The modal mu-calculus alternation hierarchy is strict

J. C. Bradfield

Laboratory for Foundations of Computer Science, Department of Computer Science, University of Edinburgh, Edinburgh, United Kingdom, EH9 3JZ (email: jcb@dcs.ed.ac.uk)

Abstract: One of the open questions about the modal mu-calculus is whether the alternation hierarchy collapses; that is, whether all modal fixpoint properties can be expressed with only a few alternations of least and greatest fixpoints. In this paper, we resolve this question by showing that the hierarchy does not collapse.

1 Introduction

The modal mu-calculus, or Hennessy–Milner logic with fixpoints, is a popular logic for expressing temporal properties of systems. It was first studied by Kozen in [Koz83], and since then there has been much work on both theoretical and practical aspects of the logic. The feature of the logic that gives it both its simplicity and its power is that it is possible to have mutually dependent minimal and maximal fixpoint operators. This makes it simple, as the fixpoints are the only non-first-order operators, and powerful, as by such nesting one can express complex properties such as 'infinitely often' and fairness. A measure of the complexity of a formula is the *alternation depth*, that is, the number of alternating blocks of minimal/maximal fixpoints. Formulae of alternation depth higher than 2 are notoriously hard to understand, and in practice one rarely produces them—not least because they are so hard to understand. It is therefore natural to wonder whether in fact higher alternation depths are needed—it could be the case that this *alternation hierarchy* collapses. Until now, the best result was that we need both min–max and max–min formulae of depth 2, which was proven by Arnold and Niwiński in [ArN90] using automata-theoretic methods and results of Rabin [Rab70].

This question is given additional spice by the consideration of complexity issues. All known algorithms for model-checking modal mu-calculus properties are exponential in the alternation depth d. The natural algorithm, by Emerson and Lei [EmL86], was $O(n^d)$; this has recently been improved to $O(n^{d/2})$ by Long *et al.*[LB+94] On the other hand, the problem is in NP (due to Emerson, Jutla and Sistla [EJS93], and more directly seen by Stirling's game-theoretic approach [Sti95]), and since the logic is closed under negation, the problem is in NP $\cap$ co-NP, which suggests that it may well be in P, even if P $\neq$ NP. If the alternation hierarchy is strict, then we know that algorithms exponential in the alternation depth cannot be made polynomial just by reducing all formulae to alternation depth 3 (say) equivalents. Of course, if the hierarchy did collapse, we would not necessarily immediately get a polynomial solution, since the reduction might involve a large blow-up in the size of the formula.

The contribution of this paper is to resolve the question by establishing the strictness of the hierarchy. The technique is slightly unusual, being not at all automata-theoretic; instead, we analyse the descriptive complexity, in the sense of effective descriptive set theory, of properties in the modal mu-calculus, and then code suitable arithmetic formulae into a certain transition system in order to achieve the upper bounds established for the complexity of properties. In previous work [Bra96], transferring standard hierarchies allowed us to re-prove Arnold and Niwiński's result, and obtain some other mildly interesting results, such as a Δ^1_2 upper bound on the complexity of modal mu-calculus properties. In this paper, we transfer a similar alternation hierarchy for arithmetic with fixpoints, and thereby show the strictness of the modal mu-calculus hierarchy.

The remainder of this paper is thus: in section 2 we define the modal mu-calculus, and arithmetic with fixpoints, and present some of the results on which we rely. In section 3, we establish the non-collapse of the 'simple' alternation hierarchy, which we extend to the real alternation hierarchy in section 4. Section 5 is the conclusion, and the appendix (section A) gives a summary of the mu-arithmetic hierarchy result that we transfer.

2 Preliminaries

2.1 Modal mu-calculus

We assume some familiarity with the modal mu-calculus, so in this section we give brief definitions to establish notations and conventions. Expository material on the modal mu-calculus may be found in [Bra91,Sti91].

The modal mu-calculus, with respect to some countable set $\mathscr{L}$ of *labels*, has formulae Φ defined inductively thus: variables Z and the boolean constants tt, ff are formulae; if Φ_1 and Φ_2 are formulae, so are $\Phi_1 \vee \Phi_2$ and $\Phi_1 \wedge \Phi_2$; if Φ is a formula and l a label, then $[l]\Phi$ and $\langle l\rangle\Phi$ are formulae; and if Φ is a formula and Z a variable, then $\mu Z.\, \Phi$ and $\nu Z.\, \Phi$ are formulae.

Note that we adopt the convention that the scope of the binding operators μ and ν extends as far as possible. For consistency, we also apply this convention to the $\forall$ and $\exists$ of first-order logic, writing $\forall x.\, (\exists y.\, P) \vee Q$ rather than the logicians' traditional $\forall x\, [\exists y\, [P] \vee Q]$.

Observe that negation is not in the language, but any closed mu-formula can be negated by using the usual De Morgan dualities—μ and ν are dual by $\neg\mu Z.\, \Phi(Z) = \nu Z.\, \neg\Phi(\neg Z)$. Where necessary, we shall assume that free variables can be negated just by adjusting the valuation. We shall use $\Rightarrow$ etc. freely, though we must ensure that bound variables only occur positively.

Given a labelled transition system $\mathscr{T} = (\mathscr{S}, \mathscr{L}, \longrightarrow)$, where $\mathscr{S}$ is a set of states, $\mathscr{L}$ a set of labels, and $\longrightarrow \subseteq \mathscr{S} \times \mathscr{L} \times \mathscr{S}$ is the transition relation (we write $s \xrightarrow{l} s'$), and given also a valuation $\mathscr{V}$ assigning subsets of $\mathscr{S}$ to variables, the denotation $\|\Phi\|^{\mathscr{T}}_{\mathscr{V}} \subseteq \mathscr{S}$ of a mu-calculus formula Φ is defined in the obvious way for the variables and booleans, for the modalities by

$$\|[l]\Phi\|^{\mathcal{T}}_{\mathcal{V}} = \{\, s \mid \forall s'.\, s \xrightarrow{l} s' \;\Rightarrow\; s' \in \|\Phi\|^{\mathcal{T}}_{\mathcal{V}} \,\}$$
$$\|\langle l\rangle\Phi\|^{\mathcal{T}}_{\mathcal{V}} = \{\, s \mid \exists s'.\, s \xrightarrow{l} s' \,\wedge\, s' \in \|\Phi\|^{\mathcal{T}}_{\mathcal{V}} \,\},$$

and for the fixpoints by

$$\|\mu Z.\, \Phi\|^{\mathcal{T}}_{\mathcal{V}} = \bigcap \{\, S \subseteq \mathcal{S} \mid \|\Phi\|^{\mathcal{T}}_{\mathcal{V}[Z:=S]} \subseteq S \,\}$$
$$\|\nu Z.\, \Phi\|^{\mathcal{T}}_{\mathcal{V}} = \bigcup \{\, S \subseteq \mathcal{S} \mid S \subseteq \|\Phi\|^{\mathcal{T}}_{\mathcal{V}[Z:=S]} \,\}\,.$$

We shall drop the $^{\mathcal{T}}$ and $_{\mathcal{V}}$ whenever they are obvious, and accordingly write $\|\Phi\|_{Z:=S}$ for the denotation of Φ when the variable Z has value S.

It is often useful to think of μZ and νZ as meaning respectively finite and infinite looping from Z back to μZ (νZ) as one 'follows a path of the system through the formula'. Examples of properties expressible by the mu-calculus are 'always (on a-paths) P', as $\nu Z.\, P \wedge [a]Z$, 'eventually (on a-paths) P', as $\mu Z.\, P \vee \langle a\rangle Z$, and 'there is an $\{a,b\}$-path along which b happens infinitely often', as $\nu Y.\, \mu Z.\, \langle b\rangle Y \vee \langle a\rangle Z$. (For the latter, we can loop around Y for ever, but each internal loop round Z must terminate.)

There are several notions of alternation. The naive notion is simply to count syntactic alternations of μ and ν, resulting in the following definition: A formula Φ is said to be in the classes $\Sigma^{S\mu}_0$ and $\Pi^{S\mu}_0$ iff it contains no fixpoint operators ('S' for 'simple' or 'syntactic'). The class $\Sigma^{S\mu}_{n+1}$ is the least class containing $\Sigma^{S\mu}_n \cup \Pi^{S\mu}_n$ and closed under the following operations: (i) application of the boolean and modal combinators; (ii) the formation of $\mu Z.\, \Phi$, where $\Phi \in \Sigma^{S\mu}_{n+1}$. Dually, to form the class $\Pi^{S\mu}_{n+1}$, take $\Sigma^{S\mu}_n \cup \Pi^{S\mu}_n$, and close under (i) boolean and modal combinators, (ii) $\nu Z.\, \Phi$, for $\Phi \in \Pi^{S\mu}_{n+1}$. Thus the examples above are in $\Pi^{S\mu}_1$, $\Sigma^{S\mu}_1$, and $\Pi^{S\mu}_2$ (but not $\Sigma^{S\mu}_2$) respectively. We shall say a formula is *strict* $\Sigma^{S\mu}_n$ if it is in $\Sigma^{S\mu}_n - \Pi^{S\mu}_n$.

