

Daniel Willems

Abstraktion zeitstetiger Markov-Ketten

DIPLOMARBEIT

MOVES: Software Modeling and Verification

Lehrstuhl II für Informatik der

Rheinisch-Westfälischen Technischen Hochschule Aachen

April 2006

Betreuung durch

Prof. Dr. Ir. Joost-Pieter Katoen

Hiermit versichere ich, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, den 14. September 2006

Inhaltsverzeichnis

1	Einleitung	2
2	Grundlagen	5
2.1	CTL Model Checking	5
2.2	Markov-Ketten	13
2.3	CSL (Continuous Stochastic Logic)	19
2.4	Transient Analyse und Uniformisierung für CTMCs	22
2.5	Abstraktion mit Bisimulationsäquivalenzklassen	26
2.6	Markov-Entscheidungsprozesse	27
2.7	Dreiwertige Logiken	28
3	Abstraktion für zeitstetige Markov-Ketten	31
3.1	Abstrakte zeitstetige Markov-Ketten	31
3.2	Scheduler und cut-Funktion	37
3.3	Probabilistische Simulation	52
3.4	Mögliche Abstraktionen	57
3.4.1	Abstraktion A	57
3.4.2	Abstraktion B	63
3.4.3	Abstraktion C	67
3.5	Uniformisierte abstrakte zeitstetige Markov-Ketten	71
3.6	Zusammenhang zwischen CTMC- und DTMC-Abstraktion	78
3.7	Wahrscheinlichkeitsräume	81
4	Model Checking für abstrakte Markov-Ketten	87
4.1	3-CSL (3-valued Continuous Stochastic Logic)	87
4.2	Quantitative, zeitabhängige Erreichbarkeit	88
4.3	3-CSL Model Checking	97
5	Zusammenfassung und Ausblick	111

1 Einleitung

Die Komplexität von Systemen vom Bereich der Softwareentwicklung über Hardwarearchitekturen bis zur Modellierung von natürlichen Systemen in Biologie und Chemie übersteigt heutzutage in der Regel das menschliche Verständnis. Auch wenn noch die großen Zusammenhänge überschaubar erscheinen, so gerät der Nachweis von bestimmten Eigenschaften eines Systems schnell zur Sisyphus-Arbeit oder ist sogar praktisch unmöglich.

Eine der ältesten Techniken, die Eigenschaften eines Systems zu prüfen, dürfte wohl das *Testen* sein. Heutzutage wird sie noch in vielen Bereichen verwendet, beispielsweise im Automobilbau (Crash-Tests) oder in der Softwareentwicklung (Test-Klassen). Die *Simulation* ist eine ans Testen angelehnte Methode, bei der ein mathematisches Modell aufgestellt und das Verhalten von Instanzen des gegebenen Problems in diesem Modell untersucht wird. Beiden Techniken ist gemein, dass kein vollständiger Nachweis der gewünschten Eigenschaften erfolgen kann. Vielmehr wird versucht schwierige Fälle zu finden, was oft noch per Hand geschieht, und für diese Fälle durch Testen oder Simulation die gesuchte Eigenschaft nachzuweisen. Durch die Beschränkung insbesondere der Ressourcen Zeit und Geld können natürlich nur relativ wenige Situationen überprüft werden.

Zwei weitere Techniken, die im Gegensatz zu den bisher genannten einen *formalen Nachweis* ermöglichen und auch bereits den Weg in die kommerzielle Anwendung gefunden haben, sind die *deduktive Verifikation* und das sogenannte *Model Checking*. In beiden Fällen muss ähnlich wie bei der Simulation zunächst ein mathematisches Modell aufgestellt werden. Einfach ausgedrückt wird bei der *Deduktion* durch logische Schlussfolgerung überprüft, ob alle Instanzen, die in einer gegebenen Spezifikation enthalten sind, auch in der Beschreibung des Modells enthalten sind. Wenn es keine Instanz gibt, die ausschließlich in der Spezifikation enthalten ist, so ist damit gezeigt dass das Modell der Spezifikation genügt. Eine anschauliche Einführung aus dem Bereich *künstliche Intelligenz* findet man unter [RN95].

In dieser Arbeit werden wir uns jedoch mit der Technik des Model Checking auseinandersetzen. Im Grunde genommen werden beim Model Checking *sämtliche Zustände* des Modells in geschickter Weise durchsucht. Falls kein Zustand gefunden werden kann, in dem die gesuchte Eigenschaft verletzt wird, so kann gefolgert werden, dass die Eigenschaft im Modell gültig ist. Da die Dauer der Suche stark von der Größe des Modells abhängig ist, gehört zum Model Checking auch praktisch immer eine Verkleinerung des Zustandsraums und ihrer Repräsentation.

Model Checking gibt es in vielen Ausprägungen, etwa für *timed automata*¹ oder *Promela* Modelle². Wir werden uns hier mit dem Model Checking Problem für *zeitstetige Markov-Ketten* beschäftigen, einer speziellen Art von *stochastischen Prozessen*,

¹Für eine Einführung in Model Checking für *timed automata* mit dem Tool *Uppaal* sei auf [BDL04] verwiesen.

²In *Promela* (*Process Meta Language*) können verteilte Systeme modelliert werden, die mit dem traditionsreichen Model Checker *SPIN* verifiziert werden können. Ein umfangreiches Handbuch zu *SPIN* findet sich unter [Hol04].

die hauptsächlich für die Untersuchung von Leistungsverhalten und Zuverlässigkeit von Systemen verwendet werden. Ein Beispiel, welches wir noch genauer betrachten werden, ist ein mehrfach redundant ausgelegtes System. Systeme werden üblicherweise dann redundant ausgelegt, wenn die Ausfallsicherheit von größter Bedeutung ist. Der Airbus A380 ist ein Paradebeispiel dafür. Der dort verwendete *fly-by-wire* Ansatz und das extreme Sicherheitsbedürfnis in der Flugzeugindustrie machte es erforderlich, praktisch alle wichtigen Subsysteme redundant auszulegen. Weitere Beispiele für den Einsatz von zeitstetigen Markov-Ketten findet man bei den Kommunikationsprotokollen in der Informatik (*Bluetooth* sei als Negativ-Beispiel genannt), mit Populationsentwicklungen in der Biologie oder mit Molekülreaktionen in der Chemie.

Weiterhin können Markov-Ketten auch als *Low-Level Konstrukt* für verschiedene höhere Beschreibungssprachen verwendet werden. Hier seien exemplarisch *stochastischen Petri-Netze* (siehe [Mol81]) und *Queueing Netze* (siehe [BGdMT98]) erwähnt. Die Zustandsräume der entsprechenden Markov-Ketten können hier schon bei recht einfachen *High-Level* Modellen unangenehme Größen erreichen. Eine geschickte Verkleinerung des Zustandsraums ist also hier auch besonders wichtig.

Für Model Checking von zeitstetigen Markov-Ketten existieren bereits mehrere Implementierungen. Das Tool PRISM (*Probabilistic Symbol Model Checker*, [HMNP06]) wurde an der *University of Birmingham* entwickelt und kann neben zeitstetigen auch zeitdiskrete Markov-Ketten sowie *Markov-Entscheidungsprozesse* behandeln. Die Praxistauglichkeit von PRISM wurde in zahlreichen Fallstudien gezeigt. Der *Erlangen Twente Markov Chain Checker* ($E \vdash MC^2$, [HKMKS]), eine niederländisch-deutsche Zusammenarbeit, ist ein weiteres Tool zum Model Checking von zeitstetigen Markov-Ketten. Mit dem Editor DaNAMiCS³ können stochastische Petri-Netze erzeugt und in das $E \vdash MC^2$ Format für zeitstetige Markov-Ketten umgewandelt werden.

Auf die Implementierung eines Model Checkers wird hier verzichtet, stattdessen wird die Frage nach möglichen Abstraktionstechniken im Vordergrund stehen. Eine bereits etablierte Technik⁴ arbeitet mit Bisimulationsäquivalenzrelationen, ähnlich wie sie auch für die Minimierung von deterministischen endlichen Automaten verwendet werden können. Diese Technik hat jedoch den Nachteil, dass schon Zustände mit kleinsten Abweichungen im Verhalten zu unterscheiden sind. An diesem Punkt werden wir ansetzen und eine auf Simulation statt auf Bisimulation basierende Abstraktion definieren⁵. Stärkere Abstraktion müssen wir dabei durch entsprechend ungenauere Ergebnisse erkaufen. Um erkennbar zu machen ob die Abstraktion zu stark ist werden wir daher eine Logik verwenden, in der Aussagen nicht nur *wahr* oder *falsch* sein können, sondern auch *unbestimmt*. Eine *unbestimmte* Aussage bezüglich einer Spezifikation ist dann so zu deuten, dass die gewählte Abstraktion des Modells für die gegebene Spezifikation nicht angemessen ist.

³Siehe [CDKN98].

⁴Siehe Kapitel 5 in [BHHK03].

⁵In [FLW06] wurde ein solcher Ansatz bereits für zeitdiskrete Markov-Ketten beschrieben.

Um Spezifikationen bezüglich der dreiwertigen Logik formulieren zu können, müssen wir CSL, eine Spezifikationslogik für zeitstetige Markov-Ketten, nur leicht modifizieren. Für abstrakte zeitstetige Markov-Ketten, mit denen wir die Abstraktionen beschreiben werden, werden wir dann ein Model Checking Verfahren für ein wichtiges Fragment der dreiwertige Variante von CSL vorstellen. Wir werden feststellen, dass sich die Gültigkeit bzw. Ungültigkeit von Spezifikationen auf das ursprüngliche Modell übertragen lassen.

Der Aufbau diese Arbeit ist wie folgt: In Kapitel 2 werden zunächst die wichtigsten Grundlagen aus den Bereichen Model Checking, Abstraktion zeitstetiger Markov-Ketten und dreiwertige Logik vermittelt. In Kapitel 3 werden abstrakte zeitstetige Markov-Ketten eingeführt und mögliche Abstraktionen besprochen. Außerdem enthält es einige Ergebnisse bezüglich *uniformer* abstrakter zeitstetiger Markov-Ketten und zum Vergleich mit der zeitdiskreten Variante dieser Abstraktionstechnik. Nach der Definition des Wahrscheinlichkeitsraums wird dann in Kapitel 4 die Spezifikationslogik 3-CSL für die dreiwertige Logik und Model Checking von 3-CSL Formeln für abstrakte zeitstetige Markov-Ketten untersucht. Dabei werden das *quantitative Erreichbarkeitsproblem* und die Frage nach der *Präservierung* von 3-CSL-Formeln im Mittelpunkt stehen. Abschließend werden in Kapitel 5 die wichtigsten Ergebnisse zusammengefasst, sowie ausstehende Fragen und weiterführende Forschungsmöglichkeiten aufgezeigt.

2 Grundlagen

In diesem Kapitel werden Grundlagen vermittelt, die im weiteren Verlauf der Arbeit noch von Bedeutung sein werden. Dies wird nicht in einer streng formalen Form geschehen, stattdessen werden insbesondere diejenigen Ideen, die später noch in ähnlicher Form auftauchen, auf verständliche Weise präsentiert. Wir beginnen mit einer grundsätzlichen Vorstellung des Model Checking Konzeptes für Transitionssysteme und die *Computation Tree Logic* (CTL).

Anschließend werden Markov-Ketten zur Modellierung von Systemen mit Wahrscheinlichkeiten definiert werden, sowie der mathematische Formalismus, der uns erlaubt, über Wahrscheinlichkeiten von Pfaden in Markov-Ketten sprechen zu können. Als Spezifikationsprache für zeitstetige Markov-Ketten wird dann die *Continuous Stochastic Logic* (CSL) als Erweiterung von CTL eingeführt werden. Ansatzweise wird auch das Thema Model Checking für zeitstetige Markov-Ketten angeschnitten, wir beschränken uns dabei jedoch weitgehend auf die so genannte *Transient Analyse* für uniformisierte zeitstetige Markov-Ketten.

Einen etablierter Ansatz zur Abstraktion von zeitstetigen Markov-Ketten verwendet die Bisimulationsäquivalenz. Da bisimulationsäquivalente Zustände dieselben Eigenschaften bezüglich CSL-Formeln aufweisen, kann man durch Zusammenfassen dieser Zustände eine kleinere (abstraktere) zeitstetige Markov-Kette erzeugen. Diesen Ansatz werden wir in diesem Kapitel vorstellen und später in Kapitel 3 noch weiter ausbauen, um stärkere Abstraktionen zu erhalten.

Markov-Entscheidungsprozesse erweitern das Konzept der Markov-Ketten um Nicht-determinismus. Die später verwendeten *abstrakten* Markov-Ketten sind mit diesen recht eng verwandt, weswegen wir kurz erläutern wollen worum es sich dabei handelt und warum im weiteren Verlauf nicht Entscheidungsprozesse verwendet werden.

Mit einer kurzen Einführung in dreiwertige Logik, die wir zur Repräsentation von unbestimmten Eigenschaften benötigen werden, schließen wir dieses Kapitel ab.

2.1 CTL Model Checking

Model Checking ist ein Verfahren zur Verifikation von Systemen. Das zu verifizierende System muss dabei in einem geeigneten mathematischen Modell vorliegen. Ein Formalismus mit dem Systeme mathematisch beschrieben werden können sind Transitionssysteme, bestehend aus Zuständen, Zustandsübergängen (oder Transitionen) und Eigenschaften, die den einzelnen Zuständen zugeordnet werden können. Die einzelnen Zustände eines Transitionssystems repräsentieren dabei Zustände oder Konfigurationen des zu modellierenden Systems. An dieser Stelle wird praktisch immer schon eine Abstraktion durchgeführt. Ist man bei der Modellierung einer Komponente beispielsweise nur an der Funktionstüchtigkeit interessiert, nicht aber an internen Vorgängen, so kann diese Komponente mit zwei Zuständen *funktionsfähig* und *defekt* modelliert werden. Die Transitionen beschreiben die möglichen Veränderungen des Systems. Bei einer Ampelschaltung wäre etwa eine Transition vom Zustand *grün* in den Zustand *gelb* anzugeben, von *grün* nach *rot* dagegen nicht.

Der zweite Schritt beim Model Checking besteht darin, die Anforderungen an das System zu formulieren. Nehmen wir als Beispiel das Modell einer Fußgängerampel. Eine sinnvolle Anforderung an das System wäre etwa, dass zu keiner Zeit Fußgänger und Straßenverkehr gleichzeitig Grün haben dürfen. Um dies mathematisch sauber zu formulieren, können wir uns der *Computation Tree Logic* bedienen, einer temporalen Erweiterung der Prädikatenlogik.

In einem dritten Schritt, der Verifikation, werden die gegebenen Spezifikationen auf Gültigkeit bezüglich des Systems überprüft. Dies ist vollautomatisch mit Rechnerunterstützung möglich. Wir werden dies nicht im Detail besprechen und stattdessen nur die Grundideen erläutern. Eine ausführlichere Beschreibung des Verfahrens ist unter anderem in [HR00] zu finden.

Die folgenden Definitionen sind so gewählt, dass sie später möglichst einfach erweitert werden können.

Definition 1 (Transitionssysteme). Sei AP die Menge aller atomaren Eigenschaften (engl. *atomic propositions*), dann bezeichnet man als Transitionssystem \mathcal{T} über AP ein Tupel $\mathcal{T} = (S, S_0, \mathbf{T}, L)$, bestehend aus:

- Zustandsmenge S ,
- Startzustandsmenge S_0 , einer Teilmenge von S ,
- Transitionsmatrix $\mathbf{T} : S \times S \mapsto \{0, 1\}$ mit $\mathbf{T}(s_i, s_j) = 1$ falls eine Transition von Zustand $s_i \in S$ nach Zustand $s_j \in S$ führt und $\mathbf{T}(s_i, s_j) = 0$ sonst,
- Beschriftungsfunktion $L : S \times AP \mapsto \mathbb{B}$, die für jeden Zustand die dort gültigen atomaren Eigenschaften $a \in AP$ festlegt⁶.

—

Definition 2 (Pfade in Transitionssystemen). Sei $\mathcal{T} = (S, S_0, \mathbf{T}, L)$ eine Transitionssystem. Ein unendlicher Pfad σ ist eine Folge von Zuständen $s_0 \longrightarrow s_1 \longrightarrow \dots$ mit $\mathbf{T}(s_i, s_{i+1}) > 0$ für alle $i \in \mathbb{N}$. Ein endlicher Pfad σ ist eine Folge von Zuständen $s_0 \longrightarrow s_1 \longrightarrow \dots \longrightarrow s_n$ mit $\mathbf{T}(s_i, s_{i+1}) = 1$ für alle $i \in \{0, \dots, n-1\}$.

—

Die Menge aller Pfade bezüglich eines Transitionssystems \mathcal{T} bezeichnen wir mit $Pfad^{\mathcal{T}}$ und die Menge aller Pfade mit Anfangszustand s mit $Pfad_s^{\mathcal{T}}$.

Die Funktion $|\cdot| : Pfad^{\mathcal{T}} \mapsto (\mathbb{N} \cup \infty)$ liefert die Länge eines Pfades. Für endliche Pfade $\sigma = s_0 \rightarrow s_1 \rightarrow \dots \rightarrow s_n$ ist die Länge gegeben mit $|\sigma| = n$, für unendliche Pfade mit ∞ . Mit $\sigma[i]$ bezeichnen wir den i -ten Zustand eines Pfades σ mit Länge $|\sigma| \geq i$.

⁶Mit $\mathbb{B} = \{\top, \perp\}$ bezeichnen wir die Menge der Wahrheitswerte *wahr* (\top) und *falsch* (\perp).

Beispiel 1 (TMR als Transitionssystem). Im Folgenden betrachten wir ein Beispiel eines dreifach redundant ausgelegten Systems (engl. *Triple Modular Redundant System*, TMR). In diesem System gibt es drei unabhängig voneinander arbeitende Einheiten, die alle dieselben Berechnungen durchführen. Eine vierte Einheit, der sogenannte Voter, legt das Ergebnis eines der funktionierenden Einheiten auf die Ausgabe. Üblicherweise ist der Voter weniger anfällig für Störungen, sodass die Komponenten in der Regel viele Male ausfallen können, bevor das System komplett ausgetauscht werden muss. Wir wollen annehmen, dass das Reperaturteam nur jeweils eine Einheit gleichzeitig reparieren kann. Sollten mehrere Einheiten defekt sein, so wird die Einheit mit dem kleinsten Index zuerst repariert. Ist der Voter defekt, so wird das gesamte System inklusive aller Komponenten ausgetauscht.

Die Zustände $u_1u_2u_31$ mit $u_1, u_2, u_3 \in \{0, 1\}$ repräsentieren die Situationen, in der alle Einheiten i mit $u_i = 1$ und der Voter intakt sind. Zustand 0000 repräsentiert die Situation, in der der Voter ausgefallen ist und damit das gesamte System ausgetauscht werden muss. Die Zustände mit i funktionsfähigen Einheiten sind hier jeweils mit up_i beschriftet und der Zustand 0000 mit *down*.

Wir abstrahieren zunächst noch von den Ausfallwahrscheinlichkeiten und Zeitspannen und erhalten nach obiger Beschreibung folgendes Transitionssystem:

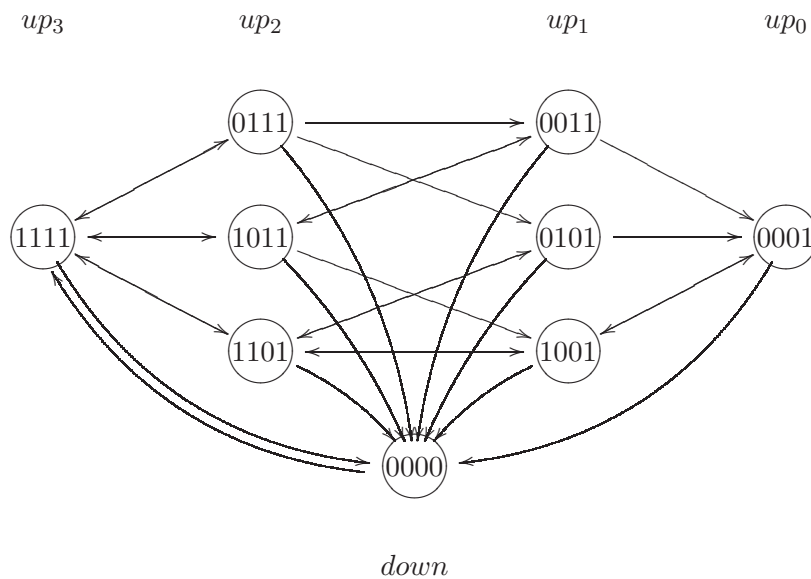


Abbildung 1: TMR als Transitionssystem

Kommen wir nun zur Spezifikation. Als mathematischen Formalismus zur Beschreibung von Spezifikationen bezüglich Transitionssystemen werden wir die CTL verwenden. Zusätzlich zu den aus der Prädikatenlogik bekannten Operatoren wie Negation, Konjunktion, etc. können Aussagen über Pfade formuliert werden, also über die zeitliche Abfolge von Zustandsübergängen. Dazu werden die Operatoren $\diamond\psi$ („es gilt in

irgendeinem Zustand des Pfades die Eigenschaft ψ “), $\Box\psi$ („es gilt in jedem Zustand des Pfades die Eigenschaft ψ “), $\mathcal{X}\psi$ („im nächsten Zustand gilt die Eigenschaft ψ “) und $\psi_1\mathcal{U}\psi_2$ („es gilt ψ_1 solange, bis in einem Zustand ψ_2 gilt“) eingeführt. Außerdem werden die Pfadquantoren $E\Psi$ („auf einem möglichen Pfad gilt die Pfadeigenschaft Ψ “) und $A\Psi$ („für alle Pfade gilt die Pfadeigenschaft Ψ “) benötigt.

Um die Beschreibung der Beispiele abzukürzen, wollen wir davon ausgehen, dass für jeden Zustand $s_i \in S$ eine Eigenschaft $at_{s_i} \in AP$ existiert, die ausschließlich in diesem Zustand gültig ist. Im Beispiel 1 könnten wir also die Forderung, dass zu keinem Zeitpunkt der Voter defekt und funktionstüchtig zugleich ist, ausdrücken mit der Formel $A\Box\neg(at_{0000} \wedge (at_{0001} \vee at_{0011} \vee \dots \vee at_{1111}))$.

Definition 3 (CTL-Formeln). Die Menge $\mathbb{F}_{\mathcal{T}}$ sei die Menge aller CTL-Formeln bezüglich eines Transitionssystems $\mathcal{T} = (S, S_0, \mathbf{T}, L)$. Als atomare CTL-Formeln in $\mathbb{F}_{\mathcal{T}}$ werden bezeichnet:

- $true \in \mathbb{F}_{\mathcal{T}}$
- $false \in \mathbb{F}_{\mathcal{T}}$
- $a \in \mathbb{F}_{\mathcal{T}}$ falls $a \in AP$

Weiterhin wird die Menge der CTL-Formeln nach folgenden Regeln induktiv aufgebaut:

- $(\neg\psi) \in \mathbb{F}_{\mathcal{T}}$ falls $\psi \in \mathbb{F}_{\mathcal{T}}$
- $(\psi_1 \wedge \psi_2) \in \mathbb{F}_{\mathcal{T}}$ falls $\psi_1, \psi_2 \in \mathbb{F}_{\mathcal{T}}$
- $(E\Box\psi) \in \mathbb{F}_{\mathcal{T}}$ falls $\psi \in \mathbb{F}_{\mathcal{T}}$
- $(E\mathcal{X}\psi) \in \mathbb{F}_{\mathcal{T}}$ falls $\psi \in \mathbb{F}_{\mathcal{T}}$
- $(E(\psi_1\mathcal{U}\psi_2)) \in \mathbb{F}_{\mathcal{T}}$ falls $\psi_1, \psi_2 \in \mathbb{F}_{\mathcal{T}}$

—

Operatoren wie Implikation, Konjunktion etc. werden in der üblichen Weise als Abkürzungen aufgefasst (z.B. $\psi_1 \rightarrow \psi_2 = \neg\psi_1 \vee \psi_2$), ebenso wie die übrigen quantifizierten Pfadformeln (siehe [Sch03]):

- $A\mathcal{X}\psi = \neg E\mathcal{X}\neg\psi$ mit $\psi \in \mathbb{F}_{\mathcal{T}}$
- $E\Diamond\psi = E(true\mathcal{U}\psi)$ mit $\psi \in \mathbb{F}_{\mathcal{T}}$
- $A\Box\psi = \neg E\Diamond\neg\psi$ mit $\psi \in \mathbb{F}_{\mathcal{T}}$
- $A\Diamond\psi = \neg E\Box\neg\psi$ mit $\psi \in \mathbb{F}_{\mathcal{T}}$
- $A(\psi_1\mathcal{U}\psi_2) = \neg E(\neg\psi_2\mathcal{U}(\neg\psi_1 \wedge \neg\psi_2)) \wedge \neg E\Box\neg\psi_2$ mit $\psi_1, \psi_2 \in \mathbb{F}_{\mathcal{T}}$

Beispiel 2. Passend zum Modell in Beispiel 1 können wir nun Spezifikationen in CTL formulieren. Eine Bedingung die das Modell offensichtlich erfüllt ist die, dass immer dann wenn sich das System in einem up_i -Zustand befindet nicht die Eigenschaft *down* erfüllt sein kann. Die entsprechende CTL-Formel dazu lautet $A\Box(up_3 \vee up_2 \vee up_1 \vee up_0) \rightarrow \neg down$.

Eine weitere interessantere Eigenschaft wäre, wenn es keine absorbierenden Zustände im System gäbe, also wenn für jeden Zustand ein Nachfolgezustand existieren würde. Dass dies gilt, kann man in diesem Beispiel noch recht einfach sehen. Wenn der Zustandsraum jedoch in mehrstellige Bereiche wächst, ist diese Eigenschaft nicht mehr so leicht per Hand nachzuweisen. Die CTL-Formel die diese Eigenschaft beschreibt lautet $A\Box \mathcal{X} \text{ true}$.

Nun da wir die nötigen Mittel zur Beschreibung des Problems haben, können wir uns mit dessen Lösung beschäftigen, also der Verifikation. Zunächst müssen wir dazu die formale Semantik von CTL festlegen, um über Gültigkeit von Formeln auf mathematische Weise sprechen zu können. Da die boolesche Algebra sehr grundlegend in der Informatik ist, dürfte hier über die Identifizierung der Symbole hinaus eigentlich keine weitere Ausführung zu diesem Thema nötig sein. Da wir jedoch später noch eine dreiwertige Logik benötigen, welche einen dritten Wahrheitswert *unbestimmt* einführt, soll hier dennoch eine kurze formale Auseinandersetzung mit dem Thema erfolgen:

Als Symbol für eine *wahre Aussage*, d.h. eine vom Modell erfüllte Formel, verwenden wir \top , für eine *falsche Aussage* schreiben wir \perp . Aus der Menge $\mathbb{B} = \{\top, \perp\}$ kann man einen vollständigen Verband bilden mit $\perp < \top$ und dem Supremum und Infimum wie in Tabelle 1. Das Supremum kann zur Definition der Semantik der Konjunktion verwendet werden (siehe Tabelle 2). Würde man die Semantik der Disjunktion direkt definieren und nicht die Disjunktion als bloße Abkürzung auffassen, so würde man sie über das Infimum definieren. Häufig werden statt \perp und \top die Zahlen 0 und 1 als Wahrheitswerte verwendet. Dann wird als Infimum das Minimum und als Supremum das Maximum gewählt.

Das Komplement α^c definieren wir durch $\perp^c = \top$ und $\top^c = \perp$. Da wir von einer *closed world assumption* ausgehen (d.h. $\llbracket s, \psi \rrbracket = \perp \Rightarrow \llbracket s, \neg\psi \rrbracket = \top$), kann das Komplement zur Definition der Semantik von $\neg\psi$ verwendet werden. Mit den Wahrheitswerten 0 und 1 wird das Komplement üblicherweise als $\alpha^c = 1 - \alpha$ definiert.

\sqcup	\perp	\top	\sqcap	\perp	\top
\perp	\perp	\top	\perp	\perp	\perp
\top	\top	\top	\top	\perp	\top

Tabelle 1: *join* (\sqcup) und *meet* (\sqcap) für \mathbb{B}

$\llbracket s, true \rrbracket$	$= \top$
$\llbracket s, a \rrbracket$	$= L(s, a)$
$\llbracket s, \neg\psi \rrbracket$	$= \llbracket s, \psi \rrbracket^c$
$\llbracket s, \psi_1 \wedge \psi_2 \rrbracket$	$= \llbracket s, \psi_1 \rrbracket \sqcap \llbracket s, \psi_2 \rrbracket$
$\llbracket \sigma, i, \psi \rrbracket$	$= \begin{cases} \llbracket \sigma[i], \psi \rrbracket & \text{falls } i < \sigma \\ \perp & \text{sonst} \end{cases}$

$\llbracket \sigma, \mathcal{X} \psi \rrbracket$	$= \llbracket \sigma, 1, \psi \rrbracket$
$\llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket$	$= \begin{cases} \top & \text{falls } \exists i : (\llbracket \sigma, i, \psi_2 \rrbracket = \top \\ & \wedge \forall 0 \leq j < i : \llbracket \sigma, j, \psi_1 \rrbracket = \top) \\ \perp & \text{falls } \forall i : (\llbracket \sigma, i, \psi_2 \rrbracket = \perp \\ & \vee \exists 0 \leq j < i : \llbracket \sigma, j, \psi_1 \rrbracket = \perp) \end{cases}$

$\llbracket \sigma, \Box\psi \rrbracket$	$= \begin{cases} \top & \text{falls } \forall 0 \leq i < \sigma : \llbracket \sigma[i], \psi \rrbracket = \top \\ \perp & \text{sonst} \end{cases}$
$\llbracket s, E\Psi \rrbracket$	$= \begin{cases} \top & \text{falls } \exists \sigma \in Pfad_s^{\mathcal{T}} : \llbracket \sigma, \Psi \rrbracket = \top \\ \perp & \text{sonst} \end{cases}$

mit $s \in S^{\mathcal{T}}$; $a \in AP^{\mathcal{T}}$; $\psi, \psi_1, \psi_2 \in \mathbb{S}^{\mathcal{T}}$; $\Psi \in \mathbb{P}^{\mathcal{T}}$; $\sigma \in Pfad^{\mathcal{T}}$; $i \in \mathbb{N}$
für eine gegebenes Transitionssystem $\mathcal{T} = (S^{\mathcal{T}}, S_0^{\mathcal{T}}, \mathbf{T}^{\mathcal{T}}, L^{\mathcal{T}})$

Tabelle 2: CTL Semantik

Die Semantik von CTL wird in der Tabelle 2 vollständig beschrieben. Eine kurze intuitive Erläuterung wurde schon vor der Definition der CTL-Formeln gegeben, daher betrachten wir im Folgenden nur noch ein paar einfache Beispiele, um den Umgang mit der Semantik zu demonstrieren. Dazu legen wir das Modell aus Beispiel 1 zugrunde:

- In Zustand 0000 gilt die atomare Eigenschaft *down*:

$$\llbracket 0000, \text{down} \rrbracket = L(0000, \text{down}) = \top$$

- In Zustand 0000 ist die Eigenschaft verletzt, dass *up*₁ oder *up*₂ gilt:

$$\begin{aligned} \llbracket 0000, \text{up}_1 \vee \text{up}_2 \rrbracket &= \llbracket 0000, \neg(\neg \text{up}_1 \wedge \neg \text{up}_2) \rrbracket \\ &= \llbracket 0000, \neg \text{up}_1 \wedge \neg \text{up}_2 \rrbracket^c \\ &= (\llbracket 0000, \neg \text{up}_1 \rrbracket \sqcap \llbracket 0000, \neg \text{up}_2 \rrbracket)^c \\ &= (\llbracket 0000, \text{up}_1 \rrbracket^c \sqcap \llbracket 0000, \text{up}_2 \rrbracket^c)^c \\ &= (L(0000, \text{up}_1)^c \sqcap L(0000, \text{up}_2)^c)^c \\ &= (\perp^c \sqcap \perp^c)^c \\ &= (\top \sqcap \top)^c \\ &= \top^c \\ &= \perp \end{aligned}$$

- Nun zeigen wir in einem etwas längeren Beispiel, wie die Semantik auf Pfaden zu verstehen ist. Wir zeigen dazu, dass für alle Zustände des Transitionssystems die Formel $A\Box(\text{down} \rightarrow A\mathcal{X}\text{up}_3)$ erfüllt ist:

$$\llbracket s, A\Box(\text{down} \rightarrow A\mathcal{X}\text{up}_3) \rrbracket = \llbracket s, A\Box(\neg \text{down} \vee A\mathcal{X}\text{up}_3) \rrbracket = \top$$

gilt genau dann, wenn

$$\forall \sigma \in \text{Pfad}_s : \llbracket \sigma, \Box(\neg \text{down} \vee A\mathcal{X}\text{up}_3) \rrbracket = \top$$

Dies wiederum gilt, falls $\llbracket \sigma[i], \neg \text{down} \vee A\mathcal{X}\text{up}_3 \rrbracket$ gültig ist für alle $0 \leq i < |\sigma|$.

Untersuchen wir zunächst $\neg \text{down} \vee A\mathcal{X}\text{up}_3$ etwas näher:

$$\begin{aligned} \llbracket s, \neg \text{down} \vee A\mathcal{X}\text{up}_3 \rrbracket &= \llbracket s, \neg(\text{down} \wedge \neg A\mathcal{X}\text{up}_3) \rrbracket \\ &= \llbracket s, \text{down} \wedge \neg A\mathcal{X}\text{up}_3 \rrbracket^c \\ &= \top \end{aligned}$$

gilt genau dann, wenn

$$\begin{aligned} \llbracket s, \text{down} \wedge \neg A\mathcal{X}\text{up}_3 \rrbracket &= \llbracket s, \text{down} \rrbracket \sqcap \llbracket s, \neg A\mathcal{X}\text{up}_3 \rrbracket \\ &= \llbracket s, \text{down} \rrbracket \sqcap \llbracket s, A\mathcal{X}\text{up}_3 \rrbracket^c \\ &= \perp \end{aligned}$$

Für den Fall dass $\llbracket s, \text{down} \rrbracket = \perp$ gilt, folgt wegen der Definition von \sqcap auch, dass $\llbracket s, \text{down} \rrbracket \sqcap \llbracket s, A\mathcal{X}\text{up}_3 \rrbracket^c = \perp$ gilt. Es bleibt also für den Fall $\llbracket s, \text{down} \rrbracket = \top$ noch zu zeigen, dass $\llbracket s, A\mathcal{X}\text{up}_3 \rrbracket^c = \perp$ beziehungsweise $\llbracket s, A\mathcal{X}\text{up}_3 \rrbracket = \top$ gilt.

Da $\llbracket s, \text{down} \rrbracket = \top$ lediglich für $s = 0000$ erfüllt ist, interessiert uns also im Folgenden nur noch, ob für alle $\sigma \in \text{Pfad}_{0000}$ gilt, dass $\llbracket \sigma, \mathcal{X} \text{up}_3 \rrbracket$:

$$\llbracket \sigma, \mathcal{X} \text{up}_3 \rrbracket = \llbracket \sigma, 1, \text{up}_3 \rrbracket = \begin{cases} \llbracket \sigma[1], \text{up}_3 \rrbracket & \text{falls } 1 < |\sigma| \\ \perp & \text{sonst} \end{cases}$$

Die Eigenschaft $\llbracket \sigma[1], \text{up}_3 \rrbracket$ ist genau dann erfüllt, wenn $\sigma[1] = 1111$, da 1111 der einzige up_3 Zustand in unserem Beispiel ist. Außerdem hat der Zustand 0000 als einzigen Nachfolger den Zustand 1111, sodass für alle Pfade mit $\sigma[0] = 0000$ und $|\sigma| > 1$ gelten muss, dass $\sigma[1] = 1111$ ist. Da 0000 nicht absorbierend ist, gilt natürlich auch, dass $|\sigma| > 1$.

Die Formel $\neg \text{down} \vee A \mathcal{X} \text{up}_3$ ist also für alle Zustände des Transitionssystems erfüllt. Damit gilt also $\forall 0 \leq i < |\sigma| : \llbracket \sigma[i], \neg \text{down} \vee A \mathcal{X} \text{up}_3 \rrbracket = \top$, da es keinen Zustand $\sigma[i]$ eines Pfades gibt, für den $\neg \text{down} \vee A \mathcal{X} \text{up}_3$ verletzt werden könnte.

Die Formel $A \Box (\text{down} \rightarrow A \mathcal{X} \text{up}_3)$ ist demnach in allen Zuständen gültig.

Da wir nun gesehen haben, wie die Gültigkeit von Spezifikationen bezüglich eines Modells per Hand überprüft werden kann, wollen wir noch einen sehr kurzen Blick auf die algorithmische Herangehensweise werfen.

Zunächst stellen wir fest, dass wir in den Beispielen die gegebene Formel immer in ihre Teilformeln zerlegt haben und für diese nach Zuständen gesucht haben, in denen sie erfüllt sind. Die einfacheren Bausteine, für die keine Betrachtung von Pfaden nötig ist, kommen aus der Aussagenlogik. Für jeden Zustand kann man auf übliche Weise die Gültigkeit der aussagenlogischen Formeln bestimmen, sobald der Wahrheitswert für die Teilformeln bekannt ist.

Zur Überprüfung der Pfadformeln müssen wir den Transitionsbaum betrachten. Bei der Rückwärtsanalyse, auf die wir zurückgreifen, müssen wir jeden Zustand pro Tiefenebene lediglich einmal betrachten, wie es auch in Abbildung 3 dargestellt ist.

Sei das Transitionssystem aus Abbildung 2 gegeben mit dem Startzustand s_0 .

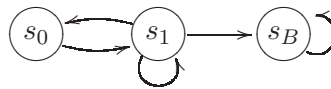


Abbildung 2: Transitionssystem

Wir interessieren uns nun für die Erreichbarkeit von Zustand s_B , die wir nach dem allgemein bekannten Rückwärtsanalyseverfahren bestimmen wollen.

Die Analyse läuft wie folgt von unten nach oben bzw. von hinten nach vorne ab: Zu Beginn, also *ganz hinten*, werden nur die Zustände markiert, die es zu erreichen gilt (hier nur s_B). Anschließend markiert man sämtliche direkten Vorgänger der markierten Zustände und wiederholt dies solange, bis man bezüglich der Markierungen einen Fixpunkt erreicht hat, also bis zwei übereinanderliegende Ebenen dieselben

Bezüglich der Zeitmessung wird zwischen zeitdiskreten und die zeitstetigen Markov-Ketten unterschieden. Bei den zeitdiskreten Markov-Ketten betrachtet man die Zeit als in diskreten (abzählbaren) Schritten ablaufend. Dabei definiert man üblicherweise die Zeit so, dass mit jedem Zustandsübergang bzw. Ereignis genau eine Zeiteinheit vergeht.

Markov-Ketten haben ihren Namen von der Markov-Eigenschaft geerbt, die auch Gedächtnislosigkeit genannt wird. Ein gedächtnisloser Prozess zeichnet sich dadurch aus, dass das zukünftige Verhalten zu jedem Zeitpunkt unabhängig vom früheren Verhalten ist. Wir bezeichnen den chronologisch i -ten Zustand des Prozesses mit der Zufallsvariablen X_i . Die Gedächtnislosigkeit ist genau dann gegeben, falls für alle $i \in \mathbb{N}_{>0}$ gilt, dass die Wahrscheinlichkeit für $X_i = s_i$ nur vom *aktuellen* Zustand $X_{i-1} = s_{i-1}$ abhängt:

$$Pr\{X_i = s_i \mid X_{i-1} = s_{i-1}, \dots, X_1 = s_1\} = Pr\{X_i = s_i \mid X_{i-1} = s_{i-1}\}$$

Definition 4 (Zeitdiskrete Markov-Ketten). Als zeitdiskrete Markov-Kette (engl.: *discrete-time Markov chain, DTMC*) bezeichnet man ein Tupel $\mathcal{M} = (S, \mathbf{P}, L)$, bestehend aus:

- Zustandsmenge S ,
- Wahrscheinlichkeitsmatrix $\mathbf{P} : S \times S \mapsto [0, 1]$ mit $\sum_{s_j \in S} \mathbf{P}(s_i, s_j) \in \{0, 1\}$ für alle Zustände $s_i \in S$,
- Beschriftungsfunktion $L : S \times AP \mapsto \mathbb{B}$, die für jeden Zustand die dort gültigen atomaren Eigenschaften $a \in AP$ festlegt.

—

Falls $\mathbf{P}(s_i, s_j) > 0$ für Zustände s_i und s_j gilt, so sagt man auch, dass eine Transition von Zustand $s_i \in S$ nach Zustand $s_j \in S$ führt. Gilt $\sum_{s_j \in S} \mathbf{P}(s_i, s_j) = 0$ für einen Zustand s_i , gibt es also keine Transitionen die aus s_i herausführen, so heißt s_i absorbierend.

Zeitdiskrete Markov-Ketten haben die Eigenschaft der Gedächtnislosigkeit, da für Zustandsübergänge, etwa von s_i nach s_j , unabhängig von der Vergangenheit immer dieselbe Wahrscheinlichkeit gegeben ist ($\mathbf{P}(s_i, s_j)$).

Definition 5 (Pfade in zeitdiskreten Markov-Ketten). Sei $\mathcal{M} = (S, \mathbf{P}, L)$ eine zeitdiskrete Markov-Kette. Ein unendlicher Pfad σ ist eine Folge von Zuständen $s_0 \longrightarrow s_1 \longrightarrow \dots$ mit $\mathbf{P}(s_i, s_{i+1}) > 0$ für alle $i \in \mathbb{N}$. Ein endlicher Pfad σ ist eine Folge von Zuständen $s_0 \longrightarrow s_1 \longrightarrow \dots \longrightarrow s_n$ mit s_n ist absorbierend und $\mathbf{P}(s_i, s_{i+1}) > 0$ für alle $i \in \{0, \dots, n-1\}$.

—

Die Notation für die Länge von Pfaden, für den i -ten Zustand eines Pfads und für Pfadmengen können wir von den Transitionssystemen übernehmen.

Eine weitergehende Einführung in das Thema zeitdiskrete Markov-Ketten soll hier nicht gegeben werden. Mehr dazu findet man beispielsweise in [Tij03]. Stattdessen werden wir für zeitstetige Markov-Ketten die Spezifikationslogik formal einführen und eine Möglichkeit erläutern, das *quantitative* Erreichbarkeitsproblem algorithmisch zu behandeln, das für das Model Checking zeitstetiger Markov-Ketten ausgenutzt werden kann. Als *quantitatives* Erreichbarkeitsproblem bezeichnen wir die Frage nach der Wahrscheinlichkeit, mit der eine Zielmenge erreicht werden kann. Im Gegensatz dazu beschäftigt sich das *qualitative* Erreichbarkeitsproblem mit der Frage, ob ein Zustand (fast) immer erreicht werden kann oder nicht (siehe [BL04]).

Bei den zeitstetigen Markov-Ketten, die in dieser Arbeit im Mittelpunkt stehen, werden in der Regel die Zeitangaben aus $\mathbb{R}_{\geq 0}$ gewählt. Zustandsänderungen können zu beliebigen Zeitpunkten erfolgen, weshalb zusätzlich zu der Wahrscheinlichkeit für den Übergang in einen bestimmten Nachfolgezustand auch noch die sogenannte *Exitrate* eine Rolle spielt, die durchschnittliche Anzahl von Transition mit der ein Zustand pro Zeiteinheit verlassen würde. Umso höher die Exitrate, umso schneller wird der Zustand verlassen.

Statt zusätzlich zu den Wahrscheinlichkeiten noch Exitraten für jeden Zustand zu notieren, schreibt man statt der Wahrscheinlichkeit die Rate, Produkt aus Exitrate und Wahrscheinlichkeit, an eine Transition. Die Raten beschreiben also, wie oft pro Zeiteinheit vom aktuellen Zustand in den Nachfolgezustand gegangen würde. Wenn im Beispiel 1 also die Zeit in Jahren gemessen würde und die einzelnen Komponenten im Durchschnitt alle drei Monate ausfallen, dann wäre etwa die Rate für den Übergang von 1111 in einen up_2 -Zustand mit $\frac{12}{3} = 4$ anzusetzen.

Die Unabhängigkeit von den früheren Zuständen ist hier ebenso wie bei den zeitdiskreten Markov-Ketten gegeben. Um in zeitstetigen Markov-Ketten die Eigenschaft der Gedächtnislosigkeit zu erhalten, darf aber auch die Zeit, die sich ein Prozess bereits in einem Zustand befindet, keine Rolle spielen. Die erwartete Dauer für einen Zustandsübergang muss also immer dieselbe sein, egal ob der Prozess gerade in diesen Zustand gewechselt ist, oder ob er sich schon seit langem dort befindet. Wenn wir nun eine Verteilung angeben wollen, die uns die Wahrscheinlichkeit liefert, dass innerhalb einer bestimmten Zeitdauer t ein Übergang von Zustand s in einen Zustand s' stattfindet, dann muss für diese Verteilung gelten, dass $Pr\{X_{t'+t} = s' \mid X_{t'} = s\} = Pr\{X_t = s' \mid X_0 = s\}$ für beliebige $t' \in \mathbb{R}_{\geq 0}$. Die einzige Wahrscheinlichkeitsverteilung, die dies leistet ist die Exponentialverteilung mit Parameter λ :

$$f_\lambda(t) = 1 - e^{-\lambda \cdot t} \text{ für } t \in \mathbb{R}_{\geq 0} \quad (1)$$

Der Erwartungswert einer λ -Exponentialverteilung ist bekanntermaßen $\frac{1}{\lambda}$. Da wir für unser Beispiel als Erwartungswert für die Dauer bis zum Ausfall einer Komponente drei Monate bzw. $\frac{1}{4}$ Jahr annehmen wollten, beschreibt also für eine zeitstetige Markov-Kette nur die Exponentialverteilung mit $\lambda = 4$ (der gegebenen Rate) diesen Sachverhalt angemessen. Für weitere Erläuterungen sei hier auf [Tij03] verwiesen.

Um die Notation des Transitionssystems so zu erweitern, dass damit zeitstetige Markov-Ketten beschrieben werden können, müssen wir die 0/1-Matrix durch eine Ratenmatrix ersetzen. Es ist zu beachten, dass die Exponentialverteilung für $\lambda = 0$ als Ergebnis immer 0 liefert. Die Wahrscheinlichkeit für einen beliebigen Zeitpunkt t mit einer Rate 0 in einen bestimmten Nachfolgezustand zu gelangen ist also ebenfalls immer 0. Die Rate 0 entspricht also einer fehlenden Kante in einem Transitionssystem.

Definition 6 (Zeitstetige Markov-Ketten). Als zeitstetige Markov-Kette (engl.: *continuous-time Markov chain, CTMC*) bezeichnet man ein Tupel $\mathcal{M} = (S, \mathbf{R}, L)$, bestehend aus:

- Zustandsmenge S ,
- Ratenmatrix $\mathbf{R} : S \times S \mapsto \mathbb{R}_{\geq 0}$ mit $\mathbf{R}(s_i, s_j) > 0$ falls eine Transition von Zustand $s_i \in S$ nach Zustand $s_j \in S$ führt und $\mathbf{R}(s_i, s_j) = 0$ sonst,
- Beschriftungsfunktion $L : S \times AP \mapsto \mathbb{B}$, die für jeden Zustand die dort gültigen atomaren Eigenschaften $a \in AP$ festlegt.

—

Die Exitrate eines Zustands s ist durch $E(s) = \sum_{s' \in S} \mathbf{R}(s, s')$ bestimmt. Sie gibt an, wie oft der Zustand über eine *beliebige* Kante pro Zeiteinheit verlassen werden würde. Gäbe es beispielsweise für einen Zustand zwei ausgehende Kanten mit Raten r_1 und r_2 , so würde der Zustand pro Zeiteinheit über die erste Kante r_1 mal verlassen werden und über die zweite Kante r_2 mal. Insgesamt würde der Zustand demnach $r_1 + r_2$ mal pro Zeiteinheit verlassen werden.

Definition 7 (Pfade in zeitstetigen Markov-Ketten). Sei $\mathcal{M} = (S, \mathbf{R}, L)$ eine zeitstetige Markov-Kette. Ein unendlicher Pfad σ ist eine Folge von Zuständen und Zeitangaben $s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} \dots$ mit $t_i \in \mathbb{R}_{>0}$ und $\mathbf{R}(s_i, s_{i+1}) > 0$ für alle $i \in \mathbb{N}$. Ein endlicher Pfad σ ist eine Folge von Zuständen und Zeitangaben $s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} \dots \xrightarrow{t_{n-1}} s_n$ mit $\mathbf{R}(s_n, s) = 0$ für alle $s \in S$ und $\mathbf{R}(s_i, s_{i+1}) > 0$ für alle $i \in \{0, \dots, n-1\}$.

—

Zusätzlich zu den in zeitdiskreten Markov-Ketten verwendeten Notationen bezüglich Pfaden definieren wir für zeitstetige Markov-Ketten noch für unendliche Pfade die Verweildauer δ im i -ten Zustand als $\delta(\sigma, i) = t_i$ und den Zustand des Pfades zur Zeit t als $\sigma@t = \sigma[s_i]$ wobei der Zustand s_i derjenige mit dem kleinsten Index i ist, für den gilt $t \leq \sum_{k=0}^i t_k$. Intuitiv gesprochen ist s_i also der erste Zustand, der nach dem Zeitpunkt t verlassen wird.

Für einen endlichen Pfad σ mit $|\sigma| = l$ ist die Verweildauer von $\sigma[l]$ nach obiger Beschreibung nicht ordentlich definiert. Wir definieren daher sinnvollerweise $\delta(\sigma, l) = \infty$, da der Zustand nicht mehr verlassen werden kann. Für $i \in \{0, \dots, l-1\}$ sei die Definition wie bei den unendlichen Pfaden. Ebenso müssen wir für $\sigma@t$ den Sonderfall behandeln, dass $t > \sum_{k=0}^{l-1} t_k$. Ist dies gegeben, so definieren wir $\sigma@t = s_l$. Für $t \leq \sum_{k=0}^{l-1} t_k$ wird $\sigma@t$ wieder wie im Fall von unendlichen Pfaden definiert.

Die zeitdiskrete Markov-Kette mit denselben Zuständen, Wahrscheinlichkeiten und Beschriftungen wird im Zusammenhang mit der gegebenen zeitstetigen Markov-Kette auch die *eingebettete Markov-Kette* genannt.

Definition 8 (Eingebettete Markov-Kette). Zu einer gegebenen zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}, L)$ heißt die zeitdiskrete Markov-Kette $\bar{\mathcal{M}} = (S, \mathbf{P}, L)$ die in \mathcal{M} *eingebettete* Markov-Kette, falls für alle $s, s' \in S$ gilt:

$$\mathbf{P}(s, s') = \begin{cases} \mathbf{R}(s, s')/E(s) & \text{falls } E(s) > 0 \\ 0 & \text{falls } E(s) = 0 \end{cases}$$

Lemma 1. Für alle Ratenvektoren $\mathbf{R}(s, \cdot)$ mit $s \in S$ sind die zugehörigen Zeilen der Matrix \mathbf{P} der eingebetteten Markov-Kette entweder Wahrscheinlichkeitsverteilungen oder Nullvektoren.

Beweis. Falls s absorbierend ist, also $\mathbf{R}(s, s') = 0$ für alle $s' \in S$, dann folgt daraus $E(s) = \sum_{s' \in S} \mathbf{R}(s, s') = 0$ und damit auch $\mathbf{P}(s, s') = 0$ für alle $s' \in S$.

