# Probabilistic Automata: Semantics, Equivalence & Minimization

## Master Thesis

by

**Fabian Klopfer**

Faculty of Sciences
Department of Computer and Information Science

**1st Reviewer:** Prof. Dr. Tatjana Petrov
**2nd Reviewer:** ToDo

Konstanz, 2020

**Abstract:**

Some New Abstract Text

# Contents

# 1 Introduction

# 2 Background

## 2.1 Category Theory

**Definition 2.1** (Category, Semicategory, final object, graph homomorphism, functor)**.** *"set of objects and arrows/morphisms", "arrows"/morphism are compositional.*
*every object needs to have an identity morphism, i.e. neutral elem wrt. compostion. morphism composition is associative*

*A Semicategory is a category that does not require an identity morphism for every object*

*A final/terminal object F is such that for any other object A there is exactly one morphism to F.*

*An initial object I is such that for each other object there is exactly one morphism from I.*

*A graph homomorphism f is a function from one graph to another s.t. if vertex u is linked to v then f(u) is linked to f(v)*

*An isomorphism f is a morphism from object A to object B, if there extists a morphism g that maps from B to A such that when applying f and then g, it is the identity morphism*

*A covariant functor F from category C to category D is such that fir each object A of C we have F(A) is in D. For each morphism f:A to B in C we have F(f):F(A) to F(B) in D such that F(g)řF(f) = F ( g ř f) and* $F(id_A) = id_F(A).AFunctorpreservesthestructureofacategorymappingitintoanothercategory$

*A contravariant functor G from category C to category D is such that fir each object A of C we have F(A) is in D. For each morphism f:A to B in C we have F(f):F(B) to F(A) in D such that F(g)řF(g) = F ( g ř f) and* $F(id_A) = id_F(A).AFunctorreversesthemorphismsofacategorywhenmappingitinto$

*The categorical product is a candidate AxB,* $pi_1 : AxB-> A, pi_2 : AxB-> Bs.t.foranyothercandidateX, f : X-> A, g : X-> Bthereeiszsauniqueh : X-> AxBs.t.pi_1h = fandpi_2h = g$

**Remark.** *All terminal objects in a category are isomorphic*

[Empty, Set, Ord, Product, Monoid, Bimonoidal/Semiring, cat & CAT] The empty category contains no objects and no morphisms.

The Set Category contains objects that are sets and morphisms that are total functions between sets.

The Ord Category contains objects that are ordered sets and the morphisms are functors between objects

The Product Category for two categories C, D is the category whose objects are ordered pairs of objects from C and D, morphisms are ordered pairs, compostition is defined component wise.

The Mon category is a category C equipped with a functor out of the product category of C with itself called tensor product, a unit object, a left unitor, a right unitor and and associator.

A Bimon or Semiring category C is a category with to symmetric monoidal structure for addition and a monoidal structure for multiplication, together with distributivity natural and absorption isomorphisms

The cat Category is the category of all small categories having functors as morphisms. The CAT Category is the category of all large objects (such as Set or Mon) having functors as morphisms

## 2.2 Algebraic Preliminaries

**Definition 2.2** (Equivalence Relation)**.**

**Definition 2.3** (Monoid).

**Definition 2.4** (Ring, Semiring).

**Theorem 2.4.1** (The Ring of Formal Power Series).

**Definition 2.5** (Coalgebra, F-Coalgebra).

## 2.3 Probability Theory

**Definition 2.6** (Sample space, events, $\sigma$-algebra, measurable space). *Let $\Omega$ be a set, namely the set of possible outcomes of a chance experiment, called* **sample space***.*

*Let $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ with $\mathcal{P}$ the powerset, a set of subsets of the sample space, whose elements are called* **events***.*

*Let $\Omega$ be a sample space, $\mathcal{F}$ a set of events. $\mathcal{F}$ is called a* **$\sigma$-algebra** *over $\Omega$, if and only if*

- *the sample space is contained in the set of events,*

$$\Omega \in \mathcal{F}$$

- *the set of events is closed under complementation,*

$$A \in \mathcal{F} \Rightarrow \Omega \setminus A \in \mathcal{F}$$

- *and the set of events is closed under countable union:*

$$\forall i \geq 0 : A_i \in F \Rightarrow ( \bigcup_{n \in \mathbb{N} A_n) \in \mathcal{F}}$$

*The pair $(\Omega, \mathcal{F})$ is called a* **measurable space***.*

*Let $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ measurable spaces. A funtion $f : \Omega_1 \to \Omega_2$ is called a* **measurable function** *if and only if for every $A \in \mathcal{F}_2$ the preimage of $A$ under $f$ is in $\mathcal{F}_1$.*

$$\forall A \in \mathcal{F}_2 : f^{-1}(A) \in \mathcal{F}_1$$

**Definition 2.7** (Probability Space, Probability Measure, Discrete and Continuous Probability Space and Measure).

**Definition 2.8** (Random Variable, Probability Distribution, Distribution Funtion and Cumulative Density Function).

**Definition 2.9** (Mean, Median, Mode, Expectation, Variance).

**Definition 2.10** (Joint Probability Distribution, Conditional Probability, Bayes Rule, Independence, Conditional Independence).

