CS 576 Lecture 8 2/9/17

J. Meseguer

1. The Boundedness Property

Recall the addition example, where we saw that the term $N+M$ did not have a finite complete set of variants. Thinking in terms of the rewriting sequences that we can associate to instances of $N+M$ by normalized substitutions, this has to do with the fact that such instances can have arbitrary length:

$$N+M\{M\mapsto 0\} = N+0 \xrightarrow[\vec{E}]{} N$$

$$N+M\{M\mapsto s(M')\} = N+s(M') \xrightarrow[\vec{E}]{} s(N+M')$$

$$N+M\{M\mapsto s(s(M''))\} = N+s(s(M'')) \xrightarrow[\vec{E}]{} s(N+s(M')) \xrightarrow[\vec{E}]{} s(s(N+M''))$$

$$\vdots$$

$$N+M\{M\mapsto s^k(M'^k)\} = N+s^k(M'') \xrightarrow[\vec{E}]{k} s^k(N+M'^k)$$

$$\cdots$$

This suggests the following notion as a rewriting-based condition that might help us get a better understanding of the finite variant property (FVP).

Def. A decomposition $(\Sigma, B, \vec{E})$ of $(\Sigma, E\cup B)$ has the boundedness property if $\forall\, t \in T_\Sigma(X)\ \exists\ b(t)\in \mathbb{N}$ s.t. $\forall\, \theta$ $\vec{E},B$-normalized $\exists$ rew. seq. $t\theta \xrightarrow[\vec{E},B]{k} t\theta!$ s.t. $k \leq b(t)$.

Here is the key theorem relating the boundedness property (BP) and FVP:

<u>Theorem</u>. Let $(\Sigma, B, \vec{E})$ be a decomposition of $(\Sigma, E \cup B)$ and let B have a finitary unification algorithm.
Then $(\Sigma, B, \vec{E})$ is BP iff it is FVP.

<u>Proof</u>. ($\Rightarrow$). By BP, for any variant (u, θ) of a term t we have a rewrite sequence $t\theta \xrightarrow{k}{}^{!}_{\vec{E},B} (t\theta)!_{\vec{E},B} =_B u$

with $k \leq b(t)$. But then the proof of the Theorem in pg 11 of Lecture 7 shows that for t $\vec{E},B$-reducible (resp. $\vec{E},B$-red.) the set $\{(u, \eta|_{vars(t)}) \mid t \overset{\nu}{\underset{\vec{E},B}{\leadsto}} u$ of length $\leq b(t)\}$
(resp. that set plus $(t\rho, \rho)$, ρ idempotent renaming of $vars(t)$)
is a complete set of variants. But all such sequences appear at depth $\leq b(t)$ in the <u>finitely branching</u> narrowing tree to t (since B-unification is finitary), which shows that we have a complete and <u>finite</u> set of $\vec{E},B$-variants of t, so $(\Sigma, B, \vec{E})$ is FVP.

($\Leftarrow$). Let $t \in T_\Sigma(X)$ be any term, and let $\{(u_1, \theta_1), \ldots, (u_n, \theta_n)\}$ be a complete finite set of $\vec{E},B$-variants of t. We then have rewrite sequences $t\theta_i \xrightarrow{m_i} !_{\vec{E},B} =_B u_i$

Let $b(t) = max\{m_i \mid 1 \leq i \leq n\}$. Now consider any $\vec{E},B$-normalized substitution γ and consider the $\vec{E},B$-variant $(t\gamma!_{\vec{E},B}, \gamma|_{vars(t)})$. We then have some (u_i,θ_i), $1 \leq i \leq n$, such that $(u_i,\theta_i) \sqsupseteq_B (t\gamma!_{\vec{E},B}, \gamma|_{vars(t)})$.

