A characterization theorem for the alternation-free fragment of the modal μ-calculus

Alessandro Facchini
U. Warsaw
Email: facchini@mimuw.edu.pl

Yde Venema
U. Amsterdam
Email: y.venema@uva.nl

Fabio Zanasi
ENS Lyon, U. Lyon, CNRS, INRIA, UCBL
Email: fabio.zanasi@ens-lyon.fr

***Abstract*—We provide a characterization theorem, in the style of van Benthem and Janin-Walukiewicz, for the alternation-free fragment of the modal μ-calculus. For this purpose we introduce a variant of standard monadic second-order logic (MSO), which we call well-founded monadic second-order logic ($WFMSO$). When interpreted in a tree model, the second-order quantifiers of $WFMSO$ range over subsets of well-founded subtrees. The first main result of the paper states that the expressive power of $WFMSO$ over trees exactly corresponds to that of weak MSO-automata. Using this automata-theoretic characterization, we then show that, over the class of all transition structures, the bisimulation-invariant fragment of $WFMSO$ is the alternation-free fragment of the modal μ-calculus. As a corollary, we find that the logics $WFMSO$ and $WMSO$ (weak monadic second-order logic, where second-order quantification concerns finite subsets), are incomparable in expressive power.**

I. Introduction

A seminal result in the theory of modal logic is van Benthem's Bisimulation Theorem [1], stating that every first-order formula $\alpha(x)$ which is invariant under bisimulations, is actually equivalent to (the standard translation of) a modal formula. Concisely, this result can be formulated as follows:

$$FO/{\leftrightarroweq} = ML. \tag{1}$$

Over the years, a wealth of variants of the Bisimulation Theorem have been obtained. For instance, Rosen proved that van Benthem's theorem is one of the few preservation results that transfers to the setting of finite models [2]; for a recent, rich source of van Benthem-style characterization results, see Dawar & Otto [3].

Of particular interest to us is the work of Janin & Walukiewicz [4], who extended van Benthem's result to the setting of fixpoint logics, by proving that the modal μ-calculus (MC) is the bisimulation-invariant fragment of monadic second-order logic (MSO):

$$MSO/{\leftrightarroweq} = MC. \tag{2}$$

The general pattern of these results takes the following shape:

$$\mathcal{L}/{\leftrightarroweq} = \mathcal{M} \quad \text{(over } \mathcal{C}\text{)}, \tag{3}$$

stating that $\mathcal{M}$ is the bisimulation-invariant fragment of $\mathcal{L}$ over a class $\mathcal{C}$ of models. Here $\mathcal{L}$ is some yardstick logic such as first-order logic (FO), monadic second-order logic (MSO) or weak monadic second-order logic ($WMSO$); $\mathcal{M}$ is some variant of modal logic such as the modal μ-calculus or one of its most important fragments: the alternation-free fragment ($AFMC$); and $\mathcal{C}$ is some class of models, such as finite, transitive, or tree models. Table I summarizes some important results following this pattern.

Structures ($\mathcal{C}$)	$\mathcal{L}$	$\mathcal{M}$	reference
TSs	FO	ML	[1]
	MSO	MC	[4]
	WMSO	?	–
	?	AFMC	–
binary trees	WMSO	AFMC	[5]
finite TSs	FO	ML	[2]
	WMSO = MSO	?	–
transitive TSs	WMSO	ML	[6]
	MSO	AFMC	[3], [7]

Table I
TS stands for 'transition system'.

Table I suggestively indicates the existence of some open problems, but let us first address the issue *why* characterization results of the form (3) are of interest, apart from their obvious importance in (finite) model theory. The point is that the mentioned logics, and the models they are interpreted in, feature prominently in the area of formal verification theory. Generally, one is interested in applications where these models are transition structures representing certain computational processes, and one usually takes the point of view that bisimilar models represent the *same* process. For this reason, properties of transition structures that are not bisimulation-invariant are simply irrelevant. Seen in this light, (3) is an *expressive completeness* result, stating that all relevant properties of $\mathcal{L}$ (which is generally some kind of expressive yardstick formalism), can already be expressed in a (usually computationally more feasible) fragment $\mathcal{M}$. Or, conversely starting from $\mathcal{M}$, according to (3), one may think of $\mathcal{L}$ as an extension of $\mathcal{M}$ that is completely covered by $\mathcal{M}$ when it comes to expressing relevant properties.

In this paper, which is based on an MSc thesis [8] written by the third author and supervised by the first two authors, we try to improve our grasp of such 'expressiveness modulo bisimilarity' results. We are particularly interested in the relation between (variants of) monadic second-order logic and modal fixpoint logics; that is, in variants of the Janin-Walukiewicz result (2). More concretely, we fill in one of the three gaps of Table I by providing a natural solution $\mathcal{L}$ to the equation

$$\mathcal{L}/{\leftrightarrows} = AFMC \quad \text{(over all TSs)}. \tag{4}$$

Naively, one might think that when considering this question in the context of all transition systems, the situation is the same as for the class of binary trees [5], so that $\mathcal{L} = WMSO$ would solve (4). But if this were the case, then $AFMC$ would also be the bisimulation-invariant fragment of $WMSO$ over trees. However, it turns out that the class of well-founded trees, which is definable in $AFMC$ by the formula $\mu x.\Box x$, is not $WMSO$-definable. The reason comes from topology: the class of well-founded trees is not Borel, whereas all $WMSO$-definable tree languages are Borel. On the other hand, it is not clear either whether the bisimulation-invariant fragment of $WMSO$ is included in $AFMC$ (or even in the modal μ-calculus itself), no matter how reasonable this may seem. The point is that, contrary to the finitely branching case, $WMSO$ is not a fragment of MSO over trees of arbitrary branching degree. (In fact the two logics are incomparable, as a consequence of the following *finite branching property* of MSO: every non-empty MSO definable tree language contains at least a finitely branching tree [9]. It follows that MSO cannot define the class of infinitely branching trees, which on the contrary is clearly $WMSO$-definable.)

It turns out that in order to solve the equation (4), we need to introduce a *new* variant of monadic second-order logic. In this variant, that we shall call *well-founded* monadic second-order logic ($WFMSO$), the second-order quantifiers range over special subsets of the transitions system that we call *noetherian*. Roughly, a subset S of a transition system $\mathbb{T}$ is noetherian if there are no infinite paths in the past of S (a more precise definition follows). Note that a subset S in a tree model $\mathbb{T}$ is noetherian iff S is a subset of a (conversely) well-founded subtree of $\mathbb{T}$ — this explains our terminology.

Theorem 1. *Let $\mathcal{L}$ a bisimulation closed class of transition systems. The following are equivalent.*

1) *$\mathcal{L}$ is $AFMC$-definable.*
2) *$\mathcal{L}$ is $WFMSO$-definable.*

As in the case of MSO and the modal μ-calculus, the result is obtained by using automata-theoretic techniques. We will work with Walukiewicz' MSO-automata [9], or more specifically, with variants in which acceptance is defined in terms of a weak parity or a Büchi condition. Restricting, as usual, to tree models, we will prove the following result.

Theorem 2. *Let $\mathcal{L}$ a tree language. The following are equivalent:*

1) *$\mathcal{L}$ is recognized by a weak MSO-automaton.*
2) *$\mathcal{L}$ is $WFMSO$-definable.*
3) *$\mathcal{L}$ and its complement $\overline{\mathcal{L}}$ are both recognized by a non-deterministic Büchi MSO-automaton.*

Both our results are generalizations to structures of arbitrary branching degree, of results known for binary trees (see Arnold & Niwiński [5] for Theorem 1, and Rabin [10] or Muller, Saoudi & Schupp [11] for Theorem 2). This should come as no surprise once we realize that $WMSO$ and $WFMSO$ are the *same* logic on finitely branching trees, and therefore a fortiori on binary trees. The key observation here is that by König's Lemma, over the class of finitely branching trees, the noetherian subsets *coincide* with the finite ones. Intuitively, the idea behind noetherian sets is that they somehow bound the set's 'vertical' dimension, whereas the branching degree concerns the 'horizontal' dimension. Perhaps this separation of dimensions can be seen as a conceptual contribution of our paper, which hopefully will further increase our understanding of the interaction between monadic second-order logics, modal fixpoint logics, and automata, both on trees and on arbitrary models.

Finally, we address the question of the relative expressive power of the logic $WFMSO$ with respect to MSO and $WMSO$. It is not hard to see that MSO has more expressive power than $WFMSO$. It then follows from Theorem 1 that on the class of all transition structures and on the class of all trees, MSO is strictly more expressive than $WFMSO$, while the logics $WMSO$ and $WFMSO$ are incomparable. This provides further evidence that a complete understanding of $WMSO$-expressivity on arbitrary trees requires a different, non-obvious variant of MSO-automata.

Due to space limitations some proofs are moved to the Appendix.

II. Preliminaries

A. Transition Systems and Trees.

Throughout this article we fix a set P of elements that will be called *proposition letters* and denoted with small Latin letters $p, q, \ldots$. We denote with C the set $\wp(P)$ of *labels* on P; it will be convenient to think of C as an *alphabet*. Given a binary relation $R \subseteq X \times Y$, for any element $x \in X$, we indicate with $R[x]$ the set $\{y \in Y \mid (x, y) \in R\}$, while R^+ and $R^\star$ are defined respectively as the transitive closure of R and the reflexive and transitive closure of R. The set $Ran(R)$ is defined as $\bigcup_{x \in X} R[x]$.

A *C-transition system* is a tuple $\mathbb{T} = \langle T, s_I, R, V\rangle$ where $\langle T, R\rangle$ is a directed graph, $s_I \in T$ is a distinguished node, and $V : T \to C$ is a labeling function. Let p be a proposition letter (not necessarily in P). A *p-variant* of a transition system $\mathbb{T} = \langle T, s_I, R, V\rangle$ is a $\wp(P \cup \{p\})$-transition system $\mathbb{T}' = \langle T, s_I, R, V'\rangle$ such that $V'(s) \setminus \{p\} = V(s)$ for all $s \in T$. Given a set $S \in \wp(T)$, we let $\mathbb{T}[p \mapsto S]$ denote the p-variant $\langle T, s_I, R, V'\rangle$ of $\mathbb{T}$ where $V'(s)$ is defined as $V(s) \cup \{p\}$ if s is in S and $V'(s) = V(s)$ otherwise. A *path* through $\mathbb{T}$ is a sequence $\pi = (u_i)_{i<\alpha}$ of elements of T, where α is either ω or a natural number, and $(u_i, u_{i+1}) \in R$ for all i with $i + 1 < \alpha$.

A *C-tree* is a C-transition system $\mathbb{T} = \langle T, s_I, R, V\rangle$ where $\langle T, R\rangle$ is a graph in which every node can be reached from s_I (that is, $R^\star[s_I] = T$), and every node, except s_I, has a unique predecessor; s_I is called the *root* of $\mathbb{T}$. A *branch* of $\mathbb{T}$ is a maximal path through $\mathbb{T}$ starting at the root; we may identify a branch with the *set* of its nodes. Each node $s \in T$ uniquely defines a subtree of $\mathbb{T}$ with carrier $R^\star[s]$ and root s, that we denote with $\mathbb{T}.s$. A tree language over P (or just a tree language if P is clear from the context) is just a class of C-trees.

Given a tree $\mathbb{T}$, we say that $G \subseteq T$ is a *frontier* of $\mathbb{T}$ if $G \cap E$ is a singleton for every branch E of $\mathbb{T}$. A set S is a *prefix* of $\mathbb{T}$ if there exists a frontier G of $\mathbb{T}$ such that $S = \{s \in T \mid sR^{\star}t \text{ for some } t \in G\}$. It is easy to see that every prefix is uniquely determined by a frontier and vice versa; if S is a prefix, we denote with $Ft(S)$ its associated frontier. It is similarly straightforward to verify that the set of prefixes is in 1-1 correspondence with the collections of well-founded subtrees of $\mathbb{T}$ that have the same root as $\mathbb{T}$. Given two frontiers G_1 and G_2 of $\mathbb{T}$, we write $G_1 < G_2$ if, for every branch E in $\mathbb{T}$, given $s_1 \in G_1 \cap E$ and $s_2 \in G_2 \cap E$, we have that $s_1 R^+ s_2$. Analogously, $G_1 \leq G_2$ holds if, for every branch E in $\mathbb{T}$, given $s_1 \in G_1 \cap E$ and $s_2 \in G_2 \cap E$, we have that $s_1 R^{\star} s_2$.

Given a transition system $\mathbb{T}$ and a subset S of T, let $\Uparrow S$ denote the set of points from which S can be reached by a finite path. More precisely, $\Uparrow S := \{t \in T \mid R^{\star}[t] \cap S \neq \emptyset\}$. Call S *noetherian* if there is no infinite path through $\Uparrow S$ (that is, no sequence $(u_i)_{i<\omega}$ such that $u_i \in \Uparrow S$ and $u_i R u_{i+1}$, for all i). It is straightforward to verify that in the case of tree models, a subset S is noetherian iff it is a subset of a prefix of the tree iff it is a subset of a well-founded subtree. We let $N(\mathbb{T})$ denote the collection of noetherian subsets of $\mathbb{T}$. A p-variant $\mathbb{T}[p \mapsto S]$ of $\mathbb{T}$ is *noetherian* if $S \in N(T)$; similarly, $\mathbb{T}[p \mapsto S]$ is a *finite* p-variant if $S \subseteq_\omega T$.

Unless explicitly specified otherwise, all transition systems $\mathbb{T}$ are considered to be C-labeled.

Convention. Throughout this paper, we will only consider transition system $\mathbb{T}$ in which $R[s]$ is non-empty, for every node $s \in T$. In particular this means that every tree we consider is *leafless*. All our results, however, can easily be lifted to the general case.

B. Board Games.

We introduce some terminology and background on infinite games. All the games that we consider involve two players called *Éloise* ($\exists$) and *Abelard* ($\forall$). In some contexts we refer to player Π, meaning that we want to specify a notion for a generic player in $\{\exists, \forall\}$.

Given a set A, by A^* and A^ω we denote respectively the set of words (finite sequences) and streams (or infinite words) over A.

A *board game* $\mathcal{G}$ is a tuple $(G_\exists, G_\forall, E, Win)$, where $G_\exists$ and $G_\forall$ are disjoint sets whose union $G = G_\exists \cup G_\forall$ is called the *board* of $\mathcal{G}$, $E \subseteq G \times G$ is a binary relation encoding the *admissible moves*, and $Win \subseteq G^\omega$ is a *winning condition*. An *initialized board game* $\mathcal{G}@u_I$ is a tuple $(G_\exists, G_\forall, u_I, E, Win)$ where $(G_\exists, G_\forall, E, Win)$ is a board game and $u_I \in G$ is the *initial position* of the game. When $\mathcal{G}$ is a parity game, i.e. Win is given by a parity function $\Omega : G \to \omega$, we sometimes write $\mathcal{G} = (G_\exists, G_\forall, E, \Omega)$.

Given a board game $\mathcal{G}$, a *match* of $\mathcal{G}$ is simply a path through the graph (G, E); a match of $\mathcal{G}@u_I$ is supposed to start at u_I. For a match $\pi = (u_i)_{i<k}$ for some finite $k < \omega$, we call $last(\pi) = u_{k-1}$ the *last position* of the match; the player Π such that $last(\pi) \in G_\Pi$ is supposed to move at this position, and if $E[last(\pi)] = \emptyset$, we say that Π *gets stuck* in π. A match π is called *total* if it is either finite, with one of the two players getting stuck, or infinite. Matches that are not total are called *partial*. Any total match π is *won* by one of the players: If π is finite, then it is won by the opponent of the player who gets stuck. Otherwise, if π is infinite, the winner is $\exists$ if $\pi \in Win$, and $\forall$ if $\pi \notin Win$.

Given a board game $\mathcal{G}$ and a player Π, let PM^G_Π denote the set of partial matches of $\mathcal{G}$ whose last position belongs to player Π. A *strategy for* Π is a function f of type $PM^G_\Pi \to G$. A match $\pi = (u_i)_{i<\alpha}$ of $\mathcal{G}$ is *f-conform* if for each $i < \alpha$ such that $u_i \in G_\Pi$ we have that $u_{i+1} = f(u_0, \ldots, u_i)$.

