Bachelor's Thesis Computing Science

Radboud University Nijmegen

Finding New Asymptotic Bounds on the Minimal Optimal Strategy Sets of Parametric MDPs

Author:
Pim Leerkes
s1060308

First supervisor/assessor:
Dr. Sebastian Junges

Second assessor:
Dr. Jurriaan Rot

February 26, 2024

Abstract

Parametric Markov decision processes (pMDPs) are an extension of classical Markov decision processes (MDPs) where we also allow for functions over a set of parameters as opposed to constant transition probabilities. Finding a precise asymptotic bound on the size of the minimal set of strategies that will be optimal for families of pMDPs, is an open problem in the domain of theoretical computer science. Information on the bounds of this set could namely be used to classify certain decision problems. We refer to this set as the minimal optimal strategy set (MOSS). This thesis has the aim of tackling the problem in question. During that process we proved that in general for families of simple, acyclic, univariate pMDPs (SAU-pMDPs), the MOSS can grow quadratically with respect to the size. This is done constructively by providing a concrete family for which it holds. To expand on this, an empirical argument is given for why it is likely that there also exists a family for which the size of the MOSS even has an exponential lower bound with respect to the number of states, which is conclusively formulated by a conjecture. Furthermore, a candidate is presented for future analysis that might yield more insights into the existence of an exponential lower bound with respect to the number of actions. Several additional ideas are presented as well that might prove useful in subsequent research.

Contents

Chapter 1

Introduction

Markov decision processes (MDPs) [Put94; How72] are probabilistic models that can be used to represent real life situations that involve elements of uncertainty and sequential decision making. They provide numerous applications in fields that deal with optimization problems. An example of this is the field of robotics, where a programmer could want to specify how a robot makes decisions in response to insecure situations in order to achieve a goal [Thr05]. In the study of MDPs, one aim we might have, is to look for an optimal strategy, also called an optimal policy. In essence, a strategy is a specification of what decision to make at each state, and the optimal strategy refers to a strategy where the chances of reaching a certain objective are either maximized or minimized. This objective might for example be to reach or avoid a certain state.

Parametric Markov decision processes (pMDPs) [HHZ11] extend regular MDPs by allowing for more generalisation. What this means is that we can model an entire set of MDPs at once where the graph structure is the same but where the probabilities may differ. The way this is achieved, is by also allowing for real/rational functions as transition probabilities in pMDPs. These functions depend on variables, which in their turn can be any value from a predefined set or interval. A specific choice of these values is called a parameter valuation, and this valuation once again induces a regular MDP. A valuation is well defined if the induced MDP does not have states where the sum of the outgoing probabilities does not equal one. All of the well defined valuations together construct the well defined parameter space, denoted by VAL.

We can still search for an optimal strategy in pMDPs, but because we use functions over parameters instead of constant values, we almost never have a single strategy that is optimal for all parameter valuations. However, there do exist strategies that are not optimal for any parameter valuations and also strategies that are optimal for one or more valuations. The latter are referred to as *somewhere optimal strategies*. Since there is usually more than one somewhere optimal strategy for a certain pMDP, we can construct an optimal strategy set (OSS). For each parameter valuation, this set contains at least one strategy that is optimal for that valuation. Furthermore, we can reduce this to a minimal optimal strategy set (MOSS), which is an OSS with minimal cardinality. In other words, a MOSS is an OSS such that it is no longer an OSS if you remove a strategy.

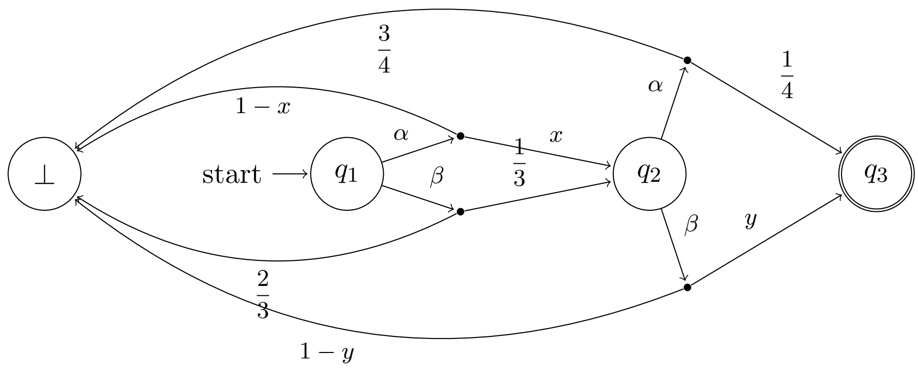


Figure 1.1: Graphical representation of an example pMDP

Motivation and related work The concept of pMDPs spawned new decision problems of which their complexity is the subject of research [JKPW19; JKPW20]. The idea of a MOSS then originated from the fact that this could help classify certain decision problems. What previous research in particular has found is that there are two decision problems called $\exists\exists \mathit{Reach}_*^{\bowtie}$ and $\exists\forall \mathit{Reach}_*^{\bowtie}$ that reside in coNP if the size of a MOSS is in general polynomially bounded [JKPW19].

Lemma 1.0.1. [1] *If the size of a MOSS on a VAL is polynomially bounded for fixed-parameter pMDPs then* $\exists\exists \mathit{Reach}_*^{\bowtie}$ *and* $\exists\forall \mathit{Reach}_*^{\bowtie}$ *are in coNP.*

The bounds on a MOSS can be expressed with for example big O notation as is also seen in the analysis of algorithm complexity. Thus far the only thing that was known about the bounds on the size of a MOSS was that it can grow exponentially in the arbitrary parameter case [Jun20], i.e., in that case there is an exponential lower bound.

Proposition 1.0.1. [2] *There exists a family* $(\mathcal{M}_n)_{n\in\mathbb{N}}$ *of simple pMDPs with* $n+2$ *states s.t.* $|\Omega^{\mathcal{M}_n}| \geq 2^n$ *for any OSS* $\Omega^{\mathcal{M}_n}$ *on VAL, i.e., the size of a MOSS can grow exponentially in the pMDPs size.*

[1] Refer to [JKPW19] for a definition of these decision problems.
[2] Refer to Chapter 2 and Chapter 3 for background information on the notation

Research question Finding new bounds on the size of a MOSS is the main purpose of this thesis, hence the research question is as follows: "What are the asymptotic bounds on the size of the minimal optimal strategy sets of pMDPs?". In this thesis, we deliver partial answers to this question in the form of a newly determined lower bound. Mathematical proofs are also given to support these new discoveries. Additionally, there is also an hypothesized stronger bound accompanied by numerical evidence. This list is not exhaustive as several other ideas and results are presented.

Structure of the thesis In Chapter 2 the necessary background information is provided on the topic of polynomials and the basics of Markov models, such as pMCs and pMDPs. Before we delve into the research portion of the thesis, the problem that inspired the research question is formalized in Chapter 3. Then, in Chapter 4 we investigate bounds on the MOSS and this is accompanied by proofs to support the findings. Following that, there is a brief reflection on the results and a discussion of additional ideas and future research in Chapter 5. Finally, the most important findings are summarized in Chapter 6.

Chapter 2

Preliminaries

2.1 Polynomials

Polynomials are algebraic expressions consisting of variables and coefficients, combined through addition, subtraction and multiplication. Polynomials can have as many different variables as needed and they do not have upper bounds on their length. In this thesis we exclusively work with polynomials consisting of only one variable. These are called univariate polynomials. For general background material on the topics in this section consult [FR97; Pra04].

Example 2.1.1. *$f(x) = x^2 - 1$, $g(x) = 5\frac{2}{3}x^3y^4 + \frac{1}{2}x^2 + y$ and $h(x) = 4x$ are all polynomials, where $h(x)$ is a monomial.*

Definition 2.1.1 (Univariate polynomial)**.** *A (real) univariate polynomial $f \in \mathbb{R}[X]$ is an expression $\sum_{i=0}^{n} a_i x^i$ with $a_i \in \mathbb{R}$ and $n \in \mathbb{N}$.*

Remark. *In this definition $\mathbb{R}[X]$ denotes the ring of all univariate polynomials with coefficients in $\mathbb{R}$. In other cases we could also have for example $\mathbb{Q}[X]$ (polynomials with rational coefficients), but in this thesis we work exclusively with polynomials with real coefficients for reasons that will later become apparent.*

We can also reason about polynomials consisting of only one term and these are appropriately called monomials.

Definition 2.1.2 (Monomial)**.** *A (real) univariate monomial is an expression ax^i with $a \in \mathbb{R}$ and $i \in \mathbb{N}$.*

The degree of a polynomial is the highest power i of all terms $a_i x^i$ with $a_i \neq 0$.

Example 2.1.2. *The degree of $f(x) = x^2 - 1$ is 2 and the degree of $g(x) = 4$ is 0.*

2.1.1 Roots of polynomials

The roots of a univariate polynomial f are all values of x such that $f(x) = 0$. The fundamental theorem of algebra states that each non constant polynomial has at least one complex root. In the research section of this thesis, roots of polynomials play an important role in proving the new theorems. While searching for these roots, it is important to note that looking for the roots of a polynomial is essentially equivalent to factorising the polynomial. This is because the roots of a polynomial are determined by the roots of its irreducible factors. We call a factor irreducible if it cannot be factored any further.

Example 2.1.3. *$f(x) = x^2 - 1$ from Example 2.1.1 has as its roots $x = 1$ and $x = -1$ because $f(1) = 0$ and $f(-1) = 0$.*

There do not exist general formulas for calculating the roots of polynomials of degree 5 or higher [Abe26], therefore we often use numerical methods to obtain/approximate roots of polynomials [DM89; McN07; MP13].

Example 2.1.4. *Consider the polynomial $f(x) = 3x^4 + 5x^3 + 10x^2 + 20x - 8$. We can factor this polynomial into $(3x - 1)(x + 2)(x^2 + 4)$, and we can find roots for each of these factors. In fact, we can see that the real roots are $x = \frac{1}{3}$ and $x = -2$.*

Two separate polynomials have a common root x when they share a common factor. The common root is then the root of this common factor. All common factors of two polynomials are the common factors of the greatest common divisor (GCD) of the two polynomials. We can compute the GCD of two polynomials by using the Euclidean algorithm or by using factorisation. If two polynomials $g(x)$ and $f(x)$ happen to not share any factors, we call these two polynomials coprime, i.e., when $gcd(f(x), g(x)) = 1$ (similar to the case of integers).

Example 2.1.5. *Consider the two polynomials $f(x) = x^2 + 5x + 4$, and $g(x) = x^2 - 2x - 3$. Then $gcd(f(x), g(x)) = (1+x)$, because if we factorise the two polynomials, we see that $f(x) = (x+1)(x+4)$ and $g(x) = (x+1)(x-3)$ and here they both share the factor $(x + 1)$. This is their greatest common factor because there is no other way of factoring the polynomials. Thus these two polynomials have the common root $x = -1$.*

2.2 Markov Chains

A Markov chain (MC) [Nor97; kS60] is a stochastic process that describes a state transition system of probabilities and can be modelled by a directional graph (digraph) [Gou12]. It is essentially a finite state machine where you do not always end up in the same next state, but where you can end up is specified by a fixed distribution of probabilities. Markov chains are applicable in a broad range of fields, such as physics [Sen16], information theory [Khi57] and many more. A concrete example is the page rank algorithm that Google uses in their search engine [PLS13]. A distinctive quality of Markov chains is that when we are in a certain state, the next state only depends on this current state and not on states we previously visited. This is called the memoryless property (or the Markov property).

Definition 2.2.1 (Markov chain)**.** *A (finite, discrete-time) Markov chain* $\mathcal{M}$ *is a tuple* (S, s_0, P, T) *where* S *is a finite, nonempty set of states,* $s_0 \in S$ *is the initial state,* $P\colon S \times S \to \mathbb{R}$ *is the transition probability function and* $T \subseteq S$ *denotes the set of target states.*

Remark. *In literature, the definitions of a Markov chain differ slightly per paper but for the purposes of this thesis we define it in this manner. It is conventional to use* $\mathbb{Q}$ *as the codomain of the transition probability function for algorithmic purposes [BK08], but as we will later see it is more useful for us to define it as* $\mathbb{R}$.

Remark. *A Markov chain induces an underlying digraph where states act as vertices, and there is an edge from* s *to* s' *if and only if* $P(s, s') > 0$.

Example 2.2.1. *Consider the Markov chain* $\mathcal{M}$ *with its corresponding digraph in Figure 2.1, with as set of states* $S = \{q_0, q_1, q_2\}$*, initial state* $s_0 = q_0$ *and transition probability function*

$$\begin{cases} P(q_0, q_1) = \frac{1}{4}, P(q_0, q_2) = \frac{3}{4}, \\ P(q_1, q_0) = \frac{2}{3}, P(q_1, q_2) = \frac{1}{3}, \\ P(q_2, q_0) = \frac{1}{3}, P(q_2, q_2) = \frac{2}{3}. \end{cases}$$

All other possible transitions have probability 0. There are no target states in this case so we specify the set of target states as the empty set, i.e., $T = \emptyset$.

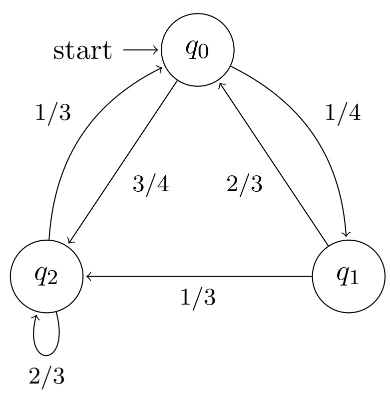


Figure 2.1: Simple Markov chain with 3 states

In probability theory, usually all the probabilities of a distribution must add up to one, hence we often also want a Markov chain to have this property. To differentiate between models where this property is satisfied and where it is not, we call the models where it is satisfied *well defined*.

Definition 2.2.2 (Well defined)**.** *A Markov chain $\mathcal{M}$ is called well defined if all transition probabilities are non-negative, and for each $s \in S$ we have that the sum of outgoing transition probabilities equals 1, i.e., $\sum_{s' \in S} P(s, s') = 1$.*

Remark. *In this thesis we only consider well defined models.*

In Markov chains we also make a distinction between acyclic and cyclic models, which are properties of its underlying digraph (Figure 2.2). Furthermore, in a lot of situations we want a certain state to be such that you cannot leave it. We refer to these states as absorbing states. See for example Figure 2.3, where a six sided die is simulated by a two sided coin [KY76]. Eventually, you end up at one of the six numbers and you never leave these states.

Definition 2.2.3 (Absorbing state)**.** *An absorbing state is a state s such that $P(s, s) = 1$.*

Remark. *An absorbing state is graphically denoted either as a state with a loop edge going to itself with transition probability 1, simply by omitting any outgoing edges or lastly a special case of an absorbing state is the sink state of which there is typically at most one per model. A sink state is denoted by $\perp$.*

Definition 2.2.4 (Acyclic)**.** *A Markov chain $\mathcal{M}$ is called acyclic when the underlying digraph is acyclic (with possible exception of absorbing states).*

Remark. *The remainder of this thesis focuses exclusively on acyclic models.*

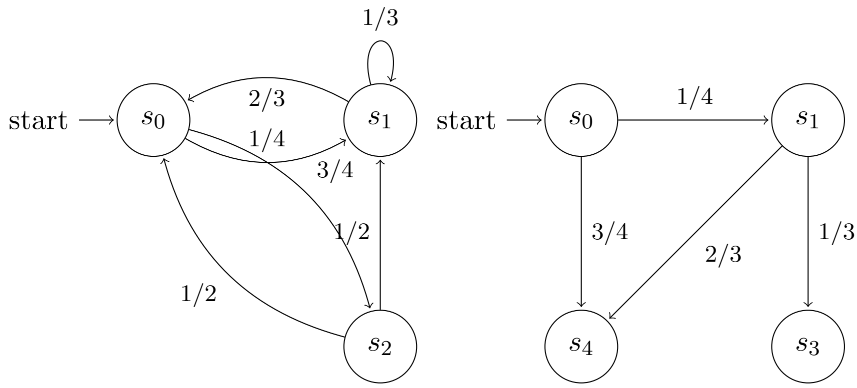


Figure 2.2: cyclic (left) and acyclic Markov chain (right)

Some definitions in literature also add a reward function that specifies a reward for certain states, or a cost function that assigns costs to transitions [BK08]. In this thesis we will be exclusively looking at Markov models without a reward or cost associated with states or transitions. In stochastic research there are several other types of Markov chains with additional properties being utilized, such as interval Markov chains [KU02], hidden Markov chains [Yoo09], continuous time Markov chains [Nor97] and more. In this thesis we mainly consider Markov decision processes (MDPs) [Put94] and more specifically the parametric kind [HHZ11], which will both be explained in the next sections.

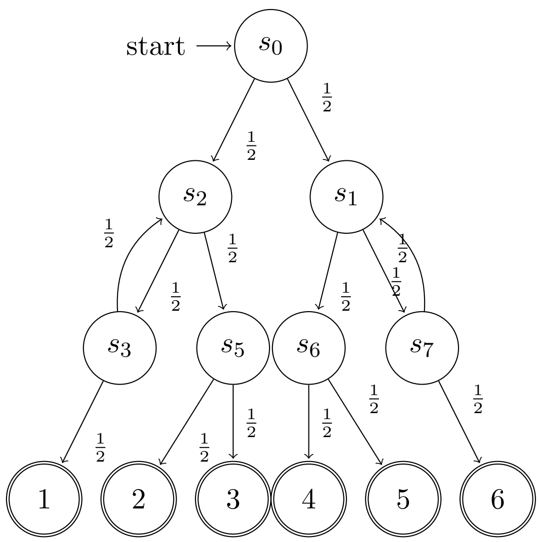


Figure 2.3: Knuth and Yao Markov chain [KY76] (two sided coin simulating a six sided die)

2.3 Markov Decision Processes

In Markov decision processes (MDPs) [Put94] we expand upon our notion of Markov chains by also adding the element of *choice*, hence MDPs can be truly nondeterministic. In MDPs, we have that in every state you can make a choice, and only after you make a choice there is a probability distribution that specifies in which state you can end up next. These choices are referred to as actions. We cannot have actions in a regular Markov chain so we need to add an extra element to our tuple, namely the *Act* set, where each element is an action. An action is a label of an outgoing edge of a state that can be chosen nondeterministically. Each state has its own set of *available* actions, and each action you then choose has its own probability distribution. Because of their nondeterministic nature, MDPs have applications in fields that contain optimization problems involving decision making, for example robotics [Thr05], machine learning [SB08] and more prominently in reinforcement learning [vOW12].

Definition 2.3.1 (Markov Decision Process)**.** *A (finite) Markov decision process* $\mathcal{M}$ *is a tuple* (S, s_0, Act, P, T) *where* S *is a finite, nonempty set of states,* $s_0 \in S$ *is the initial state,* Act *is a finite set of actions,* $P\colon S \times Act \times S \to \mathbb{R}$ *is a transition probability function and* $T \subseteq S$ *is the set of goal states.*

Remark. *An MDP can also be called well defined and this is the case when for every* $s \in S$*, we have that* $\sum_{s' \in S} P(s, \alpha, s') = 1$ *for every* $\alpha \in Act$ *(and if all transition probabilities are non-negative). We only consider well defined MDPs in this thesis.*

Remark. *The other properties and notions we defined for Markov chains also apply to MDPs in a similar way as they did to Markov chains.*

Remark. *An MDP in essence, is a more general version of a Markov chain because a Markov chain has in all states* $s \in S$ *as set of available actions the singleton set, i.e., Markov chains are a subset of MDPs.*

Example 2.3.1. *Consider the MDP* $\mathcal{M}$ *in Figure 2.4, with set of states* $S = \{q_1, q_2, \perp\}$*, set of actions* $Act = \{\alpha, \beta\}$*, initial state* $s_0 = q_1$*, transition probability function*

$$\begin{cases} P(q_0, \alpha, q_1) = \frac{3}{4},\ P(q_0, \alpha, q_2) = \frac{2}{3}, \\ P(q_0, \beta, q_1) = \frac{1}{4},\ P(q_0, \beta, q_2) = \frac{1}{3}, \end{cases}$$

and set of goal states $T = \{q_1, q_2\}$*.*

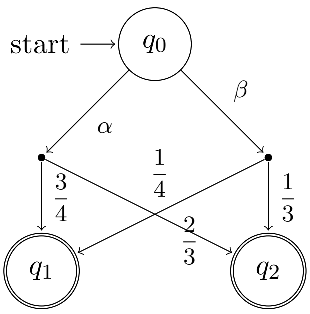


Figure 2.4: Small example MDP

For MDPs, since we have nondeterminism, we can specify for a certain MDP what action to take in each state. This mapping from states to actions is called a policy, but in literature the terms scheduler [BK08] and strategy [GOP11] are also used as synonyms. Additionally in literature there is a differentiation between memoryless policies and policies that are not memoryless (see for example [J+19] for a definition of policies that are not necessarily memoryless). In our case we only consider policies that are memoryless. Before we formally define policies we will first define the available set of actions for a certain state.

Definition 2.3.2. *We define* $Act(s) = \{\alpha \in Act \mid \exists s' \in S : P(s, \alpha, s') \neq 0\}$.

Definition 2.3.3 (Policy)**.** *A (memoryless) policy/scheduler/strategy* σ *is a function* $\sigma \colon S \to Act$ *defined by* $\sigma(s) = \alpha$ *with the condition that* $\alpha \in Act(s)$.

Remark. *We denote the set of all policies for a certain MDP* $\mathcal{M}$ *by* $\Sigma^{\mathcal{M}}$. *Since we only concider finite MDPs, this set is always finite.*

Example 2.3.2. *In Figure 2.5, we can choose for example the policy* $\sigma = \{q_0 \to \alpha, q_1 \to \beta\}$.

In Chapter 3, we will expand upon the notion of policy and introduce the concept of optimal policies.

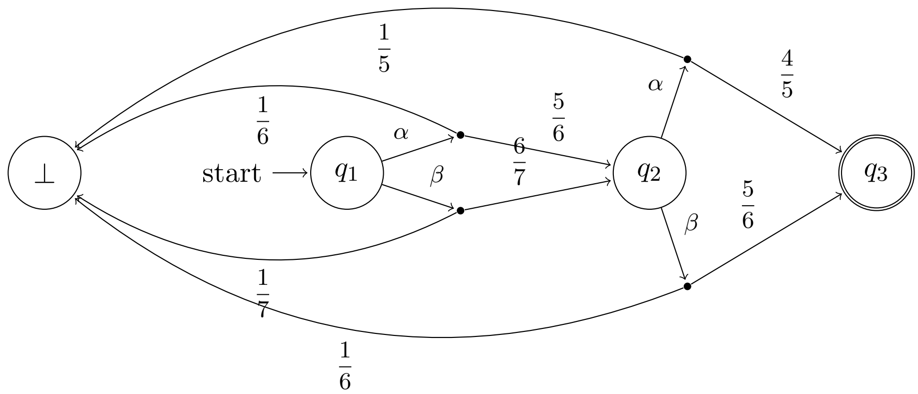


Figure 2.5: Small example MDP with sink state and 2 actions

Given a specific policy, an MDP will be transformed back into a regular Markov chain because now the actions will disappear and we will instead have regular transition probabilities. We call this the induced Markov chain.

Definition 2.3.4 (Induced Markov chain)**.** *For a MDP* $\mathcal{M}$ *with given policy* σ *the induced Markov chain* $\mathcal{M}_\sigma$ *is a Markov chain* (S, s_0, P_σ, T) *such that:*

$$\forall s, s' \in S : P_\sigma(s, s') = P(s, \sigma(s), s').$$

2.4 Parametric Models

2.4.1 Parametric Markov chains

Instead of constants as transition probabilities, we might also need functions depending on parameters that label the transitions. In practice, these are used to describe sets of systems where the graph structure is the same but where the probabilities may differ. To model this mathematically you can use an interval Markov chain for intervals of probabilities [KU02] or more generally, parametric Markov chains (pMCs) for real/rational functions [Daw04].

Example 2.4.1. *Consider again the Markov chain from Figure 2.3, but now we use the parameters x and y instead of concrete transition probabilities (Figure 2.6).*

Definition 2.4.1 (Parametric Markov Chain)**.** *A parametric Markov chain $\mathcal{M}$ is a tuple (S, s_0, P, T, X) where S is a finite, nonempty set of states, s_0 is the initial state, $X = \{x_1, ..., x_n\}$ is a finite set of parameters, T is the set of goal states and P is the transition probability function $P\colon S \times S \to \mathbb{R}[X]$.*

Remark. *The notions we defined for regular Markov chains transfer to pMCs (only well definedness is different but this will be explained later this section).*

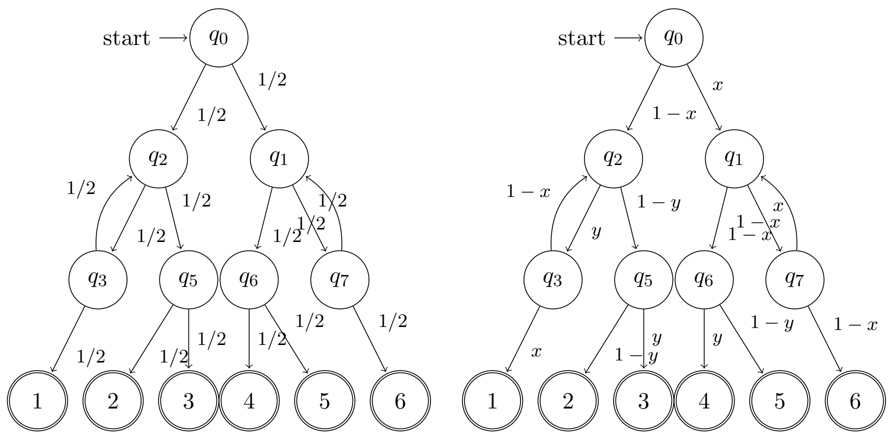


Figure 2.6: Knuth and Yao Markov chain [KY76] with concrete transition probabilities (left) and parameters (right)

In pMCs we speak of valuations, where a valuation means a specific choice of values for every parameter $x \in X$. A useful, formal way to talk about a valuation is that it is a function that assigns to each parameter a value.

Definition 2.4.2 (Valuation)**.** *A valuation is a function $val \colon X \to [0,1]$ that maps a parameter to a concrete value.*

When using parameters, the pMC describes multiple Markov chains at once because you essentially have a different Markov chain for every valuation of the parameters. In other words, when we choose a parameter valuation for a certain pMC this will again yield a regular Markov chain.

Example 2.4.2. *The Knuth and Yao Markov Chain [KY76] in Figure 2.6 on the left side is an induced Markov chain $\mathcal{M}(val)$ of the one on the right side with $val(x) = 1/2$ and $val(y) = 1/2$.*

For a pMC to be well defined, we need to have that for each valuation it induces a well defined Markov chain. All valuations that induce a well defined Markov chain together construct the well defined parameter space.

Definition 2.4.3 (Well defined parameter space)**.** *The well defined parameter space VAL for a certain pMC $\mathcal{M}$ is defined by*

$$VAL = \{val \colon X \to [0,1] : \mathcal{M}(val) \text{ is a well defined Markov chain}\}.$$

2.4.2 Parametric Markov decision processes

Just as with pMCs, we can also parameterise MDPs, making us arrive at the most important notion in this thesis, namely that of parametric Markov Decision Processes (pMDPs)[HHZ11].

Example 2.4.3. *Consider again the MDP from Figure 2.5. In Figure 2.7 we see a similar MDP but now some concrete probabilities are replaced by parameters x and y, which makes it a pMDP.*

Definition 2.4.4 (Parametric Markov Decision Process)**.** *A parametric Markov Decision Process $\mathcal{M}$ is a tuple (S, s_0, Act, P, X, T) where S is a finite set of states, $s_0 \in S$ is the initial state, Act is a finite set of actions, $X = \{x_1, ..., x_n\}$ is a finite set of parameters, $P \colon S \times Act \times S \to \mathbb{R}[X]$ is a transition probability function and T is the set of goal states.*

Remark. *A pMDP can induce both MDPs and pMCs. An MDP is induced by choosing a parameter valuation and a pMC is induced by choosing a policy. A well defined pMDP is a pMDP such that each induced MDP and each induced pMC is well defined. All other properties we defined for regular MDPs and pMCs transfer to pMDPs.*

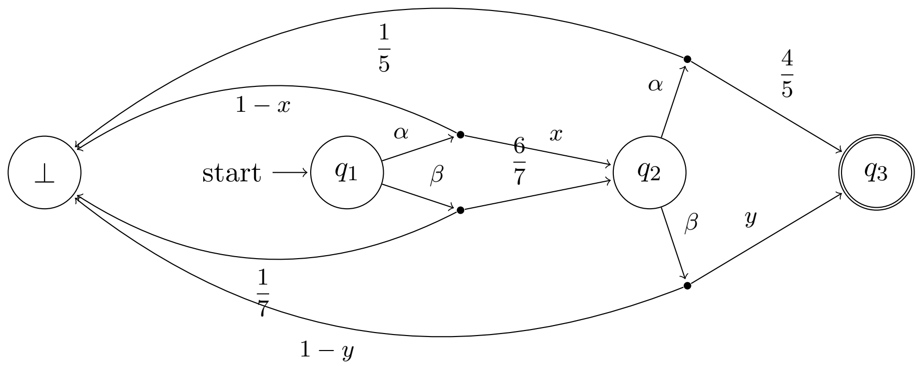


Figure 2.7: Small example pMDP with two parameters

Chapter 3

Problem Definition

This section formally defines the problem of this thesis, but before we can define the problem, an understanding of several concepts is needed. We will first briefly show how to calculate probabilities of reaching certain states and what the different kinds of solution functions are. After this we are able to explain the notion of (somewhere) optimal policies which can then be used to understand the optimal strategy sets (OSS) and minimal optimal strategy sets (MOSS) of pMDPs. Following these definitions, we will introduce some necessary constraints on pMDPs that can narrow down the focus of our research. We will refer to this special type of pMDPs as simple, acyclic, univariate pMDPs (SAU-pMDPs). At the end of this chapter the formal problem definition together with several subproblems will follow.

3.1 Solution Functions

Reachability probabilities Before we can properly introduce the solution function, we first need to understand how reachability probabilities are calculated for (p)MDPs and (p)MCs. In Markov chains we denote by: $Pr_{\mathcal{M}}(s \rightarrow \Diamond s'))$[1] the probability of eventually reaching s' when starting from s. We can calculate it by multiplying and adding certain transition probabilities on the path between s and s' as in the examples below, where we determine it for induced Markov chains. In general, the reachability probabilities can be computed by several techniques, e.g., by solving systems of linear equations [BK08].

Example 3.1.1. *Consider the induced Markov chain $\mathcal{M}(val)$ of the pMC in Figure 3.1 with $val(x) = \frac{1}{2}$. The probability of reaching q_3 starting from q_0 is calculated by: $Pr_{\mathcal{M}(val)}(q_0 \rightarrow \Diamond q_3) = \frac{1}{2} \cdot Pr_{\mathcal{M}(val)}(q_1 \rightarrow \Diamond q_3) + (1 - \frac{1}{2}) \cdot Pr_{\mathcal{M}(val)}(q_4 \rightarrow \Diamond q_3) = \frac{1}{2} \cdot \frac{1}{3} + 0 = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$.*

[1] $\Diamond s'$ is notation from linear temporal logic [HR04]. In this case it denotes reaching s'.

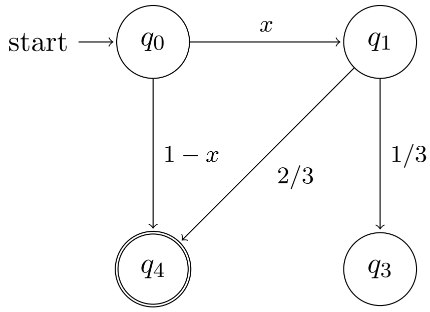


Figure 3.1: pMC for illustrating reachability probabilities

In (p)MDPs, the reachability probability depends on the chosen policy. In the next example, we first get the induced pMC and from this we derive the induced Markov chain, and calculate a reachability probability for it.