But this means that there is a ρ such that

$$\gamma|_{vars(t)} =_B (\theta_i\rho)|_{vars(t)} \quad \text{and} \quad t\gamma!_{\vec{E},B} =_B u_i\rho$$

therefore we have:

$$\begin{array}{ccc} t\theta_i\rho & \xrightarrow[\vec{E},B]{m_i} ! & u_i\rho \\ \|_B & & \|_B \\ t\gamma & \overset{m_i}{\underset{\vec{E},B}{\dashrightarrow}} & t\gamma!_{\vec{E},B} \end{array}$$

by the strict B-coherence of $(\Sigma, B, \vec{E})$, and since $m_i \leq b(t)$ this proves that $(\Sigma, B, \vec{E})$ is BP. $\square$

<u>Q</u>: Can we do better than this? Can we further simplify the conditions needed to check FVP? <u>A</u>: Yes!

Theorem A decomposition $(\Sigma, B, \vec{E})$ of $(\Sigma, E \cup B)$ is FVP iff for each $f : s_1 \cdots s_n \to s$ in Σ the term $f(x_1 : s_1, \ldots, x_n : s_n)$ has a finite complete set of variants.

Proof. $(\Rightarrow)$ is trivial. To see $(\Leftarrow)$, by the theorem in pg. 2 it is enough to show that if each $f(x_1, .., x_n)$ has a finite complete set of variants, $(\Sigma, B, \vec{E})$ is BP. This can be proved by structural induction on $t \in T_\Sigma(X)$. If $t = x \in X$, then obviously $b(t) = 0$, and if $t = a$ is a constant, we can pick as $b(a)$ the length k of any sequence $a \xrightarrow{k}!_{\vec{E},B} a!_{\vec{E},B}$.

For the induction step consider $t = f(v_1, \ldots, v_n)$. Reasoning as in the proof of $(\Leftarrow)$ in the Theorem in pg. 2, if the reduction lengths of the complete finite set of variants of $f(x_1, .., x_n)$ are $m_1, \ldots, m_\ell$, we can choose $b(f(x_1, .., x_n)) = \max\{m_i \mid 1 \leq i \leq \ell\}$ as a bound.

Let now $b(v_1), .., b(v_n)$ be the bounds given by the induction hypothesis for $v_1, \ldots, v_n$. Then we can choose $b(f(x_1,..,x_n)) + \sum_{i=1}^{n} b(v_i)$ as bound for $f(v_1,..,v_n)$.

Indeed, let θ be any $\vec{E},B$-normalized substitution.

Then we have:

$$f(v_1,..,v_n)\theta = f(v_1\theta,..,v_n\theta) \xrightarrow[\vec{E},B]{\leq \sum_{1}^{n} b(v_i)} f((v_1\theta)!_{\vec{E},B},\ldots,(v_n\theta)!_{\vec{E},B}) \xrightarrow[\vec{E},B]{\leq b(f(x_1..x_n))} (f(v_1,..,v_n)\theta)!_{\vec{E},B}$$

as desired. □

<u>Remark</u>. Since, as we shall see, if a term t has a finite complete set of $\vec{E},B$-variants and B is linear and regular and has a finitary B-unification algorithm, there is an algorithm that will <u>terminate</u> yielding a finite set of such most general $\vec{E},B$-variants of t, and this has been implemented in Maude for combinations of assoc, comm, id, (except assoc without comm), this gives us a simple and practical <u>semi-decision procedure</u> to determine whether a decomposition $(\Sigma, B, \vec{E})$ is FVP.

2. The Folding Variant Narrowing Strategy

Suppose a decomposition $(\Sigma, B, \vec{E})$ of $(\Sigma, E \cup B)$ is FVP. How can we <u>compute</u> a <u>finite</u> complete set of variants of a term t? The problem is that the set of pairs $(u, \eta|_{var(t)})$ such that $t \overset{\eta}{\rightsquigarrow}{}^{*} u$ is a complete set of variants of t, but:

1. may be infinite
2. we may never <u>know</u> when to stop (at which depth n of the narrowing tree) even if we have <u>already generated</u> a complete set at that depth.

