A tentative axiomatization of ascending sequences

Yesterday I tried to base my lecture on EWD817 "An introduction to three algorithms for sorting in situ", which I had written with A.J.M. van Gasteren in early 1982. At the time, I remember, I liked that text very much, but yesterday I learned that now, more than 11 years later, the text did not work anymore. I found it too verbose and was several times tempted to resort to pictures: I gave a fairly abominable lecture. So let me try an alternative.

We use capital letters $X, Y, Z, \ldots$ as variables of type "finite sequence", lower case letters $p, q, r, \ldots$ as variables of type "singleton sequence", and ε to denote the empty sequence. Consequently, $X := p$ and $Y := \varepsilon$ are instances of legal instantiations, $p := X$ and $q := \varepsilon$ are not. The associative operation of concatenation is — dangerously and stupidly, but let me sin for once! — denoted by juxtaposition with a binding power higher than functional application, which is denoted by an infix dot.

At least for the time being, I won't axiomatize concatenation; this will not prevent me from carrying out mathematical induction over the length or the grammar of sequences.

(This because mathematical induction over a potentially ambiguous grammar is felt to be a separate issue.) Hence I do not commit myself concerning the status — axiom or theorem — of propositions like

$$XY = X \quad \equiv \quad Y = \varepsilon$$

$$YX = X \quad \equiv \quad Y = \varepsilon$$

$$XY = \varepsilon \quad \equiv \quad X = \varepsilon \wedge Y = \varepsilon$$

$$pX \neq X \qquad Xp \neq X$$

$$pX = qY \vee Xp = Yq \quad \Rightarrow \quad p = q \wedge X = Y \text{, etc.}$$

* * *

I wish to capture "ascending sequences", defined in terms of the relation $\leq$, which is a total order on the singletons:

$$p \leq p \qquad \text{(reflexive)}$$

$$p \leq q \wedge q \leq p \Rightarrow p = q \qquad \text{(antisymmetric)}$$

$$p \leq q \wedge q \leq r \Rightarrow p \leq r \qquad \text{(transitive)}$$

$$p \leq q \vee q \leq p \qquad \text{(total)}$$

but propose to do that via a relation between sequences. Informally $X \prec Y$ says that $p \leq q$ holds for any p from X and any q from Y. The relation $\prec$ — let us pronounce it as "under" — has been introduced in the hope that it will reduce the number of universal quantifications we have to indicate explicitly. The special character $\prec$ has been introduced because $\prec$ has properties very different from $\leq$: $\prec$ is not reflexive, not antisymmetric, not transitive, and not total.

(0) $\quad X \prec \varepsilon \qquad \varepsilon \prec X$

(1) $\quad p \prec q \;\equiv\; p \leq q$

(2) $\quad XY \prec Z \;\equiv\; X \prec Z \wedge Y \prec Z$

(3) $\quad Z \prec XY \;\equiv\; Z \prec X \wedge Z \prec Y$

Relation $\prec$ is not transitive:

$$X \prec Y \wedge Y \prec Z \Rightarrow X \prec Z$$

reduces for $Y := \varepsilon$ on account of (0) to $X \prec Z$, which need not hold. Replace Y by Yq , and

(4) $\quad X \prec Yq \wedge Yq \prec Z \Rightarrow X \prec Z$.

We shall prove (4) to show the pattern. To prove (4) we observe

$$\begin{array}{ll} & X \prec Yq \wedge Yq \prec Z \\ = & \quad \{(3), \text{ twice}\} \\ & X \prec Y \wedge X \prec q \wedge Y \prec Z \wedge q \prec Z \\ \Rightarrow & \quad \{\text{pred. calc.}\} \end{array}$$

$$
\begin{array}{ll}
 & X \prec q \ \wedge \ q \prec Z \\
\Rightarrow & \quad \{(5)\} \\
 & X \prec Z
\end{array}
$$
, where

(5) $\quad X \prec q \ \wedge \ q \prec Z \ \Rightarrow \ X \prec Z$.

