MONADIC SECOND ORDER LOGIC WITH MEASURE AND CATEGORY QUANTIFIERS

MATTEO MIO, MICHAŁ SKRZYPCZAK, AND HENRYK MICHALEWSKI

CNRS and École Normale Supérieure de Lyon
e-mail address: matteo.mio@ens-lyon.fr

Institute of Informatics, University of Warsaw
e-mail address: m.skrzypczak@mimuw.edu.pl

Institute of Mathematics, Polish Academy of Sciences
e-mail address: h.michalewski@mimuw.edu.pl

ABSTRACT. We investigate the extension of Monadic Second Order logic, interpreted over infinite words and trees, with generalized "for almost all" quantifiers interpreted using the notions of Baire category and Lebesgue measure.

1. INTRODUCTION

Monadic Second Order logic (MSO) is an extension of first order logic (FO) allowing quantification over subsets of the domain ($\forall X. \phi$) and including a membership predicate between elements of the domain and its subsets ($x \in X$). Two fundamental results about MSO, due to Büchi and Rabin respectively, state that the MSO theories of $(\mathbb{N}, <)$ and of the *full binary tree* (see Section 3 for precise definitions) are decidable.

A different kind of extension of first order logic which has been studied in the literature, starting from the seminal works of Mostowski [21], is by means of *generalized quantifiers*. A typical example is the "there exists infinitely many" quantifier:

$$\mathcal{M} \models \exists^\infty x. \phi \iff \text{the set } \{m \in M \mid \phi(m)\} \text{ is infinite.}$$

The formula $\neg(\exists^\infty x. (x = x))$ holds only on finite models and is therefore not expressible in first order logic. Hence, in general, $\text{FO} \subsetneq \text{FO} + \exists^\infty$. Note however that on some fixed structures the logic FO and $\text{FO} + \exists^\infty$ may have the same expressive power. This is the case, for example, for the structure $(\mathbb{N}, <)$.

It is natural to try to merge the two approaches and consider extensions of MSO with generalized quantifiers. In this context, one can consider generalized first order quantifiers as well as generalized second order quantifiers. For example, the second order generalized quantifier $(\exists^\infty X. \phi)$ has the following meaning:

$$\mathcal{M} \models \exists^{\infty} X. \phi \iff \text{the set } \{A \subseteq M \mid \phi(A)\} \text{ is infinite.}$$

The topic of MSO with generalized second order quantifiers is not new. In particular, Bojańczyk in 2004 ([3]) has initiated an investigation of the *unbounding quantifier* $(\mathsf{U}X. \phi)$

$$\mathcal{M} \models \mathsf{U}X. \phi \iff \forall n \in \mathbb{N}. \exists A \subseteq M. \big(A \text{ is finite, } |A| \geq n, \phi(A)\big).$$

It was shown recently that the MSO $+ \mathsf{U}$ theory of $(\mathbb{N}, <)$ is undecidable [7].

Bárány, Kaiser and Rabinovich have investigated in [2] the second order quantifier $\exists^{\infty}$ mentioned above and also the cardinality quantifier $\exists^{\mathfrak{c}}$ defined as:

$$\mathcal{M} \models \exists^{\mathfrak{c}} X. \phi \iff \text{the set } \{A \subseteq M \mid \phi(A)\} \text{ has cardinality } \mathfrak{c} = 2^{\aleph_0}.$$

Interestingly it was shown that, on the structures $(\mathbb{N}, <)$ and the full binary tree, the quantifiers $\exists^{\infty}$ and $\exists^{\mathfrak{c}}$ can be eliminated. That is, to each formula of MSO $+ \exists^{\mathfrak{c}} + \exists^{\infty}$ corresponds (effectively) a semantically equivalent formula of MSO.

In a series of manuscripts in the late 70's (see [25] for an overview) H. Friedman investigated two kinds of first-order generalized quantifiers called (Baire) *Category quantifier* $(\forall^*)$ and (full probability) *Measure quantifier* $(\forall^{=1})$, respectively.

The goal of this article is to collect recent results (from [18], [19] and [20]) regarding extensions of MSO with second order Category and Measure quantifiers. The meaning of such second order quantifiers is given as follows:

$$\mathcal{M} \models \forall^* X. \phi \iff \text{the set } \{A \subseteq M \mid \phi(A)\} \text{ is comeager.}$$

and

$$\mathcal{M} \models \forall^{=1} X. \phi \iff \text{the set } \{A \subseteq M \mid \phi(A)\} \text{ has Lebesgue measure 1.}$$

We will exclusively be interested in the properties of MSO$+\forall^*$ and MSO$+\forall^{=1}$ interpreted over the two fundamental structures $(\mathbb{N}, <)$ and the full binary tree. Therefore, the domain of $\mathcal{M}$ is always countable and the space of subsets $A \subseteq M$ is a copy of the Cantor space 2^{ω} where the notions of (co)meager sets and Lebesgue (i.e., coin-flipping) measure are well defined. See Section 2 for precise definitions.

Results regarding $S1S$**.** By denoting with $S1S$ the MSO theory of $(\mathbb{N}, <)$ we prove (Theorem 4.1) that the Category quantifier $(\forall^*)$ can be (effectively) eliminated. Therefore $S1S + \forall^* = S1S$ and its theory is decidable. On the other hand, the addition of the Measure quantifier $(\forall^{=1})$ leads to an undecidable theory (Theorem 5.2). This was proved in [18] using recent results from the theory of probabilistic automata. In this article we give a different proof of undecidability of $S1S + \forall^{=1}$ based on an interpretation of Bojańczyk's undecidable theory $S1S + \mathsf{U}$.

Results regarding $S2S$**.** By denoting with $S2S$ the MSO theory of the full binary tree, the theory of $S1S + \forall^{=1}$ can be interpreted within $S2S + \forall^{=1}$. As a consequence $S2S + \forall^{=1}$ has an undecidable theory (Theorem 7.1).

In [18] it was claimed that the theory $S2S + \forall^*$ is effectively equivalent to $S2S$. The proof relied on an automata-based quantifier elimination procedure. Unfortunately, the proof contains a crucial mistake which we do not know how to fix. The mistake is caused by a wrong citation in Proposition 1: although the automaton $\mathcal{A}^{\exists}$ recognizes the language $\exists X. \phi(X, \overrightarrow{Y})$, this may not be the case for the automaton $\mathcal{A}^{\forall}$ and the language $\forall X. \phi(X, \overrightarrow{Y})$. Because of that, the automaton $\mathcal{A}_{\psi}$ constructed in Definition 6 does not fairly simulate the Banach–Mazur game in the moves of the universal player $\forall$.

The proof technique adopted in [18], while faulty in the general case, leads to the following weaker result (Theorem 6.1): if $\phi(X, \vec{Y})$ is a $S2S$ formula (with parameters X, $\vec{Y}$) definable by a *game automaton*, then the quantifier $\forall^* X. \phi(X, \vec{Y})$ can be effectively eliminated. That is, there (effectively) exists a $S2S$ formula $\psi(\vec{Y})$ equivalent to $\forall^* X. \phi(X, \vec{Y})$. Comparing to the proof presented in [18], the proof in this paper is shorter and more precise.

The question of whether the full logic $S2S + \forall^*$ has a decidable theory remains open from this work. See the list of open problems in Section 10.

Results regarding $S2S$ with generalized path quantifiers. Another interesting way to extend $S2S$ with variants of Friedman's Category and Measure quantifiers is to restrict the quantification to range over infinite branches (paths) of the full binary tree. This leads to the definition of the path-Category ($\forall_\pi^* X. \phi$) and the path-Measure ($\forall_\pi^{=1} X. \phi$) second-order quantifiers.

A path π in the full binary tree is an infinite set of vertices which can be uniquely identified with an infinite set of direction $\pi \in \{\mathtt{L}, \mathtt{R}\}^\omega$. By denoting with $\mathcal{P}$ the collection $\{\mathtt{L}, \mathtt{R}\}^\omega$ we define:

$$\forall_\pi^* X. \phi \Longleftrightarrow \{\pi \in \mathcal{P} \mid \phi(\pi)\} \text{ is comeager as a subset of } \mathcal{P}$$

and

$$\forall_\pi^{=1} X. \phi \Longleftrightarrow \{\pi \in \mathcal{P} \mid \phi(\pi)\} \text{ has measure 1 as a subset of } \mathcal{P}$$

We show that the path-Category quantifier can be (effectively) eliminated so that $S2S + \forall_\pi^* \equiv S2S$. On the other hand, we show that $S2S + \forall_\pi^{=1} \supsetneq S2S$ is a strict extension of $S2S$. The question of whether $S2S + \forall_\pi^{=1}$ has a decidable theory remains open from this work. See the list of open problems in Section 10.

Related Work. A significant amount of work has been devoted to the study of properties of regular sets of words and infinite trees with respect to Baire Category and probability. For example, a remarkable result of Staiger ([24, Theorem 4]) states that a regular set $L \subseteq \Sigma^\omega$ of ω-words is comeager if and only if $\mu(L) = 1$, where μ is the probability (coin flipping) measure on ω-words. Furthermore the probability $\mu(L)$ can be easily computed by an algorithm based on linear programming and it is always a rational number (see, e.g., Theorem 2 of [11]). Staiger's theorem was rediscovered by Varacca and Völzer and has been used to develop a theory of fairness for reactive systems [27].

In another direction, Gräedel have investigated determinacy, definability, and complexity issues of Banach–Mazur games played on graphs [13]. Our results regarding $S1S + \forall^*$ are closely related to the results of Gräedel.

The study of Category and Measure properties of regular sets of infinite trees happens to be more complicated. Indeed, due to their high topological complexity, it is not even clear if such sets are Baire-measurable and Lebesgue measurable. These properties were established only recently (see [14] for Baire measurability and [12] for Lebesgue measurability). Furthermore, it has been shown in [19] that the equivalent of Staiger's theorem fails for infinite trees. That is, there exist comeager regular sets of trees having probability 0.

The problem of computing the probability of a regular set of infinite trees is still open in the literature. An algorithm for computing the (generally irrational) probability of regular sets definable by game automata is presented in [19] but a solution for the general case is still unknown (and the problem could be undecidable). The result of Theorem 6.1 proved in

this article imply that it is possible to determine the Baire category of regular sets definable by game automata. A solution to the general case is unknown.

All the results mentioned above contribute to the theory of regular sets of ω-words ($S1S$ definable) and infinite trees ($S2S$ definable). The novelty of the approach followed in this research is instead to augment the base logic ($S1S$ or $S2S$) with generalized quantifiers capable of expressing properties related to Baire category and probability, and to study such logical extensions.

Interestingly, a different but related approach has been explored by Carayol, Haddad and Serre in [8] (see also [9, 10] and the habilitation thesis [23]). Rather than extending the base logic, it is the base (Rabin) tree automata model which is modified. The usual acceptance condition "there exists a run such that *all* paths are accepting" is replaced by the weaker "there exists a run such that *almost all* paths are accepting". When *almost all* is interpreted as "for a set of measure 1", this results in the notion of qualitative automata of [8]. The connections between this automata-oriented and our logic-oriented approach is further discussed in Section 9.

Addendum: after the submission of this article, we have been informed by M. Bojańczyk that a proof of the decidability of the theory of weak$S2S + \forall_\pi^{=1}$, the fragment of $S2S + \forall_\pi^{=1}$ where second-order quantification is restricted to finite sets, can be obtained from the results of [4, 6].

2. Technical Background

In this section we give the basic definitions regarding descriptive set theory, probability measures and tree automata that are needed in this work.

The set of natural numbers and their standard total order are denoted by the symbols ω and $<$, respectively. Given sets X and Y we denote with X^Y the space of functions $X \to Y$. We can view elements of X^Y as Y-indexed sequences $\{x_i\}_{i \in Y}$ of elements of X. Given a finite set Σ, we refer to Σ^ω as the collection of *ω-words over* Σ. The collection of *finite* sequences of elements in Σ is denoted by Σ^*. As usual we denote with ϵ the empty sequence and with ww' the concatenation of $w, w' \in \Sigma^*$. The prefix order on sequences is denoted $\preceq$ i.e., $u \preceq w$ if u is a prefix of w.

We identify finite sequences $\{\text{L}, \text{R}\}^*$ as vertices of the infinite full binary tree so that ϵ is the root and each vertex $w \in \{\text{L}, \text{R}\}^*$ has $w\text{L}$ and $w\text{R}$ as children. We denote by $\mathcal{T}_\Sigma$ the set $\Sigma^{\{\text{L},\text{R}\}^*}$ of functions labelling each vertex with an element in Σ. Elements $t \in \mathcal{T}_\Sigma$ are called *trees over* Σ.

2.1. **Games.** In this paper we work with games of perfect information and infinite duration commonly referred to as Gale–Stewart games. In what follows we provide the basic definitions. We refer to [16, Chapter 20.A] for a detailed overview.

Gale–Stewart games are played between two players: Player $\exists$ and Player $\forall$. A game is given by a set of configurations V; an initial configuration $v_0 \in V$; a winning condition $W \subseteq V^\omega$; and a definition of a *round*. A *round* starts in a configuration v_n; then players perform a fixed sequence of choices; and depending on these choices the round ends in a new configuration v_{n+1}. Each game specifies what are the exact choices available to the players and how the successive configuration v_{n+1} depends on the decisions made.

A *play* of such a game is an infinite sequence of configurations $v_0, v_1, \ldots$ where the configuration v_{n+1} is a possible outcome of the n-th round that starts in a configuration v_n.

A play is winning for Player $\exists$ if it satisfies the winning condition (i.e. $(v_0, v_1, \dots) \in W$); otherwise the play is winning for Player $\forall$. A *strategy* of a player is a function that assigns to each finite sequence of configurations $(v_0, v_1, \dots, v_n)$ a way of resolving the choices of that player in the consecutive n-th round that starts in v_n. A play is *consistent* with a strategy if it can be obtained following that strategy. A strategy of a player is *winning* if all the plays consistent with that strategy are winning for the player. We say that a game is *determined* if one of the players has a winning strategy.

2.2. **Topology.** Our exposition of topological and set–theoretical notions follows [16].

The set $\{0, 1\}^\omega$ endowed with the product topology (where $\{0, 1\}$ is given the discrete topology) is called the *Cantor space*. The Cantor space is a Polish space (i.e., a complete separable metric space) and is *zero-dimensional* (i.e., it has a basis of *clopen* sets). Given a finite set Σ with at least two elements, the spaces Σ^ω and $\mathcal{T}_\Sigma$ are both homeomorphic to the Cantor space.

