Contents

On Distributive Fixpoint Expressions

Helmut Seidl* Damian Niwiński†

June 1998

Abstract

For every fixpoint expression e of alternation-depth r, we construct a new fixpoint expression e' of alternation-depth 2 and size $\mathcal{O}(r \cdot |e|)$. Expression e' is equivalent to e whenever operators are distributive and the underlying complete lattice has a co-continuous least upper bound. We show that our transformation is optimal not only w.r.t. alternation-depth but also w.r.t. the increase in size of the resulting expression.

1 Introduction

Fixpoint calculus is a logical formalism based on explicit notation for inductive and co-inductive definitions. It is recognized as a useful framework especially for reasoning about temporal properties of finite state systems.

The role of *alternation* of least (μ) and greatest (ν) fixpoint operators as a source of a sharp expressive power for the fixpoint calculus has been recognized in the early 1980's [12, 6, 5]. In particular, D. Park in his studies on the semantics of parallelism [12] observed that the *fair merge* of two infinite sequences can be adequately characterized only using *both* extremal fixpoint operators. It was shown somewhat later that the alternation of μ and ν gives rise to a strict hierarchy of properties of infinite *trees*, corresponding to some hierarchy of Rabin automata [8, 10]. The strongest result in this direction is the infinity of the alternation hierarchy in Kozen's propositional modal μ-calculus, which has been only recently established by J. Bradfield [3] and G. Lenzi [7].

These facts can be contrasted with the early result by Park [13] stating that the hierarchy collapses to the $\nu\mu$ level in the algebra of ω–languages with the basic operations consisting of prefixing by a single letter, and binary set union[1]. In this framework the fixpoint definable ω–languages are precisely ω–regular languages, and collapsing of the hierarchy to the $\nu\mu$ level follows from the structure of acceptance condition of Büchi automata. The result was subsequently generalized in the second author's doctoral dissertation [9] to what is called there *monadic Kleene algebras*; it was however never published elsewhere.

*FB IV – Informatik, Universität Trier, D–54286 Trier, Germany, email: `seidl@uni-trier.de`

†Inst. of Informatics, University of Warsaw, Ul. Banacha 2, 02-097 Warsaw, Poland, email: `niwinski@mimuw.edu.pl`.
Supported by Polish KBN grant No. 8 T11C 002 11

[1] As this author points out, this result was obtained in collaboration with J. Tiuryn.

In recent years, an interest is coming back in fixpoint calculi *weaker* than Kozen's μ-calculus. This is because of the *model checking* problem. The general question whether a given finite state transition system satisfies or not a μ-calculus formula, appears to be relatively hard[2]. Therefore, it would be profitable to distinguish a comprehensive enough *sub-calculus* enjoying better complexity. One such successful case is a sub-logic L_2 proposed by Emerson, Jutla and Sistla [4], and proven to be strong enough to subsume, and in fact to be equivalent to, the program logic ECTL*.

Now, an observation can be made that the logic L_2 admits a collapsing property exactly as the aforementioned calculus of Park, and in fact for the same reason (the structure of Büchi automata). One is tempted to think that the collapsing of alternation hierarchy is a characteristic property that separates "easy" and "difficult" fragments of the μ-calculus, as far as complexity is concerned. In this context, we believe it worthy to understand the early result by Park in its full generality, and also to clarify its complexity aspect, i.e. the complexity of reduction of a fixpoint expression to the $\nu\mu$ level. To our best knowledge, the last question has not yet been considered in literature.

This clarification has also practical implications. Model-checkers for L_2-formulas can be constructed according to (at least) three strategies. The first one simply applies methods for general μ-calculus model-checking. The disadvantage is that efficiency problems may arise due to arbitrary nesting of fixpoints. A second alternative, as proposed by Bhat and Cleaveland [2], tries to improve upon this by exploiting special properties of L_2-formulas. The complication still is that potentially arbitrarily nested fixpoints have to be dealt with. The third strategy, therefore, relies on the collapsing property. It first eliminates deep nesting of alternating fixpoints by transforming the formula into an equivalent one of small alternation-depth (namely, 2). To the resulting formula either a specialized algorithm is applied or ordinary fixpoint iteration. The efficiency of this third approach, however, crucially depends on how small the extra overhead through transformation of the formula can be made.

To attack this question, we first give a general algebraic condition sufficient for the hierarchy of fixpoint expressions to collapse to the $\nu\mu$–level. It requires of all the basic operations (over a complete lattice) to be distributive, and of the binary least upper bound "$\sqcup$", additionally, to commute with the greatest lower bound of downward directed sets. (This generalization is stronger than the aforementioned one of [11].) Next, we show that a fixpoint expression e of the alternation-depth r can be transformed uniformly over the structures of the aforementioned property to an expression of alternation-depth 2 (in fact, $\nu\mu$) and size $O(r \cdot |e|)$. We finally prove that this estimation of the size of the resulting expression is optimal.