However, this simple notion of alternation is not what we are concerned with, since it does not capture the complexity of feedback between fixpoints: it does not distinguish these two formulae

$$\Upsilon_1 = \nu Y.\, \mu Z.\, \langle a\rangle Y \wedge (P \vee \langle b\rangle Z)$$
$$\Upsilon_2 = \nu Y.\, \langle a\rangle Y \wedge (\mu Z.\, P \vee \langle b\rangle Z)$$

—both are strict $\Pi^{S\mu}_2$, but the first is more complex, as the inner fixpoint depends on the outer, whereas in the second, the inner fixpoint is self-contained. To take account of this, we need a stronger definition, for which there is more than one candidate. The most common version is that of Emerson–Lei [EmL86]; however, a more refined notion was used by Niwiński [Niw86], and since this captures the intuitive notion better than Emerson–Lei, as well as providing a better complexity measure, we shall follow Niwiński. (For an explanation of the differences, see the end of section 4—our results trivially imply the non-collapse of the hierarchy of [EmL86].) In fact, it is possible with a very minor modification to strengthen Niwiński's definition a little further; see the end of section 4.

A formula Φ is said to be in the classes $\Sigma^{N\mu}_0$ and $\Pi^{N\mu}_0$ iff it contains no fixpoint operators. To form the class $\Sigma^{N\mu}_{n+1}$, take $\Sigma^{N\mu}_n \cup \Pi^{N\mu}_n$, and close under (i) boolean and modal combinators, (ii) $\mu Z.\, \Phi$, for $\Phi \in \Sigma^{N\mu}_{n+1}$, and (iii) substitution of $\Phi' \in \Sigma^{N\mu}_{n+1}$ for a free variable of $\Phi \in \Sigma^{N\mu}_{n+1}$ provided that no free variable of Φ' is captured by Φ; and dually for $\Pi^{N\mu}_{n+1}$.

Now we can distinguish Υ_1 and Υ_2: both are in $\Pi_2^{N\mu}$, but the 'non-alternating' Υ_2 is also in $\Sigma_2^{N\mu}$, for the following reason: $\mu Z.\, P \vee \langle b \rangle Z$ is in $\Sigma_1^{N\mu}$, and so also in $\Sigma_2^{N\mu}$; and $\nu Y.\, \langle a \rangle Y \wedge W$ is in $\Pi_1^{N\mu}$, and hence in $\Sigma_2^{N\mu}$ by rule (i); and Υ_2 is the result of substituting the former for the free variable W of the latter, and so is in $\Sigma_2^{N\mu}$ by rule (iii).

Intuitively, we are allowed to have arbitrary syntactic alternation, as long as the real semantic dependency between the various fixpoints is restricted. To take the simplest example, $\mu X_1.\, \nu X_2.\, \mu X_3.\, \ldots \nu X_{2n}.\, \mu X_{2n+1}.\, X_1 \vee X_{2n+1}$ has syntactic alternation depth $2n+1$, but its real alternation depth is just 1, since all we have is an inner minimal fixpoint depending on an outer minimal fixpoint, all the other fixpoints being vacuous. The definition via restricted substitution is the simplest way of capturing this notion.

The *(Niwiński) alternation depth* of a formula Φ is the least n such that $\Phi \in \Sigma_{n+1}^{N\mu} \cap \Pi_{n+1}^{N\mu}$.

The relationship between the simple and Niwiński hierarchies is that $\Sigma_n^{S\mu} \subseteq \Sigma_n^{N\mu}$, but $\Sigma_2^{N\mu}$ has non-empty intersection with every $\Sigma_n^{S\mu} - \Sigma_{n-1}^{S\mu}$; thus the non-collapse of the simple hierarchy is not sufficient to give the non-collapse of the (Niwiński) alternation hierarchy, which is the real question of interest.

Given a formula $\mu Z.\, \Phi(Z)$ and a model, the *approximants* $\mu^\zeta Z.\, \Phi$, for ζ an ordinal, are defined recursively by $\mu^\zeta Z.\, \Phi = \|\Phi\|_{Z:=S}$ where $S = \bigcup_{\eta<\zeta} \mu^\eta Z.\, \Phi$, and dually for $\nu Z.\, \Phi$. We may also write Φ^ζ or Z^ζ for $\mu^\zeta Z.\, \Phi$. We shall write $Z^{<\zeta}$ for $\bigcup_{\eta<\zeta} Z^\eta$. (This is the notation introduced in [Mos74]; earlier work, and also most work in the modal mu-calculus, uses a different notation, with the effect that their Z^ζ is our $Z^{<\zeta}$. For our purposes, Moschovakis' notation is a little neater.)

Since mu-calculus formulae are monotonic in their variables, for successor ordinals we have $\mu^{\zeta+1} Z.\, \Phi = \|\Phi\|_{Z:=\mu^\zeta Z.\, \Phi}$. Further, by the standard Tarski fixpoint theorem we have $\|\mu Z.\, \Phi\| = \bigcup_{\zeta \in \mathrm{Ord}} \mu^\zeta Z.\, \Phi$; the smallest κ such that $\|\mu Z.\, \Phi\| = \bigcup_{\zeta<\kappa} \mu^\zeta Z.\, \Phi$ is called the *closure ordinal* of $\mu Z.\, \Phi$.

2.2 The arithmetic mu-calculus

In [Lub93] (first presented at LICS '89), Robert Lubarsky studies the logic given by adding fixpoint constructors to first-order arithmetic. Precisely, the logic ('mu-arithmetic' for short) has as basic symbols the following: function symbols f, g, h; predicate symbols P, Q, R; first-order variables x, y, z; set variables X, Y, Z; and the symbols $\vee, \wedge, \exists, \forall, \mu, \nu$, $\neg, \in$. As with the modal mu-calculus, $\neg$ can be pushed inwards to apply only to atomic formulae, by De Morgan duality.

The language has expressions of three kinds, individual terms, set terms, and formulae. The individual terms comprise the usual terms of first-order logic. The set terms comprise set variables and expressions $\mu(x, X).\, \phi$ and $\nu(x, X).\, \phi$, where X occurs positively in ϕ. Here μ binds both an individual variable and a set variable; henceforth we shall write just $\mu X.\, \phi$, and assume that the individual variable is the lower-case of the set variable. The formulae are built by the usual first-order construction, together with the rule that if τ is an individual term and Ξ is a set term, then $\tau \in \Xi$ is a formula.

This language is interpreted over a structure $\mathscr{M}$ for its first-order part. The semantics of the first-order connectives is as usual; $\tau \in \Xi$ is interpreted naturally; and the set term $\mu X.\, \phi(x, X)$ is interpreted as the least fixpoint of the functional $\mathbf{X} \mapsto \{\, m \in \mathscr{M} \mid \mathscr{M} \models$

$\phi(m, \mathbf{X})\,\}$ (where $\mathbf{X} \subseteq \mathscr{M}$). As with the modal mu-calculus, we define approximants, and indeed the approximant approach is, of course, essential to Lubarsky's method: X^ζ is now the set $\{\, m \in \mathscr{M} \mid \mathscr{M} \vDash \phi(m, \bigcup_{\xi<\zeta} X^\xi)\,\}$, and the interpretation of $\mu X.\,\phi$ is then $\bigcup_{\zeta\in\mathrm{Ord}} X^\zeta$; and dually for the greatest fixpoint. Closure ordinals are as for the modal mu-calculus.

Henceforth we shall take $\mathscr{M}$ to be the structure $\mathbb{N}$ of first-order arithmetic with recursive functions and predicates. In particular, let $\langle -,- \rangle$, $(-)_0$ and $(-)_1$ be standard pairing and unpairing functions.

The simplest examples of mu-arithmetic just use least fixpoints to represent an inductive definition. For example, $\mu X.\, x = 0 \vee (x > 1 \wedge (x-2) \in X)$ is the set of even numbers. Of course, the even numbers are also the complement of the odd numbers: the odd numbers are defined by $\mu X.\, x = 1 \vee (x > 1 \wedge (x-2) \in X)$, so by negating we can express the even numbers as a maximal fixpoint $\nu X.\, x \neq 1 \wedge (x > 1 \Rightarrow (x-2) \in X)$. To produce natural examples involving alternating fixpoints is rather difficult, since even one induction is already very powerful, and most natural mathematical objects are simple.

Lubarsky establishes a normal form theorem. The μ-normal form is defined thus: a set term is in μ-normal form if it is a set variable, or of the form $\mu X.\,\phi$ or $\nu X.\,\phi$ with ϕ in μ-normal form. A formula is in μ-normal form if it is quantifier-free, or of the form $\tau \in \Xi$ with Ξ in μ-normal form, or of the form $\exists x.\,\phi$ or $\forall x.\,\phi$ with ϕ in μ-normal form.

In the presence of a pairing function, as in $\mathbb{N}$, it is further possible to move first-order quantifiers inside fixpoint quantifiers and produce a *pair-normal* form: a pair-normal formula is either first-order or of the form $\tau \in \Xi$ for pair-normal Ξ, and a pair-normal Ξ is $\mu X.\,\phi$ where ϕ is either first-order or $\tau' \in \nu Y.\,\psi$ for pair-normal ψ (and dually). Thus there is an alternating string of fixpoints, followed by a first-order formula.

