

Minimization of Stochastic Dynamical Systems

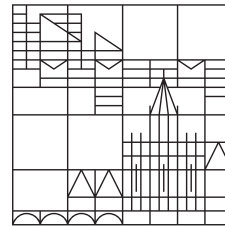
A Comparison of Models & Methods

M.Sc. Project Report

by

Fabian Klopfer

Universität
Konstanz



Faculty of Sciences

Department of Computer and Information Science

Modelling of Complex Self-Organizing Systems Group

Supervisor: Dr. Stefano Tognazzi
Professor: Prof. Dr. Tatjana Petrov

Konstanz, 2020

Contents

1. Introduction	3
2. Models	4
2.1. Automata	4
2.2. Differential Equations	4
2.2.1. Definitions	4
2.2.2. Types of Differential Equations	4
2.2.3. Extensions	4
3. Reductions & Minimization	5
4. Evaluation	6
4.1. Benchmarks	6
A. Software Requirement Specification & Design Documentation	11
A.1. Introduction	11
A.1.1. Purpose	11
A.1.2. Scope	11
A.1.3. Definitions, Acronyms, Abbreviations	11
A.1.4. Overview	11
A.2. Overall description	12
A.2.1. Product Perspective	12
A.2.2. Product Functions	12
A.2.3. User Characteristics	12
A.2.4. Constraints	12
A.2.5. Assumptions and Dependencies	12
A.3. Requirement Specification	13
A.3.1. Software System requirements	13
A.3.2. External Requirements	13
A.3.3. Functional requirements	13
A.4. Proposed Software Architecture	15
A.4.1. Overview	15
A.4.2. Subsystem decomposition	15
A.5. Algorithm Test Cases	20
A.5.1. Weighted Automata	20

1. Introduction

2. Models

2.1. Automata

2.2. Differential Equations

2.2.1. Definitions

2.2.2. Types of Differential Equations

2.2.3. Extensions

3. Reductions & Minimization

4. Evaluation

4.1. Benchmarks

Bibliography

- [1] Steven Abney, David McAllester, and Fernando Pereira. "Relating probabilistic grammars and automata". In: *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*. 1999, pp. 542–549.
- [2] Giovanni Bacci et al. "On the metric-based approximate minimization of Markov Chains". In: *Journal of Logical and Algebraic Methods in Programming* 100 (2018), pp. 36–56.
- [3] Andrea Bobbio and Kishor S. Trivedi. "An aggregation technique for the transient analysis of stiff Markov chains". In: *IEEE Transactions on computers* 9 (1986), pp. 803–814.
- [4] François Boulier et al. "Towards an automated reduction method for polynomial ODE models of biochemical reaction systems". In: *Mathematics in Computer Science* 2.3 (2009), pp. 443–464.
- [5] George Edward Briggs and John Burdon Sanderson Haldane. "A note on the kinetics of enzyme action". In: *Biochemical journal* 19.2 (1925), p. 338.
- [6] Janusz A Brzozowski. "Canonical regular expressions and minimal state graphs for definite events". In: *Mathematical theory of Automata* 12.6 (1962), pp. 529–561.
- [7] Peter Buchholz. "Bisimulation relations for weighted automata". In: *Theoretical Computer Science* 393.1-3 (2008), pp. 109–123.
- [8] Peter Buchholz. "EXACT AND ORDINARY LUMPABILITY IN FINITE MARKOV CHAINS". In: 1994.
- [9] RG Bukharaev. "Probabilistic automata". In: *Journal of Soviet Mathematics* 13.3 (1980), pp. 359–386.
- [10] Luca Cardelli et al. "Maximal aggregation of polynomial dynamical systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 114 38 (2017), pp. 10029–10034.
- [11] Charles Francis Curtiss and Joseph O Hirschfelder. "Integration of stiff equations". In: *Proceedings of the National Academy of Sciences of the United States of America* 38.3 (1952), p. 235.
- [12] Pedro R DArgenio and Joost-Pieter Katoen. "A theory of stochastic systems part I: Stochastic automata". In: *Information and computation* 203.1 (2005), pp. 1–38.
- [13] Pedro R DArgenio and Joost-Pieter Katoen. "A theory of stochastic systems. Part II: Process algebra". In: *Information and Computation* 203.1 (2005), pp. 39–74.
- [14] Peter Deuflhard and Jörg Heroth. "Dynamic dimension reduction in ODE models". In: *Scientific computing in chemical engineering*. Springer, 1996, pp. 29–43.
- [15] Manfred Droste, Werner Kuich, and Heiko Vogler. *Handbook of weighted automata*. Springer Science & Business Media, 2009.
- [16] Manfred Droste and Dietrich Kuske. "Weighted automata". In: *Handbook AutoMathA* (2013).
- [17] Albert Einstein. "On the theory of the Brownian movement". In: *Ann. Phys* 19.4 (1906), pp. 371–381.
- [18] Kousha Etesami. "Analysis of probabilistic processes and automata theory". In: *Automata: from Mathematics to Applications* (2016).