Ist s nicht absorbierend, so gilt $E(s) > 0$. Mit $\mathbf{P}(s, s') \in [0, 1]$ für alle $s' \in S$ (wegen $E(s) \geq \mathbf{R}(s, s')$) und

$$\begin{aligned} \mathbf{P}(s, S) &= \sum_{s' \in S} \mathbf{P}(s, s') \\ &= \sum_{s' \in S} (\mathbf{R}(s, s')/E(s)) \\ &= (\sum_{s' \in S} \mathbf{R}(s, s'))/E(s) \\ &= E(s)/E(s) \\ &= 1 \end{aligned}$$

folgt, dass $\mathbf{P}(s, \cdot)$ eine Wahrscheinlichkeitsverteilung ist. □

Um mit Wahrscheinlichkeiten von Pfaden, oder besser gesagt von Pfadmengen, arbeiten zu können, benötigen wir ein Wahrscheinlichkeitsmaß. Maße sind im mathematischen Sinne Bewertungsfunktionen von Messräumen mit bestimmten Eigenschaften. In unserem Fall wäre als Messraum ein Raum zu wählen, in dem sämtliche Pfadmengen und Kombinationen aus Pfadmengen repräsentiert werden, sodass wir für jede Pfadmenge ein Wahrscheinlichkeitsmaß bestimmen können.

Wir benötigen jedoch noch die Definitionen für Borel-Räume und Wahrscheinlichkeitsräume, auf denen dann das Wahrscheinlichkeitsmaß definiert werden kann.

Definition 9 (Borel-Raum). Die Menge $\mathcal{B} \subseteq 2^\Omega$ heißt Borel-Raum über Ω , falls gilt:

- $\Omega \in \mathcal{B}$
- aus $E \in \mathcal{B}$ folgt $\Omega \setminus E \in \mathcal{B}$
- aus $E_i \in \mathcal{B}$ für alle Elemente einer abzählbaren Menge $\{E_i \mid i \in \mathbb{N}\}$ folgt $\bigcup_{i \in \mathbb{N}} E_i \in \mathcal{B}$

Ein Borel-Raum kann von einer *abzählbaren* Menge \mathcal{E} erzeugt werden, indem die Menge bezüglich Komplement und Vereinigung abgeschlossen wird. Die einzelnen Teilmengen eines Borel-Raums nennen wir *meßbar*.

Definition 10 (Wahrscheinlichkeitsraum). Als Wahrscheinlichkeitsraum bezeichnen wir $\mathcal{PS} = (\Omega, \mathcal{B}, Pr)$, wobei \mathcal{B} der aus Ω erzeugte Borel-Raum ist. Die Funktion $Pr : \mathcal{B} \rightarrow [0, 1]$ nennen wir auch das Wahrscheinlichkeitsmaß. Sie ordnet jedem Element des Borel-Raums eine Wahrscheinlichkeit zu, wobei im Sinne der Kolmogorov-Axiome gelten muss, dass $Pr(\Omega) = 1$ und $Pr(\bigcup_{i \in \mathbb{N}} E_i) = \sum_{i \in \mathbb{N}} Pr(E_i)$ für eine abzählbare Menge $\{E_i \mid i \in \mathbb{N}\}$ von paarweise disjunkten Elementen des Borel-Raums.

Pfadmengen, deren Pfade durch gemeinsame Anfänge $s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} \dots \xrightarrow{t_{k-1}} s_k$ mit $t_i \in I_i$ für $i = 0, \dots, k-1$ charakterisiert sind, notieren wir mit $C(s_0, I_0, \dots, s_k)$. Diese Pfadmengen stellen die Grundereignisse des Wahrscheinlichkeitsraums dar und werden daher auch Basiszylindermengen genannt. In dem durch die Basiszylindermengen induzierten Wahrscheinlichkeitsraum können wir nun ein Wahrscheinlichkeitsmaß $Pr_{\tilde{\alpha}}$ für eine Startverteilung⁷ $\tilde{\alpha}$ induktiv definieren durch:

$$\begin{aligned} Pr_{\tilde{\alpha}}(C(s_0)) &= \alpha(s_0) \text{ und} \\ Pr_{\tilde{\alpha}}(C(s_0, I_0, \dots, s_{k-1}, I_{k-1}, s_k)) \\ &= Pr_{\tilde{\alpha}}(C(s_0, I_0, \dots, s_{k-1})) \cdot \mathbf{P}(s_{k-1}, s_k) \cdot (e^{-E(s_{k-1}) \cdot \inf I_{k-1}} - e^{-E(s_{k-1}) \cdot \sup I_{k-1}}) \\ &= Pr_{\tilde{\alpha}}(C(s_0, I_0, \dots, s_{k-1})) \cdot \mathbf{P}(s_{k-1}, s_k) \cdot X(s_{k-1}, I_{k-1}) \end{aligned}$$

für $k \geq 1$

Dabei beschreibt $X(s, I)$ die Wahrscheinlichkeit, dass die Verweildauer eines Zustands s innerhalb eines Zeitintervalls I liegt. Dies ergibt sich direkt aus der Stammfunktion der Exponentialverteilung:

$$X(s, I) = \int_I E(s) \cdot e^{-E(s) \cdot t} dt = e^{-E(s) \cdot \inf I} - e^{-E(s) \cdot \sup I}$$

⁷Statt Startzustandsmengen, wie wir sie aus den Transitionssystemen kennen, können im probabilistischen Fall auch Startverteilungen untersucht werden. Diese geben an, mit welcher Wahrscheinlichkeit sich anfangs das System in den jeweiligen Zuständen befindet.

Beispiel 3 (TMR als zeitstetige Markov-Kette). Nehmen wir nun das Beispiel 1 und fügen eine zeitliche Komponente ein. Wir geben die Ausfallrate der Komponenten mit λ an, die der votereinheit mit γ und die Reparaturzeiten mit μ für die Komponenten und δ für das gesamte System⁸, so erhalten wir die zeitstetige Markov-Kette in Abbildung 4.

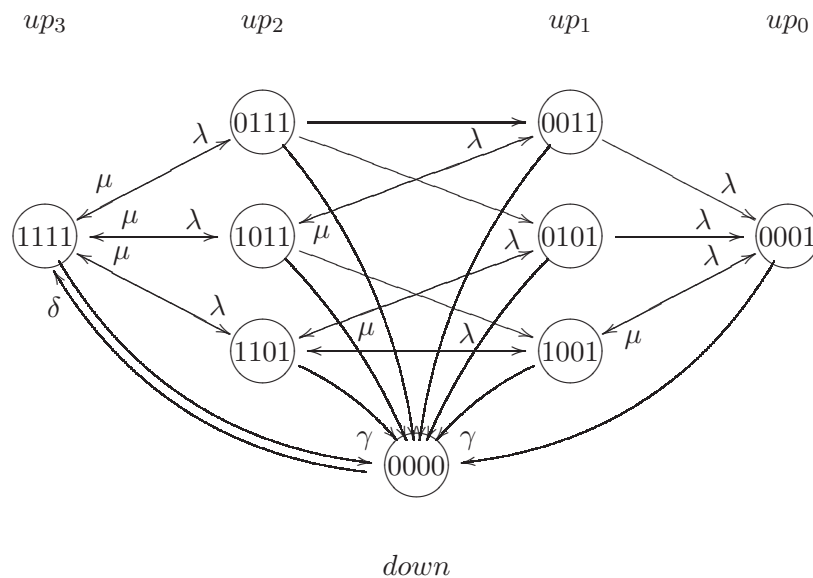


Abbildung 4: TMR als zeitstetige Markov-Kette

Das Maß für die Wahrscheinlichkeit der Pfadmengung mit gemeinsamen Pfadanzug 0000 $\xrightarrow{[0,10]}$ 1111 und der Startverteilung mit $\tilde{a}(0000) = 1$ ergibt sich dann wie folgt:

$$\begin{aligned}
 Pr_{\tilde{a}}(C(0000, [0, 10], 1111)) & \\
 &= Pr_{\tilde{a}}(C(0000)) \cdot \mathbf{P}(0000, 1111) \cdot X(0000, [0, 10]) \\
 &= 1 \cdot (\mathbf{R}(0000, 1111)/E(0000)) \cdot (e^{-E(0000) \cdot 0} - e^{-E(0000) \cdot 10}) \\
 &= 1 - e^{-\delta \cdot 10}
 \end{aligned}$$

2.3 CSL (Continuous Stochastic Logic)

Um Spezifikationen bezüglich Markov-Ketten formulieren zu können ist die *Computation Tree Logic* nicht mehr ausreichend. Ihr fehlen Möglichkeiten, quantitative Aussagen der Art zu treffen, dass die Erreichbarkeit eines Zustands mit mindestens 90 Prozent gegeben ist, oder dass mit signifikanter Wahrscheinlichkeit innerhalb der nächsten t Zeiteinheiten kein absorbierender Zustand erreicht werden kann. Um solche Spezifikationen formulieren zu können führen wir nun CSL ein, das eine probabilistische Erweiterung von CTL darstellt. Da in dieser Arbeit nicht vom *Steady-State*

⁸In der Abbildung sind der Übersichtlichkeit halber die Werte auf einigen Kanten nicht eingetragen, wo eine Nachbarkante beschriftet ist.

Operator (siehe [BHHK03]) die Rede sein wird, werden wir ihn auch schon bei der Definition von CSL auslassen.

Zunächst definieren wir die Syntax, die den Operator \mathcal{P} zur Beschreibung von Wahrscheinlichkeiten enthält und außerdem die Pfadoperatoren im Vergleich zu CTL um ein einzuhaltendes Zeitintervall I erweitert. Im Anschluß wird dann die Semantik der Erweiterungen erläutert.

Definition 11 (Continuous Stochastic Logic). Die Menge $\mathbb{F}_{\mathcal{M}}$ sei die Menge aller CSL-Formeln bezüglich einer Markov-Kette $\mathcal{M} = (S, \mathbf{R}, L)$, bestehend aus zustandsbezogenen Formeln $\mathbb{S}_{\mathcal{M}}$ und pfadbezogenen Formeln $\mathbb{P}_{\mathcal{M}}$. Als atomare zustandsbezogene CSL-Formeln werden bezeichnet:

- $true \in \mathbb{S}_{\mathcal{M}}$
- $a \in \mathbb{S}_{\mathcal{M}}$ falls $a \in AP$

Weiterhin wird die Menge der zustandsbezogenen CSL-Formeln nach folgenden Regeln mit $\boxtimes \in \{<, \leq, \geq, >\}$ induktiv aufgebaut:

- $(\neg\psi) \in \mathbb{S}_{\mathcal{M}}$ falls $\psi \in \mathbb{S}_{\mathcal{M}}$
- $(\psi_1 \wedge \psi_2) \in \mathbb{S}_{\mathcal{M}}$ falls $\psi_1, \psi_2 \in \mathbb{S}_{\mathcal{M}}$
- $(\mathcal{P}_{\boxtimes p}\Psi) \in \mathbb{S}_{\mathcal{M}}$ falls $\psi \in \mathbb{P}_{\mathcal{M}}$

Die Menge der pfadbezogenen CSL-Formeln setzt sich wie folgt induktiv zusammen:

- $(\mathcal{X}^I\psi) \in \mathbb{P}_{\mathcal{M}}$ falls $\psi \in \mathbb{S}_{\mathcal{M}}$
- $(\psi_1 \mathcal{U}^I \psi_2) \in \mathbb{P}_{\mathcal{M}}$ falls $\psi_1, \psi_2 \in \mathbb{S}_{\mathcal{M}}$

—

Die Semantik der aus CTL bekannten Operatoren bleibt gleich. Einige Operatoren aus CTL wurden jedoch nicht in der Definition übernommen, da sie obsolet geworden sind. Um dies zu erläutern benötigen wir jedoch die Semantik des Operators \mathcal{P} , mit dem man Aussagen über Pfadwahrscheinlichkeiten treffen kann, also über die Wahrscheinlichkeit, dass die Pfade ab dem gegebenen Zustand die Pfadeigenschaft der Teilformel erfüllen. Mit diesem Operator lassen sich nun die Pfadeigenschaften aus CTL als Abkürzungen auffassen.

Wir gehen im Folgenden von einer *fairen* Semantik aus. Das bedeutet, dass gewisse nicht faire Pfade von der Betrachtung ausgeschlossen werden. Beispielsweise in der Markov-Kette in Abbildung 5 ist genau ein unfairer Pfad enthalten, und zwar der unendliche Pfad $s_0 \longrightarrow s_0 \longrightarrow \dots$. Das unfaire an diesem Pfad ist, dass für den Zustandsübergang immer der *self-loop* gewählt wird, obwohl noch eine andere Transition existiert. Dieser Pfad hat jedoch lediglich eine Wahrscheinlichkeit von 0, wird also praktisch nie eingeschlagen. Da alle fairen Pfade in s_1 enden, gilt in einer fairen Semantik bezüglich dieser Markov-Kette die Formel $A \diamond at_{s_1}$ ⁹.

⁹Zu Model Checking für faire Semantiken, siehe [EL87].

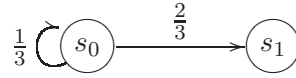


Abbildung 5: Zu fairer Semantik

Die Existenz eines Pfades mit Eigenschaften Ψ ist gegeben, wenn die Wahrscheinlichkeit echt größer als 0 ist, dass es ein solchen gibt, also $E\Psi \equiv \mathcal{P}_{>0}(\Psi)$. Die Allquantifizierung kann man dadurch ausdrücken, dass man für die Pfadwahrscheinlichkeit keinen kleineren Wert als 1 zulässt, also durch $A\Psi \equiv \mathcal{P}_{\geq 1}(\Psi)$. Für Pfade, auf denen zumindest in einem Zustand eine Zustandseigenschaft ψ im Zeitintervall I erfüllt ist, gilt, dass die Wahrscheinlichkeit einen solchen Zustand in I zu erreichen größer als 0 sein muss. Dies kann man über die Äquivalenz $\mathcal{P}_{\geq p}(\diamond^I \psi) \equiv \mathcal{P}_{\geq p}(\text{true } \mathcal{U}^I \psi)$ ausdrücken. Genau die Pfade, auf denen in jedem Zustand eine Zustandseigenschaft ψ erfüllt sein muss, erfüllen auch die Eigenschaft, dass $\neg\psi$ in keinem Zustand erfüllt sein darf. Wenn die Wahrscheinlichkeit für Pfade, in denen $\neg\psi$ in keinem Zustand erfüllt ist, größer ist als $1 - p$, dann muss also die Wahrscheinlichkeit für solche Pfade, in denen in jedem Zustand ψ erfüllt sein muss kleiner sein als p . Wir drücken dies und den Fall mit „kleiner als $1 - p$ “ also durch die Äquivalenzen $\mathcal{P}_{\leq p}(\square^I \psi) \equiv \mathcal{P}_{\geq 1-p}(\diamond^I \neg\psi)$ und $\mathcal{P}_{\geq p}(\square^I \psi) \equiv \mathcal{P}_{\leq 1-p}(\diamond^I \neg\psi)$ aus, wobei $\leq \in \{<, \leq\}$, $\geq \in \{\geq, >\}$, und $\geq \Rightarrow$ genau dann, wenn $\leq = <$ und $\geq = \geq$ genau dann, wenn $\leq = \leq$.

Die Erweiterung der Operatoren \mathcal{X} und \mathcal{U} um Zeitintervallgrenzen ist intuitiv leicht zu verstehen. Für eine Pfadformel $\mathcal{X}^I \psi$ muss die Zustandseigenschaft ψ innerhalb des gegebenen Intervalls I im Zustand $\sigma[1]$ des Pfades σ erfüllt sein. Entsprechendes gilt für den *Until*-Operator. Die CTL-Versionen von *Next* und *Until* erhält man, indem man als Intervall $[0, \infty)$ wählt.

Die formale Semantik ist in Tabelle 3 definiert, wobei die mit CTL übereinstimmenden Operatoren nicht noch einmal aufgeführt wurden. Deren Semantik kann nach wie vor in Tabelle 2 nachgeschlagen werden. In Tabelle 3 wird die Menge der Zustände, in denen eine Zustandsformel ψ erfüllt ist, mit $Sat(\psi) = \{s \mid \llbracket s, \psi \rrbracket = \top\}$ bezeichnet.

Das Model Checking für CSL kann man für den nicht-probabilistischen Teil aus CTL übernehmen. Was als Fragestellung übrig bleibt ist, wie die Erfüllbarkeit von \mathcal{P} -Formeln bestimmt werden kann. Eine Quelle zum Thema CSL Model Checking in der diese Fragestellung umfassend behandelt wird ist [BHHK03]. Für einen etwas sanfteren Einstieg in das Thema ist auch [Hav02] sehr zu empfehlen. Im folgenden Kapitel werden wir nur die *Transient Analyse* vorstellen, die den Kern der Verifikation von \mathcal{P} -Formeln darstellt und auch im weiteren Verlauf noch auf *abstrakte zeitstetige Markov-Ketten* übertragen werden wird.

$$\begin{aligned}
\llbracket \sigma, \mathcal{X}^I \psi \rrbracket &= \begin{cases} \llbracket \sigma, 1, \psi \rrbracket & \text{falls } \delta(\sigma, 0) \in I \\ \perp & \text{sonst} \end{cases} \\
\llbracket \sigma, \psi_1 \mathcal{U}^I \psi_2 \rrbracket &= \begin{cases} \top & \text{falls } \exists t \in I : (\llbracket \sigma @ t, \psi_2 \rrbracket = \top \\ & \wedge \forall t' \in [0, t) : \llbracket \sigma @ t', \psi_1 \rrbracket = \top) \\ \perp & \text{falls } \forall t \in I : (\llbracket \sigma @ t, \psi_2 \rrbracket = \perp \\ & \vee \exists t' \in [0, t) : \llbracket \sigma @ t', \psi_1 \rrbracket = \perp) \end{cases} \\
\llbracket s, \mathcal{P}_{\boxtimes p}(\Psi) \rrbracket &= \begin{cases} \top & \text{falls } Pr\{\sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, \Psi \rrbracket = \top\} \boxtimes p \\ \perp & \text{sonst} \end{cases}
\end{aligned}$$

mit $\psi, \psi_1, \psi_2 \in \mathbb{S}^{\mathcal{M}}$; $\Psi \in \mathbb{P}^{\mathcal{M}}$; $\boxtimes \in \{<, \leq, \geq, >\}$; $p \in [0, 1], I \subseteq [0, \infty)$

für eine gegebene zeitstetige Markov-Kette $\mathcal{M} = (S^{\mathcal{M}}, \mathbf{R}^{\mathcal{M}}, L^{\mathcal{M}})$

Tabelle 3: CSL Semantik

2.4 Transient Analyse und Uniformisierung für CTMCs

Eine Klasse von zeitstetigen Markov-Ketten mit besonderer Bedeutung für die Berechnung von Erreichbarkeitswahrscheinlichkeiten sind die *uniformen* CTMCs. Sie zeichnen sich dadurch aus, dass die Exitraten für alle Zustände gleich sind. Für solche Markov-Ketten kann man das um Zeitschranken erweiterte Erreichbarkeitsproblem, also die Frage, ob eine Menge von Zuständen innerhalb einer gewisser Zeitgrenze erreicht werden kann, algorithmisch effizient lösen.

Definition 12 (Uniforme zeitstetige Markov-Kette). Eine zeitstetige Markov-Kette $\mathcal{M} = (S, \mathbf{R}, L)$ heißt *uniform*, falls ein $E_{unif} \in \mathbb{R}_{\geq 0}$ existiert, sodass für alle $s \in S$ gilt $E(s) = E_{unif}$.

Die Grundidee bei der Berechnung besteht darin, für alle Pfade einer bestimmten Länge die Wahrscheinlichkeit des Erreichens der gegebenen Menge von Zuständen nach der eingebetteten *zeitdiskreten* Markov-Kette zu bestimmen. In dieser kommen lediglich die Wahrscheinlichkeiten für verschiedene Nachfolger von Zuständen zum Ausdruck, nicht jedoch die zeitlichen Eigenschaften. Auf diese Weise kann man für Pfade der Längen $l = 0, 1, 2, \dots$ die zeitunabhängigen Erreichbarkeitswahrscheinlichkeiten bestimmen. Um nun die Zeit zu berücksichtigen, in der ein Zustand der Zielmenge erreicht werden soll, bedienen wir uns der *Poisson-Verteilung* $\varphi(\lambda \cdot t, n)$. Sie ist dafür bestens geeignet, da sie die Wahrscheinlichkeit des Auftretens von n unabhängigen Ereignissen (hier Zustandsübergängen) innerhalb t Zeiteinheiten beschreibt, wobei mit λ der Erwartungswert für das Auftreten des Ereignisses in einer Zeiteinheit gegeben ist (hier durch die Raten repräsentiert). Multiplizieren wir

also zu der Wahrscheinlichkeit eines Pfades der Länge l die Wahrscheinlichkeit für l Ereignisse mit Rate E_{unif} innerhalb einer Zeiteinheit, so erhalten wir die Wahrscheinlichkeit für das Erreichen der Zielmenge mit Pfaden der Länge l . Über eine unendliche Summe über alle Pfadlängen $l \in \mathbb{N}$ kann dann die Wahrscheinlichkeit bei beliebiger Pfadlänge berechnet werden.

Da für die Poisson-Verteilung gilt, dass $\lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} \varphi(\lambda \cdot t, n) = 0$, kann die Berechnung¹⁰ bei Erreichen einer erwünschten Genauigkeit gestoppt werden. Denn selbst wenn die Wahrscheinlichkeit für Pfade der Länge k und größer immer gleich Eins ist, so ist der Gesamtbeitrag für diese Pfade durch $\sum_{n=k}^{\infty} \varphi(\lambda \cdot t, n)$ begrenzt. Bei einer gewünschten Genauigkeit ε kann die Berechnung also abgebrochen werden, sobald $\sum_{n=k}^{\infty} \varphi(\lambda \cdot t, n) = 1 - \sum_{n=0}^{k-1} \varphi(\lambda \cdot t, n) \leq \varepsilon$ erfüllt ist.

Kommen wir nun zu einem formalen Ansatz. Die Wahrscheinlichkeitsverteilung, ausgehend von einer Startverteilung \tilde{a} nach t Zeiteinheiten ist nach [Tij03] gegeben als Taylor-McLaurin Reihe:

$$\pi(\tilde{a}, t) = \tilde{a} \cdot e^{\mathbf{Q} \cdot t} = \tilde{a} \cdot \left(\sum_{i=0}^{\infty} (\mathbf{Q} \cdot t)^i / i! \right) \quad (2)$$

Die sogenannte Generatormatrix einer CTMC ist definiert durch $\mathbf{Q} = \mathbf{R} - \text{diag}(E)$, wobei die Funktion diag einen Vektor so auf eine quadratische Matrix abbildet, dass die Einträge auf der Diagonalen den Vektoreinträgen entsprechen und die übrigen Matrixeinträge gleich 0 sind. Die Generatormatrix kann auch so aufgefasst werden, dass in den Diagonaleinträgen nicht mehr die Raten des *self-loop*¹¹ stehen wie bei der Ratenmatrix, sondern der negativen Summe der in *andere* Zustände führenden Raten. Dadurch wird mit den Diagonalwerten in dem Maße ein negativer Beitrag für die Wahrscheinlichkeit geliefert, sich in einem bestimmten Zustand zu befinden, in dem dieser Zustand in Richtung anderer Zustände verlassen wird. Befindet sich ein System beispielsweise zum Zeitpunkt t sicher in einem Zustand s , mit *self-loop* mit Rate 1 und ausgehenden Kanten mit gesamter Rate 3, so wird der nächste Zustandsübergang mit Wahrscheinlichkeit $\frac{3}{4}$ aus dem Zustand herausführen und die Wahrscheinlichkeit sich weiterhin in Zustand s zu befinden wird sich entsprechend um $\frac{3}{4}$ auf $\frac{1}{4}$ verringern.

Es sei nochmals darauf hingewiesen, dass nach der Taylor-McLaurin Reihe die Wahrscheinlichkeitsverteilung bezüglich der Zuständen der Markov-Kette nach t Zeiteinheiten von der Generatormatrix \mathbf{Q} direkt abhängig ist und nicht von der Wahrscheinlichkeitsmatrix \mathbf{P} oder der Ratenmatrix \mathbf{R} .

Aufgrund der Definition der Generatormatrix können nun den einzelnen Zuständen beliebig *self-loops* hinzugefügt werden, oder die Rate des *self-loop* erhöht werden. Durch eine Erhöhung des *self-loop* eines Zustandes s_i um δ erhält man auch eine um δ erhöhte Exitrate für s_i . Die Generatormatrix verändert sich dadurch jedoch nicht:

¹⁰Die Poisson-Verteilung kann nur für verhältnismäßig kleine λ direkt berechnet werden. In [FG88] wird eine Möglichkeit für die Berechnung größerer λ beschrieben.

¹¹Als *self-loop* bezeichnet man eine Transition von einem Zustand zu sich selbst.

$$\mathbf{Q}(s_i, s_i) = (\mathbf{R}(s_i, s_i) + \delta) - (E(s_i) + \delta) = \mathbf{R}(s_i, s_i) - E(s_i)$$

Dies können wir nun ausnutzen, um eine nicht-uniforme Markov-Kette zu einer uniformen Markov-Kette mit denselben *Transient* Eigenschaften zu transformieren. Da nur eine Erhöhung der *self-loop* Raten möglich ist, müssen wir eine uniforme Exitrate $E_{unif} \geq \max\{E(s_i)\}$ wählen.

Für uniforme Markov-Ketten und nach Substitution von \mathbf{Q} durch $E_{unif} \cdot (\mathbf{U} - I)$ mit Einheitsmatrix I , lässt sich die Gleichung (2) schreiben als:

$$\begin{aligned} \pi(\tilde{a}, t) &= \tilde{a} \cdot \left(\sum_{i=0}^{\infty} e^{E_{unif} \cdot t} \cdot \frac{(E_{unif} \cdot t)^i}{i!} \cdot \mathbf{U}^i \right) \\ &= \tilde{a} \cdot \left(\sum_{i=0}^{\infty} \varphi(E_{unif} \cdot t, i) \cdot \mathbf{U}^i \right) \\ &= \sum_{i=0}^{\infty} \varphi(E_{unif} \cdot t, i) \cdot \pi_i \end{aligned}$$

Interpretiert man \mathbf{U} als Wahrscheinlichkeitsmatrix einer DTMC, so entspricht diese Formel genau der *Transient* Analyse für uniforme Markov-Ketten nach der ersten Beschreibung weiter oben.

Mit π_i bezeichnen wir die Wahrscheinlichkeitsverteilung nach i Zeiteinheiten in der zeitdiskreten Markov-Kette mit Wahrscheinlichkeitsmatrix \mathbf{U} . Der Vektor π_i kann rekursiv berechnet werden durch $\pi_0 = \tilde{a}$ und $\pi_{i+1} = \pi_i \cdot \mathbf{U}$ für $i \in \mathbb{N}$. Diese Berechnung ähnelt stark dem Markierungsalgorithmus, den wir im Zusammenhang mit den Transitionssystemen kennengelernt hatten. Statt der schrittweisen Übertragung der Markierungen führen wir hier in jedem Iterationsschritt eine Multiplikation von zwei Wahrscheinlichkeiten durch. Erstens der Wahrscheinlichkeit, von den Zuständen der Markov-Kette aus mit i Schritten einen Zielzustand zu erreichen, und zweitens der Wahrscheinlichkeit, innerhalb eines Schritts diese Zustände zu erreichen.

Beispiel 4 (Uniformisierte und eingebettete Markov-Kette). Wir bestücken nun die Kanten des Transitionssystems aus Abbildung 2 mit Raten, sodass wir eine uniforme zeitstetige Markov-Kette erhalten. Die zugehörige eingebettete Markov-Kette ergibt sich dann aus der Division der Raten durch die uniforme Exitrate $E_{unif} = 6$ (Abbildung 6).

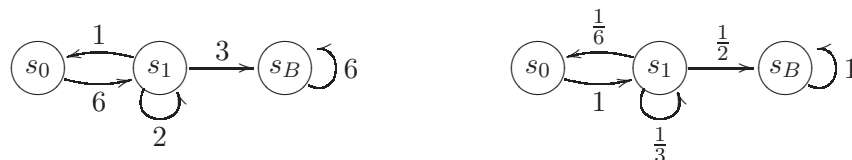


Abbildung 6: Uniformisierte (links) und eingebettete (rechts) Markov-Kette

Durch die Abwicklung und anschließende Zusammenfassung aller gleichen Zustände pro Ebene erhalten wir die eingebettete Markov-Kette mit Wahrscheinlichkeitsmatrix \mathbf{P} der Abbildung 7, in der die Wahrscheinlichkeiten π_i sich durch das Produkt der, zu den einzelnen Zustandsübergängen gehörenden, Wahrscheinlichkeiten berechnen lässt.

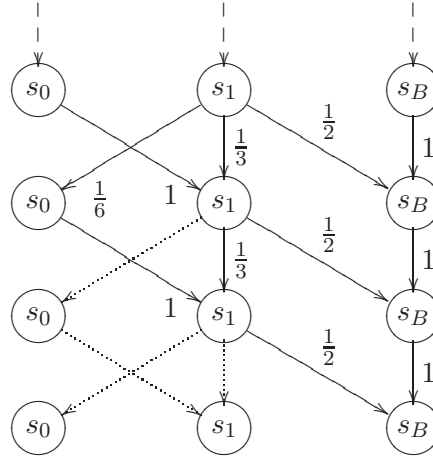


Abbildung 7: Zusammengefasste Abwicklung

Die ersten Werte für π_i berechnen sich dann wie folgt:

$$\pi_0 = \tilde{a} = (1, 0, 0)$$

$$\pi_1 = \pi_0 \cdot \mathbf{P} = (0, 1, 0)$$

$$\pi_2 = \pi_1 \cdot \mathbf{P} = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right)$$

$$\pi_3 = \pi_2 \cdot \mathbf{P} = \left(\frac{1}{6} \cdot \frac{1}{3}, \frac{1}{6} + \frac{1}{3} \cdot \frac{1}{3}, \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3}\right) = \left(\frac{1}{18}, \frac{5}{18}, \frac{2}{3}\right)$$

Damit ergeben sich die Wahrscheinlichkeiten für Pfade der Länge i von s_0 in den Zustand s_B zu gelangen als $\pi_0(s_B) = \pi_1(s_B) = 0$, $\pi_2(s_B) = \frac{1}{2}$, $\pi_3(s_B) = \frac{2}{3}$ und so weiter. Wollen wir die Wahrscheinlichkeit für Pfade unterschiedlicher Längen berücksichtigen, so müssen wir die gewünschte maximale Dauer t für gültige Pfade angeben. Für $t = \frac{1}{10}$ ergeben sich beispielsweise die folgenden Wahrscheinlichkeiten:

$$\varphi\left(\frac{1}{10} \cdot E_{unif}, 0\right) \approx 0,5488$$

$$\varphi\left(\frac{1}{10} \cdot E_{unif}, 1\right) \approx 0,3293$$

$$\varphi\left(\frac{1}{10} \cdot E_{unif}, 2\right) \approx 0,0988$$

$$\varphi\left(\frac{1}{10} \cdot E_{unif}, 3\right) \approx 0,0198$$

Da diese Wahrscheinlichkeiten für $i \leq 3$ in der Summe annähernd 1 ergeben, spielen hier längere Pfade als solche mit maximal drei Transitionen keine wesentliche Rolle¹². Man kann die Wahrscheinlichkeit von s_0 aus in $\frac{1}{10}$ Zeiteinheiten den Zustand s_B zu erreichen also approximieren durch:

$$\begin{aligned} & \left(\sum_{i=0}^3 \varphi\left(\frac{1}{10} \cdot E_{unif}, i\right) \cdot \pi_i \right) (s_B) \\ &= 0,5488 \cdot 0 + 0,3293 \cdot 0 + 0,0988 \cdot \frac{1}{2} + 0,0198 \cdot \frac{2}{3} \\ &\approx 0,079 \end{aligned}$$

¹²Veranschlagt man eine Genauigkeit im Bereich von $\varepsilon = \frac{1}{1000}$ oder genauer, so muss man entsprechend längere Pfade berücksichtigen.

2.5 Abstraktion mit Bisimulationsäquivalenzklassen

Bei der Generierung von Markov-Ketten aus höheren Beschreibungsmodellen wie kommunizierenden Prozessen tritt häufig das Problem der *Zustandsraumexplosion* auf. Um die Zustandsmengen der generierten Markov-Ketten besser in den Griff zu bekommen, kann man verschiedene Techniken zur Verkleinerung der Darstellung anwenden. Eine Möglichkeit besteht in der Abstraktion der Markov-Ketten durch die Ermittlung von bisimulationsäquivalenten Zuständen. Intuitiv kann man Bisimulationsäquivalenz (engl. auch *lumpability*) verstehen als eine Relation in der Zustände zueinander stehen, die bezüglich ihrer Eigenschaften und der nachfolgenden Zuständen nicht zu unterscheiden sind.

Definition 13 (Bisimulationsäquivalenz). Sei $\mathcal{M} = (S, \mathbf{R}, L)$ eine zeitstetige Markov-Kette, dann ist $\mathcal{R} \subseteq S \times S$ ein Bisimulationsäquivalenz, falls aus $s\mathcal{R}s'$ folgt, dass $L(s, a) = L(s', a)$ für alle $a \in AP$ und $\sum_{t \in C} \mathbf{R}(s, t) = \sum_{t \in C} \mathbf{R}(s', t)$ für alle Zustände des Zustandsquotienten $C \in S/\mathcal{R}$.

Falls eine Bisimulationsäquivalenz \mathcal{R} existiert, für die $s\mathcal{R}s'$ gilt, so schreibt man auch $s \cong s'$, sprich s und s' sind bisimilar.

—

Zwei bisimilare Zustände haben also dieselben atomaren Eigenschaften und gleiche Raten für Übergänge zu anderen Zuständen. Fasst man nun alle bisimilaren Zustände zusammen, so erhält man eine Abstraktion, in der nach [BHHK03] die Gültigkeit von CSL Formeln erhalten bleibt:

$$\begin{aligned} \mathcal{M}/\mathcal{R} &= (S/\mathcal{R}, \mathbf{R}_{\mathcal{R}}, L_{\mathcal{R}}) \\ \text{mit } \mathbf{R}_{\mathcal{R}}([s]_{\mathcal{R}}, C) &= \sum_{t \in C} \mathbf{R}(s, t) \text{ für alle } C \in S/\mathcal{R} \\ \text{und } L_{\mathcal{R}}([s]_{\mathcal{R}}, a) &= L(s, a) \text{ für alle Äquivalenzklassen } [s]_{\mathcal{R}} \text{ aus } \mathcal{R} \\ &\text{und alle atomaren Eigenschaften } a \in AP. \end{aligned}$$

Auf den Beweis für die Erhaltung der Gültigkeit von CSL Formeln werden wir hier verzichten, er ist jedoch nicht allzu schwierig und kann durch strukturelle Induktion über den Aufbau der Formeln geführt werden. Intuitiv sollte klar sein, dass jeder Zustand der abstrahierten Markov-Kette eine oder mehrere Entsprechungen in der ursprünglichen Markov-Kette hat, und dass jede Bewegung, egal in welcher der beiden Markov-Ketten, in der jeweils anderen so nachgestellt werden kann, dass die Zustände nach der Bewegung wieder bisimilar sind.

Beispiel 5 (Abstraktion mit Bisimulationsäquivalenz). Wendet man die Abstraktion auf das Beispiel 3 an, so erhält man die Markov-Kette¹³ in Abbildung 8.

Die Zustände $s_{i,1}$ repräsentieren die Situationen, in denen noch i Komponenten funktionstüchtig sind und in Zustand $s_{0,0}$ ist der Voter defekt, sodass das gesamte System erneuert werden muss.

¹³Dieses Beispiel ist aus [BHHK03] entnommen. Die angegebene Markov-Kette wurde jedoch fälschlicherweise als Abstraktion des Beispiels aus dem Kapitel über Bisimulation in [BHHK03] angegeben. Sie unterscheidet sich in einigen Raten von der korrekten ursprünglichen Markov-Kette in Beispiel 3.

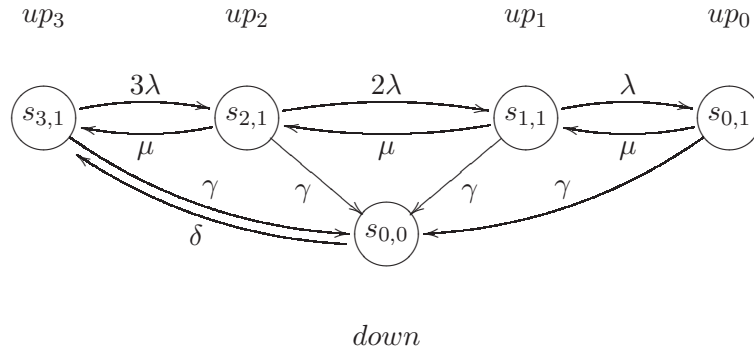


Abbildung 8: Bisimulationsäquivalente Abstraktion des TMR

Stellen wir uns vor, es gäbe leichte Variationen in der Ausfallsicherheit der einzelnen Komponenten, zum Beispiel da sie aus verschiedenen Produktionsreihen mit unterschiedlicher Fertigungsqualität stammen. Nehmen wir an, die Ausfallrate variiert bei den drei Komponenten um ε nach oben und nach unten. Dann könnte man etwa die Zustände 0011, 0101 und 1001 nicht mehr zu $s_{1,1}$ zusammenfassen.

Wenn ε relativ klein ist kann es dennoch Sinn machen diese drei Zustände als einen zu betrachten. Wir werden daher später abstrakte Markov-Ketten definieren, an deren Kanten Intervalle von Raten stehen können. In unserem Beispiel könnte man also an die Kante von $s_{1,1}$ nach $s_{0,1}$ das Intervall $[\lambda - \varepsilon, \lambda + \varepsilon]$ schreiben.

2.6 Markov-Entscheidungsprozesse

In weiteren Verlauf werden wir wie schon erwähnt abstrakte zeitstetige Markov-Ketten einführen, bei denen die Raten durch Ratenintervalle ersetzt werden. Dadurch erreichen wir eine kompakte Darstellung von unendlich vielen Raten, die in einem Intervall enthalten sein können.

Eine andere Möglichkeit wäre die Verwendung von so genannten *Markov-Entscheidungsprozessen*. Bei Entscheidungsprozessen wird in jedem Zustand eine nichtdeterministische Wahl verschiedener *Aktionen* erlaubt, in unserem Fall etwa alle möglichen Kombinationen von Raten. Für eine unendliche Menge von Raten müsste man also auch eine unendliche Menge von Aktionen festlegen.

Definition 14. Ein zeitstetiger Markov-Entscheidungsprozess $\mathcal{M} = (S, A, \mathbf{R})$ besteht aus der Zustandsmenge S , einer endlichen Menge A von Aktionen und einer dreidimensionalen Ratenmatrix $\mathbf{R} : S \times A \times S \mapsto \mathbb{R}_{\geq 0}$, die für jede Aktion aus A eine Ratenmatrix impliziert.

Der Nichtdeterminismus wird üblicherweise mit sogenannten *Schedulern* behandelt. Ein Scheduler ist eine Funktion, die für jeden Zustand eines Pfades eine Aktion, und damit den Ratenvektor für den Zustand bestimmt. Auf diese Weise wird durch

einen Scheduler eine (unendliche) zeitstetige Markov-Kette impliziert. Bei den Schemulern unterscheidet man üblicherweise zwischen verschiedenen Klassen. Je nach Klasse können etwa die bisherigen Entscheidungen auf einem Pfad berücksichtigt werden, oder die Entscheidung kann randomisiert getroffen werden. Auf eine formale Einführung soll hier verzichtet werden, mit [Put94] gibt es jedoch eine mathematisch fundierte Abhandlung zu diesem Thema.

Der Grund, warum im weiteren Verlauf nicht Entscheidungsprozesse zur Abstraktion verwendet werden ist der, dass die abstrakten Markov-Ketten eine wesentlich stärkere Abstraktion ermöglichen und damit auch den Speicherbedarf für computer-gestützte Berechnungen wesentlich stärker reduzieren können.

Beispiel 6 (Schwache Abstraktion durch Markov-Entscheidungsprozesse). Gegeben sei die zeitstetige Markov-Kette aus Abbildung 9 links. Abstrahiert man diese, indem man die Zustände s_0 bis s_2 zu s_* zusammenfasst und für jede Kante eine entsprechende Aktion r_0, r_1 und r_2 einführt, welche die jeweilige Auswahl der Rate impliziert, so erhält man den Markov-Entscheidungsprozess in der mittleren Abbildung. Wie man leicht sehen kann gewinnt man durch diese Art der Abstraktion bezüglich der Zahl der benötigten Kanteninformationen nichts. Die Abstraktion in der rechten Abbildung dagegen fasst sämtliche Kanteninformation für ein Paar von Zuständen in einem Intervall zusammen, sodass hier nur noch eine obere Schranke $r^u = \max\{r_0, r_1, r_2\}$ und eine untere Schranke $r^l = \min\{r_0, r_1, r_2\}$ angegeben werden muss. Diese Art der Abstraktion wird auch im weiteren Verlauf verwendet werden.

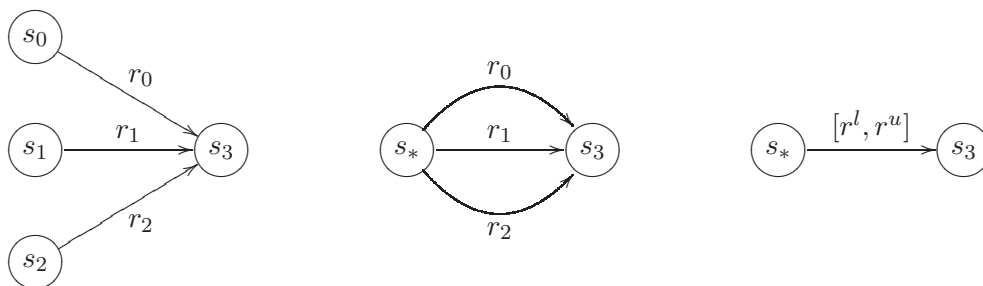


Abbildung 9: Schwache Abstraktion durch Markov-Entscheidungsprozesse

2.7 Dreiwertige Logiken

Die in der klassischen Logik verwendeten Wahrheitswerte *wahr* und *falsch* sind nicht in allen Bereichen der Logik ausreichen, etwa wenn uns nicht nur interessiert, ob eine Eigenschaft sicher erfüllt wird, sondern auch ob sie möglicherweise sicher verletzt wird. Wenn man beispielsweise über eine gewisse Grundeigenschaft keine Informationen besitzt, so kann es sein, dass andere davon abhängige Eigenschaften ebenfalls nicht mit Sicherheit erfüllt oder verletzt werden. Genau diese Situation kann in den abstrakten Markov-Ketten des folgenden Kapitels auftreten. Wir führen daher

zusätzlich zu den Wahrheitswerten der klassischen Logik \top für *gültig* und \perp für *ungültig* nun einen einen dritten Wahrheitswert $?$ für *unbestimmt* ein.

Über die Menge der Wahrheitswerte $\mathbb{B}^3 = \{\perp, ?, \top\}$ kann wie bei der klassischen Logik ein vollständiger Verband gebildet werden mit $\perp < ? < \top$. Die Tabelle 4 für *join* und *meet* in diesem Verband ist eine Erweiterung der Tabelle 1. Die Werte für Kombinationen aus \perp und \top sind identisch. Betrachten wir also die Einträge etwas näher, bei denen ein Wert mit $?$ gegeben ist. Für den *join*, ergibt sich nur im Fall, dass der andere Wert \top ist ebenfalls \top und ansonsten $?$. Dies passt zur Semantik der Disjunktion, da die Gültigkeit von einem Wert genügt, damit der gesamte Ausdruck gültig wird. Für *meet* wird dagegen der gesamte Ausdruck ungültig, sobald einer der Werte \perp ist, was der Semantik der Disjunktion entspricht. Definiert man das Komplement als $\perp^c = \top$, $?^c = ?$ und $\top^c = \perp$, so kann die Semantik der dreiwertigen Aussagenlogik genau wie die der zweiwertigen über *meet* und Komplement definiert werden.

\sqcup	\perp	$?$	\top	\sqcap	\perp	$?$	\top
\perp	\perp	$?$	\top	\perp	\perp	\perp	\perp
$?$	$?$	$?$	\top	$?$	\perp	$?$	$?$
\top	\top	\top	\top	\top	\perp	$?$	\top

Tabelle 4: *join* (\sqcup) und *meet* (\sqcap) für \mathbb{B}^3

Übertragen wir nun die Kolmogorov-Axiome auf den dreiwertigen Fall. Das Wahrscheinlichkeitsmaß sei mit $Pr : \Omega \times \mathbb{B}^3 \mapsto \mathbb{R}_{\geq 0}$ für den von der Ereignismenge $\Omega \times \mathbb{B}^3$ induzierten Wahrscheinlichkeitsraum gegeben. Dann muss das Folgende gelten:

- Ereigniswahrscheinlichkeiten sind reelle Zahlen zwischen 0 und 1:

$$Pr(E, \alpha) \in [0, 1] \text{ für } \alpha \in \mathbb{B}^3 \quad (3)$$

- Das sichere Ereignis hat die Wahrscheinlichkeit 1:

$$Pr(\Omega, \top) = 1 \quad (4)$$

- Die Wahrscheinlichkeit für die Vereinigung von unabhängigen Ereignissen ist gleich der Summe über die Wahrscheinlichkeit der einzelnen Ereignisse:

$$Pr(E_1 \cup E_2, \top) = Pr(E_1, \top) + Pr(E_2, \top) \text{ falls } Pr(E_1 \cap E_2, \perp) = 1 \quad (5)$$

Die üblichen Folgerungen, die man aus den Kolmogorov-Axiomen zieht, lassen sich nur bedingt übertragen. Beispielsweise ist die Wahrscheinlichkeit für ein Ereignis zusammen mit dem komplementären Ereignis nicht gleich 1. Dieser Umstand ergibt sich daraus, dass es einen unbestimmten Anteil für ein Ereignis geben kann. Es muss jedoch offensichtlich $Pr(E, \perp) + Pr(E, ?) + Pr(E, \top) = 1$ gelten, woraus sich nach folgendem Lemma ein vergleichbarer Schluss ziehen lässt.

Lemma 2. Wenn die Wahrscheinlichkeit des Eintretens eines Ereignisses $\neg E \in \Omega$ einen Wert $(1 - p) \in [0, 1]$ erreicht oder übersteigt, so ist die Wahrscheinlichkeit für das Eintreten des Gegenereignisses A kleiner oder gleich p .

Beweis.

$$Pr(\neg E, \top) \geq 1 - p$$

$$\Rightarrow Pr(E, \perp) \geq 1 - p \quad (\llbracket \neg E \rrbracket = \llbracket E \rrbracket^c = \top \text{ gdw. } \llbracket E \rrbracket = \perp)$$

$$\Rightarrow Pr(E, \top) + Pr(E, ?) \leq p \quad (\text{wg. 4})$$

$$\Rightarrow Pr(E, \top) \leq p \quad (\text{wg. 3})$$

□

3 Abstraktion für zeitstetige Markov-Ketten

In Abschnitt 2.5 haben wir eine Abstraktionstechnik kennengelernt, bei der je ein abstrakter Zustand eine Bisimulationsäquivalenzklasse repräsentiert. Um weitere Abstraktion, also das Zusammenfassen von Zuständen, die nicht zwingend bisimulationsäquivalent sind, zu ermöglichen, führen wir in diesem Kapitel die Notation der *abstrakten* zeitstetigen Markov-Ketten ein. Wir werden dazu verschiedene Arten von Schemulern definieren, um ähnlich wie für Markov-Entscheidungsprozesse den Nichtdeterminismus behandeln zu können. Um korrekt arbeitende Scheduler für abstrakte Markov-Ketten zu beschreiben, werden wir insbesondere eine *cut*-Funktion einführen müssen, die dazu benutzt wird, um Werte der abstrakten Markov-Kette zu entfernen, die durch den Scheduler nicht zu einer gültigen konkreten Markov-Kette vervollständigt werden können.

Nachdem wir eine *Simulationsrelation* für zeitstetige Markov-Ketten angegeben haben, werden wir verschiedene Möglichkeiten untersuchen, wie man aus einer Markov-Kette durch Zusammenfassen von Zuständen zu sogenannten Makro-Zuständen eine abstrakte Markov-Kette erzeugen kann. Die Abstraktionen sind so gewählt, dass die Makro-Zustände jeweils die zugehörigen Zustände der ursprünglichen Markov-Kette simulieren, sodass wir im nächsten Kapitel auf die Abstraktionen zurückgreifen können, um Eigenschaften für eine konkrete Markov-Kette nachzuweisen.

Für das Erreichbarkeitsproblem des nächsten Kapitels von Bedeutung ist außerdem eine Möglichkeit zur Uniformisierung. Wir werden feststellen, dass die drei Abstraktionen für uniformisierte abstrakte Markov-Ketten dieselben Ergebnisse liefern. Zusätzlich wird der *cut* durch die feste Exitrate deutlich einfacher werden. Dennoch werden die vorherigen Erkenntnisse nicht umsonst sein, da eine vorherige Uniformisierung nicht für *Next*-Formeln zulässig ist.

In einem kurzen Unterkapitel soll dann noch geklärt werden, in welchem Verhältnis die hier vorgestellte Abstraktion zeitstetiger Markov-Ketten und die Abstraktion zeitdiskreter Markov-Ketten aus [FLW06] stehen.

Abschließend definieren wir den Wahrscheinlichkeitsraum zu einer abstrakten zeitstetigen Markov-Kette, wodurch es uns möglich sein wird, über Wahrscheinlichkeiten von bestimmten Pfadmengen zu sprechen. Wir werden auch feststellen, dass es für die maximalen und minimalen Wahrscheinlichkeiten solcher Pfadmengen genügt, eine bestimmte abzählbare Menge von Schemulern zu betrachten.

3.1 Abstrakte zeitstetige Markov-Ketten

Bevor wir den Begriff der abstrakten zeitstetigen Markov-Kette formal definieren, wollen wir uns an einem kleinen Beispiel verdeutlichen worum es sich dabei handeln soll.

Beispiel 7 (TMR als abstrakte zeitstetige Markov-Kette). In Beispiel 5 hatten wir schon angedeutet, wie eine abstrakte zeitstetige Markov-Kette aussehen könnte, und zwar durch die Erweiterung der Raten auf Intervalle. Nehmen wir nun also an, dass die Ausfallraten aufgrund von Produktionsschwankungen statistisch nachweisbar voneinander abweichen. Die erste Komponente habe dabei die Rate $\lambda + \varepsilon$ und die beiden anderen jeweils λ . Damit ergibt sich zunächst eine zeitstetige Markov-Kette wie in Beispiel 3 mit entsprechend angepassten Raten. Abstrahiert man diese mit der in 2.5 vorgestellten Methode, so führt dies zu folgender Markov-Kette:

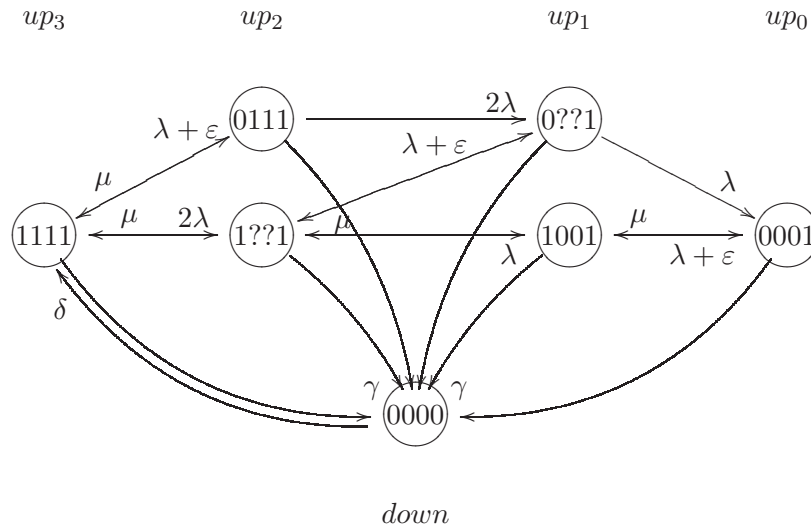


Abbildung 10: Variante des TMR als CTMC

Die Zustände $0??1$ und $1??1$ beschreiben die Situationen, dass die Funktionstüchtigkeit des Voters und der ersten Komponente bekannt ist, und dass genau eine der beiden anderen Komponenten defekt ist.

