The modal mu-calculus alternation hierarchy is strict

J. C. Bradfield
Laboratory for Foundations of Computer Science
University of Edinburgh, The King's Buildings
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 fix-point properties can be expressed with only a few alternations of least and greatest fix-points. In this paper, we resolve this question by showing that the hierarchy does not collapse.

Keywords: temporal logic, mu-calculi, hierarchies, alternation

1 Introduction.

The modal mu-calculus, or Hennessy–Milner logic with fix-points, 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 fix-point operators. This makes it simple, as the fix-point is the only non first-order operator, 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 fix-points. 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 and Jutla

[EmJ88], and more directly seen by Stirling's game-theoretic approach [Sti95]), and since the logic is closed under negation, the problem is in NP ∩ co-NP, which suggests quite strongly that it is in fact in P (even if P ≠ 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 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. 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 fix-points, and thereby show the strictness of the modal mu-calculus hierarchy.

The remainder of this paper is thus: in section 2 we introduce the modal mu-calculus, and arithmetic with fix-points, with 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. The technical report version of this paper [Bra96a] contains an appendix giving a summary of the mu-arithmetic hierarchy result that we transfer; unfortunately, page limits forbid its inclusion here.

2 Preliminaries.

2.1 Modal mu-calculus.

The modal mu-calculus, in positive form, has formulae Φ built up from variables Z by the booleans tt, ff, $\vee$ and $\wedge$, the modal connectives $[K]\Phi$ and $\langle K\rangle\Phi$ (for a set $K \subseteq \mathscr{L}$, where $\mathscr{L}$ is a countable set of *labels*), and the minimal and maximal fix-point connectives $\mu Z.\Phi(Z)$ and $\nu Z.\Phi(Z)$. We also write $[-K]\Phi$ for $[\mathscr{L}-K]\Phi$, and $[-]\Phi$ for $[\mathscr{L}]\Phi$. We further require that the label sets K are recursive (in practice, they are always small finite sets, often singletons, or the complements thereof).

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

$$\|[K]\Phi\|^{\mathscr{T}}_{\mathscr{V}} = \{\, s \mid \forall s' \,.\, \forall l \in K \,.\, s \xrightarrow{l} s' \;\Rightarrow\; s' \in \|\Phi\|^{\mathscr{T}}_{\mathscr{V}} \,\}$$

and dually for $\langle K\rangle\Phi$, and for the fix-points by

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

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

(We shall drop the $^{\mathscr{T}}$ and $_{\mathscr{V}}$ whenever they are obvious.)

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 P', as $\nu Z.P \wedge [-]Z$, 'eventually P', as $\mu Z.P \vee \langle -\rangle Z$, and 'there is a path along which c happens infinitely often', as $\nu Y.\mu Z.\langle c\rangle Y \vee \langle -c\rangle Z$. (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 $\Sigma_0^{S\mu}$ and $\Pi_0^{S\mu}$ iff it contains no fix-point operators ('S' for 'simple' or 'syntactic'). To form the class $\Sigma_{n+1}^{S\mu}$, take $\Sigma_n^{S\mu} \cup \Pi_n^{S\mu}$, and close under (i) boolean and modal combinators, (ii) $\mu Z.\Phi$, for $\Phi \in \Sigma_{n+1}^{S\mu}$. Dually, to form $\Pi_{n+1}^{S\mu}$, take $\Sigma_n^{S\mu} \cup \Pi_n^{S\mu}$, and close under (i) boolean and modal combinators, (ii) $\nu Z.\Phi$, for $\Phi \in \Pi_{n+1}^{S\mu}$. Thus the examples above are in $\Pi_1^{S\mu}$, $\Sigma_1^{S\mu}$, and $\Pi_2^{S\mu}$ (but not $\Sigma_2^{S\mu}$) respectively. We shall say a formula is *strict* $\Sigma_n^{S\mu}$ if it is in $\Sigma_n^{S\mu} - \Pi_n^{S\mu}$.

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

$$\Upsilon_1 = \nu Y.\mu Z.\langle -\rangle Y \wedge (P \vee \langle -\rangle Z)$$

$$\Upsilon_2 = \nu Y.\langle -\rangle Y \wedge (\mu Z.P \vee \langle -\rangle Z)$$

—both are in $\Pi_2^{S\mu} - \Sigma_2^{S\mu}$. 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].)

