A sequel to EWD912, i.e. to Draft Ch. 5

[I had restricted the material in EWD912 to what I thought I would need in the next chapter. Writing the latter, I found I wanted to include some program transformations for the justification of which EWD was insufficient. Hence this extension.]

Solutions of an equation of the form

$$Y: [f.Y \equiv Y]$$

are known as "fixpoints of f": for fixpoints of f, the application of f is the identity operation. The Theorem of Knaster-Tarski tells us that for monotonic f, the weakest and strongest fixpoints of f exist. For monotonic functions, taking an extreme fixpoint is order-preserving in the sense expressed precisely in the following lemma.

Lemma 5.12 Let $f0$ and $f1$ be two monotonic predicate transformers, and let the strongest solutions —or the weakest solutions— of

$$Y: [f0.Y \equiv Y] \quad \text{and} \quad Y: [f1.Y \equiv Y]$$

be $E0$ and $E1$ respectively. Then, if

(23) $[f0.Y \Rightarrow f1.Y]$ for all Y,

we have $[E0 \Rightarrow E1]$.

Proof 5.12 We shall prove the lemma for strongest solutions, the conclusion for weakest solutions being the dual one.

$$
\begin{array}{ll}
 & true \\
= & \{ \text{definition of } E1 \} \\
 & [\, f1.\ E1 \equiv E1 \,] \\
\Rightarrow & \{ (23) \text{ with } Y := E1 \} \\
 & [\, f0.\ E1 \Rrightarrow E1 \,] \\
\Rightarrow & \{ \text{def. of } E0 \text{ and Knaster-Tarski} \} \\
 & [\, E0 \Rrightarrow E1 \,] \qquad .
\end{array}
$$

(End of Proof 5.12)

As an aside we remark that the above proof is a perfect illustration of the discussion following Corollary 5.3 : when using that E1 is a solution, we take the equation with the equivalence sign, when using that E0 is an extreme solution, we take the corresponding equation with the implication.

Remark Lemma 5.12 and its proof illustrate another point. As a lemma captures the fruits of the (intellectual or formal) labour that has gone in its proof, it has been conjectured that, the more "laborious" the proof, the more "valuable" the lemma. And as a result it would be hardly worthwhile to take the trouble of formulating a theorem when its proof is very short. But the conjecture is wrong: the "value" of a theorem depends on the ease with which it can be applied. We shall use Lemma 5.12 in

our next proof. (End of Remark.)

By way of introduction to our next lemma we consider for monotonic f the equations

(i) $Y: [f.Y \equiv Y]$ and (ii) $Y: [f^2.Y \equiv Y]$.

On account of Leibniz's Rule, each solution of (i) solves (ii); for this conclusion, f's monotonicity is not needed. It raises the question, whether, conversely, each solution of (ii) solves (i). A counter-example - even with monotonic f - suffices to show that the question deserves a negative answer. (Let z be an integer coordinate of the underlying space and let f be given by $[f.Y \equiv Y_{-z}^{z}]$. In that case, f^2 is the identity function.) A more positive result, however, is that for monotonic f (i) and (ii) have the same extreme solutions; this is a consequence of the next lemma.

Lemma 5.13 For monotonic f and any natural n, the equations

(i) $Y: [(\underline{A}i : 0 \leq i : f^{i+1}.Y) \equiv Y]$

(ii) $Y: [f^{n+1}.Y \equiv Y]$

(iii) $Y: [(\underline{E}i : 0 \leq i : f^{i+1}.Y) \equiv Y]$

have the same extreme solutions; in particular, the extreme solutions of (ii) are independent of n (provided, of course, n is natural).

Proof 5.13 Since the conjugate of f is monotonic as well, it suffices to give the proof for the strongest solutions.