Ist ε so klein, dass es sich nicht oder nur kaum bemerkbar macht, so könnte man die Zustände 0111 und $1??1$ sowie 1001 und $0??1$ zusammenfassen. Da diese Zustände jeweils wegen der unterschiedlichen ausgehenden Raten nicht bisimilar sind, müssen wir hier *ein Auge zudrücken*. Wir schreiben an die ausgehenden Kanten der zusammengefassten Zustände *alle* möglichen Summen über Raten der ursprünglichen Zustände in Form eines Intervalls, als wären diese bisimilar. Die Zustände 1001 und $0??1$ können wir beispielsweise zusammenführen, indem wir die ausgehenden Kanten nach 0001 mit dem Intervall $[\lambda, \lambda + \varepsilon]$ beschreiben. Bei den Zuständen 0111 und $1??1$ müssen wir wie bei der Bisimulationsäquivalenz die Summe über die, in dieselben Makro-Zustände führenden, ausgehenden Kanten bilden. Damit erhalten wir die Markov-Kette *mit Intervallen* in Abbildung 11.

Das Auge zudrücken werden wir noch genau definieren müssen, dennoch sollte die Kernidee an diesem Beispiel deutlich geworden sein.

Bevor wir zu den formalen Definitionen kommen führen wir zunächst noch einige Schreibweisen ein, um leichter über Matrizen sprechen zu können.

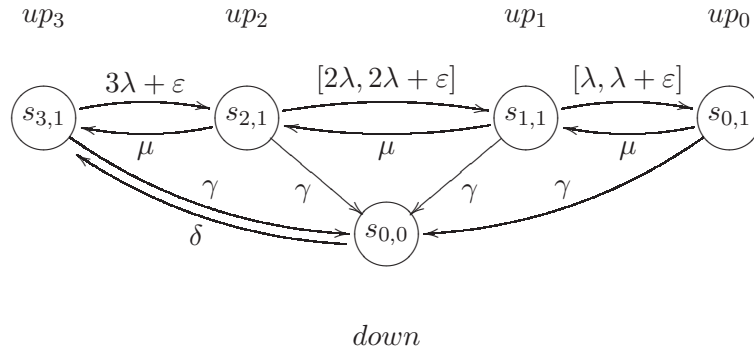


Abbildung 11: Abstraktion der TMR-Variante

Für Mengen $Y, Z \subseteq X$ und Funktionen $\mathbf{Q} : X \times X \mapsto \mathbb{R}_{\geq 0}$ sei:

$$\mathbf{Q}(Y, Z) = \sum_{y \in Y, z \in Z} \mathbf{Q}(y, z)$$

Zur Vereinfachung schreiben wir statt $\mathbf{Q}(\{s\}, Z)$ auch $\mathbf{Q}(s, Z)$. Für $\mathbf{Q} : X \times X \mapsto \mathbb{R}_{\geq 0}$ und $q : X \mapsto \mathbb{R}_{\geq 0}$ definieren wir $\mathbf{Q}(s, \cdot)$ als die zu s gehörige Zeile in \mathbf{Q} und $\mathbf{Q}(\cdot, s')$ als die zu s' gehörige Spalte:

$$\mathbf{Q}(s, \cdot) = q \text{ mit } q(s') = \mathbf{Q}(s, s') \text{ für alle } s' \in S$$

$$\mathbf{Q}(\cdot, s') = q \text{ mit } q(s) = \mathbf{Q}(s, s') \text{ für alle } s \in S$$

Für zwei Vektoren $v, v' \in X_1 \times \dots \times X_n$ und $\boxtimes \in \{<, \leq, =, \geq, >\}$ gelte $v \boxtimes v'$ gdw. $v_{x_1, \dots, x_n} \boxtimes v'_{x_1, \dots, x_n}$ für $x_1 \in X_1, \dots, x_n \in X_n$. Man beachte, dass durch die mehrfache Anwendung dieser Definition auch der komponentenweise Vergleich von Matrizen beliebiger Dimension über \boxtimes definiert ist.

Auch wenn es für zeitstetige Markov-Ketten ausreichend ist, die Raten anzugeben, definieren wir abstrakte CTMCs nicht nur über Ratenintervalle wie es im zeitdiskreten Fall möglich ist (siehe [FLW06]), sondern zusätzlich werden Wahrscheinlichkeitsintervalle und Exitratenintervalle benötigt. In den ersten beiden der drei im Anschluss vorgestellten Abstraktionen würden entweder nur Ratenintervallen oder nur Wahrscheinlichkeits- und Exitratenintervalle benötigt. Es wird sich jedoch herausstellen, dass wir für die dritte und genaueste Abstraktion alle drei Arten von Intervallen brauchen werden.

Definition 15 (Abstrakte zeitstetige Markov-Ketten). Als eine *abstrakte zeitstetige Markov-Kette* (engl. abstract continuous-time Markov chain, ACTMC) bezeichnet man ein Tupel $\mathcal{M} = (S, \mathbf{R}^l, \mathbf{P}^l, E^l, L)$, bestehend aus:

- Zustandsmenge S ,
- $\mathbf{R}^l = (\mathbf{R}^l, \mathbf{R}^u)$ wobei $\mathbf{R}^l : S \times S \mapsto \mathbb{R}_{\geq 0}$ die Matrix der unteren Ratenintervallgrenzen ist und $\mathbf{R}^u : S \times S \mapsto \mathbb{R}_{\geq 0}$ die Matrix der oberen Ratenintervallgrenzen,

- $\mathbf{P}^I = (\mathbf{P}^l, \mathbf{P}^u)$ wobei $\mathbf{P}^l : S \times S \mapsto [0, 1]$ die Matrix der unteren Wahrscheinlichkeitsintervallgrenzen ist und $\mathbf{P}^u : S \times S \mapsto [0, 1]$ die Matrix der oberen Wahrscheinlichkeitsintervallgrenzen,
- $E^I = (E^l, E^u)$ wobei $E^l : S \mapsto \mathbb{R}_{\geq 0}$ der Vektor der unteren Exitratenintervallgrenzen ist und $E^u : S \mapsto \mathbb{R}_{\geq 0}$ der Vektor der oberen Exitratenintervallgrenzen,
- Beschriftungsfunktion $L : S \times AP \mapsto \mathbb{B}^3$, die für jeden Zustand festlegt, ob eine atomare Eigenschaft $a \in AP$ dort gültig oder ungültig ist, oder ob die Gültigkeit *unbestimmt* ist.

Die Matrix der unteren Ratenintervallgrenzen muss dabei komponentenweise kleiner oder gleich der Matrix der oberen Ratenintervallgrenzen sein, also $\mathbf{R}^l \leq \mathbf{R}^u$. Ebenso muss gelten, dass $\mathbf{P}^l \leq \mathbf{P}^u$ sowie $E^l \leq E^u$.

Für die Wahrscheinlichkeitsintervalle fordern wir außerdem $\mathbf{P}^l(s_i, S) \leq 1 \leq \mathbf{P}^u(s_i, S)$ für alle $s_i \in S$, da andernfalls nicht sichergestellt ist, dass aus den Wahrscheinlichkeitsintervallen eine Verteilung wählbar ist. Desweiteren soll eine abstrakte zeitstetige Markov-Kette mindestens eine CTMC induzieren. Wir fordern daher, dass für alle Zustände $s \in S$ eine Kombination aus Raten $\mathbf{R}^l(s, s') \leq r_{s'} \leq \mathbf{R}^u(s, s')$, Exitrate $E^l(s) \leq e \leq E^u(s)$ und Wahrscheinlichkeiten $\mathbf{P}^l(s, s') \leq p_{s'} \leq \mathbf{P}^u(s, s')$ gewählt werden kann, die nicht im Widerspruch zu den Gleichungen $\sum_{s' \in S} r_{s'} = e$ und $r_{s'} = e \cdot p_{s'}$ für alle $s' \in S$ steht.

Wir bezeichnen mit $ACTMC(S, L)$ die Menge aller *abstrakten* zeitstetigen Markov-Ketten mit Zustandsmenge S und Beschriftungsfunktion L , und mit $CTMC(S, L)$ die entsprechende Menge aller zeitstetigen Markov-Ketten.

—

Wenn Intervalle die Form $[r, r]$ haben, werden wir in den weiteren Beispielen der Einfachheit halber gelegentlich auch nur r schreiben. Ebenso können die Wahrscheinlichkeits-, Raten- oder Exitratenintervalle nicht explizit notiert sein.

Da abstrakte zeitstetige Markov-Ketten als Generalisierung von konkreten zeitstetigen Markov-Ketten aufzufassen sind, geben wir nun mit *abstr* an, wie ein konkrete CTMC formal zu einer abstrakten CTMC transformiert werden kann:

$$\begin{aligned}
 \text{abstr} : CTMC(S, L) &\mapsto ACTMC(S, L) \\
 \text{abstr}(S', \mathbf{R}', L') &= (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L) \\
 \text{mit } S &= S', L = L', \\
 \mathbf{R}^l(s, s') &= \mathbf{R}^u(s, s') = \mathbf{R}'(s, s'), \\
 \mathbf{P}^l(s, s') &= \mathbf{P}^u(s, s') = \mathbf{P}'(s, s') \text{ und} \\
 E^l(s) &= E^u(s) = E'(s) \text{ für alle } s, s' \in S
 \end{aligned}$$

Legen wir nun einige Notationen fest, um den Umgang mit abstrakten Markov-Ketten zu vereinfachen. Wir schreiben für die Menge der möglichen Ratenmatrizen einer abstrakten zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$:

$$\mathbf{R} = \{\tilde{\mathbf{R}} : S \times S \mapsto \mathbb{R}_{\geq 0} \mid \mathbf{R}^l \leq \tilde{\mathbf{R}} \leq \mathbf{R}^u\}$$

Entsprechend schreiben wir für die Menge der möglichen Wahrscheinlichkeitsmatrizen von \mathcal{M} :

$$\mathbf{P} = \{\tilde{\mathbf{P}} : S \times S \mapsto [0, 1] \mid \mathbf{P}^l \leq \tilde{\mathbf{P}} \leq \mathbf{P}^u \\ \text{und } \tilde{\mathbf{P}}(s, S) \in \{0, 1\} \text{ für alle } s \in S\}$$

und für die Menge der im Intervall liegenden Exitraten:

$$E = \{\tilde{E} : S \mapsto \mathbb{R}_{\geq 0} \mid E^l \leq \tilde{E} \leq E^u\}$$

Die Menge der gültigen Ratenvektoren bzw. Ratenmatrizen, welche die Ausgangsraten bezüglich eines Zustandes $s \in S$ beschreiben, notieren wir als:

$$\text{rates}(\mathbf{R}(s, \cdot)) = \{r : S \mapsto \mathbb{R}_{\geq 0} \mid r \in \mathbf{R}(s, \cdot)\} \\ \text{rates}(\mathbf{R}) = \{r : S \times S \mapsto \mathbb{R}_{\geq 0} \mid r \in \mathbf{R}\}$$

Es ist zu beachten, dass $\text{rates}(\mathbf{R})$ und \mathbf{R} dieselben Mengen sind. Dort wo es den Lesefluß unterstützt wird jedoch auch die längere Darstellung verwendet werden.

Die Menge der, unabhängig von den Ratenintervallen, gültigen Wahrscheinlichkeitsverteilungen und -matrizen von \mathcal{M} schreiben wir als:

$$\text{distr}(\mathbf{P}(s, \cdot)) = \{p : S \mapsto [0, 1] \mid p \in \mathbf{P}(s, \cdot) \text{ und } p(S) = 1\} \\ \text{distr}(\mathbf{P}) = \{p : S \times S \mapsto [0, 1] \mid p \in \mathbf{P} \text{ und } p(s, S) = 1 \text{ für alle } s \in S\}$$

Wir werden später noch sehen, dass Raten- und Wahrscheinlichkeitsintervalle sich gegenseitig einschränken können, da sie über die totale Ausgangsraten der Zustände miteinander gekoppelt sind.

Im nächsten Unterkapitel benötigen wir die formalen Mittel, um die Genauigkeit von abstrakten Markov-Ketten miteinander vergleichen zu können. Wir werden dazu eine Halbordnung definieren, in der zwei ACTMCs in Relation stehen, wenn eine der beiden in der anderen enthalten ist.

Definition 16 (Relation \subseteq über ACTMCs). Für zwei abstrakte zeitstetige Markov-Ketten $\mathcal{M}_1 = (S_1, \mathbf{R}_1^l, \mathbf{P}_1^l, E_1^l, L_1)$ und $\mathcal{M}_2 = (S_2, \mathbf{R}_2^l, \mathbf{P}_2^l, E_2^l, L_2)$ gelte $\mathcal{M}_1 \subseteq \mathcal{M}_2$ bzw. $\mathcal{M}_2 \supseteq \mathcal{M}_1$, sprich \mathcal{M}_1 ist *feiner* als \mathcal{M}_2 oder \mathcal{M}_2 ist *gröber* als \mathcal{M}_1 , gdw.

- $S_1 = S_2$
- $L_1 = L_2$
- $\mathbf{R}_2^l \leq \mathbf{R}_1^l$ und $\mathbf{R}_1^u \leq \mathbf{R}_2^u$
- $\mathbf{P}_2^l \leq \mathbf{P}_1^l$ und $\mathbf{P}_1^u \leq \mathbf{P}_2^u$
- $E_2^l \leq E_1^l$ und $E_1^u \leq E_2^u$

Lemma 3 (Relation \subseteq über ACTMCs ist eine Halbordnung). Die Relation \subseteq bildet über der Menge aller abstrakten zeitstetigen Markov-Ketten \mathcal{D} eine Halbordnung $\mathcal{D} = \langle \mathcal{D}, \subseteq \rangle$.

Beweis. Im Folgenden seien $\mathcal{M}, \mathcal{M}_1, \mathcal{M}_2$ und \mathcal{M}_3 abstrakte zeitstetige Markov-Ketten. Um zu zeigen, dass \mathcal{D} eine Halbordnung ist, müssen die Eigenschaften der Reflexivität, Transitivität und Anti-Symmetrie nachgewiesen werden.

1. Reflexivität: $\mathcal{M} \subseteq \mathcal{M}$

Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ und $\mathcal{M}_i = (S_i, \mathbf{R}_i^I, \mathbf{P}_i^I, E_i^I, L_i)$ für $i \in \{1, 2\}$. Wir zeigen die Reflexivität, indem wir zeigen, dass $\mathcal{M} = \mathcal{M}_1 \subseteq \mathcal{M}_2 = \mathcal{M}$ gilt:

- $S = S_1 = S_2 = S$
- $L = L_1 = L_2 = L$
- $\mathbf{R}^l = \mathbf{R}_2^l \leq \mathbf{R}_1^l = \mathbf{R}^l$ und $\mathbf{R}^u = \mathbf{R}_1^u \leq \mathbf{R}_2^u = \mathbf{R}^u$
- $\mathbf{P}^l = \mathbf{P}_2^l \leq \mathbf{P}_1^l = \mathbf{P}^l$ und $\mathbf{P}^u = \mathbf{P}_1^u \leq \mathbf{P}_2^u = \mathbf{P}^u$
- $E^l = E_2^l \leq E_1^l = E^l$ und $E^u = E_1^u \leq E_2^u = E^u$

Also gilt $\mathcal{M} \subseteq \mathcal{M}$.

2. Transitivität: $\mathcal{M}_1 \subseteq \mathcal{M}_2$ und $\mathcal{M}_2 \subseteq \mathcal{M}_3 \Rightarrow \mathcal{M}_1 \subseteq \mathcal{M}_3$.

Sei $\mathcal{M}_i = (S, \mathbf{R}_i^I, \mathbf{P}_i^I, E_i^I, L)$ für $i \in \{1, 2, 3\}$. Dann gilt wegen $\mathcal{M}_1 \subseteq \mathcal{M}_2$ und $\mathcal{M}_2 \subseteq \mathcal{M}_3$, dass

- $S_1 = S_2$ und $S_2 = S_3 \Rightarrow S_1 = S_3$
- $L_1 = L_2$ und $L_2 = L_3 \Rightarrow L_1 = L_3$
- $\mathbf{R}_2^l \leq \mathbf{R}_1^l$ und $\mathbf{R}_1^u \leq \mathbf{R}_2^u$ sowie $\mathbf{R}_3^l \leq \mathbf{R}_2^l$ und $\mathbf{R}_2^u \leq \mathbf{R}_3^u$
 $\Rightarrow \mathbf{R}_3^l \leq \mathbf{R}_1^l$ und $\mathbf{R}_1^u \leq \mathbf{R}_3^u$
- $\mathbf{P}_2^l \leq \mathbf{P}_1^l$ und $\mathbf{P}_1^u \leq \mathbf{P}_2^u$ sowie $\mathbf{P}_3^l \leq \mathbf{P}_2^l$ und $\mathbf{P}_2^u \leq \mathbf{P}_3^u$
 $\Rightarrow \mathbf{P}_3^l \leq \mathbf{P}_1^l$ und $\mathbf{P}_1^u \leq \mathbf{P}_3^u$
- $E_2^l \leq E_1^l$ und $E_1^u \leq E_2^u$ sowie $E_3^l \leq E_2^l$ und $E_2^u \leq E_3^u$
 $\Rightarrow E_3^l \leq E_1^l$ und $E_1^u \leq E_3^u$

Damit gilt also auch $\mathcal{M}_1 \subseteq \mathcal{M}_3$.

3. Anti-Symmetrie: $\mathcal{M}_1 \subseteq \mathcal{M}_2$ und $\mathcal{M}_2 \subseteq \mathcal{M}_1 \Rightarrow \mathcal{M}_1 = \mathcal{M}_2$.

Sei $\mathcal{M}_i = (S, \mathbf{R}_i^I, \mathbf{P}_i^I, E_i^I, L)$ für $i \in \{1, 2\}$, dann gilt wegen $\mathcal{M}_1 \subseteq \mathcal{M}_2$ und $\mathcal{M}_2 \subseteq \mathcal{M}_1$, dass

- $S_1 = S_2$
- $L_1 = L_2$
- $\mathbf{R}_2^l \leq \mathbf{R}_1^l$ und $\mathbf{R}_1^u \leq \mathbf{R}_2^u$ sowie $\mathbf{R}_1^l \leq \mathbf{R}_2^l$ und $\mathbf{R}_2^u \leq \mathbf{R}_1^u$
 $\Rightarrow \mathbf{R}_1^l = \mathbf{R}_2^l$ und $\mathbf{R}_1^u = \mathbf{R}_2^u$

- $\mathbf{P}_2^l \leq \mathbf{P}_1^l$ und $\mathbf{P}_1^u \leq \mathbf{P}_2^u$ sowie $\mathbf{P}_1^l \leq \mathbf{P}_2^l$ und $\mathbf{P}_2^u \leq \mathbf{P}_1^u$
 $\Rightarrow \mathbf{P}_1^l = \mathbf{P}_2^l$ und $\mathbf{P}_1^u = \mathbf{P}_2^u$
- $E_2^l \leq E_1^l$ und $E_1^u \leq E_2^u$ sowie $E_1^l \leq E_2^l$ und $E_2^u \leq E_1^u$
 $\Rightarrow E_1^l = E_2^l$ und $E_1^u = E_2^u$

Daraus folgt also, dass $\mathcal{M}_1 = \mathcal{M}_2$.

□

Definition 17 (Monotonie und Beschränktheit für Folgen). Eine Folge von abstrakten zeitstetigen Markov-Ketten $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots)$ heißt *monoton fallend*, falls gilt:

$$\mathcal{M}_i \supseteq \mathcal{M}_{i+1} \text{ für alle } i \in \mathbb{N}$$

Eine Folge von abstrakten zeitstetigen Markov-Ketten $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots)$ heißt *nach unten beschränkt*, falls eine Markov-Kette \mathcal{M} existiert, für die gilt:

$$\mathcal{M}_i \subseteq \mathcal{M} \text{ für alle } i \in \mathbb{N}$$

Wenn im Folgenden von *monotonen* oder *beschränkten* Folgen von ACTMCs gesprochen wird, sind immer *monoton fallende* und *nach unten beschränkte* gemeint.

3.2 Scheduler und cut-Funktion

Ähnlich wie in Markov-Entscheidungsprozessen haben wir bei den abstrakten Markov-Ketten in jedem Zustand mögliche Nichtdeterminismen. In den Entscheidungsprozessen besteht der Nichtdeterminismus darin, dass in jedem Zustand aus einer endlichen Menge von Aktionen gewählt werden kann. Als eine Möglichkeit mit dieser Wahlfreiheit umzugehen, haben wir in Abschnitt 2.6 Scheduler kennen gelernt.

Auch wenn in abstrakten Markov-Ketten im Gegensatz zu Entscheidungsprozessen aus einer überabzählbaren Menge von möglichen Kombinationen von Raten, Wahrscheinlichkeiten und Exitraten gewählt werden kann, können wir den Nichtdeterminismus dennoch mithilfe von Schemulern auf einfache Weise behandeln. Wir geben im folgenden die Definitionen von *history dependent* Schemulern und *extremen* Schemulern im Kontext der abstrakten Markov-Ketten an.

Definition 18 (HD-Scheduler). Ein HD-Scheduler (*history dependent*) bezüglich einer abstrakten zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ ist eine Funktion

$$\begin{aligned} \eta : S^+ &\longmapsto \{(\bar{\mathbf{R}}, \bar{\mathbf{P}}, \bar{E}) \in \text{rates}(\mathbf{R}) \times \text{distr}(\mathbf{P}) \times E \\ &\quad | \bar{\mathbf{R}}(s, s') = \bar{\mathbf{P}}(s, s') \cdot \bar{E}(s) \text{ für alle } s, s' \in S\} \end{aligned} \quad (6)$$

die für alle möglichen Folgen von Zuständen s_0, \dots, s zueinander passende Ratenvektoren aus $\text{rates}(\mathbf{R}(s, \cdot))$, Verteilungen aus $\text{distr}(\mathbf{P}(s, \cdot))$ und Exitraten aus E bestimmt. Die Menge aller HD-Scheduler bezüglich einer Markov-Kette \mathcal{M} bezeichnen wir als $\mathcal{S}(\mathcal{M})$.

–

Es ist zu bemerken, dass die *history dependent* Scheduler nur von den zeitabstrakten Pfaden abhängig sind, nicht jedoch von den Verweilzeiten in den einzelnen Zuständen eines Pfades.

Definition 19 (Induzierte zeitstetige Markov-Ketten). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette und $\eta \in \mathcal{S}(\mathcal{M})$ ein *history dependent* Scheduler. Die von η induzierte zeitstetige Markov-Kette ist gegeben durch $\mathcal{M}_\eta = (S_\eta, \mathbf{R}_\eta^I, \mathbf{P}_\eta^I, E_\eta^I, L_\eta)$ mit:

- $S_\eta = S^+$ ist die (unendliche) Zustandsmenge,
- $L_\eta(s_0, \dots, s_n) = L(s_n)$ ist eine Abbildung der Beschriftungsfunktion auf die ursprüngliche Beschriftungsfunktion,
- $(\mathbf{R}_\eta^I, \mathbf{P}_\eta^I, E_\eta^I) = \eta(s_0, \dots, s_n)$ für den Zustand der induzierten Markov-Kette $(s_0, \dots, s_n) \in S^+$ sind die Raten und Wahrscheinlichkeiten, die vom Scheduler im gegebenen Zustand gewählt werden.

—

Eine induzierte zeitstetige Markov-Kette ist zwar in der Regel unendlich groß, hat aber den Vorteil deterministisch zu sein. Induzierte Markov-Ketten werden unter anderem für die Definition von Wahrscheinlichkeitsmaßen benötigt.

Eine Sorte von Schemulern, die später noch von besonderem Interesse sein wird, sind die *extremen* Scheduler. Anstatt eine beliebige passende Wahl für Verteilung, Ratenvektor und Exitrate treffen zu können, werden extreme Scheduler darauf beschränkt, nur minimale und maximale Werte eines Intervalls zu wählen. Da der Scheduler für jeden Nachfolgezustand lediglich aus zwei möglichen Optionen wählen kann und es nur endlich viele Nachfolgezustände gibt, kann er im Gegensatz zu HD-Schemulern nur aus einer *endlichen* Menge von möglichen Aktionen wählen¹⁴. Aus diesem Grund werden wir später bei der Behandlung des Erreichbarkeitsproblems extreme Scheduler statt HD-Schemuler verwenden. Bevor wir die Definition formal angeben können, müssen wir uns jedoch zunächst noch der Frage widmen, wie verhindert werden kann, dass ein Scheduler *inkonsistente* Wahrscheinlichkeiten, Raten und Exitraten wählt¹⁵.

Insbesondere ist es nicht immer möglich für Verteilung, Ratenvektor und Exitrate eines Zustandes lediglich Randwerte der Intervalle zu wählen. Beispielsweise könnte die Wahl maximaler Raten zu einer Exitrate führen, die außerhalb des Exitratenintervalls liegt. Daher muss in das Auswahlverfahren des Schemulers eine *cut*-Funktion integriert werden, die vor der Wahl jeder einzelnen Rate diejenigen Werte aus den

¹⁴Durch die Anwendung des *cut* nach jeder Wahl eines Wertes kann das Ergebnis von der Reihenfolge abhängen, in der die Nachfolgezustände abgearbeitet werden. Da es für endliche Zustandsmengen nur endlich viele verschiedene Reihenfolgen gibt, bleibt die Anzahl der möglichen extremen Scheduler auch bei Berücksichtigung dieses Umstands noch endlich.

¹⁵Die Wahl eines Schemulers bezeichnen wir dann als *inkonsistent*, wenn das Resultat keiner zeitstetigen Markov-Kette entspricht, etwa wenn eine Zeilensumme weder den Wert 0 noch den Wert 1 hat.

Intervallen entfernt, die nicht mehr zur Vervollständigung zu einer CTMC beitragen können. Nach Anwendung des *cut* kann der Scheduler aus den übrigen Werten wieder einen extremen Wert wählen, sodass man zum Schluss eine gültige Wahl für die Wahrscheinlichkeiten und Raten erhält.

Um besser kenntlich zu machen, welche Intervalle Ratenintervalle sind und welche Wahrscheinlichkeitsintervalle, werden wir in den Abbildungen der folgenden Beispiele statt $[p^l, p^u]$ für letztere $|p^l, p^u|$ schreiben.

Beispiel 8 (*cut*-Funktion 1). Sei die abstrakte zeitstetige Markov-Kette \mathcal{M} mit Zustandsmenge $\{\bar{s}_0, \bar{s}_1, \bar{s}_2\}$ aus der Abbildung 12 (links oben) gegeben. In diesem Beispiel gibt es anfangs keine Werte, die nicht zu einer CTMC vervollständigt werden könnten, also lassen wir den Scheduler gleich für den Übergang von \bar{s}_0 nach \bar{s}_1 die Wahrscheinlichkeit $\frac{1}{3}$ wählen. Wir erhalten damit die abstrakte Markov-Kette aus der Abbildung (rechts oben).

Wir wissen, dass die Wahrscheinlichkeit nach \bar{s}_2 zu gelangen genau $\frac{2}{3}$ sein muss, damit sich eine Verteilung für \bar{s}_0 ergibt. Die Exitrate ist folglich 3, da für einen größeren Wert eine größere Rate nach \bar{s}_2 erforderlich wäre als 2, oder eine größere Wahrscheinlichkeit als $\frac{2}{3}$. Die Abbildung links unten zeigt den Stand bis hierher.

Da für gültige Werte die Gleichung $\mathbf{R}(s, s') = E(s) \cdot \mathbf{P}(s, s')$ erfüllt sein muss, ergeben sich mit diesen Zahlen dann auch die Raten wie in der Abbildung rechts unten.

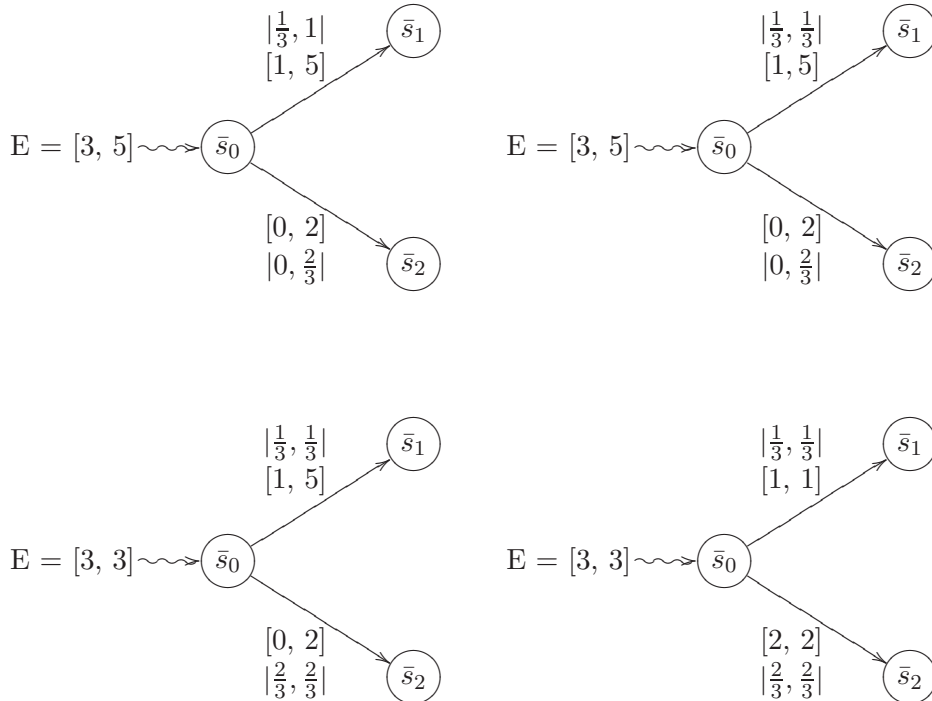


Abbildung 12: Beispiel 1 zur *cut*-Funktion

Untersuchen wir jetzt aber zunächst im Allgemeinen, wie man Raten, Wahrscheinlichkeiten und Exitraten, die nicht zu einer CTMC vervollständigt werden können, aus den Intervallen entfernt. Wir haben drei Restriktionen bezüglich der Werte in \mathbf{R} , \mathbf{P} und E :

- a) Erstens muss die Summe über alle gewählten Raten in dem Intervall der Exitraten liegen und umgekehrt müssen sich die Exitraten im Intervall als Summe über einzelne Raten ergeben können. Ist dies für bestimmte Werte nicht der Fall, so müssen diese aus den Intervallen entfernt werden.
- b) Weiterhin müssen wir Wahrscheinlichkeiten entfernen, die sich zusammen mit anderen wählbaren Wahrscheinlichkeiten zu keiner Verteilung vervollständigen lassen. Dabei ist zu beachten, dass in absorbierenden Zuständen die Summe über alle Wahrscheinlichkeiten, den Zustand zu verlassen, nicht gleich Eins ist. Dies ist also als Sonderfall zu behandeln.
- c) Zuletzt muss noch sichergestellt sein, dass keine Raten r , Wahrscheinlichkeiten p oder Exitraten e in der Markov-Kette enthalten sind, für welche jeweils die Gleichung $r = p \cdot e$ nicht durch Wahl der übrigen Werte erfüllbar ist. Ist beispielsweise eine Wahrscheinlichkeit p im Intervall enthalten, für die keine Rate r und keine Exitrate e wählbar ist, sodass die obige Gleichung gilt, so ist diese aus dem Intervall zu entfernen.

Formulieren wir diese drei Restriktionen nun als Funktionen auf Markov-Ketten. Dazu werden wir zunächst sechs Teilfunktionen definieren, die im weiteren Verlauf zum vollständigen *cut* zusammengefügt werden. Um die Lesbarkeit zu verbessern, sind im Folgenden die Teile Grau dargestellt, die sicherstellen, dass die Intervallgrenzen nicht erweitert werden, und die oberen Grenzen nicht kleinere Werte erhalten können als die unteren Grenzen.

1. Schneide alle Raten ab, die nicht so vervollständigt werden können, dass die Exitrate im entsprechenden Intervall ist (zu a):

$cut_{R_1^\gamma}^M : S \times S \mapsto \mathbb{R}_{\geq 0}$ für $\gamma \in \{u, l\}$ mit:

$$cut_{R_1^u}^M(s, s') = \max\{\mathbf{R}^l(s, s'), \min\{\mathbf{R}^u(s, s'), E^u(s) - \sum_{v \neq s'} \mathbf{R}^l(s, v)\}\}$$

$$cut_{R_1^l}^M(s, s') = \min\{\mathbf{R}^u(s, s'), \max\{\mathbf{R}^l(s, s'), E^l(s) - \sum_{v \neq s'} \mathbf{R}^u(s, v)\}\}$$

2. Schneide die Exitraten ab, die nicht aus einem Ratenvektor resultieren können (zu a):

$cut_{E_1^\alpha}^M : S \mapsto \mathbb{R}_{\geq 0}$ für $\alpha \in \{u, l\}$ mit:

$$cut_{E_1^u}^M(s) = \max\{E^l(s), \min\{E^u(s), \mathbf{R}^u(s, S)\}\}$$

$$cut_{E_1^l}^M(s) = \min\{E^u(s), \max\{E^l(s), \mathbf{R}^l(s, S)\}\}$$

3. Schneide alle Wahrscheinlichkeiten ab, die nicht so vervollständigt werden können, sodass man eine Verteilung erhält (zu b):

$cut_{P_1^\gamma}^M : S \times S \mapsto [0, 1]$ für $\gamma \in \{u, l\}$ mit:

$$cut_{P_1^u}^M(s, s') = \max\{\mathbf{P}^l(s, s'), \min\{\mathbf{P}^u(s, s'), 1 - \sum_{v \neq s'} \mathbf{P}^l(s, v)\}\}$$

für $E^u(s) > 0$ und $cut_{P_1^u}^M(s, s') = 0$ sonst,

$$cut_{P_1^l}^M(s, s') = \min\{\mathbf{P}^u(s, s'), \max\{\mathbf{P}^l(s, s'), 1 - \sum_{v \neq s'} \mathbf{P}^u(s, v)\}\}$$

für $E^l(s) > 0$ und $cut_{P_1^l}^M(s, s') = 0$ sonst.

4. Schneide alle Raten ab, die sich nicht als Produkt von gültigen Wahrscheinlichkeiten und Exitraten ergeben (zu c):

$cut_{R_2^\gamma}^M : S \times S \mapsto \mathbb{R}_{\geq 0}$ für $\gamma \in \{u, l\}$ mit:

$$cut_{R_2^u}^M(s, s') = \max\{\mathbf{R}^l(s, s'), \min\{\mathbf{R}^u(s, s'), E^u(s) \cdot \mathbf{P}^u(s, s')\}\}$$

$$cut_{R_2^l}^M(s, s') = \min\{\mathbf{R}^u(s, s'), \max\{\mathbf{R}^l(s, s'), E^l(s) \cdot \mathbf{P}^l(s, s')\}\}$$

5. Schneide alle Wahrscheinlichkeiten ab, die als Produkt mit einer Exitrate keine möglichen Raten ergeben (zu c):

$cut_{P_2^\gamma}^M : S \times S \mapsto [0, 1]$ für $\gamma \in \{u, l\}$ mit:

$$cut_{P_2^u}^M(s, s') = \max\{\mathbf{P}^l(s, s'), \min\{\mathbf{P}^u(s, s'), \mathbf{R}^u(s, s')/E^l(s)\}\}$$

für $E^u(s) > 0$ und $cut_{P_2^u}^M(s, s') = 0$ sonst,

$$cut_{P_2^l}^M(s, s') = \min\{\mathbf{P}^u(s, s'), \max\{\mathbf{P}^l(s, s'), \mathbf{R}^l(s, s')/E^u(s)\}\}$$

für $E^l(s) > 0$ und $cut_{P_2^l}^M(s, s') = 0$ sonst,

6. Schneide alle Exitraten ab, die als Produkt mit einer Wahrscheinlichkeit keine mögliche Rate ergeben (zu c):

$cut_{E_2^\gamma}^M : S \times S \mapsto [0, 1]$ für $\gamma \in \{u, l\}$ mit:

$$cut_{E_2^u}^M(s, s') = \max\{E^l(s), \min\{E^u(s), \mathbf{R}^u(s, s')/\mathbf{P}^l(s, s')\}\}$$

$$cut_{E_2^l}^M(s, s') = \min\{E^u(s), \max\{E^l(s), \mathbf{R}^l(s, s')/\mathbf{P}^u(s, s')\}\}$$

$$\text{wobei } x/y = \begin{cases} \frac{x}{y} & \text{falls } y \neq 0 \\ 0 & \text{falls } x = 0 \\ \infty & \text{falls } x \neq 0 \text{ und } y = 0 \end{cases}$$

Vergleicht man die *cut*-Funktionen in Punkten 1 und 3, so erkennt man leicht die Verwandtschaft der beiden. Im ersten Punkt betrachtet man Raten, wo im dritten Punkt die zugehörigen Wahrscheinlichkeiten stehen. Es drängt sich die Frage auf, warum kein *cut* bezüglich Wahrscheinlichkeiten benötigt wird, der dem Punkt 2 entspricht. Der Grund dafür ist, dass die Summe über die Wahrscheinlichkeiten einer Verteilung immer genau 1 sein muss. Im Gegensatz dazu gibt es für die Exitraten im

Allgemeinen, abhängig von den Ratenintervallen, mehrere mögliche Werte. Der *cut* kann also grundsätzlich das Exitratenintervall verkleinern. Bei Wahrscheinlichkeiten ist dies nicht möglich, da das *Exitwahrscheinlichkeitsintervall*¹⁶ immer $[1, 1]$ sein muss.

Wenn die ACTMC \mathcal{M} aus dem Kontext hervorgeht, werden wir sie im Folgenden als Parameter der *cut*-Funktionen weglassen.

Beispiel 9 (*cut*-Funktion 2). Im ersten Beispiel zum *cut* wurde $cut_{P_1^l}$, $cut_{E_2^u}$, $cut_{R_2^u}$ und $cut_{R_2^l}$ hintereinander ausgeführt. Legt man anfangs die Exitrate von \bar{s}_0 auf 5 fest, statt eine Wahrscheinlichkeit wählen, so benötigt man erst cut_{R_1} und anschließend cut_{P_2} um die Intervallgrenzen komplett zu beschneiden.

Für die Rate nach \bar{s}_1 können wir wegen dem Intervall $[0, 2]$ nur noch Werte zwischen 3 und 5 wählen. Außerdem müssen die Wahrscheinlichkeitsintervalle an die neuen Raten angepasst werden.

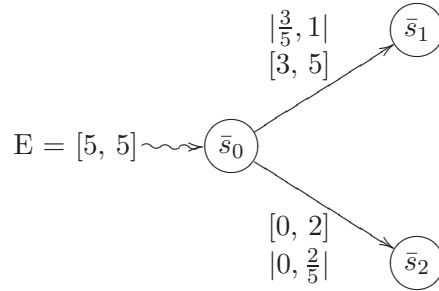


Abbildung 13: Beispiel 2 zur *cut*-Funktion

Beispiel 10 (*cut*-Funktion 3). Nun konstruieren wir noch ein Beispiel, in dem cut_{E_1} benötigt wird, um eine unzulässige Exitrate zu eliminieren. Dazu erhöhen wir in der Markov-Kette der Abbildung 13 die obere Grenze des Exitratenintervall auf 8, sodass wir die linke Markov-Kette aus Abbildung 14 erhalten.

Wendet man hier cut_{E_2} an, so verändert sich das Exitratenintervall nicht, denn:

$$\mathbf{R}^u(\bar{s}_0, \bar{s}_1) / \mathbf{P}^l(\bar{s}_0, \bar{s}_1) = 5 / \frac{3}{5} = \frac{25}{3} = 8, \bar{3} > 8 = E^u(\bar{s}_0)$$

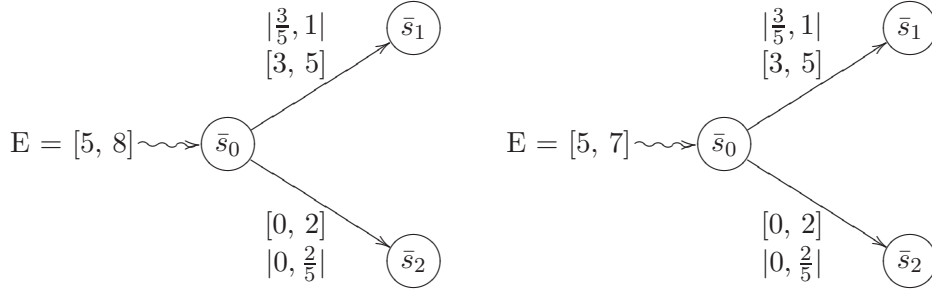
$$\mathbf{R}^l(\bar{s}_0, \bar{s}_1) / \mathbf{P}^u(\bar{s}_0, \bar{s}_1) = 3 / 1 = 3 < 5 = E^l(\bar{s}_0)$$

$$\mathbf{R}^u(\bar{s}_0, \bar{s}_2) / \mathbf{P}^l(\bar{s}_0, \bar{s}_2) = 2 / 0 = \infty > 8 = E^u(\bar{s}_0)$$

$$\mathbf{R}^l(\bar{s}_0, \bar{s}_2) / \mathbf{P}^u(\bar{s}_0, \bar{s}_2) = 0 / \frac{2}{5} = 0 < 5 = E^l(\bar{s}_0)$$

Die vorherige Ausführung der cut_R - und cut_P -Funktionen bewirkt ebenfalls keine Veränderung. Lediglich die Anwendung von cut_{E_1} ergibt einen neuen kleineren Wert für die obere Grenze des Intervalls, da $\mathbf{R}^u(\bar{s}_0, \bar{s}_1) + \mathbf{R}^u(\bar{s}_0, \bar{s}_2) = 5 + 2 = 7 \in [5, 8]$. Man erhält damit die ACTMC in Abbildung 14 rechts.

¹⁶ Aufgrund der Trivialität wurde in der Definition der abstrakten zeitstetigen Markov-Kette auf die explizite Angabe eines *Exitwahrscheinlichkeitsintervall* verzichtet. Es wäre nicht sehr hilfreich aber durchaus legitim, würde man es hinzunehmen.

Abbildung 14: Beispiel 3 zur *cut*-Funktion

Wir sehen an den drei Beispielen, dass die sechs aufgeführten *cut*-Funktionen alle benötigt werden. Um eine beliebige abstrakte Markov-Kette zu beschneiden, können wir solange die verschiedenen *cuts* anwenden, bis keine Verkleinerungen der Intervallgrenzen mehr möglich ist. Die Reihenfolge spielt dabei eigentlich keine Rolle, solange sie in einer fairen¹⁷ Folge ausgeführt werden. Wir werden die Reihenfolge dennoch willkürlich festlegen, um die Fairness zu erzwingen.

Der vollständige *cut* wird im Folgenden als Fixpunktgleichung definiert und im Anschluss wird die Konvergenz der Folge $(cut^0(\mathcal{M}), cut^1(\mathcal{M}), cut^2(\mathcal{M}), \dots)$ und damit ihre Berechenbarkeit des *cut* gezeigt. Üblicherweise wird für Fixpunktgleichungen zwar die Existenz eines kleinsten Fixpunktes über die Monotonie der Funktion auf einem vollständigen Verband gezeigt. Der hier gewählte Ansatz ist jedoch intuitiver und für unsere Zwecke ebenso geeignet.

Wir schreiben im Folgenden für die Substitution von Komponente \mathbf{X} durch die Komponente \mathbf{X}' in der Markov-Kette \mathcal{M} :

$$\mathcal{M}[\mathbf{X} \mapsto \mathbf{X}'].$$

Definition 20 (*cut*). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine endliche zeitstetige abstrakte Markov-Kette mit $S = \{s_1, \dots, s_n\}$. Dann definieren wir:

$$cut : ACTMC(S, L) \mapsto ACTMC(S, L)$$

$$cut(\mathcal{M}) = \lim_{n \rightarrow \infty} cut_{RPE}^n(\mathcal{M})$$

$$\text{mit } cut_{RPE} : ACTMC(S, L) \mapsto ACTMC(S, L)$$

$$cut_{RPE}(\mathcal{M}) = cut_R(cut_P(cut_E(\mathcal{M})))$$

$$cut_X : ACTMC(S, L) \mapsto ACTMC(S, L) \text{ für } X \in \{R, P, E\}$$

$$cut_R(\mathcal{M}) = (S, cut_{R^l}(\mathcal{M}[\mathbf{R} \mapsto cut_{R^u}(\mathcal{M})]), \mathbf{P}^I, E^I, L)$$

$$cut_P(\mathcal{M}) = (S, \mathbf{R}^I, cut_{P^l}(\mathcal{M}[\mathbf{P} \mapsto cut_{P^u}(\mathcal{M})]), E^I, L)$$

$$cut_E(\mathcal{M}) = (S, \mathbf{R}^I, \mathbf{P}^I, cut_{E^l}(\mathcal{M}[E \mapsto cut_{E^u}(\mathcal{M})]), L)$$

$$cut_X : ACTMC(S, L) \mapsto (\mathbb{R}_{\geq 0}^{|S| \times |S|}) \times (\mathbb{R}_{\geq 0}^{|S| \times |S|}) \text{ für } X \in \{R^l, R^u\}$$

$$cut_{R^l}(\mathcal{M}) = (cut'_{R^l}(\mathcal{M}), \mathbf{R}^u)$$

$$cut_{R^u}(\mathcal{M}) = (\mathbf{R}^l, cut'_{R^u}(\mathcal{M}))$$

¹⁷Eine unendliche Folge heißt fair, wenn jedes Element unendlich oft vorkommt. Hier sind mit den Elementen die sechs *cut*-Funktionen gemeint.

$$cut_X : ACTMC(S, L) \mapsto ([0, 1]^{|S| \times |S|}) \times ([0, 1]^{|S| \times |S|}) \text{ für } X \in \{P^l, P^u\}$$

$$cut_{P^l}(\mathcal{M}) = (cut'_{P^l}(\mathcal{M}), \mathbf{P}^u)$$

$$cut_{P^u}(\mathcal{M}) = (\mathbf{P}^l, cut'_{P^u}(\mathcal{M}))$$

$$cut_X : ACTMC(S, L) \mapsto (\mathbb{R}_{\geq 0}^{|S| \times |S|}) \times (\mathbb{R}_{\geq 0}^{|S| \times |S|}) \text{ für } X \in \{E^l, E^u\}$$

$$cut_{E^l}(\mathcal{M}) = (cut'_{E^l}(\mathcal{M}), E^u)$$

$$cut_{E^u}(\mathcal{M}) = (E^l, cut'_{E^u}(\mathcal{M}))$$

$$cut'_X : ACTMC(S, L) \mapsto \mathbb{R}_{\geq 0}^{|S| \times |S|} \text{ für } X \in \{R^l, R^u\}$$

$$cut'_X(\mathcal{M}) = \begin{pmatrix} cut''_X^{\mathcal{M}}(s_1, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_1, s_n) \\ \vdots & \ddots & \vdots \\ cut''_X^{\mathcal{M}}(s_n, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_n, s_n) \end{pmatrix}$$

$$cut'_X : ACTMC(S, L) \mapsto [0, 1]^{|S| \times |S|} \text{ für } X \in \{P^l, P^u\}$$

$$cut'_X(\mathcal{M}) = \begin{pmatrix} cut''_X^{\mathcal{M}}(s_1, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_1, s_n) \\ \vdots & \ddots & \vdots \\ cut''_X^{\mathcal{M}}(s_n, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_n, s_n) \end{pmatrix}$$

$$cut'_X : ACTMC(S, L) \mapsto \mathbb{R}_{\geq 0}^{|S| \times |S|} \text{ für } X \in \{E^l, E^u\}$$

$$cut'_X(\mathcal{M}) = \begin{pmatrix} cut''_X^{\mathcal{M}}(s_1, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_1, s_n) \\ \vdots & \ddots & \vdots \\ cut''_X^{\mathcal{M}}(s_n, s_1) & \cdots & cut''_X^{\mathcal{M}}(s_n, s_n) \end{pmatrix}$$

$$cut''_X : ACTMC(S, L) \times S \times S \mapsto \mathbb{R}_{\geq 0} \text{ für } X \in \{R^l, R^u\}$$

$$cut''_{R^l}(\mathcal{M}, s, s') = \max\{cut_{R_1^l}^{\mathcal{M}}(s, s'), cut_{R_2^l}^{\mathcal{M}}(s, s')\}$$

$$cut''_{R^u}(\mathcal{M}, s, s') = \min\{cut_{R_1^u}^{\mathcal{M}}(s, s'), cut_{R_2^u}^{\mathcal{M}}(s, s')\}$$

$$cut''_X : ACTMC(S, L) \times S \times S \mapsto [0, 1] \text{ für } X \in \{P^l, P^u\}$$

$$cut''_{P^l}(\mathcal{M}, s, s') = \max\{cut_{P_1^l}^{\mathcal{M}}(s, s'), cut_{P_2^l}^{\mathcal{M}}(s, s')\}$$

$$cut''_{P^u}(\mathcal{M}, s, s') = \min\{cut_{P_1^u}^{\mathcal{M}}(s, s'), cut_{P_2^u}^{\mathcal{M}}(s, s')\}$$

$$cut''_X : ACTMC(S, L) \times S \times S \mapsto \mathbb{R}_{\geq 0} \text{ für } X \in \{E^l, E^u\}$$

$$cut''_{E^l}(\mathcal{M}, s, s') = \max\{cut_{E_1^l}^{\mathcal{M}}(s, s'), cut_{E_2^l}^{\mathcal{M}}(s, s')\}$$

$$cut''_{E^u}(\mathcal{M}, s, s') = \min\{cut_{E_1^u}^{\mathcal{M}}(s, s'), cut_{E_2^u}^{\mathcal{M}}(s, s')\}$$

Eine Markov-Kette mit $\mathcal{M} = cut(\mathcal{M})$ nennen wir auch *bereinigt*, da sie keine Werte in den Intervallen enthält, die nicht zu einer CTMC vervollständigt werden könnten.

—

Wir zeigen nun, dass der Grenzwert $\lim_{n \rightarrow \infty} cut_{RPE}^n(\mathcal{M})$ existiert, also dass der cut die Intervalle einer abstrakten zeitstetigen Markov-Kette immer nur verkleinern kann. Da Punktintervalle nicht weiter verkleinert werden können, muss also die Folge $(cut_{RPE}^0(\mathcal{M}), cut_{RPE}^1(\mathcal{M}), cut_{RPE}^2(\mathcal{M}), \dots)$ letztendlich konvergieren.

Satz 1 (Konvergenz der cut -Funktion). Sei $\mathcal{M} = (S, \mathbf{R}^l, \mathbf{P}^l, E^l, L)$ eine zeitstetige abstrakte Markov-Kette, dann konvergiert die Folge:

$$(cut_{RPE}^0(\mathcal{M}), cut_{RPE}^1(\mathcal{M}), cut_{RPE}^2(\mathcal{M}), \dots)$$

Beweis. Wir werden zunächst die Monotonie der cut -Funktion nachweisen, indem wir bezüglich der Definition *bottom-up* vorgehen. Anschließend werden wir zeigen, dass die Folge beschränkt ist, woraus wir auf die Konvergenz schließen können.