A formula Φ is said to be in the classes $\Sigma_0^{N\mu}$ and $\Pi_0^{N\mu}$ iff it contains no fix-point operators. To form the class $\Sigma_{n+1}^{N\mu}$, take $\Sigma_n^{N\mu} \cup \Pi_n^{N\mu}$, and close under (i) boolean and modal combinators, (ii) $\mu Z.\Phi$, for $\Phi \in \Sigma_{n+1}^{N\mu}$, and (iii) substitution of $\Phi' \in \Sigma_{n+1}^{N\mu}$ for a free variable of $\Phi \in \Sigma_{n+1}^{N\mu}$ provided that no free variable of Φ' is captured by Φ; and dually for $\Pi_{n+1}^{N\mu}$. 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 $\Sigma_2^{N\mu}$. (Intuitively, we are allowed to have arbitrary syntactic alternation, as long as the real semantic dependency between the various fix-points is restricted. To take the simplest example, $\mu X_1.\nu X_2.\ldots.\nu X_{2n}.\mu X_{2n+1}.X_1 \vee X_{2n+1}$ has syntactic alternation depth $2n+1$, but its real alternation depth is just 1, since all we have is an inner minimal fix-point depending on an outer minimal fix-point, all the other fix-points 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 $\Phi \in \Sigma_{n+1}^{N\mu}$ and $\Phi \in \Pi_{n+1}^{N\mu}$.

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 $\bigcup_{\eta<\zeta} Z^\eta$.†

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 fix-point 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 fix-point 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

† This is the notation introduced in [Mos74]; earlier work, and also work in the modal mu-calculus, uses a different notation, with the effect that their Z^ζ is (equal to) our $Z^{<\zeta}$.

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. Note that 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 fix-point of the functional $\mathbf{X} \mapsto \{\, m \in \mathscr{M} \mid \mathscr{M} \vDash \phi(m, \mathbf{X}) \,\}$ (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 fix-point. 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 $\langle m, n \rangle$, $(n)_0$ and $(n)_1$ be standard pairing and unpairing functions.

The simplest examples of mu-arithmetic just use least fix-points 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, so can be expressed as a maximal fix-point $\nu X.x \neq 1 \wedge (x > 1 \Rightarrow (x-2) \in X)$. To produce natural examples involving alternating fix-points is rather difficult, since even one induction is already very powerful, and all 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 happens in $\mathbb{N}$, it is further possible to shuffle 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 fix-points, followed by a first order formula, which can in fact be restricted to be a boolean combination of Σ_1 and Π_1 formulae.)

Theorem 1. [Lub93] Every formula and set term is semantically equivalent to

one in pair-normal form (in a structure with pairing). □

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 $\phi \in \Sigma_{n+1}^{S\mu}$.

The following crucial lemma is also (taken as obvious) in [Lub93], modulo a technical point to be explained afterwards. It is indeed obvious from the proof of the pair-normal form theorem, but since it is only obvious from the proof, here it is worthwhile to prove the lemma directly—also, we shall use a simpler form of one of its constructions a little later. The following is a slightly condensed version of pp. 298–299 of [Lub93]; refer to the original for more explanation.

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

Proof. We transform ϕ to pair-normal form by structural induction.

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 $n \in Y(m, X)$ iff $\langle 1, \langle m, n\rangle\rangle \in W$ (remember that τ and ψ may depend on the individual variable x as well as the set variable X). Thus we translate the original term into

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

where 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. Next, to obtain pair-normal form, note that ψ' is (if not first-order) of the form $\rho \in \nu Z.\theta$; we can then pull the ν out through the rest of the formula, since $(\rho \in \nu Z.\theta) \wedge \vartheta$ is equivalent to $\rho \in \nu Z.\theta \wedge \vartheta$, etc.; and then repeat until the rest of the formula has been pushed into the inner $\Sigma_0^{S\mu}$ block. This procedure does not change the nesting of fix-points, so preserves the level in the hierarchy. (It may be noted that although Y depends on both x and X, we have only explicitly coded the dependency on x. By standard monotonicity arguments about adjacent fix-points of the same sign, the dependency on X can be ignored.)

For formulae, the booleans are dealt with by the fix-point shuffling just described. For the case $\sigma \in \Xi$, transform to $\langle 0, \sigma\rangle \in \Xi'$, where Ξ' is the construction of the previous paragraph.

For formulae $\exists x\,.\,\phi$, assume that ϕ is $\tau \in \mu Y.\psi$. The existential quantifier is pushed inside the fix-point 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; and the case of ϕ first-order is dealt with by adding a dummy fix-point of the appropriate sign.

