On computing a longest path in a tree

by

The Eindhoven Tuesday Afternoon Club[F]

Given a finite tree with all edges having positive length, we wish to compute a longest path. This can be done using the following procedure, which was invented by Edsger W. Dijkstra around the 1960s.

We build a physical model of the tree by connecting each pair of adjacent nodes by a piece of string of the given edge length. Now we pick up the physical tree at an arbitrary node U, let the contraption hang down, and determine a deepest node X. Then we pick up the tree at X and determine a deepest node Y. The claim is that the path between X and Y is a longest path in the tree.

To the best of our knowledge, we have never seen a formal proof of this claim, and the purpose of this note is to provide one.

* * *

F R.W. Bulterman, F.W. van der Sommen, G. Zwaan, A.J.M. van Gasteren, W.H.J. Feijen

We first introduce some nomenclature and notation, and give some elementary properties.

- A tree provides a unique (shortest in number of edges) connection for each pair of nodes; we refer to this connection as the path between these nodes.

- The length of the path between nodes v and w is denoted by vw : it is the sum of the lengths of the edges on the path.

- For all nodes v, w: $\quad vw = wv$.

- As a shorthand we use the phrase "m on vw", standing for "node m is on the path connecting nodes v and w".

- For all nodes m, v, and w, we have

  $\Delta\leq$: $\quad vw \leq vm + mw$, and

  $\Delta=$: $\quad m$ on vw
  $\quad\equiv$
  $\quad vw = vm + mw$.

  (In $\Delta\leq$ we use that all edges have nonnegative length; for $\Delta=$ we need that all lengths are positive.)

So much for our preliminaries.

* * *

In the following m, v, w, z range over arbitrary nodes, whereas U, X, Y are the specific nodes mentioned in the procedure. The procedure can be summarized by the following properties of U, X, and Y:

(0) $\langle \forall z :: Uz \leq UX \rangle$

(1) $\langle \forall z :: Xz \leq XY \rangle$,

and the claim is

$$\langle \forall v, w :: vw \leq XY \rangle \ .$$

We now design a calculational proof, interspersed with some heuristic remarks. For any v, w we have

$vw \leq XY$

$\Leftarrow$ { there is not much else we can do apart from using (1) and the transitivity of $\leq$ }

$\langle \exists z :: vw \leq Xz \rangle$

$\Leftarrow$ { in finding witnesses, there are only five identified nodes available, viz. v, w, U, X, and Y; of these, X makes no sense and Y is absorbed by the demonstrandum }

$vw \leq Xv \ \vee \ vw \leq Xw \ \vee \ vw \leq XU$

$\Leftarrow$ { it is quite unlikely that disjunct $vw \leq XU$ could possibly contribute to the validity of this expression:

node U is arbitrary and vw could be the length of a longest path }

(*) $vw \leq Xv \quad \vee \quad vw \leq Xw$.

And here we are left with an expression that is symmetric in v and w, so that we can afford to focus on one disjunct only.

None of our givens (0) and (1) apply to $vw \leq Xv$, and thus we fall short of manipulative freedom. Additional freedom is usually obtained by parametrization, and it is here that $\Delta\leq$ and $\Delta=$ from our little "theory of trees" come in handy, since these can introduce new nodes. By applying $\Delta\leq$ to vw and $\Delta=$ to Xv – the least committing choice for strengthening $vw \leq Xv$ ∇ –, we obtain

$vw \leq Xv$

$\Leftarrow$ { introduction of nodes m and n, $\bullet^{F}$ n on Xv }

$vm + mw \leq Xn + nv$.

But if we have to proceed from here, we had better decide on $m = n$, because the terms vm and nv then cancel.

With these considerations in mind, our calculation continues as follows:

F $\bullet$ is to be read as: "on the premisse that"

$vw \leq Xv$

$\Leftarrow$ { $\Delta\leq$ on vw, using transitivity of $\leq$; $\Delta =$ on Xv, • m on Xv }

$vm + mw \leq Xm + mv$

$\equiv$ { algebra, using $vm = mv$ }

$mw \leq Xm$

$\equiv$ { introduction of U, heading for (0), which has not been used yet }

$Um + mw \leq Xm + mU$

$\Leftarrow$ { $\Delta\leq$ on righthand side; $\Delta =$ on lefthand side, • m on Uw }

$Uw \leq XU$

$\Leftarrow$ { (0) with $z := w$ }

true .

In summary, we established

$vw \leq Xv$

$\Leftarrow$

$\langle \exists m :: (m \text{ on } Xv) \wedge (m \text{ on } Uw) \rangle$.

By symmetry between v and w, we have for the other disjunct of (*)

$vw \leq Xw$

$\Leftarrow$

$\langle \exists m :: (m \text{ on } Xw) \wedge (m \text{ on } Uv) \rangle$.

So (*) and hence the theorem has been proved whenever we can rely on

$$\langle \exists m :: (m \text{ on } Xv) \wedge (m \text{ on } Uw) \rangle \\ \vee \\ \langle \exists m :: (m \text{ on } Xw) \wedge (m \text{ on } Uv) \rangle ,$$

and it so happens that this is an instance of the general property of trees that

for all nodes A, B, C, D of a tree,

$$\langle \exists m :: (m \text{ on } AB) \wedge (m \text{ on } CD) \rangle \\ \vee \\ \langle \exists m :: (m \text{ on } AD) \wedge (m \text{ on } CB) \rangle ,$$

a property to be added to our little "theory of trees" (and to be shown by the reader).

* * *

The calculation proper consists of eight small calculation steps and apart from the exploitation of the symmetry between v and w, there is no case analysis involved. Nevertheless it took us a very long time to arrive at the above argument. It was only after we banished the use of pictures, both on the blackboard and in our minds, that our design started to converge to the above one. And it was only then that it became clear which little "theory of trees" had to support the solution to the problem proper. It was this separation that did the job.

How often do we still need to tell each other "avoid interpretational thought and the use of pictures like the plague"? Why do we still fall into this trap so often? Apparently we have not yet freed ourselves sufficiently well from our traditional mathematical education. Without our calculational style we would never have arrived at the above design, in which the algorithm's two ingredients (0) and (1) have been used exactly once.

For the ETAC,

W.H.J. Feijen
15 November 2000