Um die Beweisführung zu vereinfachen, führen wir mit $cut_{R^\gamma}^{\mathcal{M}, s, s'}(\mathbf{R}^\gamma(s, s'))$ eine neue Schreibweise für $cut_{R^\gamma}^{\mathcal{M}}(s, s')$ und $\gamma \in \{l, u\}$ ein (entsprechend für \mathbf{P} und E). Dadurch lässt sich deutlicher zeigen, dass nach der Anwendung des cut die unteren Intervallgrenzen mindestens so groß sind und die oberen höchstens so groß wie vorher:

1.) $cut_{X^l}(\mathcal{M}) \geq (\mathbf{X}^l, \mathbf{X}^u)$ folgt aus:

Für cut_{R^l} gilt:

$$\begin{aligned} cut_{R_1^l}^{\mathcal{M}, s, s'}(\mathbf{R}^l(s, s')) &= \min\{\mathbf{R}^u(s, s'), \max\{\mathbf{R}^l(s, s'), E^l(s) - \sum_{v \neq s'} \mathbf{R}^u(s, v)\}\} \\ &\geq \mathbf{R}^l(s, s') \\ cut_{R_2^l}^{\mathcal{M}, s, s'}(\mathbf{R}^l(s, s')) &= \min\{\mathbf{R}^u(s, s'), \max\{\mathbf{R}^l(s, s'), E^u(s) \cdot \mathbf{P}^u(s, s')\}\} \\ &\geq \mathbf{R}^l(s, s') \\ \Rightarrow cut_{R^l}''(\mathcal{M}, s, s') &\geq \mathbf{R}^l(s, s') \text{ für alle } s, s' \in S \\ \Rightarrow cut_{R^l}'(\mathcal{M}) &\geq \mathbf{R}^l \\ \Rightarrow cut_{R^l}(\mathcal{M}) &\geq (\mathbf{R}^l, \mathbf{R}^u) \end{aligned}$$

Für cut_{P_l} unterscheiden wir zwischen den Fällen $E^l(s) = 0$ und $E^l(s) > 0$. Im ersten Fall kann der Zustand absorbierend sein, das heißt, dass für alle $s' \in S$ die Wahrscheinlichkeit $\mathbf{P}^l(s, s') = 0$ sein muss. Die beiden Funktionen $cut_{P_1^l}$ und $cut_{P_2^l}$ liefern in diesem Sonderfall ebenfalls 0 als Ergebnis. Es gilt dann $cut_{P^l}''(\mathcal{M}, s, s') = \mathbf{P}^l(s, s')$.

Für den Fall $E^l(s) > 0$ gilt:

$$\begin{aligned} cut_{P_1^l}^{\mathcal{M}, s, s'}(\mathbf{P}^l(s, s')) &= \min\{\mathbf{P}^u(s, s'), \max\{\mathbf{P}^l(s, s'), 1 - \sum_{v \neq s'} \mathbf{P}^u(s, v)\}\} \\ &\geq \mathbf{P}^l(s, s') \\ cut_{P_2^l}^{\mathcal{M}, s, s'}(\mathbf{P}^l(s, s')) &= \min\{\mathbf{P}^u(s, s'), \max\{\mathbf{P}^l(s, s'), \mathbf{R}^l(s, s')/E^u(s)\}\} \\ &\geq \mathbf{P}^l(s, s') \end{aligned}$$

$$\Rightarrow \text{cut}_{P^l}''(\mathcal{M}, s, s') \geq \mathbf{P}^l(s, s') \text{ für alle } s, s' \in S$$

$$\Rightarrow \text{cut}_{P^l}'(\mathcal{M}) \geq \mathbf{P}^l$$

$$\Rightarrow \text{cut}_{P^l}(\mathcal{M}) \geq (\mathbf{P}^l, \mathbf{P}^u)$$

Für cut_{E^l} gilt:

$$\text{cut}_{E_1^l}^{\mathcal{M}, s}(E^l(s)) = \min\{E^u(s), \max\{E^l(s), \mathbf{R}^l(s, S)\}\} \geq E^l$$

$$\text{cut}_{E_2^l}^{\mathcal{M}, s, s'}(E^l(s)) = \min\{E^u(s), \max\{E^l(s), \mathbf{R}^l(s, s')/\mathbf{P}^u(s, s')\}\} \geq E^l$$

$$\Rightarrow \text{cut}_{E^l}''(\mathcal{M}, s, s') \geq E^l(s) \text{ für alle } s, s' \in S$$

$$\Rightarrow \text{cut}_{E^l}'(\mathcal{M}) \geq E^l$$

$$\Rightarrow \text{cut}_{E^l}(\mathcal{M}) \geq (E^l, E^u)$$

2.) $\text{cut}_{X^u}(\mathcal{M}) \leq (\mathbf{X}^l, \mathbf{X}^u)$ folgt aus:

Für cut_{R^u} gilt:

$$\begin{aligned} \text{cut}_{R_1^u}^{\mathcal{M}, s, s'}(\mathbf{R}^u(s, s')) &= \max\{\mathbf{R}^l(s, s'), \min\{\mathbf{R}^u(s, s'), E^u(s) - \sum_{v \neq s'} \mathbf{R}^l(s, v)\}\} \\ &\leq \mathbf{R}^u(s, s') \end{aligned}$$

$$\begin{aligned} \text{cut}_{R_2^u}^{\mathcal{M}, s, s'}(\mathbf{R}^u(s, s')) &= \max\{\mathbf{R}^l(s, s'), \min\{\mathbf{R}^u(s, s'), E^u(s) \cdot \mathbf{P}^u(s, s')\}\} \\ &\leq \mathbf{R}^u(s, s') \end{aligned}$$

$$\Rightarrow \text{cut}_{R^u}''(\mathcal{M}, s, s') \leq \mathbf{R}^u(s, s') \text{ für alle } s, s' \in S$$

$$\Rightarrow \text{cut}_{R^u}'(\mathcal{M}) \leq \mathbf{R}^u$$

$$\Rightarrow \text{cut}_{R^u}(\mathcal{M}) \leq (\mathbf{R}^l, \mathbf{R}^u)$$

Für cut_{P^u} unterscheiden wir wieder zwischen den Fällen $E^u(s) = 0$ und $E^u(s) > 0$. Im ersten Fall muss der Zustand absorbierend sein, das heißt, dass für alle $s' \in S$ die Wahrscheinlichkeit $\mathbf{P}^u(s, s') = 0$ sein muss. Die beiden Funktionen $\text{cut}_{P_1^u}$ und $\text{cut}_{P_2^u}$ liefern in diesem Fall ebenfalls 0 als Ergebnis. Es gilt dann $\text{cut}_{P^u}''(\mathcal{M}, s, s') = \mathbf{P}^u(s, s')$.

Für den Fall $E^u(s) > 0$ gilt:

$$\begin{aligned} \text{cut}_{P_1^u}^{\mathcal{M}, s, s'}(\mathbf{P}^u(s, s')) &= \max\{\mathbf{P}^l(s, s'), \min\{\mathbf{P}^u(s, s'), 1 - \sum_{v \neq s'} \mathbf{P}^l(s, v)\}\} \\ &\leq \mathbf{P}^u(s, s') \end{aligned}$$

$$\begin{aligned} \text{cut}_{P_2^u}^{\mathcal{M}, s, s'}(\mathbf{P}^u(s, s')) &= \max\{\mathbf{P}^l(s, s'), \min\{\mathbf{P}^u(s, s'), \mathbf{R}^u(s, s')/E^l(s)\}\} \\ &\leq \mathbf{P}^u(s, s') \end{aligned}$$

$$\Rightarrow \text{cut}_{P^u}''(\mathcal{M}, s, s') \leq \mathbf{P}^u(s, s') \text{ für alle } s, s' \in S$$

$$\Rightarrow \text{cut}_{P^u}'(\mathcal{M}) \leq \mathbf{P}^u$$

$$\Rightarrow \text{cut}_{P^u}(\mathcal{M}) \leq (\mathbf{P}^l, \mathbf{P}^u)$$

Für cut_{E^u} gilt:

$$\begin{aligned}
cut_{E_1^u}^{\mathcal{M},s}(E^u(s)) &= \max\{E^l(s), \min\{E^u(s), \mathbf{R}^u(s, S)\}\} \leq E^u(s) \\
cut_{E_2^u}^{\mathcal{M},s,s'}(E^u(s)) &= \max\{E^l(s), \min\{E^u(s), \mathbf{R}^u(s, s')/\mathbf{P}^l(s, s')\}\} \leq E^u(s) \\
&\Rightarrow cut''_{E^u}(\mathcal{M}, s, s') \leq E^u(s) \text{ für alle } s, s' \in S \\
&\Rightarrow cut'_{E^u}(\mathcal{M}) \leq E^u \\
&\Rightarrow cut_{E^u}(\mathcal{M}) \leq (E^l, E^u)
\end{aligned}$$

Auf der nächsten Ebene zeigen wir, dass das Ergebnis der Funktionen cut_R, cut_P und cut_E eine *feinere* ACTMC ist als die gegebene ACTMC. Wir verwenden hierbei die Projektion $(M_1, M_2, \dots, M_n)|_i = M_i$ auf die i -te Komponente eines Vektors:

Für alle $s, s' \in S$ gilt:

$$\begin{aligned}
&cut_{R^l}(\mathcal{M}[\mathbf{R} \mapsto cut_{R^u}(\mathcal{M})])|_1(s, s') \\
&= \max\{cut_{R_1^l}(\mathcal{M}, s, s'), cut_{R_2^l}(\mathcal{M}, s, s')\} \geq \mathbf{R}^l(s, s') \\
&cut_{R^l}(\mathcal{M}[\mathbf{R} \mapsto cut_{R^u}(\mathcal{M})])|_2(s, s') \\
&= \min\{cut_{R_1^u}(\mathcal{M}, s, s'), cut_{R_2^u}(\mathcal{M}, s, s')\} \leq \mathbf{R}^u(s, s') \\
&\Rightarrow (S, cut_{R^l}(\mathcal{M}[\mathbf{R} \mapsto cut_{R^u}(\mathcal{M})]), \mathbf{P}^l, E, L) = cut_R(\mathcal{M}) \subseteq \mathcal{M} \\
&cut_{P^l}(\mathcal{M}[\mathbf{P} \mapsto cut_{P^u}(\mathcal{M})])|_1(s, s') \\
&= \max\{cut_{P_1^l}(\mathcal{M}, s, s'), cut_{P_2^l}(\mathcal{M}, s, s')\} \geq \mathbf{P}^l(s, s') \\
&cut_{P^l}(\mathcal{M}[\mathbf{P} \mapsto cut_{P^u}(\mathcal{M})])|_2(s, s') \\
&= \min\{cut_{P_1^u}(\mathcal{M}, s, s'), cut_{P_2^u}(\mathcal{M}, s, s')\} \leq \mathbf{P}^u(s, s') \\
&\Rightarrow (S, \mathbf{R}^l, cut_{P^l}(\mathcal{M}[\mathbf{P} \mapsto cut_{P^u}(\mathcal{M})]), E, L) = cut_P(\mathcal{M}) \subseteq \mathcal{M} \\
&cut_{E^l}(\mathcal{M}[E \mapsto cut_{E^u}(\mathcal{M})])|_1(s, s') \\
&= \max\{cut_{E_1^l}(\mathcal{M}, s, s'), cut_{E_2^l}(\mathcal{M}, s, s')\} \geq E^l(s) \\
&cut_{E^l}(\mathcal{M}[E \mapsto cut_{E^u}(\mathcal{M})])|_2(s, s') \\
&= \min\{cut_{E_1^u}(\mathcal{M}, s, s'), cut_{E_2^u}(\mathcal{M}, s, s')\} \leq E^u(s) \\
&\Rightarrow (S, \mathbf{R}^l, \mathbf{P}^l, cut_{E^l}(\mathcal{M}[E \mapsto cut_{E^u}(\mathcal{M})]), L) = cut_E(\mathcal{M}) \subseteq \mathcal{M}
\end{aligned}$$

Auf der höchsten Ebene zeigen wir schließlich, dass die Monotonie bei Anwendung von cut_R, cut_P und cut_E erhalten bleibt. Durch wiederholtes Anwenden der cut -Funktionen erhalten wir also eine monotone Folge von ACTMCs.

$$\begin{aligned}
& cut_R(cut_P(cut_E(\mathcal{M}))) \\
&= cut_R(cut_P(\underbrace{(S, \mathbf{R}^I, \mathbf{P}^I, cut_{E^l}(cut_{E^u}(\mathcal{M})), L)}_{=\mathcal{M}' \subseteq \mathcal{M}}), L)) \\
&= cut_R(\underbrace{(S, \mathbf{R}^I, cut_{P^l}(cut_{P^u}(\mathcal{M}')), cut_{E^l}(cut_{E^u}(\mathcal{M})), L)}_{=\mathcal{M}'' \subseteq \mathcal{M}'}) \\
&= \underbrace{(S, cut_{R^l}(cut_{R^u}(\mathcal{M}'')), cut_{P^l}(cut_{P^u}(\mathcal{M}')), cut_{E^l}(cut_{E^u}(\mathcal{M})), L)}_{=\mathcal{M}''' \subseteq \mathcal{M}''}
\end{aligned}$$

Es gilt also $\mathcal{M}''' \subseteq \mathcal{M}'' \subseteq \mathcal{M}' \subseteq \mathcal{M}$ und aufgrund der Transitivität von \subseteq auch $\mathcal{M}''' \subseteq \mathcal{M}$. Wir haben damit gezeigt, dass die Folge $(cut_{RPE}^0(\mathcal{M}), cut_{RPE}^1(\mathcal{M}), \dots)$ tatsächlich monoton ist.

Laut [BS81] gilt, dass monotone Folgen genau dann konvergieren, wenn sie beschränkt sind. Wir müssen demnach noch die Beschränktheit des cut zeigen, also dass es eine abstrakte zeitstetige Markov-Kette \mathcal{M}_\perp gibt, für die gilt:

$$\mathcal{M}_\perp \subseteq cut_{RPE}^i(\mathcal{M}) \text{ für alle } i \in \mathbb{N}_{\geq 0}$$

Gegeben sei $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots) = (cut_{RPE}^0(\mathcal{M}), cut_{RPE}^1(\mathcal{M}), cut_{RPE}^2(\mathcal{M}), \dots)$ mit $\mathcal{M}_i = (S_i, \mathbf{R}_i^I, \mathbf{P}_i^I, E_i^I, L_i)$. Aus der Monotonie der Folge bezüglich \supseteq und der Voraussetzung für Intervalle, dass untere Grenzen stets kleiner sein müssen als obere Grenzen, können wir folgendes ableiten:

- $S_0 = S_1 = S_2 = \dots$
- $L_0 = L_1 = L_2 = \dots$
- $\mathbf{R}_0^l \leq \mathbf{R}_1^l \leq \mathbf{R}_2^l \leq \dots \leq \mathbf{R}_2^u \leq \mathbf{R}_1^u \leq \mathbf{R}_0^u$
- $\mathbf{P}_0^l \leq \mathbf{P}_1^l \leq \mathbf{P}_2^l \leq \dots \leq \mathbf{P}_2^u \leq \mathbf{P}_1^u \leq \mathbf{P}_0^u$
- $E_0^l \leq E_1^l \leq E_2^l \leq \dots \leq E_2^u \leq E_1^u \leq E_0^u$

Nun kann man eine begrenzende Markov-Kette \mathcal{M}_\perp angeben, als:

$$\begin{aligned}
\mathcal{M}_\perp &= (S_0, \mathbf{R}_\perp^I, \mathbf{R}_\perp^I, E_\perp^I, L_0) \text{ mit} \\
&\mathbf{R}_\perp^l \in \mathbf{R}_i \text{ für alle } i \in \mathbb{N}_{\geq 0} \text{ und } \mathbf{R}_\perp^u = \mathbf{R}_\perp^l \\
&\mathbf{P}_\perp^l \in \mathbf{P}_i \text{ für alle } i \in \mathbb{N}_{\geq 0} \text{ und } \mathbf{P}_\perp^u = \mathbf{P}_\perp^l \\
&E_\perp^l \in E_i \text{ für alle } i \in \mathbb{N}_{\geq 0} \text{ und } E_\perp^u = E_\perp^l
\end{aligned}$$

Die Folge $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2, \dots)$ ist demnach beschränkt. Zusammen mit der Monotonie haben wir damit auch die Konvergenz der Folge gezeigt. □

Damit wissen wir, wie der cut über ein iteratives Verfahren approximiert werden kann. Auf eine genauere Untersuchung des Konvergenzverhaltens wird an dieser Stelle verzichtet.

Wir definieren nun eine Menge von Schedulingen, von denen jeder nur aus einer endlichen Menge von Wahlmöglichkeiten pro Zustand auswählen kann. In Kapitel 3.7 werden wir sehen, dass diese Menge von Schedulingen bezüglich der Berechnung von minimalen und maximalen Wahrscheinlichkeiten weniger mächtig ist als die Menge der HD-Schedulingen. Dennoch können wir damit eine interessante Teilklasse des Erreichbarkeitsproblems behandeln.

Definition 21 (Extremer Ratenvektor). Der Ratenvektor $\mu \in \text{rates}(\mathbf{R}(s, \cdot))$ bezüglich einer Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ heißt *extrem*, falls gilt:

$$\mu \in \text{extr}(\mathbf{R}^I(s, \cdot), \mathbf{P}^I(s, \cdot), E^I(s), S)$$

wobei

$$\mu \in \text{extr}(r^I(s, \cdot), p^I(s, \cdot), e^I(s), S')$$

gdw.

$$\begin{aligned} & (\exists s' \in S' : \mu(s') \in \{r^l(s, s'), r^u(s, s')\} \\ & \quad \wedge \mu \in \text{extr}(\text{cut}^\circ(r^I(s, \cdot)[s' \mapsto \mu(s')], p^I(s, \cdot), e^I(s)), S' \setminus \{s'\})) \\ & \vee (S' = \emptyset \wedge \mu = r^l = r^u) \end{aligned}$$

wobei die Funktion $\text{cut}^\circ(r^I, p^I, e^I)$ eine abkürzende Schreibweise von $\text{cut}(\mathcal{M}^\circ)$ darstellt für $\mathcal{M}^\circ = (S, r^I, p^I, e^I, L)$ und wobei hier mit $x^I[s \mapsto x]$ die Substitution der Werte $x^l(s)$ und $x^u(s)$ durch den Wert x bezeichnet ist.

—

Da für ein $\mu \in \text{rates}(\mathbf{R}(s, \cdot))$ die Exitrate $E(s) = \mu(S)$ eindeutig ist, folgt auch die Verteilung $\bar{\mu} \in \text{distr}(\mathbf{P}(s, \cdot))$ eindeutig aus $\mu(s') = E(s) \cdot \bar{\mu}(s')$.

Beispiel 11. Sei die abstrakte zeitstetige Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ aus Abbildung 15 gegeben. Um zu bestimmen, ob beispielsweise $\mu = (0, 3, 2)$ ein extremer Ratenvektor ist, müssen wir lediglich überprüfen, ob er die Definition 21 erfüllt.

$$\mu \in \text{extr}(r^I(s, \cdot), p^I(s, \cdot), e^I(s), S')$$

gdw.

$$\begin{aligned} & (\exists s' \in S' : \mu(s') \in \{r^l(s, s'), r^u(s, s')\} \\ & \quad \wedge \mu \in \text{extr}(\text{cut}^\circ(r^I(s, \cdot)[s' \mapsto \mu(s')], p^I(s, \cdot), e^I(s)), S' \setminus \{s'\})) \\ & \vee (S' = \emptyset \wedge \mu = r^l = r^u) \end{aligned}$$

Da $\mu(s_1) = 3 \in [3, 5]$ die Existenzbedingung erfüllt, muss nur noch folgendes nachgewiesen werden:

$$\mu \in \text{extr}(\text{cut}^\circ(r^I(s, \cdot)[s_1 \mapsto 3], p^I(s, \cdot), e^I(s)), \{s_0, s_2\})$$

Werten wir nun den *cut* aus, so muss das Ratenintervall von s_0 nach s_2 sowie die Wahrscheinlichkeitsintervalle wegen der festen Bindung zu den Raten über das Punktintervall $[5, 5]$ der Exitrate von s_0 wie folgt angepasst werden:

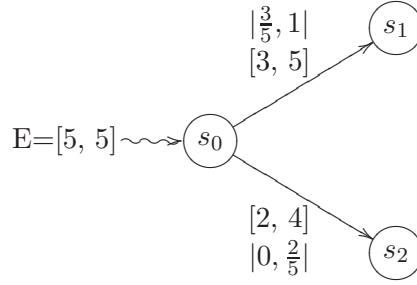


Abbildung 15: Extremer Ratenvektor

$$\begin{aligned} & cut^\circ(r^I(s, \cdot)[s_1 \mapsto 3], p^I, e^I) \\ &= (r^I(s, \cdot)[s_0 \mapsto 0][s_1 \mapsto 3, s_2 \mapsto 2], p^I(s, \cdot)[s_0 \mapsto 0][s_1 \mapsto \frac{3}{5}, s_2 \mapsto \frac{2}{5}], e^I) \end{aligned}$$

Da durch den *cut* bereits sämtliche Werte festgelegt wurden und diese nicht inkonsistent sind, erhält man durch Iteration $\mu = (0, 3, 2) = r^l = r^u$ und eine leere Menge für S' in der Definition, woraus folgt:

$$\mu \in extr(cut^\circ(r^I(s, \cdot)[s_1 \mapsto 3], p^I(s, \cdot), e^I(s)), \{s_0, s_2\})$$

Definition 22 (Extreme Scheduler). Ein extremer Scheduler η bezüglich einer abstrakten Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ ist ein HD-Scheduler, der für alle möglichen Folgen von Zuständen s_0, \dots, s_n lediglich extreme Ratenvektoren $\mu \in extr(\mathbf{R}^I(s, \cdot), \mathbf{P}^I(s, \cdot), E^I(s), S)$ mit dazu passenden Exitraten und Wahrscheinlichkeitsverteilungen wählt.

Wir schließen dieses Unterkapitel mit einigen Beobachtungen über Eigenschaften von bereinigten ACTMCs ab, die in späteren Beweisen noch benötigt werden.

Korollar 1 (Eigenschaften bereinigter Markov-Ketten). Für bereinigte abstrakte zeitstetige Markov-Ketten $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ gelten die folgenden Aussagen:

1. $\mathbf{R}^u(s, s') \leq E^u(s) - \sum_{v \neq s'} \mathbf{R}^l(s, v)$ und
 $\mathbf{R}^l(s, s') \geq E^l(s) - \sum_{v \neq s'} \mathbf{R}^u(s, v)$ für alle $s, s' \in S$.
2. $E^u(s) \leq \mathbf{R}^u(s, S)$ und
 $E^l(s) \geq \mathbf{R}^l(s, S)$ für alle $s \in S$.
3. $\mathbf{P}^u(s, s') \leq 1 - \sum_{v \neq s'} \mathbf{P}^l(s, v)$ falls $E^u(s) > 0$ und
 $\mathbf{P}^l(s, s') \geq 1 - \sum_{v \neq s'} \mathbf{P}^u(s, v)$ falls $E^l(s) > 0$
für alle $s, s' \in S$.
4. $\mathbf{R}^u(s, s') \leq E^u(s) \cdot \mathbf{P}^u(s, s')$ und
 $\mathbf{R}^l(s, s') \geq E^l(s) \cdot \mathbf{P}^l(s, s')$ für alle $s, s' \in S$.

5. $\mathbf{P}^u(s, s') \leq \mathbf{R}^u(s, s')/E^l(s)$ falls $E^u(s) > 0$ und
 $\mathbf{P}^l(s, s') \geq \mathbf{R}^l(s, s')/E^u(s)$ falls $E^l(s) > 0$
für alle $s, s' \in S$.
6. $E^u(s) \leq \mathbf{R}^u(s, s')/\mathbf{P}^l(s, s')$ und
 $E^l(s) \geq \mathbf{R}^l(s, s')/\mathbf{P}^u(s, s')$ für alle $s, s' \in S$.

Beweis. Wir zeigen für die oberen Intervallgrenzen jeweils über einen kurzen indirekten Beweis, dass die aufgeführten Aussagen gelten.

1. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $\mathbf{R}^u(s, s') > E^u(s) - \sum_{v \neq s'} \mathbf{R}^l(s, v)$. Durch Anwendung von cut_{R_1} würde $\mathbf{R}^u(s, s')$ genau auf $E^u(s) - \sum_{v \neq s'} \mathbf{R}^l(s, v)$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.
2. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $E^u(s) > \mathbf{R}^u(s, S)$. Durch Anwendung von cut_{E_1} würde $E^u(s, s')$ genau auf $\mathbf{R}^u(s, S)$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.
3. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $\mathbf{P}^u(s, s') > 1 - \sum_{v \neq s'} \mathbf{P}^l(s, v)$ und $E^u(s) > 0$. Durch Anwendung von cut_{P_1} würde $\mathbf{P}^u(s, s')$ genau auf $1 - \sum_{v \neq s'} \mathbf{P}^l(s, v)$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.
4. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $\mathbf{R}^u(s, s') > E^u(s) \cdot \mathbf{P}^u(s, s')$. Durch Anwendung von cut_{R_2} würde $\mathbf{R}^u(s, s')$ genau auf $E^u(s) \cdot \mathbf{P}^u(s, s')$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.
5. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $\mathbf{P}^u(s, s') > \mathbf{R}^u(s, s')/E^l(s)$ für $E^u(s) > 0$. Durch Anwendung von cut_{P_2} würde $\mathbf{P}^u(s, s')$ auf $\mathbf{R}^u(s, s')/E^l(s)$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.
6. Angenommen in \mathcal{M} gäbe es $s, s' \in S$ mit $E^u(s) > \mathbf{R}^u(s, s')/\mathbf{P}^l(s, s')$. Durch Anwendung von cut_{E_2} würde $E^u(s)$ genau auf $\mathbf{R}^u(s, s')/\mathbf{P}^l(s, s')$ verringert werden. Dies steht aber im Widerspruch zur Voraussetzung, dass \mathcal{M} bereinigt ist.

Für die unteren Intervallgrenzen verlaufen die Beweise analog.

□

3.3 Probabilistische Simulation

Im Beispiel 7 mussten wir *ein Auge zudrücken*, um Zustände zusammenzufassen, die nicht bisimilar waren. Da wir nun definiert haben, was eine abstrakte zeitstetige Markov-Kette ist, können wir auch genau beschreiben, wann wir *ein Auge zudrücken* dürfen und wann nicht. Die grundsätzliche Idee dabei ist zu fordern, dass nur der Makro-Zustand aus der Abstraktion die durch ihn repräsentieren Zustände simuliert und nicht zusätzlich auch umgekehrt wie bei der Bisimulationsäquivalenz. Damit werden wir später immer noch folgern können, dass alle Eigenschaften, die in einer Abstraktion gelten, auch in der ursprünglichen Markov-Kette gelten müssen.

Um die Definition der probabilistischen Simulation verstehen zu können, benötigen wir noch den Begriff der Gewichtung. Im Allgemeinen bezeichnet man als Gewichtung über S und S' bezüglich einer Relation $\mathcal{R} \subseteq S \times S'$ die Lösung des *maximalen Flußproblems* in einem Graphen mit Zustandsmenge $S \cup S'$ und der Transitionsmatrix $\Delta \subseteq (S \cup S') \times (S \cup S')$ mit $(s, s') \in \Delta$ gdw. $s \in S$ und $s' \in S'$ (siehe [Bai96]).

Wir verwenden Gewichtungen um zu beschreiben, welche Verteilungen über die Zustände einer Abstraktion den Verteilungen bezüglich der ursprünglichen Markov-Kette entsprechen. Dabei verlangen wir, dass Zustände nur auf ihre Makro-Zustände abgebildet werden dürfen. Das heißt, dass es von Zustand $s \in S$ nach Zustand $s' \in S'$ nur Kanten mit Gewicht größer 0 geben darf, wenn s durch s' *repräsentiert* bzw. *simuliert* wird. Folgendes Beispiel wird dies verdeutlichen.

Beispiel 12 (Gewichtung). Betrachten wir die Abstraktionen aus Beispiel 7. Die Zustandsmenge der ersten Abstraktion war $\{1111, 0111, 1??1, 1001, 0??1, 0001, 0000\}$ und die der zweiten $\{s_{3,1}, s_{2,1}, s_{1,1}, s_{0,1}, s_{0,0}\}$. Nehmen wir an, dass wir für eine bestimmte Situation mit $\mu = (\frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10}, 0)$ die Wahrscheinlichkeiten dafür kennen, in den einzelnen Zuständen der ersten Abstraktion zu sein. Uns interessiert nun, wie die Wahrscheinlichkeitsverteilung für dieselbe Situation in der zweiten Abstraktion aussehen sollte.

Wir erhalten durch Lösen des abgebildeten vereinfachten¹⁸ Flussproblems für die entsprechende Verteilung der zweiten Abstraktion den Vektor $\mu' = (\frac{1}{10}, \frac{5}{10}, \frac{3}{10}, \frac{1}{10}, 0)$, wie man an Abbildung 16 leicht nachvollziehen kann.

Die gleiche Situation, die in der ersten Abstraktion zu der Verteilung μ führt, müsste also in der zweiten Abstraktion die Verteilung μ' ergeben.

¹⁸Hier sind nur Kanten von s nach s' abgebildet, falls s von s' repräsentiert wird. Die nachfolgende Definition der probabilistischen Simulation fordert, dass eine Kante einer Gewichtung von Zustand s nach Zustand s' nur dann größer als 0 sein darf, falls s und s' in Relation stehen. Bezüglich dieser Definition ist das vereinfachte maximale Flussproblem zu betrachten.

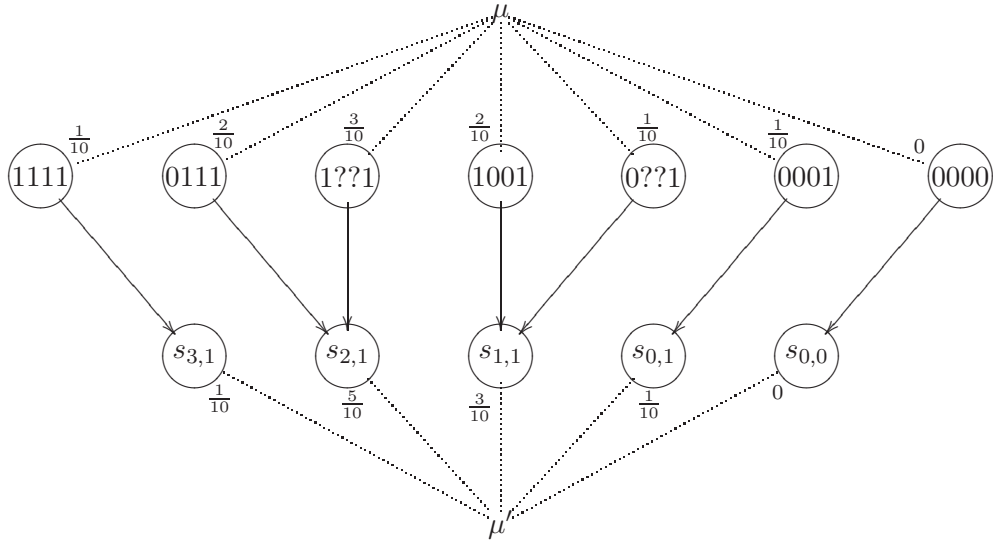


Abbildung 16: Gewichtung

Definition 23 (Vereinigung). Die Vereinigung $\mathcal{M} = \mathcal{M}_1 \uplus \mathcal{M}_2$ zweier abstrakter Markov-Ketten $\mathcal{M}_1 = (S_1, \mathbf{R}_1^I, \mathbf{P}_1^I, E_1^I, L_1)$ und $\mathcal{M}_2 = (S_2, \mathbf{R}_2^I, \mathbf{P}_2^I, E_2^I, L_2)$ sei definiert durch $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ mit:

- $S = S_1 \cup S_2$, wobei o.B.d.A. von $S_1 \cap S_2 = \emptyset$ ausgegangen werden soll,

- $$\mathbf{Q}(s, s') = \begin{cases} \mathbf{Q}_1(s, s') & \text{falls } s, s' \in S_1 \\ \mathbf{Q}_2(s, s') & \text{falls } s, s' \in S_2 \\ 0 & \text{sonst} \end{cases} \quad \text{für } \mathbf{Q} \in \{\mathbf{R}^l, \mathbf{R}^u, \mathbf{P}^l, \mathbf{P}^u\}$$

- $$\mathbf{Q}(s) = \begin{cases} \mathbf{Q}_1(s) & \text{falls } s \in S_1 \\ \mathbf{Q}_2(s) & \text{falls } s \in S_2 \end{cases} \quad \text{für } \mathbf{Q} \in \{E^l, E^u\}$$

- $$L(s, a) = \begin{cases} L_1(s, a) & \text{falls } s \in S_1 \\ L_2(s, a) & \text{falls } s \in S_2 \end{cases}$$

Wir verwenden im Folgenden eine leicht abgewandelte Definition für probabilistische Simulation, wie sie in [BHKW05] zu finden ist. Statt zu fordern, dass ein simulierender Zustand schneller oder gleich schnell wie der simulierte Zustand verlassen werden kann, fordern wir die stärkere Bedingung, dass simulierender und simulierter Zustand genau die gleiche totale Ausgangsrate haben müssen.

Definition 24 (Probabilistische Simulation). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette, dann heißt $\mathcal{R} \subseteq S \times S$ probabilistische Simulationsrelation, falls aus $s\mathcal{R}s'$ folgt:

1. Für alle $a \in AP$ gilt $(L(s', a) \neq ?) \rightarrow (L(s, a) = L(s', a))$, d.h. falls die Gültigkeit von a in s' nicht unbestimmt ist, gilt a in s gdw. a auch in s' gilt.
2. Für alle Verteilungen $\bar{\mu} \in \text{distr}(\mathbf{P}(s, \cdot))$ und alle (nach Anwendung des *cut* wählbaren) Exitraten $E(s)$ bezüglich \mathcal{M} existiert eine Verteilung $\bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))$ mit entsprechender Exitrate $E(s')$, so dass gilt:
 - 2.1. Es existiert eine Gewichtung $\Delta : S \times S \mapsto [0, 1]$ mit:
 - 2.1.1. $\Delta(u, v) > 0 \Rightarrow u\mathcal{R}v$,
 - 2.1.2. $\Delta(u, S) = \bar{\mu}(u)$,
 - 2.1.3. $\Delta(S, v) = \bar{\mu}'(v)$.
 - 2.2. $E(s) = E(s')$, d.h. Zustände s' und s haben gleiche totale Ausgangsraten.

—

Falls eine probabilistische Simulationsrelation \mathcal{R} existiert, für die gilt $s\mathcal{R}s'$, dann schreiben wir auch $s \preceq s'$ (s wird simuliert von s').

Beispiel 13 (Probabilistische Simulation). Betrachten wir den Begriff der probabilistischen Simulation aus spieltheoretischer Sicht, so sagen wir, dass ein Zustand s von einem anderen Zustand t simuliert wird, falls ein Spieler B, der sich in t befindet, auf jeden möglichen *Zug* eines Spielers A in Zustand s reagieren kann. Dazu betrachten wir nun die ACTMC¹⁹ aus Abbildung 17 und die abgewinkelte Darstellung in Abbildung 18. Die einzelnen Zustände sind pro Ebene leicht versetzt worden, um die Zusammenhänge zwischen s -Zuständen und t -Zuständen besser verdeutlichen zu können. Die Zustände s_1 und t_1 , s_2 und t_2 sowie s_3 und t_2 stehen in probabilistischer Simulationsrelation. Im Folgenden werden wir dies nicht beweisen²⁰, stattdessen betrachten wir ein beispielhaft die ersten Züge eines Spiels zwischen Spielern A und B.

Wir gehen von s_1 und t_1 als Startzustände für Spieler A und Spieler B aus. Spieler A darf also nur in den s -Zuständen und Spieler B nur in den t -Zuständen ziehen. Da s_1 und t_1 in Relation stehen, sind die Startzustände also so gewählt, dass Spieler B eine Chance hat zu gewinnen²¹.

Zieht nun Spieler A mit einem Anteil $\frac{1}{2}$ nach s_1 und dem restlichen Anteil $\frac{1}{2}$ nach s_2 (mit Exitrate E), so muss der Spieler B die Möglichkeit haben, mit den gleichen

¹⁹In diesem Beispiel ist die Exitrate immer gleich 1, weswegen die Raten- und die Wahrscheinlichkeitsintervallmatrix gleich sind.

²⁰Den Anteil der Markov-Kette mit t -Zuständen kann man durch die in den folgenden Abschnitten vorgestellten Abstraktionen A, B und C gleichermaßen aus dem Anteil mit s -Zuständen bestimmen. Dazu werden s_1 zu t_1 sowie s_2 und s_3 zu t_2 zusammengefasst. Aus den Sätzen 2, 3 bzw. 4 folgt dann die Behauptung.

²¹Spieler B gewinnt, falls Spieler A niemals einen Zug machen kann, der von Spieler B nicht bedient werden kann.

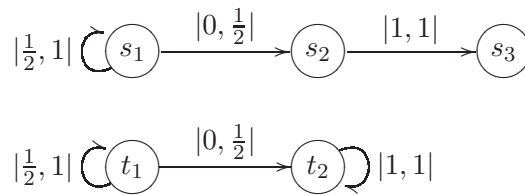


Abbildung 17: Probabilistische Simulation

Anteilen und der gleichen Exitrate in Zustände t_1 und t_2 zu ziehen, für die $s_1 \preceq t_1$ und $s_2 \preceq t_2$ gilt. In der mittleren Abbildung ist das zu lösende (vereinfachte) maximale Flussproblem durch geschlängelte Linien und Pfeile hervorgehoben. Für Spieler B wurde jeweils die Wahrscheinlichkeit $\frac{1}{2}$ gewählt. Für diese Wahl existiert eine Lösung des Flußproblems, die durch die Beschriftung der geschlängelten Linien zwischen den s -Zuständen und den t -Zuständen angedeutet wurde.

Nun befinden sich die beiden Spieler anteilsweise in verschiedenen Zuständen, und zwar in s_1 und s_2 bzw. t_1 und t_2 . Dies kann im Sinne der probabilistischen Simulation so interpretiert werden, dass Spieler A für jeden Zustand indem er sich befindet einen Zug macht, auf den Spieler B entsprechend reagieren muss. In der unteren Abbildung sind nun also zwei maximale Flussprobleme zu betrachten. Das eine ist wieder durch geschlängelte, das andere durch gestrichelte Linien und Pfeile hervorgehoben. Wie in der mittleren Abbildung wurde auch diese Abbildung jeweils mit einem gültigen Ergebnis beschriftet.

Es ist zu beachten, dass in der abstrakten zeitstetigen Markov-Kette in Abbildung 17 die Zustände s_2 und t_2 sowie s_3 und t_2 *nicht* bisimilar sind. Diese Tatsache ist einfach nachzuvollziehen, indem man sich dem Problem wieder aus spieltheoretischer Sicht nähert.

Beginne Spieler A in Zustand t_2 und Spieler B in s_2 . Spieler A kann in diesem Fall nur den Übergang nach t_2 wählen, sowie Spieler B nur den Übergang nach s_3 wählen kann. Soweit gibt es noch kein Problem, wenn Spieler A jedoch wiederum den Übergang von t_2 nach t_2 wählt, so kann Spieler B in s_3 keine vergleichbare Wahl treffen, da s_3 absorbierend ist. Damit sind die Zustände s_2 und t_2 sowie s_3 und t_2 also nicht bisimilar. Ebenso sind auch s_1 und t_1 nicht bisimilar.

—

Die probabilistische Simulation spielt im folgenden Abschnitt eine wesentliche Rolle. Es ist also empfehlenswert die obigen Definitionen und Beispiele genauestens zu studieren.

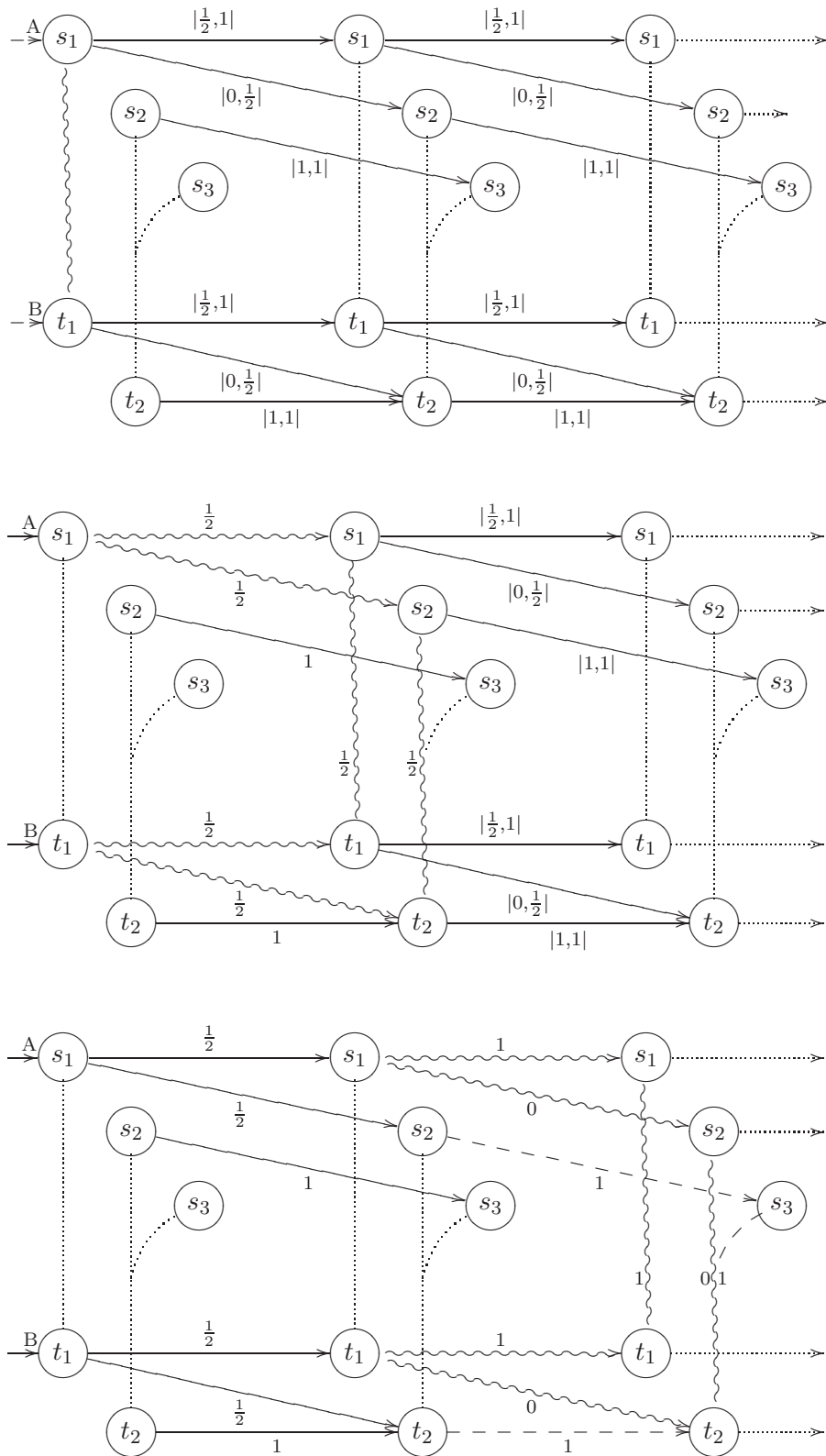


Abbildung 18: Probabilistische Simulation (Abwicklung)

3.4 Mögliche Abstraktionen

In der Regel sprechen wir von einer abstrakten Markov-Kette im Zusammenhang mit einer *gegebenen* oder *ursprünglichen* Markov-Kette. Daher wollen wir nun definieren, wie man durch das Zusammenfassen von Zuständen abstrakte Markov-Ketten erhält. Das *Zusammenfassen von Zuständen* werden wir mathematisch als Partitionierung²² beschreiben. Dabei werden alle Zustände die in einer Partition enthalten sind zu einem *Makro-Zustand* zusammengefasst.

Im Folgenden stellen wir drei mögliche Abstraktionen vor und zeigen, dass die Zustände der resultierenden abstrakten Markov-Ketten die Zustände der ursprünglichen Markov-Kette simulieren. Spezifikationen, die in der abstrakten Markov-Kette als gültig verifiziert werden können, werden sich also in der ursprünglichen Markov-Kette verifizieren lassen.

3.4.1 Abstraktion A

Für den ersten Ansatz einer Abstraktion zeitstetiger Markov-Ketten betrachten wir primär die Raten. Werden durch eine Partitionierung mehrere Zustände zusammengefasst, so sollte die abstrakte Markov-Kette diese Zustände wie bei der Abstraktion durch Bisimulationsäquivalenzklassen mit einem Makro-Zustand beschreiben.

Bei den Beschriftungen können wir diejenigen Eigenschaften als wahr (oder als falsch) im Makro-Zustand übernehmen, die in allen zugehörigen Zuständen der ursprünglichen Markov-Kette auch schon wahr (oder falsch) waren. Ist die Eigenschaft in einigen ursprünglichen Zuständen wahr, in anderen jedoch unwahr, so muss im Makro-Zustand der Wert *unbestimmt* (?) eingetragen werden, da der Zustand die ursprünglichen Zustände sonst nicht simuliert.

Bei den ausgehenden Raten wählen wir für den Makro-Zustand die kleinsten und größten Werte der ursprünglichen Zustände, um der Simulationsrelation Rechnung zu tragen. Würden wir die Grenzen enger wählen, so gäbe es in der konkreten Markov-Kette einen Zustand, für den die ausgehenden Raten nicht durch die abstrakte Markov-Kette simuliert werden könnte. Formal zeigen wir dies in Satz 2. Für die Wahrscheinlichkeiten und Exitraten wählen wir die Grenzen so, dass zu Beginn durch Ausführung des *cut* keinerlei Raten von den Intervallen abgeschnitten und alle Werte aus den Intervallen zu CTMCs vervollständigt werden können.

Definition 25 (Abstraktion A). Eine Partitionierung $\mathcal{A} = \{A_1, \dots, A_n\}$ des Zustandsraumes einer (abstrakten) zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ induziert eine Abstraktion $abstractA(\mathcal{M}, \mathcal{A}) = (\tilde{S}, \tilde{\mathbf{R}}^I, \tilde{\mathbf{P}}^I, \tilde{E}^I, \tilde{L})$, mit

²² $\mathcal{A} = \{A_1, \dots, A_n\}$ heißt Partitionierung des Zustandsraums S , falls
a) für alle $i \in \{1, \dots, n\}$ gilt, dass $A_i \subseteq S$ und $A_i \neq \emptyset$,
b) für alle Zustände $s \in S$ eine Partition A_i existiert mit $s \in A_i$ und
c) $A_i \cap A_j = \emptyset$ für $i, j \in \{1, \dots, n\}$ und $i \neq j$.

- $\tilde{S} = \mathcal{A}$ ist die Menge der abstrakten Zustände,
- $\tilde{\mathbf{R}}^l(A_i, A_j) = \min_{s \in A_i} \mathbf{R}^l(s, A_j)$ ist die untere Schranke für die *Geschwindigkeit* von A_i nach A_j zu wechseln,
- $\tilde{\mathbf{R}}^u(A_i, A_j) = \max_{s \in A_i} \mathbf{R}^u(s, A_j)$ ist die obere Schranke für die *Geschwindigkeit* von A_i nach A_j zu wechseln,
- $\tilde{\mathbf{P}}^l(A_i, A_j) = \tilde{\mathbf{R}}^l(A_i, A_j) // (\tilde{\mathbf{R}}^l(A_i, A_j) + \sum_{A_k \neq A_j} \tilde{\mathbf{R}}^u(A_i, A_k))$,
- $\tilde{\mathbf{P}}^u(A_i, A_j) = \tilde{\mathbf{R}}^u(A_i, A_j) // (\tilde{\mathbf{R}}^u(A_i, A_j) + \sum_{A_k \neq A_j} \tilde{\mathbf{R}}^l(A_i, A_k))$,
- $\tilde{E}^l(A_i) = \tilde{\mathbf{R}}^l(A_i, S)$,
- $\tilde{E}^u(A_i) = \tilde{\mathbf{R}}^u(A_i, S)$,
- $\tilde{L}(A_i, a) = \begin{cases} \top & \text{falls für alle } s \in A_i \text{ gilt, dass } L(s, a) = \top \\ \perp & \text{falls für alle } s \in A_i \text{ gilt, dass } L(s, a) = \perp \\ ? & \text{sonst} \end{cases}$

wobei $x // y = \begin{cases} x/y & \text{falls } y \neq 0 \\ 0 & \text{falls } y = 0 \end{cases}$

—

Korollar 2 (Gültige Wahrscheinlichkeiten). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette und $\mathcal{M} = \text{abstractA}(\mathcal{M}, \mathcal{A})$ eine Abstraktion bezüglich der Partitionierung $\mathcal{A} = \{A_1, \dots, A_n\}$. Für die mit Abstraktion \mathcal{A} berechneten Wahrscheinlichkeitsintervalle gilt dann $[\tilde{\mathbf{P}}^l(A_i, A_j), \tilde{\mathbf{P}}^u(A_i, A_j)] \subseteq [0, 1]$.

Beweis. Für $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ und $\mathcal{A} = \{A_1, \dots, A_n\}$ ist zu zeigen, dass die untere Wahrscheinlichkeitsintervallgrenze der Abstraktion $\text{abstractA}(\mathcal{M}, \mathcal{A})$ größer oder gleich Null ist und die obere kleiner oder gleich Eins:

$$\tilde{\mathbf{P}}^l(A_i, A_j) = \underbrace{\tilde{\mathbf{R}}^l(A_i, A_j)}_{\geq 0} // \left(\underbrace{\tilde{\mathbf{R}}^l(A_i, A_j)}_{\geq 0} + \sum_{A_k \neq A_j} \underbrace{\tilde{\mathbf{R}}^u(A_i, A_k)}_{\geq 0} \right) \geq 0$$

Für die obere Intervallgrenze müssen wir zwischen absorbierendem Zustand und nicht absorbierendem unterscheiden. Für absorbierende Zustände sind alle Ratenintervalle der Form $[0, 0]$ und wegen der Definition von $x // y$ sinnigerweise auch die Wahrscheinlichkeitsintervalle.

Für nicht absorbierende Zustände können ähnliche Fälle auftreten, für die der Nenner bei der Berechnung der Wahrscheinlichkeiten Null wird. Diese werden wie bei den absorbierenden Zuständen behandelt. Andernfalls ist zumindest einer der Summanden im Nenner größer Null:

$$\tilde{\mathbf{P}}^u(A_i, A_j) = \tilde{\mathbf{R}}^u(A_i, A_j) / \underbrace{\left(\underbrace{\tilde{\mathbf{R}}^u(A_i, A_j)}_{\geq 0} + \sum_{A_k \neq A_j} \underbrace{\tilde{\mathbf{R}}^l(A_i, A_k)}_{\geq 0} \right)}_{> 0}$$

Aus $\tilde{\mathbf{R}}^u(A_i, A_j) \leq \tilde{\mathbf{R}}^u(A_i, A_j) + \sum_{A_k \neq A_j} \tilde{\mathbf{R}}^l(A_i, A_k)$ folgt dann $\tilde{\mathbf{P}}^u(A_i, A_j) \leq 1$. \square

Bevor wir zum Beweis kommen, dass Makro-Zustände nach Abstraktion A die zugehörigen Zustände der ursprünglichen Markov-Kette simulieren, wollen wir uns vorher an einem Beispiel ansehen, wie die probabilistische Simulation im Zusammenhang mit Abstraktion zu verstehen ist.

Beispiel 14 (Probabilistische Simulation). Wir abstrahieren die zeitstetige Markov-Kette in der Abbildung 19 zu der rechten abstrakten Markov-Kette, indem wir die Zustände s_i und s'_i für $i \in \{0, 1\}$ zu \bar{s}_i zusammenfassen und die Intervallgrenzen nach Definition berechnen.

