3.2a Congruence Closure (Part 1)

Donnerstag, 29. Oktober 2015 13:00

Special Case: sets of equations $\mathcal{E}$ without variables

Here, it will turn out that $\equiv_{\mathcal{E}}$ is decidable.

Applications:

- compiler construction: optimization of code
  e.g. Common Subexpression problem (identify subexpressions that are "equal" w.r.t. their occurrence in the program). Then these expressions only have to be evaluated once
- program verification
  corresponds to the question of validity of universally quantified formulas of pred. logic

etc.

Def 3.2.1 (Ground Identity)

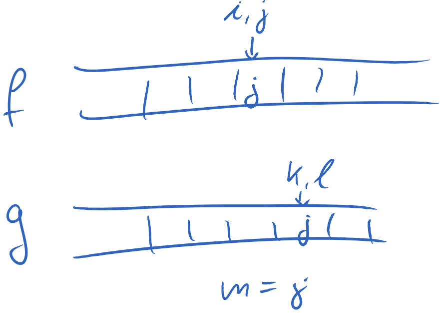
An equation $s \equiv t$ is a ground identity iff $\mathcal{V}(s) = \mathcal{V}(t) = \emptyset$.

Ex 3.2.2 Imperative program with variables

i, j, k, l, m variables for numbers

f, g variables for arrays

```
⋮
i = j;
k = l;
f[i] = g[k];
if (j == f[j]) {
```



$m = g[l];$

$\vdots$ ⟵ Does $f[m] = g[k]$ hold here?
(Yes)

}

So we want to find out whether $f(m) \equiv_{\mathcal{E}} g(k)$ holds for
$\mathcal{E} = \{i \equiv j,\ k \equiv l,\ f(i) \equiv g(k),\ j \equiv f(j),\ m \equiv g(l)\}$.

Here, $i, j, \ldots$ are not variables from $\mathcal{V}$, but constants from Σ_0.

By Birkhoff's Theorem, instead of $f(m) \equiv_{\mathcal{E}} g(k)$, we can check $f(m) \overset{*}{\longleftrightarrow}_{\mathcal{E}} g(k)$.

Goal: Develop an efficient decision procedure that checks $s \equiv_{\mathcal{E}} t$ if $\mathcal{E}$ doesn't contain variables.

Idea: Given $\mathcal{E}$, compute all equations $u \equiv v$ that are entailed by $\mathcal{E}$ (i.e., all equations with $u \equiv_{\mathcal{E}} v$). Then check whether $s \equiv t$ is among these equations.

To compute all equations entailed by $\mathcal{E}$, consider

- reflexivity
- symmetry
- transitivity
- congruence/monotonicity

$\Big\}$ of $\equiv_{\mathcal{E}}$

Def 323 (Direct Consequences of $\mathcal{E}$)

For a set of ground identities $\mathcal{E}$ over a signature Σ, we define:

- R [illegible]

derive:

- $R = \{ t \equiv t \mid t \in \mathcal{T}(\Sigma) \}$
- $S(\mathcal{E}) = \{ t \equiv s \mid s \equiv t \in \mathcal{E} \}$
- $T(\mathcal{E}) = \{ s \equiv r \mid$ there exists a $t \in \mathcal{T}(\Sigma)$ such that $s \equiv t, t \equiv r \in \mathcal{E} \}$
- $C(\mathcal{E}) = \{ f(s_1, \dots, s_n) \equiv f(t_1, \dots, t_n) \mid f \in \Sigma,$ $s_i \equiv t_i \in \mathcal{E}$ for all $1 \leq i \leq n \}$

Ex 324

Let $\mathcal{E} = \{ i \equiv j, k \equiv l, f(i) \equiv g(k), j \equiv f(j), m \equiv g(l) \}$

$R = \{ i \equiv i, j \equiv j, f(i) \equiv f(i), f(g(k)) \equiv f(g(k)), \dots \}$

$S(\mathcal{E}) = \{ j \equiv i, l \equiv k, g(k) \equiv f(i), f(j) \equiv j, g(l) \equiv m \}$

$T(\mathcal{E}) = \{ i \equiv f(j) \}$

$C(\mathcal{E}) = \{ f(i) \equiv f(j), g(i) \equiv g(j), \dots \} \leftarrow$ 10 equations

Clearly, $\mathcal{E} \cup R \cup S(\mathcal{E}) \cup T(\mathcal{E}) \cup C(\mathcal{E})$ only contains equations that are entailed by $\mathcal{E}$, but it does not yet contain all of these equations.