At this point, we should mention that we are modifying slightly two of the definitions in [Lub93]: the pair-normal form definition above is not precisely that in [Lub93], and the definition of alternation in [Lub93] is not exactly the same as that we give below. However, all the proofs in [Lub93] work also with our definitions. These modifications are discussed in the Appendix. That said, we now have the following theorem from [Lub93]:

Lemma 1 *Every formula and set term is semantically equivalent to one in pair-normal form.* □

One can define the syntactic alternation classes for arithmetic just as for the modal mu-calculus: First-order formulae are $\Sigma_0^{S\mu}$ and $\Pi_0^{S\mu}$, as are set variables. The $\Sigma_{n+1}^{S\mu}$ formulae and set terms are formed from the $\Sigma_n^{S\mu} \cup \Pi_n^{S\mu}$ formulae and set terms by closing under (i) the first-order connectives and (ii) forming $\mu X.\,\phi$ for $\phi \in \Sigma_{n+1}^{S\mu}$.

We can now strengthen the preceding lemma to say that the conversion to pair-normal form does not change the alternation class. This is a crucial result, and we therefore sketch the proof, taking into account the minor modifications mentioned above. We shall be terse, and omit details; for some elaboration, see pp. 298–299 of [Lub93].

Lemma 2 *If ϕ is $\Sigma_n^{S\mu}$ ($\Pi_n^{S\mu}$), it is semantically equivalent to a pair-normal formula that is also $\Sigma_n^{S\mu}$ ($\Pi_n^{S\mu}$).*

Proof. We proceed by induction on n, and by structural induction on formulae and set terms.

For a set term $\mu X.\,\phi$, we assume inductively that ϕ is pair-normal; then we are already pair-normal unless ϕ is $\tau \in \mu Y.\,\psi$. In that case, the translation pairs up X and Y into W in the natural way, so that $m \in X$ iff $\langle 0, m\rangle \in W$ and $n \in Y$ iff $\langle 1, \langle x, n\rangle\rangle \in W$ (remember that τ, ψ and Y may depend on the individual variable x as well as the set variable X). Note that although Y depends on both x and X, we have only explicitly coded the dependency on x. By standard monotonicity arguments about adjacent fixpoints of the same sign, the dependency on X can be ignored. Thus we translate the original term into

$$\mu W.\,((w)_0 = 0 \wedge \langle 1, \langle (w)_1, \tau'\rangle\rangle \in W) \vee ((w)_0 = 1 \wedge \psi')$$

where τ' is obtained from τ by replacing every occurrence of x by $(w)_1$, and ψ' is obtained from ψ by replacing every '$\rho \in X$' by '$\langle 0, \rho\rangle \in W$', and every '$\rho \in Y$' by '$\langle 1, \langle x, \rho\rangle\rangle \in W$', and then every x by $((w)_1)_0$ and every y by $((w)_1)_1$. This procedure clearly preserves the level in the hierarchy. Now, as ϕ was pair-normal, its body ψ was a $\Pi^{S\mu}_{n-1}$ formula; hence the body of $\mu W.\,\ldots$ is $\Pi^{S\mu}_{n-1}$, and by induction can be transformed into a $\Pi^{S\mu}_{n-1}$ pair-normal formula, and we are done.

Now we consider formulae. For the case $\tau \in \Xi$, inductively transform Ξ to its pair-normal form Ξ', as in the previous paragraph. Note that if the pairing of adjacent fixpoints above is required, then we need to write $\langle 0, \tau\rangle \in \Xi'$, as τ is supposed to be in X, not W.

The booleans are easy, since $(\tau \in \mu Z.\,\phi) \wedge \psi$ is equivalent to $\tau \in \mu Z.\,\phi \wedge \psi$. A little care is needed, though: if we have the conjunction of two fixpoints, one μ and the other ν, we need to put the μ on the outside if we're trying to make it $\Sigma^{S\mu}_n$, and the ν if we're trying to make it $\Pi^{S\mu}_n$. Thus a formula that is both $\Sigma^{S\mu}_n$ and $\Pi^{S\mu}_n$ has a $\Sigma^{S\mu}_n$ pair-normal form and also a $\Pi^{S\mu}_n$ pair-normal form, but does not have a pair-normal form that is both $\Sigma^{S\mu}_n$ and $\Pi^{S\mu}_n$.

For formulae $\exists x.\,\phi$, assume that ϕ is $\tau \in \mu Y.\,\psi$. The existential quantifier is pushed inside the fixpoint by a similar construction to that used in the case of set terms: let W be a new variable, and build ψ' from ψ exactly as before. Then the set term

$$\mu W.\,(w = \langle 0, 0\rangle \wedge \exists x.\,\langle 1, \langle x, \tau\rangle\rangle \in W) \vee ((w)_0 = 1 \wedge \psi')$$

contains $\langle 0, 0\rangle$ iff $\exists x.\,\phi$. Now the case of ϕ being $\tau \in \nu Y.\,\psi$ is similar.

Similarly for formulae $\forall x.\,\phi$.

So we see that the transformation makes no change to the $\Sigma^{S\mu}_n$ level, as claimed. $\square$

This is quite a complex construction, and some simple examples may be helpful. Firstly, consider 't is even and t is not a multiple of three'. If we use the inductive definition of multiples that we had before, we get

$$\begin{gathered}(t \in \mu X.\, x = 0 \vee (x > 1 \wedge (x-2) \in X)) \\ \wedge\, (t \in \nu Y.\, y \neq 0 \wedge (y > 2 \Rightarrow y - 3 \in Y))\,.\end{gathered}$$

As the two fixpoints are independent, we can move one inside the other to get

$$\begin{aligned}t \in \mu X.\, t \in \nu Y.\ &(x = 0 \vee (x > 1 \wedge (x-2) \in X)) \\ &\wedge\, (y \neq 0 \wedge (y > 2 \Rightarrow y - 3 \in Y))\,.\end{aligned}$$

As an example of the treatment of first-order quantifiers, consider 't is a composite number'. Let us again, for the purposes of exposition, use an inductive definition of multiple, but use an existential quantifier over possible factors, that is to say 'there is an $x > 1$ such that t is a multiple (> 1) of x':

$$\exists x.\, x > 1 \wedge t \in \mu Y.\, y = 2x \vee (y > 2x \wedge (y - x \in Y)).$$

Applying the construction given above yields, where for readability we write w_{10} etc. for $((w)_1)_0$ etc.:

$$\begin{aligned}\langle 0,0\rangle \in \mu W.\ (w = \langle 0,0\rangle \wedge \exists x.\, \langle 1, \langle x, t\rangle\rangle \in W)\\ \vee\ (w_0 = 1 \wedge w_{10} > 1 \wedge (w_{11} = 2w_{10} \vee\\ (w_{11} > 2w_{10} \wedge \langle 1, \langle w_{10}, w_{11} - w_{10}\rangle\rangle \in W)))\,.\end{aligned}$$

Here the meat of the inductive definition is the same as before, but it's now being carried on in the $(()_1)_1$ component of W, which is parametrized by the $(()_1)_0$ component representing x. The first line says, effectively, that the flag value $\langle 0,0\rangle$ is in W only if $\exists x.\, t \in \mu Y.\, \ldots$, and the second and third lines compute Y as the last component of W, with the constraint on x included in this computation.

Having established the normal form results, Lubarsky then proves

Theorem 3 *The hierarchy of the sets of integers definable by $\Sigma_n^{S\mu}$ formulae of the arithmetic mu-calculus is a strict hierarchy.* $\square$

The theorem is actually that a set of integers is $\Sigma_n^{S\mu}$ definable iff it is Σ_1 over the (least n-reflecting admissible ordinal)-th level of the constructible universe, but all we need is the existence and strictness of the hierarchy. We include in the appendix a summary of the notions and proof ideas required for this very interesting theorem.

3 Transferring the simple hierarchy

Our aim is to transfer Lubarsky's result to the Niwiński alternation hierarchy for the modal mu-calculus. However, we shall start with the simple hierarchy, since a simple coding trick will then extend it to the Niwiński hierarchy. In order to establish our results, we shall work with a particular class of transition systems.

We define a *recursively presented transition system (r.p.t.s.)* to be a labelled transition system $(\mathscr{S}, \mathscr{L}, \longrightarrow)$ such that $\mathscr{S}$ is (recursively codable as) a recursive set of integers, $\mathscr{L}$ likewise, and $\longrightarrow$ is recursive. Henceforth we consider only recursively presented transition systems, with recursive valuations for the free variables.

The first result is simple:

Theorem 4 *For a modal mu-calculus formula $\Phi \in \Sigma_n^{S\mu}$, the denotation $\|\Phi\|$ in any r.p.t.s. is a $\Sigma_n^{S\mu}$ definable set of integers.*

Proof. All we have to do is translate the semantics of the modal mu-calculus into arithmetic. We translate $s \in \|\Phi\|$ into a mu-arithmetic formula $\phi(s)$ by induction on Φ. For variables X that were bound in the top-level formula, we translate $s \in \|X\|$ to $s \in X$; for a free variable P of the top-level formula (which has a recursive valuation by assumption), we translate to $P(s)$, where P is the appropriate recursive predicate. The booleans are obvious. For the modal operators, we have $s \in \|[a]\Psi\|$ iff $\forall t \in \mathscr{S}.\,(s \xrightarrow{a} t) \Rightarrow \psi$ where ψ is the translation of $t \in \|\Psi\|$, and dually. Finally, for the fixpoint operators, we translate $s \in \|\mu X.\,\Psi\|$ to $s \in \mu X.\,\psi$, where ψ is the translation of $x \in \|\Psi\|$, and dually.

