CS 624 - HW 4

Max Weiss

24 February 2020

1. Toward a contradiction, suppose that some edge occurs more than once in P. Then P must contain a subpath of the form $u \to v \to \cdots \to u \to v$. We must have $u \neq v$, since otherwise, u will occur at least four times in P, contradicting the definition of loop-simplicity. But then, more than one vertex occurs more than once in P, and this too contradicts the definition.

2. Suppose that G is a tree in which x, y are distinct vertices. Since G is connected, G must contain at least one path from x to y. It therefore suffices to show that G contains at most one simple path from x to y.

Toward a contradiction, suppose that P and Q are distinct simple paths from x to y in G. Let $\bar{Q}$ be the reversal of Q, and let $P\bar{Q}$ be the concatenation of P with $\bar{Q}$. I will argue, by induction on the length of $P\bar{Q}$, that the graph G must contain a simple loop.

Since P and Q are distinct, at least three vertices of G must have occurrences in $P\bar{Q}$, so that this must have length ≥ 4. Assume, for a basis step, that the length is just 4. Then $P\bar{Q}$ must have one of the forms $x \to z \to y \to x$ or $x \to y \to z \to x$, where x, y, z are distinct. Each of these is a simple loop itself, so this completes the basis step.

Now suppose the P, Q are distinct simple paths from x to y, such that the length of $P\bar{Q}$ is some $k > 4$. Further suppose this claim to hold for all distinct simple paths P', Q' with same source and target, such that $P'\bar{Q}'$ has length less than k. If $P\bar{Q}$ is itself a simple loop then we are done, so suppose it is not. Since $k > 4$, it follows from the definition of simple loop that in $P\bar{Q}$, either (i) x occurs not exactly twice, or (ii) some vertex other than x occurs more than once.

However, case (i) is impossible: x cannot occur less than twice since $P\bar{Q}$ is a loop; but x also cannot occur more than twice, since paths P, Q are simple.

So we must be in case (ii), and some vertex $z \neq x$ must occur more than once in $P\bar{Q}$. By symmetry, we may also assume that $z \neq y$. Since paths P, Q are simple, z cannot occur more than once in either. Therefore, P and Q must have forms P_1P_2 and Q_1Q_2, where P_1, Q_1 each are simple paths from x to z and P_2, Q_2 are simple paths from z to y. Moreover, since $P \neq Q$, it follows that either $P_1 \neq P_2$

or $Q_1 \neq Q_2$. In either case, $P_i\bar{Q}_i$ has length $< k$, and the result now follows from the induction hypothesis.

Lineages. From exercise (2), it follows that for any node x of a rooted tree, there is a unique simple path from the root to x. I will refer to this path the *lineage* of x.

By definition, y is an ancestor of x iff there is a simple path from the root through y to x. It follows that y is an ancestor of x iff y occurs in the lineage of x.

3a. Toward a contradiction, suppose that node x has parents y and z. Then x has lineages P and Q whose penultimate nodes are y and z respectively. Since x has at most one lineage, we have $P = Q$ and therefore $y = z$.

3b. By part (a), it suffices to show that every node other than the root has at least one parent. If x is not the root, then its lineage has the form $\cdots y \to x$ for some y, and this y must be a parent of x.

4. Suppose that x is an ancestor of y, and vice versa. Then x has a lineage P through y, and y has a lineage Q through x. By definition, P must contain a simple path from y to x, and Q must contain a simple path from x to y. By exercise (2) it follows that $x = y$.

5. Suppose that x has a lineage P through a, and a lineage Q through b. Then P must have the form P_1P_2, where P_1 is a lineage of a, and P_2 is a path from a to x. Now if b occurred neither in P_1 nor in P_2, then b would not occur in P. Since b does occur in Q we would then have $P \neq Q$, and x would have two lineages wihch is impossible.

Therefore, either b occurs in P_1 or b occurs in P_2. If b occurs in P_1, then P_1 is a lineage of a through b, so that b is an ancestor of a. So suppose that b occurs in P_2. Then P_2 has an initial segment Q which is a simple path from a to x. Consequently, the concatenation of P_1 with Q is a lineage of b through a so that a is an ancestor of b.