$\Rightarrow$ One has to repeat the application of S, T, and C (potentially infinitely often).

$\Rightarrow$ This results in the congruence closure of $\mathcal{E}$.

Def 325 (Congruence Closure)

For a set of ground identities $\mathcal{E}$ over a signature Σ, we define:

- $\mathcal{E}_0 = \mathcal{E} \cup R$
- $\mathcal{E}_{i+1} = \mathcal{E}_i \cup S(\mathcal{E}_i) \cup T(\mathcal{E}_i) \cup C(\mathcal{E}_i)$

  for all $i \in \mathbb{N}$.

The congruence closure of $\mathcal{E}$ is the "limes" of the sequence $\mathcal{E}_0, \mathcal{E}_1, \mathcal{E}_2, \dots$, i.e.,

$$CC(\mathcal{E}) = \bigcup_{i \in \mathbb{N}} \mathcal{E}_i$$

To check whether $s \equiv_{\mathcal{E}} t$ holds, we now start computing the congruence closure $CC(\mathcal{E})$. If $s \equiv t \in CC(\mathcal{E})$, we return "yes" and otherwise, we return "no".

Questions

1. Is this correct (i.e., does "$s \equiv_{\mathcal{E}} t$ iff $s \equiv t \in CC(\mathcal{E})$" hold?) Yes, see Thm 327

2. In general, the iteration $\mathcal{E}_0, \mathcal{E}_1, \dots$ does not terminate (i.e., $\mathcal{E}_0 \subsetneq \mathcal{E}_1 \subsetneq \mathcal{E}_2 \subsetneq \dots$). Therefore this procedure doesn't terminate if $s \equiv t \notin CC(\mathcal{E})$.

Ex 326: $f(m) \equiv g(k) \in \mathcal{E}_g \subseteq CC(\mathcal{E})$

Thm 327 (Congruence Closure is Sound + Complete)
Let $\mathcal{E}$ be a set of ground identities over Σ, let $s, t \in \mathcal{T}(\Sigma)$. Then: $s \equiv_{\mathcal{E}} t$ iff $s \equiv t \in CC(\mathcal{E})$.

Proof
"$\Leftarrow$" (Soundness)
We have to show that $s \equiv t \in \mathcal{E}_i$ implies $s \equiv_{\mathcal{E}} t$ for all $i \in \mathbb{N}$. We use induction on i.

Ind. Base: $i = 0$
Since $\mathcal{E}_0 = \mathcal{E} \cup R$, this follows from reflexivity of $\equiv_{\mathcal{E}}$ (Lemma 3.1.10).

Ind. Step: $s \equiv t \in \mathcal{E}_{i+1} = \mathcal{E}_i \cup S(\mathcal{E}_i) \cup T(\mathcal{E}_i) \cup C(\mathcal{E}_i)$.
By the induction hypothesis, we have $u \equiv_{\mathcal{E}} v$ for all $u \equiv v \in \mathcal{E}_i$. Then we also have $s \equiv_{\mathcal{E}} t$, since $\equiv_{\mathcal{E}}$ is a congruence relation (Lemma 3.1.10).

"$\Rightarrow$" (Completeness): We postpone this proof, because we will show a stronger statement in Thm 3.2.17.

$\boxed{\mathcal{E}}$

Problem of Question 2 remains:

In general, $\mathcal{E}_i \subsetneq \mathcal{E}_{i+1}$ and thus, the iteration to compute $CC(\mathcal{E})$ does not stop.

$$i \equiv j \in \mathcal{E}_0$$

$$f(i) \equiv f(j) \in \mathcal{E}_1 \setminus \mathcal{E}_0$$

$$f(f(i)) \equiv f(f(j)) \in \mathcal{E}_2 \setminus \mathcal{E}_1$$

$$\vdots$$

$$f^n(i) \equiv f^n(j) \in \mathcal{E}_n \setminus \mathcal{E}_{n-1}$$