Restricted Optimal Unification

Wojciech Buszkowski
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Poznań Poland

Barbara Dziemidowicz
Institute of Mathematics and Computer Science
University of Warmia and Mazury
Olsztyn Poland

Abstract

In this paper we present a more general approach to restricted optimal unification, introduced and appplied to learning algorithms for categorial grammars in [1] and further developed in [6, 7, 3, 4]. In particular, we solve a rather general problem of finding minimal unifiers with respect to some preordering relation between substitutions.

1 Introduction and preliminaries

The notion of an optimal unifier of a finite family of finite sets of terms has been introduced in [1] as a tool for determining minimal categorial grammars compatible with a given sample of (structured) expressions of a language. In [6, 7] there are defined and studied optimal unifiers restricted to a set of substitutions; the goal is to make the procedure of determining categorial grammars sensitive to certain negative constraints upon the form of final outcomes. In [3, 4] the (unrestricted) procedures from [1] are used to define computable learning functions for categorial grammars.

In this paper we simplify and generalize the framework of restricted optimal unification. Instead of finite families of finite sets of terms we consider finite binary relations between terms (to be thought of as finite systems of equations). We find a simple characterization of optimal unifiers as most general unifiers of maximal unifiable subrelations of the given relation, which provides a straightforward algorithm for determining them. We also state quite general conditions for a preordering relation on the set of substitutions which are sufficient to solve the problem of finding minimal unifiers in a general setting. This result can directly be applied to justify some procedures of determining categorial grammars which are minimal with respect to different cardinality postulates.

Let L be a first-order language. By TER(L) we denote the set of all terms of L. VAR denotes the (countable) set of individual variables, and CON(L) denotes the set of all individual constants of L. *A substitution* is a mapping α :VAR$\mapsto$TER(L). Each substitution α is uniquely extended to a homomorphism α :TER(L)$\mapsto$TER(L), by setting: (i) $a\alpha = a$, for a $\in$CON(L), (ii) $f(t_1, \ldots, t_n)\alpha = f(t_1\alpha, \ldots, t_n\alpha)$. $\alpha\beta$ denotes the composition of α and β: $t\alpha\beta = (t\alpha)\beta$, for any term t. A substitution α is said to be *finite*, if $x\alpha = x$, for all but finitely many x $\in$VAR. Terms s, t are called *variants*, if there are substitutions α, β such that $s\alpha = t$ and $t\beta = s$ (then, s is equal to t up to renaming variables). ε denotes the identity substitution: $x\varepsilon = x$, for all variables x.

A set T $\subseteq$TER(L) is said to be *unifiable*, if there is a substitution α such that $s\alpha = t\alpha$, for all $s, t \in T$; α is called *a unifier* of T. A substitution α is called *a most general unifier* (m.g.u.) of T, if α is a unifier of T and, for every unifier β of T, there is a substitution γ such that $\beta = \alpha\gamma$. The well-known *unification algorithm* [5, 2] takes an arbitrary finite set of terms as an input and returns an m.g.u. of T, if T is unifiable; otherwise, the algorithm replies negatively.

For our purposes, it is more expedient to consider unifiers of relations. Let R be a finite binary relation on TER(L). A substitution α is called *a unifier* of R, if $s\alpha = t\alpha$, for any pair $(s, t) \in R$. The relation R is said to be *unifiable*, if there exists a unifier of R. The notion of an m.g.u. of R is defined as above. It is easy to adapt the unification algorithm for binary relations as inputs. Let $R = \{(s_1, t_1), \ldots, (s_n, t_n)\}$. Take a new $n-$ary function symbol F and consider the set T consisting of two terms: $F(s_1, \ldots, s_n)$ and $F(t_1, \ldots, t_n)$. Then, unifiers of R are precisely unifiers of T (not using F), and consequently, the unification algorithm applied to T provides an m.g.u. of R, if R is unifiable, and replies negatively, otherwise.