Given a position $u \in G$ and a strategy $f : PM^G_\Pi \to G$, consider the following two conditions.

1) For each f-conform partial match π of $\mathcal{G}@u$, if $last(\pi)$ is in G_Π then $f(\pi)$ is legitimate, i.e., $(last(\pi), f(\pi)) \in E$.
2) Π wins each f-conform total match of $\mathcal{G}@u$.

If f respects the first condition, we say that f is a *surviving strategy* for Π in $\mathcal{G}@u$, and if it satisfies both conditions, then we call f a *winning strategy* for Π in $\mathcal{G}@u$. In the latter case we say that u is a *winning position* for Π in $\mathcal{G}$. We denote with $Win_\Pi(\mathcal{G})$ the set of positions of $\mathcal{G}$ that are winning for Π. A strategy $f : PM^G_\Pi \to G$ is called *positional* if $f(\pi) = f(\pi')$ for each π and π' in $Dom(f)$ with $last(\pi) = last(\pi')$. A board game $\mathcal{G}$ with board G is *determined* if $G = Win_\exists(\mathcal{G}) \cup Win_\forall(\mathcal{G})$, that is, each $u \in G$ is a winning position for one of the two players.

Fact 1 (Positional Determinacy of Parity Games, [12], [13])**.** *For each parity game $\mathcal{G}$, there are positional strategies $f_\exists$ and $f_\forall$ respectively for player $\exists$ and $\forall$, such that for every position $u \in G$ there is a player Π such that f_Π is a winning strategy for Π in $\mathcal{G}@u$.*

From now on, we always assume that each strategy we work with in parity games is positional.

C. Monadic Second-Order Logics.

We define three variants of monadic second-order logic: *(standard) monadic second-order logic* (MSO_P), *weak monadic second-order logic* ($WMSO_P$) and *well-founded monadic second-order logic* ($WFMSO_P$). We omit the subscript P when the set of proposition letters is clear from the context. These logics share the same syntax: formulas of the *monadic second-order language* on P are defined by the following grammar:

$$\varphi \quad ::= \quad p \sqsubseteq q \mid S(p) \mid R(p,q) \mid \neg\varphi \mid \varphi \vee \varphi \mid \exists p.\varphi,$$

where p and q are letters from P. We adopt the standard convention that no letter is both free and bound in φ.

The three logics are distinguished by their semantics. Given a transition system $\mathbb{T} = \langle T, s_I, R, V \rangle$, the interpretation of the atomic formulas and the boolean connectives is fixed and standard, e.g.:

$$\begin{array}{lll} \mathbb{T} \models p \sqsubseteq q & \text{iff} & \forall s \in T. p \in V(s) \Rightarrow q \in V(s) \\ \mathbb{T} \models R(p,q) & \text{iff} & \forall s \in T. p \in V(s) \Rightarrow \exists t \in R[s]. q \in V(t) \\ \mathbb{T} \models S(p) & \text{iff} & \forall s \in T. p \in V(s) \Rightarrow s = s_I. \end{array}$$

The interpretation of the existential quantifier is that $\mathbb{T} \models \exists p.\varphi$ if and only if

(MSO)	$\mathbb{T}[p \mapsto S] \models \phi$ for some $S \subseteq T$
(WMSO)	$\mathbb{T}[p \mapsto S] \models \phi$ for some finite $S \subseteq T$
(WFMSO)	$\mathbb{T}[p \mapsto S] \models \phi$ for some noetherian $S \subseteq T$.

Let $\varphi \in MSO$ be a formula. We denote with $\|\varphi\|_P$ the set of C-transition structures $\mathbb{T}$ such that $\mathbb{T} \models \varphi$. The subscript P is omitted when the set P of proposition letters is clear from the context. A class $\mathcal{L}$ of transition systems is *MSO-definable* if there is a formula $\varphi \in MSO$ such that $\|\varphi\| = \mathcal{L}$. We define the analogous notions for $WMSO$ and $WFMSO$ in the same way.

D. The Modal μ-Calculus.

The language of the modal μ-calculus (MC) on P is given by the following grammar:

$$\varphi \ ::= \ p \mid \neg p \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \Diamond\varphi \mid \Box\varphi \mid \mu q.\varphi \mid \nu q.\varphi,$$

where $p, q \in P$, and in the clauses for $\mu q.\varphi$ and $\nu q.\varphi$, q does not occur in the scope of $\neg$.

The semantics of this language is completely standard. Let $\mathbb{T} = \langle T, s_I, V, R\rangle$. We inductively define the *meaning* $\|\varphi\|^{\mathbb{T}}$ which includes the following clauses for the least (μ) and greatest (ν) fixpoint operators:

$$\begin{array}{lll} \|\mu p.\psi\|^{\mathbb{T}} & := & \bigcap\{S \subseteq T \mid S \supseteq \|\psi\|^{\mathbb{T}[p\mapsto S]}\} \\ \|\nu p.\psi\|^{\mathbb{T}} & := & \bigcup\{S \subseteq T \mid S \subseteq \|\psi\|^{\mathbb{T}[p\mapsto S]}\} \end{array}$$

We say that φ is *true* in $\mathbb{T}$ - notation $\mathbb{T} \models \varphi$ - if $s_I \in \|\varphi\|^{\mathbb{T}}$. As for the case of MSO, $\|\varphi\|_P$ denotes the class of C-transition systems $\mathbb{T}$ where φ is true.

Formulae of the modal μ-calculus are classified according to their *alternation depth*, which roughly is given as the maximal length of a chain of nested alternating least and greatest fixpoint operators [14]. In particular, we are interested in the *alternation-free fragment* of the modal μ-calculus ($AFMC$) which is the collection of MC-formulae without nesting of least and greatest fixpoint operators. It is well known that over transition systems there is a MC-formula φ such that $\|\varphi\|_P$ is not $AFMC$-definable [15].

E. Bisimulation.

Bisimulation is a notion of behavioral equivalence between processes. For the case of transition systems, it is formally defined as follows.

Definition 1. *Given C-transition systems $\mathbb{T} = \langle T, s_I, R, V\rangle$ and $\mathbb{T}' = \langle T', s'_I, R', V'\rangle$, a* bisimulation *is a relation $Z \subseteq T \times T'$ such that for all $(t, t') \in Z$ the following holds:*

- *$V(t) = V'(t')$;*
- *for all $s \in R[t]$ there is $s' \in R'[t']$ such that $(s, s') \in Z$;*
- *for all $s' \in R'[t']$ there is $s \in R[t]$ such that $(s, s') \in Z$.*

The transition systems $\mathbb{T}$ and $\mathbb{T}'$ are bisimilar *if there is a bisimulation $Z \subseteq T \times T'$ containing (s_I, s'_I). We write $\mathbb{T} \leftrightarroweq \mathbb{T}'$ to indicate that $\mathbb{T}$ and $\mathbb{T}'$ are bisimilar.*

The tree unraveling of a transition system $\mathbb{T}$ is denoted by $\mathbb{T}^e$.

Fact 2. *$\mathbb{T}$ and $\mathbb{T}^e$ are bisimilar, for every transition system $\mathbb{T}$.*

A class of transition systems $\mathcal{L}$ is *bisimulation closed* if $\mathbb{T} \leftrightarroweq \mathbb{T}'$ implies that $\mathbb{T} \in \mathcal{L} \leftrightarroweq \mathbb{T}' \in \mathcal{L}$ for each $\mathbb{T}$ and $\mathbb{T}'$. A formula φ is *bisimulation-invariant* if $\mathbb{T} \leftrightarroweq \mathbb{T}'$ implies that $\mathbb{T} \models \varphi \leftrightarroweq \mathbb{T}' \models \varphi$ for each $\mathbb{T}$ and $\mathbb{T}'$.

Fact 3. *Each MC-definable class of transition systems is bisimulation closed.*

The Janin-Walukiewicz theorem can be formulated as follows.

Fact 4 ([4])**.** *Let $\mathcal{L}$ be a bisimulation closed class of transition systems. The following are equivalent.*

1) *$\mathcal{L}$ is MC-definable.*
2) *$\mathcal{L}$ is MSO-definable.*

III. Automata for $WFMSO$

A. Automata over trees.

In this section we work with a restricted class of MSO-automata, called *weak MSO*-automata. Intuitively, an MSO-automaton is weak if the reachability relation on its states induced by the transition function 'respects' the parity map.

First, we present a first-order logic on a signature given by a set of unary predicates A that will be used to define the transition function of automata. We define $For^+(A)$ as the set of monadic first-order formulae with identity ($\approx$), where negation can only occur in front of atomic formulae of the kind $x \approx y$. Given a subset S of A, we introduce the notation

$$\tau^+_S(x) \quad := \quad \bigwedge_{a\in S} a(x).$$

The formula $\tau^+_S(x)$ is called a *(positive) A-type*. We use the convention that, if S is the empty set, then $\tau^+_S(x)$ is $\top$ and we call it an *empty* A-type. Given a set $Y \subseteq For^+(A)$ of formulae, $Disj(Y) = \{\bigvee X \mid X \subseteq_\omega Y\}$ is the collection of all finite disjunctions of formulae in Y. We indicate with $FO^+(A)$ the set of *sentences* from $For^+(A)$. A sentence $\varphi \in FO^+(A)$ is in *basic form* if it is of shape

$$\begin{aligned} \varphi \quad = \quad & \exists x_1 \ldots \exists x_k \Big(diff(\bar{x}) \wedge \bigwedge_{1\leq i\leq k} \tau^+_{B_i}(x_i) \\ & \wedge \forall z \ (diff(\bar{x}, z) \to \bigvee_{1\leq l\leq j} \tau^+_{C_l}(z))\Big), \end{aligned}$$

where each $\tau^+_{B_i}(x_i)$ and $\tau^+_{C_l}(z)$ is an A-type, $diff(y_1, \ldots, y_n) := \bigwedge_{1\leq m<m'<n}(y_m \not\approx y_{m'})$, and the conditional subformula is defined as expected. We denote with $BF^+(A)$ the set of all sentences from $FO^+(A)$ that are in

basic form. A sentence $\varphi \in BF^+(A)$ is in *functional basic form* if, for each non-empty A-type $\tau^+_S(x)$ occurring in it, S is a singleton. If φ is in functional basic form and no empty type occurs in it then we say that φ is in *special basic form*. We denote with $FBF^+(A)$ and $SBF^+(A)$ respectively the set of all sentences in $BF^+(A)$ which are in functional basic form and in special basic form.

Turning to the semantics, given a set X, a function $m : A \to \wp(X)$ and a valuation $v : Var \to X$, we define the notion of a formula $\varphi \in For^+(A)$ being *true* in (X, m, v) in the obvious way. In this setting, we call the function m a *marking*.

Definition 2. *An MSO*-automaton *on alphabet C is a tuple $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ where:*

- *A is a finite set of states, $a_I \in A$ is the initial state,*
- *$\Delta : A \times C \to Disj(FO^+(A))$,*
- *$\Omega : A \to \omega$ is a parity function,*

Let $\mathbb{A}$ be an MSO-automaton. Given two states $a, b \in A$, we say that b is reachable from a if there is a sequence $a_0, \dots, a_n$ of states in A such that $a_0 = a$, $a_n = b$ and for every $i < n$, a_{i+1} occurs in $\Delta(a_i, c)$, for some $c \in C$. An MSO-automaton is called *weak* if for every $a, b \in A$, a is reachable from b and b is reachable from a, then $\Omega(a) = \Omega(b)$. It is called *non-deterministic* if Δ has type $A \times C \to Disj(FBF^+(A))$.

Given a tree $\mathbb{T}$, the *acceptance game* $\mathcal{A}(\mathbb{A}, \mathbb{T})$ of $\mathbb{A}$ on $\mathbb{T}$ is the parity game defined according to the rules of table II. Finite matches of $\mathcal{A}(\mathbb{A}, \mathbb{T})$ are lost by the player who gets stuck. An infinite match of $\mathcal{A}(\mathbb{A}, \mathbb{T})$ is won by $\exists$ if and only if the *minimum* parity occurring infinitely often is even. The tree $\mathbb{T}$ is *accepted* by $\mathbb{A}$ if and only if $\exists$ has a winning strategy in $\mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$. The tree language accepted by $\mathbb{A}$ is denoted by $\mathcal{L}(\mathbb{A})$.

Remark 1. It is easy to see that every weak MSO-automaton can be seen as having a parity function ranging only over priorities $\{0, 1\}$. Intuitively, states with priority 0 are the accepting states, whereas states with priority 1 are the rejecting state. This is because a weak MSO-automaton accepts a tree iff in the corresponding acceptance game, Éloise can always force a play to finally stay in an even (i.e. accepting) strongly connected component of the automaton.

Fact 5 ([9])**.** *For every $\varphi \in MSO$, there is an effectively constructible MSO-automaton $\mathbb{A}_\varphi$ such that on tree languages $\|\varphi\| = \mathcal{L}(\mathbb{A}_\varphi)$.*

In what follows, we show that the analogon of the previous theorem also holds for $WFMSO$ and weak MSO-automata. The argument proceeds by induction on φ. We focus on the inductive case of $WFMSO$ existential quantification, which is the only non-trivial part of the proof. For this purpose, we define a closure operation on tree languages corresponding to the semantics of $WFMSO$ existential quantification.

Definition 3. *Let $\mathbb{T}$ be a tree and p a propositional letter (not necessarily in P). Let $\mathcal{L}$ be a tree language. The* noetherian projection *of $\mathcal{L}$ over p is the language $\exists_W p.\mathcal{L}$ defined as the class of trees $\mathbb{T}$ such that there is a noetherian p-variant $\mathbb{T}'$ of $\mathbb{T}$, with $\mathbb{T}' \in \mathcal{L}$.*

A class $\mathcal{C}$ of tree languages is closed under noetherian projection over p *if, for any language $\mathcal{L}$ in $\mathcal{C}$, also the language $\exists_W p.\mathcal{L}$ is in $\mathcal{C}$.*

B. The Two-Sorted Construction.

Our goal is to provide a *projection construction* that, given a weak MSO-automaton $\mathbb{A}$, provides a weak MSO-automaton $\exists_W p.\mathbb{A}$ recognizing $\exists_W p.\mathcal{L}(\mathbb{A})$.

The idea is to proceed by analogy with the construction showing that the tree languages recognized by MSO-automata are closed under projection. In the case of MSO-automata, the proof crucially uses the fact that every MSO-automaton can be simulated by a *non-deterministic* MSO-automaton.

Fact 6 (Simulation Theorem [9], [16])**.** *For every MSO-automaton $\mathbb{A}$, there is an effectively constructible non-deterministic MSO-automaton $\mathbb{A}^n$ which is equivalent to $\mathbb{A}$.*

Unfortunately, the proof of this result does not transfer to the setting of *weak MSO*-automata, in the sense that starting with a weak automaton $\mathbb{A}$ one does not necessarily end up with an automaton $\mathbb{A}^n$ which is also weak. This means that we cannot use the full power of non-determinism in the projection construction for weak MSO-automata. This notwithstanding, in the sequel we show how a restricted version of non-determinism suffices for our purposes.

Let $\mathbb{A}$ be a weak MSO-automaton, $\mathbb{T}$ a tree and f a winning strategy for $\exists$ in $\mathcal{G}_A = \mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$. It is not difficult to verify that non-determinism corresponds to the property that any marking suggested by f assigns *at most one* state to the successors of the node under consideration. If this is the case, we say that f is *functional*. The nice thing about this property is that it propagates, in the sense that if $\exists$ plays a functional strategy f in $\mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$, then for any node s in $\mathbb{T}$ there is at most one state a of the automaton such that the position (a, s) can be reached, in any match that is consistent with f. This is particularly helpful when, in order to define a p-variant of $\mathbb{T}$ that is accepted by the projection construction over $\mathbb{A}$, we need to decide whether such a node s should be labeled with p or not.