Of course if we <u>know</u> for each $t \in T_\Sigma(X)$ a bound $b(t)$ on the length of normalizing sequences for each $t\theta$ with θ normalized, we can safely stop with a complete set of variants by considering all $(u, \eta|_{var(t)})$ such that $t \overset{\eta}{\rightsquigarrow}{}^{k} u$ with $k \leq n$. However,

a. we may not know $b(t)$ for each t, since this is a theory-specific bound that may require careful analysis of $\vec{E}, B$

b. even if we were to know, this set may be much bigger than necessary, so computing it would be inefficient, and

c. we do not need to know, since we can do much better by:

c.1 Computing a complete set of most general variants, for each t as explained below, and then

c.2 Computing $b(t)$ for each t by (i) first computing a complete set of most general variants of the $f(x_1 .. x_n)$, $f \in \Sigma$, and (ii) then computing $b(t)$ recursively as done in the proof of the Theorem in pg. 4.

The key idea, called folding variant narrowing, consists on filtering the narrowing tree of t to make it much smaller. Let us define:

$$NT^{0}_{\vec{E},B}(t) = \{ t \overset{\rho_0}{\underset{\vec{E},B}{\leadsto}} t\rho \}$$ where ρ is an idempotent variable

renaming (we choose this instead of $t \overset{id_0}{\underset{\vec{E},B}{\leadsto}} t$ to make the substitution idempotent) with $dom(\rho) = vars(t)$

$$NT^{n+1}_{\vec{E},B}(t) = \left\{ t \overset{\eta\mu_{n+1}}{\underset{\vec{E},B}{\leadsto}} u \;\middle|\; t \overset{\eta_n}{\underset{\vec{E},B}{\leadsto}} v \in NT^{n}(t) \wedge t \overset{\eta_n}{\underset{\vec{E},B}{\leadsto}} v \overset{\mu}{\underset{\vec{E},B}{\leadsto}} u \right\}$$

correct narrowing sequence

Likewise, we define $NT^{\leq n}_{\vec{E},B}(t) = \{ t \overset{\eta\ *}{\underset{\vec{E},B}{\leadsto}} v \mid k \leq n \}$, and

$$NT_{\vec{E},B}(t) = \{ t \overset{\eta\ n}{\underset{\vec{E},B}{\leadsto}} v \mid n \in \mathbb{N} \}.$$

we define the following (folding) preorder $\sqsupseteq_B$ on $NT_{\vec{E},B}(t)$:

$$\left(t \overset{\alpha\ *}{\underset{\vec{E},B}{\leadsto}} u \right) \sqsupseteq_B \left(t \overset{\beta\ *}{\underset{\vec{E},B}{\leadsto}} v \right)$$

iff $\exists \rho$ with $dom(\rho) \subseteq ran(\alpha)$ s.t.

1. $(\alpha\rho)|_{vars(t)} =_B \beta|_{vars(t)}$

2. $u\rho =_B v$

We then say that $t \overset{\beta}{\underset{\vec{E},B}{\leadsto}}{}^{*} v$ can be *folded* into $t \overset{\alpha}{\underset{\vec{E},B}{\leadsto}}{}^{*} u$, and that $t \overset{\alpha}{\underset{\vec{E},B}{\leadsto}}{}^{*} u$ is *more general* than $t \overset{\beta}{\underset{\vec{E},B}{\leadsto}}{}^{*} v$ modulo B.

Given a finite subset $\mathcal{D} \subseteq NT_{\vec{E},B}(t)$, if B has a B-matching algorithm we can *effectively compute* a subset $\max_{\sqsupseteq_B}(\mathcal{D}) \subseteq \mathcal{D}$ such that:

(i) $\forall\ t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v \in \mathcal{D}\ \exists\ t \overset{\alpha}{\underset{\vec{E},B}{\leadsto}} u \in \max_{\sqsupseteq_B}(\mathcal{D})$

s.t. $\left(t \overset{\alpha}{\underset{\vec{E},B}{\leadsto}} u\right) \sqsupseteq_B \left(t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v\right)$

(ii) If $t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v\ ;\ t \overset{\gamma}{\underset{\vec{E},B}{\leadsto}} w \in \max_{\sqsupseteq_B}(\mathcal{D})$

and $\left(t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v\right) \sqsupseteq \left(t \overset{\gamma}{\underset{\vec{E},B}{\leadsto}} w\right)$, then

$\left(t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v\right) = \left(t \overset{\gamma}{\underset{\vec{E},B}{\leadsto}} w\right)$.

as a choice of a set of $\sqsupseteq_B$-maximal elements by weeding out non-maximal ones.

