## Program Verification: Lecture 8

José Meseguer

Computer Science Department
University of Illinois at Urbana-Champaign

## Overlaps as Unification Problems

We reduced confluence (under the termination assumption) to joinability of context-free nested simplifications with overlap. But note that we can have a context-free overlap situation with equations $u=v$ and $u^{\prime}=v^{\prime}$ (again, with disjoint variables) if and only if there is a nonvariable position $p$ in $u$ and a substitution $\theta$ such that,

$$
(\dagger) \quad u_{p} \theta=u^{\prime} \theta
$$

Therefore, finding all possible context-free nested simplifications with overlap can be reduced to finding, for all pairs of equations $u=v$ and $u^{\prime}=v^{\prime}$ in $E$ and all nonvariable positions $p$ in $u$, all solutions to $(\dagger)$. Problems of the form ( $\dagger$ ) are called unification problems.

## Unification

In general, the unification problem consists in, given terms $t$ and $t^{\prime}$ whose sorts are in the same connected component, finding a substitution $\theta$ that makes them equal, so that we have identical terms, $t \theta=t^{\prime} \theta$. The substitution $\theta$ is then called a unifier of $t$ and $t^{\prime}$.

Under very reasonable conditions on $\Sigma$, such as finiteness, this problem is decidable in a very strong sense. Namely, we can effectively find a finite set of unifiers $\left\{\theta_{1}, \ldots \theta_{n}\right\}$, that are the most general possible, in the sense that for any other substitution $\mu: \operatorname{vars}\left(t=t^{\prime}\right) \longrightarrow T_{\Sigma}(V)$ such that $t \mu=t^{\prime} \mu$, we can always find a $\theta_{i}$, say, $\theta_{i}: \operatorname{vars}\left(t=t^{\prime}\right) \longrightarrow T_{\Sigma}(X)$, and a substitution $\rho: X \longrightarrow T_{\Sigma}(V)$ such that for each $x \in \operatorname{vars}\left(t=t^{\prime}\right)$ we have $x \mu=x \theta_{i} \rho$.

## $B$-Unification

The standard unification problem is to try to unify two terms. But we have already encountered situations, such as the relation $\longrightarrow_{E / B}$, in which it is very useful to deal not with terms, but with equivalence classes of terms modulo some equational axioms $B$.

Therefore, it is natural, given a set of equational axioms $B$, such as the associativity, commutativity, and identity of some operators, to generalize the unification problem to the following $B$-unification problem: given an equation $t=t^{\prime}$ are there substitutions $\theta$ such that

$$
t \theta={ }_{B} t^{\prime} \theta
$$

## $B$-Unification (II)

For $B$ any combination of associativity, commutativity, and identity axioms, there are known algorithms that can find a family of most general unifiers for each given unification problem $t=t^{\prime}$. However, for the case of associativity alone, or of associativity and identity alone, this family of most general unifiers may be infinite.

In particular, for $\Sigma$ a finite signature, if we choose $B$ to be any combination of associativity, commutativity and identity axioms for different (subsort-overloaded) binary operators in $\Sigma$, except associativity without commutativity, there is indeed an algorithm that, given an $B$-unification problem, either declares the problem unsolvable, or finds a finite set of most general unifiers solving it. Such a $B$-unification algorithm is used by the Church-Rosser Checker.

## More on Unification

So far we have said nothing about unification algorithms, that can effectively find a set of most general unifiers or declare the corresponding problem unsolvable.

Unification is indeed a vast research area, and the more we can do in this course is to give a flavor for the key ideas. This can be done quite well by considering the simplest version of the unification problem, for which, if the given equation has a solution, then it has a unique most general unifier.

## More on Unification (II)

This simplest version is the case of a sensible many-sorted signature $\Sigma$ without ad-hoc overloading.

The key idea of a unification algorithm is to transform the original equation we want to solve into a set of equations equivalent to the original equation, in the sense that both sets have the same solutions.