**Definition 2.11** (Stochastic Process, Bernouilli & Binomial Process).

**Definition 2.12** (Geometric Distribution).

**Definition 2.13** ((Negative) Exponential Distr.).

**Theorem 2.13.1** (Memoryless property).

**Definition 2.14** (Markov Property, Markov Process, Time Homogeneity).

## 2.4 Automata Theory

**Definition 2.15** (Transition System)**.**

**Definition 2.16** (Labelled Transition System)**.**

**Definition 2.17** (Path, Trace, Cylinder Sets, Prefix, Postfix)**.**

**Definition 2.18** (Determinism and Non-Determinism, Internal vs. External, Adversaries/Policies)**.**

**Definition 2.19** (Weighted Automata)**.**

**Definition 2.20** (Probabilistic Automata, Initial Distribution, transition probability function, stochastic matrix, transition probability matrix)**.**

**Remark** (Probabilistic vs. Non-Deterministic Choice)**.**

**Remark** (Sokolova's System Types Paper and focus on MC-like models)**.**

**Definition 2.21** (Markov Chains, Discrete-Time, Continuous-Time, Labelled, (MDP?, Generalized Stochastic Petri Nets?))**.**

**Definition 2.22** (Bisimulation Relations, Strong, weak, exact, ordinary, Prob., Buchholz)**.**

# 3 Semantics, Equivalence & Minimization

The general case, may help with some things, e.g. if proven independent of the semiring used. So for each subsection here starting with WAs independent of the Semiring that is used, continue with PA results that are as independent as possible from the concrete transition structure. Finally apply the aforementioned results to LMCs/MCs. Also discriminate between DTMC, MDP, PA model(s) I.e. per subsection apply the folling structure

**WA - General Case for arbitrary Semirings**

**PA - Results for other transition structures**

**MC-like models - Application of above and other literature on specific example of LMC**

## 3.1 Semantics

### 3.1.1 Parametrization and Initialization

### 3.1.2 Trace Semantics

Execution as Sparse Matrix Multiplication, (constrained) reachability (pr.), path/trace distributions, ergodicity, state residency time, Uniformization

### 3.1.3 Transient Semantics (incl. Reward/weighted semantics)

wrt. probabilities?!

**Definition 3.1** (Transient probability distribution)**.**

### 3.1.4 Threshold semantics

### 3.1.5 Funtion Transient Semantics (?)

word functions instead of probability functions?

## 3.2 Equivalence & Tanja's Draft Equivalences

exact/ordinary(=strong bisim)/strict lumping, forward/backward/strong bisimulation, weak bisimulation, trace Equivalence, transient equivalence, branching bisimulation (considers internal actions), rooted & Divergence preserving branching bisimulation, finite-horizon bisimulation (bisimulation up to time-step k)

### 3.2.1 Trace semantics is decidable in P for all

paz p.36, Kiefer , Tzeng, Schützenberger, Bollig & Zeitoun, Kiefer WA should hold for PA when using sufficient eps (theoretically unclean), Doyen,
Let automata $A_i = \left(S_i, \Sigma_i, M^{(i)}, \pi_i, \eta_i\right)$ for $i \in \{1, 2\}$ Equivalence of initialized automata:

$$A_1(\pi_1)\ A_2(\pi_2) \Leftrightarrow \forall \sigma \in \Sigma :\ \pi_1 \eta_1 = \pi_2 \eta_2 \wedge \pi_1 M_\sigma^{(1)} = \pi_2 M_\sigma^{(2)}$$

How about uninitialized?

Problem with proof in [1]: don't they construct the powerset (sum-sustituting recursively, i.e. for each letter in alphabet replace in last formula), negating this by "there are only $n_1 + n_2$ linear independent formulas with $n_1 + n_2$ variables...?

#### 3.2.2 Transient semantics is lower NP, upper EXP(?)

As we need to construct the postfix space which is a cylinder set, transfers also to reward/weighted transient

What is T here? A Trace with a certain property (e.g. ending with $\sigma$ or up to $k$ steps?)

$$A_1(\pi_1)\ A_2(\pi_2) \Leftrightarrow \forall \sigma \in \Sigma :\ \pi_1 M_\sigma^{(1)} \eta_1 = \pi_2 M_\sigma^{(2)} \eta_2 \wedge \pi_1 T_1 = \pi_2 T_2$$

#### 3.2.3 Word Function-based

Same as above, just that lower bound is less; stop on first false; maybe eq of kiefer helps here

$$A_1(\pi_1)\ A_2(\pi_2) \Leftrightarrow \forall \sigma \in \Sigma :\ \pi_1 M_\sigma^{(1)} \eta_1 \mu_\sigma = \pi_2 M_\sigma^{(2)} \eta_2 \mu_\sigma \wedge \pi_1 T_1 = \pi_2 T_2$$

### 3.3 Minimization + wrt. Draft Eqs.

**Definition 3.2** (Lumping/lumpable).

#### 3.3.1 Approaches

##### 3.3.1.1 Partition Refinement - Coalgebraic Approach

Proof: PA minimization is in P for uninitialized.
Deiffel, WiSSmann, Paz p.24ff (IntroToProbabilisticAutomata) for all in O(n log n) as partition refinement is minimal (see sources above) for uninitialized Automata.

##### 3.3.1.2 Schützenberger's Construction $\pm$ Arnoldi Iteration with Housholder Reflectors &

Show either that is also in P oder correct Kiefers runtime analysis. For uninitialized see kiefer;

#### 3.3.2 Wrt. different semantics

##### 3.3.2.1 Decidability

trace semantics: is decidable, see e.g. Kiefer or Mateus, Qiu, Li;;;; Trace equivalence PSpace Complete (according to TU Eindhoven's j.f. Groote
Transient semantics ? finite horizon bisimulation minimization using partition refinement
weighted transient semantics?

##### 3.3.2.2 Complexity

trace semantics: TODO figure out if kiefers reduction is flawed or if the runtime analysis of his algo is flawed.
Transient semantics? n log m ?
weighted transient semantics?

# 4 Conclusions