CS 576 Lecture 31 5/2/17

J. Meseguer

1. Convex Theories

The only source of non-determinism in the NO inference system is the <u>split</u> rule. Convex theories make this rule unnecessary because in such theories if

$$T_i \models \varphi_i \wedge \bigwedge E \Rightarrow x_1 = y_1 \vee \ldots \vee x_n = y_n$$

then there is a k, $1 \leq k \leq n$ such that $T_i \models \varphi_i \wedge \bigwedge E \Rightarrow x_k = y_k$.

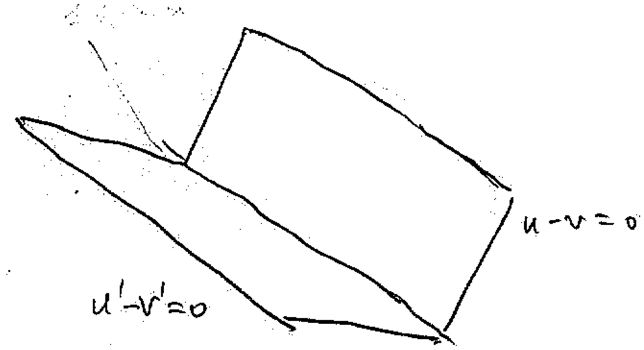
The importance of convex theories was identified by Oppen in his seminal paper, "Complexity, Convexity and Combination of Theories," TCS, (12) 291-302 (1980). Here is the general definition:

<u>Definition</u>. Let Σ be an OS-FO theories which, WLOG we may assume of the form $\Sigma = ((S, \leq), F, \emptyset)$. Let $\Sigma_0 \subseteq \Sigma$ be a subsignature, and $\mathcal{F} \subseteq \bigwedge Lit(\Sigma)$, $\mathcal{G} \subseteq At(\Sigma_0)$. A theory T with $sig(T) = \Sigma$ is called $\mathcal{F}/\mathcal{G}$-<u>convex</u> iff $\forall \varphi \in \mathcal{F}$ $\forall$ $u_i = v_i \in \mathcal{G}$

$$T \models \varphi \Rightarrow u_1 = v_1 \vee \ldots \vee u_n = v_n \text{ implies } (\exists k, 1 \leq k \leq n)\ T \models \varphi \Rightarrow u_k = v_k.$$

When $\mathcal{F} = \bigwedge Lit(\Sigma)$ and $\mathcal{G} = At(\Sigma_0)$, T <u>is</u> called Σ_0-<u>convex</u>.

A well-known example of a convex theory is $(\{+, -, 0, 1, \geq\}, \{\mathbb{Q}_{+,-}\})$, the theory of linear arithmetic, where "convex" has a geometrical meaning in the vector space $\mathbb{Q}^m$, where m is the number of variables in the formula, and where each linear equation $u = v$ is equivalent to the hyperplane equation $u - v = 0$, so that, geometrically, the disjunction $u=v \vee u'=v'$ denotes the union of hyperplanes:



Two well-known counterexamples are:

1. Presburger arithmetic, $(\{+, 0, 1, \geq\}, \{\mathbb{N}_{+,\geq}\})$, where for
$\varphi \equiv 1 \leq z \wedge z \leq 2 \wedge x = 1 \wedge y = 2$ we have
$\mathbb{N}_{+,\geq} \models \varphi \Rightarrow z = x \vee z = y$
but $\mathbb{N}_{+,\geq} \not\models \varphi \Rightarrow z = x$, and $\mathbb{N}_{+,\geq} \not\models z = y$.

2. The theory of extensional arrays, where if
$\varphi \equiv A[i := v][j] = v$ we have
$T_{Array} \models \varphi \Rightarrow (i = j \vee A[j] = v)$, but
$T_{Array} \not\models \varphi \Rightarrow i = j$, and $T_{Array} \not\models \varphi \Rightarrow A[j] = v$.

The NO Inference System Revisited

For T_1, T_2 with signatures Σ_1, Σ_2 optionally intersecting on $\Sigma_1 \cap \Sigma_2 = (S_0, =_0)$ with $S_0 = \{s_1, \dots, s_n\}$ and $s_1, \dots, s_n$ stably infinite sorts in T_1 and T_2, if at least one of them is S_0-convex, the non-determinism of the split rule can be reduced, and if both are S_0-convex split can be eliminated altogether. The idea is that, instead of a pair $\{\varphi_1, \varphi_2\}$, we can keep a pair of pairs:

$$\{(\varphi_1, b_1), (\varphi_2, b_2)\}, \quad \text{where } b_i = \begin{cases} 0 & \text{if } T_i \text{ not } S_0\text{-convex} \\ 1 & \text{if } T_i \text{ is } S_0\text{-convex} \end{cases}$$

