State Space Reduction For Parity Automata

Christof Löding and Andreas Tollkötter

RWTH Aachen, Lehrstuhl 7 für Informatik

Abstract. Exact minimization of ω-automata is a difficult problem and heuristic algorithms are a subject of current research. We establish a framework to generalize the known notion of quotient automata and uniformly describe such algorithms. We investigate several approaches to reduce the state space of deterministic parity automata. These are based on extracting information from structures within the automaton, such as strongly connected components, coloring of the states, and equivalence classes of given relations, to determine states that can safely be merged. The description of these procedures consists of a theoretical analysis as well as data collected from experiments.

Keywords: foo · bar

1 Introduction

Finite automata are a long established computation model that dates back to sources such as [9] and [12]. A known problem for finite automata is state space reduction, referring to the search of a language-equivalent automaton which uses fewer states than the original object. For deterministic finite automata (DFA), not just reduction but minimization was solved in [6]. Regarding nondeterministic finite automata (NFA), [7] proved the PSPACE-completeness of the minimization problem, which is why reduction algorithms such as [3] and [1] are a popular alternative.

In his prominent work [2], Büchi introduced the model of Büchi automata (BA) as an extension of finite automata to read words of one-sided infinite length. As these ω-automata tend to have higher levels of complexity in comparison to standard finite automata, the potential gain of state space reduction is even greater. Similar to NFAs, exact minimization for deterministic Büchi automata was shown to be NP-complete in [13] and spawned heuristic approaches such as [13], [8], or [4].

As [14] displays, deterministic Büchi automata are a strictly weaker model than nondeterministic Büchi automata. It is therefore interesting to consider different models of ω-automata in which determinism is possible while maintaining enough power to describe all ω-regular languages. Parity automata (PA) are one such model, a combination of Büchi automata and Moore automata ([10]), that use a parity function, assigning a number (called priority or color) to each state. The convention used in this paper is that even priorities correspond to "good" states while odd priorities correspond to "bad" states. The smallest priority that is seen infinitely often during a run defines its acceptance; if it is even, the run will be accepting, and if it is odd, the run will be rejecting. [11], [15] showed that deterministic parity automata are indeed sufficient to recognize all ω-regular languages. As for DBAs, the exact minimization problem for DPAs is NP-complete ([13]).

In this paper, we investigate multiple techniques of heuristic state space reduction for DPAs. We present a general framework to uniformly describe these algorithms in a transparent manner. As opposed to a black box which provides as output a reduced automaton for each input, the framework aims towards giving more precise information about how the reduced DPA comes to

be. This in theory makes it easier to find structures in the DPAs which are responsible for the reduction being applicable. The base of the framework was the notion of quotient automata, which, for a given congruence relation on the states, merges all states of each equivalence class into one single representative. For example, building the quotient automaton of a DFA with the relation defined by the Myhill-Nerode theorem [] will yield the minimal DFA for that language.

Our proposal for this framework are merger functions which generalize quotient automata. Merger functions map sets of states in the original DPA, called the merge set, to other sets of states, called the candidate set. The easiest interpretation, which we refer to as representative merge, allows us to merge all states from the merge set into any single representative that is chosen from the candidate set.

The most basic merge simply adapts the algorithm from [6]. Every parity automaton can be interpreted as a Moore automaton which can then be minimized using said algorithm. In this context, we call the equivalence relation which considers two states to be equivalent if they are merged by this algorithm the *Moore equivalence*.

A simple merger function that is introduced is the *skip merger* which takes an equivalence relation that implies language equivalence on the states as a parameter and from that builds a merger function. The idea of using one given equivalence relation on the states and refining it is used several times in the paper. This merge decides on one particular strongly connected component (SCC) of the automaton and removes all states of an equivalence class that do not lie in this SCC, essentially "skipping" all other SCCs.

We adapt the works of [5], who considered alternating parity automata, to our case of deterministic parity automata and find that there are sufficient differences to warrant a separate analysis. The *delayed simulation merger* considers two states to be equal, if on every run from those states on a shared word, if one run visits a priority at some position, the other run must visit a priority at most as high at some point in the future. It then holds that both runs will see the same smallest priority infinitely often and therefore either both accept or both reject. It suffices to choose one representative of each such equivalence class that has minimal priority and build the quotient automaton with those representatives.

We bring up merging via *path refinement*, which uses an existing equivalence class of some congruence relation and refines it to a point where states can safely be merged. This refinement occurs so that two states from the class are equivalent if on each path from that state back to the class, both states visit the same minimal priority.

Another merger function is the *threshold Moore merger*, which again refines an existing relation. Two states are considered to be equivalent under the refinement if a relaxation of Moore equivalence, that considers all priorities greater than some k to be equal, matches them. In this case we require the two states to have equal priority and choose k to be that exact value.

A similar but slightly different approach is the *LSF* merger function. It removes the need for two states to have equal priority and instead takes the value k as a parameter. On the other side it adds a requirement similar to that of the skip merger, in that the candidates of the merge are those that all lie in one single SCC if we modify the automaton to only contain those states of priority at least k.

2 Merger Functions

Our definition of DPAs is a quadruple (Q, Σ, δ, c), where Q is the finite set of states, Σ is the finite set of symbols (the alphabet), $\delta : Q \times \Sigma \to Q$ is the deterministic transition function, and $c : Q \to \mathbb{N}$ is the priority function. The set of ω-words over Σ is then denoted by Σ^ω.