- [19] Bernhard C Geiger et al. "Optimal kullback–leibler aggregation via information bottleneck". In: *IEEE Transactions on Automatic Control* 60.4 (2014), pp. 1010–1022.
- [20] Stuart Geman and Mark Johnson. "Probabilistic grammars and their applications". In: *International Encyclopedia of the Social & Behavioral Sciences* 2002 (2002), pp. 12075–12082.
- [21] G Grinstein, C Jayaprakash, and Yu He. "Statistical mechanics of probabilistic cellular automata". In: *Physical review letters* 55.23 (1985), p. 2527.
- [22] Jack K Hale et al. *Introduction to functional differential equations*. Vol. 99. Springer Science & Business Media, 1993.
- [23] Thomas Henzinger, Barbara Jobstmann, and Verena Wolf. "Formalisms for Specifying Markovian Population Models". In: vol. 22. Sept. 2009. DOI: 10.1007/978-3-642-04420-5_2.
- [24] John Hopcroft. "An $n \log n$ algorithm for minimizing states in a finite automaton". In: *Theory of machines and computations*. Elsevier, 1971, pp. 189–196.
- [25] Tao Jiang and Bala Ravikumar. "Minimal NFA problems are hard". In: *SIAM Journal on Computing* 22.6 (1993), pp. 1117–1141.
- [26] Tsunehiko Kameda and Peter Weiner. "On the state minimization of nondeterministic finite automata". In: *IEEE Transactions on Computers* 100.7 (1970), pp. 617–627.
- [27] Michael N Katehakis and Laurens C Smit. "A successive lumping procedure for a class of Markov chains". In: *Probability in the Engineering and Informational Sciences* 26.4 (2012), pp. 483–508.
- [28] Stefan Kiefer and Björn Wachter. "Stability and complexity of minimising probabilistic automata". In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2014, pp. 268–279.
- [29] Stefan Kiefer and Björn Wachter. "Stability and complexity of minimising probabilistic automata". In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2014, pp. 268–279.
- [30] Stefan Kiefer et al. "On the Complexity of Equivalence and Minimisation for Q-weighted Automata". In: *Logical Methods in Computer Science* 9 (2013).
- [31] Peter Kunkel and Volker Mehrmann. *Differential-algebraic equations: analysis and numerical solution*. Vol. 2. European Mathematical Society, 2006.
- [32] Wilhelm Kutta. "Beitrag zur naherungsweise integration totaler differentialgleichungen". In: *Z. Math. Phys.* 46 (1901), pp. 435–453.
- [33] Pierre-Simon Laplace. *Philosophical Essay on Probabilities: Translated from the fifth French edition of 1825*. Vol. 13. Springer Science & Business Media, 1998.
- [34] Cristopher Moore and James P Crutchfield. "Quantum automata and quantum grammars". In: *Theoretical Computer Science* 237.1-2 (2000), pp. 275–306.
- [35] Isaac Newton. "Methodus fluxionum et serierum infinitarum". In: *Opuscula mathematica, philosophica et philologica* 1 (1774), p. 1774.
- [36] Bernt Øksendal. "Stochastic differential equations". In: *Stochastic differential equations*. Springer, 2003, pp. 65–84.
- [37] Stephan Olariu and Albert Y Zomaya. "Cellular Automata, PDEs, and Pattern Formation". In: *Handbook of Bioinspired Algorithms and Applications*. Chapman and Hall/CRC, 2005, pp. 291–302.
- [38] Alan V Oppenheim. *Discrete-time signal processing*. Pearson Education India, 1999.
- [39] Steven A Orszag. "Numerical methods for the simulation of turbulence". In: *The Physics of Fluids* 12.12 (1969), pp. II–250.
- [40] Azaria Paz. *Introduction to probabilistic automata*. Academic Press, 2014.

