# Using Grammars for Finite Domain Evaluation 

Robert Matzinger*<br>Technische Universität Wien

October 7, 1997


#### Abstract

In [8] we investigated representing Herbrand models via context-free grammars and found the representation power of this method to be exactly the finite models. Based on these observations we now present a clause set evaluation algorithm that operates directly on grammars, avoiding the exponential blow-up from the number of nonterminals in the grammar to the number of elements in the finite domain of the corresponding model, ending up with a not-so-obvious evaluation procedure for arbitrary clause sets over finite interpretations (specified via grammars).


## 1 Introduction

Currently there is considerable ongoing work in the field of automated model building (see e.g. $[2-5,10,11]$ ), a subfield of automated theorem proving that attempts to design algorithms for finding models of satisfiable first order logic formulas. Note that a model of a formula $\neg A$ is nothing else than a counterexample of $A$, which may be of enormous help for finding the reason why $A$ failed to be valid. Clearly, the ability to represent models of first order logic formulas in a computationally feasible way is a necessary prerequisite for trying to build them. However, representation mechanisms for particular models of first order logic formulas play an important role in many other fields too, e.g. semantic resolution, model checking, etc. That's why we considered it worthwhile to investigate mechanisms for representing interpretations of first order logic formulas symbolically by their own, aiming to accompany work in the automated model building field.

From a theoretical point of view it makes sense to restrict ourselves to Herbrand models of skolemized formulas (i.e. clause sets), but this is also justified from a practical viewpoint to utilize the intuitive requirement of understandability, because in Herbrand models the domain and the interpretation of the function symbols are clear, fixed and intuitive. Still we know there exists a Herbrand model for any (skolemized) satisfiable formula.

To describe a certain Herbrand model over a fixed signature we just have to find a method for specifying potentially infinite sets of (true) ground atoms, i.e. sets of terms (or strings, depending on how we want to look at them). This viewpoint reveals the (to our opinion) most interesting aspect of our approach: We are lead to investigating model properties in terms of syntactical properties of the corresponding true-ground-atom set.

In $[7,8]$ we investigated the use of context-free grammars to represent the true ground atom set of an interpretation ${ }^{1}$ in the following sense: Consider a fixed signature $\Sigma$.

[^0]$\mathrm{HD}_{\Sigma}$ denotes the set of terms over this signature (i.e. the Herbrand domain). Let $G=$ $\left(\mathcal{A}_{\Sigma}, \mathcal{P} \cup \mathcal{N}, \Pi\right)$ be a grammar where $\mathcal{A}_{\Sigma}$ is the set of constant symbols plus the function symbols plus the symbols ${ }^{\prime}\left({ }^{\prime},{ }^{\prime}\right)^{\prime}$, and ${ }^{\prime},{ }^{\prime} . \mathcal{P}$ is the set of predicate symbols in the signature (used as nonterminals in the grammar), $\mathcal{N}$ is an appropriate set of additional nonterminal symbols and $\Pi$ is a set of productions. We omit the notation of a starting symbol with the grammar, writing it as an index of the language instead, e.g. $L_{P}$ denotes the language derived by $G$ with starting symbol $P$. We may restrict ourselves to flat term grammars i.e. grammars in which each production is either of the form $N \longrightarrow a$ with $a$ being a constant or is of the form $N \longrightarrow f\left(N_{1}, \ldots, N_{n}\right)$ with $f$ being a function symbol of arity $n$ and $N, N_{1}, \ldots, N_{n} \in \mathcal{P} \cup \mathcal{N}$ being nonterminals. Note that we do not lose generality, because any context-free ${ }^{2}$ subset of $\mathrm{HD}_{\Sigma}$ can be generated also by a flat term grammar, a fact we proved in [9].

Now we can simply define the Herbrand interpretation $\Gamma_{G}$ by stating that a ground atom $P(t)$ shall be interpreted true iff $t \in L_{P}$. We call any Herbrand interpretation a context-free interpretation (or context-free model) iff we can define it via a grammar in the above sense. Clearly flat term grammars are just another notation for indeterministic bottom-up tree automata (see [6]), which (beside other desirable properties) can always be translated to deterministic ones. This and the fact that regular tree languages are closed under inversion, intersection and complement enables to show that
Theorem. For any context-free model of a skolemized monadic formula (with arbitrary function symbols) there is an equivalent finite model and every finite model contains a submodel that is equivalent to a context-free model.
Corollary. For any finite model of a skolemized formula (with arbitrary function symbols) the set of true ground atoms is a context-free language. A formula has a finite model iff it has a context-free one.

