6 Parity formulas & model checking

In this chapter we introduce *parity formulas* — structures that also go under the name of *alternating tree automata*, and that are closely related to various kinds of *modal equation systems*. Parity formulas are like ordinary (modal) formulas, with the difference that the underlying structure of a parity formula can be an arbitrary directed graph, not only a tree; and that one adds an explicit priority labelling to this underlying graph, to ensure a well-defined game-theoretical semantics. Parity formulas thus provide an alternative, bound-variable-free way of presenting the modal μ-calculus.

The advantage of parity formulas over ordinary ones comes out if one has an interest in computational aspects of the modal μ-calculus. At the heart of this algorithmic side lie the *parity games* that were introduced in the previous chapter, and the natural way of looking at evaluation games for the modal μ-calculus as parity games is *via* parity formulas. Parity formulas themselves have a completely straightforward connection with parity games, which allows for an unambiguous definition of the most relevant complexity measures of parity formulas. In particular, the *size* of a parity formula is simply defined as the number of vertices of its underlying graph, and its *index* corresponds to the maximal length of some naturally defined alternating chain of states.

This algorithmic transparency makes parity formulas very suitable as *yardsticks* for comparing various complexity measures of ordinary μ-calculus formulas. Taking this approach, we can further develop the complexity-theoretical framework of the μ-calculus by investigating the links between μ-calculus formulas and their representation as parity formulas. As an example we mention the analysis of the alternation depth of a formula that we undertake in Section 6.5.

In short, parity formulas are graph-based modal formulas with an added parity condition, that will allow us to view the evaluation games of μ-calculus formulas as instances of parity games. Providing a link between the world of μ-calculus formulas and that of parity games, they illuminate the complexity-theoretic analysis of various problems related to the modal μ-calculus, in particular model checking.

Parity formulas can also be studied in their own right, as an interesting generalisation of the ordinary (tree-based) μ-calculus formulas.

6.1 Parity formulas

We start with the basic definition of a parity formula. Recall that, given a set P of proposition letters, we define the sets $\mathtt{Lit}(\mathsf{P})$ and $\mathtt{At}(\mathsf{P})$ of *literals* and *atomic formulas* over P by setting $\mathtt{Lit}(\mathsf{P}) := \{p, \overline{p} \mid p \in \mathsf{P}\}$ and $\mathtt{At}(\mathsf{P}) := \mathtt{Lit}(\mathsf{P}) \cup \{\top, \bot\}$, respectively.

Definition 6.1 Let P be a finite set of proposition letters. A *parity formula over* P is a quintuple $\mathbb{G} = (V, E, L, \Omega, v_I)$, where

a) (V, E) is a finite, directed graph, with $|E[v]| \leq 2$ for every vertex v;[4]

b) $L : V \to \mathtt{At}(\mathsf{P}) \cup \{\wedge, \vee, \Diamond, \Box, \varepsilon\}$ is a labelling function;

[4] When discussing *disjunctive* parity formulas we will drop this requirement, allowing $E[v]$ to be of arbitrary finite size.

c) $\Omega : V \overset{\circ}{\to} \omega$ is a partial map, the *priority* map of $\mathbb{G}$; and

d) v_I is a vertex in V, referred to as the *initial* node of $\mathbb{G}$;

such that

1) $|E[v]| = 0$ if $L(v) \in \mathtt{At}(\mathsf{P})$, and $|E[v]| = 1$ if $L(v) \in \{\Diamond, \Box\} \cup \{\varepsilon\}$;

2) every cycle of (V, E) contains at least one node in $\mathsf{Dom}(\Omega)$.

A node $v \in V$ is called *silent* if $L(v) = \varepsilon$, *constant* if $L(v) \in \{\top, \bot\}$, *literal* if $L(v) \in \mathtt{Lit}(\mathsf{P})$, *atomic* if it is either constant or literal, *boolean* if $L(v) \in \{\wedge, \vee\}$, and *modal* if $L(v) \in \{\Diamond, \Box\}$. We say that a proposition letter q *occurs* in $\mathbb{G}$ if $L(v) \in \{q, \overline{q}\}$ for some $v \in V$.

Elements of $\mathsf{Dom}(\Omega)$ will be called *states*, and we refer to (V, E) as the *(underlying) graph* of the parity formula. A parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ is *balanced* if its states coincide with its silent nodes, that is, if $\mathsf{Dom}(\Omega) = \{v \in V \mid L(v) = \varepsilon\}$. ◁

Remark 6.2 Parity formulas share many characteristics of *automata*, and indeed they also go under the name of *alternating tree automata*. We will say a bit more about this at the end of this section. Since we decided to use the term 'formulas' to describe the objects in Definition 6.1, it will be useful to have an adjective to describe the 'regular' formulas, that is, the elements of the language $\mu\mathtt{ML}$. We will refer to these as 'ordinary' formulas. ◁

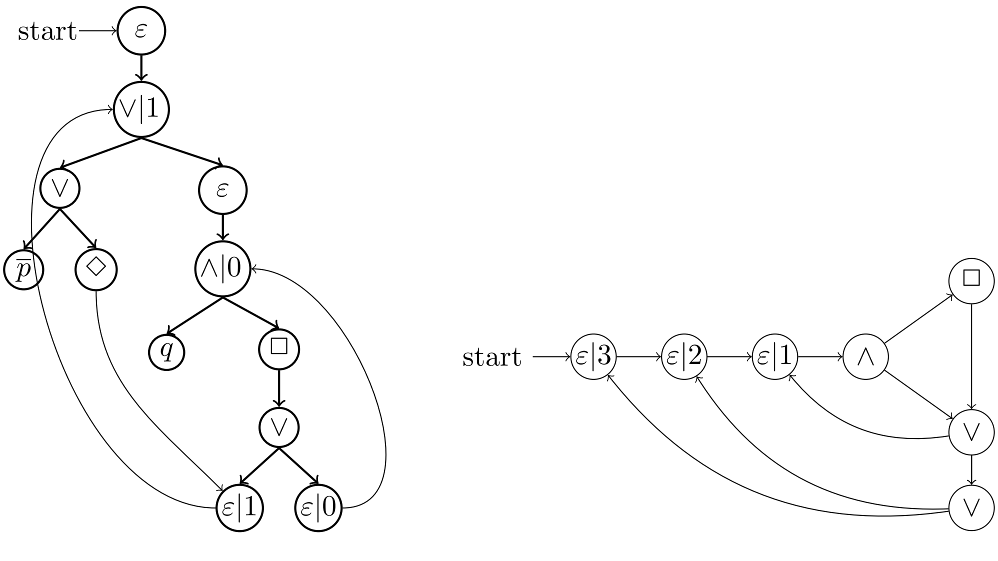


Figure 8: Two parity formulas

Example 6.3 In Figure 8 we give two examples of parity formulas. The picture on the left displays a parity formula that is directly based on the μ-calculus formula $\xi = \mu x.(\overline{p} \vee \Diamond x) \vee \nu y.(q \wedge \Box(x \vee y))$, by adding *back edges* to the subformula dag of ξ. The picture on the right displays a parity formula that is based on the closure graph of the formula ξ_1 of Example 2.38. In both cases we display the label and (if defined) the priority of a node inside its representing circle. Note that the parity formula to the right is balanced, but the one to the left is not. ◁

Position	Player	Admissible moves
(v, s) with $L(v) = p$ and $s \in V(p)$	$\forall$	$\varnothing$
(v, s) with $L(v) = p$ and $s \notin V(p)$	$\exists$	$\varnothing$
(v, s) with $L(v) = \overline{p}$ and $s \in V(p)$	$\exists$	$\varnothing$
(v, s) with $L(v) = \overline{p}$ and $s \notin V(p)$	$\forall$	$\varnothing$
(v, s) with $L(v) = \bot$	$\exists$	$\varnothing$
(v, s) with $L(v) = \top$	$\forall$	$\varnothing$
(v, s) with $L(v) = \varepsilon$	-	$E[v] \times \{s\}$
(v, s) with $L(v) = \vee$	$\exists$	$E[v] \times \{s\}$
(v, s) with $L(v) = \wedge$	$\forall$	$E[v] \times \{s\}$
(v, s) with $L(v) = \Diamond$	$\exists$	$E[v] \times R[s]$
(v, s) with $L(v) = \Box$	$\forall$	$E[v] \times R[s]$

Table 8: The evaluation game $\mathcal{E}(\mathbb{G}, \mathbb{S})$

The definition of parity formulas needs little explanation. Condition 2) says that every cycle must pass through at least one state; this is needed to provide a winner for infinite matches of the evaluation games that we use to define the semantics of parity formulas. The rules (admissible moves) in this evaluation game are completely obvious.

Definition 6.4 Let $\mathbb{S} = (S, R, V)$ be a Kripke model for a set P of proposition letters, and let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be a parity P-formula. The *evaluation game* $\mathcal{E}(\mathbb{G}, \mathbb{S})$ is the parity game (G, E', Ω') of which the board consists of the set $V \times S$, the priority map $\Omega' : V \times S \to \omega$ is given by

$$\Omega'(v, s) := \begin{cases} \Omega(v) & \text{if } v \in \mathsf{Dom}(\Omega) \\ 0 & \text{otherwise,} \end{cases}$$

and the game graph is given in Table 8. Note that we do not need to assign a player to positions that admit a single move only. $\lhd$

Definition 6.5 We say that a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ *holds at* or *is satisfied by* a pointed Kripke model $(\mathbb{S}, s)$, notation: $\mathbb{S}, s \Vdash \mathbb{G}$, if the pair (v_I, s) is a winning position for $\exists$ in $\mathcal{E}(\mathbb{G}, \mathbb{S})$. Analogously to ordinary formulas, we define the *meaning* of a parity formula $\mathbb{G}$ in a Kripke model $\mathbb{S}$ as follows:

$$[\![\mathbb{G}]\!]^{\mathbb{S}} := \{s \in S \mid \mathbb{S}, s \Vdash \mathbb{G}\},$$

and, for any propositional variable x, the map $\mathbb{G}^{\mathbb{S}}_x : \wp(S) \to \wp(S)$ by setting

$$\mathbb{G}^{\mathbb{S}}_x(A) := [\![\mathbb{G}]\!]^{\mathbb{S}[x \mapsto A]}.$$

We call two parity formulas $\mathbb{G}$ and $\mathbb{G}'$ *equivalent* if they have the same meaning in any model, notation: $\mathbb{G} \equiv \mathbb{G}'$. We will use the same terminology and notation to compare parity formulas with standard formulas. $\lhd$

The two key complexity measures of a parity formula, viz., size and index, both have perspicuous definitions. We will introduce these measures here, together some other useful notions pertaining to parity formulas.

Definition 6.6 The *size* of a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ is defined as its number of nodes: $|\mathbb{G}| := |V|$. ◁

Next to size, as the second fundamental complexity measure for a parity formula we need is its *index*, which corresponds to the alternation depth of ordinary formulas. It concerns the degree of alternation between odd and even positions in an infinite match of the evaluation game, and it is thus closely related to the range of the priority map of the formula. The most straightforward approach would be to define the index of a parity formula as the size of this range; a slightly more sophisticated approach is a *clusterwise* version of this.

Definition 6.7 Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be a parity formula, and let u and v be vertices in V. We say that v is *active* in u if E^+uv, and we let $\bowtie_E \subseteq V \times V$ hold between u and v is u is active in v and vice versa, i.e., $\bowtie_E := E^+ \cap (E^{-1})^+$. We let $\equiv_E$ be the equivalence relation generated by $\bowtie_E$; the equivalence classes of $\equiv_E$ will be called *clusters*. A cluster C is called *transient* if it is a singleton $\{v\}$ such that v is not active in itself, and *proper* otherwise.

The collection of clusters of a parity formula $\mathbb{G}$ is denoted as $Clus(\mathbb{G})$, and we say that a cluster C is *higher* than another cluster C' if there are some $u \in C$ and $u \in C'$ with E^+uu'. ◁

Note that a cluster is transient iff there is a nontrivial path between any pair of vertices. Intuitively, vertices belong to the same (proper) cluster if they can jointly occur infinitely often in some infinite match of some acceptance game for the formula. Furthermore, note that the notion of one cluster being higher than another could have been defined in many different (but equivalent) ways. For instance, C is higher than C' iff for every $u \in C$ there is a $u' \in C'$ with E^+uu', iff E^+uu' holds for every $u \in C$ and $u' \in C$, etc. Finally, observe that the 'higher than' relation between clusters is a partial order.

Proposition 6.8 *Let $\tau = (t_n)_{n\in\omega}$ be an infinite path through the graph of a parity formula $\mathbb{G}$. Then $\mathbb{G}$ has a unique cluster C such that, for some k, all t_n with $n > k$ belong to C. This cluster is proper.*

As a corollary of this, the relative priorities of states only matter if we stay in the same cluster.

Definition 6.9 Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be a parity formula. Given a cluster C of $\mathbb{G}$, we define its *index* as the number of priorities reached by states in C, that is, $ind_{\mathbb{G}}(C) = |\mathsf{Ran}(\Omega\restriction_C)|$. The *index of* $\mathbb{G}$ is given as the maximal index of its proper clusters, $ind(\mathbb{G}) := \max\{ind_{\mathbb{G}}(C) \mid C \in Clus(\mathbb{G}) \text{ proper}\}$. ◁

Note that a proper cluster must have at least one state. Consequently, a parity formula has a nonzero index iff it has proper clusters.

The notion of balance comes into the picture when we consider the operations of transforming ordinary formulas into parity formulas and vice versa. The following proposition, the proof of which is left to the reader, will then become important.

Proposition 6.10 *Let* $\mathbb{G}$ *be a parity formula. Then there is an equivalent balanced parity formula* $\mathbb{G}'$ *of size* $|\mathbb{G}'| \leq 2 \cdot |\mathbb{G}|$ *and index* $ind(\mathbb{G}') = ind(\mathbb{G})$.

Alternative presentation: alternating tree automata and equation systems

As mentioned earlier, in the literature one may find our parity formulas (or very similar structures) under the names of *alternating tree automata* or *(hierarchical or modal) equation systems*. We discuss these alternative perspectives here.

An *alternating tree automaton* or ATA over a set P of proposition letters is a quadruple $\mathbb{A} = (A, \Theta, \Omega, a_I)$, where A is a finite set of objects called *states*, $a_I \in A$ is an element of A called the *initial state*, $\Omega : A \to \omega$ is a (total) priority function, and $\Theta : A \to \mathtt{TC}(\mathsf{P}, A)$ is the *transition map* of the automaton. Here $\mathtt{TC}(\mathsf{P}, A)$ is a logical language of usually rather simple modal formulas that are called *transition conditions* over (P and) A. The exact definition of this language comes with some variations; here we define the set $\mathtt{TC}(\mathsf{P}, A)$ via the following grammar:

$$\alpha ::= \bot \mid \top \mid p \mid \overline{p} \mid a \mid \Diamond a \mid \Box a \mid a \wedge b \mid a \vee b, \tag{60}$$

where $p \in \mathsf{P}$ and $a, b \in A$.

The (operational) semantics of alternating tree automata is formulated in terms of an acceptance game $\mathcal{A}(\mathbb{A}, \mathbb{S})$, for an ATA $\mathbb{A}$ and a Kripke model $\mathbb{S} = (S, R, V)$. The positions of this game are given as the set

$$\{(\alpha, s) \in A \times S \mid a \in A \text{ or } \alpha = \Theta(a), \text{ for some } a \in A\},$$

while its rules and winning conditions are listed in Table 9.

Position	Player	Admissible moves
(a, s) with $\Theta(a) = \bot$	$\exists$	$\varnothing$
(a, s) with $\Theta(a) = \top$	$\forall$	$\varnothing$
(a, s) with $\Theta(a) = p$ and $s \in V(p)$	$\forall$	$\varnothing$
(a, s) with $\Theta(a) = p$ and $s \notin V(p)$	$\exists$	$\varnothing$
(a, s) with $\Theta(a) = \neg p$ and $s \in V(p)$	$\exists$	$\varnothing$
(a, s) with $\Theta(a) = \neg p$ and $s \notin V(p)$	$\forall$	$\varnothing$
(a, s) with $\Theta(a) = b$	-	$\{(b, s)\}$
(a, s) with $\Theta(a) = (b_0 \vee b_1)$	$\exists$	$\{(b_0, s), (b_1, s)\}$
(a, s) with $\Theta(a) = (b_0 \wedge b_1)$	$\forall$	$\{(b_0, s), (b_1, s)\}$
(a, s) with $\Theta(a) = \Diamond b$	$\exists$	$\{(a, t) \mid sRt\}$
(a, s) with $\Theta(a) = \Box b$	$\forall$	$\{(a, t) \mid sRt\}$

Table 9: The acceptance game $\mathcal{A}(\mathbb{A}, \mathbb{S})$

From this definition it is fairly obvious that we may identify parity formulas and ATAs: with an alternating tree automaton $\mathbb{A}$ we may associate the parity formula $\mathbb{G}_{\mathbb{A}} = (A, E, L, \Omega, a_I)$ given by

$$E(a) := \begin{cases} \varnothing \\ \{b\} \\ \{b\} \\ \{a_0, a_1\} \end{cases} \quad \text{and} \quad L(a) := \begin{cases} \Theta(a) & \text{if } \Theta(a) \in \mathtt{At}(\mathsf{P}) \\ \varepsilon & \text{if } \Theta(a) = b, \text{ for some } b \in A \\ \heartsuit & \text{if } \Theta(a) = \heartsuit b, \text{ for some } b \in A \\ \odot & \text{if } \Theta(a) = a_0 \odot a_1, \text{ for some } a_0, a_1 \in A \end{cases}$$

Conversely, it is equally simple to turn a parity formula into an alternating tree automaton which is based on the set of vertices of the formula. Note that if we simply take the size of an automaton to be its number of states, the constructions that we just outlined respect size in either direction.

Another perspective on alternating tree automata is that of *hierarchical equation systems* or HESs. We need not go into detail here, but to get the basic idea, fix an ATA $\mathbb{A} = (A, \Theta, \Omega, a_I)$. Think of the states in A as *variables*, then with each state $a \in A$ we may associate an *equation* of the form $a = \Theta(a)$, and make this equation inherit the priority $\Omega(a)$. Then we may organise this collection of equations by grouping equations of the same priority together in so-called *blocks*, and order these blocks according to their priorities. The result of this operation is known as a hierarchical equation system. It should be clear from this description that there is a 1-1 correspondence between alternating tree automata and these hierarchical equations systems.