It follows immediately from the definitions that $s \in \|\Phi\|$ iff $\phi(s)$, and that ϕ has the same $\Sigma^{S\mu}$ complexity as Φ. $\square$

The converse, showing that there are models of the modal mu-calculus with arbitrarily complex mu-arithmetic translations, is conceptually quite straightforward: we just define a suitable (and rather powerful!) transition system to code the evaluation of the target formula.

Theorem 5 *Let $\phi(z)$ be a $\Sigma_n^{S\mu}$ formula of mu-arithmetic. There is a r.p.t.s. $\mathscr{T}$ with recursive valuation $\mathscr{V}$ and a $\Sigma_n^{S\mu}$ formula Φ of the modal mu-calculus such that $\phi((s)_0)$ iff $s \in \|\Phi\|^{\mathscr{T}}_{\mathscr{V}}$. (Thus if ϕ is not $\Sigma_{n-1}^{S\mu}$-definable, neither is $\|\Phi\|$.)*

Proof. We assume that ϕ is alpha-converted so that all variables are distinct, and if there are any free set variables, we replace them with predicate symbols. Our transition system has as its states tuples of integers, one for each individual term in ϕ. Let s_τ denote the τ component of a state s, for a term τ. We shall construct Φ such that $s \in \|\Phi\|$ iff $\phi(s_z)$ (and in general if ϕ has multiple free individual variables $z_1, \ldots, z_k$, then $s \in \|\Phi\|$ iff $\phi(s_{z_1}, \ldots, s_{z_k})$.)

We shall write corner quotes $\ulcorner\ \urcorner$ to turn pieces of arithmetic syntax into modal mu-calculus variable and label symbols.

For every atomic formula $P(\tau)$ occurring in ϕ, we equip the modal mu-calculus with a variable $\ulcorner P(\tau)\urcorner$ such that $s \in \mathscr{V}(\ulcorner P(\tau)\urcorner)$ iff $P(s_\tau)$, and similarly for n-ary predicates.

For every individual term τ occurring in ϕ that has the form $f(\tau_1, \ldots, \tau_k)$, we equip $\mathscr{T}$ with a label $\ulcorner f(\tau_1, \ldots, \tau_k)\urcorner$, and the transitions given by $s \xrightarrow{\ulcorner f(\tau_1,\ldots,\tau_k)\urcorner} t$ iff $t_\tau = f(s_{\tau_1}, \ldots s_{\tau_k})$ and $t_{\tau''} = s_{\tau''}$ for every $\tau'' \neq \tau$.

For every first-order quantifier $\forall x$ or $\exists x$ in ϕ, we equip $\mathscr{T}$ with a label $\ulcorner x\urcorner$ and the transitions given by $s \xrightarrow{\ulcorner x\urcorner} t$ iff $t_\tau = s_\tau$ for every $\tau \neq x$.

This is now sufficient to deal with the first-order part of mu-arithmetic. Before considering the fixpoints, let us set down the construction of Φ for the first-order part:

If $\phi(z)$ is $\phi_1 \vee \phi_2$, then Φ is $\Phi_1 \vee \Phi_2$, and similarly for $\wedge$. If ϕ is $\forall x.\,\psi$, then Φ is $[\ulcorner x\urcorner]\Psi$, and if ϕ is $\exists x.\,\psi$, then Φ is $\langle\ulcorner x\urcorner\rangle\Psi$.

If ϕ is an atomic formula $P(\tau)$, then Φ is $\langle\tau*\rangle\ulcorner P(\tau)\urcorner$, where $\langle\tau*\rangle\Upsilon$ is defined thus: $\langle x*\rangle\Upsilon$ is just Υ, and $\langle f(\tau_1, \ldots \tau_k)*\rangle\Upsilon$ is $\langle\tau_1*\rangle\cdots\langle\tau_k*\rangle\langle\ulcorner f(\tau_1, \ldots, \tau_k)\urcorner\rangle\Upsilon$—that is, we compute the arguments of f and then f. Let $\xrightarrow{\tau*}$ be the corresponding sequence of transitions. (We could, of course, wrap all this computation up into one transition; it makes no difference.)

A simple induction now shows that $s \in \|\Phi\|_{\mathscr{V}}$ iff $\phi(s_z)$.

Thus we have used the transition system to code up all the computation in arithmetic, and the modal connectives to code the first-order quantifiers. To finish the job, we need to translate the fixpoint operators of arithmetic into the fixpoint operators of the modal mu-calculus.

For every fixpoint μX or νX occurring in ϕ, and for every formula $\tau \in X$ in ϕ, we equip $\mathscr{T}$ with a label $\ulcorner x(\tau)\urcorner$ (recall that $\mu X.\,\psi$ is short for $\mu(x, X).\,\psi$). The transitions are given by $s \xrightarrow{\ulcorner x(\tau)\urcorner} t$ iff $t_x = s_\tau$ and $t_{\tau'} = s_{\tau'}$ for every $\tau' \neq x$. (That is, we copy the value of τ to the 'input variable' x of the fixpoint.)

We now complete the translation: if ϕ is $\tau \in X$ for X a set variable, then Φ is $\langle\tau *\rangle\langle\ulcorner x(\tau)\urcorner\rangle X$. If ϕ is $\tau \in \mu X.\,\psi$, then Φ is $\langle\tau *\rangle\langle\ulcorner x(\tau)\urcorner\rangle\mu X.\,\Psi$; and similarly for ν.

The simple induction now extends in the obvious way. Since it is a little more complex than the first-order part, we give details.

We add to the inductive hypothesis the clause that if X is a free set variable of ϕ with valuation $\mathbf{X}$, then $s \in \|X\|$ iff $s_x \in \mathbf{X}$. Now if ϕ is $\tau \in X$, then $\phi(s_z)$ iff $\tau \in \mathbf{X}$ iff $\exists t.\, t_x = \tau \wedge t_x \in \mathbf{X}$ iff $\exists t.\, s \xrightarrow{\tau *} \xrightarrow{\ulcorner x(\tau)\urcorner} t \wedge t_x \in \mathbf{X}$ iff $s \in \|\langle\tau *\rangle\langle\ulcorner x(\tau)\urcorner\rangle X\|$, as required.

For the fixpoints themselves, we work by induction on the approximants. Consider the set term $\mu X.\,\psi$, and suppose by induction that $s \in \mu^\eta X.\,\Psi$ iff $s_x \in \mu^\eta X.\,\psi$, for $\eta < \zeta$. Then $s \in \mu^\zeta X.\,\Psi$ iff $s \in \|\Psi\|_{X:=X^{<\zeta}}$ iff $\psi(s_z, s_x)$ (where X is valued at $\mathbf{X} = X^{<\zeta}$ in arithmetic) iff $s_x \in \mu^\zeta X.\,\psi$. Thus the modal and arithmetic approximants correspond, and then so do the limits.

Finally, observe that the fixpoint structure of ϕ is preserved in the translation to Φ. $\square$

To illustrate this construction, consider the mu-arithmetic definition of the even numbers that was given above; we can rewrite this as

$$\mu(x, X).\, P(x) \vee (Q(x) \wedge f(x) \in X)$$

where $P(x)$ iff $x = 0$, $Q(x)$ iff $x > 1$, and $f(x) = x - 2$. The individual terms of this formula are x and $f(x)$, so the states of the constructed transition system are pairs (m, n) of integers. We equip the model with two atomic propositions $\ulcorner P(x)\urcorner$ and $\ulcorner Q(x)\urcorner$ such that $\|\ulcorner P(x)\urcorner\| = \{(0, n)\}$ and $\|\ulcorner Q(x)\urcorner\| = \{\,(m, n) \mid m > 1\,\}$. We also have a label $\ulcorner f(x)\urcorner$ such that $(m, n) \xrightarrow{\ulcorner f(x)\urcorner} (m, m-2)$. Finally, we have a label $\ulcorner x(f(x))\urcorner$ such that $(m, n) \xrightarrow{\ulcorner x(f(x))\urcorner} (n, n)$. Part of this machine is illustrated in the figure. The translation into modal logic is then

$$\mu X.\, \ulcorner P(x)\urcorner \vee (\ulcorner Q(x)\urcorner \wedge \langle\ulcorner f(x)\urcorner\rangle\langle\ulcorner x(f(x))\urcorner\rangle X)$$

and a state (m, n) satisfies this formula just in case m is even.