6. Let's say that z is a *common ancestor* of x, y if it is an ancestor of each of x and y. Also, let's say that a node x is *minimal* in a set X of nodes provided that x is the only descendant of x in X. Finally, a *least common ancestor*[1] of x, y is a node which is minimal in the set of common ancestors of x, y.

To show that any nodes x, y have a unique least common ancestor, it suffices to show three things: (i) x, y have at least one common ancestor; (ii) any nonempty

[1] An alternative formulation would be this: r is a least common ancestor of x, y iff r is a least element in the intersection of the lineages of each of x and y; this is equivalent because the common ancestry of x, y is just the intersection of their lineages.

set of nodes has a minimal element; and (iii) the set of common ancestors of x, y has most one minimal element.

Part (i) is clear, since the root is an ancestor of all nodes in the tree.

As for part (ii), recall that a partial order is a transitive antisymmetric relation. I will argue that if X is a nonempty but finite, partially ordered set, then X has a minimal element. Suppose, to the contrary, that X does not have a minimal element. I will argue that X must be infinite, because for all n, it contains a descending chain of n distinct elements. This is clear for $n = 1$. So suppose that $x_1 \geq \cdots \geq x_n$ is a chain of n distinct elements. By hypothesis, x_n is not minimal in X, so there must be a $y \in X$ with $y \leq x_n$ and $y \neq x_n$. If $y = x_i$ for some $i < n$, then by transitivity of $\leq$ we would have $x_n \leq y$ while $y \leq x_n$, so that $y = x_n$ by antisymmetry; and this is a contradiction.

Let's now work to apply this lemma in the case at hand. Where x, y are nodes of a tree, I'll here write $x \leq y$ to mean that x is a descendant of y. I will argue that $\leq$ is indeed a partial order. Antisymmetry follows from exercise (4). As for transitivity, suppose that $x \leq y$ and that $y \leq z$. Then y has a lineage P through x, and z has a lineage Q through y. In particular, Q has the form Q_1Q_2, where Q_1 is a simple path from the root to y, and Q_2 is a simple path from y to z. By exercise (2), we must have that $Q_1 = P$. Hence $Q_1 = P_1P_2$, where P_1 is a lineage for x, and P_2 is a simple path from x to y. It follows that $Q = P_1(P_2Q_2)$ is a lineage for z through x, so that x is indeed an ancestor of z.

Finally as for (iii), let u and v each be a common ancestor of x, y. Then each is an ancestor of x. So by exercise (5), it follows that either u is an ancestor of v, or v is an ancestor of u. Without loss of generality, assume that u is an ancestor of v. Then there is a simple path from the root through u to v; hence either $v = u$, or $v \neq u$ and v is not minimal. In any case, no distinct common ancestors are minimal, so at most one minimal common ancestor exists.

7. First some remarks about "order determined by the tree". In the notes, the successor of a node in a binary search tree is said to be the node which is "the successor in the order determined by the tree." Intuitively, this order is the smallest reflexive relation R such that for every node x, all descendants of the left child of x bear R to x, while x bears R to all descendants of its right child.

It will be useful to have a somewhat closer analysis. Let $\leq_1$ and $\leq_2$ be relations[2] on the sets X_1 and X_2. I'll write $\leq_1 \circ \leq_2$ for the relation $\leq_1 \cup \leq_2 \cup (X_1 \times X_2)$ on $X_1 \cup X_2$. Clearly $\circ$ is associative. It is also clear that if each of $\leq_1, \leq_2$ is a total order and if X_1, X_2 are disjoint, then $\leq_1 \circ \leq_2$ is a total order on $X_1 \cup X_2$. We can then define, by induction on nodes x, the ordering determined by the (sub-) tree rooted at x:

$$\leq_x \;=\; \leq_x^L \;\circ\; \{\langle x, x\rangle\} \;\circ\; \leq_x^R \tag{1}$$

[2] In the set-of-ordered-pairs construal of "relation".

where if x has a left child y, then $\leq_x^L$ is $\leq_y$, and otherwise is the empty set; similarly for $\leq_x^R$ and the right child of x. Since subtrees rooted at distinct children must be disjoint, $\circ$ preserves the property of being a total order, and so it follows by induction that $\leq_x$ is a total order on the tree rooted at x.

