Solutions to Quiz 1

Problem 1 (20 points). We have explained that, given any proof-checking program for a sound proof system for arithmetic inequalities over the integers, we can construct a valid inequality which has no proof in the system. But Ben Bitdiddle (remember him from 6.001? `:-)` asks, "If the inequality has no proof, how can we possibly know it is valid?"

Provide a brief explanation to clear up Ben's confusion.

Solution. Hold on Ben! Nobody said the inequality that isn't provable in a particular sound system has no *mathematical* proof at all. In fact, as soon as you show me a valid inequality that can't be proved in some particular formal proof system, I can actually show you *another* sound formal proof system in which the inequality *can* be proved—namely the original system augmented with the unprovable inequality as an axiom! So there is no such thing as an *inherently unprovable* valid inequality.

Of course there is another valid inequality that is still not provable in the augmented system; so we could also add this new unprovable inequality as another axiom, and so on. But however many times we augment, there will remain a valid inequality unprovable in the extended proof system.

Comment (Going beyond the problem solution.)

OK, so given a sound formal proof system, how do we analyze its limitations well enough to find a valid inequality that is not provable in that particular system? We'll answer this question in a lecture later in the course, using a short elegant argument explaining why the inequality is both unprovable and valid. The argument will be short because it begins with the powerful hypothesis that the given formal proof system is *sound*, *i.e.*, that all the inequalities formally provable in the system are themselves valid.

So how we can know that a given proof system is sound? That is a much harder question in general, and it will bear further discussion later in the course. ■

Problem 2 (20 points). We consider proofs using the standard equational proof rules (Table 1 in the Appendix) for terms over a signature with two symbols, f and g, both of arity 2, a single constant, c. The sole axiom is $f(x, y) = g(x, y)$.

Let F_0 be the term $f(c, c)$, and define $F_{n+1} ::= f(F_n, F_n)$; likewise for G_n. (That is, G_n is the same as F_n with all f's replaced by g's.)

(a) (5 points) Explain how to construct a *sequence-of-equations proof* of length $O(n)$ for the equation $F_n = G_n$.

Solution. From a sequence-of-equations proof of $F_n = G_n$, we can construct a sequence-of-equations proof of $F_{n+1} = G_{n+1}$ by adding three lines:

$$
\begin{aligned}
f(F_n, F_n) &= g(F_n, F_n) && \text{(the (axiom))} \\
g(F_n, F_n) &= g(G_n, G_n) && \text{(by (congruence) from } F_n = G_n) \\
f(F_n, F_n) &= g(G_n, G_n) && \text{(by (transitivity))}
\end{aligned}
$$

The last equation is precisely the desired equation $F_{n+1} = G_{n+1}$. So the length of a proof of $F_n = G_n$ is at most $3n = O(n)$. ■

(b) (15 points) Let $l(n)$ be the length of the *shortest substitution proof* of $F_n = G_n$. Prove that $2^n = O(l(n))$.

Solution. In a substitution proof with this one axiom, only one symbol can change between successive terms — either "g" becomes "f" or vice-versa. Since F_n has 2^n occurrences of "f" and G_n has none, it requires at least 2^n successive terms to arrive at G_n starting from F_n. ■

Problem 3 (30 points). As in Assignment 4, we consider terms with constants R, F, D and a binary operation symbol, $\circ$, with the model, $\mathcal{A}$, being the automorphisms of the square. That is, the domain of $\mathcal{A}$ is the eight automorphisms of the square, and the constants R, F, D mean 90^0 clockwise rotation, reflection about a vertical axis, and reflection about an upper-left/lower-right diagonal, respectively, and the operation symbol, $\circ$, means function composition.

Define another model, $\mathcal{A}'$, whose elements will be pairs (A, s), where A is an automorphism of the square and s is a binary string. The $\circ$ symbol in $\mathcal{A}'$ means the operation on pairs that composes the first coordinates and concatenates the second coordinates. For example,

$$(\text{vertical reflection}, 010) \circ_{\mathcal{A}'} (\text{diagonal reflection}, 11) ::= (90^o \text{ rotation}, 01011).$$

The meanings of the constants R, F, D in $\mathcal{A}'$ will be the pairs $(90^o \text{ rotation}, \lambda)$, $(\text{vertical reflection}, \lambda)$, and $(\text{diagonal reflection}, \lambda)$, respectively, where λ is the empty string.