Finally, the definition of HESs and ATAs can be modified/generalised in two directions. First of all, one may vary the set of admissible transition conditions. For instance, a so-called *modal equation system* or MES corresponds to an alternating tree automaton $\mathbb{A} = (A, \Theta, \Omega, a_I)$ where the transition map Θ maps every state $a \in A$ to an arbitrary modal formula as given by the following grammar

$$\alpha, \beta ::= \bot \mid \top \mid p \mid \overline{p} \mid a \mid \Diamond\alpha \mid \Box\alpha \mid \alpha \wedge \beta \mid \alpha \vee \beta, \tag{61}$$

where $p \in \mathsf{P}$ and $a \in A$. The acceptance game associated with these automata are obtained via a minor modification of the one given in Table 9. Note, however, that in this case one should not simply take the size of the ATA to be its number of states; one has to take the size of the formulas in the range of the transition map into account as well.

Second, one may change the nature of the transition map so that it shifts the propositions from its output (transition conditions) to the input. That is, given a finite set P of proposition letters, think of its power set $\wp\mathsf{P}$ as an alphabet of *colours*. The transition map of an ATA can then be given as a map $\Theta : A \times \wp\mathsf{P} \to \mathtt{TC}(A)$ mapping pairs of states and colours to transition conditions that now may only involve the states of the automata as propositional variables. The resulting *modal* automata will be discussed in detail in Chapter 10.

Exercises

Exercise 6.1 Provide an effective procedure transforming an arbitrary parity formula $\mathbb{G}$ into an equivalent balanced parity formula $\mathbb{G}^b$ such that $|\mathbb{G}^b| \leq 2 \cdot |\mathbb{G}|$ and $ind(\mathbb{G}^b) = ind(\mathbb{G})$.

6.2 Basics

Model checking for parity formulas

Since the evaluation game for parity formulas is given as a parity game, we immediately get a quasi-polynomial upper bound on the time complexity of the *model checking* problem for parity formulas. Recall that the size of a (pointed) labelled transition system is simply defined as the number of points in the model.

Definition 6.11 The *model checking problem* for parity formulas is the problem to compute whether $\mathbb{S}, s \Vdash \mathbb{G}$, where $\mathbb{S}$ is a (finite) labelled transition system, and $\mathbb{G}$ is a parity formula. $\lhd$

Theorem 6.12 *Assume that the problem of determining the winning regions of a parity game $\mathbb{G}$ can be solved in time $f(n, d)$, where n and d are, respectively, the size and the index of $\mathbb{G}$. Then the model checking problem for parity formulas can be solved in time $f(m \cdot n, d + 1)$, where m is the size of the labelled transition system, and n and d are the size and index of the parity formula, respectively.*

Remark 6.13 In the above theorem we state the model checking problem as being solvable in time $f(m \cdot n, d + 1)$, and not $f(m \cdot n, d)$. The reason for adding the '+1' is a technicality: the index of a parity game is usually defined as the *global* range of the priority map, whereas for parity formulas we define the index as the *clusterwise* size of the priority map. This makes a difference of 1 in the case of a parity formula of index d which has a cluster C with $\mathsf{Ran}(\Omega\restriction_C) = \{0, \dots, d-1\}$ and another cluster C' with $\mathsf{Ran}(\Omega\restriction_{C'}) = \{1, \dots, d\}$. $\lhd$

In the next section we will see how we can use this result to analyse the computational complexity of the model checking problem for ordinary formulas.

Operations on parity formulas

Parity formulas are interesting logical objects in their own right, and so one might want to develop their theory. To start with, it is fairly easy to define various operations on parity formulas. The following proposition covers the boolean operations (including negation), the modalities, and substitution.

Proposition 6.14 *Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ and $\mathbb{G} = (V', E', L', \Omega', v'_I)$ be two parity formulas.*

1) *There is a parity formula $\mathbb{H}$ such that $|\mathbb{H}| \leq 1+|\mathbb{G}|+|\mathbb{G}'|$ and $ind(\mathbb{H}) \leq \max(ind(\mathbb{G}), ind(\mathbb{G}'))$, while for any pointed Kripke model $(\mathbb{S}, s)$ we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}, s \Vdash \mathbb{G} \textit{ and } \mathbb{S}, s \Vdash \mathbb{G}'.$$

2) *There is a parity formula $\mathbb{H}$ such that $|\mathbb{H}| \leq 1+|\mathbb{G}|+|\mathbb{G}'|$ and $ind(\mathbb{H}) \leq \max(ind(\mathbb{G}), ind(\mathbb{G}'))$, while for any pointed Kripke model $(\mathbb{S}, s)$ we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}, s \Vdash \mathbb{G} \textit{ or } \mathbb{S}, s \Vdash \mathbb{G}.$$

3) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| = |\mathbb{G}|$ *and* $ind(\mathbb{H}) \leq ind(\mathbb{G})$*, while for any pointed Kripke model* $(\mathbb{S}, s)$ *we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}, s \nVdash \mathbb{G}.$$

4) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| \leq 1 + |\mathbb{G}|$ *and* $ind(\mathbb{H}) \leq ind(\mathbb{G})$*, while for any pointed Kripke model* $(\mathbb{S}, s)$ *we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}, t \Vdash \mathbb{G} \textit{ for some } t \in R[s].$$

5) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| \leq 1 + |\mathbb{G}|$ *and* $ind(\mathbb{H}) \leq ind(\mathbb{G})$*, while for any pointed Kripke model* $(\mathbb{S}, s)$ *we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}, t \Vdash \mathbb{G} \textit{ for all } t \in R[s].$$

6) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| \leq |\mathbb{G}| + |\mathbb{G}'|$ *and* $ind(\mathbb{H}) \leq \max(ind(\mathbb{G}), ind(\mathbb{G}'))$*, while for any pointed Kripke model* $(\mathbb{S}, s)$ *we have*

$$\mathbb{S}, s \Vdash \mathbb{H} \textit{ iff } \mathbb{S}[x \mapsto [\![\mathbb{G}']\!]^{\mathbb{S}}], s \Vdash \mathbb{G}.$$

Proof. We only cover the case of conjunctions, leaving the other cases as exercises to the reader. Given the parity formulas $\mathbb{G}$ and $\mathbb{G}$, we let $\mathbb{G} \wedge \mathbb{G}' := (V'', E'', L'', \Omega'', v''_I)$ be the structure given as follows. We let V'' be the disjoint union of V, V' and some singleton $\{v''\}$, and we define $E'' := E \cup E' \cup \{(v'', v_I), (v'', v'_I)\}$. Furthermore we put $L''(v'') := \wedge$, while we let L'' agree with L on V and with L' on V'. Similarly, we leave Ω'' undefined on v'' and let Ω'' agree with Ω on V and with Ω' on V'. Finally, we define $v''_I := v''$. It is then completely straightforward to show that $\mathbb{G} \wedge \mathbb{G}'$ is a parity formula indeed, and that it is equivalent to the conjunction of $\mathbb{G}$ and $\mathbb{G}'$. QED

The proposition below covers the least and greatest fixpoint operation. We leave its proof as an exercise for the reader.

Proposition 6.15 *Let* $\mathbb{G} = (V, E, L, \Omega, v_I)$ *be a parity formula, and let* x *be a proposition letter that only occurs positively in* $\mathbb{G}$*. Then the following hold:*

1) *The map* $\mathbb{G}^{\mathbb{S}}_x : \wp(S) \to \wp(S)$ *is monotone, for any Kripke model based on* S*.*
2) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| \leq |\mathbb{G}|$ *and* $ind(\mathbb{H}) \leq ind(\mathbb{G})$*, while for any Kripke model* $\mathbb{S}$ *we have*

$$[\![\mathbb{H}]\!]^{\mathbb{S}} = \mathrm{LFP}.\mathbb{G}^{\mathbb{S}}_x.$$

3) *There is a parity formula* $\mathbb{H}$ *such that* $|\mathbb{H}| \leq |\mathbb{G}|$ *and* $ind(\mathbb{H}) \leq ind(\mathbb{G})$*, while for any Kripke model* $\mathbb{S}$ *we have*

$$[\![\mathbb{H}]\!]^{\mathbb{S}} = \mathrm{GFP}.\mathbb{G}^{\mathbb{S}}_x.$$

Morphisms and equivalence notions between parity formulas

Furthermore, we will need various *structural* notions of equivalence between parity formulas. A simple but very useful concept is that of two parity formulas being *parity variants*.

Definition 6.16 A *parity variant* of a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ is a parity formula $\mathbb{G} = (V, E, L, \Omega', v_I)$ such that (i) $\Omega(v) \equiv_2 \Omega'(v)$, for all v, and (ii) $\Omega(u) < \Omega(v)$ iff $\Omega'(u) < \Omega'(v)$, for all u and v that belong to the same cluster but have different parity. $\lhd$

It is easy to see that parity variants are semantically equivalent, and have the same size (but not necessarily the same index). From this it follows that there are certain normal forms for parity formulas. Recall that we define $[k, n] := \{i \in \omega \mid k \leq i \leq n\}$.

Definition 6.17 A parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ is called *lean* if Ω is injective, and *tight* if for any cluster C, the range of Ω on C is connected, that is, of the form $\mathsf{Ran}(\Omega{\restriction_C}) = [k, n]$ for some natural numbers k, n with $k \leq n$. $\lhd$

It is not hard to see that every parity formula can be effectively transformed into either a lean or a tight parity variant.

Turning to the notion of a *morphism* between parity formulas, we have various options. Given the coalgebraic flavour of parity formula, it should come as no surprise that the definition we take integrates elements of the notion of a bounded morphism between Kripke models. The definition also incorporates the notion of a parity variant of Definition 6.16.

Definition 6.18 Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ and $\mathbb{G}' = (V', E', L', \Omega', v'_I)$ be two parity formulas. A *morphism* from $\mathbb{G}$ to $\mathbb{G}'$ is a map $f : V \to V'$ satisfying the following conditions, for all $u, v \in V$:

1) $L(u) = L'(f(u))$
2) if Euv then $E'f(u)f(v)$
3) if $E'f(u)v'$ then Euv for some v with $f(v) = v'$
4) (a) $\Omega(v) \equiv_2 \Omega'(fv)$, for all v, and (b) $\Omega(u) < \Omega(v)$ iff $\Omega'(fu) < \Omega'(fv)$, for all u and v that belong to the same cluster but have different parity
5) $f(v_I) = v'_I$.

We write $f : \mathbb{G} \twoheadrightarrow \mathbb{G}'$ to denote that f is a surjective morphism from $\mathbb{G}$ to $\mathbb{G}'$. $\lhd$

As an example, here is the parity formula version of a subformula.

Definition 6.19 Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be a parity formula, and let v be a vertex in V. We let $\mathbb{G}\langle v\rangle := (V, E, L, \Omega, v)$ denote the variant of $\mathbb{G}$ that takes v as its initial node; we define V_v to be the smallest subset of V which contains v and is closed under taking successors, and we call $\mathbb{G}_v := (V_v, E{\restriction_{V_v}}, L{\restriction_{V_v}}, \Omega{\restriction_{V_v}}, v)$ the *subformula* of $\mathbb{G}$ that is *generated* from v. $\lhd$

The following proposition will be needed further on; we omit its proof, since it is straightforward.

Proposition 6.20 *Let $f : \mathbb{G} \to \mathbb{G}'$ be a morphism of parity formulas. Then for any node v in $\mathbb{G}$ it holds that*

$$\mathbb{G}\langle v\rangle \equiv \mathbb{G}'\langle f(v)\rangle.$$

In particular, for every node $v \in V$ we have

$$\mathbb{G}\langle v\rangle \equiv \mathbb{G}_v.$$

In the same way that one may generalize bounded morphisms to bisimulations, we can generalize parity formula morphisms to parity formula bisimulations.

Definition 6.21 Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ and $\mathbb{G}' = (V', E', L', \Omega', v'_I)$ be two parity formulas. A *parity formula bisimulation* between $\mathbb{G}$ and $\mathbb{G}'$ is a relation $Z \subseteq V \times V'$ satisfying the following conditions, for all $u, v \in V$ and $u', v' \in V'$ such that $(u, u'), (v, v') \in Z$:

1) $L(u) = L'(f(u))$
2) if Euv then there is some $w' \in V'$ such that $(v, w') \in Z$ and $Eu'w'$;
3) if $E'u'v'$ then there is some $w \in V$ such that $(w, v') \in Z$ and Euw;
4) (a) $\Omega(u) \equiv_2 \Omega'(u')$ and (b) $\Omega(u) < \Omega(v)$ iff $\Omega'(u') < \Omega'(v')$, if u and v in $\mathbb{G}$, and u' and v' in $\mathbb{G}'$ belong to the same cluster but have different parity;
5) $(v_I, v'_I) \in Z$.

We write $Z : \mathbb{G} \sim \mathbb{G}'$ to denote that Z is a parity formula bisimulation between $\mathbb{G}$ and $\mathbb{G}'$. ◁

Proposition 6.22 *Let $Z : \mathbb{G} \sim \mathbb{G}'$ be a parity formula bisimulation. Then we have*

$$\mathbb{G}\langle v\rangle \equiv \mathbb{G}'\langle v'\rangle, \text{ whenever } (v, v') \in Z.$$

► `Somewhere: expansion map is a parity formula morphism;`

6.3 From ordinary formulas to parity formulas

In this section we will see how to represent a ordinary formula as an equivalent parity formula. In order to use the complexity result for the model checking problem for parity formulas to make similar observations on the model checking of ordinary formulas, we obviously want to minimimize the *size* and the *index* of the representation. It should come as no surprise that the index of this parity formula will somehow correspond to the alternation depth of the formula, while the size of the parity formula will clearly depend on the graph structure that we pick to represent the original formula.

Suppose that we are looking for a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ representing the ordinary formula ξ. It seems that there are three natural candidates for the underlying graph (V, E) of $\mathbb{G}$: we could take the syntax tree of ξ, its subformula graph, or its closure graph. Note that each of these three structures induces a natural *size measure* of μ-calculus formulas, respectively length, subformula-size, and (closure-)size. The tree representation has the advantage of being immediately available for *any* $\xi \in \mu\mathtt{ML}$, whereas the subformula and closure graph are defined only for the clean and the tidy formulas, respectively. Because of the unwieldy size of syntax trees, however, we focus here on the other two kinds of representations.

Definition 6.23 The *subformula dag* of a clean formula ξ is the graph $(Sf(\xi), \triangleright_0)$, where $\triangleright_0$ is the converse of the direct subformula relation $\triangleleft_0$. We obtain the *subformula graph* $(Sf(\xi), \triangleright_0 \cup B)$ from this by adding the set $B := \{(x, \delta_x) \mid x \in BV(\xi)\}$ of *back edges* to this dag. The *closure graph* of a tidy formula $\xi \in \mu\mathtt{ML}$ is the structure $\mathbb{C}_\xi = (Cl(\xi), \rightarrow_C)$, where $\rightarrow_C$ is the trace relation (restricted to the closure of ξ). $\triangleleft$

In the remainder of this section we will see how to expand either of these two graphs into a full parity formula structure. In both cases it is obvious how to define the labelling map L, and which vertex to take as the initial one. It is also more or less clear what the *states* should be, i.e., on which vertices the priority map Ω should be defined: In the case of the subformula graph of a clean formula ξ we will take the set $BV(\xi)$ of *bound variables* of ξ, while in the closure graph of a tidy formula ξ we will consider the set of all *fixpoint formulas* in the closure of ξ.

For the actual *values* of the priority maps, there is some choice. Observe, however, that in both versions of the evaluation game, the winning conditions are defined in terms of some *priority order* on the set of states, in combination with a fixed assignment of a parity/player to each state. For instance, in the case of the subformula game for a clean formula ξ, we used the unfolding order $\trianglelefteq_\xi$ on $BV(\xi)$, together with the partition of $BV(\xi)$ into μ- and ν-variables. To assign a winner to an infinite match π, we consider the $\trianglelefteq_\xi$-maximal element of the set $Unf^\infty(\pi)$ (consisting of those bound variables that are unfolded infinitely often during π), and the winner of π is then determined by the nature of this variable. Similarly, in the case of the closure game for ξ, we looked at the fixpoint formulas that occur infinitely often in an infinite match of the game, and observed that this set has a smallest element with respect to the (free) subformula ordering.

It therefore makes sense to discuss how to transform a priority order on the set of states into a suitable priority map, *in some generality.* Playing with the shape of the priority order may have an effect on the *index* of the associated priority map, and one priority order may yield a lower index than another. Interestingly, both for the subformula game and for the closure game, we will see that the order used in Chapter 2 does *not* give an optimal index.

6.3.1 From parity posets to priority maps

In this subsection we will see how to represent a priority order on some (partitioned) set of states by a suitable priority map. We first develop some terminology. Throughout this subsection the reader should think of Z as the set of states in (some cluster of) a parity formula that is either based on the subformula graph or on the closure graph of a formula ξ.

Definition 6.24 A *parity poset* is a structure $\mathbb{Z} = (Z, \preccurlyeq, p)$, where $(Z, \preccurlyeq)$ is a finite poset and p is a *parity map* on P, that is, a map $p : Z \to \{0, 1\}$. We write $z \prec z'$ if $z \preccurlyeq z'$ and $z \neq z'$. ◁

Example 6.25 Where ξ is a clean formula, take $Z := BV(\xi)$, and let p be the function mapping μ-variables to 1 and ν-variables to 0. The most obvious priority order on $BV(\xi)$ is the relation $\trianglelefteqslant_\xi$ given by

$$x \trianglelefteqslant_\xi y \text{ if } \delta_x \trianglelefteqslant \delta_y.$$

For a tidy formula ξ we may take Z to be the set of fixpoint formulas belonging to $Cl(\xi)$, and let p be the function mapping μ-formulas to 1 and ν-formulas to 0. The most obvious priority order on this set is the relation $\trianglerighteqslant_f$, that is, the converse of the free subformula relation. ◁

To give a precise formulation of the required connection between parity posets and priority maps, we need the notion of an *alternating chain* in a parity poset, which is a good measure of its complexity.

Definition 6.26 Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ be a parity poset. An *alternating chain* in $\mathbb{Z}$ of *length* k in $\mathbb{Z}$ is a finite sequence $z_1 \cdots z_k$ of states such that, for all $i < k$, $z_i \prec z_{i+1}$ while z_i and z_{i+1} have different parity. The *alternating chain depth* $\mathsf{acd}(\mathbb{Z})$ of Z is the maximal length of an alternating chain in $\mathbb{Z}$. ◁

Since the parity map on a set of states is usually fixed, with a slight abuse of notation we will often write $\mathsf{acd}(\preccurlyeq)$ rather than $\mathsf{acd}(Z, \preccurlyeq, p)$; this notation is particularly useful if we compare distinct priority relations on a fixed set Z (with a fixed parity map p).

