Introduction to the theory of computation
week 3

20th July 2005

Tree example re-visited

Last week we began an example about full binary trees. We need some definitions.

A TREE is a directed graph with the property that one of its nodes (if it has ≥ 1 nodes) is distinguished as the ROOT, and for every node there is a unique (one and only one) path from the root to that node.

THE PARENT of node v is the node u if (u, v) is an edge of the tree. In this case v is u's CHILD. A childless node is called a LEAF, and a node with at least one child is called an INTERNAL node.

A BINARY TREE is one where each node has, at most, two children, and (if they exist) the children are labelled "left" or "right" (but not both). Each parent has no more than one left or right child. In practice we draw the children to the left or right of their parent, rather than labelling them.

A FULL BINARY TREE is a binary tree where every node has either 0 or two children (equivalently, every internal node has two children).

CLAIM: Let $P(n)$ be "If a full binary tree has n nodes, then n is odd." Then for all $n \in \mathbf{N} - \{0\}, P(n)$.

PROOF (COMPLETE INDUCTION ON n): [1]

Base cases other than 0

In the last example our claim was not true for $n = 0$, but true for every other natural number. A small modification of induction allows you to use induction when you start from some natural number $n' > 0$. Here's a recipe for showing that induction beginning from $n' > 0$ is equivalent to starting from 0. Create a new predicate, $S(n) = P(n + n')$. With this notation you must show

1. $P(n')$ is true (equivalent to showing that $S(0)$ is true).

2. For an arbitrary $n \geq n'$, $P(n)$ implies $P(n + 1)$ (equivalent to showing that for an arbitrary $m \in \mathbf{N}$, $S(m)$ implies $S(m + 1)$).

Now you can conclude that $P(n)$ is true for all $n \geq n', n \in \mathbf{N}$. You may also use more than one base case, since there may be several initial values of n where $P(n)$ is true, but they can't be derived from smaller values. Here's an example that is analogous to Example 1.12 in the Course Notes.

Another postage stamp problem

Claim: There exists a natural number n_0 such that for all natural numbers $n \geq n_0$, postage of n cents can be formed using 5- and 3-cent stamps.

Proof: There's two parts to this claim. The first part involves finding a suitable n_0, which you can probably achieve through trial-and-error.[2] Now we proceed by complete induction on n, starting from $n_0 = ?$

Claim: Let $P(n)$ be "Postage of n cents can be formed using 5- and 3-cent stamps." Then for all natural numbers $n \geq ?, P(n)$.

Proof (complete induction on n): [3]

A set of n elements has 2^n subsets

Claim: Let $P(n)$ be "A set of n elements has exactly 2^n subsets." Then $\forall n \in \mathbf{N}, P(n)$.

Proof (complete induction on n): [4]

All hexagons have the same number of sides

This result may seem so obvious that it doesn't require a proof, but humour me. We'll turn induction loose on it, and see where we get.

Claim: Let $P(n)$ be "No set of n hexagons contains a pair of hexagons with different numbers of sides." Then $\forall n \in \mathbf{N}, P(n)$.

Proof (complete induction on n): [5]

Be very suspicious of this "proof." Although the conclusion is certainly true, the way we got there is not valid. Take a pencil and go back over the proof, replacing each occurrence of the string "hexagon" with "polygon." Good — now you've proved that a triangle has the same number of sides as a hexagon. What went wrong?

Check the base cases.[6] What is the smallest counter-example you can construct, that is the smallest set of polygons that contradict $P(n)$?[7] Go back through the induction step with this set to see what goes wrong.

The sum $\sum_{t=0}^{n} 2^t = 2^{n+1}$, doesn't it?

In most proofs by induction, the real work is in the induction step, proving that $P(n) \Rightarrow P(n+1)$, and verifying the base cases is trivial. So in the next proof we'll jump right into the induction step.

Claim: Let $P(n)$ be "The sum $\sum_{t=0}^{n} 2^t = 2^{n+1}$." Then $\forall n \in \mathbf{N}, P(n)$.

Proof (induction on n): [8]

Oh-oh, we've just "proved" something that that isn't true (find a counter-example). The induction step is certainly valid, but we never bothered with the base case. A bit of checking will show that $P(n)$ is false for any $n \in \mathbf{N}$ (if it were true for any $n \in \mathbf{N}$, it would we true for all n's successors). So we really need to check the base case.

The body of this proof will work for a modified predicate. Try $P(n)$: "The sum $\sum_{t=0}^{n} 2^t = 2^n - 1$." Now $\forall n \in \mathbf{N}, P(n)$. Just plug this predicate into the "proof" above, and you will prove something that's actually true!