A subset $A \subseteq X$ of a topological space is *nowhere dense* if the interior of its closure is the empty set, that is $(\texttt{int}(\texttt{cl}(A)) = \emptyset$. A set $A \subseteq X$ is of *(Baire) first category* (or *meager*) if A can be expressed as countable union of nowhere dense sets. A set $A \subseteq X$ which is not meager is *of the second (Baire) category*. The complement of a meager set is called *comeager*. Recall that a set $A \subseteq \{0, 1\}^\omega$ is a G_δ *set* if it can be expressed as a countable intersection of open sets. The following property of comeager sets will be useful. A subset $B \subseteq \{0, 1\}^\omega$ is comeager if and only if there exists a dense G_δ set $A \subseteq B$ which is comeager.

A set $B \subseteq X$ has the *Baire property* if $X = U \triangle M$, for some open set $U \subseteq X$ and meager set $M \subseteq X$, where $\triangle$ is the operation of symmetric difference $X \triangle Y = (X \cup Y) \setminus (X \cap Y)$.

We now describe Banach–Mazur game (see [16, Chapter 8.H] for a detailed overview) which characterizes Baire category. We specialize the game to zero–dimensional Polish spaces, such as Σ^ω and $\mathcal{T}_\Sigma$, since we only deal with this class of spaces in this work.

Definition 2.1 (Banach–Mazur Game). Let X be a zero–dimensional Polish space. Fix a payoff set $A \subseteq X$. The configurations of the game $\mathbf{BM}(X, A)$ are non-empty clopen sets $U \subseteq X$; the initial configuration is X. Consider a round starting in a configuration U_n. In such a round, first Player $\forall$ chooses a non-empty clopen set $U_n' \subseteq U_n$ and then Player $\exists$ chooses a non-empty clopen set $U_{n+1} \subseteq U_n'$. The consecutive configuration is U_{n+1}. An infinite play of that game is won by Player $\exists$ if $\bigcap_{n \in \omega} U_n \cap A \neq \emptyset$ and Player $\forall$ wins otherwise.

The standard way of defining that game is to represent it as a form of a Gale–Stewart game.

Player $\forall$	U_0'		U_1'		$\dots$
Player $\exists$		U_1		U_2	$\dots$

For a proof of the following result see, e.g., Theorem 8.33 in [16].

Theorem 2.2. *Let X be a zero–dimensional Polish space. If $A \subseteq X$ has the Baire property, then $\mathbf{BM}(X, A)$ is determined and Player $\exists$ wins iff A is comeager.*

If X is the space of ω-words or infinite trees over a finite alphabet Σ, the above game can be simplified as in Definitions 2.3 and 2.4 below.

Definition 2.3. Fix a language $L \subseteq \Sigma^\omega$ of ω-words over Σ. The *Banach–Mazur game* on L (denoted $\mathbf{BM}(L)$) is the following game with configurations of the form (s, r) where $s \in \Sigma^*$

is a finite word and $r \in \{\exists, \forall\}$. The initial configuration (s_0, r_0) is $(\epsilon, \forall)$. Consider a round $n = 0, 1, \ldots$ that starts in a configuration (s_n, r_n). In this round, the player r_n chooses a finite word s_{n+1} that contains s_n as a strict prefix and the round ends in the configuration $(s_{n+1}, \bar{r_n})$ where $\bar{r_n}$ is the opponent of r_n.

An infinite play of this game is winning for $\exists$ if $\alpha \in L$, where $\alpha = \bigcup_{n \in \omega} s_n$ is the unique ω-word that has all s_n as prefixes.

Now we move to the tree variant of this game. A *tree-prefix* is a partial function $s \colon \{\texttt{L}, \texttt{R}\}^* \rightharpoonup \Sigma$ where the domain $\text{dom}(s)$ is finite, prefix-closed, and each node of $\text{dom}(s)$ has either zero (a leaf) or two (an internal node) children in $\text{dom}(s)$. We say that s' *extends* s if $s' \neq \emptyset$, $s \subseteq s'$, and each leaf of s is an internal node of s'.

Definition 2.4. Fix a language $L \subseteq \mathcal{T}_\Sigma$ of trees over Σ. The *Banach–Mazur game* on L (denoted $\mathbf{BM}(L)$) is the following game with configurations of the form (s, r) where s is a tree-prefix and $r \in \{\exists, \forall\}$. The initial configuration (s_0, r_0) is $(\emptyset, \forall)$. Consider a round $n = 0, 1, \ldots$ that starts in a configuration (s_n, r_n). In this round, the player r_n chooses a tree-prefix s_{n+1} that extends s_n and the round ends in the configuration $(s_{n+1}, \bar{r_n})$ where $\bar{r_n}$ is the opponent of r_n. An infinite play of this game is winning for Player $\exists$ if $t \in L$, where $t = \bigcup_{n \in \omega} s_n$.

Theorem 2.5. *A language L of ω-words or trees over Σ with the Baire property is comeager if and only if $\exists$ has a winning strategy in the Banach–Mazur game on L.*

2.3. **Probability Measures.** We summarize below the basic concepts related to Borel measures. For more details see, e.g, [16, Chapter 17].

In what follows, let X be a zero–dimensional Polish space. The smallest σ-algebra of subsets of X containing all open sets is denoted by $\mathcal{B}$ and its elements are called *Borel sets*. A *Borel probability measure* on X is a function $\mu \colon \mathcal{B} \to [0, 1]$ such that: $\mu(\emptyset) = 0$, $\mu(X) = 1$ and, if $\{B_n\}_{n \in \omega}$ is a sequence of disjoint Borel sets, $\mu\left(\bigcup_n B_n\right) = \sum_n \mu(B_n)$.

A subset $A \subseteq X$ is a μ-*null set* if $A \subseteq B$ for some Borel set $B \subseteq X$ such that $\mu(B) = 0$.

A subset $A \subseteq X$ is a F_σ *set* if it can be expressed as a countable union of closed sets. Every Borel measure μ on a X is *regular*: for every Borel set B there exists an F_σ set $A \subseteq B$ such that $\mu(A) = \mu(B)$.

Lebesgue measure on ω-words. For a fixed finite non-empty set $\Sigma = \{a_1, \ldots, a_k\}$, we will be exclusively interested in one specific Borel measure μ_ω on the space Σ^ω which we refer to as *Lebesgue measure* (on ω-words). This is the unique Borel measure satisfying the equality $\mu_w(B_{n=a}) = \frac{1}{k}$ where $B_{n=a} = \{(a_i)_{i \in \omega} \mid a_n = a\}$, for each $n \in \mathbb{N}$ and $a \in \Sigma$. Intuitively, the Lebesgue measure on ω-words generates an infinite sequence $(a_0, a_1, \ldots)$ by labelling the n-th entry with some letter $a \in \Sigma$ chosen randomly and uniformly. This is done independently for each position $n \in \omega$.

Lebesgue measure on trees. Similarly, for a finite non-empty set $\Sigma = \{a_1, \ldots, a_k\}$, the *Lebesgue measure* on trees $\mathcal{T}_\Sigma$ is the unique Borel measure, denoted by μ_t, satisfying $\mu_t(B_{v=a}) = \frac{1}{k}$ where $B_{v=a} = \{t \in \Sigma^{\{\texttt{L}, \texttt{R}\}^*} \mid t(v) = a\}$, for $v \in \{\texttt{L}, \texttt{R}\}^*$ and $a \in \Sigma$. Once again, the Lebesgue measure on trees generates an infinite tree over Σ by labelling each vertex v with a randomly chosen letter $a \in \Sigma$. This is done independently for each vertex $v \in \{\texttt{L}, \texttt{R}\}^*$.

2.4. **Large Sections Projections.** Let X and Y be Polish spaces. We denote with $X \times Y$ the product space endowed with the product topology. Given a subset $A \subseteq X \times Y$ and an element $x \in X$ we denote with $A_x \subseteq Y$ the *section of A at x* defined as: $A_x \stackrel{\text{def}}{=} \{y \in Y \mid (x, y) \in A\}$.

A fundamental operation in descriptive set theory, corresponding to the logical operation of universal quantification, is that of taking full sections: given $A \subseteq X \times Y$ we denote with $\forall_Y A \subseteq X$ the set

$$\forall_Y A \stackrel{\text{def}}{=} \{x \in X \mid A_x = Y\}$$

In other words, $\forall_Y A$ is the collection of those $x \in X$ such that for all $y \in Y$ the pair (x, y) is in A, see Figure 1.
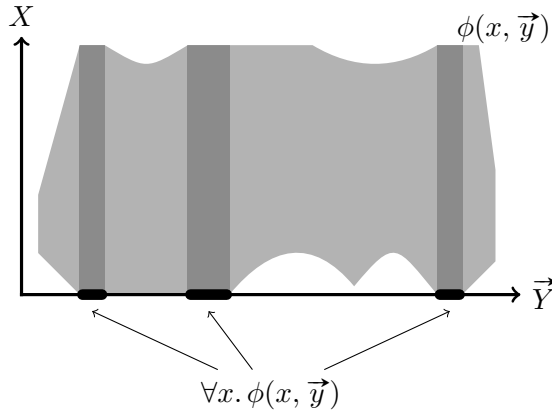

Figure 1: The universal quantifier $\forall$. A formula $\forall x. \phi(x, \vec{y})$ holds for parameters $\vec{y} \in \vec{Y}$ if for every choice of $x \in X$ we have $\phi(x, \vec{y})$. The full sections selected by this quantifier are marked in gray.

Definition 2.6 (Projective sets)**.** The family of projective sets is the smallest family containing Borel sets and closed under Boolean operations and full sections.

The content of this article is based on other operations studied in descriptive set theory (see, e.g., [16, § 29.E]) based on the idea of taking *large sections*, rather than full sections, see Figure 2. From a logical point of view, these operations corresponds to some kind of weakened "for almost all" quantification.

The two notions of "largeness" that are relevant in this work are that of being *comeager* and having *Lebesgue measure* 1, see Figure 2.
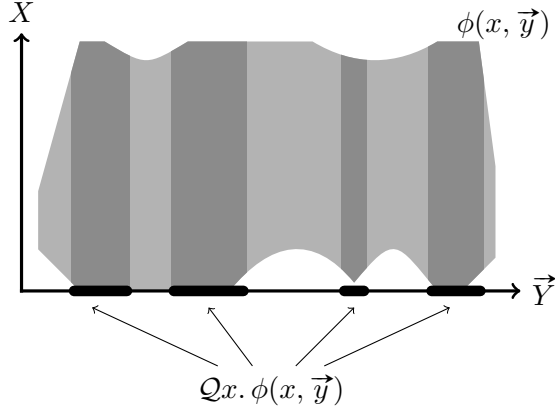
Figure 2: A large section quantifier $\mathcal{Q}$, for instance $\mathcal{Q}$ can be $\forall$, $\forall^*$, $\forall^{=1}$, etc. A formula $\mathcal{Q}x.\,\phi(x, \overrightarrow{y})$ holds for parameters $\overrightarrow{y} \in \overrightarrow{Y}$ if there are in some sense many (but possibly not all) values $x \in X$ such that $\phi(x, \overrightarrow{y})$ holds. The large sections selected by this quantifier are marked in gray.

Correspondingly, given $A \subseteq X \times Y$, we denote with $\forall_Y^* A \subseteq X$ and $\forall_Y^{=1} A \subseteq X$ the sets defined as:

$$\forall_Y^* A \overset{\text{def}}{=} \{x \in X \mid A_x \text{ is comeager}\}$$

$$\forall_Y^{=1} A \overset{\text{def}}{=} \{x \in X \mid A_x \text{ has Lebesgue measure 1}\}$$

If the space Y is known from the context, we skip the subscript Y. We refer to the operations $\forall^*$ and $\forall^{=1}$ as the *(Baire) Category quantifier* and the *(Lebesgue) Measure quantifier*, respectively.

Remark 2.7. It is consistent with ZFC that, e.g., there exists a projective set A (even a Σ_2^1 set) which does not have the Baire property and is not Lebesgue measurable. However, the sets $\forall_Y^{=1} A$ and $\forall_Y^*$ are well-defined even if the set A is not Lebesgue measurable or it does not have the Baire property.

The following classical result states that the above operations preserve the class of Borel sets.

Theorem 2.8. *If $A \subseteq X \times Y$ is a Borel set then $\forall_Y^* A$ and $\forall_Y^{=1} A$ are Borel sets.*

A proof of the above theorem can be obtained from [16, Theorem 29.22] for $\forall^*$ and [16, Theorem 29.26] for $\forall^{=1}$, which are more generally stated for analytic sets, and an application of Souslin theorem [16, Theorem 14.11].

2.5. **Alternating automata.** Given a finite set X, we denote with $\mathcal{DL}(X)$ the set of expressions e generated by the grammar

$$e ::= x \mid e \wedge e \mid e \vee e \quad \text{for } x \in X$$

An *alternating tree automaton* over a finite alphabet Σ is a tuple $\mathcal{A} = \langle \Sigma, Q, q_{\mathrm{I}}, \delta, \mathcal{F} \rangle$ where Q is a finite *set of states*, $q_{\mathrm{I}} \in Q$ is the *initial state*, $\delta \colon Q \times \Sigma \to \mathcal{DL}(\{\mathtt{L}, \mathtt{R}\} \times Q)$ is the *alternating*

transition function, $\mathcal{F} \subseteq \mathcal{P}(Q)$ is a set of subsets of Q called the *Muller condition*. The Muller condition $\mathcal{F}$ is called a *parity condition* if there exists a *parity assignment* $\pi \colon Q \to \omega$ such that: $\mathcal{F} = \{F \subseteq Q \mid (\max_{q \in F} \pi(q))$ is even$\}$.

An alternating automaton $\mathcal{A}$ over an alphabet Σ defines, or "accepts", a set of Σ-trees. The *acceptance* of a tree $t \in \mathcal{T}_\Sigma$ is defined via a two-player ($\exists$ and $\forall$) game of infinite duration denoted as $\mathcal{A}(t)$. The configurations of the game $\mathcal{A}(t)$ are of the form $\langle u, q\rangle$ or $\langle u, e\rangle$ with $u \in \{\texttt{L}, \texttt{R}\}^*$, $q \in Q$, and $e \in \mathcal{DL}(\{\texttt{L}, \texttt{R}\} \times Q)$ (we can implicitly restrict to the set of sub-formulae of the formulae appearing in the alternating transition function δ).