This theorem, together with the previous theorem, gives us non-collapse:

Theorem 6 *The simple hierarchy in the modal mu-calculus does not collapse.*

Proof. Use the theorem to code an arithmetic strict $\Sigma_n^{S\mu}$ set of integers by a strict $\Sigma_n^{S\mu}$ modal mu-formula Φ on a r.p.t.s. $\mathscr{T}$; by the previous theorem, no $\Sigma_{n-1}^{S\mu}$ modal formula can have the same denotation in $\mathscr{T}$, and so no $\Sigma_{n-1}^{S\mu}$ modal formula is logically equivalent to Φ. $\square$

Some remarks are in order at this point. Readers who think of the modal mu-calculus as being really about finite systems may be feeling slightly uneasy about the transition

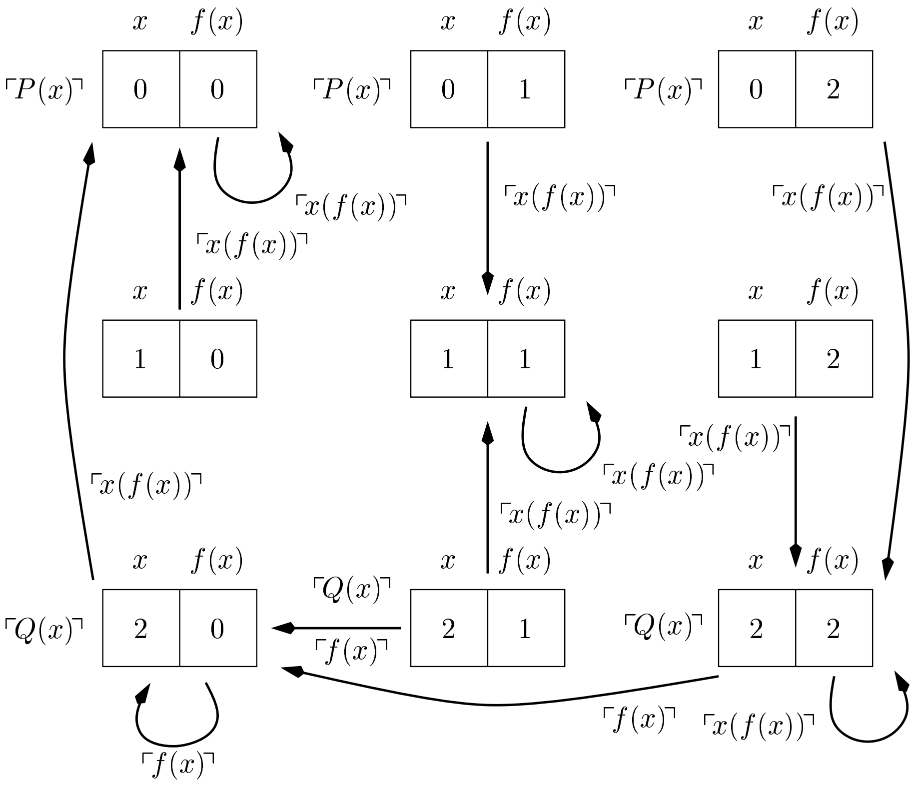

Fig. 1. Part of a transition system representing a μ-arithmetic formula

systems built above—is it not cheating to use infinite (and worse, infinite-branching) systems with arbitrary recursive transition relations? Further, should we not be dealing only with pure sentences, without arbitrary recursive predicates in the logic?

To deal with the second point first, the predicates can be easily replaced by transitions coding their characteristic functions.

As for the first point, we have also the theorem

Theorem 7 *The simple hierarchy is strict on the class of finite models.*

Proof. Let Φ be a strict $\Sigma_n^{S\mu}$ formula; then for every $\Sigma_{n-1}^{S\mu}$ formula Ψ, we have that $\neg(\Phi \Leftrightarrow \Psi)$ is satisfiable, by the main theorem. But the modal mu-calculus has the finite model property [Koz88,StE89], so $\neg(\Phi \Leftrightarrow \Psi)$ is satisfiable within the class of finite models. Hence Φ is not equivalent to any $\Sigma_{n-1}^{S\mu}$ formula even on finite models. $\square$

Another slight strengthening concerns the number of labels. The coding above assumes the availability of an unlimited number of labels, to code symbols of arithmetic formulae; however, this can easily be reduced to two, just by using a binary code for the labels, and using intermediate states. For example, one might code the $\exists$ variable x by the sequence of labels aab, in which case the transition system would have dummy states after the a transitions, and would branch on the b transition.

In application work, the modal mu-calculus is often extended to allow modalities to be indexed by sets of labels, which allows the convenient expression of properties such as 'a happens infinitely often' (see, for example, [Bra91]). The hierarchy theorem applies also to this extension, provided that the index sets are restricted to be recursive: in practice, they are usually finite or cofinite.

A more significant extension would be to determine whether the hierarchy is strict on trees of bounded degree, the models originally studied by Niwiński. It is possible that on trees of a fixed degree n, the hierarchy does collapse. However, this question remains open.

4 The non-collapse of the alternation hierarchy

In order to extend this result to the Niwiński alternation hierarchy, we need to do a little more work along the lines of the normal form theorems. It would be possible to do this stage of the work in the modal mu-calculus, using the theory of simultaneous fixpoints [Bek84] and extending the foregoing, but we choose instead to work in arithmetic, as the necessary coding is easier there. So we shall prove that $\Sigma_i^{N\mu}$ formulae are equivalent (in mu-arithmetic) to $\Sigma_i^{S\mu}$ formulae.

It is possible to define a notion of the Niwiński hierarchy for arithmetic that correctly captures the idea of genuine semantic alternation, but because of the existence of individual variables in arithmetic, the definition is a little complex, and additional minor complexities arise in the subsequent proof. As will become apparent, in the presence of pairing the Niwiński and simple classes are equal, up to logical equivalence, so there is no intrinsic interest in defining it generally. Hence, to avoid unnecessary work, we shall restrict ourselves to formulae of a well-behaved form, sufficient for our purposes. We say that a formula ϕ of mu-arithmetic is *nice* if no set term $\mu X.\, \psi$ or $\nu X.\, \psi$ occurring in ϕ contains any free individual variable. Observe that the translation of the modal mu-calculus into mu-arithmetic, given in the proof of Theorem 4, produces only nice formulae. We then define the Niwiński hierarchy $\Sigma_n^{N\mu}$ for nice formulae of mu-arithmetic exactly as for the modal mu-calculus, where the variable capture constraint refers to set variables. To be precise, the $\Sigma_0^{N\mu} = \Pi_0^{N\mu}$ formulae are the first-order formulae and the $\Sigma_0^{N\mu} = \Pi_0^{N\mu}$ set terms are the set variables; $\Sigma_{n+1}^{N\mu}$ is formed by taking $\Sigma_n^{N\mu} \cup \Pi_n^{N\mu}$ and closing under (i) first-order connectives and $\in$; (ii) $\mu X.\, \psi$ for $\psi \in \Sigma_{n+1}^{N\mu}$, provided that ψ has at most x as a free individual variable; (iii) substitution of a $\Sigma_{n+1}^{N\mu}$ set term Ξ for a free set variable of $\psi \in \Sigma_{n+1}^{N\mu}$ provided that no free set variable of Ξ is captured by ψ. Hence the translation takes modal $\Sigma_n^{N\mu}$ to arithmetic $\Sigma_n^{N\mu}$.

Theorem 8 *If ϕ is a (nice, by definition) $\Sigma_n^{N\mu}$ formula of mu-arithmetic, it is equivalent to some $\Sigma_n^{S\mu}$ formula, which moreover is nice and has μ as its top-level connective (i.e. is of the form $\sigma \in \mu X.\, \psi$).*

Proof. The proof is firstly by induction on n, and secondly by induction on the construction of formulae according to the rules (i)–(iii) of the alternation hierarchy. The case $n = 0$ is trivial. So assume n, and prove $n + 1$.

The base of the inner induction is the $\Sigma_n^{N\mu}$ and $\Pi_n^{N\mu}$ formulae, which by the outer induction are $\Sigma_n^{S\mu}$ and $\Pi_n^{S\mu}$ respectively. Any first-order combination ϕ of such formulae can be made $\Sigma_{n+1}^{S\mu}$ by wrapping a dummy fixpoint round it—if ϕ has free variable x, transform it to $x \in \mu X.\, \phi$ (and then alpha-convert if desired), and if there is more than one free variable, use pairing to code them into one. Similarly, given $\Sigma_{n+1}^{S\mu}$ formulae in the required form, a first-order combination of them can be wrapped in a dummy fixpoint to have the required form.

For case (ii), if ϕ is $\Sigma^{N\mu}_{n+1}$, inductively it is equivalent to some $\Sigma^{S\mu}_{n+1}$ formula, and then wrapping a μ round it keeps it in $\Sigma^{S\mu}_{n+1}$ of the desired form.

The non-trivial case is the substitution rule (iii). Let ψ, Ξ be the $\Sigma^{N\mu}_{n+1}$ formula and term such that ϕ is the result of substituting Ξ for the free variable Z of ψ. Inductively, ψ and Ξ are equivalent to formulae $\sigma \in \mu X.\, \psi_1$ and $\mu Y.\, \psi_2$ of the required form (and by niceness, the individual variables in σ are exactly the free individual variables of ψ). All we need do now is to combine the inductive generation of X and Y along the lines of the pair normal form theorem. By the side condition of (iii), Ξ contains no reference to any bound set variable of ψ, and by niceness no reference to any individual variable of ψ. We can therefore pull it out to the same level as X. In detail, we use a new variable W, with the intent that $\langle 0, n\rangle \in W \Leftrightarrow n \in X$ and $\langle 1, n\rangle \in W \Leftrightarrow n \in Y$. So we use the formula

$$\langle 0, \sigma\rangle \in \mu W.\, ((w)_0 = 0 \wedge \psi_1') \vee ((w)_0 = 1 \wedge \psi_2')$$

where ψ_1' is formed from ψ_1 by replacing '$\rho \in X$' by '$\langle 0, \rho\rangle \in W$', replacing '$\rho \in Z$' by '$\langle 1, \rho\rangle \in W$', and x by $(w)_1$; and similarly for ψ_2'.