Definition 6.27 Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ be a parity poset, and let $\Omega : Z \to \omega$ be some priority map. Then we say that Ω *represents* $\mathbb{Z}$ if it satisfies the following conditions, for all $x, z \in Z$:
1) $p(z) = \Omega(z) \mod 2$;
2) $x \preccurlyeq z$ implies $\Omega(x) \leq \Omega(z)$. ◁

The conditions 1 and 2 constitute some kind of soundness condition: if Ω represents $\mathbb{Z}$, then it yields the same winning conditions as $\mathbb{Z}$. This is the content of the next Proposition, the proof of which is trivial.

Proposition 6.28 *Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ and $\Omega : Z \to \omega$ be, respectively, a parity poset and a priority map representing it. Furthermore, let $\zeta = (z_n)_{n<\omega}$ be an infinite sequence of elements in Z, and assume that the set $Inf(\zeta)$ has a $\preccurlyeq$-greatest element z. Then $p(z) = 0$ iff $\max\big(\Omega[Inf(\zeta)]\big)$ is even.*

It is important to keep the range $\mathsf{Ran}(\Omega)$ of priorities small, given its contribution to the complexity of the model checking of parity formulas. It is not hard to see that the elements of an alternating chain all need to have a different priority, which means that $\mathsf{acd}(\mathbb{Z})$ is a lower bound for $|\mathsf{Ran}(\Omega)|$. The example below shows that in general we cannot even require that $|\mathsf{Ran}(\Omega)| = \mathsf{acd}(\mathbb{Z})$:

Where $\mathbb{Z}$ is the parity preorder in the picture, it is impossible to represent $\mathbb{Z}$ by a priority map Ω with $|\mathsf{Ran}(\Omega)| = 2$. However, if we partition $\mathbb{Z}$ into components and consider the alternating chain depth piecewise, a precise match is possible. For instance, in the case where Z is the closure set of some formula ξ, one may take the *clusterwise* converse subformula ordering; that is, only rank ψ higher than φ if (ψ is a subformula of φ and) there are traces from φ to ψ and vice versa.

Definition 6.29 Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ be a parity poset. A partition $(Z_i)_{i\in I}$ is called *disconnected* if there are no elements z_i, z_j in different components such that $z_i \preccurlyeq z_j$. In this setting we let $\mathbb{Z}_i$ denote the substructure of $\mathbb{Z}$ based on the set Z_i.

We call a map $\Omega : Z \to \omega$ representing $\mathbb{Z}$ a *tight* representation if $\mathbb{Z}$ admits a disconnected partition $(Z_i)_{i\in I}$ such that $\mathsf{acd}(\mathbb{Z}_i) = |\mathsf{Ran}(\Omega{\restriction}_{Z_i})|$, for each component $\mathbb{Z}_i$. ◁

Example 6.30 Figure 9 below depicts a parity poset (the picture to the left) and two priority maps representing it; the priority map to the right is tight, the other one is not. ◁

The key result here is that, under a mild condition that is met in the relevant cases, every parity poset is indeed tightly represented by some priority map.

Definition 6.31 We say that $\mathbb{Z}$ is *weakly directed* if Z admits a disconnected partition consisting of *directed* subsets. ◁

It will be convenient for us to define a *canonical* priority map for this purpose.

Definition 6.32 Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ be a parity poset; partition $\mathbb{Z}$ into maximal connected components, and assume that every component of $\mathbb{Z}$ has finite alternation depth.

Given a point $z \in Z$ we define $h^{\uparrow}(z)$ (respectively, $h^{\downarrow}(z)$) as the maximal length of an alternating chain starting at z (ending at z, respectively). We define the following map $\Omega_{\mathbb{Z}} : Z \to \omega$; given $z \in Z$, let $\mathbb{Z}_z$ be the component to which z belongs, and define

$$\Omega_{\mathbb{Z}}(z) := \begin{cases} \mathsf{acd}(\mathbb{Z}_z) - h^{\uparrow}(z) & \text{if } \mathsf{acd}(\mathbb{Z}_z) - h^{\uparrow}(z) \equiv_2 p(z) \\ \mathsf{acd}(\mathbb{Z}_z) - h^{\uparrow}(v) + 1 & \text{if } \mathsf{acd}(\mathbb{Z}_z) - h^{\uparrow}(z) \not\equiv_2 p(z), \end{cases} \tag{62}$$

and we will call this map the *priority map induced by* $\mathbb{Z}$. ◁

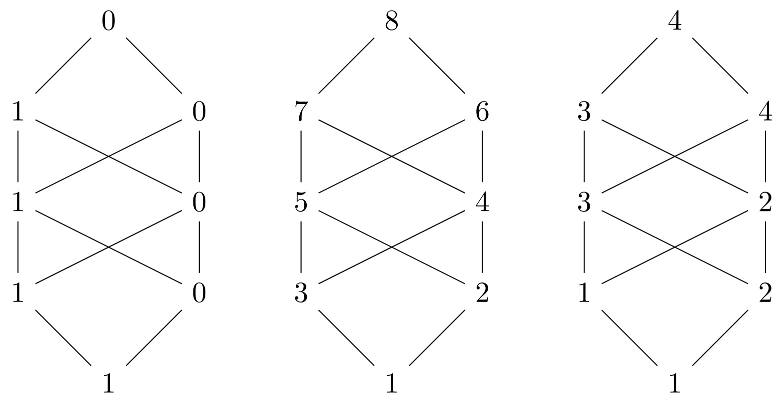


Figure 9: From a parity poset to a priority map

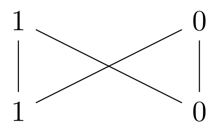
With a minor abuse of notation, we will often denote the priority map $\Omega_{\mathbb{Z}}$ as $\Omega_{\preccurlyeq}$.

Theorem 6.33 *Every weakly directed parity poset $\mathbb{Z} = (Z, \preccurlyeq, p)$ is tightly represented by its induced priority map $\Omega_{\mathbb{Z}}$.*

Proof. To avoid clutter we write Ω for $\Omega_{\mathbb{Z}}$. It is not very difficult to see that Ω represents $\mathbb{Z}$: the parity condition is straightforward, and for the order preservation it suffices to observe that $x \preccurlyeq y$ implies $h^{\uparrow}(x) \geq h^{\uparrow}(y)$, which in its turn implies $\Omega(x) \leq \Omega(y)$. For tightness, without loss of generally we may assume that $\mathbb{Z}$ is connected, which by (weak) directedness implies the existence of a $\preccurlyeq$-maximum m. Earlier on we already discussed that we always have $\mathsf{acd}(\mathbb{Z}) \leq |\mathsf{Ran}(\Omega)|$ — we need to prove the oppositve inequality.

To see that $|\mathsf{Ran}(\Omega)| \leq \mathsf{acd}(\mathbb{Z})$, take an arbitrary $z \in Z$; it is clear that $h^{\uparrow}(z) \in [1, \mathsf{acd}(\mathbb{Z})]$, so that we find $\mathsf{acd}(\mathbb{Z}) - h^{\uparrow}(z) \in [0, \mathsf{acd}(\mathbb{Z}) - 1]$. The second observation is then that, in the above definition of $\Omega(z)$, the elements of Z *uniformly* fall in one of the two mentioned cases. This means that we either have $\mathsf{Ran}(\Omega) \subseteq [0, \mathsf{acd}(\mathbb{Z}) - 1]$ or $\Omega[Z] \subseteq [1, \mathsf{acd}(\mathbb{Z})]$. In both cases this means that $|\mathsf{Ran}(\Omega)| \leq \mathsf{acd}(\mathbb{Z})$ as required. QED

Example 6.34 The condition of weak directedness in Theorem 6.33 is necessary, as the following simple example witnesses:



This parity poset clearly has alternation depth 2, but it cannot be represented by any priority map with range of size ≤ 2. $\triangleleft$

6.3.2 Parity formulas on the subformula graph

The following theorem shows that for a clean formula, we can indeed obtain an equivalent parity formula which is based on its *subformula graph*, which we defined as the subformula dag, augmented with *back edges*.

Theorem 6.35 *There is an algorithm that constructs, for a clean formula $\xi \in \mu\mathtt{ML}(\mathsf{P})$, an equivalent parity formula $\mathbb{H}_\xi$ over P, based on the subformula graph of ξ, which satisfies $|\mathbb{H}_\xi| = |\xi|^s$ and $ind(\mathbb{H}_\xi) = \mathsf{ad}(\xi)$.*

The basic idea underlying the proof of Theorem 6.35 is to view the evaluation games for clean formulas in $\mu\mathtt{ML}$ as instances of parity games. Given an arbitrary formula $\xi \in \mu\mathtt{ML}$, we then need to see which modifications are needed to turn the subformula dag $(Sf(\xi), \triangleright_0)$ into a parity formula $\mathbb{H}_\xi$ such that, for any model $\mathbb{S}$, the evaluation games $\mathcal{E}(\xi, \mathbb{S})$ and $\mathcal{E}(\mathbb{H}_\xi, \mathbb{S})$ are more or less identical. Clearly, the fact that the *positions* of the evaluation game $\mathcal{E}(\xi, \mathbb{S})$ are given as the pairs in the set $Sf(\xi) \times S$, means that we can take the set

$$V_\xi := Sf(\xi)$$

as the carrier of $\mathbb{H}_\xi$ indeed.

Looking at the admissible moves in the two games, it turns out that we cannot just take the converse direct subformula relation $\triangleright_0$ as the edge relation of $\mathbb{H}_\xi$: we need to add all *back edges* from the set

$$B_\xi := \{(x, \delta_x) \mid x \in BV(\xi)\},$$

where, as usual, we let δ_x denote the unique formula such that, for some $\eta \in \{\mu, \nu\}$ the formula $\eta x.\delta_x$ is a subformula of ξ. In fact, if we write D_ξ for the relation $\triangleright_0$, restricted to $Sf(\xi)$, then we can take

$$E_\xi := D_\xi \cup B_\xi,$$

as the edge relation of $\mathbb{H}_\xi$. Furthermore, the labelling map L_ξ is naturally defined via the following case distinction:

$$L_\xi(\varphi) := \begin{cases} \varphi & \text{if } \varphi \in \{\top, \bot\} \cup \{p, \overline{p} \mid p \in FV(\xi)\} \\ \odot & \text{if } \varphi \text{ is of the form } \varphi_0 \odot \varphi_1 \text{ with } \odot \in \{\wedge, \vee\} \\ \heartsuit & \text{if } \varphi \text{ is of the form } \heartsuit\psi \text{ with } \heartsuit \in \{\Diamond, \Box\} \\ \varepsilon & \text{if } \varphi \text{ is of the form } \eta_x x.\delta_x \text{ with } \eta \in \{\mu, \nu\} \\ \varepsilon & \text{if } \varphi \in BV(\xi). \end{cases}$$

With this definition, it is easy to see that the *boards* of the two evaluation games $\mathcal{E}(\xi, \mathbb{S})$ and $\mathcal{E}(\mathbb{H}_\xi, \mathbb{S})$ are *isomorphic* (in fact, identical), for any labeled transition system $\mathbb{S}$. As the initial node v_ξ of $\mathbb{H}_\xi$ we simply take

$$v_\xi := \xi.$$

In order to finish the definition of the parity formula $\mathbb{H}_\xi$ it is then left to come up with a suitable priority map Ω_ξ on V_ξ. For this we base ourselves on the discussion in the previous subsection, but rather than the unfolding order $\trianglelefteq_\xi$ we will consider the *depency order* $\preccurlyeq_\xi$

defined below. Intuitively, $x \preccurlyeq_\xi y$ indicates not only that δ_x is a subformula of δ_y, but also that y is 'more significant' than x, in the sense that the meaning of x/δ_x (in principle) depends on the meaning of y/δ_y. The key situation where this happens is when y occurs freely in δ_x.

Definition 6.36 Given a clean formula ξ, we define a *dependency* or *subordination order* $\preccurlyeq_\xi$ on the set $BV(\xi)$, saying that y *ranks higher* than x if $x \preccurlyeq_\xi y$. The relation $\preccurlyeq_\xi$ is defined as the least partial order containing all pairs (x, y) such that $y \trianglelefteq \delta_x \trianglelefteq \delta_y$. $\lhd$

For an example where the two orders give a different alternation depth, we refer to the formula ξ_3 in Example 2.52. The following Proposition shows that $\preccurlyeq_\xi$ is a suitable alternative to $\trianglelefteq_\xi$, in the sense that we will assign the same winner to any infinite match of the evaluation game, whether we base ourselves on $\preccurlyeq_\xi$ or on $\trianglelefteq_\xi$.

Proposition 6.37 *Let ξ be a clean $\mu\mathtt{ML}_\mathsf{D}$-formula, and let π be an infinite match of the evaluation game $\mathcal{E}(\xi, \mathbb{S})$, for some model $\mathbb{S}$. Then*

$$\max_{\preccurlyeq_\xi}(\mathit{Unf}^\infty(\pi)) = \max_{\trianglelefteq_\xi}(\mathit{Unf}^\infty(\pi)).$$

Proof. Abbreviate $U := \mathit{Unf}^\infty(\pi)$, and let $y \in U$ be the maximum of U in terms of $\trianglelefteq_\xi$ — such a y exists by Proposition 2.24.

We claim that

$$x \preccurlyeq_\xi y, \text{ for all } x \in U. \tag{63}$$

Suppose for contradiction that there is a variable $x \in U$ with $x \not\preccurlyeq_\xi y$. It follows from the assumption on y that $\delta_x \trianglelefteq \delta_y$, and without loss of generality we may assume x to be such that δ_x is a *maximal* subformula of δ_y such that $x \not\preccurlyeq_\xi y$ (in the sense that $z \preccurlyeq_\xi y$ for all $z \in U$ with $\delta_x \lhd \delta_z$). In particular then we have $y \notin FV(\delta_x)$. But since y is unfolded infinitely often, there must be a variable $z \in FV(\delta_x)$ which allows π to 'leave' δ_x infinitely often; this means that $z \in U$, $\delta_x \trianglelefteq \delta_z$ but $\delta_z \not\trianglelefteq \delta_x$. From this it is immediate that $x \preccurlyeq_\xi z$, while from $z \in U$ we obtain $\delta_z \trianglelefteq \delta_y$. It now follows from our maximality assumption on x that $z \preccurlyeq_\xi y$. But then by transitivity of $\preccurlyeq_\xi$ we find that $x \preccurlyeq_\xi y$. In other words, we have arrived at the desired contradiction.

This shows that (63) holds indeed, and from this the Proposition is immediate. QED

For our definition of $\mathbb{H}_\xi$ we will take the priority map Ω_ξ given by

$$\Omega_\xi := \Omega_{\preccurlyeq_\xi}.$$

Note that Ω_ξ, as a map on $Sf(\xi)$, is partial indeed: it is only defined for those subformulas of ξ that belong to its bound variables.

Summarizing, we define

$$\mathbb{H}_\xi := (V_\xi, E_\xi, L_\xi, \Omega_\xi, \xi).$$

Proof of Theorem 6.35. In the light of the above discussion, the equivalence of ξ and $\mathbb{H}_\xi$ follows from the Propositions 6.28 and 6.37 (and the easily verified fact that Ω_ξ satisfies the conditions 1 and 2 of Definition 6.27). It is immediate by the definitions that $|\mathbb{H}_\xi| = |Sf(\xi)| = |\xi|^s$. Finally, it follows from Theorem 6.33 that the index of $\mathbb{H}_\xi$ is given by the alternating chain depth of $\preccurlyeq_\xi$. In order to get an *exact* match of the index of $\mathbb{H}_\xi$ and the inductively defined alternation depth of ξ we need to work a bit harder: in Section 6.5 below we will prove Proposition 6.58 stating that $\mathsf{ad}(\xi) = \mathsf{acd}(\preccurlyeq_\xi)$. QED

6.3.3 Parity formulas on the closure graph

The next theorem states that for a tidy formula, we can find an equivalent parity formula that is based on the formula's *closure graph*, and has an index which is bounded by the alternation depth of the formula. This result applies to *arbitrary* tidy formulas, that is: no renaming of the formula's bound variables or other kinds of preprocessing the formula are needed.

Theorem 6.38 *There is a construction transforming an arbitrary tidy formula $\xi \in \mu\mathtt{ML}$ into an equivalent balanced parity formula $\mathbb{G}_\xi$ which is based on the closure graph of ξ, so that $|\mathbb{G}| = |\xi|$; in addition we have $ind(\mathbb{G}_\xi) \leq \mathsf{ad}(\xi)$.*

When it comes to complexity issues, this is in fact the main result that bridges the gap between the world of formulas and that of automata and parity games.

Theorem 6.39 *Assume that the problem of determining the winning regions of a parity game $\mathcal{G}$ can be solved in time $f(n, d)$, where n and d are, respectively, the size and the index of $\mathcal{G}$. Then the model checking problem for parity formulas can be solved in time $f(m \cdot n, d)$, where m is the size of the labelled transition system, and n and d are the size and alternation depth of the formula, respectively.*

In particular, as an immediate corollary of Theorem 6.38 and the quasi-polynomial time complexity result on the model checking problem for parity formulas (Theorem 6.12), we find that model checking for μ-calculus formulas can be solved in quasi-polynomial time.

▶ `COROLLARY: give concrete upper bound for complexity of model checking using Theorem 6.39 and Theorem 5.40.`

The priority map Ω_C that we will define on the closure graph of a tidy formula is in fact *global* in the sense that it can be defined uniformly for all (tidy) formulas, independently of any ambient formula. Similar to the case of the subformula game, the free subformula relation $\trianglelefteq_f$ used in Chapter 2 does *not* give an optimal result. We will rather base the map Ω_C on a partial order of fixpoint formulas, the *closure priority relation* $\preccurlyeq_C$, that we will introduce now.

Definition 6.40 We let $\equiv_C$ denote the equivalence relation generated by the trace relation $\rightarrow_C$, in the sense that: $\varphi \equiv_C \psi$ if $\varphi \twoheadrightarrow_C \psi$ and $\psi \twoheadrightarrow_C \varphi$. We will refer to the equivalence classes of $\equiv_C$ as *(closure) clusters*, and denote the cluster of a formula φ as $(\!|\varphi|\!)_C$.

Furthermore, we define the *closure priority relation* $\preccurlyeq_C$ on fixpoint formulas by putting $\varphi \preccurlyeq_C \psi$ precisely if $\psi \twoheadrightarrow_C^{\psi} \varphi$, and we write $\varphi \prec_C \psi$ if $\varphi \preccurlyeq_C \psi$ and $\psi \not\preccurlyeq_C \varphi$. ◁

In words: $\varphi \preccurlyeq_C \psi$ if there is a trace from ψ to φ such that ψ is a (free) subformula of every formula on the trace. This definition is somewhat involved, but this seems to be unavoidable if we want an optimal complexity of model checking.