The game $\mathcal{A}(t)$ starts in the configuration $\langle \epsilon, q_\text{I}\rangle$. Game configurations of the form $\langle u, q\rangle$, including the initial state, have only one successor configuration, to which the game progresses automatically. The successor state is $\langle u, e\rangle$ with $e = \delta(q, a)$, where $a = t(u)$ is the labelling of the vertex u given by t. The dynamics of the game at configurations $\langle u, e\rangle$ depends on the possible shapes of e. If $e = e_1 \vee e_2$, then Player $\exists$ moves either to $\langle u, e_1\rangle$ or $\langle u, e_2\rangle$. If $e = e_1 \wedge e_2$, then Player $\forall$ moves either to $\langle u, e_1\rangle$ or $\langle u, e_2\rangle$. If $e = (\texttt{L}, q)$ then the game progresses automatically to the state $\langle u\texttt{L}, q\rangle$. Lastly, if $e = (\texttt{R}, q)$ the game progresses automatically to the state $\langle u\texttt{R}, q\rangle$. Thus a play in the game $\mathcal{A}(t)$ is a sequence Π of configurations, that can look like

$$\Pi = (\langle \epsilon, q_\text{I}\rangle, \ldots, \langle \texttt{L}, q_1\rangle, \ldots, \langle \texttt{LR}, q_2\rangle, \ldots, \langle \texttt{LRL}, q_3\rangle, \ldots, \langle \texttt{LRLL}, q_4\rangle, \ldots),$$

where the dots represent part of the play in configurations of the form $\langle u, e\rangle$. Let $\inf(\Pi)$ be the set of automata states $q \in Q$ occurring infinitely often in the configurations of the form $\langle u, q\rangle$ of Π. We then say that the play Π of $\mathcal{A}(t)$ is winning for $\exists$, if $\inf(\Pi) \in \mathcal{F}$. The play Π is winning for $\forall$ otherwise. The *language* of Σ-trees defined by $\mathcal{A}$ denoted $\text{L}(\mathcal{A})$ is the collection $\{t \in \mathcal{T}_\Sigma \mid \exists$ has a winning strategy in the game $\mathcal{A}(t)\}$.

An alternating automaton for ω-words is defined analogously, except that $\delta \colon Q \times \Sigma \to \mathcal{DL}(Q)$. In that case the the first coordinate of configurations of the game $\mathcal{A}(\alpha)$ is always a natural number, the game moves from a configuration (n, q) to the configuration $(n+1, e)$ with $e = \delta(q, a)$, where $a = \alpha(n)$ is the labelling of the position n given by α. In that case $\text{L}(\mathcal{A}) \subseteq \Sigma^\omega$.

Types of automata. An alternating automaton $\mathcal{A}$ is *non-deterministic* if all the transitions of $\mathcal{A}$ are of the form

$$\big((\texttt{L}, q_{1,\texttt{L}}) \wedge (\texttt{R}, q_{1,\texttt{R}})\big) \ \vee \ \big((\texttt{L}, q_{2,\texttt{L}}) \wedge (\texttt{R}, q_{2,\texttt{R}})\big) \ \vee \ \ldots \ \vee \ \big((\texttt{L}, q_{k,\texttt{L}}) \wedge (\texttt{R}, q_{k,\texttt{R}})\big).$$

An alternating automaton $\mathcal{A}$ is a *game automaton* if all the transitions of $\mathcal{A}$ are of the form $(\texttt{L}, q_\texttt{L}) \, \chi \, (\texttt{R}, q_\texttt{R})$ with $\chi = \vee$ or $\wedge$. An alternating automaton $\mathcal{A}$ is *deterministic* if it is a game automaton with $\chi = \wedge$ in all transitions.

3. Syntax and Semantics of Monadic Second Order Logic

In this section we define the syntax and the semantics of the Monadic Second-order logic interpreted over the linear order of natural numbers (denoted $S1S$) and over the full binary tree (denoted $S2S$). This material is entirely standard. A detailed exposition can be found in [26].

3.1. **Logic $S1S$ and ω-words.** The syntax of $S1S$ formulas is given by the grammar

$$\phi ::= x < y \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\, \phi \mid \forall X.\, \phi$$

where x, y and X, Y range over countable sets of *first-order variables* and *second-order variables*, respectively.

The semantics of $S1S$ formulas on the structure $(\omega, <)$ is defined as expected by interpreting the formula $\forall X.\, \phi$ as quantifying over all subsets $X \subseteq \omega$ and the relation $(\in)$ as the membership relation between elements $x \in \omega$ and sets $X \subseteq \omega$.

For example the formula

$$\exists X.\, \forall x.\, \big((x \in X) \Rightarrow (\exists y.\, y \in X \wedge x < y) \big)$$

expresses the existence of an unbounded set X of natural numbers.

We write $\phi(\vec{x}, \vec{X})$ to specify that the first-order and second-order variables of ϕ are $\vec{x}$ and $\vec{X}$, respectively.

We can always assume that the free variables of a formula ϕ are all second-order variables, i.e., $\phi = \phi(\emptyset, \vec{X})$. Indeed if $\phi(\{x_1, \ldots, x_n\}, \vec{Y})$ is not of this form, we can consider the semantically equivalent formula $\psi(\emptyset, \{X_1, \ldots, X_n\} \cup \vec{Y})$, having second-order $\vec{X}$ variables (constrained to be singletons) simulating first-order ones.

Each set $X \subseteq \omega$ can be identified as the corresponding ω-word over the alphabet $\Sigma = \{0, 1\}$. Hence a $S1S$ formula $\phi(X_1, \ldots, X_n)$ defines a collection of n-tuples of ω-words over $\Sigma = \{0, 1\}$ or, equivalently, a collection of ω-words over the alphabet $\Gamma = \{0, 1\}^n$. The *language defined by ϕ*, denoted by $\mathrm{L}(\phi) \subseteq \Gamma^\omega$ is the set of ω-words over Γ satisfying ϕ.

The following theorem and proposition are well known.

Theorem 3.1. *The theory of $S1S$ is decidable. The expressive power of $S1S$ is effectively equivalent with alternating / non-deterministic / deterministic Muller automata.*

Proposition 3.2. *For every $S1S$ formula ϕ, the language $\mathrm{L}(\phi)$ is a Borel subset of Γ^ω.*

Hence for every $S1S$ formula ϕ, the language $\mathrm{L}(\phi)$ has the Baire property and is Lebesgue measurable.

3.2. **Logic $S2S$ and infinite trees.** We now introduce, following a similar presentation, the syntax and the semantics of the MSO logic on the full binary tree ($S2S$).

The *full binary tree* is the structure $(\{\mathtt{L}, \mathtt{R}\}^*, \mathrm{succ}_L, \mathrm{succ}_R)$ where succ_L and succ_R are binary relations on $\{\mathtt{L}, \mathtt{R}\}^*$ relating a vertex with its left and right child, respectively:

$$\mathrm{succ}_L = \{(w, w\mathtt{L}) \mid w \in \{\mathtt{L}, \mathtt{R}\}^*\} \qquad \mathrm{succ}_R = \{(w, w\mathtt{R}) \mid w \in \{\mathtt{L}, \mathtt{R}\}^*\}$$

The syntax of $S2S$ formulas is given by the grammar

$$\phi ::= \mathrm{succ}_L(x, y) \mid \mathrm{succ}_R(x, y) \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\, \phi \mid \forall X.\, \phi$$

where x, y and X, Y range over countable sets of *first-order variables* and *second-order variables*, respectively.

The semantics of $S2S$ formulae on the full binary tree is defined by interpreting the formula $\forall X.\, \phi$ as quantifying over all subsets $X \subseteq \{\mathtt{L}, \mathtt{R}\}^*$ and $x \in X$ as the membership relations between vertices $x \in \{\mathtt{L}, \mathtt{R}\}^*$ and sets $X \subseteq \{\mathtt{L}, \mathtt{R}\}^*$.

For example the formula

$$\exists X.\, \forall x.\, \big((x \in X) \Rightarrow (\exists y.\, y \in X \wedge \mathrm{succ}_L(x, y)) \big)$$

expresses the existence of a set of vertices which is closed under taking left children.

As for $S1S$, we can assume without loss of generality that the free variables of a $S2S$ formula ϕ are all second-order.

Each set $X \subseteq \{L, R\}^*$ can be identified as a tree $t \in \mathcal{T}_\Sigma$ over the alphabet $\Sigma = \{0, 1\}$. Hence a $S2S$ formula $\phi(X_1, \ldots, X_n)$ defines a collection of n-tuples of trees in $\mathcal{T}_{\{0,1\}}$ or, equivalently, a collection of trees in $\mathcal{T}_\Gamma$ over the alphabet $\Gamma = \{0, 1\}^n$. The set of $t \in \mathcal{T}_\Gamma$ satisfying the formula ϕ is called the *language defined by* ϕ and is denoted by $\mathrm{L}(\phi)$.

The following theorem and proposition are well known.

Theorem 3.3. *The theory of $S2S$ is decidable. The expressive power of* MSO *over infinite trees is effectively equivalent with alternating / non-deterministic Muller automata. The expressive power of deterministic automata is strictly weaker.*

Proposition 3.4. *For every $S2S$ formula ϕ, the language $\mathrm{L}(\phi)$ is a $\mathbf{\Delta}^1_2$ subset of $\mathcal{T}_\Gamma$.*

It has recently been shown that $\mathrm{L}(\phi)$ is always Lebesgue measurable and has the Baire property [12].

Proposition 3.5. *For every $S2S$ formula ϕ the language $\mathrm{L}(\phi)$ has the Baire property and is Lebesgue measurable.*

4. $S1S$ EXTENDED WITH THE CATEGORY QUANTIFIER

In this section we define the syntax and semantics of the logic $S1S$ extended with the *category quantifier* ($\forall^*$). The syntax of $S1S + \forall^*$ extends that of $S1S$ with the new second-order quantifier $\forall^*$ as follows:

$$\phi \quad ::= \quad x < y \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\,\phi \mid \forall X.\,\phi \mid \forall^* X.\,\phi$$

The semantics of the category quantifier is specified as follows.

$$\mathrm{L}\big(\forall^* X.\,\phi(X, Y_1, \ldots, Y_n)\big) = \Big\{ w \in \Gamma^\omega \mid \{w' \in \{0,1\}^\omega \mid \phi(w', w)\} \text{ is comeager} \Big\}$$

where $\Gamma = \{0, 1\}^n$.

In other words, an n-tuple $w = (w_1, \ldots, w_n)$, with $w_i \in \{0,1\}^\omega$, satisfies the formula $\forall^* X.\,\phi(X, Y_1, \ldots, Y_n)$ if for a large (comeager) collection of $w' \in \{0,1\}^\omega$, the extended tuple (w', w) satisfies $\phi(X, Y_1, \ldots, Y_n)$. Informally, w satisfies $\forall^* X.\,\phi(X, \overrightarrow{Y})$ if "for almost all" $w' \in \{0,1\}^\omega$, the tuple (w', w) satisfies ϕ. The set defined by $\forall^* X.\,\phi(X, \overrightarrow{Y})$ can be illustrated as in Figure 2, as the collection of tuples w having a *large* (comeager) section $\phi(X, w)$.

The main result about $S1S + \forall^*$ is the following quantifier elimination theorem which appears to be folklore. We provide a complete proof of this theorem for two reasons: first, instead of just simulating the Banach–Mazur game over ω-words, we want to inductively remove all the meager quantifiers from the formula; second, the proof presented here is a good base for understanding the more difficult proof from Section 6.

Theorem 4.1. *For every $S1S + \forall^*$ formula ϕ one can effectively construct a semantically equivalent $S1S$ formula ψ.*

Corollary 4.2. *The theory of $S1S + \forall^*$ is decidable.*

The rest of this section is devoted to an inductive proof of Theorem 4.1. For that sake it is enough to show how to eliminate one application of the quantifier $\forall^*$. Consider a formula of $S1S + \forall^*$ of the form $\forall^* X.\, \psi$, where ψ is a formula of $S1S$. We will prove that there exists a formula ψ' of $S1S$ that is equivalent to $\forall^* X.\, \psi$. We start by translating ψ into an equivalent deterministic Muller automaton $\mathcal{A}$. Then we apply the following technical proposition that provides us an alternating Muller automaton $\mathcal{B}$, recognising ω-words that satisfy $\forall^* X.\, \psi$. Finally, we can turn $\mathcal{B}$ into a non-deterministic Muller automaton and express $\mathrm{L}(\mathcal{B})$ by an $S1S$ formula ψ'.

Proposition 4.3. *Let $\mathcal{A}$ be a deterministic Muller automaton over ω-words, with the alphabet $\Sigma \times \Gamma$. Consider the language $\forall^*\mathrm{L}(\mathcal{A})$ that contains an ω-word $\alpha \in \Sigma^\omega$ if and only if the set*

$$\left\{ \beta \in \Gamma^\omega \mid (\alpha, \beta) \in \mathrm{L}(\mathcal{A}) \right\} \tag{4.1}$$

is comeager.

Then, $\forall^*\mathrm{L}(\mathcal{A})$ is ω-regular and one can effectively construct an alternating Muller automaton $\mathcal{B}$ for this language. Additionally, the size of the automaton $\mathcal{B}$ is polynomial in the size of $\mathcal{A}$.

Fix a deterministic automaton $\mathcal{A}$ over a product alphabet $\Sigma \times \Gamma$. Let Q be the set of states of $\mathcal{A}$ and $\delta^{\mathcal{A}} \colon Q \times \Sigma \times \Gamma \to Q$ be the transition function: for a triple $(q, a, b) \in Q \times \Sigma \times \Gamma$ it assigns the successive state $\delta^{\mathcal{A}}(q, a, b) \in Q$ that should be reached after reading a letter (a, b) from the state q.

We will construct an alternating automaton $\mathcal{B}$ over the alphabet Σ. Intuitively, $\mathcal{B}$ will simulate the Banach–Mazur game over $\forall^*\mathrm{L}(\mathcal{A})$ in which a new ω-word $\beta \in \Gamma^\omega$ is constructed. During that play, the automaton $\mathcal{A}$ will be run over the pair of ω-words $(\alpha, \beta) \in (\Sigma \times \Gamma)^\omega$ to verify if this pair belongs to $\mathrm{L}(\mathcal{A})$. To achieve this, the set of states of $\mathcal{B}$ will keep track of both, the current state of $\mathcal{A}$ and the player in charge of construction of β. More formally, let the set of states of $\mathcal{B}$ be $Q \times \{\exists, \forall\}$.

The initial state of $\mathcal{B}$ is $(q_\mathrm{I}, \forall)$. Consider a state (q, r) of $\mathcal{B}$ and a letter $a \in \Sigma$. The transition $\delta^{\mathcal{B}}(q, r, a)$ is supposed to allow the player r to choose a letter $b \in \Gamma$ and the next player $r' \in \{\exists, \forall\}$. After that choices are done, the successive state of the automaton should be (q, r') where $q = \delta^{\mathcal{A}}(q, a, b)$. The structure of the automaton $\mathcal{B}$ is depicted in Figure 3. More formally, we achieve that by putting:

$$\delta^{\mathcal{B}}(q, r, a) \overset{\mathrm{def}}{=} \bigchi_{b \in \Gamma} \bigchi_{r' \in \{\exists, \forall\}} (q, r'), \tag{4.2}$$

where χ is $\vee$ if $r = \exists$ and $\wedge$ otherwise, and $\delta^{\mathcal{A}}(q, a, b) = q$.