Provided that $\mu X.\, \psi_1$ and $\mu Y.\, \psi_2$ are $\Sigma^{S\mu}_{n+1}$, which they are by induction, this formula is also $\Sigma^{S\mu}_{n+1}$, so we are done. □

Corollary 9 *The denotation of a modal $\Sigma^{N\mu}_n$ formulae is arithmetic $\Sigma^{S\mu}_n$ definable.* □

By combining this result with Theorem 4 and Theorem 5, and the fact that every $\Sigma^{S\mu}_n$ formula is also $\Sigma^{N\mu}_n$, we immediately obtain our desired

Theorem 10 *The modal mu-calculus alternation hierarchy is strict.* □

For those who know only the notion of alternation depth defined in [EmL86], it may be worth explaining the differences between that notion and the one we are using (noted briefly in [Kai95]). In [EmL86], alternation depth is directly defined by an inductive definition on formulae; however, it can easily be cast into our framework. In fact, the definition of [EmL86] contains a minor error, as noted in [And93]; we assume the corrected version.

Recall the definition of the classes $\Sigma^{N\mu}_n$; we can define Emerson–Lei versions $\Sigma^{EL\mu}_n$ of these classes by modifying clause (iii) as follows: (iii′) substitution of $\Phi' \in \Sigma^{EL\mu}_{n+1}$ for a free variable of $\Phi \in \Sigma^{EL\mu}_{n+1}$ provided that Φ' is a closed formula. A simple induction now shows the following

Lemma 11 *A formula Φ has Emerson–Lei alternation depth $\leq n$ iff $\Phi \in \Sigma^{EL\mu}_{n+1} \cap \Pi^{EL\mu}_{n+1}$, for $n \geq 0$.* □

Since it is immediate from the definitions that $\Sigma^{S\mu}_n \subseteq \Sigma^{EL\mu}_n \subseteq \Sigma^{N\mu}_n$, we obtain as a corollary of our results that

Corollary 12 *The alternation depth hierarchy of [EmL86] is strict.* □

In his thesis [And93], Andersen presents a somewhat complex improvement to the direct definition of alternation depth in a way that provides tighter complexity bounds on the algorithm of [EmL86], and better reflects intuition. In fact, Andersen is bringing the direct definition closer to the Niwiński notion: his definition satisfies the "$\Rightarrow$" direction of the above lemma replacing $\Sigma^{EL\mu}$ by $\Sigma^{N\mu}$. The examples Andersen provides (p. 28) of the difference between the original definition and his improvement also serve as examples of the differences between $\Sigma^{EL\mu}$ and $\Sigma^{N\mu}$; for example,

$$\mu X.\, \nu Z.\, \mu U.\, \nu Y.\, Y \wedge X$$

is in $\Sigma_2^{N\mu}$, but only in $\Sigma_4^{EL\mu}$.

As mentioned earlier, there is a slight strengthening of the Niwiński definition, to address the following minor irritation. Although at higher levels, $\Sigma^{N\mu}$ ignores vacuous fixpoints, it doesn't do so at the bottom of the hierarchy: the formula $\mu X.\, \nu Y.\, [a]X$ is $\Sigma_2^{N\mu}$ but not $\Sigma_1^{N\mu}$, even though the inner fixpoint is vacuous. One can fix this by simply defining $\Sigma_0^{N\mu}$ to be closed under vacuous fixpoint formation, i.e. forming $\mu Z.\, \Phi$ and $\nu Z.\, \Phi$ where Z is not free in Φ (see [Mad96] for details). All our results (and indeed all the complexity results on alternation) carry through with this modification; whether it is worth doing, is a matter of taste.

5 Conclusion

The results of this paper solve the alternation hierarchy problem by a relatively simple reduction to a known hierarchy problem. This is a fairly powerful technique: it will apply to any mu-calculus whose models are powerful enough to allow the coding of arithmetic. There is, however, a drawback: we should like to have simple explicit examples of strict $\Sigma_n^{N\mu}$ modal mu-calculus formulae. Although the arithmetic examples are constructible in principle, the complexity of the proofs in [Lub93] means that the examples are not practically presentable, and so neither are the translations into the modal mu-calculus.

Giacomo Lenzi has recently produced an independent proof of the non-collapse of the Emerson–Lei alternation hierarchy in a closely related mu-calculus. The technique is a very delicate topological analysis of the finite models of formulae of the logic, and it does produce explicit and simple examples of strict formulae—exactly the examples one would expect, in fact. The reader is referred to Lenzi's paper [Len96].

Finally, I should like to thank Colin Stirling and other members of the Concurrency Club at the LFCS for valuable discussions and pointers; Robert Lubarsky for some additional remarks on his work; anonymous referees for their useful suggestions, and Perdita Stevens, for suggesting further improvements in the exposition.

A preliminary version of this paper appeared in CONCUR '96 [Bra96a]; a final version will appear in *Theoretical Computer Science*.

References

[And93] H. R. Andersen, *Verification of Temporal Properties of Concurrent Systems*, DAIMI PB – 445, Computer Science Dept, Aarhus University, 1993.

[ArN90] A. Arnold and D. Niwinski, Fixed point characterization of Büchi automata on infinite trees, *J. Inf. Process. Cybern.* **EIK 26** (1990) 451–459.

[Bar75] J. Barwise, *Admissible Sets and Structures* (Springer, Berlin, 1975).

[Bek84] H. Bekič, Definable operations in general algebras, and the theory of automata and flow charts, in: H. Bekič (ed. C. B. Jones) *Programming Languages and Their Definition*, LNCS **177** (Springer, Berlin, 1984).

[Bra91] J. C. Bradfield, *Verifying Temporal Properties of Systems* (Birkhäuser, Boston, 1991).

[Bra96] J. C. Bradfield, On the expressivity of the modal mu-calculus, in: C. Puech and Rüdiger Reischuk, eds., *Proc. STACS '96*, LNCS **1046** (Springer, Berlin, 1996) 479–490.

[Bra96a] J. C. Bradfield, The modal mu-calculus alternation hierarchy is strict, in: U. Montanari and V. Sassone, eds., *Proc. CONCUR '96*, LNCS **1119** (Springer, Berlin, 1996) 233–246.

[EJS93] E. A. Emerson, C. Jutla and A. P. Sistla, On model-checking for fragments of μ-calculus, in: C. Courcoubetis, ed., *Proc. CAV '93*, LNCS **697** (Springer, Berlin, 1993) 385–396.

[EmL86] E. A. Emerson and C.-L. Lei, Efficient model checking in fragments of the propositional mu-calculus, in: *Proc. 1st LICS* (IEEE, Los Alamitos, CA, 1986) 267–278.

[Kai95] R. Kaivola, On modal mu-calculus and Büchi tree automata, *Inf. Proc. Letters* **54** (1995) 17–22.

[Koz83] D. Kozen, Results on the propositional mu-calculus, *Theoret. Comput. Sci.* **27** (1983) 333–354.

[Koz88] D. Kozen, A finite model theorem for the propositional μ-calculus, *Studia Logica* **47** (1988) 233–241.

[Len96] G. Lenzi, A hierarchy theorem for the mu-calculus, in: F. Meyer auf der Heide and B. Monien, eds., *Proc. ICALP '96*, LNCS **1099** (Springer, Berlin, 1996) 87–109.

[LB+94] D. Long, A. Browne, E. Clarke, S. Jha and W. Marrero, An improved algorithm for the evaluation of fixpoint expressions, in: D. L. Dill, ed., *Proc. CAV '94*, LNCS **818** (Springer, Berlin, 1994) 338–350.

[Lub93] R. S. Lubarsky, μ-definable sets of integers, *J. Symbolic Logic* **58** (1993) 291–313.

[Lub97] R. S. Lubarsky, personal communication (1997).

[Mad96] A. Mader, *Verification of Modal Properties Using Boolean Equation Systems*, doctoral dissertation, Institut für Informatik, Technische Universität München (1996).

[Mos74] Y. N. Moschovakis, *Elementary Induction on Abstract Structures* (North-Holland, Amsterdam, 1974).

[Niw86] D. Niwiński, On fixed point clones, in: L. Kott, ed., *Proc. 13th ICALP*, LNCS **226** (Springer, Berlin, 1986) 464–473.

[Rab70] M. O. Rabin, Weakly definable relations and special automata, in: Y. Bar-Hillel (ed.), *Mathematical Logic and Foundations of Set Theory* (North-Holland, Amsterdam, 1970) 1–23.

[RiA74] W. Richter and P. Aczel, Inductive definitions and reflecting properties of admissible ordinals, in: J. E. Fenstad and P. G. Hinman, eds., *Generalized Recursion Theory* (North-Holland, Amsterdam, 1974) 301–381.

[StE89] R. S. Streett and E. A. Emerson, An automata theoretic decision procedure for the propositional mu-calculus. *Information and Computation* **81** (1989) 249–264.

[Sti91] C. P. Stirling, Modal and temporal logics, in: S. Abramsky, D. Gabbay and T. Maibaum, eds., *Handbook of Logic in Computer Science*, Vol. 2 (Oxford University Press, 1991) 477–563.

[Sti95] C. Stirling, Local model checking games, in: I. Lee and S. A. Smolka, eds., *Proc. Concur '95*, LNCS **962** (Springer, Berlin, 1995) 1–11.