Assume that B has a finitary B-unification algorithm, and therefore a B-matching algorithm as well. Then note that we can <u>compute</u> the <u>finite</u> sets

$NT^{n}_{\vec{E},B}(t)$ and $NT^{\leq n}_{\vec{E},B}(t)$, and that we <u>can also compute</u> the subsets

$$FVN^{n}_{\vec{E},B}(t) \subseteq NT^{n}_{\vec{E},B}(t), \quad FVN^{\leq n}_{\vec{E},B}(t) \subseteq NT^{\leq n}_{\vec{E},B}(t),$$ and $FVN_{\vec{E},B}(t) \subseteq NT_{\vec{E},B}(t)$

defined as follows:

- $$FVN^{0}_{\vec{E},B}(t) = FVN^{\leq 0}_{\vec{E},B}(t) = \{t \overset{\rho}{\rightsquigarrow}_{\vec{E},B} t\rho\}$$
- $$FVN^{\leq n}_{\vec{E},B}(t) = \bigcup_{0 \leq i \leq n} FVN^{i}_{\vec{E},B}(t)$$
- $$FVN^{n+1}_{\vec{E},B}(t) = \max_{\sqsupseteq_B} \left\{ t \overset{\eta\gamma}{\rightsquigarrow}{}^{n+1}_{\vec{E},B} w \;\middle|\; \begin{array}{l} (\eta\gamma)|_{vars(t)}\ \vec{E},B\text{ normal} \wedge \\ t \overset{\eta}{\rightsquigarrow}{}^{n}_{\vec{E},B} u \in FVN^{n}_{\vec{E},B}(t) \wedge \\ t \overset{\eta}{\rightsquigarrow}{}^{n}_{\vec{E},B} u \overset{\gamma}{\rightsquigarrow}_{\vec{E},B} w \text{ correct} \\ \text{narr. seq.} \wedge \\ \nexists\, t \overset{\alpha}{\rightsquigarrow}{}^{*}_{\vec{E},B} v \in FVN^{\leq n}_{\vec{E},B}(t) \\ \text{s.t. } (t \overset{\alpha}{\rightsquigarrow}_{\vec{E},B} v) \sqsupseteq_B (t \overset{\eta\gamma}{\rightsquigarrow}{}^{n+1}_{\vec{E},B} w) \end{array} \right\}$$
- $$FVN(t) = \bigcup_{n \in \mathbb{N}} FVN^{n}(t)$$

Completeness Lemma. For $(\Sigma, B, \vec{E})$ as above, $t \in T_\Sigma(X)$, θ an $\vec{E},B$-normalized substitution with $dom(\theta) \subseteq vars(t)$, $n \in \mathbb{N}$, and $t\theta \xrightarrow[\vec{E},B]{n} w$ there exists $\left(t \overset{\alpha}{\underset{\vec{E},B}{\rightsquigarrow^*}} u\right) \in FVN^{\leq n}_{\vec{E},B}(t)$

and normalized substitution γ with $dom(\gamma) \subseteq ran(\alpha)$ s.t.

(i) $(\alpha\gamma)|_{vars(t)} =_B \theta$

(ii) $u\gamma =_B w$

Proof. By induction on n. For $n = 0$ we can just choose $\gamma = \rho^{-1}\theta$ for $t \overset{\rho}{\rightsquigarrow}{}^0 t\rho$ since then $(\rho\rho^{-1}\theta)_{vars(t)} = \theta$, and $t\rho\rho^{-1}\theta = t\theta$. Suppose we have $t\theta \xrightarrow[R,B]{n} v \xrightarrow[R,B]{} w$.