Now, in the case of weak MSO-automata we are interested only in *noetherian* p-variants: the main idea is that guessing a noetherian p-variant only requires f to be functional in a finite initial segment (i.e. a partial match) π_F of each f-conform match π of $\mathcal{G}_A$. This amounts to say that $\mathbb{A}$ behaves as a non-deterministic automaton as far as the match is played along π_F. We call this behavior the *non-deterministic mode* of $\mathbb{A}$. During the remaining part of the match, in which f is no longer required to be functional, we say that $\mathbb{A}$ has entered the *alternating mode*. This distinction induces a well-founded subtree $\mathbb{W}$ of $\mathbb{T}$, consisting of the nodes from which f is functionally defined. A noetherian p-variant of $\mathbb{T}$ is built by allowing only nodes in $\mathbb{W}$ to be labeled with p.

To formalize this argument, which goes back to [11], we first show that every weak MSO-automaton $\mathbb{A}$ can be turned into

Position	Player	Admissible moves	Parity
$(a,s) \in A \times S$	$\exists$	$\{m : A \to \wp(R[s]) \mid (R[s], m) \models \Delta(a, V(s))\}$	$\Omega(a)$
$m : A \to \wp(R[s])$	$\forall$	$\{(b,t) \mid t \in m(b)\}$	$Max(\Omega[A])$

Table II
Acceptance game for MSO-automata

an equivalent weak MSO-automaton $\mathbb{A}^{2S}$, which we call *two-sorted* since its carrier is split into an initial non-deterministic and a subsequent alternating part. For the precise definition of the non-deterministic part of $\mathbb{A}^{2S}$ we need the following proposition.

Proposition 1. *For every MSO-automaton $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$, there is an effectively constructible automaton $\mathbb{A}^\sharp = \langle A^\sharp, a_I^\sharp, \Delta^\sharp, NBT\rangle$ which is non-deterministic, i.e. $\Delta^\sharp$ has type $A^\sharp \times C \to Disj(FBF^+(A^\sharp))$, is based on the set $A^\sharp = \wp(A \times A)$ of binary relations over A, takes the singleton set $a_I^\sharp = \{(a_I, a_I)\}$ as its starting state, and has the property that for any binary relation $Q \subseteq A \times A$, and any tree $\mathbb{T}$:*

$$\mathbb{A}_Q^\sharp \text{ accepts } \mathbb{T} \text{ iff } \mathbb{A}_a \text{ accepts } \mathbb{T}, \text{ for all } a \in Ran(Q),$$

where $\mathbb{A}_a = \langle A, a, \Delta, \Omega\rangle$ denotes the variant of the automaton $\mathbb{A}$ that takes a as its starting state, and similarly for $\mathbb{A}_Q^\sharp$.

In particular, the automaton $\mathbb{A}^\sharp$ itself is equivalent to $\mathbb{A}$.

We call $\mathbb{A}^\sharp$ the *refined powerset construct* over $\mathbb{A}$. Note that $A^\sharp$ is *almost* a non-deterministic MSO-automaton, the only difference being that the acceptance condition is not given by a parity condition.

We now turn to the definition of the automaton $\mathbb{A}^{2S}$, which we call *two-sorted*, because it roughly consists of a copy of $\mathbb{A}^\sharp$ 'followed by' a copy of $\mathbb{A}$. As we observed, $\mathbb{A}^\sharp$ is a non-deterministic automaton, whereas $\mathbb{A}$ generally is not. Thus, given a tree $\mathbb{T}$, the idea is to make any match π of $\mathcal{A}(\mathbb{A}^{2S}, \mathbb{T})$ consist of two parts:

- **(Non-deterministic mode)** During finitely many steps, π can be seen as a match of the acceptance game of $A^\sharp$ on $\mathbb{T}$, where any winning strategy for $\exists$ can be assumed to be functional;
- **(Alternating mode)** At a certain stage, π abandons the non-deterministic part of $\mathbb{A}^{2S}$ and turns into a match of the acceptance game of $\mathbb{A}$ on $\mathbb{T}$.

The definition of $\mathbb{A}^{2S}$ will guarantee the correctness of this construction, making $\mathbb{A}^{2S}$ equivalent to the original automaton $\mathbb{A}$.

Definition 4. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be a weak MSO-automaton and $A^\sharp = \langle A^\sharp, a_I^\sharp, \Delta^\sharp, NBT\rangle$ its refined powerset construct. The weak MSO-automaton $\mathbb{A}^{2S} = \langle A^{2S}, a_I^{2S}, \Delta^{2S}, \Omega^{2S}\rangle$ is defined as follows.*

$$\begin{aligned}
A^{2S} &:= A \cup A^\sharp \\
a_I^{2S} &:= a_I^\sharp \\
\Delta^{2S}(a,c) &:= \Delta(a,c) \\
\Delta^{2S}(R,c) &:= \Delta^\sharp(R,c) \vee \bigwedge_{a \in Ran(R)} \Delta(a,c) \\
\Omega^{2S}(a) &:= \Omega(a) \\
\Omega^{2S}(R) &:= 1
\end{aligned}$$

Here a and R denote arbitrary states in A and $A^\sharp$, respectively. The automaton $\mathbb{A}^{2S}$ is called the two-sorted construct over $\mathbb{A}$.

Then we can prove a version of a simulation theorem that will suffice for our purposes.

Proposition 2. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be a weak MSO-automaton and $\mathbb{A}^{2S}$ the two-sorted construct on $\mathbb{A}$. Then $\mathcal{L}(\mathbb{A}^{2S}) = \mathcal{L}(\mathbb{A})$.*

C. Closure under noetherian projection.

We are now ready to show the main result of this section: the class of tree languages recognized by weak MSO-automata is closed under noetherian projection. The argument is analogous to the one showing that MSO-automata are closed under projection, but we use the two-sorted construction instead of the refined powerset construction. The p-variant induced by the projection automaton will be guaranteed to be noetherian because all nodes labeled with p are visited when the automaton is in non-deterministic mode.

Definition 5. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be a weak MSO-automaton on alphabet $\wp(P \cup \{p\})$. Let $\mathbb{A}^{2S}$ denote its two-sorted construct. We define the automaton $\exists_W p.\mathbb{A} = \langle A^{2S}, a_I^{2S}, \widetilde{\Delta}, \Omega^{2S}\rangle$ on alphabet C by putting*

$$\begin{aligned}
\widetilde{\Delta}(a,c) &:= \Delta^{2S}(a,c) \\
\widetilde{\Delta}(R,c) &:= \Delta^{2S}(R,c) \vee \Delta^{2S}(R, c \cup \{p\}).
\end{aligned}$$

The automaton $\exists_W p.\mathbb{A}$ is called the two-sorted projection construct of $\mathbb{A}$ over p.

Proposition 3. *For each weak MSO-automaton $\mathbb{A}$ on alphabet $\wp(P \cup \{p\})$, we have that $\mathcal{L}(\exists_W p.\mathbb{A}) = \exists_W p.\mathcal{L}(\mathbb{A})$.*

As mentioned, the above proposition takes care of the only non-trivial induction case in the inductive proof of the following analogon of Fact 5:

Theorem 3. *For every $\varphi \in WFMSO$, there is an effectively constructible weak MSO-automaton $\mathbb{A}_\varphi$ such that on tree languages $\|\varphi\| = \mathcal{L}(\mathbb{A}_\varphi)$.*

Remark 2. Given any non-deterministic MSO-automaton $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ where $\Delta : A \times C \to Disj(FBF^+(A))$ we can construct an equivalent non-deterministic MSO-automaton $\mathbb{A}' = \langle A', a_I, \Delta', \Omega'\rangle$ with Δ' of type $A' \times C \to Disj(SBF^+(A'))$. That is, we may replace an arbitrary 'FBF-automaton' $\mathbb{A}$ with an equivalent 'SBF-automaton' $\mathbb{A}'$. This automaton $\mathbb{A}'$ is based on carrier $A \cup \{a^\top\}$, where $a^\top \notin A$ acts as a 'bin state' always leading to the acceptance of the input tree. For each $a \in A$ and $c \in C$, we can replace the empty A-types $\tau_S^+(x) = \top$ occurring in $\Delta(a, c)$ with $a^\top(x)$. This leads to the definition of a transition function Δ' associated only with sentences in *special* basic form. It is readily seen that any winning strategy for $\exists$ in the acceptance game for $\mathbb{A}'$ and some input tree $\mathbb{T}$ can be assumed to mark each node of $\mathbb{T}$ with *exactly one* state of $\mathbb{A}'$. This strengthening of the functionality condition conveniently simplifies the constructions presented in the next section.

IV. From weak *MSO*-Automata to *WFMSO*

In this section we discuss the converse statement of Theorem 3. Using an argument going back to Rabin [10], we will prove that for every weak *MSO*-automaton $\mathbb{A}$ there is a formula $\varphi_{\mathbb{A}} \in$ *WFMSO* which is equivalent to $\mathbb{A}$.

A. From Weak MSO to Büchi Automata.

The first idea would be to construct, for a weak *MSO*-automaton $\mathbb{A}$, a formula $\varphi_{\mathbb{A}}$ that expresses, when interpreted in a given tree $\mathbb{T}$, the existence of a winning strategy f for $\exists$ in $\mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$. This encoding would go smoothly if we could assume that f marks each node with *exactly one* state of $\mathbb{A}$.

For this purpose, by Theorem 6 we can construct a non-deterministic *MSO*-automaton $\mathbb{A}^n$ which is equivalent to $\mathbb{A}$. However, as observed already, the automaton $\mathbb{A}^n$ is not generally weak. This means that different parities can occur infinitely often in the same match of $\mathcal{A}(\mathbb{A}^n, \mathbb{T})$. Intuitively, this implies that we cannot give an account of the winning conditions of this acceptance game by referring only to well-founded subtrees of $\mathbb{T}$. The quantification of *WFMSO* is too restrictive and it would seem that we need instead the full generality of *MSO* quantifiers.

We overcome this difficulty by showing that, because it is weak, $\mathbb{A}$ can be turned into an equivalent non-deterministic *Büchi* automaton (abbreviated *NDB*), i.e. a non-deterministic *MSO*-automaton $\mathbb{B}$ where the parity map $\Omega_B : B \to \omega$ can be assumed to range only over $\{0, 1\}$. The acceptance game associated with such a $\mathbb{B}$ turns out to be essentially simpler than the one for arbitrary *MSO*-automata. The states of $\mathbb{B}$ can be divided into *accepting states* - the ones with parity 0 - and *rejecting states* - the ones with parity 1. It should be clear that $\exists$ wins a match if and only if at least one accepting state occurs infinitely often along the play. This constitutes a Büchi acceptance condition, and in fact we can simply describe $\mathbb{B}$ as an automaton where the acceptance condition is given by a set F of accepting states, instead of a parity map Ω. It turns out that Büchi acceptance conditions can be described in terms of well-founded trees, so that we can express them by means of *WFMSO*-formulae without requiring the full expressiveness of *MSO* quantifiers. This is the key observation leading to the logical characterization of non-deterministic Büchi automata and weak *MSO*-automata.

Definition 6 (Büchi powerset construction)**.** *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be a weak MSO-automaton. We can assume that $Ran(\Omega)$ is a subset of $\{0, 1\}$. Let $A^\sharp = \langle A^\sharp, a_I^\sharp, \Delta^\sharp, NBT\rangle$ be the refined powerset construct over $\mathbb{A}$. We define an NDB automaton $\mathbb{A}^B = \langle A^\sharp, a_I^\sharp, \Delta^\sharp, F_\Omega\rangle$ by putting*

$$F_\Omega \quad := \quad \{R \in A^\sharp \mid \Omega(a) = 0 \textit{ for all } a \in Ran(R)\}.$$

We say that $\mathbb{A}^B$ is the Büchi powerset construct *over $\mathbb{A}$.*

We can now verify the following.

Proposition 4. *Let $\mathbb{A}$ be a weak MSO-automaton and $\mathbb{A}^B$ the Büchi powerset construct over $\mathbb{A}$. We have that $\mathcal{L}(\mathbb{A}) = \mathcal{L}(\mathbb{A}^B)$.*

B. The Bounded Information Property.

We now formalize two key intuitions about non-deterministic Büchi automata:

1) checking whether a non-deterministic Büchi automaton $\mathbb{B}$ accepts a tree $\mathbb{T}$ reduces to verifying a condition on prefixes of $\mathbb{T}$ (Proposition 5);
2) checking whether the intersection of the languages of two non-deterministic Büchi automata is non-empty can proceed via the construction of a finite sequence of well-founded trees with certain properties (Proposition 6).

The idea is that a run of a non-deterministic Büchi automaton on a tree $\mathbb{T}$ can be split into several tasks concerning well-founded subtrees (and prefixes, which are just a particular kind of well-founded subtrees) of $\mathbb{T}$, and there is never the need to consider $\mathbb{T}$ as a whole.

Definition 7. *Let $\mathbb{B} = \langle B, b_I, \Delta, F\rangle$ be a non-deterministic Büchi automaton and $\mathbb{T}$ a tree. Let f be a surviving strategy for $\exists$ in $\mathcal{A}(\mathbb{B}, \mathbb{T})@(b_I, s_I)$. Let $\gamma \leq \omega$ be an ordinal.* A γ-accepting sequence for f over $\mathbb{B}$ and $\mathbb{T}$ *is a sequence $(E_i)_{i<\gamma}$ such that, for all $i < \gamma$:*

1) *E_i is a prefix of $\mathbb{T}$;*
2) *$Ft(E_i) < Ft(E_{i+1})$;*
3) *for each s in the frontier of E_i, there is a unique $a \in A$ such that f is defined on position (a, s); in addition, a is in F.*

Intuitively, for $k < \omega$, a k-accepting sequence for a surviving strategy f witnesses the fact that f 'behaves as' a winning strategy for $\exists$ in the prefix E_k of $\mathbb{T}$. For each prefix E_i in the sequence, the condition that each $s \in Ft(E_i)$ is associated with a *unique* accepting state is motivated by the fact that f can be assumed to be functional, $\mathbb{B}$ being non-deterministic.

Proposition 5. *Let $\mathbb{B} = \langle B, b_I, \Delta, F\rangle$ be a non-deterministic Büchi automaton and $\mathbb{T}$ a tree. The following are equivalent.*

- *Player $\exists$ has a winning strategy in $\mathcal{A}(\mathbb{B}, \mathbb{T})@(b_I, s_I)$.*
- *Player $\exists$ has a surviving strategy f in $\mathcal{A}(\mathbb{B}, \mathbb{T})@(b_I, s_I)$ and there is an ω-accepting sequence for f over $\mathbb{B}$ and $\mathbb{T}$.*

For non-deterministic Büchi automata $\mathbb{B}_1$ and $\mathbb{B}_2$ and a tree $\mathbb{T} \in L(\mathbb{B}_1) \cap L(\mathbb{B}_2)$, let $(G_i^1)_{i<\omega}$ and $(G_i^2)_{i<\omega}$ be ω-accepting sequences respectively for $\mathbb{B}_1$ and $\mathbb{B}_2$ on $\mathbb{T}$. We introduce the notion of *k-trap* for $\mathbb{B}_1$ and $\mathbb{B}_2$. The idea is that a k-trap is a finite sequence $(E_i)_{i\leq k}$ witnessing some kind of interleaving of the sequences $(G_i^1)_{i<\omega}$ and $(G_i^2)_{i<\omega}$ up to level k.

In this aim, we first introduce the following auxiliary notion. Let $\mathbb{B}$ be a NDB-automaton and $\mathbb{T}$ a tree. Given a set of nodes $N \subset T$, we say that a strategy f for player $\exists$ in $\mathcal{A}(\mathbb{B},\mathbb{T})@(b_I,s_I)$ is *surviving in* N if, for each basic position $(b,s) \in B \times N$ that is reached in some f-conform match, the marking m suggested by f makes $\Delta(b, V(s))$ true in $R[s]$.