2 Distributive Functions

Assume D_1, D_2 are complete lattices. Function $f : D_1 \to D_2$ is called *distributive* iff $f\,(d_1 \sqcup d_2) = f\,d_1 \sqcup f\,d_2$.

Fact 1 *for every non-empty finite set V, $f : (V \to D_1) \to D_2$ is distributive iff*

$$f = \bigsqcup_{x \in V} f_x \circ \pi_x$$

[2] Model checking for Kozen's μ-calculus is in $NP \cap co\text{-}NP$ (even in $UP \cap co\text{-}UP$), but it is not known whether it is in P.

for unary distributive functions $f_x : D_1 \to D_2$ *where for* $x \in V$, $\pi_x : (V \to D_1) \to D_1$ *is given by* $\pi_x \rho = \rho\, x$.

Proof. Define $f_x\, d = f\,\langle x, d\rangle$ where $\langle x, d\rangle : V \to D_1$ maps x to d and all $y \neq x$ to $\perp$. □

Therefore, we usually consider just *unary* distributive functions.

Assume D is a complete lattice. For monotonic $f : D \to D$, function f^* is defined as the least fixpoint

$$f^* = \mu\, F.f \circ F \sqcup I$$

Whenever f is distributive, then f^* is distributive as well. Non-empty subset $X \subseteq D$ is called *co-directed* iff $x_1, x_2 \in X$ implies that X also contains some z with $z \sqsubseteq x_1$ and $z \sqsubseteq x_2$. Least upper bound "$\sqcup$" is called *co-continuous* if it commutes with co-directed greatest lower bounds, i.e.,

$$d \sqcup \bigsqcap X = \bigsqcap_{x \in X} (d \sqcup x)$$

for every $d \in D$ and co-directed subset $X \subseteq D$.

The last condition holds in particular in any complete Boolean algebra [14]. Especially, this is the case for $D = 2^S$ (ordered by "$\subseteq$") for any base set S.

Our considerations rely onto the following key observation:

Theorem 1 *Assume* D *is a complete lattice, and* $f, g : D \to D$ *are monotonic functions where* $g\, x = f\, x \sqcup d$ *for some* $d \in D$. *Then*

1. $\mu\, g = f^*\, d$;
2. *If binary* "$\sqcup$" *is co-continuous and* f *is distributive, then*

$$\nu\, g = \nu\, f \sqcup \mu\, g$$

Here, $\mu\, g$ *and* $\nu\, g$ *denote the least and greatest fixpoints of* g, *respectively.*

For a proof see appendix A.

3 Fixpoint Expressions

We fix a countably infinite set of variables $\mathcal{X}$, a set of constant symbols $\mathbf{C}$, and a set of unary function symbols $\mathbf{F}$ ($\mathbf{C}$ and $\mathbf{F}$ are not necessarily finite).

A *fixpoint expression* e is given by the following grammar:

$$e \quad ::= \quad c \mid x \mid f\, e \mid \mu\, x.e \mid \nu\, x.e \mid e_1 \sqcup e_2$$

where $c \in \mathbf{C}$ and $f \in \mathbf{F}$.

A *interpretation* is a complete lattice D given together with an element c^D, for each $c \in \mathbf{C}$, and a distributive function $f^D : D \to D$, for each $f \in \mathbf{F}$.

The *meaning* $[\![e]\!]_D$ of a fixpoint expression e is a function $(\mathcal{X} \to D) \to D$ mapping variable assignments to values. It is defined inductively as follows. (We usually omit the subscript D for simplicity. The symbols $\bigsqcap, \bigsqcup, \sqcup$ refer to the corresponding operations in D.)

$$\begin{array}{rcl} [\![c]\!]\,\rho & = & c^D \\ [\![x]\!]\,\rho & = & \rho\, x \\ [\![f\, e']\!]\,\rho & = & f^D\,([\![e']\!]\,\rho) \\ [\![\mu\, x.e']\!]\,\rho & = & \bigsqcap\{d \mid G\,d \sqsubseteq d\} \\ [\![\nu\, x.e']\!]\,\rho & = & \bigsqcup\{d \mid G\,d \sqsupseteq d\} \\ [\![e_1 \sqcup e_2]\!]\,\rho & = & [\![e_1]\!]\,\rho \sqcup [\![e_2]\!]\,\rho \end{array}$$

where $G : D \to D$ is given by

$$G\,d \quad = \quad [\![e']\!]\,(\rho + \{x \mapsto d\})$$

Proposition 2 *If all f^D are distributive and the operation $\sqcup$ is co-continuous in D, then $[\![e]\!]_D$ is distributive.*

In fact, it is this proposition which allows us to call e itself distributive and thus justifies the title of our paper.

