Discrete Mathematics 342 (2019) 152–167

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Existential monadic second order logic on random rooted trees

Alexander E. Holroyd [a], Avi Levy [b], Moumanti Podder [d,*], Joel Spencer [c]

[a] Microsoft Research, 1 Microsoft Way, Redmond, WA 98052, United States
[b] University of Washington, Box 354350, Seattle, WA 98195-4350, United States
[c] Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, United States
[d] University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195, United States

ARTICLE INFO

Article history:
Received 26 July 2017
Received in revised form 9 September 2018
Accepted 10 September 2018
Available online xxxx

Keywords:
Galton–Watson tree
Finiteness of tree
Existential monadic second order properties
Almost sure expressibility

ABSTRACT

We address questions of logic and expressibility in the context of random rooted trees. Infiniteness of a rooted tree is not expressible as a first order sentence, but is expressible as an existential monadic second order sentence (EMSO). On the other hand, finiteness is not expressible as an EMSO. For a broad class of random tree models, including Galton–Watson trees with offspring distributions that have full support, we prove the stronger statement that finiteness does not agree up to a null set with any EMSO. We construct a finite tree and a non-null set of infinite trees that cannot be distinguished from each other by any EMSO of given parameters. This is proved via set-pebble Ehrenfeucht games (where an initial colouring round is followed by a given number of pebble rounds).

Published by Elsevier B.V.

1. Introduction

The problem of expressibility in a given language parametrized by mathematical logic has been one of the classically studied problems. In our paper, the setting is the space of all rooted trees, and the language is the set of all existential monadic second order sentences on rooted trees. The property we specifically focus on is the finiteness of the rooted tree.

Before we go into the details of the questions and how we seek to answer them, we point out here that such questions have important implications in descriptive complexity theory. Descriptive complexity measures the syntactic complexity of formulae that express a certain property, instead of its computation complexity. See [5,8] and [10] for more general discussions on this theory. A fundamental result in this area is the well-known *Fagin's theorem*, which states that a property is in the class NP (non-deterministic polynomial time computability) if and only if it is describable as an existential second-order logical sentence (see [3]). If one can prove that the class of all existential second-order sentences is not closed under negation, one shall establish that NP $\neq$ co-NP and therefore P $\neq$ NP.

It is possible to express the property of infiniteness of the rooted tree, in a simple way, as an existential monadic second order sentence (EMSO) (see [1,2,4] and [11] for more on EMSO). This naturally raises the question as to whether finiteness of the rooted tree, the negation of infiniteness, can also be expressed as an EMSO. The objective of this paper is to tie in probability with this question. We ask if it is possible, under a measure μ that satisfies certain naturally occurring conditions, that finiteness is expressible as an EMSO on all but a subset of trees of measure 0. We answer this question in the negative, and the result is stated in Theorem 1.1. As a straightforward consequence of Theorem 1.1, one can conclude that finiteness is not expressible tautologically as an EMSO.

* Correspondence to: 5215 15th Avenue NE, Apartment 16, Seattle, WA 98105, United States.
E-mail addresses: holroyd@uw.edu (A.E. Holroyd), avius@uw.edu (A. Levy), moumantip3@gmail.com (M. Podder), spencer@cims.nyu.edu (J. Spencer).

https://doi.org/10.1016/j.disc.2018.09.012
0012-365X/Published by Elsevier B.V.

Henceforth, $\mathcal{T}$ denotes the set of all rooted trees on a set of nodes either enumerated by $\{1, \ldots, n\}$ for any $n \in \mathbb{N}$, or by $\mathbb{N}$, such that every vertex in the tree has finite degree. The notations used in the statement of Theorem 1.1 are explained in Section 1.1.

Theorem 1.1. *Let μ be a probability measure on $\mathcal{T}$ that assigns positive probability to every set of the form $\{T : T|_n = T'\}$, where $n \in \mathbb{N}$ and T' is a finite tree. Then, the property of finiteness of rooted trees is not* almost expressible *as an existential monadic second order sentence under the measure μ. That is, there do not exist a μ-null set of infinite trees $\mathcal{T}_0$ and an EMSO A, such that every finite tree satisfies A and every infinite tree in $\mathcal{T}_0^c$ satisfies the negation of A.*

Such a measure assigns positive weight to every finite tree. In particular, we can consider the measure induced by the well-known Galton–Watson branching process with an offspring distribution χ that is supported on all of $\mathbb{N}_0$, the set of non-negative integers, with expectation greater than 1 (i.e. the supercritical regime). An example is the Poisson distribution with expectation greater than 1. In the following subsection, we set down the notations we use throughout the paper.