- [41] Tatjana Petrov. "Markov chain aggregation and its application to rule-based modelling". In: *CoRR* abs/1812.09774 (2018). arXiv: 1812.09774. URL: <http://arxiv.org/abs/1812.09774>.
- [42] Linda Petzold. "Differential/Algebraic Equations Are Not ODEs". In: *SIAM J. Sci. Stat. Comput.* 3.3 (Sept. 1982), pp. 367–384. ISSN: 0196-5204. DOI: 10.1137/0903023. URL: <https://doi.org/10.1137/0903023>.
- [43] Michael O Rabin. "Probabilistic automata". In: *Information and control* 6.3 (1963), pp. 230–245.
- [44] John Rhodes, Chrystopher L Nehaniv, and Morris W Hirsch. *Applications of automata theory and algebra: via the mathematical theory of complexity to biology, physics, psychology, philosophy, and games*. World Scientific, 2010.
- [45] Jean-Pierre Richard. "Time-delay systems: an overview of some recent advances and open problems". In: *Automatica* 39.10 (2003), pp. 1667–1694. ISSN: 0005-1098. DOI: [https://doi.org/10.1016/S0005-1098\(03\)00167-5](https://doi.org/10.1016/S0005-1098(03)00167-5). URL: <http://www.sciencedirect.com/science/article/pii/S0005109803001675>.
- [46] Claus Scheiderer. *Script: Linear Algebra I*. Department of Mathematics and Statistics, University of Konstanz, 2017.
- [47] Uwe Schöning. *Theoretische Informatik-kurz gefasst*. BI-Wiss.-Verlag, 1992.
- [48] Marcel Paul Schützenberger. "On the definition of a family of automata". In: *Information and control* 4.2-3 (1961), pp. 245–270.
- [49] C. E. Shannon and J. McCarthy. *Automata Studies.(AM-34)*. Vol. 34. Princeton University Press, 1956.
- [50] George F Simmons. *Differential equations with applications and historical notes*. CRC Press, 2016.
- [51] Ana Sokolova and Erik P De Vink. "Probabilistic automata: system types, parallel composition and comparison". In: *Validation of Stochastic Systems*. Springer, 2004, pp. 1–43.
- [52] Ana Sokolova and Erik P De Vink. "Probabilistic automata: system types, parallel composition and comparison". In: *Validation of Stochastic Systems*. Springer, 2004, pp. 1–43.
- [53] Mariëlle Stoelinga. "An introduction to probabilistic automata". In: *Bulletin of the EATCS* 78.176-198 (2002), p. 2.
- [54] Stephen Strogatz. "Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering (studies in nonlinearity)". In: (2001).
- [55] Robert Tarjan. "Dept-first search and linear graph algorithms". In: *Siam Journal on Computing - SIAMCOMP* 1 (Jan. 1997).
- [56] Antti Valmari and Giuliana Franceschinis. "Simple $O(m \log n)$ time Markov chain lumping". In: *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer. 2010, pp. 38–52.
- [57] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier, 1992.
- [58] Vito Volterra. "Theory of functionals and of integral and integro-differential equations". In: (1959).
- [59] John Von Neumann. *Mathematical foundations of quantum mechanics: New edition*. Princeton university press, 2018.
- [60] John Von Neumann, Arthur W Burks, et al. "Theory of self-reproducing automata". In: *IEEE Transactions on Neural Networks* 5.1 (1966), pp. 3–14.

- [61] Thorsten Wissmann et al. "From Generic Partition Refinement to Weighted Tree Automata Minimization". In: *arXiv preprint arXiv:2004.01250* (2020).
- [62] Daniel Zwillinger. *Handbook of differential equations*. Vol. 1. Gulf Professional Publishing, 1998.

A. Software Requirement Specification & Design Documentation

A.1. Introduction

A.1.1. Purpose

The purpose of this Software Requirements Specification is to describe the features, constraints and demands of a tool for model reduction and minimization of stochastic dynamical systems in detail. This document is intended for both the supervisors and the developers of the tool.