Example 3.1.2. *Consider the pMDP $\mathcal{M}$ in Figure 3.2. Let the induced pMC be $\mathcal{M}_\sigma$ with policy $\sigma = \{q_1, q_2 \rightarrow \alpha\}$, and let this pMC induce a Markov chain $\mathcal{M}_\sigma(val)$ with $val(x) = \frac{1}{3}, val(y) = \frac{1}{2}$. The probability of reaching q_3 is then calculated by: $Pr_{\mathcal{M}_\sigma(val)}(q_1 \rightarrow \Diamond q_3) = \frac{1}{3} \cdot Pr_{\mathcal{M}_\sigma(val)}(q_2 \rightarrow \Diamond q_3) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$.*

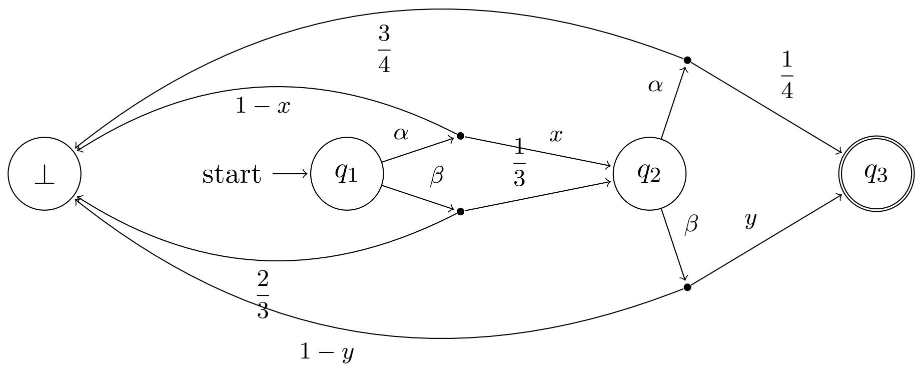


Figure 3.2: Small pMDP to illustrate reachability probabilities

The solution function Usually, the most important reachability probability in (parametric) Markov models is the probability of reaching the goal states T. In (p)MCs, we call this specific probability the *solution.* As we already saw, in pMCs the reachability probabilities can depend on the valuation. This is why for the case of pMCs, we will express this as a function that we will call the solution function [Jun20], which takes as input a valuation and gives as output the probability of reaching the goal state. In pMDPs there is more to solutions than in pMCs because the probability of reaching the goal state depends on the policy we choose. To resolve this, in the case of pMDPs, we talk about the minimal solution and the maximal solution functions. These functions can be used to reason about the lowest probability with which we can possibly reach the goal states given all policies, and the highest probability of reaching the goal states given all policies respectively. The functions that map the valuations to these probabilities are the *minsol* and the *maxsol* functions.

Example 3.1.3. *In Figure 3.1, we have that the solution function with respect to initial state q_0 is $1 - \frac{1}{3}x$ and in Figure 3.2, we have that the maximal solution function with respect to the initial state q_1 is*

$$\begin{cases} xy & \text{if } val(x) \geq \frac{1}{3} \text{ and } val(y) \geq \frac{1}{4}, \\ \frac{1}{3}y & \text{if } val(x) < \frac{1}{3} \text{ and } val(y) \geq \frac{1}{4}, \\ \frac{1}{4}x & \text{if } val(x) \geq \frac{1}{3} \text{ and } val(y) < \frac{1}{4}, \\ \frac{1}{12} & \text{if } val(x) < \frac{1}{3} \text{ and } val(y) < \frac{1}{4}. \end{cases}$$

Definition 3.1.1 (Solution function)**.** *For a pMC $\mathcal{M}$ and state s, let the solution function $sol_s^{\mathcal{M}}\colon VAL \to [0,1]$ be defined as*

$$sol_s^{\mathcal{M}}(val) = Pr_{\mathcal{M}(val)}(s \to \Diamond T).$$

Definition 3.1.2 (Minimal solution function)**.** *For a certain pMDP $\mathcal{M}$, the minimal solution is a function $minsol_s^{\mathcal{M}}\colon VAL \to [0,1]$ which is defined as*

$$minsol_s^{\mathcal{M}}(val) = min_{\sigma \in \Sigma^{\mathcal{M}}} sol_s^{\mathcal{M}_\sigma}(val).$$

Definition 3.1.3 (maximal solution function)**.** *For a certain pMDP $\mathcal{M}$, the maximal solution is a function $maxsol_s^{\mathcal{M}}\colon VAL \to [0,1]$ which is defined as*

$$maxsol_s^{\mathcal{M}}(val) = max_{\sigma \in \Sigma^{\mathcal{M}}} sol_s^{\mathcal{M}_\sigma}(val).$$

Remark. *We are allowed to take the minimum and maximum over all policies because $\Sigma^{\mathcal{M}}$ is always a finite set.*

Similarly to the regular reachability probabilities in Markov chains, for (p)MDPs the maximal and minimal probabilities of reaching a goal state can in general be calculated explicitly by solving a system of equations [BK08].

3.2 Optimal Policies

Now that we have defined the several types of solution functions, we can build upon both these definitions and the definition of policies from last chapter and define what an optimal policy is. Informally, an optimal policy (in our case) is a policy that reaches the goal state with the highest probability. An important thing to note as well is that there can be more than one optimal policy for the same valuation. For pMDPs, this occurs in the case where for a certain valuation, two or more policies result in induced MCs with the same probability of reaching the goal state that is also the maximal probability of reaching the goal state over all policies in the pMDP.

Definition 3.2.1 (Optimal policy)**.** *For a pMDP $\mathcal{M}$ and a certain $val \in VAL$ the optimal policy is the policy σ such that*

$$sol_s^{\mathcal{M}_\sigma}(val) = maxsol_s^{\mathcal{M}}(val).$$

Example 3.2.1. *In the pMDP in Figure 3.2, we have for a val with $val(x) = 1/2, val(y) = 4/5$ that the optimal policy is $\sigma = \{q_1 \to \alpha, q_2 \to \beta\}$. To verify this we can calculate $sol_s^{\mathcal{M}_\sigma}(val) = Pr_{\mathcal{M}_\sigma(val)}(s \to \Diamond T) = \frac{1}{2} \cdot \frac{4}{5} = \frac{2}{5}$. It can be verified that the other policies yield solutions lower than this so $sol_s^{\mathcal{M}_\sigma}(val) = maxsol_s^{\mathcal{M}}(val)$, hence it is the optimal policy by definition.*

If our (p)MDP is sufficiently small we can find an exact optimal policy efficiently by methods that make use of dynamic programming techniques, for example value iteration, policy iteration or linear programming [Bel58; BK08].

3.2.1 Optimal strategy sets

In a pMDP, one can have policies that are nowhere optimal, and policies that are somewhere optimal. A nowhere optimal policy is a policy that is not optimal for any valuation, and a somewhere optimal policy is a policy for which there exists at least one valuation such that it is optimal. It trivially follows from this that we can construct a finite set out of all the somewhere optimal policies together that is a subset of the set of all policies. Such a set is called an optimal strategy set (OSS) [Jun20; JKPW19].

Example 3.2.2. *In the pMDP $\mathcal{M}$ in Figure 3.3, we have that the policy $\{q_1 \to \alpha\}$ is optimal when $val(x) \geq \frac{1}{3}$ and $\{q_1 \to \beta\}$ is optimal when $val(x) \leq \frac{1}{3}$, so a somewhere optimal policy set is $\Omega^{\mathcal{M}} = \{q_1 \to \alpha\ , q_1 \to \beta\}$.*

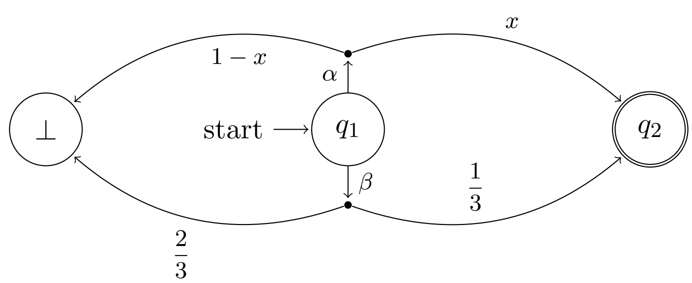


Figure 3.3: pMDP with two somewhere optimal policies

Definition 3.2.2 (OSS)**.** *For a pMDP $\mathcal{M}$, an optimal strategy set (OSS) $\Omega^{\mathcal{M}} \subseteq \Sigma^{\mathcal{M}}$ is a set such that*

$$\forall val \in VAL : \exists \sigma \in \Omega^{\mathcal{M}} : sol_s^{\mathcal{M}_\sigma}(val) = maxsol_s^{\mathcal{M}}(val).$$

3.2.2 Minimal optimal strategy sets

An OSS can grow exponentially in simple cases (see Example 3.2.3) because we allow for too many *unnecessary* policies. We are more interested in a set with reduced cardinality, namely the minimal optimal strategy set (MOSS) [Jun20; JKPW19]. The MOSS is an OSS such that if you remove any policy, it is no longer an OSS. We will provide a few examples followed by a formal definition.

Example 3.2.3. *In the pMDP $\mathcal{M}$ in Figure 3.4, we have that $|\Omega^{\mathcal{M}}| = 2^4$ because when $val(x) = \frac{1}{2}$, all policies result in $sol_{s_0}^{\mathcal{M}_\sigma}(val) = (\frac{1}{2})^4 = \frac{1}{16}$. However, $|MOSS^{\mathcal{M}}| = |\{q_0, q_1, q_2, q_3 \to \alpha\} \cup \{q_0, q_1, q_2, q_3 \to \beta\}| = 2$.*

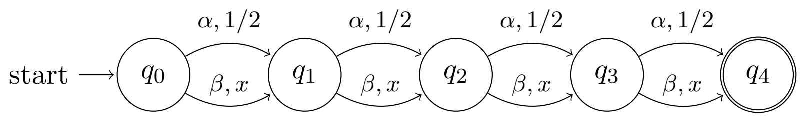


Figure 3.4: pMDP with $|\Omega^{\mathcal{M}}| = 2^4$ and $|MOSS^{\mathcal{M}}| = 2$ (sink edges omitted)

Example 3.2.4. *in the pMDP $\mathcal{M}$ in Figure 3.1 we see that $\Sigma^{\mathcal{M}} = 2^4$, but if we count optimal policies we find $|\Omega^{\mathcal{M}}| = 5$.*

For $val(x) < \frac{1}{5}$ only the policy $\{q_0, q_1, q_2, q_3 \to \alpha\}$ is optimal.
For $\frac{1}{5} < val(x) < \frac{1}{4}$ we have that only $\{q_0, q_1, q_2 \to \alpha, q_3 \to \beta\}$ is optimal.
For $\frac{1}{4} < val(x) < \frac{1}{3}$ we have that only $\{q_0, q_1 \to \alpha, q_2, q_3 \to \beta\}$ is optimal.
For $\frac{1}{3} < val(x) < \frac{1}{2}$ we have that only $\{q_0 \to \alpha, q_1, q_2, q_3 \to \beta\}$ is optimal, and for $\frac{1}{2} < val(x)$ we have that only $\{q_0, q_1, q_2, q_3 \to \beta\}$ is optimal.
In the cases where val(x) equals one of these probabilities exactly, no new policies become optimal.

It follows that in this case $|MOSS^{\mathcal{M}}| = |\Omega^{\mathcal{M}}| = 5$.

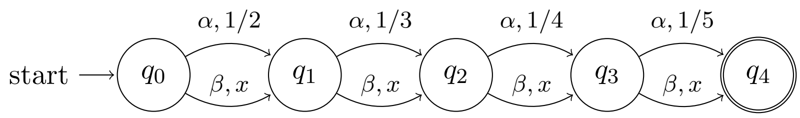


Figure 3.5: pMDP with $|MOSS| = 5$ (sink edges omitted)

Definition 3.2.3 (MOSS)**.** *For a pMDP $\mathcal{M}$, the minimal optimal strategy set (denoted $MOSS^{\mathcal{M}}$) is an OSS $\Omega^{\mathcal{M}}$ such that for each element $\sigma \in \Omega^{\mathcal{M}}$ it holds that $\Omega^{\mathcal{M}} \setminus \{\sigma\}$ is not an OSS.*

Remark. *The $\mathcal{M}$ is omitted from $MOSS^{\mathcal{M}}$ when it is clear from context to which pMDP we refer.*

3.3 SAU-pMDPs

The problem of this thesis is defined for a specific type of pMDP, which we will call SAU-pMDPs. Here, SAU stands for simple, acyclic and univeriate. We will only consider pMDPs for which these three properties hold. This choice was made because it would make the focus of our research more narrow and hence more organized. We can make this restriction because our solutions for the problems defined for this specific type of pMDPs also hold for the more general case mentioned as assumption in Lemma 1.0.1. This is because we only discover new *lower* bounds, and relaxing our constraints will not remove policies.

The notion acyclic has already been defined (Definition 2.2.4). A univariate pMDP is a pMDP where the total number of parameters is one. The definition of a simple pMDP is slightly more extensive but in summary it comes down to a well defined pMDP where the transition probabilities are only allowed to be constants, a parameter or the complement of a parameter. In literature this notion is defined more in depth [Jun20].

Definition 3.3.1. *A univariate pMDP $\mathcal{M}$ is a pMDP with $|X| = 1$.*

Definition 3.3.2. *A simple pMDP $\mathcal{M}$ is a well defined pMDP where the codomain of the transition probability function P is defined as:*

$$([0,1] \cap \mathbb{R}) \cup X \cup \{1 - x \mid x \in X\}.$$

Definition 3.3.3 (SAU-pMDP)**.** *A SAU-pMDP $\mathcal{M}$ is a pMDP which is simple, acyclic and univariate.*

Remark. *All pMDPs seen in this chapter are examples of SAU-pMDPs.*

3.4 Formal Problem Statement

An open question regarding the previously established notions is what the asymptotic bounds are on the size of the MOSS, i.e., how fast does $|MOSS^{\mathcal{M}_n}|$ grow as the size grows of an arbitrary family of (SAU-)pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$ [Jun20; JKPW19]. We will first define the meaning of these terms more precisely.

Definition 3.4.1. *We define the size of a pMDP $\mathcal{M}$ by $|S| + |Act|$.*

Remark. *Note that the definition of size varies per paper. In the paper where Lemma 1.0.1 originates from [JKPW19], the total number transition probabilities greater than zero also counts towards the size. For simplicity we use the alternative definition in our case.*

Definition 3.4.2. *A family of pMDPs is a sequence $(\mathcal{M}_n)_{n\in\mathbb{N}}$ of pMDPs where for each pMDP in the sequence, zero or more elements of the tuple depend on the index n.*

Remark. *We denote by S_n, Act_n and X_n the set of states, actions and parameters respectively for instance n of a family of pMDPs.*

So far we know that a MOSS can at least grow exponentially in the arbitrary parameter case [Jun20], but it remains open what further bounds are. This is the problem that this thesis addresses and the answer comes in the form of a function $f(m)$ that will bound the size of the set.

Problem 1. *Given an arbitrary family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$, find a function $f(k)$ such that $|MOSS^{\mathcal{M}_n}| \in \Theta(f(k))$, where k is the size of SAU-pMDP $\mathcal{M}_n$.*

This problem can be broken down into smaller problems in multiple ways. It can be split into a questions regarding only a lower or only an upper bound (O and Ω), but we can also split the size of the pMDP into states and actions and search for bounds with respect to those more specific properties. We will define the latter option explicitly.

- **Bound in States:** Given an arbitrary family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$, find a function $f(|S_n|)$ such that $|MOSS^{\mathcal{M}_n}| \in \Theta(f(|S_n|))$.
- **Bound in Actions:** Given an arbitrary family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$, find a function $f(|Act_n|)$ such that $|MOSS^{\mathcal{M}_n}| \in \Theta(f(|Act_n|))$.

We could even drop the univariate condition and study how the MOSS grows when the set of parameters grows. The next chapter will not involve attempts to solve this problem.

Chapter 4

Finding the Asymptotic Bounds

In this chapter, we provide partial solutions to the problems stated in the previous chapter, supported by mathematical proofs. We first explain how polynomials relate to SAU-pMDPs, which is necessary because it builds the foundation of the results. We then use this to provide an example of a family of SAU-pMDPs that has an quadratically growing MOSS with respect to size. After that, we expand on this by providing another family accompanied by numerical evidence which hints at an exponential lower bound with respect to states. At the end of the chapter, a candidate is presented for further analysis that could provide insight into the existence of a new exponential lower bound with respect to actions.

4.1 Polynomials in Relation to SAU-pMDPs

When looking at large and complicated SAU-pMDPs, it becomes increasingly difficult to formally analyze the size of the MOSS. So it is helpful to represent a SAU-pMDP in a more simple way, in our case with a set of polynomials. Polynomials are a well-researched topic that we can hence use to do calculations with in a straightforward manner [Pra04; FR97]. The main idea is that we first translate from SAU-pMDPs to sets of polynomials, then derive the necessary results for these sets, and finally we translate back to SAU-pMDPs. We will first explain how to convert between sets of polynomials and SAU-pMDPs, and following this we will introduce some important additional definitions for these sets of polynomials such as the equal maxima set and the unique maxima quantity. The theory in this section is presented along with several results which we will use while proving the new bounds in the later sections.

4.1.1 Conversion between polynomials and SAU-pMDPs

We can convert between polynomials and SAU-pMDPs in two directions. We can generate from a SAU-pMDP a set of polynomials where each polynomial corresponds to a policy namely, the polynomial equals the solution function that results from a policy induced pMC. The other way around means constructing a SAU-pMDP from a set of polynomials which allows for more freedom as we will see.

From SAU-pMDPs to polynomials We first explain how to construct a set of polynomials from a concrete SAU-pMDP by an example.

Example 4.1.1. *Consider the SAU-pMDP $\mathcal{M}$ in Figure 4.1. In this case $\Sigma^{\mathcal{M}} = 2^4$ and each $\sigma \in \Sigma^{\mathcal{M}}$ can be represented by a polynomial, for example the policy $\sigma = \{q_0, q_1, q_2, q_3 \to \beta\}$ can be represented by $sol_{s_0}^{\mathcal{M}_\sigma}(val) = x^4$. This way we can represent the entire policy space of the SAU-pMDP with this set of polynomials:*

$$\{x^4, c_1x^3, c_1x^3, c_2x^3, c_3x^3, c_0c_1x^2, c_0c_2x^2, c_0c_3x^2, c_1c_2x^2, c_1c_3x^2,$$
$$c_2c_3x^2, c_0c_1c_2x, c_0c_1c_3x, c_0c_2c_3x, c_1c_2c_3x, c_1c_2c_3c_4\}.$$

We can do it this way because we have as a corollary of a result by [JKPW19] that the solution function of an induced pMC $\mathcal{M}_\sigma$ of a SAU-pMDP $\mathcal{M}$ is always a polynomial. In the general case we create the set of polynomials corresponding to a SAU-pMDP using the following set construction.

Definition 4.1.1. *The set of polynomials $\mathcal{P}$ corresponding to a SAU-pMDP $\mathcal{M}$ is defined as*

$$\mathcal{P} = \{sol_{s_0}^{\mathcal{M}_\sigma}(val) \mid \sigma \in \Sigma^{\mathcal{M}}\}.$$

Remark. *We will sometimes use x as a substitute for val, and $f(x)$ as a substitute for $sol_{s_0}^{\mathcal{M}_\sigma}(val)$.*

We can also represent an entire family of SAU-pMDPs with a set of polynomials, and this is done by constructing a sequence of sets of polynomials.

Definition 4.1.2. *The sequence of sets of polynomials $(\mathcal{P}_n)_{n\in\mathbb{N}}$ corresponding to a family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$ is defined by*

$$\mathcal{P}_n = \{sol_{s_0}^{(\mathcal{M}_n)_\sigma}(val) \mid \sigma \in \Sigma^{\mathcal{M}_n}\}.$$

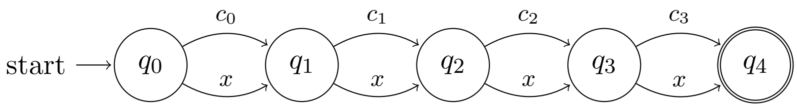


Figure 4.1: small example pMDP (sink edges omitted)

From polynomials to SAU-pMDPs Now that we have established how to construct a set of polynomials given a SAU-pMDP or a family of SAU-pMDPs, we continue by explaining how this can be done the other way around. This procedure provides much more freedom as there are usually multiple SAU-pMDPs that map to the same set of polynomials $\mathcal{P}$ as is seen in Example 4.1.2. This is why we do not know of a general procedure apart from choosing a SAU-pMDP followed by checking if the set of polynomials generated by this SAU-pMDP is equal to the set of polynomials you started with.

Example 4.1.2. *Consider the set of polynomials* $\mathcal{P} = \{(1-x)^i x^j \mid i, j \in \mathbb{N} \wedge i + j \leq 3\}$. *This set is generated by the SAU-pMDP* $\mathcal{M}$ *in Figure 4.2. In Figure 4.3 we have another SAU-pMDP* $\mathcal{M}$ *that also generates* $\mathcal{P}$.

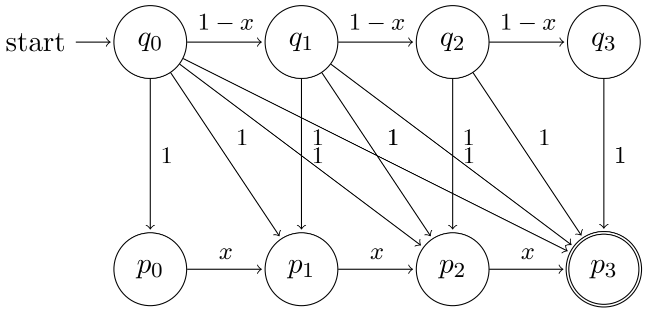


Figure 4.2: SAU-pMDP that generates $\mathcal{P}$ (sink edges omitted)

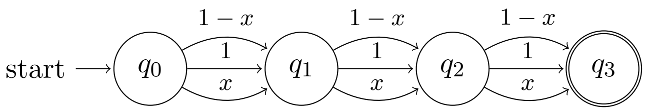


Figure 4.3: Another SAU-pMDP that also generates $\mathcal{P}$ (sink edges omitted)

Adequate polynomials We conclude this subsection by providing a pair of statements. One will give us an indication for when we can for a certain set of polynomials $\mathcal{P}$ produce a SAU-pMDP $\mathcal{M}$ that generates this set, and the other provides insight into what the elements of $\mathcal{P}$ might look like. Before that however, it is needed to define what adequate polynomials are.

This term will also reappear several times in the next sections. In essence, an adequate polynomial is a polynomial f for which $f(val)$ stays between the boundaries of 0 and 1 when val is between 0 and 1.

Definition 4.1.3 (Adequate polynomials)**.** *A polynomial* $f \in \mathbb{Q}[X]$ *is adequate if* $0 < f(val) < 1$ *for all* $val : X \to (0, 1)$*, and* $0 \leq f(val) \leq 1$ *for all* $val \colon X \to [0, 1]$*.*

The theorem below then states that all adequate polynomials are the solution function of a certain pMC. Since for any pMC there exists a pMDP where choosing a policy induces it[1], we can conclude that for all sets of adequate, univariate polynomials, we can construct a SAU-pMDP that generates it.

Theorem 4.1.1 (Winkler's trick)**.** *[Jun20] Let* $f \in \mathbb{Q}[X]$ *be a univariate, adequate polynomial. There exists a simple, acyclic pMC* $\mathcal{M}$ *with a target state* T *such that* $f = sol_s^{\mathcal{M}}$*.*

Remark. *The definition of adequate polynonials and Winkler's trick will also hold for* $\mathbb{R}[X]$ *instead of* $\mathbb{Q}[X]$*.*

We can also deduce as a result of a lemma from [Jun20], that the other way around, each polynomial from a set of polynomials $\mathcal{P}$ generated by a SAU-pMDP is adequate, constantly zero or constantly one.

Lemma 4.1.1. *Let* $\mathcal{M}$ *be a simple, acyclic pMC. The function* $sol_s^{\mathcal{M}}$ *is adequate, constantly zero or constantly one.*

As any pMC can be induced from choosing a policy in a pMDP, and since SAU-pMDPs are a subset of simple, acyclic pMDPs, we can derive our desired result.

4.1.2 Somewhere optimal polynomials

Now that we understand how to make a mapping between sets of polynomials and SAU-pMDPs, we would like there to be a connection between the notion of optimal policies and a comparable property in polynomials. For this purpose, we introduce the notion of somewhere optimal polynomials. Given a set of polynomials $\mathcal{P}$, a somewhere optimal polynomial is a polynomial f for which there is at least one point $x \in (0, 1)$ such that $f(x)$ is greater or equal to all values $g(x)$ of the other polynomials g from the same set. We can directly make a connection between somewhere optimal polynomials and somewhere optimal policies as is seen in Lemma 4.1.2.

[1]This trivially follows from the fact that pMCs are a subset of SAU-pMDPs, namely in pMCs the set of available actions is the singleton set in each state.

Example 4.1.3. *In Figure 4.4, we see a SAU-pMDP that generates* $\mathcal{P} = \{x, 1-x, \frac{2}{3}\}$. *If we plot these polynomials we see that they each become somewhere optimal on a different interval on the x-axis.* $1-x$ *is optimal for* $val(x) \in [0, \frac{1}{3}]$, $\frac{2}{3}$ *is optimal for* $val(x) \in [\frac{1}{3}, \frac{2}{3}]$ *and* x *is optimal when* $val(x) \in [\frac{2}{3}, 1]$.

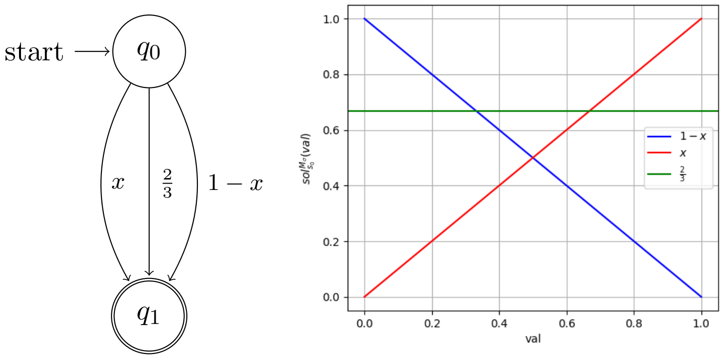


Figure 4.4: SAU-pMDP (sink edges omitted) and its corresponding set of polynomials plotted

Definition 4.1.4 (Somewhere optimal polynomial)**.** *For a certain set* $\mathcal{P}$ *of polynomials, a somewhere optimal polynomial* $f \in \mathcal{P}$ *is a polynomial such that*

$$\exists x \in [0,1] : \forall g \in \mathcal{P} : f(x) \geq g(x).$$

For the value of x *at which this condition holds, we say that* f *is optimal in* x.

Sometimes we want to reason about a stronger notion called somewhere super optimal polynomials.

Definition 4.1.5 (Somewhere super optimal polynomial)**.** *For a certain set* $\mathcal{P}$ *of polynomials, a super somewhere optimal polynomial* $f \in \mathcal{P}$ *is a polynomial such that*

$$\exists x \in [0,1] : \forall g \in \mathcal{P} : f(x) > g(x).$$

For the value of x *at which this condition holds, we say that* f *is super optimal in* x.

Lemma 4.1.2. *For a SAU-pMDP* $\mathcal{M}$ *and its generated set of polynomials* $\mathcal{P}$*, a polynomial* $sol_{s_0}^{\mathcal{M}_\sigma}(val) \in \mathcal{P}$ *is a somewhere optimal polynomial if and only if* σ *is a somewhere optimal policy for* $\mathcal{M}$.

Proof. Take an arbitrary SAU-pMDP $\mathcal{M}$ and its corresponding set of polynomials $\mathcal{P}$. Then, a polynomial $sol_{s_0}^{\mathcal{M}_\sigma}(val) \in \mathcal{P}$ for a certain policy $\sigma \in \Sigma^{\mathcal{M}}$ is a somewhere optimal polynomial (by definition) if and only if

$$\exists val \in \text{VAL} : \forall sol_{s_0}^{\mathcal{M}_{\sigma'}}(val) \in \mathcal{P} : sol_{s_0}^{\mathcal{M}_\sigma}(val) \geq sol_{s_0}^{\mathcal{M}_{\sigma'}}(val).$$

This is equivalent to saying that

$$\exists val \in \text{VAL} : \forall \sigma' \in \Sigma^{\mathcal{M}} : sol_{s_0}^{\mathcal{M}_\sigma}(val) \geq sol_{s_0}^{\mathcal{M}_{\sigma'}}(val)$$

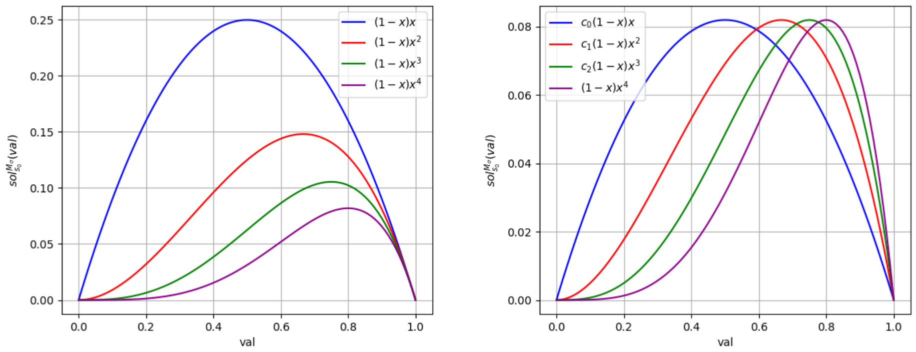
by definition of the set of polynomials, which holds if and only if

$$\exists val \in \text{VAL} : sol_{s_0}^{\mathcal{M}_\sigma}(val) = maxsol_{s_0}^{\mathcal{M}}(val)$$

by definition of the maxsol function. To conclude, we note that by definition of an optimal policy, the former is true if and only if σ is an optimal policy for $\mathcal{M}$ for some valuation $val \in \text{VAL}$, which is equal to saying that σ is a somewhere optimal policy for $\mathcal{M}$. □

Making polynomials somewhere (super) optimal Before we try to make a connection from a set of polynomials to a MOSS we will first explain how we can make more polynomials from a set somewhere (super) optimal. This is useful because when we try to prove the new lower bounds we try to make the number of policies as high as possible in order to rule out certain upper bounds.

If we plot the polynomials from the set $\mathcal{P} = \{(1-x)x^i \,|\, 1 \leq i \leq 4\}$ (see Figure 4.5), for example, we see that only $(1-x)x$ is a somewhere super optimal polynomial and the others are not. By a simple procedure however, we can let each polynomial become somewhere super optimal. This procedure consists of multiplying each polynomial $f \in \mathcal{P}$ with a scalar, such that the value of $f(x)$ of their maxima between 0 and 1 will end up at the same exact value. The way these scalars are created is by dividing for a certain polynomial $f \in \mathcal{P}$ the value $g(x)$, the lowest local maximum of all polynomials $g \in \mathcal{P}$ between 0 and 1, by the $f(x)$ value of the local maximum of f between 0 and 1. It can easily be deduced from Rolle's Theorem and elementary calculus that each adequate polynomial will have a maximum between 0 and 1 [Ste08].



(a) Without scalars

(b) With scalars

Figure 4.5: Polynomials of $\mathcal{P} = \{(1-x)x^i \,|\, 1 \le i \le 4\}$ without (left) and with additional scalars (right)

In general, once we have multiplied each polynomial from the set by this scalar, we end up with a new set of polynomials which we will call the equal maxima set. For a certain set of adequate polynomials $\mathcal{P}$ we denote this set by $\mathcal{P}'$. We will first give a formal definition of this set, followed by an example of how such a set $\mathcal{P}'$ can be computed given a set of polynomials $\mathcal{P}$.

Definition 4.1.6 (Equal maxima set)**.** *The equal maxima set $\mathcal{P}'$ of a set $\mathcal{P}$ of adequate polynomials, is the set*

$$\mathcal{P}' = \{c \cdot f \,|\, f \in \mathcal{P},\, c = \frac{max_l}{max_f}\},$$

where $max_l = min\{max\{g(x) \,|\, x \in [0,1]\} \,|\, g \in \mathcal{P}\}$, and $max_f = max\{f(x) \,|\, x \in [0,1]\}$.

Remark. *Note that in the above definition we always have that $c \in [0,1]$.*

Example 4.1.4. *Consider the set $\mathcal{P} = \{(1-x)x^i \mid 1 \le i \le 4\}$ once again. We construct the equal maxima set $\mathcal{P}'$ by first taking the derivative of each polynomial:*

$$\frac{d}{dx}(1-x)x = 1 - 2x,$$
$$\frac{d}{dx}(1-x)x^2 = 2x - 3x^2,$$
$$\frac{d}{dx}(1-x)x^3 = 3x^2 - 4x^3,$$
$$\frac{d}{dx}(1-x)x^4 = 4x^3 - 5x^4.$$

Setting them equal to 0 *and solving for* x *gives:*

$$x = \frac{1}{2}, x = \frac{2}{3}, x = \frac{3}{4}, x = \frac{4}{5}.$$

Substituting back in the formulas gives the y *coordinates of the maxima:*

$$\frac{1}{4}, \frac{4}{27}, \frac{27}{256}, \frac{256}{3125}.$$

The lowest is $\frac{256}{3125}$*, so we multiply* $(1-x)x$ *with* $\frac{\frac{256}{3125}}{\frac{1}{4}}$*,* $(1-x)x^2$ *with* $\frac{\frac{256}{3125}}{\frac{4}{27}}$*,* $(1-x)x^3$ *with* $\frac{\frac{256}{3125}}{\frac{27}{256}}$*, and we multiply* $(1-x)x^4$ *with* $\frac{\frac{256}{3125}}{\frac{256}{3125}} = 1$*.*

Then we get a set of polynomials for which all elements are somewhere super optimal, as seen in Figure 4.5.

The following lemma and its corollary are important as they insure that necessary conditions for later statements are satisfied.

Lemma 4.1.3. *For every set* $\mathcal{P}$ *of univariate, adequate polynomials, and its equal maxima set* $\mathcal{P}'$*, we have that there exists a value* $y \in \mathbb{R}$ *such that for all* $f \in \mathcal{P}'$*, the value of* $f(x)$ *for the maximum of* f *in* $(0, 1)$ *equals* y*.*

Proof. When an adequate, univariate polynomial f is multiplied by a scalar to obtain a new polynomial $g = c \cdot f$ for $c \in [0, 1]$, then for all $x \in dom(f)$[2], we have that $g(x) = c \cdot f(x)$ so also for all x such that $f(x)$ is a local maximum (of which there is always at least one between 0 and 1 since f is adequate).

If $c = \frac{max_l}{max_f}$ then $c \cdot max_f = max_l$ so each newly obtained polynomial g will have the y coordinate of its maximum between $(0, 1)$ end up at max_l. □

Corollary 4.1.3.1. *For a set of univariate, adequate polynomials* $\mathcal{P}$*, each polynomial* $f \in \mathcal{P}'$ *is a somewhere optimal polynomial.*

Proof. Each polynomial $f \in \mathcal{P}'$ has a maximum at the same value of $f(x)$ by Lemma 4.1.3. This means that for a certain polynomial f, it holds that at the x coordinate of its maximum, we have that $f(x) \geq g(x)$ for all other polynomials g. Therefore, f is a somewhere optimal polynomial by definition. □

[2] We use $dom(f)$ to indicate the domain of f.

Counting unique x coordinates of maxima To finally make a useful connection between sets of polynomials and the MOSS, we must observe that in some cases the maxima belonging to separate polynomials from the same set $\mathcal{P}$ have the exact same x coordinate. When counting the number of somewhere super optimal polynomials from a certain set $\mathcal{P}'$, we only count one of them in this case. This is similar to the definition of a MOSS, where in most cases we also discard policies when multiple are optimal for the same valuation. Establishing how many somewhere super optimal polynomials there are is generally difficult and requires an analysis of the patterns of local maxima. Nonetheless, this brings us to the unique maxima quantity. Here, we define a function that returns the quantity of different local maxima for a given set of polynomials.

Definition 4.1.7 (Unique maxima quantity)**.** *The unique maxima quantity is a function* $\mathcal{U}_I : \mathscr{P} \to \mathbb{N}$ *defined by*

$$\mathcal{U}_I(\mathcal{P}) = |\{x \in I : \exists f \in \mathcal{P} : \frac{d}{dx}f(x) = 0 \wedge \frac{d^2}{dx^2}f(x) < 0\}|$$

where $\mathscr{P}$ *denotes the set of all sets of univariate polynomials.*

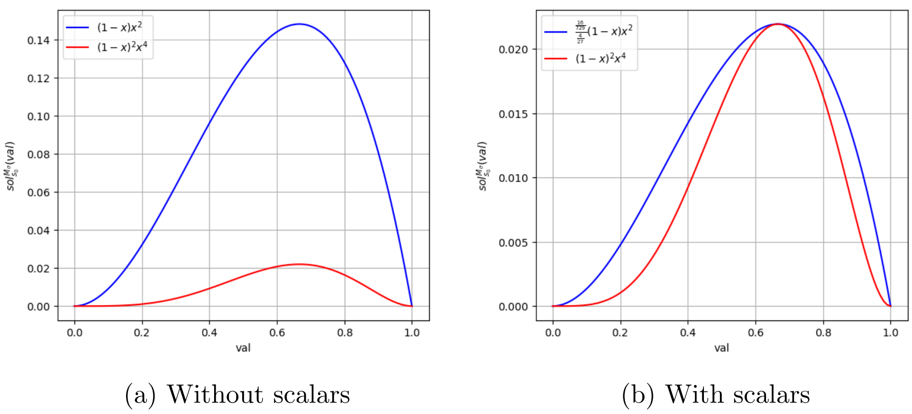
Remark. *In our case we only consider* $I = [0, 1]$.

Example 4.1.5. *For the set of polynomials* $\mathcal{P} = \{(1-x)x^i \,|\, 1 \leq i \leq 4\}$ *displayed in Figure 4.5, we have that* $\mathcal{U}_{[0,1]}(\mathcal{P}) = |\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}\}| = 4$.

Example 4.1.6. *In Figure 4.6, we see the polynomials* $(1-x)x^2$ *and* $(1-x)^2x^4$ *which both have their maximum at* $x = \frac{2}{3}$*, hence*

$$\mathcal{U}_{[0,1]}(\{(1-x)x^2, (1-x)^2x^4\}) = 1$$

even though the number of polynomials is 2.



(a) Without scalars (b) With scalars

Figure 4.6: Two different polynomials with a maximum at $x = \frac{2}{3}$ without and with additional scalars

From our definition of the unique maxima quantity, we can finally make the important connection to the MOSS using a theorem that we later use to prove the new bounds.

Theorem 4.1.2. *For a SAU-pMDP $\mathcal{M}$ with its generated set $\mathcal{P}$ of polynomials such that $\mathcal{P} = \mathcal{P}'$ and for all $f \in \mathcal{P}$ we have that f has at most one local maximum in $[0, 1]$, it holds that $|MOSS^{\mathcal{M}}| = \mathcal{U}_{[0,1]}(\mathcal{P}')$.*

Proof. Take an arbitrary SAU-pMDP $\mathcal{M}$ with its corresponding set of polynomials $\mathcal{P}$ that is equal to $\mathcal{P}'$. By Lemma 4.1.1, each polynomial in this set is adequate, constantly zero or constantly one, and since we have as a condition that each polynomial has at most one local maximum in $[0, 1]$, it cannot be constantly zero or constantly one, so it must be adequate. Furthermore, because a SAU-pMDP is univariate, each polynomial in $\mathcal{P}$ is univariate.

It follows from this that by Lemma 4.1.2, we have for all $f \in \mathcal{P}'$ that their corresponding σ is a somewhere optimal policy, since all polynomials in this set are somewhere optimal by Corollary 4.1.3.1.

All these policies together construct an OSS $\Omega^{\mathcal{M}}$ with $|\Omega^{\mathcal{M}}| = |\mathcal{P}'|$. If $\Omega^{\mathcal{M}}$ is not a MOSS we are able to remove elements from $\Omega^{\mathcal{M}}$ such that it is still an OSS by definition. If we remove such an element this equates to removing a policy such that after removal, we still have that for all *val* there exists a policy that is optimal for that *val* by definition. This translates to removing a f from $\mathcal{P}'$ such that at the points $x \in [0, 1]$ where it is optimal, there is also another polynomial which is optimal at that point. In other words, f is not somewhere super optimal, so its maximum in $[0, 1]$ (of which there is one by assumption) is the maximum of another polynomial in $\mathcal{P}'$ as well.

Starting with the equation $|\Omega^{\mathcal{M}}| = |\mathcal{P}'|$, if we remove all σ from $\Omega^{\mathcal{M}}$ until it is a MOSS and these corresponding polynomials f from $\mathcal{P}'$ we end up with $|MOSS^{\mathcal{M}}|$ on the left side and $\mathcal{U}_{[0,1]}(\mathcal{P}')$ on the right side by their definitions, i.e., we then have that $|MOSS^{\mathcal{M}}| = \mathcal{U}_{[0,1]}(\mathcal{P}')$. □

Remark. *The condition that a polynomial should have at most one maximum in $[0, 1]$ is necessary because otherwise you could have a case where for a SAU-pMDP $\mathcal{M}$ and its generated set of polynomials $\mathcal{P}$, a polynomial $f \in \mathcal{P}'$ has two maxima with the exact same value of $f(x)$. In that case $|MOSS^{\mathcal{M}}| < \mathcal{U}_{[0,1]}(\mathcal{P}')$.*

4.2 A Quadratic Lower Bound

Now that we have established all the necessary results for polynomials, we can start with investigating new bounds on the MOSS. The way we did this is by first ruling out a linear upper bound by finding a specific family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$ for which $|MOSS^{\mathcal{M}_n}| \notin O(n)$. In particular, it is the case that there exists a family for which the MOSS grows quadratically. We will first give a sequence of sets of polynomials and later present an example of a SAU-pMDP family that corresponds to it where this bound holds, which we will prove rigorously.

The sequence of sets of polynomials The first step is to find a useful sequence of sets of polynomials $(\mathcal{P}_n)_{n\in\mathbb{N}}$ for which the cardinality of its sets is not linearly bounded from above so that we can later construct a SAU-pMDP family $(\mathcal{M}_n)_{n\in\mathbb{N}}$ where its policies correspond to the polynomials in the equal maxima sets of this sequence $(\mathcal{P}'_n)_{n\in\mathbb{N}}$. In order to prove that $|MOSS^{\mathcal{M}_n}|$ grows superlinearly we will make sure that $\mathcal{U}_{[0,1]}(\mathcal{P}_n)$ does (as $\mathcal{U}_{[0,1]}(\mathcal{P}'_n) = \mathcal{U}_{[0,1]}(\mathcal{P}_n)$, since only y values change in the transformation from $\mathcal{P}$ to $\mathcal{P}'$) and the rest will follow from Theorem 4.1.2, since we make sure that its conditions are satisfied.

Definition 4.2.1 (The quadratic sequence)**.** *Consider the sequence of sets of polynomials $(\mathcal{P}^q_n)_{n\in\mathbb{N}}$ defined by $(n \geq 2)$:*

$$\mathcal{P}^q_n = \{(1-x)^i x^j \mid 1 \leq i < j \leq n \text{ where } i, j \in \mathbb{N}\}.$$

We call this sequence the quadratic sequence.

Remark. *Letting $i < j$ is a deliberate choice that was made to simplify the analysis of the sets graphically. To achieve the result of finding a family with a superlinear growing MOSS, this restriction is not needed. By symmetry all the results also hold for a set without this constraint, but the number of polynomials is only roughly cut in half in our case.*

Remark. *Note that $\mathcal{P}^q_n$ is undefined for $n = 1$ and $n = 0$.*

Example 4.2.1. $\mathcal{P}^q_3 = \{(1-x)x^2, (1-x)x^3, (1-x)^2x^3\}$.

The quadratic sequence is useful because it is easy to show that its cardinality grows quadratically when n increases and every polynomial in any of its sets is adequate. The following two lemmas support these claims.

Lemma 4.2.1. $|\mathcal{P}^q_n| = \frac{1}{2}n^2 - \frac{1}{2}n$.

Proof. We use proof by induction.

Base case $(n = 2)$:

$$\mathcal{P}^q_2 = \{(1-p)p^2\} \longrightarrow |\mathcal{P}^q_2| = 1 = \frac{1}{2} \cdot 2^2 - \frac{1}{2} \cdot 2.$$

Inductive step: For size n we have $|\mathcal{P}_n^q| = \frac{1}{2}n^2 - \frac{1}{2}n$. This is the induction hypothesis. We now prove it for $n+1$:

$$\begin{aligned}|\mathcal{P}_{n+1}^q| &= \text{(there are } n \text{ polynomials } (1-x)^i x^{n+1} \text{ with } 0 < i < n+1)\\ &\quad |\mathcal{P}_n^q| + n\\ &\stackrel{IH}{=} \frac{1}{2}n^2 - \frac{1}{2}n + n\\ &= \frac{1}{2}n^2 + n + \frac{1}{2} - \left(\frac{1}{2}n + \frac{1}{2}\right)\\ &= \frac{1}{2}(n^2 + 2n + 1) - \left(\frac{1}{2}n + \frac{1}{2}\right)\\ &= \frac{1}{2}(n+1)^2 - \frac{1}{2}(n+1).\end{aligned}$$

□

Before we state the next lemma, first note that in general the derivative of a polynomial $f \in \mathcal{P}_n^q$ is given by

$$\frac{d}{dx}f(x) = (1-x)^{i-1}x^{j-1}(j(1-x) - ix). \tag{4.1}$$

Lemma 4.2.2. *Every polynomial $f \in \mathcal{P}_n^q$ has exactly one maximum with its x coordinate in $(0,1)$ and $f(x)$ in $(0,1)$.*

Proof. Let f be an arbitrary polynomial in $\mathcal{P}_n^q$. It holds that

$$(1-x)^i x^j = 0 \longrightarrow 1 - x = 0 \vee x = 0 \longrightarrow x = 1 \vee x = 0,$$

and then it follows from Rolle's Theorem [Ste08] that f attains at least one local extreme point with its x coordinate between 0 and 1.

Now take the derivative $\frac{d}{dx}f(x) = (1-x)^{i-1}x^{j-1}(j(1-x) - ix)$. We see that all irreducible factors of this derivative are of the form $(1-x)$, x and $(j(1-x) - ix)$, therefore the roots of the derivative are of the form $x = 1$, $x = 0$ and $x = \frac{j}{i+j}$ and we know that $\frac{j}{i+j} \in (0,1)$. Substituting yields $(1 - (\frac{j}{i+j}))^i(\frac{j}{i+j})^j$ which is also clearly in $(0,1)$. We know now that there is only one extreme point in $(0,1)$ and that it is a maximum. □

Corollary 4.2.2.2. *Each polynomial $f \in \mathcal{P}_n^q$ is adequate.*

Proof. It immediately follows from Lemma 4.2.2 that each polynomial from $f \in \mathcal{P}_n^q$ is adequate, since it follows from the reasoning in the proof that the extreme points at $x = 0$ and $x = 1$ are not maxima, because if they were maxima there must have been minima between them and $\frac{j}{i+j}$. We can see that the values of $f(x)$ of the local minima at 0 and 1 are precisely 0, and the $f(x)$ value of the only maximum in $(0,1)$ is in $(0,1)$. The conditions for a polynomial being adequate are therefore satisfied. □

Concurrent x coordinates of maxima For the quadratic sequence we can generate its corresponding sequence of equal maxima sets $(\mathcal{P}_n^{q\prime})_{n\in\mathbb{N}}$ where for a certain size n, each polynomial $f \in \mathcal{P}_n^{q\prime}$ has the $f(x)$ value of its maximum between 0 and 1 at the same value by Lemma 4.1.3. The more important question is how often maxima occur at the same x coordinates, i.e., what is $\mathcal{U}_{[0,1]}(\mathcal{P}_n^q)$ for arbitrary n. We see that two maxima are on the same position on the x-axis for example for the polynomials $(1-x)x^2$ and $(1-x)^2x^4$, namely in both cases it is at $x = \frac{2}{3}$. If we investigate into why that is, we see that $(1-x)^2x^4 = (1-x)x^2 \cdot (1-x)x^2 = ((1-x)x^2)^2$. Generally speaking, it holds that two polynomials from $\mathcal{P}_n^q$ have a maximum on the same x value when one is a power of the other.

Proposition 4.2.1. *Two polynomials $f, g \in \mathcal{P}_n^q$ attain a maximum on the same $x \in (0,1)$ if and only if $f(x) = (g(x))^c$ or $g(x) = (f(x))^c$ for some $c \in \mathbb{N}$.*

Proof. We will prove both implications separately. For better readability we use m and n instead of i and j, and we use k as the index of the set of polynomials.
$(\longrightarrow)$:

Take two arbitrary polynomials $f, g \in \mathcal{P}_k^q$ and assume they attain a maximum on the same x value for $x \in (0,1)$ then this implies that their derivatives share a common factor with a root in $(0,1)$, i.e., $(1-x)^{m_1-1}x^{n_1-1}(n_1(1-x) - m_1x)$ and $(1-x)^{m_2-1}x^{n_2-1}(n_2(1-x) - m_2x)$ (4.1) share a common factor with a root in $(0,1)$.

Since all factors of these derivatives are of the form: $(1-x)$, x or $(n(1-x) - mx)$ the former implies that the common factor with a root in $(0,1)$ must be a factor of $(n_2(1-x) - m_2x)$ and $(n_1(1-x) - m_1x)$ since $(1-x)$ has as root $x = 1$ and x has as root $x = 0$ which are both not in $(0,1)$.

Since the factors $(n_2(1-x) - m_2x)$ and $(n_1(1-x) - m_1x)$ are both irreducible we have that the former implies that one of these factors divides the other, i.e., $(n_2(1-x) - m_2x) = c \cdot (n_1(1-x) - m_1x) = (cn_1(1-x) - cm_1x)$ or $(n_1(1-x) - m_1x) = c \cdot (n_2(1-x) - m_2x) = (cn_2(1-x) - cm_2x)$ for some $c \in \mathbb{N}$. This implies that $(n_1, m_1) = c \cdot (n_2, m_2)$ or $(n_2, m_2) = c \cdot (n_1, m_1)$ for some $c \in \mathbb{N}$.

So when we look at the entire derivatives again, we have that either $(1-x)^{m_1-1}x^{n_1-1}(n_1(1-x) - m_1x) = (1-x)^{cm_2-1}x^{cn_2-1}(cn_2(1-x) - cm_2x)$ or $(1-x)^{m_2-1}x^{n_2-1}(n_2(1-x) - m_2x) = (1-x)^{cm_1-1}x^{cn_1-1}(cn_1(1-x) - cm_1x)$. The polynomials on the right side are precisely the derivatives of $(1-x)^{cm_1}x^{cn_1}$ and $(1-x)^{cm_2}x^{cn_2}$ respectively, and these are equal to $((1-x)^{m_1}x^{n_1})^c$ and $((1-x)^{m_2}x^{n_2})^c$, so if we integrate on both sides we get $q(x) = (p(x))^c$ or

$p(x) = (q(x))^c$ for some $c \in \mathbb{N}$ (plus integration constant, but this must equal zero since none of the polynomials in $\mathcal{P}_k^q$ contain a constant term).

$(\longleftarrow)$:

Take two arbitrary polynomials $f, g \in \mathcal{P}_k^q$ and assume $f(x) = (g(x))^c$ or $g(x) = (f(x))^c$ for some $c \in \mathbb{N}$. If we take the derivatives on both sides we get $(1-x)^{m_1-1}x^{n_1-1}(n_1(1-x)-m_1x) = c\cdot(1-x)^{cm_2-1}x^{cn_2-1}(n_2(1-x)-m_2x)$ or $(1-x)^{m_2-1}x^{n_2-1}(n_2(1-x)-m_2x) = c\cdot(1-x)^{cm_1-1}x^{cn_1-1}(n_1(1-x)-m_1x)$ (4.1), therefore it must be the case that the factor $(n_2(1-x)-m_2x)$ divides $(1-x)^{m_1-1}x^{n_1-1}(n_1(1-x)-m_1x)$ or that the factor$(n_1(1-x)-m_1x)$ divides $(1-x)^{m_2-1}x^{n_2-1}(n_2(1-x)-m_2x)$.

It cannot be the case that a factor of the form $(n(1-x)-mx)$ divides one of the form $(1-x)$ or x so we must have that $(n_2(1-x)-m_2x)$ divides $(n_1(1-x)-m_1x)$ or that the factor$(n_1(1-x)-m_1x)$ divides $(n_2(1-x)-m_2x)$. This implies that the derivatives share a common factor of the form $(n(1-x)-mx)$, and thus a common root of the form $x = \frac{n}{n+m}$.

Since this root clearly lies in $(0,1)$, we have that the original polynomials f and g have a common extreme point that lies in $(0,1)$. It follows from Lemma 4.2.2 that this must be a maximum and not a minimum, so these two polynomials have a maximum on the same $x \in (0,1)$. □

Given this mandatory condition for when two polynomials have a maximum on the same x coordinate between 0 and 1 this gives us enough information to be able to construct a precise formula for $\mathcal{U}_{[0,1]}(\mathcal{P}_n^q)$. Namely for a number n it equals the number of polynomials that are not a power of another polynomial from the same set, and this is precisely the case for a polynomial $(1-x)^i x^j$ when i and j are coprime so when $\gcd(i,j) = 1$. To count the number of polynomials for which this holds, we use the Euler's totient function because this function returns for a number n precisely the amount of numbers $i < j$ such that i and j are coprime.

Definition 4.2.2 (Euler's totient function)**.** *The Euler's totient function* $\varphi\colon \mathbb{N} \to \mathbb{N}$ *is defined by:*

$$\varphi(n) = |\{1 \leq m \leq n : \gcd(m,n) = 1\}|.$$

Remark. *For more background reading consult [Apo76].*

Proposition 4.2.2. $\mathcal{U}_{[0,1]}(\mathcal{P}_n^q) = \sum_{i=2}^{n} \varphi(i)$.

Proof. By Proposition 4.2.1, we know that two polynomials in $\mathcal{P}_n^q$ share a maximum if and only if one is a power of the other, so if a polynomial $(1-x)^{i_1}x^{j_1}$ is not a power of any other polynomial $(1-x)^{i_2}x^{j_2}$ then this polynomial has a new maximum. This is precisely the case when $(i_1, j_1) \neq c(i_2, j_2)$ for any $c, i_2, j_2 \in \mathbb{N}$ (exponent rules). This means that $\gcd(i_1, j_1) = 1$ because no number c besides 1 divides both.

For a certain size n we can count the number of times that $\gcd(i_1, j_1) = 1$ with $i_1, j_1 \leq n$ by using the Euler's totient function for each i lower than n starting from 2, and we obtain

$$\varphi(2) + \varphi(3) + ... + \varphi(n) = \sum_{i=2}^{n} \varphi(i).$$

□

The corresponding family of SAU-pMDPs From the sequence of sets of somewhere optimal polynomials we now have (written as $(\mathcal{P}_n^{q\prime})_{n\in\mathbb{N}}$), we construct a family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$ where for every size n, each policy corresponds to exactly one of the polynomials in the set $\mathcal{P}_n^{q\prime}$ (Theorem 4.1.1 together with Corollary 4.2.2.2 ensure that we can construct such a family). In order to satisfy our goal of showing that there exists a family of SAU-pMDPs for which the MOSS grows superlinearly with respect to states, we can choose our family in such a way that it generates precisely this sequence of polynomials, because as we will see $\sum_{i=2}^{n} \varphi(i)$ grows superlinearly. We give a few concrete examples of members of this family. The general idea is that we must first take a $1-x$ transition probability because every polynomial in $\mathcal{P}_n^{q\prime}$ contains this. From then we choose how many more $1-x$ transitions we take before we *step out* by choosing an action that brings us to a new location (with transition probability defined by the scalars to obtain $\mathcal{P}_n^{q\prime}$ from $\mathcal{P}_n^{q}$) in the next row where there are only transitions forward with probability x. However, when we are still in the $1-x$ row we can only choose actions that brings us to states where the remaining number of x transitions will be higher than the number of $1-x$ transitions that we have travelled through. This is because of the $i < j$ condition in the set construction.

Definition 4.2.3 (The quadratic family). *Consider the family of SAU-pMDPs $(\mathcal{M}_n^q)_{n\in\mathbb{N}}$ with set of states $S = \{q_0, q_1, ..., q_{n-1}, p_0, p_1, ..., p_n\} \cup \{\bot\}$, initial state $s_0 = q_0$, set of actions $Act = \{\alpha_i \mid 0 \leq i < n\}$, goal state $T = \{p_n\}$ and transition probability function*

$$\begin{cases} P(q_i, \alpha_0, q_{i+1}) = 1 - x, \\ P(p_i, \alpha_0, p_{i+1}) = x, \\ P(q_i, \alpha_j, p_j) = c_{i,j} \;\; \text{if } c_{i,j}(1-x)^i x^{n-j} \in \mathcal{P}_n^{q\prime}. \end{cases}$$

Each transition probability has a complementary transition probability for the sink state and the unspecified transitions have probability 0.

Remark. *Note that $|S|$ and $|Act|$ clearly grow linear with respect to n, hence the size does as well.*

In Figure 4.7 and Figure 4.8, we can see a visualization of how the members of the family change from $n = 2$ to $n = 4$. It can easily be checked that the quadratic family conveniently generates the sequence of sets of polynomials $(\mathcal{P}_n^{q\prime})_{n\in\mathbb{N}}$ by definition of its transition probability function.

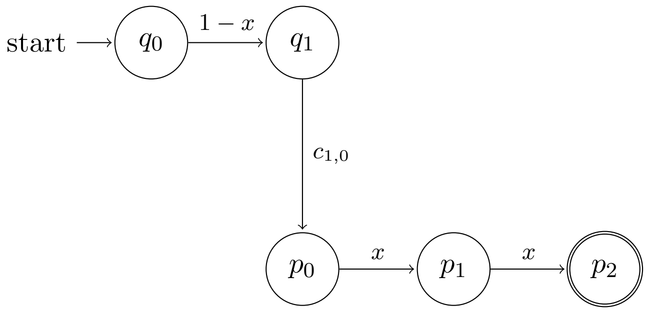


Figure 4.7: Graphical representation of $\mathcal{M}_2^q$ (sink edges omitted)

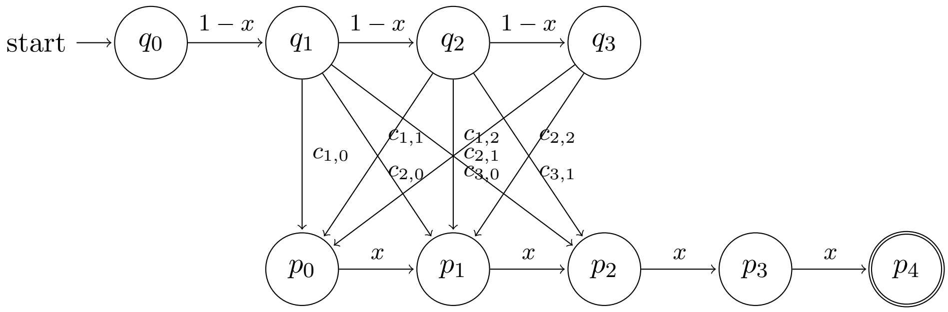


Figure 4.8: Graphipcal representation of $\mathcal{M}_4^q$ (sink edges omitted)

We now arrive at the most important result in this section. We can deduce from previously established results that the growth rate of $|MOSS^{\mathcal{M}_n^q}|$ must be precisely equal to $\sum_{i=2}^{n} \varphi(i)$, and for this function the asymptotic growth rate is known.

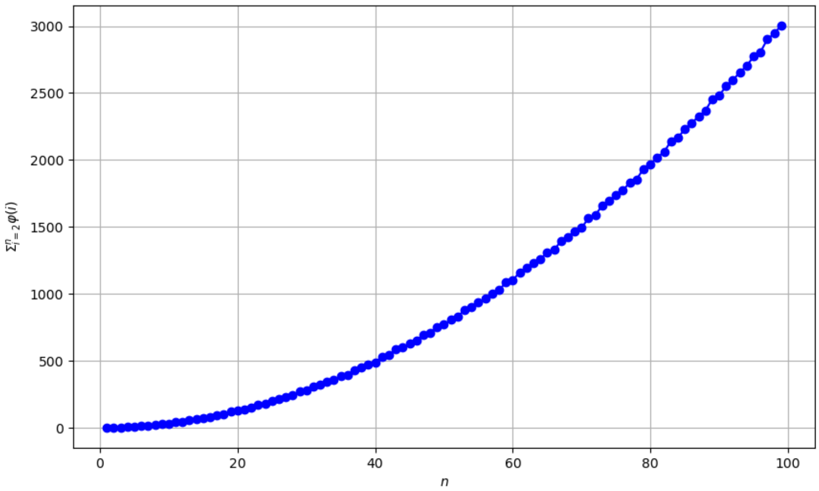


Figure 4.9: $\sum_{i=2}^{n} \varphi(i)$ plotted to visualize its clear superlinear growth

Theorem 4.2.1. *In the quadratic family we have that* $|MOSS^{\mathcal{M}_n^q}| = \sum_{i=2}^{n} \varphi(i)$.

Proof. Each polynomial in $\mathcal{P}_n^{q\prime}$ has roots at $x = 0$ and $x = 1$ so it follows from this and Lemma 4.2.2 that each polynomial in $\mathcal{P}_n^{q\prime}$ has exactly one maximum in $[0, 1]$. Therefore, we may apply Theorem 4.1.2 and we get $|MOSS^{\mathcal{M}_n^q}| = \mathcal{U}_{[0,1]}(\mathcal{P}_n^{q\prime})$ which equals $\mathcal{U}_{[0,1]}(\mathcal{P}_n^{q})$ in our case, and according to Proposition 4.2.2, this is equal to $\sum_{i=2}^{n} \varphi(i)$. □

From this we can derive the result that in general the size of a MOSS has a quadratic lower bound with respect to size. More specifically we know that in the quadratic family, we have a quadratically growing MOSS (both upper and lower bound) with respect to size.

Corollary 4.2.1.1. *Given an arbitrary family of SAU-pMDPs* $(\mathcal{M}_n)_{n\in\mathbb{N}}$*, we have that* $|MOSS^{\mathcal{M}_n}| \in \Omega(k^2)$ *with* k *the size of* $\mathcal{M}_n$*.*

Proof. By Theorem 4.2.1, we have for $(\mathcal{M}_n^q)_{n\in\mathbb{N}}$, that $|MOSS^{\mathcal{M}_n^q}| = \sum_{i=2}^{n} \varphi(i)$, and it is known (see [BDH+19]) that

$$\sum_{i=1}^{n} \varphi(i) = \frac{n^2}{2\zeta(2)} + O(n(log(n))^{\frac{2}{3}}(log(log(n)))^{\frac{4}{3}}),$$

so

$$\sum_{i=2}^{n} \varphi(i) = \frac{n^2}{2\zeta(2)} + O(n(log(n))^{\frac{2}{3}}(log(log(n)))^{\frac{4}{3}}) - 1,$$

and clearly

$$\frac{n^2}{2\zeta(2)} + O(n(log(n))^{\frac{2}{3}}(log(log(n)))^{\frac{4}{3}}) - 1 \in \Theta(n^2),$$

so consequently $|MOSS^{\mathcal{M}_n^q}| \in \Theta(n^2)$, hence in general for a family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$, we have that $|MOSS^{\mathcal{M}_n}| \in \Omega(k^2)$ where k equals the size of $\mathcal{M}_n$ (since the size has a linear relationship with n in the quadratic family). □

4.3 A Potentially Exponential Bound

The question that remains is if a quadratic, or more generally a polynomial upper bound could also be ruled out. For this purpose, we try to construct a new sequence of sets of polynomials with its corresponding family of SAU-pMDPs, for which the size of its MOSS would grow exponentially. In particular, we construct a sequence of sets of polynomials with an exponential cardinality with respect to n, and based on numeric evidence we have strong reasons to believe that its unique maxima quantity also grows exponentially. However, to prove this formally remains an open problem as it may require very nontrivial algebraic methods. If each polynomial were to be adequate, Theorem 4.1.1 would then imply that our sequence can be generated using a family of SAU-pMDPs. This family could then have an exponentially growing MOSS. However, this does not hold with respect to actions, but only states, and as a consequence also not with respect to size. For a potentially exponential lower bound with respect to actions, an alternative candidate family is presented following our conjecture about the bound with respect to states.

4.3.1 Bound with respect to states

We first define our sequence of sets of polynomials which we will refer to as the exponential sequence.

