Frame definability for classes of trees in the μ-calculus

Gaëlle Fontaine[1] and Thomas Place[2] ⋆

[1] ILLC, Universiteit van Amsterdam, gaelle.fontaine@uva.nl
[2] INRIA and ENS Cachan, place@lsv.ens-cachan.fr

Abstract. We are interested in frame definability of classes of trees, using formulas of the μ-calculus. In this set up, the proposition letters (or in other words, the free variables) in the μ-formulas correspond to second order variables over which universally quantify. Our main result is a semantic characterization of the **MSO** definable classes of trees that are definable by a μ-formula. We also show that it is decidable whether a given **MSO** formula corresponds to a μ-formula, in the sense that they define the same class of trees.

Basic modal logic and μ-calculus can be seen as logical languages for talking about *Kripke models* and *Kripke frames*. On Kripke models every modal formula is equivalent to a first order formula in one free variable and every μ-calculus formula is equivalent to a monadic second order formula in one free first order variable. On Kripke frames, we universally quantify over the free propositional variables occurring in the formulas and each modal formula or μ-formula is equivalent to a sentence of monadic second order logic. For example, the modal formula $p \to \Diamond p$ corresponds locally on Kripke models to the first order formula $\alpha(u, P) = P(u) \to \exists v(uRv \wedge P(v))$ (where P is a unary predicate corresponding to p, R is the binary relation of the model and u is a point of the model). The same modal formula corresponds globally on Kripke frames to the second order sentence $\forall P \forall u \alpha(u, P)$, which happens to be equivalent to the first order sentence $\forall u, uRu$.

The expressive power of modal logic from both perspectives (models and frames) has been extensively studied. For Kripke models, Johan van Benthem characterized modal logic semantically as the bisimulation invariant fragment of first order logic [vB76]. The problem whether a formula of first order logic in one free variable has a modal correspondent on the level of models, is undecidable [vB96].

The expressive power of modal logic on Kripke frames has been studied since the 1970s and this study gave rise to many key results in the modal logic area. When interpreted on frames, modal logic corresponds to a fragment of monadic

⋆ The research of the first author has been made possible by VICI grand 639.073.501 of the NWO. We would like to thank Balder ten Cate for inspiring us and for helping us continuously during the research and redaction of this paper. We also thank Johan van Benthem, Diego Figueira, Luc Segoufin and Yde Venema for helpful insights.

second order logic, because the definition of validity involves quantifying over all the proposition letters in the formulas. However, most work has concentrated on the first order aspect of modal definability.

A landmark result is the Goldblatt-Thomason Theorem [GT75] which gives a characterization of the first order definable classes of frames that are modally definable, in terms of semantic criteria. It is undecidable whether a given first order sentence corresponds to a modal formula, in the sense that they define the same class of frames.

On the level of Kripke models, the expressive power of the μ-calculus is well understood. In [JW96], David Janin and Igor Walukiewicz showed that the μ-calculus is the bisimulation invariant fragment of **MSO**. It is undecidable whether a class of Kripke models definable in **MSO** is definable by a formula of the μ-calculus. For classes of trees, the problem becomes decidable (see [JW96]).

About the expressive power of the μ-calculus on the level of Kripke frames, nothing is known. This paper contributes to a partial solution of this question by giving a characterization of the **MSO** definable classes of trees that are definable by a μ-formula. Our main result states that an **MSO** definable class of trees is definable in the μ-calculus iff it is closed for subtrees and p-morphic images. We also show that given an **MSO** formula, it is decidable whether there exists a μ-formula which defines the same class of trees as the **MSO** formula.

The proof is in three steps. First, we use the connection between **MSO** and the graded μ-calculus proved by Igor Walukiewicz [Wal02] and establish a correspondence between the **MSO** formulas that are preserved under p-morphic images and a fragment that is between the μ-calculus and the graded μ-calculus (the fragment with a counting $\Box$ operator and a usual $\Diamond$ operator). We call this fragment the $\Box$-graded μ-calculus.

The second step consists in showing that each $\Box$-graded μ-formula φ can be translated into a μ-formula ψ such that locally, the truth of φ (on trees seen as Kripke models) corresponds to the validity of ψ (on trees seen as Kripke frames). So this step is a move from the model perspective to the frame perspective. The last step consists in shifting from the local perspective to the global one (that is, we are interested in validity at all points, not at a given point).

1 Preliminaries

μ-calculus The set of formulas of the *μ-calculus* (over a set *Prop* of proposition letters and a set *Var* of variables) is given by

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

where p ranges over the set *Prop*, x ranges over the set *Var*.

A *Kripke frame* over a set *Prop* is a pair (W, R), where W is a set and R a binary relation on W. A *Kripke model* over *Prop* is a triple (W, R, V) where (W, R) is a Kripke frame and $V : \mathit{Prop} \to \mathcal{P}(W)$ a valuation.

Given a formula φ, a Kripke model $\mathcal{M} = (W, R, V)$ and an assignment $\tau : \mathit{Var} \to \mathcal{P}(W)$, we define a subset $[\![\varphi]\!]_{\mathcal{M},\tau}$ that is interpreted as the set of points at which φ is true. We only recall that

$$[\![\mu x.\varphi]\!]_{\mathcal{M},\tau} = \bigcap\{U \subseteq W \mid [\![\varphi]\!]_{\mathcal{M},\tau[x:=U]} \subseteq U\},$$
$$[\![\nu x.\varphi]\!]_{\mathcal{M},\tau} = \bigcup\{U \subseteq W \mid U \subseteq [\![\varphi]\!]_{\mathcal{M},\tau[x:=U]}\},$$

where $\tau[x := U]$ is the assignment τ' such that $\tau'(x) = U$ and $\tau'(y) = \tau(y)$, for all $y \neq x$. The set $[\![\mu x.\varphi]\!]_{\mathcal{M},\tau}$ is the least fixpoint of the map $\varphi_x : \mathcal{P}(W) \to \mathcal{P}(W)$ defined by $\varphi_x(U) := [\![\varphi]\!]_{\mathcal{M},\tau[x:=U]}$, for all $U \subseteq W$.

In case $w \in [\![\varphi]\!]_{\mathcal{M},\tau}$, we write $\mathcal{M}, w \Vdash_\tau \varphi$ and we say that φ is *true* at w. If all the variables in φ are bound, we simply write $\mathcal{M}, w \Vdash \varphi$. A formula φ is *true* in $\mathcal{M}$, notation: $\mathcal{M} \Vdash \varphi$, if for all $w \in W$, we have $\mathcal{M}, w \Vdash \varphi$. Two formulas φ and ψ are *equivalent* if for all models $\mathcal{M}$ and for all $w \in \mathcal{M}$, $\mathcal{M}, w \Vdash \varphi$ iff $\mathcal{M}, w \Vdash \psi$.

If (W, R) is a Kripke frame and w belongs to W, we say that φ is *valid* at w if for all valuations V, φ is true at w in (W, R, V). We use the notation $(W, R), w \Vdash \varphi$. Finally, φ is *valid* in (W, R), notation: $(W, R) \Vdash \varphi$, if φ is valid at w, for all w in W.

Trees Our characterizations apply only to classes of trees, not classes of arbitrary Kripke frames.

