PFPL Supplement: Weakening in Statics[*]

Robert Harper

Fall, 2023

The proof of substitution for typing in Chapter 4 of **PFPL** involves an easily overlooked subtlety. Recall, the substitution lemma for the statics states that if $\Gamma \vdash e : \tau$ and $\Gamma, x : \tau \vdash e' : \tau'$, then $\Gamma \vdash \{e/x\}e' : \tau'$. Consider the case of the typing rule for `let`, which is as follows:

$$\frac{\Gamma \vdash e_1 : \tau_1 \qquad \Gamma, y : \tau_1 \vdash e_2 : \tau_2}{\Gamma \vdash \texttt{let}\, y\, \texttt{be}\, e_1\, \texttt{in}\, e_2 : \tau_2}\;\textsc{let}$$

The proof of the substitution lemma proceeds by induction on the second premise. In the case of the LET rule the context Γ is of the form $\Gamma', x : \tau$, where x is the variable being substituted, and we are given $\Gamma' \vdash e : \tau$. Let us assume, for now, that the variable y is distinct from the variable x, though in fact it is this very assumption that gives rise to another difficulty. By induction on the first premise of the above rule we have $\Gamma' \vdash \{e/x\}e_1 : \tau_1$ and it is tempting, but incorrect, to also apply induction to the second premise, to obtain $\Gamma', y : \tau_1 \vdash \{e/x\}e_2 : \tau_2$, from which the intended result follows directly. *The problem is that the inductive hypothesis does not apply to the second premise!* Why? For that to apply we must have that $\Gamma', y : \tau_1 \vdash e : \tau$, but we are only given $\Gamma' \vdash e : \tau$. This gives rise to the need for *weakening*, a structural property of typing, that allows this inference, allowing us to complete the proof of substitution for this case.

To be precise weakening is the statement that, if $\Gamma \vdash e : \tau$ and $x \notin \Gamma$, then $\Gamma, x : \rho \vdash e : \tau$, for any type ρ. Intuitively, the assumption ensures that x is not free in e, and so the addition of x to the context cannot disrupt typing. Although morally correct, the proof of this statement is rather delicate, as was pointed out by Aydemir et al. (2008). The obvious proof method is to proceed by induction on the derivation of $\Gamma \vdash e : \tau$, extending Γ with $x : \rho$ throughout, and concluding the desired result. The proof is entirely straightforward, except for one case, when the variable y in the LET rule is in fact the same as the variable x being added to the context. In that case *it does not make sense to add an additional assumption governing the variable* x! There is no issue if $y \neq x$, but in the special case that $y = x$, the proof breaks down.

Given that abt's are identified up to renaming of bound variables, one might think that the infelicitous choice of y to be x can be avoided by α-renaming of the `let` expression before applying the typing rule, so that no conflict can arise. *But this does not work!* The problem is that the derivation "already" made the choice of y before we came along wanting to weaken by the typing for x, so that there is no opportunity to evade the difficult case! If the derivation in question were altered after the fact to choose $y \neq x$, then the problem is avoided, *except that* the inductive hypothesis does not apply to the altered derivation, only to the given one! And therein lies the rub.

What to do? There are two approaches. The one advocated by Aydemir et al. (2008) is to change the typing rule so that it has *infinitely many* premises, one for each choice of variable x to be used in the typing of the second premise. The choices are limited by the need to avoid variables that are already in use. For this Aydemir et al. (2008) maintain a finite *avoidance set*, and state the LET rule to have a sub-derivation of the body for each of infinitely many variables other than those in the avoidance set. By adding the weakening variable to the avoidance set, that premise is excluded from consideration, and the inductive hypothesis applies to all remaining choices, of which there are infinitely (cofinitely) many.

Another approach in the same spirit is to associate a finite set of variables with each typing derivation consisting of those variables that are introduced by the choice of α-variant of the binders involved. It is then clear that a given derivation may be weakened by any variable not already in Γ and that is not in that set. Finally, to weaken by a variable that lies in that set, simply note that by choosing different α-variants of terms within the derivation it is always possible to avoid the need to avoid that variable, so that the derivation may be weakened as desired. A further variation is to divide the context into two parts, the given free variables, and a deBruijn-form context that is referenced using his eponymous indices. Entering the scope of a binder extends the deBruijn context, and thus avoids interfering with any attempt to extend the free variable context.

An alternative is to accept that the given derivation makes a particular choice of bound variable name that was freely available when the derivation was constructed, but which may no longer be available once the context is weakened by an additional assumption. If it happens that there is no conflict, the weakening proof goes through without change. But what if the chosen variable is precisely the one for which weakening is to be proved? As remarked, renaming the variable ruins the induction. However, one cannot help but notice that the renamed derivation is utterly identical to the given one but for that choice of name. In particular it has the same *height*, when viewed as a tree structure, which suggests changing the proof to proceed by induction on the height, rather than the structure, of the derivation.

But then how is the height to be defined, and is it invariant under renaming? Intuitively, if the height is defined in terms of the number of rule instances in a derivation, taking the max when there are multiple sub-derivations, then it is clearly independent of the choice of variable names in the context, and is thus well-defined and admits weakening. But how can this be made precise? One approach is to index derivations by their height, specifying it to be $n + 1$ if n is the maximum height of the derivations of its premises; uses of assumptions are deemed to be of height 0. One may then prove weakening by induction on the intrinsic height of the derivation, which is independent of the variable context, then define the un-indexed form by existentially quantifying over the index. Note that any un-indexed derivation can be decorated with indices that conform to the above requirements, so that those rules can be used in all other situations besides the proof of weakening. It is a bit heavy-handed to do this in every situation; it is enough to do it once, then consider that an indexed formulation can be given, solely for the sake of proving weakening, with the understanding that the resulting derivation may have certain internal name choices changed for the sake of the induction. This is no problem as far as the derived judgment is concerned; it is only a matter of the route taken to add the additional hypothesis.

References

Brian E. Aydemir, Arthur Charguéraud, Benjamin C. Pierce, Randy Pollack, and Stephanie Weirich. Engineering formal metatheory. In *POPL '08*, 2008.

Robert Harper. *Practical Foundations for Programming Languages*. Cambridge University Press, Cambridge, England, Second edition, 2016.