Proof. We proceed by induction on the structure of e. The only non-trivial case is where e equals a fixpoint expression $\sigma x.e'$ ($\sigma \in \{\mu, \nu\}$). Let $\mathcal{Y}$ denote the set of free variables of e, and $\rho : \mathcal{Y} \to D$. By fact 1, $[\![e']\!]\,(\rho + \{x \mapsto d\}) = f_x\, d \sqcup f'\,\rho$ for distributive functions $f_x : D \to D$ and $f' : (\mathcal{Y} \to D) \to D$.

If e is a least-fixpoint expression (i.e., $\sigma = \mu$), then by theorem 1,

$$[\![e]\!] = [\![\mu x.e']\!] = f_x^* \circ f'$$

Since distributive functions are closed under composition and "$*$", the assertion follows.

Accordingly, if e is a greatest-fixpoint expression (i.e., $\sigma = \nu$), then by theorem 1,

$$[\![e]\!] = [\![\nu x.e']\!] = (\lambda\rho.\, \nu f_x) \sqcup (f_x^* \circ f')$$

Since constant functions are distributive and distributive functions are closed under "$\sqcup$", the assertion follows. □

4 Alternation-Depth

For closed fixpoint expression e, we need to determine the maximal number of nesting of greatest and least fixpoints. This is achieved by the following top-down traversal through the syntax tree of e which assigns a *level* $\alpha\, e' \geq 1$ to every subexpression e' of e. α is determined according to the following two rules:

1. $\alpha\, y$ equals the level of the fixpoint expression where y is bound;
2. Let $\mathcal{Y}$ denote the set of free variables of e'. If $\mathcal{Y} = \emptyset$, then $\alpha\, e' = 1$. Otherwise, assume we are given levels $\alpha\, y$ for all $y \in \mathcal{Y}$ with $r = \bigsqcup_{y \in \mathcal{Y}} \alpha\, y$ and $\alpha\, y' = r$. We distinguish two cases.
   - e' is not a fixpoint expression. Then $\alpha\, e' = r$.
   - $e' \equiv \sigma x.e_1$. Then $\alpha\, e' = r$ if either both x and y' are μ-variables or both are ν-variables. Otherwise, $\alpha\, e' = r + 1$.

This definition consistently allows to assign levels to all subexpressions of a given expression – provided we are given the levels of the free variables. The *alternation-depth* of closed expression e then is defined as the maximal level of a subterm of e. (This definition is consistent with the alternation hierarchy of fixpoint terms considered in [11].) Expressions of alternation-depth 1 are also called *alternation-free.*

We illustrate this definition by the following example.

Example 1 Let

$$e \equiv \mu x.\nu y.f\, y \sqcup (\mu z.g\, z \sqcup x)$$

Since e is closed, variable x receives level 1. Variable y receives level 2 since x occurs free in $\nu y.f\, y \sqcup (\mu z.g\, z \sqcup x)$. Variable z, however, again receives level 1, since the only free variable in subexpression $\mu z.g\, z \sqcup x$ is x which is also a μ-variable and of level 1. Overall, we find that e is of alternation-depth 2. □

The level numbers of subexpressions can then be used to construct a suitable vectorial fixpoint expression. Essentially, variables of the same level can be grouped into the same vector.

Theorem 2 *For every expression e, an expression $\tilde{e}$ can be constructed such that the following holds.*

1. $[\![e]\!]_D = [\![\tilde{e}]\!]_D$ *for every interpretation D with distributive operators, provided that operation "$\sqcup$" is co-continuous.*
2. *$\tilde{e}$ is of alternation-depth at most 2 where every greatest fixpoint subexpression has level 1.*
3. *$\tilde{e}$ has size $\mathcal{O}(r \cdot |e|)$ where r is the alternation-depth of e.* □

Expressions with property (2) are also said to be of class $\nu\mu$.

Proof. We call greatest-fixpoint expression $e \equiv \nu x.e_1$ *simple* if

- e is closed;
- e_1 does not contain closed subexpressions;
- all ν-variables are of level 1.

Thus, expression $\nu\, x.f\, x \sqcup (\mu\, y.g\, y \sqcup x)$ is simple whereas expression $\nu\, x.f\, x \sqcup (\mu\, y.g\, y)$ is not, since subexpression $\mu\, y.g\, y$ is closed.

Expression e is called *r-clean* if it is obtained from an expression without ν-variables at levels $\geq r$ by replacing some constant symbols with simple expressions. Clearly, simple expressions are of alternation-depth at most 2. Since simple expressions are closed, we conclude that closed expressions which are clean up to level 1 are also of alternation-depth at most 2.

Assume we are given an arbitrary expression $\bar{e}$. W.l.o.g. we assume that all bound variables in $\bar{e}$ are distinct, and all fixpoint subexpressions are non-trivial, i.e., fixpoint variables occur in the respective bodies at least once. In order to obtain a sufficiently small transformed expression, we proceed by removing several greatest fixpoints *simultaneously*. Assume $r \geq 1$ is the maximal level of ν-variables in $\bar{e}$. By definition, $\bar{e}$ is $(r+1)$-clean. Our goal is to transform $\bar{e}$ into expression $\tilde{e}$ which is r-clean.