Let (W, R) be a Kripke frame. A point r in W is a *root* if for all w in W, there is a sequence $w_0, \ldots, w_n$ such that $w_0 = r$, $w_n = w$ and (w_i, w_{i+1}) belongs to R, for all $i \in \{0, \ldots, n-1\}$. The frame (W, R) is a *tree* if it has a root, every point distinct from the root has a unique predecessor and the root has no predecessor. If (W, R, V) is a Kripke model over a set *Prop* and (W, R) is a tree, we say that (W, R, V) is a *tree Kripke model* over *Prop* or simply a tree over *Prop* or a tree model. Two formulas φ and ψ over *Prop* are *equivalent on tree models* if for all trees t over *Prop* with root r, $t, r \Vdash \varphi$ iff $t, r \Vdash \psi$.

If the frame (W, R) is a tree, v is *child* of w if $(w, v) \in R$ and we write $Child(w)$ to denote the children of w. A *subtree* of a tree t is a tree consisting of a node in t and all of its descendants in t. If t is a tree and u is a node of t, we let $t|_u$ denote the subtree of t at position u.

A class of trees L over *Prop* is a *regular class of trees* if there exists an **MSO** formula α such that for all trees t, t belongs to L iff α is valid on t. When this happens, we say that α *defines* L.

2 μ-definability on trees

We are interested in characterizing the regular classes of trees (seen as Kripke frames) that are definable using μ-formulas. The characterization we propose, is very natural and only involves two well-known notions of modal logic: subtree and p-morphism. We recall these notions together with the notion of definability and state our main result (Theorem 1).

μ-definability A class of trees L is *μ-definable* if there exists a μ-formula φ such that L is exactly the class of trees which make φ valid.

p-morphisms Let $\mathcal{M} = (W, R, V)$ and $\mathcal{M}' = (W', R', V')$ be two models. A map $f : W \to W'$ is a *p-morphism* between $\mathcal{M}$ and $\mathcal{M}'$ if the two following conditions hold. For all $w, v \in W$ such that wRv, $f(w)R'f(v)$. For all $w \in W$ and $v' \in W'$ such that $f(w)R'v'$, there exists $v \in W$ such that $f(v) = v'$ and wRv.

A class L of Kripke models is *closed under p-morphic images* if for all Kripke models $\mathcal{M} \in L$ and and for all Kripke models $\mathcal{M}'$ such that there is a surjective p-morphism between $\mathcal{M}$ and $\mathcal{M}'$, we have that $\mathcal{M}'$ belongs to L. An **MSO** formula α is *preserved under p-morphic images for tree models* if the class of tree models defined by α is closed under p-morphic images.

Closure for subtrees A class of trees L is *closed for subtrees* if for all $t \in L$ and all $u \in t$, we have that $t|_u$ belongs to L.

Theorem 1. *A regular class of trees is μ-definable iff it is closed under p-morphic images for tree models and closed for subtrees.*

The rest of this paper is devoted to the proof of Theorem 1. First we characterize the class of regular classes of trees which are preserved under p-morphic images. It corresponds to some fragment of the graded μ-calculus (roughly, the fragment where we allow counting with the $\Box$ operator, but not the $\Diamond$ operator). Next we prove that this fragment defines the regular classes of trees which are of the form $\{t \text{ tree} \mid \text{ for all } V : \mathit{Prop}' \to \mathcal{P}(t), (t, V), r \Vdash \varphi\}$, where r is the root of t and φ is a formula of the μ-calculus. Finally we show how to derive Theorem 1.

3 Graded μ-calculus: connection with MSO and disjunctive normal form

An important tool for characterizing the class of regular tree languages which are preserved under p-morphic images, is the connection between **MSO** and graded μ-calculus. We also use fact that there is a normal form for the graded μ-formulas (when they are expressed using a ∇-like operator).

Graded μ-calculus The set μGL of formulas of the *graded μ-calculus* (over a set *Prop* of proposition letters and a set *Var* of variables) is given by

$$\varphi \ ::= \ \top \mid p \mid x \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \Diamond^k \varphi \mid \Box^k \varphi \mid \mu x.\varphi \mid \nu x.\varphi,$$

where p ranges over the set *Prop*, x ranges over the set *Var* and k is a natural number. Given a formula φ, a model $\mathcal{M} = (W, R, V)$, an assignment $\tau : \mathit{Var} \to \mathcal{P}(W)$ and $w \in W$, the relation $\mathcal{M}, w \Vdash \varphi$ is defined by induction as in the case of the μ-calculus with the extra conditions:

$\mathcal{M}, w \Vdash_\tau \Diamond^k \varphi$	if	there exist successors $w_0, \dots, w_k$ of w s.t. for all $i \neq j$, $w_i \neq w_j$ and for all $i \in \{0, \dots, k\}$, $\mathcal{M}, w_i \Vdash \varphi$
$\mathcal{M}, w \Vdash_\tau \Box^k \varphi$	if	w has no successors or there exist successors $w_1, \dots, w_k$ of w s.t. for all $w \notin \{w_1, \dots, w_k\}$, $\mathcal{M}, w \Vdash \varphi$.

We define a ∇-like operator corresponding to graded μ-calculus (inspired by [Wal02]). The ∇-formulas of the graded μ-calculus correspond exactly to the formulas of the graded μ-calculus.

∇ operator for the graded μ-calculus Given a multiset Φ of formulas, the *multiplicity* of a formula φ in Φ is the number of occurrences of φ in Φ. The total number of elements in a multiset, including repeated memberships, is the *cardinality* of the multiset. We denote by $card(\Phi)$, the cardinality of Φ. A *literal* over a set *Prop* is a proposition letter in *Prop* or the negation of a proposition letter.

The set $\mu\mathrm{GL}^{\nabla}$ of ∇-formulas of the graded μ-calculus (over a set *Prop* of proposition letters and a set *Var* of variables) is given by:

$$\varphi \ ::= \ x \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \Pi \bullet \nabla^{g}(\Phi;\Psi) \mid \mu x.\varphi \mid \nu x.\varphi,$$

where x ranges over the set *Var*, Π is a conjunction literals or $\Pi = \top$, Φ is a multiset of formulas and Ψ is a finite set of formulas.

Given a formula φ, a model $\mathcal{M} = (W, R, V)$, an assignment $\tau : Var \to \mathcal{P}(W)$ and $w \in W$, the relation $\mathcal{M}, w \Vdash_{\tau} \varphi$ is defined by induction as in the case of the μ-calculus with the extra condition: $\mathcal{M}, w \Vdash_{\tau} \Pi \bullet \nabla^{g}(\Phi, \Psi)$ iff $\mathcal{M}, w \Vdash_{\tau} \Pi$ and for some $\{w_{\varphi}$ successor of $w \mid \varphi \in \Phi\}$, we have

1. the size of the set $\{w_{\varphi} \mid \varphi \in \Phi\}$ is equal to $card(\Phi)$,
2. $\mathcal{M}, w_{\varphi} \Vdash_{\tau} \varphi$,
3. for all successors u of w such that $u \notin \{w_{\varphi} \mid \varphi \in \Phi\}$, $\mathcal{M}, u \Vdash_{\tau} \bigvee \Psi$.

A map $m : \mu\mathrm{GL}^{\nabla} \to \mathcal{P}(R[w])$ is a ∇^{g}*-marking* for (Φ, Ψ) if there exists a set $\{w_{\varphi} \mid \varphi \in \Phi\}$ of size $card(\Phi)$ such that $w_{\varphi} \in m(\varphi)$ and for all successors u of w such that $u \notin \{w_{\varphi} \mid \varphi \in \Phi\}$, there is $\psi \in \Psi$ such that $u \in m(\psi)$.