We then stop either with failure, or with a set of equations in solved form, that is, equations having the shape, $\left\{x_{1}=t_{1}, \ldots, x_{n}=t_{n}\right\}$, where the $x_{i}$ do not appear in the $t_{j}$. But this is just another garb for a substitution $\theta=\left\{\left(x_{1}, t_{1}\right), \ldots,\left(x_{n}, t_{n}\right)\right\}$.

## The Unification Algorithm

We can describe the unification algorithm, à la
Martelli-Montanari, as a set of inference rules, that
transform a set of equations $E$ into another set of equations that is equivalent to it from the solvability point of view, or into the constant failure. The following inference rules assume a many-sorted signature, make the equality symbol commutative, and use a global set $V$ of variables:

- Delete:

$$
\frac{\{E, t=t\}}{\{E\}}
$$

- Decompose:

$$
\frac{\left\{E, f\left(t_{1}, \ldots, t_{n}\right)=f\left(t_{1}^{\prime}, \ldots, t_{n}^{\prime}\right)\right\}}{\left.\left\{E, t_{1}=t_{1}^{\prime}, \ldots, t_{n}=t_{n}^{\prime}\right)\right\}}
$$

## The Unification Algorithm (II)

- Conflict:

$$
\frac{\left\{E, f\left(t_{1}, \ldots, t_{n}\right)=g\left(t_{1}^{\prime}, \ldots, t_{m}^{\prime}\right)\right\}}{\text { failure }}
$$

$$
\text { if } f \neq g
$$

- Coalesce:

$$
\begin{aligned}
& \frac{\{E, x=y\}}{\{E\{x \mapsto y\}, x=y\}} \\
& \text { if } x, y \in \operatorname{vars}(E), x \neq y
\end{aligned}
$$

- Check:

$$
\frac{\{E, x=t\}}{\text { failure }}
$$

if $x \in \operatorname{vars}(t), x \neq t$

## The Unification Algorithm (III)

- Eliminate:

$$
\begin{array}{r}
\frac{\{E, x=t\}}{\{E\{x \mapsto t\}, x=t\}} \\
\text { if } x \notin \operatorname{vars}(t), t \notin V, x \in \operatorname{vars}(E)
\end{array}
$$

We can illustrate the use of these rules by finding the most general unifier for a relatively simple, yet nontrivial, unification problem, namely, solving the equation,

$$
f(g(x, h(y)), z)=f(z, g(k(u), v))
$$

for which the above rules give us the following transformations:

$$
\{f(g(x, h(y)), z)=f(z, g(k(u), v))\} \longrightarrow(\text { Decompose })
$$

## The Unification Algorithm (IV)

$\{g(x, h(y))=z, z=g(k(u), v)\} \longrightarrow($ Eliminate $)$
$\{g(x, h(y))=g(k(u), v), z=g(k(u), v)\} \longrightarrow($ Decompose)
$\{x=k(u), v=h(y), z=g(k(u), v)\} \longrightarrow($ Eliminate $)$
$\{x=k(u), v=h(y), z=g(k(u), h(y))\}$,
which is the desired most general unifier, yielding the identity,
$f(g(k(u), h(y)), g(k(u), h(y)))=f(g(k(u), h(y)), g(k(u), h(y)))$.

## Unification Modulo Commutativity

To illustrate the case of $B$-unification in a many-sorted signature $\Sigma$, let us assume that $B=C o m m$ is a collection of commutativity axioms for some binary symbols $\Sigma_{c o m m} \subseteq \Sigma$. The inference rules for unification modulo commutativity are:

- Delete:

$$
\frac{\{E, t=t\}}{\{E\}}
$$

- Decompose: $\left(f \in\left(\Sigma-\Sigma_{\text {comm }}\right)\right)$

$$
\frac{\left\{E, f\left(t_{1}, \ldots, t_{n}\right)=f\left(t_{1}^{\prime}, \ldots, t_{n}^{\prime}\right)\right\}}{\left.\left\{E, t_{1}=t_{1}^{\prime}, \ldots, t_{n}=t_{n}^{\prime}\right)\right\}}
$$