1.1. Some notation

We denote by T_μ the random rooted tree which follows the measure μ. For any tree $T \in \mathcal{T}$, we let $V(T)$ denote its set of nodes. For any $v \in V(T)$, we let $d(v)$ denote the depth of v in T, where the root, usually denoted ϕ, has depth $d(\phi) = 0$. For $v \in V(T)$, let $T(v)$ denote the subtree of T that is rooted at v, i.e. it is the subtree of T that consists of v and all the descendants of v in T. When v is a child of the root, we call $T(v)$ a *principal branch* of T. Let $\pi(v)$ denote the parent of v, for any $v \in V(T) \setminus \{\phi\}$. For a positive integer n and $T \in \mathcal{T}$, let $T|_n$ denote the truncation of T, consisting of all nodes of depth at most n. For two given rooted trees T and T', with roots ϕ and ϕ' respectively, the statement $T = T'$ implies that there exists an isomorphism between T and T', i.e. a mapping $\varphi : V(T) \to V(T')$ such that $\varphi(\phi) = \phi'$ and for every $u, v \in V(T)$, we have $\pi(v) = u$ if and only if $\pi(\varphi(v)) = \varphi(u)$. For a positive integer k, we set $[k] = \{0, 1, \ldots, k\}$.

1.2. EMSO on trees

Existential monadic second order (EMSO) sentences on $\mathcal{T}$ are of the form

$$\exists S_1, \ldots, \exists S_n[P]$$

where $S_1, \ldots, S_n$ are subsets of nodes of the tree, and P is a first-order sentence that involves the root as a constant symbol and the relations $=$ (equality of nodes), π (parent–child relationship) and $\in$ (inclusion in one of the subsets $S_1, \ldots, S_n$). Often, it is more easily visualizable if we identify the subsets $S_1, \ldots, S_n$ with colours, i.e. we partition the set of all rooted trees into n colour classes. A classical example would be the infiniteness of the tree, which is expressible as follows:

$$\exists S\Big[[\phi \in S] \wedge [\forall u \in S\, [\exists v \in S[\pi(v) = u]]]\Big]. \tag{1.1}$$

In words, this asserts that there exists a set S of nodes containing the root, such that every element u of S has a child v in S. As mentioned earlier, this paper is concerned with showing that the complementary event i.e. that the tree dies out, is not expressible as an EMSO almost surely under any probability measure μ on $\mathcal{T}$ that satisfies the hypotheses of Theorem 1.1. Our proof technique relies on a suitable version of the well-known Ehrenfeucht games, the set-pebble Ehrenfeucht games (see Definition 2.2). For a discussion of Ehrenfeucht games, we refer the readers to [9] and [7].

2. Rooted colourings, set-pebble Ehrenfeucht and types games

In this paper, we fix an arbitrary positive integer r and consider a set $\Sigma = \{\mathrm{col}_0, \ldots, \mathrm{col}_r\}$ of $r + 1$ colours. Later on, we shall consider a set $\overline{\Sigma}$ of "augmented" colours that is derived from Σ.

Definition 2.1 (*Rooted Colouring*)**.** Given the set Σ of colours, and a tree $T \in \mathcal{T}$, we call a colouring $\sigma : V(T) \to \Sigma$ a (Σ, col_0)-*rooted colouring* of T if $\sigma(v) = \mathrm{col}_0 \Leftrightarrow v = \phi$, for all $v \in V(T)$.

We insist upon assigning a unique colour to the root because it is a constant symbol in our language. Given a $T \in \mathcal{T}$ and a (Σ, col_0)-rooted colouring σ of T, we shall talk of the pair (T, σ) in the subsequent exposition, and call it a (Σ, col_0)-*coloured tree*.

As previously mentioned, the set-pebble Ehrenfeucht game, defined below, will be our main tool in proving that indeed, there is a set of trees of positive measure on which finiteness is not an EMSO. A generalized version of this game is described in Definition 7.17 of [6], where it is referred to as the *Fagin game*.

In all games considered in this paper, when we say that a particular player wins, we mean with optimal play by both the players. In other words, the player who wins has a strategy that guarantees a win regardless of the other player's moves.

Definition 2.2 (*The Set-Pebble Ehrenfeucht Game*)**.** This game is played between two players, Spoiler and Duplicator. They are given two trees T_1, T_2 with roots ϕ_1, ϕ_2, and a positive integer k. The game consists of $k + 1$ rounds, and each round consists of a move by Spoiler and a subsequent move by Duplicator. These rounds can be divided into the following: