A symmetry in the sorting task

That an integer sequence $f.i\ (0 \le i)$ is ascending is usually stated as

$$(\underline{A}\, i,j : i < j : f.i \le f.j)$$

We can reformulate this in terms of $<$ only as

$$(\underline{A}\, i,j : i < j : \neg(f.j < f.i))$$

or – trading and de Morgan – more symmetrically as

$$(\underline{A}\, i,j :: \neg(i < j \ \wedge\ f.j < f.i)) \quad .$$

Calling the inner conjunction the characterization of a "conflict", we see that "being sorted" can be described as "having no conflicts".

Some remarks are in order.

- The ordering between i,j could differ from the ordering between the f-values, which could be reals, words, figures or what have you. Well, we knew this.

- We went for an ascending sequence rather than an increasing one because we wished to allow $f.i = f.j$ for $i \ne j$. But equality is a very strong notion, closely associated with the rule of Leibniz, and we can get the same freedom by introducing among the f-values a partial order:

what used to be equality can then be weakened to "ordered in neither direction": for two values a and b, $\neg(a<b \vee b<a)$ then expresses that we don't care about their relative order in which they occur in a sorted sequence.

• We shall use —as we just did— $<$ to denote a nonreflexive partial order, i.e. we have the reflexivity

(0) $\qquad \neg\ x<x$

and the transitivity

(1) $\qquad x<y \wedge y<z \Rightarrow x<z$.

Because it is always a good idea to draw trivial conclusions immediately, we deduce

(2) $\qquad \neg(x<y \wedge y<x)$.

• The last remark is that the i,j need not be taken from the integers but may be taken from some partially ordered set. A sequence being sorted can then be generalized to a one-to-one correspondence being free of conflicts. For the definition of a conflict:

let i,j belong to the one set, and
let a,b belong to the other set, and
let (i,a) and (j,b) be a pair of elements of a one-to-one correspondence; then, this

pair presenting a conflict means

$$(i<j \wedge b<a) \vee (j<i \wedge a<b) \quad .$$

We may expect that it is possible to remove a conflict by "twisting", i.e by replacing a conflicting pair (i,a) and (j,b) by the pair (i,b) and (j,a). Furthermore — thanks to the transitivity of $<$ — we may expect that a local removal of a conflict by a twist reduces the total number of conflicts. The big question is: how do we prove all this without destroying the symmetry between the two sets, without case analyses, and without all sorts of duplication in the argument?

For the sake of brevity

- we denote sequences of elements from the original sets by juxtaposition
- $<$ between two sequences of equal length denotes the sequence of truth values obtained by element-wise application of $<$
- the conjunction of a sequence of truth-values is denoted by surrounding the sequence by a pair of square brackets.

e.g. $\quad [ib<ja] \equiv i<j \wedge b<a$

The important properties are illustrated by

$$[ib < ja] \equiv [bi < aj]$$

$$[x < y] \equiv x < y$$

$$[xx < yy] \equiv [x < y]$$

$$[ib < ja] \wedge [k < c] \equiv [ibk < jac] .$$

As a consequence of (0) we mention

(4) $\neg[xy < xz] \wedge \neg[yx < zx]$,

and as a consequence of (0) and (1)

(5) $\neg[abc < bca] \wedge \neg[abc < cab]$.

Denoting the presence of a conflict between the elements (i,a) and (j,b) of a one-to-one correspondence by $(i,a)\underline{\text{con}}(j,b)$, we can now define in our new notation $\underline{\text{con}}$ by

(6) $(i,a)\underline{\text{con}}(j,b) \equiv [ib < ja] \vee [ja < ib]$.

Its evident properties are

$$(i,a)\,\underline{\text{con}}\,(j,b) \equiv (j,b)\,\underline{\text{con}}\,(i,a)$$
$$(i,a)\,\underline{\text{con}}\,(j,b) \equiv (a,i)\,\underline{\text{con}}\,(b,j)$$
$$\neg(j,a)\,\underline{\text{con}}\,(i,a) \wedge \neg(j,a)\,\underline{\text{con}}\,(j,b) .$$

In order to show that removal of a conflict by "twisting" works and reduces the total number of conflicts, we prove two theorems.

Theorem 0

$$(i,a)\underline{con}(j,b) \wedge (k,c)\underline{con}(i,b) \wedge (k,c)\underline{con}(j,a)$$
$$\Rightarrow \quad (k,c)\underline{con}(i,a) \wedge (k,c)\underline{con}(j,b)$$

Proof We first establish (i)

$$\begin{array}{ll}
 & (k,c)\underline{con}(i,a) \wedge (k,c)\underline{con}(j,b) \\
= & \quad \{(6), \text{def. of } \underline{con}\} \\
 & ([ka<ic] \vee [ic<ka]) \wedge ([kb<jc] \vee [jc<kb]) \\
= & \quad \{\wedge \text{ distributes over } \vee\} \\
 & [kakb<icjc] \vee [kajc<ickb] \vee \\
 & [ickb<kajc] \vee [icjc<kakb]
\end{array}$$

Next we establish (ii)

$$\begin{array}{ll}
 & [ib<ja] \wedge (k,c)\underline{con}(i,b) \wedge (k,c)\underline{con}(j,a) \\
= & \quad \{(i) \text{ with } a,b := b,a\} \\
 & [ib<ja] \wedge \\
 & ([kbka<icjc] \vee [kbjc<icka] \vee \\
 & \;[icka<kbjc] \vee [icjc<kbka]) \\
\Rightarrow & \quad \{\wedge \text{ distributes over } \vee \text{ ; absorption}\} \\
 & [kbka<icjc] \vee [icjc<kbka] \vee \\
 & [ibkbjc<jaicka] \vee [ibicka<jakbjc] \\
= & \quad \{(5)\} \\
 & [kbka<icjc] \vee [icjc<kbka] \\
= & \quad \{\text{symmetry of } \wedge\} \\
 & [kakb<icjc] \vee [icjc<kakb] \\
\Rightarrow & \quad \{(i)\} \\
 & (k,c)\underline{con}(i,a) \wedge (k,c)\underline{con}(j,b)
\end{array}$$

From (ii) with $i,b,j,a := j,a,i,b$ and the symmetry of $\wedge$ we derive the same implication, except that the first conjunct of the antecedent is replaced by $[ja < ib]$. Disjunctive combination of (ii) and the latter implication yields Theorem (0).