Definition 8. *Let* $\mathbb{B}_1 = \langle B_1, b_I^1, \Delta_1, F_1\rangle$ *and* $\mathbb{B}_2 = \langle B_2, b_I^2, \Delta_2, F_2\rangle$ *be NDB automata and let* $\mathbb{T}$ *be a tree. Given some fixed* $k < \omega$, *let* $(E_i)_{i\leq k}$ *be a sequence of prefixes of* $\mathbb{T}$ *such that* $E_0 = \{s_I\}$ *and* $E_i \subsetneq E_{i+1}$ *for all* $i \leq k$.

We say that $\mathbb{T}$ *and* $(E_i)_{i\leq k}$ *constitute a* k-trap for $\mathbb{B}_1$ and $\mathbb{B}_2$ *if there exist*

1) *a strategy* f_1 *for* $\exists$ *in* $\mathcal{A}(\mathbb{B}_1,\mathbb{T})@(b_I^1,s_I)$ *which is surviving in* E_k,
2) *a strategy* f_2 *for* $\exists$ *in* $\mathcal{A}(\mathbb{B}_2,\mathbb{T})@(b_I^2,s_I)$ *which is surviving in* E_k,
3) *a k-accepting sequence* $(G_i^1)_{i\leq k}$ *for* f_1 *over* $\mathbb{B}^1$ *and* $\mathbb{T}$,
4) *a k-accepting sequence* $(G_i^2)_{i\leq k}$ *for* f_2 *over* $\mathbb{B}^2$ *and* $\mathbb{T}$,

such that, for all $i < k$, *the following conditions hold:*

- $Ft(E_i) \leq Ft(G_i^1) < Ft(E_{i+1})$;
- $Ft(E_i) \leq Ft(G_i^2) < Ft(E_{i+1})$.

We say that the strategies f_1 *and* f_2 witness *the k-trap for* $\mathbb{B}_1$ *and* $\mathbb{B}_2$.

Proposition 6 ([10])**.** *Let* $\mathbb{B}_1$ *and* $\mathbb{B}_2$ *be NDB-automata and let* m *be the product of the cardinalities of their carriers. If there exists an* m*-trap for* $\mathbb{B}_1$ *and* $\mathbb{B}_2$ *then* $\mathcal{L}(\mathbb{B}_1)\cap\mathcal{L}(\mathbb{B}_2) \neq \emptyset$.

C. Non-Deterministic Büchi Automata versus $WFMSO$.

We are now ready to prove the main result of this section.

Theorem 4. *For any weak MSO-automaton* $\mathbb{A}$ *there is a formula* $\varphi \in WFMSO$ *such that over tree languages* $\|\varphi\| = \mathcal{L}(\mathbb{A})$.

Proof: Let $\mathbb{A}$ be a weak MSO-automaton and $\mathbb{B}$ an NDB-automaton which is equivalent to $\mathbb{A}$, as in Proposition 4. Clearly then it suffices to come up with a formula ϕ in $WFMSO$ that holds in a tree $\mathbb{T}$ if and only if $\mathbb{B}$ accepts $\mathbb{T}$. Since weak MSO-automata are closed under complementation, we are also provided with a weak MSO-automaton $\overline{\mathbb{A}}$ recognizing the complement of $\mathcal{L}(\mathbb{A})$, and consequently with an NDB-automaton $\overline{\mathbb{B}}$ which is equivalent to $\overline{\mathbb{A}}$. Our formula $\varphi = \varphi_{\mathbb{B},\overline{\mathbb{B}}}$ depends on both $\mathbb{B}$ and $\overline{\mathbb{B}}$.

More concretely, let m be the product of the cardinalities of B and $\overline{B}$. The formula $\varphi_{\mathbb{B},\overline{\mathbb{B}}} \in WFMSO$ will express the existence of a strategy f for $\exists$ and an $m+1$-accepting sequence $(E_i)_{i\leq m+1}$ such that f is functional and surviving in E_{m+1}. The key observation is that the encoding of $(E_i)_{i\leq m+1}$ and the associated surviving strategy into a formula only requires variables for noetherian sets of nodes. For this, the expressive power of $WFMSO$ will suffice.

Proposition 5 will help showing one direction of the equivalence, namely that, given a tree $\mathbb{T}$ and a winning strategy f for $\exists$ in $\mathcal{A}(\mathbb{B},\mathbb{T})$, the formula $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ is true in $\mathbb{T}$. For the converse direction, we use the automaton $\overline{\mathbb{B}}$ accepting the complement of the language of $\mathbb{B}$. The idea is to suppose by way of contradiction that $\overline{\mathbb{B}}$ accepts a tree $\mathbb{T}$ where $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ is true. Then by Proposition 5 there is an ω-accepting sequence $(E_i^\delta)_{i<\omega}$ witnessing the fact that $\mathbb{T}$ is in $\mathcal{L}(\overline{\mathbb{B}})$. The ω-accepting sequence $(E_i^\delta)_{i<\omega}$ contains an $m+1$-accepting sequence $(E_i^\delta)_{i\leq m+1}$. By the fact that $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ is true, we also have an $m+1$-accepting sequence $(E_i)_{i\leq m+1}$. Then we can show that the two sequences witness a trap for $\mathbb{B}$ and $\overline{\mathbb{B}}$ as in Definition 8. But by Proposition 6 this means that the intersection of $\mathcal{L}(\mathbb{B})$ and $\mathcal{L}(\overline{\mathbb{B}})$ is non-empty, contradicting the fact that $\overline{\mathbb{B}}$ accepts the complement of $\mathcal{L}(\mathbb{B})$.

The definition of $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ essentially follows the same line of reasoning as in [10]. Given any state $b \in B$, we define by induction on $i < \omega$ a formula $K_i^b(x) \in WFMSO$, where no variable different from x occurs free. For the base case, we put $K_0^b(x) := \top$. Inductively, $K_{i+1}^b(x)$ is given as a formula expressing the following situation (relative to a tree $\mathbb{T}$):

- Given a node s on which x is being evaluated, for each prefix E of $\mathbb{T}.s$, there is a prefix E' of $\mathbb{T}.s$ including E, and a function $m_p : B \to \wp(E)$, such that $\exists$ has a functional strategy f in $\mathcal{A}(\mathbb{B},\mathbb{T}.s)@(s,b)$, which is surviving in E' and has the following properties:
  - from each basic position (b',t) on which it is defined, the strategy f suggests to $\exists$ the restriction of m_p to a marking $m_{p,t} : B \to \wp(R[t])$;
  - for each node t on the frontier of E', let $b_t \in B$ be the unique state in B such that (b_t,t) is a reachable position in an f-conform match. Then b_t is an accepting state in F, and the formula $K_i^{b_t}(y)$ is true in $\mathbb{T}$ for y evaluated on t.

Given a formula $Root(y)$ stating that y is the root of the tree, we define $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ as $\exists y\ (Root(y) \wedge K_{m+1}^{b_I}(y))$. Then the key lemma underlying the proof of Theorem 4 states that

$$\mathcal{L}(\mathbb{B}) = \|\varphi_{\mathbb{B},\overline{\mathbb{B}}}\|.$$

This finishes the proof (sketch) of Theorem 4. ∎

As an immediate corollary we obtain the following characterization of $WFMSO$ which generalizes Rabin's automata-theoretic characterization of $WMSO$ on binary trees [10].

Corollary 1. *A tree language* $\mathcal{L}$ *is WFMSO-definable if and only if there are non-deterministic Büchi automata* $\mathbb{B}$ *and* $\overline{\mathbb{B}}$ *such that* $\mathcal{L} = \mathcal{L}(\mathbb{B})$ *and* $\overline{\mathcal{L}} = \mathcal{L}(\overline{\mathbb{B}})$.

D. Proof of Theorem 2

Finally, Theorem 2 is now immediate by the Theorems 3 and 4 and Corollary 1.

V. A Characterization Theorem for $AFMC$

Theorem 4 states that over transition systems, the modal μ-calculus (MC) is as expressive as the bisimulation-invariant fragment of MSO. In this section we consider the same question for the bisimulation-invariant fragment of $WFMSO$. It turns out that $WFMSO$ is still weaker than MSO in this respect, being as expressive as the alternation-free fragment of the modal μ-calculus ($AFMC$). This outcome is coherent with the perspective on $WFMSO$ and weak MSO-automata. Indeed, there is a tight connection between fixpoint operators of the μ-calculus and parities occurring infinitely often in parity games [17]. The absence of alternation in formulae of $AFMC$ intuitively corresponds to at most one parity occurring infinitely often along infinite matches of a parity game (cf. Remark 1).

Turning to the proof of our main result, Theorem 1, we once again use an automata-theoretic argument. Roughly, the idea is that automata for $AFMC$ over trees are the weak counterpart of automata for MC, just as automata for $WFMSO$ are the weak version of MSO-automata. Then the argument, used by Janin & Walukiewicz [4] to show that automata for MC and MSO have the same expressive power modulo bisimulation, can be restricted to show an analogous result for the weak counterparts.

In the sequel we use the translation introduced in [4], which transforms sentences in special basic form into sentences of $FO^+(A)$ without equality. These will provide the first-order language associated with automata for MC.

Definition 9 (Modal Translation)**.** *Given a set A of unary predicates, let $\varphi \in SBF^+(A)$ be a sentence in special basic form of shape*

$$\varphi \quad = \quad \exists x_1 \dots \exists x_k \Big(\mathit{diff}(\bar{x}) \wedge \bigwedge_{1 \leq i \leq k} a_i(x_i) \wedge \forall z (\mathit{diff}(\bar{x}, z) \rightarrow \bigvee_{1 \leq l \leq j} b_l(z)) \Big).$$

We define its modal translation φ^∇ by putting

$$\varphi^\nabla \quad := \quad \exists x_1 \dots \exists x_k \bigwedge_{1 \leq i \leq k} a_i(x_i) \wedge \forall z \bigvee_{1 \leq l \leq j} b_l(z).$$

We denote with $SBF^\nabla(A)$ the set $\{\varphi^\nabla \mid \varphi \in SBF^+(A)\}$.

Remark 3. Our terminology stems from the observation that the formula ϕ^∇ corresponds to the modal formula $\bigwedge_{1 \leq i \leq k} \Diamond a_i \wedge \Box \bigvee_{1 \leq l \leq j} b_l$.

The modal μ-calculus is characterized by a class of automata which are defined as non-deterministic MSO-automata but for the transition function, which ranges over sentences from $Disj(SBF^\nabla(A))$ instead of from $Disj(SBF^+(A))$ [18]. If we restrict to the alternation-free fragment, then a weaker version of these automata suffices [19]. We use the name *modal non-deterministic Büchi automata* to emphasize their connection with non-deterministic Büchi automata.

Definition 10. *A* modal non-deterministic Büchi automaton *on alphabet C is an MSO-automaton $\mathbb{B} = \langle B, b_I, \Delta, \Omega \rangle$ with $\Delta : B \times C \rightarrow Disj(SBF^\nabla(B))$ and $\Omega : B \rightarrow \{0, 1\}$.*

In [19] an automata-theoretic characterization of $AFMC$ in terms of modal non-deterministic Büchi automata is provided.

Fact 7 ([19])**.** *Let $\mathcal{L}$ be a tree language. The following are equivalent.*

- *There exists $\varphi \in AFMC$ such that $\mathcal{L} = \|\varphi\|$.*
- *There are modal non-deterministic Büchi automata $\mathbb{M}$ and $\overline{\mathbb{M}}$ such that $\mathcal{L} = \mathcal{L}(\mathbb{M})$ and $\overline{\mathcal{L}} = \mathcal{L}(\overline{\mathbb{M}})$.*

We introduce a translation from NDB-automata to modal NDB-automata. Let $\mathbb{B}$ be a NDB-automaton. In analogy with Janin and Walukiewicz's argument, we are going to show that, if $L(\mathbb{B})$ is closed under bisimulation, then the modal NDB automaton $\mathbb{B}^\nabla$ that we obtain from $\mathbb{B}$ through the translation is such that $\mathbb{B} \equiv \mathbb{B}^\nabla$. This is the content of Proposition 7.

Definition 11. *Let $\mathbb{B} = \langle B, b_I, \Delta, \Omega \rangle$ be an NDB automaton. We define an automaton $\mathbb{B}^\nabla = \langle B, b_I, \Delta^\nabla, \Omega \rangle$ by putting*

$$\Delta^\nabla(a, c) \quad := \quad \bigvee \{\varphi^\nabla \mid \varphi \text{ is a disjunct of } \Delta(a, c)\}.$$

By Definition 9 the transition function Δ^∇ has type $B \times C \rightarrow Disj(SBF^\nabla(B))$, meaning that $\mathbb{B}^\nabla$ is a modal NDB-automaton.

Proposition 7. *Let $\mathbb{B}$ be an NDB automaton and $\mathbb{B}^\nabla$ the modal NDB automaton constructed from $\mathbb{B}$ as in Definition 11. If $\mathcal{L}(\mathbb{B})$ is closed under bisimulation, then $\mathcal{L}(\mathbb{B}) = \mathcal{L}(\mathbb{B}^\nabla)$.*

This proposition provides the key result to prove that $AFMC$ is at least as expressive as the bisimulation-invariant fragment of $WFMSO$. The converse statement is in fact the easy direction of Theorem 1, being essentially a corollary of the automata-theoretic characterization of $AFMC$ over trees provided in [19].

Fact 8 ([19])**.** *Let $\varphi \in AFMC$ be a sentence. There is a weak MSO-automaton $\mathbb{A}_\varphi$ such that on tree languages $\|\varphi\| = \mathcal{L}(\mathbb{A}_\varphi)$.*

Proof of Theorem 1: Because of Proposition 2, it is enough to prove the claim for tree languages. Thus, let $\mathcal{L}$ be a tree language that is closed under bisimulation. The direction $(1 \Rightarrow 2)$ follows by Proposition 8 and Theorem 4. The proof of direction $(2 \Rightarrow 1)$ is obtained as follows. Assume that there is a formula $\varphi \in WFMSO$ such that $\|\varphi\| = \mathcal{L}$. By Theorem 2, there are NDB-automata $\mathbb{B}, \overline{\mathbb{B}}$ such that $\mathcal{L}(\mathbb{B}) = \mathcal{L}$ and $\mathcal{L}(\overline{\mathbb{B}}) = \overline{\mathcal{L}}$. By Proposition 7, this implies that there are modal NDB-automata $\mathbb{B}^\nabla, \overline{\mathbb{B}}^\nabla$ such that $\mathcal{L}(\mathbb{B}^\nabla) = \mathcal{L}$ and $\mathcal{L}(\overline{\mathbb{B}}^\nabla) = \overline{\mathcal{L}}$. Finally, Proposition 7 yields a formula $\varphi_1 \in AFMC$ such that $\|\varphi_1\| = \mathcal{L}$. ∎

As a corollary of Theorem 1, we obtain an incomparability result for $WFMSO$ and $WMSO$. We can see this as a strengthening of the incomparability between $WMSO$ and MSO, $WFMSO$ being strictly weaker than MSO.

Corollary 2. *The collection of WMSO-definable classes of transition systems and the collection of WFMSO-definable classes of transition systems are incomparable.*

VI. Conclusion

A. Overview.

In this work we have presented two main contributions. The first one concerns the connection between automata and logic and establishes a logical characterization of weak MSO-automata on trees of arbitrary branching degree. For this purpose we introduce a new variant of MSO which we call well-founded monadic second-order logic ($WFMSO$), and prove that for tree languages, being $WFMSO$-definable and being accepted by a weak MSO-automata coincide. The proof passes through non-deterministic Büchi automata, that generalize Rabin's 'special automata' [10] working on binary trees. We give a second characterization for $WFMSO$ in connection with this class of automata: a tree language $\mathcal{L}$ is $WFMSO$-definable if and only if both $\mathcal{L}$ and its complement are recognized by non-deterministic Büchi automata. This generalizes of an analogous result of Rabin for $WMSO$ on binary trees [10].

The second main contribution is the modal characterization of the bisimulation-invariant fragment of $WFMSO$, which is proven to be as expressive as the alternation-free fragment of the modal μ-calculus. This result somehow completes the net of correspondences between $WFMSO$ and MSO, the bisimulation-invariant fragment of MSO being as expressive as the modal μ-calculus [4]. As expected, this implies that $WFMSO$ and $WMSO$ have incomparable expressive power.