The acceptance condition of the automaton $\mathcal{B}$ will be the following Muller condition: a set of states $F \subseteq Q \times \{\exists, \forall\}$ belongs to $\mathcal{F}^{\mathcal{B}}$ if either:

(1) all the states $(q, r) \in F$ satisfy $r = \forall$ (i.e. from some point on $\forall$ has always chosen $r' = r = \forall$,

(2) or F contains both states of the form $(q, \exists)$ and $(q', \forall)$; and additionally the set of states $\{q \in Q \mid (q, r) \in F \text{ for some } r \in \{\exists, \forall\}\}$ is accepting for $\mathcal{A}$ (i.e. belongs to $\mathcal{F}^{\mathcal{A}}$).

In other words, if any of the players from some point on chooses $r' = r$ then he or she looses. Otherwise, the game is won by $\exists$ if the sequence of states visited is accepting in $\mathcal{A}$.

The following lemma proves correctness of our construction and thus concludes the proof of Theorem 6.1.
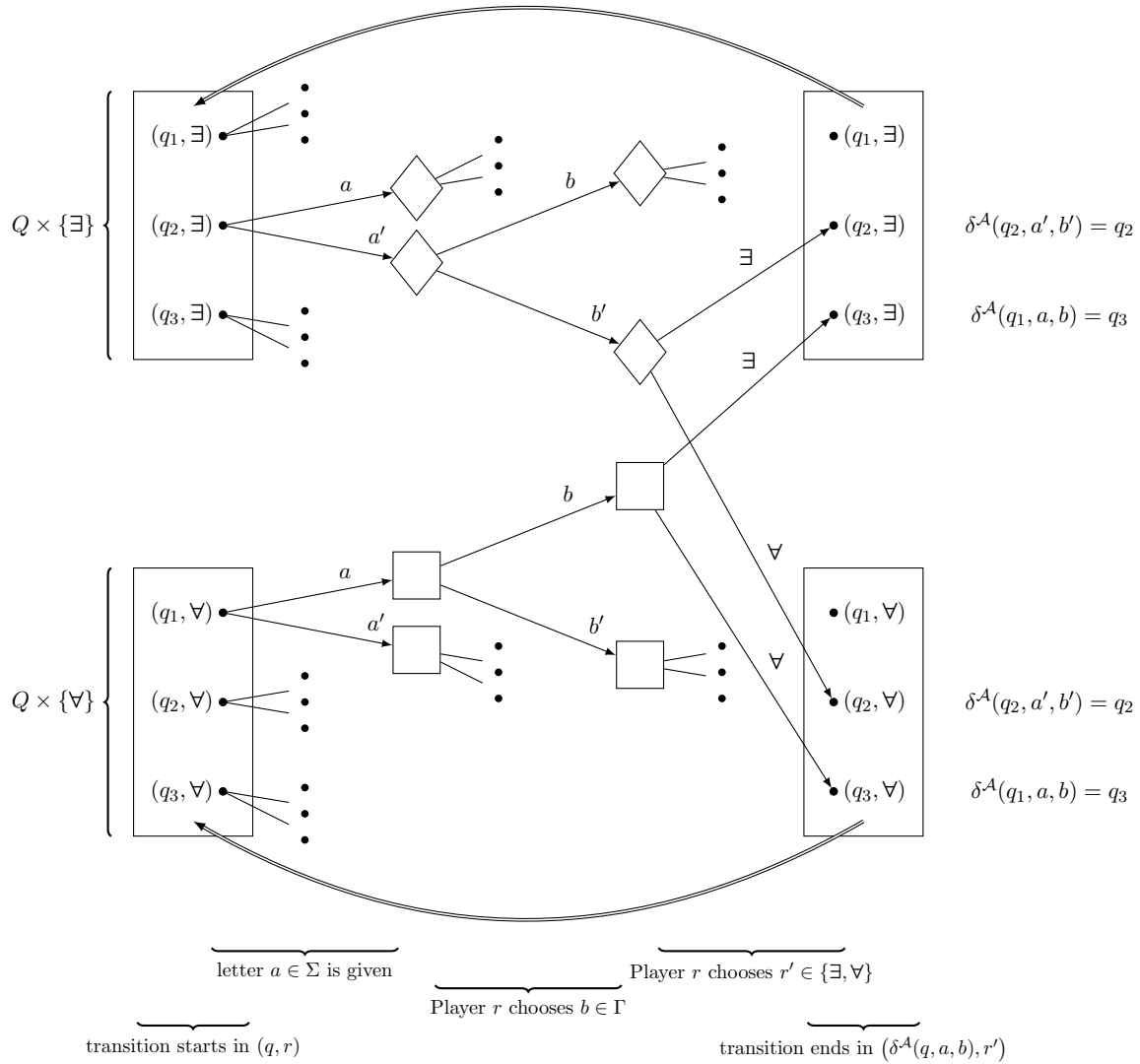
Figure 3: The structure of the automaton $\mathcal{B}$. The transitions are depicted with the convention that choices of $\exists$ (i.e., $\vee$) are represented by diamonds; and the choices of $\forall$ (i.e., $\wedge$) are represented by squares. The states of the automaton are drawn twice, on the left and right edge of the picture for the sake of readability, the double arrows show that we actually go back to the left copy. The phases of each transition (i.e. rounds of the acceptance game) are explained by the braces below the picture.

Lemma 4.4. *The automaton $\mathcal{B}$ accepts an ω-word $\alpha \in \Sigma^\omega$ if and only if the language from (4.1) is comeager.*

Informally, the core of the above lemma is the observation that for a fixed $\alpha \in \Sigma^\omega$ there is a bijection between:

- the plays of the Banach–Mazur game over the language (4.1) (denoted **BM**);
- the plays of the acceptance game of $\mathcal{B}$ over α, in which infinitely many times $r' \neq r$.

To make this relationship more formal, we will show how to simulate plays of **BM** using the acceptance game $\mathcal{B}(\alpha)$. Consider an ω-word $\alpha \in \Sigma^\omega$ and assume that $P \in \{\exists, \forall\}$ wins the acceptance game of $\mathcal{B}$ over α (i.e. $\alpha \in \mathrm{L}(\mathcal{B})$ iff $P = \exists$). We will prove that P wins **BM**.

Fix a strategy σ_P of P in the acceptance game $\mathcal{B}(\alpha)$. We will inductively describe how P can simulate this strategy in the game **BM**. Assume that we are at the beginning of the nth round of **BM** and the current configuration of this game is (s_n, r_n).

Claim 4.5. *During the simulated play of **BM** the following invariant will be preserved: there exists a finite play of $\mathcal{B}(\alpha)$, that is consistent with σ_P and in this play, after reading the first $|s_n|$ symbols of α the reached configuration has the form $\langle |s_n|, (q_n, r_n) \rangle$ for some $q_n \in Q$ (and r_n as above).*

We will now show how to play one round of **BM** while still preserving the invariant from Claim 4.5. Consider the following two possibilities for $r_n \in \{\exists, \forall\}$.

The case $r_n \neq P$. In that case it is the opponent of P (denoted $\bar{P}$) that chooses the successive word s_{n+1} in **BM**. Assume that $\bar{P}$ has played $s_{n+1} = s_n w$ for some non-empty word $w = b_0 b_1 \dots b_k$. Consider the $k+1$ rounds of $\mathcal{B}(\alpha)$ starting from $\langle |s_n|, (q_n, r_n) \rangle$ in which $\bar{P}$ plays the successive letters $b_0, b_1, \dots, b_k$. In the rounds $0, 1, \dots, k-1$ he chooses $r' = r = \bar{P}$ and in the last round he chooses $r' \neq r$ such that $r' = P$. After that $k+1$ rounds the simulated play of $\mathcal{B}(\alpha)$ has reached a configuration of the form $\langle |s_{n+1}|, (q_{n+1}, r_{n+1}) \rangle$ with $r_{n+1} = P$, and the invariant is satisfied.

The case $r_n = P$. Now consider the more involved case where this is Player P who chooses a word s_{n+1}. Consider the successive rounds of the game $\mathcal{B}(\alpha)$ after the round that ended in $\langle |s_n|, (q_n, r_n) \rangle$. Since $r_n = P$ we know that it will be Player P who will propose successive letters $b_0, b_1, \dots$ and players r' in each round. Let $b_0, b_1, \dots$ be the finite or infinite sequence of letters played by P according to σ_P until P chooses to set $r' \neq r = P$ for the first time. Since σ_P is winning, we know that P cannot keep $r' = r = P$ forever, otherwise P would lose the infinite play. Thus, the sequence $b_0, b_1, \dots, b_k$ is finite and after playing b_k Player P has chosen $r' \neq r = P$. Let $s_{n+1} = s_n b_0 b_1 \cdots b_k$ be the response of P in the game **BM**. After that, the configuration of **BM** is $(s_{n+1}, \bar{P})$ and the configuration of $\mathcal{B}(\alpha)$ is of the form $\langle |s_{n+1}|, (q_{n+1}, r_{n+1}) \rangle$ with $r_{n+1} = \bar{P}$, so the invariant is satisfied.

This finishes the proof of Claim 4.5. What remains is to prove that the described simulation strategy is winning for P in **BM**.

Why P wins. We need to prove that the produced word $\beta = \bigcup_{n \in \omega} s_n$ satisfies $(\alpha, \beta) \in \mathrm{L}(\mathcal{A})$ if and only if $P = \exists$. But we know that the strategy σ_P is winning for P, therefore the simulated play of $\mathcal{B}(\alpha)$ must be winning for P. Since we have guaranteed that infinitely many times $r' \neq r$, it means that the sequence of visited states of $\mathcal{A}$ is accepting iff $P = \exists$. Thus, $(\alpha, \beta) \in \mathrm{L}(\mathcal{A})$ iff $P = \exists$, what means that P has won the considered play of **BM**.

This concludes the proof of Lemma 4.4 as well as Theorem 4.1.

5. S1S EXTENDED WITH THE MEASURE QUANTIFIER

In this section we define the syntax and semantics of the logic $S1S$ extended with the *measure quantifier* ($\forall^{=1}$). The syntax of $S1S + \forall^{=1}$ extends that of $S1S$ with the new second-order quantifier $\forall^{=1}$ as follows:

$$\phi \quad ::= \quad x < y \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\,\phi \mid \forall X.\,\phi \mid \forall^{=1} X.\,\phi$$

The semantics of the measure is specified as follows.

$$\mathrm{L}\Big(\forall^{=1} X.\,\phi(X, Y_1, \ldots, Y_n)\Big) = \Big\{w \in \Gamma^\omega \mid \mu_w\big(\{w' \in \{0,1\}^\omega \mid \phi(w', w)\ \text{holds}\}\big) = 1\Big\}$$

where $\Gamma = \{0,1\}^n$ and μ_ω is the Lebesgue measure on $\{0,1\}^\omega$.

In other words, a n-tuple $w = (w_1, \ldots, w_n)$, with $w_i \in \{0,1\}^\omega$, satisfies the formula $\forall^* X.\,\phi(X, Y_1, \ldots, Y_n)$ if for a large (having measure 1) collection of $w' \in \{0,1\}^\omega$, the extended tuple (w', w) satisfies $\phi(X, Y_1, \ldots, Y_n)$.

Once again, informally, w satisfies $\forall^{=1} X.\,\phi(X, \vec{Y})$ if "for almost all" $w' \in \{0,1\}^\omega$, the tuple (w', w) satisfies ϕ. In the logic $S1S + \forall^{=1}$ the "almost all" means for all but a negligible (having measure 0) set.

The main result of this section is that, unlike the case of $S1S + \forall^*$, the theory of $S1S + \forall^{=1}$ is undecidable (Theorem 5.2 below). This result was first proved in [20] by means of a reduction to the emptiness problem of *probabilistic Büchi automata* which is known to be undecidable from [1].

Here we propose a different proof based on the recent result of [7] stating the logic $S1S + \mathsf{U}$, introduced by Bojańczyk in [3], has an undecidable theory. We recall that the syntax of $S1S + \mathsf{U}$ extends that of $S1S$ with the new second-order *unbounding quantifier* ($\mathsf{U}X.\,\phi$) whose semantics is specified as follows:

$$\mathrm{L}\Big(\mathsf{U}X.\,\phi(X, Y_1, \ldots, Y_n)\Big) = \Big\{w \in \Gamma^\omega \mid \forall n.\exists w' \in \{0,1\}^\omega.\big(n < |w'| < \infty \wedge \phi(w', w)\ \text{holds}\big)\Big\}$$

where $\Gamma = \{0,1\}^n$ and $n < |w'| < \infty$ means that w', seen as a subset of $\mathbb{N}$, is finite and has at least n elements.

Theorem 5.1. *For every formula φ of $S1S + \mathsf{U}$ there effectively exists a formula φ' of $S1S + \forall^{=1}$ such that $\mathrm{L}(\varphi) = \mathrm{L}(\varphi')$.*

The above reduction, together with [7, 15], give us the following corollary.

Corollary 5.2. *The logic $S1S + \forall^{=1}$ has an undecidable theory. Furthermore, for each positive number $k \in \omega$ there exists a $S1S + \forall^{=1}$ formula $\phi(X)$ such that $\mathrm{L}(\phi) \subseteq \{0,1\}^\omega$ does not belong to the $\mathbf{\Delta}^1_k$ class of the projective hierarchy.*

The rest of this section is devoted to proving Theorem 5.1. We will assume that using $S1S$ logic we can freely quantify over ω-words over fixed alphabets — this can be simulated by quantifying over tuples of sets representing positions that bear a given letter.

The first step of our proof is standard: we prove that instead of using the U quantifier in $S1S + \mathsf{U}$, one can use a specific predicate.

Definition 5.3. Let $\alpha \in \{0, 1, R\}^\omega$ be an infinite word. We say that α is *unbounded* (denoted $U(\alpha)$) if α contains infinitely many letters R and, for every n there exists a pair of consecutive letters R such that in-between them there are at least n letters 1.

Before moving forward, we will introduce a couple of useful concepts.

Definition 5.4. Fix $\alpha \in \{0, 1, R\}^\omega$ and assume that α contains infinitely many letters R.
(1) A set B of consecutive positions $\{i, i+1, \ldots, j\}$ of α is called a *block* if: α contains no R at positions $\{i, \ldots, j-1\}$; $i = 0$ or $\alpha(i-1) = R$; and $\alpha(j) = R$. In other words, a block is a maximal set of successive positions of α such that the only R that appears among them is the last position in the block.
(2) Since blocks are disjoint, we can identify an arbitrary set of blocks with its union S — a subset of ω.
(3) If B is a block, by the *value of B* we denote the number v of positions of α in B that are labeled by 1. v is always a non-negative integer.