Then equality propagation and sat-check rules are slightly modified to disregard the flags b_1, b_2, and split rule now takes the form:

split

$$[\{(\varphi_i,0),(\varphi_j,b_j)\}, E, X^{1,2}] \to [\{[(\varphi_i,0)],[(\varphi_j,b_j)]\}, E \cup \{x_1 = y_1\}, X^{1,2}] \vee \dots \vee [\{(\varphi_i,0),(\varphi_j,b_j)\}, E \cup \{x_n = y_n\}, X^{1,2}]$$

if $\exists\, s_1, \dots, s_n \in S_0$, $x_i = y_i \in X^{1,2}_{s_i}$, $1 \leq i \leq n$, $x_i \not\equiv y_i$, $x_i = y_i \notin E$

$$T_i \models \varphi_i \wedge \bigwedge E \Rightarrow x_1 = y_1 \vee \dots \vee x_n = y_n$$

This rule: (i) agrees with the original split rule when $b_1 = b_2 = 0$;
(ii) remains sound and complete if some $b_j = 1$,

since then $T_j \models \varphi_j \wedge \bigwedge E \Rightarrow x_1 = y_1 \vee \cdots \vee x_n = y_n$

implies $T_j \models \varphi_j \wedge \bigwedge E \Rightarrow x_k = y_k$ for some $1 \leq k \leq n$,

so that equality propagation can be applied.

Note that this also means that if $[\{(\varphi_i, b_i), (\varphi_j, b_j)\}, E, X^{1,2}]$ is in canonical form by the equations equality propagation and sat-check and $b_j = 1$, we must have

$$T_j \not\models \varphi_j \wedge \bigwedge E \Rightarrow x_1 = y_1 \vee \cdots \vee x_n = y_n$$

for any such disjunction of equalities between S_0-variables, with $x_\ell = y_\ell \notin E$, $1 \leq \ell \leq n$;

(iii) If $b_1 = b_2 = 1$, split never has to be applied, so that the NO inference system consists in computing the canonical form of the original problem $[\{(\varphi_1, 1), (\varphi_2, 1)\}, \emptyset, X^{1,2}]$ with the equations equality propagation and sat-check, which can in fact be combined into a single, more efficient equation using an if-then-else operator.

It is of course of great value to identify classes of theories that are convex. I shall prove in detail convexity for two classes:

(1) Horn Theories (a result due to Tinelli (see readings))

(2) Initial Algebras of FVP theories with OS-compact constructors (a restricted result for the constructor reduct).

For both cases, the following technical lemma, proved in the <u>Appendix</u> of the lecture, will be very useful:

<u>Lemma</u> <u>(Substitution Elimination)</u>. Let T be an OS-FO theory with $sig(T) = \Sigma$, and

$$(\hat{\theta} \wedge \psi) \Rightarrow \eta$$

a QF Σ-formula with $vars(\psi \Rightarrow \eta) = X \stackrel{+}{\cup} Y$, $\theta = \{(x_1, t_1), \ldots, (x_n, t_n)\}$ an <u>idempotent</u> substitution with $dom(\theta) = \{x_1, \ldots, x_n\} = X$, and $ran(\theta) \subseteq Y$, and $\hat{\theta}$ the conjunction of equation $\hat{\theta} = \bigwedge_{1 \leq i \leq n} x_i = t_i$, i.e., we regard the substitution as a conjunction.

Then

$$T \models (\hat{\theta} \wedge \psi) \Rightarrow \eta \quad \text{iff} \quad T \models (\psi \Rightarrow \eta)\theta . \quad \square$$

1.2 Horn Theories are Convex

Again, WLOG we may assume that any Horn theory is of the form $T = (\Sigma, E)$, where $\Sigma = ((S, \leq), F, \phi)$, and E is a set of <u>conditional equations</u>. In order to prove that any such T is Σ_0-convex for any

$\Sigma_0 \subseteq \Sigma$, we first prove a well-known lemma about Horn theories.

Lemma (McKinsey's lemma). Let $T = (\Sigma, E)$ be an OS conditional equational theory, and $\{u_i = v_i\}_{i \in I}$ a collection of (ground) Σ-equations such that for each $\mathcal{A} \in \underline{Alg}(\Sigma, E)$ there exist an $i \in I$, such that $\mathcal{A} \models u_i = v_i$. Then there exist $k \in I$ such that $T \models u_k = v_k$.