So we can cover exactly the finite models with this representation mechanism, that well satisfies the requirements for model representations as raised in [4]: The atom test (i.e. evaluating $P(t)$ for a ground term $t$ ) is nothing else than a syntax check for $t$ with respect to the term grammar, which can be solved in linear time (w.r.t. the length of the atom). The equivalence check (i.e. whether or not two grammars represent the same model) can be shown to be decidable with some knowledge in tree automata theory, but can also be solved e.g. by a translation to clause sets and resolution refinements (see [8]). Note also that in many practical situations specifying a model by a recursive definition of properties (under the closed world assumption), which is nothing else than a grammar's rule set, may be preferable among the explicit notation of a much bigger finite model. Still there's a very natural relation between the (indeterministic) grammar and the finite model. Although the equivalence with a finite model does yield a clause set evaluation procedure by enabling us to construct the equivalent finite model and evaluate all possible instances of every clause $C$ in a clause set $\mathcal{C}$, this procedure is extremely inefficient due to the exponential blow-up from grammar nonterminals to finite domain members. ${ }^{3}$ That's
of arbitrary arity). Note that this is no loss of generality, as we always get a sat-equivalent formula via the translation $P\left(t_{1}, \ldots, t_{n}\right)$ to $T\left(p\left(t_{1}, \ldots, t_{n}\right)\right.$ ) ( $p$ being a new n-place function symbol for every predicate $P$ and $T$ being an entirely new predicate). Additionally any model of the $T(p(\ldots))$-formula can be trivially translated to one of the original formula.
${ }^{2}$ i.e. generated by an arbitrary context-free (string) grammar. Example: The language $\{f(a)\}$ generated by $S \longrightarrow A B, A \longrightarrow f(, B \longrightarrow a)$ can also be generated by $S \longrightarrow f(Q), Q \longrightarrow a$.
${ }^{3}$ Recall that in the usual construction there are exponentially many nonterminals in the deterministic grammar and there may be exponentially many finite domain members than there were nonterminals in the original grammar.
why developing more efficient procedures remained an interesting target.

## 2 Clause Evaluation On Context-Free Interpretations

To avoid the explicit construction of the finite model we asked ourselves whether we could evaluate clause sets directly on the grammar. Difficulties arise from the fact that the grammar resembles an indeterministic tree automaton, for which we want to avoid explicit determination. However we can utilize the connections between syntactical properties of the interpretation and the truth of formulas to design the following evaluation procedure:

Let $G=\left(\mathcal{A}_{\Sigma}, \mathcal{P} \cup \mathcal{N}, \Pi\right)$ be a fixed flat term grammar. $\Pi[P, f]$ is defined to be the set of productions $\subseteq \Pi$ that have the form $P \longrightarrow f(\ldots)$, where $P \in \mathcal{P} \cup \mathcal{N}$ is a nonterminal and $f$ is a function symbol. We analogously define $\Pi[P, a]$ for constant symbols $a$.

