Notre Dame Journal of Formal Logic
Volume 28, Number 2, April 1987

On n-Equivalence of Binary Trees

KEES DOETS*

Summary and introduction This note presents a simple characterization of the class of all trees which are n-elementary equivalent with B_m: the binary tree with one root all of whose branches have length m (for each pair of positive integers n and m). Section 1 contains some preliminaries. Section 2 introduces the class $Q(n)$ of binary trees and proves that every tree in it is n-equivalent with B_m whenever $m \geq 2^n - 1$. Section 3 shows that, conversely, each n-equivalent of a B_m with $m \geq 2^n - 1$ belongs to $Q(n)$. Finally, all n-equivalents of B_m for $m < 2^n - 1$ are isomorphic to B_m.

1 Preliminaries Define the relation $\equiv^n$ between models of the same finite vocabulary (not containing function-symbols) using induction on n by

(1) $A \equiv^0 B$ iff A and B have the same true atomic sentences
(2) $A \equiv^{n+1} B$ iff both
  (i) $\forall a \in A \exists b \in B (A, a) \equiv^n (B, b)$
  (ii) $\forall b \in B \exists a \in A (A, a) \equiv^n (B, b)$.

Also, when $\underline{a} \in A^k$, define the first-order (!) formula $\sigma^n_{\underline{a}}(x_0, \ldots, x_{k-1})$ of quantifier rank n by

(1′) $\sigma^0_{\underline{a}}$ is the conjunction of all formulas with at most $x_0, \ldots, x_{k-1}$ free satisfied by $\underline{a}$ in A which are either atomic or negated atomic
(2′) $\sigma^{n+1}_{\underline{a}}$ is $\forall x_k \bigvee_{b \in A} \sigma^n_{\underline{a}^\frown\langle b\rangle} \wedge \bigwedge_{b \in A} \exists x_k \sigma^n_{\underline{a}^\frown\langle b\rangle}$.

For a definition of the Ehrenfeucht-game and a proof of the next lemma (be it in the context of linear orderings) I refer to [1], pp. 93–96, 247–252 and 359–361.

*I thank Piet Rodenburg for communicating his question (answered by 2.5 below) on which this note is a digression, and Prof. Specker for a lecture featuring Ehrenfeucht-games.

Received January 18, 1985

1.1 Lemma *The following are mutually equivalent:*
(1) $A \equiv^n B$
(2) *A and B have the same true sentences of quantifier rank* $\leq n$
(3) *Player II has a winning strategy in the Ehrenfeucht-n-game between A and B*
(4) $B \models \sigma^n_{\varnothing}$.

1.2 Lemma *Suppose that A and B are finite linear orderings. Then* $A \equiv^n B$ *iff* $|A| = |B|$ *or* $|A|, |B| \geq 2^n - 1$.

Proof: cf. [1], Corollary 6.9, p. 99 and exercise 6.10, p. 100.

2 Binary trees

2.1 Definitions, notations Assume that $\leq$ partially orders the nonempty set T.

(1) T is a *tree* if, for all $x \in T$, $x\downarrow = \{y \in T | y < x\}$ is a finite set linearly ordered by $\leq$. The *height* of x, $h(x)$, is the order type of $x\downarrow$.
(2) The tree T is *binary* if it has a least element (its *root*) and every non-maximal element has exactly two immediate successors.
(3) A *branch through* T is a maximal linearly ordered set. A *branch above* $x \in T$ is a branch in the subtree $x\uparrow = \{y \in T | x < y\}$. The order type of a branch is called its *length*.

The characterization promised is contained in the following

2.2 Definition Let $n \geq 1$. The binary tree T *satisfies* $Q(n)$ iff the following conditions are met:

$Q.1(n)$ If $n = 2$ then $\exists x \forall y \geq x (y = x)$; if $n \geq 3$ then $\forall x \exists y \geq x \forall z \geq y (z = y)$.
$Q.2(n)$ Every branch through T has length $\geq 2^n - 2$.
$Q.3(n)$ Some branch through T has length $\geq 2^n - 1$.
$Q.4(n)$ For all $x \in T$ and $m < 2^{n-1} - 1$: if some branch above x has length m then every branch above x has length m.