## Unification Modulo Commutativity (II)

- Decompose-C: $\left(f \in \Sigma_{\text {comm }}\right)$

$$
\frac{\left\{E, f\left(t_{1}, t_{2}\right)=f\left(t_{1}^{\prime}, t_{2}^{\prime}\right)\right\}}{\left.\left.\left\{E, t_{1}=t_{1}^{\prime}, t_{2}=t_{2}^{\prime}\right)\right\} \vee\left\{E, t_{1}=t_{2}^{\prime}, t_{2}=t_{1}^{\prime}\right)\right\}}
$$

- Conflict:

$$
\underline{\left\{E, f\left(t_{1}, \ldots, t_{n}\right)=g\left(t_{1}^{\prime}, \ldots, t_{m}^{\prime}\right)\right\}}
$$

failure
if $f \neq g$

- Coalesce:

$$
\frac{\{E, x=y\}}{\{E\{x \mapsto y\}, x=y\}}
$$

if $x, y \in \operatorname{vars}(E), x \neq y$

## Unification Modulo Commutativity (III)

- Check:

$$
\frac{\{E, x=t\}}{\text { failure }}
$$

$$
\text { if } x \in \operatorname{vars}(t), x \neq t
$$

- Eliminate:

$$
\begin{array}{r}
\frac{\{E, x=t\}}{\{E\{x \mapsto t\}, x=t\}} \\
\text { if } x \notin \operatorname{vars}(t), t \notin V, x \in \operatorname{vars}(E)
\end{array}
$$

Note that now, because of Rule Decompose-C, there can be several solutions to a unification problem. Also, we define failure as an identity element for _ $V_{\text {_ }}$.

We can illustrate the use of these rules by finding the most general unifiers modulo commutativity when $\Sigma_{\text {comm }}=\{g\}$.

## Unification Modulo Commutativity (IV)

Let us apply these rules to solve the equation,

$$
\begin{aligned}
& f(g(h(y), x), z)=f(z, g(k(u), v)) \\
& \{f(g(h(y), x), z)=f(z, g(k(u), v))\} \longrightarrow(\text { Decompose }) \\
& \{g(h(y), x)=z, z=g(k(u), v)\} \longrightarrow(\text { Eliminate }) \\
& \{g(h(y), x)=g(k(u), v), z=g(k(u), v)\} \longrightarrow(\text { Decompose-C) } \\
& \{x=v, k(u)=h(y), z=g(k(u), v)\} \vee\{x=k(u), v=h(y), z= \\
& g(k(u), v)\} \longrightarrow(\text { Conflict } \vee \text { Eliminate })
\end{aligned}
$$

failure $\vee\{x=k(u), v=h(y), z=g(k(u), h(y))\}=$ $\{x=k(u), v=h(y), z=g(k(u), h(y))\}$
applying the resulting unifier we obtain the identity, $f(g(h(y), k(u)), g(h(y), k(u)))={ }_{c o m m} f(g(k(u), h(y)), g(k(u), h(y)))$.

## Where to Go from Here

We can only sketch how to go from here to more general unification algorithms, such as $B$-unification in an order-sorted signature $\Sigma$.

First of all, note that the presence of overloading and subsorts will typically move us from a single most general unifier to a finite set of them. This is because of the presence of subsort overloaded operators, which may lead to several different solutions. Note also that, in the presence of subsorts, even apparently innocent equations such as $x: s=y: s^{\prime}$ may lead to failure, because the sorts $s$ and $s^{\prime}$ may not have any common subsort. For example, in an INT specification, an equation $\mathrm{X}=\mathrm{Y}$, with X of sort NzNat, and Y of sort NzNeg will fail.

## Where to Go from Here (II)

Chapter 15.1 of the Maude book gives a detail presentation of the inference rules for order-sorted $C$-unification and gives an implementation that you can use in Maude for experimentation.