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

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

Before quoting Lubarsky's main theorem, we need to address one technical point. Lubarsky uses a notion of alternation depth that is even weaker than $\Sigma_n^{S\mu}$, for reasons that are not clear to the present author (perhaps it is required for the mu-calculus over structures with no coding ability at all). Namely, the class Σ_n is closed under μ and under first-order quantifiers, but not under booleans: he has, rather, that if ϕ_1 and ϕ_2 are in $\Sigma_n \cup \Pi_n$, then $\phi_1 \vee \phi_2$ and $\phi_1 \wedge \phi_2$ are in $\Sigma_{n+1} \cap \Pi_{n+1}$ (and also all first-order formulae are $\Sigma_0 \cap \Pi_0$). However, everything in [Lub93] remains true (more simply, if anything) with the $\Sigma_n^{S\mu}$ hierarchy; in any case, the two notions coincide on formulae in pair-normal form. Henceforth we silently substitute $\Sigma_n^{S\mu}$ for the Σ_n of [Lub93].

Lubarsky then proves

Theorem 3. [Lub93] The hierarchy of the sets of integers definable by $\Sigma_n^{S\mu}$ formulae of the arithmetic mu-calculus is a strict hierarchy. □

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. The full paper contains an appendix with 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.

A *recursively presented transition system (r.p.t.s.)* is 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 $s \in \|[K]\Psi\|$ iff $\forall t \in \mathscr{S} \,.\, \forall a \in K \,.\, (s \xrightarrow{a} t) \Rightarrow \psi$ where ψ is the translation of $t \in \|\Psi\|$, and dually. Finally, for the fix-point 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. □

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 $s \in \|\Phi\|_{\mathscr{V}}^{\mathscr{T}}$. (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 $\phi(s_{z_1}, \ldots, s_{z_k})$.)

For every atomic formula $P(\tau)$ occurring in ϕ, we equip the modal mu-calculus with a variable P_τ such that $s \in \mathscr{V}(P_\tau)$ 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 $f_{(\tau_1,\ldots,\tau_k)}$, and the transitions given by $s \xrightarrow{f_{(\tau_1,\ldots,\tau_k)}} t$ 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 x and the transitions given by $s \xrightarrow{x} t$ 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 fix-points, 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 dually for $\wedge$. If ϕ is $\forall x \,.\, \psi$, then Φ is $[x]\Psi$, and if ϕ is $\exists x \,.\, \psi$, then Φ is $\langle x \rangle \Psi$.

If ϕ is an atomic formula $P(\tau)$, then Φ is $\langle \tau * \rangle P_\tau$, 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 $\langle \tau_1 * \rangle \cdots \langle \tau_k * \rangle \langle f_{(\tau_1,\ldots\tau_k)} \rangle \Upsilon$—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 fix-point operators of arithmetic into the fix-point operators of the modal mu-calculus.

For every fix-point μX or νX occurring in ϕ, and for every formula $\tau \in X$ in ϕ, we equip $\mathscr{T}$ with a label x_τ (recall that $\mu X.\psi$ is short for $\mu(x, X).\psi$). The transitions are given by $s \xrightarrow{x_\tau} t$ 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 fix-point.)

We now complete the translation: if ϕ is $\tau \in X$ for X a set variable, then Φ is $\langle \tau * \rangle \langle x_\tau \rangle X$. If ϕ is $\tau \in \mu X.\psi$, then Φ is $\langle \tau * \rangle \langle x_\tau \rangle \mu X.\Psi$; 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 $\exists t \,.\, s \xrightarrow{\tau *} \xrightarrow{x_\tau} t \wedge t_x \in \mathbf{X}$ iff $s \in \|\langle \tau * \rangle \langle x_\tau \rangle X\|$, as required.

For the fix-points 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 fix-point structure of ϕ is preserved in the translation to Φ. □

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 P_x and Q_x such that $\|P_x\| = \{(0, n)\}$ and $\|Q_x\| = \{\, (m, n) \mid m > 1 \,\}$. We also have a label f_x such that $(m, n) \xrightarrow{f_x} (m, m-2)$. Finally, we have a label $x_{f(x)}$ such that $(m, n) \xrightarrow{x_{f(x)}} (n, n)$. The translation into modal logic is then

$$\mu X.P_x \vee (Q_x \wedge \langle f_x \rangle \langle x_{f(x)} \rangle X)$$

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

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

Corollary 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

$\Sigma^{S\mu}_{n-1}$ modal formula can have the same denotation in $\mathscr{T}$, and so no $\Sigma^{S\mu}_{n-1}$ modal formula is logically equivalent to Φ. □

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 queasy about the transition 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. For the first point, note that the modal mu-calculus has the finite model property [StE89], so if a hierarchy is strict on any class of models, it is also strict on the class of finite models. Furthermore, although it looks as if we have relied on the existence of lots of labels, many of these can also be coded away using simple techniques, such those employed in chapter 5 of [Bra91] and in [Bra96]. The only critical labels are those for the first-order variables.

