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)
  \[ f(f(x), f(y)) \]
  \[ f_n_3 \]
  \[ f_n_2 \]
  \[ f_n_1 \]
  \[ f_n \]
  \[ f \]
  \[ f \]
  \[ f \]
  \[ f \]
also a representation of \( f(n_3) \)
with \( n_3 = n_1 - n_2 \)
\( n_1 = f(x) \)
\( n_2 = f(y) \)
- equalities in the E-DAG - as separate edges
\( f(a, b) = a \)
- equalities define a relation \( R \) on DAG nodes
• equivalence closure of a relation $R$
  - the smallest relation $RE$ such that $R \subseteq RE$
    and
    • for all nodes $n$ $(n, m) \in RE$
    • $(u_1, u_2) \in RE \Rightarrow (u_2, u_1) \in RE$
    • $(u_1, u_2), (u_2, u_3) \in RE \Rightarrow (u_1, u_3) \in RE$

  - we draw only the edges of $R$ but we will talk about $RE$ all the time

• we choose arbitrarily representatives of equivalence classes.
  $n^*$ is the representative of $n$
  under the relation $RE$

  $(u_1, u_2) \in RE$ iff $n_1^* = u_2^*$ (some representative)

• congruence closure of an equivalence relation is the smallest $RC \supseteq RE$ such that
  • for all nodes $f(n_1, \ldots, n_k)$ and $f(m_1, \ldots, m_k)$ in the E-DAG we have such that $(n_i, m_i) \in RC$
    \[(f(n_1, \ldots, n_k), f(m_1, \ldots, m_k)) \in RC\]

  • we will add edges to the E-DAG to represent $RC$

• Note: $RC$ must also be an equivalence relation.
Example

\[ f(a, b) = b \]

\[ \text{inference} \quad f(f(a, b), b) = a \]

\* 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 \( f(t_1, \ldots, t_k)^* \equiv n^* \).
The sat. proc. for Equality

- Given $F = \bigwedge_i t_i = t_i' \land \bigwedge_j u_j \neq u_j'$

- Represent all terms in the E-DAG

- Create $R = \{ (t_i, t_i') \}$

- close $R$ under equivalences and congruences $\rightarrow R^e$

- pick representatives for each class $u_i R^e$

- $F$ is sat $\iff \forall j \quad u_j^* \neq u_j'$

Proof of soundness (sketch)

$\Leftarrow$. Must show a universe and an interpretation $\Psi$ 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

- $\Psi(f(n_1^*, ..., n_k^*)) = \left\{ \begin{array}{ll} f(n_1^*, ..., n_k^*) & \text{if } f(n_1, ..., n_k) \text{ is represented by the E-DAG} \\ \text{arbitrary otherwise.} & \end{array} \right.$

$\Psi(f^*)$ is well-defined

- because $R^e$ is closed under congruences
clearly \( \psi(t_i) = \psi(t_i') \) because \( t_i^* = t_i^* \)

because \((t_i, t_i') \in R\)

\( \psi(u_j) \neq \psi(u_j') \) because \( u_j^* \neq u_j'^* \)

(by hypothesis)

- not that \( t_i, t_i', u_j, u_j' \) are repeated
by construction of E-DAG

\[
\begin{align*}
\text{Proof} & \implies \text{by induction on the \# of steps in the} \\
\text{construction of the congruence closure.} \\
\text{- Let } \psi \text{ an interpretation that satisfies } F \text{ then } t_i^* = t_i'^* \implies \psi(t) = \psi(t') \\
\text{- Proof by induction on the \# of steps in the} \\
\text{construction of the congruence closure} \\
\text{ - Base step: } \ (t_i^* = t_i'^*) \in R \implies \text{by hypothesis. } \psi(t) = \psi(t') \\
\text{ - Inductive step} \\
\text{ Assume } \forall (u, u') \in R' \implies \psi(u) = \psi(u') \\
\text{ Case Let } R'' = R' \cup \{(u_1, u_3) \mid (u_1, u_2) \in R', (u_2, u_3) \in R'^1\}. \\
\text{ Show } \forall (u, u') \in R'' \implies \psi(u) = \psi(u') \checkmark \\
\text{ Case Let } R'' = R' \cup \{(f(t), f(t')) \mid (t, t') \in R'^2\}. \\
\psi(f(t)) = \psi(f)(\psi(t)) = \psi(f)(\psi(t')) = \psi(f(t')) \checkmark
\end{align*}
\]
We can now prove that \textit{Equality} is convex.

Assume not.

Let $E$ a conjunction of equalities.

and assume $E \models E_1 \vee \ldots \vee E_n$ (a disjunction of equalities)

then $E \land \neg E_1 \ldots \land \neg E_n$ is unsat.

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

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

\textbf{Implementation}

\begin{itemize}
  \item add some fields to nodes
    \begin{itemize}
      \item root field $\rightarrow$ points to the representative node
      \item class field $\rightarrow$ points to the set of equivalent nodes.
      \item parent field $\rightarrow$ points to a set of nodes that preced the node in E-DAG (or equiv. nodes)
      \item forbid field $\rightarrow$ points to a set of nodes that are known to be $\neq$ with the current node
    \end{itemize}
  \item We keep an undo stack whose members are 
    \begin{align*}
      & (t_1 = t_2, \text{ pf } (t_1 = t_2)) \quad \text{or} \quad (t_1 \neq t_2, \text{ pf } (t_1 \neq t_2))
    \end{align*}
\end{itemize}