Definition 4.3.1 (The exponential sequence)**.** *Consider the sequence of sets of polynomials $(\mathcal{P}_n^e)_{n\in\mathbb{N}}$ defined by $(n \geq 1)$:*

$$\mathcal{P}_n^e = \{\frac{1}{k}(1-x)(b_1x^1 + b_2x^2 + ... + b_nx^n) \,|\, b_i \in \{0,1\},\, k = |\{b_i : b_i = 1\}|\},$$

with the additional constraint that at least one b_i is equal to 1.

Remark. $\mathcal{P}_0^e$ *is undefined.*

Example 4.3.1. $\mathcal{P}_2^e = \{(1-x)(x), (1-x)(x^2), \frac{1}{2}(1-x)(x+x^2)\}$.

This sequence is useful because each element seems to be an adequate polynomial with exactly one local maximum in $(0,1)$, and it is easy to prove that the cardinality of the sets grows exponentially when n increases[3].

[3]There is no rigorous proof as well for the claim that each polynomial has one local maximum in $(0,1)$, but refer to Figure 4.11 for an numerical argument.

Lemma 4.3.1. $|\mathcal{P}_n^e| = 2^n - 1$.

Proof. If we combinatorically analyze the set construction of $\mathcal{P}_n^e$ we see that there is a case distinction. We can choose one b_i to be equal to 1, or two terms b_i, b_j to be both equal to 1, etc., up to choosing n terms $b_1, ..., b_n$ to be all equal to 1. Each of those cases can be done in $\binom{n}{i}$ ways with i the number of terms. So added together we have $\sum_{i=1}^{n} \binom{n}{i}$ ways and we know that $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ (follows from the binomial theorem [HK16]). Choosing 0 terms can only be done in one way, so we subtract it since we must at least choose one b_i to be equal to 1, and we get $\sum_{i=1}^{n} \binom{n}{i} = 2^n - 1$. □

Resulting family of SAU-pMDPs Because of Lemma 4.1.3, we can acquire the equal maxima set $\mathcal{P}_n^{e\prime}$ for each n. A family of SAU-pMDPs that generates the sequence of equal maxima sets of the exponential sequence $(\mathcal{P}_n^{e\prime})_{n\in\mathbb{N}}$ is $(\mathcal{M}_n^e)_{n\in\mathbb{N}}$, which we call the exponential family. We make use of actions and their flexibility in how many outgoing transitions they can have to construct it. Again, we can easily see that this family generates $(\mathcal{P}_n^{e\prime})_{n\in\mathbb{N}}$ because of how the transition probability function is defined.

Definition 4.3.2 (The exponential family)**.** *Consider the family of SAU-pMDPs $(\mathcal{M}_n^e)_{n\in\mathbb{N}}$, with set of states $S = \{q_0, q_1, p_0, p_1, ..., p_n\} \cup \{\bot\}$, initial state $s_0 = q_0$, goal state $T = \{p_n\}$, set of actions $Act = \{\alpha\} \cup \{\beta_c \,|\, cf \in \mathcal{P}_n^{e\prime}$ for some $f \in \mathcal{P}_n^e\}$ and transition probability function*

$$\begin{cases} P(q_0, \alpha, q_1) = 1 - x, \\ P(p_i, \alpha, p_{i+1}) = x, \\ P(q_1, \beta_c, p_i) = \frac{c}{k} \text{ if } b_{n-i} = 1 \text{ in } f \text{ s.t. } cf \in \mathcal{P}_n^{e\prime}. \end{cases}$$

Furthermore, each transition probability has a complementary probability of going to the sink state and all non specified transitions have probability 0. Note that in the transition probability function, k in $\frac{c}{k}$ means k in f (s.t. $cf \in \mathcal{P}_n^{e\prime}$) as in Definition 4.3.1.

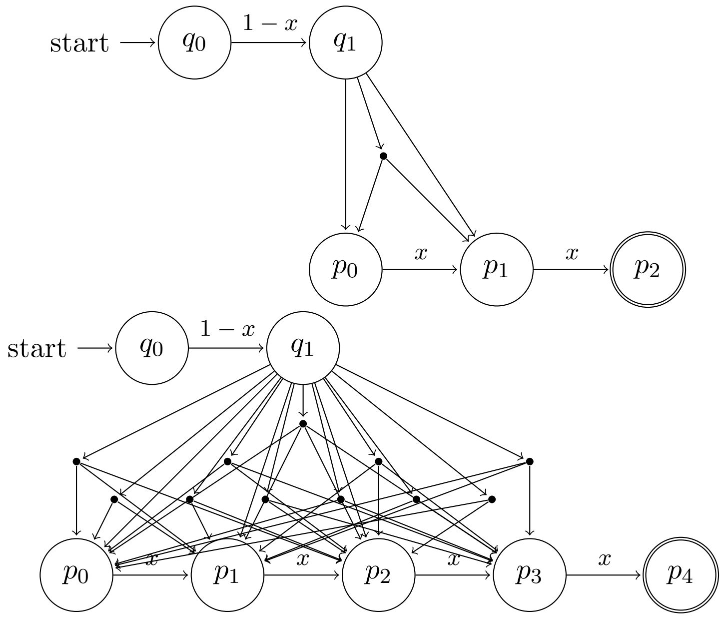


Figure 4.10: SAU-pMDPs (sink edges omitted) from the exponential family with size 2 (above) and 4 (below)

Remark. *This definition is the reason we use* $\mathbb{R}$ *instead of the conventional* $\mathbb{Q}$ *in the definition of the transition probability function (Definition 2.4.4 and Definition 2.4.1). For example* $\frac{1}{2}(1-x)(x+x^2)$ *has as its local maximum in* $(0,1)$ *the coordinate* $(\frac{1}{3\sqrt{3}}, \frac{1}{\sqrt{3}})$*, so in order to create* $\mathcal{P}_n^{e}{}'$ *one must use irrational numbers since* $\sqrt{3} \notin \mathbb{Q}$*.*

Factorisation The problem however, is that it is very complicated to algebraically compute the value of $\mathcal{U}_{[0,1]}(\mathcal{P}_n^e)$ (which is equal to $\mathcal{U}_{[0,1]}(\mathcal{P}_n^{e}{}')$). This arises from the fact that for the set of derivatives of the polynomials (needed to find the maxima), we have not yet found a general method for factorising them. Consequently, we cannot easily find a pattern for when two polynomials have a common factor with its roots lying in $(0,1)$ like we did with Proposition 4.2.1. Or differently put, a condition for when the GCD of two polynomials equals 1. This makes it hard to show when maxima occur at the same x value.

Numerical evidence Nonetheless, when calculating the roots through numerical methods by using a python program, it becomes evident that concurring roots in $[0,1]$ of the derivative set of the polynomials are very rare. In particular, it occurs only 3 times at $n = 20$. Judging from the table below, it appears that $\mathcal{U}_{[0,1]}(\mathcal{P}_n^e)$ grows exponentially and thus $|MOSS^{\mathcal{M}_n^e}|$ as well by Theorem 4.1.2. We can possibly apply this theorem, because judging from the table, it is likely that all of the polynomials obtain exactly one local maximum in $[0, 1]$ (which is a condition). We can observe this, because each polynomial is adequate, so they have at least one maximum (between 0 and 1), and since the unique maxima quantity is less than or equal to the number of polynomials it does not exceed 1.

Growth of $\mathcal{U}_{[0,1]}(\mathcal{P}_n^e)$ with respect to n						
n	2	3	4	10	15	20
$\lvert\mathcal{P}_n^e\rvert$	3	7	15	1023	32767	1048574
$\lvert\{p \in \mathcal{P}_n^e : p \textit{ is adequate}\}\rvert$	3	7	15	1023	32767	1048574
$\mathcal{U}_{[0,1]}(\mathcal{P}_n^e)$	3	7	15	1023	32765	1048571
$\lvert\mathcal{P}_n^e\rvert - \mathcal{U}_{[0,1]}(\mathcal{P}_n^e)$	0	0	0	0	2	3

Figure 4.11: Table showing that the number of x values of maxima that are shared between multiple polynomials is negligible up to $n = 20$. It also shows that all polynomials are adequate and contain one local maximum in $[0, 1]$ up to $n = 20$.

Since in Definition 4.3.2, it is clear that the cardinality of the set of states has a linear relationship with n and the cardinality of the set of actions has an exponential relationship with n, we conclude with a conjecture regarding only a new bound in states.

Conjecture 4.3.1. *For the exponential family $(\mathcal{M}_n^e)_{n\in\mathbb{N}}$ we have that $|MOSS^{\mathcal{M}_n^e}| \in \Theta(2^{|S_n|})$, and consequently, that the size of a MOSS is in general exponentially lower bounded with respect to states.*

4.3.2 Bound with respect to actions

Because in Definition 4.3.2 the size of the set of actions has an exponential relationship with n, we can see that with respect to actions, the MOSS in the previous case still grows only linearly. One possible route of continuation now is to try to reduce the number of actions for the exponential family and still generate the same sequence $(\mathcal{P}_n^e)_{n\in\mathbb{N}}$ (or a similar sequence). The goal is to let different policies use the same action as often as possible. In this section we provide an (incomplete) alternative family that generates a similar sequence to the exponential sequence $(\mathcal{P}_n^e)_{n\in\mathbb{N}}$, which in the future could lead to a configuration of transition probabilities being found for this family such that the MOSS grows exponentially with respect to actions.

Alternative exponential family Consider again the exponential sequence $(\mathcal{P}_n^e)_{n\in\mathbb{N}}$, but now construct a family of SAU-pMDPs (that generates a sequence of supersets of it) in a different manner (Figure 4.12). Here only linearly many actions remain. We call this family the alternative exponential family.

Definition 4.3.3 (The alternative exponential family)**.** *Consider the family of SAU-pMDPs* $(\mathcal{M}^e_n*)_{n\in\mathbb{N}}$ *with set of states* $S = \{s_0, s_1, s_2\} \cup \{\bot\} \cup \bigcup_{i,j=0}^{n} q_{i,j}$*, initial state* s_0*, goal state* $T = \{s_2\}$*, set of actions* $Act = \{\alpha_i \,|\, 1 \leq i \leq n\}$ *and transition probability function*

$$\begin{cases} P(s_0, \alpha, s_1) = 1 - x, \\ P(s_1, \alpha, q_{i,0}) = \frac{1}{n}, \\ P(q_{i,n}, \alpha, s_2) = x, \\ P(q_{i.j}, \alpha, q_{i,j+1}) = x \text{ if } j > 0, \\ P(q_{i,0}, \alpha_j, q_{i,j}) = c_{i,j} \text{ s.t. } c_{i,j} \in \mathbb{R}. \end{cases}$$

Each transition probability has a complementary probability of going to the sink state, unspecified transitions have probability 0 *and the values of the constants* $c_{i,j}$ *are left undetermined as of now.*

Remark. *When each c is equal to* 1*, this generates a sequence of supersets of* $(\mathcal{P}^e_n)_{n\in\mathbb{N}}$.

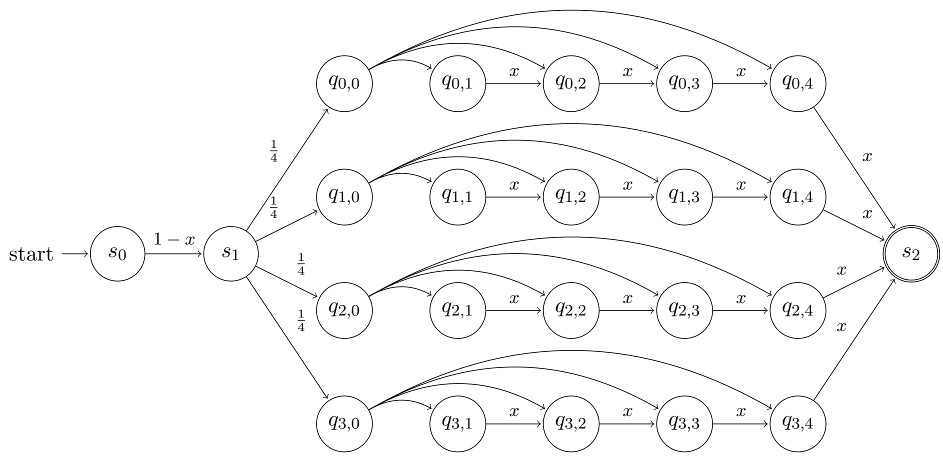


Figure 4.12: Grahpical representation of $\mathcal{M}^e_4*$ (sink edges omitted)

Lemma 4.3.2. *In the alternative exponential family, we have that* $\Sigma^{\mathcal{M}^e_n*} = n^n$.

Proof. There are n states $(q_{i,0})$ where we are allowed to choose between n different actions each. This results in $n \cdot n \cdot ... \cdot n$ (n times) $= n^n$ policies in total. □

It might be possible to choose the not yet decided transition probabilities in Definition 4.3.3 in such a way that we generate a set of polynomials $\mathcal{P}^e_n*$, such that exponentially many polynomials from this set become somewhere optimal with respect to actions. To find this transition probability function, it is likely that we need to relax the requirement that all maxima of the polynomials need to obtain the exact same y value. See Appendix A for an example configuration of transition probabilities for $n = 3$ that would still result in 7 elements in the MOSS, which is equal to $|MOSS^{\mathcal{M}_3^e}|$. We conclude this section with the following statements.

Proposition 4.3.1. *If there exists a transition probability function P such that $\mathcal{M}_n^e*$ generates a set $\mathcal{P}_n^e*$ for which it holds that*

$$|\{f \in \mathcal{P}_n^e* : f \text{ is somewhere optimal}\}| \in \Theta(2^n),$$

then

$$\exists \Omega^{\mathcal{M}_n^e*} \subseteq \Sigma^{\mathcal{M}_n^e*} : |\Omega^{\mathcal{M}_n^e*}| \in \Theta(2^{|Act_n|}).$$

Proof. (We abbreviate $\{f \in \mathcal{P}_n^e* : f$ is somewhere optimal$\}$ to $\mathcal{P}_n^e *+$) Assume that such a transition probability function P exists. Since a somewhere optimal polynomial corresponds to a somewhere optimal policy by Lemma 4.1.2, we can construct an OSS $\Omega^{\mathcal{M}_n^e*}$ that consists of all of these optimal policies corresponding to the polynomials in $\mathcal{P}_n^e *+$. Since $\mathcal{P}_n^e *+$ has an exponential cardinality and $|\mathcal{P}_n^e *+| = |\Omega^{\mathcal{M}_n^e*}|$, it holds that $|\Omega^{\mathcal{M}_n^e*}| \in \Theta(2^n)$. Since $|Act_n|$ has a linear relationship with n, we have that $|\Omega^{\mathcal{M}_n^e*}| \in \Theta(2^{|Act_n|})$. □

Conjecture 4.3.2. *If there exists a transition probability function P such that $\mathcal{M}_n^e*$ generates a set $\mathcal{P}_n^e*$ for which it holds that*

$$|\{f \in \mathcal{P}_n^e* : f \text{ is somewhere optimal}\}| \in \Theta(2^n),$$

then we have that

$$|MOSS^{\mathcal{M}_n^e*}| \in \Theta(2^{|Act_n|})$$

Remark. *The reason this is a conjecture and not a proposition is because we do not know for certain that the number of different x coordinates for which a maxima is obtained is growing exponentially, and if the number of maxima that each polynomial has between* 0 *and* 1 *is exactly one. Figure 4.11 provides some evidence for this however, since the polynomials in $\mathcal{P}_n^e*$ are very similar to the polynomials in $\mathcal{P}_n^e$.*

If this conjecture, Conjecture 4.3.1 and their assumptions are all true then that would imply that for an arbitrary family of SAU-pMDPs, a MOSS has an exponential lower bound with respect to size. The alternative exponential family seems to be a very promising candidate because with n^n total policies (Lemma 4.3.2), which is greater than 2^n, we have a lot of free space.

Chapter 5

Discussion

As the bound of a MOSS has a newly proven quadratic lower bound with respect to the size of a pMDP, our research question is partly answered. However, a complete answer that also includes an upper bound has not been found yet. Aside from proving an upper bound, there remain other parts that are still open such as a proof for the exponential lower bound with respect to states, or any information on an exponential bound with respect to actions. Several questions were not considered in this thesis, such as what the bound with respect to parameters is or, how a MOSS behaves when the pMDP is not univariate. While not considered, these questions might also be the focus of future exploration. In this chapter we will discuss several of these questions, whether our work has implications and parts that are still left to be completed after our research.