Now let's start with a clause set $\mathcal{C}$ for which we want to know whether it evaluates to true in the model specified by the grammar $G$ (in the sense described before) or not, i.e. we want to know whether $\Gamma_{G} \models \mathcal{C}$. Now observe that if $\Pi[P, f]=$ $\left\{P \longrightarrow f\left(Q_{1}^{1}, Q_{2}^{1}, \ldots Q_{n}^{1}\right), \ldots, P \longrightarrow f\left(Q_{1}^{k}, Q_{2}^{k}, \ldots Q_{n}^{k}\right)\right\}$ this means that $\Gamma_{G} \models P\left(f\left(x_{1}, x_{2}\right.\right.$, $\left.\left.\ldots x_{n}\right)\right) \leftrightarrow\left(\left(Q_{1}^{1}\left(x_{1}\right) \wedge \ldots \wedge Q_{n}^{1}\left(x_{n}\right)\right) \vee \ldots \vee\left(Q_{1}^{k}\left(x_{1}\right) \wedge \ldots \wedge Q_{n}^{k}\left(x_{n}\right)\right)\right)$. We can denote this formula negatively too: $\Gamma_{G} \models \neg P\left(f\left(x_{1}, x_{2}, \ldots x_{n}\right)\right) \leftrightarrow\left(\left(\neg Q_{1}^{1}\left(x_{1}\right) \vee \ldots \vee \neg Q_{n}^{1}\left(x_{n}\right)\right) \wedge \ldots \wedge\right.$ $\left.\left(\neg Q_{1}^{k}\left(x_{1}\right) \vee \ldots \vee \neg Q_{n}^{k}\left(x_{n}\right)\right)\right)$. Now take these formulas as rewrite rules (to be used from the left to the right) that let us reduce the literals in a clause set $\mathcal{C}$. (Clearly we have to use the law of distributivity after every application of a rewrite rule to get a clause set again). We write $\mathcal{C} \mapsto \mathcal{C}^{\prime}$, if $\mathcal{C}^{\prime}$ is the clause form of the result of applying one of above mentioned rewrite rules to $\mathcal{C}$. Clearly $\Gamma_{G} \models \mathcal{C} \Leftrightarrow \Gamma_{G} \models \mathcal{C}^{\prime}$. As we reduce the term-depth of the literals in each step, we must reach an irreducible form when repeatedly applying the rules. Since we have only one applicable rule for every literal this irreducible form is unique. ${ }^{4}$ Let's call it $\mathcal{C}^{0}$. Because every literal in $\mathcal{C}^{0}$ is irreducible, it is either of the form $(\neg) P(x)$ where $x$ is a variable or $(\neg) P(a)$ where $a$ is a constant symbol. Clearly we can decide whether $\Gamma_{G}=\mathcal{C}^{0}$, if we can decide whether $\Gamma_{G}=C$ for a single clause $C \in \mathcal{C}^{0}$. Let's write $C[x]$ for the set of all literals on the variable $x$ in $C$. $C$ [const] shall denote the set of ground literals in $C$. Thus $C=C\left[x_{1}\right] \vee C\left[x_{2}\right] \vee \ldots \vee C\left[x_{m}\right] \vee C[$ const $]$. Clearly $\Gamma_{G}=C$ iff $\Gamma_{G}$ entails at least one of the $C\left[x_{i}\right]$ or $\Gamma_{G}$ entails at least one of the literals in $C[$ const $]$. Now recall the definition of $\Gamma_{G}$ to see that a ground literal $P(a)$ is true in $\Gamma_{G}$ iff $a \in L_{P}$, which can easily be checked in the grammar. Thus it remains to evaluate whether $\Gamma_{G} \vDash C[x]$. Observe that $C[x]$ corresponds to the formula $\forall x C[x]$. Thus $C[x]$ evaluates to true iff every ground term makes at least one literal true in $C[x]$. In general some literals are positive and some are negative in $C[x]$, i.e. $C[x]=\neg P_{1}(x) \vee \ldots \vee \neg P_{l}(x) \vee P_{l+1}(x) \vee \ldots \vee P_{m}(x)$. According to the definition of the interpretation $\Gamma_{G}$ every ground term makes at least one of the $(\neg) P_{i}(x)$ true iff $\left(\mathrm{HD}_{\Sigma} \backslash L_{P_{1}}\right) \cup \ldots \cup\left(\mathrm{HD}_{\Sigma} \backslash L_{P_{l}}\right) \cup L_{P_{l+1}} \cup \ldots \cup L_{P_{m}}=\mathrm{HD}_{\Sigma}$, which is the case iff $L_{P_{1}} \cap \ldots \cap L_{P_{l}} \subseteq L_{P_{l+1}} \cup \ldots \cup L_{P_{m}}$. If there are either no negative or no positive literals the situation is pretty much the same: the corresponding conditions are $\mathrm{HD}_{\Sigma} \subseteq L_{P_{1}} \cup \ldots \cup L_{P_{m}}$ or $L_{P_{1}} \cap \ldots \cap L_{P_{m}} \subseteq \emptyset$, respectively.
Example. Let the Herbrand interpretation $\Gamma_{G}$ be given by the following grammar: $G=\left\{\left\{f, a,^{\prime}\left({ }^{\prime},{ }^{\prime}\right)^{\prime},^{\prime}{ }^{\prime},\right\},\{P, Q\},\{P \longrightarrow f(Q), Q \longrightarrow f(P), P \longrightarrow a\}\right\}$, i.e. the interpretation where $\left\{P\left(f^{2 i}(a)\right), Q\left(f^{2 i+1}(a)\right) \mid i \geq 0\right\}$ is true, the rest is false. Assume we want to evaluate $C=P(x) \vee P(f(x))$ in this interpretation. The grammar rule $P \longrightarrow f(Q)$ corresponds to the fact that $P(f(x)) \leftrightarrow Q(x)$, which we use as a rewrite rule (from left to right)