B. Future Work.

The original driving motivation of our work was the observation that weak MSO-automata do not characterize $WMSO$ on all trees, meaning that $AFMC$ is not the bisimulation-invariant fragment of $WMSO$. A natural continuation would be thence to provide a different class of automata, which characterizes $WMSO$. The crux of the matter is to understand how to define these automata, in such a way that their expressive power is incomparable with respect to MSO-automata. In particular they should lack the finite branching property. But then a problem arises, for all the projection constructions that we considered so far are tightly connected to such property. In order to give a projection construction corresponding to $WMSO$-quantification, essentially different methods seem to be proposed.

A second natural line of research concerns the bisimulation-invariant fragment of $WMSO$. This investigation is motivated by the fact that all $WMSO$-definable tree languages are Borel. If the bisimulation-invariant fragment of $WMSO$ is strictly weaker than the modal μ-calculus, then it would correspond to a sort of 'Borelian' fragment, providing a better understanding of the topological complexity of modal fixpoint logics. In fact there are reasons to believe that this is the case. To the best of our knowledge, all examples of tree languages that are $WMSO$-definable but not MSO-definable are not bisimulation closed. This motivates the conjecture that the bisimulation-fragment of $WMSO$ 'collapses inside' the modal μ-calculus, and particularly its alternation-free fragment for the intuitive reason that $WMSO$ is not stronger than $WFMSO$ in expressing properties on the vertical dimension of trees.

ACKNOWLEDGMENT

The first author is supported by the *Expressiveness of Modal Fixpoint Logics* project realized within the 5/2012 Homing Plus programme of the Foundation for Polish Science, co-financed by the European Union from the Regional Development Fund within the Operational Programme Innovative Economy ("Grants for Innovation").

REFERENCES

[1] J. van Benthem, "Modal correspondence theory," Ph.D. dissertation, Universiteit van Amsterdam, 1977.
[2] E. Rosen, "Modal logic over finite structures," *Journal of Logic, Language and Information*, vol. 6, pp. 427–439, 1997.
[3] A. Dawar and M. Otto, "Modal characterisation theorems over special classes of frames," *Ann. Pure Appl. Logic*, vol. 161, no. 1, pp. 1–42, 2009.
[4] D. Janin and I. Walukiewicz, "On the expressive completeness of the propositional μ-calculus with respect to monadic second order logic," in *Proceedings of the 7th International Conference on Concurrency Theory*, ser. CONCUR '96. London, UK: Springer-Verlag, 1996, pp. 263–277. [Online]. Available: http://portal.acm.org/citation.cfm?id=646731.703838
[5] A. Arnold and D. Niwinski, "Fixed point characterization of weak monadic logic definable sets of trees," in *Tree Automata and Languages*, 1992, pp. 159–188.
[6] B. ten Cate and A. Facchini, "Characterizing EF over infinite trees and modal logic on transitive graphs," in *MFCS*, 2011, pp. 290–302.
[7] L. Alberucci and A. Facchini, "The modal μ-calculus over restricted classes of transition systems," *J. Symb. Log.*, vol. 74, no. 4, pp. 1367–1400, 2009.
[8] F. Zanasi, "Expressiveness of monadic second order logics on infinite trees of arbitrary branching degree," Master's thesis, ILLC, Universiteit van Amsterdam.
[9] I. Walukiewicz, "Monadic second order logic on tree-like structures," in *STACS*, 1996, pp. 401–413.
[10] M. O. Rabin, "Weakly definable relations and special automata," in *Proceedings of the Symposium on Mathematical Logic and Foundations of Set Theory (SMLFST'70)*, Y. Bar-Hillel, Ed. North-Holland, 1970, pp. 1–23.
[11] D. E. Muller, A. Saoudi, and P. E. Schupp, "Alternating automata, the weak monadic theory of trees and its complexity," *Theor. Comput. Sci.*, vol. 97, no. 2, pp. 233–244, 1992.
[12] E. A. Emerson and C. S. Jutla, "Tree automata, μ-calculus and determinacy (extended abstract)," in *FOCS*, 1991, pp. 368–377.
[13] A. Mostowski, "Games with forbidden positions," University of Gdansk, Tech. Rep. 78, 1991.
[14] D. Niwinski, "On fixed-point clones (extended abstract)," in *ICALP*, ser. Lecture Notes in Computer Science, L. Kott, Ed., vol. 226. Springer, 1986, pp. 464–473.
[15] D. Park, "On the semantics of fair parallelism," in *Proceedings of the Abstract Software Specifications, 1979 Copenhagen Winter School*. London, UK, UK: Springer-Verlag, 1980, pp. 504–526. [Online]. Available: http://dl.acm.org/citation.cfm?id=647448.727229
[16] S. Safra, "On the complexity of ω-automata," in *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, ser. SFCS '88. Washington, DC, USA: IEEE Computer Society, 1988, pp. 319–327. [Online]. Available: http://dx.doi.org/10.1109/SFCS.1988.21948
[17] T. Wilke, "Alternating tree automata, parity games, and modal μ-calculus," *Bull. Soc. Math. Belg*, vol. 8, 2001.
[18] D. Janin and I. Walukiewicz, "Automata for the modal μ-calculus and related results," in *MFCS*, 1995, pp. 552–562.
[19] O. Kupferman and M. Y. Vardi, "$\Pi_2 \cap \Sigma_2 \equiv AFMC$," in *ICALP*, 2003, pp. 697–713.
[20] Y. Venema, "Lectures on the modal μ-calculus," 2008, lecture Notes, Institute for Logic, Language and Computation, University of Amsterdam.
[21] E. Grädel, W. Thomas, and T. Wilke, Eds., *Automata, Logics, and Infinite Games: A Guide to Current Research*, ser. Lecture Notes in Computer Science, vol. 2500. Springer, 2002.

Appendix A
The refined powerset construction

In this appendix we recall the refined powerset construction [9] [20] which plays an essential role in the two-sorted construction. Given an MSO-automaton $\mathbb{A}$, the powerset automaton $\mathbb{A}^\sharp$ will be based on *macro-states*, mimicking multiple runs of $\mathbb{A}$ in parallel. In order to give a correct notion of acceptance condition, we need to keep track of the structure of each run of $\mathbb{A}$ which is simulated by $\mathbb{A}^\sharp$ in parallel. For this reason, macro-state are not given as sets of states of $\mathbb{A}$, as one could expect, but as *binary relations* $Q \subseteq A \times A$ between them. Runs of $\mathbb{A}$ will then be recovered as *traces* through a run of $\mathbb{A}^\sharp$.

Definition 12. *Let A be a finite set of states and let $\rho \in (\wp(A \times A))^\omega$ be a $\wp(A \times A)-$stream.*

$$\rho \quad := \quad Q_0, Q_1, \ldots, Q_n, \ldots$$

A trace $\alpha \in A^\omega$ *through ρ is an $A-$stream such that $a_i Q_{i+1} a_{i+1}$ for all $i < \omega$.*

$$\alpha \quad := \quad a_0, a_1, \ldots, a_n, \ldots$$

Definition 13. *Let A be a finite set of states and $\Omega : A \to \omega$ a parity map. We say that a trace $\alpha \in A^\omega$ is* good *if the minimum parity occurring infinitely often along α is even, and* bad *otherwise. The set $NBT \subseteq (\wp(A \times A))^\omega$ is defined as*

$$NBT \;:=\; \{\rho \in (\wp(A \times A))^\omega \mid \textit{every trace through } \rho \textit{ is good}\}.$$

The last component we need to consider before giving the formal definition of $\mathbb{A}^\sharp$ is the transition function. For this purpose, first we state a normal form theorem for sentences associated to the transition function of $\mathbb{A}$. Using Ehrenfeucht-Fraïssé Games it is possible to show that every sentence $\varphi \in FO^+(A)$ can be rewritten into an equivalent disjunction of sentences in $BF^+(A)$.

Proposition 8 ([9] - Lemma 38, [21] - Lemma 16.23)**.** *Let $\varphi \in FO^+(A)$ be a sentence. There is a sentence $\varphi' \in Disj(BF^+(A))$ such that $\varphi \equiv \varphi'$.*

The next step is to perform a 'change of base' on the transition function of $\mathbb{A}$, passing from first-order sentences on signature A to first-order sentences on signature $\wp(A)$.

Definition 14 (Change of base)**.** *Fix a set A of unary predicates. Let $\varphi \in BF^+(A)$ be a sentence in basic form depending on sequences $B_1 \ldots B_k$ and $C_1 \ldots C_j$ of subsets of A. For each subset S in the sequence, we put $\tau^\wp_S(x) := S(x)$ if $S \neq \emptyset$ and $\tau^\wp_S(x) := \top$ otherwise. We denote with $\varphi^\wp$ the sentence given as follows:*

$$\varphi^\wp \;= \exists x_1 \ldots x_k \,\Big(\mathit{diff}(\bar{x}) \wedge \bigwedge_{1\leq i\leq k} \tau^\wp_{B_i}(x_i) \\ \wedge \forall z\;(\mathit{diff}(\bar{x}, z) \to \bigvee_{1\leq l\leq j} \tau^\wp_{C_l}(z))\Big).$$

In order to define the acceptance conditions of $\mathbb{A}^\sharp$ in terms of traces, we need to shift the same argument to the signature $\wp(A \times A)$ of binary relations on A. For this purpose we introduce an auxiliary translation, transforming first-order sentences on signature A into first-order sentences on signature $A \times A$.

Definition 15. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be an MSO-automaton. Fix $a \in A$ and $c \in C$. The sentence $\Delta^\star(a, c)$ is defined as*

$$\Delta^\star(a, c) \quad := \quad \Delta(a, c)[(a, b) \setminus b \mid b \in A],$$

where $\Delta(a, c)[(a, b) \setminus b \mid b \in A]$ denotes the sentence in $FO^+(A \times A)$ obtained by replacing each occurrence of an unary predicate $b \in A$ in $\Delta(a, c)$ with the unary predicate $(a, b) \in A \times A$. ◁

Finally we put together definition 14 and 15 to characterize a transition function ranging over sentences on signature $\wp(A \times A)$.

Definition 16. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be an MSO-automaton. Let $c \in C$ be a label and $Q \in \wp(A\times A)$ a binary relation on A. By proposition 8 there is a sentence $\Psi'_{Q,c} \in Disj(BF^+(A\times A))$ such that*

$$\bigwedge_{a\in Ran(Q)} \Delta^\star(a, c) \quad \equiv \quad \Psi'_{Q,c}.$$

By definition $\Psi'_{Q,c}$ is of the form $\bigvee_{1\leq i\leq k} \varphi_i$. We put $\Psi_{Q,c} := \bigvee_{1\leq i\leq k} (\varphi_i)^\wp$, where the translation $(-)^\wp$ is given as in definition 14.

Now we have all the ingredients to provide the definition of $\mathbb{A}^\sharp$.

Definition 17. *Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be an MSO-automaton. The automaton $\mathbb{A}^\sharp = \langle A^\sharp, a^\sharp_I, \Delta^\sharp, Acc\rangle$ is defined as follows.*

$$\begin{aligned} A^\sharp &:= \wp(A \times A) \\ a^\sharp_I &:= \{a_I, a_I\} \\ \Delta^\sharp(Q, c) &:= \Psi_{Q,c} \\ Acc &:= NBT \end{aligned}$$

Here $\Psi_{Q,c}$ is given according to proposition 16 and NBT is given according to definition 13. The automaton $\mathbb{A}^\sharp$ is called the refined powerset construct *over $\mathbb{A}$.*

Observe that by definition each value $\Delta^\sharp(Q, c)$ of the transition function of $\mathbb{A}^\sharp$ is an element of $Disj(FBF^+(A^\sharp))$, meaning that $\mathbb{A}^\sharp$ is a non-deterministic automaton.

Appendix B
Proofs of section III

In this appendix we give a more detailed proof of proposition 2, 3 and theorem 3.

Proposition 2. Let $\mathbb{A} = \langle A, a_I, \Delta, \Omega\rangle$ be a weak MSO-automaton and $\mathbb{A}^{2S}$ the two-sorted construct on $\mathbb{A}$. Then $\mathcal{L}(\mathbb{A}^{2S}) = \mathcal{L}(\mathbb{A})$.

Proof: **(⇒)** Let $\mathbb{T}$ be a tree which is accepted by $\mathbb{A}^{2S}$, with f^{2S} some winning strategy for player $\exists$ in the game $\mathcal{G}^{2S} = \mathcal{A}(\mathbb{A}^{2S}, \mathbb{T})@(a^\sharp_I, s_I)$. We want to define a strategy f

that is winning for $\exists$ in $\mathcal{G} = \mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$. For such purpose, we construct in stages an f-conform match π of $\mathcal{G}$, while playing an f^{2S}-conform shadow match π^{2S} of $\mathcal{G}^{2S}$. For each round z_i, we want to keep the following relation between the two matches:

1) either basic positions of the form $(Q, s) \in A^\sharp \times T$ and $(a, s) \in A \times T$ occur respectively in π^{2S} and π, with $a \in Ran(Q)$,
2) or the same basic position of the form $(a, s) \in A \times T$ occurs in both matches. (‡)

The key observation is that, because f^{2S} is winning, a basic position of the form $(Q, s) \in A^\sharp \times T$ can occur only for finitely many initial rounds $z_0, \ldots, z_n$ that are played in π^{2S}, whereas for all successive rounds $z_n, z_{n+1}, \ldots$ only basic positions of the form $(a, s) \in A \times T$ are encountered. Operationally, this means that in order to accept $\mathbb{T}$ the automaton $\mathbb{A}^{2S}$ eventually has to switch from the non-deterministic mode (corresponding to positions of type $A^\sharp \times T$) to the alternating mode (corresponding to positions of type $A \times T$). Indeed, if this was not the case then either $\exists$ would get stuck or the minimum parity occurring infinitely often would be odd, since all parities of states from $A^\sharp$ are equal to 1 according to the definition of $\mathbb{A}^{2S}$.

It follows that enforcing a relation between the two matches as in (‡) suffices to prove that the defined strategy f is winning for $\exists$ in π. For this purpose, consider any round z_i that is played in π and π^{2S}, respectively with basic positions $(a, s) \in A \times T$ and $(q, s) \in A^{2S} \times T$. In order to define the suggestion of f in π, we distinguish two cases.

1) First suppose that (q, s) is of the form $(Q, s) \in A^\sharp \times T$. By inductive hypothesis we assume that a is in $Ran(Q)$ - which is indeed the case at the initial round, where $R = a_I^\sharp = \{a_I, a_I\}$. Let $m_{Q,s} : A^{2S} \to \wp(R[s])$ be the marking suggested by f^{2S}, verifying the sentence $\Delta^{2S}(Q, V(s))$. We distinguish two further cases, depending on which disjunct of $\Delta^{2S}(Q, V(s))$ is made true by $m_{Q,s}$.
   a) If $(R[s], m_{Q,s}) \models \bigwedge_{b \in Ran(Q)} \Delta(b, V(s))$, then we let $\exists$ pick the marking $m_{Q,s}$.
   b) If $(R[s], m_{Q,s}) \models \Delta^\sharp(Q, V(s))$, we let $\exists$ pick a marking $m_{a,s} : A \to \wp(R[s])$ defined by putting
   $$m_{a,s}(b) := \bigcup_{b \in Ran(Q')} \{t \in R[s] \mid t \in m_{Q,s}(Q')\}.$$

   It can be readily checked that the suggested move is admissible for $\exists$ in π, i.e. it makes $\Delta(a, V(s))$ true in $R[s]$. In particular, for case (b) this follows by definition of $\Delta^\sharp$ in terms of Δ. In order to show that we can maintain the relation (‡) for another round, observe that case (a) corresponds to $\mathbb{A}^{2S}$ switching to the alternating mode, whereas case (b) corresponds to $\mathbb{A}^{2S}$ continuing in non-deterministic mode. In the former situation, any next position $(b, t) \in A \times T$ that can be picked by player $\forall$ in π is also available for $\forall$ in π^{2S}, and we end up in case (‡.2). In the latter situation, by definition of $m_{a,s}$, given the choice (b, t) of $\forall$, there is some $Q' \in A^\sharp$ such that b is in $Ran(Q')$ and (Q', t) is an available choice for $\forall$ in the shadow match π^{2S}. By letting π^{2S} advance at round z_{i+1} with such move, we are able to maintain (‡.1) also in z_{i+1}.
2) In the remaining case, inductively we are given with the same basic position $(a, s) \in A \times T$ both in π and in π^{2S}. The marking suggested by f^{2S} in π^{2S} verifies $\Delta^{2S}(a, s) = \Delta(a, s)$, so we let it be also the choice of $\exists$ in the match π. It is immediate that any next move of $\forall$ in π can be mirrored by the same move in π^{2S}, meaning that we are able to maintain the same position - whence the relation (‡) - also in the next round.