A Appendix: summary of the mu-arithmetic hierarchy theorem

As is clear from the foregoing, there is no need to understand Lubarsky's hierarchy result in order to apply it. However, since the modal mu-calculus is essentially about the power of mixed induction and co-induction, and [Lub93] answers that question, it is well worth while to gain (at least a superficial) understanding of it.

An obstacle to this, and one that perhaps explains the surprising fact that it has taken six years for the observations of this paper to be noticed, is that [Lub93] is a highly technical paper, written for an audience well versed in admissible set theory and definability theory, whereas those interested in the modal mu-calculus tend to be concurrency theorists or automata theorists. We therefore now give a quick summary of the material.

A.1 Preliminaries

The mathematics underlying the result is the theory of admissible sets and the theory of inductive definability, the main development of which (up to the point required here) occurred around 1960–1975. The standard reference on admissible sets is [Bar75], and on inductive definitions is [Mos74].

In the (first-order) language of set theory (where $\in$ is the membership relation), we say that a formula is Δ_0 ($= \Sigma_0 = \Pi_0$) if it contains no unbounded quantifiers—for example, $\forall x \in a. \exists y \in b. x = y$ is Δ_0, but $\exists a. \forall x, y \in a. x = y$ is not. A Σ *formula* may also have unbounded existential quantifiers; if there is exactly one such, and it is at the front, the formula is Σ_1. As usual, Π is the dual of Σ.

Kripke–Platek set theory (KP) is a theory in which there is no Powerset axiom, and Separation and Replacement are restricted. In fact, they are restricted to Δ_0 and Σ_1 formulas respectively—one can then derive them for Δ and Σ respectively, where a formula is Δ if it is Σ and also equivalent to a Π formula. The intuition for this is that instead of being able to form arbitrary subsets and use arbitrary functions, one can only use 'recursive' subsets and 'recursive' functions, where 'recursive' is relative to the universe, not just to the integers.

One of the most important theorems of KP (which in fact Kripke took as basic) is the 'Σ Reflection Principle', that for any Σ formula ϕ, we have $\mathrm{KP} \vdash \phi \Leftrightarrow \exists a. \phi^{(a)}$, where $^{(a)}$ is the operation of replacing unbounded quantifers by quantifiers bounded in a. Thus, if a Σ formula is true in some model of KP, that model contains a set (a witness) within which the formula is true. Therefore, Σ formulae are *persistent*; that is, if true in a model, they are true in larger models (in an appropriate sense). Stronger reflection properties are the key to [Lub93]'s results.

It is useful to introduce *ur-elements* (primitive elements with no membership structure); in our case the integers $\mathbb{N}$ will be ur-elements (and so they are *not* the same as the finite ordinals!). The appropriate extension of KP is called KPU.

A set A (in the normal set-theoretic universe plus ur-elements) is *admissible* (over a collection M) if it is a model of KPU—that is, it is a big enough fragment of the universe to 'do computable things' in. If moreover $M \in A$ then we say A is admissible *above* M—and then quantification over ur-elements becomes bounded quantification. An ordinal α is admissible (over/above M) if it is the ordinal of some admissible set (including/containing M).

Gödel's constructible universe L is given (in one version) by iterating over the ordinals the operation of defining subsets by one of a certain number of primitive operations which are sufficient to give all of KPU—for each stage L_α (or $\mathrm{L}_\alpha(M)$ if we start with M rather than $\emptyset$), $\mathrm{L}_{\alpha+1}$ adds the sets definable by one use of a primitive operation. (Accordingly, $\mathrm{L}_{\alpha+\omega}$ adds the sets that are first-order definable over L_α.) If α is admissible above M, then $\mathrm{L}_\alpha(M)$ is an admissible set above M, and is in fact the smallest with ordinal α.

One point that may cause confusion should be mentioned. The first admissible ordinal is ω (note that KPU does not have the Infinity axiom; the associated constructible level L_ω is the set of hereditarily finite sets). However, if we have a countably infinite collection M of ur-elements, such as $\mathbb{N}$, the first admissible *above* M is the second pure admissible, which turns out to be (not at all by coincidence!) ω_1^{CK}, the first non-recursive ordinal.

The rest of this subsection gives notation and definitions from [Lub93] §1.

We have mentioned Σ reflection; the standard general definition is that if α is an ordinal, Γ a class of formulas, and X a class of ordinals, α is Γ-reflecting on X if for all $\phi \in \Gamma$ (with parameters from L_α), if $\mathrm{L}_\alpha \vDash \phi$ then $\mathrm{L}_\beta \vDash \phi$ for some $\beta \in \alpha \cap X$.

Now we define *gap reflection*. For $\alpha \in X$, let α^{+X} be the next member of X beyond α. We say α is Γ gap-reflecting on X if for all $\phi \in \Gamma$ (with parameters from L_α and a constant symbol L_α), if $\mathrm{L}_{\alpha^{+X}} \vDash \phi[L_\alpha := \mathrm{L}_\alpha]$ then $\mathrm{L}_{\beta^{+X}} \vDash \phi[L_\alpha := \mathrm{L}_\beta]$ for some $\beta < \alpha$. This is rather hard to give intuition for; the best the present author can suggest is that in the same sense as a Γ-reflecting ordinal is big enough that it 'doesn't tell you anything you didn't already know' about Γ at your current level of the universe, so a Γ gap-reflecting ordinal is so big that it doesn't tell you anything you couldn't already find out about Γ by looking X-far further up in the universe—not very helpful, but perhaps better than nothing.

Now α is said to be a 1-reflecting admissible if it is admissible, and $(n+1)$-reflecting if it is Π_1 gap-reflecting on the collection of n-reflecting admissibles. We write α^{+n} for the next n-reflecting admissible beyond α, and α^+ for α^{+1}.

§2 of [Lub93] establishes various results about n-reflecting admissibles, which we use as required, noting them thus (§2).

A.2 Modifications to [Lub93]

We mentioned before Lemma 1 that we have made two slight modifications to the definitions in Lubarsky's paper, and we should now explain these.

The first modification is to the definition of 'pair-normal'. We define it so that the first-order matrix is an arbitrary first-order formula. Lubarsky imposes the additional restriction that the first-order matrix is a boolean combination of Σ_1^0 and Π_1^0 (in the usual sense of Kleene) formulae. The proof of the pair-normal form theorem now contains an additional clause to reduce any first-order formula to pair-normal form. This is done by the introduction of vacuous fixpoints: for example, $\forall x.\, \exists y.\, \forall z.\, \phi$ is equivalent to $\forall x'.\, x' \in \mu X.\, \exists y'.\, y' \in \mu Y.\, \forall z.\, \phi$, and then the construction we gave in the proof can be applied, and results in the production of a pair-normal formula in Lubarsky's sense. Since the sign of the vacuous fixpoints is irrelevant, it can be chosen so as not to affect the alternation depth of any formula with an existing fixpoint. However, it is not necessary to do this; none of the subsequent theorems make any use of the restriction on the form of the first-order matrix. Lubarsky chose [Lub97] to impose this additional restriction for aesthetic reasons, to simplify as much as possible the matrix, and because it provides the natural generalization of Kleene's classical result that any first-order inductive definition on $\mathbb{N}$ is equivalent to a Π^0_1 inductive definition. For our purposes, it is cleaner to allow an arbitrary first-order matrix.

The second modification concerns the alternation classes. Lubarsky uses the plain notation Σ_n for his classes; these are defined as for our $\Sigma_n^{S\mu}$, except that (i) instead of closure under conjunction and disjunction, we have: if $\phi = \psi_1 \wedge \psi_2$ and each of ψ_i is Σ_n or Π_n, then ϕ is Σ_{n+1} and Π_{n+1}; *but also* (ii) if ϕ is first-order, ϕ is Σ_0 and Π_0. This has the consequence that according to the definition $(\tau_1 \in \mu X_1.\, \phi_1) \wedge (\tau_2 \in \mu X_2.\, \phi_2)$, for first-order ϕ_i, is only Σ_2, whereas its pair-normal form is Σ_1; in our formulation, the formula is $\Sigma_1^{S\mu}$ to start with. There is no reason to prefer the definition of [Lub93], and it may have that form simply owing to an oversight [Lub97]. Indeed, the details of §4 of [Lub93] itself, if they were written out in full, would be rather more complicated with Lubarsky's definition than with ours.

A.3 The upper bound theorem

The easier half of the main result of [Lub93] is the upper bound theorem:

Theorem 13 ([Lub93] §4) *If $\phi(z)$ is a $\Sigma_n^{S\mu}$ formula of mu-arithmetic, the set defined by ϕ is Σ_1 over the least n-reflecting admissible (above $\mathbb{N}$).*

Proof. We assume that ϕ is in pair-normal form (in our sense).

The proof is by induction on n. We first strengthen the theorem to take account of set-valued parameters: given a set $\mathbf{W}$ of integers, generalize the definition of n-reflecting admissible to n–$\mathbf{W}$-reflecting admissible, by restricting to the admissibles above $\mathbf{W}$. Then if $\phi(z)$ contains a free set variable W, the set defined by ϕ is Σ_1 over the least n–$\mathbf{W}$-reflecting admissible, when W is interpreted as $\mathbf{W}$. (The case of more than one free set variable merely requires a little more coding; and similarly for free individual variables.) The notation α^{+n} is extended to mean 'the next n–$\mathbf{W}$-admissible after α', where $\mathbf{W}$ will be understood from context (and similarly, α may depend on $\mathbf{W}$).