[^1]on $P(f(x))$ to obtain $C^{0}=P(x) \vee Q(x)$. Thus $C$ evaluates to true iff $L_{P} \cup L_{Q}=\mathrm{HD}_{\Sigma}$. We enhance the grammar by the rules $T \longrightarrow a, T \longrightarrow f(T)$ (i.e. $L_{T}=\mathrm{HD}_{\Sigma}$ ) and check for $L_{T} \subseteq L_{P} \cup L_{Q}$. On the other hand if we want to evaluate the clause $C^{\prime}=P(x) \vee P\left(f^{2}(x)\right)$, we end up with $C^{\prime 0}=P(x) \vee P(x)$, so we have to check whether or not $L_{T} \subseteq L_{P}$.

Thus we are done, if we have procedures that decide (efficiently) whether or not $L_{P_{1}} \cap$ $\ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$ for regular tree languages $L_{P_{i}}, L_{Q_{j}}$. It is well known that this problem is decidable, but we are interested particularly in a solution that avoids the exponential blow-up of the state-space that comes from determination of the corresponding tree automata. In the next section we give an algorithm, that works directly on the indeterministic automata (grammar).

## 3 Calculating with Regular Tree Languages

So we need a procedure to decide for a given flat term grammar $G$ whether or not $L_{P_{1}} \cap$ $\ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$. The idea of the procedure is to algorithmically construct a "proof by induction on term depth" for the truth or falsehood of the questioned property. We use the following rules to unfold expressions of the form $L_{P} \cap \ldots \subseteq L_{Q} \cup \ldots$ :
(i): $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}} \Leftrightarrow \wedge_{f} L_{P_{1}}[f] \cap \ldots \cap L_{P_{n}}[f] \subseteq L_{Q_{1}}[f] \cup \ldots \cup L_{Q_{m}}[f]$, where $L[f]$ denotes the restriction of the language $L$ to $f$-terms and $\wedge_{f}$ ranges over all function symbols and constant symbols,
(ii): If $\Pi[P, f]=\left\{P \longrightarrow f\left(P_{1}, \ldots\right), \ldots, P \longrightarrow f\left(P_{n}, \ldots\right)\right\}$, then $L_{P}[f]=f\left(L_{P_{1}}, \ldots\right) \cup$ $\ldots \cup f\left(L_{P_{n}}, \ldots\right){ }^{5}$
(iii): $L_{A} \cup L_{B} \subseteq L_{C} \Longleftrightarrow L_{A} \subseteq L_{C} \wedge L_{B} \subseteq L_{C}$ and
(iv): $f\left(L_{A}, \ldots\right) \cap f\left(L_{B}, \ldots\right)=f\left(L_{A} \cap L_{B}, \ldots\right)$,

Clearly these rules are correct. See that they enable us to unfold $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq$ $L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$ to a boolean formula consisting only of subformulas of the form $f\left(L_{A_{1}} \cap\right.$ $\left.\ldots \cap L_{A_{k}}, \ldots, L_{Z_{1}} \cap \ldots \cap L_{Z_{k}}\right) \subseteq f\left(L_{A_{1}^{\prime}}, \ldots, L_{Z_{1}^{\prime}}\right) \cup \ldots \cup f\left(L_{A_{l}^{\prime}}, \ldots, L_{Z_{l}^{\prime}}\right)$, connected by $\wedge$ and $\vee$.