represents our remaining proof obligation, which we meet by mathematical induction over (the length of) Z . For the base we observe ($Z := \varepsilon$)

$$
\begin{array}{ll}
 & X \prec \varepsilon \\
= & \quad \{(0)\} \\
 & true \\
\Leftarrow & \quad \{\text{pred. calc.}\} \\
 & X \prec q \ \wedge \ q \prec \varepsilon
\end{array}
$$

For the step we observe ($Z := rZ$)

$$
\begin{array}{ll}
 & X \prec rZ \\
= & \quad \{(3)\} \\
 & X \prec r \ \wedge \ X \prec Z \\
\Leftarrow & \quad \{\text{ex hypothesi, i.e. } (5)\} \\
 & X \prec r \ \wedge \ X \prec q \ \wedge \ q \prec Z \\
\Leftarrow & \quad \{(6)\} \\
 & X \prec q \ \wedge \ q \prec r \ \wedge \ q \prec Z \\
= & \quad \{(3)\} \\
 & X \prec q \ \wedge \ q \prec rZ
\end{array}
$$
where

(6) $\quad X \prec q \ \wedge \ q \prec r \ \Rightarrow \ X \prec r$

represents our remaining proof obligation, which can be met by induction over (the length of) X. For the base $(X := \varepsilon)$ we observe

$$\begin{array}{ll} & \varepsilon \prec q \ \wedge \ q \prec r \ \Rightarrow \ \varepsilon \prec r \\ = & \quad \{ (0), \text{pred. calc} \} \\ & \text{true} \end{array}$$

,

for the step we observe $(X := pX)$

$$\begin{array}{ll} & pX \prec r \\ = & \quad \{ (2) \} \\ & p \prec r \ \wedge \ X \prec r \\ \Leftarrow & \quad \{ \text{ex hypothesi, } (6) \} \\ & p \prec r \ \wedge \ X \prec q \ \wedge \ q \prec r \\ = & \quad \{ (1) \} \\ & p \leq r \ \wedge \ X \prec q \ \wedge \ q \prec r \\ \Leftarrow & \quad \{ \leq \text{ is transitive} \} \\ & p \leq q \ \wedge \ q \leq r \ \wedge \ X \prec q \ \wedge \ q \prec r \\ = & \quad \{ (1) \text{ and pred. calc.} \} \\ & p \prec q \ \wedge \ X \prec q \ \wedge \ q \prec r \\ = & \quad \{ (2) \} \\ & pX \prec q \ \wedge \ q \prec r \end{array}$$

,

with which our proof obligations have been fulfilled. I trust that the above proof structure is typical whenever induction is needed.

* * *

I would like to define the function "asc" of type: sequence $\rightarrow$ bool by

(7) $\quad asc.p$

(8) $\quad asc.XY \equiv asc.X \wedge asc.Y \wedge X \prec Y$,

but I am not quite sure about my proof obligations that this definition of asc makes sense, i.e. does not lead to contradiction. My gut feeling is that I have done my duty when I have shown that "asc" is a function, i.e. satisfies Leibniz's principle

$$X = Y \Rightarrow (asc.X \equiv asc.Y)$$

for all possible ways in which $X = Y$ can hold, where the possible ways are listed by the axioms about equality of sequences.

So we would have to show that (7) ∧ (8) is compatible with

$$asc.X \equiv asc.X\varepsilon \quad \text{and}$$
$$asc.X \equiv asc.\varepsilon X \quad ,$$

the obligation rising from $X = X\varepsilon$ and $X = \varepsilon X$. The obligation is met by defining (concluding?) $asc.\varepsilon$, i.e. the empty sequence is ascending.

So we have to show that (7) ∧ (8) is compatible with

(9) $\quad asc.X(YZ) \equiv asc.(XY)Z$,

the obligation rising from the associativity of concatenation; I expect to prove (9) in view of (8) thanks to the associativity of $\wedge$.

$$
\begin{aligned}
& \quad asc.X(YZ) \\
= & \quad \{(8) \text{ with } Y:=YZ\} \\
& \quad asc.X \wedge asc.YZ \wedge X \prec YZ \\
= & \quad \{(8) \text{ with } X,Y:=Y,Z;\ (3) \text{ with } X,Y,Z:=Y,Z,X\} \\
& \quad asc.X \wedge asc.Y \wedge asc.Z \wedge Y \prec Z \wedge X \prec Y \wedge X \prec Z \\
= & \quad \{(8),(2)\} \\
& \quad asc.XY \wedge asc.Z \wedge XY \prec Z \\
= & \quad \{(8) \text{ with } X,Y:=XY,Z\} \\
& \quad asc.(XY)Z \quad .
\end{aligned}
$$

The above proof is of an almost embarrassing triviality, but it is nice that we did not need mathematical induction.

