Homework #2

CS231

Due by the end of the day on October 14. You should submit three files:
`hw2a.ml`, `hw2b.ml`, and `hw2.pdf`.

Remember that you are encouraged to work in pairs on homework assignments. See the course syllabus for details. Also remember the course's academic integrity policy. In particular, you must credit other people and other resources that you consulted. Again, see the syllabus for details.

1. Like C, OCaml includes terms of the form $\texttt{t}_1$ `&&` $\texttt{t}_2$, which represent short-circuited "and" for booleans: the second boolean expression is evaluated only if necessary. In this problem we add terms of this form to our language of booleans and numbers (see the Homework #2 Cheat Sheet).

   (a) Add rules to the small-step operational semantics to support the new kind of term. You should give a *direct* semantics for this new term rather than simply "desugaring" it to some other kind of term (we'll do that later).

   $$\frac{}{\texttt{true \&\& } \texttt{t}_2 \longrightarrow \texttt{t}_2} \qquad \text{(E-AndTrue)}$$

   $$\frac{}{\texttt{false \&\& } \texttt{t}_2 \longrightarrow \texttt{false}} \qquad \text{(E-AndFalse)}$$

   $$\frac{\texttt{t}_1 \longrightarrow \texttt{t}_1'}{\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2 \longrightarrow \texttt{t}_1' \texttt{ \&\& } \texttt{t}_2} \qquad \text{(E-And)}$$

   (b) Add rules to the type system to support the new operator. Give a name to each new rule.

   $$\frac{\texttt{t}_1\texttt{:Bool} \qquad \texttt{t}_2\texttt{:Bool}}{\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2\texttt{:Bool}} \qquad \text{(T-And)}$$

   (c) Provide only the cases specific to the new $\texttt{t}_1$ `&&` $\texttt{t}_2$ term for the proof of the Progress theorem:

   Theorem: If $\texttt{t} : \texttt{T}$, then either $\texttt{t}$ is a value or there exists some term $\texttt{t}'$ such that $\texttt{t} \longrightarrow \texttt{t}'$.

   Assume that the proof is performed by induction on the derivation of $\texttt{t} : \texttt{T}$. If your proof uses a Canonical Forms lemma, state the lemma clearly before your proof (but you need not prove the lemma).

   Lemma (Canonical Forms): If $\texttt{v}\texttt{:Bool}$, then either $\texttt{v} = \texttt{true}$ or $\texttt{v} = \texttt{false}$.

   Case T-And: Then $\texttt{t} = \texttt{t}_1$ `&&` $\texttt{t}_2$ and $\texttt{t}_1\texttt{:Bool}$ and $\texttt{t}_2\texttt{:Bool}$ and $\texttt{T} = \texttt{Bool}$. By induction, either $\texttt{t}_1$ is a value or there exists some $\texttt{t}_1'$ such that $\texttt{t}_1 \longrightarrow \texttt{t}_1'$. By the Canonical Forms lemma above, if $\texttt{t}_1$ is a value then it is either `true` or `false`. We have several cases:

   - Case $\texttt{t}_1 \longrightarrow \texttt{t}_1'$: Then $\texttt{t}$ steps by E-And.
   - Case $\texttt{t}_1 = \texttt{true}$. Then $\texttt{t}$ steps by E-AndTrue.

- Case $\texttt{t}_1 = \texttt{false}$. Then $\texttt{t}$ steps by E-AndFalse.

(d) Provide only the cases specific to the new $\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2$ term for the proof of the Preservation theorem:

Theorem: If $\texttt{t} : \texttt{T}$ and $\texttt{t} \longrightarrow \texttt{t}'$, then $\texttt{t}' : \texttt{T}$.

Assume that the proof is performed by induction on the derivation of $\texttt{t} : \texttt{T}$.

Case T-And: Then $\texttt{t} = \texttt{t}_1 \texttt{ \&\& } \texttt{t}_2$ and $\texttt{t}_1 : \texttt{Bool}$ and $\texttt{t}_2 : \texttt{Bool}$ and $\texttt{T} = \texttt{Bool}$. Case analysis on the last rule in the derivation of $\texttt{t} \longrightarrow \texttt{t}'$.

- Case E-And: Then $\texttt{t}' = \texttt{t}'_1 \texttt{ \&\& } \texttt{t}_2$ and $\texttt{t}_1 \longrightarrow \texttt{t}'_1$, and by the inductive hypothesis we have $\texttt{t}'_1 : \texttt{Bool}$. Then by T-And also $\texttt{t}'_1 \texttt{ \&\& } \texttt{t}_2 : \texttt{Bool}$, so the result follows.
- Case E-AndTrue: Then $\texttt{t}_1 = \texttt{true}$ and $\texttt{t}' = \texttt{t}_2$. Then the result follows by the fact shown earlier that $\texttt{t}_2 : \texttt{Bool}$.
- Case E-AndFalse: Then $\texttt{t}_1 = \texttt{false}$ and $\texttt{t}' = \texttt{false}$. Then the result follows by T-False.