By SUB(L) we denote a set of all substitutions α :VAR$\mapsto$TER(L). Following [6, 7], a set X $\subseteq$SUB(L) is said to be *hereditary* if, for all $\alpha, \beta \in X$, if $\alpha\beta \in X$, then $\alpha \in X$. Notice that a hereditary set X is nonempty iff $\varepsilon \in X$. Further, hereditary sets of substitutions are closed with respect to arbitrary joins and meets, that means, the family of all hereditary sets of substitutions is a complete lattice of sets.

Hereditary sets of substitutions are a useful tool to make unification sensitive to some negative constraints. Let $s_1, \ldots, s_n$ be some terms, and let $t_1, \ldots, t_n$ be some ground terms (i.e. terms containing no variable). Then, the set of all substitutions α such that $s_i\alpha \neq t_i$, for all $1 \leq i \leq n$, is hereditary. Another example of a hereditary set is the set of all α such that each term $t_i\alpha$ contains at most ten function symbols. Since hereditary sets form a complete lattice of sets, then different negative constraints can be joined in conjunctions and disjunctions, and the resulting compound constraint again gives rise to a hereditary set.

In what follows, X always denotes a hereditary set of substitutions. We also assume that the problem $\alpha \in X$?, where α can be an arbitrary finite substitution, is decidable.

A relation R is said to be *unifiable in* X, if there exists a unifier α of R such that $\alpha \in X$.

Proposition 1 *If R is unifiable in X, then every m.g.u. of R belongs to X.*

PROOF. Let $\alpha \in X$ be a unifier of R, and let σ be an m.g.u. of R. Then, $\alpha = \sigma\gamma$, for some γ. Since X is hereditary, then $\sigma \in X$. $\Box$

Consequently, the unification algorithm can easily be modified to check if the given relation is unifiable in X and, if so, to return an m.g.u. of R which belongs to X. First, the standard unification algorithm checks if R is unifiable. If not, R is not unifiable in X. If so, then the algorithm returns a (finite) m.g.u. of R. One checks if this m.g.u. belongs to X. If not, then R is not unifiable in X. If so, then this m.g.u. is an m.g.u. of R which belongs to X.

At the end of this section, we note that the notion of a unifier of a relation R and all related notions are actually different formulations of concepts defined in [1] in terms of families of sets of terms. Let $\mathcal{T}$ be a finite family of finite sets of terms. A substitution α is called *a unifier* of $\mathcal{T}$, if α is a unifier of every set $T \in \mathcal{T}$ (but not necessarily of the union of $\mathcal{T}$). For a family $\mathcal{T}$, we define a relation $R(\mathcal{T})$ as follows: $(s,t) \in R(\mathcal{T})$ iff, for some $T \in \mathcal{T}$, both $s \in T$ and $t \in T$. Then, unifiers of $\mathcal{T}$ are precisely unifiers of $R(\mathcal{T})$, and similarly for other notions, defined in terms of unifiers. Conversely, unifiers of R are precisely unifiers of the family of all sets $\{s,t\}$ such that $(s,t) \in R$. Analogous reformulations are possible for all notions considered later on. The formulation in terms of relations is more expedient for the treatment of optimal unifiers, given in the next section.

2 Restricted optimal unifiers

If the relation R is not unifiable, then it is reasonable to look for best partial unifiers of R. A natural application of this framework is categorial grammar. Given a language sample, one looks for a most economical categorial grammar compatible with this sample [1]. More on this and other applications will be said in the next section.

For a substitution α, the relation $\mathrm{Ker}\alpha$ (*the kernel* of α) is defined as follows: for $s,t \in$TER(L), $(s,t) \in\mathrm{Ker}\alpha$ iff $s\alpha = t\alpha$. For a relation R between terms, we set $\mathrm{Ker}_R\alpha = R\cap\mathrm{Ker}\alpha$. Clearly, α is a unifier of $\mathrm{Ker}_R\alpha$, for any R. Also, $\mathrm{Ker}_R\alpha \subseteq\mathrm{Ker}_R\alpha\beta$, for all R, α and β.