Definition 1. *Let $\mu : D \to (2^Q \setminus \{\emptyset\})$ be a function for some $D \subseteq 2^Q$. We call μ a* merger function *if all sets in D are pairwise disjoint and for all sets $M \in D$, $\mu(M) \cap (\bigcup D \setminus M) = \emptyset$.*

A representative merge *of $\mathcal{A}$ w.r.t. μ is constructed by choosing a representative $r_M \in \mu(M)$ for all $M \in D$ and then removing all states in $M \setminus \{r_M\}$. Transitions that originally lead to one of the removed states are redirected to the representative r_M instead.*

While merger functions are a generalization of the more restrictive combination of congruence relation and quotient automaton, we still often build up our various merger functions from the basis of an equivalence relation. We briefly go over a few cases of relations that are of interest in this context before moving on to the first reduction technique.

We consider several types of different relations, mostly over the state domain Q. A relation R is a preorder if it is reflexive and transitive. R is an equivalence relation if it is a symmetric preorder. R is a congruence relation if it is an equivalence relation that is compatible with δ, i.e. if $(p, q) \in R$, then also $(\delta(p, a), \delta(q, a)) \in R$ for all $a \in \Sigma$.

If $\sim$ is an equivalence relation and $\mathcal{A}$ is a DPA, we write $\mathfrak{C}(\sim) \subseteq 2^Q$ for the set of equivalence classes in $\mathcal{A}$.

Definition 2. *The* language equivalence relation *is defined by $p \equiv_L q$ iff reading every ω-word α from either p or q gives the same acceptance.*

The priority almost equivalence relation *is defined by $p \equiv_\dagger q$ iff reading every ω-word α from either p or q yields two runs that differ in priorities at only finitely many positions.*

The Moore equivalence relation *is defined by $p \equiv_M q$ iff reading every finite word w from either p or q ends up in states with the same priority.*

All three of these equivalence relations imply language equivalence between states. However, only Moore equivalence is strong enough of a contract to allow for immediate merging of states. Merging states according to $\equiv_\dagger$ or $\equiv_L$ can change the language of the DPA. The idea of using already defined relations that imply language equivalence but are not strong enough to allow merging on their own and then refining those relations into finer equivalence classes can be seen in multiple of our techniques, such as the skip merger, path refinement, and LSF merger.

In fact, building the quotient automaton w.r.t. $\equiv_M$ is the canonical way to minimize a deterministic Moore automaton. We can express the same by a merger function to reduce the state space of a DPA.

Definition 3. *The* Moore merger function *is defined as $\mu_M : \mathfrak{C}(\equiv_M) \to 2^Q$ with $\mu_M(\kappa) = \kappa$.*

Theorem 1. *A representative merge of a DPA w.r.t. μ_M is language equivalent to the original.*

3 Schewe Merge

While the focus of the reduction analysis lies on the different techniques to generate merger functions, one can also try to improve the actual merging step performed. The representative merge

defined above is the most intuitive and direct option when considering how to apply a merger function to a given DPA but we give another example in this section that potentially allows for better overall reduction.

Based on [13], the Schewe merge works rather similar to the representative merge. In addition to merging states from the merge sets into the chosen representative, it also redirects some transitions to the candidate set. While this does not remove additional states on its own, it simplifies the structure of the automaton to potentially improve the reduction of further reduction algorithms that are applied after the first.

Definition 4. *Let μ be a merger function. A* Schewe merge *of a DPA $\mathcal{A}$ w.r.t. μ is constructed first by building a representative merge. Then, for all merge sets M in μ and all transitions $\delta(p, a) = q$ in the original automaton, if $q \in \mu(M)$ and p is not reachable from q, then the transition is redirected to r_M instead.*

A Schewe merge differs from the representative merge if there are more than one state from the candidate set remaining and the states in a class are distributed over multiple SCCs. Whenever a transition would move the automaton to a candidate while changing SCC at the same time, that transition is instead redirected to the chosen representative state. One can imagine that, for example, this potentially enhances the reduction of a consecutive Moore merger, as more states now uniformly target the same representative.

It is not obvious if one can simply replace the representative merge with the Schewe merge and still keep the same properties such as preservation of language. We can classify a set of requirements that merger functions have to satisfy to be compatible to the Schewe merge.

Definition 5. *For a representative merge $\mathcal{A}'$ of $\mathcal{A}$ w.r.t. μ, we define the candidate relation $\sim_{\mathcal{C}}^{\mu}$ as $p \sim_{\mathcal{C}}^{\mu} q$ if and only if there is a $C \in \mu(D)$ with $p, q \in C$.*

We call μ Schewe suitable *if for all representative merges $\mathcal{A}'$, $\mu_{\mathcal{C}}^{\mu}$ is a congruence relation, it implies language equivalence, and the reachability order restricted to any equivalence class of $\mu_{\mathcal{C}}^{\mu}$ is symmetric.*

Theorem 2. *Let μ be a Schewe suitable merger function. For a DPA $\mathcal{A}$, if a representative merge and a Schewe merge of $\mathcal{A}$ are built with the same choices for the representative states, then these two merge DPAs are language equivalent.*

Proof. Let $\mathcal{A}'$ be the representative merge and $\mathcal{A}''$ be the Schewe merge. Let q_0 be some starting state for the three runs ρ, ρ', ρ'' of the three automata on some word α. We claim that ρ' and ρ'' have the same acceptance status.

