, Grenoble.