A formula in $\mu\mathrm{GL}^{\nabla}$ is in *disjunctive normal form* if its only subformulas of the form $\varphi_0 \wedge \varphi_1$ are such that φ_0 and φ_1 are literals or conjunctions of literals.

The next result is proved by a standard (although a bit tedious) induction on the complexity of the formulas.

Proposition 1. *Each formula in* $\mu\mathrm{GL}$ *is equivalent to a formula in* $\mu\mathrm{GL}^{\nabla}$*. Each formula in* $\mu\mathrm{GL}^{\nabla}$ *is equivalent to a formula in* $\mu\mathrm{GL}$*.*

In [Wal02], Igor Walukiewicz showed that on trees, **MSO** is equivalent to first order logic extended with the unary fixpoint operator. By adapting[1] the proof of Lemma 44 in [Wal02], we can obtain the following result (see also [JL03]).

Theorem 2 (from [Wal02]). *For every **MSO** formula α, there is a graded μ-formula φ such that for all trees t with root r, α is valid on t iff φ is true at r. For every graded μ-formula φ, there is an **MSO** formula α such that for all trees t with root r, φ is true at r iff α is valid on t.*

[1] This adaptation is mainly based on the following observation (which is immediate from Proposition 1): For all formulas $\varphi \in \mathrm{DBF(n)}$ (as defined in [Wal02]), there is $\psi \in \mu\mathrm{GL}$ such that for all trees t, for all nodes u in t, ψ is true at u iff the formula obtained from φ by relativizing all the quantifiers to the children of u, holds.

As mentioned earlier, a key result for one of our proofs is the fact that the graded μ-calculus has a normal form. This follows from a result proved by David Janin in [JW95].

Theorem 3. *Each formula of the graded μ-calculus is equivalent to a formula of the graded μ-calculus in disjunctive normal form.*

4 $\Box$-graded μ-calculus and preservation under p-morphic images

We establish a correspondence between the **MSO** formulas that are preserved under p-morphic images and some set of formulas, that is in between the μ-calculus and the graded μ-calculus. We call this set the set of $\Box$-graded formulas, as we are only allowed to count with the $\Box$ operator. For this set of $\Box$-graded formulas, we also introduce a ∇-like operator, that we write ∇'.

$\Box$-graded μ-calculus The set $\mu\mathrm{GL}^{\Box}$ of fixpoint $\Box$-graded formulas (over a set *Prop* of proposition letters and a set *Var* of variables) is given by

$$\varphi \ ::= \ \top \mid p \mid x \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \Diamond\varphi \mid \Box^k\varphi \mid \mu x.\varphi \mid \nu x.\varphi,$$

where p ranges over the set *Prop*, x ranges over the set *Var* and k is a natural number. Given a formula φ, a model $\mathcal{M} = (W, R, V)$, an assignment $\tau : Var \to \mathcal{P}(W)$ and $w \in W$, the relation $\mathcal{M}, w \Vdash_\tau \varphi$ is defined by induction as in the case of the graded μ-calculus, with $\Diamond = \Diamond^0$.

The set $\mu\mathrm{GL}^{\nabla'}$ of ∇'-formulas of the graded μ-calculus are given by:

$$\varphi \ ::= \ x \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \Pi \bullet \nabla'(\Phi; \Psi) \mid \mu x.\varphi \mid \nu x.\varphi,$$

where x ranges over the set *Var* of variables, Π is a conjunction of literals or $\Pi = \top$, Φ is a multiset of formulas and Ψ is a finite set of formulas.

Given a formula φ, a Kripke model $\mathcal{M} = (W, R, V)$, an assignment $\tau : Var \to \mathcal{P}(W)$ and a point $w \in W$, the relation $\mathcal{M}, w \Vdash_\tau \varphi$ is defined by induction as in the case of the μ-calculus with the extra condition: $\mathcal{M}, w \Vdash_\tau \Pi \bullet \nabla^g(\Phi, \Psi)$ iff $\mathcal{M}, w \Vdash_\tau \Pi$ and for some $\{w_\varphi$ successor of $w \mid \varphi \in \Phi\}$, we have

1. $\mathcal{M}, w_\varphi \Vdash_\tau \varphi$,
2. for all successors u of w such that $u \notin \{w_\varphi \mid \varphi \in \Phi\}$, $\mathcal{M}, u \Vdash_\tau \bigvee \Psi$.

A map $m : \mu\mathrm{GL}^{\nabla'} \to \mathcal{P}(R[w])$ is a ∇'-*marking* for (Φ, Ψ) if the set there exists a set $\{w_\varphi \mid \varphi \in \Phi\}$ such that $w_\varphi \in m(\varphi)$ and for all successors u of w such that $u \notin \{w_\varphi \mid \varphi \in \Phi\}$, there is $\psi \in \Psi$ such that $u \in m(\psi)$.

The only difference between a ∇'-marking and a ∇^g-marking is that the set of successors associated to a ∇'-marking might not contain points which are pairwise distinct (as in the case for the ∇^g-marking).

A formula $\varphi \in \mathrm{GL}^{\Box} \cup \mu\mathrm{GL}^{\nabla'}$ is *equivalent on tree models* to an **MSO** formula α if for all tree models t, α is valid on t iff φ is true at the root of t.

We also introduce a game semantic for the languages $\mu\mathrm{GL}^{\nabla}$ and $\mu\mathrm{GL}^{\nabla'}$. The fact the the existence of a winning strategy for a player in the game corresponds to the truth of a formula at a given point, is proved using classical methods (as in the case of μ-calculus).

Game semantics Let φ be a formula in $\mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'}$ such that each variable in φ is bound. Without loss of generality, we may assume that for all $x \in \mathit{Var}$ which occurs in φ, there is a unique subformula of φ, which is of the form $\eta x.\delta_x$, where $\eta \in \{\mu, \nu\}$. We also fix a model $\mathcal{M} = (W, R, V)$. We define the *evaluation game* $\mathcal{E}(\mathcal{M}, \varphi)$ as a parity game between two players, $\forall$ and $\exists$. The rules of the game are given in the table below.

Position	Player	Possible moves
$(\top, w)$	$\forall$	$\emptyset$
(x, w)	-	$\{(\delta_x, w)\}$
$(\varphi_1 \wedge \varphi_2, w)$	$\forall$	$\{(\varphi_1, w), (\varphi_2, w)\}$
$(\varphi_1 \vee \varphi_2, w)$	$\exists$	$\{(\varphi_1, w), (\varphi_2, w)\}$
$(\eta x.\psi, w)$	-	$\{(\psi, w)\}$
$(\Pi \bullet \nabla^g(\Phi, \Psi), w)$	$\exists$	$\{m : \mu\mathrm{GL}^{\nabla^g} \to \mathcal{P}(R[w]) \mid \mathcal{M}, w \Vdash \Pi$ and m ∇^g-marking for $(\Phi, \Psi)\}$
$(\Pi \bullet \nabla'(\Phi, \Psi), w)$	$\exists$	$\{m : \mu\mathrm{GL}^{\nabla'} \to \mathcal{P}(R[w]) \mid \mathcal{M}, w \Vdash \Pi$ and m ∇'-marking for $(\Phi, \Psi)\}$
$m : \mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'} \to \mathcal{P}(R[w])$	$\forall$	$\{(\psi, u) \mid u \in m(\psi)\}$

where w belongs to W, x belongs to Var, η belongs to $\{\mu, \nu\}$, φ_1, φ_2 and ψ belongs to $\mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'}$, Π is a conjunction of literals or $\Pi = \top$, Ψ is a finite subset of $\mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'}$, Φ is a finite multiset of formulas in $\mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'}$.