A substitution α is called *an optimal unifier* of a relation R, if it satisfies the following conditions:

(OU.1) for every substitution β, $\mathrm{Ker}_R\alpha =\mathrm{Ker}_R\alpha\beta$,

(OU.2) α is an m.g.u. of $\mathrm{Ker}_R\alpha$.

Proposition 2 *Assuming (OU.2), the condition (OU.1) is equivalent to:* *(OU.1') there is no substitution* γ *such that* $Ker_R\alpha \subset Ker_R\gamma$.

PROOF. It is obvious that (OU.1') entails (OU.1). Assume (OU.2) and the negation of (OU.1'). Then, $\mathrm{Ker}_R\alpha \subset \mathrm{Ker}_R\gamma$, for some γ. Since γ is a unifier of $\mathrm{Ker}_R\alpha$, then, by (OU.2), there exists β such that $\gamma = \alpha\beta$. Consequently, $\mathrm{Ker}_R\alpha \neq \mathrm{Ker}_R\alpha\beta$, hence (OU.1) does not hold. □

Accordingly, an optimal unifier of R is a substitution which unifies R as far, as possible, and is most general in a sense. The next theorem provides a simple characterization of optimal unifiers.

Theorem 1 *Optimal unifiers of* R *are precisely m.g.u.'s of maximal unifiable subrelations of* R.

PROOF. Assume α be an optimal unifier of R. Then, α is an m.g.u. of $\mathrm{Ker}_R\alpha$. We show that $\mathrm{Ker}_R\alpha$ is a maximal unifiable subrelation of R. It is a subrelation of R. Assume $\mathrm{Ker}_R\alpha \subset S \subseteq R$ and S be unifiable. Let γ be a unifier of S. Then, $S \subseteq \mathrm{Ker}_R\gamma$, and consequently, $\mathrm{Ker}_R\alpha \subset \mathrm{Ker}_R\gamma$, which contradicts (OU.1').

Assume α be an m.g.u. of a relation $S \subseteq R$ which is a maximal unifiable subrelation of R. Then, $S \subseteq \mathrm{Ker}_R\alpha$, which yields $S = \mathrm{Ker}_R\alpha$, since $\mathrm{Ker}_R\alpha \subseteq R$ is unifiable, and S is maximal unifiable. That proves (OU.2). (OU.1') also holds, by the maximality of S. □.

Corollary 1 *If* R *is unifiable, then optimal unifiers of* R *are precisely m.g.u.'s of* R.

Corollary 1 shows that the notion of an optimal unifier is a natural generalization of the notion of an m.g.u. for nonunifiable relations. Theorem 2 justifies a nondeterministic procedure for determining all optimal unifiers of R (up to variants). One finds all maximal unifiable subrelations of R and m.g.u.'s of these subrelations, by means of the unification algorithm. Another procedure has been described in [1].

Let X be a hereditary set of substitutions. A substitution α is called *an optimal unifier* of R *restricted* to X, if it satisfies the following conditions:

(ROU.0) $\alpha \in X$,

(ROU.1) for all β, if $\alpha\beta \in X$, then $\mathrm{Ker}_R\alpha = \mathrm{Ker}_R\alpha\beta$,

(ROU.2) α is an m.g.u. of $\mathrm{Ker}_R\alpha$.

For $X = \mathrm{SUB(L)}$, the above notion is the same as the notion of an optimal unifier of R. One can easily modify the proofs of all results from this section to obtain their versions for restricted optimal unifiers.

Proposition 3 *Assuming (ROU.0) and (ROU.2), (ROU.1) is equivalent to:* *(ROU.1') there is no substitution* $\gamma \in X$ *such that* $Ker_R\alpha \subset Ker_R\gamma$.

