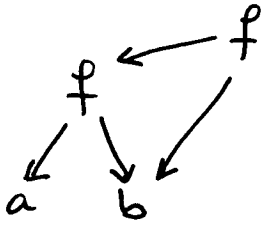
Sat. Procedure for Equality Using Congruence Closure

- First, how do we represent terms, literals
  - as a DAG

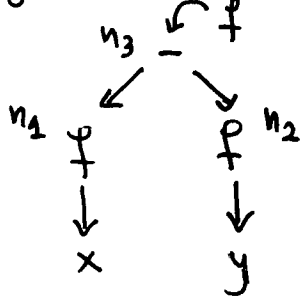
$f(f(a,b),b)$ ---



  - share common subexpressions
  - nodes become convenient names for subexpressions

(useful for separating literals per theory)

$f(f(x) - f(y))$



also a representation of $f(n_3)$

with $n_3 = n_1 - n_2$

$n_1 = f\,x$

$n_2 = f$

  - <u>equalities in the E-DAG</u> — as separate edges

$f(a,b) = a$

  - equalities define a relation R on DAG nodes

- equivalence closure of a relation R
  - the smallest relation R^E such that $R \subseteq R^E$ and
    - for all nodes n $\quad (n,n) \in R^E$
    - $(u_1, u_2) \in R^E \Rightarrow (u_2, u_1) \in R^E$
    - $(u_1, u_2), (u_2, u_3) \in R^E \Rightarrow (u_1, u_3) \in R^E$
  - we draw only the edges of R but we will talk about R^E all the time
- we choose arbitrarily representatives of equivalence classes.

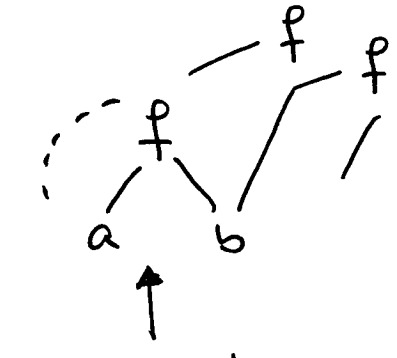
  n^* is the representative of n under the relation R^E

  $(n_1, n_2) \in R^E$ iff $n_1^* \equiv n_2^*$ (same representative)

- congruence closure of an equivalence relation is the smallest $R_E^C \supseteq R_E$ such that
  - for all nodes $f(u_1, \ldots, u_k)$ and $f(m_1, \ldots, m_k)$ in the E-DAG ~~we have~~ such that $(n_i, m_i) \in R^C$

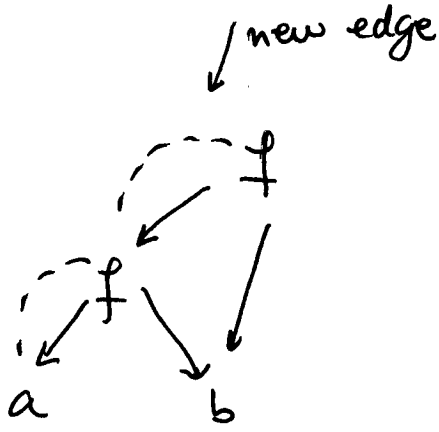
    $\left(f(n_1, \ldots, n_k), f(u_1, \ldots, m_k)\right) \in R^C$
  - we will add edges to the E-DAG to represent R^C
  - Note: R^C must also be an equivalence relation.

Example



$f(a,b) = b$

congruence closure →



infers → $f(f(a,b), b) = a$

- **Note**
  - Congruence closure is an inference procedure for equality
  - Congruence closure always terminates because we do not add nodes
  - We say that $f(t_1, \ldots, t_k)$ is represented in the E-DAG if there is a node $n \equiv f(n_1, \ldots, n_k)$ such that $n_i^* \equiv t_i^*$
  - we further say that $f(t_1, \ldots, t_k)^* \stackrel{\text{def}}{=} n^*$

The sat. proc. for Equality

- Given $F = \bigwedge_i t_i = t_i' \wedge \bigwedge_j u_j \neq u_j'$
- Represent all terms in the E-DAG
- Create $R = \{(t_i, t_i')\}$
- close R under equivalences and congruences $\rightarrow R^c$
- pick representatives for each class in R^c
- F is sat $\iff \forall j \quad u_j^* \not\equiv u_j'^*$

Proof of soundness (sketch)

$\Leftarrow$. Must show a universe and an interpretation ψ such that $\psi(t_i) = \psi(t_i')$ and $\psi(u_j) \neq \psi(u_j')$

- universe is the set of representatives in the E-DAG
- $\psi(t) = t^*$ if t is represented in the E-DAG
- ~~$\psi(f(t_1, \ldots, t_k))$~~

$$\psi(f)(n_1^*, \ldots, n_k^*) = \begin{cases} f(n_1^*, \ldots, n_k^*)^* & \text{if } f(n_1^*, \ldots, n_k^*) \text{ is represented by the E-DAG} \\ \text{arbitrary otherwise.} \end{cases}$$

$\psi(f)$ is well-defined

- because R^c is closed under congruences
- ~~because if t is not represented, then~~

clearly $\psi(t_i) = \psi(t_i')$ because $t_i^* \equiv t_i'^*$
because $(t_i, t_i') \in R$

$\psi(u_j) \neq \psi(u_j')$ because $u_j^* \not\equiv u_j'^*$
(by hypothesis)

- not that t_i, t_i', u_j, u_j' are represented by construction of E-DAG

Proof $\Rightarrow$ ~~by induction on the # of steps in the construction of the congruence closure.~~

- Let ψ an interpretation that satisfies F
then $t_i^* \equiv t'^* \Rightarrow \psi(t) = \psi(t')$
- ~~Proof~~ by induction on the # of steps in the construction of the congruence closure
  - Base step: $(t^* \equiv t'^*) \in R \Rightarrow$ by hypothesis. $\psi(t) = \psi(t')$
  - Inductive step
    Assume $\forall (u, u') \in R' \Rightarrow \psi(u) = \psi(u')$

Case Let $R'' = R' \cup \{(u_1, u_3) \mid (u_1, u_2) \in R', (u_2, u_3) \in R'\}$

Show $\forall (u, u') \in R'' \Rightarrow \psi(u) = \psi(u')$ ✓

Case Let $R'' = R' \cup \{(f(t), f(t')) \mid (t, t') \in R'\}$

$\psi(f(t)) = \psi(f)(\psi(t)) = \psi(f)(\psi(t')) = \psi(f(t'))$ ✓

We can now prove that Equality is convex

- Assume not.
  Let E a conjunction of Equalities.
  and assume $E \vdash E_1 \vee \ldots \vee E_n$ (a disjunction of equalities)

then $E \wedge \neg E_1 \ldots \wedge \neg E_n$ is unsat

but the congruence closure sat. proc. will find one E_i such that $E \wedge \neg E_i$ is unsat

which means that $E \vdash E_i \Rightarrow$ convex theory.

Implementation

- add some fields to nodes
  - root field → points to the representative node
  - class field → points to the set of equivalent nodes.
  - parents field → points to a set of nodes that precede the node in E-DAG (or equiv. nodes)
  - forbid field → points to a set of nodes that are known to be $\neq$ with the current node
- We keep an undoStack whose members are
  $(t_1 = t_2, \text{pf}(t_1 = t_2))$ or
  $(t_1 \neq t_2, \text{pf}(t_1 \neq t_2))$