A.1.2. Scope

This software system shall implement a tool for model reduction and minimization of stochastic dynamical systems, that provides support for differential equations and automata, equivalent benchmark examples for each, at least one algorithm for the reduction of each — namely [30] and [10]. An optional feature is the implementation of a conversion algorithm between automata and differential equations, when theoretically possible.

A.1.3. Definitions, Acronyms, Abbreviations

Differential equation (DE) An equation relating a function to its derivative with respect to one or more independent variable.

Weighted automata (WA) An automaton consisting of states and a transition function which associates a weight to each transition

Possible DE extensions Extensions of the differential equation framework. For details on the specifics, see section 2.2.3

Isomorphism A structure-preserving, revertable, i.e. bijective morphism or mapping from a domain to a codomain

A.1.4. Overview

In the second chapter several conditions, assumptions and circumstances will be mentioned, that help characterizing the software's special use case. In the third chapter the concrete requirements are listed.

A.2. Overall description

A.2.1. Product Perspective

The software system shall be the practical and evaluative basis for the underlying theory, described in the previous report. Especially it shall provide an empirical basis for the analysis of equivalence of reduction techniques for DEs and the minimization of WAs, i.e. if reducing a DE using an algorithm, does one arrive at the same complexity as when minimising a weighted automaton representing the same system as the DE. For that it shall offer scalable benchmark examples, that is an example DE with an equivalent WA. Further it shall provide a base for implementing further algorithms or new methods to be developed (e.g. if an isomorphism is found between DEs and WAs a conversion tool).

A.2.2. Product Functions

The system shall take a DE or a WA as input and apply one selected algorithm onto those and then output the minimized model. Further it shall be able to compare the results of a minimisation in terms of model complexity and equivalence, i.e. answer the question whether DE reduction and WA minimisation yield different representations of the same model or if one model is more complex.

A.2.3. User Characteristics

For now the only user is going to be the developer in the course of the master project and thesis. When successful enough this may or may not change. This implies that maintainability and extensibility is important, if it's ever handed-over to another developer or even extended to be usable by non-developers.

A.2.4. Constraints

- Run time is crucial for these algorithms to be feasible. The system shall not impose significant overhead and the implementations shall be as fast as possible
- Memory safety is important as inputs may induce infinite memory requirements that must not execute if so (else the system will crash)
- Correctness is the most important aspect of the implementation as the results will be unusable if minor errors are introduced that corrupt the reduction or minimization processes as well as comparisons and benchmarks

A.2.5. Assumptions and Dependencies

to be set when dependencies are fixed

CMake, C++,

one of OpenCL & cuda, **ViennaCL**, LAPACK, BLAS, libflame, BLIS, Eigen, Boost, Amardillo, SuiteSparse, CSpase, Blaze, MKL, GNU Scientific lib, dlib, GLM, gotoblas, atlas, PETSc, **MAGMA**, plasma ...

A.3. Requirement Specification

A.3.1. Software System requirements

The system shall be implemented with focus on the following traits:

- S 1. Correctness
- S 2. Extensibility
- S 3. Maintainability
- S 4. Lightweightness
- S 5. As fast as possible
- S 6. Memory safe for all inputs
- S 7. Cross-compileable.

A.3.2. External Requirements

The system shall:

- E 1. to be set when dependencies are fixed

A.3.3. Functional requirements

- F 1. Control flow system

The system shall

- 1. provide a mechanism for applying the methods depending on the specified model, i.e. apply WA validation checks, algorithm and evaluation for WA input and the respective equivalent for DEs.
- 2. provide a file based input and output mechanism.
- 3. provide a user interface based input and output mechanism.
- 4. validate the inputs before algorithm execution

- F 2. User interface

The system shall

- 1. offer 3 different methods to be applied: WA minimization, DE reduction and Benchmarking.
- 2. provide a way to specify which input method is to be used, either file-based or user interface-based.
- 3. provide a way to specify which output method is to be used.
- 4. provide the possibility to select an algorithm from a set of available ones for both models.
- 5. display the input, the output and summaries and statistics provided by the model specific modules

- F 3. WA minimization

The system shall

- 1. specify an input format for WA.
- 2. provide a validation mechanism to ensure the well-formedness of the input.
- 3. implement the algorithm specified in [30].
- 4. use the algorithm above on an example specified in A.5.1 and work deterministically correct with a fixed set of random vectors.
- 5. summarize the results and statistics on the difference between input and output.

