On the Iterative Solution of the Towers of Hanoi Problem

Maarten M. Fokkinga
Dept. INF, University of Twente,
fokkinga@cs.utwente.nl

Version of 25th May 2000

Abstract. We give an elegant proof (formal, calculational, readable, machine verifiable, without case distinctions) of the property that in the Towers of Hanoi solution all disks cycle.

Introduction. Recall the well-known problem of the Towers of Hanoi and its recursive solution. Given are three places A, B, C and N disks, numbered $0..N-1$ in order of increasing size, and initially they are stacked in decreasing size (smallest disk at the top) at place A. The goal is to move the tower to place B, according to the following rules:

only one disk is moved at a time,
a larger disk is never placed on a smaller one, and
the disks are stacked only at the three given places.

Denoting a direct move of disk n from place a to b not touching c by $(n, (a, b, c))$, the shortest sequence of moves to bring a tower of N disks from place A to B using place C along the way, is expressed by $Hanoi_N(A, B, C)$, where:

$$\begin{aligned} Hanoi_0(a, b, c) &= \langle\ \rangle \\ Hanoi_{n+1}(a, b, c) &= Hanoi_n(a, c, b) \frown \langle(n, (a, b, c))\rangle \frown Hanoi_n(c, b, a) \end{aligned}$$

The definition is non-trivially *recursive*, and is therefore not suitable as a recipe for human beings. There exists an iterative solution, which is easy to execute by humans:

If N is odd, *cycle* the smallest disk through A, B, C ("clockwise"), and after every move of the smallest disk, do *one* other move (for which there is only one possibility according to the rule that a larger disk may not be placed on a smaller one).

If N is even, the smallest disk should be cycled through A, C, B ("anti-clockwise").

We set out to formally prove that a crucial property of this iterative solution follows from *Hanoi*'s definition: the smallest disk cycles in the way indicated. Actually, we prove in one go, without extra effort, that all disks cycle. We do not prove that the smallest disk is moved alternately with the others.

Anybody familiar with the principle of induction is able to prove the property themself (try it!); so the existence of a proof is not at all surprising. What is surprising, however, is the *elegance* of our proof: we use only elementary, everyday set calculus (expressed in the Z notation [2]) and no case distinctions, and our proof is a calculation of a few lines long, and consists of steps that are both "human readable" and "machine verifiable". (Also, neither powers-of-two nor lengths-of-sequences occur anywhere in our proof.) It was Backhouse's paper [1] (with a quite different formulation and proof of the same property) that prompted us to do so.

Abstraction. As a preparation, we abstract from the places and keep only the *direction* of the moves. To this end, we first write the definition of *Hanoi* as follows:

$$\begin{aligned} Hanoi_0(d) &= \langle\, \rangle \\ Hanoi_{n+1}(d) &= Hanoi_n(f\ d) \frown \langle (n, d) \rangle \frown Hanoi_n(g\ d) \end{aligned}$$

Here, d ranges over triples (a, b, c), and $f = \lambda\, a, b, c \bullet (a, c, b)$ and $g = \lambda\, a, b, c \bullet (c, b, a)$.

Now, identify (A, B, C), (B, C, A), (C, A, B) with each other and denote these with "direction 0", and similarly for (A, C, B), (C, B, A), (B, A, C) and "direction 1". Observe that, under this identification, f and g become 'the successor/predecessor modulo 2', denoted with a postfixed $\oplus 1$. Thus, we find the definition:

$$\begin{aligned} H_0(d) &= \langle\, \rangle \\ H_{n+1}(d) &= H_n(d \oplus 1) \frown \langle (n, d) \rangle \frown H_n(d \oplus 1) \end{aligned}$$

Here, and in the sequel, d varies over directions $0\,..\,1$, and we let $\oplus$ denote addition modulo 2. We shall use the following laws from 'mod 2' calculus:

$\oplus$ is associative and commutative

$$\begin{aligned} i \oplus n \oplus d &= d \qquad \text{if } i = n \\ (n-1) \oplus d \oplus 1 &= n \oplus d \end{aligned}$$

Some set notation. (Here and in the remainder of the paper one may specialise "arbitrary i" to just $i = 0$; our more general treatment is for free.)

We wish to talk about the set of directions of disk i in the sequence $H_n\, d$ of moves: our claim will be that this set is a *singleton* set if $i < n$ (that is, "disk i *cycles* if it is one of the disks $0\,..\,n-1$") and empty otherwise. More specifically, the singleton set contains direction $i \oplus (n-1) \oplus d$, that is, "it contains direction d iff i and $n-1$ have the same parity". In order to express that singleton or empty set, recall Z's set notation $\{D \mid P \bullet E\}$, where D is a collection of variable declarations, P is a predicate, and E is an expression; the set consists of all values E where the variables range over sets as declared by D but only as far as they satisfy predicate P. The set notation $\{\mid P \bullet E\}$ is a special case (neither tricky nor abuse of notation, just a special case): it is equal to $\{E\}$ if P holds and $\varnothing$ otherwise, and saves us from explicitly making case distinctions. In particular, the set of interest is expressed by:

$$\{\mid i<n \bullet\ i \oplus (n-1) \oplus d\}$$

It remains to define the set $D_i(S)$ of directions d that appear in an item (i, d) in sequence S. Using the Z convention that a sequence is a function mapping indices to items, and that a function is a set of (argument, result)-pairs, the definition reads:

$$D_i(S) = \{d \mid (i, d) \in \operatorname{ran} S\}$$

$$= \text{ran} \ (\{i\} \lhd \ \text{ran}\, S)$$

Here $X \lhd Y$ is the Z notation for 'domain restriction': $X \lhd Y = \{u, v \mid u \in X \wedge (u, v) \in Y\}$.