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. In fact, we shall prove that $\Sigma^{N\mu}_i$ formulae are equivalent (in mu-arithmetic) to $\Sigma^{S\mu}_i$ 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 $\Sigma^{N\mu}_n$ 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 $\Sigma^{N\mu}_0 = \Pi^{N\mu}_0$ formulae are the first-order formulae and the $\Sigma^{N\mu}_0 = \Pi^{N\mu}_0$ set terms are the set variables; $\Sigma^{N\mu}_{n+1}$ is formed by taking $\Sigma^{N\mu}_n \cup \Pi^{N\mu}_n$ and closing under (i) first-order connectives and $\in$; (ii) $\mu X.\psi$ for $\psi \in \Sigma^{N\mu}_{n+1}$, provided that ψ has at most x as a free individual variable; (iii) substitution of a $\Sigma^{N\mu}_{n+1}$ set term Ξ for a free set variable of $\psi \in \Sigma^{N\mu}_{n+1}$ provided that no free set variable of Ξ is captured by ψ. Hence the translation takes modal $\Sigma^{N\mu}_n$ to arithmetic $\Sigma^{N\mu}_n$.

Theorem 7. 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 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 fix-point 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 fix-point to have the required form.

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

The non-trivial case is the substitution rule (iii). Let ψ, Ξ be the $\Sigma_{n+1}^{N\mu}$ 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 $\Sigma_{n+1}^{S\mu}$, which they are by induction, this formula is also $\Sigma_{n+1}^{S\mu}$, so we are done. □

Corollary 8. The denotation of a modal $\Sigma_n^{N\mu}$ formulae is arithmetic $\Sigma_n^{S\mu}$-definable. □

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

Theorem 9. The alternation hierarchy for the modal mu-calculus 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 $\Sigma_n^{N\mu}$; we can define Emerson–Lei versions $\Sigma_n^{EL\mu}$ of these classes by modifying clause (iii) as follows: (iii′) substitution of $\Phi' \in \Sigma_{n+1}^{EL\mu}$ for a free variable of $\Phi \in \Sigma_{n+1}^{EL\mu}$ provided that Φ' is a closed formula. A simple induction now shows the following

Lemma 10. A formula Φ has Emerson–Lei alternation depth $\leq n$ iff $\Phi \in \Sigma_{n+1}^{EL\mu} \cap \Pi_{n+1}^{EL\mu}$, for $n \geq 0$. □

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

Corollary 11. 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 "⇒" direction of the above lemma replacing $\Sigma^{EL\mu}$ by $\Sigma^{N\mu}$, and probably also the "⇐" direction, though this remains to be proved. 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 $\Sigma_4^{EL\mu}$.

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, in his case, though the argument should work for Niwiński) alternation hierarchy. The technique is a very delicate topological analysis of the finite models of formulae of a logic closely related to the modal mu-calculus, 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 forthcoming 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 helpful additional remarks on his work; and the CONCUR referees for useful suggestions.

6 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**, 451–459 (1990).

[Bar75] J. Barwise, *Admissible sets and structures*, Springer-Verlag, Berlin/New York (1975).

[Bra91] J. C. Bradfield, *Verifying Temporal Properties of Systems.* Birkhäuser, Boston, Mass. ISBN 0-8176-3625-0 (1991).

[Bra96] J. C. Bradfield, On the expressivity of the modal mu-calculus. *Proc. STACS '96*, LNCS **1046**, 479–490 (1996).

[Bra96a] J. C. Bradfield, The modal mu-calculus alternation hierarchy is strict. Online via the Web page `http://www.dcs.ed.ac.uk/home/jcb/` or by ftp at `ftp://ftp.dcs.ed.ac.uk/export/jcb/Research/althi.ps.gz`.

[EmJ88] E. A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. Extended version from *FOCS '88.* (1988).

[EmL86] E. A. Emerson and C.-L. Lei, Efficient model checking in fragments of the propositional mu-calculus. *Proc. First IEEE Symp. on Logic in Computer Science* 267–278 (1986).

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

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

[Len96] G. Lenzi, A hierarchy theorem for the mu-calculus. To appear in *Proc. ICALP '96.*

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

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

[Mos74] Y. N. Moschovakis, *Elementary induction on abstract structures*, North-Holland, Amsterdam (1974).

[Niw86] D. Niwiński, On fixed point clones. *Proc. 13th ICALP*, LNCS **226** 464–473 (1986).

[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 Fenstad and Hinman (ed.), *Generalized recursion theory*, North-Holland, Amsterdam 301–381 (1974).

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

[Sti95] C. Stirling, Local model checking games. *Proc. Concur '95*, LNCS **962**, 1–11 (1995).

[Tak86] M. Takashashi, The greatest fixed-points and rational omega-tree languages. *Theor. Comput. Sci.* **44** 259–274 (1986).