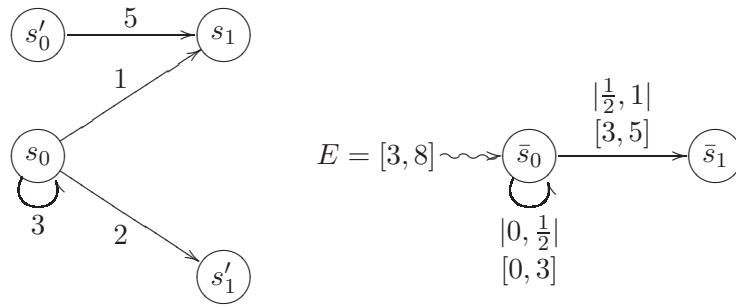


Abbildung 19: CTMC (links) und Abstraktion (rechts)

Wenn etwa der Zustand s_0 durch den Zustand \bar{s}_0 simuliert werden soll, so muss für alle möglichen Verteilungen in $\text{distr}(\mathbf{P}(s_0, \cdot))$ eine Gewichtung Δ und eine Verteilung aus $\text{distr}(\mathbf{P}(\bar{s}_0, \cdot))$ existieren, welche die Bedingungen aus Definition 24 erfüllen. Überprüfen wir dies für $\bar{\mu} = (\frac{3}{6}, 0, \frac{1}{6}, \frac{2}{6}) \in \text{distr}(\mathbf{P}(s_0, \cdot))$, um ein besseres Gefühl für die Bezeichnungen zu bekommen und das Verständnis der folgenden Beweise zu erleichtern.

Da nach Definition 24.1.1 für $\Delta(s, s')$ nur positive Werte erlaubt sind wenn $s \mathcal{R} s'$ gilt, können wir für alle Kombinationen von s und s' für die nicht $s \in \{s_0, s'_0, s_1, s'_1\}$ ist und $s' \in \{\bar{s}_0, \bar{s}_1\}$ den Wert von $\Delta(s, s')$ auf 0 festlegen. Diese sind in der Abbildung grau dargestellt. Die Gewichtung Δ muss nach Definition 24.1.2 in der Zeilensumme die Werte von μ ergeben und nach 24.1.3 ergeben sich aus der Spaltensumme dann die Werte von μ' . Mit der Gewichtung in Abbildung 20 erhält man eine zulässige Verteilung $\mu' = (\frac{3}{6}, \frac{3}{6}) = (\frac{1}{2}, \frac{1}{2}) \in \text{distr}(\mathbf{P}(\bar{s}_0, \cdot))$.

Wir werden später nach solchen Gewichtungen nicht mehr suchen müssen, sondern zeigen stattdessen, dass für die Abstraktionen eine solche immer existiert.

Δ	s_0	s'_0	s_1	s'_1	\bar{s}_0	\bar{s}_1	Σ
s_0	0	0	0	0	$\frac{3}{6}$	0	$\frac{3}{6}$
s'_0	0	0	0	0	0	0	0
s_1	0	0	0	0	0	$\frac{1}{6}$	$\frac{1}{6}$
s'_1	0	0	0	0	0	$\frac{2}{6}$	$\frac{2}{6}$
\bar{s}_0	0	0	0	0	0	0	0
\bar{s}_1	0	0	0	0	0	0	0
Σ	0	0	0	0	$\frac{3}{6}$	$\frac{3}{6}$	1

Abbildungung 20: Gewichtung und Verteilung für probabilistische Simulation

Satz 2 (Simulation durch Makro-Zustände nach Abstraktion A). Gegeben sei eine abstrakte zeitstetige Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ und $abstractA(\mathcal{M}, \mathcal{A})$, eine durch Partition \mathcal{A} induzierte Abstraktion von \mathcal{M} . Dann gilt:

$$s \in A \Rightarrow s \preceq A \text{ für alle } s \in S \text{ und } A \in \mathcal{A},$$

d.h. ein Zustand s in der ursprünglichen Markov-Kette wird durch den zugehörigen Makro-Zustand A der Abstraktion simuliert.

Beweis. Sei $\mathcal{M} = (\bar{S}, \bar{\mathbf{R}}^I, \bar{\mathbf{P}}^I, \bar{E}^I, \bar{L})$ eine abstrakte zeitstetige Markov-Kette, \mathcal{A} eine Partitionierung und $abstract(\mathcal{M}, \mathcal{A}) = (\mathcal{A}, \tilde{\mathbf{R}}^I, \tilde{\mathbf{P}}^I, \tilde{E}^I, \tilde{L})$ eine Abstraktion von \mathcal{M} . Die Relation \mathcal{R} definieren wir als $\mathcal{R} = \{(s, A) \mid s \in A, A \in \mathcal{A}\}$, d.h. \mathcal{R} ist durch die Partitionierung \mathcal{A} vollständig bestimmt. Um zu zeigen, dass \mathcal{R} eine Simulationrelation ist, zeigen wir, dass die Bedingungen aus Definition 24 erfüllt sind. Dazu betrachten wir im Folgenden die Vereinigung $\mathcal{M} \uplus abstract(\mathcal{M}, \mathcal{A}) = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$.

1. Für die erste Bedingung unterscheiden wir die Fälle $L(s, a) = ?$ und $L(s, a) \in \mathbb{B}$ mit $a \in AP$. Für $L(s, a) = ?$ ist die Bedingung trivialerweise erfüllt. Für den zweiten Fall bleibt zu zeigen, dass für alle $(s, s') \in \mathcal{R}$ gilt: $L(s, a) = L(s', a)$. Da per Definition in \mathcal{R} nur Paare (s, A) mit $s \in A$ enthalten sind und nach Definition 25 für alle $s \in A$ und $\alpha \in \mathbb{B}$ gilt, dass $L(A, a) = \alpha \Rightarrow L(s, a) = \alpha$, ist dieser Teil der Bedingung ebenfalls erfüllt.
2. Für die zweite Bedingung betrachten wir zunächst die Ratenvektoren und zeigen, dass die Exitraten von s und s' gleich sind. Da wir uns für diese Bedingung jedoch für die Verteilungen interessieren, müssen wir noch die Beobachtungen bezüglich der Ratenvektoren, die mit der Exitrate an die Wahrscheinlichkeiten gekoppelt sind, auf Verteilungen übertragen.

Legen wir zu jedem $\mu \in rates(\mathbf{R}(s, \cdot))$, ein passendes $\mu' \in rates(\mathbf{R}(s', \cdot))$ fest:

$$\mu'(A') = \begin{cases} \sum_{u \in A'} \mu(u) & \text{falls } A' \in \mathcal{A} \\ 0 & \text{sonst} \end{cases} \quad (7)$$

Verifizieren wir, dass μ' tatsächlich aus $rates(\mathbf{R}(s', \cdot))$ ist:

Zunächst stellen wir fest, dass wegen der Vorbedingung $s\mathcal{R}s'$ und aufgrund der Definition der Relation \mathcal{R} gilt, dass $s' \in \mathcal{A}$. Außerdem gilt für $u \in \bar{S}$ wegen $\mathbf{R}^l(s', u) = 0 = \mathbf{R}^u(s', u)$ und $\mu'(u) = 0$, dass $\mu'(u) \in rates(\mathbf{R}(s', u))$. Es bleibt als nicht trivialer Fall lediglich $\mu'(A')$ mit $A' \in \mathcal{A}$ übrig. Wir zeigen nun, dass $\mathbf{R}^l(A, A') \leq \mu'(A') \leq \mathbf{R}^u(A, A')$ für $A, A' \in \mathcal{A}$:

$$\begin{aligned}
\mathbf{R}^l(A, A') &= \min_{\hat{s} \in \mathcal{A}} \mathbf{R}^l(\hat{s}, A') && \text{(Def. 25)} \\
&\leq \mathbf{R}^l(s, A') = \sum_{s' \in \mathcal{A}'} \mathbf{R}^l(s, s') \\
&\leq \sum_{s' \in \mathcal{A}'} \mu(s') && \text{(wegen } \mu \in rates(\mathbf{R}(s, \cdot))) \\
&= \mu'(A') && \text{(wegen 7)} \\
&\leq \sum_{s' \in \mathcal{A}'} \mathbf{R}^u(s, s') = \mathbf{R}^u(s, A') && \text{(wegen } \mu' \in rates(\mathbf{R}(s', \cdot))) \\
&\leq \max_{\hat{s} \in \mathcal{A}} \mathbf{R}^u(\hat{s}, A') \\
&= \mathbf{R}^u(A, A') && \text{(Def. 25)}
\end{aligned}$$

Damit gilt also $\mu' \in rates(\mathbf{R}(s', \cdot))$.

- 2.2. Wir ziehen die Bedingung 2.2 vor, da wir sie für 2.1 verwenden wollen, um die Beobachtungen von Ratenvektoren auf Wahrscheinlichkeitsverteilungen zu übertragen. Dafür zeigen wir, dass $E(s) = E(s')$ bezüglich der Ratenvektoren μ und μ' gilt:

$$\begin{aligned}
E(s) &= \sum_{u \in S} \mu(u) && \text{(wegen } \mu \in rates(\mathbf{R}(s, \cdot))) \\
&= \sum_{u \in \bar{S}} \mu(u) \\
&= \sum_{u \in \mathcal{A}', A' \in \mathcal{A}} \mu(u) \\
&= \sum_{A' \in \mathcal{A}} \mu'(A') && \text{(wegen 7)} \\
&= \sum_{A' \in \mathcal{A}} \mu'(A') + \sum_{v \in \bar{S}} \mu'(v) && \text{(wegen 7)} \\
&= \sum_{v \in S} \mu'(v) \\
&= E(s') && \text{(wegen } \mu' \in rates(\mathbf{R}(s', \cdot)))
\end{aligned}$$

Da für μ und μ' also $E(s) = E(s')$ gilt, erhalten wir aus (7) eine entsprechende Definition für $\bar{\mu}'$ bezüglich $\bar{\mu}$:

$$\bar{\mu}'(A') = \begin{cases} \sum_{u \in \mathcal{A}'} \bar{\mu}(u) & \text{falls } A' \in \mathcal{A} \\ 0 & \text{falls } A' \in \bar{S} \end{cases} \quad (8)$$

Dass die beiden Definitionen äquivalent sind, folgt aus:

$$\begin{aligned}
&\bar{\mu}'(A') = \sum_{u \in \mathcal{A}'} \bar{\mu}(u) \\
gdw. \quad E(s') \cdot \bar{\mu}'(A') &= E(s') \cdot \sum_{u \in \mathcal{A}'} \bar{\mu}(u) && \text{(wegen } E(s) = (E(s'))) \\
gdw. \quad E(s') \cdot \bar{\mu}'(A') &= E(s) \cdot \sum_{u \in \mathcal{A}'} \bar{\mu}(u) \\
gdw. \quad E(s') \cdot \bar{\mu}'(A') &= \sum_{u \in \mathcal{A}'} (E(s) \cdot \bar{\mu}(u)) \\
gdw. \quad \mu'(A') &= \sum_{u \in \mathcal{A}'} \mu(u) && \text{(wegen } \mu(s) = E(s) \cdot \bar{\mu}(s))
\end{aligned}$$

Da $\bar{\mu}$ eine Verteilung ist und außerdem gilt

$$\sum_{s' \in S} \bar{\mu}'(s') = \sum_{A' \in \mathcal{A}} \bar{\mu}'(A') = \sum_{A' \in \mathcal{A}, s' \in A'} \bar{\mu}(s') = \sum_{s' \in S} \bar{\mu}(s') = 1$$

wissen wir, dass auch $\bar{\mu}'$ eine Verteilung ist. Diese Verteilungen werden verwendet, um den ersten Teil der zweiten Bedingung zu zeigen.

2.1. Nun müssen wir eine Gewichtung Δ finden, die den gegebenen Anforderungen gerecht wird. Dazu definieren wir:

$$\Delta(u, A') = \begin{cases} \bar{\mu}(u) & \text{falls } (u, A') \in \mathcal{R} \\ 0 & \text{sonst} \end{cases} \quad (9)$$

Die Bedingung 2.1.1 ist für dieses Δ trivialerweise erfüllt. Für Bedingungen 2.1.2 und 2.1.3 greifen wir wieder auf die Tatsache zurück, dass \mathcal{R} von der Partitionierung \mathcal{A} vollständig bestimmt wird. Zu Bedingung 2.1.2:

$$\begin{aligned} \Delta(u, S) &= \sum_{s \in S} \Delta(u, s) \\ &= \sum_{s \in \bar{S}} \Delta(u, s) + \sum_{A \in \mathcal{A}} \Delta(u, A) \\ &= \Delta(u, A_u) && \text{(für } A_u \in \mathcal{A} \text{ mit } u \in A_u) \\ &= \bar{\mu}(u) && \text{(wegen 9)} \end{aligned}$$

Die Bedingung 2.1.3 ist für $u \in \bar{S}$ trivialerweise erfüllt, da $\Delta(S, u) = 0 = \bar{\mu}'(u)$. Es bleibt noch zu zeigen, dass die Bedingung auch für $A' \in \mathcal{A}$ erfüllt ist:

$$\begin{aligned} \Delta(S, A') &= \sum_{u \in S} \Delta(u, A') \\ &= \sum_{u \in A'} \Delta(u, A') \\ &= \sum_{u \in A'} \bar{\mu}(u) && \text{(wegen 9)} \\ &= \bar{\mu}'(A') && \text{(wegen 8)} \end{aligned}$$

Damit ist gezeigt, dass der Teil 2.1 der Definition erfüllt ist. Wir haben also gezeigt, dass alle Bedingungen einer probabilistischen Simulation nach Definition 24 erfüllt sind.

□

Sehen wir uns ein Beispiel für die Anwendung der Abstraktion A an.

Beispiel 15 (Schwachstellen der Abstraktion A). Gegeben sei die zeitstetige Markov-Kette \mathcal{M} aus Abbildung 21 mit Zustandsmenge $S = \{s_0, s_1, s_2, s'_0, s'_1, s'_2\}$.

Wir führen nun die Abstraktion durch, indem wir jeweils s_i und s'_i zu \bar{s}_i zusammenfassen. Daraus erhalten wir die abstrakte zeitstetige Markov-Kette in der Abbildung unten.

Man sieht, dass in der ursprünglichen CTMC nur die Wahrscheinlichkeiten $\frac{1}{3}$ und $\frac{2}{3}$ vorkommen. Die Intervalle der Wahrscheinlichkeiten $[\frac{1}{5}, \frac{1}{2}]$ und $[\frac{1}{2}, \frac{4}{5}]$ sind also unnötig groß. Wir wollen versuchen, diese Beobachtung in eine zweite Abstraktion einfließen zu lassen.

—

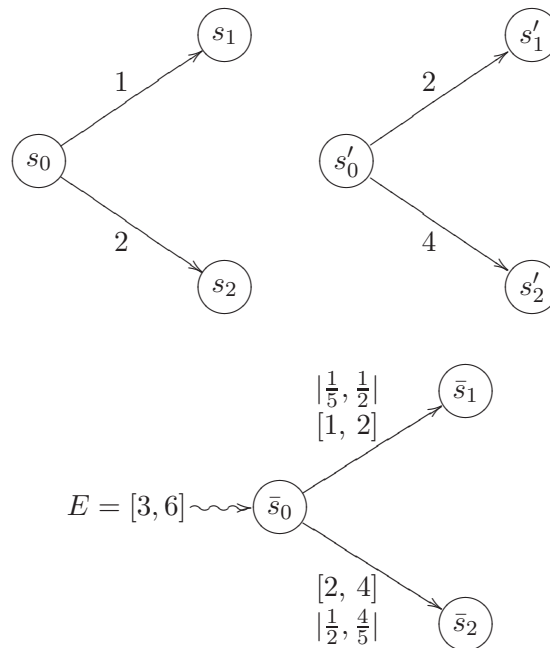


Abbildung 21: Abstraktion A

3.4.2 Abstraktion B

Die Ratendarstellung der zeitstetigen Markov-Ketten kann man als effiziente Darstellung der Wahrscheinlichkeiten und der zugehörigen Exitrate auffassen. Sieht man es auf diese Weise, sind zeitstetige Markov-Ketten eine Erweiterung der zeitdiskreten Markov-Ketten bei denen eine durchschnittliche Verweildauer $E(s)^{-1}$ für jeden Zustand eingeführt wurde. In unserem zweiten Ansatz, zeitstetige Markov-Ketten zu abstrahieren, wollen wir uns auf diese Sichtweise stützen.

Anstatt die Ratenintervalle anhand der ursprünglichen Markov-Kette zu generieren und anschließend die Wahrscheinlichkeitsintervalle und Exitratenintervalle so zu wählen, dass sie zu den Ratenintervallen passen, gehen wir jetzt umgekehrt vor. Wir bestimmen die Grenzen der Wahrscheinlichkeitsintervalle eines Makrozustands nach den maximalen bzw. minimalen Werten der zugehörigen Zustände der ursprünglichen Markov-Kette. Die Grenzen des Exitratenintervalls eines Makrozustands erhält man auf die gleiche Weise, abgesehen davon, dass bei den Wahrscheinlichkeiten ausgeschlossen werden muss, dass die Intervallgrenzen größer 1 werden.

Wir untersuchen in folgendem Beispiel, warum bei der Bildung der Wahrscheinlichkeitsintervalle auf maximale Wahrscheinlichkeiten 1 geprüft werden muss:

Beispiel 16 (Wahrscheinlichkeitsintervalle ohne Überprüfung). Bestimmen wir wie oben beschrieben die Abstraktion B der folgenden Markov-Kette durch Zusammenfassen von Zuständen s_1 und s_2 zu \bar{s}_1 ohne eine Überprüfung der Wahrscheinlichkeitsintervalle, so erhalten wir hier *keine* gültige ACTMC:

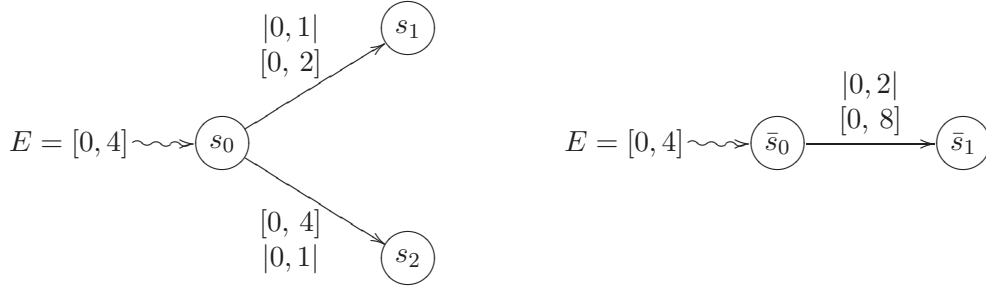


Abbildung 22: Abstraktion B (keine Überprüfung der Wahrscheinlichkeitswerte)

Wird dagegen das Wahrscheinlichkeitsintervall auf $[0, 1]$ verkleinert, so erhalten wir die gewünschte abstrakte Markov-Kette mit dem Ratenintervall $[0, 4]$.

Definition 26 (Abstraktion B). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetigen Markov-Kette. Eine Partitionierung $\mathcal{A} = \{A_1, \dots, A_n\}$ des Zustandsraumes von \mathcal{M} induziert eine Abstraktion $abstractB(\mathcal{M}, \mathcal{A}) = (\tilde{S}, \tilde{\mathbf{R}}^I, \tilde{\mathbf{P}}^I, \tilde{E}^I, \tilde{L})$, mit

- $\tilde{S} = \mathcal{A}$ ist die Menge der abstrakten Zustände,
- $\tilde{\mathbf{R}}^l(A_i, A_j) = \tilde{\mathbf{P}}^l(A_i, A_j) \cdot \tilde{E}^l(A_i)$,
- $\tilde{\mathbf{R}}^u(A_i, A_j) = \tilde{\mathbf{P}}^u(A_i, A_j) \cdot \tilde{E}^u(A_i)$,
- $\tilde{\mathbf{P}}^l(A_i, A_j) = \min\{1, \min_{s \in A_i} \mathbf{P}^l(s, A_j)\}$ ist die untere Schranke für die *Wahrscheinlichkeit* von A_i nach A_j zu wechseln,
- $\tilde{\mathbf{P}}^u(A_i, A_j) = \min\{1, \max_{s \in A_i} \mathbf{P}^u(s, A_j)\}$ ist die obere Schranke für die *Wahrscheinlichkeit* von A_i nach A_j zu wechseln,
- $\tilde{E}^l(A_i) = \min_{s \in A_i} E^l(s)$ ist die untere Schranke für die Exitrate des Makro-Zustands A_i .
- $\tilde{E}^u(A_i) = \max_{s \in A_i} E^u(s)$ ist die obere Schranke für die Exitrate des Makro-Zustands A_i .
- $\tilde{L}(A_i, a) = \begin{cases} \top & \text{falls für alle } s \in A_i \text{ gilt, dass } L(s, a) = \top \\ \perp & \text{falls für alle } s \in A_i \text{ gilt, dass } L(s, a) = \perp \\ ? & \text{sonst} \end{cases}$

Hierbei ist zu beachten, dass die $\tilde{\mathbf{R}}^I$ so gewählt wurden, dass es keine Einschränkung in Bezug auf die Wahl der Wahrscheinlichkeitsverteilungen und Exitraten durch einen Scheduler gibt, ähnlich wie bei $\tilde{\mathbf{P}}^I$ in Abstraktion A.

Auch für Abstraktion B müssen wir zeigen, dass die Zustände der ursprünglichen Markov-Kette durch die Makro-Zustände der Abstraktion simuliert werden. Der Beweis ist zu dem bezüglich Abstraktion A jedoch recht ähnlich.

Satz 3 (Simulation durch Makro-Zustände nach Abstraktion B).

Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette, dann gilt für die durch Partitionierung \mathcal{A} induzierte Abstraktion $abstractB(\mathcal{M}, \mathcal{A})$:

$$s \in A \Rightarrow s \preceq A \text{ für alle } s \in S \text{ und } A \in \mathcal{A}.$$

Beweis. Sei $\mathcal{M} = (\bar{S}, \bar{\mathbf{R}}^I, \bar{\mathbf{P}}^I, \bar{E}^I, \bar{L})$ eine zeitstetige Markov-Kette, \mathcal{A} eine Partitionierung und $abstractB(\mathcal{M}, \mathcal{A}) = (\mathcal{A}, \tilde{\mathbf{R}}^I, \tilde{\mathbf{P}}^I, \tilde{E}^I, \tilde{L})$ eine Abstraktion von \mathcal{M} . Die Relation \mathcal{R} wird wieder definiert als $\mathcal{R} = \{(s, A) \mid s \in A, A \in \mathcal{A}\}$. Um zu zeigen, dass die Bedingungen aus Definition 24 erfüllt sind, betrachten wir die Vereinigung $\mathcal{M} \uplus abstractB(\mathcal{M}, \mathcal{A}) = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$.

Für die erste Bedingung verweisen wir auf den Beweis von Satz 2, da die Bedingung sich im Vergleich zu Abstraktion A nicht verändert hat.

Für die zweite Bedingung legen wir wieder genau wie im Beweis zu Satz 2 zunächst für jedes $\bar{\mu} \in distr(\mathbf{P}(s, \cdot))$, ein passendes $\bar{\mu}' \in distr(\mathbf{P}(s', \cdot))$ fest (vgl. (8)):

$$\bar{\mu}'(A') = \begin{cases} \sum_{u \in A'} \bar{\mu}(u) & \text{falls } A' \in \mathcal{A} \\ 0 & \text{falls } A' \in \bar{S} \end{cases}$$

In Satz 2 wurde auch bereits gezeigt, dass $\bar{\mu}$ und $\bar{\mu}'$ Verteilung sind. Insbesondere ist zu beachten, dass wir in diesem Fall nicht den Umweg über die Raten gehen müssen, da die probabilistische Simulation nur über Wahrscheinlichkeiten und Exitraten definiert wurde.

Wir ziehen nun wieder die Bedingung 2.2 vor. Für Zustände s, s' mit $s \preceq s'$, folgt ihre Gültigkeit direkt aus der Definition 26, da die Exitrate bezüglich s' aus dem Intervall $[E^l(s'), E^u(s')]$ gewählt werden kann:

$$\begin{aligned} E^l(s') &= \min_{\hat{s} \in A} E^l(\hat{s}) && \text{(Def. 26)} \\ &\leq E^l(s) \\ &\leq E^u(s) \\ &\leq \max_{\hat{s} \in A} E^u(\hat{s}) \\ &= E^u(s') && \text{(Def. 26)} \end{aligned}$$

Das heißt, dass die Exitratenintervalle von Makro-Zuständen Obermengen der Exitratenintervalle der zugehörigen ursprünglichen Zustände sind. Bei jeder Wahl einer Exitrate $E \in [E^l(s), E^u(s)]$ in der ursprünglichen Markov-Kette kann also in der abstrakten Markov-Kette im Zustand s' dieselbe Wahl getroffen werden.

Den ersten Teil der zweiten Bedingung kann man wieder wie im Beweis von Satz 2 zeigen. Da dort ebenfalls mit Wahrscheinlichkeiten gearbeitet wurde, sind keinerlei Anpassungen notwendig.

Wir haben damit gezeigt, dass alle Bedingungen einer probabilistischen Simulation nach Definition 24 auch für Abstraktion B erfüllt sind.

□

Beispiel 17 (Stärken der Abstraktion B). Betrachten wir die Abstraktion B für das Negativ-Beispiel von Abstraktion A, so stellen wir fest, dass das gewünschte Resultat erreicht wurde. Die Wahrscheinlichkeitsintervalle sind auf $[\frac{1}{3}, \frac{1}{3}]$ und $[\frac{2}{3}, \frac{2}{3}]$ zusammengeschrumpft, während die übrigen Werte gleichgeblieben sind.

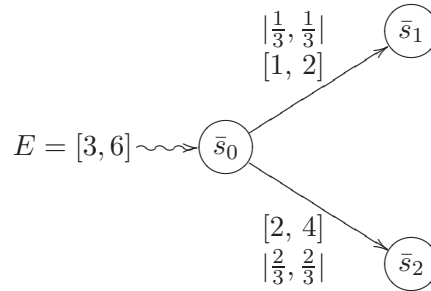


Abbildung 23: Abstraktion B

Beispiel 18 (Schwachstellen der Abstraktion B). Leider müssen wir feststellen, dass Abstraktion B keineswegs immer besser ist als Abstraktion A. Dazu betrachten wir die zeitstetige Markov-Kette in Abbildung 24 oben.

Wir bilden die Abstraktionen A (links) und B (rechts) durch Zusammenfassen der Zustände s_0 und s'_0 zu \bar{s}_0 :

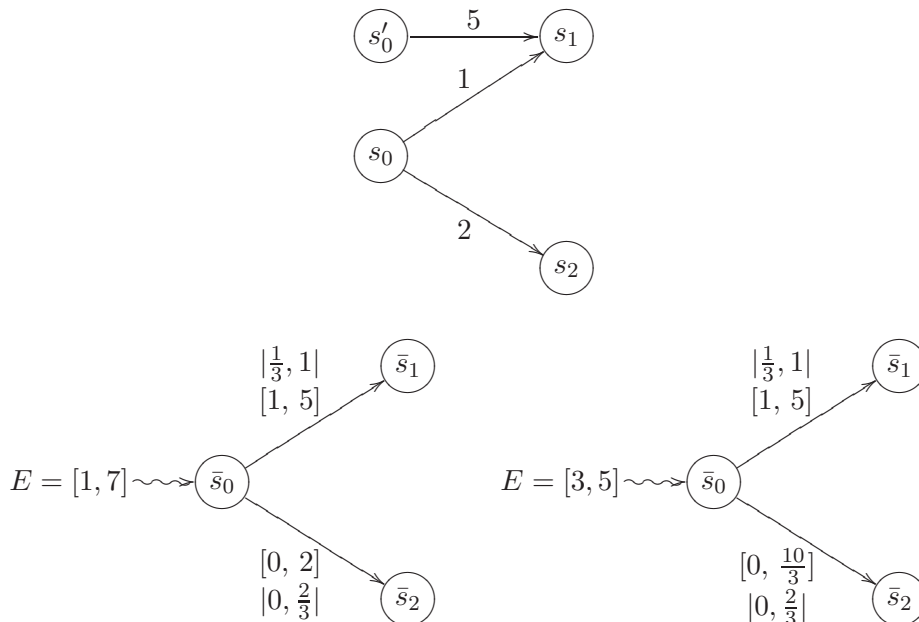


Abbildung 24: Schwächen der Abstraktion B (rechts)

Zwar hat die Abstraktion A das größere Exitratenintervall, bei den Ratenintervallen schneidet jedoch Abstraktion B schlechter ab, da $[0, \frac{10}{3}] \supseteq [0, 2]$. In einer dritten

und letzten Abstraktion wollen wir versuchen die Vorteile der Abstraktionen A und B zu vereinigen.

3.4.3 Abstraktion C

Wir haben in zwei Beispielen gesehen, dass bei abstrakten Markov-Ketten trotz aller Zusammenhänge die Raten, Wahrscheinlichkeiten und Exitraten getrennt voneinander betrachtet werden sollten. Es wurde schon gesagt, dass wir in Abstraktion C die Vorteile der beiden vorigen Abstraktionen vereinigen wollen. Es ist jedoch besser davon zu sprechen, die Nachteile der beiden Abstraktionen zu eliminieren, da das Ergebnis der Abstraktion $abstractC(\mathcal{M}, \mathcal{A})$ sich auch aus dem Schnitt²³ der Markov-Ketten $abstractA(\mathcal{M}, \mathcal{A})$ und $abstractB(\mathcal{M}, \mathcal{A})$ ergibt.

Definition 27 (Abstraktion C). Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette. Eine Partitionierung $\mathcal{A} = \{A_1, \dots, A_n\}$ des Zustandsraumes von \mathcal{M} induziert eine Abstraktion $abstractC(\mathcal{M}, \mathcal{A}) = (\tilde{S}, \tilde{\mathbf{R}}^I, \tilde{\mathbf{P}}^I, \tilde{E}^I, \tilde{L})$, mit \tilde{S} , $\tilde{\mathbf{R}}^I$ und \tilde{L} wie in Abstraktion A und $\tilde{\mathbf{P}}^I$ und \tilde{E}^I wie in Abstraktion B.

Die Zustandsmenge \tilde{S} und die Beschriftungsfunktion \tilde{L} sind in allen drei Abstraktionen gleich. Dies ist auch nicht verwunderlich, da es keinerlei Abhängigkeiten zu anderen Teilen der abstrakten Markov-Kette gibt.

Lemma 4 (Abstraktion C ist Schnitt aus Abstraktionen A und B). Sei \mathcal{M} eine zeitstetige Markov-Kette und $\mathcal{M}_A = abstractA(\mathcal{M}, \mathcal{A})$, $\mathcal{M}_B = abstractB(\mathcal{M}, \mathcal{A})$ und $\mathcal{M}_C = abstractC(\mathcal{M}, \mathcal{A})$ abstrakte zeitstetige Markov-Ketten. Dann gilt $\mathcal{M}_C \subseteq \mathcal{M}_A$ und $\mathcal{M}_C \subseteq \mathcal{M}_B$ und es existiert keine abstrakte zeitstetige Markov-Kette \mathcal{M}_D mit $\mathcal{M}_D \subseteq \mathcal{M}_A$, $\mathcal{M}_D \subseteq \mathcal{M}_B$ und $\mathcal{M}_C \subset \mathcal{M}_D$.

Wir schreiben dann auch $\mathcal{M}_C = \mathcal{M}_A \cap \mathcal{M}_B$.

Beweis. Seien $\mathcal{M}_X = (S, \mathbf{R}_X^I, \mathbf{P}_X^I, E_X^I, L)$ für $X \in \{A, B, C, D\}$ abstrakte zeitstetige Markov-Ketten, wobei $\mathcal{M}_A = abstractA(\mathcal{M}, \mathcal{A})$, $\mathcal{M}_B = abstractB(\mathcal{M}, \mathcal{A})$ und $\mathcal{M}_C = abstractC(\mathcal{M}, \mathcal{A})$ für eine zeitstetige Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$.

Da alle drei Abstraktionen dieselben Zustandsmengen und Beschriftungsfunktionen erzeugen, müssen wir also nur noch zeigen, dass aus den Raten-, Wahrscheinlichkeits- und Exitratenintervalle $\mathcal{M}_C \subseteq \mathcal{M}_A$ und $\mathcal{M}_C \subseteq \mathcal{M}_B$ folgt. Dazu betrachten wir die drei Fälle einzeln. In jedem der drei Teile werden wir über die Eigenschaften aus Korollar 1 argumentieren, dass es keine Rate, Wahrscheinlichkeit oder Exitrate geben darf, die nicht zu einer CTMC vervollständigt werden kann. Wäre beispielsweise die kleinste wählbare Rate r_{min} kleiner als das Produkt der dazu passenden kleinsten Wahrscheinlichkeit p_{min} und der kleinsten Exitrate e_{min} , so wäre es nicht möglich, eine Wahrscheinlichkeit $p \geq p_{min}$ und eine Exitrate $e \geq e_{min}$ zu finden, sodass die Gleichung $r_{min} = p \cdot e$ erfüllt wäre.

²³Im Sinne von $abstractC(\mathcal{M}, \mathcal{A}) \subseteq abstractA(\mathcal{M}, \mathcal{A})$ und $abstractC(\mathcal{M}, \mathcal{A}) \subseteq abstractB(\mathcal{M}, \mathcal{A})$.

- Da ein Scheduler bei der Wahl von $\mathbf{R}^l(s, A_j)$ sonst keine Wahrscheinlichkeiten und Exitraten wählen könnte, die zu einer korrekten CTMC führen, gilt:

$$\begin{aligned}
\mathbf{R}_C^l(A_i, A_j) &= \mathbf{R}_A^l(A_i, A_j) \\
&= \min_{s \in A_i} \mathbf{R}^l(s, A_j) \\
&\geq \min_{s \in A_i} (\mathbf{P}^l(s, A_j) \cdot E^l(s, A_j)) && \text{(Kor. 1.4)} \\
&\geq (\min_{s \in A_i} \mathbf{P}^l(s, A_j)) \cdot (\min_{s \in A_i} E^l(s)) \\
&= (\min\{1, \min_{s \in A_i} \mathbf{P}^l(s, A_j)\}) \cdot (\min_{s \in A_i} E^l(s)) && \text{(Kor. 2)} \\
&= \mathbf{P}_B^l(A_i, A_j) \cdot E_B^l(A_i) \\
&= \mathbf{R}_B^l(A_i, A_j)
\end{aligned}$$

In Abstraktion A ist die untere Intervallgrenze größer, also werden die Raten nach unten genauer abgeschätzt.

Ebenso gilt für $\mathbf{R}^u(s, A_j)$:

$$\begin{aligned}
\mathbf{R}_C^u(A_i, A_j) &= \mathbf{R}_A^u(A_i, A_j) \\
&= \max_{s \in A_i} \mathbf{R}^u(s, A_j) \\
&\leq \max_{s \in A_i} (\mathbf{P}^u(s, A_j) \cdot E^u(s, A_j)) && \text{(Kor. 1.4)} \\
&\leq (\max_{s \in A_i} \mathbf{P}^u(s, A_j)) \cdot (\max_{s \in A_i} E^u(s)) \\
&\leq (\min\{1, \max_{s \in A_i} \mathbf{P}^u(s, A_j)\}) \cdot (\max_{s \in A_i} E^u(s)) && \text{(Kor. 2)} \\
&= \mathbf{P}_B^u(A_i, A_j) \cdot E_B^u(A_i) \\
&= \mathbf{R}_B^u(A_i, A_j)
\end{aligned}$$

Zusammengenommen folgt also, dass $\mathbf{R}_C^I = \mathbf{R}_A^I \subseteq \mathbf{R}_B^I$. Es lässt sich also offensichtlich kein \mathbf{R}_D^I finden mit $\mathbf{R}_C^I \subset \mathbf{R}_D^I \subseteq \mathbf{R}_A^I \subseteq \mathbf{R}_B^I$.

- Wir unterscheiden nun zwischen den Fällen $E^l(s) = 0$ und $E^l(s) > 0$. Im Fall $E^l(s) = 0$ kann der Zustand s absorbierend sein, also müssen die Wahrscheinlichkeiten $\mathbf{P}_X^l(A_i, A_j)$ für $X \in \{A, B, C\}$ jeweils gleich 0 sein.

Betrachten wir nun den zweiten Fall mit $E^l(s) > 0$. Da ein Scheduler bei der Wahl von $\mathbf{P}^l(s, A_j)$ sonst keine Raten und Exitraten wählen könnte, die zu einer korrekten CTMC führen, gilt:

$$\begin{aligned}
\mathbf{P}_C^l(A_i, A_j) &= \mathbf{P}_B^l(A_i, A_j) \\
&= \min\{1, \min_{s \in A_i} \mathbf{P}^l(s, A_j)\} \\
&\geq \min\{1, \min_{s \in A_i} (\mathbf{R}^l(s, A_j)/E^u(s, A_j))\} \quad (\text{Kor. 1.5}) \\
&\geq \min\{1, (\min_{s \in A_i} \mathbf{R}^l(s, A_j)) \\
&\quad /((\min_{s \in A_i} \mathbf{R}^l(s, A_j)) \\
&\quad + (\sum_{A_k \neq A_j} \max_{s \in A_i} \mathbf{R}^u(s, A_k)))\} \\
&= \min\{1, \mathbf{R}_A^l(A_i, A_j)/(\mathbf{R}_A^l(A_i, A_j) \\
&\quad + \sum_{A_k \neq A_j} \mathbf{R}_A^u(A_i, A_k))\} \\
&= \min\{1, \mathbf{P}_A^l(A_i, A_j)\} \\
&= \mathbf{P}_A^l(A_i, A_j) \quad (\text{Kor. 2})
\end{aligned}$$

Für $\mathbf{P}^u(s, A_j)$ müssen wir zwischen den Fällen $E^u(s) = 0$ und $E^u(s) > 0$ unterscheiden. Im Fall $E^u(s) = 0$ muss der Zustand s absorbierend sein, also müssen auch die Wahrscheinlichkeiten $\mathbf{P}_X^u(A_i, A_j)$ für $X \in \{A, B, C\}$ jeweils gleich 0 sein.

Den zweiten Fall können wir ähnlich wie $\mathbf{P}^l(s, A_j)$ abhandeln:

$$\begin{aligned}
\mathbf{P}_C^u(A_i, A_j) &= \mathbf{P}_B^u(A_i, A_j) \\
&= \min\{1, \max_{s \in A_i} \mathbf{P}^u(s, A_j)\} \\
&\leq \min\{1, \max_{s \in A_i} (\mathbf{R}^u(s, A_j)/E^l(s, A_j))\} \quad (\text{Kor. 1.5}) \\
&\leq \min\{1, (\max_{s \in A_i} \mathbf{R}^u(s, A_j)) \\
&\quad /((\max_{s \in A_i} \mathbf{R}^u(s, A_j)) \\
&\quad + (\sum_{A_k \neq A_j} \min_{s \in A_i} \mathbf{R}^l(s, A_k)))\} \\
&= \min\{1, \mathbf{R}_A^u(A_i, A_j)/(\mathbf{R}_A^u(A_i, A_j) \\
&\quad + \sum_{A_k \neq A_j} \mathbf{R}_A^l(A_i, A_k))\} \\
&= \min\{1, \mathbf{P}_A^u(A_i, A_j)\} \\
&= \mathbf{P}_A^u(A_i, A_j) \quad (\text{Kor. 2})
\end{aligned}$$

Zusammengenommen folgt also, dass $\mathbf{P}_C^I = \mathbf{P}_B^I \subseteq \mathbf{P}_A^I$. Auch hier lässt sich kein \mathbf{P}_D^I finden mit $\mathbf{P}_C^I \subset \mathbf{P}_D^I \subseteq \mathbf{P}_B^I \subseteq \mathbf{P}_A^I$.

- Da ein Scheduler bei der Wahl von $E^l(s)$ sonst keine Raten und Wahrscheinlichkeiten wählen könnte, die zu einer korrekten CTMC führen, gilt:

$$\begin{aligned}
E_C^l(A_i) &= E_B^l(A_i) \\
&= \min_{s \in A_i} E^l(s) \\
&\geq \min_{s \in A_i} \mathbf{R}^l(s, \mathcal{A}) \quad (\text{Kor. 1.2}) \\
&= \mathbf{R}_A^l(A_i, A_j) \\
&= E_A^l(A_i)
\end{aligned}$$

Ebenso gilt für $E^u(s)$:

$$\begin{aligned}
E_C^l(A_i) &= E_B^u(A_i) \\
&= \max_{s \in A_i} E^u(s) \\
&\leq \max_{s \in A_i} \mathbf{R}^u(s, \mathcal{A}) \quad (\text{Kor. 1.2}) \\
&= \mathbf{R}_A^u(A_i, A_j) \\
&= E_A^u(A_i)
\end{aligned}$$

Zusammengenommen folgt also, dass $E_C^I = E_B^I \subseteq E_A^I$. Wie auch in den anderen beiden Fällen, lässt sich hier kein E_D^I finden mit $E_C^I \subseteq E_D^I \subseteq E_B^I \subseteq E_A^I$.

Damit ist die Behauptung bewiesen, die Abstraktion C ergebe einen Schnitt aus den Ergebnissen der Abstraktionen A und B. □

Der folgende Satz zur Simulation durch Makro-Zustände nach Abstraktion C könnte zwar auch auf ähnliche Weise wie die Sätze 2 und 3 gezeigt werden, wir werden dies jedoch hier über einen kürzeren Beweis zeigen, der die Tatsache ausnutzt, dass die Abstraktion den Schnitt der beiden vorigen Abstraktionen liefert.

Satz 4 (Simulation durch Makro-Zustände nach Abstraktion C).

Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette, dann gilt für die durch Partition \mathcal{A} induzierte Abstraktion $abstractC(\mathcal{M}, \mathcal{A})$:

$$s \in A \Rightarrow s \preceq A \text{ für alle } s \in S \text{ und } A \in \mathcal{A}.$$

Beweis. Sei M_{sound} die Menge aller CTMCs, die von der ursprünglichen Markov-Kette induziert werden. Sei außerdem $M_{abstract}$ die Menge der CTMCs, die von einer abstrakten Markov-Kette $abstract(\mathcal{M}, \mathcal{A})$ induziert werden. Für alle korrekten Abstraktionen muss $M_{sound} \subseteq M_{abstract}$ gelten.

Wir haben in Satz 2 und 3 bewiesen, dass Abstraktionen A und B korrekt sind, indem wir gezeigt haben, dass in den abstrakten Markov-Ketten ein Scheduler bezüglich der ursprünglichen Markov-Kette simuliert werden kann. Bilden wir den Schnitt über die von Abstraktionen A und B induzierten CTMCs, so erhalten wir wiederum eine korrekte Abstraktion, da $M_{sound} \subseteq M_{abstractC}$:

$$\begin{aligned}
M_{abstractC} &= M_{abstractA} \cap M_{abstractB} \\
&= (M_{sound} \cup M_{abstractA}) \cap (M_{sound} \cup M_{abstractB}) \\
&= M_{sound} \cup (M_{abstractA} \cap M_{abstractB})
\end{aligned}$$

□

Beispiel 19 (Abstraktion C). Wie erwartet erhalten wir für die beiden Beispiele zu Abstraktionen A und B mit Abstraktion C feinere Markov-Ketten:

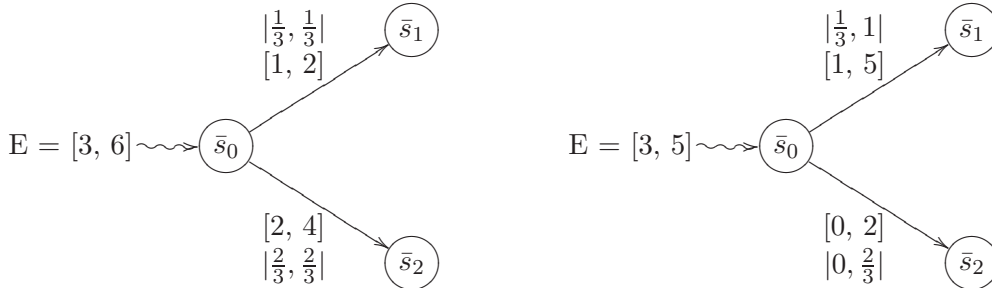


Abbildung 25: Abstraktion C

Beispiel 20 (Abstraktion bereinigter Markov-Ketten). Dass man bei der Anwendung von Abstraktion C auf eine bereinigte Markov-Kette nicht zwingendermaßen wieder eine bereinigte Markov-Kette erhält wollen wir uns am Beispiel der Abbildung 26 klar machen. Fasst man dort die Zustände s_1 und s'_1 zu \bar{s}_1 zusammen, so erhält man eine nicht bereinigte ACTMC mit einer zu hohen oberen Ratenintervallgrenze:

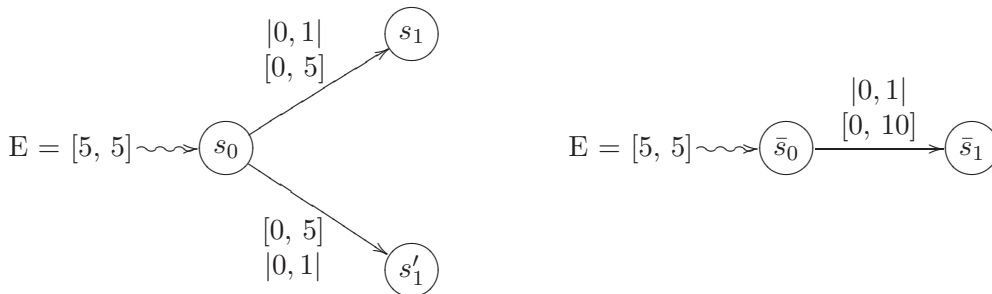


Abbildung 26: Abstraktion bereinigter Markov-Ketten

3.5 Uniformisierte abstrakte zeitstetige Markov-Ketten

Im weiteren Verlauf werden wir noch das Erreichbarkeitsproblem für abstrakte zeitstetige Markov-Ketten untersuchen. Im Grundlagenkapitel wurde dazu für den Fall der CTMCs die Technik der Uniformisierung eingeführt. Diese wollen wir abgewandelt später auch für ACTMCs verwenden, weswegen wir uns im Folgenden der Frage widmen, wie uniforme abstrakte Markov-Ketten aus nicht-uniformen Markov-Ketten erzeugt werden können. Zunächst benötigen wir jedoch noch die formale Definition für *uniforme abstrakte zeitstetige Markov-Ketten*.

Uniforme Markov-Ketten haben die spezielle Eigenschaft, dass die Exitrate für alle Zustände gleich ist. Diese Eigenschaft werden wir noch zur Behandlung des Erreichbarkeitsproblems ausnutzen. Da verschiedene Wahlmöglichkeiten bestehen würden, wenn wir bei den abstrakten zeitstetigen Markov-Ketten nur gleiche Intervalle bei den Exitraten fordern würden, definieren wir uniforme ACTMCs so, dass die Exitratenintervalle Punktintervalle sein müssen, also die Form $[e, e]$ haben.

Definition 28 (Uniforme abstrakte zeitstetige Markov-Kette). Eine abstrakte zeitstetige Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ heißt *uniform*, falls ein $E_{unif} \in \mathbb{R}_{\geq 0}$ für alle $s \in S$ existiert, sodass gilt $E^l(s) = E_{unif} = E^u(s)$.

Wir untersuchen nun, inwiefern sich die Abstraktion von uniformen Markov-Ketten zur Abstraktion im Allgemeinen unterscheidet.

Beispiel 21 (Abstraktion uniformer Markov-Ketten). Betrachten wir wieder die Markov-Kette aus Beispiel 18, diesmal jedoch uniformisieren wir sie zunächst nach der üblichen Methode für CTMCs mit Exitrate $E_{unif} = 5$ (Abbildung 27 links). Wie sich leicht nachprüfen läßt, erhält man für alle drei Abstraktionsmethoden, die wir im letzten Abschnitt behandelt haben, die abstrakte zeitstetige Markov-Kette rechts in der Abbildung.

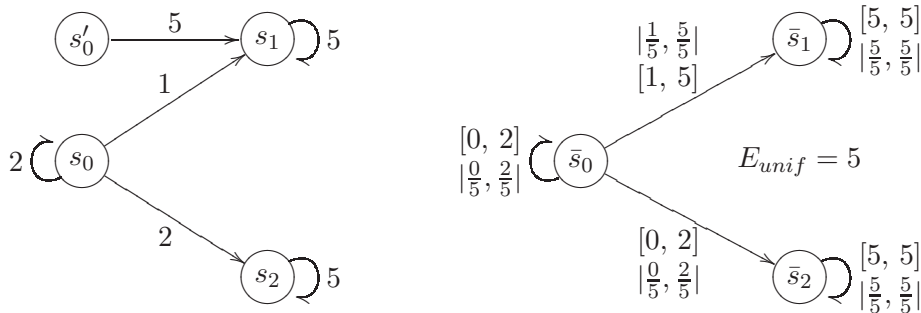


Abbildung 27: Abstraktion uniformer Markov-Ketten

Dieses Beispiel lässt erahnen, dass sich bei der Abstraktion uniformer Markov-Ketten durch die feste Exitrate, die Intervallgrenzen der Wahrscheinlichkeitsintervalle immer als Division der Ratenintervallgrenzen mit der Exitrate ergeben. Dass dies tatsächlich der Fall ist und sich daraus auch schließen lässt, dass die drei Abstraktionen immer dieselben Ergebnisse liefern, werden wir im folgenden Lemma zeigen.

Lemma 5 (Gleichheit der Abstraktionen bei uniformen Markov-Ketten). Sei \mathcal{M}' eine uniforme abstrakte zeitstetige Markov-Kette mit einer festen Exitrate für alle Zustände und \mathcal{A} eine Partitionierung des Zustandsraums. Dann gilt für die bereinigte uniforme CTMC $\mathcal{M} = cut(\mathcal{M}') = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$, dass:

$$abstractA(\mathcal{M}, \mathcal{A}) = abstractB(\mathcal{M}, \mathcal{A}) = abstractC(\mathcal{M}, \mathcal{A}).$$

Beweis. Wir verlangen, dass die gegebenen Markov-Kette mit dem *cut*-Operator von nicht wählbaren Kombinationen befreit wurde. Außerdem wissen wir, dass die Exitrate E_{unif} für alle Zustände dieselbe ist.