Let $\mathcal{E}$ be the equational axioms

$$x \circ (y \circ z) = (x \circ y) \circ z, \qquad \text{(associativity)}$$
$$F^2 \circ x = x \qquad \text{(left identity)}$$
$$x \circ F^2 = x \qquad \text{(right identity)}$$

and *all* equations between variable-free terms that are true in $\mathcal{A}$, for example,

$$R^4 = F^2,$$
$$R^4 = D^2,$$
$$FR = D,$$
$$R^3 F = D,$$
$$\vdots$$

(a) (5 points) Explain why $\mathcal{A}' \not\models x^5 = x$.

Solution. Let x be $(R, 0)$. Then $x^5 = (R, 00000) \neq x$, so The equation is not valid. ■

(b) (15 points) Explain why $\mathcal{A}' \models \mathcal{E}$.

Solution. Let $\mathcal{B}$ be the model whose domain is binary strings, the meaning of the symbol $\circ$ is string concatenation, and the meaning of each of the constants R, F, D is the empty string. So the domain of $\mathcal{A}'$ is the product of the domains of $\mathcal{A}$ and $\mathcal{B}$, and the operations of the two parts of $\mathcal{A}'$ work independently.

It follows that

$$\mathcal{A}' \models M = N \quad \longleftrightarrow \quad \mathcal{A} \models M = N \text{ and } \mathcal{B} \models M = N. \tag{1}$$

(Full credit given for stating (1); a rigorous proof was not expected[1].)

Now the axioms were chosen to ensure that $\mathcal{A} \models \mathcal{E}$. But $\mathcal{B} \models \mathcal{E}$ also: (associativity) holds because string concatentation is associative, and all the other axioms are valid in $\mathcal{B}$ because the constants all denote the empty string. So by (1), we conclude that $\mathcal{A}' \models \mathcal{E}$. ■

(c) (10 points) Conclude that $\mathcal{E} \not\vdash x^5 = x$.

---

[1] More precisely, for any term, M, and $\mathcal{A}'$-valuation, V,

$$[\![M]\!]_{\mathcal{A}'} V = ([\![M]\!]_{\mathcal{A}} V_1, [\![M]\!]_{\mathcal{B}} V_2), \tag{2}$$

where V_1 and V_2 are, respectively, the unique $\mathcal{A}$ and $\mathcal{B}$ valuations such that

$$V(x) = (V_1(x), V_2(x))$$

for all variables, x. This follows by an easy structural induction on M.

From (2), we have

$$V \models_{\mathcal{A}'} M = N \quad \longleftrightarrow \quad V_1 \models_{\mathcal{A}} M = N \text{ and } V_2 \models_{\mathcal{B}} M = N,$$

which immediately implies (1).

Solution. Since $\mathcal{A}' \models \mathcal{E}$ by part (b), and the equation $x^5 = x$ is not valid in $\mathcal{A}'$ by part (a), it follows that $\mathcal{E}$ does not semantically imply ($\models$) the equation $x^5 = x$. However, by soundness of the proof system, all the equations provable from $\mathcal{E}$ are semantically implied by $\mathcal{E}$. So $x^5 = x$ cannot be proved from the axioms $\mathcal{E}$. ■

Problem 4 (30 points). Consider ae's extended to include applications

$$((\lambda(x)e)f)$$

A *free occurrence* of a variable x, in an ae, a, is an occurrence of x that is not in a subexpression of the form $(\lambda(x)\dots)$. For example, we highlight in boldface all the free occurrences of variables in the ae

$$\begin{array}{l}([\lambda(y)\\ \quad (\ (\lambda(x)(x+y))\\ \qquad ((y \cdot \mathbf{w}) \cdot \mathbf{x})\ )]\\ ((\mathbf{y} - ((\lambda(z)z)\ \mathbf{x})) \cdot ((\lambda(x)(x-\mathbf{y}))\ 7)))\end{array}$$

(We used square brackets],[instead of parentheses to make it easier to see the scope of $\lambda(y)$.)