Theorem 2 *Optimal unifiers of R restricted to X are precisely m.g.u.'s of maximal unifiable in X subrelations of R.*

Corollary 2 *If R is unifiable in X, then optimal unifiers of R restricted to X are precisely m.g.u.'s of R.*

We note that optimal unifiers of R restricted to X are not simply those optimal unifiers of R which belong to X. Take $R = \{(x, f(y)\}$. Let X consist of all substitutions α such that $x\alpha$ is a variable. Then, X is hereditary. The substitution $\{x/f(y)\}$ is an m.g.u. of R, hence it is an optimal unifier of R, but there is no optimal unifier of R which belongs to X. The identity substitution ε is an optimal unifier of R restricted to X.

Consider the relation $R = \{(x, f(y)), (x, y)\}$. R is not unifiable. Two optimal unifiers of R are $\{x/f(y)\}$ and $\{x/y\}$.

3 Minimal unifiers

In some applications one is interested in determining all substitutions which are minimal in a sense, say, the image $T\alpha$ is of a least cardinality. We prove a theorem which solves this problem in a rather general form.

Recall that a preordering relation on a set X is reflexive and transitive. If $\leq$ is a preordering relation on X, then we write $a \sim b$ if $a \leq b$ and $b \leq a$, and $a < b$ if $a \leq b$ and $b \not\leq a$, for $a, b \in X$.

Let $X \subseteq$SUB(L) be hereditary. Let R be a finite binary relation between terms of L. A preordering relation $\leq$ on X is said to be *sound* for R, if it satisfies the following conditions:

(C1) if $\mathrm{Ker}_R\alpha \subset \mathrm{Ker}_R\beta$, then $\beta < \alpha$,

(C2) if $\mathrm{Ker}_R\alpha = \mathrm{Ker}_R\beta$, then $\alpha \sim \beta$,

for all $\alpha, \beta \in X$. (C1) means that a substitution β which unifies R stronger than α must be smaller than α with respect to $\leq$. (C2) means that two substitutions which unify R in the same way are equivalent with respect to $\leq$. Notice that (C1) and (C2) entail: (C3) if $\mathrm{Ker}_R\alpha \subseteq \mathrm{Ker}_R\beta$, then $\beta \leq \alpha$, for all $\alpha, \beta \in X$.

Take $X =$SUB(L). Define $\alpha \leq_0 \beta$ iff $\mathrm{Ker}_R\beta \subseteq \mathrm{Ker}_R\alpha$. Then, $\leq_0$ is sound for R. Let $T \subseteq$TER(L) be finite. Define $\alpha \leq_1 \beta$ iff $\mathrm{card}(T\alpha) \leq \mathrm{card}(T\beta)$. This preordering is sound for $R(\{T\})$. Now, let $\mathcal{T} = \{T_1, \ldots, T_n\}$ be a finite family of finite sets of terms of L. We consider the standard product ordering on $n-$tuples of integers: $(k_1, \ldots, k_n) \leq (l_1, \ldots, l_n)$ iff $k_i \leq l_i$, for all $1 \leq i \leq n$. Define $\alpha \leq_2 \beta$ iff $(k_1, \ldots, k_n) \leq (l_1, \ldots, l_n)$, where k_i is the cardinality of $T_i\alpha$ and l_i is the cardinality of $T_i\beta$, for all $1 \leq i \leq n$. Then, $\leq_2$ is sound for $R(\mathcal{T})$. We can also consider the lexicographical ordering $\leq'$ on $n-$tuples of integers: $(k_1, \ldots, k_n) \leq' (l_1, \ldots, l_n)$ iff either both tuples are equal, or $k_j < l_j$, for the least $1 \le j \le n$ such that $k_j \ne l_j$. Define $\le_3$ in a similar way as $\le_2$ except that $\le$ is replaced by $\le'$. Again, $\le_3$ is sound for $R(\mathcal{T})$.