For this consider greatest-fixpoint expression $e = \nu x_1.e'$ of level r. Let $\mathcal{X} = \{x_1, \ldots, x_m\}$ and $\mathcal{Y} = \{y_1, \ldots, y_k\}$ denote the set of greatest-fixpoint variables of e of level r and the set of free variables, respectively.

Let us furthermore assume that

1. e is not simple,
2. $\alpha\, y_i < r$ for all i; and
3. e is minimal with this property.

Let $e_s, s = m+1, \ldots, n$, denote the maximal subexpressions of e' which only contain free variables from $\mathcal{Y}$. By construction, none of these may contain greatest-fixpoint subexpressions of level r. We aim at replacing e by expression $e_1 \sqcup e_2$ with "nicer" properties.

Expression e_1 is obtained by replacing all subexpressions $e_s, s > m$, with $\perp$ and then absorbing lubs $e_0 \sqcup \perp$ resp. $\perp \sqcup e_0$ into e_0. Instead, expression e_2 is constructed by replacing in e all occurrences of binders νx_j, $j = 1, \ldots, m$, with μx_j, respectively.

Let $\tilde{e}$ be obtained from $\bar{e}$ by replacing e with $e_1 \sqcup e_2$. Let D be an interpretation where operation "$\sqcup$" is co-continuous. We claim:

1. $[\![e]\!] = [\![e_1 \sqcup e_2]\!]$;
2. e_1 is simple;
3. all variables from $\mathcal{X}$ in e_2 receive level $r-1$ (relative to $\tilde{e}$);
4. all remaining ν-variables in e_2 (relative to $\tilde{e}$) are of level at most $r-1$.

We only prove claim 1. Assume $\nu x_i.e'_i$ is the subexpression of e' where x_i is bound. Thus, especially, $e'_1 = e'$. Since all these greatest-fixpoint expressions have identical level r, their fixpoints can be computed jointly. So, assume e''_i is obtained from e'_i by replacing occurrences of greatest-fixpoint expressions $\nu x_j.e'_j, j \neq i$, with x_j, respectively. Then

$$[\![e]\!]\, \rho = (\nu G)\, x_1$$

where $G : (\mathcal{X} \to D) \to \mathcal{X} \to D$ is defined by

$$G\, \delta\, x_i = [\![e''_i]\!]\, (\delta + \rho)$$

By distributivity, function G can be decomposed. For $i, j \in \{1, \ldots, m\}$, define expression e'_{ij} by replacing in e''_i all $x_{j'}, j' \neq j$, as well as all expressions $e_s, s > m$, with $\perp$. Additionally for $i = 1, \ldots, m$, define expression $e^{\mathcal{Y}}_i$ by replacing in e''_i all variables x_j, $j = 1, \ldots, m$, with $\perp$.

By this construction, e'_{ij} contains x_j as only free variable whereas $e^{\mathcal{Y}}_i$ contains free variables only from $\mathcal{Y}$, and we find:

$$[\![e''_i]\!]\, (\delta + \rho) = \bigsqcup_{j=1}^{m} [\![e'_{ij}]\!]\, \{x_j \mapsto \delta\, x_j\} \sqcup [\![e^{\mathcal{Y}}_i]\!]\, \rho$$

for $i = 1, \ldots, m$. Therefore, $G\,\delta = F\,\delta \sqcup H$ where $F : (\mathcal{X} \to D) \to \mathcal{X} \to D$ and $H : \mathcal{X} \to D$ are given by

$$\begin{array}{rcl} F\,\delta\,x_i & = & \bigsqcup_{j=1}^{m} [\![e'_{ij}]\!]\,\{x_j \mapsto \delta\,x_j\} \\ H\,x_i & = & [\![e_i^y]\!]\,\rho \end{array}$$

Consequently,

$$\nu G = \nu\,\delta.F\,\delta \sqcup H = \nu F \sqcup \mu G$$

By construction, $(\nu F)\,x_1$ equals $[\![e_1]\!]\,\emptyset = [\![e_1]\!]\,\rho$ (remember that e_1 is closed) whereas $(\mu G)\,x_1$ equals $[\![e_2]\!]\,\rho$.

This finishes the proof of our claim.

Applying this transformation successively to the greatest fixpoint expressions in $\bar{e}$ of level r, we obtain a r-clean expression equivalent to $\bar{e}$. All the newly created expressions are simple and jointly bounded in size by the size of the original expression. Since the addition of simple subterms may occur at most at r levels, we obtain the claimed upper bound for the size of the resulting expression. □

Remark By fact 1, the above result can be easily extended to the case where the expressions involve function symbols of arbitrary finite arity, provided that the interpretation $g^D : D^k \to D$ of a k–ary symbol g is distributive over D^k.