Da durch cut_{R_2} und cut_{P_2} alle Raten r und Wahrscheinlichkeiten p eliminiert wurden, welche die Gleichung $r = p \cdot E_{unif}$ nicht erfüllen, können die oberen und unteren Grenzen der Ratenintervalle und der Wahrscheinlichkeitsintervalle nur um den Faktor E_{unif} verschieden sein:

$$\mathbf{P}^\gamma(s, A_j) \cdot E_{unif} = \mathbf{R}^\gamma(s, A_j) \text{ für alle } s \in S, A_j \in \mathcal{A} \text{ und } \gamma \in \{l, u\}$$

Ansonsten gäbe es eine Rate oder eine Wahrscheinlichkeit, die nicht zu einer zulässigen Markov-Kette vervollständigt werden könnte. Damit können wir nun zeigen, dass die Raten-, Wahrscheinlichkeits- und Exitratenintervalle der drei Abstraktionen immer übereinstimmen:

- Für die Ratenintervalle gilt:

$$\begin{aligned} \mathbf{R}_C^l(A_i, A_j) &= \mathbf{R}_A^l(A_i, A_j) \\ &= \min_{s \in A_i} \mathbf{R}^l(s, A_j) \\ &= \min_{s \in A_i} (\mathbf{P}^l(s, A_j) \cdot E_{unif}) \\ &= (\min_{s \in A_i} \mathbf{P}^l(s, A_j)) \cdot E_{unif} \\ &= (\min\{1, \min_{s \in A_i} \mathbf{P}^l(s, A_j)\}) \cdot E_{unif} \quad (\text{Kor. 2}) \\ &= \mathbf{P}_B^l(A_i, A_j) \cdot E_B^l(A_i) \\ &= \mathbf{R}_B^l(A_i, A_j) \end{aligned}$$

$$\begin{aligned} \mathbf{R}_C^u(A_i, A_j) &= \mathbf{R}_A^u(A_i, A_j) \\ &= \max_{s \in A_i} \mathbf{R}^u(s, A_j) \\ &= \max_{s \in A_i} (\mathbf{P}^u(s, A_j) \cdot E_{unif}) \\ &= (\max_{s \in A_i} \mathbf{P}^u(s, A_j)) \cdot E_{unif} \\ &= (\min\{1, \max_{s \in A_i} \mathbf{P}^u(s, A_j)\}) \cdot E_{unif} \quad (\text{Kor. 2}) \\ &= \mathbf{P}_B^u(A_i, A_j) \cdot E_B^u(A_i) \\ &= \mathbf{R}_B^u(A_i, A_j) \end{aligned}$$

$$\Rightarrow \mathbf{R}_A^I = \mathbf{R}_B^I = \mathbf{R}_C^I$$

- Für die Wahrscheinlichkeitsintervalle gilt:

$$\begin{aligned} \mathbf{P}_A^l(A_i, A_j) &= \mathbf{R}_A^l(A_i, A_j) // (\mathbf{R}_A^l(A_i, A_j) + \sum_{A_k \neq A_j} \mathbf{R}_A^u(A_i, A_k)) \\ &= \mathbf{R}_B^l(A_i, A_j) // (\mathbf{R}_B^l(A_i, A_j) + \sum_{A_k \neq A_j} \mathbf{R}_B^u(A_i, A_k)) \\ &= \mathbf{P}_B^l(A_i, A_j) \\ &= \mathbf{P}_C^l(A_i, A_j) \end{aligned}$$

$$\begin{aligned}
\mathbf{P}_A^u(A_i, A_j) &= \mathbf{R}_A^u(A_i, A_j) // (\mathbf{R}_A^u(A_i, A_j) + \sum_{A_k \neq A_j} \mathbf{R}_A^l(A_i, A_k)) \\
&= \mathbf{R}_B^u(A_i, A_j) // (\mathbf{R}_B^u(A_i, A_j) + \sum_{A_k \neq A_j} \mathbf{R}_B^l(A_i, A_k)) \\
&= \mathbf{P}_B^u(A_i, A_j) \\
&= \mathbf{P}_C^u(A_i, A_j)
\end{aligned}$$

$$\Rightarrow \mathbf{P}_A^I = \mathbf{P}_B^I = \mathbf{P}_C^I$$

- Für die Exitratenintervalle gilt:

$$E_A^l(A_i) = \mathbf{R}_A^l(A_i, S) = \mathbf{R}_B^l(A_i, S) = E_B^l(A_i) = E_C^l(A_i)$$

$$E_A^u(A_i) = \mathbf{R}_A^u(A_i, S) = \mathbf{R}_B^u(A_i, S) = E_B^u(A_i) = E_C^u(A_i)$$

$$\Rightarrow E_A^I = E_B^I = E_C^I$$

Die Zustandsmengen und Beschriftungsfunktionen sind im Allgemeinen für alle drei Abstraktionen gleich, folglich liefern die drei Abstraktionen für bereinigte uniforme ACTMCs die gleichen Resultate.

□

Widmen wir uns nun dem eigentlichen Grund für dieses Unterkapitel, nämlich der Frage nach der Uniformisierung bezüglich abstrakter zeitstetiger Markov-Ketten.

Beispiel 22 (Uniformisierung abstrakter zeitstetiger Markov-Ketten). Betrachten wir die Markov-Kette aus Beispiel 15. In einem ersten Versuch führen wir zunächst die Abstraktion C durch. Auf naive Weise führen wir dann eine Uniformisierung mit Exitrate $E_{unif} = 6$ durch, indem entsprechende *self-loops* hinzugefügt werden. Diese Herangehensweise scheitert jedoch daran, dass der *self-loop* nur noch ein Wahrscheinlichkeitsintervall $[0, 0]$ erhalten könnte, was aber im Widerspruch zur Gleichung $r = p \cdot e$ steht. Durch Ausführung des *cut* würde das Ratenintervall des *self-loop* also sofort wieder auf $[0, 0]$ reduziert werden, und damit komplett entfernt. Durch das Hinzufügen von neuen Ratenintervallen über die *self-loops*, werden die Wahrscheinlichkeitsintervalle also offensichtlich ungültig. Wir erhalten jedoch abgesehen davon eine *uniforme* abstrakte Markov-Kette, also eine ACTMC, dessen sämtlichen induzierten CTMCs uniform sind. Wie wir im Beweis zu Lemma 5 schon gesehen haben, gilt für uniforme ACTMCs, dass sich die Intervallgrenzen der Wahrscheinlichkeiten als Division der Ratenintervallgrenzen durch die Exitrate ergeben. Berechnen wir auf diese Weise die Wahrscheinlichkeitsintervalle neu, so erhalten wir die uniforme abstrakte zeitstetige Markov-Kette in Abbildung 28 rechts.

Dies deckt sich hier mit einer zweiten Variante, zuerst die Uniformisierung der zeitstetigen Markov-Kette durchzuführen und anschließend erst die Abstraktion. Dafür müssen wir natürlich von einer CTMC ausgehen. Die Uniformisierung der ursprünglichen Markov-Kette aus dem Beispiel mit Exitrate $E_{unif} = 6$ ergibt die folgende Markov-Kette in Abbildung 29.

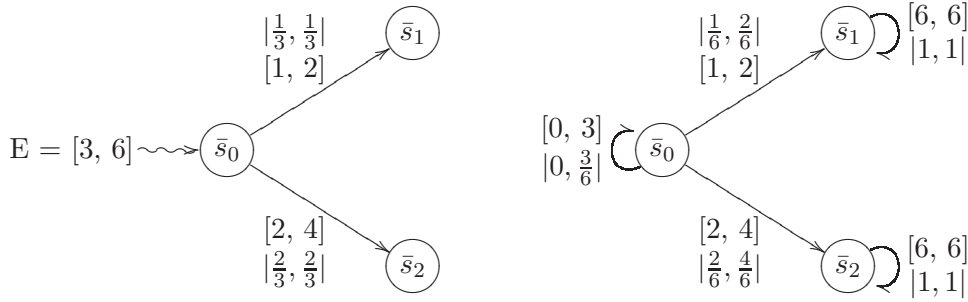


Abbildung 28: Uniformisierung von ACTMCs

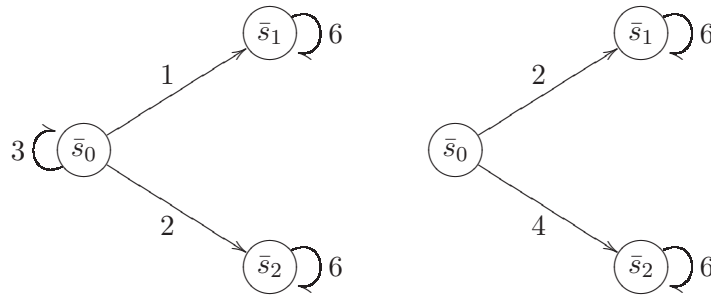


Abbildung 29: Uniformisierte CTMC

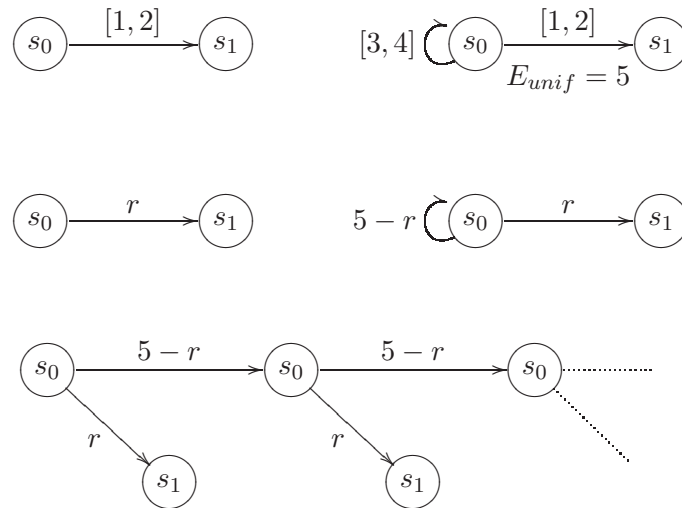
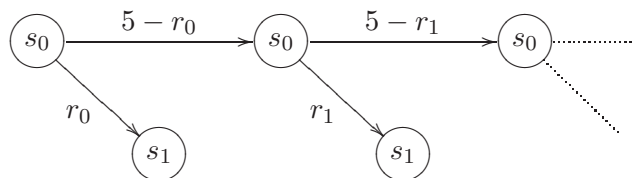
Bilden wir nun die Abstraktion nach Beispiel 15, so erhalten wir wieder dieselbe uniforme abstrakte zeitstetige Markov-Kette wie bei der ersten Methode. Dass die beiden Methoden sich aber dennoch unterscheiden, werden wir in einem weiteren Beispiel sehen.

Beispiel 23 (Probleme bei der Uniformisierung von ACTMCs). Für das Beispiel 22 konnten wir zwar sehen, dass dort die Reihenfolge von Uniformisierung und Abstraktion nicht von Bedeutung ist, im Allgemeinen ist dies jedoch nicht der Fall. Dazu untersuchen wir die Markov-Kette oben links in Abbildung 30 bezüglich des *Transient*-Verhaltens bei Uniformisierung mit Exitrate $E_{unif} = 5$ (oben rechts) etwas genauer.

Zunächst bemerken wir, dass die gegebene abstrakte Markov-Kette eine unendliche Menge von Markov-Ketten \mathcal{M}_r induziert mit $r \in [1, 2]$ (mitte links). Durch Uniformisierung mit Exitrate E_{unif} ergeben sich daraus jeweils Markov-Ketten wie sie mittig rechts abgebildet sind. Eine anschließende Abwicklung ergibt dann die Darstellung unten in der Abbildung.

Soweit zu den *uniformisierten induzierten* Markov-Ketten. Betrachten wir nun die induzierten CTMCs der uniformisierten abstrakten Markov-Kette oben rechts. Da der Scheduler in jedem Schritt einen Ratenvektor wählen kann, ergeben sich im Allgemeinen die CTMCs aus Abbildung 31 mit $r_i \in [1, 2]$ für $i \in \mathbb{N}$.

Beispielsweise durch die Wahl von $r_i = 1$ für $0 \leq i < 10$ und $r_i = 2$ für $i \geq 10$ erhält

Abbildung 30: *Transient*-Verhalten in uniformisierten induzierten Markov-KettenAbbildung 31: *Transient*-Verhalten in induzierten Markov-Ketten der UACTMCs

man also eine Markov-Kette, für die es bezüglich des *Transient*-Verhaltens keine Entsprechung in der Menge der uniformisierten induzierten Markov-Ketten gibt.

Über die Poissonverteilung $\varphi(E \cdot t, n)$ bestimmen wir, wie groß die Wahrscheinlichkeit für n Transitionen in t Zeiteinheiten bei einer gegebenen Exitrate E ist. Da wir in unserem Beispiel in den ersten neun Transitionen eines Pfads auf jeden Fall ein anderes Verhalten gegeben haben als auf den längeren Pfaden, bei denen zu Beginn zehn Mal oder öfter der *self-loop* gewählt wurde, wollen wir exemplarisch für die Zeitdauern $\frac{1}{5}$, 2 und 10 die Wahrscheinlichkeit für maximal neun Transitionen bestimmen:

$$\begin{aligned}\sum_{n=0}^9 \varphi(1, n) &\approx 0.99999988857 \\ \sum_{n=0}^9 \varphi(10, n) &\approx 0.45792971447 \\ \sum_{n=0}^9 \varphi(50, n) &\approx 0.00000000750\end{aligned}$$

Bei Zeitdauer $\frac{1}{5}$ tragen Pfade mit mehr als neun Transitionen nahezu gar nicht zum Ergebnis bei. Damit ist das Verhalten ähnlich zur dem der uniformisierten induzierten Markov-Kette \mathcal{M}_1 . Wie wir sehen können, wird das Verhalten der Markov-Kette bei zunehmender Zeitdauer immer stärker von längeren Pfaden und damit auch vom hinteren Teil geprägt. Dort ist die Wahrscheinlichkeit in den einzelnen Zuständen für einen Übergang zum Zustand s_1 größer als am Anfang. Daher entfernt sich das *Transient*-Verhalten zunehmend von dem der Markov-Kette \mathcal{M}_1 .

Um diesem Problem aus dem Weg zu gehen, werden wir uns bei der Uniformisierung darauf beschränken, diese auf CTMCs anzuwenden und die resultierenden uniformisierten Markov-Ketten zu abstrahieren. Dazu ist noch zu zeigen, dass Abstraktionen uniformer Markov-Ketten ebenfalls uniform sind.

Lemma 6. Sei \mathcal{M} eine zeitstetige Markov-Kette und \mathcal{A} eine Partitionierung des Zustandsraums von \mathcal{M} , dann ist die *uniformisierte abstrakte* zeitstetige Markov-Kette $abstractC(abstr(unif(\mathcal{M})), \mathcal{A})$ uniform²⁴.

Beweis. Gegeben sei $\mathcal{M} = (S, \mathbf{R}, L)$ und die Partitionierung \mathcal{A} . Damit erhalten wir als uniformisierte zeitstetige Markov-Kette:

$$\mathcal{M}_{unif} = unif(\mathcal{M}) = (S, \mathbf{R}_{unif}, L)$$

Für die Exitraten der uniformisierten Markov-Kette gilt per Definition von *unif*, dass $E_{unif}(s) = \mathbf{R}_{unif}(s, S) = E_{unif}$ für alle $s \in S$ und ein passendes $E_{unif} \in \mathbb{R}_{\geq 0}$. Da *unif* als Ergebnis wieder eine konkrete zeitstetige Markov-Kette liefert, wandeln wir diese mit *abstr* zu einer abstrakten zeitstetigen Markov-Kette um.

Bei der Abstraktion der ACTMC $abstr(\mathcal{M}_{unif})$ bezüglich der Partitionierung \mathcal{A} werden nun die Minima und Maxima der oberen und unteren Exitratenintervallgrenzen

²⁴Die Funktion *abstr* wurde in Abschnitt 3.1 so definiert, dass sie für eine gegebene zeitstetige Markov-Kette \mathcal{M} diejenige abstrakte zeitstetige Markov-Kette liefert, die alleine die gegebene CTMC \mathcal{M} induziert.

von Zuständen jeder Partition bestimmt. Da alle Exitratenintervalle von \mathcal{M}_{unif} die Form $[E_{unif}, E_{unif}]$ haben, erhält man für die Exitraten der abstrakten uniformisierten Markov-Kette ebenfalls die Intervalle $[E_{unif}, E_{unif}]$ für alle $s \in \mathcal{A}$. Damit ist gezeigt, dass uniformisierte abstrakte zeitstetige Markov-Ketten uniform sind. \square

Da wir bei uniformisierten abstrakten zeitstetigen Markov-Ketten (oder UACTMCs) in jedem Zustand ein festes Exitratenintervall $[E_{unif}, E_{unif}]$ haben, vereinfacht sich auch der *cut* enorm. Erstens fällt natürlich die Funktion cut_E weg, da ein Punktintervall nicht mehr verkleinert werden kann.

Zweitens sind die Werte in \mathbf{R}^I und \mathbf{P}^I durch E_{unif} stark aneinander gekoppelt. Gehen wir im Folgenden davon aus, dass die gegebene UACTMC bereits mit dem *cut* bereinigt wurde. Bei der Wahl einer Rate r durch einen Scheduler wissen wir dann, dass wegen der Kopplung über die feste Exitrate E_{unif} , die entsprechende Wahrscheinlichkeit p festgelegt werden muss, als $p = r/E_{unif}$. Entsprechend bestimmt die Wahl einer Wahrscheinlichkeit p die zugehörige Rate r durch $r = p \cdot E_{unif}$. Bei der Anpassung der übrigen Raten- und Wahrscheinlichkeitsintervallen nach $cut_{R_1^\gamma}$ und $cut_{P_1^\gamma}$ darf nach derselben Argumentation nur einer der beiden *cuts* durchgeführt werden, wenn man den anderen Wert entsprechend über das Verhältnis $r = p \cdot E_{unif}$ bestimmt.

Desweiteren kann die Funktion cut_{P_2} wegfallen, da sich die Wahrscheinlichkeiten der aus *abstr* resultierenden ACTMC aus der Division der Raten durch die feste Exitrate E_{unif} ergeben. Mit cut_{P_2} werden nur solche Wahrscheinlichkeiten entfernt, für die keine passende Exitraten und Raten existieren, weswegen dieser Teil des *cuts* überflüssig ist.

Was nun übrig bleibt ist im Wesentlichen cut_{P_1} . Dies ist auch der Teil, der im Fall von zeitdiskreten Markov-Ketten benötigt wird. Die übrigen Teile des *cut* fallen entweder ganz weg oder können durch Neuberechnung der Raten beziehungsweise Wahrscheinlichkeiten über die Verbindung durch die Exitrate bestimmt werden. Das Problem, dass die Uniformisierung von Markov-Ketten nicht für *Next*-Formeln geeignet ist bleibt jedoch bestehen, sodass bei solchen Formeln der aufwändige *cut* auf nicht-uniformen abstrakten zeitstetigen Markov-Ketten weiterhin benötigt wird.

3.6 Zusammenhang zwischen CTMC- und DTMC-Abstraktion

Nun, da wir die grundlegenden Definitionen und Eigenschaften abstrakter zeitstetiger Markov-Ketten eingeführt haben, wollen wir diese mit denen von abstrakten abstrakten *zeitdiskreten* Markov-Ketten nach [FLW06] vergleichen. Dort werden abstrakte DTMCs definiert als $\mathcal{M}_D = (S, \mathbf{P}^I, L)$ mit Zustandsmenge S , Wahrscheinlichkeitsintervallen \mathbf{P}^I sowie der Beschriftungsfunktion L . Die Abstraktion $\tilde{\mathcal{M}}_D = abstract_{DTMC}(\mathcal{M}_D, \mathcal{A}) = (\tilde{S}, \tilde{\mathbf{P}}^I, \tilde{L})$ einer solchen abstrakten zeitdiskreten Markov-Kette bezüglich einer Partitionierung $\mathcal{A} = \{A_1, \dots, A_n\}$ ist definiert durch:

- $\tilde{S} = \mathcal{A}$
- $\tilde{\mathbf{P}}^l(A_i, A_j) = \min\{1, \min_{s \in A_i} \mathbf{P}^l(s, A_j)\}$
- $\tilde{\mathbf{P}}^u(A_i, A_j) = \min\{1, \max_{s \in A_i} \mathbf{P}^u(s, A_j)\}$
- $\tilde{L}(A_i, a) = \begin{cases} \top & \text{falls } L(s, a) = \top \text{ für alle } s \in A_i \\ \perp & \text{falls } L(s, a) = \perp \text{ für alle } s \in A_i \\ ? & \text{sonst} \end{cases}$

Es ist leicht zu sehen, dass diese mit der Definition von Abstraktion B nahezu identisch ist. Die Zustandsmenge, die Wahrscheinlichkeitsintervallgrenzen und die Beschriftungsfunktion der Abstraktion werden auf exakt dieselbe Weise bestimmt. Die Raten- und Exitratenintervalle sind im zeitdiskreten Fall natürlich nicht definiert.

Da es auch bei abstrakten zeitstetigen Markov-Ketten ungültige Kombinationen von Wahrscheinlichkeiten geben kann, wird auch dort ein *cut* verwendet. Dieser entfernt alle Wahrscheinlichkeiten die nicht zu einer Verteilung vervollständigt werden können, was dem Anteil von cut_{P_1} am cut für den zeitstetigen Fall entspricht. Da abstrakte zeitdiskrete Markov-Ketten keine Raten oder Exitraten benötigen, werden weder die Teile cut_R oder cut_E benötigt, noch cut_{P_2} , da kein abgleich mit Raten und Exitraten wie im Fall von ACTMCs notwendig ist.

Nun betrachten wir die Eigenschaften von abstrakten Markov-Ketten, ohne die Verweildauern in den Zuständen zu berücksichtigen. Dazu benötigen wir die Definition der *eingebetteten abstrakten* zeitstetigen Markov-Kette, die sich wesentlich von der von eingebetteten zeitstetigen Markov-Ketten unterscheidet. Für CTMCs musste die Matrix der Wahrscheinlichkeitsintervallgrenzen über die Ratenintervallgrenzen und die eindeutigen Exitraten berechnet werden. In ACTMCs dagegen ist die Wahrscheinlichkeitsmatrix explizit gegeben. Damit in den eingebetteten abstrakten Markov-Ketten keine Wahrscheinlichkeitsverteilungen enthalten sein können, die aufgrund der Raten- und Exitratenintervalle in den ursprünglichen ACTMCs nicht wählbar gewesen wären, wird die eingebettete Markov-Kette einer ACTMC über die *bereinigten* Wahrscheinlichkeitsintervalle definiert.

Definition 29 (Eingebettete abstrakte zeitdiskrete Markov-Kette). Für eine abstrakte zeitstetige Markov-Kette \mathcal{M} heißt $\mathcal{M}_{emb} = (S_{emb}, \mathbf{P}_{emb}^I, L_{emb})$ die *eingebettete* abstrakte zeitdiskrete Markov-Kette, falls für $\mathcal{M}' = cut(\mathcal{M}) = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ gilt, dass $S_{emb} = S$, $\mathbf{P}_{emb}^I = \mathbf{P}^I$ und $L_{emb} = L$.

—

Lemma 7 (Abstraktion eingebetteter Markov-Ketten). Seien \mathcal{M}_{del} eine bereinigte ACTMC und \mathcal{A} eine Partitionierung über den Zustandsraum von \mathcal{M}_{del} . Die bereinigte Abstraktion der eingebetteten Markov-Kette $\mathcal{M}_{emb,abstr,del}$ bezüglich \mathcal{A} ist dann *größer* als die abstrakte zeitdiskrete Markov-Kette $\mathcal{M}_{abstrC,del,emb}$, der eingebetteten Markov-Kette der bezüglich \mathcal{A} abstrahierten und bereinigten Markov-Kette $\mathcal{M}_{abstrC,del}$ (siehe Abb. 32).

$$\begin{array}{ccccc}
\mathcal{M}_{del} & \xrightarrow[\text{CTMC}]{\text{abstr.}} & \mathcal{M}_{abstrC} & \xrightarrow{\text{cut}} & \mathcal{M}_{abstrC,del} \\
\text{emb.} \downarrow & & & & \downarrow \text{emb.} \\
\mathcal{M}_{emb} & \xrightarrow[\text{DTMC}]{\text{abstr.}} & \mathcal{M}_{emb,abstr} & \xrightarrow{\text{cut}} & \mathcal{M}_{emb,abstr,del} \supseteq \mathcal{M}_{abstrC,del,emb}
\end{array}$$

Abbildung 32: Abstraktion eingebetteter Markov-Ketten

Beweis. Im Folgenden seien die Matrizen der Wahrscheinlichkeitsintervallgrenzen der verschiedenen Markov-Ketten durch die Indizes der zugehörigen Markov-Kette identifiziert. Beispielsweise bezeichnet \mathbf{P}_{emb}^I die Wahrscheinlichkeitsintervalle der Markov-Kette \mathcal{M}_{emb} .

Wir können zunächst feststellen, dass aufgrund der Definition einer eingebetteten abstrakten Markov-Kette gilt, dass $\mathbf{P}_{emb}^I = \mathbf{P}_{del}^I$ und $\mathbf{P}_{emb,abstr,del}^I = \mathbf{P}_{abstrC,del}^I$ gilt. Da die Abstraktion bezüglich zeitdiskreter Markov-Ketten bis auf die fehlenden Raten und Exitraten genau mit der für zeitstetige Markov-Ketten übereinstimmt und die gegebenen Matrizen \mathbf{P}_{emb}^I und \mathbf{P}_{del}^I genau gleich sind, gilt außerdem auch $\mathbf{P}_{abstrC}^I = \mathbf{P}_{emb,abstr}^I$.

Damit die Behauptung aufgeht, muss nun noch gezeigt werden, dass der *cut* auf der ACTMC \mathcal{M}_{abstrC} mindestens die Werte der Wahrscheinlichkeitsintervalle entfernt wie auch bezüglich der abstrakten DTMC $\mathcal{M}_{emb,abstr}$.

Wie wir schon festgestellt haben beschränkt sich der *cut* bezüglich abstrakter zeitdiskreter Markov-Ketten auf cut_{P_1} . Die Iteration von cut_{P_1} bezüglich zeitdiskreter Markov-Ketten ist konvergent (siehe [FLW06]). Wendet man zunächst nur cut_{P_1} auf die zeitstetige Markov-Kette \mathcal{M}_{abstrC} an bis der Fixpunkt erreicht ist²⁵, so erhält man dieselben Wahrscheinlichkeitsintervalle wie in $\mathcal{M}_{emb,abstr,del}$. Die dadurch entstandene ACTMC ist im Allgemeinen noch nicht bereinigt, weswegen sie noch immer mit dem *cut* zu behandeln ist. Da der *cut* eine monotone Funktion ist (siehe Beweis zu Satz 1), folgt damit $\mathbf{P}_{abstrC,del}^I \subseteq \mathbf{P}_{emb,abstr,del}^I$. Wegen $\mathbf{P}_{emb,abstr,del}^I = \mathbf{P}_{abstrC,del}^I$ gilt also auch $\mathbf{P}_{abstrC,del,emb}^I \subseteq \mathbf{P}_{emb,abstr,del}^I$ und damit die Behauptung des Lemmas. \square

Es bleibt an dieser Stelle die interessante Frage offen, ob auch die umgekehrte Inklusion gilt oder nicht.

Betrachten wir abschließend noch die Simulationsrelation im zeitstetigen und zeitdiskreten Fall. In [FLW06] wird die folgende Definition für eine probabilistische Simulationsrelation verwendet:

Definition 30 (Probabilistische Simulation für DTMCs). Sei $\mathcal{M} = (S, \mathbf{P}^I, L)$ eine abstrakte zeitdiskrete Markov-Kette, dann heißt $\mathcal{R} \subseteq S \times S$ probabilistische Simulationsrelation, falls aus $s\mathcal{R}s'$ folgt:

²⁵Bei der Definition des *cut* wurde zwar die Reihenfolge der Teilfunktionen festgelegt, es wurde aber dort schon angemerkt, dass es genügt, wenn die Abfolge der Teilfunktionen fair ist.

1. Für alle $a \in AP$ gilt $(L(s', a) \neq ?) \rightarrow (L(s, a) = L(s', a))$.
2. Für alle Verteilungen $\bar{\mu} \in \text{distr}(\mathbf{P}(s, \cdot))$ existiert eine entsprechende Verteilung $\bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))$ und eine Gewichtung $\Delta : S \times S \mapsto [0, 1]$, sodass für alle $u, v \in S$ gilt:
 - 2.1. $\Delta(u, v) > 0 \Rightarrow u\mathcal{R}'v$,
 - 2.2. $\Delta(u, S) = \bar{\mu}(u)$,
 - 2.3. $\Delta(S, v) = \bar{\mu}'(v)$.

Existiert eine solche Simulationsrelation \mathcal{R} mit $s\mathcal{R}s'$, so schreiben wir $s \preceq' s'$.

Diese Definition ist, abgesehen von den fehlenden Bedingungen für Exitraten, genau dieselbe, die wir im zeitstetigen Fall verwenden. Das bedeutet also, dass Zustände, die bezüglich einer abstrakten CTMC in Simulationsrelation stehen, automatisch auch in der eingebetteten abstrakten Makrov-Kette in Simulationsrelation stehen.

Lemma 8. Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine bereinigte abstrakte zeitstetige Markov-Kette und $\mathcal{M}_{emb} = (S_{emb}, \mathbf{P}_{emb}^I, L_{emb})$ die zugehörige eingebetteten Markov-Kette. Dann gilt für Zustände $s, s' \in S$:

$$s \preceq s' \text{ in } \mathcal{M} \Rightarrow s \preceq' s' \text{ in } \mathcal{M}_{emb}$$

Beweis. Aus der Definition der eingebetteten Markov-Kette folgt, dass $S_{emb} = S$, $\mathbf{P}_{emb}^I = \mathbf{P}^I$ sowie $L_{emb} = L$ gelten muss. Da sich die beiden Simulationsrelationen \mathcal{R} und \mathcal{R}' in den Punkten 1 und 2 nicht unterscheiden, folgt die Behauptung unmittelbar:

Falls $s \preceq s'$ gilt, so existiert eine Relation \mathcal{R} nach Definition 24 mit $s\mathcal{R}s'$. Bezüglich dieser Relation gilt wegen 24.1 und 24.2 auch $s\mathcal{R}'s'$ nach der Definition 30. Damit folgt dann, dass auch $s \preceq' s'$ gilt, womit die Behauptung des Lemmas gezeigt wäre. \square

3.7 Wahrscheinlichkeitsräume

Im nachfolgenden Kapitel interessieren wir uns insbesondere für Wahrscheinlichkeiten von Pfadmengen. Um dies behandeln zu können, müssen wir zunächst festlegen, wie der Wahrscheinlichkeitsraum aussieht und wie die Wahrscheinlichkeiten in diesem Raum gemessen werden sollen. Bei den Definitionen orientieren wir uns an [BHHK03] und [FLW06]. Im Gegensatz zu [BHHK03] werden jedoch keine beliebigen Zeitintervalle für die Verweildauer in einem bestimmten Zustand eines Pfades zugelassen, da die Berechnung der maximierenden und minimierenden Raten sonst zu schwierig wird. Betrachten wir zunächst ein Beispiel dazu:

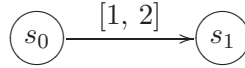


Abbildung 33: Abstrakte zeitstetige Markov-Kette

Beispiel 24 (Wahrscheinlichkeitsmaß für Pfade mit beliebigen Intervallen). Wir möchten in der folgenden abstrakten Markov-Kette die maximale Wahrscheinlichkeit bestimmen, in frühestens $\frac{3}{10}$ und spätestens 1 Zeiteinheiten von s_0 nach s_1 zu gelangen.

Dies können wir berechnen mit der Formel:

$$\mathbf{P}(s_0, s_1) \cdot \max_{r \in [1, 2]} (e^{-\frac{3}{10} \cdot r} - e^{-1 \cdot r}) = 1 \cdot \max_{r \in [1, 2]} (e^{-\frac{3}{10} \cdot r} - e^{-1 \cdot r}).$$

Da $e^{-\frac{3}{10} \cdot r} - e^{-1 \cdot r}$ die Dichtefunktion einer Exponentialverteilung ist, gibt es für $r \geq 0$ ein eindeutiges Maximum. Dieses Maximum können wir wie üblich über die Nullstelle der ersten Ableitung bzgl. r bestimmen:

$$\begin{aligned} e^{-1 \cdot r} - \frac{3}{10} \cdot e^{-\frac{3}{10} \cdot r} &= 0 \\ \text{gdw.} \quad e^{-1 \cdot r} &= \frac{3}{10} \cdot e^{-\frac{3}{10} \cdot r} \\ \text{gdw.} \quad e^{-1 \cdot r} \cdot e^{\frac{3}{10} \cdot r} &= \frac{3}{10} \\ \text{gdw.} \quad e^{-\frac{7}{10} \cdot r} &= \frac{3}{10} \\ \text{gdw.} \quad r &= \ln \frac{3}{10} \cdot \left(-\frac{10}{7}\right) \approx 1,72 \end{aligned}$$

Für beliebige Zeit-Intervalle ist die maximale Wahrscheinlichkeit also nicht zwangsläufig in einem Randwert des Ratenintervalls zu finden. Da wir uns im weiteren Verlauf auf Scheduler beschränken werden, die nur extreme Intervallgrenzen wählen, beschränken wir uns auch bereits jetzt schon auf $[0, t]$ -Intervalle. Dann erhalten wir in diesem Beispiel die Formel:

$$\mathbf{P}(s, s') \cdot \max_{r \in [1, 2]} (e^0 - e^{-t \cdot r}) = \max_{r \in [1, 2]} (1 - e^{-t \cdot r})$$

Um das maximierende r für $1 - e^{-t \cdot r}$ zu finden, genügt es natürlich, das minimierende r für $e^{-t \cdot r}$ zu bestimmen. Da $e^{-t \cdot r}$ bekanntermaßen streng monoton fallend bezüglich $r \in \mathbb{R}_{\geq 0}$ ist, muss für r ein größtmöglicher Wert gewählt werden, also die obere Schranke des Intervalls. Wir werden im Folgenden nur $[0, t]$ -Intervalle zulassen, sodass die maximalen und minimalen Wahrscheinlichkeiten einfacher bestimmt werden können.

Sei $\mathcal{I} = \{[0, t] \mid t \in \mathbb{R}_{>0}\}$, dann bezeichnen wir bezüglich einer gegebenen zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, L)$ die Ereignismenge Ω als die Menge aller möglichen Pfadanfänge in \mathcal{M} . Wie schon für CTMCs erhalten wir wieder als Erzeugnis aus der Menge aller Pfadanfänge \mathcal{B}_b einen Borel-Raum, über den wir Zugang zu den *Maßen* von Zylindermengen, d.h. zu den Pfadwahrscheinlichkeiten erhalten. Die Zylindermenge $C(s_0, I_0, \dots, s_{k-1}, I_{k-1}, s_k)$ mit $s_0, \dots, s_k \in S$ und $I_0, \dots, I_{k-1} \in \mathcal{I}$

repräsentiert alle Pfade $\sigma \in Pfad(s_0)$ mit $\sigma[i] = s_i$ für $i \in \{0, \dots, k\}$ und $\delta(\sigma, i) \in I_i$ für $i \in \{0, \dots, k-1\}$.

In abstrakten Markov-Ketten sprechen wir von Pfadwahrscheinlichkeiten zunächst einmal nur im Zusammenhang mit den induzierten Markov-Ketten. Ein Scheduler $\eta \in \mathcal{S}(\mathcal{M})$ induziert eine Markov-Kette mit dem zugehörigen Wahrscheinlichkeitsraum $\mathcal{PS}^\eta = (\Omega, \mathcal{B}, Pr^\eta)$, sodass Pr^η für einen Startzustand s_0 eindeutig bestimmt ist durch:

$$Pr^\eta(C(s_0, I_0, \dots, s_{k-1}, I_{k-1}, s_k)) = (\mathbf{P}(s_0, s_1) \cdot X(s_0, I_0)) \cdot \dots \cdot (\mathbf{P}(s_{k-1}, s_k) \cdot X(s_{k-1}, I_{k-1})) \quad (10)$$

mit der Wahrscheinlichkeit $X(s, I) = \int_I E(s) \cdot e^{-E(s) \cdot t} dt = (e^{-E(s) \cdot \inf I} - e^{-E(s) \cdot \sup I})$, im Intervall I den Zustand s zu verlassen.

Betrachten wir nun Supremum und Infimum von messbaren Mengen in *uniformisierten* ACTMCs, so stellen wir fest, dass es ausreicht nur die extremen Scheduler zu untersuchen. Dies ist ein signifikanter Vorteil, da wir statt mit überabzählbar vielen nur noch mit abzählbar vielen Schemulern arbeiten, denn in jedem Schritt haben solche Scheduler nur endlich viele Wahlmöglichkeiten, sodass man nach Art einer Breitensuche im Entscheidungsbaum der Scheduler eine Nummerierung angeben kann.

Satz 5. Für eine uniforme abstrakte zeitstetige Markov-Kette \mathcal{M} gilt für alle messbaren Mengen Q des induzierten Wahrscheinlichkeitsraums:

$$\begin{aligned} \sup_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(Q) &= \sup_{\eta \in \mathcal{S}(\mathcal{M})} Pr^\eta(Q) \\ \inf_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(Q) &= \inf_{\eta \in \mathcal{S}(\mathcal{M})} Pr^\eta(Q) \end{aligned}$$

Beweis. Wir zeigen im Folgenden per struktureller Induktion, dass für beliebige Zylindermengen extreme Scheduler zur Berechnung des Supremums ausreichen. Außerdem zeigen wir, dass dies für die Vereinigungen von Zylindermengen und für die Komplementbildung ebenfalls gilt.

Beginnen wir mit der Maximierung des Maßes für Zylindermengen bezüglich \mathcal{M} als Induktionsanfang. Dazu konstruieren wir $\eta \in \mathcal{S}(\mathcal{M})$ als einen Scheduler mit

$$\eta(s_0, \dots, s_i) = (\mu_i, \bar{\mu}_i, E_i) \in (rates(\mathbf{R}(s_i, \cdot)), distr(\mathbf{P}(s_i, \cdot)), E(s_i)),$$

der das maximale Maß für eine Zylindermenge $C(s_0, I_0, \dots, s_k)$ liefert.

Die Gleichung (10) wird durch den Scheduler η maximiert, falls die einzelnen Faktoren $(\mathbf{P}(s_i, s_{i+1}) \cdot X(s_i, I_i))$ maximiert werden. In *uniformen* zeitstetigen Markov-Ketten ist die Wahrscheinlichkeit, innerhalb eines Intervalls $I \in \mathcal{I}$ den Zustand s zu verlassen, bestimmt durch $X(s, I) = \int_I E(s) \cdot e^{-E(s) \cdot t} dt = (1 - e^{-E_{unif} \cdot \sup I})$ und damit unabhängig von der Wahl des Schemulers. Es bleibt noch, für jedes Zustandspaar s_i, s_{i+1} der Zylindermenge die Übergangswahrscheinlichkeit $\mathbf{P}(s_i, s_{i+1}) = \bar{\mu}_i(s_{i+1})$

zu maximieren. Die zugehörige Rate ergibt sich dann eindeutig als $\mu_i(s_{i+1}) = \bar{\mu}_i(s_{i+1}) \cdot E_{unif}$ und ist nach Korollar 1.4 die maximale wählbare Rate.

Die übrigen Wahrscheinlichkeiten und Raten zu Zuständen $s_j \neq s_{i+1}$ liefern keinen Beitrag zum Wahrscheinlichkeitsmaß der gegebenen Zylindermenge und könnten eigentlich beliebig gewählt werden. Um einen extremen Scheduler zu erhalten geben wir jedoch vor, dass die übrigen Raten bzw. Wahrscheinlichkeiten vom Scheduler η immer maximal gewählt werden. Wir wissen damit also, dass immer ein extremer Scheduler konstruiert werden kann, der das Maß von Basiszylindermengen maximiert.

Nun folgt der Induktionsschluß: Unser Wahrscheinlichkeitsraum wurde als ein, durch die Menge aller Basiszylindermengen der Markov-Kette \mathcal{M} erzeugter, Borel-Raum definiert. Daher zeigen wir nun, wie sich maximierende extreme Scheduler für Vereinigung und Komplement von Zylindermengen finden lassen, für die wir bereits maximierende extreme Scheduler kennen. Daraus folgt mit dem Induktionsanfang, dass wir für alle Elemente des Borel-Raums einen maximierenden extremen Scheduler konstruieren können.

Betrachten wir nun die Vereinigung zweier Zylindermengen $C(s_0, I_0, \dots, s_k)$ und $C(s'_0, I'_0, \dots, s'_l)$, wobei o.B.d.A. $k \leq l$ gelte. Teilen wir das Problem in disjunkte Vereinigung und nicht-disjunkte Vereinigung auf:

1. Die Mengen der Pfade, die von den beiden Zylindermengen repräsentiert werden, sind genau dann nicht disjunkt, wenn $s_i = s'_i$ für alle $i \in \{0, \dots, k\}$. Der Grund dafür ist, dass es in diesem Fall immer einen nicht-leeren Schnitt der beiden Zylindermengen geben muss²⁶:

$$\begin{aligned} & C(s_0, I_0, \dots, s_k) \cap C(s'_0, I'_0, \dots, s'_l) \\ &= C(s_0, I_0 \cap I'_0, s_1, I_1 \cap I'_1, \dots, s_k, I'_k, s'_{k+1}, I'_{k+1}, \dots, s'_l) \end{aligned}$$

Außerdem kann es keine nicht-disjunkten Zylindermengen geben mit $s_i \neq s'_i$ für ein $i \in \{0, \dots, k\}$, da gilt $\sigma[i+1] = s_i \neq s'_i = \sigma'[i+1]$ für alle $\sigma \in C(s_0, I_0, \dots, s_k)$ und alle $\sigma' \in C(s'_0, I'_0, \dots, s'_l)$.

Ein Scheduler η kann in *uniformen* Markov-Ketten nur durch die Wahl von Wahrscheinlichkeiten Einfluß auf das Maß ausüben. Die Exitrate ist festgelegt und damit die Geschwindigkeit für das Verlassen eines Zustandes, wie wir schon früher beobachtet haben. Da die Zylindermenge $C(s'_0, I'_0, \dots, s'_l)$ lediglich eine Teilmenge der Zylindermenge $C(s_0, I_0, \dots, s_k)$ darstellt, muss ein maximierender Scheduler für die kleinere Zylindermenge $C(s'_0, I'_0, \dots, s'_l)$ in den ersten k Entscheidungen auch das Maß für die größere Zylindermenge maximieren. Für die Vereinigung der Zylindermengen liefert der maximierende Scheduler

²⁶An dieser Stelle verwenden wir bereits die Beschränkung auf Intervalle der Form $[0, t]$. Gäbe es an dieser Stelle die Beschränkung noch nicht, so müsste ein weiterer Fall behandelt werden, in dem die Zustände s_i und s'_i für $i \in \{0, \dots, k\}$ zwar gleich, die Zylindermengen aber wegen disjunkter Zeitintervalle in einem Zustand aber dennoch disjunkt sind.

der kleineren Zylindermenge also ebenfalls ein maximales Maß. Nach Induktionsvoraussetzung ist dieser ein extremer Scheduler.

2. Die Mengen der Pfade, die von zwei Zylindermengen repräsentiert werden sind disjunkt, falls $s_i \neq s'_i$ für ein $i \in \{0, \dots, k\}$. Wir können für diesen Fall einen extremen Scheduler konstruieren, der das maximale Maß für die Vereinigung liefert.

Dazu betrachten wir den letzten gemeinsamen Zustand s_{i-1} , für den $s_j = s'_j$ für alle $j \in \{0, \dots, i-1\}$ gilt. Ein Scheduler, der das maximale Maß für die Vereinigung bestimmt, muss die Summe über die beiden zu vereinigenden Anteile maximieren. Den anfänglichen Teil bis einschließlich Zustand s_{i-1} können wir dabei auslassen, da ein maximierender Scheduler in $C(s_0, I_0, \dots, s_k)$ ebenfalls maximierend ist für $C(s_0, I_0, \dots, s_{i-1})$ und für $C(s'_0, I'_0, \dots, s'_{i-1})$. Es bleibt also Folgendes zu maximieren:

$$\begin{aligned} & \max(\mathbf{P}(s_{i-1}, s_i) \cdot X(s_0, I_0) \cdot \dots \cdot \mathbf{P}(s_{k-1}, s_k) \cdot X(s_{k-1}, I_{k-1})) \\ & \quad + (\mathbf{P}(s'_{i-1}, s'_i) \cdot X(s'_{i-1}, I'_{i-1}) \cdot \dots \cdot \mathbf{P}(s'_{l-1}, s'_l) \cdot X(s'_{l-1}, I'_{l-1})) \\ & = \max((\mathbf{P}(s_{i-1}, s_i) \cdot \dots \cdot \mathbf{P}(s_{k-1}, s_k)) + (\mathbf{P}(s'_{i-1}, s'_i) \cdot \dots \cdot \mathbf{P}(s'_{l-1}, s'_l))) \\ & = \max(\mathbf{P}(s_{i-1}, s_i) \cdot p + \mathbf{P}(s'_{i-1}, s'_i) \cdot p') \end{aligned}$$

Die Faktoren p und p' können mit den nach Induktionsvoraussetzung gegebenen maximierenden extremen Schemulern bestimmt werden.

Als erstes muss der zu konstruierende Scheduler in s_{i-1} demnach die Wahrscheinlichkeit $\bar{\mu}(s_i)$ maximal wählen, falls $p > p'$ und die Wahrscheinlichkeit $\bar{\mu}(s'_i)$ sonst. Dadurch wird die Zylindermenge mit dem größeren Maß bei den Übergangswahrscheinlichkeiten bevorzugt. Nach der Ausführung des *cut* sollte der Scheduler dann die jeweils andere Wahrscheinlichkeit maximal wählen, womit alle interessanten Pfadmengen abgedeckt wären. Für die übrigen Wahrscheinlichkeiten, die zu nicht relevanten Zylindermengen führen, lassen wir den Scheduler aus den verbleibenden möglichen Werten minimal mögliche Werte wählen.

Das Ergebnis ist ein Scheduler, der bis zum Zustand s_{i-2} und ab Zustand s_i bzw. s'_i immer dieselben Entscheidungen trifft wie die per Induktion gegebenen extremen Scheduler. In Zustand s_{i-1} wird wie oben beschrieben ebenfalls eine extreme Wahrscheinlichkeitsverteilung und damit ein extremer Ratenvektor gewählt. Der resultierende Scheduler ist also wieder ein extremer Scheduler.

Die Maximierung des Maßes des Komplements einer Zylindermenge C können wir leicht auf die Minimierung dessen Maßes zurückführen. Da der gesamte Wahrscheinlichkeitsraum Ω das Maß Eins hat und C Teilmenge von Ω ist, gilt:

$$Pr^\eta(C^c) = Pr^\eta(\Omega \setminus C) = Pr^\eta(\Omega) - Pr^\eta(C) = 1 - Pr^\eta(C).$$

Daraus folgt:

$$\sup_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(C^c) = \sup_{\eta \in \mathcal{ES}(\mathcal{M})} (1 - Pr^\eta(C)) = 1 - \inf_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(C)$$

Dass auch $\inf_{\eta \in \mathcal{ES}(s)} Pr^\eta(Q) = \inf_{\eta \in \mathcal{S}(s)} Pr^\eta(Q)$ gilt, kann analog zu obigem Beweis gezeigt werden. Damit ist also auch die Maximierung des Komplements der Zylindermenge mit extremen Schedulern möglich.

Damit ist der Induktionsbeweis abgeschlossen und wir wissen, dass für Supremum und Infimum von Wahrscheinlichkeitsmaßen die Betrachtung aller extremen Scheduler statt der Betrachtung aller HD-Scheduler genügt.

□

4 Model Checking für abstrakte Markov-Ketten

In diesem Kapitel beschäftigen wir uns mit Model Checking für die in Kapitel 3 eingeführten abstrakten zeitstetigen Markov-Ketten. Die *Continuous Stochastic Logic*, die auf der zweiwertigen Logik aufgebaut ist, ist nicht ausreichend um Eigenschaften von ACTMCs überprüfen zu können, die per Definition atomare Eigenschaften mit dem Wahrheitswert *unbestimmt* enthalten können. Wir müssen also zunächst eine dreiwertige Variante von CSL definieren.

Bevor wir uns anschließend dem Model Checking für dreiwertiges CSL auf abstrakten zeitstetigen Markov-Ketten zuwenden, werden wir noch das quantitative Erreichbarkeitsproblem isoliert betrachten, wobei wir weitgehend dem Ansatz aus [BHKH04] folgen werden. Dabei werden wir voraussetzen, dass die gegebene Markov-Kette uniform ist. Anschließend beschäftigen wir uns mit der Frage, wie das Erreichbarkeitsproblem für die Verifikation von Spezifikationen verwendet werden kann und zeigen, dass Spezifikationen, die für eine Abstraktion erfüllbar oder unerfüllbar sind, auch in der zugehörigen ursprünglichen Markov-Kette erfüllbar bzw. unerfüllbar sein müssen. Diese Beobachtung erlaubt uns dann, in der Abstraktion Eigenschaften zu überprüfen und daraus Rückschlüsse auf die gegebene Markov-Kette ziehen zu können.

4.1 3-CSL (3-valued Continuous Stochastic Logic)

Die Menge der 3-CSL Formeln ist syntaktisch genau wie die der CSL Formeln induktiv definiert. Bei der Semantik muss jedoch berücksichtigt werden, dass die Gültigkeit einer Formel auch unbestimmt sein kann. Für den aussagenlogischen Anteil haben wir schon in Kapitel 2.7 festgestellt, dass die Semantik wie im zweiwertigen Fall über Komplement und *meet* definiert werden kann. Die Semantik des *Next*-Operators reduziert sich auf die Semantik von Nachfolgezuständen, falls es solche gibt, und muss ebenfalls nicht neu definiert werden. Die Definition des *Until*-Operators muss nur geringfügig erweitert werden. Wenn weder die Bedingung für Gültigkeit noch die für Ungültigkeit erfüllt werden kann, so muss das Ergebnis *unbestimmt* sein. Wie der Wahrheitswert der *Until*-Formel für einen Pfad algorithmisch bestimmt werden kann, wird später in Kapitel 4.3 erläutert werden.

Für den Operator $\mathcal{P}_{\boxtimes p}$ muss im Grunde eine ähnliche Modifikation durchgeführt werden wie bei *Until*. In der Definition der CSL Semantik wurde der Fall der Unerfüllbarkeit jedoch nicht explizit festgelegt. Die Voraussetzung für die Erfüllbarkeit von $\mathcal{P}_{\geq p}(\psi)$ war, dass die Wahrscheinlichkeit für Pfade mit Eigenschaft ψ bezüglich des gegebenen Zustands den Wert p nicht unterschreitet. Damit diese Voraussetzung mit Sicherheit verletzt ist, müsste die Wahrscheinlichkeit größer sein als $1 - p$, dass auf den Pfaden die Eigenschaft ψ definitiv nicht gilt (siehe Tabelle 5). Dann bleibt für die Wahrscheinlichkeit von erfüllbaren und unbestimmten Pfaden bezüglich ψ nur noch weniger als Wahrscheinlichkeit p übrig.

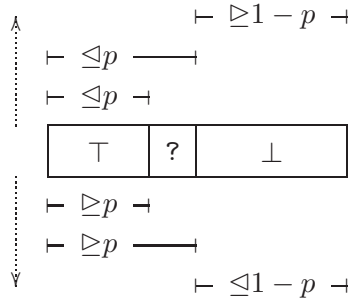


Abbildung 34: Zusammenhang dreiwertiger Wahrscheinlichkeiten

Wir verwenden als Abkürzungen für obere und untere Schranken für Wahrscheinlichkeitsmaßen bezüglich einer Pfadmenge M im Folgenden:

$$Pr^l(M) = \inf_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(M)$$

$$Pr^u(M) = \sup_{\eta \in \mathcal{ES}(\mathcal{M})} Pr^\eta(M)$$

Für einen Zustand s , eine Pfadformel Ψ und $\alpha \in \mathbb{B}^3$ schreiben wir als Abkürzungen:

$$Pr^l(s, \Psi, \alpha) = Pr^l(\{\sigma \in Pfad_s \mid \llbracket \sigma, \Psi \rrbracket = \alpha\})$$

$$Pr^u(s, \Psi, \alpha) = Pr^u(\{\sigma \in Pfad_s \mid \llbracket \sigma, \Psi \rrbracket = \alpha\})$$

Für $\mathcal{P}_{\leq p}(\psi)$ ergibt sich die Semantik fast analog. Um nicht obere Grenze von Wahrscheinlichkeiten Pr^u verwenden zu müssen, schreiben wir die Bedingung ein wenig um, sodass über Pr^l argumentiert werden kann. Wenn die Wahrscheinlichkeit für die Erfüllbarkeit von ψ höchstens den Wert p haben soll, so ist dies gesichert, falls die Wahrscheinlichkeit, dass ψ sicher verletzt wird, mindestens $1 - p$ beträgt. Entsprechend ist $\mathcal{P}_{\leq p}(\psi)$ definitiv nicht erfüllbar, wenn die minimale Wahrscheinlichkeit, dass ψ erfüllt ist, mindestens p beträgt (siehe Abb. 34). Die in Tabelle 5 verwendete Menge $Sat(\psi) = \{s \mid \llbracket s, \psi \rrbracket = \top\}$ bezeichnet die Menge aller Zustände, in denen eine Zustandsformel ψ erfüllt ist. Die Menge der Zustände, in denen eine Zustandsformel ψ nicht erfüllen kann, bezeichnen wir mit $Fls(\psi) = \{s \mid \llbracket s, \psi \rrbracket = \perp\}$.

Nun ist es uns möglich zu abstrakten zeitstetigen Markov-Ketten Spezifikationen anzugeben. Da sich diese kaum vom denen in CSL unterscheiden, verzichten wir darauf dies an einem Beispiel zu demonstrieren. Interessanter wird die Frage nach einem Model Checking Verfahren für 3-CSL, was wir in Kapitel 4.3 behandeln werden.