The latest version of Maude provides general order-sorted unification algorithm modulo any combination of $C$ and/or $A$ and/or $U$ axioms. Since the set of $A$-unifiers of an equation can be infinite, Maude provides a finite set and a warning if more solutions may exist.

The Church-Rosser Checker uses unification modulo any combinations of associativity, commutativity, and identity axioms; but may not generate all critical pairs [giving a warning] for associativity without commutativity axioms.

## Where to Go from Here (III)

For a survey of unification algorithms modulo axioms see:
J.-P. Jouannaud and C. Kirchner, "Solving Equations in Abstract Algebras," in J.-L. Lassez and G.Plotkin, eds., Computational Logic: Essays in Honor of Alan Robinson.

For order-sorted $B$-unification see:
J. Meseguer, J.A. Goguen, and G. Smolka, "Order-Sorted Unification," J. Symbolic Computation, Volume 8, 1989, pages 383-413.
J. Hendrix and J. Meseguer, "Equational Order-Sorted Unification Revisited," Electr. Notes Theor. Comput. Sci., Vol. 290, 2012, pages 37-50.

## What Are Critical Pairs?

Theorem: Let $(\Sigma, B \cup E)$ be an order-sorted equational theory, with $\Sigma B$-preregular, $\vec{E}$ sort-decreasing, and $\longrightarrow_{E / B}$ terminating. Let $V=\operatorname{vars}(E)$, and $\gamma: V \rightarrow V^{\prime}$ a bijective, sort-preserving renaming of variables with $V \cap V^{\prime}=\emptyset$. Then, $\vec{E}$ is confluent modulo $B$ iff, for each pair of equations ${ }^{\text {a }}$ $u=v$ in $E$ and $u^{\prime}=v^{\prime}$ in $E \gamma$ (including $\left(u^{\prime}=v^{\prime}\right) \equiv(u \gamma=v \gamma)$ ) for each nonvariable position $p$ in $u$, and for each most general order-sorted $B$-unifier $\theta$ such that $u_{p} \theta={ }_{B} u^{\prime} \theta$, we have,

$$
\text { (b) } \quad v \theta \downarrow_{E / B} u\left[v^{\prime}\right]_{p} \theta \text {. }
$$

The corresponding equations $v \theta=u\left[v^{\prime}\right]_{p} \theta$, are called the critical pairs of the equations $E$ modulo $B$.

[^0]
## What Are Critical Pairs? (II)

Proof: (For $B=\emptyset$ ). We had already reduced checking confluence to checking that, for each pair of equations $(u=v) \in E$ and $\left(u^{\prime}=v^{\prime}\right) \in E \gamma$, for each nonvariable position $p$ in $u$, and for each order-sorted unifier $\mu$ such that $u_{p} \mu=u^{\prime} \mu$, we have,

$$
\text { (b) } v \mu \downarrow_{E} u\left[v^{\prime}\right]_{p} \mu
$$

But if $\operatorname{Unif}\left(u_{p}=u^{\prime}\right)=\left\{\theta_{1}, \ldots, \theta_{n}\right\}$ is the set of most general order-sorted $B$-unifiers for the equation $u_{p}=u^{\prime}$, then we can find a $\theta_{i} \in \operatorname{Unif}_{B}\left(u_{p}=u^{\prime}\right)$ and a substitution $\rho$ such that $\mu=\theta_{i} \rho$.

Since we know that there is a $w$ such that $v \theta_{i} \rightarrow_{E}^{*} w$ and $u\left[v^{\prime}\right]_{p} \theta_{i} \rightarrow_{E}^{*} w$, we will be done if we prove the following:

## What Are Critical Pairs? (III)

Substitution Lemma: If $t \xrightarrow{*}_{E} t^{\prime}$ and $\rho$ is a substitution, then $t \rho \xrightarrow{*}_{E} t^{\prime} \rho$.