If a match is finite, the player who get stuck, loses. If a match is infinite, we let *Inf* be the set of variables x such that there are infinitely many positions of the form (x, w) in the match. Let x_0 be a variable in *Inf* such that for all variables $x \in \mathit{Inf}$, δ_x is a subformula of δ_{x_0}. If x_0 is bound by a μ-operator, then $\forall$ wins. Otherwise $\exists$ wins.

The notions of *strategy* and *winning strategy* for a player are defined as usual. If h is a strategy for a player P, a match during which P plays according to h is called an *h-conform match*.

Proposition 2. *Let φ be a formula in $\mu\mathrm{GL}^{\nabla} \cup \mu\mathrm{GL}^{\nabla'}$. For all Kripke models $\mathcal{M} = (W, R, V)$ and all $w \in W$, $\mathcal{M}, w \Vdash \varphi$ iff $\exists$ has a winning strategy in the game $\mathcal{E}(\mathcal{M}, \varphi)$ with starting position (w, φ).*

We are now ready to show that modulo equivalence on trees, the **MSO** formulas preserved under p-morphic images are exactly the $\Box$-graded formulas.

Proposition 3. *Let α be an **MSO** formula. The following are equivalent:*

- *α is equivalent on tree models to a $\Box$-graded formula,*

– α *is equivalent on tree models to a formula in* $\mu\mathrm{GL}^{\nabla'}$,
– α *is preserved under* p*-morphic images for tree models* .

Proof. We only give details for the hardest implication which is that an **MSO** formula which is preserved under p-morphic images for tree models is equivalent on tree models to a ∇'-graded fixpoint formula. Let α be an **MSO** formula that is preserved under p-morphic images for tree models. By Theorem 2, there is a graded μ-formula χ such that for all tree models t with root r, α is valid on t iff χ is true at r. By Proposition 3, we may assume the formula χ to be in disjunctive normal form.

Now let δ be the formula χ in which we replace each operator ∇^g by ∇'. We show that under the assumption that α is preserved under p-morphic images for trees, χ and δ are equivalent on tree models. It is easy to check that for all tree models, if χ holds at a node, then δ also holds at the node.

For the other direction, let t be a tree model with root r and suppose that δ is true at r. We have to show that χ is true at r. Since δ is true at r, $\exists$ has a winning strategy h in the evaluation game with starting position (r, δ). We start by fixing some notation. We denote by $\mathbb{N}^*$ the set of finite sequences over $\mathbb{N}$. The empty sequence ϵ belongs to $\mathbb{N}^*$. If φ is a ∇'-formula, we write φ^g for the formula obtained by replacing ∇' by ∇^g in φ. If Φ is set (multiset) of formulas, we write Φ^g for the set (multiset) $\{\varphi^g \mid \varphi \in \Phi\}$. We also let A be the set $\{(u, \boldsymbol{n}) \mid u \text{ is a node of } t \text{ and } \boldsymbol{n} \in \mathbb{N}^*)\}$.

The idea is to define a new tree model t' the domain of which is a subset of A and the child relation of which is such that for all $(u_1, (n_1, \dots, n_k))$ and $(u_2, \boldsymbol{m})$ in t', $(u_2, \boldsymbol{m})$ is a child of $(u_1, (n_1, \dots, n_k))$ iff u_2 is a child of u_1 in t and there is a natural number n_{k+1} such that $\boldsymbol{m} = (n_1, \dots, n_k, n_{k+1})$. The depth of a node $(u_1, (n_1, \dots, n_k))$ in t' is $k+1$. The tree t' will also be such that its root is (r, ϵ). Finally, we define a positional strategy h' for the evaluation game $\mathcal{E}(t', \chi)$ with starting position $((r, \epsilon), \chi)$ which satisfies the following conditions. For all $(u, \boldsymbol{n}) \in t$, there is exactly one match during which $(u, \boldsymbol{n})$ occurs. Moreover, a position of the form $((u, \boldsymbol{n}), \varphi^g)$ is reached in an h'-conform match iff the position (u, φ) is reached in an h-conform match.

The definitions of t' and h' will be by induction. More precisely, at stage i of the induction, we specify which are the nodes of t' of depth i and we also define $\exists$'s answer (according to h') when a position of the form $((u, \boldsymbol{n}), \varphi^g)$ is reached, where the depth of $(u, \boldsymbol{n})$ in t' is $i - 1$.

For the basic case, the only node of depth 1 in t' is the node (r, ϵ). For the induction step, take $i > 1$ and suppose that we know already which are the nodes in t' of depth at most i and that we also have defined the strategy h' for all positions of the form $((u, \boldsymbol{n}), \varphi)$, where the depth of $(u, \boldsymbol{n})$ in t' is at most $i - 1$. We have to specify which points of the form $(u, (n_1, \dots, n_i))$ belongs to t' and what is the strategy for the points of depth i. Let $(u, \boldsymbol{n}) = (u, (n_1, \dots, n_{i-1}))$ be a node in t' of depth i (note that if $i = 1$, then $(n_1, \dots, n_{i-1})$ is ϵ). Suppose that in a (partially defined) h'-conform match π', a position of the form $((u, \boldsymbol{n}), \varphi^g)$ is reached and if $\boldsymbol{n} = \epsilon$, we can assume that $\varphi = \delta$ and $\varphi^g = \chi$. By induction hypothesis, such a match π' is unique. Moreover, we know (by induction hypothesis if $i > 1$ or trivially if $i = 1$) that the position (u, φ) is reached in an h-conform match π. Now, there are different possibilities depending on the shape of φ. First, suppose that φ is a disjunction $\varphi_1 \vee \varphi_2$. Then, in the h-conform match π, the position following (u, φ) is of the form (u, ψ), where ψ is either φ_1 or φ_2. We define h' such that the position following $((u, \boldsymbol{n}), \varphi^g)$ is $((u, \boldsymbol{n}), \psi^g)$. Finally, suppose that φ^g is of the form $\Pi \bullet \nabla^g(\Phi^g; \Psi^g)$. Then, in the h-conform match π, the position following (u, φ) is a marking $m : \Phi \cup \Psi \to \mathcal{P}(Child(s))$ such that m is a ∇'-marking for (Φ, Ψ). In the h'-conform match π', we first define which are the children of $(u, \boldsymbol{n})$ in t' and then, we give a ∇^g-marking $m^g : Child(u, \boldsymbol{n}) \to \mathcal{P}(\Psi^g \cup \Gamma^g)$ for (Φ^g, Ψ^g).