4.2 Quantitative, zeitabhängige Erreichbarkeit

Betrachten wir nun das quantitative, zeitabhängige Erreichbarkeitsproblem für abstrakte zeitstetige Markov-Ketten, also die Frage nach der größten bzw. kleinsten Wahrscheinlichkeit, von einem Startzustand s aus in höchstens t Zeiteinheiten einen Zustand der Zielmenge B zu erreichen. Für abstrakte zeitstetige Markov-Ketten kann es je nach Scheduler verschiedene Wahrscheinlichkeiten für die Erreichbarkeit einer Zielmenge geben.

$$\begin{aligned}
\llbracket \sigma, \psi_1 \mathcal{U}^I \psi_2 \rrbracket &= \begin{cases} \top & \text{falls } \exists t \in I : (\llbracket \sigma @ t, \psi_2 \rrbracket = \top \\ & \wedge \forall t' \in [0, t) : \llbracket \sigma @ t', \psi_1 \rrbracket = \top) \\ \perp & \text{falls } \forall t \in I : (\llbracket \sigma @ t, \psi_2 \rrbracket = \perp \\ & \vee \exists t' \in [0, t) : \llbracket \sigma @ t', \psi_1 \rrbracket = \perp) \\ ? & \text{sonst} \end{cases} \\
\llbracket s, \mathcal{P}_{\supseteq p}(\Psi) \rrbracket &= \begin{cases} \top & \text{falls } Pr^l \{ \sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, \Psi \rrbracket = \top \} \supseteq p \\ \perp & \text{falls } Pr^l \{ \sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, \Psi \rrbracket = \perp \} \supset 1 - p \\ ? & \text{sonst} \end{cases} \\
\llbracket s, \mathcal{P}_{\leq p}(\Psi) \rrbracket &= \begin{cases} \top & \text{falls } 1 - p \leq Pr^l \{ \sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, \Psi \rrbracket = \perp \} \\ \perp & \text{falls } p \triangleleft Pr^l \{ \sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, \Psi \rrbracket = \top \} \\ ? & \text{sonst} \end{cases}
\end{aligned}$$

mit $\psi, \psi_1, \psi_2 \in \mathbb{S}^{\mathcal{M}}$; $\Psi \in \mathbb{P}^{\mathcal{M}}$; $p \in [0, 1]$, $I \in \mathcal{I}$;

$$\leq \in \{<, \leq\}, \triangleleft = \begin{cases} < & \text{falls } \leq = \leq \\ \leq & \text{falls } \leq = < \end{cases}$$

$$\supseteq \in \{<, \leq\}, \triangleright = \begin{cases} > & \text{falls } \supseteq = \geq \\ \geq & \text{falls } \supseteq = > \end{cases}$$

für eine gegebene abstrakte zeitstetige Markov-Kette \mathcal{M}

Tabelle 5: 3-CSL Semantik

Da wir im weiteren Verlauf immer minimale Wahrscheinlichkeiten verwenden werden, behandeln wir im Folgenden nur diesen Fall. Der Fall mit maximalen Wahrscheinlichkeiten kann auf ähnliche Weise behandelt werden. Wir interessieren uns also nun für die Wahrscheinlichkeit, die gerade so klein ist, dass sie von keinem Scheduler mehr unterboten werden kann, also die größte Wahrscheinlichkeit, die *unabhängig vom Scheduler* die Eigenschaft erfüllt, mindestens so klein zu sein, wie die Wahrscheinlichkeit, einen Zustand in B innerhalb von t Zeiteinheiten zu erreichen:

$$\inf_{\eta \in \mathcal{S}(\mathcal{M})} Pr^\eta(\text{Reach}_{\leq t}(s, B))$$

Hierbei sei $\text{Reach}_{\leq t}(s, B) = \{\sigma \in \text{Pfad}_s^{\mathcal{M}} \mid \sigma @ t' \in B \text{ für ein } t' \in [0, t]\}$ die Menge aller Pfade einer Markov-Kette \mathcal{M} die von s aus in maximal t Zeiteinheiten einen Zustand aus B erreichen. Mit Pr^η bezeichnen wir das Wahrscheinlichkeitsmaß für eine Pfadmenge in der durch Scheduler η induzierten zeitstetigen Markov-Kette.

Wie wir aus dem Grundlagenkapitel wissen, sind uniforme CTMCs für die Betrachtung der *Transient* Eigenschaften geeignet. Dass uniforme ACTMCs ausschließlich uniforme CTMCs induzieren, können wir im Folgenden für die *Transient* Analyse ausnutzen. Wir gehen im folgenden also von uniformen ACTMCs aus, wobei wir mit E_{unif} immer die uniforme Exitrate bezeichnen werden.

Wir betrachten nun also Pfade, die in einen Zustand aus der Menge B führen. Um die Berechnung zu vereinfachen, führen wir eine Transformation für Markov-Ketten ein, durch die die Wahrscheinlichkeiten dieser Pfade nicht verändert werden. Die Transformation ist relativ simpel und besteht darin, für alle $s_B \in B$ die ausgehenden Kanten zu entfernen, und sie stattdessen mit *self-loops* mit Wahrscheinlichkeiten Eins auszustatten. Dadurch müssen wir nicht mehr alle Positionen eines Pfads nach Zuständen in B absuchen, sondern es genügt die jeweils letzten Zustände der Pfade zu betrachten, da durch die Transformation sichergestellt wurde, dass nach Erreichen eines Zustandes aus B kein anderer Zustand mehr folgen kann²⁷.

Stellen wir nun eine Gleichung auf, für die Wahrscheinlichkeit einen Zustand aus B zu erreichen. Dazu betrachten wir die *induzierten* Markov-Ketten. Für uniforme CTMCs wissen wir, dass der Vektor $\pi(\tilde{a}, t) = \alpha \cdot \sum_{n=0}^{\infty} \varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}^n$ die Wahrscheinlichkeiten liefert, von einer Startverteilung α in genau t Zeiteinheiten in die verschiedenen Zustände in S zu gelangen. Statt der Wahrscheinlichkeitsmatrix \mathbf{P} haben wir bei der bezüglich B transformierten und durch einen Scheduler η induzierten Markov-Kette die Matrix $\mathbf{P}_{\eta, B}$. Da uns nur diejenigen Pfade interessieren, die in Zuständen aus B enden, fügen wir noch den Vektor i'_B hinzu, der für Zustände aus B den Wert Eins hat und ansonsten Null. Pfade, die in einem Zustand aus $S \setminus B$ enden, werden sozusagen durch eine Multiplikation mit Null aus dem Ergebnis entfernt.

Da die konkrete zeitstetige Markov-Kette, die durch den Scheduler η induziert wird, in der Regel einen unendlichen Zustandsraum S^+ und damit eine unendlich große

²⁷In einem solchen Zustand ist die Eigenschaft $at_B \wedge (A \mathcal{X} \text{ true})$ erfüllt. Vergleiche dazu Punkt 4.2 in [EL87]

Wahrscheinlichkeitsmatrix \mathbf{P} hat, erhalten wir zunächst einen unendlichen Vektor von Wahrscheinlichkeiten, von einem gegebenen Zustand $\sigma \in S^+$ der induzierten Markov-Kette aus in einen B -Zustand zu gelangen:

$$(Pr^\eta(\text{Reach}_{\leq t}(s, B)))_{\sigma \in S^+} = \sum_{n=0}^{\infty} \varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}_{\eta, B}^n \cdot i'_B$$

$$\text{mit } i'_B = (i_B(\sigma))_{\sigma \in S^+} \text{ und } i_B(\sigma) = \begin{cases} 1 & \text{falls } \sigma \in (S^* \cdot B) \\ 0 & \text{sonst} \end{cases}$$

$$\text{sowie } \mathbf{P}_{\eta, B}(\sigma, \sigma') = \begin{cases} \mathbf{P}_\eta(\sigma, \sigma') & \text{falls } \sigma \notin (S^* \cdot B) \\ 1 & \text{falls } \sigma \in (S^* \cdot B) \text{ und } \sigma'[\|\sigma'\|] = \sigma[\|\sigma\|] \\ 0 & \text{sonst} \end{cases}$$

Da wir von einer festen uniformen Exitrate E_{unif} ausgehen, ergibt sich die transformierte Ratenmatrix wie folgt:

$$\mathbf{R}_{\eta, B}(\sigma, \sigma') = \begin{cases} \mathbf{R}_\eta(\sigma, \sigma') & \text{falls } \sigma \notin (S^* \cdot B) \\ E_{unif} & \text{falls } \sigma \in (S^* \cdot B) \text{ und } \sigma'[\|\sigma'\|] = \sigma[\|\sigma\|] \\ 0 & \text{sonst} \end{cases}$$

Für die induzierte uniforme CTMC können wir die Wahrscheinlichkeiten des Erreichbarkeitsproblems wie üblich berechnen, indem wir aus dem Vektor den Wert für unseren Startzustand s auswählen²⁸:

$$Pr^\eta(\text{Reach}_{\leq t}(s, B)) = \left(\sum_{n=0}^{\infty} \varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}_{\eta, B}^n \cdot i_B \right) (s)$$

Da wir die unendliche Summen im Allgemeinen nicht numerisch berechnen können, beschränken wir uns hier wie schon bei den CTMCs in Kapitel 2.4 auf die Approximation durch:

$$Pr_k^\eta(\text{Reach}_{\leq t}(s, B)) = \left(\sum_{n=0}^k \varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}_{\eta, B}^n \cdot i_B \right) (s)$$

Der Parameter k lässt sich unabhängig von Zustand und Scheduler abschätzen, sodass eine gewünschte Genauigkeit $\varepsilon \in (0, 1)$ eingehalten werden kann: Da wir wissen, dass alle Einträge der Wahrscheinlichkeitsmatrix und die Werte des Vektors i_B im Intervall $[0, 1]$ liegen folgt, dass sich als Ergebnis der Matrixmultiplikation $\mathbf{P}_{\eta, B}^n \cdot i_B$ wieder ein Vektor mit Werten aus $[0, 1]$ ergibt. Diesen Vektor schätzen wir durch den Eins-Vektor i_S nach oben ab:

$$\begin{aligned} \left(\sum_{n=k+1}^{\infty} \varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}_{\eta, B}^n \cdot i_B \right) (s) &\leq \left(\sum_{n=k+1}^{\infty} \varphi(E_{unif} \cdot t, n) \cdot i_S \right) (s) \\ &= \sum_{n=k+1}^{\infty} \varphi(E_{unif} \cdot t, n) \\ &\leq \varepsilon \end{aligned}$$

Wir wissen, dass für die Poisson-Verteilung $\sum_{n=0}^{\infty} \varphi(\lambda, n) = 1$ gilt. Da außerdem $\varphi(\lambda, n) > 0$ für alle $n \in \mathbb{N}$ gilt, ist $f(\lambda, k) = \sum_{n=0}^k \varphi(\lambda, n)$ eine streng monoton

²⁸Für eine Startverteilung \tilde{a} erhält man die Wahrscheinlichkeiten, indem man für die Startverteilung und die unendliche Summe eine Matrixmultiplikation durchführt. Der Einfachheit halber beschränken wir uns hier auf einen Startzustand s .

steigende Funktion über \mathbb{N} . Um ein zu ε passendes k zu berechnen, müssen wir lediglich ein k mit $f(\lambda, k) \geq 1 - \varepsilon$ bestimmen, denn:

$$\begin{aligned} \sum_{n=k+1}^{\infty} \varphi(\lambda, n) &= (\sum_{n=0}^{\infty} \varphi(\lambda, n)) - (\sum_{n=0}^k \varphi(\lambda, n)) \\ &= 1 - f(\lambda, k) \\ &\leq \varepsilon \\ \text{gdw. } 1 - \varepsilon &\leq f(\lambda, k) \end{aligned}$$

Dies kann effizient implementiert werden, indem man die Poisson-Verteilung durch $\varphi(\lambda, n) = \varphi(\lambda, n-1) \cdot \frac{\lambda}{n}$ für alle $n > 0$ und $\varphi(\lambda, 0) = e^{-\lambda}$ rekursiv berechnet. Summiert man parallel zur Berechnung von $\varphi(\lambda, k)$ die Zwischenergebnisse $\varphi(\lambda, n)$ für $n = 0, \dots, k$ auf, so erhält man in jeder Rekursionsstufe den entsprechenden Wert $f(\lambda, n)$.

Die Approximation über eine feste maximale Pfadlänge ist auch aus einem weiteren Grund vorteilhaft. Für unendliche Pfade ist die Menge der extremen Scheduler wie schon früher erwähnt wurde abzählbar. Für abzählbar unendliche Mengen ist die Existenz eines Maximums oder eines Minimums jedoch nicht zwingend gegeben, weswegen wir etwa in Abschnitt 3.7 von Supremum und Infimum sprechen mussten. Durch die Beschränkung der maximalen Pfadlänge auf k wird es nun unnötig Scheduler zu unterscheiden, die in den ersten k Schritten dieselben Entscheidungen treffen. Es genügt also für eine Menge solcher Scheduler lediglich einen Repräsentanten zu betrachten. Die Menge der benötigten extremen Scheduler bezüglich der maximalen Pfadlänge k , die wir im Folgenden mit \mathcal{ES}_k bezeichnen wollen, ist dann nur noch endlich. Dann können wir auch statt von Supremum und Infimum wieder von Maximum und Minimum sprechen.

Bevor wir zum Algorithmus zur Berechnung der Wahrscheinlichkeiten für das quantitative Erreichbarkeitsproblem kommen, wollen wir zunächst noch einmal ein Beispiel betrachten. Dieses Beispiel werden wir anschließend zu einer ACTMC abändern und untersuchen, wie wir darin die Berechnungen vernünftigerweise durchführen können.

Beispiel 25 (Erreichbarkeit bei CTMCs). Sei die uniforme zeitstetige Markov-Kette \mathcal{M} mit $E_{unif} = 1$ wie in Abbildung 35 gegeben, s_0 der Startzustand und $B = \{s_B\}$. Zur Vereinfachung nehmen wir an, dass die gepunkteten Pfeile niemals in den Zustand s_B führen können. Die dadurch angedeuteten Pfade können wir damit bei der Betrachtung des Erreichbarkeitsproblems außen vor lassen.

Im diesem Beispiel bezeichnen wir mit q_i^s die Wahrscheinlichkeit für einen Pfad der Länge i der bisher in s endet, noch mit maximal $3 - i$ Schritten von Zustand s aus in den Zustand s_B zu gelangen. \mathbf{P}_B bezeichnet die Wahrscheinlichkeitsmatrix der bezüglich B transformierten Markov-Kette. Die Werte lassen sich rekursiv berechnen durch:

$$q_{i-1}^s = \begin{cases} \mathbf{P}_B \cdot q_i^s & \text{falls } s \in S \setminus B \\ \varphi(E_{unif} \cdot t, i-1) + q_i^s & \text{falls } s \in B \end{cases} \quad \text{für } i \in \{1, 2, 3\}$$

und $q_4^s = 0$ für alle $s \in S$

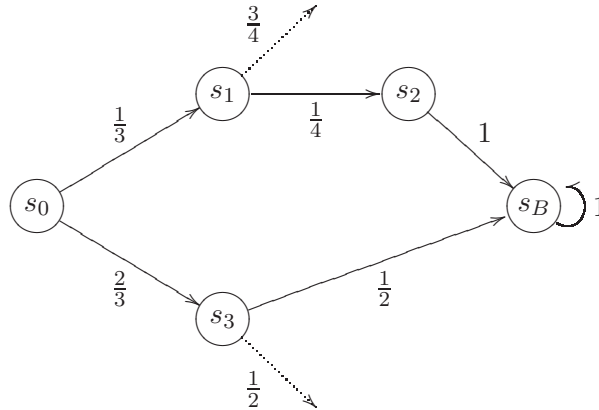


Abbildung 35: Erreichbarkeitsanalyse bei CTMCs

Wir benutzen wegen der besseren Lesbarkeit nun die folgende Abkürzungen:

$$\varphi_j^k = \sum_{i=j}^k \varphi(E_{unif} \cdot t, i) \text{ und } \varphi_j = \varphi_j^j$$

$$\begin{aligned} q_i &= (& q_i^{s_0}, & q_i^{s_1}, & q_i^{s_2}, & q_i^{s_3}, & q_i^{s_B}, & \dots) \\ q_3 &= (& 0, & 0, & 0, & 0, & \varphi_3, & \dots) \\ q_2 &= (& 0, & 0, & 1 \cdot \varphi_3, & \frac{1}{2} \cdot \varphi_3, & \varphi_2^3, & \dots) \\ q_1 &= (& \frac{2}{3} \cdot \frac{1}{2} \cdot \varphi_3, & \frac{1}{4} \cdot 1 \cdot \varphi_3, & 1 \cdot \varphi_2^3, & \frac{1}{2} \cdot \varphi_2^3, & \varphi_1^3, & \dots) \\ q_0 &= (& \frac{2}{3} \cdot \frac{1}{2} \cdot \varphi_2^3 + \frac{1}{3} \cdot \frac{1}{4} \cdot 1 \cdot \varphi_3, & \frac{1}{4} \cdot 1 \cdot \varphi_2^3, & 1 \cdot \varphi_1^3, & \frac{1}{2} \cdot \varphi_1^3, & \varphi_0^3, & \dots) \end{aligned}$$

In diesem Beispiel würde sich das Ergebnis bei der Betrachtung weiterer Schritte *nur* noch um den Faktor φ_4^k verändern. Umso größer die betrachtete Zeitspanne ist, umso größer wird die Wahrscheinlichkeit, dass die längeren Pfade signifikante Beiträge zum Ergebnis liefern. Die kumulierten Poisson-Wahrscheinlichkeiten liefern wie zuvor beschrieben ein Abbruchkriterium, das von der gewünschten Genauigkeit abhängig ist.

Beispiel 26 (Erreichbarkeit bei ACTMCs). Ändern wir das Beispiel wie folgt zu einer abstrakten Markov-Kette mit Exitraten $[1, 1]$ in allen Zuständen ab²⁹:

Hier betrachten wir also eine abstrakte uniforme Markov-Kette, in der ein Scheduler ausschließlich in s_0 eine echte Wahl treffen kann. Um einen Scheduler η zu finden, der die Wahrscheinlichkeit minimiert, von den Zuständen $s \in S \setminus B$ aus s_B zu erreichen, müssen wir jeweils bei der Berechnung von q_{i-1} die Gleichung $\sum_{s' \in S} \mathbf{P}_{\eta, B}(s, s') \cdot q_i^{s'}$ minimieren. In diesem Beispiel ergeben sich damit neue Werte für $q_1^{s_0}$ und $q_0^{s_0}$. Da $\frac{1}{2} \cdot \varphi_2^3 = q_1^{s_3} > q_1^{s_1} = \frac{1}{4} \cdot 1 \cdot \varphi_3$ und $\frac{1}{2} \cdot \varphi_3 = q_2^{s_3} > q_2^{s_1} = 0$, werden $q_0^{s_0}$ und $q_1^{s_0}$

²⁹Der Übersichtlichkeit halber wurden die Punktintervalle nicht ausgeschrieben. Korrekterweise müssten alle Kanten mit Intervallen beschriftet werden. Ebenso fehlt die Beschriftung mit Wahrscheinlichkeitsintervallen. Da die Exitrate mit $E_{unif} = 1$ gegeben ist, stimmen jedoch die Wahrscheinlichkeiten mit den Raten überein.

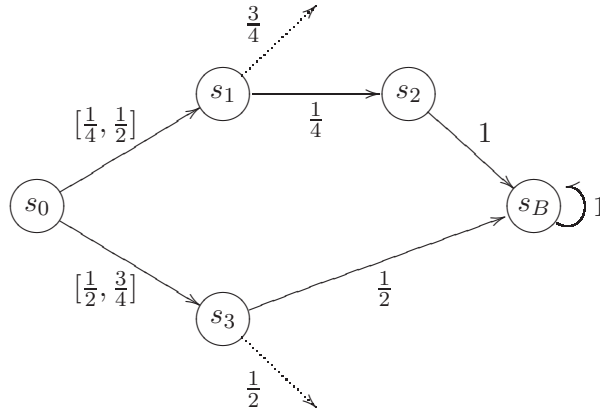


Abbildung 36: Erreichbarkeitsanalyse bei ACTMCs

minimiert, wenn mit möglichst großer Wahrscheinlichkeit der Pfad über s_1 und s_2 gewählt wird:

$$\begin{aligned}
 q_1 &= (\quad \quad \quad \frac{1}{2} \cdot \frac{1}{2} \cdot \varphi_3, \quad \frac{1}{4} \cdot 1 \cdot \varphi_3, \quad 1 \cdot \varphi_2^3, \quad \frac{1}{2} \cdot \varphi_2^3, \quad \varphi_1^3, \quad \dots) \\
 q_0 &= (\quad \frac{1}{2} \cdot \frac{1}{2} \cdot \varphi_2^3 + \frac{1}{2} \cdot \frac{1}{4} \cdot 1 \cdot \varphi_3, \quad \frac{1}{4} \cdot 1 \cdot \varphi_2^3, \quad 1 \cdot \varphi_1^3, \quad \frac{1}{2} \cdot \varphi_1^3, \quad \varphi_0^3, \quad \dots)
 \end{aligned}$$

Wie das Beispiel nahelegt werden wir nun einen Algorithmus entwickeln, der auf einer Rückwärtsanalyse basiert und in jedem Iterationsschritt die Wahrscheinlichkeit minimiert, in einem Schritt in einen Zustand zu gelangen, von dem aus mit möglichst hoher Wahrscheinlichkeit ein Zustand in der Zielmenge erreicht werden kann. Durch die Rückwärtsanalyse sind letztere Wahrscheinlichkeiten immer bekannt.

Grundsätzlich suchen wir nach einem HD-Scheduler, der die Wahrscheinlichkeit minimiert, in höchstens $k \in \mathbb{N}$ Schritten einen Zustand in B zu erreichen. Wie wir in Satz 5 gesehen haben, ist dafür die Menge der extremen Scheduler ebensogut geeignet, wie die Menge aller HD-Scheduler. Da Scheduler zu einer gegebenen Markov-Kette \mathcal{M} in jedem Schritt andere Werte aus den Intervallen wählen können, wollen wir eine Bezeichnung für die im i -ten Schritt gewählte Wahrscheinlichkeitsmatrix eines minimalen Schedulers einführen.

Definition 31 (Wahrscheinlichkeitsmatrizen eines minimalen Schedulers).

Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine zeitstetige abstrakte Markov-Kette und $k \in \mathbb{N}$ die maximale betrachtete Pfadlänge. Dann ist \mathbf{P}_i^B die im i -ten Schritt von einem minimierenden Scheduler $\eta_B \in \mathcal{ES}_k(\mathcal{M})$ gewählte Wahrscheinlichkeitsmatrix, wobei η_B die Wahrscheinlichkeit minimiert, in maximal k Schritten eine Zielmenge $B \subseteq S$ zu erreichen.

Wenn aus dem Kontext hervorgeht, welche Menge die Zielmenge darstellt, so schreiben wir statt \mathbf{P}_i^B auch einfach \mathbf{P}_i .

Damit können wir nun eine formale Definition von q_i bezüglich einer maximalen Pfadlänge k geben, wobei Pfade der Länge 0 aufgrund der fehlenden Transition gesondert behandelt werden:

$$\begin{aligned}
q_0 &= \varphi_0 \cdot i_B + q_1 \\
q_i &= \sum_{n=i}^k (\varphi_n \cdot \mathbf{P}_i \cdot \dots \cdot \mathbf{P}_n \cdot i_B) \\
&= \varphi_i \cdot \mathbf{P}_i \cdot i_B + \mathbf{P}_i \cdot \sum_{n=i+1}^k (\varphi_n \cdot \mathbf{P}_{i+1} \cdot \dots \cdot \mathbf{P}_n \cdot i_B) \\
&= \varphi_i \cdot \mathbf{P}_i \cdot i_B + \mathbf{P}_i \cdot q_{i+1} \\
&\text{für } i \in \{1, \dots, k\} \text{ und } q_{k+1}(s) = 0 \text{ für alle } s \in S.
\end{aligned}$$

Der erste Summand in der Definition von q_i entspricht dabei der Wahrscheinlichkeit für einen Pfad der Länge $i - 1$ im nächsten Schritt und innerhalb von t Zeiteinheiten in einen Zustand $s_B \in B$ zu gelangen. Der zweite Summand drückt die Wahrscheinlichkeit aus, von den möglichen Nachfolgezuständen aus einen B -Zustand zu erreichen. In $q_0(s)$ können wir dann die Wahrscheinlichkeit ablesen, von s aus in maximal t Zeiteinheiten einen Zustand aus B zu erreichen. Da bei der Berechnung von einer maximalen Pfadlänge k ausgegangen wird, wird auf diese Weise der Wert jedoch nur approximiert.

Die Frage die sich nun stellt ist die, wie wir an die Matrizen \mathbf{P}_i gelangen können. Dazu untersuchen wir den Vektor q_i etwas genauer bezüglich der einzelnen Komponenten:

$$\begin{aligned}
q_i(s) &= (\varphi_i \cdot \mathbf{P}_i \cdot i_B + \mathbf{P}_i \cdot q_{i+1})(s) \\
&= (\varphi_i \cdot (\mathbf{P}_i \cdot i_B))(s) + (\mathbf{P}_i \cdot q_{i+1})(s) \\
&= (\varphi_i \cdot (\mathbf{P}_i(s, \cdot) \cdot i_B) + (\mathbf{P}_i(s, \cdot) \cdot q_{i+1})) \\
&= \min_{\eta \in \mathcal{ES}_k(\mathcal{M})} (\varphi_i \cdot \mathbf{P}_{\eta, B}(s, \cdot) \cdot i_B + \mathbf{P}_{\eta, B}(s, \cdot) \cdot q_{i+1}) \quad (\text{Def. 31}) \\
&= \min_{\eta \in \mathcal{ES}_k(\mathcal{M})} (\sum_{s' \in S} \varphi_i \cdot \mathbf{P}_{\eta, B}(s, s') \cdot i_B(s') \\
&\quad + \sum_{s' \in S} \mathbf{P}_{\eta, B}(s, s') \cdot q_{i+1}(s')) \\
&= \min_{\eta \in \mathcal{ES}_k(\mathcal{M})} (\sum_{s' \in S} (\varphi_i \cdot \mathbf{P}_{\eta, B}(s, s') \cdot i_B(s') \\
&\quad + \mathbf{P}_{\eta, B}(s, s') \cdot q_{i+1}(s'))) \\
&= \min_{\eta \in \mathcal{ES}_k(\mathcal{M})} (\sum_{s' \in S} \mathbf{P}_{\eta, B}(s, s') \cdot \underbrace{(\varphi_i \cdot i_B(s') + q_{i+1}(s'))}_{=p(s')}) \\
&= \min_{\eta \in \mathcal{ES}_k(\mathcal{M})} (\mathbf{P}_{\eta, B}(s, \cdot) \cdot p)
\end{aligned}$$

Der Vektor p ist zum Zeitpunkt der Berechnung bekannt und enthält die Wahrscheinlichkeiten, von den einzelnen Zuständen aus in einen B -Zustand zu gelangen. Geht man davon aus, dass die in p verwendeten Wahrscheinlichkeiten durch einen minimalen Scheduler gewählt wurden, so ist es nun für jeden Zustand $s \in S$ möglich, einen minimalen Scheduler für $\mathbf{P}_{\eta, B}(s, \cdot) \cdot p$ zu berechnen. Dazu teilt man den Nachfolgezuständen $s' \in S$ mit den größten Werten $p(s')$ die kleinstmöglichen Wahrscheinlichkeiten zu. Würde man eine Wahrscheinlichkeit von $\delta > 0$ dem Nachfolgezustand $s' \in S$ zuordnen, obwohl die untere Intervallgrenze des Zustands s' damit überschritten würde, so würde man die Wahrscheinlichkeit einen Zustand aus der Menge B zu erreichen stärker vergrößern, als wenn die Wahrscheinlichkeit einem anderen Nachfolger $s'' \in S$ mit $p(s'') > p(s')$ zugeordnet würde³⁰. Damit sind

³⁰Es ist hierbei zu beachten, dass die Verwendung von Schemulern bezüglich abstrakten zeitsteti-

wir also nicht mehr auf einen *brute-force* Ansatz angewiesen, sondern können einen minimierenden Scheduler direkt berechnen.

Wir wollen nun ein Ergebnis festhalten, welches im Grunde klar sein sollte wegen der Tatsache, dass wir eine Rückwärtsanalyse verwenden. Dabei handelt es sich um die Feststellung, dass man einen Scheduler der minimierend ist für Pfade mit maximaler Länge $k + 1$ aus einem Scheduler konstruieren kann, der minimierend ist für Pfade mit maximaler Länge k . Wir werden dieses Ergebnis im Beweis zum Hauptsatz dieses Kapitels noch benötigen, weswegen wir es in folgendem Lemma festhalten.

Lemma 9 (Rekursive Bestimmung minimierender Scheduler).

Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette und die Scheduler $\eta_B^k \in \mathcal{ES}_k(\mathcal{M})$ für $k \in \mathbb{N}$ seien solche, die die Wahrscheinlichkeit minimieren, die Menge $B \subseteq S$ mit maximaler Pfadlänge k zu erreichen. Die im i -ten Schritt von einem Scheduler η_B^k gewählte Wahrscheinlichkeitsmatrix sei \mathbf{P}_i und die für einen Scheduler η_B^{k+1} sei \mathbf{P}'_i . Dann gilt aufgrund der rekursiven Definition, dass es für k und $k + 1$ minimierende Scheduler η_B^k und η_B^{k+1} gibt, sodass gilt:

$$\mathbf{P}_i = \mathbf{P}'_{i+1} \text{ für alle } i \in \{1, \dots, k\}$$

Beweis. Wir führen den Beweis per Induktion über k . Für den Induktionsanfang $k = 0$ gibt es keine zu erfüllende Bedingung, da kein $i \in \{\}$ existiert.

Für den Induktionsschritt $k \rightarrow k+1$ betrachten wir nun q_0 und q'_0 , die per Definition gegeben sind als:

$$\begin{aligned} q'_0 &= \varphi_0 \cdot i_B + \sum_{n=1}^{k+1} (\varphi_n \cdot \mathbf{P}'_1 \cdot \dots \cdot \mathbf{P}_n \cdot i_B) \\ &= \varphi_0 \cdot i_B + \varphi_1 \cdot \mathbf{P}'_1 \cdot i_B + \mathbf{P}'_1 \cdot \sum_{n=2}^{k+1} (\varphi_n \cdot \mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_n \cdot i_B) \\ &= \varphi_0 \cdot i_B + \mathbf{P}'_1 \cdot (\varphi_1 \cdot i_B + \sum_{n=2}^{k+1} (\varphi_n \cdot \mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_n \cdot i_B)) \\ &= \varphi_0 \cdot i_B + \mathbf{P}'_1 \cdot \underbrace{(\varphi_1 \cdot i_B + \varphi_2 \cdot \mathbf{P}'_2 \cdot i_B + \dots + \varphi_{k+1} \cdot \mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{k+1} \cdot i_B)}_{=a} \\ q_0 &= \varphi_0 \cdot i_B + \sum_{n=1}^k (\varphi_n \cdot \mathbf{P}_1 \cdot \dots \cdot \mathbf{P}_n \cdot i_B) \\ &= \varphi_0 \cdot i_B + \varphi_1 \cdot \mathbf{P}_1 \cdot i_B + \dots + \varphi_k \cdot \mathbf{P}_1 \cdot \dots \cdot \mathbf{P}_k \cdot i_B \end{aligned}$$

Der Summand $\varphi_0 \cdot i_B$ ist unabhängig vom Scheduler und kann daher hier vernachlässigt werden. Ein minimierender Scheduler muss das Matrixprodukt $\mathbf{P}'_1 \cdot a$ minimieren. Wir haben bereits gesehen, wie dieses Matrixprodukt minimiert werden kann, sofern a bekannt ist und von einem minimierenden Scheduler bestimmt wurde. Vergleicht man q_0 mit a so fällt sofort ins Auge, dass beide Terme abgesehen von den um 1 verschobenen Indizes der Matrizen \mathbf{P}_i mit $i \in \{1, \dots, k\}$ und \mathbf{P}'_i mit $i \in \{2, \dots, k+1\}$ genau gleich sind. Da wir als Induktionsvoraussetzung von q_0 verlangen, dass es durch einen minimierenden Scheduler bestimmt wurde, können wir jeweils \mathbf{P}'_{i+1} durch \mathbf{P}_i ersetzen. Die Wahrscheinlichkeitsmatrix \mathbf{P}'_1 muss dann nur noch so bestimmt werden, dass das Matrixprodukt $\mathbf{P}'_1 \cdot a$ minimal ist.

gen Markov-Kette immer auch die Anwendung des *cut* impliziert.

Damit wissen wir also, wie ein minimierender Scheduler für maximale Pfadlänge $k + 1$ angegeben werden kann, der in den letzten k Schritten gleich ist zu einem minimalen Scheduler für maximale Pfadlänge k . Die Behauptung des Lemmas ist also gültig. □

4.3 3-CSL Model Checking

Model Checking für 3-CSL auf abstrakten zeitstetigen Markov-Ketten stellt eine deutlich schwierigere Aufgabe dar als das für CSL auf CTMCs. In [BHHK03] wird gezeigt, wie letzteres effizient gelöst werden kann. Leider ist dieses Verfahren nicht ohne weiteres auf Model Checking für ACTMCs übertragbar. Im Moment ist noch nicht klar, wie man für Intervalle der Form $[t, t']$ mit $0 < t \leq t'$ die minimale oder maximale Wahrscheinlichkeiten dafür bestimmen kann, eine bestimmte Zielmenge zu erreichen.

Auch mit Markov-Entscheidungsprozessen kommt man hier vorerst nicht weiter: Wie wir schon im vorigen Kapitel festgestellt haben genügen extreme Scheduler, falls nur maximale und minimale Wahrscheinlichkeiten von Interesse sind. Zusammen mit der Endlichkeit der Zustandsmenge der Markov-Ketten können von extremen Schemulern nur endlich viele verschiedene Kombinationen von Raten gewählt werden. Betrachten wir nun jede Wahl einer dieser Kombinationen als eine Aktion, so kann man abstrakte zeitstetige Markov-Ketten als zeitstetige Markov-Entscheidungsprozesse auffassen (siehe Kapitel 2.6). Model Checking für ACTMCs ließe sich also (zumindest teilweise) auf Model Checking für CTMDPs reduzieren. Abgesehen vom Erreichbarkeitsproblem für uniformisierte CTMDPs scheint bisher jedoch nur wenig in diesem Bereich erforscht zu sein, sodass eine Reduktion auf CTMDPs zu diesem Zeitpunkt nicht sehr hilfreich ist.

Aufbauend auf den Erkenntnissen aus dem vorigen Kapitel werden wir nun ein Model Checking Verfahren für 3-CSL entwickeln. Wir beschränken uns dabei wegen des Algorithmus für das Erreichbarkeitsproblem auf uniforme ACTMCs. Weiterhin beschränken wir uns auf *Until* mit Intervallen der Form $[0, t]$ mit $t \in \mathbb{R}_{>0}$. Diese Beschränkung ist deswegen nötig, da die extremen Scheduler ansonsten nicht ausreichend wären, wie wir in Beispiel 24 gesehen haben.

Der aussagenlogischen Teil des Model Checking verläuft auch hier wie üblich, nur dass die Semantik der dreiwertigen Logik verwendet werden muss, um mit atomaren Eigenschaften mit unbestimmter Erfüllbarkeit umgehen zu können. Model Checking für den *Next*-Operator ist relativ simpel, da die Semantik sich weitestgehend auf die der möglichen Nachfolgezustände zurückführen lässt. Als Spezialfall müssen absorbierende Zustände s behandelt werden, von denen aus keine Pfade $\sigma \in Pfd_s$ mit Länge $|\sigma| \geq 1$ existieren, weswegen auch kein solcher Pfad eine *Next*-Eigenschaft erfüllen kann. Ansonsten bestimmt man die Wahrscheinlichkeit für Pfade mit Eigenschaft $\mathcal{X}^I \psi$ auf die übliche Weise, indem man alle Zustände mit der Eigenschaft ψ bestimmt und anschließend die Vorgänger aller Zustände betrachtet. Abhängig

vom Intervall I und der Exitrate des Vorgängerzustands s , kann dann die Wahrscheinlichkeit $X(s, I)$ berechnet werden, den Zustand s zu verlassen. Zusammen mit der kumulierten Wahrscheinlichkeit zu einem nachfolgenden Zustand v überzugehen in dem ψ gilt, erhält man dann die Wahrscheinlichkeit bezüglich der ausgehenden Pfade von s , dass diese $\mathcal{X}^I\psi$ erfüllen. Die vom Scheduler η gewählte Verteilung in s bezeichnen wir mit $\bar{\mu} \in \text{distr}(\mathbf{P}, \cdot)$ und erhalten damit:

$$Pr^\eta\{\sigma \in \text{Pfad}_s \mid \llbracket \sigma, \mathcal{X}^I\psi \rrbracket = \top\} = X(s, I) \cdot \sum_{v \in \text{Sat}(\psi)} \bar{\mu}(v)$$

Da extreme Scheduler zur Berechnung der unteren und oberen Wahrscheinlichkeitsgrenzen ausreichen, kann man diese entweder mit dem *brute-force* Ansatz oder über das in Kapitel 4.2 beschriebene Verfahren bestimmen.

Korollar 3. Sei $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ eine abstrakte zeitstetige Markov-Kette. Mit $Pr^l(s, \mathcal{X}^I\psi, \alpha)$ und $Pr^u(s, \mathcal{X}^I\psi, \alpha)$ mit $s \in S$, $\psi \in \mathbb{S}_{\mathcal{M}}$, $\alpha \in \mathbb{B}$ und $I \in \mathcal{I}$ sind die Grenzen der Wahrscheinlichkeit für die Erfüllbarkeit bzw. Unerfüllbarkeit von $\mathcal{X}^I\psi$ bezeichnet. Dann gilt für alle $\bar{\mu} \in \text{distr}(\mathbf{P}(s, \cdot))$:

$$\begin{aligned} Pr^l(s, \mathcal{X}^I\psi, \top) &\leq X(s, I) \cdot \sum_{v \in \text{Sat}(\psi)} \bar{\mu}(v) \leq Pr^u(s, \mathcal{X}^I\psi, \top) \\ Pr^l(s, \mathcal{X}^I\psi, \perp) &\leq X(s, I) \cdot \sum_{v \in \text{Fls}(\psi)} \bar{\mu}(v) \leq Pr^u(s, \mathcal{X}^I\psi, \perp) \end{aligned}$$

Für Model Checking von *Until*-Formeln wollen wir die Situation etwas genauer untersuchen: *Until*-Eigenschaften können wie im Fall von CTMCs auf das Erreichbarkeitsproblem reduziert werden. Ist die Formel von der Form $\text{true } \mathcal{U}^I\psi$, so ist die Reduktion offensichtlich. Es muss lediglich als Zielmenge $\text{Sat}(\psi)$ gewählt werden. Die Zustandseigenschaft *true* gilt per Definition in allen Zuständen, sodass bezüglich eines Zustands s für die Wahrscheinlichkeit von Pfaden $\sigma \in \text{Pfad}_s$ mit $\llbracket \sigma, \text{true } \mathcal{U}^I\psi \rrbracket = \top$ gilt:

$$Pr^l(s, \text{true } \mathcal{U}^I\psi, \top) = \inf_{\eta \in \mathcal{S}(\mathcal{M})} Pr^\eta(\text{Reach}_{\leq \sup I}^{\mathcal{M}}(s, \text{Sat}(\psi)))$$

Nun führen wir den Fall für *Until*-Formeln der Form $\psi_1 \mathcal{U}^I \psi_2$ durch eine Transformation der gegebenen Markov-Kette auf den obigen Fall zurück:

Damit die *Until*-Formel für einen gegebenen Pfad erfüllt ist, muss in den Zuständen des Pfades vor dem ersten ψ_2 -Zustand immer ψ_1 gelten. Dass es keine anderen Pfade geben kann, die zu einem ψ_2 Zustand führen, kann dadurch sichergestellt werden, dass von Zuständen s' der gegebenen Markov-Kette \mathcal{M} mit $\llbracket s', \psi_1 \rrbracket \neq \top$ und $\llbracket s', \psi_2 \rrbracket \neq \top$ sämtliche ausgehenden Kanten entfernt werden. Alle Zustände von denen aus $\psi_1 \mathcal{U} \psi_2$ nicht erfüllt werden kann werden also zu absorbierenden Zuständen, sodass auf allen Pfaden der transformierten Markov-Kette \mathcal{M}' , abgesehen vom letzten Zustand des Pfades, entweder ψ_1 oder ψ_2 gelten muss. Falls der gesamte Pfad keinen Zustand mit ψ_2 enthält, so muss die *Until*-Formel auf diesem Pfad ungültig sein.

Bezüglich eines Zustands s entspricht die Wahrscheinlichkeit von Pfaden $\sigma \in \text{Pfad}_s^{\mathcal{M}}$ mit $\llbracket \sigma, \psi_1 \mathcal{U}^I \psi_2 \rrbracket = \top$ also der von Pfaden $\sigma' \in \text{Pfad}_s^{\mathcal{M}'}$ mit $\llbracket \sigma', \text{true } \mathcal{U}^I \psi_2 \rrbracket = \top$ in der transformierten Markov-Kette:

$$Pr^l(s, \psi_1 \mathcal{U}^I \psi_2, \top) = \inf_{\eta \in \mathcal{S}(\mathcal{M}')} Pr^\eta(Reach_{\leq \sup I}^{\mathcal{M}'}(s, Sat(\psi)))$$

Wir können dieses Konzept algorithmisch umsetzen, indem wir die Zustände bezüglich der Kombinationen von Wahrheitswerten für ψ_1 und ψ_2 zunächst in verschiedene Mengen einordnen, dann die Transformation bezüglich dieser Mengen definieren und anschließend eine passende Erreichbarkeitsanalyse starten, wobei wir ebenfalls diese Mengen zur Hilfe nehmen werden.

Mit $W^{\top+}$ bezeichnen wir die Menge aller Zustände, in denen die *Until*-Formel definitiv erfüllt ist und mit $W^{\top-}$ die Menge derjenigen Zustände, in denen die Erfüllbarkeit nicht mehr zweifelsfrei nachgewiesen werden kann. Entsprechend bezeichnet $W^{\perp+}$ die Menge aller Zustände in denen die *Until*-Formel definitiv verletzt wird und $W^{\perp-}$ die Menge der Zustände, in denen eine Verletzung der Formel nicht mehr mit Sicherheit vorkommen kann. Im Fall einer zweiwertigen Logik würden $W^{\top+}$ und $W^{\perp-}$ sowie $W^{\top-}$ und $W^{\perp+}$ zusammenfallen. Wie wir nun sehen werden, unterscheiden sie sich jedoch in der hier verwendeten dreiwertigen Logik. Die Zuordnungen in den folgenden Abschnitten sind in Tabelle 6 zusammengefasst.

Betrachten wir zunächst die einfacheren Fälle: Falls ψ_2 in einem Zustand erfüllt ist, so ist natürlich auch $\psi_1 \mathcal{U}^I \psi_2$ erfüllt. Solche Zustände sind also in den beiden Mengen $W^{\top+}$ und $W^{\perp-}$ enthalten. Ein zweiter sehr einfacher Fall ist der mit $\llbracket s, \psi_1 \rrbracket = \llbracket s, \psi_2 \rrbracket = \perp$ für Zustand s . In diesem Fall wird die *Until*-Formel definitiv verletzt, sodass solche Zustände den Mengen $W^{\top-}$ und $W^{\perp+}$ zuzuordnen sind. Soweit stimmen die Mengen $W^{\top+}$ und $W^{\perp-}$ sowie $W^{\top-}$ und $W^{\perp+}$ wie im zweiwertigen Fall überein.

Wenn in einem Zustand ψ_2 ungültig oder unbestimmt ist und ψ_1 nicht definitiv erfüllt wird, ist die Erfüllbarkeit der *Until*-Formel nicht mehr mit Sicherheit nachweisbar, da der Fall ψ_1 *ist unbestimmt* im Zweifelsfall auch als unerfüllbar behandelt werden kann. Daher sind solche Zustände in der Menge $W^{\top-}$ enthalten. Außer wenn ψ_1 und ψ_2 definitiv nicht erfüllbar sind, was dem zweiten einfachen Fall entspricht, fallen die Mengen $W^{\top-}$ und $W^{\perp+}$ ab hier jedoch nicht mehr zusammen.

Für Zustände s mit $\llbracket s, \psi_2 \rrbracket = ?$ kann man im Zweifelsfall auch annehmen, dass ψ_2 erfüllbar ist, wodurch diese Zustände der Menge $W^{\perp-}$ zuzuordnen sind. Für $\llbracket s, \psi_2 \rrbracket = ?$ und $\llbracket s, \psi_1 \rrbracket \neq \top$ kann also keinerlei gesicherte Aussage über die Erfüllbarkeit oder die Unerfüllbarkeit der *Until*-Formel bezüglich Zustand s gemacht werden, da der Zustand dann sowohl $W^{\top-}$ als auch $W^{\perp-}$ zugeordnet ist.

Für die übrigen Kombinationen können keine weiteren Zugehörigkeiten festgestellt werden. Ist ψ_1 erfüllt und ψ_2 unbestimmt oder verletzt, so hängt die Erfüllbarkeit von $\psi_1 \mathcal{U}^I \psi_2$ von den nachfolgenden Zuständen ab. Ist ψ_1 erfüllt oder unbestimmt aber ψ_2 definitiv verletzt, so müssen zum Nachweis der Unerfüllbarkeit der *Until*-Formel ebenfalls die Nachfolger betrachtet werden.

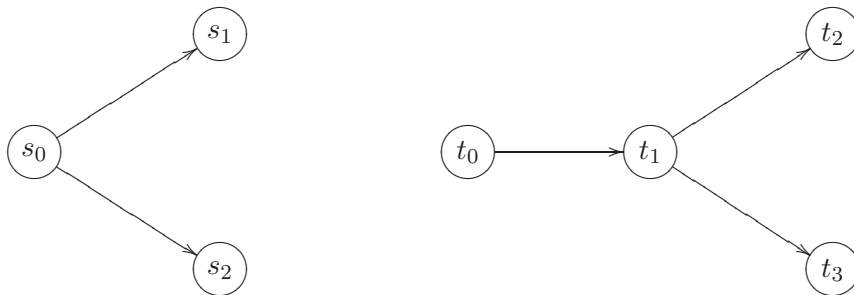
Da für einen einzigen gegebenen Pfad die Nachfolger und die Verweildauern jedes Zustands des Pfades bekannt sind, können wir den Pfad in den Zeitintervallgrenzen chronologisch nach Zuständen der W -Mengen durchsuchen. Da die Intervalle als untere Grenze den Wert 0 haben, muss man also immer beim Startzustand des Pfades

	ψ_2					
	\perp		$?$		\top	
\perp	\top^-	\perp^+	\top^-	\perp^-	\top^+	\perp^-
$\psi_1 ?$	\top^-		\top^-	\perp^-	\top^+	\perp^-
\top				\perp^-	\top^+	\perp^-

Tabelle 6: Zugehörigkeit zu W -Mengen für $Until$ -Formel $\psi_1 \mathcal{U}^I \psi_2$.

beginnen. Sobald man einen Zustand gefunden hat, der in einer dieser Mengen enthalten ist, kann man Aussagen über die Erfüllbarkeit bzw. über die Unerfüllbarkeit der $Until$ -Formel machen.

Beispiel 27 ($Until$ auf einzelnen Pfaden). Wir wollen uns nun anhand von zwei Beispielen verdeutlichen, wie die W -Mengen zur Berechnung von $Until$ -Eigenschaften verwendet werden können. Der Einfachheit halber verzichten wir auf ein Zeitintervall, wodurch auch die Raten vernachlässigt werden können. Was wir hier betrachten sind im Grunde also nur Transitionssysteme. Die Gültigkeit der Eigenschaften ψ_1 und ψ_2 in den jeweiligen Zuständen ist wie in der Tabelle in Abbildung 37 gegeben. Daraus lassen sich die Zugehörigkeiten zu den W -Mengen nach Tabelle 6 bestimmen.



s	s_0	s_1	s_2	t_0	t_1	t_2	t_3
$\llbracket s, \psi_1 \rrbracket$	\top	\perp	$?$	\top	$?$	\perp	$?$
$\llbracket s, \psi_2 \rrbracket$	$?$	\top	$?$	\perp	\perp	\perp	\top
$s \in W^{\top^+}$		✓					✓
$s \in W^{\top^-}$			✓		✓	✓	
$s \in W^{\perp^+}$						✓	
$s \in W^{\perp^-}$	✓	✓	✓				✓

Abbildung 37: $Until$ -Eigenschaft bezüglich Pfaden

Wir wollen nun für verschiedene Pfade σ_i die Eigenschaft $\psi_1\mathcal{U}\psi_2$ untersuchen:

- $\sigma_1 = s_0 \longrightarrow s_1$:

Zunächst stellen wir fest, dass der Pfad bereits aufgrund von $s_0 \in W^{\perp-}$ nicht mit Sicherheit die Eigenschaft $\psi_1\mathcal{U}\psi_2$ verletzen kann. Dies liegt daran, dass der Wert von ψ_2 in s_0 unbestimmt ist.

Da s_0 weder in $W^{\top+}$ noch in $W^{\top-}$ enthalten ist, müssen wir also den nachfolgenden Zustand s_1 des Pfades betrachten. Dieser ist in $W^{\top+}$ enthalten, woraus folgt, dass $\llbracket\sigma_1, \psi_1\mathcal{U}\psi_2\rrbracket = \top$

- $\sigma_2 = s_0 \longrightarrow s_2$:

Dass die erste Feststellung bezüglich Pfad σ_1 nicht automatisch bedeutet, dass $\llbracket\sigma, \psi_1\mathcal{U}\psi_2\rrbracket = \top$ für alle Pfade σ mit $\sigma[0] = s_0$ gilt, zeigen wir nun anhand dieses Pfades. Wir untersuchen wieder den Nachfolger von s_0 , der diesmal s_2 ist, und stellen fest, dass er in der Menge $W^{\top-}$ enthalten ist. Insgesamt ergibt sich also $\llbracket\sigma_2, \psi_1\mathcal{U}\psi_2\rrbracket = ?$.

- $\sigma_3 = t_0 \longrightarrow t_1 \longrightarrow t_2$:

Der Zustand t_0 ist in keiner W -Menge enthalten. Die Semantik von $\llbracket\sigma_3, \psi_1\mathcal{U}\psi_2\rrbracket$ entspricht also genau der von $\llbracket t_1 \longrightarrow t_2, \psi_1\mathcal{U}\psi_2\rrbracket$ ³¹. Da ψ_2 in t_1 nicht gilt und ψ_2 unbestimmt ist, ist die Erfüllung der *Until*-Formel nicht möglich. Was jedoch möglich ist, ist die sichere Verletzung von $\psi_1\mathcal{U}\psi_2$. Wäre ψ_1 in t_1 verletzt, so wäre $\psi_1\mathcal{U}\psi_2$ automatisch auch verletzt. Ist das Gegenteil der Fall, so müssen wir den nächsten Zustand des Pfades betrachten. Da dieser in $W^{\perp+}$ enthalten ist, wissen wir also, dass die *Until*-Formel für den Pfad σ_3 nicht erfüllt ist.

- $\sigma_4 = t_0 \longrightarrow t_1 \longrightarrow t_3$:

Die Beobachtungen bezüglich des Pfadanfangs aus σ_3 gelten natürlich auch hier. Da die *Until*-Formel schon ab t_1 nicht mehr erfüllt werden kann, hilft es auch nichts, dass t_3 in $W^{\top+}$ enthalten ist. Da der Zustand aber auch in $W^{\perp-}$ enthalten ist, gilt im Gegensatz zu σ_3 für diesen Pfad nicht $\llbracket\sigma_4, \psi_1\mathcal{U}\psi_2\rrbracket = \perp$ sondern $\llbracket\sigma_4, \psi_1\mathcal{U}\psi_2\rrbracket = ?$.