Since the left-hand sides of the equations are all three monotonic, the three strongest solutions exist; let them be, in order, P, Q, and R. Since the left-hand sides of the three equations are weakening in the order given, we have on account of Lemma 5.12

$$[P \Rrightarrow Q] \wedge [Q \Rrightarrow R] \quad .$$

The proof is completed by showing $[R \Rrightarrow P]$ as follows, the range "$0 \leq i$" uniformly having been omitted for brevity's sake.

$$\begin{aligned}
& \text{true} \\
= & \{ f \text{ is monotonic} \} \\
& [f.(\underline{A}i :: f^i.P) \Rrightarrow (\underline{A}i :: f.(f^i.P))] \\
= & \{ \text{definition of functional iteration and renaming the left dummy} \} \\
& [f.(P \wedge (\underline{A}i :: f^{i+1}.P)) \Rrightarrow (\underline{A}i :: f^{i+1}.P)] \\
= & \{ P \text{ solves (i); idempotency of } \wedge \} \\
& [f.P \Rrightarrow P] \\
= & \{ \text{monotonicity of } f \text{ and induction over } i \} \\
& (\underline{A}i :: [f^{i+1}.P \Rrightarrow P]) \\
= & \{ \text{interchange of quanitifications and } \Rrightarrow \text{ "distribution"} \} \\
& [(\underline{E}i :: f^{i+1}.P) \Rrightarrow P] \\
= & \{ R \text{ is strongest solution of (iii) and Knaster-Tarski} \} \\
& [R \Rrightarrow P] \quad .
\end{aligned}$$

(End of Proof 5.13)

Remark Consider the equation

$$Y: [(\underline{Q}\, i: i \in V: f^{i+1}.Y) \equiv Y]$$

with $\underline{Q}$ either $\underline{A}$ or $\underline{E}$ and V a non-empty bag of natural numbers. The predicate on its left is weaker ~~than that of~~ than that of (i) and stronger than that of (iii). With Lemmata 5.12 and 5.13 we conclude that the above equation has the same extreme solutions as (i) and (iii). (End of Remark.)

Are the preceding lemmata rather general, the next one is very specific. It is, in fact, too special to our taste; the reason that we include it nevertheless is that we need it in the next chapter in which the semantics of the repetition will be defined in terms of extreme solutions of an indeed rather specific equation.

Lemma 5.14 Consider for monotonic p and q the equation

$$(24) \qquad Y: [\, p.X \wedge q.Y \equiv Y\,]$$

in which p is, furthermore, weakening and idempotent, i.e

$$[X \Rrightarrow p.X] \quad \text{and} \quad [p^2.X \equiv p.X] \text{ for all } X,$$

then the extreme solutions $g.X$ and $h.X$ of (24) are idempotent.

Proof 5.14 To begin with, we remark, on account of $[g^2.X \equiv g.(g.X)]$ and $g.X$ being the strongest solution of (24), that

$g^2.X$ is the strongest solution of
(25) $\quad Y: [p.(g.X) \wedge q.Y \equiv Y]$

and, similarly,

$h^2.X$ is the weakest solution of
(26) $\quad Y: [p.(h.X) \wedge q.Y \equiv Y]$.

Our proof proceeds in three parts: the first part shows $[g^2.X \Rrightarrow g.X]$ and $[h^2.X \Rrightarrow h.X]$, the two remaining parts will each demonstrate one of the inverse implications.

(i) $[g^2.X \Rrightarrow g.X]$ and $[h^2.X \Rrightarrow h.X]$.

This part of the proof does not rely on the assumption that p is weakening. To begin with we observe for any predicates X, Y, and Z

$$
\begin{array}{ll}
 & [p.X \wedge q.Z \equiv Z] \\
\Rightarrow & \{ \text{predicate calculus} \} \\
 & [Z \Rrightarrow p.X] \\
\Rightarrow & \{ p \text{ is monotonic} \} \\
 & [p.Z \Rrightarrow p^2.X] \\
= & \{ p \text{ is idempotent} \} \\
 & [p.Z \Rrightarrow p.X] \\
\Rightarrow & \{ \text{predicate calculus} \} \\
 & [p.Z \wedge q.Y \Rrightarrow p.X \wedge q.Y]
\end{array}
$$

$\Rrightarrow$ { Lemma 5.12 . and (24) }
[an extreme solution of $Y: [\, p.Z \wedge q.Y \equiv Y \,] \Rrightarrow$
the corresponding extreme solution of (24)] .