- F 4. DE reduction

The system shall

1. specify an input format for arbitrary DEs and possible extensions, specifically stochastic differential, functional differential, integro-differential and differential-algebraic equations.
2. provide a validation mechanism to ensure the well-formedness of the expression.
3. implement the algorithm specified in [10]
4. summarize the results and statistics on the difference between inputs and outputs

F 5. Benchmark The system shall

1. provide example WA and DEs, which are
 1. theoretically proven to be equivalent, see section 4.1
 2. scalable by a parameter, defining the number of states or the number of variables
2. implement a method to compare the results in terms of
 1. run time,
 2. memory
 3. resulting model complexity

F 6. Conversion or Equivalence The system shall

1. be able to convert the benchmark examples from one model to the other or to prove that instances of distinct models are equivalent
2. (Optional) be able to convert all instances of one model to the other model

A.4. Proposed Software Architecture

A.4.1. Overview

Main Module Entry point of the process, handles for the control flow.

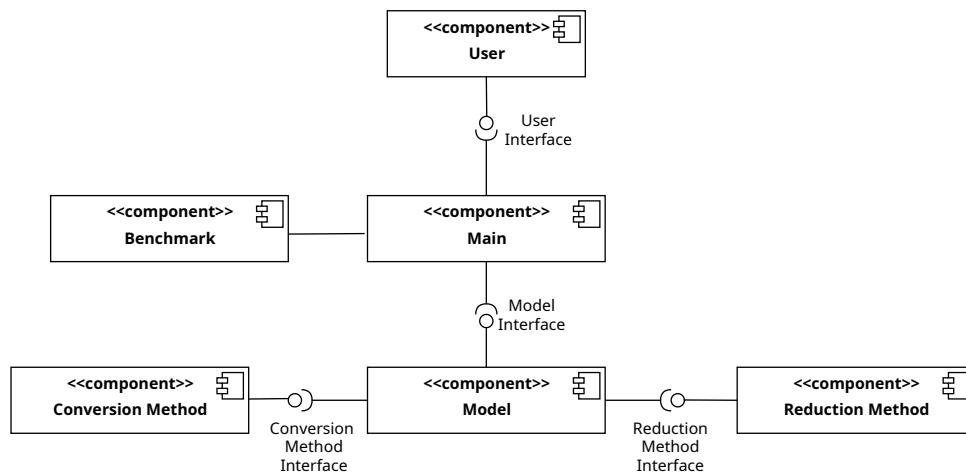
User Interface Lets the user control the main module.

Model Interface Defines a model including format.

Reduction Method Interface Defines a reduction method for a specific model.

Conversion/Equivalence Interface Defines the conversion or equivalence notion from a specific model to a specific model.

Benchmark Module Defines benchmark examples and generation processes for those.



A.4.2. Subsystem decomposition

Main Module The main module is the main entry point of the process, responsible for executing the user interface, the algorithms and utilities from the model specific modules, the conversions, as well as dealing with benchmark flows and loading the necessary examples. There are two mandatory basic control flows and one optional one that depends on the theoretical feasibility that is to be proven:

1. Standard model reduction for a single model with user specified input
2. Benchmark flow using the provided examples, resulting in two sequential model reduction executions. They can not be parallelized, as they would interfere with the caches and slow each other down. After the execution it converts one instance to the other model preserving model complexity or checks if they are equivalent.

User Interface The User Interface lets the user control the main module, provides means of input to the main module and displays the output from the main module. Also contains usage information.

A text-based user interface is to be implemented (also called command-line interface). The user interface must contain a menu letting the user specify what control flow is to be executed, thus what task the user wants to perform. It must be able to either read the data from file or from the standard input of the terminal and display the input, output and summary on the standard output of the terminal, as well as possible errors.