By the induction hypothesis we then have $\left(t \overset{\alpha}{\underset{\vec{E},B}{\rightsquigarrow}} u\right) \in FVN^{\leq n}_{\vec{E},B}(t)$, say $\left(t \overset{\alpha}{\underset{\vec{E},B}{\rightsquigarrow}}{}^k u\right) \in FVN^{k}_{\vec{E},B}(t)$ for some $k \leq n$. But since γ is normalized and $u\gamma =_B v$ we then also have a narrowing sequence

$u \overset{\eta}{\underset{\vec{E},B}{\rightsquigarrow}} w_0$ and a normalized substitution δ with $dom(\delta) \subseteq ran(\eta)$

such that $t \overset{\alpha\eta}{\underset{\vec{E},B}{\rightsquigarrow}}{}^{k+1} w_0$ is a correct narrowing sequence,

$(\alpha\eta\delta)|_{vars(t)} =_B \theta$, and $w_0\delta =_B w$.

But then, since θ is $\vec{E},B$ normalized, $(\alpha\eta)|_{vars(t)}$ is also $\vec{E},B$ normalized. Therefore two things can happen:

(i) $\exists\ t \overset{\beta}{\underset{\vec{E},B}{\rightsquigarrow}}{}^{*} v \in FVN^{\leq k}_{\vec{E},B}(t)$ s.t.

$(t \overset{\beta}{\underset{\vec{E},B}{\rightsquigarrow}} v) \sqsupseteq_B (t \overset{\alpha\eta}{\rightsquigarrow}{}^{k+1} w_0)$ so that we

have ρ' with $(\beta\rho')|_{vars(t)} =_B (\alpha\eta)|_{vars(t)}$,

$dom(\rho') \subseteq ran(\beta)$, and $v\rho' =_B w_0$, (which by $(\alpha\eta)|_{vars(t)}$ $\vec{E},B$-normalized makes ρ' $\vec{E},B$-normalized, and therefore $\rho'\delta$ is also normalized and we have

$$(\beta(\rho'\delta))|_{vars(t)} = (\alpha\eta\delta)|_{vars(t)} =_B \theta$$

and $v(\rho'\delta) =_B w_0\delta =_B w$, proving the lemma, or

(ii) otherwise, we must have

$$t \overset{\beta}{\underset{\vec{E},B}{\leadsto}}{}^{*} v \in FVN^{k+1}(t) \quad \text{such that, again,}$$

$$\left(t \overset{\beta}{\underset{\vec{E},B}{\leadsto}} v\right) \sqsupseteq_B \left(t \overset{\alpha\eta}{\underset{\vec{E},B}{\leadsto}}{}^{k+1} w_0\right), \text{ which by}$$

the exact same reasoning as in case (i) again proves the lemma. $\square$

The key property about FVN is the following

<u>Termination Theorem</u>. For $(\Sigma, B, \vec{E})$ a decomposition of $(\Sigma, E \cup B)$ with B having a finitary B-unification algorithm, if $t \in T_\Sigma(X)$ has a finite complete set of variants, $(u_1, \theta_1), \ldots, (u_n, \theta_n)$ and m_k is the biggest length of the length-minimal sequences $t\theta_i \xrightarrow[\vec{E},B]{m_i} t\theta_i!_{E,B} =_B u_i$, then $FVN^{m_k+1}(t) = \emptyset$.