Remark 6.41 The definition of the priority order $\preccurlyeq_C$ may look overly complicated. In fact, simpler definitions would suffice if we are only after the equivalence of a tidy formula with an associated parity formula that is based on its closure graph, i.e., if we do not need an exact match of index and alternating-chain depth.

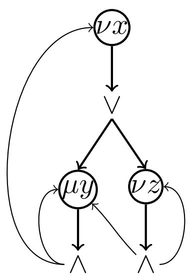
▶ `rewrite next paragraph`

In particular, we could have introduced an alternative priority order $\preccurlyeq'_C$ by putting $\varphi \preccurlyeq'_C \psi$ if $\varphi \equiv_C \psi$ and $\psi \trianglelefteq_f \varphi$. If we would base a priority map Ω'_C on this priority order instead of on $\preccurlyeq_C$, then we could prove the equivalence of any tidy formula ξ with the associated parity formula $\mathbb{G}'_\xi := (\mathbb{C}_\xi, \Omega'_C \restriction_{Cl(\xi)}, \xi)$. However, we would not be able to prove that the index of $\mathbb{G}'_\xi$ is bounded by the alternation depth of ξ.

To see this, consider the following formula:

$$\alpha_x := \nu x.\big((\mu y.x \wedge y) \vee \nu z.(z \wedge \mu y.x \wedge y)\big).$$

(For simplicity we use a formula without modalities to make the argument — nothing hinges on this.) We leave it for the reader to verify that α_x has alternation depth two, and that its closure graph looks as follows:



Let α_y and α_z be the other two fixpoint formulas in the cluster of α_x, that is, let $\alpha_y := \mu y.\alpha_x \wedge y$ and $\alpha_z := \nu z.z \wedge \alpha_y$. (These formulas correspond to the nodes in the graph that are labelled μy and νz, respectively.) Now observe that we have $\alpha_x \triangleleft_f \alpha_y \triangleleft_f \alpha_z$, so that this cluster has an alternating $\preccurlyeq'_C$-chain of length *three*: $\alpha_z \prec'_C \alpha_y \prec'_C \alpha_x$. Note however, that any trace from α_y to α_z must pass through α_x, the $\preccurlyeq_C$-maximal element of the cluster. In particular, we do *not* have $\alpha_z \preccurlyeq_C \alpha_y$, so that there is *no* $\preccurlyeq_C$-chain of length three in the cluster. $\triangleleft$

Here are some basic observations on the relation $\preccurlyeq_C$ and its connection with the closure equivalence relation $\equiv_C$.

Proposition 6.42 *1) The relation $\preccurlyeq_C$ is a partial order.*

2) The relation $\preccurlyeq_C$ is included in the closure equivalence relation: $\varphi \preccurlyeq_C \psi$ implies $\varphi \equiv_C \psi$.

3) The relation $\preccurlyeq_C$ is included in the converse free subformula relation: $\varphi \preccurlyeq_C \psi$ implies $\psi \trianglelefteq_f \varphi$.

4) Every cell of $\equiv_C$ contains a unique fixpoint formula $\xi = \eta x.\chi$ such that $\xi \notin Cl(\chi)$. This formula is the $\preccurlyeq_C$-maximum element of its cluster.

Proof. For item 1) we need to show that $\preccurlyeq_C$ is reflexive, transitive and antisymmetric. Reflexivity is obvious, and antisymmetry follows from 3). For transitivity assume that $\varphi \preccurlyeq_C \psi$ and $\psi \preccurlyeq_C \chi$ hold. By definition this means that $\psi \twoheadrightarrow^\psi_C \varphi$ and $\chi \twoheadrightarrow^\chi_C \psi$. The latter entails that $\chi \trianglelefteq_f \psi$ and the former means that there is some $\to_C$-trace from ψ to φ such that ψ is a free subformula of every formula along this trace. Because $\chi \trianglelefteq_f \psi$ and $\trianglelefteq_f$ is transitive it then also holds that χ is a free subformula of every formula on the trace from ψ to φ. Composing this trace with the one from χ to ψ we obtain a trace from χ to φ such that χ is a free subformula of all formulas along this trace. Hence $\chi \twoheadrightarrow_C^{\chi} \varphi$ and so $\varphi \preccurlyeq_C \chi$.

For item 2) we assume that $\varphi \preccurlyeq_C \psi$ and need to show that $\varphi \twoheadrightarrow_C \psi$ and $\psi \twoheadrightarrow_C \varphi$. The assumption $\varphi \preccurlyeq_C \psi$ means that $\psi \twoheadrightarrow_C^{\psi} \varphi$ which clearly entails $\psi \twoheadrightarrow_C \varphi$. But, as already observed above, $\psi \twoheadrightarrow_C^{\psi} \varphi$ also entails that $\psi \trianglelefteq_f \varphi$, from which $\varphi \twoheadrightarrow_C \psi$ follows by Proposition 2.58.

Item 3) is immediate by the definition of $\preccurlyeq_C$.

Finally, consider item 4). We first argue that

$$\text{every cluster of } \to_C \text{ contains a fixpoint formula } \xi = \eta x.\chi \notin Cl(\chi). \tag{64}$$

To this aim we claim that if $\xi = \eta x.\chi$ and $\xi \in Cl(\chi)$ then there is some fixpoint formula $\psi \in (\!|\xi|\!)_C$ with $\xi \prec_C \psi$. To see this, consider some $\xi = \eta x.\chi$ with $\xi \in Cl(\chi)$. Then we have a trace $\chi \twoheadrightarrow_C \xi$ to which we can apply Proposition 2.39 to obtain a fixpoint formula ψ with $\chi \twoheadrightarrow_C^{\psi} \psi$ and $\psi \twoheadrightarrow_C^{\psi} \xi$. From the latter we have that $\xi \preccurlyeq_C \psi$ and from the former we get $\psi \trianglelefteq_f \chi$ which entails $\psi \neq \xi$ because $\xi = \eta x.\chi$.

But if a cluster would only contain fixpoint formulas $\xi = \eta x.\chi$ with ξ *in* $Cl(\chi)$ then this would allow us to construct an infinite $\prec_C$-chain in the cluster, which is impossible as all clusters are finite. This proves (64).

Uniqueness of the fixpoint formula ξ in (64) is immediate by the following claim:

$$\text{if } \xi = \eta x.\chi \text{ and } \xi \notin Cl(\chi) \text{ then } \xi \trianglelefteq_f \rho, \text{ for all } \rho \in (\!|\xi|\!)_C. \tag{65}$$

For the proof of (65), take an arbitrary $\rho \in Cl(\xi)$. It follows by Proposition 2.44 5) that either $\rho = \xi$, in which case we are done, or $\rho = \gamma[\xi/x]$ for some $\gamma \in Cl(\chi)$. Now $x \in FV(\gamma)$ because otherwise $\chi \twoheadrightarrow_C \gamma = \gamma[\xi/x] = \rho \twoheadrightarrow_C \xi$ contradicting $\xi \notin Cl(\chi)$. But if $\rho = \gamma[\xi/x]$ and $x \in FV(\gamma)$ then by definition we have $\xi \trianglelefteq_f \rho$. This finishes the proof of (65).

For the $\preccurlyeq_C$-maximality of ξ consider an arbitrary fixpoint formula $\psi \in (\!|\xi|\!)_C$. Then we have $\xi \twoheadrightarrow_C \psi$, and since every formula ρ on this trace $\xi \twoheadrightarrow_C \psi$ is in $(\!|\xi|\!)_C$ it follows from (65) that $\xi \trianglelefteq_f \rho$ and thus $\xi \twoheadrightarrow_C^{\xi} \psi$. By definition this means $\psi \preccurlyeq_C \xi$. QED

Since $\preccurlyeq_C$ is a partial order, we may use Definition 6.27 to base a priority map on it. We are now ready for the definition of the parity formula $\mathbb{G}_\xi$ corresponding to a tidy formula ξ.

Definition 6.43 Fix some tidy formula ξ. We define $\mathbb{C}_\xi$ to be the closure graph $(Cl(\xi), \to_C)$ of ξ, expanded with the natural labelling L_C given by

$$L_C(\varphi) = \begin{cases} \varphi & \text{if } \varphi \in \mathtt{At}(\mathsf{P}) \\ \heartsuit & \text{if } \varphi = \heartsuit\psi, \quad \text{with } \heartsuit \in \{\Diamond, \Box\} \\ \odot & \text{if } \varphi = \psi_0 \odot \psi_1, \quad \text{with } \odot \in \{\wedge, \vee\} \\ \varepsilon & \text{if } \varphi = \eta x.\psi, \quad \text{with } \eta \in \{\mu, \nu\} \end{cases}$$

Finally, we let $\mathbb{G}_\xi$ be the parity formula

$$\mathbb{G}_\xi := (\mathbb{C}_\xi, \Omega_C{\restriction}_{Cl(\xi)}, \xi),$$

where Ω_C is the priority map that is induced by the partial order $\preccurlyeq_C$; in particular, Ω_C is a partial map on tidy formulas that is only defined for fixpoint formulas. $\lhd$

Proof of Theorem 6.38. It is obvious that $\mathbb{G}_\xi$ is based on the closure graph of ξ, and it follows from the discussion in subsection 6.3.1 that the index of $\mathbb{G}_\xi$ is equal to the alternating chain depth of the relation $\preccurlyeq_C$. It then follows from Proposition 6.61 that $ind(\mathbb{G}_\xi) \leq \mathtt{ad}(\xi)$. The remaining statement of the Theorem, that is, the equivalence of $\mathbb{G}_\xi$ and ξ, will be proved below, cf. Proposition 6.47. QED

The remainder of this section is devoted to proving the equivalence of a tidy formula ξ to its representing parity formula $\mathbb{G}_\xi$. For this purpose we need some auxiliary technical results. The following proposition states that the global priority map indeed captures the right winner of infinite matches of the evaluation game. Recall that the winner of an infinite match π of the closure evaluation game is given by the parity of the most significant formula $\mathtt{msf}(\pi_L)$ on the trace part of π (that is, the left projection π_L of π.

Proposition 6.44 *Let $\tau = (\xi_n)_{n\in\omega}$ be an infinite trace of tidy formulas, and that $\xi = \eta x\,\chi$ is the most significant formula of τ. Then*

1) ξ is the unique formula which occurs infinitely often on τ and satisfies $\xi_n \preccurlyeq_C \xi$ for cofinitely many n.
2) $\max\Big(\{\Omega(\varphi) \mid \varphi$ occurs infinitely often on $\tau\}\Big)$ is even iff $\eta = \nu$.

Proof. Part 1) is more or less immediate by the Propositions 2.39 and 2.58. From this it follows by Proposition 6.33 that $\Omega_C(\varphi) \leq \Omega_C(\xi)$ for all φ that occur infinitely often on τ, and that $\Omega_C(\xi)$ has the right parity. This proves Part 2). QED

The following proposition describes the relation between the substitution operation and the free subformula relation.

Proposition 6.45 *Let φ, ψ and ξ be formulas in μ*ML *such that $x \in FV(\varphi)$, and ξ is free for x in both φ and ψ. Then*
1) $\varphi \trianglelefteq_f \psi$ implies $\varphi[\xi/x] \trianglelefteq_f \psi[\xi/x]$;
2) $\varphi[\xi/x] \trianglelefteq_f \psi[\xi/x]$ implies $\varphi \trianglelefteq_f \psi$, provided that $\xi \not\trianglelefteq_f \varphi, \psi$.

Proof. For part 1), assume that $\varphi \trianglelefteq_f \psi$, then $\psi = \psi'[\varphi/y]$ for some formula ψ' such that $y \in FV(\psi')$ and φ is free for y in ψ'. Without loss of generality we may assume that y does not occur in ψ. Then by Proposition 2.55 we obtain that $\varphi[\xi/x]$ is free for y in $\psi'[\xi/x]$, and that $\psi'[\xi/x][\varphi[\xi/x]/y] = \psi[\xi/x]$. This means that $\varphi[\xi/x] \trianglelefteq_f \psi[\xi/x]$ indeed.

For part 2), clearly it suffices to show that

$$\text{if } \psi[\xi/x] = \rho\big[\varphi[\xi/x]/y\big] \text{ with } y \in FV(\rho) \text{ then } \psi = \rho'[\varphi/y], \text{ for some } \rho' \text{ with } y \in FV(\rho'). \tag{66}$$

We will prove (66) by induction on the *maximal depth $d_y(\rho)$ of y in ρ* which we define as follows:

$$\begin{array}{lcll}
d_y(\rho) & := & \left\{\begin{array}{ll} 0 & \text{if } \rho = y \\ -\infty & \text{if } \rho \text{ is atomic but } \rho \neq y \end{array}\right. & \\
d_y(\rho_0 \star \rho_1) & := & 1 + \max(d_y(\rho_0), d_y(\rho_1)) & (\star \in \{\vee, \wedge\}) \\
d_y(\heartsuit\rho) & := & 1 + d_y(\heartsuit\rho) & (\heartsuit \in \{\Diamond, \Box\}) \\
d_y(\eta z\,\rho) & := & \left\{\begin{array}{ll} 1 + d_y(\heartsuit\rho) & \text{if } y \neq z \\ -\infty & \text{if } y = z \end{array}\right. &
\end{array}$$

As the basis of our induction we take the cases where $d_y(\rho) \leq 0$. First assume $d_y(\rho) = -\infty$, which means that $y \notin FV(\rho)$, so that $\rho\big[\varphi[\xi/x]/y\big] = \rho$. In this case (66) holds vacuously.

In the other base case we assume $d_y(\rho) = 0$, which means that $\rho = y$, so that $\rho\big[\varphi[\xi/x]/y\big] = \varphi[\xi/x]$. But then the equation of (66) reads $\psi[\xi/x] = \varphi[\xi/x]$, so that we find $\psi = \varphi$ by Proposition 2.59; that is, we may take $\rho' := y$.

In the induction step of the proof we have $d_y(\rho) > 0$, from which we readily infer that ρ is non-atomic and that $y \in FV(\rho)$. From this it is immediate that ξ is a *proper* subformula of $\rho\big[\varphi[\xi/x]/y\big]$. Now make a case distinction as to the nature of ρ, and confine attention to the following two cases.

First assume that $\rho = \rho_0 \wedge \rho_1$, so that the equation in (66) reads $\psi[\xi/x] = \rho_0\big[\varphi[\xi/x]/y\big] \wedge \rho_1\big[\varphi[\xi/x]/y\big]$. It follows that ψ must be of the form $\psi_0 \wedge \psi_1$ with $\psi_i[\xi/x] = \rho_i\big[\varphi[\xi/x]/y\big]$ for $i = 0, 1$ — for otherwise we would have $\psi = x$, which would imply $\psi[\xi/x] = \xi$ and thus contradict our observation that $\xi \triangleleft \rho\big[\varphi[\xi/x]/y\big]$. But then inductively there are formulas ρ'_i such that $\psi_i = \rho'_i[\varphi/x]$, with $y \in FV(\rho_i)$ iff $y \in FV(\rho'_i)$, for $i = 0, 1$. It is then straightforward to verify that $\rho' := \rho'_0 \wedge \rho'_1$ meets the requirement.

Now assume that $\rho = \eta z\, \rho_0$. Then it must be the case that $\psi = \eta z\, \psi_0$, with $\psi_0[\xi/x] = \rho_0\big[\varphi[\xi/x]/y\big]$. (The only other option would be that $\psi = x$, which would mean that the equation in (66) would read $\xi = \rho\big[\varphi[\xi/x]/y\big]$, again contradicting the fact that ξ is a proper subformula of $\rho\big[\varphi[\xi/x]/y\big]$.) By the inductive hypothesis there is a formula ρ'_0 with $y \in FV(\rho_0)$ such that $\psi_0 \equiv \rho'_0[\varphi/y]$. Now define $\rho' := \eta z\, \rho'_0$, then we find $y \in FV(\rho')$, and $\rho'[\varphi/y] = \eta z\, \rho'_0[\varphi/y] = \eta z\, \psi_0 = \psi$. QED

▶ `Exercise: show that conditions` $\xi \ntrianglelefteq_f \varphi, \psi$ `are both needed.`

The next proposition is the key technical ingredient in proving the equivalence of ξ and $\mathbb{G}_\xi$, in the inductive case where ξ is of the form $\eta x.\chi$. Roughly, it states that the substitution ξ/x is 'almost an isomorphism' between $\mathbb{G}_\chi$ and $\mathbb{G}_\xi$; note, however, that actually, rather than χ we consider its variant $\chi' := \chi[x'/x]$ — this guarantees tidyness. Recall that the alternation height $h^{\downarrow}(\xi)$ of a formula ξ was introduced in Definition 6.32

Proposition 6.46 *Let $\xi = \eta x.\chi$ be a tidy fixpoint formula such that $x \in FV(\chi)$ and $\xi \notin Cl(\chi)$. Furthermore, let $\chi' := \chi[x'/x]$ for some fresh variable x'. Then χ' is tidy and the following hold.*

1) the substitution ξ/x' is a bijection between $Cl(\chi')$ and $Cl(\xi)$.

Let $\varphi, \psi \in Cl(\chi')$. Then we have

2) if $\varphi \neq x'$, then $\varphi \to_C \psi$ iff $\varphi[\xi/x'] \to_C \psi[\xi/x']$;

3) if $\varphi \neq x'$, then $L_C(\varphi) = L_C(\varphi[\xi/x'])$;

4) if $x' \in FV(\varphi)$ then $\varphi \trianglelefteq_f \psi$ iff $\varphi[\xi/x'] \trianglelefteq_f \psi[\xi/x']$;

5) if φ and ψ are fixpoint formulas then $\psi \preccurlyeq_C \varphi$ iff $\psi[\xi/x'] \preccurlyeq_C \varphi[\xi/x']$;

6) if $(\varphi_n)_{n\in\omega}$ is an infinite trace through $Cl(\chi')$, then $(\varphi_n)_{n\in\omega}$ has the same winner as $(\varphi_n[\xi/x'])_{n\in\omega}$.

Proof. Let $\xi = \eta x.\chi$ be a tidy fixpoint formula such that $x \in FV(\chi)$ and $\xi \notin Cl(\chi)$, and let $\chi' := \chi[x'/x]$ for some fresh variable x'. We leave it for the reader to verify that χ' is tidy, and first make the following technical observation:

$$\text{if } \varphi \in Cl(\chi') \text{ then } \xi \notin Cl(\varphi) \text{ and } \xi \not\trianglelefteq_f \varphi. \tag{67}$$

Leaving the proof of (67) to the reader, we turn to the respective items of the Proposition.