(End of Proof.)

Theorem 1

$$(i,a)\,\underline{con}\,(j,b) \;\wedge\; (k,c)\,\underline{con}\,(i,b) \;\Rightarrow$$
$$(k,c)\,\underline{con}\,(i,a) \;\vee\; (k,c)\,\underline{con}\,(j,b)$$

Proof

$$\begin{array}{ll}
 & (i,a)\,\underline{con}\,(j,b) \;\wedge\; (k,c)\,\underline{con}\,(i,b) \\
= & \quad \{\text{def. of } \underline{con}\text{, i.e. (6)}\} \\
 & ([ib < ja] \vee [ja < ib]) \wedge ([kb < ic] \vee [ic < kb]) \\
= & \quad \{\text{distribution of } \wedge \text{ over } \vee\} \\
 & [ibkb < jaic] \;\vee\; [ibic < jakb] \;\vee \\
 & [jakb < ibic] \;\vee\; [jaic < ibkb] \\
\Rightarrow & \quad \{\text{transitivity and absorption}\} \\
 & [kb < jc] \;\vee\; [ic < ka] \;\vee \\
 & [ka < ic] \;\vee\; [jc < kb] \\
= & \quad \{\text{(6), i.e. def. of } \underline{con}\} \\
 & (k,c)\,\underline{con}\,(i,a) \;\vee\; (k,c)\,\underline{con}\,(i,b)
\end{array}$$

(End of Proof.)

<u>Corollary</u>

$$(i,a)\ \underline{\mathrm{con}}\ (j,b) \Rightarrow \neg\ (j,a)\ \underline{\mathrm{con}}\ (i,b)$$

<u>Proof</u> Theorem 1 with $k,c := j,a$. (End of Proof.)

* * *

Consider now a one-to-one correspondence containing the pair X given by $X = \{(i,a), (j,b)\}$ and let X contain a conflict, i.e let $(i,a)\ \underline{\mathrm{con}}\ (j,b)$. Let us derive a new one-to-one correspondence by twisting the elements of X, i.e. by replacing X by Y, given by $Y = \{(i,b), (j,a)\}$.

From the Corollary we conclude that Y contains no conflict, i.e. that the twist successfully removed the conflict in X.

The other elements, which occur in both one-to-one correspondences, have at most two conflicts with X and at most two with Y. According to Theorem 0, an element having both conflicts with Y has both conflicts with X. According to Theorem 1, an element having a conflict with Y has a conflict with X.

We conclude that replacing X by Y has reduced the total number by at least one.

* * *

To claim that I have been successful at removing all duplication, would be an exaggeration. In retrospect I could have tried the introduction of a symmetric operator, § say, between sequences of equal length. Giving § a higher binding power than the logical operators we could have written

(7) $(i,a)\ \underline{\text{con}}\ (j,b) \equiv ib \S ja$.

We would then have developed a §-calculus. The symmetry and associativity of $\wedge$ would be expressed by

(8) $X \S Y \equiv p.X \S p.Y$ for any permutation p.

The most complicated rule would be

(9) $X \S Y \wedge R \S S \equiv XR \S YS \vee XS \S YR$

The analogue of (5) would have been – with lower case letters for single elements –

(10) $\neg (xYZ \S YxZ)$.

I might have saved a factor of 2 , and I think it would have been an improvement. Let us investigate how we would then have proved Theorem 0 at a single go.

$\quad$ antecedent of Theorem 0

$= \quad \{(7)\}$

$\quad ib \S ja \wedge kb \S ic \wedge ka \S jc$

$= \quad \{(9)\}$

$\quad ib \S ja \;\wedge\; (kbka \S icjc \;\vee\; kbjc \S icka)$

$\Rightarrow \quad \{\wedge \text{ distributes over } \vee \text{ ; absorption}\}$

$\quad kbka \S icjc \;\vee\; (ib \S ja \;\wedge\; kbjc \S icka)$

$= \quad \{(9)\}$

$\quad kbka \S icjc \;\vee\; (ibkbjc \S jaicka \;\vee\; ibicka \S jakbjc)$

$= \quad \{(10)\}$

$\quad kbka \S icjc$

$= \quad \{(8)\}$

$\quad kakb \S icjc$

$\Rightarrow \quad \{\text{absorption}\}$

$\quad kakb \S icjc \;\vee\; kajc \S ickb$

$= \quad \{(9)\}$

$\quad ka \S ic \;\wedge\; kb \S jc$

$= \quad \{(7)\}$

$\quad$ consequent of Theorem 0 .

The experiment shows a considerable improvement; I still hesitate about its moral. Note that the first two "evident properties" of <u>con</u> are shared by § ! In view of the above proof, the obvious advantage of § is that it can be applied to sequences of any length.

Austin, 12 Feb. 1988

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