Lemma 1 *Let X be hereditary. For every substitution $\alpha \in X$, there exists an optimal unifier β of R restricted to X such that $Ker_R\alpha \subseteq Ker_R\beta$.*

PROOF. Let $\alpha \in X$. Then $\mathrm{Ker}_R\alpha$ is unifiable in X. There exists a maximal unifiable in X relation $S \subseteq R$ such that $\mathrm{Ker}_R\alpha \subseteq S$. Let β be an m.g.u. of S. By proposition 1, $\beta \in X$. By theorem 2, β is an optimal unifier of R restricted to X. Clearly, $\mathrm{Ker}_R\alpha \subseteq \mathrm{Ker}_R\beta$. □

Lemma 2 *Let $\le$ be a preordering on X sound for R. Then, for every $\alpha \in X$, there exists an optimal unifier β of R restricted to X such that $\beta \le \alpha$.*

PROOF. Use lemma 1 and (C3). □

Let $\le$ be a preordering on X. A substitution $\alpha \in X$ is said to be *minimal* for $\le$, if there is no $\beta \in X$ such that $\beta < \alpha$. An optimal unifier α of R restricted to X is called *a minimal optimal unifier* of R restricted to X with respect to $\le$, if there is no optimal unifier β of R restricted to X such that $\beta < \alpha$. We are ready to prove the main result of this section.

Theorem 3 *Let X be hereditary, R be a finite relation between terms of L. Let $\le$ be a preordering on X sound for R. Then, for every $\alpha \in X$, the following conditions are equivalent: (i) α is minimal for $\le$, (ii) there exist a minimal optimal unifier β of R restricted to X with respect to $\le$ and a substitution γ such that $\alpha = \beta\gamma$.*

PROOF. Assume (i). By lemma 1, there exists an optimal unifier β of R restricted to X such that $\mathrm{Ker}_R\alpha \subseteq \mathrm{Ker}_R\beta$. Suppose $\mathrm{Ker}_R\alpha \subset \mathrm{Ker}_R\beta$. Then, $\beta < \alpha$, by (C1), which contradicts the minimality of α. Accordingly, $\mathrm{Ker}_R\alpha = \mathrm{Ker}_R\beta$, hence $\alpha \sim \beta$, by (C2). Suppose $\beta' < \beta$, for some optimal unifier β' of R restricted to X. Then, $\beta' < \alpha$, against the minimality of α. Consequently, β is a minimal optimal unifier of R restricted to X with respect to $\le$. Since β is an m.g.u. of $\mathrm{Ker}_R\beta$, by (RC2), and α unifies $\mathrm{Ker}_R\beta$, then there exists γ such that $\alpha = \beta\gamma$. We have shown (ii).

Assume (ii). First, $\mathrm{Ker}_R\alpha = \mathrm{Ker}_R\beta$, by (ROU.1), hence $\alpha \sim \beta$. Suppose $\alpha' < \alpha$, for some $\alpha' \in X$. Then, $\alpha' < \beta$. By lemma 2, there exists an optimal unifier β' of R restricted to X such that $\beta' \le \alpha'$. Then, $\beta' < \beta$, which contradicts the assumption that β is a minimal optimal unifier of R restricted to X with respect to $\le$. Consequently, α is minimal for $\le$. We have shown (i). □

For $X =$SUB(L), we obtain the following version of theorem 3:

Theorem 4 *Let $\le$ be a preordering on SUB(L) sound for a finite relation R. Then, for every substitution α, the following conditions are equivalent: (i) α is minimal for $\le$, (ii) there exist a minimal optimal unifier of R with respect to $\le$ and a substitution γ such that $\alpha = \beta\gamma$.*

As a consequence, we can describe a general method of determining all substitutions minimal for $\leq$, where $\leq$ is a preordering (on X) compatible with a finite relation R. We must only assume that the condition $\alpha \leq \beta$ is decidable for finite substitutions α, β (in X). First, we find all (up to variants) optimal unifiers of R (restricted to X), by the methods of section 2. Then, we determine all minimal optimal unifiers of R (restricted to X) with respect to $\leq$. They are all "most general" substitutions minimal for $\leq$.