Lemma 5.5. *The logic $S1S+\mathsf{U}$ is effectively expressively equivalent to the logic $S1S$ equipped with the U predicate defined above.*

The above lemma is a standard technical adjustment of the U quantifier, see for instance the notion of *sequential witness* from Lemma 5.5 of [5]. For the sake of completeness, we sketch a proof of it.

Proof. Clearly the U predicate can easily be expressed by the U quantifier. Now consider an application of a U quantifier of the form $\mathsf{U}X.\,\phi(X)$. We claim that the following formula is equivalent to $\mathsf{U}X.\,\phi(X)$

$$\psi \stackrel{\text{def}}{=} \exists \alpha.\, U(\alpha) \wedge \forall B.\, \big(B \text{ is a block of } \alpha\big) \Rightarrow$$
$$\Big[\exists X.\, X \text{ is finite}, \phi(X), \text{ and } \forall x \in B.\, \alpha(x) = 1 \Rightarrow x \in X\Big].$$

Clearly, if ψ holds then $\mathsf{U}X.\,\phi(X)$ holds as well (the positions labeled 1 in the blocks of α witness that). The other direction is also easy, it is enough to construct a witness α. We proceed inductively, at nth step taking a finite set X of cardinality large enough to guarantee that it has at least n elements greater in the order $\leq$ from all the previously considered positions. These n new positions will be labeled 1 in the nth block of α. $\qquad\square$

In the presence of the above lemma, Theorem 5.1 follows directly from the following lemma.

Lemma 5.6. *There exists a formula ψ_U of $S1S + \forall^{=1}$ over the alphabet $\{0, 1, R\}$ such that $\alpha \in \{0, 1, R\}^\omega$ satisfies ψ if and only if $U(\alpha)$ holds.*

Proof. Consider the following formula ψ_U:

$$\psi_U(\alpha) \stackrel{\text{def}}{=} \exists S.\, S \text{ is an infinite set of blocks in } \alpha \wedge$$
$$\neg\big[\forall^{=1}X.\, \exists B.\, B \subseteq S \text{ and } B \text{ is a block in } \alpha \wedge$$
$$\forall x.\, \text{if } x \in B \text{ and } \alpha(x) = 1 \text{ then } x \in X\big].$$

Claim 5.7. *The formula $\psi_U(\alpha)$ holds if and only if $U(\alpha)$ holds.*

Without loss of generality we can restrict to ω-words α containing infinitely many letters R (otherwise both properties are false).

To simplify our notation, let us denote by $\varphi(X, B)$ the property that: for every $x \in B$ such that $\alpha(x) = 1$ we have $x \in X$. Consider a fixed block B of α that has value v and let $X \subseteq \omega$ be a randomly chosen set. Then, the probability that $\varphi(X, B)$ holds is equal to 2^{-v}. Now, if S is an infinite set of blocks of α with values $v_0, v_1, \ldots$, the probability that for a
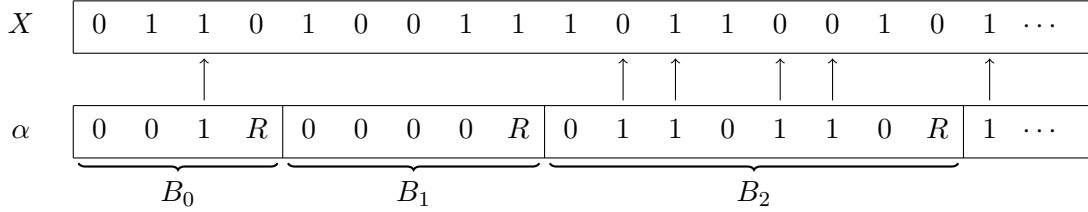
| X | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | $\cdots$ |

| α | 0 | 0 | 1 | R | 0 | 0 | 0 | 0 | R | 0 | 1 | 1 | 0 | 1 | 1 | 0 | R | 1 | $\cdots$ |

$$\underbrace{}_{B_0} \quad \underbrace{}_{B_1} \quad \underbrace{}_{B_2}$$

Figure 4: A word α divided into blocks B_0, B_1, $\ldots$ The values of the blocks are $v_0 = 1$, $v_1 = 0$, $v_2 = 4$, $\ldots$ The formula $\varphi(X, B)$ holds for B_0 and B_1 but not for B_2.

random set $X \subseteq \omega$ some block of S satisfies $\varphi(X, B)$ is equal to

$$1 - \prod_{n \in \omega} \left(1 - 2^{-v_n}\right) \tag{5.1}$$

For an illustration of an ω-word α and its division into blocks, see Figure 4.

We are now in position to prove Claim 5.7. First consider α such that it is not the case that $U(\alpha)$ holds. Therefore, there is a global bound M such that if B is a block in α then the value of B is at most M. Let S be any infinite set of blocks in α as in the formula ψ_U. By (5.1), the probability that some block B of S satisfies $\varphi(X, B)$ equals $1 - \prod_{n \in \omega} \left(1 - 2^{-v_n}\right) \geq 1 - \prod_{n \in \omega} \left(1 - 2^{-M}\right) = 1 - 0 = 1$. Therefore, $\psi_U(\alpha)$ is false.

Now consider the opposite case that $U(\alpha)$ holds. We will prove that $\psi_U(\alpha)$ holds. By the assumption, for every M there exists a block B in α of value greater than M. Let $0 < c_0 < c_1 < \ldots < 1$ be a sequence of numbers such that $\prod_{n \in \omega} c_n > 0$. For each n, there exists a block B_n of α of value v_n such that $1 - 2^{-v_n} > c_n$. Let S be the union of the blocks B_n, clearly S is infinite. Take a random set $X \subseteq \omega$. By (5.1), the probability that some block B of S satisfies $\varphi(X, B)$ equals $1 - \prod_{n \in \omega} \left(1 - 2^{-v_n}\right) \leq 1 - \prod_{n \in \omega} c_n < 1$. Therefore, $\psi_U(\alpha)$ holds.

This concludes the proof of Claim 5.7 and also the proof of Theorem 5.1. $\qquad \square$

6. S2S EXTENDED WITH THE CATEGORY QUANTIFIER

In this section we consider the extension of the monadic second order logic of the full binary tree ($S2S$) with the category quantifier.

The syntax of $S2S + \forall^*$ extends that of $S2S$ with the new second-order quantifier $\forall^*$ as follows:

$$\phi ::= \mathrm{succ}_L(x, y) \mid \mathrm{succ}_R(x, y) \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\, \phi \mid \forall X.\, \phi \mid \forall^* X.\, \phi$$

The semantics of the category quantifier is specified as follows.

$$\mathrm{L}\left(\forall^* X.\, \phi(X, Y_1, \ldots, Y_n)\right) = \left\{ t \in \mathcal{T}_\Gamma \mid \text{ the set } \{t' \in \mathcal{T}_{\{0,1\}} \mid \phi(t', t) \text{ holds}\}) \text{ is comeager} \right\}$$

where $\Gamma = \{0, 1\}^n$.

Our main theorem is the following result, showing that the category quantifier preserves $S2S$-definability when the inner formula can be recognised by a game automaton.

Theorem 6.1. *Let $\mathcal{A}$ be a game automaton over an alphabet $\Sigma \times \Gamma$. Consider the language $\forall^* \mathrm{L}(\mathcal{A})$ that contains a tree $t \in \mathcal{T}_\Sigma$ if and only if the set*

$$\{ t' \in \mathcal{T}_\Gamma \mid (t, t') \in \mathrm{L}(\mathcal{A}) \} \tag{6.1}$$

is comeager.

  Then, $\forall^* \mathrm{L}(\mathcal{A})$ is regular and one can effectively construct an alternating Muller tree automaton $\mathcal{B}$ for this language. Additionally, the size of the automaton $\mathcal{B}$ is polynomial in the size of $\mathcal{A}$.

  It is quite clear that if L is a regular language of trees, then the Banach–Mazur game $\mathbf{BM}(L)$ (see Definition 2.4) cannot be directly encoded in the $S2S$ logic because a strategy of a player in this game is a complex object (at least it assigns prefixes to prefixes).

  Fix a game automaton $\mathcal{A}$ over a product alphabet $\Sigma \times \Gamma$. Let Q be the set of states of $\mathcal{A}$ and $\delta^{\mathcal{A}} \colon Q \times \Sigma \times \Gamma \to \mathcal{DL}(\{\mathtt{L}, \mathtt{R}\} \times Q)$ be the transition function: for a triple $(q, a, b) \in Q \times \Sigma \times \Gamma$ it assigns an expression of the form $(\mathtt{L}, q_\mathtt{L}) \, \chi \, (\mathtt{R}, q_\mathtt{R})$ with χ either $\vee$ (a transition of $\exists$) or $\wedge$ (a transition of $\forall$).

  We will construct an alternating automaton $\mathcal{B}$ over the alphabet Σ that will read a tree $t \in \mathcal{T}_\Sigma$ and verify that the language (6.1) is comeager. The construction follows the same lines as the proof of Proposition 4.3. The additional difficulty here is that we deal with a model of automata that is a bit stronger than deterministic ones. Therefore, we need to take some additional care to resolve the game arising from the semantics of these automata.

  Intuitively, $\mathcal{B}$ will simulate synchronously two games: the Banach–Mazur game over $\forall^* \mathrm{L}(\mathcal{A})$ in which a new tree $t' \in \mathcal{T}_\Gamma$ is constructed; and the acceptance game of $\mathcal{A}$ over the constructed pair of trees (t, t'). This parallel execution of the two games will be visible in the transitions of $\mathcal{B}$. Let the set of states of $\mathcal{B}$ keep track of both, the current state of $\mathcal{A}$ and the player in charge of construction of the tree, i.e., the set of states of $\mathcal{B}$ is $Q \times \{\exists, \forall\}$.

  The initial state of $\mathcal{B}$ is $(q_\mathrm{I}, \forall)$. Consider a state (q, r) of $\mathcal{B}$ and a letter $a \in \Sigma$. The transition $\delta^{\mathcal{B}}(q, r, a)$ is supposed to give the following choices to the players:

- Player r chooses a letter $b \in \Gamma$,
- Player r chooses the next player, $r' \in \{\exists, \forall\}$,
- the owner of the transition $\delta^{\mathcal{A}}(q, a, b) = (\mathtt{L}, q_\mathtt{L}) \, \psi \, (\mathtt{R}, q_\mathtt{R})$ in $\mathcal{A}$ resolves this formula by choosing a direction $d \in \{\mathtt{L}, \mathtt{R}\}$.

After that choices are done, the transition that is taken should be $\langle d, (q_d, r') \rangle$, i.e. we should move in the direction d to the new state (q_d, r'). More formally, we achieve that by putting:

$$\delta^{\mathcal{B}} \langle (q, r), a \rangle \overset{\mathrm{def}}{=} \underset{b \in \Gamma}{\chi} \; \underset{r' \in \{\exists, \forall\}}{\chi} \; \underset{d \in \{\mathtt{L}, \mathtt{R}\}}{\psi} \; \langle d, (q_d, r') \rangle, \tag{6.2}$$

where χ is $\vee$ if $r = \exists$ and $\wedge$ otherwise, and $\delta^{\mathcal{A}} \langle q, (a, b) \rangle = (\mathtt{L}, q_\mathtt{L}) \, \psi \, (\mathtt{R}, q_\mathtt{R})$ (i.e. the last boolean operator in $\delta^{\mathcal{B}} \langle (q, r), a \rangle$ is the same as the one in $\delta^{\mathcal{A}} \langle q, (a, b) \rangle$).

  The acceptance condition of the automaton $\mathcal{B}$ will be the same as in Section 4: a set of states $F \subseteq Q \times \{\exists, \forall\}$ belongs to $\mathcal{F}^{\mathcal{B}}$ if either:

(1) all the states $(q, r) \in F$ satisfy $r = \forall$ (i.e. from some point on $\forall$ has always chosen $r' = r = \forall$,

(2) or F contains both states of the form $(q, \exists)$ and $(q', \forall)$; and additionally the set of states $\{q \in Q \mid (q, r) \in F \text{ for some } r \in \{\exists, \forall\}\}$ is accepting for $\mathcal{A}$ (i.e. belongs to $\mathcal{F}^{\mathcal{A}}$).

In other words, if any of the players from some point on chooses $r' = r$ then he or she looses. Otherwise, the game is won by $\exists$ if the sequence of states visited is accepting in $\mathcal{A}$.

The following lemma proves correctness of our construction and thus concludes the proof of Theorem 6.1. Notice that, except the choice of directions d, the construction is almost identical to the one presented in Section 4.

Lemma 6.2. *The automaton $\mathcal{B}$ accepts a tree $t \in \mathcal{T}_\Sigma$ if and only if the language from (6.1) is comeager.*

Before giving a formal (but technical) proof of this lemma, we will provide an overview. Assume that a player P has a winning strategy σ_P in the acceptance game $\mathcal{B}(t)$ of $\mathcal{B}$ over a tree $t \in \mathcal{T}_\Sigma$. By the definition we know that $t \in L(\mathcal{B})$ iff $P = \exists$. Our aim is to simulate the strategy σ_P as a strategy of P in the Banach–Mazur game over the language (6.1) (we denote this game **BM** for the rest of the proof). This will imply that: $t \in L(\mathcal{B})$ iff $P = \exists$ iff $\exists$ wins **BM** iff the language (6.1) is comeager.

Since the domain of all our trees is the same — $\{\text{L}, \text{R}\}^*$, we can imagine that during a play of **BM**, the players write the respective prefixes $s \colon \{\text{L}, \text{R}\}^* \rightharpoonup \Gamma$ on top of the given tree $t \in \mathcal{T}_\Sigma$. More formally, for $s \colon \{\text{L}, \text{R}\}^* \rightharpoonup \Gamma$ let us define a prefix $t \otimes s \colon \{\text{L}, \text{R}\}^* \rightharpoonup \Sigma \times \Gamma$ as follows: let $\text{dom}(t \otimes s) \stackrel{\text{def}}{=} \text{dom}(s)$ and for $u \in \text{dom}(s)$ let $\big(s \otimes t\big)(u) \stackrel{\text{def}}{=} \big(t(u), s(u)\big)$.

We will use the strategy σ_P to help the player P win **BM** by the following two requirements:

(1) Player P will provide the letters of s_n and the moments to finish the current prefix according to his choices in the transitions of $\mathcal{B}$,

(2) Player P will separately store some information about his choices regarding directions in a data structure called τ.

Then, we will prove that a play of **BM** is won by Player P by noticing that the finally obtained data structure τ is a winning strategy of P in the acceptance game $\mathcal{A}(t, t')$ of the game automaton $\mathcal{A}$ over the product tree (t, t') with $t' = \bigcup_{n \in \omega} s_n$.

The following claim describes the invariants of our construction, see Figure 5 for an illustration.