Let K be the set of positions where ρ'' uses a transition that does not exist in ρ'. We can observe that for every equivalence class κ of $\sim_{\mathcal{C}}^{\mu}$, there is at most one k_κ in K. If there would be two such positions k_κ and l_κ, then $\rho''(l_\kappa - 1)$ would be reachable from $\rho''(k_\kappa)$ which contradicts the requirement for the redirection of that edge in the first place.

As $K = \{k_1, \ldots, k_n\}$ is finite, ρ'' eventually only uses transitions that are also present in ρ'. By induction on i, we can show that $\rho'(k_i + 1) \sim_{\mathcal{C}}^{\mu} \rho''(k_i + 1)$, in particular for $i = n$. As $\sim_{\mathcal{C}}^{\mu}$ implies language equivalence by assumption, that means ρ' and ρ'' must have the same acceptance status.

4 Skip Merger

The skip merger is the first reduction algorithm we want to introduce. It can be seen more of a first proof of concept rather than a practical novelty, as the underlying idea is neither complex nor had it great effect during empirical tests.

The skip merger considers the different strongly connected components (SCCs) of the automaton. An SCC is a set of states in which each element can reach every other via some path. We sometimes speak of the "deepest" SCCs with a certain property, which are those SCCs such that no other SCC with that property is reachable anymore.

If we have an equivalence relation that implies language equivalence but is not strong enough to warrant a merge of states on its own, each equivalence class of that relation only requires its states in the deepest SCC. We can therefore redirect any transition that would move the automaton to a state to a representative in a deepest SCC instead.

To formally capture this idea, the *reachability preorder* is defined as $p \preceq_{\text{reach}} q$ if and only if q is reachable by some path from p. Reachability can be computed in $\mathcal{O}(|Q|^3)$ which, in general, is too high of a complexity to efficiently deal with. It is sufficient to use a total extension of reachability though, which is a minimal superset of the reachability preorder that is a total preorder itself. Such a relation can be computed in linear time by a topological sorting on the SCCs of the automaton.

Definition 6. *Let $\sim$ be a congruence relation on Q that implies language equivalence. Let $\preceq$ be a total extension of reachability. For each $\kappa \in \mathfrak{C}$, we define $C_\kappa \subseteq \kappa$ to be the set of $\preceq$-maximal states and $M_\kappa = \kappa \setminus C_\kappa$.*

We define the skip merger function $\mu^{\sim}_{skip} : \{M_\kappa \mid \kappa \in \mathfrak{C}(\sim)\} \to 2^Q$ *with* $\mu^{\sim}_{skip}(M_\kappa) = C_\kappa$.

Theorem 3. *A representative merge of a DPA w.r.t. $\mu^{\sim}_{skip}$ is language equivalent to the original.*

Proof. If we have two runs, π and ρ, of the original DPA and the merged DPA on the same word α, we can observe that at every position, $\pi(i)$ and $\rho(i)$ will be $\sim$-equivalent, as $\sim$ is a congruence relation. Furthermore, as there are only finitely many SCCs in an automaton, ρ will eventually reach a point j from which on only transitions are taken that also exist in the original DPA. As $\pi(j) \sim \rho(j)$, that means that the two runs must have the same status of acceptance. □

Theorem 4. *For a given $\sim$ in a suitable data structure, $\mu^{\sim}_{skip}$ can be computed in $\mathcal{O}(|Q|)$.*

5 Delayed Simulation

Based on [5], we consider reduction via delayed simulation. The original paper uses alternating parity automata, a much stronger model compared to our DPAs. Limiting ourselves to this special case not only reduces the run time of the algorithm but allows for a fundamentally better approach to improve the amount of reduction.

Similarly to other simulation techniques for reduction, the concept of delayed simulation is to consider two states to be mergeable if both "simulate" the behavior of the respective other. Delayed simulation in particular considers simulation in the sense that on every path, both states eventually see the same smallest priority.

Definition 7. *We define the* delayed simulation equivalence relation *as $p \equiv_{de} q$ if and only if the following property holds for all $w \in \Sigma^*$: Let $p' = \delta^*(p, w)$ and $q' = \delta^*(q, w)$. Every run in the automaton that starts in p' or q' eventually sees a priority less than or equal to* $\min\{c(p'), c(q')\}$.

$\equiv_{\text{de}}$ is a congruence relation that implies language equivalence but states that are $\equiv_{\text{de}}$-equivalent do in general not have the same priority. It is therefore not trivial to see that equivalent states can be merged. In fact, the merger function is not as simple as a quotient automaton. Rather, we have to add the additional requirement for the chosen representative state to have minimal priority.

Definition 8. *We define the* delayed simulation merger function $\mu_{de} : \mathfrak{C}(\equiv_{de}) \to 2^Q$ *with* $\mu_{de}(\kappa) = \{q \in \kappa \mid c(q) = \min c(\kappa)\}$.

Theorem 5. *A representative merge of a DPA w.r.t.* μ_{de} *is language equivalent to the original.*

Proof. The main idea of the proof is to show that for two states $p \equiv_{\text{de}} q$ with $c(p) < c(q)$, we can set the priority of q to $c(p)$ without changing the language of the automaton. Then we can build an equivalent DPA to the original in which $\equiv_{\text{de}}$ implies priority equivalence, which shows that the quotient automaton of that DPA is language equivalent to the original. That quotient automaton is equivalent to our merger function μ_{de}.