The order $\leq_r$ determined by a tree with root r is in turn supposed to generate a notion of "successor" of a given node x. Suppose that x is not the largest node in the tree. Now consider the set of all nodes which are greater than x. By the lemma to exercise 6 this must contain at least one minimal element. And since indeed $\leq_r$ is a total order, there can be at most one. I will refer to this least node greater than x as the *successor* of x.

The foregoing stuff about order is pure theory of binary trees. Now a *labeled* binary tree is a binary tree plus a function which maps nodes to keys. I will write k_x for the key of node x. Finally a *binary search tree* is a labeled search tree such that the key of every node of the left subtree of each node x is no greater than x, and the key of every node of the right subtree of x is no smaller than x.

In this question we are asked to assume that the tree contains no duplicate keys. We then get the following useful fact:

(*) if y is a descendant of x, then y is in the left subtree of x iff $k_y < k_x$, and y is in the right subtree of x iff $k_y > k_x$.

Another useful fact is this.

(**) Suppose the key function maps each node to its ordinal position in order determined by the tree itself, as this is defined above. The result is a binary search tree.

In other words, any descendant of the left child of some node x is less than x in the order determined by the tree, and similarly for descendants of the right child.

7a. I will first argue that if a is an ancestor of x and $k_a > k_x$, then x is in the left subtree of a. Since $x \neq a$, therefore x must be in either the left or the right subtree of x. But x cannot be in the right subtree of a since $k_a > k_x$. Therefore, x is in the left subtree of a.

It is clear that if d is a descendant of x and if x is in the left subtree of a, then d is in the left subtree of a. So by fact (*), $k_d < k_a$.

7b. Suppose that a is the successor of x. Consider the set of all nodes y such that $x \leq_y a$. By the lemma to exercise (6), this must have a minimal element. We can now complete the proof by enumerating cases of the definition for $x \leq_y a$.

- $x \leq_y^L a$, or $x \leq_y^R a$. This contradicts the minimality of y.

- x descends from the left child of y, and a descends from the right. Then $x < y < a$ and so a is not the successor; hence this is impossible.
- x descends from the left child of y and $y = a$. Then a is an ancestor of x.
- a descends from the right child of y and $y = x$. Then x is an ancestor of a.

7c. Suppose that the right subtree of x is nonempty, so that x has a right child b. Invoking the analysis from (7b), we find that if x descends from the left child of a, then $x < b < a$, contradicting the hypothesis that a is the successor of x. Hence, a must descend from the right child of x. Since all descendants of a's right child are greater than x, therefore x must be the least of them.

It now suffices to prove the following lemma: that the least node b in a tree with root r is not the descendant of any right child. If to the contrary it were the descendant of some right child c, then where d is the parent of c we would have $d <_d z$ for all descendants z of c and in particular $d <_d b$; hence $d <_r b$ which contradicts minimality of b.

7d. Assume x has an empty right subtree, and that a is the successor of x. Then, the last case of (7b) cannot arise. So we must be in the penultimate case, where x descends from the left child of a. It remains to show that a is the least node whose left child is an ancestor of x. Suppose to the contrary that x descended from the left child of some b, which in turn descends from the left child of x. Well, any descendant of the left child of something is less than it, and so we would have $x < b < a$, contradicting the hypothesis that a is the successor of x.

7e. Suppose that x has a right child c. Then the right subtree of x is nonempty. So by (7c), the successor a of x is the leftmost descendant of c. Since a is leftmost, it (I suppose by definition) does not have a left child.

7f. If x is not the least node in the tree, then a *predecessor* of x is the greatest node less than x. Existence and uniqueness of predecessors is similar to that for successors (using a symmetric form of the lemma for question (6)). We can then develop a symmetric form of (7b), which in turn can be used to deduce the following counterpart of (7c): if the left subtree of x is nonempty, then the predecessor of x is just the rightmost node in the left subtree. Symmetrically with 7e, we conclude that if x has a left child c, then the predecessor of x is the rightmost descendant of c.