The UserInterface provides functionality to control the program and to deliver outputs to the user. Therefore most functions are parameterless and take their input from the user, so those must ensure the validity of the user input. The only special case is the actual input data which is validated by a model specific function and the output functions which do not take user input.

```

Interface UserInterface {
    public Task select_task() {
        ensure Task exists
    }

    public ModelInterface select_model() {
        ensure model exists
    }

    public ReductionMethodInterface select_method(ModelInterface model) {
        require model not null
        ensure method exists && method compatible to model
    }

    public IOMethod select_input_method() {
        ensure input method exists
    }

    public FILE* file_input() {
        ensure file exists && file is closed after read
    }

    public char* stdin_input() {
        require limited input
        ensure input is read completely or rejected
    }

    public IOMethod select_output_method() {
        ensure output method exists
    }

    public FILE* set_output_location() {
        ensure file exists or create new file && containing directory exists
    }

    public void display_exec_file(FILE* output_destination) {
        require output_destination not null
        assure file is closed after write
    }

    public void display_exec_stdout(char* output) {
        require output not null && output limited
        ensure output is less than default terminal buffer size || output is paged
    }
}

```

Model Interface The model interface specifies the details of a model including the representation format, input validation functionality, concrete instances will hold an array of function pointers of the reduction method interface type for the concrete model. Depending on the task that the user wants to perform either one or another instance is used to execute the control flow. The data flow from the user via the main module is passed to the model interface instance which does the actual processing using the defined functions and one of the functions from the array holding the

model reduction method pointers. Instances of this interface must also specify how to summarize and compute statistics over the result, i.e. the differences in the model and in model complexity. To be implemented are instances for DEs and WAs.

```
Interface ModelInterface {
    ModelRepresentation representation require not null
    ReductionInterface*[] ReductionMethodInterface[]
    ConversionMethodInterface*[] conversions

    public ModelRepresentation validate_model_instance(char* input) {
        require input not null
        ensure model is valid wrt. representation
    }

    public char* summarize_reduction(ModelRepresentation input, ModelRepresentation reduced) {
        require input not null && reduced not null
    }
}
```

Model Representation Interface As the data needs to be validated, this interface shall specify a structure for holding an instance of the model.

```
Interface ModelRepresentationInterface {
    struct ModelRepresentation require not null
}
```

Reduction Method Interface This interface is used to implement concrete reduction methods for a specific model. For each models there may be multiple reduction methods. To be implemented are the reduction methods specified in [10, 30].

```
Interface ReductionMethodInterface {
    public ModelRepresentation reduce(ModelRepresentation input) {
        require input not null
        ensure return value valid and not null
    }
}
```

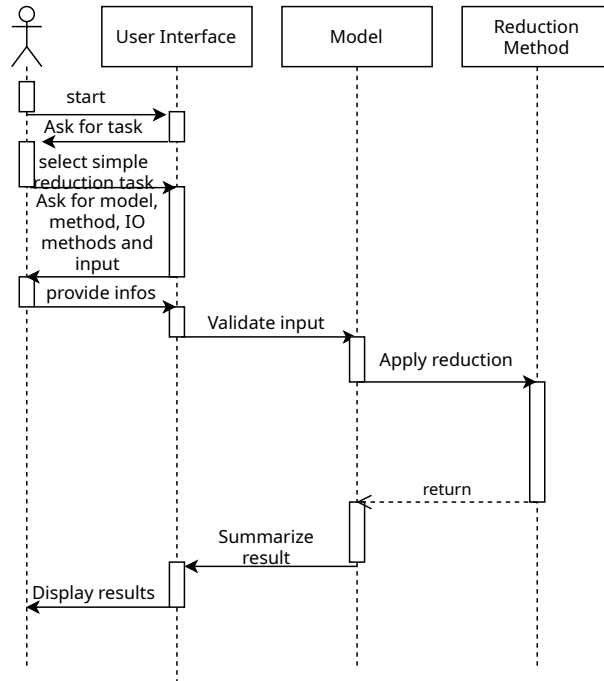
Conversion/Equivalence Interface Using the model interfaces, implementations of this interface shall provide a method to convert an instance of a model to another model or to prove the equivalence of two instances of different models. As of now this is a question to be further investigated in terms of theory. One open question is if there exists an isomorphism between the to be implemented methods and if so how to do the conversion or equivalence check without also transforming the model complexity (e.g. from DEs to WAs, how not to produce more or less states than there are implicitly in the DE representation). The interface should optimally provide methods for conversion in both direction, i.e. a programatically equivalence proof and shall be attached to both models that it converts as pointer.