Proof: It is enough to prove the case for $t \longrightarrow_{E} t^{\prime}$, since then the case $t \xrightarrow{*}_{E} t^{\prime}$ follows easily by induction on the number of steps. But $t \longrightarrow_{E} t^{\prime}$ means that there is an equation $(u=v) \in E$, position $q$ and a substitution $\theta$ such that $t=t[u \theta]_{q}$ and $t^{\prime}=t[v \theta]_{q}$.

But note that, by the definition of the function $\quad \rho$, we can easily prove that we have, $t \rho=t \rho[u \theta \rho]_{q}$, and $t^{\prime} \rho=t \rho[v \theta \rho]_{q}$. Therefore, $t \rho \longrightarrow_{E} t^{\prime} \rho$ holds by applying $u=v$ at position $q$ with substitution $\theta ; \_$, as desired. q.e.d.

This finishes the proof of the Theorem (for $B=\emptyset$ ). q.e.d.

## In Summary

What the Church-Rosser Checker does is:

- it checks that the equations $E$ are sort-decreasing;
- it forms all the critical pairs for the equations $E$ and tries to join them;
- it returns as proof obligations those equation specializations that it could not prove sort-decreasing, and those simplified critical pairs that it could not join.

The arguments in Lecture 7 and in this lecture have shown that this method is correct for checking confluence, under the termination assumption.

## Checking Sufficient Completeness

We need methods to check that an equational theory $(\Sigma, E)$ is sufficiently complete. For arbitrary equational theories sufficient completeness is in general undecidable. This is not so bad: it just means that we may have to do some inductive theorem proving.

Sufficient completeness is decidable for a very broad class of order-sorted theories, namely, unconditional theories of the form $(\Sigma, E \cup B)$ with: (iv) $B$ a set of axioms for operators allowing any combination of associativity and/or commutativity and/or identity, except associativity without commutativity, and $E$ : (i) left-linear; (ii) sort-decreasing; and (iii) terminating.

## Checking Sufficient Completeness (II)

Furthermore, even for cases satisfying the above requirements (i)-(iii), but where $B$ includes operators that are only associative, or associative and identity, sufficient completeness, although undecidable in theory, becomes decidable in practice for many specifications of interest using specialized heuristic algorithms.

Left-linearity (i) means that if $t=t^{\prime} \in E$, then $t$ has no repeated variables. This fails, e.g., for the idempotency equation $x \cup x=x$. Properties (ii)-(iii) (modulo $B$ ) we are alredy familiar with.

## Tree Automata

For equational theories satisfying the above requirements (i)-(iii) we can use decidability results from tree automata theory to cast the sufficient completeness problem into a tree automata problem and decide the problem that way.

An ordinary finite-state automaton $\mathcal{A}$ has a finite set $Q$ of states and accepts strings of inputs when they lead the automaton to a subset $Q_{0} \subseteq Q$ of accepting states. The language $L(\mathcal{A})$ of the automaton is the the set of all strings accepted by $\mathcal{A}$. Such languages are called regular Ianguages and have nice decidablity results: they are closed under Boolean operations (we can construct automata for each such operation); and we can decide whether $L(\mathcal{A})$ is empty.

## Tree Automata (II)

All this is generalized by finite-state tree automata, which accept terms in an unsorted term algebra $\mathbb{T}_{\Sigma}$ instead of just strings. A tree automaton is a tuple $\mathcal{A}=\left(\Sigma, Q, Q_{0}, R\right)$ with $\Sigma$ an unsorted signature, $Q$ a set of extra constants not in $\Sigma$ called states, $Q_{0} \subseteq Q$ a subset of accepting states, and $R$ a set of transition rules, which can be of two forms:

- $f\left(q_{1}, \ldots, q_{n}\right) \rightarrow q$, with $q_{1}, \ldots, q_{n}, q \in Q, f \in \Sigma$ and $f\left(q_{1}, \ldots, q_{n}\right) \in T_{\Sigma(Q)}$ (for $n=0 f$ can be a constant)
- $q \rightarrow q^{\prime}$, with $q, q^{\prime} \in Q$ (epsilon transition)