In both cases, the suggestion of strategy f was a legitimate move for $\exists$ maintaining the relation (‡) between the two matches for any next round z_{i+1}. It follows that f is a winning strategy for $\exists$ in $\mathcal{G}$.

($\Leftarrow$) Given a tree $\mathbb{T}$ that is accepted by $\mathbb{A}$, let f be a winning strategy f for $\exists$ in $\mathcal{G}$. The idea is to let $\exists$ play the very same strategy in the game $\mathcal{G}^{2S}$: this intuitively corresponds to $\mathbb{A}^{2S}$ immediately entering the alternating mode. For this purpose, suppose to initialize an f-conform shadow match π and a match π^{2S}, respectively from positions (a_I, s_I) and $(a_I^\sharp, s_I)$. The marking m suggested by f makes $\Delta(a_I, V(s_I))$ true. But this is just the formula $\bigwedge_{a \in Ran(a_I^\sharp)} \Delta(a, V(s_I))$, meaning that $(R[s_I], m) \models \Delta^{2S}(a_I^\sharp, V(s_I))$. It is straightforward to check that from this round on the two matches can be maintained with the same basic positions, implying that f is a winning strategy for $\exists$ in $\mathcal{G}^{2S}$. ■

Proposition 3. For each weak MSO-automaton $\mathbb{A}$ on alphabet $\wp(P \cup \{p\})$, we have that $\mathcal{L}(\exists_W p.\mathbb{A}) = \exists_W p.\mathcal{L}(\mathbb{A})$.

Proof: We fix $\mathbb{A} = \langle A, a_I, \Delta, \Omega \rangle$ and its two-sorted projection $\exists_W p.\mathbb{A} = \langle A^{2S}, a_I^{2S}, \widetilde{\Delta}, \Omega^{2S} \rangle$. The claim is that for any tree $\mathbb{T}$:

$$\exists_W p.\mathbb{A} \text{ accepts } \mathbb{T} \quad \text{iff} \quad \text{there is a noetherian } p\text{-variant } \mathbb{T}' \text{ of } \mathbb{T} \text{ such that } \mathbb{A} \text{ accepts } \mathbb{T}'.$$

($\Rightarrow$) Let $\mathbb{T}$ be a tree and $\tilde{f}$ a winning strategy for $\exists$ in $\tilde{\mathcal{G}} = \mathcal{A}(\exists_W p.\mathbb{A}, \mathbb{T})@(a_I^\sharp, s_I)$. Our goal is to provide a noetherian p-variant $\mathbb{T}'$ of the tree $\mathbb{T}$ and a winning strategy for $\exists$ in $\mathcal{G}^{2S} = \mathcal{A}(\mathbb{A}^{2S}, \mathbb{T}')@(a_I^\sharp, s_I)$. The proof will be completed by applying proposition 2.

For this purpose, the key idea is to exploit the fact that a winning strategy associated with a non-deterministic automaton can be assumed to be *functional*. By definition of the two-sorted automaton $\exists_W p.\mathbb{A}$, there is a prefix W of $\mathbb{T}$ where $\exists_W p.\mathbb{A}$ behaves as a non-deterministic automaton. Then we can assume that $\tilde{f}$ is a functional strategy on partial matches played in W: this provides enough information to label with p the nodes in W.

More formally, for each node $s \in W$, by functionality of $\tilde{f}$ we can fix a unique $q_s \in A^{2S}$ such that s is visited along $\tilde{f}$-conform matches of $\tilde{\mathcal{G}}$ only when associated with q_s in a position (q_s, s). Then we can also associate s with a marking $\widetilde{m}_s$, which is the suggestion of $\tilde{f}$ from position (q_s, s) - we let m_s be arbitrarily defined if s is not visited along any $\tilde{f}$-conform match. Since W is always traversed when $\mathbb{A}^{2S}$ is still in the non-deterministic mode, we can assume that $q_s = Q_s$ for some $Q_s \in A^\sharp$. By these observation we define a subset X_p of $\mathbb{T}$ as follows.

$$X_p \;:=\; \{s \in W \mid (R[s], \widetilde{m}_s)) \models \Delta^{2S}(Q_s, V(s) \cup \{p\})\} \quad (5)$$

We define $\mathbb{T}'$ as the p-variant of $\mathbb{T}$ where the nodes labeled with p are exactly the members of X_p, that is, $\mathbb{T}' = \mathbb{T}[p \mapsto X_p]$. Since W is a prefix, we are guaranteed that $\mathbb{T}'$ is a noetherian p-variant. What remains to show is that player $\exists$ has a winning strategy in the game $\mathcal{G}^{2S}$. In fact we claim that the very same strategy $\tilde{f}$ is also winning in $\mathcal{G}^{2S}$. In order to see that, let us construct in stages an $\tilde{f}$-conform match π^{2S} of $\mathcal{G}^{2S}$ and an $\tilde{f}$-conform shadow match $\tilde{\pi}$ of $\tilde{\mathcal{G}}$. The inductive hypothesis we want to bring from one round to the next is that the same basic position occurs in both matches, as this suffices to prove that $\tilde{f}$ is winning for $\exists$ in $\mathcal{G}^{2S}$.

First we consider a basic position $(q, s) \in A^{2S} \times W$ occurring in the two matches at the same round z, while playing on the prefix W. By definition of W we know that the automaton $\exists_W p.\mathbb{A}$ is in non-deterministic mode at round z, meaning that (q, s) is in fact of the form $(Q, s) \in A^\sharp \times W$. By assumption $\tilde{f}$ provides a marking $\widetilde{m}_s$ that makes $\widetilde{\Delta}(Q, V(s))$ true in $R[s]$. By definition of $\widetilde{\Delta}$ this means that $\widetilde{m}_s$ verifies either $\Delta^{2S}(Q, V(s))$ or $\Delta^{2S}(Q, V(s) \cup \{p\})$. Now, the match π^{2S} is played on the p-variant $\mathbb{T}'$, where the labeling $V'(s)$ is decided by the membership of s to X_p. According to (5), if $\widetilde{m}_s$ verifies $\Delta^{2S}(Q, V(s) \cup \{p\})$ then s is in X_p, meaning that it is labeled with p in $\mathbb{T}'$, i.e. $V'(s) = V(s) \cup \{p\}$. Therefore $\widetilde{m}_s$ also verifies $\Delta^{2S}(Q, V'(s))$ and it is a legitimate move for $\exists$ in match π^{2S}. In the remaining case, $\widetilde{m}_s$ verifies $\Delta^{2S}(Q, V(s))$ but falsifies $\Delta^{2S}(Q, V(s) \cup \{p\})$, implying by definition that s is not in X_p. This means that s is not labeled with p in $\mathbb{T}'$, i.e. $V'(s) = V(s)$. Thus again $\widetilde{m}_s$ verifies $\Delta^{2S}(Q, V'(s))$ and it is a legitimate move for $\exists$ in match π^{2S}.

It remains to consider the case of a basic position $(q, s) \in A^{2S} \times T$ occurring in the two matches at the same round, while playing outside the prefix W. By definition of W, the automaton $\exists_W p.\mathbb{A}$ is in alternating mode while playing round z, meaning that the basic position that we are considering is of the form $(a, s) \in A \times T$. By definition $\widetilde{\Delta}(a, V(s))$ is just $\Delta^{2S}(a, V(s))$. Since the node s is assumed to be outside W, we also know that it is not labeled with p in the p-variant $\mathbb{T}'$, meaning that $V'(s) = V(s)$. Thus $\tilde{f}$ makes $\Delta^{2S}(a, V'(s)) = \Delta^{2S}(a, V(s))$ true in $R[s]$ and its suggestion is a legitimate move for $\exists$ in match π^{2S}. In order to conclude the proof, observe that for all positions that we consider the same marking is suggested to $\exists$ in both games: this means that any next position that is picked by player $\forall$ in π^{2S} is also available for $\forall$ in the shadow match $\tilde{\pi}$.

($\Leftarrow$) Let $\mathbb{T}$ be a tree and $\mathbb{T}'$ a noetherian p-variant of $\mathbb{T}$ that is accepted by $\mathbb{A}$, with f a winning strategy for $\exists$ in $\mathcal{G} = \mathcal{A}(\mathbb{A}, \mathbb{T}')@(a_I, s_I)$. As usual we want to provide a winning strategy $\tilde{f}$ for $\exists$ in $\tilde{\mathcal{G}} = \mathcal{A}(\exists_W p.\mathbb{A}, \mathbb{T})@(a_I^\sharp, s_I)$, by constructing in stages an $\tilde{f}$-conform match of $\tilde{\mathcal{G}}$. The argument will be similar to the one showing the equivalence between $\mathbb{A}$ and $\mathbb{A}^{2S}$, the automaton $\exists_W p.\mathbb{A}$ being almost defined as $\mathbb{A}^{2S}$. The difference is that now we have to take care of nodes that are labeled with p in $\mathbb{T}'$ but not in $\mathbb{T}$. For this reason, we cannot simply let $\exists_W p.\mathbb{A}$ immediately enter the alternating mode, because the definition of $\widetilde{\Delta}$ only takes the letter p into account while $\exists_W p.\mathbb{A}$ runs in non-deterministic mode. Instead, the idea is to keep $\exists_W p.\mathbb{A}$ in non-deterministic mode until a node $s^\star$ is encountered, such that the induced subtree $\mathbb{T}'_{s^\star}$ of $\mathbb{T}'$ does not contain any node labeled with p (in this case we say that $\mathbb{T}'_{s^\star}$ is *p-free*). The key observation is that such condition is verified in finitely many rounds, $\mathbb{T}'$ being a *noetherian* p-variant of $\mathbb{T}$. Once the two matches are played in the subtree $\mathbb{T}'_{s^\star}$, no further node labeled with p is encountered and we make $\exists_W p.\mathbb{A}$ enter the alternating mode, behaving as $\mathbb{A}$ from that stage on.

We now proceed with the definition of $\tilde{f}$, formalizing the argument given above. In its first segment the $\tilde{f}$-conform match $\tilde{\pi}$ will consists of positions of the form $(Q, s) \in A^\sharp \times T$, corresponding to $\exists_W p.\mathbb{A}$ being in the non-deterministic mode. The intuition - actually the core idea of the refined powerset construction - is that (Q, s) is a 'macro-position', representing the situation in which $\exists$ has to play from basic positions $\{(a, s)\}_{a \in Ran(Q)}$. For this reason, instead of a single shadow match, we want to maintain a *bundle* $\mathcal{B}$ of f-conform shadow matches $\{\pi_a\}_{a \in Ran(Q)}$, defining $\tilde{f}$ from position (Q, s) in terms of the suggestion of f from each position (a, s) with $a \in Ran(Q)$. More formally, we are going to define $\tilde{f}$ in such a way that, for any round z_i, we are given with our match $\tilde{\pi}$ and a bundle $\mathcal{B}_i$ of f-conform shadow matches of $\mathcal{G}$, with the following relation enforced between them.

1) If the current (i.e. at round z_i) basic position in $\tilde{\pi}$ is of the form $(Q, s) \in A^\sharp \times T$, then for each $a \in Ran(Q)$ there is an f-conform (partial) shadow match π_a in $\mathcal{B}_i$ with current basic position $(a, s) \in A \times T$. Either $\mathbb{T}'_s$ is not p-free or s has some sibling t such that $\mathbb{T}'_t$ is not p-free.
2) Otherwise, the current basic position in $\tilde{\pi}$ is of the form $(a, s) \in A \times T$, with s a node such that $\mathbb{T}'_s$ is p-free. The bundle $\mathcal{B}_i$ only consists of a single f-conform match π_a whose current basic position is also (a, s).

(‡)

We initialize $\tilde{\pi}$ from position $(a_I^\sharp, s) \in A^\sharp \times T$ and the bundle $\mathcal{B}$ as $\mathcal{B}_0 = \{\pi_{a_I}\}$, with π_{a_I} the partial f-conform match consisting only of the position $(a_I, s) \in A \times T$. In this way the relation (‡.1) holds at the initial stage of the construction. Inductively, suppose that we are given at round z_i with a position $(q, s) \in A^{2S} \times T$ in π^{2S} and a bundle $\mathcal{B}_i$ as in (‡).

We distinguish two cases.

1) If (q, s) is of the form $(Q, s) \in A^\sharp \times T$, by inductive hypothesis we are given with f-conform shadow matches $\{\pi_a\}_{a \in Ran(Q)}$ in $\mathcal{B}_i$. For each match π_a in the bundle, we are provided with a marking $m_{a,s} : A \to \wp(R[s])$ making $\Delta(a, V'(s))$ true. Then we further distinguish the following two cases.

   a) If $\mathbb{T}'_s$ is not p-free, we keep $\exists_W p.\mathbb{A}$ in the non-deterministic mode. This means that, for each $Q' \in A^\sharp$ we put

   $$\widetilde{m}_{Q,s}(Q') := \bigcap_{\substack{(a,b) \in Q', \\ a \in Ran(Q)}} \{t \in R[s] \mid t \in m_{a,s}(b)\}.$$

   By definition of $\Delta^\sharp$ it is routine to check that $\widetilde{m}$ makes $\Delta^\sharp(Q, V'(s))$ - and then also its logical consequence $\Delta^{2S}(Q, V'(s))$ - true in $R[s]$. In order to be a legitimate move for $\exists$ in $\tilde{\pi}$, the marking $\widetilde{m}$ should make $\widetilde{\Delta}(Q, V(s))$ true: this is the case, for $\Delta^{2S}(Q, V'(s))$ is either equal to $\Delta^{2S}(Q, V(s))$, if $p \notin V'(s)$, or to $\Delta^{2S}(Q, V(s) \cup \{p\})$ otherwise. In order to check that we can maintain the relation (‡), let $(Q', t) \in A^\sharp \times T$ be any next position picked by $\forall$ in $\tilde{\pi}$ at round z_{i+1}. The new bundle $\mathcal{B}_{i+1}$ is given in terms of the bundle $\mathcal{B}_i$: for each $\pi_a \in \mathcal{B}_i$ with $a \in Ran(Q)$, we look if for some $b \in Ran(Q')$ the position (b, t) is a legitimate move for $\forall$ at round z_{i+1}; if so, then we bring π_a to round z_{i+1} at position (b, t) and put the resulting (partial) shadow match π_b in $\mathcal{B}_{i+1}$. Observe that, if $\forall$ is able to pick such position (Q', t) in $\tilde{\pi}$, then by definition of $\widetilde{m}$ the new bundle $\mathcal{B}_{i+1}$ is non-empty and consists of an f-conform (partial) shadow match π_b for each $b \in Ran(Q')$. In this way we are able to keep the relation (‡.1) at round z_{i+1}.

   b) Otherwise, $\mathbb{T}'_s$ is p-free and we let $\exists_W p.\mathbb{A}$ enter the alternating mode, by guessing some $a \in Ran(Q)$ from which $\exists_W p.\mathbb{A}$ starts to behave as $\mathbb{A}$. This amounts to define the suggestion of $\tilde{f}$ as the marking $\widetilde{m}_{Q,s}$ assigning to each node $t \in R[s]$ all states in $\bigcup_{a \in Ran(Q)} \{b \in A \mid t \in m_{a,s}(b)\}$. It can be checked that this marking makes $\bigwedge_{a \in Ran(Q)} \Delta(a, V'(s))$ - and then also its logical consequence $\Delta^{2S}(Q, V'(s))$ - true in $R[s]$, . As observed for the case $(1.a)$, it follows that $\widetilde{m}$ also makes $\widetilde{\Delta}(Q, V(s))$ true, whence it is a legitimate choice for $\exists$ in $\tilde{\pi}$. Any next basic position $(b, t) \in A \times T$ picked by $\forall$ in $\tilde{\pi}$ is an available move for $\forall$ in some shadow match π_a in the bundle $\mathcal{B}_i$. We dismiss the bundle - i.e. make it a singleton - and bring only π_a to the next round in the same position (b, t).