Item 1): Injectivity can be proved using Proposition 2.59, surjectivity using Proposition 2.44.

Item 2): This follows immediately from the 'back-and-forth lemma' (Proposition 2.56) and item 1).

Item 3): The claim that $L_C(\varphi) = L_C(\varphi[\xi/x'])$ is rather trivial.

Item 4): This is Proposition 6.45. The assumptions $\xi \not\trianglelefteq_f \psi$ and $\xi \not\trianglelefteq_f \varphi$ follow from (67).

Item 5): For the left-to-right direction assume that $\psi \preccurlyeq_C \varphi$. By definition there is some trace $\varphi = \rho_0 \to_C \rho_1 \to_C \ldots \to_C \rho_n = \psi$ such that $\varphi \trianglelefteq_f \rho_i$ for all $i \in [0, n]$. It is clear that none of the ρ_i is equal to x because x has no outgoing $\to_C$-edges and $\psi \neq x'$. Thus we can use item 2) to obtain a trace $\varphi[\xi/x'] = \rho_0[\xi/x'] \to_C \rho_1[\xi/x'] \to_C \ldots \to_C \rho_n[\xi/x'] = \psi[\xi/x']$. By Proposition 6.45 it follows from $\varphi \trianglelefteq_f \rho_i$ that $\varphi[\xi/x'] \trianglelefteq_f \rho_i[\xi/x']$, for all $i \in [0, n]$. That is, we have shown that $\varphi[\xi/x'] \twoheadrightarrow_C^{\varphi[\xi/x']} \psi[\xi/x']$.

Before we turn to the opposite direction we show that, for all $\rho, \sigma \in Cl(\chi')$, we have

$$\text{if } \rho[\xi/x'] \twoheadrightarrow_C \sigma[\xi/x'] \text{ and } x' \in FV(\sigma) \text{ then } x' \in FV(\rho). \tag{68}$$

This claim holds because, since ξ is free for x' in σ, by definition of $\trianglelefteq_f$ it follows from $x' \in FV(\sigma)$ that $\xi \trianglelefteq_f \sigma[\xi/x']$, and thus we find $\sigma[\xi/x'] \twoheadrightarrow_C \xi$ by Proposition 2.58. If it were the case that $x' \notin FV(\rho)$ then we would have that $\rho = \rho[\xi/x'] \twoheadrightarrow_C \sigma[\xi/x'] \twoheadrightarrow_C \xi$, contradicting (67).

Turning to the right-to-left direction of item 5), assume that $\psi[\xi/x'] \preccurlyeq_C \varphi[\xi/x']$. This means that there is a trace $\varphi[\xi/x'] = \rho'_0 \to_C \ldots \to_C \rho'_m = \psi[\xi/x']$ with $\varphi[\xi/x'] \trianglelefteq_f \rho'_i$ for all $i \in [0, m]$. By Proposition 6.42 we have $\psi[\xi/x'] \equiv_C \varphi[\xi/x']$. It follows from (68) and $\psi[\xi/x'] \equiv_C \varphi[\xi/x']$ that x' is either free in both φ and ψ, or free in neither of the two formulas. In the second case we obtain $\varphi = \varphi[\xi/x']$ and $\psi = \psi[\xi/x']$, so that the statement of this item holds trivially.

We now focus on the case where $x' \in FV(\varphi) \cap FV(\psi)$. Our first claim is that $\rho'_i \neq \xi$ for all $i \in [0, m]$. This follows from the fact that $\varphi[\xi/x'] \trianglelefteq_f \rho'_i$, which holds by assumption, and the observation that ξ is a proper free subformula of $\varphi[\xi/x']$, which holds since φ is a fixpoint formula and hence, distinct from x'. But if $\rho'_i \neq \xi$ for all $i \in [0, m]$, we may use the items 1) and 2) to obtain a trace $\varphi = \rho_0 \to_C \ldots \to_C \rho_m = \psi$ such that $\rho_i[\xi/x'] = \rho'_i$ for all $i \in [0, m]$. Furthermore, by Proposition 2.42 it follows from $x' \in FV(\psi)$ that $x' \in FV(\rho_i)$, and so we may use item 4) to obtain $\varphi \trianglelefteq_f \rho_i$, for all $i \in [0, m]$. This suffices to show that $\psi \preccurlyeq_C \varphi$.

Item 6): This observation is immediate by item 5) and Proposition 6.44. QED

Proposition 6.47 *Let ξ be a tidy μ-calculus formula. Then $\xi \equiv \mathbb{G}_\xi$.*

Proof. It will be convenient to consider the *global* formula graph $\mathbb{G} := (\mu\mathtt{ML}^t, \to_C, L_C, \Omega_C)$, where $\mu\mathtt{ML}^t$ is the set of all tidy formulas using a fixed infinite set of variables, and L_C is the obviously defined global labelling function. We may assign a semantics to this global graph using an equally obvious definition of an acceptance game, where the only non-standard aspect is that the carrier set of this 'formula' is infinite. For each tidy formula φ we may then consider the structure $\mathbb{G}\langle\varphi\rangle := (\mu\mathtt{ML}^t, \rightarrow_C, L_C, \Omega_C, \varphi)$ as an initialised (generalised) parity formula. Note that all structures of this form have the same (infinite) set of vertices, but that the only vertices that are accessible in $\mathbb{G}\langle\varphi\rangle$ are the formulas in the (finite) set $Cl(\varphi)$. It is then easy to see that $\mathbb{G}\langle\varphi\rangle \equiv \mathbb{G}_\xi\langle\varphi\rangle$, for any pair of tidy formulas φ, ξ such that $\varphi \in Cl(\xi)$.

In order to prove the Proposition, it therefore suffices to show that every tidy formula ξ satisfies the following:

$$\mathbb{G}\langle\varphi\rangle \equiv \varphi, \text{ for all } \varphi \in Cl(\xi). \tag{69}$$

We will prove (69) by induction on the length of ξ. In the base step of this induction we have $|\xi|^\ell = 1$, which means that ξ is an atomic formula. In this case it is easy to see that (69) holds.

In the induction step of the proof we assume that $|\xi|^\ell > 1$, and we make a case distinction. The cases where ξ is of the form $\xi = \xi_0 \odot \xi_1$ with $\odot \in \{\wedge, \vee\}$ or $\xi = \heartsuit\xi_0$ with $\heartsuit \in \{\Diamond, \Box\}$, are easy and left as exercises for the reader.

In the case where ξ is of the form $\xi = \eta x.\chi$ with $\eta \in \{\mu, \nu\}$ we make a further case distinction. If ξ belongs to the closure set of χ, then we have $Cl(\xi) \subseteq Cl(\chi)$, so that (69) immediately follows from the induction hypothesis, applied to the formula χ.

This leaves the case where ξ is of the form $\eta x.\chi$, while $\xi \notin Cl(\chi)$. Let x' be some fresh variable, then obviously we may apply the induction hypothesis to the (tidy) formula $\chi' := \chi[x'/x]$. The statement that $\xi \equiv \mathbb{G}\langle\xi\rangle$ now follows by a routine argument, based on the observations in Proposition 6.46. QED •

6.4 From parity formulas to ordinary formulas

In section 6.3 we saw constructions that, for a given ordinary formula, produce equivalent parity formulas based on, respectively, the subformula graph and the closure graph of the original formula. We will now move in the opposite direction: we will give a construction that turns an arbitrary parity formula $\mathbb{G}$ into an equivalent ordinary formula $\xi_{\mathbb{G}} \in \mu\mathtt{ML}$. Basically this construction takes a parity formulas as a system of equations, and it solves these equations by a Gaussian elimination of variables.

Interestingly, we encounter a significant difference between the two size measures introduced in Definition 2.48: whereas the closure-size of the resulting formula $\xi_{\mathbb{G}}$ is *linear* in the size of $\mathbb{G}$, its number of subformulas is only guaranteed to be exponential. And in fact, Proposition 6.55 shows that there is a family of parity formulas for which the translation actually reaches this exponential subformula-size.

The main result in this section is the following.

Theorem 6.48 *There is an effective procedure which transforms a parity formula* $\mathbb{G} = (V, E, L, \Omega, v_I)$ *over some set* P *of proposition letters, into a* $\mu\mathtt{ML}$*-formula* $\xi_{\mathbb{G}}$ *such that* $\xi_{\mathbb{G}} \equiv \mathbb{G}$*,* $|\xi_{\mathbb{G}}| \leq 2 \cdot |\mathbb{G}|$ *and* $\mathsf{ad}(\xi_{\mathbb{G}}) \leq ind(\mathbb{G})$*.*

In order to prove Theorem 6.48 we will supply a map

$$\mathtt{tr}_{\mathbb{G}} : V \to \mu\mathtt{ML}(\mathsf{P})$$

for every balanced parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ over some set P of proposition letters. For an arbitrary parity formula $\mathbb{G}$ we then obtain the formula $\xi_{\mathbb{G}}$ by taking some balanced version $\mathbb{G}'$ of $\mathbb{G}$ and apply the translation $\mathtt{tr}_{\mathbb{G}'}$ to the initial vertex of $\mathbb{G}'$.

The definition of the map $\mathtt{tr}_{\mathbb{G}}$, just like most definitions and proofs in this section will proceed by an induction on a certain complexity measure of parity formulas that we shall refer to as *weight*.

Definition 6.49 We define the *weight* of a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ as the pair $(|\mathsf{Dom}(\Omega)|, |\mathbb{G}|)$ consisting of, respectively, the number of states and the size of $\mathbb{G}$. Pairs of this form will be ordered lexicographically. ◁

Definition 6.50 Consider a balanced parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$, and let T be the top cluster of $\mathbb{G}$, that is, the cluster of the initial state v_I. We make the following case distinction.

Case 1: T is transient. In this case we must have $T = \{v_I\}$, with $v_I \notin \mathsf{Ran}(E)$, and for every $u \neq v_I$ we may consider the parity formula $G^{-}\langle u \rangle$ obtained by removing v_I from $\mathbb{G}$ and making u the initial node. Each parity formula $\mathbb{G}^{-}\langle u \rangle$ has the same number of states as $\mathbb{G}$, but less vertices, so that inductively we may assume that we have defined some formula $\mathtt{tr}_{\mathbb{G}^{-}}(u)$ for each $u \neq v_I$. (Here we write $\mathtt{tr}_{\mathbb{G}^{-}}$ instead of $\mathtt{tr}_{\mathbb{G}^{-}\langle u \rangle}$; this is justified since the translation we define does not depend on the initial node.)

We define

$$\mathtt{tr}_{\mathbb{G}}(u) := \mathtt{tr}_{\mathbb{G}^{-}}(u)$$

for $u \neq v_I$, while for v_I we set[5]

$$\mathtt{tr}_{\mathbb{G}}(v_I) := \begin{cases} L(v_I) & \text{if } L(v_I) \in \mathtt{At}(\mathsf{P}) \\ \heartsuit \mathtt{tr}_{\mathbb{G}^-}(u) & \text{if } L(v) = \heartsuit \in \{\Diamond, \Box\} \text{ and } E(v) = \{u\} \\ \bigodot\{\mathtt{tr}_{\mathbb{G}^-}(u) \mid u \in E(v)\} & \text{if } L(v) = \odot \in \{\wedge, \vee\} \end{cases}$$

Case 2: T is proper. In this case we have $T \cap \mathsf{Dom}(\Omega) \neq \varnothing$. Take an arbitrary state $z \in T$ such that $\Omega(z)$ equals the maximal priority reached on T. Since $\mathbb{G}$ is balanced, z is silent (that is, $L(z) = \varepsilon$); let z' be its unique successor.

Consider the parity formula $\mathbb{G}^- = (V, E^-, L^-, \Omega^-, v_I)$, which is characterised by the definition of E^-:

$$E^- := E \setminus \{(z, z')\},$$

that is, we obtain E^- from E by cutting the edge from z to z'. Furthermore, we define

$$L^-(v) := \begin{cases} L(v) & \text{if } v \neq z \\ z & \text{if } v = z, \end{cases}$$

that is, $\mathbb{G}^-$ sees z as a proposition letter. Finally, we set $\Omega^- := \Omega\restriction_{V\setminus\{z\}}$, so that $\mathbb{G}^-$ has one state less than $\mathbb{G}$.

Inductively, then, we may assume that for all $v \in V$, some formula $\mathtt{tr}_{\mathbb{G}^-}(v)$ has been defined, and observe that $\mathtt{tr}_{\mathbb{G}^-}(z) = z$. We now define

$$\mathtt{tr}_{\mathbb{G}}(v) := \mathtt{tr}_{\mathbb{G}^-}(v)[\eta_z z.\mathtt{tr}_{\mathbb{G}^-}(z')/z]$$

where $\eta_z := \mu$ if $\Omega(z)$ is odd, and $\eta_z := \nu$ if $\Omega(z)$ is even. ◁

▶ `Example to be supplied!`

Remark 6.51 Before we turn to more interesting properties we note that the translation $\mathtt{tr}_{\mathbb{G}}$ is *well-defined*.

▶ `some detail to be supplied`

Note in particular that, although the definition of the translation map $\mathtt{tr}_{\mathbb{G}}$ involves many substitution operations, it does *not* involve any renaming of variables. ◁

Definition 6.52 Let $\mathbb{G}$ be an arbitrary parity formula. If $\mathbb{G}$ is balanced, we define

$$\xi_{\mathbb{G}} := \mathtt{tr}_{\mathbb{G}}(v_I),$$

where v_I is the initial vertex of $\mathbb{G}$. If $\mathbb{G}$ is not balanced, we put $\xi_{\mathbb{G}} := \xi_{\mathbb{G}^b}$, where $\mathbb{G}^b$ is some effectively obtainalbe balanced formula which is equivalent to $\mathbb{G}$ and satisfies $|\mathbb{G}^b| \leq 2 \cdot |\mathbb{G}|$ and $ind(\mathbb{G}^b) = ind(\mathbb{G})$. ◁

[5]Note that the formulation of the boolean clause of our definition (i.e., the case where $L(v) = \odot \in \{\wedge, \vee\}$) is a bit sloppy, since our language only has binary conjunctions and disjunctions, no conjunctions or disjunctions over finite sets. A more precise definition can be given as follows; assume that $\odot = \wedge$, the case where $\odot = \vee$ is treated analogously. Assume that we have some arbitrary but fixed ordering of the vertices in $\mathbb{G}$. We put $\mathtt{tr}_{\mathbb{G}}(v) = \top$ if $E(v) = \varnothing$, $\mathtt{tr}_{\mathbb{G}}(v) = \mathtt{tr}_{\mathbb{G}}(u)$ if $E(v) = \{u\}$, and $\mathtt{tr}_{\mathbb{G}}(v) = \mathtt{tr}_{\mathbb{G}}(u_0) \wedge (\ldots \wedge (\mathtt{tr}_{\mathbb{G}}(u_{k-1}) \wedge \mathtt{tr}_{\mathbb{G}}(u_k))$ if $E(v) = \{u_0, \ldots, u_k\}$ where each u_i comes before u_{i+1} in the mentioned ordering.

The crucial technical lemma in the proof of Theorem 6.48 is the following proposition.

Proposition 6.53 *Let* $\mathbb{G}$ *be a balanced parity formula. Then the map* $\mathtt{tr}_{\mathbb{G}}$ *is a surjective parity formula morphism:*

$$\mathtt{tr}_{\mathbb{G}} : \mathbb{G} \twoheadrightarrow \mathbb{G}_{\xi_{\mathbb{G}}}.$$

Proof. We will prove the proposition via a series of claims. Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be a balanced parity formula. First we show that $\mathtt{tr}_{\mathbb{G}}$ is a surjection from V to $Cl(\xi_{\mathbb{G}})$.

CLAIM 1 $\mathtt{tr}_{\mathbb{G}}[V] = Cl(\xi_{\mathbb{G}})$.

PROOF OF CLAIM By induction on the weight of $\mathbb{G}$ we will prove the following two statements:
1) $\mathtt{tr}_{\mathbb{G}}(v) \in Cl(\xi_{\mathbb{G}})$, for all $v \in V$;
2) $Cl(\xi_{\mathbb{G}}) \subseteq \mathtt{tr}_{\mathbb{G}}[V]$.

As in Definition 6.50, we let T be the top cluster of $\mathbb{G}$, and make a case distinction. Leaving the case where T is transient as an exercise, we focus on the case where T is proper, and we let z and $\mathbb{G}^-$ be as in Definition 6.50. We abbreviate $\tau := [\mathtt{tr}_{\mathbb{G}}(z)/z]$, and we define the *closure of* $\mathbb{G}$ as follows:

$$Cl(\mathbb{G}) := \bigcup \big\{ Cl(\mathtt{tr}_{\mathbb{G}}(v)) \mid v \in V \big\}.$$

For part 1) we first observe that $\xi_{\mathbb{G}} = \mathtt{tr}_{\mathbb{G}}(v_I) = \mathtt{tr}_{\mathbb{G}^-}(v_I)[\tau] = \xi_{\mathbb{G}^-}[\tau]$. From this it follows by Proposition 2.44 that $Cl(\xi_{\mathbb{G}}) \supseteq \{\varphi[\tau] \mid \varphi \in Cl(\xi_{\mathbb{G}^-})\}$. Now take an arbitrary vertex $v \in V$. Then by the inductive hypothesis we have $\mathtt{tr}_{\mathbb{G}^-}(v) \in Cl(\xi_{\mathbb{G}^-})$, so that $\mathtt{tr}_{\mathbb{G}^-}(v)[\tau] \in \{\varphi[\tau] \mid \varphi \in Cl(\xi_{\mathbb{G}^-})\} \subseteq Cl(\xi_{\mathbb{G}})$ as required.