We shall use the following elementary properties about these notations:

$$\begin{array}{lcl}
\text{ran}\ \langle x \rangle & = & \{x\} \\
\text{ran}\ (S \frown T) & = & \text{ran}\, S \cup \text{ran}\, T \\
X \lhd\ (Y \cup Z) & = & X \lhd Y \ \cup\ X \lhd Z \\
\{i\} \lhd \{(n, d)\} & = & \{\mid i = n \bullet (n, d)\} \\
\text{ran}\{\mid P \bullet (E, E')\} & = & \{\mid P \bullet E'\} \\
\{\mid P \bullet E\} \cup \{\mid Q \bullet E\} & = & \{\mid P \vee Q \bullet E\}
\end{array}$$

Theorem. For all naturals i and n:

$$D_i(H_n d) \quad = \quad \{\mid i < n \bullet i \oplus (n-1) \oplus d\}$$

Proof.
We use induction on n. For $n = 0$ we calculate:

$$\begin{array}{lr}
D_i(H_0 d) & \\
= \text{ran}\ (\{i\} \lhd \ \text{ran}\langle\ \rangle) & \text{defs } D_i,\ H_0 \\
= \varnothing & \text{ran on } \langle\ \rangle,\ X \lhd \text{ on } \varnothing,\ \text{ran on } \varnothing \\
= \{\mid i < 0 \bullet i \oplus (0-1) \oplus d\} & i \geq 0
\end{array}$$

For $n+1 > 0$, not applying any other set theoretic law than mentioned above:

$$\begin{array}{lr}
D_i(H_{n+1} d) & \\
= \text{ran}\ (\{i\} \lhd\ (\text{ran}\ \ (H_n(d \oplus 1) \frown \langle (n, d) \rangle \frown H_n(d \oplus 1)))) & \text{defs } D_i,\ H_{n+1} \\
= \text{ran}\ (\{i\} \lhd\ (\text{ran}\, H_n(d \oplus 1)\ \cup\ \text{ran}\langle (n, d) \rangle\ \cup\ \text{ran}\, H_n(d \oplus 1))) & \text{ran on } \frown \\
= \text{ran}\ (\{i\} \lhd\ (\text{ran}\, H_n(d \oplus 1)\ \cup\ \text{ran}\langle (n, d) \rangle)) & \text{idempotency } \cup \\
= \text{ran}\ (\{i\} \lhd\ (\text{ran}\, H_n(d \oplus 1)\ \cup\ \{(n, d)\})) & \text{ran on } \langle\ \rangle \\
= \text{ran}\ (\{i\} \lhd (\text{ran}\, H_n(d \oplus 1))\ \cup\ \{i\} \lhd \{(n, d)\}) & \{i\} \lhd \text{ on } \cup \\
= \text{ran}\ (\{i\} \lhd (\text{ran}\, H_n(d \oplus 1))\ \cup\ \{\mid i = n \bullet (n, d)\}) & \{i\} \lhd \text{ on } \{_\} \\
= \text{ran}(\{i\} \lhd (\text{ran}\, H_n(d \oplus 1)))\ \cup\ \text{ran}\{\mid i = n \bullet (n, d)\} & \text{ran on } \cup \\
= D_i(H_n(d \oplus 1))\ \cup\ \{\mid i = n \bullet d\} & \text{def } D_i,\ \text{ran on } \{\mid\ \} \\
= \{\mid i < n \bullet i \oplus (n-1) \oplus d \oplus 1\}\ \cup\ \{\mid i = n \bullet d\} & \text{induction hypothesis} \\
= \{\mid i < n \bullet i \oplus n \oplus d\}\ \cup\ \{\mid i = n \bullet i \oplus n \oplus d\} & \text{`mod 2' calculus} \\
= \{\mid i < n \vee i = n \bullet i \oplus n \oplus d\} & \cup \text{ of similar sets} \\
= \{\mid i < n+1 \bullet i \oplus ((n+1)-1) \oplus d\} & \text{arithmetic}
\end{array}$$

This completes the proof.

As corollary we have, for odd N, that $D_0(H_N 0) = \{0\}$: the smallest disk cycles in direction 0 ("clockwise" through A, B, C), and similarly for even $N > 0$ that $D_0(H_N(0)) = \{1\}$.

References

[1] R.C. Backhouse. The associativity of equivalence and the Tower of Hanoi problem. `http://www.cs.nott.ac.uk/~rcb/papers`, version of Febr 2000.

[2] J.M. Spivey. *The Z notation: a reference manual (2nd edition)*. Prentice Hall International, UK, 1992.