To see that the priority of q can be changed, assume there is a run ρ of $\mathcal{A}$ on some ω-word α such that changing the priority of q to $c(p)$ also changes the acceptance status of ρ. As the argument works symmetrically, we can assume that $c(p)$ is even. We call the DPA with modified priorities $\mathcal{A}'$ with function c'. As the priority of every state in $\mathcal{A}$ is at least as good as in $\mathcal{A}'$, we can assume that $c(\rho)$ is accepting and $c'(\rho)$ is rejecting.

ρ must visit q infinitely often and in $c'(\rho)$, $c'(q)$ must be the smallest priority. Otherwise, the two runs would have the same smallest priority that occurs infinitely often. Hence, there is a word w such that reading w from q moves back to q and only priorities greater than $c'(q)$ are seen in between.

Now consider the two runs π_p and π_q on the word w^ω starting in p and q respectively. By choice of w, $c(\pi_q)$ only visits priorities strictly greater than $c'(q)$. On the other hand, π_p starts at state p with $c(p) = c'(q)$. As $p \equiv_{\text{de}} q$, $c(\pi_q)$ must see a priority of at most $c(p)$ on all paths eventually. This is a contradiction and the run ρ cannot exist. □

Computation of delayed simulation is not trivial, especially if it is to be done efficiently. Taken from the original paper [5], our approach is to use a Büchi automaton and reduce $\equiv_{\text{de}}$ in the original DPA to language acceptance in that automaton.

The Büchi automaton, called $\mathcal{G}_{\text{de}}$, is a product automaton with an additional component that it uses to track the visited priorities. This third component, called "obligations", makes sure that whenever a priority occurs in the first component, the second component eventually has to simulate one at most as large. For each state pair p, q we can then use $\mathcal{G}_{\text{de}}$ to determine whether q simulates p. If this holds in both directions, the two states are $\equiv_{\text{de}}$-equivalent.

Definition 9. *We define the deterministic Büchi automaton* $\mathcal{G}_{de} = (Q_{de}, \Sigma, \delta_{de}, F_{de})$ *as*

- $Q_{de} = Q \times Q \times (c(Q) \cup \{\checkmark\})$
- $\delta_{de}((p, q, k), a) = (p', q', \gamma(c(p'), c(q'), k))$, *where* $p' = \delta(p, a)$ *and* $q' = \delta(q, a)$
- $F_{de} = Q \times Q \times \{\checkmark\}$.

The obligation function γ *is defined as*

$$\gamma(i, j, k) = \begin{cases} \checkmark & \text{if } j \leq i \text{ and } j \leq_\checkmark k \\ \min_{\leq_\checkmark}\{i, k\} & \text{else} \end{cases}$$

where $0 \leq_\checkmark 1 \leq_\checkmark 2 \leq_\checkmark \cdots \leq_\checkmark \checkmark$.

Now using this automaton, we can relate delayed simulation to the question of universal language. A state is *language universal* if starting from it, every path is accepted.

Theorem 6. *For two states p and q, let $q^0_{de}(p,q) = (p, q, \gamma(c(p), c(q), \checkmark))$. Then $p \equiv_{de} q$ if and only if $q^0_{de}(p,q)$ and $q^0_{de}(q,p)$ are language universal states.*

The proof of this theorem is just a technical comparison between the theoretical definition of $\equiv_{\mathrm{de}}$ and the algorithmic definition of $\mathcal{G}_{\mathrm{de}}$. We omit it, as no particular insight is gained.

Theorem 7. *μ_{de} can be computed in $\mathcal{O}(|Q|^2 \cdot |c(Q)|)$.*

Proof. Assuming that we can compute $\equiv_{\mathrm{de}}$ in a suitable data structure in the described time, building μ_{de} from that is rather trivial. To see how we compute $\equiv_{\mathrm{de}}$, observe that the size of $\mathcal{G}_{\mathrm{de}}$ is $\mathcal{O}(|Q|^2 \cdot |c(Q)|)$. The set of language universal states in a DBA can be computed in linear time: we are looking for loops in the subgraph that only consists of the non-accepting states. Then every state from which such a loop is reachable is not language universal. These operations can all be done with classic graph operations such as depth first search. □

6 Path Refinement

The upcoming technique, called congruence path refinement or just path refinement, is again one which takes in an already defined relation on the states and refines it to a point where merging classes is a valid operation. To be precise, we define the PR-equivalence for each equivalence class independently. Given such a class, which we call λ, we consider "loops" which move from λ back to some state in λ via some finite path. If two states in λ see the same smallest priority on every such loop and we can guarantee that all states they will reach via the loop have the same property, then every run that visits λ infinitely often can be divided into path segments on which runs starting in these two states see the same smallest priority. As this then holds for every segment in the run, the set of infinitely often occurring priorities must then also have the same minimum.

Formally, we describe these loops in a set called $L_{\lambda\hookleftarrow}$. This is the set of all finite words, such that reading them from λ will move back to λ at the end and never before.