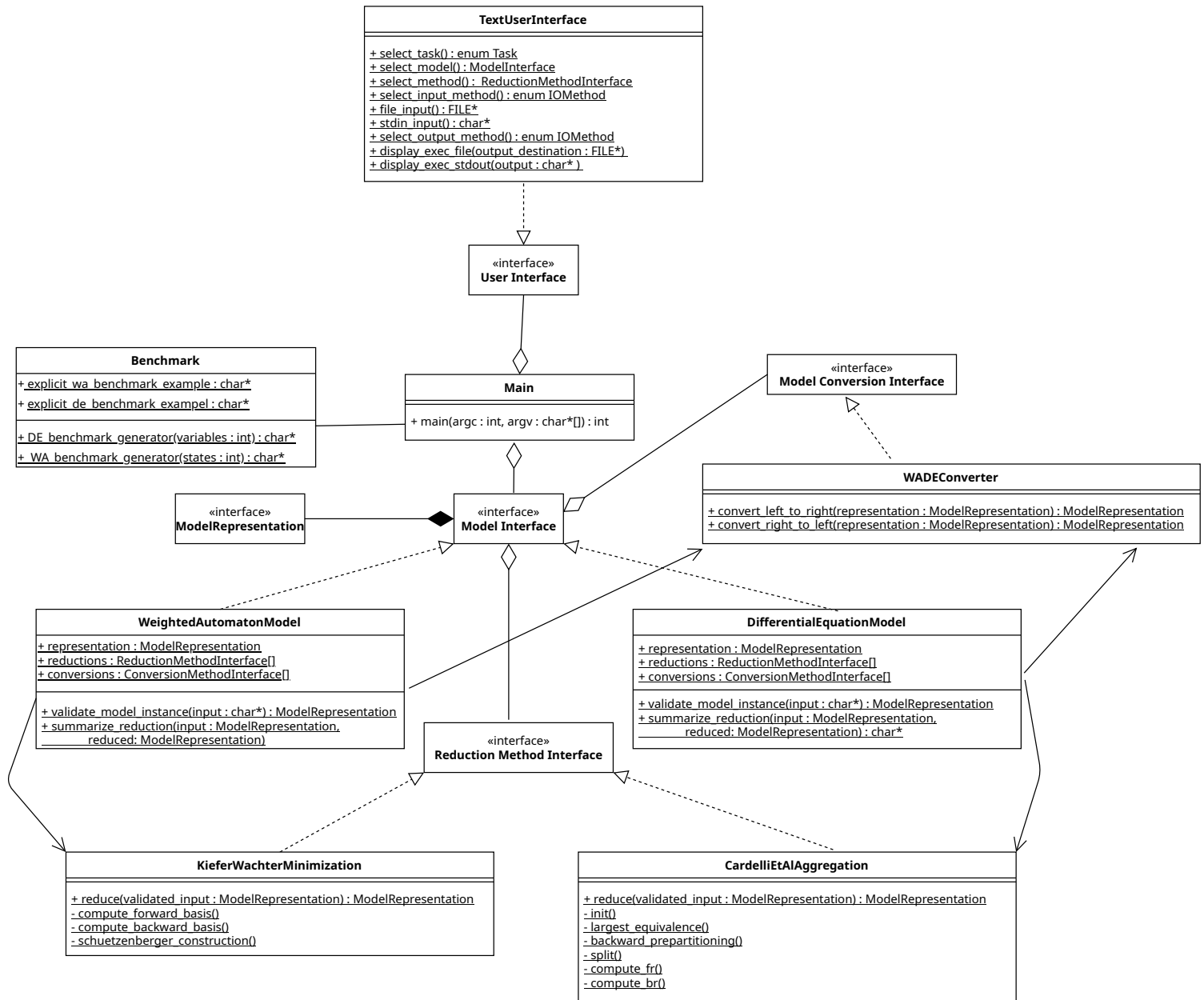
```
Interface ModelConversionInterface {
    public ModelRepresentation convert_left_to_right(ModelRepresentation left) {
        require left not null
        ensure output is valid in rights's representation
    }
}
```

```

ModelRepresentation convert_right_to_left(ModelRepresentation right) {
    require right not null
    ensure output is valid in left's representation
}
}