Die Erweiterung der Pfade durch Verweildauern und der *Until*-Formeln durch Intervalle ist nur um weniges aufwändiger. Da nur Intervalle der Form $[0, t]$ mit $t \in \mathbb{R}_{>0}$ zugelassen sind, kann die Vorgehensweise von den Transitionssystemen übernommen werden. Man muss lediglich den gegebenen Pfad hinter dem letzten Zustand abschneiden, der noch vor dem Zeitpunkt t besucht wird. An den abgeschnittenen Pfad fügt man einen *virtuellen* Zustand³² an, der die abgeschnittenen Zustände

³¹Achtung: Betrachtet man zeitstetige Markov-Ketten, so ist dies nicht zwingend der Fall. Pfade haben in diesem Zusammenhang zusätzlich pro Zustand (außer dem letzten) noch Verweildauern, sodass bezüglich einer intervallbeschränkten *Until*-Formel die Semantik von σ_3 und von $t_1 \longrightarrow t_2$ unterschiedlich sein kann.

³²Mit einem virtuellen Zustand bezeichnen wir hier einen Zustand, der in der ursprünglichen Markov-Kette nicht enthalten ist und ausschließlich zur Vereinfachung der Verifikation benötigt wird.

repräsentiert. In diesem Zustand darf ψ_1 und ψ_2 nicht erfüllt sein. Wenn also nicht bereits früher auf dem Pfad die Erfüllbarkeit der *Until*-Formel entschieden wurde, wird sie wie gewünscht als unerfüllbar erkannt, sobald der neu angehängte letzte Zustand des Pfades erreicht wurde.

—

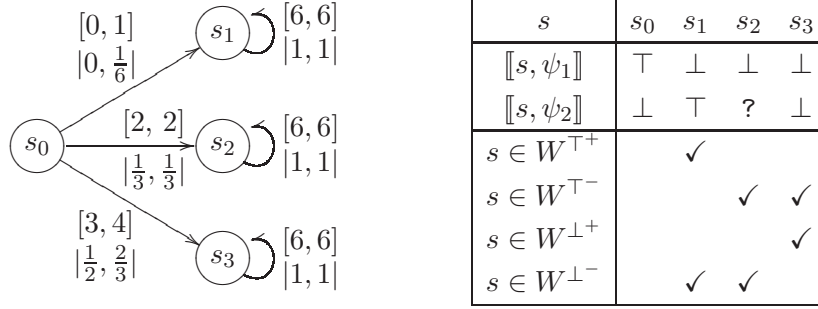
Statt Aussagen zu einzelnen Pfaden wollen wir uns nun Aussagen über alle Pfade, die von einem gegebenen Zustand ausgehen, zuwenden. Dazu werden wir eine Kombination aus W -Mengen und Erreichbarkeit in abstrakten zeitstetigen Markov-Ketten betrachten.

Für die Erfüllbarkeit einer *Until*-Formel bezüglich Pfaden sind nur die W^\top -Mengen von Bedeutung, bei der Unerfüllbarkeit dagegen sind es die W^\perp -Mengen. Mit den Mengen $W^{\top-}$ und $W^{\perp-}$ kann man die schon beschriebene Transformation durchführen, die bezüglich der gegebenen *Until*-Formel dieselben Ergebnisse liefert. Betrachten wir Erfüllbarkeit, so kann man für Zustände in denen die Eigenschaft verletzt wird, also die Zustände der Menge $W^{\top-}$, sämtliche ausgehenden Kanten entfernen. Bei der Unerfüllbarkeit entfernt man stattdessen die Zustände der Menge $W^{\perp-}$. Zustände, in denen die Erfüllbarkeit bzw. Unerfüllbarkeit sofort gegeben ist, also diejenigen, die in den Mengen $W^{\top+}$ bzw. $W^{\perp+}$ enthalten sind, können wie in Kapitel 4.2 beschrieben mit einem *self-loop* mit Wahrscheinlichkeit Eins ausgestattet werden nachdem ebenfalls alle ausgehenden Kanten entfernt wurden.

Sofern die dadurch erhaltene abstrakte zeitstetige Markov-Kette uniform ist, kann man nun die Erreichbarkeitsanalyse mit der Zielmenge $W^{\top+}$ bzw. $W^{\perp+}$ verwenden, um die minimalen Wahrscheinlichkeiten über der Menge aller ausgehenden Pfade des Zustands zu bestimmen, dass $\psi_1 \mathcal{U}^I \psi_2$ erfüllt ist. Wenn bereits die gegebene abstrakte zeitstetige Markov-Kette uniform war, so kann man durch Anfügen von *self-loops* an die absorbierenden Zustände und durch entsprechende Wahl der Rateintervalle aller durch die Transformation erhaltenen *self-loops* als $[E_{unif}, E_{unif}]$ wieder eine uniforme abstrakte Markov-Kette erhalten. Durch die Ausstattung der absorbierenden Zustände mit *self-loops* ändert man die Semantik bezüglich der gegebenen *Until*-Formel natürlich nicht, da sich der Wahrheitswert der Formel durch Zustandsübergänge über *self-loops* nicht mehr ändern kann.

Damit haben wir im Grunde auch schon die Frage nach einem Model Checking Algorithmus für $\mathcal{P}_{\boxtimes p}$ -Formeln geklärt. Einziger noch fehlender Baustein ist der Vergleich mit der gegebenen Wahrscheinlichkeit p . Dieser muss nach der Semantiktabelle 5 durchgeführt werden. Um die Prozedur zu veranschaulichen betrachten wir nun ein einfaches Beispiel.

Beispiel 28 (*Until*-Eigenschaften bezüglich Zuständen). Um den Umfang des Beispiels im Rahmen zu halten, betrachten wir eine uniforme ACTMC mit $E_{unif} = 6$, für die zum Nachweis der Gültigkeit bzw. Ungültigkeit der Spezifikation $\mathcal{P}_{\leq p}(\psi_1 \mathcal{U} \psi_2)$ jeweils *keine* Transformation durchgeführt werden muss. Die Gültigkeit der Eigenschaften ψ_1 und ψ_2 in den jeweiligen Zuständen ist wie in der Tabelle gegeben. Daraus lassen sich wieder die Zugehörigkeiten zu den W -Mengen nach Tabelle 6 bestimmen.

Abbildung 38: *Until*-Eigenschaften bezüglich Zuständen

Nach der Semantik von 3-CSL gilt:

$$\llbracket s_0, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U} \psi_2) \rrbracket = \begin{cases} \top & \text{falls } 1 - p \leq Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \perp\} \\ \perp & \text{falls } p < Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \top\} \\ ? & \text{sonst} \end{cases}$$

Wir müssen also die Wahrscheinlichkeiten $Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \perp\}$ und $Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \top\}$ berechnen. Da keine Transformation der Markov-Kette nötig ist, muss für den ersten Teil nur eine Erreichbarkeitsanalyse mit Zielmenge $W^{\perp+} = \{s_3\}$ durchgeführt werden. Ein Scheduler der die Wahrscheinlichkeit minimiert, in einem ersten Zustandsübergang nach s_3 zu gelangen, muss die Verteilung $(0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}) \in \text{distr}(\mathbf{P}(s_0, \cdot))$ wählen. Für $i \in \{1, 2, 3\}$ gibt es jeweils nur die eine Möglichkeit, $\bar{\mu}_{s_i} \in \text{distr}(\mathbf{P}(s_i, \cdot))$ mit $\bar{\mu}_{s_i}(s_i) = 1$ zu wählen. Zusammen ergeben die Verteilungen $\bar{\mu}_{s_i}$ für $i \in \{0, 1, 2, 3\}$ die Matrix \mathbf{P}_1 des minimalen Schedulers.

Da für $\bar{\mu}_{s_3}$ immer nur $(0, 0, 0, 1)$ gewählt werden kann, gilt also $\mathbf{P}_i(s_3) = (0, 0, 0, 1)$ für $i \in \mathbb{N}_{>0}$. Damit ergibt sich die Wahrscheinlichkeit $Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \perp\}$ über die unendliche Summen wie folgt:

$$\begin{aligned} & (\varphi(\infty, 0) \cdot i_{s_3} + \sum_{n=1}^{\infty} \varphi(\infty, n) \cdot \mathbf{P}_1 \cdot \dots \cdot \mathbf{P}_n \cdot i_{s_3})(s_0) \\ &= \varphi(\infty, 0) \cdot 0 + (\sum_{n=1}^{\infty} \varphi(\infty, n) \cdot \mathbf{P}_1 \cdot \dots \cdot \mathbf{P}_n \cdot i_{s_3})(s_0) \\ &= \sum_{n=1}^{\infty} \varphi(\infty, n) \cdot (0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}) \cdot \mathbf{P}_2 \cdot \dots \cdot \mathbf{P}_n \cdot (0, 0, 0, 1)^T \\ &= \sum_{n=1}^{\infty} \varphi(\infty, n) \cdot (0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}) \cdot (0, 0, 0, 1)^T \\ &= \sum_{n=1}^{\infty} \varphi(\infty, n) \cdot \frac{1}{2} \\ &= \frac{1}{2} \cdot \sum_{n=1}^{\infty} \varphi(\infty, n) \\ &= \frac{1}{2} \cdot (-\varphi(\infty, 0) + \sum_{n=0}^{\infty} \varphi(\infty, n)) \\ &= \frac{1}{2} \cdot \sum_{n=0}^{\infty} \varphi(\infty, n) \\ &= \frac{1}{2} \end{aligned}$$

Die Wahrscheinlichkeit $Pr^l\{\sigma \in Pfad_{s_0}^M \mid \llbracket \sigma, \psi_1 \mathcal{U} \psi_2 \rrbracket = \top\}$ erhält man auf gleiche Weise, nur dass die Zielmenge $W^{\top+} = \{s_1\}$ ist und der minimierende Scheduler daher die Verteilung $(0, 0, \frac{1}{3}, \frac{2}{3}) \in \text{distr}(\mathbf{P}(s_0, \cdot))$ wählt. Das Ergebnis der Berechnung ist 0, sodass gefolgert werden kann, dass:

$$\llbracket s_0, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U} \psi_2) \rrbracket = \begin{cases} \top & \text{falls } 1 - p \leq \frac{1}{2} \\ \perp & \text{falls } p < 0 \\ ? & \text{sonst} \end{cases} = \begin{cases} \top & \text{falls } p \geq \frac{1}{2} \\ ? & \text{falls } p < \frac{1}{2} \end{cases}$$

Nun wissen wir also, wie Model Checking für abstrakten zeitstetigen Markov-Kette durchgeführt werden kann. Wir wissen jedoch noch nicht, wie aus Spezifikationen bezüglich einer Abstraktion in Form einer ACTMC Rückschlüsse auf die ursprüngliche Markov-Kette gezogen werden können. Dass, sofern uns kein unbestimmtes Ergebnis vorliegt, die Gültigkeit und die Ungültigkeit einer Formel in der Abstraktion auch die Gültigkeit bzw. Ungültigkeit der Formel bezüglich der simulierten Zustände in der ursprünglichen Markov-Kette folgt, werden wir nun zeigen.

Da jedoch das in Kapitel 4.2 behandelte Verfahren für die quantitative, zeitabhängige Erreichbarkeitsanalyse wegen der Bestimmbarkeit von minimalen und maximalen Schedulingen nur Zeitintervalle der Form $[0, t]$ zulässt, müssen wir uns bei dem folgenden Satz auf *zeitbegrenzte* 3-CSL beschränken.

Definition 32 (Zeitbegrenzte 3-CSL Formeln). Die Menge $\mathbb{F}_{\mathcal{M}}^{\leq t}$ aller zeitbegrenzten 3-CSL Formeln ist eine Teilmenge der Menge aller 3-CSL Formeln $\mathbb{F}_{\mathcal{M}}$ bezüglich einer abstrakten zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$. Die induktive Definition der zustandsbezogenen Formeln in beiden Fällen gleich ist, die Menge der pfadbezogenen *zeitbegrenzten* 3-CSL-Formeln setzt sich wie folgt induktiv zusammen:

- $(\mathcal{X}^{[0,t]}\psi) \in \mathbb{F}_{\mathcal{M}}$ falls $\psi \in \mathbb{S}_{\mathcal{M}}$
- $(\psi_1 \mathcal{U}^{[0,t]}\psi_2) \in \mathbb{F}_{\mathcal{M}}$ falls $\psi_1, \psi_2 \in \mathbb{S}_{\mathcal{M}}$

Satz 6 (Preservation von zeitbegrenzten 3-CSL-Formeln). Seien s und s' Zustände einer abstrakten zeitstetigen Markov-Kette $\mathcal{M} = (S, \mathbf{R}^I, \mathbf{P}^I, E^I, L)$ mit $s \preceq s'$, dann gilt für zeitbegrenzte 3-CSL-Formeln $\psi \in \mathbb{F}_{\mathcal{M}}^{\leq t}$:

$$\llbracket s', \psi \rrbracket \neq ? \Rightarrow \llbracket s, \psi \rrbracket = \llbracket s', \psi \rrbracket \quad (11)$$

Beweis. Wir zeigen per struktureller Induktion über den Formelaufbau, dass (11) aus $s \preceq s'$ folgt.

Für den Induktionsanfang betrachten wir zunächst die atomaren Formeln $\psi \in \{\text{true}\} \cup \{a \mid a \in AP\}$ in \mathbb{F}_{AP} .

$$\llbracket s', \text{true} \rrbracket = \top = \llbracket s, \text{true} \rrbracket \text{ für alle } s, s' \in S \quad \checkmark$$

Für $a \in AP$ und $\llbracket s', a \rrbracket \neq ?$ gilt nach Definition 24.1:

$$\llbracket s', a \rrbracket = L(s', a) = L(s, a) = \llbracket s, a \rrbracket \quad \checkmark$$

Induktionsannahme (IA) sei im Folgenden: Für alle direkten Teilformeln ψ von $\tilde{\psi}$ sei $\llbracket s', \psi \rrbracket = ?$ oder $\llbracket s', \psi \rrbracket = \llbracket s, \psi \rrbracket$. Dies impliziert:

$$\llbracket s', \psi \rrbracket = \alpha \Rightarrow \llbracket s, \psi \rrbracket = \alpha \text{ für } s \preceq s' \text{ und } \alpha \in \mathbb{B}$$

Fall 1: $\tilde{\psi} \equiv \neg\psi$

Ist $\llbracket s', \psi \rrbracket = ?$, so ist $\llbracket s', \neg\psi \rrbracket = \llbracket s', \psi \rrbracket^c = ?^c = ?$. Andernfalls gilt:

$$\llbracket s', \neg\psi \rrbracket = \llbracket s', \psi \rrbracket^c \stackrel{IA}{=} \llbracket s, \psi \rrbracket^c = \llbracket s, \neg\psi \rrbracket \quad \checkmark$$

Fall 2: $\tilde{\psi} \equiv \psi_1 \wedge \psi_2$

Genau dann wenn $\llbracket s', \psi_1 \rrbracket = ?$ und $\llbracket s', \psi_2 \rrbracket \neq \perp$, oder wenn $\llbracket s', \psi_1 \rrbracket \neq \perp$ und $\llbracket s', \psi_2 \rrbracket = ?$, dann gilt nach Tabelle 4, dass $\llbracket s', \psi_1 \wedge \psi_2 \rrbracket = ?$. Andernfalls gilt (ebenfalls nach Tabelle 4):

$$\llbracket s', \psi_1 \wedge \psi_2 \rrbracket = \llbracket s', \psi_1 \rrbracket \sqcup \llbracket s', \psi_2 \rrbracket = \llbracket s, \psi_1 \rrbracket \sqcup \llbracket s, \psi_2 \rrbracket \stackrel{IA}{=} \llbracket s, \psi_1 \wedge \psi_2 \rrbracket$$

Hierzu betrachten wir den Fall $\llbracket s', \psi_1 \rrbracket = \perp$ und $\llbracket s', \psi_2 \rrbracket = ?$ etwas näher. Lösen wir zunächst die Gleichung $\llbracket s', \psi_1 \rrbracket \sqcup \llbracket s', \psi_2 \rrbracket \stackrel{IA}{=} \llbracket s, \psi_1 \rrbracket \sqcup \llbracket s, \psi_2 \rrbracket$ etwas weiter auf:

$$\llbracket s', \psi_1 \rrbracket \sqcup \llbracket s', \psi_2 \rrbracket = \perp \sqcup ? = \perp \stackrel{IA}{=} \perp \sqcup \llbracket s, \psi_2 \rrbracket = \llbracket s, \psi_1 \rrbracket \sqcup \llbracket s, \psi_2 \rrbracket$$

Über $\llbracket s, \psi_2 \rrbracket$ können wir nichts weiter sagen, außer dass es ein Wert aus \mathbb{B} sein muss. Da jedoch $\perp \sqcup \perp = \perp = \perp \sqcup \top$, ist die Gleichung in jedem Fall gültig. Für $\llbracket s', \psi_1 \rrbracket = ?$ und $\llbracket s', \psi_2 \rrbracket = \perp$ gilt dieselbe Überlegung und die übrigen Fälle sind trivialerweise erfüllt. \checkmark

Fall 3: $\tilde{\psi} \equiv \mathcal{P}_{\leq p}(\mathcal{X}^I \psi)$

Falls s' absorbierend ist, gilt $E(s') = 0$. Nach Definition 24.2.2 folgt damit, dass auch $E(s) = 0$, also s ebenfalls absorbierend ist. Da für alle Pfade σ die in absorbierenden Zuständen beginnen $|\sigma| = 0$ gilt, erhält man für alle $\psi \in \mathbb{S}_{AP}$ bezüglich solcher Pfade $\llbracket \sigma, \mathcal{X}^I \psi \rrbracket = \perp$. Demnach folgt direkt aus Tabellen 3 und 5, dass $Pr(s, \mathcal{X}^I \psi, \top) = 0 = Pr(s', \mathcal{X}^I \psi, \top)$ und damit gilt $\llbracket s, \mathcal{P}_{\leq p}(\mathcal{X}^I \psi) \rrbracket = \llbracket s', \mathcal{P}_{\leq p}(\mathcal{X}^I \psi) \rrbracket$.

Betrachten wir den Fall, dass s' nicht absorbierend und $\llbracket s', \mathcal{P}_{\leq p}(\mathcal{X}^I \psi) \rrbracket = \top$ ist. Wie wir in Kapitel 4.1 festgestellt haben, ist dies gegeben, wenn mindestens mit Wahrscheinlichkeit $1 - p$ im nächsten Zustand ψ *nicht* erfüllt ist (siehe Tabelle 5 auf Seite 89). Aus Korollar 3 wissen wir, dass für alle Wahrscheinlichkeitsverteilungen $\bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))$ gilt:

$$Pr^l(s', \mathcal{X}^I \psi, \perp) \leq X(s', I) \cdot \sum_{v \in Fls(\psi)} \bar{\mu}'(v) \leq Pr^u(s', \mathcal{X}^I \psi, \perp)$$

Wählen wir nun einen Ratenvektor $\mu_{min} \in \text{rates}(\mathbf{R}(s, \cdot))$, für dessen zugehörige Verteilung $\bar{\mu}_{min} \in \text{distr}(\mathbf{P}(s, \cdot))$ gilt $X(s, I) \cdot \sum_{u \in Fls(\psi)} \bar{\mu}_{min}(u) = Pr^l(s, \mathcal{X}^I \psi, \top)$, dann existiert nach Definition 24.2 ein Ratenvektor $\mu' \in \text{rates}(\mathbf{R}(s', \cdot))$ mit der zugehörigen Wahrscheinlichkeitsverteilung $\bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))$ und eine Gewichtung Δ , so dass:

$$\begin{aligned}
1 - p &\triangleq Pr^l(s', \mathcal{X}^I \psi, \perp) \\
&\leq X(s', I) \cdot \sum_{v \in Fls(\psi)} \bar{\mu}'(v) \\
&= (e^{-E(s') \cdot t_l} - e^{-E(s') \cdot t_u}) \cdot \sum_{v \in Fls(\psi)} \bar{\mu}'(v) \quad (I = [t_l, t_u]) \\
&= (e^{-E(s) \cdot t_l} - e^{-E(s) \cdot t_u}) \cdot \sum_{v \in Fls(\psi)} \bar{\mu}'(v) \quad (\text{Def. 24.2.2}) \\
&= X(s, I) \cdot \sum_{v \in Fls(\psi)} \bar{\mu}'(v) \\
&= X(s, I) \cdot \sum_{v \in Fls(\psi), u \in S} \Delta(u, v) \quad (\text{Def. 24.2.1.3}) \\
&= X(s, I) \cdot \sum_{v \in Fls(\psi), u: u \preceq v} \Delta(u, v) \quad (\text{Def. 24.2.1.1}) \\
&= X(s, I) \cdot \sum_{u, v \in Fls(\psi), u: u \preceq v} \Delta(u, v) \quad (\text{IA}) \\
&\leq X(s, I) \cdot \sum_{u \in Fls(\psi), v \in S} \Delta(u, v) \\
&= X(s, I) \cdot \sum_{u \in Fls(\psi)} \bar{\mu}_{min}(u) \\
&= Pr^l(s, \mathcal{X}^I \psi, \perp)
\end{aligned}$$

Das bedeutet also, dass die Wahrscheinlichkeit für Pfade mit Eigenschaft $\mathcal{X}^I \psi$ in Zustand s' höchstens so groß ist wie in Zustand s .

Der Fall mit s' ist nicht absorbierend und $\llbracket s', \mathcal{P}_{\preceq p}(\mathcal{X}^I \psi) \rrbracket = \perp$ verläuft analog zu obigem Fall. Man muss lediglich \top und \perp miteinander vertauschen und $Fls(\psi)$ durch $Sat(\psi)$ sowie $1 - p \triangleq$ durch $p \triangleleft$ (im Sinne von Tabelle 5) ersetzen.

✓

Fall 4: $\tilde{\psi} \equiv \mathcal{P}_{\preceq p}(\psi_1 \mathcal{U}^I \psi_2)$

Wir zeigen im Folgenden, dass $\llbracket s', \mathcal{P}_{\preceq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top \Rightarrow \llbracket s, \mathcal{P}_{\preceq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$. Wegen der Induktionsvoraussetzung gilt $\llbracket s', \psi_i \rrbracket = \alpha \Rightarrow \llbracket s, \psi_i \rrbracket = \alpha$ für $i \in \{1, 2\}$ und $\alpha \in \mathbb{B}$. Da die Gültigkeit von ψ_2 in s' auch die von $\tilde{\psi}$ in s' impliziert und selbiges auch für s gilt, ist im Fall von $\llbracket s', \psi_2 \rrbracket = \top$ die Behauptung offensichtlich korrekt. Ein weiterer trivialer Fall ist der, dass $\llbracket s', \psi_1 \rrbracket = \llbracket s', \psi_2 \rrbracket = \perp$. In einem solchen Zustand s' ist natürlich auch die Formel $\tilde{\psi}$ unerfüllbar. Da für alle $s \preceq s'$ gilt, dass $\llbracket s, \psi_1 \rrbracket = \llbracket s, \psi_2 \rrbracket = \perp$, gilt folglich auch $\llbracket s, \tilde{\psi} \rrbracket = \perp$.

Da es für $\llbracket s, \tilde{\psi} \rrbracket = ?$ nichts zu zeigen gibt, bleibt als nicht-trivialer Fall lediglich noch der mit $\llbracket s, \psi_1 \rrbracket = \top$ und $\llbracket s, \psi_2 \rrbracket \neq \top$ übrig³³. Aus Kapitel 4.2 wissen wir, dass die minimale Wahrscheinlichkeit für das Erreichbarkeitsproblem für eine Zielmenge B und Zeitschranke t in einer uniformen zeitstetigen Markov-Kette mit Exitrate E_{unif} über die folgende Summe berechnet werden kann:

$$\begin{aligned}
q_0 &= \varphi(E_{unif} \cdot t, 0) \cdot i_B + q_1 \\
&= \underbrace{\varphi(E_{unif} \cdot t, 0) \cdot i_B}_{\pi_{E_{unif} \cdot t}^{(0)}} + \sum_{n=1}^{\infty} \underbrace{(\varphi(E_{unif} \cdot t, n) \cdot \mathbf{P}_1 \cdot \dots \cdot \mathbf{P}_n \cdot i_B)}_{\pi_{E_{unif} \cdot t}^{(n)}}
\end{aligned}$$

Mit den einzelnen Summanden $\pi_{E_{unif} \cdot t}^{(i)}$ bezeichnen wir im Folgenden die Wahrscheinlichkeiten des Erreichbarkeitsproblems für Pfade der Länge $i \in \mathbb{N}_{\geq 0}$. Mit B'

³³Die nicht-trivialen Fälle entsprechen den Zellen der W -Mengen-Tabelle auf Seite 100, in denen weder \top^+ noch \top^- eingetragen ist.

bezeichnen wir die Menge der Makro-Zustände für die gilt, dass jeweils alle repräsentierten Zustände der ursprünglichen Markov-Kette in der Zielmenge B enthalten waren.

Wir werden durch Induktion über die Pfadlänge i zeigen, dass $\pi'_\lambda^{(i)}(s') \leq \pi_\lambda^{(i)}(s)$ für alle $s, s' \in S$ mit $s \preceq s'$ und alle $i \in \mathbb{N}_{\geq 0}$ gilt. Diese Erkenntnis werden wir anschließend verwenden um zu zeigen, dass die minimalen Wahrscheinlichkeit für die gesuchten Pfade beliebiger Länge bezüglich der Abstraktion nur maximal so groß sein können, wie Pfade bezüglich der ursprünglichen abstrakten Markov-Kette.

Induktionsanfang: ($i = 0$)

$$\begin{aligned} \pi'_\lambda^{(0)}(s') &= (\varphi(\lambda, 0) \cdot i_{B'}) (s') \\ &= (\varphi(\lambda, 0) \cdot i_B) (s) \quad (\text{wg. } s \preceq s' \Rightarrow (s \in B \text{ gdw. } s' \in B')) \\ &= \pi_\lambda^{(0)}(s) \end{aligned}$$

Induktionsschluß: ($i \rightarrow i + 1$)

Die Menge aller Zustände, in denen ψ_1 oder ψ_2 erfüllt ist, also solche, in denen die Gültigkeit von $\psi_1 \mathcal{U}^I \psi_2$ noch möglich ist³⁴, notieren wir mit \bar{S} . Für Zustände aus \bar{S} gilt insbesondere aufgrund der äußeren Induktionsvoraussetzung:

$$(u \preceq v \wedge v \in \bar{S}) \Rightarrow u \in \bar{S}$$

Wir zeigen nun, dass die Ungleichung aus der Behauptung für $i + 1$ gilt, falls sie bereits für i gilt:

$$\begin{aligned} \pi'_\lambda^{(i+1)}(s') &= \varphi(\lambda, i + 1) \cdot (\sum_{v \in S} (\mathbf{P}'_1 \cdot \mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{n+1} \cdot i_{B'})(v)) (s') \\ &\leq \varphi(\lambda, i + 1) \cdot \sum_{u \in \bar{S}} ((\mathbf{P}_1 \cdot \mathbf{P}_2 \cdot \dots \cdot \mathbf{P}_{n+1} \cdot i_B)(u)) (s) \\ &= \pi_\lambda^{(i+1)}(s) \end{aligned}$$

Wie im *Next*-Fall betrachten wir wieder einen Ratenvektor $\mu_{min} \in \text{rates}(\mathbf{R}(s, \cdot))$, der zu einer minimalen Wahrscheinlichkeit $Pr^l(s, \psi_1 \mathcal{U}^I \psi_2, \top)$ gehört. Eine solcher Ratenvektor impliziert $\bar{\mu}_{min} \in \text{distr}(\mathbf{P}(s, \cdot))$. Nach Definition 24.2 existiert ein entsprechender Ratenvektor $\mu' \in \text{rates}(\mathbf{R}(s', \cdot))$ mit der zugehörigen Wahrscheinlichkeitsverteilung $\bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))$ und eine Gewichtung Δ , mit denen wir zeigen können, dass die folgende Ungleichung gilt:

$$\begin{aligned} &(\sum_{v \in \bar{S}} (\mathbf{P}'_1 \cdot \mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{n+1} \cdot i_{B'})(v)) (s') \\ &= \sum_{v \in \bar{S}} \bar{\mu}'(v) \cdot (\mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{n+1} \cdot i_{B'})(v) \quad (\text{wg. } \bar{\mu}' \in \text{distr}(\mathbf{P}(s', \cdot))) \\ &= \sum_{v \in \bar{S}} \Delta(S, v) \cdot (\mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{n+1} \cdot i_{B'})(v) \quad (\text{Def. 24.2.1.3}) \\ &= \sum_{u, v \in \bar{S}, u:u \preceq v} \Delta(u, v) \cdot (\mathbf{P}'_2 \cdot \dots \cdot \mathbf{P}'_{n+1} \cdot i_{B'})(v) \quad (\text{Def. 24.2.1.1}) \\ &\leq \sum_{u, v \in \bar{S}, u:u \preceq v} \Delta(u, v) \cdot (\mathbf{P}''_1 \cdot \dots \cdot \mathbf{P}''_n \cdot i_B)(u) \quad (\text{IA, Lemma 9}) \\ &\leq \sum_{u \in \bar{S}} \Delta(u, S) \cdot (\mathbf{P}''_1 \cdot \dots \cdot \mathbf{P}''_n \cdot i_B)(u) \\ &= \sum_{u \in \bar{S}} \bar{\mu}_{min}(u) \cdot (\mathbf{P}''_1 \cdot \dots \cdot \mathbf{P}''_n \cdot i_B)(u) \\ &= (\sum_{u \in \bar{S}} (\mathbf{P}_1 \cdot \mathbf{P}_2 \cdot \dots \cdot \mathbf{P}_{n+1} \cdot i_B)(u)) (s) \quad (\text{Lemma 9}) \end{aligned}$$

³⁴Diese Zustände entsprechen genau denen, die nicht in der Menge $W^{\top-}$ enthalten sind.

Aus der inneren Induktion folgt also:

$$\begin{aligned} \pi'_\lambda{}^{(i)}(s') &\leq \pi_\lambda{}^{(i)}(s) && \text{für } s \preceq s' \text{ und alle } i \in \mathbb{N}_{\geq 0} \\ \Rightarrow \sum_{i=0}^{\infty} \pi'_\lambda{}^{(i)}(s') &\leq \sum_{i=0}^{\infty} \pi_\lambda{}^{(i)}(s) && \text{für } s \preceq s' \\ \Rightarrow \pi'_\lambda(s') &\leq \pi_\lambda(s) && \text{für } s \preceq s' \end{aligned}$$

Das bedeutet, dass die Wahrscheinlichkeit für Pfade mit der Eigenschaft $\psi_1 \mathcal{U}^I \psi_2$ in Zustand s' höchstens so groß ist wie in Zustand s . Falls also $\llbracket s', \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$ gilt, so gilt auch $\llbracket s, \mathcal{P}_{\leq p'}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$ für alle $p' \in [p, 1]$ und daher auch $\llbracket s, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$.

Für $\llbracket s', \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \perp \Rightarrow \llbracket s, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \perp$ muss die Beweisführung nur leicht modifiziert werden. Auf die Ausführung dieses zweiten Teils werden wir hier jedoch verzichten.

✓

Der letzte Fall mit $\tilde{\psi} \equiv \mathcal{P}_{\geq p}(\psi_1 \mathcal{U}^I \psi_2)$ verläuft analog zu Fall 4. Anhand der Semantiktabelle 5 sieht man, dass die Bedingungen für $\llbracket s, \mathcal{P}_{\geq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$ fast genau denen für $\llbracket s, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \perp$ entsprechen und die für $\llbracket s, \mathcal{P}_{\geq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \perp$ fast genau denen für $\llbracket s, \mathcal{P}_{\leq p}(\psi_1 \mathcal{U}^I \psi_2) \rrbracket = \top$.

Damit ist die Preservation von CSL Formeln des gegebenen Fragments bewiesen.

□

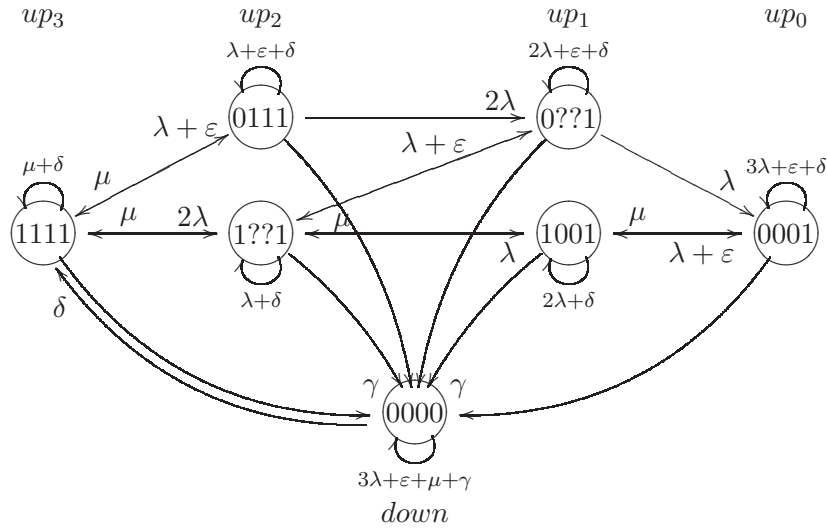
Beispiel 29 (Model Checking in Abstraktionen zeitstetiger Markov-Ketten).

Betrachten wir nun noch einmal abschließend das *Triple Redundant Modular System* in der abgewandelten Version aus Beispiel 7. Wir interessieren uns im Folgenden ob die Wahrscheinlichkeit einen Wert p überschreitet, dass innerhalb von t Zeiteinheiten eine Erneuerung des gesamten Systems notwendig wird. Dies lässt sich mathematisch beschreiben als:

$$\mathcal{P}_{>p}(\neg \text{down } \mathcal{U}^{[0,t]} \text{down})$$

Statt bezüglich der zeitdiskreten Markov-Kette soll nun diese Eigenschaft auf der Basis einer Abstraktion des Modells nachgewiesen werden. Um in abstrakten Markov-Ketten probabilistische Aussagen über Pfadeigenschaften machen zu können, müssen wir zunächst eine Uniformisierung durchführen. Dazu benötigen wir zunächst eine uniforme Exitrate E_{unif} bestimmen, für die die folgenden Ungleichungen erfüllt sind:

$$\begin{aligned} E(1111) &= 3\lambda + \varepsilon + \gamma \leq E_{unif} \\ E(0111) &= 2\lambda + \gamma + \mu \leq E_{unif} \\ E(1??1) &= 2\lambda + \varepsilon + \gamma + \mu \leq E_{unif} \\ E(1001) &= \lambda + \varepsilon + \gamma + \mu \leq E_{unif} \\ E(0??1) &= \lambda + \gamma + \mu \leq E_{unif} \\ E(0001) &= \gamma + \mu \leq E_{unif} \\ E(0000) &= \delta \leq E_{unif} \end{aligned}$$

Abbildung 39: Uniformisiert mit $E_{unif} = 3\lambda + \epsilon + \gamma + \mu + \delta$.

$E_{unif} = 3\lambda + \epsilon + \gamma + \mu + \delta$ erfüllt diese Ungleichungen für alle $\lambda, \epsilon, \gamma, \mu, \delta \in \mathbb{R}_{>0}$. Die Uniformisierung mit E_{unif} führt zur zeitstetigen Markov-Kette in Abbildung 39.

Abstrahieren wir die so erhaltene uniforme zeitstetige Markov-Kette bezüglich der Partitionierung $\mathcal{A} = \{\{0000\}, \{0111, 1??1, 0??1, 1001, 0001, 1111\}\}$, wobei wir den Zustand $\{0000\}$ umbenennen zu s_0 und den anderen zu s_1 , so erhalten wir die abstrakte zeitstetige Markov-Kette³⁵ \mathcal{M} in Abbildung 40 links. Wir wählen hier eine noch stärkere Abstraktion als in Beispiel 7 angedeutet wurde, um die Verifikation der Spezifikation so einfach wie möglich zu gestalten.

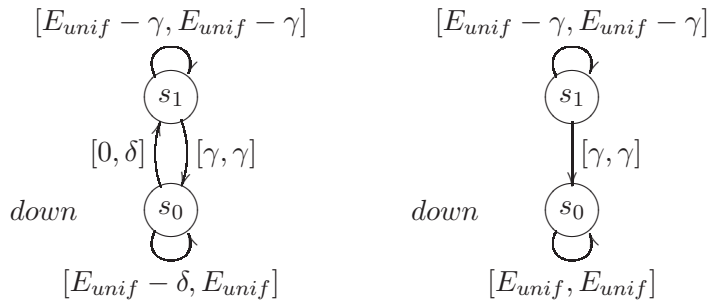


Abbildung 40: Abstraktion (links) und transformierte ACTMC (rechts)

Um nun die quantitative Erreichbarkeitsanalyse zur Bestimmung der unteren Wahrscheinlichkeitsschranke $Pr^l\{\sigma \in Pfad_s \mid \llbracket \sigma, \neg down \mathcal{U}^{[0,t]} down \rrbracket = \top\}$ für Zustände s anwenden zu können, müssen wir die Mengen $W^{\top+}$ und $W^{\top-}$ bestimmen. Natürlich ist $W^{\top+} = \{s_0\}$ die Menge der *down*-Zustände. In $W^{\top-}$ sind nach Tabelle 6 genau die Zustände s enthalten, für die gilt $\llbracket s, down \rrbracket \neq \top$ und $\llbracket s, down \rrbracket \neq \perp$. Da in der Abstraktion keine Zustände enthalten sind, für die $\llbracket s, down \rrbracket = ?$ gilt, sind also

³⁵Der Übersichtlichkeit halber wurden die Wahrscheinlichkeitsintervalle nicht notiert.

in $W^{\top-}$ keine Zustände enthalten. Die Transformation bezüglich der W -Mengen besteht also darin, die ausgehende Kante von Zustand s_0 zu entfernen und die Rate des *self-loop* auf den Wert E_{unif} festzulegen. Die Transformation ergibt demnach die ACTMC \mathcal{M}' in Abbildung 40 rechts, für die gilt:

$$\begin{aligned} & Pr^l(\sigma \in Pfad_s^{\mathcal{M}'} \mid \llbracket \sigma, true \mathcal{U}^{[0,t]} down \rrbracket) \\ &= Pr^l(\sigma \in Pfad_s^{\mathcal{M}} \mid \llbracket \sigma, -down \mathcal{U}^{[0,t]} down \rrbracket). \end{aligned}$$

Die erhaltene abstrakte zeitstetige Markov-Kette \mathcal{M}' lässt keinen Spielraum für den Scheduler, sodass die \mathbf{P}_i für alle $i \in \mathbb{N}_{>0}$ gegeben sind mit:

$$\mathbf{P}_i = \begin{pmatrix} 1 - \mu/E_{unif} & \mu/E_{unif} \\ 0 & E_{unif} \end{pmatrix}$$

Wie wir aus Kapitel 4.2 wissen, erhält man eine Approximation des Ergebnisses durch Lösen des folgenden Gleichungssystems nach q_0 :

$$\begin{aligned} q_0 &= \varphi(E_{unif} \cdot t, n) \cdot i_{\{s_1\}} + q_1 \\ q_i &= \varphi(E_{unif} \cdot t, i) \cdot \mathbf{P}_i \cdot i_{\{s_1\}} + \mathbf{P}_i \cdot q_{i+1} \\ &\text{für } i \in \{1, \dots, k\} \text{ und } q_{k+1}(s) = 0 \text{ für alle } s \in S. \end{aligned}$$

Eine geeignete maximale Pfadlänge k lässt sich wie in Kapitel 4.2 beschrieben durch Vergleichen der kumulierten Poisson-Wahrscheinlichkeiten mit der gewünschten Genauigkeit bestimmen.

Sofern das Ergebnis der Berechnung für s_1 größer ist als p ist die Spezifikation bezüglich s_1 also gültig. In diesem Beispiel eine eher negativ zu bewertende Eigenschaft, da in der Regel eine häufige Erneuerung des Systems nicht gewünscht ist. Wegen Satz 6 überträgt sich diese Eigenschaft von s_1 auf alle durch s_1 simulierten Zustände der ursprünglichen Markov-Kette, hier also auf alle Zustände, außer 0000.

—

5 Zusammenfassung und Ausblick

In dieser Arbeit wurde eine neue Technik zur Abstraktion für zeitstetige Markov-Ketten eingeführt. Statt über eine Bisimulationsrelation Zustände zu finden die sich gleich verhalten, wollten wir eine Möglichkeit haben, auch Zustände mit ähnlichem Verhalten zusammenfassen zu können. Was dabei als *ähnliches Verhalten* aufzufassen ist, hängt von der jeweiligen Anwendung ab und wurde hier nicht weiter untersucht. Da zu starke Abstraktion dazu führen kann, dass die gegebenen Spezifikationen weder nachgewiesen noch widerlegt werden können, haben wir eine dreiwertige Logik eingeführt, mit der die *Unbestimmtheit* einer Aussage ausgedrückt werden kann.

Beim Modellierungsformalismus haben wir uns gegen Markov-Entscheidungsprozesse und für die abstrakten zeitstetigen Markov-Ketten entschieden, da bei den Entscheidungsprozessen eine Abstraktion mehrerer Kanten von einem Zustand zu einem anderen nicht möglich ist. Die Scheduler als Hilfsmittel zur formalen Behandlung des Nicht-Determinismus in abstrakten zeitstetigen Markov-Ketten wurden dagegen praktisch von den Markov-Entscheidungsprozessen übernommen. Wir mussten feststellen, dass durch die Zusammenhänge, die zwischen Raten, Wahrscheinlichkeiten und Exitraten bestehen, nicht beliebige Werte aus den Intervallen der abstrakten Markov-Kette von einem Scheduler gewählt werden können. Um Werte zu entfernen, die von einem Scheduler nicht zu einer gültigen zeitstetigen Markov-Kette vervollständigt werden können, haben wir daher den *cut* eingeführt. Dieser ist deutlich aufwändiger als der für zeitdiskrete Markov-Ketten, weswegen wir lediglich die Konvergenz nachgewiesen haben. Die Frage, ob die Berechenbarkeit nachgewiesen werden kann oder wie ein Verfahren zur Approximation bezüglich einer gegebenen Genauigkeit aussehen könnte, wurde offen gelassen.

Die probabilistische Simulation, die wir später in erster Linie dazu verwendet haben, die Präservation von Spezifikationen bezüglich der Abstraktion nachzuweisen, stellte sich im Vergleich zur Bisimulation als recht kompliziert heraus. Da die nachfolgenden Abstraktionsmethoden jedoch die probabilistische Simulation der einzelnen Zustände durch ihre Makro-Zustände implizierten, war im weiteren Verlauf keinerlei Bestimmung oder Überprüfung einer solchen probabilistischen Simulationsrelation mehr nötig. Bei den Abstraktionsmethoden A und B, wo versucht wurde mit möglichst wenig Aufwand eine Abstraktion zu bestimmen, mussten wir feststellen, dass diese nicht die erhoffte Genauigkeit boten. Mit der Abstraktion C gelang es dann den Schnitt aus den ersten beiden Abstraktionen zu bilden, wodurch deren Schwächen jeweils eliminiert werden konnten.

Da die uniformen Markov-Ketten bei der Bestimmung der *Transient* Eigenschaften von Bedeutung sind, haben wir diese etwas genauer untersucht. Ein auf den ersten Blick erstaunliches Ergebnis war, dass für uniforme Markov-Ketten die drei Abstraktionen dieselben Ergebnisse liefern³⁶. Leider war nicht klar, wie nicht-uniforme abstrakte Markov-Ketten uniformisiert werden können. Für nicht-abstrakte zeitstetige

³⁶Wegen des *Next*-Operators wird die Abstraktion C dennoch nicht obsolet.

Markov-Ketten dagegen ist die Uniformisierung ein einfaches und bekanntes Verfahren. Wir konnten nachweisen, dass für nicht-abstrakte zeitstetige Markov-Ketten die Uniformisierung und anschließende Abstraktion zu einer uniformen abstrakten Markov-Kette führt.

Ein Vergleich der vorgestellten Abstraktion mit der bezüglich zeitstetiger Markov-Ketten zeigte, dass ausgehend von einer abstrakten zeitstetigen Markov-Kette die Abstraktion auf der zeitstetigen Ebene mit anschließender Bereinigung mindestens so genau ist wie die Abstraktion auf zeitdiskreter Ebene bezüglich der eingebetteten Markov-Kette. Ob auch die umgekehrte Inklusion gilt, also ob die Abstraktion auf zeitdiskreter Ebene mindestens so genau ist wie die auf zeitstetiger Ebene, hatten wir dabei nicht betrachtet. Ein weiteres Ergebnis bezüglich des Vergleichs der beiden Arten von Markov-Ketten war, dass Zustände, die in probabilistischer Simulationsrelation stehen, auch in der eingebetteten Markov-Kette in Simulationsrelation bezüglich zeitdiskreter Markov-Ketten stehen.

Vorbereitend für das Kapitel über Model Checking von abstrakten zeitstetigen Markov-Ketten haben wir dann ein Wahrscheinlichkeitsmaß definiert, das uns erlaubt über Wahrscheinlichkeiten von Pfadmengen zu sprechen. Ein für die Erreichbarkeitsanalyse in zeitstetigen Markov-Ketten wichtiges Ergebnis ist, dass die Betrachtung der extremen Scheduler bei der Betrachtung von Infimum und Supremum von Pfadwahrscheinlichkeiten ausreichend ist. Dadurch konnte später die Wahl des Schedulers innerhalb eines Zustandes auf eine endliche Zahl von Alternativen begrenzt werden.

Bevor Model Checking für abstrakte zeitstetige Markov-Ketten behandelt werden konnte, musste noch die Spezifikationslogik CSL auf die dreiwertige Logik angehoben werden. Anschließend betrachteten wir die quantitative Erreichbarkeitsanalyse für uniforme abstrakte zeitstetige Markov-Ketten, die den Kern des Model Checking für *Until*-Formeln bildet. Über eine Transformation des Modells bezüglich einer gegebenen *Until*-Formel und der zugehörigen W -Mengen konnte dann das Model Checking für *Until* auf die Erreichbarkeitsanalyse zurückgeführt werden. Das letzte Ergebnis dieser Arbeit handelte von der Präservation des untersuchten 3-CSL Fragments. Falls das Verfahren für Model Checking von abstrakten zeitstetigen Markov-Ketten nicht ein *unbestimmtes* Ergebnis liefert, so lässt sich die Gültigkeit bzw. Ungültigkeit der Spezifikation auf die ursprüngliche Markov-Kette übertragen.

Das Thema dieser Diplomarbeit lässt noch weiterführende Forschung in verschiedene Richtungen zu. So wäre es noch interessant zu sehen, ob und gegebenenfalls wie Model Checking für weitere Teile von 3-CSL durchgeführt werden könnte. Die Hinzunahme des *Steady-State* Operators könnte noch untersucht werden und auch eine Vergrößerung der Menge von zulässigen Zeitintervallen für *Until* wäre wünschenswert.

Ein Punkt, der insbesondere auch für den praktischen Einsatz von Interesse sein dürfte, ist die Verfeinerung von Abstraktionen. Erhält man durch das Model Checking Verfahren ein unbestimmtes Ergebnis, so wäre eine automatische oder halb-automatische Verfeinerung der Abstraktion von großer Hilfe. Dazu müssten erstens Makro-Zustände identifiziert werden, die für die Unbestimmtheit des Ergeb-

nisses verantwortlich sind, und zweitens eine ausreichend feine Abstraktion für die Zustände gefunden werden, die durch den Makro-Zustand repräsentiert werden.

Weiterhin wäre es interessant zu sehen, von welchem Nutzen die vorgestellte Abstraktionstechnik in der Praxis ist. Dazu müsste zunächst ein Model Checker implementiert werden. Die Untersuchung einiger Fallstudien sollte dann über die Praxistauglichkeit Aufschluss geben. Im Rahmen der Fallstudien wären auch Möglichkeiten für die automatische oder halbautomatische Generierung von sinnvollen Abstraktionen bzw. Partitionierungen des Zustandsraums zu erforschen.

Literatur

- [Bai96] Christel Baier. Polynomial Time Algorithms for Testing Probabilistic Bisimulation and Simulation. In Rajeev Alur and Thomas A. Henzinger, editors, *CAV*, volume 1102 of *Lecture Notes in Computer Science*, pages 50–61. Springer, 1996.
- [BDL04] G. Behrmann, A. David, and K.G. Larsen. A Tutorial on Uppaal. In *Formal Methods for the Design of Real-Time Systems: 4th International School on Formal Methods for the Design of Computer, Communication, and Software Systems*, number 3185 in LNCS, pages 200–236. Springer-Verlag, 2004.
- [BGdMT98] G. Bolch, S. Greiner, H. de Meer, and K.S. Trivedi. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. Wiley-Interscience, 1998.
- [BHHK03] C. Baier, B. R. Haverkort, H. Hermanns, and J.P. Katoen. Model-Checking Algorithms for Continuous-Time Markov chains. *IEEE Transactions on Software Engineering*, 29(6):524–541, 2003.
- [BHKH04] C. Baier, H. Hermanns, J.P. Katoen, and B.R. Haverkort. Efficient Computation of Time-Bounded Reachability Probabilities in Uniform Continuous-Time Markov Decision Processes. In Kurt Jensen and Andreas Podelski, editors, *TAPAS*, volume 2988 of *Lecture Notes in Computer Science*, pages 61–76. Springer, 2004.
- [BHKW05] C. Baier, H. Hermanns, J.P. Katoen, and V. Wolf. Comparative Branching-Time Semantics for Markov Chains. *Information and Computation*, 200 (2):149–214, 2005.
- [BL04] B. Bollig and M. Leucker. Verifying Qualitative Properties of Probabilistic Programs. *Lecture Notes in Computer Science*, 2925:124–146, 2004.
- [BS81] I. N. Bronstein and K. A. Semendjajew. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 1981. 20. Auflage. Herausgegeben von G. Grosche und V. Ziegler.
- [CDKN98] B. Changuion, I. Davies, P.S. Kritzinger, and M.A. Nelte. DaNaMiCS - Modelling Concurrent Systems. In *First Annual South African Telecommunications, Networks and Applications Conference*, pages 606–610, 1998.
- [CGP99] E. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT Press, 1999.

- [EL87] E.A. Emerson and C.L. Lei. Modalities for Model Checking: Branching Time Logic Strikes Back. *Science of Computer Programming*, 8 (3):275–306, 1987.
- [FG88] B.L. Fox and P.W. Glynn. Computing Poisson Probabilities. *Communications of the ACM*, 31:440–445, 1988.
- [FLW06] H. Fecher, M. Leucker, and V. Wolf. Don't Know in Probabilistic Systems. In *SPIN*, Lecture Notes in Computer Science. Springer, 2006.
- [Hav02] B.R. Haverkort. Markovian Models for Performance and Dependability Evaluation. *Lectures on Formal Methods and Performance Analysis: First EEF/Euro Summer School on Trends in Computer Science*, 1:38–83, 2002.
- [HKMKS] H. Hermanns, J.P. Katoen, J. Meyer-Kayser, and M. Siegle. A Tool for Model-Checking Markov Chains. *Software Tools for Technology Transfer*, 4 (2):153–172.
- [HMNP06] A. Hinton, M.K., G. Norman, and D. Parker. PRISM: A Tool for Automatic Verification of Probabilistic Systems. In *Proceedings of 12th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, 2006.
- [Hol04] G.J. Holzmann. *The SPIN Model Checker: Primer and Reference Manual*. Addison Wesley, 2004.
- [HR00] M.R.A. Huth and M.D. Ryan. *Logic in Computer Science: Modelling and Reasoning about Systems*. Cambridge University Press, Cambridge, England, 2000.
- [Mol81] M.K. Molloy. *On the Integration of Delay and Throughput Measures in Distributed Processing Models*. PhD Thesis (University of California at Los Angeles), 1981.
- [Put94] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.
- [RN95] S. Russel and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 1995.
- [Sch03] K. Schneider. *Verification of Reactive Systems: Formal Methods and Algorithms*. Springer-Verlag, 2003.
- [Tij03] H.C. Tijms. *A First Course in Stochastic Models*. John Wiley and Sons, 2003.