We start our procedure with an empty set $H$ of induction hypothesis on an expression $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$. First we unfold it according to above rule set. We get a boolean expression containing subformulas of the type $f(\ldots) \subseteq f(\ldots) \cup \ldots \cup f(\ldots)$. In the simplest case the subformula is just $f\left(L_{A}, \ldots, L_{Z}\right) \subseteq f\left(L_{A^{\prime}}, \ldots, L_{Z^{\prime}}\right)$. We solve this by recursively applying our method on each $L_{A} \subseteq L_{A^{\prime}}, \ldots, L_{Z} \subseteq L_{Z^{\prime}}$, with the enhanced induction hypothesis $H \cup\left\{L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}\right\}$ (because we stepped one step downward in the term depth; this corresponds to the induction step). Any more complicated case of $f\left(\ldots L_{A_{i}} \ldots\right) \subseteq f\left(\ldots L_{B_{j}} \ldots\right) \cup \ldots \cup f\left(\ldots L_{C_{k}} \ldots\right)$ is similarly solved by reducing it to a boolean expression over language inclusion properties of the $L_{A_{i}}, L_{B_{j}}, L_{C_{k}}$. See appendix Appendix A for technical details. We can once again use our idea of "stepping downward" and reapply our procedure on the subproblems with an enhanced induction hypothesis set. We can stop recursing and return true or false if we come down to constant symbols (this check is trivial; it corresponds to the induction base), or we can return true if the inclusion we have to test is a consequence of the induction hypothesis, i.e. if we have to test whether $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$ and there's already a formula $L_{P_{1}} \cap \ldots \cap L_{P_{i}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{j}}$ in the induction hypothesis set $H$ for some $i \leq n, j \leq m$.

[^2]So regardless whether $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq L_{Q_{1}} \cup \ldots \cup L_{Q_{m}}$ is true or false, the procedure's run can be seen as an inductive proof that this is the case, from which we can conclude the correctness of the procedure. However, as our grammar has only finitely many nonterminal symbols, there are only finitely many possible induction hypothesis. As we always enlarge the set of possible induction hypothesis, we can guarantee termination of the procedure. Note that evaluations in the different recursion branches are totally independent, giving way for various search algorithms and heuristics as well as for parallelization. Note furthermore that the grammar we evaluate upon remains fixed, so every proposition about it, once proven, remains true. It's an obvious and substantial improvement to maintain a database of $L_{P} \cap \ldots \subseteq L_{Q} \cup \ldots$-propositions that are already proven and to cut off recursive calls that would otherwise just re-prove known facts. In many practical examples the algorithm terminates quickly on the constant symbol check. However a worst-case exponential blow-up seems to be unavoidable due to the inherent difficulty of the problem - note that the language inclusion problem "is $\Sigma^{*} \subseteq L_{R}$ ?" for regular expressions $R$ (which resemble grammars with only unary function symbols) is PSPACE-complete (see e.g. [1]) and the language intersection nonemptiness problem for regular tree languages (i.e. whether $L_{P_{1}} \cap \ldots \cap L_{P_{n}} \subseteq \emptyset$ ) was shown to be EXPTIME-complete (e.g. in [12]).

Example. (continued) We denote our procedure as $L_{A} \nsucceq L_{B} \mid H$, if we check whether or not $L_{A} \subseteq L_{B}$ under the induction hypothesis $H$. To check whether or not $L_{T} \subseteq L_{P} \cup L_{Q}$ we call $L_{T} \overparen{Z} L_{P} \cup L_{Q}\left|\emptyset=f\left(L_{T}\right) \nsucceq f\left(L_{Q}\right) \cup f\left(L_{P}\right)\right| \emptyset \wedge\{a\} ?\{a\} \mid \emptyset=$ $=L_{T} \overparen{?} L_{Q} \cup L_{P} \mid\left\{L_{T} \subseteq L_{P} \cup L_{Q}\right\} \wedge$ true $=$ true $\wedge$ true $=$ true. On the other hand if we want to check $L_{T} \subseteq L_{P}$, we get $L_{T} \nsupseteq L_{P}\left|\emptyset=f\left(L_{T}\right) \nsucceq f\left(L_{Q}\right)\right| \emptyset \wedge\{a\} \nsucceq\{a\}\left|\emptyset=L_{T} \nsucceq L_{Q}\right|\left\{L_{T} \subseteq L_{P}\right\} \wedge$ true $=$ $=f\left(L_{T}\right) \nsucceq f\left(L_{P}\right)\left|\left\{L_{T} \subseteq L_{P}\right\} \wedge\{a\} \nsucceq \emptyset\right|\left\{L_{T} \subseteq L_{P}\right\} \wedge$ true $=$
$=L_{T} \nsupseteq L_{P} \mid\left\{L_{T} \subseteq L_{P}, L_{T} \subseteq L_{Q}\right\} \wedge$ false $\wedge$ true $=$ true $\wedge$ false $\wedge$ true $=$ false.
With this algorithm we conclude our evaluation procedure for clause sets on a given context-free interpretation. Knowing that every context-free interpretation is equivalent to a finite one and vice-versa, we got an interesting and not-so-obvious method for evaluating arbitrary clause sets on finite interpretations. Furthermore we think that our algorithm will be of benefit for a complexity analysis of the problem of interpreting clauses over context-free interpretations because of its close relation to regular tree languages for which many complexity results are known.