The *free variables*, FV(e), of an ae, e, are those variables which have one or more free occurrences in e. For example, letting e_0 be the ae above, we have FV$(e_0) = \{x, y, w\}$.

(a) (10 points) Define FV(e) recursively on the structure of e.

Solution.

$$\begin{aligned}\text{FV}(c) &::= \emptyset,\\ \text{FV}(x) &::= \{x\},\\ \text{FV}(e+f) &::= \text{FV}(e) \cup \text{FV}(f),\\ &\quad \text{likewise for } \cdot \text{ and } -,\\ \text{FV}(((\lambda(x)e)f)) &::= \text{FV}(f) \cup (\text{FV}(e) - \{x\}).\end{aligned}$$

■

(b) (20 points) Prove that if V_1, V_2 are valuations such that

$$V_1(x) = V_2(x) \qquad \text{for all } x \in \text{FV}(e),$$

then

$$[\![e]\!]V_1 = [\![e]\!]V_2. \tag{3}$$

Solution. The proof is by structural induction on e.

Base case (e is a constant, c) We have $[\![c]\!]V_1 = [\![c]\!]_0 = [\![c]\!]V_2$ by definition of $[\![c]\!]$, proving that (3) holds when e is c.

Base case (e is a variable, x) We have $[\![x]\!]V_i = V_i(x)$ for $i = 1, 2$ by definition of $[\![x]\!]$. But $x \in \text{FV}(x)$, so $V_1(x) = V_2(x)$ by hypothesis. This proves that (3) holds when e is x.

Structural induction case (e is $(e_1 + e_2)$) Since $\text{FV}(e_i) \subseteq \text{FV}(M)$, we know that V_1 and V_2 agree on the free-variables of each e_i, so by induction we may assume that $[\![M_i]\!]V_1 = [\![M_i]\!]V_2$ for $i = 1, 2$. Now

$$\begin{aligned}
[\![e_1 + e_2]\!]V_1 &= [\![e_1]\!]V_1 + [\![e_2]\!]V_1 && \text{(def of } [\![e]\!]\text{)} \\
&= [\![e_1]\!]V_2 + [\![e_2]\!]V_2 && \text{(ind. hypothesis)} \\
&= [\![e_1 + e_2]\!]V_2 && \text{(def of } [\![M]\!]\text{).}
\end{aligned}$$

proving that (3) holds when e is $(e_1 + e_2)$.

Structural induction case (e is $(e_1 \cdot e_2)$ or $-e_1$) Essentially the same as for $+$.

Structural induction case (e is $((\lambda(x)f)\ g)$) By the definition of $\text{FV}(e)$, we know that V_1 and V_2 agree on $\text{FV}(g)$ and on $\text{FV}(f) - \{x\}$. So by induction hypothesis for g, we have

$$[\![g]\!]V_1 = [\![g]\!]V_2.$$

Further, $V_1[x \leftarrow n]$ and $V_2[x \leftarrow n]$ agree on $\text{FV}(f)$ for any integer, n. So by induction hypothesis for f, we also have

$$[\![f]\!](V_1[x \leftarrow n]) = [\![f]\!](V_2[x \leftarrow n]). \tag{4}$$

So

$$\begin{aligned}
[\![((\lambda(x)f)\ g)]\!]V_1 &= [\![f]\!](V_1[x \leftarrow [\![g]\!]V_1]) && \text{(def. of } [\![\text{application}]\!]V_1\text{)} \\
&= [\![f]\!](V_1[x \leftarrow [\![g]\!]V_2]) && \text{(ind. hypothesis for } g\text{)} \\
&= [\![f]\!](V_2[x \leftarrow [\![g]\!]V_2]) && \text{(by (4))} \\
&= [\![((\lambda(x)f)\ g)]\!]V_2 && \text{(def. of } [\![\text{application}]\!]V_2\text{)}
\end{aligned}$$

proving that (3) holds in the final case when e is $((\lambda(x)f)\ g)$.

■