Claim 6.3. Assume that the current configuration of **BM** is (s_n, r_n). In that case, there must exist: a set $\tau_n \subseteq \{\text{L}, \text{R}\}^*$ of nodes and an assignment π_n that for every node $u \in \tau_n$ gives a finite play $\pi_n(u)$ of the acceptance game of $\mathcal{A}$ over $t \otimes s_n$. The set τ_n and the assignment π_n need to extend the previous ones τ_{n-1} and π_{n-1} and additionally the following invariants are required:

(1) $\epsilon \in \tau_n$; τ_n is prefix-closed; and $\pi_n(\epsilon)$ consists of the initial configuration $\langle \epsilon, (q_\text{I}, \forall) \rangle$ of $\mathcal{A}(t \otimes s_n)$.

(2) If $u \in \tau_n$ and $u \notin s_n$ then no sequence extending u belongs to τ_n (we call such u an *end-point* of τ_n).

(3) If $ud \in \tau_n$ for some u, d, then the play $\pi_n(ud)$ extends the play $\pi_n(u)$ by a round played according to σ_P in which the chosen letter b is $s_n(u)$; the chosen player r' equals the previous player r iff ud is not an end-point for one of $\tau_0, \tau_1, \ldots, \tau_n$; and the chosen direction is d.

(4) Take $u \in \tau_n$. Then the play $\pi_n(u)$ needs to end in a configuration $\langle u, (q, r) \rangle$.

(5) Take $u \in \tau_n$ that is not an end-point of τ_n and assume that the play $\pi_n(u)$ ends in a configuration of the form $\langle u, (q, r) \rangle$. If $\delta^\mathcal{A}\big(q, t(u), s_n(u)\big)$ is a transition of Player $\bar{P}$ then both $u\text{L}$ and $u\text{R}$ belong to τ_n; otherwise exactly one of $u\text{L}$, $u\text{R}$ belongs to τ_n (the choice of d will depend on the strategy σ_P of P).
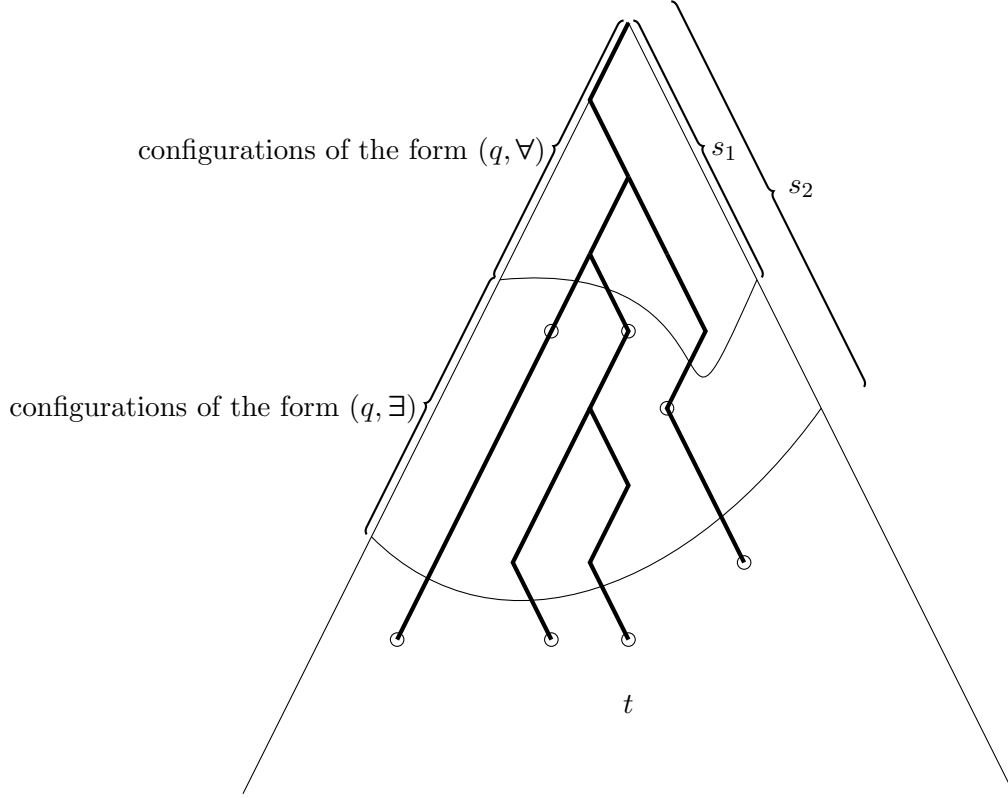
Figure 5: An illustration of our simulation procedure. We are right after the second round
of **BM**. The currently played prefixes are $s_0 = \emptyset$, s_1 and s_2. The boldfaced
subtree of t is our data structure τ_2 (it contains τ_1). The nodes in circles are the
end-points of τ_1 and τ_2. While we are still in s_1, all the configurations visited
in the simulated plays of $\mathcal{A}(t, t')$ are of the form $\langle u, (q, \forall) \rangle$. Then, within s_2 we
ensure that the configurations have the form $\langle u, (q, \exists) \rangle$ etc. The branching in τ_1
and τ_2 occurs exactly in those places where the respective transition of the game
automaton $\mathcal{A}$ is controlled by the opponent of our Player P. When the respective
transition belongs to P, the choice of direction is resolved immediately. In that
case we do not care about the content of the played prefixes in the skipped subtree
(outside the respective τ_n).

Notice that Invariants (3) and (4) guarantee that the last configuration of the play $\pi_n(u)$
for u an end-point of τ_n is of the form $\langle u, (q, r_n) \rangle$ — the position in the tree is u and the
players switch exactly when moving to an end-point of one of $\tau_0, \ldots, \tau_n$.

Initially, for $n = 0$, we have $s_0 = \emptyset$ and all the invariants are satisfied by $\tau_0 = \{\epsilon\}$ and
π_0 that maps ϵ to the initial configuration of the acceptance game of $\mathcal{A}$ (i.e. $\langle \epsilon, (q_I, \forall) \rangle$).

We will now describe how to inductively preserve the invariants from Claim 6.3. Consider
an nth round of the game **BM**. Its initial configuration is (s_n, r_n) with a set τ_n and assignment
π_n. There are two cases depending whether $r_n = P$ or not.

Simulation: the case of $r_n = \bar{P}$. First assume that $r_n \neq P$, i.e. the considered round of
BM is controlled by our opponent. Assume that in this round the player r_n plays a prefix

s_{n+1} that extends s_n. Thus, the round ends in the configuration (s_{n+1}, P). We construct τ_{n+1} and π_{n+1} as follows: start from $\tau_{n+1} := \tau_n$, $\pi_{n+1} := \pi_n$ and inductively extend τ_{n+1} and π_{n+1} for every end-point u of τ_{n+1} that belongs to s_{n+1}. Consider a round of the acceptance game of $\mathcal{A}$ over $t \otimes s_{n+1}$ right after the last configuration of $\pi_{n+1}(u)$. We know that this configuration is of the form $\langle u, (q, \bar{P}) \rangle$.

Assume that in this round $\bar{P}$ chooses as the letter $b = s_{n+1}(u)$ and as the successive player $r' = \bar{P}$ if u is not a leaf of s_{n+1} and $r' = P$ otherwise. Now consider the following cases for the rest of this round:

- If the transition $\delta^{\mathcal{A}}(q, t(u), s_{n+1}(u))$ belongs to $\bar{P}$ then for both $d \in \{\text{L}, \text{R}\}$ we add $ud \in \tau_{n+1}$ and define $\pi_{n+1}(ud)$, assuming that $\bar{P}$ played $d = \text{L}$ and $d = \text{R}$ respectively.
- Otherwise, the strategy σ_P chooses some direction $d \in \{\text{L}, \text{R}\}$. We then add $ud \in \tau_{n+1}$ and define $\pi_{n+1}(ud)$, assuming that P played d. In that case $u\bar{d} \notin \tau_{n+1}$.

Clearly by the definition all the invariants are satisfied in this case.

Simulation: the case of $r_n = P$. Now take the more involved case when $r_n = P$, i.e. the considered round of **BM** is controlled by Player P which strategy σ_P we simulate. We will define s_{n+1}, τ_{n+1}, and π_{n+1} inductively, starting from the end-points of τ_n. For nodes outside the constructed set τ_{n+1} the letters of s_{n+1} are arbitrary, i.e. we can assume that for every ud such that $u \in \tau_{n+1}$ or $u \in s_n$ but $ud \notin \tau_{n+1}$ we let $s_{n+1}(ud) = b_0$ for some fixed letter $b_0 \in \Gamma$. In that case ud will be a leaf of s_{n+1}.

We start from $\tau_{n+1} := \tau_n$, $\pi_{n+1} := \pi_n$. Let u be an end-point of τ_{n+1} such that the last configuration of the play $\pi_{n+1}(u)$ is of the form $\langle u, (q, r) \rangle$. If $r \neq P$ then we finish this branch of construction, letting $u \notin s_{n+1}$ be an end-point of τ_{n+1}. If $r = P$ it means that u will not be an end-point of τ_{n+1}. Consider the successive round after the play $\pi_{n+1}(u)$ in which P plays according to σ_P. First, P chooses a letter $b \in \Gamma$ and a player $r' \in \{\exists, \forall\}$. We can immediately define $s_{n+1}(u) = b$. Now consider the following cases for the rest of this round:

- If the transition $\delta^{\mathcal{A}}(q, t(u), b)$ belongs to $\bar{P}$ then for both $d \in \{\text{L}, \text{R}\}$ we add $ud \in \tau_{n+1}$ and define $\pi_{n+1}(ud)$, assuming that $\bar{P}$ played $d = \text{L}$ and $d = \text{R}$ respectively.
- Otherwise, the strategy σ_P chooses some direction $d \in \{\text{L}, \text{R}\}$. We then add $ud \in \tau_{n+1}$ and define $\pi_{n+1}(ud)$, assuming that P played d. In that case $u\bar{d} \notin \tau_{n+1}$.

If the above inductive procedure ends after finitely many steps, we obtain a finite prefix s_{n+1} together with τ_{n+1} and π_{n+1}. Notice that for every end-point u of τ_{n+1} the last configuration of the play $\pi_{n+1}(u)$ is of the form $\langle u, (q, \bar{P}) \rangle$. Thus, the invariant is satisfied.

Consider the opposite case that the procedure runs indefinitely. In that case, by König's Lemma, there is an infinite path in the constructed set τ_{n+1}. This path corresponds to an infinite play of the acceptance game of $\mathcal{B}$ over t in which Player P keeps $r' = r = P$ constantly equal P from some point on. This contradicts the assumption that P plays according to a winning strategy σ_P, as such a play is losing for Player P.

This way we have managed to play a consecutive round of **BM** while preserving the invariants. Therefore, by induction on n we can construct a strategy of Player P in **BM**.

Why P wins? What remains to prove is that any play of **BM** in which P plays according to this simulation strategy is winning. Take such a play and consider $t' = \bigcup_{n \in \omega} s_n$, $E = \bigcup_{n \in \omega} \tau_n$, $\Pi = \bigcup_{n \in \omega} \pi_n$. It remains to prove that $(t, t') \in \mathrm{L}(\mathcal{A})$ if and only if $P = \exists$. We achieve that by proving that E encodes a winning strategy of P in the acceptance game of $\mathcal{A}$ over (t, t'). Clearly, by the structure of all the sets τ_n, their union E encodes the following strategy of

P: stay in the nodes of (t, t') that belong to E. Take any infinite branch β contained in E. Notice that there is a unique play π of the acceptance game of $\mathcal{B}$ over t that is the limit of the plays $\Pi(u)$ for $u \prec \beta$. Thus, since we considered plays according to a winning strategy σ_P of P, this play π must be winning for P. Clearly, the sequence of players r in this play does not have a limit. Thus, the sequence of visited states q must be either accepting if $P = \exists$ or rejecting otherwise. This concludes the proof that the strategy represented by E is winning for P. Therefore, $(t, t') \in \mathrm{L}(\mathcal{A})$ iff $P = \exists$ and thus P wins **BM**.

7. S2S EXTENDED WITH THE MEASURE QUANTIFIER

In this section we consider the extension of the monadic second order logic of the full binary tree $(S2S)$ with the measure quantifier.

The definitions of the syntax and semantics of $S2S + \forall^{=1}$ are similar to those given for $S1S + \forall^{=1}$. The syntax of $S2S + \forall^{=1}$ extends that of $S2S$ with the new second-order quantifier $\forall^{=1}$ as follows:

$$\phi ::= \mathrm{succ}_L(x, y) \mid \mathrm{succ}_R(x, y) \mid x \in X \mid \neg\phi \mid \phi_1 \vee \phi_2 \mid \forall x.\, \phi \mid \forall X.\, \phi \mid \forall^{=1}X.\, \phi$$

The semantics of the measure quantifier is specified as follows.

$$\mathrm{L}\left(\forall^{=1}X.\, \phi(X, Y_1, \dots, Y_n)\right) = \left\{ t \in \mathcal{T}_\Gamma \mid \mu_t\left(\{t' \in \mathcal{T}_{\{0,1\}} \mid \phi(t', t) \text{ holds}\}\right) = 1 \right\}$$

where $\Gamma = \{0, 1\}^n$ and μ_t is the Lebesgue measure on $\mathcal{T}_{\{0,1\}}$.

Once again, informally, t satisfies $\forall^{=1}X.\, \phi(X, \vec{Y})$ if "for almost all" t', the tuple (t', t) satisfies ϕ, where "almost all" means for all but a negligible (having measure 0) set.

Our main result regarding the logic $S2S + \forall^{=1}$ is the following.

Theorem 7.1. *The logic $S2S + \forall^{=1}$ has an undecidable theory.*

This is obtained as a direct corollary of Theorem 5.1 by interpreting the logic $S1S + \forall^{=1}$ within $S1S + \forall^{=1}$.

Recall that a standard interpretation of $S1S$ within $S2S$ is based on the identification of $(\mathbb{N}, <)$ with the set of vertices $\{\mathtt{L}\}^*$ (belonging to the leftmost branch of the full binary tree) ordered by the prefix relation, which are both easily definable in $S2S$. For example, the $S1S$ formula $\forall X.\phi(X)$ is translated to the $S2S$ formula

$$\forall X.\left(\overline{\phi}(X \cap \{\mathtt{L}\}^*)\right)$$

where $\overline{\phi}(Y)$ is the translation of the simpler formula $\phi(Y)$.