## Tree Automata (III)

Notice that we can view the transition rules $R$ as ground rewrite rules and can use them to rewrite terms in the term algebra $\mathbb{T}_{\Sigma(Q)}$. Notice also that we have an inclusion $T_{\Sigma} \subseteq T_{\Sigma(Q)}$. We then define the language $L(\mathcal{A})$ as the subset $L(\mathcal{A}) \subseteq T_{\Sigma}$ of those $t \in T_{\Sigma}$ such that there is a $q \in Q_{0}$ such that $t \rightarrow_{R}^{*} q$. A subset $L \subseteq T_{\Sigma}$ is called regular iff there is a finite-state tree automaton $\mathcal{A}$ such that $L=L(\mathcal{A})$.

Tree automata have the same decidablity results as string automata: they are closed under Boolean operations (we can construct automata for each such operation); and we can decide whether $L(\mathcal{A})$ is empty.

## Tree Automata as Order-Sorted Signatures

Automata are labeled graphs. Tree automata are labeled multigraphs, actually, order-sorted signatures. Any tree automaton $\mathcal{A}=\left(\Sigma, Q, Q_{0}, R\right)$ is, graphically, the order-sorted signature: $\Sigma_{\mathcal{A}}=\left((Q,<), F_{\mathcal{A}}\right)$, where, by definition,

$$
q<q^{\prime} \Leftrightarrow q \rightarrow_{R_{\epsilon}}^{+} q^{\prime}
$$

where $R_{\epsilon}$ are the epsilon transitions (w.l.o.g. may assume $R_{\epsilon}$ terminating). And for each $q_{1}, \ldots, q_{n}, q \in Q$ we have:

$$
f\left(q_{1}, \ldots, q_{n}\right) \rightarrow q \text { in } R \quad \Leftrightarrow \quad f: q_{1}, \ldots, q_{n} \rightarrow q \text { in } F_{\mathcal{A}}
$$

If the accepting states are $Q_{0}=\left\{q_{1}, \ldots, q_{k}\right\}$, then the regular language $L(\mathcal{A})$ is:

$$
L(\mathcal{A})=T_{\Sigma_{\mathcal{A}, q_{1}}} \cup \ldots \cup T_{\Sigma_{\mathcal{A}, q_{k}}}
$$

This is because, for each $q \in Q, t \in T_{\Sigma_{\mathcal{A}, q}}$ iff $t \rightarrow_{R}^{*} q$

## Tree Automata for Sufficient Completeness

The key observation is that, for theories $(\Sigma, E \cup B)$ satisfying conditions (i)-(iii), the following sets of ground $\Sigma$-terms are regular sets:

- the set $D_{s}$ of terms of sort $s$ having a defined symbol on top and constructor terms as arguments;
- the set $C_{s}$ of constructor terms of sort $s$; and
- the set Red of terms reducible by the equations $E$ (modulo B), i.e., terms not in normal form.

Under conditions (i)-(iii) $(\Sigma, E \cup B)$ is sufficiently complete iff for each sort $s$ we have $D_{s}-\left(\operatorname{Red} \cup C_{s}\right)=\emptyset$, which can be decided by deciding emptiness of the corresponding tree automaton.

## Tree Automata Modulo $B$

Of course, in general we need to consider tree automata modulo $B$, that is, tuples $\mathcal{A}=\left(\Sigma, B, Q, Q_{0}, R\right)$, where $\left(\Sigma, Q, Q_{0}, R\right)$ is an ordinary tree automaton, and $B$ is a set of equational $\Sigma$-axioms such as associativity, commutativity, and identity of some symbols. The language of $\mathcal{A}$ is then a subset $L(\mathcal{A}) \subseteq T_{\Sigma}$, now defined by rewriting with $R$ modulo $B$. That is, $L(\mathcal{A}) \subseteq T_{\Sigma}$ is the set of those $t \in T_{\Sigma}$ such that there is a $q \in Q_{0}$ such that $t \rightarrow_{R / B}^{*} q$.