For part 2) our key observation is that

$$Cl(\mathbb{G}) \subseteq \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\}, \tag{70}$$

where τ is the substitution $\tau = [\mathtt{tr}_{\mathbb{G}}(z)/z]$. For a proof of (70), we have to show that

$$Cl(\mathtt{tr}_{\mathbb{G}}(v)) \subseteq \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\},$$

for every vertex v in V. To show this we make a case distinction. In case $v = z$, we have $\mathtt{tr}_{\mathbb{G}}(z) = \eta_z z.\mathtt{tr}_{\mathbb{G}^-}(z')$, and so we find

$$\begin{aligned}
Cl(\mathtt{tr}_{\mathbb{G}}(z)) &= \{\mathtt{tr}_{\mathbb{G}}(z)\} \cup \{\varphi[\tau] \mid \varphi \in Cl(\mathtt{tr}_{\mathbb{G}^-}(z'))\} && \text{(Proposition 2.44)}\\
&\subseteq \{z[\tau]\} \cup \{\varphi[\tau] \mid \varphi \in Cl(\mathtt{tr}_{\mathbb{G}^-}(z'))\} && \text{(obvious)}\\
&\subseteq \{z[\tau]\} \cup \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\} && \text{(definition } Cl(\mathbb{G}^-))\\
&\subseteq \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\} && (z \in Cl(\mathbb{G}^-))
\end{aligned}$$

On the other hand, if $v \neq z$, we have $\mathtt{tr}_{\mathbb{G}}(v) = \mathtt{tr}_{\mathbb{G}^-}(v)[\mathtt{tr}_{\mathbb{G}}(z)/z]$, and so here we obtain

$$\begin{aligned}
Cl(\mathtt{tr}_{\mathbb{G}}(v)) &= \{\varphi[\tau] \mid \varphi \in Cl(\mathtt{tr}_{\mathbb{G}^-}(v))\} \cup Cl(\mathtt{tr}_{\mathbb{G}}(z)) && \text{(Proposition 2.44)}\\
&= \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)(v))\} \cup Cl(\mathtt{tr}_{\mathbb{G}}(z)) && \text{(definition } Cl(\mathbb{G}^-))\\
&\subseteq \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\} && \text{(just proved)}
\end{aligned}$$

Now that we have established (70), the remainder of the proof is straightforward:

$$\begin{aligned} Cl(\mathbb{G}) &\subseteq \{\varphi[\tau] \mid \varphi \in Cl(\mathbb{G}^-)\} && \text{(just proved)} \\ &= \{\mathtt{tr}_{\mathbb{G}^-}(v)[\tau] \mid v \in V\} && \text{(induction hypothesis)} \\ &= \{\mathtt{tr}_{\mathbb{G}}(v) \mid v \in V\} && \text{(definition } \mathtt{tr}_{\mathbb{G}}) \end{aligned}$$

This suffices to prove the Claim. ◀

The next claim states that the map $\mathtt{tr}_{\mathbb{G}}$ has the back-and-forth property; we leave its proof as an exercise for the reader.

Claim 2 Let v be some vertex in $\mathbb{G}$. Then the following hold:
1) if Evw then $\mathtt{tr}_{\mathbb{G}}(v) \to_C \mathtt{tr}_{\mathbb{G}}(w)$;
2) if $\mathtt{tr}_{\mathbb{G}}(v) \to_C \varphi$ then there is a vertex $w \in V$ such that Evw and $\mathtt{tr}_{\mathbb{G}}(w) = \varphi$.

To establish the morphism conditions on the priority map we formulate the following claim.

Claim 3 The map $\mathtt{tr}_{\mathbb{G}}$ satisfies condition (4)) of Definition 6.18.

Proof of Claim It is immediate by the definitions that $\mathtt{tr}_{\mathbb{G}}(v)$ is a fixpoint formula whenever v is a *state* in $\mathbb{G}$, and that v and $\mathtt{tr}_{\mathbb{G}}(v)$ have the same parity. Given the definition of the parity formula $\mathbb{G}_{\xi_{\mathbb{G}}}$, it is left to show that

$$\Omega(v) < \Omega(w) \text{ iff } \mathtt{tr}_{\mathbb{G}}(v) \prec_C \mathtt{tr}_{\mathbb{G}}(w), \tag{71}$$

for any pair of states v, w of distinct parity that belong to the same cluster of $\mathbb{G}$. For the proof from left to right, the key observation is that $\Omega(v) < \Omega(w)$ means that w is 'processed' later than v in the definition of $\mathtt{tr}_{\mathbb{G}}$, while v and w belonging to the same cluster implies that w is a free variable in $\mathtt{tr}_{\mathbb{G}'}(v)$ for any intermediate-stage graph $\mathbb{G}'$ in which $\mathtt{tr}_{\mathbb{G}'}(v)$ has been defined already. Further details are left to the reader.

For the right-to-left direction of (71) we reason by contraposition. Suppose that $\Omega(v) \not< \Omega(w)$, then since v and w have different parity we must have $\Omega(w) < \Omega(v)$, so that by the left-to-right direction of (71) we find $\mathtt{tr}_{\mathbb{G}}(w) \prec_C \mathtt{tr}_{\mathbb{G}}(v)$. But then clearly we cannot have $\mathtt{tr}_{\mathbb{G}}(v) \prec_C \mathtt{tr}_{\mathbb{G}}(w)$. ◀

Finally, since we have $\mathtt{tr}_{\mathbb{G}}(v_I) = \xi_G$, it is immediate by the definitions that the map $\mathtt{tr}_{\mathbb{G}}$ satisfies condition (5)) of the definition or a parity formula morphism. QED

We now turn to the proof of the final item of Theorem 6.48.

Proposition 6.54 *For any parity formula $\mathbb{G}$ and for any vertex v in $\mathbb{G}$ we have* $\mathsf{ad}(\mathtt{tr}_{\mathbb{G}}(v)) \leq$ *ind*$(\mathbb{G})$.

Proof. Let C be a cluster of $\mathbb{G}$, and let η be either μ or ν. An alternating η-chain in C of length k is a sequence $v_1 \cdots v_k$ of states such that, for all $i < k$ we have $\Omega(v_i) < \Omega(v_k)$, while v_i and v_{i+1} have different parities. We define $\mathsf{acd}^{\eta}_{\Omega}(C)$ as the maximal length of such a chain, with $\mathsf{acd}^{\eta}_{\Omega}(C) := 0$ if C has no such chains. Our key observation is the following claim.

Claim 1 Let d and η be such that $\mathsf{acd}^{\eta}_{\Omega}(C) \leq d$ for every cluster C of $\mathbb{G}$. Then $\mathtt{tr}_{\mathbb{G}}(v) \in \Theta^{\eta}_{d}$.

Proof of Claim We prove the claim by induction on the weight $(|\mathsf{Dom}(\Omega)|, |\mathbb{G}|)$ of $\mathbb{G}$. Let T be the top cluster of $\mathbb{G}$, and make a case distinction. We leave the case where T is transient as an exercise, and focus on the case where T is proper. Let, as in Definition 6.50, z be a state in T of maximal priority.

Let $\mathbb{G}_T = (V, E_T, L_T, \Omega_T, v_I)$ be the parity formula given by $E_T := E \cap (T \times V)$, $L_T := L{\restriction_T} \cup \{(u,u) \mid u \in V \setminus T\}$ and $\Omega_T := \Omega{\restriction_T}$. In words, $\mathbb{G}_T$ is the parity formula we obtain from $\mathbb{G}$ by focusing on the top cluster T, replacing, for every vertex $u \notin T$, the generated subgraph G^u with the 'atomic' parity formula representing the atom u. It is not hard to see that, for all $u \in V \setminus T$ we have

$$\mathtt{tr}_{\mathbb{G}}(u) = \mathtt{tr}_{\mathbb{G}^u}(u), \tag{72}$$

while the point of the construction is that for every $t \in T$ we get:

$$\mathtt{tr}_{\mathbb{G}}(t) = \mathtt{tr}_{\mathbb{G}_T}(t)[\mathtt{tr}_{\mathbb{G}^u}(u)/u \mid u \in V \setminus T]. \tag{73}$$

Now suppose that we can prove, for all $t \in T$, that

$$\mathtt{tr}_{\mathbb{G}_T}(t) \in \Theta^{\eta}_{d}. \tag{74}$$

Note that by the induction hypothesis, applied to the parity formulas $\mathbb{G}^u$ with $u \in V \setminus T$, we have $\mathtt{tr}_{\mathbb{G}^u}(u) \in \Theta^{\eta}_{d}$. Then we may use clause (4) of Definition 2.50 to derive from (73) and (74) that $\mathtt{tr}_{\mathbb{G}}(v) \in \Theta^{\eta}_{d}$ as required.

It is thus left to prove (74), and for this purpose we shall apply the induction hypothesis to the parity formula $\mathbb{G}^{-}_{T}$. Let λ be the parity of z. We will make use of the following relation between $\mathtt{tr}_{\mathbb{G}}$ and $\mathtt{tr}_{\mathbb{G}_T}$, which is not hard to prove:

$$\text{if } \mathsf{Ran}(\mathtt{tr}_{\mathbb{G}^-}) \subseteq \Theta^{\overline{\lambda}}_{e} \text{ then } \mathsf{Ran}(\mathtt{tr}_{\mathbb{G}}) \subseteq \Theta^{\overline{\lambda}}_{e}. \tag{75}$$

Turning to the proof of (74), we make a case distinction, as to the nature of λ. Our reasoning will be slightly different in either case.

First consider the case where $\lambda = \eta$. This implies that every cluster D of $\mathbb{G}^{-}_{T}$ satisfies $ind_{\overline{\eta}}(D) \leq d-1$. Then by the induction hypothesis we find that $\mathtt{tr}_{\mathbb{G}^{-}_{T}}(v) \in \Theta^{\overline{\eta}}_{d-1}$, for all $v \in T$. From this it follows by (75) that $\mathsf{Ran}(\mathtt{tr}_{\mathbb{G}}) \subseteq \Theta^{\overline{\lambda}}_{d-1} = \Theta^{\overline{\eta}}_{d-1}$, which means that we are done since $\Theta^{\overline{\eta}}_{d-1} \subseteq \Theta^{\eta}_{d}$.

If, on the other hand, we have $\lambda = \overline{\eta}$, then we reason as follows. Clearly, every cluster D of $\mathbb{G}^{-}_{T}$ satisfies $ind_{\eta}(D) \leq d$. It follows by the induction hypothesis that $\mathtt{tr}_{\mathbb{G}^{-}_{T}}(v) \in \Theta^{\eta}_{d}$, for all $v \in T$. But then by (75) every formula of the form $\mathtt{tr}_{\mathbb{G}_T}(t)$ belongs to Θ^{η}_{d} as required. ◀

Finally, it is not hard to derive the Proposition from Claim 1. With $d := ind(\mathbb{G})$ and C a cluster of $\mathbb{G}$, one easily derives from the definitions that $\mathsf{acd}^{\mu}_{\Omega}(C), \mathsf{acd}^{\nu}_{\Omega}(C) \leq d$. From this it is immediate by the Claim that $\mathtt{tr}_{\mathbb{G}}(v) \in \Theta^{\mu}_{d} \cap \Theta^{\nu}_{d}$, for any v in $\mathbb{G}$. From this we obtain $\mathsf{ad}(\mathtt{tr}_{\mathbb{G}}(v)) \leq d$ as required. QED

Proof of Theorem 6.48. First let $\mathbb{G}$ be an arbitrary balanced parity formula. By Proposition 6.53 and Theorem 6.38 we find that, respectively, $\mathbb{G} \equiv \mathbb{G}_{\xi_{\mathbb{G}}}$ and $\mathbb{G}_{\xi_{\mathbb{G}}} \equiv \xi_{\mathbb{G}}$. From this it is immediate that $\mathbb{G} \equiv \xi_{\mathbb{G}}$. Furthermore, it follows from Claim 1 in the proof of Proposition 6.53 that $|\xi_{\mathbb{G}}| \leq |\mathbb{G}|$, and from Proposition 6.54 that $\mathtt{ad}(\mathtt{tr}_{\mathbb{G}}(v)) \leq ind(\mathbb{G})$.

From this and the definition of $\xi_{\mathbb{G}}$ the Theorem is immediate. QED

The construction of the Definitions 6.50 and 6.52 provides, for every parity formula $\mathbb{G}$, an equivalent formula $\xi_{\mathbb{G}}$ of size *linear* in the size of $\mathbb{G}$, at least, if we define the size of $\xi_{\mathbb{G}}$ as it *closure-size*. If we measure $\xi_{\mathbb{G}}$ by its number of subformulas, the best upper bound that we can obtain is *exponential*, see Exercise 6.2. (Note that we cannot state that the *subformula-size* of $\mathtt{tr}_{\mathbb{G}}$ is at most exponential in the size of $\mathbb{G}$ since the formula $\xi_{\mathbb{G}}$ will generally not be clean, and so its subformula-size may not be defined.)

The next proposition reveals that the translation given in Definition 6.50 may actually produce formulas with exponentially many subformulas, relative to the size of the parity formula. The 'culprit' here is the application of the substitution operation in the inductive step of the definition, since this may double the number of subformulas each time it is applied. The proposition thus provides an example witnessing that the closure size of a formula can be *exponentially* smaller than its number of subformulas.

Proposition 6.55 *There is a family $(\mathbb{F}_n)_{n\in\omega}$ such that for every n it holds that $|\mathbb{F}_n| \leq 2n+2$, which implies that $|\xi_{\mathbb{F}_n}|$ is linear in n, while $|Sf(\xi_n| \geq 2^n$.*

Proof. For some arbitrary but fixed number n, consider the parity formula $\mathbb{F} = (V, E, L, \Omega, v_I)$ given by

$$\begin{array}{lcl} V & := & \{s_i, v_i \mid 0 \leq i \leq n\} \\ E & := & \{(s_i, v_i) \mid 0 \leq i \leq n\} \cup \{(s_{i+1}, s_i) \mid 0 \leq i \leq n-1\} \\ & & \cup\, \{(v_0, s_n)\} \cup \{(v_i, s_i) \mid 0 < i \leq n\} \\ L & := & \{(s_i, \wedge), (v_i, \diamond) \mid 0 \leq i \leq n\} \\ \Omega & := & \{(v_i, i) \mid 0 \leq i \leq n\} \\ v_I & := & v_0. \end{array}$$

In Figure 10 we display a picture of the parity formula $\mathbb{F}$, for $n = 4$.

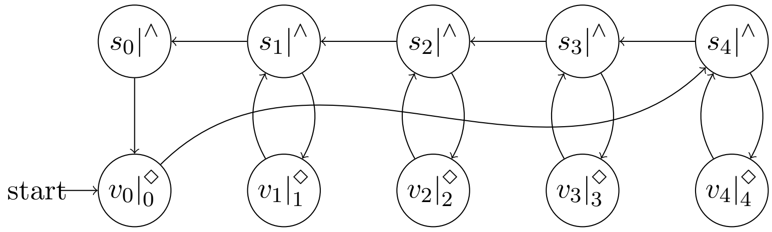


Figure 10: the parity formula $\mathbb{F}$

Our claim is that

$$|Sf(\mathtt{tr}_{\mathbb{F}}(v_0))| \geq 2^n, \tag{76}$$

and in order to prove (76), we will use the notion of *fixpoint depth* of a formula. Recall that we define $fd(\varphi) := 0$ if φ is atomic, $fd(\varphi_0 \odot \varphi_1) := \max(fd(\varphi_0), fd(\varphi_1))$, $fd(\heartsuit\varphi) := fd(\varphi)$, and $fd(\eta x.\varphi) := 1 + fd(\varphi)$. It is an easy exercise to verify that any μ-calculus formula ξ satisfies $|Sf(\xi)| \geq fd(\xi)$, so that, in order to prove (76), it suffices to show that

$$fd(\mathtt{tr}_{\mathbb{F}}(v_0)) \geq 2^n. \tag{77}$$

To calculate $\mathtt{tr}_{\mathbb{F}}(v_0)$ it will be useful to introduce some auxiliary structures. For $k \in [0, n]$, we let $\mathbb{F}_k$ denote the formula $(V_k, E_k, L_k, \Omega_k, s)$ given by

$$\begin{array}{lcl}
V_k & := & V \cup \{u_i \mid k \leq i \leq n\} \\
E_k & := & \{(s_i, v_i) \mid 0 \leq i < k\} \cup \{(s_j, u_j) \mid k \leq j \leq n\} \cup \{(s_{i+1}, s_i) \mid 0 \leq i \leq n-1\} \\
 & & \cup\, \{(v_0, s_n)\} \cup \{(v_i, s_i) \mid 0 < i \leq n\} \\
L_k & := & L \cup \{(u_j, v_j) \mid k \leq j \leq n\} \\
\Omega_k & := & \{(v_i, i) \mid 0 \leq i \leq k\}
\end{array}$$

For an example, see Figure 11, which contains a picture of the formula $\mathbb{F}_2$ in the case where $n = 4$. Using the notation of Definition 6.50 (but writing u_i for v_i^*), we have $\mathbb{F} = \mathbb{F}_n$

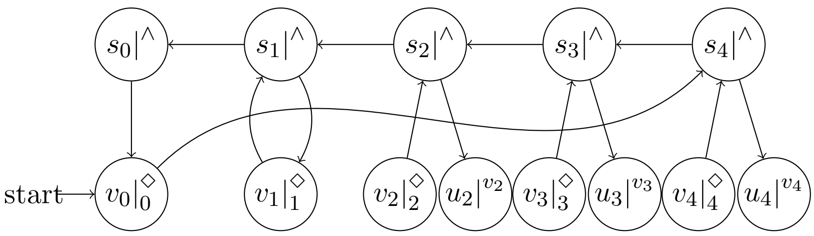


Figure 11: the parity formula $\mathbb{F}_2$

and $\mathbb{F}_{k+1}^- = \mathbb{F}_k$, for all $k \in [0, n-1]$.

We now turn to the translation maps associated with these parity formulas. Observe that it follows from the definitions that $M_{\mathbb{F}_k} = \{v_k\}$, so that we obtain the following definitions (where to avoid clutter we write $\mathtt{tr}_k$ rather than $\mathtt{tr}_{\mathbb{F}_k}$, and omit brackets in conjunctions):

$$\begin{array}{lcll}
\mathtt{tr}_0(v_0) & := & \Diamond \bigwedge_{0 \leq i \leq n} v_i & \\
\mathtt{tr}_0(v_{m)} & := & \Diamond \bigwedge_{0 \leq i \leq m} v_i & \text{for all } m \in [1, n] \\
\mathtt{tr}_{k+1}(v_k) & := & \eta_k v_k.\mathtt{tr}_k(v_k) & \text{for all } k \in [0, n] \\
\mathtt{tr}_{k+1}(v_\ell) & := & \mathtt{tr}_k(v_\ell)[\mathtt{tr}_{k+1}(v_k)/v_k] & \text{for all } k \in [0, n] \text{ and all } \ell \neq k)
\end{array}$$

In order to prove (76), we need an auxiliary notion of (relative) fixpoint depth. Given a formula φ and variable x, we let $fd(x, \varphi)$, the *fixpoint depth of x in φ*, denote the maximum number of fixpoint operators that one may meet on a path from the root of the syntax tree of φ to a free occurrence of x in φ, with $fd(x, \varphi) = -\infty$ if no such occurrence exists. Formally, we set

$$
\begin{array}{lcll}
fd(x,\varphi) & := & \begin{cases} 0 & \text{if } \varphi = x \\ -\infty & \text{if } \varphi \text{ is atomic, but } \varphi \neq x \end{cases} & \\
fd(x,\varphi_0 \odot \varphi_1) & := & \max\big(fd(x,\varphi_0), fd(x,\varphi_1)\big) & \text{where } \odot \in \{\wedge,\vee\} \\
fd(x,\heartsuit\varphi) & := & fd(x,\varphi) & \text{where } \heartsuit \in \{\Diamond,\Box\} \\
fd(x,\eta y.\varphi) & := & \begin{cases} -\infty & \text{if } x = y \\ 1 + fd(x,\varphi) & \text{if } x \neq y \end{cases} & \text{where } \eta \in \{\mu,\nu\}
\end{array}
$$

Without proof we mention that, provided $x \neq y$ and y is free for y in φ:

$$fd(x,\varphi[\psi/y]) = \max\big(fd(x,\varphi), fd(y,\varphi) + fd(x,\psi)\big).$$

From this we immediately infer that

$$fd(x,\varphi[\psi/y]) \geq fd(y,\varphi) + fd(x,\psi), \tag{78}$$

which is in fact the crucial observation in the proof: here we see that the translation doubles the fixpoint depth of the formulas in every step.