The first line of the above holds for Z any solution of (24), in particular for $[Z \equiv g.X]$ and for $[Z \equiv h.X]$. The conclusion of the above is then about extreme solutions of (25) and (26) respectively, in particular $[g^2.X \Rrightarrow g.X]$ and $[h^2.X \Rrightarrow h.X]$. (The other possible conclusions $[g.(h.X) \Rrightarrow g.X]$ and $[h.(g.X) \Rrightarrow h.X]$ are now less interesting.)

(ii) $[g.X \Rrightarrow g^2.X]$

This part of the proof relies directly on p being weakening and indirectly – by an appeal to (i) – on p being idempotent. The proof is by showing that $g^2.X$ solves an equation of which $g.X$ is the strongest solution.

true
$=$ { $g^2.X$ solves (25) }
$[\, p.(g.X) \wedge q.(g^2.X) \equiv g^2.X \,]$
$\Rrightarrow$ { p is weakening }
$[\, g.X \wedge q.(g^2.X) \Rrightarrow g^2.X \,]$
$=$ { $g.X$ solves (24) and Leibniz's Rule }
$[\, p.X \wedge q.(g.X) \wedge q.(g^2.X) \Rrightarrow g^2.X \,]$
$=$ { $[g^2.X \Rrightarrow g.X]$ – see (i) – and q is monotonic }
$[\, p.X \wedge q.(g^2.X) \Rrightarrow g^2.X \,]$
$\Rrightarrow$ { $g.X$ strongest solution of (24) and Knaster-Tarski }
$[g.X \Rrightarrow g^2.X]$.

(iii) $[h.X \Rrightarrow h^2.X]$

This part of the proof does not rely on p being idempotent. The proof is by showing that hX solves an equation of which h^2X is the weakest solution.

true

$= \{$ p is weakening ; $h.X$ solves (24) $\}$

$[h.X \Rrightarrow p.(h.X)] \wedge [h.X \Rrightarrow q.(h.X)]$

$\Rightarrow \{$ predicate calculus $\}$

$[h.X \Rrightarrow p.(h.X) \wedge q.(h.X)]$

$\Rightarrow \{ h^2.X$ weakest solution of (26) ; Knaster-Tarski $\}$

$[h.X \Rrightarrow h^2.X]$.

(End of Proof 5.14)

The previous lemma gives circumstances under which we can conclude that extreme solutions are idempotent. The next lemma gives us a conclusion we may draw if extreme solutions are idempotent.

Lemma 5.15 Let monotonic f be such that for each monotonic, idempotent predicate transformer ip, the extreme solutions of

$$Y: [f.(X, ip.Y) \equiv Y]$$

are idempotent. Let

(27) $$Y: [f.(X, Y) \equiv Y]$$

have $g.X$ and $h.X$ as its strongest and weakest solutions respectively. Then $g.X$ is the strongest

solution of

(28) $Y: [f.(X, g.Y) \equiv Y]$

and $h.X$ is the weakest solution of

(29) $Y: [f.(X, h.Y) \equiv Y]$.

Proof 5.15. Let $e.X$ be a monotonic, idempotent solution of (27); then we have

(30) $[f.(X, e.X) \equiv e.X]$;

let $k.X$ be an idempotent solution of

(31) $Y: [f.(X, e.Y) \equiv Y]$,

so that we have

(32) $[f.(X, e.(k.X)) \equiv k.X]$

We observe firstly

$$
\begin{array}{ll}
 & \text{true} \\
= & \{\text{(30) and idempotency of } e\} \\
 & [f.(X, e.(e.X)) \equiv e.X] \\
= & \{(31)\} \\
 & (e.X \text{ solves } (31))
\end{array}
$$

(1)

We observe secondly

$$
\begin{array}{ll}
 & \text{true} \\
= & \{(32) \text{ with } X := k.X\} \\
 & [f.(k.X, e.(k^2.X)) \equiv k^2.X] \\
= & \{\text{idempotency of } k\}
\end{array}
$$

$$
\begin{array}{ll}
 & [f.(k.X, e.(k.X)) \equiv k.X] \\
= & \{(30) \text{ with } X := k.X\} \\
 & [e.(k.X) \equiv k.X] \\
\Rightarrow & \{\text{Leibniz}\} \\
 & [f.(X, e.(k.X)) \equiv f.(X, k.X)] \\
= & \{(32)\} \\
 & [k.X \equiv f.(X, k.X)] \\
= & \{(27)\}
\end{array}
$$

(ii) ($k.X$ solves (27)) .