Notice that every binary tree satisfies $Q(1)$: $Q.1(n)$–$Q.3(n)$ only demand something if $n \geq 2$; and $Q.4(n)$ is nontrivial for $n \geq 3$ only.

2.3 Theorem *If the binary trees* T^1 *and* T^2 *satisfy* $Q(n)$ *then* $T^1 \equiv^n T^2$.

Proof: Induction on n. The case $n = 1$ is trivial. Assuming 2.3 to hold for n we check it for $n + 1$ using Ehrenfeucht-games. Thus, suppose $t \in T^1$ is the first move of player I in the $(n + 1)$-game between T^1 and T^2. There are three cases to consider:

(1) $h(t) < 2^n - 1$.
Decompose T^1 in:

  (i) two top-trees t^1 and t^2, final sections of T^1 the roots of which are the two minimal elements of $t\uparrow$

(ii) the linear ordering $t\!\downarrow$ of type $h(t)$
(iii) the trees t_a (where $a < t$; there are none if $h(t) = 0$): the root of t_a being the immediate successor of $a \in t\!\downarrow$ not below t.

Since T^1 satisfies $Q(n+1)$, it is clear that all trees in this decomposition satisfy $Q(n)$. For instance, $Q.1(n)$ is inherited from $Q.1(n+1)$ by final sections (this is true even if $n = 1$ or $n = 2$!). $Q.4(n+1)$ implies $Q.4(n)$; and if T^1 satisfies $Q.4(n+1)$ then so do its final sections.

By $Q.2(n+1)$, each branch in, say, t^1 has length $\geq 2^{n+1} - 2 - h(t) - 1 > 2^{n+1} - 2 - (2^n - 1) - 1 = 2^n - 2$, i.e., has length $\geq 2^n - 1$. Thus, t^1 has $Q.2(n)$ and $Q.3(n)$. The same goes for the other subtrees.

Now player II answers t with some $s \in T^2$ for which $h(s) = h(t)$. Let j be the isomorphism between $t\!\downarrow$ and $s\!\downarrow$. s induces a decomposition of T^2 similar to the one described for t in which all trees satisfy $Q(n)$. By induction-hypothesis, corresponding trees in the decompositions are n-equivalent. Therefore, II can win the remaining n-game using the following strategy: above t or s he uses winning strategies between t^i and s^i ($i = 1, 2$). Below t or s he answers using the isomorphism j. Finally, a move in some $t_a (a < t)$ by I is answered using a winning strategy between t_a and $s_{j(a)}$ and vice versa.

This strategy is clearly winning for II since the union of partial isomorphisms between corresponding substructures in the decompositions is a partial isomorphism between T^1 and T^2.

(2) There is no branch of length $\geq 2^n - 1$ above t.

By $Q.4(n+1)$ there exists $u \leq t$ such that all branches above u have length $2^n - 2$. Hence, u is the root of a final section B_u of T^1 in which all branches have length $2^n - 1$.

Since T^2 satisfies $Q.1(n+1)$ (for, $n + 1 \geq 2$) and $Q.4(n+1)$, there exists $v \in T^2$ which is the root of a final section B_v of T^2 isomorphic to B_u. (1)

By $Q.2(n+1)$, $u\!\downarrow$ and $v\!\downarrow$ have order types $\geq 2^{n+1} - 2 - (2^n - 1) = 2^n - 1$; hence $u\!\downarrow \equiv^n v\!\downarrow$ by Lemma 1.2. (2)

If a < u, branches above a through u have length $\geq 2^n - 1$; by $Q.4(n+1)$ therefore, all branches above a have length $\geq 2^n - 1$; in particular, all branches through u_a have length $\geq 2^n - 1$. Thus, u_a satisfies $Q(n)$. The same goes for the $v_b (b < v)$.