Since m is a ∇'-marking for (Φ, Ψ), there exists $\{u_\varphi \mid \varphi \in \Phi\}$ such that the two following conditions holds. For all $\varphi \in \Phi$, u_φ is a child of u and u_φ belongs to $m(\varphi)$. For all children v of u such that v does not belong to $\{u_\varphi \mid \varphi \in \Phi\}$, there exists $\psi \in \Psi$ such that v belongs to $m(\psi)$. We let $u_1, \ldots, u_k$ be the children of u such that $\{u_1, \ldots, u_k\} = \{u_\psi \mid \psi \in \Psi\}$ and $u_i \neq u_j$, if $i \neq j$. Fix i in $\{1, \ldots, k\}$. Let Φ_i be the biggest submultiset of Φ such that for all φ in Φ_i, $u_\varphi = u_i$. We let $k(i) + 1$ be the size of Φ_i and we fix an arbitrary bijection f_i between Φ_i and the set $\{0, \ldots, k(i)\}$. Now, we add to t' the set of nodes

$$\{(u_i, (n_1, \ldots, n_{i-1}, j)) \mid i \in \{1, \ldots, k\}, 0 \leq j \leq k(i)\} \cup$$
$$\{(v, (n_1, \ldots, n_{i-1}, 0)) \mid v \text{ child of } u, v \notin \{u_1, \ldots, u_k\}\}.$$

These points are the children of $(u, \boldsymbol{n})$ in t'.

We are now going to define a ∇^g-marking $m^g : \Phi^g \cup \Psi^g \to \mathcal{P}(Child(s, \boldsymbol{n}))$ for (Φ^g, Ψ^g). Fix a formula φ^g in Φ^g. We define $m^g(\varphi^g)$ as $\{(u_i, (n_1, \ldots, n_{i-1}, j))\}$, if φ belongs to Φ_i and $f_i(\varphi) = j$. Next fix a formula ψ^g in Ψ^g. We define m^g such that $m^g(\psi^g) = \{(v, (n_1, \ldots, n_{i-1}, 0)) \mid v \notin \{u_1, \ldots, u_k\}, v \in m(\psi)\}$. The proofs that m^g is a ∇^g-marking for (Φ^g, Ψ^g) and that the induction hypothesis remain true are standard. This finishes the definition of t' and h'.

It is easy to check that the strategy h' is winning for $\exists$ in the evaluation game $\mathcal{E}(t', \chi)$ with starting position $((r, \epsilon), \chi)$. Therefore, the formula χ is true at the root of t'. Now the map which sends a node $(u, \boldsymbol{n})$ to u is a surjective p-morphism between t' and t. So t is a p-morphic image of t'. Since α is preserved under p-morphic images for trees, χ is also true at the root of t and this finishes the proof that χ and δ are equivalent on tree models. It follows that χ and α are also equivalent on trees.

5 μ-definability at the root and μ-definability

μ-definability at the root A tree language L over *Prop* is *μ-definable at the root* if there are a set *Prop'* of proposition letters and a μ-formula φ over $Prop \cup Prop'$ such that L is equal to $\{t \text{ tree} \mid \text{ for all } V : Prop' \to \mathcal{P}(t), (t, V), r \Vdash \varphi\}$, where r is the root of t.

Proposition 4. *A regular class of trees over Prop is μ-definable at the root iff it is closed under p-morphic images for tree models.*

Proof. The only difficult direction is from right to left. Let L be a regular class of trees over *Prop* that is closed under p-morphic images for trees. By Proposition 3, there is a $\Box$-graded formula φ over *Prop* such that for all trees t over *Prop* with root r, φ is true at r iff t belongs to L. Now we show by induction on the complexity of φ that there exist a set of propositions $Prop'$ and a μ-formula ψ over $Prop \cup Prop'$ such that for all trees t over *Prop* with root r and for all assignments $\tau : Var \rightarrow \mathcal{P}(t)$,

$$t, r \Vdash_\tau \varphi \quad \text{iff} \quad \text{for all valuations } V : Prop' \rightarrow \mathcal{P}(t),\ (t, V), r \Vdash_\tau \psi. \qquad (1)$$

When this happens, we will say that ψ is a *μ-translation* of φ. Moreover, we prove that for all trees t over *Prop* and for all assignments $\tau : Var \rightarrow \mathcal{P}(t)$, there is a valuation $V_\psi(\tau) : Prop' \rightarrow \mathcal{P}(t)$ such that for all nodes u in t,

$$(t, V_\psi(\tau)), u \Vdash_\tau \psi \quad \text{iff} \quad \text{for all valuations } V : Prop' \rightarrow \mathcal{P}(t),\ (t, V), u \Vdash_\tau \psi. \ (2)$$

When a valuation $V_\psi(\tau) : Prop' \rightarrow \mathcal{P}(t)$ satisfies condition (2), we say that $V_\psi(\tau)$ is a *distinctive valuation* for t, τ and ψ.

We only treat the most two difficult cases where φ is a formula of the form $\Box^k \varphi_1$ or of the form $\eta x.\varphi_1$, where η belongs to $\{\mu, \nu\}$.

Suppose that φ is of the form $\Box^k \varphi_1$. By induction hypothesis, there is a set $Prop'_1$ and there is a μ-formula ψ_1 over $Prop \cup Prop'_1$ such that ψ_1 is a μ-translation of φ_1. We let $p_0, \dots, p_k$ be fresh proposition letters and we define $Prop'$ as $Prop'_1 \cup \{p_0, \dots, p_k\}$. We let ψ be the formula

$$\bigvee \{\Box(\neg p_i \vee \psi_1) \mid 0 \le i \le k\} \vee \bigvee \{\Diamond(p_i \wedge p_{i'}) \mid 0 \le i, i' \le k, i \neq i'\}.$$

Fix a tree model t and an assignment $\tau : Var \rightarrow \mathcal{P}(t)$. We define $V_\psi(\tau)$. Let V_0 be the assignment $V_{\psi_1}(\tau)$ and let u be a point in t. We define $U = \{u_j \mid j \in J\}$ as the biggest set of successors of u such that for all $j \in J$, we have $(t, V_0), u_j \nVdash \psi_1$. Suppose first that the size of U is less or equal to k. So U is a set of the form $\{u_1, \dots, u_n\}$, where $n \le k$. Then we can fix a valuation V_u which satisfies the following. For all $i \in \{0, \dots, k\}$, $V_u(p_i)$ is a set of the form $\{u_i\}$, where $u_i \in U$. Moreover, for all $v \in U$, there is $i \in \{0, \dots, k\}$ such that $V_u(p_i) = \{v\}$.

Next suppose that U is infinite or U is a finite set of size strictly greater than k. Then we can fix an arbitrary valuation V_u such that the following hold. For all $i \in \{0, \dots, k\}$, $V_u(p_i)$ is a set of the form $\{u_i\}$, where $u_i \in U$. Moreover, for all $i, i' \in \{0, \dots, k\}$, if $i \neq i'$, then $u_i \neq u_{i'}$. That is, for all $i, i' \in \{0, \dots, k\}$, if $i \neq i'$, then $V_u(p_i) \cap V_u(p_j) = \emptyset$.

We are now ready to define $V_\psi(\tau)$. For all propositions $p' \in Prop'$, we have

$$V_\psi(\tau)(p') = \begin{cases} \bigcup\{V_u(p_i) \mid u \in t\} & \text{if } p' = p_i \text{ for some } i \in \{0, \dots, k\}, \\ V_0(p') & \text{otherwise.} \end{cases}$$

The proofs ψ is a μ-translation of φ and that $V_\psi(\tau)$ is a distinctive valuation for t, are left as an exercise to the reader.

Next, suppose that φ is a formula of the form $\eta x.\varphi_1$, where η belongs to $\{\mu, \nu\}$. We will assume that $\eta = \mu$, but the proof can be easily adapted to the case where $\eta = \nu$. By induction hypothesis, there is a set $Prop'_1$ and there is a μ-formula ψ_1 over $Prop \cup Prop'_1$ such that ψ_1 is a μ-translation of φ_1. Moreover, for all trees t and all assignments $\tau : Var \to \mathcal{P}(t)$, there is a distinctive valuation $V_{\psi_1}(\tau)$ for t, τ and ψ_1.