<u>Proof</u>. Suppose $t \overset{\alpha\eta}{\underset{\vec{E},B}{\leadsto}}{}^{m_k+1} u \in FVN^{m_k+1}(t)$. Then $(\alpha\eta)|_{vars(t)}$ is $\vec{E},B$-normalized by construction and we have a rewrite sequence $t\alpha\eta \xrightarrow[\vec{E},B]{\leq m_k} u$. Therefore, by the Completeness Lemma we must have $\left(t \overset{\beta}{\underset{\vec{E},B}{\leadsto}}{}^{*} v\right) \in FVN^{\leq m_k}_{\vec{E},B}(t)$

and a normalized substitution γ with $dom(\gamma) \subseteq ran(\beta)$ such that

$$\text{(i)} \quad (\beta\gamma)|_{vars(t)} = (\alpha\eta)|_{vars(t)}$$

$$\text{(ii)} \quad v\gamma =_B u$$

which shows that $\left(t \overset{\beta i}{\underset{E,B}{\leadsto}} v\right) \sqsupseteq_B \left(t \overset{\alpha\eta\; m_k+1}{\underset{\vec{E},B}{\leadsto}} u\right)$,

contradicting $\left(t \overset{\alpha\eta\; m_k+1}{\underset{\vec{E},B}{\leadsto}} u\right) \in FVN^{m_k+1}(t)$. $\square$

<u>Theorem (Completeness)</u>. Given $(\Sigma, B, \vec{E})$ decomposition of $(\Sigma, E \cup B)$, where B has a finitary B-unification algorithm.

$$\langle\!\langle t \rangle\!\rangle_{\vec{E},B} = \{ (u, \eta|_{vars(t)}) \mid (t \overset{\eta\,*}{\leadsto} u) \in FVN_{\vec{E},B}(t) \wedge u = u!_{\vec{E},B} \}$$ is

a complete set of $\vec{E},B$-variants of t. Furthermore, t has <u>finite</u> complete set of variants iff there exists a k s.t. $FVN^{k+1}_{\vec{E},B}(t) = \emptyset$ so that $FVN_{\vec{E},B}(t) = FVN^{k}_{\vec{E},B}(t)$. In particular, $(\Sigma, B, \vec{E})$ is in FVP iff $\forall (f : s_1 \ldots s_n \to s) \in \Sigma$ $FVN_{\vec{E},B}(f(x_1{:}s_1, \ldots, x_n{:}s_n))$ always terminates after a finite number of iterations, so that $\langle\!\langle f(x_1 \ldots x_n) \rangle\!\rangle_{\vec{E},B}$ is finite.

Proof. Apply the Completeness Lemma to the rewrite

$t\theta \rightarrow^{!}_{E,B} (t\theta)!_{E,B} =_B w$ for any variant (w,θ) of t

we find $(t \overset{\alpha}{\rightsquigarrow}{}^{*}_{\vec{E},B}\, u) \in FVN_{\vec{E},B}(t)$ such that $u = u!_{\vec{E},B}$

and $(u, \alpha|_{vars(t)}) \sqsupseteq_B (w,\theta)$. (Note that $\alpha|_{vars(t)}$ is $\vec{E},B$-normalized, both by def. of FVN and because $(\alpha\gamma)|_{vars(t)} =_B \theta$ and θ is $\vec{E},B$-normalized).

By the Termination Theorem, if t has a finite complete set of variants, then there is a k such that $FVN^{k+1}_{\vec{E},B}(t) = \emptyset$, so $FVN_{\vec{E},B}(t)$ terminates as $FVN^{k}_{\vec{E},B}(t)$. Conversely, if $FVN_{\vec{E},B}(t)$ terminates at k, then it is a finite set and, by the above argument it generates a finite, complete set of variants $\langle\!\langle t\rangle\!\rangle_{\vec{E},B}$.

Therefore, $(\Sigma, B, \vec{E})$ is FVP iff $FVN_{\vec{E},B}(t)$ terminates for all t, which by the theorem in pg. 4 is equiv. to $FVN_{\vec{E},B}(f(\vec{x}))$ terminating for each $f \in \Sigma$. □

Corollary. Let $\langle\!\langle t\rangle\!\rangle_{\vec{E},B}$ be finite, and define $[\![t]\!]_{\vec{E},B} = \max_{\sqsupseteq_B}(\langle\!\langle t\rangle\!\rangle_{\vec{E},B})$. Then $[\![t]\!]_{\vec{E},B}$ is a complete set of most general variants of t.