By induction-hypothesis, $u_a \equiv^n v_b$ whenever $a < u$ and $b < v$. (3) Now II uses the following strategy. First, he answers t using the isomorphism (1). The remaining n-game is dealt with as follows. Between B_u and B_v, II goes on using the isomorphism (1). Below u or v he uses the winning strategy (2). If I makes a move x in some $u_a (a < u)$ for the first time while a has not been played yet, I is granted the *extra move* a as well. Then II answers a by some $b < v$ using (2) and next answers x by some $y \in v_b$ using (3).

Of course, if a has been played before, b has been fixed already and no extra move is granted (this occurs in particular when x isn't the first move in u_a by either player).

(3) $h(t) \geq 2^n - 1$ and some branch above t has length $\geq 2^n - 1$.

By $Q.4(n+1)$ then, all branches above t have length $\geq 2^n - 1$. Hence, in the decomposition described under (1) above, t^1 and t^2 satisfy $Q(n)$. If $a <$

t, branches above a through t and, hence, all branches above a, have length $\geq 2^n - 1$; thus t_a satisfies $Q(n)$.

Since T^2 satisfies $Q.3(n+1)$ and $2^{n+1} - 1 = 2(2^n - 1) + 1$, II can find $s \in T^2$ such that $h(s) = 2^n - 1$ while some branch above s has length $\geq 2^n - 1$. It follows that s^1, s^2 and all $s_b (b < s)$ satisfy $Q(n)$. For the remaining n-game, II uses a strategy similar to the one used under (2) above; except that above s or t he uses that $s^i \equiv^n t^i$ $(i = 1, 2)$.

2.4 Examples The following trees satisfy $Q(n)$.

1. The binary tree B_m all of whose branches have length $m \geq 2^n - 1$.
2. Infinite binary trees provided that, along every infinite branch, all finite side-trees are of type 1. Moreover, such finite side-trees have to occur infinitely often on each infinite branch.

2.5 Corollary *Finiteness of trees is not a first-order property on the class of all binary trees.*

2.6 Corollary *"Every branch has length $\geq 2^n - 1$" and its negation "Some branch has length $\leq 2^n - 2$" $(n > 1)$ cannot be expressed by first-order sentences of quantifier rank n on the class of (finite) binary trees.*

3 $Q(n)$ in first-order terms By 2.4, B_m satisfies $Q(n)$ whenever $m \geq 2^n - 1$; hence 2.3 gives one half of the following

3.1 Theorem *Let $m \geq 2^n - 1$. A binary tree T satisfies $Q(n)$ iff $T \equiv^n B_m$.*

The other half is established by Propositions 3.2–3.4 below. These results (together with 1.1) show that $Q(n)$ can be expressed by a first-order sentence of quantifier rank n which therefore, together with some first-order quantifier rank-4 axiomatization of binary trees (instead of requiring that each $x\downarrow$ be finite in 2.1.1 we merely demand it to be a discrete linear ordering with first and last element), is logically equivalent to the sentence $\sigma^n_\varnothing$ described in Section 1 (where $A = B_m$) (when $n \geq 4$ and $m \geq 2^n - 1$).

Notice that, by definition, $Q.1(n)$ has been expressed by a first-order sentence of quantifier rank $\leq n$. $Q.2(n)$–$Q.4(n)$ are dealt with by 3.4, 3.2, and 3.3, respectively.

In the sequel, $\phi^{<x}$ and $\phi^{>x}$ denote the formulas obtained from ϕ by restricting quantifiers to the sets $\{y | y < x\}$ and $\{y | x < y\}$, respectively.

3.2 Proposition *Define the sentences ϕ_n by:*

ϕ_1 *is* $\exists x(x = x)$

ϕ_{n+1} *is* $\exists x(\phi_n^{<x} \wedge \phi_n^{>x})$.

Then ϕ_n has quantifier rank n and it holds in a tree iff there is a branch of length $\geq 2^n - 1$.

Proof: Obvious.

For the next propositions, t^1 and t^2 are defined as in case (1) of the proof of 2.3.

In view of 2.6, the next result is not entirely trivial.

