Satisfiability Procedures Based on Congruence Closure

Consider the theory of arrays with McCarthy's axioms

- `symbols`: `sel`, `upd`, `=`, `≠`
  - add uninterpreted function symbols since we are going to extend congruence closure

- `axioms`
  
  \[ \forall i \in \text{set}(\text{upd}(a, i, e), i) = e \]  
  \[ \forall i \neq j \Rightarrow \text{set}(\text{upd}(a, i, e), j) = \text{set}(a, j) \]  

Example of a theorem

\[ \text{ upd}(a, i, e) = \text{ upd}(b, i, e) \land \text{ sel}(a, x) \neq \text{ sel}(b, x) \implies x = i \]

Again, like for lists just doing closure for equivalence and axioms is not enough.

\[ \text{closed for Eq, Ax} \]

\[ \text{but seems satisfiable} \]
Again, like for lists we look at two possible solutions.
1) extend the set of axioms
2) add more nodes to the graph

Try 2.
we have an E-DAG that denotes a partial interpretation \( \psi_0 \)
- universe is the set of representatives
  - if it is represented then \( \psi_0(t) = \top^* \)
- \( \psi_0 \) satisfies Eq, Ax
- but is partial

Everything is OK if we can extend \( \psi_0 \) to be total without running into contradictions.

This is tricky. E.g., the graph before we cannot define
\[
\psi(\text{sel}(\text{upd}(a, i, e), x))
\]
because it would have to be equal with
\[
\psi(\text{sel}(a, x)) \quad \text{and} \quad 
\psi(\text{sel}(\text{upd}(b, i, e), x)) \quad \text{and} \quad 
\psi(\text{sel}(b, x))
\]
Impossible, contradicts \( \text{sel}(a, x) \neq \text{sel}(b, x) \).

See, we will have to add nodes (or axioms) for solution.
Incremental extension

- Assume \( \text{upd}(a, i, e) \) is not represented, but \( a \) and \( i, e \) are represented

- There are no new instantiations of the axioms
  - We can assign any value to \( \Psi(\text{upd}(a, i, e)) \) without violating any of the axioms

- Assume \( \text{sel}(a, i) \) is not represented but \( a \) and \( i \) are represented

This suggests adding the axiom

\[ \forall b, i, e. \text{upd}(a, i, e) = \text{upd}(b', i, e') \implies e = e' \]
Another way to fix this problem is to ensure that whenever \( \text{upd}(b, i, e) \) is represented, \( \text{sel}(\text{upd}(b, i, e), i) \) is also represented.

This can be achieved by adding one \( \text{sel} \) node for each \( \text{upd} \) node.

This will prevent the scenario above.

Again, \( \text{sel}(a, i) \) is not represented, but \( a \) and \( i \) are represented.

This can be fixed with axiom

\[
i \neq j \land i \neq j' \Rightarrow \text{sel}(\text{upd}(a, j, e), i) = \text{sel}(\text{upd}(a, j', e'), i)
\]
Another way to fix this problem is to ensure that

**Rule 2**

whenever \( \text{sel}(\text{upd}(a, j, e), i) \) is represented and \( i^* \neq j^* \) then \( \text{sel}(a, i) \) is also represented.

- For each \( \text{sel} \) adds at most one other \( \text{sel} \).
- Yet another possible problem when adding \( \text{sel}(a, i) \).

To fix this we can add the axiom

\[
\text{upd}(b, j, e) = \text{upd}(b', j', e') \land i \neq j \land i \neq j'
\]

\[
\implies \text{sel}(b, i) = \text{sel}(b', i)
\]
Again, we can fix this problem by adding nodes to ensure that

$$\text{whenver } \text{upd}(b,j,e) \text{ and } \text{sel}(b,i) \text{ are represented and } i^* \neq j^* \text{ then}$$

$$\text{sel}\left(\text{upd}(b,j,e), i\right) \text{ is also represented.}$$

We can now show that if we have enough nodes (as specified above) then we can extend \( Y_G \) on unrepresented nodes:

- To add \( \text{upd}(a, i, e) \)
  - Let \( t_1, \ldots, t_n \) be the nodes labelled sel and first successors equivalent to \( a \) and second successors not equivalent to \( i \)
  - Rule 1 \( \Rightarrow \) add \( \text{sel}(\text{upd}(a, i, e), i) = e \)
  - Rule 3 \( \Rightarrow \) add \( \text{sel}(\text{upd}(a, i, e), i_k) = \text{sel}(a, i_k) \)
• to add \( \text{sel}(a, i) \)
  
  • no Rule 1 extensions
  
  • we will add a bunch of \( \text{sel}(b, i) \)
  
  • Rule 3 extensions cannot trigger Rule 2 extensions
  
  • do Rule 2 extensions first

\[
\text{sel}(\text{upd}(\text{upd}(\ldots \text{upd}(a, i_1, e_1), i_2, e_2)\ldots), i_n, e_n), i)
\]

add

\[
\text{sel}(a, i)
\]

\[
\text{sel}(\text{upd}(a, i_1, e_1), i)
\]

\[
\vdots
\]

\[
\text{sel}(\text{upd}(\text{upd}(\ldots)), i)
\]

• do Rule 3 extensions now

This is an example section theory

It can be seen already from Axiom 2

\[
\text{sel}(a, i) = \text{sel}(\text{upd}(a, i, e), j) = \text{sel}(a, j)
\]

• The max number of sel nodes added is \( O(n^2) \)

• for each \( \text{upd}(a, j, e) \) and \( \text{sel}(a, i) \) add

\[
\text{sel}(\text{upd}(b, j, e), i)
\]
The theory of arrays is non-convex

- can be seen already from Axiom 2
  \[ \Rightarrow i = j \lor \text{sel}(\text{upd}(a, i_1, e), j) = \text{sel}(a, j) \]

- Counter also
  \[
  \text{upd}(\text{upd}(\ldots \text{upd}(a, i_1, x), i_2, x), \ldots, i_n, x) = \\
  \text{upd}(\text{upd}(\ldots \text{upd}(b, j_1, x), j_2, x), \ldots, j_{n-1}, x) \
  \]

This literal implies
  \[ \text{sel}(a, i_1) \neq x \land \ldots \land \text{sel}(a, i_n) \neq x \]

This literal implies
  \[ \bigvee_{i_k = i_e} \]

Since the array on the right of equality differs
from a universe places only.

Thus a literal of length \(n\) might cause \(O(n^2)\) splits

- Algorithm
  - add nodes to graph \(O(n^2)\) nodes
  - will have \(O(n^2)\) disequations to consider
  - consider all combinations of disequations
    - if a disequation is assumed to hold
      then the added sel nodes in Rule 2 and 3
      are merged with the sel node that
      triggered the rule

\[ \Rightarrow O(2^{n^2}) \rightarrow \text{Bad} \]

\[ \text{(Bad)} \]