Now the hard work has been done. Function f being monotonic, it is -Lemma 3.19- monotonic in its second component as well and -Lemma 5.4- its extreme solutions exist and -Lemma 5.8- are monotonic. Choosing for ip the identity function, we conclude that g and h are furthermore idempotent.

Choosing for e either g or h - e.g. Lemmata 3.14 and 3.20 - the left-hand side of (31) is monotonic as well, the extreme solutions of (31) exist, and - choose monotonic, idempotent e for ip - are idempotent as well. So we may choose for $k.X$ either the strongest or the weakest solution of (31)

Choose g for e ; equation (31) becomes (28), and let $k.X$ be its strongest solution. From (i) $[k.X \Rightarrow g.X]$, from (ii) $[g.X \Rightarrow k.X]$, hence (27) and (28) have the same strongest solution $g.X$.

Choose h for e; equation (31) becomes (29). and let $k.X$ be its weakest solution. From (i) $[h.X \Rightarrow k.X]$, from (ii) $[k.X \Rightarrow h.X]$, hence (27) and (29) have the same weakest solution $h.X$.

(End of Proof 5.15).

Remarks. Since $h^*.X$ is the strongest solution of the conjugate equation of (27), and the weakest solution of (29) is the conjugate of the strongest solution of (29)'s conjugate $Y: [f^*.(X, h^*.Y) \equiv Y]$, the introduction of e could have been avoided: the second half of the conclusion is the dual of the first half.

It should be noted, however, that the above proof does not depend on the Theorem of Knaster-Tarski in the sense that in equations (27) and (31) the equivalence is not replaced by an implication. This circumstance seduced us not to destroy the symmetry and to carry out the formal work in terms of e.

As last remark, we would like to add a word of reassurance for those readers who, at first reading of our latter proofs, have been left with somewhat uneasy or disturbed feelings. They might, for instance wonder what made us conjecture the lemmata in the first place; the answer to this question may be found in the next chapter. They

also might wonder how one ever finds such a proof. Well, the answer is of course: rarely immediately and never immediately in the form in which it sees the light of day.

When you come to think of it, the overall structure of the proofs - i.e. the choice of the subgoals - is not too surprising. The detailed formal calculations are another matter, because the formulae we may use always contain free variables for which you are free to substitute whatever suits your argument. In recognition of the fact that theorem proving is a goal-directed activity one often begins by working backwards; the search of which formulae to use in one's substitutions is often narrowed down by the knowledge of a counterexample from which one can conclude that such and such condition in the antecedent is essential and, hence, has to be used somewhere. And, once a proof has been found, there comes the long process of polishing and editing. One tries to reduce the calculation to its bare essentials, one has to decide whether to capture a pattern of argument into an explicit sublemma, one has to choose one's free variables very carefully, so as to minimize the amount of renaming, one has to decide upon the order and which formulae to give a number for future reference, etc. (A detailed example of such considerations: should you write down - and give a number to - (30) and (32)?

After all, we all know what it means to be a solution to (27) or to (31), don't we? But then you consider, firstly, that in your preliminary explorations you wrote them down yourself - so as not to get confused - and, secondly, that it had been a good thing that you had done so because in some of their usages you substituted something for X. And as it is a good principle never to demand of your reader what you could not do yourself - in this case: substituting in a formula that is nowhere shown - you decide to display the "redundant" (30) and (32) explicitly.) And only when all such decisions have been taken, one can start to write (and even then one sometimes discovers having started to early). (End of Remarks.)

Austin, 5 April 1985

prof. dr. Edsger W. Dijkstra,
Department of Computer Sciences,
The University of Texas at Austin,
Austin, TX 78712-1188
United States of America