A simple lemma is

(10) $\quad asc.XYZ \Rightarrow asc.XZ$

which is proved in the same vein as the previous one:

$$
\begin{aligned}
& \quad asc.XYZ \\
= & \quad \{\text{see above}\} \\
& \quad asc.X \wedge asc.Y \wedge asc.Z \wedge Y \prec Z \wedge X \prec Y \wedge X \prec Z \\
\Rightarrow & \quad \{\text{pred. calc.}\} \\
& \quad asc.X \wedge asc.Z \wedge X \prec Z \\
= & \quad \{(8)\} \\
& \quad asc.XZ \quad .
\end{aligned}
$$

Repeated application of (10) tells us that

(11) any subsequence of an ascending subsequence is ascending

Slightly more ambitious is

(12) $\quad asc.XpY \equiv asc.Xp \wedge asc.pY$.

We observe for any X, p, Y,

$$
\begin{array}{ll}
& asc.Xp \wedge asc.pY \\
= & \quad \{(8)\ \text{twice}\} \\
& asc.X \wedge asc.p \wedge asc.Y \wedge \\
& X \prec p \wedge p \prec Y \\
= & \quad \{(4)\} \\
& asc.X \wedge asc.p \wedge asc.Y \wedge \\
& X \prec p \wedge p \prec Y \wedge X \prec Y \\
= & \quad \{(8)\ \text{and}\ (2)\ \text{or}\ (3)\} \\
& asc.XpY \quad .
\end{array}
$$

<u>Remark</u> It was a surprise for me that nothing is gained by proving (12) with a ping-pong argument. (End of Remark.)

And now we are ready to prove the beautiful

(13) $\quad asc.XY \Rightarrow asc.Xp \vee asc.pY$.

Regrettably, I cannot avoid case analysis (i) $Y = \varepsilon$ and (ii) $Y = qZ$.

(i) Since $Y = \varepsilon \Rightarrow asc.pY$ on account of (7) lemma (13) has been proved in this case.

(ii) In this case we rewrite our demonstrandum (13) with $Y := qZ$ by shunting as

$$asc.XqZ \wedge \neg asc.pqZ \Rightarrow asc.Xp$$

and observe for arbitrary p, q, X, Z

$$\begin{array}{ll} & asc.XqZ \wedge \neg asc.pqZ \\ = & \{(12)\text{, twice, and de Morgan}\} \\ & asc.Xq \wedge asc.qZ \wedge (\neg asc.pq \vee \neg asc.qZ) \\ \Rightarrow & \{\text{pred. calc.}\} \\ & asc.Xq \wedge \neg asc.pq \\ \Rightarrow & \{\text{since } -(1),(7),(8)-\ \neg asc.pq \Rightarrow asc.qp\} \\ & asc.Xq \wedge asc.qp \\ = & \{(12)\} \\ & asc.Xqp \\ \Rightarrow & \{(10)\} \\ & asc.Xp \quad , \end{array}$$

which completes the proof of (13) .

* * *

The next theorem to be proved is (of course)

(14) $asc.Z \Rightarrow \langle \exists X,Y : XY = Z : asc.Xp \wedge asc.pY \rangle$

or, in view of (12), equivalently

(14') $asc.Z \Rightarrow \langle \exists X,Y : XY = Z : asc.XpY \rangle$,

a lemma that can be viewed as underlying the feasibility of "insertion sort".