CLAIM 1 For all $k \in [1,n]$, and all $\ell, m \geq k$ we have that $fd(v_\ell, \mathtt{tr}_k(v_m)) \geq 2^k - 1$.

PROOF OF CLAIM We prove the claim by induction on k. For the base step of the induction, where $k = 1$, it suffices to observe that $fd(v_\ell, \mathtt{tr}_1(v_m)) = 1$, for all $\ell, m \geq 1$. But this is obvious by the observation that for all $m \geq 1$ we may calculate $\mathtt{tr}_1(v_m) = \Diamond\big(\nu v_0.\Diamond \bigwedge_{0 \leq i \leq n} v_i\big) \wedge \bigwedge_{1 \leq i \leq \ell} v_i$.

For the induction step, we consider the case for $k+1$. Taking arbitrary numbers $\ell, m \geq k+1$, we reason as follows:

$$
\begin{aligned}
fd(v_\ell, \mathtt{tr}_{k+1}(v_m)) &= fd(v_\ell, \mathtt{tr}_k(v_m)[\mathtt{tr}_{k+1}(v_k)/v_k]) && \text{(definition } \mathtt{tr}_{k+1}(v_m)\text{)} \\
&\geq fd(v_k, \mathtt{tr}_k(v_m)) + fd(v_\ell, \mathtt{tr}_{k+1}(v_k)) && \text{(equation (78))} \\
&= fd(v_k, \mathtt{tr}_k(v_m)) + fd(v_\ell, \eta_k v_k.\mathtt{tr}_k(v_k))) && \text{(definition } \mathtt{tr}_{k+1}(v_k)\text{)} \\
&\geq fd(v_k, \mathtt{tr}_k(v_m)) + 1 + fd(v_\ell, \mathtt{tr}_k(v_k))) && \text{(definition } fd(\cdot),\ \ell > k\text{)} \\
&\geq (2^k - 1) + 1 + (2^k - 1) && \text{(induction hypothesis, twice)} \\
&= 2^{k+1} - 1.
\end{aligned}
$$

Clearly this finishes the proof of the claim. ◀

Finally, it is easy to see how (77) follows from the Claim. QED

Exercise 6.2 Let $\mathbb{G}$ be an arbitrary balanced parity formula. Prove that the number of subformulas of $\xi_{\mathbb{G}}$ is at most exponential in the size of $\mathbb{G}$.

6.5 Alternation depth

in Chapter 2 we formally defined the notion of alternation depth by means of an inductive definition. We also mentioned that, intuitively, the alternation depth corresponds to the maximal length of certain alternating chains. In this section we will make this relation precise by proving the following results:

- If ξ is clean, then its alternation depth corresponds to the maximal length of an alternating chain of bound variables of ξ that are ordered by the dependency order $\preccurlyeq_\xi$:
$$\mathsf{ad}(\xi) = \mathsf{acd}(\preccurlyeq_\xi). \tag{79}$$
- if ξ is tidy, then its alternation depth corresponds to the maximal length of an alternating chain of fixpoint formulas in $Cl(\xi)$ that are ordered by the closure priority order $\preccurlyeq_C$:
$$\mathsf{ad}(\xi) = \mathsf{acd}(\preccurlyeq_C\restriction_\xi). \tag{80}$$

Unfortunately the proofs of these results are rather finicky, especially the one of (80).

First of all, we need to refine some of our earlier definitions on alternating chains.

Definition 6.56 Let $\mathbb{Z} = (Z, \preccurlyeq, p)$ be some parity preorder, and let $\vec{z} = z_1 \cdots z_k$ be some alternating chain in $\mathbb{Z}$. We call $\vec{z}$ a μ-chain if $p(z_k) = 1$, and a ν-chain if $p(z_k) = 0$.

For $\eta \in \{\mu, \nu\}$ we define the *alternating η-depth* $\mathsf{acd}_\eta(\mathbb{Z})$ of $\mathbb{Z}$ as the maximal length of an alternating η-chain in $\mathbb{Z}$. $\lhd$

Clearly then we have $\mathsf{acd}(\preccurlyeq) = \max(\mathsf{acd}_\mu(\preccurlyeq), \mathsf{acd}_\nu(\preccurlyeq))$. Note as well that if $\mathbb{Z}$ has no states of odd (even) parity, then we set $\mathsf{acd}_\mu(\preccurlyeq) = 0$ ($\mathsf{acd}_\nu(\preccurlyeq) = 0$, respectively).

For an intuitive understanding of the following technical observation, recall that clause (4) of Definition 2.50 states that the classes Θ^η_n are closed under substitution. Proposition 6.57 states a kind of converse to this. We leave the proof of this result as an exercise for the reader.

Proposition 6.57 *Let ξ and χ be μ-calculus formulas such that ξ is free for x in χ. If $\chi[\xi/x] \in \Theta^\eta_k$ then $\chi \in \Theta^\eta_k$. Furthermore, if $x \in FV(\chi)$ then we also have $\xi \in \Theta^\eta_k$.*

In the case of a *clean* formula there is a simple characterisation of alternation depth, linking it to the length of alternating chains of the dependency order $\preccurlyeq_\xi$ on the bound variables of ξ.

Proposition 6.58 *Let ξ be a clean formula. Then for any $k \in \omega$ and $\eta \in \{\mu, \nu\}$ we have*
$$\xi \in \Theta^\eta_k \text{ iff } \mathsf{acd}_\eta(\preccurlyeq_\xi) \leq k, \tag{81}$$
As a corollary, the alternation depth of ξ is equal to the length of its longest alternating dependency chain.

One of the key insights in the proof of this Proposition is that, with ψ free for x in φ, any dependency chain in $\varphi[\psi/x]$ originates entirely from either φ or ψ. Recall from Definition 2.2 that we write $\overline{\mu} = \nu$ and $\overline{\nu} = \mu$.

Proof of Proposition 6.58. For the sake of a concise notation we will write $d_\eta(\xi) := \mathsf{acd}_\eta(\preccurlyeq_\xi)$.

We prove the implication from left to right in (81) by induction on the derivation that $\xi \in \Theta^\eta_k$. In the base step of this induction (corresponding to clause (1) in the definition of alternation depth) ξ is atomic, so that we immediately find $d_\eta(\xi) = 0$ as required.

In the induction step of the proof, we make a case distinction as to the last applied clause in the derivation of $\xi \in \Theta^\eta_k$, and we leave the (easy) cases, where this clause was either (2) or (3), for the reader.

Suppose then that $\xi \in \Theta^\eta_k$ on the basis of clause (4). In this case we find that $\xi = \xi'[\psi/z]$ for some formulas ξ', ψ such that ψ is free for z in ξ' and $\xi', \psi \in \Theta^\eta_k$. By the 'key insight' mentioned right after the formulation of the Proposition, any η-chain in the formula ξ is a η-chain in either ξ' or ψ. But then by the induction hypothesis it follows that the length of any such chain must be bounded by k.

Finally, consider the case where $\xi \in \Theta^\eta_k$ on the basis of clause (5). We make a further case distinction. If $\xi \in \Theta^\eta_{k-1}$, then by the induction hypothesis we may conclude that $d_\eta(\xi) \leq k-1$, and from this it is immediate that $d_\eta(\xi) \leq k$. If, on the other hand, $\xi \in \Theta^{\overline{\eta}}_{k-1}$ then the induction hypothesis yields $d_{\overline{\eta}}(\xi) \leq k-1$. But since $d_\eta(\xi) \leq d_{\overline{\eta}}(\xi) + 1$ we obtain $d_\eta(\xi) \leq k$ indeed.

The opposite, right-to-left, implication in (81) is proved by induction on k. In the base step of this induction we have $d_\eta(\xi) = 0$, which means that ξ has no η-variables; from this it is easy to derive that $\xi \in \Theta^\eta_0$.

For the induction step, we assume as our induction hypothesis that (81) holds for $k \in \omega$, and we set out to prove the same statement for $k+1$ and an arbitrary $\eta \in \{\mu, \nu\}$:

$$\text{if } d_\eta(\xi) \leq k+1 \text{ then } \xi \in \Theta^\eta_{k+1}. \tag{82}$$

We will prove (82) by an 'inner' induction on the length of ξ. The base step of this inner induction is easy to deal with: if $|\xi|^\ell = 1$ then ξ must be atomic so that certainly $\xi \in \Theta^\eta_{k+1}$.

In the induction step we are considering a formula ξ with $|\xi|^\ell > 1$. Assume that $d_\eta(\xi) \leq k+1$. We make a case distinction as to the shape of ξ. The only case of interest is where ξ is a fixpoint formula, say, $\xi = \eta x.\chi$ or $\xi = \overline{\eta}x.\chi$. If $\xi = \overline{\eta}x.\chi$, then obviously we have $d_\eta(\xi) = \delta_\eta(\chi)$, so by the inner induction hypothesis we find $\chi \in \Theta^\eta_{k+1}$. From this we immediately derive that $\xi = \overline{\eta}x.\chi \in \Theta^\eta_{k+1}$ as well.

Alternatively, if $\xi = \eta x.\chi$, we split further into cases: If χ has an $\overline{\eta}$-chain $y_1 \cdots y_{k+1}$ of length $k+1$, then obviously we have $x \notin FV(\delta_{k+1})$ (where we write δ_{k+1} instead of $\delta_{y_{k+1}}$), for otherwise we would get $x >_\xi y_{k+1}$, so that we could add x to the $\overline{\eta}$-chain $y_1 \cdots y_{k+1}$ and obtain an η-chain $y_1 \cdots y_{k+1}x$ of length $k+2$. But if $x \notin FV(\delta_{k+1})$ we may take some fresh variable z and write $\xi = \xi'[\overline{\eta}y_{k+1}.\delta_{k+1}/z]$ for some formula ξ' where the formula $\overline{\eta}y_{k+1}.\delta_{k+1}$ is free for z. By our inner induction hypothesis we find that both ξ' and $\eta y_{k+1}.\delta_{k+1}$ belong to Θ^η_{k+1}. But then by clause (4) of Definition 2.50 the formula ξ also belongs to the set Θ^η_{k+1}.

If, on the other hand, χ has *no* $\overline{\eta}$-chain of length $k+1$, then we clearly have $d_{\overline{\eta}}(\chi) \leq k$. Using the outer induction hypothesis we infer $\chi \in \Theta^{\overline{\eta}}_k$, and so by clause (3) of Definition 2.50 we also find $\xi = \eta x.\chi \in \Theta^{\overline{\eta}}_k$. Finally then, clause (5) gives $\xi \in \Theta^\eta_{k+1}$. QED

Turning to the analogous result for tidy formulas, we need some further auxiliary results. The first of these states that, when analysing the alternation depth of a tidy formula of the form $\chi[\xi/x]$, we may without loss of generality assume that ξ is not a free subformula of χ.

Proposition 6.59 *Let ξ and χ be μ-calculus formulas such that ξ is free for x in χ, $x \in FV(\chi)$, $|\xi|^\ell > 1$, and $\chi[\xi/x]$ is tidy. Then there is a tidy formula χ' such that ξ is free for x' in χ', $\chi'[\xi/x'] = \chi[\xi/x]$, $|\chi'|^\ell \leq |\chi|^\ell$, $\mathsf{ad}_\eta(\chi') \leq \mathsf{ad}_\eta(\chi)$ for $\eta \in \{\mu,\nu\}$, and $\xi \not\trianglelefteq_f \chi'$.*

Proof. The proof proceeds by induction on the length of χ. In case $\xi \not\trianglelefteq_f \chi$ we are done immediately, since then we can simply take $\chi' := \chi$. We will therefore assume that $\xi \trianglelefteq_f \chi$. Then by definition we have $\chi = \varphi[\xi/y]$ for some tidy formula φ such that $y \in FV(\chi)$ and ξ is free for y in φ. Since $|\xi|^\ell > 1$ we find that $|\varphi|^\ell < |\chi|^\ell$, and so we may apply the induction hypothesis to φ. That means that we can find a tidy formula φ' such that $\chi = \varphi'[\xi/y']$ for some variable $y \in FV(\varphi')$, ξ is free for y' in φ', $|\varphi'|^\ell \leq |\varphi|^\ell$, $\mathsf{ad}_\eta(\varphi') \leq \mathsf{ad}_\eta(\varphi)$ and $\xi \not\trianglelefteq_f \varphi'$.

Now define $\chi' := \varphi'[x'/x, x'/y']$, for some fresh variable x'. Then clearly we have that $|\chi'|^\ell = |\varphi'|^\ell \leq |\varphi|^\ell < |\chi|^\ell$ and $\mathsf{ad}_\eta(\chi') = \mathsf{ad}_\eta(\varphi') \leq \mathsf{ad}_\eta(\varphi) \leq \mathsf{ad}_\eta(\chi)$, where the last inequality follows from Proposition 6.57 and the fact that $\chi = \varphi[\xi/y]$. We leave it for the reader to convince themselves that $\xi \not\trianglelefteq_f \varphi'$ entails $\xi \not\trianglelefteq_f \chi'$. Finally, it is not hard to verify that $\chi'[\xi/x'] = \varphi'[x'/x, x'/y'][\xi/x'] = \varphi'[\xi/y'][\xi/x] = \chi[\xi/x]$ as required. QED

Our main auxiliary proposition concerns the relation between parity formulas of the form $\mathbb{G}_\chi$ and $\mathbb{G}_{\chi[\xi/x]}$, respectively. Roughly, it states that under some mild conditions, the substitution ξ/x is a 'local isomorphism' between these two structures, i.e., an isomorphism at the level of certain clusters. Recall that $(\!|\psi|\!)_C$ denotes the $\equiv_C$-cluster of a formula ψ.

Proposition 6.60 (substitution as local isomorphism) *Let ξ and χ be formulas such that ξ is free for x in χ and $\chi[\xi/x]$ is tidy. Furthermore, assume that $\xi \notin Cl(\chi)$, and $x \notin FV(\xi)$. Then the following hold:*

1) the substitution $\xi/x : Cl(\chi) \to Cl(\chi[\xi/x])$ is injective;

2) if $\psi \in Cl(\chi)$ is such that $\psi[\xi/x] \notin Cl(\chi) \cup Cl(\xi)$, then ξ/x is an isomorphism between the structures $((\!|\psi|\!)_C, L_C, \to_C, \trianglelefteq_f, \preccurlyeq_C)$ and $((\!|\psi[\xi/x]|\!)_C, L_C, \to_C, \trianglelefteq_f, \preccurlyeq_C)$.

We omit the proof of this Proposition, which is similar to that of Proposition 6.46.

Proposition 6.61 *For any tidy formula ξ and $\eta \in \{\mu,\nu\}$, we have*

$$cd_\eta(\xi) \leq n \text{ iff } \xi \in \Theta^\eta_n. \tag{83}$$

As a corollary, the alternation depth of ξ is equal to the length of its longest alternating $\prec_C$-chain.

Proof. In this proof we shall be reasoning about the clusterwise ordering of fixpoints formulas, and so the following notation will be handy. Given a closure cluster D of ξ we let, for $\eta \in \{\mu,\nu\}$, $cd_\eta(D) := \mathsf{acd}_\eta(\prec_C\restriction_D)$ denote the maximal length of an alternating $\prec_C$-chain *in* D leading up to an η-formula, and similarly we write $cd(D) := \mathsf{acd}(\prec_C\restriction_D)$ and $cd(\xi) := \mathsf{acd}(\prec_C\restriction_\xi)$. Clearly then we have $cd(C) = \max(cd_\mu(C), cd_\nu(C))$, $cd(\xi) = \max(cd_\mu(\xi), cd_\nu(\xi))$, etc.

For the proof of the left-to-right direction of (83), we proceed by an outer induction on n, and an inner induction on the length $|\xi|^\ell$ of the formula ξ. We focus on the outer inductive case, leaving the base case, where $n = 0$, to the reader.

First of all, it is easy to see that every fixpoint formula ξ' in the cluster of ξ satisfies $cd_\eta(\xi') = cd_\eta(\xi)$, while it follows from Proposition 2.54 that $\xi' \in \Theta^\eta_n$ iff $\xi \in \Theta^\eta_n$. For this reason we may, without loss of generality, confine our attention to the case where ξ is the $\prec_C$-maximal element of its cluster. This means that ξ is a fixpoint formula, and so we may distinguish cases, as to its parity.

First we consider the case where ξ is of the form $\xi = \overline{\eta}x.\chi$. Let

$$\eta_1 x_1.\psi_1 \prec_C \eta_2 x_2.\psi_2 \prec_C \cdots \prec_C \eta_k x_k.\psi_k$$

be a maximal alternating η-chain in $Cl(\chi)$. Then

$$\eta_1 x_1.\psi_1[\xi/x] \prec_C \eta_2 x_2.\psi_2[\xi/x] \prec_C \cdots \prec_C \eta_k x_k.\psi_k[\xi/x]$$

is an alternating η-chain in $Cl(\xi)$, and so we have $k \leq n$. It then follows by the inner induction hypothesis that $\chi \in \Theta^\eta_n$, and so by definition of the latter set we find $\xi = \overline{\eta}x.\chi \in \Theta^\eta_n$, as required.

The other case to be discussed is where ξ is of the form $\xi = \eta x.\chi$. Now let

$$\eta_1 x_1.\psi_1 \prec_C \eta_2 x_2.\psi_2 \prec_C \cdots \prec_C \eta_k x_k.\psi_k$$

be a maximal alternating $\overline{\eta}$-chain in $Cl(\chi)$.