Now we define ψ as $\mu x.\psi_1$ and given a tree model t and an assignment $\tau : Var \to \mathcal{P}(t)$, we define $V_\psi(\tau)$ as the valuation $V_{\psi_1}(\tau[x \mapsto U_0])$, where $U_0 := [\![\mu x.\varphi_1]\!]_{t,\tau}$ $(= [\![\varphi]\!]_{t,\tau})$. We have to show that $V_\psi(\tau)$ is a distinctive valuation for t, τ and ψ and that ψ is a μ-translation of φ. It is sufficient to prove that for all trees t, all assignments $\tau : Var \to \mathcal{P}(t)$ and all $u \in t$, we have

$$t, u \Vdash_\tau \varphi \quad \Rightarrow \quad \text{for all valuations } V : Prop' \to \mathcal{P}(t),\ (t, V), u \Vdash_\tau \psi, \tag{3}$$

$$(t, V_\psi(\tau)), u \Vdash_\tau \psi \quad \Rightarrow \quad t, u \Vdash_\tau \varphi. \tag{4}$$

First, we show that condition (3) holds. So suppose that $t, u \Vdash_\tau \mu x.\varphi_1$ and let $V : Prop' \to \mathcal{P}(t)$ be a valuation. We have to verify that $(t, V), u \Vdash_\tau \mu x.\psi_1$. That is, for all subsets U of t such that $[\![\psi_1]\!]_{(t,V),\tau[x\mapsto U]} \subseteq U$, we have $u \in U$.

So fix a subset U of t such that $[\![\psi_1]\!]_{(t,V),\tau[x\mapsto U]} \subseteq U$. Since ψ_1 is a μ-translation of φ_1, we know that $[\![\varphi_1]\!]_{t,\tau[x\mapsto U]}$ is a subset of $[\![\psi_1]\!]_{(t,V),\tau[x\mapsto U]}$. It follows that $[\![\varphi_1]\!]_{t,\tau[x\mapsto U]}$ is a subset of U. That is, U is a pre-fixpoint of $(\varphi_1)_x$ in t under the assignment τ. Since $t, u \Vdash_\tau \mu x.\varphi_1$, this implies that u belongs to U and this finishes the proof of implication (3).

Next we show that implication (4) is true. Assume that $(t, V_\psi(\tau)), u \Vdash_\tau \mu x.\psi_1$. We have to prove that u belongs to $[\![\varphi]\!]_{t,\tau}$. That is, u belongs to U_0. Since $(t, V_\psi(\tau)), u \Vdash_\tau \mu x.\psi_1$, u belongs to all the pre-fixpoints of the map $(\psi_1)_x$ in the tree $(t, V_\psi(\tau))$ under the assignment τ. So it is sufficient to show that the set U_0 is a pre-fixpoint of the map $(\psi_1)_x$ in the tree $(t, V_\psi(\tau))$ under the assignment τ. That is,

$$[\![\psi_1]\!]_{(t,V_\psi(\tau)),\tau_0} \subseteq U_0, \tag{5}$$

where τ_0 is the assignment $\tau[x \mapsto U_0]$. By definition of $V_\psi(\tau)$, we have that $[\![\psi_1]\!]_{(t,V_\psi(\tau)),\tau_0}$ is equal to $[\![\psi]\!]_{(t,V_{\psi_1}(\tau_0)),\tau_0}$. Recall that for all assignments $\tau' : Var \to \mathcal{P}(t)$, we have that $[\![\psi_1]\!]_{(t,V_{\psi_1}(\tau')),\tau'}$ is equal to $[\![\varphi_1]\!]_{t,\tau'}$, as $V_{\psi_1}(\tau')$ is a distinctive valuation for t, τ' and ψ_1. In particular, $[\![\psi_1]\!]_{(t,V_{\psi_1}(\tau_0)),\tau_0}$ is equal to $[\![\varphi_1]\!]_{t,\tau_0}$. Since $U_0 = [\![\mu x.\varphi_1]\!]_{t,\tau}$, we also have that $[\![\varphi_1]\!]_{t,\tau_0}$ is equal to $[\![\mu x.\varphi_1]\!]_{t,\tau}$. That is, $[\![\varphi_1]\!]_{t,\tau_0}$ is equal to U_0. Putting everything together, we obtain that $[\![\psi_1]\!]_{(t,V_\psi(\tau)),\tau_0}$ is equal to U_0. Condition (5) immediately follows.

We can now prove Theorem 1.

Proof. We concentrate on the hardest direction, which is from left to right. Let L be a regular tree language, which is closed under p-morphic images for tree models and closed for subtrees. It follows from Proposition 4 that there is a μ-formula φ such that for all trees t, t belongs to L iff the formula φ is valid at the root of t. Now we prove that for all trees t, t belongs to L iff the formula φ is valid at all the nodes of t. It will immediately follow that L is μ-definable.

For the direction from right to left, let t be a tree such that φ is valid at all the nodes of t. In particular, φ is valid at the root. Therefore, t belongs to L.

For the other direction, let t be a tree in L. We have to show that for all nodes u in t, φ is valid at u. Fix a node u in t. Since t belongs to L and since L is closed under subtrees, the tree $t|_u$ belongs to L. That is, the formula φ is valid at the root of $t|_u$. It follows that the formula φ is valid at u in t.

Corollary 1. *It is decidable whether a regular class of trees is μ-definable.*

In order to derive this corollary from Theorem 1, it is sufficient to show that both closure for subtrees and closure under p-morphic images are decidable properties of regular classes of trees. Given a regular class of trees L, it is possible to find an **MSO** formula α such that L is closed for subtrees iff α is valid on trees. Decidability of closure for subtrees follows then from the decidability of **MSO** on trees. Decidability of closure under p-morphic images follows from a careful inspection of the proof of Proposition 3 together with the decidability of **MSO** on trees.

6 Discussion

A natural further question is to investigate the μ-definability for classes of frames, not only classes of trees. Unlike on trees, graded μ-calculus does not have the same expressive power as **MSO** on models: it corresponds to the fragment of **MSO** invariant under counting bisimilation (see [JL03] and [Wal02]). Moreover, the proof of Proposition 3 does not work for classes of frames, as it relies on the fact that given a disjunctive formula, a strategy for $\exists$ in the game associated to the formula and a tree t, there is at most one match conform to the strategy during which a given node occurs.

References

[GT75] R. Goldblatt and S. Thomason. Axiomatic classes in propositional modal logic. In *Algebra and Logic*, volume 450 of *Lecture Notes in Mathematics*, pages 163–173. Springer berlin / Heidelberg, 1975.

[Jan97] David Janin. Automata, tableaus and a reduction theorem for fixpoint calculi in arbitrary complete lattices. In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science (LICS'97)*, 1997.

[JL03] David Janin and Giacomo Lenzi. On the relationship between moandic and weak monadic second order logic on arbitrary trees, with application to the mu-calculus. *Fudamenta Informaticae*, 61:247–265, 2003.

[JW95] David Janin and Igor Walukiewicz. Automata for the mu-calculus and related results. In *MFCS'95*, 1995.

[JW96] David Janin and Igor Walukiewicz. On the expressive completeness of the propositional modal mu-calculus and related results. In *CONCUR'96*, 1996.