The case $n = 1$ is actually obvious from established results such as Gandy's theorem. However, it can be proved directly.

Since we are assuming pair-normal form, work with $\mu X.\, \phi$ for ϕ first-order, with parameter $\mathbf{W}$. Now consider the approximants $\mathbf{X}^\beta = \mu^\beta X.\, \phi(\mathbf{W})$. $\mathbf{X}^{<0} = \varnothing$ is definable over $\mathrm{L}_\omega(\mathbf{W})$, and similarly each $\mathbf{X}^\beta = \{\, x \mid \phi(x, \mathbf{X}^{<\beta}) \,\}$ is definable over $\mathrm{L}_{\omega\beta}(\mathbf{W})$. Now if α is the least $\mathbf{W}$-admissible, we claim that (i) $\mathbf{X}^{<\alpha}$ is Σ_1 over $\mathrm{L}_\alpha(\mathbf{W})$ and (ii) $\mathbf{X}^\alpha = \mathbf{X}^{<\alpha}$ (so $= \mathbf{X}^\infty$).

(i) is true because $z \in \mathbf{X}^{<\alpha}$ iff

$$\begin{aligned}\mathrm{L}_\alpha(\mathbf{W}) \vDash \exists f.\, \exists \beta.\ &(\operatorname{dom} f = \beta + 1)\\ &\wedge (f(\gamma) = \{\, x \mid \phi(x, (\bigcup_{\delta<\gamma} f(\delta))\,\}) \wedge z \in f(\beta)\end{aligned}$$

—that is, f enumerates the approximants of $\mathbf{X}$ up as far as β, and z is in $f(\beta)$. This is a Σ formula, so equivalent to a Σ_1 formula.

(ii) is true because $z \in \mathbf{X}^\alpha$ iff $\phi(z, \mathbf{X}^{<\alpha})$. Working in $\mathrm{L}_\alpha(\mathbf{W})$, replace $\rho \in \mathbf{X}^{<\alpha}$ by $\exists\beta.\, \rho \in \mathbf{X}^{<\beta}$; bring the quantifiers to the front (which we can do in an admissible set), and take sups, so getting $\exists\beta.\, \phi(z, \mathbf{X}^{<\beta})$; then $z \in \mathbf{X}^\beta$, so $z \in \mathbf{X}^{<\alpha}$.

Now consider the case $n+1$. Let α be the least $(n+1)$–$\mathbf{W}$-reflecting admissible. Our formula is now of the form $\mu X.\, \sigma \in \nu Y.\, \phi$. By induction, given parameters $\mathbf{W}$ and $\mathbf{X}$, $\nu Y.\, \phi$ is Π_1 over the least n–$\langle \mathbf{W}, \mathbf{X} \rangle$-reflecting admissible. Thus, $\mathbf{Y}_0 = \nu Y.\, \phi(\mathbf{X}^{<0})$ is Π_1 over the least n–$\mathbf{W}$-admissible, and so $\mathbf{X}^1$ is definable over it. Then $\mathbf{Y}_1 = \nu Y.\, \phi(\mathbf{X}^{<1})$ is Π_1 over the least n–$\langle \mathbf{W}, \mathbf{X}^{<1} \rangle$-reflecting admissible, which is at worst the second n–$\mathbf{W}$-admissible. This continues; and since (§2) there are α-many n–$\mathbf{W}$-admissibles below α, we get as above that $\mathbf{X}^\alpha$ is Σ_1 over $\mathrm{L}_\alpha(\mathbf{W})$.

The part where gap reflection comes in is proving that $\mathbf{X}^\alpha = \mathbf{X}^{<\alpha}$. To evaluate $\mathbf{X}^\alpha$, we need to consider $\mathbf{Y}_\alpha = \nu Y.\, \phi(\mathbf{X}^{<\alpha})$. If we consider the approximants $(\mathbf{Y}_\alpha)^\beta$, by the same argument as before, they close by at worst the next n–$\mathbf{W}$-reflecting admissible after α, that is, by α^{+n}, so that $\mathbf{Y}_\alpha = \bigcap_{\beta<\alpha^{+n}} (\mathbf{Y}_\alpha)^\beta$.

Therefore $z \in \mathbf{Y}_\alpha$ iff $\mathrm{L}_{\alpha^{+n}} \models \forall f.\, (\mathrm{dom}\, f = \gamma \wedge$ 'f enumerates $(\mathbf{Y}_\alpha)^\beta$'$) \Rightarrow \forall \beta < \gamma.\, z \in f(\beta)$. This formula is, once the implications have been untangled, Π_1 in $\mathrm{L}_{\alpha^{+n}}$, and it carries a parameter L_α, since to define $\mathbf{X}^{<\alpha}$, which is referred to in the definition of $\mathbf{Y}_\alpha$, we need to quantify over ordinals in L_α. Hence, by the definition of α, if it's true, it's true in some $\mathrm{L}_{\beta^{+n}}$ for $\beta < \alpha$, and so $z \in \mathbf{Y}_\beta$, so $z \in \mathbf{X}^{<\beta+1}$, so $z \in \mathbf{X}^{<\alpha}$. $\square$

The reader is warned that a number of details have been quietly ignored in this condensation, but the essential ideas are still there.

A.4 Attaining the upper bound

The second half of the theorem, that if a set is Σ_1 over ω^{+n}, the least n-reflecting admissible, then it is $\Sigma_n^{S\mu}$-definable, is much more complex, occupying some ten pages of coding even at a fairly abstract level of description. It is not really possible to abbreviate this while preserving detail, so we just give a rough outline of the idea. (We'll write ω_n for the n'th admissible, so $\omega_1 = \omega_1^{\mathrm{CK}}$.)

The idea is to use the fixpoints to encode into $\mathbb{N}$ methods of giving notations for all the ordinals up to α, and to encode the true Σ_1 statements about L_α.

In the base case $n = 1$, this is established material (see [Bar75]). Using induction over a first-order definition, we can construct the (codes of the) true Σ statements about L_{ω_1}, and this is expressible with one least fixpoint μX, say.

To go further, we need also to generate the true Π_1 sentences about L_{ω_1} as a greatest fixpoint νY. This is done by a similar technique: each step in the (co-)induction throws out some ill-formed or false sentences, and by the time we close at ω_1, we have only the true Π_1 sentences left.

Now for $n = 2$, we embed μX and νY inside a minimal fixpoint μW, say. Now the first approximant of W will code both the Σ_1 and Π_1 statements about L_{ω_1}. The second time round, we have access to information about L_{ω_1} itself, so we end up with the true Σ_1 and Π_1 statements about L_{ω_2}, and so on. Then, with some extensions to the control structure (encoded in the bodies of the fixpoints), we find that we eventually close at the Σ_1 statements about $L_{\omega^{+2}}$, the first 2-reflecting admissible.

For $n > 2$, the procedure is generalized appropriately.

A.5 Final remarks

This section has, we hope, given a flavour of the result on which this paper's result relies. One question that has been entirely unanswered so far is, what do these strict $\Sigma_n^{S\mu}$ formulae look like, and so, what do the examples showing the strictness of the modal mu-calculus look like? At first sight, looking at the ten pages so hastily summarized above, the answer is "an unbelievably complicated mess". At second sight, one tends to think that it's not so bad: although there is a colossal amount of coding going on, most of that should be codable in the transition system of Theorem 5, leaving, ultimately, quite simple formulae. In [Niw86], the alternation hierarchy was shown to be strict for a mu-calculus without conjunction (and without the quantifying power of the modal box), and the examples were of a very simple form; and Lenzi's proof [Len96] of an alternation hierarchy for a closely related mu-calculus also gives simple examples. It would be interesting to see whether our strict $\Sigma_n^{N\mu}$ formulae can be reduced to this form; but this is a daunting exercise.

Finally, what is an n-reflecting admissible anyway? The answer is that ω^{+2}, the first 2-reflecting admissible, is huge; so huge that nobody but a specialist is likely to have any comprehension of its magnitude (although, of course, it is countable, like everything we deal with). To give an idea of just how big it is, recall that α is recursively inaccessible if it is an admissible limit of admissibles; this is equivalent to saying that $\omega_\alpha = \alpha$. Writing ρ_β for the βth recursive inaccessible, we then say α is 1-hyperinaccessible if $\rho_\alpha = \alpha$; then we can define β-hyperinaccessible for all β. Then we continue with the least α such that α is α-hyperinaccessible, and so on, and so on. (Along the way, we pass such exotic beasts as the recursively Mahlo ordinals.) All of these are smaller than ω^{+2}; in fact, according to Aczel and Richter in [RiA74], ω^{+2} "appears to be greater than any 'reasonable' iteration into the transfinite of this diagonalization process" (a remark which is actually made about something very much smaller than ω^{+2}). The first well-known ordinal that is bigger than ω^{+2} is actually bigger than all the ω^{+n}: it is the first stable ordinal σ_0. (σ is stable if the truth of Σ formulae in L_σ is the same as their truth in L.) It turns out that σ_0 is the least ordinal not the order-type of a Δ_2^1 well-ordering of $\mathbb{N}$; this implies that the μ-definable sets of integers are actually a strict subset of the Δ_2^1 sets of integers, which in turn implies that the Δ_2^1 upper bound on the complexity of modal mu-calculus formulae, produced in [Bra96] by elementary means, is actually strict.