Hendrix, Ohsaki, and Viswanathan show that the tree automata decidablity results generalize to the modulo $B$ case, for $B$ any combination of associativity and/or commutativity and/or identity, except associativity without commutativity.

## Tree Automata Modulo $A$ (II)

Even for the case of associativity alone, or associativity and identity alone, for which some tree automata questions like emptiness become undecidable, the sufficient completeness problem can still be decided in practice for many cases of interest by specialized heuristic algorithms (Hendrix, Ohsaki, and Viswanathan, Proc. RTA 2006, Springer LNCS).

All this means that in practice we can decide the sufficient completeness of most left-linear unconditional order-sorted specifications of interest.

## An Example

To see how the desired tree automata needed to decide sufficient completeness can be built, we can use a simple example, our usual unsorted specification for addition for the Peano natural numbers with a single sort Nat, with 0 and $s$ as constructors, and with equations $x+0=x$ and $x+s(y)=s(x+y)$. This specification satisfies conditions (i)-(iii), since it is left-linear, confluent, sort-decreasing, and terminating.

To recognize each of the regular sets $R e d, D_{N a t}$, and $C_{N a t}$ we need three tree automata $\mathcal{A}_{R e d}, \mathcal{A}_{D_{N a t}}$, and $\mathcal{A}_{C_{N a t}}$.

## An Example (II)

The tree automata $\mathcal{A}_{\text {Red }}, \mathcal{A}_{D_{\text {Nat }}}$, and $\mathcal{A}_{\text {Ctor }}$ have the same signature $\Sigma$ (that of the natural numbers), set of states
$Q=\left\{\right.$ Nat, Red, $D_{N a t}$, Ctor, Zero, NzCtor $\}$, and transitions $R$ :

- $s(N a t) \rightarrow N a t, N a t+N a t \rightarrow N a t$, and ( $\epsilon$-transitions): Ctor $\rightarrow$ Nat, Red $\rightarrow$ Nat, $D_{\text {Nat }} \rightarrow$ Nat, Zero $\rightarrow$ Ctor, NzCtor $\rightarrow$ Ctor.
- (constructor transitions): $0 \rightarrow$ Zero, $s($ Ctor $) \rightarrow$ NzCtor
- (defined function transition): Ctor + Ctor $\rightarrow D_{\text {Nat }}$
- (reducibility transitions): Ctor + Zero $\rightarrow$ Red, Ctor + NzCtor $\rightarrow$ Red .

They only differ in their respective accepting state: Red, $D_{\text {Nat }}$, and Ctor.

## An Example (III)

The point now is that each Boolean operation on regular tree languages has a corresponding operation on their associated tree automata. Therefore, out of the automata $\mathcal{A}_{\text {Red }}, \mathcal{A}_{D_{\text {Nat }}}$, and $\mathcal{A}_{\text {Ctor }}$ we can construct an automaton that recognizes the Ianguage $D_{N a t}-(R e d \cup C t o r)$. Let us call this automaton $\mathcal{A}_{D_{N a t}-(\text { Red } \cup C t o r)}$. Now we know that under conditions (i)-(iii) our specification is sufficiently complete iff $D_{N a t}-(R e d \cup C t o r)=\emptyset$. Therefore, we can decide this property by testing $\mathcal{A}_{D_{N a t}-(R e d \cup C t o r)}$ for emptiness. If the test (as for this example) succeeds, we are done. If it doesn't, we get very useful counterexample terms, showing us where sufficient completeness fails.

## The Maude SCC Tool

The Maude Sufficient Completeness Checker (SCC) is a tool developed by Joseph Hendrix at UIUC. It uses a library of tree automata modulo $B$ operations also developed by him, and reduces the sufficient completeness problem of specification ( $\Sigma, E \cup B$ ) satisfying conditions (i)-(iii) to the emptiness problem for the tree automaton $\mathcal{A}_{D_{s}-\left(\text { Red } \cup C_{s}\right)}$ for each sort $s$ in $\Sigma$. It outputs either "success" or a set of counterexample terms.