Notes

[1]Suppose a complete binary tree has $n \geq 1$ nodes, and assume that $P(\{1,\ldots,n-1\})$ is true. I must show that this assumption implies $P(n)$. If $n = 1$, then $P(n)$ holds, since 1 is well-known to be odd. Otherwise, if $n > 1$, there are nodes other than the root, hence the root has children and (by the definition of a full binary tree) exactly two children. It is easy to verify that these children are themselves the roots of full binary subtrees, with n_L and n_R nodes, respectively. Since the subtrees have fewer than n, but at least 1 node, I can use the induction hypothesis and n_L is odd, as is n_R. Counting the root plus the nodes in its two subtrees yields $n_R + n_L + 1$ — an odd number. Thus $P(\{1,\ldots,n-1\})$ implies $P(n)$.
Conclude that $P(n)$ is true for every positive natural number n. QED.

[2]$n_0 = 8$ should work. Is it the smallest?

[3]For an arbitary natural number ≥ 8, assume that $P(\{8,\ldots,n-1\})$ is true. I must show that this implies $P(n)$. If $n \in \{8, 9, 10\}$, it is easy to verify that we can form postage of n cents using 5- and 3-cent stamps. If $n > 10$, then the IH says we can form postage of $n - 3$ cents using 5- and 3-cent stamps. Take that postage and add one 3-cent stamps, and this provides postage of n cents using 5- and 3-cent stamps. So $P(\{8,\ldots,n-1\}) \Rightarrow P(n)$. Conclude that P(n) is true for all natural numbers $n \geq 8$. QED.

[4]$P(0)$ states that a set with zero elements, the empty set, has $2^0 = 1$ subset. This is true since $\emptyset \subseteq \emptyset$, and that is its only subset.

Induction step: I want to show that $P(n) \Rightarrow P(n+1)$, so I assume that $P(n)$ is true for some arbitrary natural number n. Let $X = \{x_1,\ldots,x_{n+1}\}$ be an arbitrary set of $n+1$ elements. In order to count the subsets of X more easily, I partition them into mutually exclusive collections X_L contains all the subsets of X that do not contain the element x_{n+1}, and X_R contains all the subsets of X that do contain the element x_{n+1}. By the inductive hypothesis I already know that X_L contains 2^n subsets, since it contains all the subsets of $X - \{x_{n+1}\}$, a set of size n. I dream up the bijection:

$$f : X_L \to X_R \qquad f(s) = s \cup \{x_{n+1}\}$$

Function f takes subsets in collection X_L to subsets in collection X_R by adding the element x_{n+1}, and it is easy to verify that f is a bijection. This means that X_L and X_R have the same number of elements, so in total X has $2^n + 2^n = 2^{n+1}$ subsets, and $P(n) \Rightarrow P(n+1)$.
We conclude that $P(n)$ is true for all $n \in \mathbf{N}$. QED.

[5]$P(0)$ states that the empty set doesn't contain a pair of hexagons with a different number of sides, which is clearly true. $P(1)$ states that a set of 1 hexagon doesn't contain a pair of hexagons with different numbers of sides, which is also clearly true, since the hexagon must have the same number of sides as itself. So the base cases $P(0)$ and $P(1)$ hold.

Induction step: I want to show that $P(n) \Rightarrow P(n+1)$, so I assume that $P(n)$ is true for some arbitrary natural number $n \geq 1$. Now I consider a set $H = \{h_1,\ldots,h_n,h_{n+1}\}$, of $n+1$ hexagons. I can break H up into two subsets of n hexagons each, $H_L = \{h_1,\ldots,h_n\}$ and $H_R = \{h_2,\ldots,h_n,h_{n+1}\}$. Since H_L and H_R have n hexagons each, the inductive hypothesis applies, so by $P(n)$ these sets have no pairs of hexagons with different numbers of sides. Let (h_i, h_j) be a pair of hexagons from our original set H. If h_i and h_j are either both in H_L or H_R, by the inductive hypothesis they have the same number of sides. Otherwise, WLOG $h_i \in H_L$ and $h_j \in H_R$, and by the inductive hypothesis h_i has the same number of sides as h_n (which is in H_L), which in turn has the same number of sides as h_j (because h_n is also in H_R). Thus h_i has the same number of sides as h_j, and $P(n) \Rightarrow P(n+1)$.

We conclude that $P(n) \Rightarrow P(n+1)$, so any finite set of hexagons never contains a pair of hexagons with different numbers of sides. QED.

[6]They're okay.

[7]How about $\{triangle, hexagon\}$, a set of two polygons.

[8]I want to show that $P(n) \Rightarrow P(n+1)$, so I assume $P(n)$ for some arbitrary $n \in \mathbf{N}$. Now I consider the sum $\sum_{t=0}^{n+1} 2^t$, and I break it into two parts:

$$\sum_{t=0}^{n+1} 2^t = \left(\sum_{t=0}^{n} 2^t\right) + 2^{n+1}$$

By the inductive hypothesis I know that $\sum_{t=0}^{n} 2^t = 2^{n+1}$, so altogether I've got $2^{n+1} + 2^{n+1} = 2^{n+2}$, and $P(n+1)$ holds. Thus $P(n) \Rightarrow P(n+1)$.
I conclude that $P(n)$ is true for all $n \in \mathbf{N}$.