5.1 More Parameters

As the previous results were all about SAU-pMDPs, it was always the case that $|X| = 1$ by definition. The lemma (Lemma 1.0.1) that spawned the research question did not require this. This still leaves us to wonder how fast a MOSS grows when for example $|X| = 2$ (bivariate pMDPs). If Conjecture 4.3.1 were to be proven that would imply that for any fixed number of parameters the MOSS can at least grow exponentially with respect to states since if the lower bound holds for $|X| = 1$, it must also hold for $|X| > 1$.

Bound with respect to parameters If instead of looking at fixed parameter cases, we look at families of well-defined, acyclic pMDPs where the number of parameters grows as a function of n, we can perhaps also find a bound for the MOSS with respect to the number of parameters. Our research has been solely dedicated to finding bounds with respect to states and actions because the lemma that inspired the search for such bounds (Lemma 1.0.1) specifically mentions fixed parameter cases. Nonetheless, it is likely that the lower bound with respect to parameters is exponential, and this can be illustrated by an example (Figure 5.1). It is then only left to rigorously prove that this is true.

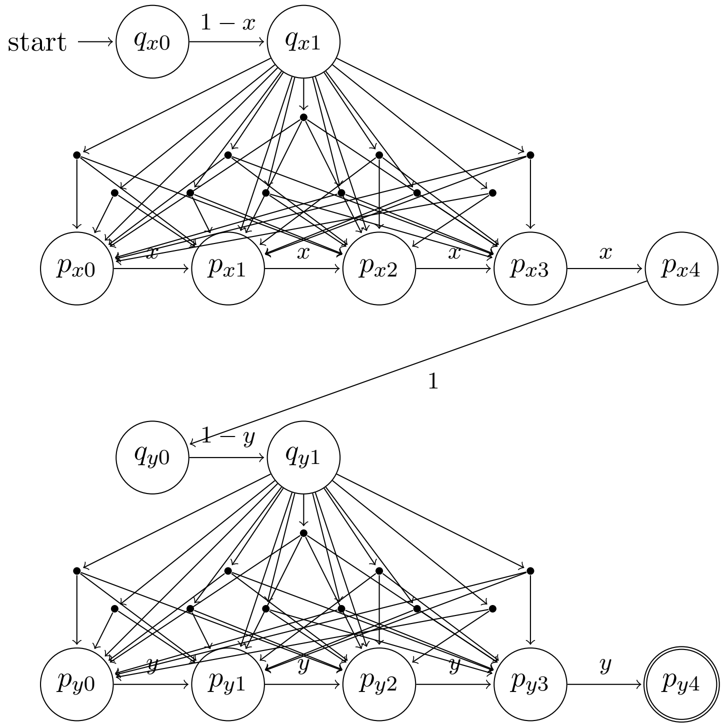


Figure 5.1: Member of a family of well-defined, acyclic pMDPs (sink edges omitted) for which its MOSS might grow exponentially with respect to $|X|$. In this family, every time we add a parameter, we add a copy of the initial structure.

5.2 Proving Further Bounds

A quadratic lower bound with respect to our definition of size has been rigorously proved. However, despite efforts, we have not been able to mathematically prove that there exists a family of SAU-pMDPs where its MOSS has an exponential lower bound. Nonetheless, it is very likely that it is the case (with respect to states), as is shown in Figure 4.11. Proving this exponential lower bound rigorously would likely require advanced algebra techniques which we have not been able to find/apply during this research. This is because the problem likely comes down to analyzing the general factorisation of a complicated set of polynomials. Namely the set of derivatives of polynomials in $\mathcal{P}_n^e$. The proof might involve the use of the Euclidean algorithm [Pra04], as this can be used to compute the GCD for concrete polynomials. In the arbitrary case however (so for not precisely known polynomials), it seems very complicated for this set.

Changing the conditions Factorising the polynomials might however not even be necessary in order to prove that the exponential family yields an exponential lower bound on the size of its MOSS with respect to states. This is the case, since this is built on the assumption that all the values of $f(x)$ of the maxima of the polynomials in $\mathcal{P}_n^e$ need to be the exact same value. If we relax this condition and add a very small number $\varepsilon > 0$ to the denominator of the scalars, we can solve the issue of concurring maxima on the x-axis. What this means is, we can make two polynomials with their maximum on the same x coordinate both somewhere super optimal, as can be seen in Figure 5.2.

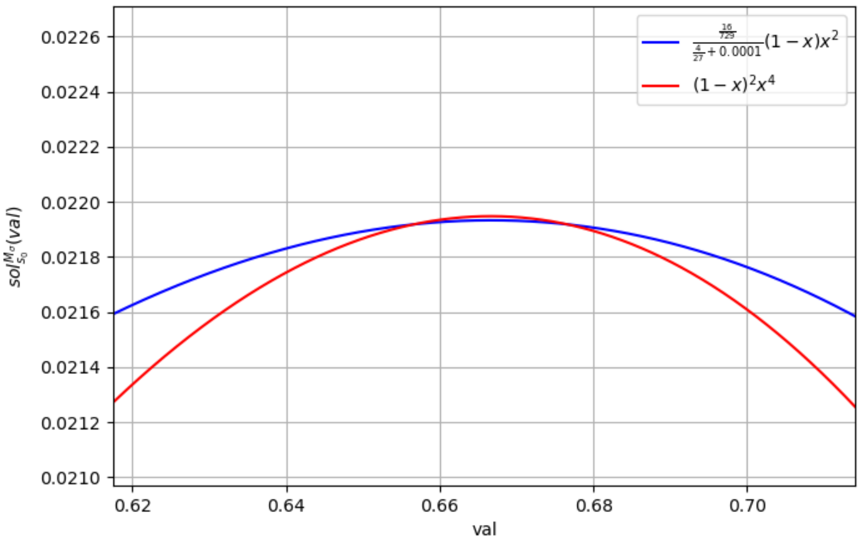


Figure 5.2: Two polynomials with their local maximum between 0 and 1 at the same x coordinate but with *epsilon adjusted scalar*. For a small interval the red polynomial has a larger value of $sol_{s_0}^{\mathcal{M}_\sigma}(val)$ than the blue polynomial.

Further research in this direction is still open for future consideration. A downside of this method is that it can become increasingly complicated when the number of polynomials with their maxima at the same x coordinate becomes higher than 2.

Super exponential bounds While Conjecture 4.3.1 mentions a new exponential lower bound, this does not mean that there does not exist a family of SAU-pMDPs for which we have, for example $n!$ growth or n^n. Finding such a SAU-pMDP family is also something that can be tried in future research. Another way of stating this is that there is as of yet no clear new upper bound.

Exponential bound with respect to actions For the exponential lower bound with respect in actions, there is still a lot unknown after our research but a promising candidate family of SAU-pMDPs is presented nonetheless. Meaning, if an exponential lower bound exists it could potentially be found in this particular family. This could perhaps be achieved by building on the pattern in Appendix A. The family we presented does not have to be the only candidate however.

5.3 Implications

If a transition probability function for the alternative exponential family could be found then that could mean that in general for a family of SAU-pMDPs $(\mathcal{M}_n)_{n\in\mathbb{N}}$ that $|MOSS^{\mathcal{M}_n}| \notin \Theta(poly(n))$. This would then not satisfy the conditions of Lemma 1.0.1 because we could then have that the MOSS also grows exponentially with respect to how size is defined for that case, because the total number of transition probabilities grows only polynomially. This still leaves the possibility open that $\exists\exists Reach_*^{\bowtie}$ and $\exists\forall Reach_*^{\bowtie}$ are not in class coNP because the size of the MOSS is not polynomially bounded for fixed-parameter pMDPs [JKPW19]. With regards to the proven quadratic bound, for how size is defined in the case of Lemma 1.0.1, it would still only be a linear bound (as the number of non-zero transitions grows quadratically in the quadratic family). This would still satisfy the conditions of the lemma because a linear bound is also a polynomial bound. Therefore, the restriction we made in the definition of size does not make a huge difference.

Chapter 6

Conclusions

In this thesis, we have proven that there exists a family of SAU-pMDPs $(\mathcal{M}_n^q)_{n\in\mathbb{N}}$ for which it holds that $|MOSS^{\mathcal{M}_n^q}|$ grows quadratically with respect to size. This proves that in general, the size of a MOSS has a quadratic lower bound for an arbitrary family of SAU-pMDPs. Numerical evidence shows why it is likely that there also exists a family of SAU-pMDPs $(\mathcal{M}_n^e)_{n\in\mathbb{N}}$ for which $|MOSS^{\mathcal{M}_n^e}|$ has an exponential growth rate with respect to the number of states. In particular, the data shows that the number of polynomials from the corresponding set of polynomials that do not become somewhere optimal is negligible up to $n = 20$. A formal proof for this exponential lower bound with respect to states has not been found during our research and thus, remains open. The question whether an exponential lower bound on the MOSS with respect to actions also exists is unanswered as well, but a promising candidate family of SAU-pMDPs is presented that could lead to an exponential lower bound being found if it exists. This thesis might provide a basis for future research with regards to the open questions.

Bibliography

[Abe26] N. H. Abel. *Beweis der Unmöglichkeit, algebraische Gleichungen von höheren Graden als dem vierten allgemein aufzulösen.* Journal für die reine und angewandte Mathematik, 1826.

[Apo76] Tom M. Apostol. *Introduction to Analytic Number Theory*, chapter 2. Springer, 1976.

[BDH$^+$19] Olivier Bordelles, Lixia Dai, Randell Heyman, Hao Pan, and Igor e. Shparlinksi. *On a Sum Involving the Euler Function.* 2019.

[Bel58] Richard Bellman. *Dynamic programming and stochastic control processes.* Information and Control, 1(3):228–239, 1958.

[BK08] Christel Baier and Joost-Pieter Katoen. *The principles of model checking.* MIT press, 2008.

[Daw04] Conrado Daws. *Symbolic and Parametric Model Checking of Discrete-Time Markov Chains.* Springer, 2004.

[DM89] Peter Doyle and Curt McMullen. *Solving the quintic by iteration.* Acta Math, 1989.

[FR97] Benjamin Fine and Gerhard Rosenberger. *The Fundamental Theorem of Algebra.* Springer, 1997.

[GOP11] Hugo Gimbert, Youssouf Oualhadj, and Soumya Paul. *Computing Optimal Strategies for Markov Decision Processes with Parity and Positive-Average Conditions.* 2011.

[Gou12] Ronald Gould. *Graph Theory.* Dover, 2012.

[HHZ11] Ernst Moritz Hahn, Tingting Han, and Lijun Zhang. *Synthesis for PCTL in Parametric Markov Decision Processes.* Springer, 2011.

[HK16] H. S. Hall and S. R . Kight. *Higher Algebra 6th edition.* Arihat Publication, 2016.

[How72] R. A. Howard. *Dynamic Probabilistic Systems, volume 2: Semi-Markov and Decision Processes.* John Wiley & Sons, 1972.

[HR04] Michael Huth and Mark Ryan. *Logic in Computer Science: Modelling and Reasoning about Systems.* Cambridge University Press, 2004.

[JKPW19] Sebastian Junges, Joost-Pieter Katoen, Guillermo A. Pérez, and Tobias Winkler. *On the Complexity of Reachability in Parametric Markov Decision Processes.* 2019.

[JKPW20] Sebastian Jungesa, Joost-Pieter Katoen, Guillermo A. Pérez, and Tobias Winkler. *The Complexity of Reachability in Parametric Markov Decision Processes.* 2020.

[Jun20] Sebastian Junges. *Parameter Synthesis in Markov Models.* PhD thesis, Aachen University, 2020.

[J$^+$19] Sebastian Junges, Erika Ábrahám, Christian Hensel, Nils Jansen, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. *Parameter Synthesis for Markov Models.* 2019.

[Khi57] A. Ya. Khinchin. *Mathematical Foundations of Information Theory.* Dover, 1957.

[kS60] J. kemeny and J. Snell. *Finite Markov Chains.* Springer, 1960.

[KU02] Igor O. Kozine and Lev V. Utkin. *Interval-Valued Finite Markov Chains.* Reliable Computing, 2002.

[KY76] D. Knuth and A. Yao. *Algorithms and Complexity: New Directions and Recent Results*, chapter The complexity of nonuniform random number generation. Academic Press, 1976.

[McN07] J.M. McNamee. *Numerical Methods for Roots of Polynomials - Part 1.* Elsevier, 2007.

[MP13] J.M. McNamee and Victor Pan. *Numerical Methods for Roots of Polynomials - Part 2.* Elsevier, 2013.

[Nor97] J. R. Norris. *Makov Chains.* Cambridge University Press, 1997.

[PLS13] Ravi Kumar P, Alex Goh Kwang Leng, and Ashutosh Kumar Singh. *Application of Markov Chain in the PageRank Algorithm.* Pertanika Journal of Science & Technology, 2013.

[Pra04] Victor V. Prasolov. *Polynomials.* Springer, 2004.

[Put94] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* John Wiley Sons, 1994.

[SB08] Olivier Sigaud and Olivier Buffet. *Markov Decision Processes in Artificial Intelligence*. John Wiley Sons, 2008.

[Sen16] Eugene Seneta. *Markov Chains as Models in Statistical Mechanics*. Institute of Mathematical Statistics, 2016.

[Ste08] James Stewart. *Calculus Early Trancendentals*, chapter 4. Thomson Brooks/Cole, sixth edition, 2008.

[Thr05] Sebastian Thrun. *Probabilistic Robotics*. MIT Press, 2005.

[vOW12] Martijn van Otterlo and Marco Wiering. *Reinforcement Learning and Markov Decision Processes*. Springer Berlin Heidelberg, 2012.

[Yoo09] Byung-Jun Yoon. *Hidden Markov Models and Their Applications in Biological Sequence Analysis*. Curr Genomics, 2009.

Appendix A

Appendix

Example transition probability configuration for the alternative exponential family (Definition 4.3.3)
For the instance $n = 3$ consider the probability transition function:

$$\begin{cases} \mathrm{P}(\mathrm{q}_{i,0}, \alpha, q_{i,1}) = 1 \text{ for } i \in [0, 2], \\ \mathrm{P}(\mathrm{q}_{i,0}, \alpha, q_{i,3}) = \frac{\frac{27}{256}}{\frac{1}{4}} \text{ for } i \in [0, 2], \\ \mathrm{P}(\mathrm{q}_{0,0}, \alpha, q_{0,2}) = \frac{\frac{27}{256}}{\frac{4}{27}}, \\ \mathrm{P}(\mathrm{q}_{1,0}, \alpha, q_{1,2}) = \frac{\frac{27}{256}}{\frac{4}{27-0.001}}, \\ \mathrm{P}(\mathrm{q}_{2,0}, \alpha, q_{2,2}) = \frac{\frac{27}{256}}{\frac{4}{27-0.0001}}. \end{cases}$$

The other transition probabilities are as specified in Definition 4.3.3.

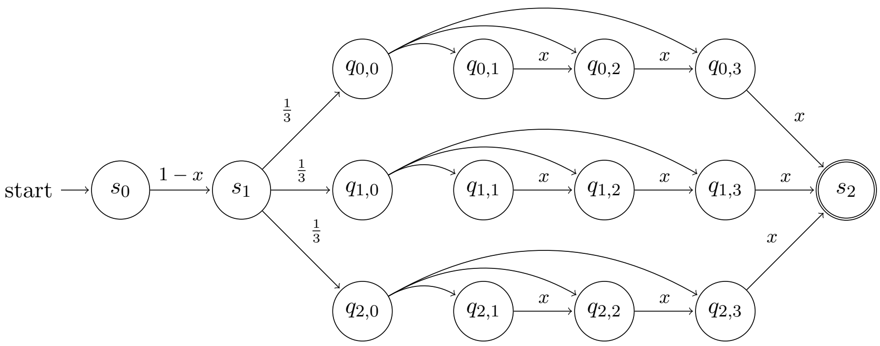


Figure A.1: Alternative SAU-pMDP that generates a superset of $\mathcal{P}_3^e$ (sink edges omitted)

We now have that for the set of polynomials $\mathcal{P}$ that this generates, that $\mathcal{U}_{[0,1]}(\mathcal{P}) = 7$ which is equal to $\mathcal{U}_{[0,1]}(\mathcal{P}_3^e)$.