Problem Solving in Computer Science · Summer 2007 — Tom Henzinger

1 Defining Correctly

Scribe: Dietmar Berwanger — 13. March 2007

Our handout shows the proof of Theorem 20 in John L. Kelley, *General Topology*, Springer-Verlag 1955. The theorem is correct but there's a bug in the proof, and we shall try to find it.

> THEOREM *If there is a one-to-one function on a set A to a subset of a set B and there is also a one-to-one function on B to a subset of A, then A and B are equipollent.*

1 Definitions

To explain the notions in this statement we proceed backwards, first defining the terms that appear in the theorem, then those used to define these terms, and so on.

Principle Definitions should be structured and simple. They should give clear answers to the questions: What is your goal? What are your assumptions? When we define a term, it may not be necessary to call the statement „Definition“ and assign it a number. Nevertheless, the defined name should be emphasised by setting it in italics or in boldface. The LaTeX-command `\emph{emphasised text}` does just this. Use it wherever you define something and use it sparingly anywhere else.

Intuitively, two sets are equipollent if they have the same number of elements. When we talk about infinite sets we require that their elements can be matched.

> Two sets A and B are *equipollent* $\left\{\begin{array}{c}\text{if}\\ \text{iff}\end{array}\right\}$ there exists a function on A to B $\left\{\begin{array}{c}\text{that}\\ \text{which}\end{array}\right\}$ is one-to-one and onto.

Whether to say "if" or "iff" (if and only if) in definitions is a matter of taste. Just make your choice and stick to it when writing a paper. However, "that" and "which" need in general to be distinguished. Use "that" if what follows defines the object ("the homework that was assigned yesterday"), and "which" if it specifies an additional property ("homework 3, which was assigned yesterday") — these which-clauses need to be separated by commas. In case a phrase occurs between the object to be defined and the clause in question, it is correct to use "which" — then there's no comma ("the homework on graphs which was assigned yesterday"). Hence, any of the four readings of our definition is correct. By naming the function, we get a simple alternative to our initial formulation.

> Two sets A and B are *equipollent* if there exists a function f on A to B such that f is one-to-one and onto.

Principle After every sentence, paragraph, or sequence of paragraphs you have written, ask yourself: Is there any unnecessary word that I can strike out? Won't it sound better if I turned the phrasing around?

You will structure your definitions bottom-up rather than top-down as we do here. Just make sure not to leave terms undefined, and do not forget to remove dangling definitions after eliminating parts of the text.

Here is a first attempt to define one-to-one (not correct yet).

> A function f is *one-to-one* if for all $x, y \in A$, $f(x) \neq f(y)$.

It is crucial to choose intuitive (but short) names for variables, and to maintain a consistent naming scheme. If at all possible, do not use more than two different fonts.

In the above statement $a, b \in A$ would be an awkward choice, because b suggest membership in B. Subscripts a_1, a_2 sometimes help, but they also complicate the notation. Finally, the names x, y may turn out to be more useful somewhere else. To ease renaming we can use macros. Then we shall be very careful to keep the use consistent, especially when writing with co-authors. Inconsistent use of macros (e.g., macros not used for all occurrences of a symbol; multiple macros for the same symbol) are more harmful than not using macros at all.

In the above formulation we have two mathematical expressions standing one next to the other separated only by a comma — this is hard to parse. One way is to say "for all ..., we have ...," but that is not really elegant. By turning around the sentence we can often do better: "$f(x) \neq f(y)$ for all $x, y \in A$". In this case, however, to get the definition right we still need to add $x \neq y$, and this comes to our rescue.

> A function f is *one-to-one* if for all $x, y \in A$, if $x \neq y$, then $f(x) \neq f(y)$.

Q: Why not express the condition in first-order logic?
A: Some readers tend to skip formulas. On the other hand, in text we can also add comments, e.g., "for all points $x, y \in A$". Such comments are often helpful.

> A function f on A to B is *onto* if for all $z \in B$ there exists $x \in A$ such that $f(x) = z$.

We decided to express membership and equality in symbols rather than spelling them out. Tom prefers "for all $x, y \in A$" over both "for all x and y in A" (too verbose) and "$\forall x, y \in A$" (too cryptic). There is a place for formal notion, however, if text would become too cumbersome and ambiguous, e.g., if there are more than two quantifier switches. In this question, too, it is important to be consistent.

2 Debug the proof

The proof in the book rests on the definition of an ancestor:

> A point x of either A or B is an *ancestor* of a point y if y can be obtained from x by successive application of f and g (or g and f).

Is this definition unambiguous?

In the scope of the definition, the functions $f : A \to B$ and $g : B \to A$ are one-to-one, and the sets A and B are assumed to be disjoint. This is how the proof goes. First, the set A is decomposed into three subsets A_E, A_O and A_I according to the following rule. A point x is in A_E, A_O, or A_I if x has an even, odd, or infinite number of ancestors, respectively. Likewise, the set B is decomposed into B_E, B_O, and B_I. Then, the idea is to construct a function $h : A \to B$ which is one-to-one and onto. The author claims that the restrictions $f : A_E \to B_O$ and $f : A_I \to B_I$ together with $g^{-1} : A_O \to B_E$ are onto (and also ono-to-one) mappings, and that collecting them yields the desired function h.

Tom claims this is not true. Our task for the next lecture is to resolve the issue.