We illustrate the construction underlying theorem 2 by the following example.

Example 2 Consider

$$e \equiv \mu\,x_3.f\,(\nu\,x_2.g\,(\mu\,x_1.h_1\,x_1 \sqcup h_2\,x_2 \sqcup h_3\,x_3) \sqcup d)$$

The only greatest-fixpoint expression occurs at level 2. It is given by:

$$\nu\,x_2.g\,(\mu\,x_1.h_1\,x_1 \sqcup h_2\,x_2 \sqcup h_3\,x_3) \sqcup d$$

which is transformed into $e_1 \sqcup e_2$ with

$$\begin{array}{rcl} e_1 & \equiv & \nu\,z.g\,(\mu\,x.h_1\,x \sqcup h_2\,z) \\ e_2 & \equiv & \mu\,x_2.g\,(\mu\,x_1.h_1\,x_1 \sqcup h_2\,x_2 \sqcup h_3\,x_3) \sqcup d \end{array}$$

Note that, in the newly generated greatest-fixpoint expression e_1, we could eliminate not only constants but also closed subexpressions and even subexpressions which become closed if all global variables are replaced with $\perp$.

Thus, finally expression e is equivalent to:

$$\begin{array}{l} \mu\,x_3.f\,(\nu\,z.g\,(\mu\,x.h_1\,x \sqcup h_2\,z) \sqcup \\ \quad \mu\,x_2.g\,(\mu\,x_1.h_1\,x_1 \sqcup h_2\,x_2 \sqcup h_3\,x_3) \sqcup d) \end{array}$$

The remaining ν-variable z is indeed of level 1. □

Our theorem 2 results in an alternative proof that fragment L_2 of the propositional μ-calculus (interpreted over finite or infinite transition systems) is no more expressive than its alternation-depth-2 fragment.

If $\Phi(m)$ is the complexity of evaluating expressions of size at most m and alternation-depth ≤ 2 in D with co-continuous "$\sqcup$" and distributive operators, this results in a procedure evaluating expressions of alternation-depth $r > 0$. We obtain:

Corollary 3 *Given a complete lattice D with co-continuous "$\sqcup$" and distributive operators, every closed fixpoint expression e of alternation-depth $r > 0$ can be evaluated in time $\mathcal{O}(r \cdot \Phi(|e|))$.* □

In the case of finite-state model-checking, theorem 3 recovers the complexity bounds given by Bhatt and Cleaveland in [2] from the methods of Andersen and Vergauwen [1] for alternation-depth 2.

5 Optimality

In terminology of [11] theorem 2 implies that the alternation-depth hierarchy of fixpoint expressions collapses for lattices D with co-continuous "$\sqcup$" and distributive operators not only at alternation-depth 2 but, more precisely, at the class $\nu\mu$. One may wonder whether this result could be even further improved upon, e.g., whether or not class $\mu\nu$ might be sufficient as well. Formally, expression e is in class $\mu\nu$ iff all μ-variables are at level 1. For such expressions, however, the transformation of theorem 2 reveals:

Fact 3 *Every fixpoint expression e in $\mu\nu$ is equivalent to an expression e' with the following properties:*

1. *e' is alternation-free;*
2. *$|e'| = \mathcal{O}(|e|)$.* □

Thus, class $\mu\nu$ is no more expressive than alternation-free fixpoint expressions.

Example 3 Consider expression

$$e \equiv \mu x.\nu y.f\, y \sqcup (\mu z.g\, z \sqcup x)$$

of example 1 which is contained in class $\mu\nu$. Applying the transformation from theorem 2, we obtain

$$\mu x.(\nu y.f\, y) \sqcup (\mu y.f\, y \sqcup (\mu z.g\, z \sqcup x))$$

which is alternation-free. □

It remains to prove that alternation-depth 2 fixpoint expressions in general are more expressive than alternation-free ones. Let Σ be a finite alphabet and let $D_\Sigma = 2^{\Sigma^\omega}$ be the complete lattice of all subsets of infinite words over Σ, ordered by subset inclusion. Let $\mathbf{C} = \emptyset$ and $\mathbf{F} = \Sigma$. We interpret a symbol $\sigma \in \Sigma$, as left-multiplication with σ, i.e., the operation $\lambda L \in D_\Sigma.\sigma L$. Then we have the following.

Fact 4 *For $L \subseteq \Sigma^\omega$ the following statements are equivalent:*

1. *$L = [\![e]\!]\,\emptyset$ for some closed alternation-free fixpoint expression e over D_Σ.*
2. *L can be accepted by a non-deterministic finite Büchi automaton where each strong component* either *consists of accepting states* or *of non-accepting states.* □

Optimality of our construction w.r.t. alternation-depth therefore follows from the following example.