Similarly, we define the translation of $S1S + \forall^{=1}$ formulas of the form $\forall^{=1}X.\, \phi$ as:

$$\forall^{=1}X.\left(\overline{\phi}(X \cap \{\mathtt{L}\}^*)\right)$$

To check that this translation is correct it is sufficient to prove that the function π

$$\pi(X) = X \cap \{\mathtt{L}\}^*$$

mapping subsets of the full binary tree to subsets of the leftmost branch (which can be identified as ω-words over the alphabet $\{0, 1\}$) is Lebesgue measure preserving. That is, we need to show that π is a continuous surjection, and this is obvious, and that for every Borel

set $A \subseteq \{0,1\}^\omega$ it holds that $\mu_t(\pi^{-1}(A)) = \mu_w(A)$. By regularity of the Lebesgue measure it is sufficient to prove that this property holds for arbitrary basic clopen sets A. These are sets of the form $A = U_{\vec{n}=\vec{b}}$, for tuples $\vec{n} = (n_1, \ldots, n_k) \in \mathbb{N}^k$ and $(b_1, \ldots, b_k) \in \{0,1\}^k$ where:

$$U_{\vec{n}=\vec{b}} = \big\{ w \in \{0,1\}^\omega \mid \bigwedge_{i=1}^{k} w(n_i) = b_i \big\}$$

By definition we have that $\mu_w(U_{\vec{n}=\vec{b}}) = \frac{1}{2}^k$.

The preimage $\pi^{-1}(A)$ is the clopen set $B = \{t \in \mathcal{T}_{\{0,1\}} \mid \bigwedge_{i=1}^{k} t(\mathtt{L}_i^n) = b_i\}$ and, by definition of Lebesgue measure on trees, it holds that $\mu_t(B) = (\frac{1}{2})^k$, as desired.

8. S2S EXTENDED WITH THE PATH-CATEGORY QUANTIFIER

As anticipated in the introduction, another interesting way to extend $S2S$ with variants of Friedman's Category and Measure quantifiers is to restrict the quantification to range over infinite branches (paths) of the full binary tree.

In this section we consider the extension of $S2S$ with the category quantifier restricted to path, henceforth denoted by $\forall_\pi^*$.

We first recall the definition of paths in the full binary tree.

Definition 8.1. A set $X \subseteq \{\mathtt{L}, \mathtt{R}\}^*$ of vertices in the full binary tree is a *path* if and only if:
(1) X contains the root ϵ, and
(2) if $v \in X$ and w is a prefix of v then $v \in X$,
(3) if $v \in X$ then either $v\mathtt{L} \in X$ or $v\mathtt{R} \in X$, but not both.
We denote with $\mathcal{P}$ the collection of paths in the full binary tree.

Since every path is uniquely determined by an infinite sequence of directions ($\mathtt{L}$ or $\mathtt{R}$), there is a one-to-one correspondence between $\mathcal{P}$ and the space $\{\mathtt{L}, \mathtt{R}\}^\omega$ which is homeomorphic to the Cantor space.

It is easy to verify that $\mathcal{P} \subseteq \mathcal{T}_{\{0,1\}}$ is Lebesgue null ($\mu_t(\mathcal{P}) = 0$) and it is meager as a subset of $\mathcal{T}_{\{0,1\}}$. However, since $\mathcal{P}$ is homeomorphic to the Cantor space, it makes sense to consider the probability (i.e., Lebesgue measure μ_w) and Baire category of subsets $A \subseteq \mathcal{P} \subseteq \mathcal{T}_{\{0,1\}}$ relative to $\mathcal{P}$.

This leads to the following definition of the Category-path quantifier $\forall_\pi^*$.

$$\mathrm{L}\Big(\forall_\pi^* X. \phi(X, Y_1, \ldots, Y_n)\Big) = \Big\{ t \in \mathcal{T}_\Gamma \mid \text{ the set } \{p \in \mathcal{P} \mid \phi(p, t') \text{ holds}\}) \text{ is comeager in } \mathcal{P} \Big\}$$

where $\Gamma = \{0,1\}^n$.

Our main result regarding the logic $S2S + \forall_\pi^*$ is the following quantifier elimination theorem.

Theorem 8.2. *For every $S2S + \forall_\pi^*$ formula ϕ one can effectively construct an semantically equivalent $S2S$ formula ψ.*

Proof. A set of paths $B \subseteq \{\mathtt{L}, \mathtt{R}\}^\omega$ is comeager if and only if there is a G_δ set G that is dense in $\{\mathtt{L}, \mathtt{R}\}^\omega$ and $G \subseteq B$. A simple argument (see, e.g., [22]) shows that $G \subseteq \{\mathtt{L}, \mathtt{R}\}^\omega$ is a G_δ set if and only if G is of the form $[X]$ for a set $X \subseteq \{\mathtt{L}, \mathtt{R}\}^*$ where:

$$[X] \overset{\mathrm{def}}{=} \{\pi \in \{\mathtt{L}, \mathtt{R}\}^\omega \mid \text{for infinitely many vertices } v \in \pi \text{ we have } v \in X\}.$$

Therefore, $\forall_\pi^* \pi. \varphi(\pi)$ holds if and only if there exists a set of nodes X of the full binary tree such that:

- for every node v there exists a node w such that $v \preceq w$ and $w \in X$ (i.e. $[X]$ is dense in $\{\text{L}, \text{R}\}^\omega$),
- for every infinite branch π, if there are infinitely many $v \in X$ such that $v \prec \pi$ then $\varphi(\pi)$ holds (i.e., $[X] \subseteq \text{L}(\varphi)$)

The whole above property is easily $S2S$-definable. $\square$

9. $S2S$ EXTENDED WITH THE PATH-MEASURE QUANTIFIER

Following the ideas presented in the previous section, we now study the extension of $S2S$ with the path-measure quantifier $\forall_\pi^{=1}$ whose semantics is defined as follows:

$$\text{L}\left(\forall_\pi^{=1} X. \phi(X, Y_1, \ldots, Y_n)\right) = \left\{t \in \mathcal{T}_\Gamma \mid \mu_w\left(\{p \in \mathcal{P} \mid \phi(p, t) \text{ holds}\}\right) = 1\right\}$$

where $\Gamma = \{0, 1\}^n$.

Therefore, the property of $\forall_\pi^{=1} X. \phi$ holds if the property ϕ holds for almost all pahts p, where "for almost all" means having probability 1 with respect to the Lebesgue measure μ_w on $\{\text{L}, \text{R}\}^\omega$ (and thus not with respect to μ_t on $\mathcal{T}_{\{0,1\}}$).

It is not immediately clear from the previous definition if the quantifier $\forall_\pi^{=1}$ can be expressed in $S2S + \forall^{=1}$. Indeed, as already observed, the set $\mathcal{P}$ of paths is Lebesgue null with respect to the Lebesgue measure μ_t on $\mathcal{T}_{\{0,1\}}$, i.e., $\mu_t(\mathcal{P}) = 0$. Therefore the naive definition

$$\forall_\pi^{=1} X. \phi(X) = \forall^{=1} X. \left(\text{``}X \text{ is a path''} \wedge \phi(X)\right)$$

does not work. Indeed the $S2S + \forall^{=1}$ formula on the right always defines the empty set because the collection of $X \in \mathcal{T}_{\{0,1\}}$ satisfying the conjunction is a subset of $\mathcal{P}$ and therefore has μ_t measure 0.

Nevertheless the quantifier $\forall_\pi^{=1}$ can be expressed in $S2S + \forall^{=1}$ with a more elaborate encoding (Theorem 9.3 below). What is needed is a $S2S$ definable continuous and measure preserving function f mapping trees $X \in \mathcal{T}_{\{0,1\}}$ to a paths $f(X) \in \mathcal{P}$. We now prove that such definable function exists.

Definition 9.1. Define the binary relation $f(X, Y)$ on $\mathcal{T}_{\{0,1\}}$ by the following $S2S$ formula:

$$\text{``}Y \text{ is a path''} \wedge \forall y \in Y. \exists z. \left(\text{Succ}_L(y, z) \wedge (z \in Y \Leftrightarrow y \in X)\right).$$

Lemma 9.2. *For every $X \in \mathcal{T}_{\{0,1\}}$ there exists exactly one $Y \in \mathcal{P} \subseteq \mathcal{T}_{\{0,1\}}$ such that $f(X, Y)$. Hence the relation f is a function $f : \mathcal{T}_{\{0,1\}} \to \mathcal{P}$. Furthermore f satisfies the following properties:*

(1) *f is a continuous and surjective function,*
(2) *f is measure preserving, i.e., for every Borel set $B \subseteq \mathcal{P}$ it holds that $\mu_t(f^{-1}(B)) = \mu_w(B)$.*

Proof. The mapping f is well defined in the sense that Y is fully determined by the $\{0, 1\}$-labeled tree X. Indeed, from the condition

$$\forall y \in Y. \exists z. \left(\text{Succ}_L(y, z) \wedge (z \in Y \Leftrightarrow y \in X)\right)$$

in Definition 9.1 it follows that for every vertex $y \in Y$, the set X determines if the unique successor of y in Y is $y\text{L}$ or $y\text{R}$ (since Y is a path, it contains either $y\text{L}$ or $y\text{R}$ and not both)

depending on whether $y \in X$ or $y \notin X$, respectively. Since a finite prefix of Y is determined by a finite subset of X, the mapping f is continuous. Furthermore, the function is surjective, because for every path $Y \in \mathcal{P}$ we can easily find some X such that $f(X) = Y$.

We now show that f is measure preserving, i.e., that $\mu_t(f^{-1})(B) = \mu_w(B)$. Since Lebesgue measures are regular, it is sufficient to prove this property for basic clopen sets B.

Basic clopen sets are of the form $U_v \subseteq \mathcal{P}$, for some $v \in \{\text{L},\text{R}\}^*$, consisting of all infinite paths extending the finite prefix v. We now show that $\mu_t(f^{-1}(U_v)) = \mu_w(U_v)$ by induction of the length $|v| = n$ of the prefix v.

If $n = 0$ then $v = \epsilon$ and $U_v = \mathcal{P}$ and therefore $f^{-1}(U_v) = \mathcal{T}_{\{0,1\}}$. Hence we have $\mu_t(f^{-1}(U_v)) = \mu_w(U_v) = 1$.

If $n > 0$ assume that $v = w\text{L}$ (the case $v = w\text{R}$ is similar) and that the inductive hypothesis holds on U_w. By unfolding the definition of f we get that the preimage $f^{-1}(U_v)$ is the clopen set

$$f^{-1}(U_v) = f^{-1}(U_w) \cap U_{v=0}$$

where $U_{v=0} = \{t \mid t(v) = 0\}$. By definition of Lebesgue measure μ_t it holds that $\mu_t(U_{v=0}) = \frac{1}{2}$. Furthermore, since $f^{-1}(U_w)$ and $U_{v=0}$ are independent events, it holds that

$$\mu_t(f^{-1}(U_w) \cap U_{v=0}) = \mu_t(f^{-1}(U_w)) \cdot \frac{1}{2}$$

By inductive hypothesis on w we have that $\mu_t(f^{-1}(U_w)) = \mu_w(U_w)$. Lastly, since $v = w\text{L}$, we have $\mu_w(U_v) = \mu_w(U_w) \cdot \frac{1}{2}$. Therefore the desired equality

$$\mu_t(f^{-1}(U_v)) = \mu_w(U_v)$$

holds. $\qquad\square$

We are now ready to prove the following theorem.

Theorem 9.3. *For every $S2S + \forall_\pi^{=1}$ formula $\phi(Z_1, \ldots, Z_n)$ there exists a $S2S + \forall^{=1}$ formula $\phi'(Z_1, \ldots, Z_n)$ such that ϕ and ϕ' denote the same set.*

Proof. The proof goes by induction on the complexity of ψ with the interesting case being $\phi(Z_1, \ldots, Z_n) = \forall_\pi^{=1}Y. \psi(Y, Z_1, \ldots, Z_n)$. By induction hypothesis on ψ, there exists a $S2S + \forall^{=1}$ formula ψ' defining the same set as ψ. Then the $S2S + \forall^{=1}$ formula ϕ' corresponding to ϕ is defined as follows:

$$\phi'(Z_1, \ldots, Z_n) = \forall^{=1}X. \left(\exists Y. \left(f(X, Y) \wedge \psi'(Y, Z_1, \ldots, Z_n)\right)\right)$$

We now show that ϕ and ϕ' indeed define the same set. The following are equivalent:

(1) The tuple $(t_1, \ldots, t_n)$ satisfies the formula $\forall_\pi^{=1}Y.\psi(Y, Z_1, \ldots, Z_n)$,
(2) (by the definition of $\forall_\pi^{=1}X$) The set $A = \{t \in \mathcal{P} \mid \psi(t, t_1, \ldots, t_n)\}$ is such that $\mu_w(A) = 1$,
(3) (by Lemma 9.2) The set $B \subseteq \mathcal{T}_{\{0,1\}}$, defined as $B = f^{-1}(A)$, i.e., as

$$B = \{X \in \mathcal{T}_{\{0,1\}} \mid \exists Y. \left(f(X, Y) \wedge \psi(Y, t_1, \ldots, t_n)\right)\}$$

is such that $\mu_{\mathcal{T}_{\{0,1\}}}(B) = 1$.
(4) (by the definition of $\forall^{=1}$ and assumption $\psi = \psi'$) the tuple $(t_1, \ldots, t_n)$ satisfies

$$\forall^{=1}X. \left(\exists Y. \left(f(X, Y) \wedge \phi'(Y, \vec{Z})\right)\right). \qquad\square$$

The result of the previous theorem can be simply stated as $S2S + \forall_\pi^{=1} \subseteq S2S + \forall^{=1}$. The next result states that $S2S + \forall_\pi^{=1}$ is strictly more expressive that $S2S$.

Theorem 9.4. *The strict inequality $S2S \subsetneq S2S + \forall_\pi^{=1}$ holds.*

Before proving this result, as a preliminary step, we introduce the following $S2S + \forall_\pi^{=1}$ definable language of trees.

Definition 9.5. Let $\mathcal{U}^{=1} \subseteq \mathcal{T}_{\{0,1\}}$ be the set of $\{0,1\}$-labeled trees t satisfying the formula $\mathcal{U}^{=1}(X)$ defined by:
$$\mathcal{U}^{=1}(X) = \forall_\pi^{=1} Y.(Y \cap X \text{ is finite}).$$

In other words $\mathcal{U}^{=1}$ is the collection of subsets X of the full binary tree such that the set of paths having only finitely many vertices in X has Lebesgue measure 1 in $\mathcal{P}$.

The following proposition states that every $S2S + \forall_\pi^{=1}$ formula is equivalent to a $S2S$ formula which, additionally, can use the additional predicate over trees $U^{=1}$.

Proposition 9.6. *The equality $S2S + \forall_\pi^{=1} = S2S + \mathcal{U}^{=1}$ holds. That is, every $S2S + \forall_\pi^{=1}$ formula ϕ is effectively equivalent to a $S2S + \mathcal{U}^{=1}$ formula.*

Proof. The proof goes by induction on the structure of $\phi(X_1, \ldots, X_n)$. The only interesting case is with φ of the form $\forall_\pi^{=1} Y.\psi(Y, X_1, \ldots, X_n)$.