2) In the remaining case, (q, s) is of the form $(a, s) \in A \times T$ and by inductive hypothesis we are given with a bundle $\mathcal{B}_i$ consisting of a single f-conform (partial) shadow match π_a at the same position (a, s). Since by assumption s is also p-free, we have that $V'(s) = V(s)$. It follows that $\widetilde{\Delta}(a, V(t))$ is just $\Delta(a, V(t)) = \Delta(a, V'(t))$ and the same marking suggested by f in π_a is a legitimate choice for $\exists$ in $\tilde{\pi}$. By letting $\exists$ choose such marking, it follows that any next move made by $\forall$ in $\tilde{\pi}$ can be mirrored by $\forall$ in the shadow match π_a.

As explained above, since $\mathbb{T}'$ is a noetherian p-variant, then (‡.1) holds for finitely many stages of construction of $\tilde{\pi}$, whereas (‡.2) holds for all the remaining stages, by construction of $\tilde{f}$. It follows that this strategy is winning for $\exists$ in $\tilde{G}$. ■

Theorem 3. For every $\varphi \in WFMSO$, there is an effectively constructible weak MSO-automaton $\mathbb{A}_\varphi$ such that on tree languages $\|\varphi\| = \mathcal{L}(\mathbb{A}_\varphi)$.

Proof: The proof is by induction on φ. Atomic and boolean cases are handled with the same argument supplied for MSO and MSO-automata [9], reflecting the fact that the semantics of MSO and $WFMSO$ coincides on these cases. For the case $\varphi = \exists p.\psi$, by inductive hypothesis we are given with a weak MSO-automaton $\mathbb{A}_\psi$ that is equivalent to ψ. Let $\exists_W p.\mathbb{A}_\psi$ be the weak MSO-automaton obtained from $\mathbb{A}_\psi$ as in definition 5. The following derivation shows that $\exists_W p.\mathbb{A}_\psi$ is equivalent to $\exists p.\psi$.

$$\begin{array}{lcl}
\exists_W p.\mathbb{A}_\psi \text{ accepts } \mathbb{T}. & \overset{(prop.3)}{\iff} & \text{There is } X_p \in N(\mathbb{T}) \text{ s.t.} \\
 & & \mathbb{A}_\psi \text{ accepts } \mathbb{T}[p \mapsto X_p]. \\
 & \overset{(ind.hyp.)}{\iff} & \text{There is } X_p \in N(\mathbb{T}) \text{ s.t.} \\
 & & \mathbb{T}[p \mapsto X_p] \models \psi. \\
 & \overset{(WFMSO)}{\iff} & \mathbb{T} \models \exists p.\psi.
\end{array}$$

■

Appendix C
Proofs of section IV

In this appendix we give a more detailed proof of proposition 4 and theorem 4.

Proposition 4. Let $\mathbb{A}$ be a weak MSO-automaton and $\mathbb{A}^B$ the Büchi powerset construct over $\mathbb{A}$. We have that $\mathcal{L}(\mathbb{A}) = \mathcal{L}(\mathbb{A}^B)$.

Proof: By proposition 1 it suffices to prove that $\mathbb{A}^\sharp$ is equivalent to $\mathbb{A}^B$. For this purpose, we fix a tree $\mathbb{T}$.

$(\Rightarrow)$ Let $f^\sharp$ be a winning strategy for $\exists$ in $\mathcal{G}^\sharp = \mathcal{A}(\mathbb{A}^\sharp, \mathbb{T})@(a_I^\sharp, s_I)$. We claim that $f^\sharp$ is also winning for $\exists$ in $\mathcal{G}^B = \mathcal{A}(\mathbb{A}^B, \mathbb{T})@(a_I^\sharp, s_I)$. In order to show that, we first observe that any $f^\sharp$-conform match of $\mathcal{G}^B$ has the same basic positions of an $f^\sharp$-conform match of $\mathcal{G}^\sharp$, being $\mathbb{A}^B$ and $\mathbb{A}^\sharp$ based on the same set of states, the same initial state and the same transition function. What remains to show is that any infinite $f^\sharp$-conform match π^B of $\mathcal{G}^B$ is winning for $\exists$. For this purpose, let $\rho = a_I^\sharp, Q_1, \ldots, Q_n, \ldots$ be the sequence of states of $\mathcal{G}^B$ visited along the play. As we observed, there is an $f^\sharp$-conform infinite match $\pi^\sharp$ of $\mathcal{G}^\sharp$ with the same sequence of states $\rho = a_I^\sharp, Q_1, \ldots, Q_n, \ldots$ visited along the play. Since

$f^\sharp$ is winning for $\exists$ in $\mathcal{G}^\sharp$, all traces through ρ are *good* according to definition 13. This means that every trace through ρ corresponds to a match of $\mathcal{G} = \mathcal{A}(\mathbb{A}, \mathbb{T})@(a_I, s_I)$ that is won by $\exists$. Let π be one such match associated with a trace through ρ. Since $\mathbb{A}$ is weak, by remark 1 we can assume that $\Omega : A \to \omega$ ranges over $\{0, 1\}$ and then *exactly one* parity between 0 and 1 occurs infinitely often along π. By assumption $\exists$ wins π, meaning that such parity is 0. Since π corresponds to an arbitrary trace through ρ, it follows that after finitely many steps only positions with parity 0 occur on each trace through ρ. This means that, for some $k < \omega$, there is a state $Q_k \in A^\sharp$ occurring in ρ after which all the states occurring in ρ belong to the set

$$F_\Omega = \{Q \in A^\sharp | \; \Omega(a) = 0 \text{ for all } a \in Ran(Q)\}.$$

Since F_Ω is finite, there is at least one state in F_Ω occurring infinitely often along ρ. It follows that player $\exists$ wins the match π^B.

($\Leftarrow$) The argument showing this direction is entirely analogous to the one provided for direction ($\Rightarrow$), so we just sketch the main steps. Given a winning strategy f^B for $\exists$ in $\mathcal{G}^B$, the same strategy f^B can be shown to be winning for $\exists$ in $\mathcal{G}^\sharp$. The key observation is that, since f^B is winning for $\exists$ in $\mathcal{G}^B$, for each infinite f^B-conform match $\pi^\sharp$ of $\mathcal{G}^\sharp$, there is some $Q \in F_\Omega$ occurring infinitely often along the play. Let ρ be the sequence of states $a_I^\sharp, Q_1, \dots, Q_n, \dots$ visited along the play in $\pi^\sharp$. Every trace α through ρ encounters some $a \in Ran(Q)$ infinitely often. By definition of F_Ω, this means that basic positions with parity 0 occur infinitely often in the match π_α of $\mathcal{G}$ associated with α. It follows that π_α is won by $\exists$. Since α was an arbitrary trace through ρ, then every trace through ρ is *good* and $\exists$ also wins $\pi^\sharp$. ■

Definition 18. *Let $B = \{b_1, \dots, b_k\}$ and $P = \{p_1, \dots, p_j\}$ be two finite collection of set variables, representing respectively the states of an NDB automaton and the propositional letters forming the labels of a C-labeled tree. In table III we fix some abbreviations for WFMSO formulae, expressing concepts which are easily seen to be definable in this logic. All valuations (for which we use the notation $\|\cdot\|$) are referred to a fixed tree $\mathbb{T}$. Observe that for each $p_i \in P$ the set $\|p_i\|$ is determined by the labeling function $V : T \to C$ of $\mathbb{T}$.*

Formula	Meaning
$Pfix_z(p)$	The set $\|p\|$ is a prefix of the subtree $\mathbb{T}.\|z\|$.
$Root(x)$	The node $\|x\|$ is the root of $\mathbb{T}$.
$Front(p)$	The set $\|p\|$ is a frontier of $\mathbb{T}$.
$State_{a,B}(x)$	a is the only set variable in B such that the node $\|x\|$ is in $\|a\|$ (we say that *the state a marks $\|x\|$*).
$Part_B(p)$	Each node in the set $\|p\|$ is marked with a unique $a \in B$.
$Trans_{B,C}(p)$	For each node $\|x\| \in \|p\|$, state $a \in B$, label $c \in C$, if a marks $\|x\|$ and $\|x\|$ is labeled with c then $\Delta(a, c)$ holds in $R[\|x\|]$ (this latter condition is rendered by relativizing all quantifiers of $\Delta(a, c)$ to elements of $R[\|x\|]$).

Table III
Abbreviations for *WFMSO* formulae

We also define $Surv_{B,C}(p)$ as $Part_B(p) \wedge Trans_{B,C}(p)$, where $Part_B(p)$ and $Trans_{B,C}(p)$ are given as in table III. Intuitively, if $Surv_{B,C}(p)$ holds then player $\exists$ is guaranteed to have a legitimate move available from any node $\|x\|$ in $\|p\|$, assigning exactly one state to each $t \in R[\|x\|]$.

Definition 19. *Let $\mathbb{B}$ and $\overline{\mathbb{B}}$ be NDB-automata, with $B = \{b_1, \dots, b_k\}$ and $F \subseteq B$ respectively the set of states and of accepting states of $\mathbb{B}$. For each $b \in B$, we define by induction a sequence of formulae $K_i^b(x)$. Put $K_0^b(x) := \top$. The formula $K_{i+1}^b(x)$ is given as follows:*

$$\begin{aligned} K_{i+1}^b(x) \;:=\; & \forall p\, \exists p'\, \exists b_1 \dots \exists b_k \Bigg(Pfix_x(p) \to \\ & \Bigg(p \subseteq p' \wedge Pfix_x(p') \wedge Surv_{B,C}(p') \wedge \\ & State_{b,B}(x) \wedge \Bigg(\forall y \; \Big(y \in Front(p') \to \\ & \Big(\bigvee_{b' \in F} (State_{b',B}(y) \wedge K_i^{b'}(y)) \Big)\Big)\Bigg)\Bigg)\Bigg). \end{aligned}$$

Let m be the product of the cardinalities of the carriers of $\mathbb{B}$ and $\overline{\mathbb{B}}$. The formula $\varphi_{\mathbb{B},\overline{\mathbb{B}}} \in WFMSO$ is defined by putting

$$\varphi_{\mathbb{B},\overline{\mathbb{B}}} \;:=\; \exists y \; (Root(y) \wedge K_{m+1}^{b_I}(y)).$$

Observe that, for any $k < \omega$, $K_k^b(x)$ is a formula of *WFMSO*. We refer to section IV for an intuitive reading of the semantics of $K_{i+1}^b(x)$.

Our next goal is to show the main result of section IV.

Theorem 4. For any weak *MSO*-automaton $\mathbb{A}$ there is a formula $\varphi \in WFMSO$ such that over tree languages $\|\varphi\| = \mathcal{L}(\mathbb{A})$.

In this aim, we fix the following terminology concerning a winning strategy f for $\exists$ in some acceptance game $\mathcal{G}$: a basic position (a, s) is *f-admissible* if there is some f-conform match of $\mathcal{G}$ where (a, s) occurs.

In the next we fill in the details of the proof sketch provided in section IV. As stated there, we can assume to work with *NDB* automata $\mathbb{B}$ and $\overline{\mathbb{B}}$, equivalent respectively to $\mathbb{A}$ and the weak *MSO*-automaton $\overline{\mathbb{A}}$ recognizing the complement $\overline{\mathcal{L}(\mathbb{A})}$ of the tree language $\mathcal{L}(\mathbb{A})$. Then theorem 4 is an immediate consequence of the following proposition.

Proposition 9. *Let $\mathbb{B} = \langle B, b_I, \Delta, F\rangle$ and $\overline{\mathbb{B}} = \langle \overline{B}, \overline{b}_I, \overline{\Delta}, \overline{F}\rangle$ be NDB automata such that $\mathcal{L}(\overline{\mathbb{B}}) = \overline{\mathcal{L}(\mathbb{B})}$. Let $\varphi_{\mathbb{B},\overline{\mathbb{B}}} \in WFMSO$ be given in terms of $\mathbb{B}$ and $\overline{\mathbb{B}}$ according to definition 19. Over tree languages, we have that $\mathcal{L}(\mathbb{B}) = \|\varphi_{\mathbb{B},\overline{\mathbb{B}}}\|$.*

Proof: ($\Rightarrow$) Let $\mathbb{T}$ be a tree and f be a winning strategy for $\exists$ in $\mathcal{G} = \mathcal{A}(\mathbb{B}, \mathbb{T})@(b_I, s_I)$. Since $\mathbb{B}$ is non-deterministic, we can assume f to be functional. Since f is winning then we are provided with an ω-accepting sequence $(E_i)_{i<\omega}$ for f over $\mathbb{B}$ and $\mathbb{T}$, according to proposition 5. Our goal is to show that $\mathbb{T} \models \varphi_{\mathbb{B},\overline{\mathbb{B}}}$. In fact, it suffices to show the following statement.

Claim 1. *For each $i < \omega$, for each $(b, s) \in B \times T$, if (b, s) is a winning position for $\exists$ in $\mathcal{G}$, then $\mathbb{T} \models K_i^b(x)$, with $\|x\| = s$.*

Proof of Claim 1: We proceed by induction on i. Since $K_0^b(x) = \top$, the base case is trivial. Inductively, let (b, s) be a winning position for $\exists$ in $\mathcal{G}$. We put $\|x\| = s$ and we claim that $\mathbb{T} \models K_{i+1}^b(x)$. Following the syntactic shape of $K_{i+1}^b(x)$, we let $\|p\|$ be an arbitrary prefix E of $\mathbb{T}.s$. By definition of the sequence $(E_i)_{i<\omega}$, for each $i < \omega$ we have that $Ft(E_i) < Ft(E_{i+1})$, implying that there is some prefix E_n in the sequence such that $E \subseteq E_n$. We pick $E_n \cap T.s$ as the witness for the set-variable p' in $K_{i+1}^b(x)$.

We still need to provide witnesses for set-variables $b_1, \ldots, b_k$ occurring in $K_{i+1}^b(x)$. The idea is to let them be suggested by the strategy f. Since f is functional, any node $s \in \mathbb{T}$ is associated with a unique $b_s \in B$ and a unique f-admissible basic position (b_s, s). For each b_j in $\{b_1, \ldots, b_k\}$, we define its valuation by putting

$$\|b_j\| \quad := \quad \{s \in (E_n \cap T.s) \mid b_j = b_s\}. \tag{6}$$

Since $E_n \cap T.s$ is well-founded then $\|b_j\|$ is noetherian, so that it is a legitimate witness for b_j according to the semantics of WFMSO.

The subformula $Surv_{B,C}(p')$ of $K_{i+1}^b(x)$ holds because the strategy f is assumed to be functional and winning for $\exists$, so in particular it is functional and surviving for $\exists$ in $E_n \cap T.s = \|p'\|$. Concerning the subformula $State_{b,B}(x)$, by assumption (b, s) is a winning position for $\exists$. This means that b is the unique set-variable marking $s = \|x\|$ according to (6), so that $State_{b,B}(x)$ holds. It remains to show the statement

$$\forall y \ (y \in Front(p') \rightarrow (\bigvee_{b \in F} State_{b,B}(y) \wedge K_i^b(y))). \tag{7}$$

For this purpose, let $\|y\|$ be some node on the frontier of $E_n = \|p'\|$. By (6) and the fact that f is functional, there is a unique set-variable $\|b\|$ marking $\|y\|$, such that $(b, \|y\|)$ is f-admissible. Therefore $(b, \|y\|)$ is a winning position for $\exists$ in $\mathcal{G}$, and $K_i^b(y)$ holds by inductive hypothesis. The fact that b is in F follows from properties of the frontier of E_n as in definition 7. ■

By applying claim 1 to the winning position (b_I, s_I) we have that $\mathbb{T} \models K_n^{b_I}(x)$ for each $n < \omega$, with x witnessed by s_I. Then in particular $\mathbb{T} \models \exists x \ (Root(x) \wedge K_{m+1}^{b_I}(x))$. This completes the proof of direction $(\Rightarrow)$.