Example 4 Consider expression

$$e \equiv \nu x_2.\mu x_1.1\, x_1 \sqcup 2\, x_2$$

of alternation-depth 2 (more precisely, of class $\nu\mu$). When interpreted over D_Σ (where $\Sigma = \{1, 2\}$), expression e denotes the language

$$\nu(1^* \circ 2) = (1^*2)^\omega$$

i.e., the set of all infinite words containing infinitely many 2. It is not difficult to see that every finite Büchi automaton accepting this language has a strong component which both contains accepting and non-accepting states. Therefore by fact 4, this language cannot be expressed by any alternation-free fixpoint expression. □

Thus, we have proved:

Theorem 4 *When interpreted over D_Σ, alternation-depth-2 fixpoint expressions (from the class $\nu\mu$) are strictly more expressive than alternation-free fixpoint expressions.* □

Recall that a slightly more complicated candidate for non-equality of alternation-depths 1 and 2 has been suggested by Park in [12], however, without actual proof.

A generalization of the expression in example 4 also allows us to show that the size of the $\nu\mu$–level expression provided by theorem 2 is optimal (up to a constant factor).

Let $\Sigma_r = \{0, 1, 2, \ldots, r\}$, where $r = 2k$. Let $L_{r,0} \subseteq (\Sigma_r \backslash \{0\})^\omega$ be the set of infinite words u such that $\limsup_{j\to\infty} u(j)$ is even, in other words, the highest number occurring infinitely often in u is even. (Here and further we identify an infinite word u with the sequence of its letters $u(0), u(1), \ldots$.) Finally, let $L_{r,n}$ denote the language of infinite words obtained from $L_{r,0}$ by inserting 0^n before every letter $s > 0$. It is well known that $L_{r,n}$ can be defined by a fixpoint expression

$$\begin{array}{rcl} e_{r,n} & = & \nu x_r.\mu x_{r-1}.\ldots\nu x_2.\mu x_1. \\ & & \quad 0^n\,(1\, x_1 \sqcup \ldots \sqcup r\, x_r) \end{array}$$

(see, e.g., [11], Section 5). Note that the above expression has alternation-depth r and is of size $\mathcal{O}(n + r)$. Hence theorem 2 gives us an equivalent $\nu\mu$ expression of size $\mathcal{O}(r \cdot (n + r))$. We will argue that this size is indeed required. Please note that this is slightly more than just to say that the estimation in theorem 2, given in terms of two parameters r and $|e|$, is exact. The example shows that the size factor $|e|$ matters also for fixed r.

Theorem 5 *Every fixpoint expression for $L_{r,n}$ of class $\nu\mu$ has size $\Omega(r \cdot (n + r))$.*

Proof. In order to prove the claim, we will use the correspondence between $\nu\mu$ expressions and Büchi automata. A linear translation of $\nu\mu$ expressions over D_Σ into Büchi automata was presented already by Park [12]. A construction described in [11] translates a fixpoint expression e (with arbitrary nesting of μ and ν) into a Rabin automaton with no more than $|e|$ states, and if e is of class $\nu\mu$, the result is a Büchi automaton.

However, the number of states is not convenient measure for us here, and neither would be the number of transitions. Instead, we introduce another simple concept.

Let A be a Büchi automaton with the set of states Q and the transition table given by relation $Q \times \Sigma \times Q$. An *arrow* of A is any pair (σ, p) such that, for some $q \in Q$, the triple (q, σ, p) is a transition of A. It is plain to see that the construction of [11] translates a fixpoint expression e into an automaton with no more than $|e| + |\Sigma|$ arrows. The component $|\Sigma|$ is required because the resulting automaton may need a subprogram recognizing the set of *all* infinite words over Σ which may be described by a constant-length fixpoint expression $\nu x.x$. Clearly, one such subprogram in an automaton suffices to do the job, and it costs precisely $|\Sigma|$ arrows.

In our case, this means that any $\nu\mu$ expression e defining $L_{r,n}$ is equivalent to a Büchi automaton with no more than $|e| + r$ arrows. On the other hand, we will show that any nondeterministic Büchi automaton recognizing $L_{r,n}$ must have at least $k \cdot (n + \frac{k+1}{2})$ arrows, where $r = 2k$. From there, the postulated lower bound for $|e|$ will clearly follow.

Let A be a nondeterministic Büchi automaton recognizing $L_{r,n}$. Our proof proceeds in two steps. First, we claim that, for any even $s \in \Sigma_r$, there must be transitions (q_s, s, p_s), (q_{s+2}, s, p_{s+2}), $\ldots$, (q_r, s, p_r), where the states $p_s, p_{s+2}, \ldots, p_r$ are all different. Intuitively, the state $p_{s+2\ell}$ "waits for the letter $s + 2\ell$".