2. (a) It turns out that terms of the form $\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2$ can be treated as "syntactic sugar," or shorthands for other terms in the language of booleans and numbers. Demonstrate this by providing a new operational semantics for these terms, which consists of a single rule with no premises.

$$\frac{}{\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2 \longrightarrow \texttt{if } \texttt{t}_1 \texttt{ then } \texttt{t}_2 \texttt{ else false}} \quad \text{(E-And)}$$

(b) Consider an eager version of `&&`, in which both operands are evaluated (in order from left to right) before producing the overall value of the term. Is this version still a syntactic sugar? If so, provide the semantics. If not, just say so.

$$\frac{}{\texttt{t}_1 \texttt{ \&\& } \texttt{t}_2 \longrightarrow \texttt{if } \texttt{t}_1 \texttt{ then } \texttt{t}_2 \texttt{ else if } \texttt{t}_2 \texttt{ then false else false}} \quad \text{(E-And)}$$

3. Consider each of the following changes to the simple language of booleans and integers. For each change, say whether it invalidates Progress, Preservation, both, or neither. Also provide a counterexample to each invalidated theorem.

(a) Remove the rule E-IfFalse.

Progress is invalidated. A counterexample is `if false then 0 else 0`, which is well-typed but is stuck.

(b) Add the following axiom to the type system:

$$\frac{}{\texttt{0} : \texttt{Bool}}$$

Progress is invalidated. A counterexample is `if 0 then true else false`, which is well-typed but is stuck.

(c) Add the following axiom to the operational semantics:

$$\frac{}{\texttt{if } \texttt{t}_1 \texttt{ then } \texttt{t}_2 \texttt{ else } \texttt{t}_3 \longrightarrow \texttt{t}_2}$$

Neither theorem is invalidated.

(d) Add the following rules:

$$\frac{}{\texttt{false + false} \longrightarrow \texttt{false}}$$

$$\frac{}{\texttt{true + true} \longrightarrow \texttt{true}}$$

$$\frac{\texttt{t}_1\texttt{:Bool} \qquad \texttt{t}_2\texttt{:Bool}}{\texttt{t}_1 \texttt{ + } \texttt{t}_2\texttt{:Int}}$$

Progress is invalidated. A counterexample is `false + true`, which is well-typed but is stuck. Preservation is also invalidated. A counterexample is `true + true`, which has type `Int` and steps to `true`. However, `true` does not have type `Int`.

(e) Add the following rules:

$$\frac{}{\texttt{if 0 then t}_2 \texttt{ else t}_3 \longrightarrow \texttt{t}_2}$$

$$\frac{\texttt{t}_1\texttt{:Int} \qquad \texttt{t}_2\texttt{:T} \qquad \texttt{t}_3\texttt{:T}}{\texttt{if t}_1 \texttt{ then t}_2 \texttt{ else t}_3\texttt{:T}}$$

Progress is invalidated. A counterexample is `if 1 then true else false`, which is well-typed but is stuck.

4. Consider again the language of just booleans and integers.

   (a) Consider this "reverse progress" theorem:
   Theorem: If $\texttt{t}'\texttt{:T}$, then there exists some term $\texttt{t}$ such that $\texttt{t} \longrightarrow \texttt{t}'$.
   Is this a true theorem? If so, prove it. If not, give a counterexample.
   Proof: Let $\texttt{t}$ be `if true then` $\texttt{t}'$ `else true`. The result follows by E-IfTrue.

   (b) Consider this "reverse preservation" theorem:
   Theorem: If $\texttt{t}'\texttt{:T}$ and $\texttt{t} \longrightarrow \texttt{t}'$, then $\texttt{t}\texttt{:T}$.
   Is this a true theorem? If so, prove it. If not, give a counterexample.
   It is not a true theorem. Let $\texttt{t}' = \texttt{0}$ and $\texttt{T} = \texttt{Int}$ and $\texttt{t} = \texttt{if true then 0 else true}$.

5. Implement the call-by-value lambda calculus (see the Homework #2 Cheat Sheet) in OCaml. See the file `hw2a.ml` for details.

6. Implement the call-by-value lambda calculus in OCaml again, now using a fancy technique called *higher-order abstract syntax*. See the file `hw2b.ml` for details.

7. In class we saw the rules for multi-step evaluation:

$$\frac{}{\texttt{t} \longrightarrow^* \texttt{t}} \qquad \text{(E-Refl)}$$

$$\frac{\texttt{t} \longrightarrow \texttt{t}'}{\texttt{t} \longrightarrow^* \texttt{t}'} \qquad \text{(E-Step)}$$

$$\frac{\texttt{t} \longrightarrow^* \texttt{t}'' \qquad \texttt{t}'' \longrightarrow^* \texttt{t}'}{\texttt{t} \longrightarrow^* \texttt{t}'} \qquad \text{(E-TRANS)}$$

Prove the following theorem for the call-by-value lambda calculus, which says that the term known as Ω only multi-steps to itself.

Theorem: If `((function x -> x x) (function x -> x x))` $\longrightarrow^*$ `t`,
then `t` = `((function x -> x x) (function x -> x x))`.

Proof: By induction on the derivation of `((function x -> x x) (function x -> x x))` $\longrightarrow^*$ `t`. Case analysis of the last rule used in the derivation.

- E-REFL: Then Ω and `t` are the same term, so the result follows.
- E-STEP: Then $\Omega \longrightarrow$ `t`. Case analysis on the last rule in the derivation of this step:
  - Case E-APP1: Then `(function x -> x x)` steps, which is a contradiction.
  - Case E-APP2: Then `(function x -> x x)` steps, which is a contradiction.
  - Case E-APPBETA: Then `t` = [`x` $\mapsto$ `(function x -> x x)`]`(x x)` = Ω.
- E-TRANS Then $\Omega \longrightarrow^*$ `t`$'$ and `t`$' \longrightarrow^*$ `t`. By the induction hypothesis on the first multistep, `t`$' = \Omega$. Then by the induction hypothesis on the second multistep, `t` = Ω.