## Appendix A

We illustrate that any expression of the form $f\left(\ldots L_{A_{i}} \ldots\right) \subseteq f\left(\ldots L_{B_{j}} \ldots\right) \cup \ldots \cup f\left(\ldots L_{C_{k}} \ldots\right)$ can be reduced to a boolean expression over language inclusion properties of the $L_{A_{i}}, L_{B_{j}}, L_{C_{k}}$ by solving the following example: We show that $f\left(L_{A}, L_{A^{\prime}}\right) \subseteq f\left(L_{C}, L_{C^{\prime}}\right) \cup f\left(L_{D}, L_{D^{\prime}}\right)$ iff $\left(L_{A} \subseteq L_{C} \cup L_{D} \vee L_{A^{\prime}}=\emptyset\right) \wedge\left(L_{A} \subseteq L_{D} \vee L_{A^{\prime}} \subseteq L_{C^{\prime}}\right) \wedge\left(L_{A} \subseteq L_{C} \vee L_{A^{\prime}} \subseteq L_{D^{\prime}}\right) \wedge\left(L_{A}=\right.$ $\left.\emptyset \vee L_{A^{\prime}} \subseteq L_{C^{\prime}} \cup L_{D^{\prime}}\right)$.
Proof. $f\left(L_{A}, L_{A^{\prime}}\right) \subseteq f\left(L_{C}, L_{C^{\prime}}\right) \cup f\left(L_{D}, L_{D^{\prime}}\right)$ holds iff for all ground terms $t, t^{\prime}$ it holds that $t \in L_{A} \wedge t^{\prime} \in L_{A^{\prime}} \rightarrow\left(t \in L_{C} \wedge t^{\prime} \in L_{C^{\prime}}\right) \vee\left(t \in L_{D} \wedge t^{\prime} \in L_{D^{\prime}}\right)$.
For notational convenience let us abbreviate "for all ground terms $t$ it holds that" by $\forall t$ and let's write $X$ for $t \in L_{X}$ and $X^{\prime}$ for $t^{\prime} \in L_{X^{\prime}}$. Thus the previous proposition reads $\forall t \forall t^{\prime} A \wedge A^{\prime} \rightarrow\left(C \wedge C^{\prime}\right) \vee\left(D \wedge D^{\prime}\right)$, which is the case iff $\forall t \forall t^{\prime} \neg A \vee \neg A^{\prime} \vee\left(C \wedge C^{\prime}\right) \vee\left(D \wedge D^{\prime}\right)$, which we can transform to the equivalent expression $\forall \forall \forall t^{\prime}\left(\neg A \vee \neg A^{\prime} \vee C \vee D\right) \wedge\left(\neg A \vee \neg A^{\prime} \vee C \vee D^{\prime}\right) \wedge\left(\neg A \vee \neg A^{\prime} \vee C^{\prime} \vee D\right) \wedge\left(\neg A \vee \neg A^{\prime} \vee C^{\prime} \vee D^{\prime}\right)$.