3.3 Proposition *Let k be any integer ≥ 1 and T a binary tree such that $T \equiv^{n+1} B_k$. Then T satisfies $Q.4(n+1)$.*

Proof: We may assume $n > 1$ since otherwise $Q.4(n+1)$ is trivially satisfied. Suppose α is a branch of minimal length $m < 2^n - 1$ above some $t \in T$ such that some branch above t has length $>m$. Assume α is a branch through t^1. Let u be the root of t^1. Then $\alpha - \{u\}$ is a branch of length $m - 1$ above u and hence, by minimality of m, all branches through t^1 have length m; and all branches through t^2 have length $\geq m$. Choose $x \in B_k$ such that $(T, u) \equiv^n (B_k, x)$. (1) By (a variation on) Lemma 1.2 it follows that all branches through $B_x = \{y \in B_k | x \leq y\}$ have length m. Let β be a branch through t^2 of length $\ell > m$.

We may assume that β is finite since T satisfies $Q.1(3)$ (this is a quantifier rank-3 sentence true in B_k and $n + 1 \geq 3$). Furthermore, let $y \in B_k$ be the element $\neq x$ with the same predecessors as x. Notice that if $s \in \beta$ and $(T, u, s) \equiv^{n-1} (B_k, x, z)$ then $z \geq y$, since $n - 1 \geq 1$ and $s \not\leq u \wedge \forall w < u \ (w < s)$ is a quantifier rank-1 sentence true in (T, u, s).

The proof is finished by indicating how I can defeat II in the n-game between (T, u) and (B_k, x), contradicting (1). If $m < 2^{n-1}$ then I, by picking the largest element s of β, wins the n-game: II has to answer with a maximal element $z \geq y$, whence there remain $m - 1 < 2^{n-1} - 1$ elements in $\{w < z | y \leq w\} = \{w < z | w \not\leq x\}$ and I can defeat II in $n - 1$ more moves by playing on β below s (use 1.2). If $2^{n-1} \leq m$ then $2^{n-1} < \ell$ and I picks $s \in \beta$ such that $\{v \in \beta | s < v\}$ has $2^{n-1} - 1$ elements. On penalty of losing (cf. 3.2) II must answer with a $z \geq y$ above which there are branches of length $\geq 2^{n-1} - 1$. But then $\{w < z | y \leq w\}$ has $\leq m - 2^{n-1} < 2^n - 1 - 2^{n-1} = 2^{n-1} - 1$ elements left and I needs only $n - 1$ more moves on β below s to defeat II.

3.4 Proposition *Suppose $k \geq 2^{n+1} - 2$. Let T be a binary tree such that $T \equiv^{n+1} B_k$. Then T satisfies $Q.2(n+1)$.*

Proof: Suppose that some branch α through T has length $\ell < 2^{n+1} - 2$. Since the quantifier rank-$(n+1)$-sentence

$$\forall x(\neg \phi_n^{<x} \rightarrow \exists y(x < y))$$

(ϕ_n defined in 3.2) holds in B_k (for $2^n \leq 2^{n+1} - 2$), α has an element t of height $2^n - 2$.

Now $\{s \in \alpha | t < s\}$ is a branch above t of length $\ell - (2^n - 1) < 2^{n+1} - 2 - (2^n - 1) = 2^n - 1$; hence, by 3.3, every branch above t has length $\ell - (2^n - 1)$. Now the quantifier rank-$(n+1)$-sentence $\forall x(\phi_n^{<x} \vee \phi_n^{>x})$ is satisfied in B_k; on the other hand, $x = t$ is a counterexample in T.

3.5 Proposition *For each $m < 2^n - 1$ there is a quantifier rank $\leq n$-sentence ϕ_m^n such that for all trees T: $T \models \phi_m^n$ iff all branches through T have length m.*

Proof: Left to the reader.

3.6 Corollary *If $m < 2^n - 1$ and T is a binary tree such that $T \equiv^n B_m$ then $T \cong B_m$.*

3.7 Corollary *$B_m \equiv^n B_k$ iff $m = k$ or $m, k \geq 2^n - 1$.*

REFERENCE

[1] Rosenstein, J. G., *Linear Orderings*, Academic Press, New York, 1982.

Mathematics Institute
University of Amsterdam
Amsterdam, The Netherlands