Our second claim just considers the distinct states $p_2, p_4, \ldots, p_r$ constructed for letter $s = 2$ according to the first claim. It states that there must exist additional states $p_{j,m}$, for $j = 2, 4, \ldots, r$ and $m = 1, \ldots, n$, such that the automaton has transitions $(p_j, 0, p_{j,1})$, and $(p_{j,m-1}, 0, p_{j,m})$ ($j = 2, 4, \ldots, r$, $m = 1, \ldots, n$), and all these transitions are pairwise distinct. While our first claim gives us $\frac{k \cdot (k+1)}{2}$ transitions for even letters $s > 0$, our second claim guarantees additional $k \cdot n$ transitions for letter 0.

Here, we give a sketch how the first claim can be proven. Assume $r \geq 6$ and $n = 0$. We show that there must be at least three different transitions of the form $(q, 2, p)$. We leave an easy generalization of this fact to the reader.

Fix a constant K greater than the number of states of A. Recall that a run of a Büchi automaton on an infinite word u can be presented as an infinite word, r say, over the automaton's states, such that $(r(t), u(t), r(t+1))$ is a transition, for $t = 0, 1, \ldots$. A run is accepting if some of (especially designated) accepting states occur infinitely often [15].

Consider infinite words

$$u = 12\,12\,12\,12\ \ldots$$

and

$$v = (123)^K\, 4\, (123)^K\, 4\ \ldots$$

Fix some accepting runs of A on u and v, say r_u and r_v, respectively. Pick a state p such that, for infinitely many t's it holds that $r_u(t+1) = p$ while $u(t) = 2$ (we say that p occurs *after reading* 2). Next, by construction of v, we can find positions $i \leq \ell < \ell + 1 < j \leq 3K$, such that $r_v(i) = r_v(j)$,

$v(\ell) = 2$, and $v(\ell+1) = 3$. We claim that $q = r_v(\ell+1)$ must be different from p, and so the transition $(r_v(\ell), 2, q)$ must different from any transition $(r_u(t), 2, r_u(t+1))$ with $r_u(t+1) = p$. For, suppose the contrary. We can find positions $t_1 < t_2$ in the run r_u such that $r_u(t_1) = r_u(t_2) = p$, and some accepting state occurs between t_1 and t_2. Using the assumption that $p = q$, we can substitute the segment cut out of r_u by the positions t_1 and t_2 at the position ℓ in the run r_v. Note that in the resulting run (and the corresponding infinite word), we have two positions (specifically, i and $j + (t_2 - t_1)$) with the same state, and such that an accepting state occurs between these positions while the highest number in underlying word is 3. Then, we can pump this segment infinitely many times, and eventually obtain an accepting run of A on an infinite word with lim sup equal to 3, a contradiction.

Note that the same argument would apply if, instead of selecting the positions i, ℓ, j in the first block (i.e. $\leq 3K$), we would select them in any other block between $m \cdot (3K+1)$ and $(m+1) \cdot (3K+1) - 1$.

Now let

$$w = (12345)^K \, 6 \, (12345)^K \, 6 \, \ldots$$

and let r_w be an accepting run on w. Pick a state q that occurs infinitely often in the run r_v after reading 2, and between some two repeated states within the same block $(123)^K$. Thus, in their neighbourhoods, these occurrences of q are analogous to the position ℓ discussed in the preceding paragraph. By the argument there, we already know that $q \neq p$.

We claim that a third state, different from p and q, must occur after reading 2 in the run r_w. To show this, we can apply to w the trick previously applied to v. That is, we select positions $i \leq \ell < \ell + 3 < j \leq 5K$, such that $r_w(i) = r_w(j)$, $w(\ell) = 2$, and $w(\ell + 3) = 5$. Then, if $r_w(\ell)$ were equal to p or q, we would be able to construct accepting runs on words with lim sup equal to 5, by an argument analogous to the above. □

6 Conclusion

We have presented a transformation on fixpoint expressions which allows to reduce the alternation-depth to 2 in case all operators are distributive and binary "$\sqcup$" is co-continuous. More precisely, we have proved that under this proviso the alternation-depth hierarchy collapses at the level $\nu\mu$. We have also shown that our transformation is not only optimal w.r.t. the resulting alternation-depth, but it is also optimal (up to a constant factor) w.r.t. the size of the transformed expression.

Acknowledgement: We like to thank Thomas Wilke for interesting remarks on classes of Büchi automata.

References

[1] Hendrik Reif Andersen and Bart Vergauwen. Efficient Checking of Behavioral Relations and Modal Assertions Using Fixed-Point Inversion. In *7th International Conference on Computer-Aided Verification (CAV)*, pages 142–154. LNCS 939, 1995.

[2] Girish Bhat and Rance Cleaveland. Efficient Local Model Checking for Fragments of the Modal μ-Calculus. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, pages 107–126. LNCS 1055, 1996.

