9 Semantics of Programming Languages (nk480)

Many languages (like C and Java) support coercions, in which values of one datatype (e.g., machine integers) can be used where values of another datatype (e.g., floating point numbers) are expected, by having the compiler silently insert code to convert from one type to another. Suppose we have a language with the following grammar of types:

\[ T ::= \text{int} \mid \text{bool} \mid \text{string} \mid T \times T' \mid T \to T' \]

Suppose we then define a subtyping relation as follows:

\[
\begin{align*}
T \leq T' & \quad T \leq T' \quad T' \leq T'' & \quad T \leq T'' \\
T_1 \leq T_1' \quad T_2 \leq T_2' & \quad T_1' \leq T_1 \quad T_2 \leq T_2' & \quad T_1 \to T_2 \leq T_1' \to T_2' \\
\text{bool} \leq \text{string} & \quad \text{int} \leq \text{string} & \quad \text{bool} \leq \text{int}
\end{align*}
\]

(a) Assuming the existence of functions \texttt{bool\_to\_string}, \texttt{int\_to\_string}, and \texttt{bool\_to\_int}, adapt the relation above to define a new relation \( T \leq T' \sim e \), where \( e \) is a coercion, a closed function of type \( T \to T' \). (You may use ML or lambda-calculus notation to define the coercions \( e \).) [10 marks]

(b) Explain what this relation could be used for in a language implementation. [2 marks]

(c) Give definitions of \texttt{bool\_to\_string} and \texttt{bool\_to\_int}, and then use the relation you defined to give two subtyping derivations \texttt{bool} \leq \texttt{string} \sim e_1 \text{ and } \texttt{bool} \leq \texttt{string} \sim e_2 \text{ such that } e_1 \text{ and } e_2 \text{ have different behaviour.} [5 marks]

(d) What problem would this lead to in a language implementation? [3 marks]