Proof. Pick $k \in I$ such that $\mathcal{T}_{\Sigma/E} \models u_i = v_k$. Then For each $\mathcal{A} \in \underline{Alg}(\Sigma, E)$ we must have $u_{k_{\mathcal{A}}} = v_{k_{\mathcal{A}}}$ by initiality, so that $\mathcal{A} \models u_k = v_k$, and therefore $T \models u_k = v_k$.

Theorem (Tinelli). Let $T = (\Sigma, E)$ be an OS conditional equational theory. Then T is Σ_0-convex for any $\Sigma_0 \subseteq \Sigma$.

Proof. Suppose $\varphi \in \bigwedge Lit(\Sigma)$, i.e., $\varphi = \bigwedge G \wedge \bigwedge D$ with G Σ-equations and D Σ-disequalities, and we have

$$(\star)\quad T \models \varphi \Rightarrow u_1 = v_1 \vee \cdots \vee u_n = v_n$$

we then have to prove that there is a k, $1 \leq k \leq n$ such that $T \models \varphi \Rightarrow u_k = v_k$.

But by the Theorem of constants ($\circledast$) is equivalent to

$$(\otimes) \quad (\Sigma(X), E) \models \varphi \Rightarrow u_1 = v_1 \vee \dots \vee u_n = v_n$$

where $X = vars(\varphi \Rightarrow u_1 = v_1 \vee \dots \vee u_n = v_n)$, and, WLOG, $X \cap vars(E) = \emptyset$.

Of course, if φ is T-unsatisfiable the theorem trivially follows. Assume, therefore, that φ is T-satisfiable. This exactly means that $\bigwedge G \wedge \bigwedge D$ as a ground formula is $(\Sigma(X), E)$-satisfiable. But, reasoning as in our study of OS-congruence closure (the fact that here the equations are conditional does not matter) we must have:

$$\mathcal{T}_{\Sigma(X)/E \cup G} \models \bigwedge D$$

since otherwise there would be some $w_j \neq w'_j \in D$ such that $\mathcal{T}_{\Sigma(X)/E \cup G} \models w_j = w'_j$ and this would make φ T-unsatisfiable by initiality.

Therefore, (by $\otimes$) $\mathcal{T}_{\Sigma(X)/E \cup G} \models u_1 = v_1 \vee \dots \vee u_n = v_n$,

and therefore we must have a k, $1 \leq k \leq n$ such that $\mathcal{T}_{\Sigma(X)/E\cup G} \models u_K = v_K$, so that

$(\Sigma(X), E\cup G) \models u_K = v_K$.

We will be done if we prove that for each $\mathcal{A} \in Alg(\Sigma, E)$ and $a \in [X \to A]$ such that $(\mathcal{A}, a) \models \wedge G \wedge \wedge D$ that $\mathcal{A} \models u_K = v_K$. But note that $(\mathcal{A}, a) \in Alg(\Sigma(X), E\cup G)$. Therefore, $(\mathcal{A}, a) \models u_K = v_K$, as desired. $\square$

<u>Corollary</u>. For $(\Sigma, \emptyset)$, and (Σ, AC_Δ), the theories of uninterpreted function symbols, resp. function symbols modulo AC_Δ for $\Delta \subseteq \Sigma$, are convex theories.

1.3 Convexity for Var Sat Theories' Constructor Part

Let $(\Sigma, E\cup B)$ have an FVP decomposition $(\Sigma, B, \vec{E})$ protecting an OS-compact constructor decomposition $(\Omega, B_\Omega, \vec{E}_\Omega)$. Recall that in the semantic theory $(\Sigma, \{\mathcal{T}_{\Sigma/E\cup B}\})$ a conjunction $\varphi = \wedge G \wedge \wedge D$ of Σ-equations G and Σ-disequalities D is equivalent to the disjunction of Ω-conjunctions:

$$\circledast \quad \bigvee_{\substack{\theta \in \mathrm{Unif}^{\Omega}_{E \cup B}(\bigwedge G) \\ (\bigwedge D', \rho) \in [\![\bigwedge D \theta]\!]^{\Omega}_{E,B}}} \widehat{(\theta\rho|_{\bar{E},B}\tau)}!_{E,B} D \wedge D' \qquad \tau \in [\mathrm{ran}(\theta\rho) \to T_{\Omega}(X)]$$

in $(\Omega, E_{\Omega} \cup B_{\Omega})$ implies s infinite

where each $\theta\rho\tau$ is an idempotent substitution, $x{:}s \in \mathrm{vars}(D')$, $D' =(D'!_{E_{\Omega},B_{\Omega}})$ is B_{Ω}-consistent and $\mathrm{vars}(D') \subseteq \mathrm{ran}(\theta\rho)$. Let $X = \mathrm{vars}(\varphi)$. If $\mathrm{vars}(G) \subsetneq X$ we may add to G the trivial equations $x = x$ for each $x \in X - \mathrm{vars}(G)$ to make sure $\mathrm{dom}(\theta\rho) = X$. We now need a technical condition:

the OS-compact theory

<u>Definition</u>. An infinite sort s in $(\Omega, E_{\Omega} \cup B_{\Omega})$ is called a <u>hereditarily infinite</u> (H^{∞}) sort iff for each $t \in T_{\Omega}(X)_s$, and each $x{:}s' \in \mathrm{vars}(t)$, s' is an infinite sort in $(\Omega, E_{\Omega} \cup B_{\Omega})$.

Let $\bigwedge \neg H^{\infty}(\Omega)$ denote the conjunctions $\bigwedge_{j \in J} v_j \neq w_j$ such that $ls(v_j)$ and $ls(w_j)$ are H^{∞} sorts, and let $H^{\infty}At(\Omega)$ be the set of all $u = v$ such that $ls(u)$ and $ls(v)$ are H^{∞} sorts.

Here is the main theorem:

<u>Theorem</u>. Under the above assumptions on the OS-compact theory $(\Omega, \{\mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega}\})$, $(\Omega, \{\mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega}\})$ is $\bigwedge \neg H^\infty(\Omega) / H^\infty At(\Omega)$-convex.

<u>Proof</u>. Suppose

$$(\dagger) \quad \Omega, \quad \mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega} \models \left(\bigwedge_{j \in J} v_j \neq w_j \right) \Rightarrow \bigvee_{i \in I} u_i = v_i$$

with $\bigwedge_{j \in J} v_j \neq w_j \in \bigwedge \neg H^\infty(\Omega)$, and $u_i = v_i \in H^\infty At(\Omega)$, $i \in I$. If $\left(\bigwedge_{j \in J} v_j \neq w_j \right) !_{E_\Omega, B_\Omega}$ is not B_Ω-consistent, the compactness result follows trivially. Therefore, we may assume all $v_j \neq w_j$ in E_Ω, B_Ω-canonical form and B_Ω-consistent. But $(\dagger)$ is equivalent to

$$\bigwedge_{j \in J} v_j \neq w_j \wedge \bigwedge_{i \in I} u_i \neq v_i \quad \mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega}\text{-unsatisfiable},$$

by OS-compactness which can only happen if for some $k \in I$ we have $u_k !_{E_\Omega, B_\Omega} = v_k !_{E_\Omega, B_\Omega}$, i.e., if $\mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega} \models u_k = v_k$ and, a fortiori,

$$(\Omega, \{\mathfrak{T}_{\Omega/E_\Omega \cup B_\Omega}\}) \models \left(\bigwedge_{j \in J} v_i = w_j \right) \Rightarrow u_K = v_K. \quad \square$$

A Modified NO Inference System Supporting VarSat (Sketch)

Because of the Presburger arithmetic example, even though an OS-compact $(\Omega, \{\mathcal{T}_{\Omega/E_\Omega \cup B_\Omega}\})$ is convex in the above H^∞ sense, $(\Sigma, \mathcal{T}_{\Sigma/E \cup B})$ is not convex for the corresponding FVP data type. However, the NO inference system can be modified to take full advantage of the convexity of $(\Omega, \{\mathcal{T}_{\Omega/E_\Omega \cup B_\Omega}\})$ as follows:

1. Instead of configurations $[\{(\varphi_1, b_1), (\varphi_2, b_2)\}, E_0, X^{1,2}]$ we can keep configurations $[\{(\varphi_1', b_1), (\varphi_2', b)\}, \emptyset, X^{1,2}]$ where $\varphi_i' = \varphi_i \wedge \bigwedge E_0$

2. If the FVP theory is $(\Sigma_1, E \cup B)$ with OS-compact constructor subsignature $(\Omega, E_\Omega \cup B_\Omega)$ and $\Sigma_1 \cap \Sigma_2 = (S_0, =_{S_0})$ with each $s \in S_0$ an H^∞-sort, what we can do is to keep φ_1' in its "Ω-normal form" (*) (pg. 9), i.e.,

$$\varphi_1' = \bigvee_{i \in I} \hat{\mu}_i \wedge \bigwedge D_i \quad \text{where } \mu_i = \mu_i|_{E_\Omega, B_\Omega} \text{ is } \Omega\text{-subst.}$$

and $\bigwedge D_i = (\bigwedge D_i)!_{E_\Omega, B_\Omega} \in \bigwedge \neg H^\infty(\Omega)$ is

B_Ω-consistent

3. Then, in each application of <u>equality propagation</u> to $(\varphi'_{1}, 0)$ observe that $\mathcal{T}_{\Sigma/E \cup B} \models \varphi'_1 \Rightarrow x = y$

iff $\forall i \in I \quad \mu_i(x) =_{E_\Omega} \mu_i(y)$, since otherwise

$\hat{\mu}_i \wedge \bigwedge D_i \wedge \mu_i(x) \neq \mu_i(y)$ would become $\mathcal{T}_{\Sigma/E \cup B}$-satisfiable, so that $\mathcal{T}_{\Sigma/E \cup B} \not\models \varphi'_1 \Rightarrow x = y$.