One way of proving this is by mathematical induction over the length of Z. For the base we have to show

$$asc.\varepsilon \Rightarrow \langle \exists U,V :\ UV = \varepsilon :\ asc.UpV \rangle \quad ;$$

the witness $U=\varepsilon, V=\varepsilon$ demonstrates the truth of its consequent. For the step it suffices to show

$$asc.Zq \Rightarrow \langle \exists U,V :\ UV = Zq :\ asc.UpV \rangle \quad ;$$

- if $q \le p$ we can take as witness $U = Zq$, $V = \varepsilon$; $asc.UpV$ reduces to $asc.Zqp$ or (12) to $asc.Zq \wedge asc.qp$, which is implied;
- if $p \le q$, we take XY satisfying (14)

(15) $\quad XY = Z \quad , \quad asc.Xp \quad , \quad asc.pY$

and can take as witness $U = X$, $V = Yq$, and observe, (assuming $asc.Zq \wedge p \le q$)

$$\begin{aligned}
& asc.UpV \\
= \quad & \{(12)\} \\
& asc.Up \ \wedge\ asc.pV \\
= \quad & \{\text{def. of } U,V\} \\
& asc.Xp \ \wedge\ asc.pYq \\
= \quad & \{(15):\ asc.Xp\} \\
& asc.pYq \\
= \quad & \{(8)\} \\
& asc.p \ \wedge\ asc.Yq \ \wedge\ p \prec Yq \\
\Leftarrow \quad & \{(7),(8),(3)\} \\
& asc.XYq \ \wedge\ p \prec Y \ \wedge\ p \prec q
\end{aligned}$$

$\Leftarrow$ $\{(15): XY = Z \ , (8), (1)\}$
$$\text{asc}.Zq \ \wedge \ \text{asc}.pY \ \wedge \ p \leq q$$
$=$ $\{(15): \text{asc}.pY\}$
$$\text{asc}.Zq \ \wedge \ p \leq q \quad ,$$
which was our assumption.

For the sake of completenes we give a different phrasing of virtually the same proof of (14)

Let Y be the <u>shortest</u> sequence such that X and Y furthermore satisfy

(16) $\qquad XY = Z \ \wedge \ \text{asc}.Xp \quad .$

[Such a shortest sequence exists, for (16) has at least 1 solution: $X = \varepsilon \ , Y = Z$] .
Our proof obligation is now -see (14)-

(17) $\qquad \text{asc}.pY \ .$

If $Y = \varepsilon$, (17) holds because of (7).
Otherwise, we write $Y = qU$, and have to conclude

(17′) $\qquad \text{asc}.pqU$

from

(16′) $\qquad XqU = Z \ \wedge \ \text{asc}.Xp$
(18) $\qquad \text{asc}.Z$
(19) $\qquad \neg\, \text{asc}.Xqp$

where (18) is a reminder of the antecedent of (14) and (19) expresses that Y was the <u>shortest</u> sequence meeting the requirement. We observe, using (16'), (18)

$$
\begin{array}{ll}
 & true \\
\Rightarrow & \{(16') \wedge (18)\} \\
 & asc.XqU \ \wedge \ asc.Xp \\
\Rightarrow & \{(8)\} \\
 & asc.Xq \ \wedge \ asc.Xp \\
= & \{(19)\} \\
 & asc.Xq \ \wedge \ asc.Xp \ \wedge \ \neg\, asc.Xqp \\
= & \{(8),(7)\} \\
 & asc.Xq \ \wedge \ asc.Xp \ \wedge \ (\neg\, asc.Xq \ \vee \ \neg\, Xq \prec p) \\
\Rightarrow & \{\text{pred. calc.}\} \\
 & asc.Xp \ \wedge \ \neg\, Xq \prec p \\
\Rightarrow & \{(8), (2), (1)\} \\
 & X \prec p \ \wedge \ (\neg\, X \prec p \ \vee \ \neg\, q \le p) \\
\Rightarrow & \{\text{pred. calc.}\} \\
 & p < q \\
\Rightarrow & \{(1), (7), (8) ; (16'), (18), (8)\} \\
 & asc.pq \ \wedge \ asc.qU \\
= & \{(12)\} \\
 & asc.pqU \ .
\end{array}
$$

I have no distinct preference for the latter phrasing.

Austin, 27 Oct. - 10 Nov. 1993

prof. dr. Edsger W. Dijkstra, Dept. of Computer Sciences
The University of Texas at Austin
Austin, TX 78712-1188, USA