[3] J. C. Bradfield. The modal mu-calculus alternation hierarchy is strict. In *7th International Conference on Concurrency Theory (CONCUR)*, pages 233–246. LNCS 1119, 1996.

[4] E. Allen Emerson, C.S. Jutla, and A.P. Sistla. On Model-Checking for Fragments of μ-Calculus. In *5th International Conference on Computer-Aided Verification (CAV)*, pages 385–396. LNCS 697, 1993.

[5] E.A. Emerson and E.M. Clark. Characterizing correctness properties of parallel programs using fixpoints. In *ICALP'80*, pages 169–181. LNCS 85, 1980.

[6] L. Flon and N. Suzuki. Consistent and Complete Proof Rules for the Total Correctness of Parallel Programs. In *19th IEEE Symp. on Foundations of Computer Science (FOCS)*, 1978.

[7] G. Lenzi. A hierarchy theorem for the μ–calculus. In *23rd International Colloquium on Automata, Languages and Programming (ICALP)*, pages 87–109. LNCS 1099, 1996.

[8] Damian Niwiński. On fixed point clones. In *13th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 464–473. LNCS 226, 1986.

[9] Damian Niwiński. *Hierarchy of objects definable in the fixed point calculus,* in Polish. PhD thesis, University of Warsaw, 1987.

[10] Damian Niwiński. Fixed points vs infinite generation. In *3rd Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 402–409. IEEE, 1988.

[11] Damian Niwiński. Fixed Point Characterization of Infinite Behavior of Finite State Systems. *Theoretical Computer Science (TCS)*, 189:1–69, 1997.

[12] David M.R. Park. On the Semantics of Fair Parallelism. In *Abstract Software Specification*, pages 504–526. LNCS 86, 1980.

[13] David M.R. Park. Concurrency and Automata on Infinite Sequences. In *Theoretical Computer Science*, pages 167–183. LNCS 104, 1981.

[14] H. Rasiowa and R. Sikorski. *Mathematics of metamathematics.* PańStowe Wydawnictwo Naukowe, 1970.

[15] W. Thomas.

A Proof of Theorem 1

To prove the first assertion, consider g^τ and F^τ, τ an ordinal, defined by

- $g^0 = \bot$, $F^0 = \lambda x.\bot$;
- $g^{\tau+1} = g\, g^\tau$, $F^{\tau+1} = f \circ F^\tau \sqcup I$; and
- for τ a limit ordinal, $g^\tau = \bigsqcup_{\tau'<\tau} g^{\tau'}$, $F^\tau = \bigsqcup_{\tau'<\tau} F^{\tau'}$.

Then $\mu\, g = g^\sigma$ as well as $f^* = F^\sigma$ for some ordinal σ. Therefore, it suffices to verify through transfinite induction that, for every ordinal τ, $g^\tau = F^\tau\, d$.

Now consider the second assertion. By definition, $\mu\, g \sqsubseteq \nu\, g$. Since $f \sqsubseteq g$, we also have $\nu\, f \sqsubseteq \nu\, g$. Therefore, $\nu\, g \sqsupseteq \nu\, f \sqcup \mu\, g$, and it remains to prove the reverse inclusion. Therefore define

- $g^0 = \top = f^0$;
- $g^{\tau+1} = g\, g^\tau$, $f^{\tau+1} = f\, f^\tau$; and
- for τ a limit ordinal, $g^\tau = \bigsqcap_{\tau'<\tau} g^{\tau'}$, $f^\tau = \bigsqcap_{\tau'<\tau} f^{\tau'}$.

Then $\nu\, g = g^\sigma$ and $\nu\, f = f^\sigma$ for some ordinal σ, and it suffices to prove that, for every ordinal τ, $g^\tau \sqsubseteq f^\tau \sqcup \mu\, g$. Again, we proceed by transfinite induction.

The case $\tau = 0$ trivially holds.
If $\tau = \tau' + 1$, then

$$\begin{array}{rcl} g^\tau & = & g\; g^{\tau'} \\ & \sqsubseteq & g\; (f^{\tau'} \sqcup \mu\, g) \\ & = & f\, f^{\tau'} \sqcup g\, (\mu\, g) \\ & = & f^\tau \sqcup \mu\, g \end{array}$$

by ind. hypothesis and distributivity of f.

If τ is a limit ordinal, then, applying the induction hypothesis to all $\tau' < \tau$, we obtain:

$$\begin{array}{rcl} g^\tau & = & \bigsqcap_{\tau'<\tau} g^{\tau'} \\ & \sqsubseteq & \bigsqcap_{\tau'<\tau} (f^{\tau'} \sqcup \mu\, g) \\ & = & (\bigsqcap_{\tau'<\tau} f^{\tau'}) \sqcup \mu\, g \\ & = & f^\tau \sqcup \mu\, g \end{array}$$

and the assertion follows. □