4. If, instead the theory T_2 posted an equality $x=y$ after equality propagation and we have a configuration

$$[\{[(\varphi'_1, 0)], (\varphi'_2, b_2)\}, x = y, X^{1,2}]$$

the way <u>sat check</u> is performed is to compute

$$\varphi''_1 = \bigvee_{i \in I} \ \bigvee_{\gamma \in Unif_{E_\Omega \cup B_\Omega}(\mu_i(x) = \mu_i(y))} \widehat{(\mu_i \gamma !_{E_\Omega, B_\Omega})} \wedge (D_i \gamma)!$$

and test whether $\varphi''_1 = \bot$ by checking the Ω-normal form of φ''_1. The resulting conf. is then $[\{(\varphi''_1, 0), (\varphi_2, b_2)\}, \phi, X^{1,2}]$.

4. $\mathcal{J}_{\Sigma/E\cup B} \models \varphi_1' \Rightarrow x_1 = y_1 \vee \cdots \vee x_n = y_n$

iff for each $\forall i \in I$ there is a k_i, $1 \leq k_i \leq n$ such that $\mu_i(x_{k_i}) =_{E_R} \mu_i(y_{k_i})$.

If this test succeeds, then the split rule takes the form:

$$[\{(\varphi_1', 0), (\varphi_2, b_2)\}, \phi, X^{1,2}]$$
$$\to [\{[(\varphi_1', 0)], [(\varphi_2', b_2)]\}, x_1 = y_1, X^{1,2}] \vee \cdots \vee [[(\varphi_1, 0)], [(\varphi_2, b_2)], x_n = y_n, X^{1,2}]$$

where now sat-check has to be applied twice before removing $x_i = y_i$, and, if successful, in both cases, will replace each φ_1' by $\varphi_1''^{(i)}$ as explained in (3).

Appendix: Proof of the Substitution Elimination Lemma

Proof. Let $\mathcal{M} = (M, -_{\mathcal{M}}) \in mod(T)$.

Note that, since the restriction function

$$[X \dot{\cup} Y \to M] \ni a \mapsto a|_Y \in [Y \to M]$$

is surjective an, furthermore, for θ idempotent with $dom(\theta) = X$ and $ran(\theta) = Y$ we have $(\theta; -b)|_Y = b$ for each $b \in [Y \to M]$. Furthermore,

$$\{a \in [X \dot{\cup} Y \to M] \mid a = \theta; -a|_Y\} =$$

$$= \{\theta; -b \in [X \dot{\cup} Y \to M] \mid b \in [Y \to M]\}.$$

We then use these facts as follows:

$$T \models (\hat{\theta} \wedge \psi) \Rightarrow \eta$$

$$\Updownarrow$$

$$T \models \hat{\theta} \Rightarrow (\psi \Rightarrow \eta)$$

$$\Updownarrow$$

$$\forall \mathcal{M} \in mod(T)\ \forall a \in [X \dot{\cup} Y \to M] \quad \mathcal{M}, a \models \hat{\theta} \Rightarrow \mathcal{M}, a \models (\psi \Rightarrow \eta)$$

$$\Updownarrow$$

$$\forall \mathcal{M} \in mod(T)\ \forall a \in [X \dot{\cup} Y \to M] \quad a = \theta; -a|_Y \Rightarrow \mathcal{M}, a \models (\psi \Rightarrow \eta)$$

$$\Updownarrow$$

$$\forall \mathcal{M} \in mod(T)\ \forall b \in [Y \to M] \quad \theta; -b = \theta; -b \Rightarrow \mathcal{M}, \theta; -b \models \psi \Rightarrow \eta$$

$$\Updownarrow$$

$$\forall \mathcal{M} \in mod(T)\ \forall b \in [Y \to M] \quad \mathcal{M}, b \models (\psi \Rightarrow \eta)\theta$$

$$\Updownarrow$$

$$T \models (\psi \Rightarrow \eta)\theta$$, as desired. $\square$