[vB76] Johan van Benthem. *Modal correspondance theory*. PhD thesis, Mathematisch Instituut en Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.

[vB96] Johan van Benthem. *Exploring logical dynamics*. CSLI publications, Stanford, 1996.
[Wal02] Igor Walukiewicz. Monadic second order logic on tree-like structures. *Theoretical Computer Science*, 275:311 – 346, 2002.

7 Appendix

7.1 Proof of Proposition 1

Proposition. *Each formula of the graded μ-calculus is equivalent to a ∇-formula of the graded μ-calculus. Each ∇-formula of the graded μ-calculus is equivalent to a graded μ-calculus formula.*

Proof. First, we show that each formula φ of the graded μ-calculus is equivalent to a ∇-formula of the graded μ-calculus. The proof is by induction on the complexity of φ. We only treat the cases when φ is of the form $\Diamond^k\psi$ or $\Box^k\psi$ (which are the only ones not straightforward). One can check that $\Diamond^k\psi$ is equivalent to $\top \bullet \nabla^g(\{\varphi, \ldots, \varphi\}, \{\top\})$, where the multiplicity of φ is $k+1$. The formula $\Box\psi$ is equivalent to $\top \bullet \nabla^g(\{\top, \ldots, \top\}, \{\psi\})$, where the multiplicity of $\top$ is k.

Second, we have to verify that each ∇-formula of the graded μ-calculus is equivalent to a graded μ-calculus formula. This can be achieved by showing that each ∇-formula of the graded μ-calculus is equivalent to an **MSO** formula and using the fact that **MSO** and graded μ-calculus have the same expressive power on trees (seen as Kripke models) [Wal02].

7.2 Proof of Theorem 3

Theorem. *Each formula of the graded μ-calculus is equivalent to a formula of the graded μ-calculus in disjunctive normal form.*

This follows from an application of a result proved by David Janin in [Jan97]. We recall the definitions required to state the main result of [Jan97].

Fixpoint algebras A *signature* is a set S of function symbols equipped with an arity function $\rho : S \to \mathbb{N}$. Over a signature S a *fixpoint algebra* is a complete lattice $(M, \vee_M, \wedge_M)$ with bottom and top element denoted by $\bot_M$ and $\top_M$ together with, for any symbol $f \in S$, a monotonic increasing function $f_M; M^{\rho(f)} \to M$. we let $\mathcal{F}_S$ denote the set of syntactic functions one can build from signature S and composition.

Over a signature S and a set Var of variables, the *fixpoint formulas* are defined by

$$\varphi \ ::= x \mid f(\varphi, \ldots, \varphi) \mid \mu x.\varphi \mid \nu x.\varphi,$$

where $x \in Var$ and $f \in S$. The notion of *truth* of a formula at a point in a fixpoint algebra M under an assignemnt $\tau : Var \to M$ and *equivalence* of two formulas over a class of fixpoint algebras are defined in a standard way, similarly to the same notions for μ-calculus (for more details, see [Jan97]).

Given a classs of fixpoint algebras $\mathcal{C}$, we say that the meet operator $\wedge$ *commutes* with S on $\mathcal{C}$ when, for any finite multiset $\{f_i \mid i \in \{1, \dots, n\}\}$ of functional symbols of $S \backslash \{\wedge\}$, there exists a function $G \in \mathcal{F}_S$ built without the symbol $\wedge$ such that an equation of the form:

$$\bigwedge \{f_i(\boldsymbol{u}_i) \mid i \in \{1, \dots, n\}\} = G(\bigwedge \boldsymbol{v}_1, \dots, \bigwedge \boldsymbol{v}_m).$$

holds on $\mathcal{C}$, where:

- $\boldsymbol{u}_i$s are vectors of distinct variables of the appropriate length,
- $\boldsymbol{v}_j$s are vectors of distinct variables taken amon those appearing in $\boldsymbol{u}_i$s,
- $\bigwedge \boldsymbol{v_j}$s denote the greastest lower bound applied to the set of all variables occurring in $\boldsymbol{v}_j$.

Theorem 4. *[Jan97] When the meet operator $\wedge$ commutes with S on $\mathcal{C}$, any closed fixpoint formula is equivalent over $\mathcal{C}$ to a formula built without the symbol $\wedge$.*

Theorem 3 follows from the Theorem above together with the following Lemma.

Lemma. *Let φ be a formula of the form $\bigwedge \{\Pi_i \bullet \nabla^g(\Phi_i, \Psi_i) \mid 1 \leq i \leq l\}$, where for all $i \in \{1, \dots, n\}$, Π_i is a conjunction of literals, Φ_i is a multi-set of graded μ-formulas and Ψ_i is a set of graded μ-formulas. Then φ is equivalent to the formula ψ defined as*

$$\bigwedge \{\Pi_i \mid 1 \leq i \leq l\} \wedge \bigvee \{\nabla^g(\Phi_R, \Psi) \mid R \subseteq \Phi_1^* \times \cdots \times \Phi_l^* \text{ and }$$
$$\text{for all } i \in \{1, \dots, n\} \text{and all } \varphi \in \Phi_i, \varphi \text{ appears exactly once in } R\},$$

where Ψ is the formula $\bigwedge \{\bigvee \Psi_i \mid 1 \leq i \leq l\}$ and Φ_R is the multiset of formulas

$$\{\bigwedge \{\varphi_i \mid i \in \{1, \dots, n\}, \varphi_i \neq *\} \wedge$$
$$\bigwedge \{\bigvee \Psi_i \mid i \in \{1, \dots, l\}, \varphi_i = *\} \mid (\varphi_1, \dots, \varphi_l) \in R\}.$$

Proof. We only give a sketch of the proof of the hardest implication; that is, φ implies ψ. Let $\mathcal{M} = (W, R, V)$ be a Kripke model, let $\tau : Var \to \mathcal{P}(W)$ be an assignment and let $w \in W$ be such that $\mathcal{M}, w \Vdash_\tau \varphi$. So for all $i \in \{1, \dots, l\}$, $\mathcal{M}, w \Vdash_\tau \Pi_i \bullet \nabla^g(\Phi_i, \Psi_i)$. In particular, there exists a set $W_i = \{w^i_\varphi \text{ successor of } w \mid \varphi \in \Phi_i\}$ such that the size of W_i is $card(\Phi_i)$ and for all $\varphi \in \Phi_i$, $\mathcal{M}, w_\varphi \Vdash_\tau \varphi$. Moreover, for all successors u of w such that u does not belong to W_i, we have $\mathcal{M}, u \Vdash_\tau \bigvee \Psi_i$. Now, to each $u \in \bigcup \{W_i \mid i \in \{1, \dots, l\}\}$, we define a l-tuple $f(u) = (\varphi_1, \dots, \varphi_l)$ such that for all $i \in \{1, \dots, l\}$, we have $\varphi_i = \varphi$ if $w^i_\varphi = u$ and $\varphi_i = *$, if there is no φ such that $w^i_\varphi = u$. The l-tuple $f(u)$ is uniquely defined as the size of W_i is $card(\Phi_i)$. Finally, we define the relation R by $\{f(u) \mid wRu\}$. Using the properties of the sets $W_1, \dots, W_l$, we can show that for all $i \in \{1, \dots, n\}$ and all $\varphi \in \Phi_i$, φ appears exactly once in R. To finish the proof that φ implies ψ, it is sufficient to show that $\mathcal{M}, w \Vdash_\tau \bigwedge \{\Pi_i \mid 1 \leq i \leq l\} \wedge \nabla^g(\Phi_R, \Psi)$. This is left to the reader.