Instructions to acces SCC can be found in the course web page. Its use is essentially very simple. One: (1) loads the module scc.maude; (2) loads the module to be checked, say FOO; (3) types "select SCC-LOOP ." and "loop init-scc ." and (4) gives to the SCC the command "(scc FOO .)".

## The Maude SCC Tool (II)

We can illustrate the use of the SCC with some examples already encountered previously in the course. Consider the module

```
fmod NATURAL is
sort Nat .
op 0 : -> Nat [ctor] .
op s : Nat -> Nat [ctor] .
op _+_ : Nat Nat -> Nat .
vars X Y : Nat .
eq X + O = X .
eq X + s(Y) = s(X + Y).
endfm
```


## The Maude SCC Tool (III)

This module is indeed successfully checked by SCC:
Maude> load scc.
Maude> in natural .
fmod NATURAL
Maude> select SCC-LOOP .
Maude> loop init-scc .
Starting the Maude Sufficient Completeness Checker.
Maude> (scc NATURAL .)
Checking sufficient completeness of NATURAL ...
Warning: This module has equations that are not
left-linear. The sufficient completeness checker will
rename variables as needed to drop the non-linearity
conditions.
Success: NATURAL is sufficiently complete under the
assumption that it is weakly-normalizing, confluent, and sort-decreasing.

## The Maude SCC Tool (IV)

## Consider the module

```
fmod MY-LIST is
    protecting NAT .
    sorts NzList List .
    subsorts Nat < NzList < List .
    op _;_ : List List -> List [assoc] .
    op _;_ : NzList NzList -> NzList [assoc ctor] .
    op nil : -> List [ctor] .
    op rev : List -> List .
    eq rev(nil) = nil .
    eq rev(N:Nat) = N:Nat .
    eq rev(N:Nat ; L:List) = rev(L:List) ; N:Nat .
endfm
```


## The Maude SCC Tool (V)

when checked by the SCC gives us the counterexample
Maude> load scc
Maude> in mylist
fmod MY-LIST
Maude> select SCC-LOOP .
Maude> loop init-scc .
Starting the Maude Sufficient Completeness Checker.
Maude> (scc MY-LIST .)
Checking sufficient completeness of MY-LIST ...
Warning: This module has equations that are not
left-linear. The sufficient completeness checker will rename variables as needed to drop the non-linearity conditions.
Failure: The term 0 ; nil is a counterexample as it is a irreducible term with sort List in MY-LIST that does not have sort List in the constructor subsignature.

## The Maude SCC Tool (VI)

We can correct this problem revising our module:

```
fmod MY-LIST2 is
    protecting NAT .
    sorts NzList List .
    subsorts Nat < NzList < List .
    op _;_ : List List -> List [assoc] .
    op _;_ : NzList NzList -> NzList [assoc ctor] .
    op nil : -> List [ctor] .
    op rev : List -> List .
    eq rev(nil) = nil .
    eq rev(N:Nat) = N:Nat .
    eq rev(N:Nat ; L:List) = rev(L:List) ; N:Nat .
    eq nil ; L:List = L:List .
    eq L:List ; nil = L:List .
endfm
```


## The Maude SCC Tool (VII)

which is now successfully checked by SCC:
Maude> load scc
Maude> in mylist2
fmod MY-LIST2
Maude> select SCC-LOOP .
Maude> loop init-scc .
Starting the Maude Sufficient Completeness Checker.
Maude> (scc MY-LIST2 .)
Checking sufficient completeness of MY-LIST2 ...
Warning: This module has equations that are not
left-linear. The sufficient completeness checker will
rename variables as needed to drop the non-linearity
conditions.
Success: MY-LIST2 is sufficiently complete under the assumption that it is weakly-normalizing, confluent, and sort-decreasing.


[^0]:    ${ }^{\text {a }}$ If $B$ contains associativity axioms, equations $E$ should first be generalized to B-match "with extension," Cf. §4.8 in All About Maude.