Definition 10. *Let $\sim$ be a congruence relation that implies language equivalence and let $\lambda \in \mathfrak{C}(\sim)$ be an equivalence class. We define a relation R_λ on λ as $(p,q) \in R_\lambda$ if and only if for all $w \in L_{\lambda\hookleftarrow}$, the smallest priority seen on the path induced by w is the same starting from p and from q.*

We define path refinement equivalence *$\equiv^\lambda_{PR}$ on λ as the largest subset of R_λ such that $p \equiv^\lambda_{PR} q$ if and only if for all $w \in L_{\lambda\hookleftarrow}$, $\delta^*(p,w) \equiv^\lambda_{PR} \delta^*(q,w)$.*

Again, note that this relation is only defined on λ. Using the merger function below, one can perform this reduction independently on every class in $\mathfrak{C}(\sim)$.

Definition 11. *We define the* path refinement merger function *$\mu^\lambda_{PR} : \mathfrak{C}(\equiv^\lambda_{PR}) \to 2^Q$ with $\mu^\lambda_{PR}(\kappa) = \{q \in \kappa \mid c(q) = \min c(\kappa)\}$.*

Theorem 8. *A representative merge of a DPA w.r.t. μ^λ_{PR} is language equivalent to the original.*

Proof. Let $\mathcal{A}'$ be the representative merge. Assume there is a starting state $q_0 \in Q'$ and a word α such that the acceptance of the runs ρ and ρ' of the two DPAs differs. We will bring this assumption to a contradiction.

First, note that at every position i, $\rho(i)$ and $\rho'(i)$ must be $\sim$-equivalent, as $\sim$ is a congruence relation. If in these runs, λ is visited only finitely often, there is a position j at which it is visited for the last time. Then from j on, ρ' only uses transitions that also exist in the original DPA $\mathcal{A}$. As $\rho(j) \sim \rho'(j)$, they must be language equivalent and therefore have the same acceptance status. This contradicts the assumption.

Otherwise, λ is visited infinitely often. However, for two consecutive positions k and k' at which λ is seen, we can show that the smallest priorities in $c(\rho(k)), \ldots, c(\rho(k'))$ and $c'(\rho'(k)), \ldots, c'(\rho'(k'))$ are the same. Then it easily extends that the entire runs share the same smallest priority that is seen infinitely often.

To observe that the two run segments see the same minimal priority, first observe that $\rho(k) \equiv^{\lambda}_{\mathrm{PR}} \rho'(k)$ by induction on k. If k is the first position at which λ is visited, then $\rho'(k)$ is the representative of the equivalence class of $\rho(k)$ and therefore $\equiv^{\lambda}_{\mathrm{PR}}$-equivalent to $\rho(k)$. Then by definition of the path refinement equivalence, the same holds for k' and therefore all following positions.

Now that we have established $\rho(k) \equiv^{\lambda}_{\mathrm{PR}} \rho'(k)$, it follows directly from the definition of PR-equivalence that the smallest priorities in $c(\rho(k)), \ldots, c(\rho(k'))$ and $c'(\rho'(k)), \ldots, c'(\rho'(k'))$ are equal. □

As for delayed simulation, the definition of path refinement is more theoretical and less constructive. We have to dedicate some additional thought to the question of how to actually compute $\mu^{\lambda}_{\mathrm{PR}}$. We use a similar approach as before and reduce the computation of PR-equivalence to a known automata problem, which in this case is the Moore equivalence on DPAs.

A direct translation of the definition to an algorithm would be a similar automaton as for delayed simulation. One can build a deterministic finite product automaton with an additional third component that tracks the smallest priority so far and which component it was seen in, and at every visit to λ makes sure that the tracked values coincide.

The *visit graph* that we now define instead is a less intuitive solution but has a size only linear in $|Q|$ instead of quadratic. It also uses a "tracker" of the smallest priorities between one visit to λ and the next but only does so for each state individually instead of tracking each state pair.

Definition 12. *The visit graph is a DPA* $(Q^{\lambda}_{visit}, \Sigma, \delta^{\lambda}_{visit}, c^{\lambda}_{visit})$ *defined by*

- $Q^{\lambda}_{visit} = Q \times c(Q) \times (c(Q) \cup \{\bot\})$
- $delta^{\lambda}_{visit}((q,k,k'),a) = \begin{cases} (q', \min\{c(q'),k\}, \bot) & \text{if } q' \notin \lambda \\ (q', c(q'), \min\{c(q'),k\}) & \text{if } q' \in \lambda \end{cases}$, *where* $q' = \delta(q,a)$
- $c_{visit}((q,k,k')) = k'$

Theorem 9. *For a state* $q \in Q$*, let* $\iota_q = (q, c(q), \max c(Q)) \in Q^{\lambda}_{visit}$*. Then* $p \equiv^{\lambda}_{PR} q$ *if and only if* $\iota_p \equiv_M \iota_q$*.*

Proof. Our first observation is that for any state $p \in \lambda$, reading some $w \in L_{\lambda\hookleftarrow}$ from $(p, c(p), k)$ ends in $(q, c(q), k')$ where k' is the smallest priority that occurs on the run segment.

If $p \not\equiv^{\lambda}_{\mathrm{PR}} q$, then there is a $w \in L_{\lambda\hookleftarrow}$ such that either the smallest priority when reading w from p and q differs, or reading w moves to non-PR-equivalent states. If the former is true, then reading w from ι_p and ι_q brings the visit graph to state with different priorities and therefore $\iota_p \not\equiv_M \iota_q$. If

the former is false and the latter is true, then one has to repeatedly apply this argument until at some point a state pair is reached at which the first case is violated.