We now make a further case distinction. If x is a free variable of some formula in this chain, it is in fact a free variable of every formula in the chain; from this it follows that

$$\eta_1 x_1.\psi_1[\xi/x] \prec_C \eta_2 x_2.\psi_2[\xi/x] \prec_C \cdots \prec_C \eta_k x_k.\psi_k[\xi/x] \prec_C \xi$$

is an alternating η-chain in the cluster $(\!|\xi|\!)_C$ of ξ. Since this chain has length $k+1$, it follows by our assumption on ξ that $k + 1 \leq n$, and so $k \leq n - 1$. Alternatively, if x is not a free variable of any formula in this chain, then the chain is itself an alternating $\overline{\eta}$-chain in $Cl(\xi)$, and from this and the assumption that $cd_\eta(\xi) \leq n$ it readily follows that $k \leq n - 1$.

In both cases we find that $k \leq n - 1$, which means that $cd_{\overline{\eta}}(\chi) \leq n - 1$. By the outer induction hypothesis we thus find that $\chi \in \Theta^{\overline{\eta}}_{n-1}$. From this it is then easy to derive that $\xi = \eta x.\chi \in \Theta^\eta_n$.

For a proof of the opposite, right-to-left direction '$\Leftarrow$' of (83), the argument proceeds by induction on the length of φ. In the base case φ is atomic and hence the claim is trivially true.

In the inductive step we make a case distinction depending on the clause of Definition 2.50 that was applied in the last step of the derivation of $\varphi \in \Theta^\eta_k$. We leave the easy cases, for the clauses 1 and 2, to the reader.

If clause 3 is used to derive $\varphi \in \Theta^\eta_n$ then $\varphi = \overline{\eta}x.\chi$ for some $\chi \in \Theta^\eta_n$. First define $\chi' = \chi[x'/x]$ for an x' that is fresh for χ and φ. Note that the length of χ' is equal to the length of χ, which is shorter than the length of φ. By Proposition 6.57 we also have that

$\chi' \in \Theta^\eta_n$. Moreover, χ' is tidy because φ is tidy, $BV(\chi') = BV(\chi) \subseteq BV(\varphi)$, $FV(\chi') = (FV(\chi) \setminus \{x\}) \cup \{x'\} \subseteq FV(\varphi) \cup \{x'\}$, and x' is fresh for φ. This means that we can apply the inductive hypothesis to χ', obtaining that $cd_\eta(\chi') \leq n$

We then distinguish cases depending on whether $\varphi \in Cl(\chi)$ or not.

If $\varphi \in Cl(\chi)$ then it is not hard to prove that $\varphi \in Cl(\chi')$ as well. It is then easy to see that every alternating chain in $\mathbb{G}_\varphi$ also exists in $\mathbb{G}_{\chi'}$, and thus it follows that $cd_\eta(\varphi) \leq n$.

If $\varphi \notin Cl(\chi)$ we distinguish further cases depending on whether $x \in FV(\chi)$. If this is not the case then $\chi' = \chi$ and $\mathbb{G}_\varphi$ is just like $\mathbb{G}_\chi$ with an additional vertex for φ that forms a transient cluster on its own and is connected just with an outgoing $\to_C$-edge to the vertex of χ' in $\mathbb{G}_{\chi'}$. Thus, every alternating chain in a cluster of $\mathbb{G}_\varphi$ also exists in $\mathbb{G}_{\chi'}$ and thus $cd_\eta(\varphi) \leq n$ follows from $cd_\eta(\chi') \leq n$.

The last case is where $\varphi \notin Cl(\chi)$ and $x \in FV(\chi)$. To prove $cd_\eta(\varphi) \leq n$ consider an alternating $\prec_C$-chain $\eta_1 x_1.\rho_1 \prec_C \cdots \prec_C \eta_m x_m.\rho_m$, of length m and with $\eta_m = \eta$ in some cluster of $\mathbb{G}_\varphi$. We now argue that $m \leq n$. Because $\eta_i x_i.\rho_i \in Cl(\varphi)$ for all $i \in [1,m]$ it follows by Proposition 6.42(4) that the only possibility for φ to be among the $\eta_i x_i.\rho_i$ in this chain is if $\varphi = \eta_m x_m.\rho_m$. This would lead to a contradiction however, because $\eta_m = \eta$ while we assumed that $\varphi = \overline{\eta} x.\chi$. We may therefore conclude that φ is not among the $\eta_i x_i.\rho_i$ for $i \in [1,m]$. By the items 1), 3) and 5) of Proposition 6.46 it follows that there is an alternating $\prec_C$-chain $\eta_1 x_1.\sigma_1 \prec_C \cdots \prec_C \eta_m x_m.\sigma_m$ in $Cl(\chi')$ such that $(\eta_i x_i.\sigma_i)[\xi/x'] = \eta_i x_i.\rho_i$ for all $i \in [1,m]$. Because $cd_\eta(\chi') \leq n$ it follows that $m \leq n$.

If clause 4 is used to derive $\varphi \in \Theta^\eta_n$ then φ is of the form $\varphi = \chi[\xi/x]$ such that $\chi, \xi \in \Theta^\eta_n$. First observe that we may assume that $x \in FV(\chi)$ and $|\xi|^\ell > 1$, otherwise the claim trivialises. Furthermore, because of Proposition 6.59 we may without loss of generality assume that in addition χ is tidy as well, that x is fresh for ξ, and that $\xi \not\trianglelefteq_f \chi$. Finally, since $|\xi|^\ell > 1$ we find that the length of χ is smaller than that of $\varphi = \chi[\xi/x]$, so that we may apply the inductive hypothesis, which gives that $cd_\eta(\chi) \leq n$ and $cd_\eta(\xi) \leq n$.

To show that $cd_\eta(\chi[\xi/x]) \leq n$ consider a fixpoint formula $\eta y.\rho \in Cl(\chi[\xi/x])$ that is at the top of a maximal alternating $\prec_C$-chain in $\mathbb{G}_{\chi[\xi/x]}$. Recall from Definition 6.32 that $h^\downarrow(\eta y.\rho)$ denotes the maximal length of an alternating $\prec_C$-chain leading up to $\eta y.\rho$. In order to show that $h^\downarrow(\eta y.\rho) \leq n$, we claim that

$$h^\downarrow(\eta y.\rho) = h^\downarrow(\eta y.\rho') \text{ for some } \eta y.\rho' \in Cl(\chi) \cup Cl(\xi). \tag{84}$$

To see this, first note that we may assume that $\eta y.\rho \notin Cl(\chi) \cup Cl(\xi)$ because otherwise we can just set $\rho' := \rho$. By Proposition 2.44 we obtain that

$$Cl(\chi[\xi/x]) = \{\psi[\xi/x] \mid \psi \in Cl(\chi)\} \cup Cl(\xi).$$

Therefore, since $\eta y.\rho \in Cl(\chi[\xi/x])$, and we assume that $\eta y.\rho \notin Cl(\xi)$, it follows that $\eta y.\rho = \psi[\xi/x]$ for some $\psi \in Cl(\chi)$. We are thus in a position to apply Proposition 6.60, which describes how the $\to_C$-cluster of ψ relates under the substitution ξ/x to the $\to_C$-cluster of $\eta y.\rho = \psi[\xi/x]$. Note that $\psi \neq x$ because otherwise we would have $\eta y.\rho = \xi$, contradicting the assumption that $\eta x.\rho \notin Cl(\xi)$. This means that $\psi = \eta y.\rho'$ for some formula ρ', since the substitution ξ/x preserves the main connective of formulas other than x. Finally, it follows from item 2) of Proposition 6.60 that $h^\downarrow(\eta y.\rho) = h^\downarrow(\eta y.\rho')$.

As an immediate consequence of (84) we obtain that $h^{\downarrow}(\eta y.\rho') \leq n$ because $\eta y.\rho'$ is either in $\mathbb{G}_{\chi}$ or in $\mathbb{G}_{\xi}$, where the inductive hypothesis applies. This finishes the proof for the case of clause 4.

We leave the last case, where clause 5 is used to derive that $\varphi \in \Theta^{\eta}_{n}$, to the reader. QED

6.6 Guarded transformation

As an example of an important construction on parity formulas, we consider the operation of guarded transformation. Recall from Definition 2.19 that a μ-calculus formula is *guarded* if every occurrence of a bound variable is in the scope of a modal operator which resides inside the variable's defining fixpoint formula. Intuitively, the effect of this condition is that, when evaluating a guarded formula in some model, between any two iterations of the same fixpoint variable, one has make a transition in the model. Many constructions and algorithms operating on μ-calculus formulas presuppose that the input formula is in guarded form, which explains the need for low-cost *guarded transformations*, that is, efficient procedures for bringing a μ-calculus formula into an equivalent guarded form.

It is easy to translate the notion of guardedness to parity formulas, but in fact we will need something stronger in the next chapter, when we present the automata-theoretic perspective on the modal μ-calculus.

Definition 6.62 A path $\pi = v_0 v_1 \cdots v_n$ through a parity formula is *unguarded* if $n \geq 1$, $v_0, v_n \in \mathsf{Dom}(\Omega)$ while there is no i, with $0 < i \leq n$, such that v_i is a modal node. A parity formula is *guarded* if it has no unguarded cycles, and *strongly guarded* if it has no unguarded paths. $\lhd$

In words, a parity formula is strongly guarded if every path, leading from one state (node in $\mathsf{Dom}(\Omega)$) to another contains at least one modal node (occurring after the path's starting state). The following theorem states that on arbitrary parity formulas, we can give an exponential-size guarded transformation; note that the index of the formula does not change. At the time of writing it is not known whether every parity formula can be transformed into a guarded equivalent of *subexponential size.*

Theorem 6.63 *There is an algorithm that transforms a parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ into a strongly guarded parity formula $\mathbb{G}^g$ such that*

1) $\mathbb{G}^g \equiv \mathbb{G}$;
2) $|\mathbb{G}^g| \leq 2^{1+|\mathsf{Dom}(\Omega)|} \cdot |\mathbb{G}|$;
3) $ind(\mathbb{G}^g) \leq ind(\mathbb{G})$;

We will prove Theorem 6.63 via a construction that step by step improves the 'degree of guardedness' of the parity formula. In the intermediate steps we will be dealing with a modified notion of guardedness.

Definition 6.64 A parity formula $\mathbb{G} = (V, E, L, \Omega, v_I)$ is *strongly k-guarded* if it every unguarded path $\pi = v_0 v_1 \cdots v_n$ satisfies $\Omega(v_n) > k$. $\lhd$

Clearly, a parity formula is (strongly) guarded iff it is (strongly) m-guarded, where m is the maximum priority value of the formula. Hence, we may prove Theorem 6.63 by successively applying the following proposition. Recall that a parity formula is called *lean* if its priority map is injective. We say that a parity formula has *silent states only* if each of its states is labelled ε.

Proposition 6.65 *Let $\mathbb{G}$ be a lean, strongly k-guarded parity formula with silent states only. Then we can effectively obtain a lean, $k+1$-guarded parity formula $\mathbb{G}'$, with silent states only, and such that $\mathbb{G}' \equiv \mathbb{G}$, $|\mathbb{G}'| \leq 2 \cdot |\mathbb{G}|$ and $ind(\mathbb{G}') \leq ind(\mathbb{G})$.*

Proof. Let $\mathbb{G} = (V, E, L, \Omega, v_I)$ be an arbitrary lean, strongly k-guarded parity formula with silent states, that is, $\mathsf{Dom}(\Omega) \subseteq L^{-1}(\varepsilon)$. Without loss of generality we may assume that in fact $\mathsf{Dom}(\Omega) = L^{-1}(\varepsilon)$. If $\mathbb{G}$ happens to be already $k+1$-guarded, then there is nothing to do: we may simply define $\mathbb{G}' := \mathbb{G}$.

On the other hand, if $\mathbb{G}$ is $k+1$-unguarded, then in particular there must be a state $z \in V$ such that $\Omega(z) = k+1$. By injectivity of Ω, z is unique with this property. In this case we will build the parity formula $\mathbb{G}'$, roughly, on the disjoint union of $\mathbb{G}$, a copy of a part of $\mathbb{G}$ that is in some sense generated from z, and an additional copy of z itself.

For the definition of $\mathbb{G}'$, let W^z be the smallest set $W \subseteq V$ containing z, which is such that $E[w] \subseteq W$ whenever $w \in W$ is boolean. Now define

$$V' := \big(V \times \{0\}\big) \cup \big(W^z \times \{1\}\big) \cup \big(\{z\} \times \{2\}\big).$$

In the sequel we may write u_0 instead of $(u, 0)$, for brevity. Furthermore, recall that we use V_m to denote the set of *modal* vertices of $\mathbb{G}$. The edge relation E' is now given as follows:

$$\begin{array}{rll} E' := & \big\{(u_0, v_0) \mid (u,v) \in E \text{ and } v \neq z\big\} & \cup \;\; \big\{(u_1, v_1) \mid (u,v) \in E \text{ and } v \neq z\big\} \\ \cup & \big\{(u_0, z_1) \mid (u,z) \in E\big\} & \\ \cup & \big\{(u_1, v_0) \mid (u,v) \in E \text{ and } u \in V_m\big\} & \cup \;\; \big\{(u_1, u_0) \mid u \in \mathsf{Dom}(\Omega) \text{ and } \Omega(u) > k+1\big\} \\ \cup & \big\{(u_1, z_2) \mid (u,z) \in E \text{ and } u \notin V_m\big\} & \end{array}$$

To understand the graph (V', E'), it helps, first of all, to realise that the set W^z provides a subgraph of (V, E), which forms a dag with root z and such that every 'leaf' is either a modal or propositional node, or else a state $v \in \mathsf{Dom}(\Omega)$ with $\Omega(v) > k$. (It cannot be the case that $\Omega(v) \leq k$ due to the assumed k-guardedness of $\mathbb{G}$.) Second, it is important to realise that the only way to move from the V-part of V' to the W^z-part is via the root z_1 of the W^z-part, while the only way to move in the converse direction is either directly following a modal node, or else by making a dummy transition from some vertex u_1 to its counterpart u_0 for any $u \in W^z$ with $\Omega(u) > k$. Finally, we add a single vertex z_2 to V'.

Furthermore, we define the labelling L' and the priority map Ω' of $\mathbb{G}'$ by putting

$$L'(u_i) := \begin{cases} L(u) & \text{if } i = 0, 1 \\ \widehat{z} & \text{if } u_i = z_2 \end{cases}$$

where we recall that $\widehat{z} = \bot$ if $\Omega(z)$ is odd and $\widehat{z} = \top$ if $\Omega(z)$ is even, and

$$\Omega'(u_i) := \begin{cases} \Omega(u) & \text{if } i = 0 \text{ and } u \in \mathsf{Dom}(\Omega) \\ \uparrow & \text{otherwise.} \end{cases}$$

In words, the label of a node (v, i) in $\mathbb{G}'$ is identical to the one of v in $\mathbb{G}$, with the sole exception of the vertex $(z, 2)$. To explain the label of the latter node, note that by construction, any unguarded E'-path from z_1 to z_2 projects to an unguarded $k+1$-cycle from

z to z in $\mathbb{G}$. If $\Omega(z) = k+1$ is odd, any such cycle represents (tails of) infinite matches that are lost by $\exists$; for this reason we may label the 'second' appearance of z in the E'-path, i.e., as the node z_2, with $\bot$.

We now turn to the proof of the proposition. It is not hard to show that $\mathbb{G}'$ is lean and that $|\mathbb{G}'| \leq 2 \cdot |\mathbb{G}|$.

To show that $ind(\mathbb{G}') \leq ind(\mathbb{G})$, note that obviously, the projection map $u_i \mapsto u$ preserves the cluster equivalence relation, i.e., $u_i \equiv_{E'} v_j$ implies $u \equiv_E v$. Hence, the image of any cluster C' of $\mathbb{G}'$ under this projection is part of some cluster C of $\mathbb{G}$. But then by definition of Ω' it is easy to see that $ind(C') \leq ind(C)$. From this it is immediate that $ind(\mathbb{G}') \leq ind(\mathbb{G})$.

To see why $\mathbb{G}'$ is $k+1$-guarded, suppose for contradiction that it has a $k+1$-unguarded path $\pi = (v_0, i_0)(v_1, i_1)\cdots(v_n, i_n)$. It is easy to see that this implies that the *projection* $v_0 v_1 \cdots v_n$ of π is an unguarded path in $\mathbb{G}$ (here we ignore possible dummy transitions of the form (u_1, u_0)), and so by assumption on $\mathbb{G}$ it must be the case that $\Omega'(v_n, i_n) = \Omega(v_n) = k+1$. This means that $(v_n, i_n) = (z, 0)$; but the only way to arrive at the node $(z, 0)$ in (V', E') is directly following a modal node (in $W^z \times \{1\}$), which contradicts the unguardedness of the path π.

In order to finish the proof of the Proposition, we need to prove the equivalence of $\mathbb{G}'$ and $\mathbb{G}$; but this can be established by a relatively routine argument of which we skip the details. QED

Proof of Theorem 6.63. Let $\mathbb{G}$ be an arbitrary parity formula; without loss of generality we may assume that $\mathbb{G}$ is lean, i.e., Ω is injective. Let $\mathsf{Ran}(\Omega) = \{k_1, \ldots, k_n\}$; then $|\mathsf{Dom}(\Omega)| = n$. To ensure that all states are silent, we may have to duplicate some vertices; that is, we continue with a version $\mathbb{H}$ of $\mathbb{G}$ that has at most twice as many vertices, but the same index, the same number of states, and silent state only.

By a straightforward induction we apply Proposition 6.65 to construct, for every $i \in [1, n]$, a k_i-guarded parity automaton $\mathbb{H}^i$ with silent states only, and such that $\mathbb{H}^i \equiv \mathbb{G}$, $|\mathbb{H}^i| \leq 2^{i+1} \cdot |\mathbb{G}|$, and $ind(\mathbb{H}^i) = ind(\mathbb{G})$. Clearly then we find that $\mathbb{H}^n$ is the desired strongly guarded equivalent of $\mathbb{G}$; and since $n = |\mathsf{Dom}(\Omega)|$ we find that $|\mathbb{H}^n| \leq 2^{1+n} \cdot |\mathbb{G}|$ as required. QED

Remark 6.66 On a closer inspection of the construction in the proof of Proposition 6.65, the reader may observe that inductively, we may assume that for every i, every predecessor of a state in $\mathbb{H}^i$ with priority at most k_i is in fact a modal node. From this, it follows that we may impose, in the formulation of Theorem 6.63, an additional constraint on $\mathbb{G}^g$, namely, that every predecessor of a state is a modal node, more formally, that $(E^g)^{-1}[\mathsf{Dom}(\Omega] \subseteq V^g_m$. $\lhd$