We can assume, from the inductive hypothesis on ψ, that $\psi(Y, X_1, \ldots, X_n)$ is expressible in $S2S + \mathcal{U}^{=1}$.

By regularity of measures on Polish spaces a set of paths $B \subseteq \{\text{L}, \text{R}\}^\omega$ has measure 1 if and only if there is an F_σ set $F \subseteq B$ such that F has measure 1. A standard argument (cf. proof of Theorem 8.2, see also [22]) shows, that $F \subseteq \{\text{L}, \text{R}\}^\omega$ is a F_σ set if and only if F is of the form $[Y]$ for a set $Y \subseteq \{\text{L}, \text{R}\}^*$ with

$$[Y] \stackrel{\text{def}}{=} \{\pi \in \{\text{L}, \text{R}\}^\omega \mid \text{there are only finitely many } v \in \pi \text{ such that } v \in Y\}.$$

Therefore, $\forall_\pi^{=1} \pi. \psi(\vec{X})$ holds if and only if there exists a set of vertuces Y such that:
- the set of paths $[Y]$ has measure 1, and
- $[Y] \subseteq \text{L}(\psi(\vec{X}))$.

Notice that the first property is expressible as $\mathcal{U}^{=1}(Y)$ and the second is easily $S2S$-definable. $\square$

We now show that the language $\mathcal{U}^{=1}$ is not definable in $S2S$. The proof is obtained by adapting the argument of Theorem 21 in [8].

Proposition 9.7. *The language $\mathcal{U}^{=1}$ is not $S2S$ definable.*

Proof. A key property of $S2S$ definable sets $A \subseteq \mathcal{T}_{\{0,1\}}$ is that A is not empty if and only if A contains a regular tree, that is a tree having only infinitely many subtrees up-to isomorphism.

To prove that $\mathcal{U}^{=1}$ is not $S2S$ definable we will specify a $S2S$ definable language L and show that $\mathcal{U}^{=1} \cap L$ is not empty and does not contain any regular tree. This of course implies that $\mathcal{U}^{=1}$ is not $S2S$ definable because $S2S$ definable sets are closed under Boolean operations.

We define $L \subseteq \mathcal{T}_{\{0,1\}}$ as the set of trees over the alphabet $\Sigma = \{0, 1\}$ satisfying the $S2S$ formula $\phi(X)$ defined as:

$$\phi(X) = \forall x.\exists y.\big(x \leq y \wedge y \in X\big)\}$$

In other words, a set of vertices $t \in \mathcal{T}_{\{0,1\}}$ of the full binary tree satisfies $\phi(X)$ if from every vertex v there exists a descendant vertex w such that $w \in t$.

We now show that $\mathcal{U}^{=1} \cap L$ is not empty and does not contain a regular tree.

Claim 9.8. $\mathcal{U}^{=1} \cap L$ is not empty.

Proof. We exhibit a concrete tree $t \in \mathcal{T}_{\{0,1\}}$ in $\mathcal{U}^{=1} \cap L$. To do this, fix any mapping $f : \mathbb{N} \to \mathbb{N}$ such that $f(0) = 0$ and for all $n > 0$ holds $f(n) > n + \sum_{i=0}^{n-1} f(i)$. We say that a vertex $v \in \{0,1\}^*$ of the full binary tree belongs to the n-th block if its depth $|v|$ is such that $f(n) \leq |v| < f(n+1)$. Each block can be seen as a forest of finite trees (see Figure 6) of depth $f(n+1) - f(n)$. We now describe the tree t. For each n, all nodes of the n-th block are labeled by 0 except the leftmost vertices of each (finite) tree in the block (seen as a forest), which are labeled by 1. Figure 6 illustrates this idea. Clearly t is in L.

Let E_n be the random event (on the space $\mathcal{P}$ of infinite branches of the full binary tree) of a path having the $f(n+1)$-th vertex labeled by 1. Then, by construction of t, the (Lebesgue measure μ_w) probability of E_n is exactly $\frac{1}{2^{f(n+1)-f(n)}}$.

This implies that $\mu(E_0) + \mu(E_1) + \ldots \leq \sum_{n=0}^{\infty} \frac{1}{2^{f(n+1)-f(n)}} \leq \frac{1}{2} + \ldots + \frac{1}{2^n} \leq 1$. The Borel-Cantelli lemma implies that the probability of infinitely many events E_n happening is 0. Hence the probability of the set of paths having infinitely many 1's is 0. Therefore $t \in \mathcal{U}^{=1}$ and thus $t \in L \cap \mathcal{U}^{=1}$. $\qquad\square$
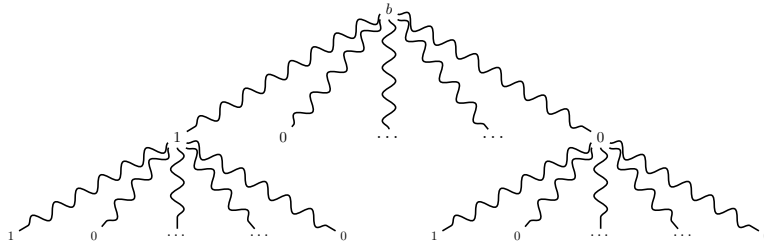

Figure 6: A prefix of a tree $t \in L \cap \mathcal{U}^{=1}$ up to the level $f(2)$.

Claim 9.9. $L \cap \mathcal{U}^{=1}$ does not contain any regular tree t.

Proof. Indeed, let G be the finite graph whose vertices are labeled by 0 or 1 and where each vertex can reach exactly two vertices. Then t represents a regular tree $t \in \mathcal{T}_{\{0,1\}}$. We can view G as a finite Markov chain where all edges have probability $\frac{1}{2}$. From the assumption that $t \in L$, we know that every vertex in G can reach a vertex labeled 1. By elementary results of Markov chains, a random infinite path in G will almost surely visit infinitely many times states labeled by 1 and this is a contradiction with the hypothesis that $t \in \mathcal{U}^{=1}$. $\quad\square$

The proofs of the above two Claims finish the proof of the Proposition. $\qquad\square$

The results of Proposition 9.6 and Proposition 9.6 together prove the claim of Theorem 9.4.

On Qualitative Automata of Carayol, Haddad, and Serre. In a recent paper [8] Carayol, Haddad, and Serre have considered a probabilistic interpretation of standard (Rabin) nondeterministic tree automata. Below we briefly discuss this interpretation referring to [8] (see also [23, Chapter 8]) for more details.

The classical interpretation from [22] of a nondeterministic tree automaton $\mathcal{A}$ over the alphabet Σ is the set $\mathrm{L}(A) \subseteq \mathcal{T}_\Sigma$ of trees $t \in \mathcal{T}_\Sigma$ such that there exists a *run* ρ of t on $\mathcal{A}$ such that for *all* paths π in ρ, the path π is accepting. The probabilistic interpretation in [8] associates to each nondeterministic tree automaton the language $\mathrm{L}^{=1}(\mathcal{A}) \subseteq \mathcal{T}_\Sigma$ of trees $t \in \mathcal{T}_\Sigma$

such that there exists a *run* ρ of X on $\mathcal{A}$ such that for *almost all* paths π in ρ, the path π is accepting, where "almost all" means having Lebesgue measure (μ_w relative to the space $\mathcal{P}$ of paths) equal to 1 .

It is clear that for every automaton $\mathcal{A}$ the language $L^{=1}(\mathcal{A})$ can be defined in the logic $S2S + \forall_\pi^{=1}$. Specifically, if $\Sigma = \{1, \ldots, 2^n\}$, the formula $\psi_\mathcal{A}(X_1, \ldots, X_n)$ defined as

$$\psi_A(\overrightarrow{X}) = \exists \overrightarrow{Y}.\big(\text{``}\overrightarrow{Y} \text{ is a run of } \overrightarrow{X} \text{ on } \mathcal{A}\text{''} \wedge \ \forall_\pi^{=1} Z.(\text{``}Z \text{ is an accepting path of } \overrightarrow{Y}\text{''})\big)$$

defines $L^{=1}(\mathcal{A})$, where the subformulas "$\overrightarrow{Y}$ is a run of $\overrightarrow{X}$ on $\mathcal{A}$" and "$\overrightarrow{Y}$ is a run of $\overrightarrow{X}$ on $\mathcal{A}$" are defined as expected (see, e.g., [26]).

On the other hand, Carayol, Haddad and Serre have shown in [8, Example 7] that there are regular (i.e., $S2S$ definable) sets of trees which are not of the form $L^{=1}(\mathcal{A})$. Therefore there are $S2S + \forall_\pi^{=1}$ languages not definable by automata with probabilistic interpretation.

As we have shown in Proposition 9.6, the (effective) equality $S2S + \forall_\pi^{=1} = S2S + \mathcal{U}^{=1}$ holds. Interestingly, the (not regular) language $\mathcal{U}^{=1}$ is definable by tree automata with probabilistic interpretation of [8]. Hence, $S2S + \forall_\pi^{=1}$ can be understood as the minimal extension of $S2S$ which is sufficiently expressive to define all the languages definable by tree automata with probabilistic interpretation.

A crucial property of languages definable by tree automata with probabilistic interpretation is that if $\mathcal{L}^{=1}(\mathcal{A}) \neq \emptyset$ then $\mathcal{L}^{=1}(\mathcal{A})$ contains a regular tree. As we have shown in the proof of Theorem 9.7 (claim 2), this useful property does not hold for $S2S + \forall_\pi^{=1}$ definable languages.

10. Open Problems

In this concluding section we present a list of problems left open from this work. This is by no means exhaustive and simply reflects the authors' view on what might be the most important, difficult and far-reaching questions encountered in the process of working on these topics.

Problem 1. Is the theory of $S2S + \forall^*$ decidable?
The stronger statement $S2S = S2S + \forall^*$ was formulated as a theorem in [18] but, as discussed in the Introduction and Section 6), the proof contained a mistake. The proof is correct only for languages defined by game automata.

Problem 2. It is well known that $S2S$ definable sets always belong to the $\mathbf{\Delta}_2^1$ class of the projective hierarchy. Is this also true for $S2S + \forall^*$ definable sets?

Problem 3. In terms of reverse mathematics as shown in [17], decidability of the theory $S2S$ is equivalent to determinacy of all Gale–Stewart games with the winning condition being a Boolean combination of F_σ sets (see [17] for a precise formulation). Is this characterization also valid for $S2S + \forall^*$?

Problem 4. Is the theory $S2S + \forall_\pi^{=1}$ decidable?

Problem 5. Are $S2S + \forall_\pi^{=1}$ definable sets always contained in the $\mathbf{\Delta}_2^1$ class of the projective hierarchy. All the examples of $S2S + \forall_\pi^{=1}$ definable sets which are not $S2S$ definable we have considered (e.g., the language $\mathcal{U}^{=1}$) are contained in the $\mathbf{\Delta}_2^1$ class.

Problem 6. Design an algorithm which for a given $S2S$ definable set computes its measure.

References

[1] C. Baier, M. Größer, and N. Bertrand. Probabilistic ω-automata. *Journal of the ACM*, 59(1):1:1–1:52, 2012.

[2] V. Bárány, Ł. Kaiser, and A. Rabinovich. Cardinality quantifiers in MLO over trees. In *CSL*, pages 117–131, 2009.

[3] M. Bojańczyk. A bounding quantifier. In *CSL*, pages 41–55, 2004.

[4] M. Bojańczyk. Thin MSO with a probabilistic path quantifier. In *ICALP (2)*, pages 96:1–96:13, 2016.

[5] M. Bojańczyk and T. Colcombet. Bounds in ω-regularity. In *LICS*, pages 285–296, 2006.

[6] M. Bojańczyk, H. Gimbert, and E. Kelmendi. Emptiness of zero automata is decidable. In *ICALP (2)*, pages 106:1–106:13, 2017.

[7] M. Bojańczyk, P. Parys, and S. Toruńczyk. The MSO+U theory of (N, <) is undecidable. In *STACS*, pages 21:1–21:8, 2016.

[8] A. Carayol, A. Haddad, and O. Serre. Randomization in automata on infinite trees. *ACM Transactions on Computational Logic*, 15(3):24:1–24:33, 2014.

[9] A. Carayol and O. Serre. How good is a strategy in a game with nature? In *LICS*, pages 609–620, 2015.

[10] A. Carayol and O. Serre. Counting branches in trees using games. *Information and Computation*, 252:221–242, 2017.

[11] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *SODA*, pages 121–130, 2004.

[12] T. Gogacz, H. Michalewski, M. Mio, and Michał Skrzypczak. Measure properties of regular sets of trees. *Information and Computation*, 256:108–130, 2017.

[13] E. Grädel. Banach-Mazur games on graphs. In *FSTTCS*, pages 364–382, 2008.

[14] G. Hjorth, B. Khoussainov, A. Montalbán, and A. Nies. From automatic structures to Borel structures. In *LICS*, pages 431–441, 2008.

[15] S. Hummel and M. Skrzypczak. The topological complexity of MSO+U and related automata models. *Fundamenta Informaticae*, 119(1):87–111, 2012.

[16] A. S. Kechris. *Classical Descriptive Set Theory*. Springer-Verlag, New York, 1994.

[17] L. A. Kołodziejczyk and H. Michalewski. How unprovable is Rabin's decidability theorem? In *LICS*, pages 788–797, 2016.

[18] H. Michalewski and M. Mio. Baire category quantifier in monadic second order logic. In *ICALP (2)*, pages 362–374, 2015.

[19] H. Michalewski and M. Mio. On the problem of computing the probability of regular sets of trees. In *FSTTCS*, pages 489–502, 2015.

[20] H. Michalewski and M. Mio. Measure quantifier in monadic second order logic. In *LFCS*, pages 267–282, 2016.

[21] A. Mostowski. On a generalization of quantifiers. *Fundamenta Mathematicae*, 44:12–36, 1957.

[22] M. O. Rabin. Decidability of second-order theories and automata on infinite trees. *Transactions of the AMS*, 141:1–35, 1969.

[23] O. Serre. *Playing with Trees and Logic*. PhD thesis, Université Paris Diderot (Paris 7), 2015.

[24] L. Staiger. Rich omega-words and monadic second-order arithmetic. In *CSL*, pages 478–490, 1997.

[25] C. I. Steinhorn. *Borel Structures and Measure and Category Logics*, volume 8 of *Perspectives in Mathematical Logic*, chapter XVI, pages 579–596. Springer-Verlag, 1985.

[26] W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, pages 389–455. Springer, 1996.

[27] H. Völzer and D. Varacca. Defining fairness in reactive and concurrent systems. *Journal of the ACM*, 59(3):13:1–13:37, 2012.