For the other direction, if $\iota_p \not\equiv_M \iota_q$, there must be a $w \in \Sigma^*$ such that the priority differs when reading w from ι_p and ι_q. As all states not in λ have the same priority, we can split $w = v_1 \dots v_n$ such that all v_i are words in $L_{\lambda\leftarrow}$. Then on the last segment, reading v_n sees different minimal priorities from the initial states and therefore p and q cannot be PR-equivalent.

Theorem 10. *$\equiv^{\lambda}_{PR}$ can be computed in $\mathcal{O}(|Q| \cdot |c(Q)|^2 \cdot \log |Q|)$.*

Proof. Moore equivalence can be computed in time $\mathcal{O}(n \log n)$ of its automaton ([6]). The visit graph has size $n \in \mathcal{O}(|Q| \cdot |c(Q)|^2)$. As $|c(Q)|$ is always at most $|Q|$, this gives us the desired complexity.

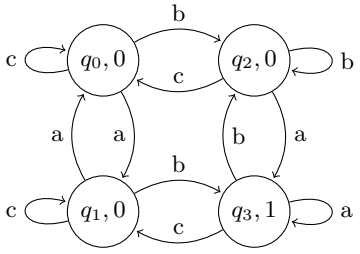

Fig. 1: Example PR

7 Threshold Moore

The following approach is based on the idea of intersecting two equivalence relations which by themselves are too weak to allow for merging of states, but together then give a strong enough guarantee of similarity between the states. One of the two relations is one that implies language equivalence, similar to the previous path refinement algorithm. The second relation is a modification of Moore equivalence that only considers state priorities up to a certain upper limit.

Definition 13. *We define the* threshold Moore equivalence relation *as $p \equiv^{\leq k}_{M} q$ if and only if for all finite words w, $\delta^*(p, w)$ and $\delta^*(q, w)$ have the same priority or both priorities are greater than k.*

Let $\sim$ be an equivalence relation that implies language equivalence. We define the TM equivalence relation *as $p \equiv^{\sim}_{TM} q$ if and only if $p \sim q$, $c(p) = c(q)$, and $p \equiv^{\leq c(p)}_{M} q$.*

We make the implicit assumption here that the relation $\sim$ is well-defined overspanning different automata: for two DPAs $\mathcal{A}$ and $\mathcal{A}'$ so that Q' is a subset of Q and for all states $q \in Q'$, both automata accept the same language starting from q, the relation $\sim$ considers two states to be equivalent in one automaton if and only if it does so in the other.

This assumption makes sure that if we merge some states while preserving language in the automaton, the $\sim$ relation on the untouched states does not change.

For the TM relation, the merger function is actually quite simple and simply merges classes of $\equiv^{\sim}_{\mathrm{TM}}$. Note that this is not, however, a quotient automaton, as the TM relation is in general not a congruence relation.

Definition 14. *We define the* TM merger function $\mu^{\sim}_{TM} : \mathfrak{C}(\equiv^{\sim}_{TM}) \to 2^Q$ *as* $\mu^{\sim}_{TM}(\kappa) = \kappa$.

Lemma 1. *Let $\mathcal{A}$ be a DPA and let $\mathcal{A}'$ be a representative merge w.r.t. some equivalence class $\kappa \in \mathfrak{C}(\mu^{\sim}_{TM})$. Let k be the priority of all states in κ. For all states p and q in $\mathcal{A}'$, the two states are $\equiv_L$-equivalent in $\mathcal{A}$ if and only if they are in $\mathcal{A}'$. If $k \geq c(p), c(q)$, then the same holds for $\equiv^{\leq k}_M$.*

Proof. We focus on the language equivalence first. Let ρ and ρ' be the runs of the two automata on some word α starting in some state q_0. We show that these two runs have the same acceptance status. Then the claim of the Lemma follows by transitivity of $\equiv_L$.

$\equiv_L$ is a congruence relation, so we have $\rho(i) \equiv_L \rho'(i)$ for all positions i. The same is true for $\equiv^{\leq k}_M$.

If $c(\rho)$ visits infinitely priorities of at most k, then the two runs will see the same smallest priority $l < k$ infinitely often, as $c(\rho(i)) = l$ if and only if $c'(\rho'(i)) = l$. Thus, they must have the same acceptance status.

If $c(\rho)$ only visits finitely many priorities of at most k, then from some point j on in ρ', only transitions that also exist in $\mathcal{A}$ are taken. As $\rho(j) \equiv_L \rho'(j)$, that means the two runs have the same acceptance status.

Regarding the second part of the Lemma, let q be a state with $k \geq c(q)$ and let ρ and ρ' be the runs of the two automata on some α starting in q. As $\equiv^{\leq k}_M$ is a congruence relation, $\rho(i) \equiv^{\leq k}_M \rho'(i)$ for all positions i.