7.3 Details for the proof of Proposition 3

Proposition. *Let α be an **MSO** formula. The following are equivalent:*

- *α is equivalent on trees to a $\Box$-graded formula,*
- *α is equivalent on trees to a formula in $\mu GL^{\nabla'}$,*
- *α is preserved under p-morphic images for trees.*

Proof. First, we show that each $\Box$-graded formula is equivalent on trees to a fixpoint ∇'-formula. For, it is enough to prove that the operators $\Box^k$ and $\Diamond$ can be expressed using the boolean operators and the operator ∇'. Fix a $\Box$-graded formula φ. It is routine to check that $\Diamond\varphi$ is equivalent to $\nabla'(\{\varphi\};\{\top\})$. If $k=0$, then $\Box^k\varphi$ is equivalent to $\bot$. If $k=1$, $\Box^k\varphi$ is equivalent to $\nabla'(\emptyset,\{\varphi\})$. Finally, if $k>1$, $\Box^k\varphi$ is equivalent to $\nabla'(\{\top,\dots,\top\};\{\varphi\})$, where the mutiplicity of $\top$ is equal to $k-1$.

Second, we prove that each fixpoint ∇'-formula is equivalent to a $\Box$-graded formula. It is sufficient to show that each formula $\nabla'(\Phi;\Psi)$ can be expressed using the operators $\Diamond$ and $\Box^k$, the boolean operators and the formulas in Φ and Ψ.

So let $\Phi=\{\varphi_1,\dots,\varphi_k\}$ be a multi-set of formulas and let $\Psi=\{\psi_1,\dots,\psi_l\}$ be a set of formula. We start by fixing some notations. If i is a natural number greater or equal to 1, we denote by $[i]$ the set $\{1,\dots,i\}$. Given a number $k'\leq k$, a number $i\leq k'$, a surjective map $f:[k]\to[k']$ and a map $g:[k']\to\mathcal{P}[l]$, we define the formula $\delta(i,k',f,g)$ as the following formula:

$$\bigwedge\{\varphi_j \mid f(j)=i\}\wedge\bigwedge\{\psi_j \mid j\in g(i)\}.$$

We also define the formula $\theta(k',f,g)$ as the formula :

$$\bigwedge\{\Diamond\delta(i,k',f,g) \mid 1\leq i\leq k'\}.$$

Finally, we define the natural number $n(k',g)$ as the number $k'-|\{j:g(j)\neq\emptyset\}|$. Now we will prove that the formula $\nabla'(\Phi;\Psi)$ is equivalent to the formula θ given by:

$$\bigvee\left\{\theta(k',f,g)\wedge\Box^{n(k',g)}\bigvee\Psi \mid k'\leq k, f:[k]\to[k']\text{ surjective}, g\mid[k']\to\mathcal{P}[l]\right\}.$$

First, we show that $\nabla'(\Phi;\Psi)$ implies θ. Fix a point u in a model over *Prop* and suppose that $\nabla'(\Phi;\Psi)$ is true at u. By definition of ∇', there exist children $u_1,\dots,u_k$ of u such that for all i in $\{1,\dots,k\}$, we have that φ_i is true at u_i and for all children v of u such that v does not belong to $\{u_1,\dots,u_n\}$, $\bigvee\Psi$ is true at v. We let $u'_1,\dots,u'_{k'}$ be children of u such that $\{u_1,\dots,u_k\}=\{u'_1,\dots,u'_{k'}\}$ and for all $i\neq j$, u'_i and u'_j are distinct. Now we let $f:[k]\to[k']$ be the map such that for all i and j, $f(i)=j$ iff $u_i=u'_j$. We also let $g:[k']\to\mathcal{P}[l]$ be the map such that for all i, $g(i)=\{j\mid\psi_j\text{ is true at }u'_i\}$.

With these definitions of k', f and g, it is possible to show that $\theta(k',f,g)\wedge\Box^{n(k',g)}(\bigvee\Psi)$ is true at u. Fix i in $[l]$. It is routine to check that the formula

$\delta(i,k',f,g)$ is true at the point u'_i. Therefore, $\Diamond\delta(i,k',f,g)$ is true at u. It follows from the definition of $\theta(k',f,g)$ that $\theta(k',f,g)$ is true at u. It remains to check that $\Box^{n(k',g)}(\bigvee\Psi)$ is true at u. That is, there are at most $n(k',g)$ children of u, at which $\bigvee\Psi$ is false. We know that for all children v of u which are not in $\{u'_1,\dots,u'_{k'}\}$, $\bigvee\Psi$ is true at v. Next, for the children $u'_1,\dots,u'_{k'}$ of u, it follows from the definition of g that $\bigvee\Psi$ is true at u'_i iff $g(i)\neq\emptyset$. Putting that together with the definition of $n(k',g)$, we obtain that there are exactly $n(k',g)$ points in $\{u'_1,\dots,u'_{k'}\}$ at which $\bigvee\Psi$ is not true and this finishes the proof that $\nabla'(\Phi;\Psi)$ implies θ.

For the other direction, suppose that θ is true at a point u in a model over *Prop*. We have to show that $\nabla'(\Phi,\Psi)$ is true at u. First, we prove that for all φ in Φ, there is a child u_φ of r, at which φ is true. Fix a formula φ in Φ. There is a natural number j such that $\varphi=\varphi_j$. By definition of θ, there are $k'\leq k$, a surjective map $f:[k]\to[k']$ and a map $g:[k']\to\mathcal{P}[l]$ such that $\bigwedge\{\Diamond\delta(i,k',f,g)\mid 1\leq i\leq k'\}\wedge\Box^{n(k',g)}(\bigvee\Psi)$ is true at u. Fix i in $[k']$. Since $\Diamond\delta(i,k',f,g)$ is true at r, there is a child v_i of u at which $\delta(i,k',f,g)$ is true. In particular, φ_j is true at $v_{f(j)}$. So we can define u_φ as $v_{f(j)}$.

In order to prove that $\nabla'(\Phi,\Psi)$ is true at u, it remains to check that for all v children of u such that v does not belong to $\{u_\varphi\mid\varphi\in\Phi\}$, $\bigvee\Psi$ is true at v. By definition of the u_φ's, this is equivalent to prove that for v children of u such that v does not belong to $\{v_i\mid i\in[k']\}$, $\bigvee\Psi$ is true at v. As before, we have that $\bigvee\Psi$ is true at v_i iff $g(i)\neq\emptyset$. Putting that together with the definition of $n(k',g)$, we know that there are exactly $n(k',g)$ points in the set $\{v_i\mid i\in[k']\}$ at which $\bigvee\Psi$ is false. Now, recall that $\Box^{n(k',g)}(\bigvee\Psi)$ is true at u. That is, u has at most $n(k',g)$ points at which $\bigvee\Psi$ is false. Gathering everything together, we obtain that $\bigvee\Psi$ is true at all points not in $\{v_i\mid i\in[k']\}$ and this finishes the proof that each fixpoint ∇'-formula is equivalent to a $\Box$-graded formula.

We showed earlier that an **MSO** formula which is preserved under p-morphic images for tree is equivalent on trees to a ∇'-graded fixpoint formula. Moreover, it is routine to show that a $\Box$-graded fixpoint formula is preserved under p-morphic image. This finishes the proof.