Let us consider a simple example. Take $T = \{x, f(x), y, f(y)\}$. We consider the preordering $\leq_1$, defined above. The optimal unifiers of T (that means, of $R(\{T\})$) are: $\{x/y\}$, $\{x/f(y)\}$ and $\{y/f(x)\}$. The first one of them is the only minimal optimal unifier of T with respect to $\leq_1$. By theorem 4, the minimal substitutions for $\leq_1$ are of the form $\{x/y\}\gamma$, for any substitution γ.

A direct application of this method can be found in the theory of categorial grammars. One considers a language sample $\mathcal{L}$, consisting of finitely many expressions, given in the form of functor-argument structures, e.g. (Joan (smiles charmingly)$_2$)$_2$ (the numerical subscript indicates which constituent is the functor; thus, 'charmingly' is the functor in 'smiles charmingly', and 'smiles charmingly' is the functor in 'Joan smiles charmingly'). A categorial grammar G is said to be compatible with $\mathcal{L}$, if $\mathcal{L} \subseteq L(G)$; here $L(G)$ denotes the (structured) language generated by G. In [1], there is given a construction of a categorial grammar $\mathrm{GF}(\mathcal{L})$, called the general form for $\mathcal{L}$, which satisfies the condition: for any categorial grammar G, G is compatible with $\mathcal{L}$ iff there exists a substitution α such that $\mathrm{GF}(\mathcal{L})\alpha \subseteq G$. Here $G\alpha$ denotes the categorial grammar which arises from G by replacing each lexical assignment $v \mapsto t$ by $v \mapsto t\alpha$ (v is a lexical atom, and t is a type).

We want to find all categorial grammars compatible with $\mathcal{L}$ such that the set of types assigned to lexical atoms is of least cardinality. We proceed in the following way. First, we find $\mathrm{GF}(\mathcal{L})$. Let T be the set of all types assigned to lexical atoms in $\mathrm{GF}(\mathcal{L})$. The required grammars are precisely the grammars $\mathrm{GF}(\mathcal{L})\alpha$ such that α is minimal for $\leq_1$. Consequently, it is enough to find all minimal optimal unifiers β of $R(\{T\})$ with respect to $\leq_1$, and the required grammars are the grammars $\mathrm{GF}(\mathcal{L})\beta\gamma$ with γ ranging over all substitutions. Other problems of minimization can be solved in a similar way.

We have consciously abandoned details of these applications, since a full presentation requires a long list of definitions and results concerning categorial grammars. A thorough presentation of restricted optimal unification applied to categorial grammars will be given in another paper.

References

[1] W. Buszkowski and G. Penn, Categorial grammars determined from linguistic data by unification, *Studia Logica* 49 (1990), 431-454.

[2] K. Doets, *From Logic to Logic Programming*, MIT Press, Cambridge Mass., 1994.

[3] M. Kanazawa, Identification in the limit of categorial grammars, *Journal of Logic, Language and Information* 5 (1996), 115-155.

[4] M. Kanazawa, *Learnable Classes of Categorial Grammars*, CSLI Publications, FoLLI, Stanford, 1998.

[5] J.W. Lloyd, *Foundations of Logic Programming*, Springer, Berlin, 1987.

[6] J. Marciniec, Learning categorial grammars by unification with negative constraints, *Journal of Applied Non-Classical Logics* 4 (1994), 181-200.

[7] J. Marciniec, The method of unification in procedures of determining categorial grammars, (in Polish), *Investigationes Linguisticae* 2 (1997), 45-96.