Let p be a second state with $k \geq c(p)$ and let π and π' be the corresponding runs. Towards a contradiction, assume that the two states are $\equiv^{\leq k}_M$-equivalent in $\mathcal{A}$ but not in $\mathcal{A}'$. The other direction works symmetrically. Then there is a position j such that $c'(\pi'(j)) \neq^{\leq k} c'(\rho'(j))$ but $c(\pi(j)) =^{\leq k} c(\rho(j))$. Since $\pi(j) \equiv^{\leq k}_M \pi'(j)$ and $\rho(j) \equiv^{\leq k}_M \rho'(j)$, this violates transitivity of $\equiv^{\leq k}_M$ and finishes our contradiction. □

Theorem 11. *A representative merge of a DPA w.r.t. $\mu^{\sim}_{TM}$ is language equivalent to the original.*

Proof. Let $\kappa_1, \ldots, \kappa_m$ be an enumeration of the equivalence classes in $\mu^{\sim}_{\mathrm{TM}}$ sorted by descending priority. With Lemma 1 and our assumption that $\sim$ behaves well with state merges, merging the states in κ_i will not change the equivalence classes $\kappa_{i+1}, \ldots, \kappa_n$. It is therefore a language preserving operation to merge all equivalence classes in the given order. The resulting automaton is the same as a representative merge w.r.t. $\mu^{\sim}_{\mathrm{TM}}$. □

The computation of $\mu^{\sim}_{\mathrm{TM}}$ is rather straight forward.

Theorem 12. *For a given $\sim$ in a suitable data structure, $\mu^{\sim}_{TM}$ can be computed in $\mathcal{O}(|Q| \cdot |c(Q)| \cdot \log |Q|)$.*

Proof. Assuming that $\equiv^{\sim}_{\mathrm{TM}}$ is known, computing $\mu^{\sim}_{\mathrm{TM}}$ is easy. $\equiv^{\sim}_{\mathrm{TM}}$ is an intersection of three equivalence relations, so the total run time is the sum of the run time to compute each of the three parts.

$\sim$ is already given, so no time is needed. Computing the relation of equal priorities is also doable in $\mathcal{O}(|Q|)$, assuming suitable data structures in the automaton. Finally, $\equiv^{\leq k}_M$ can be computed with just a slight adaption of usual algorithms for Moore equivalence in time $\mathcal{O}(|Q| \cdot \log |Q|)$. This needs to be done for every k, so $|c(Q)|$ times.

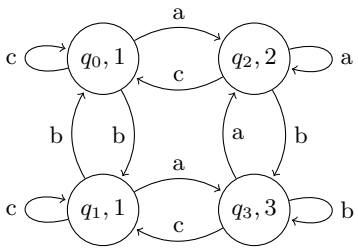

Fig. 2: Example TM

8 Labeled SCC Filter

The labeled SCC filter technique (LSF) is similar to the previous TM algorithm. Where the former required equivalent states to have equal priority, which limits its capabilities, LSF improves in that field by adding a different constraint instead; namely, reachability in the subautomaton that is constructed by removing all priorities below a certain threshold.

For every priority k, we deine $\mathcal{A} \restriction^{c}_{\geq k}$ as the subautomaton of $\mathcal{A}$ that contains only states with priority greater than k. Furthermore, we built total extensions of the reachability preorders for each such k and call them $\preceq_k$.

Definition 15. *Let $k \in \mathbb{N}$ and let $\sim$ be an equivalence relation that implies language equivalence. We define the* LSF equivalence relation $\equiv^{k,\sim}_{LSF}$ *as the intersection of $\equiv^{\leq k}_{M}$ and $\sim$.*

LSF is an example where most of the computational logic is not part of the underlying equivalence relation but rather part of the merger function. For each equivalence class κ of $\equiv^{k,\sim}_{\text{LSF}}$, we define $C^k_\kappa = \{r \in \kappa \mid c(r) > k \text{ and } r \text{ is } \preceq_k \text{-maximal among } \kappa\}$. We also set $M^k_\kappa = \kappa \setminus C^k_\kappa$.

Definition 16. *We define the* LSF merger function $\mu^{k,\sim}_{LSF}$ *as follows: for each equivalence class κ of $\equiv^{k,\sim}_{LSF}$, we map $\mu^{k,\sim}_{LSF}(M^k_\kappa) = C^k_\kappa$.*

The merging process works similar to that of the TM merger function. We also assume here that the relation $\sim$ behaves well regarding language preserving merges.

Lemma 2. *Let $\mathcal{A}$ be a DPA and let $\mathcal{A}'$ be a representative merge w.r.t. some equivalence class $\kappa \in \mathfrak{C}(\mu^{k,\sim}_{LSF})$. For all states p and q in $\mathcal{A}'$, the two states are $\equiv^{k,\sim}_{LSF}$-equivalent in $\mathcal{A}$ if and only if they are in $\mathcal{A}'$.*

Proof. We show that the statement holds for $\equiv_L$ and for $\equiv^{\leq k}_{M}$. Assuming that $\sim$ is well behaved, it then translates to $\equiv^{k,\sim}_{\text{LSF}}$.

Regarding $\equiv^{\leq k}_{M}$, let q be a state with $k \geq c(q)$ and let ρ and ρ' be the runs of the two automata on some α starting in q. As $\equiv^{\leq k}_{M}$ is a congruence relation, $\rho(i) \equiv^{\leq k}_{M} \rho'(i)$ for all positions i.

Let p be a second state with $k \geq c(p)$ and let π and π' be the corresponding runs. Towards a contradiction, assume that the two states are $\equiv^{\leq k}_{M}$-equivalent in $\mathcal{A}$ but not in $\mathcal{A}'$. The other direction works symmetrically. Then there is a position j such that $c'(\pi'(j)) \neq^{\leq k} c'(\rho'(j))$ but $c(\pi(j)) =^{\leq k} c(\rho(j))$. Since $\pi(j) \equiv^{\leq k}_{M} \pi'(j)$ and $\rho(j) \equiv^{\leq k}_{M} \rho'(j)$, this violates transitivity of $\equiv^{\leq k}_{M}$ and finishes our contradiction.