$(\Leftarrow)$ Let $\mathbb{T}$ be a tree where $\varphi_{\mathbb{B},\overline{\mathbb{B}}}$ is true. We need to show that $\mathbb{T}$ is accepted by $\mathbb{B}$.

The idea of the proof is as follows. Suppose by way of contradiction that $\mathbb{B}$ does not accept $\mathbb{T}$. Then the tree $\mathbb{T}$ is accepted by $\overline{\mathbb{B}}$. Let $\overline{f}$ be a functional winning strategy for $\exists$ in $\mathcal{A}(\overline{\mathbb{B}}, \mathbb{T})$. Suppose that we can prove from the previous assumptions the existence of an m-trap for $\mathbb{B}$ and $\overline{\mathbb{B}}$. Then by proposition 6 we have that $L(\mathbb{B}) \cap L(\overline{\mathbb{B}}) \neq \emptyset$, contradicting the fact that $L(\overline{\mathbb{B}}) = \overline{L(\mathbb{B})}$.

In order to complete the proof of direction $(\Leftarrow)$, it remains to verify the following claim.

Claim 2. *There exists an m-trap for $\mathbb{B}$ and $\overline{\mathbb{B}}$.*

Proof of Claim 2: By definition 8, we have to to provide the following components:

1) a strictly increasing sequence $(E_i)_{i \leq m}$ of prefixes of $\mathbb{T}$, with $E_0 = \{s_I\}$;
2) a strategy f^B for $\exists$ in $\mathcal{G} = \mathcal{A}(\mathbb{B}, \mathbb{T})@(b_I, s_I)$ which is surviving for $\exists$ in E_m;
3) a strategy $f^{\overline{B}}$ for $\exists$ in $\overline{\mathcal{G}} = \mathcal{A}(\overline{\mathbb{B}}, \mathbb{T})@(\overline{b}_I, s_I)$ which is surviving for $\exists$ in E_m;
4) an m-accepting sequence $(G_i^B)_{i \leq m}$ for f^B over $\mathbb{B}$ and $\mathbb{T}$;
5) an m-accepting sequence $(G_i^{\overline{B}})_{i \leq m}$ for $f^{\overline{B}}$ over $\overline{\mathbb{B}}$ and $\mathbb{T}$.

Moreover, $(E_i)_{i \leq m}$, $(G_i^B)_{i \leq m}$ and $(G_i^{\overline{B}})_{i \leq m}$ have to present the interleaving behavior described in definition 8.

We put the strategy $\overline{f}$ as witness for $f^{\overline{B}}$. By assumption $\overline{f}$ is a winning strategy for $\exists$ in $\overline{\mathcal{G}}$. Then, by proposition 5, we are also given with an ω-accepting sequence $(E_i^{\overline{f}})_{i<\omega}$ for $\overline{f}$ over $\overline{\mathbb{B}}$ and $\mathbb{T}$.

It remains to define the other components of the m-trap, which is what we do next. The idea is to define the surviving strategy f^B, the sequences $(E_i)_{i \leq m}$ and $(G_i^B)_{i \leq m}$ by using the assumption that $\mathbb{T} \models \varphi_{\mathbb{B},\overline{\mathbb{B}}}$. The last component, namely the sequence $(G_i^{\overline{B}})_{i \leq m}$, will be defined from $(E_i^{\overline{f}})_{i<\omega}$.

The construction of the strategy f^B and the sequences $(E_i)_{i \leq m}$, $(G_i^B)_{i \leq m}$ and $(G_i^{\overline{B}})_{i \leq m}$ proceeds in stages, by induction on $i \leq m$. In particular, f^B will be defined as the last element f_m^B in a sequence of strategies $(f_i^B)_{i \leq m}$.

Given $i \leq m$, the inductive hypothesis that we want to maintain along the construction can be expressed as follows.

(‡)

1) If $i < m$ then $Ft(E_i) \leq Ft(G_i^B) < Ft(E_{i+1})$. Otherwise $i = m$ and $Ft(E_i) \leq Ft(G_i^B)$.
2) If $i < m$ then $Ft(E_i) \leq Ft(G_i^{\overline{B}}) < Ft(E_{i+1})$. Otherwise $i = m$ and $Ft(E_i) \leq Ft(G_i^{\overline{B}})$.
3) The sets E_i, G_i^B $G_i^{\overline{B}}$ are prefixes of $\mathbb{T}$.
4) The function f_i^B is a strategy $\exists$ in $\mathcal{G}$ which is functional and surviving in G_i^B. If $i \geq 1$, then f_i^B extends f_{i-1}^B.
5) For each node $s \in Ft(G_i^{\overline{B}})$, there is a unique $\overline{b}_s \in \overline{B}$ such that the position $(\overline{b}_s, s)$ is $\overline{f}$-admissible; in addition, $\overline{b}_s$ is in $\overline{F}$.
6) For each node $s \in Ft(G_i^B)$, there is a unique $b_s \in B$ such that the formula $K_{m-i}^{b_s}(x)$ holds for $s = \|x\|$. The position (b_s, s) is f_i^B-admissible; in addition, b_s is in F.

Let us first show why the different components form an m-trap if condition (‡) can be maintained. By (‡.4) the strategy $f^B = f_m^B$ for $\exists$ in $\mathcal{G}$ would be functional and surviving in G_m^B. By (‡.1) we have that $Ft(E_m) \leq Ft(G_m^B)$, meaning that f^B is also surviving in E_m, as requested by point 1 of the definition of m-trap (definition 8). For $f^{\overline{B}} = \overline{f}$, we know by assumption that $\overline{f}$ is functional and winning for $\exists$ in $\overline{\mathcal{G}}$. Since E_m is a subset of T, then $f^{\overline{B}}$ is also surviving in E_m, as requested by point 2 of definition 8.

For points 3 and 4 of definition 8, we have to check that $(G_i^B)_{i\leq m}$ and $(G_i^{\overline{B}})_{i\leq m}$ are m-accepting sequences respectively for f^B and $f^{\overline{B}}$. For this purpose, there are three conditions to check according to the definition of accepting sequence (definition 7). The first condition is that $(G_i^B)_{i\leq m}$ and $(G_i^{\overline{B}})_{i\leq m}$ are sequences of prefixes, which is given by (‡.3). The second condition, on the relation between frontiers of each G_i^B, $G_i^{\overline{B}}$ and E_i, is given by (‡.1) and (‡.2). Concerning the third condition of definition 7, for each $i \leq m$, the requirements on $Ft(G_i^B)$ and $Ft(G_i^{\overline{B}})$ are fulfilled by (‡.5) and (‡.6).

The last two points of definition 8, concerning the interleaving of the frontiers of $(E_i)_{i\leq m}$, $(G_i^B)_{i\leq m}$ and $(G_i^{\overline{B}})_{i\leq m}$, just correspond to (‡.1) and (‡.2). Therefore what we obtain is indeed an m-trap, provided that we are able to maintain condition (‡).

Now we proceed with the inductive construction. For the base case, let $E_0 := \{s_I\}$. We define the first element $G_0^{\overline{B}}$ in the sequence $(G_i^{\overline{B}})_{i\leq m}$ as the smallest prefix in the sequence $(E_i^{\overline{f}})_{i<\omega}$ such that $E_0 \subseteq G_0^{\overline{B}}$, that is simply $E_0^{\overline{f}}$ because $(E_i^{\overline{f}})_{i<\omega}$ is monotone.

In order to define G_0^B, we observe that the unique witness for x in $\exists x\ (Root(x) \wedge K_{m+1}^{b_I}(x))$ must be s_I. Then, by putting E_0 as the witness of the variable p in $K_{m+1}^{b_I}(x)$, we are provided with a prefix G_0^B witnessing the variable p' in $K_{m+1}^{b_I}(x)$. We let such G_0^B be the first element in the sequence $(G_i^B)_{i\leq m}$.

In order to define the first surviving strategy f_0^B in the sequence $(f_i^B)_{i\leq m}$, we fix valuations $\|p\| = E_0$ and $\|p'\| = G_0^B$ in the formula $K_m^{b_I}(x)$ and we consider the witnesses for set-variables $b_1, \ldots, b_k$ in $K_{m+1}^{b_I}(x)$. By definition of $K_{m+1}^{b_I}(x)$, for each node s in G_0^B there is a unique $b_s \in \{b_1, \ldots, b_k\}$ such that $s \in \|b_s\|$. This yields a strategy f_0^B for $\exists$ in $\mathcal{G}$ (actually, for partial matches which are played along nodes of G_0^B), which we define as follows:

1) f_0^B is defined at the basic position (b_I, s_I);
2) given a basic position $(b_s, s) \in B \times G_0^B$ with $s \notin Ft(G_0^B)$, we let f_0^B suggest to $\exists$ a marking assigning b_t to t, for each $t \in R[s]$;
3) we leave f_0^B undefined on all other basic positions from $B \times T$.

Given prefixes $G_0^{\overline{B}}$ and G_0^B as above, we define E_1 to be the smallest prefix of $\mathbb{T}$ such that $Ft(G_0^B) < Ft(E_1)$ and $Ft(G_0^{\overline{B}}) < Ft(E_1)$.

It remains to check that conditions $1-5$ in (‡) hold for the base case. Condition (‡.1), (‡.2) and (‡.3) are clear by construction of E_0, $G_0^{\overline{B}}$, G_0^B and E_1. For condition (‡.4), by assumption we have that $State_{b_I,B}(y)$ and $Surv_{B,P}(p')$ hold, being subformulae of $K_m^{b_I}(x)$, with $\|p'\| = G_0^B$ and $\|y\| = s_I$. By construction of the strategy f_0^B, this means that f_0^B is functional and surviving for $\exists$ in G_0^B. Analogously, (‡.5) holds because the subformula of $K_m^{b_I}(x)$ given as in (7) is true, meaning that every node on the frontier of G_0^B is associated with a unique accepting state of $\mathbb{B}$ according to f_0^B. In order to fulfill condition (‡.6), we observe that, by definition of $K_{m+1}^{b_I}(x)$, every node $s \in Ft(G_0^B)$ is associated with a basic position $(b_s, s) \in B \times T$, such that $b_s \in F$ and $K_m^{b_s}(x)$ holds for $s = \|x\|$.

Inductively, we consider the stage $j+1 \leq m$ of the construction. By inductive hypothesis, we are given with sequences $(E_i)_{i\leq j+1}$, $(G_i^B)_{i\leq j}$, $(G_i^{\overline{B}})_{i\leq j}$ and a strategy f_j^B for $\exists$ as in (‡).

Analogously to the base case, we define $G_{j+1}^{\overline{B}}$ as smallest prefix in the sequence $(E_i^{\overline{f}})_{i<\omega}$ which contains E_{j+1}. For the definition of G_{j+1}^B and f_{j+1}^B, the key observation is that, by inductive hypothesis, for each node $s \in Ft(G_j^B)$ we can make the following assumptions:

1) the formula $K_{m-j}^{b_s}(x)$ holds, with $s = \|x\|$;
2) the position (b_s, s) is f_j^B-admissible.

We let $T.s \cap E_{j+1}$ be the witness for the set-variable p occurring in $K_{m-j}^{b_s}(x)$. Then by definition of $K_{m-j}^{b_s}(x)$ we are provided with a prefix $G_{j+1}^{B,s}$ of $\mathbb{T}.s$ witnessing the variable p', such that $T.s \cap E_{j+1} \subseteq G_{j+1}^{B,s}$. Also we are provided with noetherian sets of nodes witnessing variables $b_1, \ldots, b_k$. Analogously to the base case, this yields a strategy $f_{j+1}^{B,s}$ for $\exists$ in $\mathcal{G}$, which is defined as follows:

1) f_{j+1}^B is defined at the basic position (b_s, s);
2) for each basic position $(b_t, t) \in B \times G_{j+1}^{B,s}$ with $t \notin Ft(G_{j+1}^{B,s})$, we let f_{j+1}^B suggest to $\exists$ a marking assigning b_r to r, for each $r \in R[t]$.
3) we leave f_{j+1}^B undefined on all other basic positions from $B \times T$.

In other words, $f_{j+1}^{B,s}$ is a strategy for $\exists$ in partial matches of $\mathcal{A}(\mathbb{A}, \mathbb{T})@(b_s, s)$, which is functional, surviving in $G_{j+1}^{B,s}$ and marks each node $t \in Ft(G_{j+1}^{B,s})$ with a unique state from F.

We define G_{j+1}^B by putting

$$G_{j+1}^B := G_j^B \cup \bigcup_{s\in Ft(G_j^B)} G_{j+1}^{B,s}.$$

Since G_j^B is a prefix of $\mathbb{T}$ and for each $s \in Ft(G_j^B)$ the set $G_{j+1}^{B,s}$ is a prefix of $\mathbb{T}.s$, we have that G_{j+1}^B is a prefix of $\mathbb{T}$. Next, we define f_{j+1}^B by putting

$$f_{j+1}^B := f_j^B \cup \bigcup_{s\in Ft(G_j^B)} f_{j+1}^{B,s},$$

where the union of strategies just means the union of their graphs. In order to check that f_{j+1}^B is indeed a function, observe that by inductive hypothesis f_j^B is defined on basic positions in $B \times (G_j^B \setminus Ft(G_j^B))$. By construction, for each $s \in Ft(G_j^B)$, the strategy $f_{j+1}^{B,s}$ is defined on the union of $\{(b_s, s)\}$ and $B \times (G_{j+1}^{B,s} \setminus (G_j^B \cup Ft(G_{j+1}^{B,s})))$. Since $(b_s, s) \in Ft(G_j^B)$, then the domains of f_j^B and each $f_{j+1}^{B,s}$ are all disjoints. Therefore f_{j+1}^B is uniquely defined on each basic position in its domain.

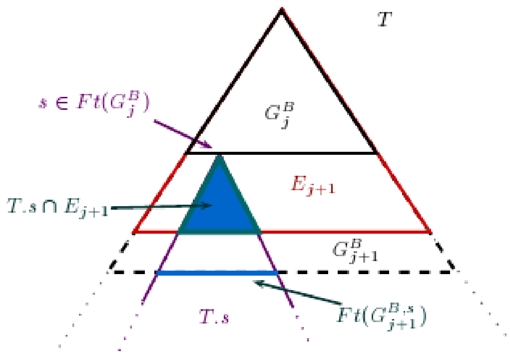

Figure 1. Construction of G^B_{j+1}

Given $G^{\overline{B}}_{j+1}$ and G^B_{j+1} as above, if $j+1 < m$ then we define E_{j+2} to be the smallest prefix of $\mathbb{T}$ such that $Ft(G^B_{j+1}) < Ft(E_{j+2})$ and $Ft(G^{\overline{B}}_{j+1}) < Ft(E_{j+2})$. The check that all conditions in (‡) hold for G^B_{j+1}, f^B_{j+1} and E^B_{j+2} is completely analogous to the base case.

We have just defined a strategy f^B, sequences $(E_i)_{i\leq m}$, $(G^B_i)_{i\leq m}$ and $(G^{\overline{B}}_i)_{i\leq m}$, such that for each $i \leq m$ condition (‡) is respected. It follows that $\overline{f}$ and f^B witness a trap for $\mathbb{B}$ and $\overline{B}$ according to definition 8. This concludes the proof of the claim. ■

The proof of claim 2 completes the proof of direction $(\Leftarrow)$. ■