We get an equivalent expression when we shift the quantifiers over the $\wedge$ and $\vee$, so we obtain $\left(\left(\forall t^{\prime} \neg A^{\prime}\right) \vee(\forall t \neg A \vee C \vee D)\right) \wedge\left((\forall t \neg A \vee C) \vee\left(\forall t^{\prime} \neg A^{\prime} \vee D^{\prime}\right)\right) \wedge\left((\forall t \neg A \vee D) \vee\left(\forall t^{\prime} \neg A^{\prime} \vee\right.\right.$ $\left.\left.C^{\prime}\right)\right) \wedge\left((\forall t \neg A) \vee\left(\forall t^{\prime} \neg A^{\prime} \vee C^{\prime} \vee D^{\prime}\right)\right)$,
which expresses nothing else than $\left(L_{A^{\prime}}=\emptyset \vee L_{A} \subseteq L_{C} \cup L_{D}\right) \wedge\left(L_{A} \subseteq L_{C} \vee L_{A^{\prime}} \subseteq\right.$ $\left.L_{D^{\prime}}\right) \wedge\left(L_{A} \subseteq L_{D} \vee L_{A^{\prime}} \subseteq L_{C^{\prime}}\right) \wedge\left(L_{A}=\emptyset \vee L_{A^{\prime}} \subseteq L_{C^{\prime}} \cup L_{D^{\prime}}\right)$.

## References

[1] A.V. Aho, J.E. Hopcroft, and J.D. Ullman. The design and analysis of computer algorithms. AddisonWesley, Reading, Ma., 1974.
[2] R. Caferra and N. Peltier. Decision procedures using model building techniques. In Computer Science Logic (9th Int. Workshop CSL'95), pages 131-144, Paderborn, Germany, 1995. Springer Verlag. LNCS 1092.
[3] R. Caferra and N. Zabel. A method for simultanous search for refutations and models by equational constraint solving. Journal of Symbolic Computation, 13(6):613-641, June 1992.
[4] C. Fermüller and A. Leitsch. Hyperresolution and automated model building. J. of Logic and Computation, 6(2):173-203, 1996.
[5] C. Fermüller and A. Leitsch. Decision procedures and model building in equational clause logic. Journal of the IGPL, 1997. to appear.
[6] F. Gécseg and M. Steinby. Tree Automata. Akadémiai Kiadó, Budapest, 1984.
[7] R. Matzinger. Comparing computational representations of Herbrand models. In A. Leitsch, editor, Computational Logic and Proof Theory, 5th Kurt Gödel Colloquium, KGC'97, volume 1289 of $L N C S$, pages 203-218, Vienna, 1997. Springer.
[8] R. Matzinger. Computational representations of Herbrand models using grammars. In D.v. Dalen, editor, Computer Science Logic, 10th International Workshop, CSL'96, volume 1258 of LNCS, pages 334-348, Utrecht, 1997. Springer.
[9] R. Matzinger. Context free term sets are regular - and some applications to logic. Technical Report TR-WB-Mat-97-2, TU Wien, Vienna/Austria, 1997.
[10] J. Slaney. FINDER (finite domain enumerator): Notes and guide. Technical Report TR-ARP-1/92, Australien National University Automated Reasoning Project, Canberra, 1992.
[11] T. Tammet. Using resolution for deciding solvable classes and building finite models. In Baltic Computer Science, pages 33-64. Springer Verlag, 1991. LNCS 502.
[12] M. Veanes. On Simultaneous Rigid E-Unification. PhD thesis, Computing Science Department, Uppsala University, 1997.


[^0]:    *Institut für Informationssysteme, Abteilung für Wissensbasierte Systeme; Treitlstr. 3/E184-3, A-1040 Wien/Austria/Europe; phone: $+43(1) 58801 / 8196$, telefax: $+43(1) 5817966$, email: matzi@kr.tuwien.ac.at, URL: http://www.kr.tuwien.ac.at/~matzi/
    ${ }^{\dagger}$ This work was supported by the Austrian Science Foundation under FWF grant P11624-MAT.
    ${ }^{1}$ For didactical reasons we consider formulas with only monadic predicate symbols (but function symbols

[^1]:    ${ }^{4}$ Modulo the AC-property of the connectives and modulo renaming of the variables.

[^2]:    ${ }^{5}$ With $f\left(L_{A}, L_{B}, \ldots\right)$ we denote the language that consists of all $f$ terms that have a member of $L_{A}$ on the first place, a member of $L_{B}$ on the second place and so forth.