Now regarding $\equiv_L$, we show that runs ρ and ρ' of $\mathcal{A}$ and $\mathcal{A}'$ always have the same acceptance status when starting in the same state. We can observe that $\rho(i) \equiv_L \rho'(i)$ and $\rho(i) \equiv_M^{\leq k} \rho'(i)$ at all positions i, as both of these are congruence relations. Thus, if ρ' from some point j on only uses transitions in $\mathcal{A}$, then the two runs have the same acceptance status, as $\rho(j) \equiv_L \rho'(j)$.

If ρ' visits infinitely many states with priority k or less, then ρ and ρ' see that priority at the same positions and therefore share the same acceptance.

Finally, assume that ρ' uses infinitely many redirected edges but only sees states of priority greater than k from some point on. To use a redirected transition, the automaton has to visit a state in M_κ^k and then moves to a state in C_κ^k instead. From there, however, it is only possible to reach M_κ^k again via some states of priority k or less (by definition of the candidate set). Hence, these two assumptions contradict another.

Theorem 13. *A representative merge of a DPA w.r.t. $\mu_{LSF}^{k,\sim}$ is language equivalent to the original.*

Proof. Let $\kappa_1, \ldots, \kappa_m$ be an enumeration of the equivalence classes of $\equiv_{\mathrm{LSF}}^{k,\sim}$. When merging $M_{\kappa_i}^k$ into $C_{\kappa_i}^k$, the equivalence classes $\kappa_{i+1}, \ldots, \kappa_m$ do not change by Lemma 2. Also the candidate sets $C_{\kappa_j}^k$ themselves do not change, so we can safely merge all $M_{\kappa_i}^k$ into $C_{\kappa_i}^k$. This is the same operation as performed by the merger function.

Computation of the LSF merger is more complicated compared to the TM merger but still rather straight forward.

Theorem 14. *For a given $\sim$ in a suitable data structure, $\mu_{LSF}^{k,\sim}$ can be computed in $\mathcal{O}(|Q| \cdot \log |Q|)$.*

Proof. $\sim$ is already given and $\equiv_M^{\leq k}$ can be computed in $\mathcal{O}(|Q| \cdot \log |Q|)$. Building $\equiv_{\mathrm{LSF}}^{k,\sim}$ is an easy linear time intersection operation.

The second step to building $\mu_{\mathrm{LSF}}^{k,\sim}$ is to compute C_κ^k for each κ. For that, it suffices to find the order $\preceq_k$ and then select all the maximal elements from each equivalence class. This order can be computed by a topological sorting on the SCCs of $\mathcal{A} \restriction_{>k}^{\mathrm{c}}$ which one can construct in linear time.

References

[1] Filippo Bonchi and Damien Pous. Checking nfa equivalence with bisimulations up to congruence. In *Proceedings of the 40th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL '13, pages 457–468, New York, NY, USA, 2013. ACM.

[2] Julius Richard Büchi. On a decision method in restricted second order arithmetic. 1966.

[3] J.-M. Champarnaud and F. Coulon. Nfa reduction algorithms by means of regular inequalities. *Theoretical Computer Science*, 327(3):241 – 253, 2004. Developments in Language Theory.

[4] Kousha Etessami, Thomas Wilke, and Rebecca A. Schuller. Fair simulation relations, parity games, and state space reduction for büchi automata. In *Automata, Languages and Programming*, pages 694–707, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

[5] Carsten Fritz and Thomas Wilke. Simulation relations for alternating büchi automata. *Theor. Comput. Sci.*, 338(1-3):275–314, June 2005.

[6] John Hopcroft. An n log n algorithm for minimizing states in a finite automaton. *An N Log N Algorithm for Minimizing States in A Finite Automaton*, page 15, 01 1971.

[7] Tao Jiang and B. Ravikumar. Minimal nfa problems are hard. In *Automata, Languages and Programming*, pages 629–640, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[8] Richard Mayr and Lorenzo Clemente. Advanced automata minimization. In *POPL 2013*, Oct 2012.

[9] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. 1943. *Bulletin of mathematical biology*, 52 1-2:99–115; discussion 73–97, 1990.

[10] Edward F. Moore. Gedanken-experiments on sequential machines. In Claude Shannon and John McCarthy, editors, *Automata Studies*, pages 129–153. Princeton University Press, Princeton, NJ, 1956.

[11] A.W. Mostowski. Hierarchies of weak automata and weak monadic formulas. *Theoretical Computer Science*, 83(2):323 – 335, 1991.

[12] Michael O Rabin and Dana Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3:114–125, 04 1959.

[13] Sven Schewe. Beyond Hyper-Minimisation—Minimising DBAs and DPAs is NP-Complete. In Kamal Lodaya and Meena Mahajan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010)*, volume 8 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 400–411, Dagstuhl, Germany, 2010. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[14] Wolfgang Thomas. Handbook of theoretical computer science (vol. b). chapter Automata on Infinite Objects, pages 133–191. MIT Press, Cambridge, MA, USA, 1990.

[15] Wolfgang Thomas. Handbook of formal languages, vol. 3. chapter Languages, Automata, and Logic, pages 389–455. Springer-Verlag New York, Inc., New York, NY, USA, 1997.