```





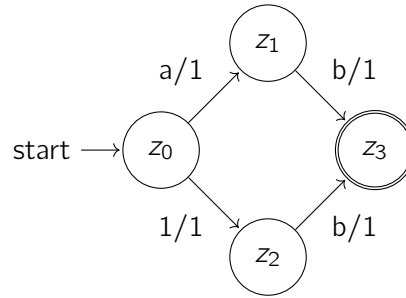
A.5. Algorithm Test Cases

A.5.1. Weighted Automata

Weighted automaton $\mathcal{A} = (n, \Sigma, \mu, \alpha, \eta)$ with

- $n = 3$,
- $\Sigma = \{a, b\}$,
- $\alpha = (1, 0, 0, 0)$,
- $\eta = (0, 0, 0, 1)$,

$$\begin{aligned} - \mu(a) &= \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ - \mu(b) &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$



All matrix multiplications were performed using Wolframalpha, all other results were verified using Wolframalpha.

Restrictions: Those are actually obvious from the automaton. To be on the safe side the “proofs” are stated explicitly here.

- We only need to evaluate words starting with “a” and ending with “b”.

$$\alpha\mu(a) = (0, 1, 1, 0); \quad \alpha\mu(b) = (0, 0, 0, 0); \quad \mu(a)\eta = (0, 0, 0, 0)^T; \quad \mu(b)\eta = (0, 1, 1, 0)^T$$

- Words containing two consecutive letters do not need to be considered:

$$\mu(a)^2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad \mu(b)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Words may only contain “ab” as “ba” yields the zero matrix and matrix multiplication is associative (i.e. $(\mu(a)\mu(b))\mu(a) = \mu(a)(\mu(b)\mu(a)) = \mu(a) \cdot 0 = 0$)

$$\mu(a)\mu(b) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mu(b)\mu(a) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Thus only words of length smaller or equal to 2 must be considered, starting with a or ending with b. This follows from the above, namely that only words starting with “a”, not having two consecutive letters or “ba” in them yield non-zero vectors or matrices.

Forward reduction, i.e. applying [30] Figure 3: The only words that need to be considered due to the above restrictions are the words “a” and “ab”.

Random vectors are generated using Wolframalpha and $K = 3$, i.e. $r^i \in \{1, \dots, 12\}^{2 \times 4}$:

$$r^{(1)} = \begin{pmatrix} 9 & 5 & 5 & 7 \\ 6 & 11 & 2 & 1 \end{pmatrix}; \quad r^{(2)} = \begin{pmatrix} 2 & 3 & 1 & 2 \\ 12 & 3 & 9 & 4 \end{pmatrix} \quad r^{(3)} = \begin{pmatrix} 2 & 7 & 9 & 10 \\ 1 & 11 & 2 & 6 \end{pmatrix} \quad r^{(4)} = \begin{pmatrix} 4 & 5 & 2 & 10 \\ 5 & 9 & 5 & 5 \end{pmatrix}$$

As the two words are the same for all v_1, \dots, v_n we have:

$$\rho(r^{(i)}) = \sum_{k=0}^n \sum_{w \in \Sigma^k} \alpha \sum_{j=1}^{|w|} \mu(w_j) r_{w_j j}^{(i)} = \alpha \mu(a) r_{a,1}^{(i)} + \alpha \mu(a) \mu(b) r_{a,1}^{(i)} r_{b,2}^{(i)} = (0, 1, 1, 0) r_{a,1}^{(i)} + (0, 0, 0, 2) r_{a,1}^{(i)} r_{b,2}^{(i)}$$

Thus:

$$v_1 = 9 \cdot (0, 1, 1, 0) + 9 \cdot 11 \cdot (0, 0, 0, 2) = (0, 9, 9, 198)$$

$$v_2 = 2 \cdot (0, 1, 1, 0) + 2 \cdot 3 \cdot (0, 0, 0, 2) = (0, 2, 2, 12)$$

$$v_3 = 2 \cdot (0, 1, 1, 0) + 2 \cdot 11 \cdot (0, 0, 0, 2) = (0, 2, 2, 44)$$

$$v_4 = 4 \cdot (0, 1, 1, 0) + 4 \cdot 9 \cdot (0, 0, 0, 2) = (0, 4, 4, 72)$$

According to Wolframalpha the maximally linearly independent subset is $\{v_1, v_2\}$ and by that

$$\vec{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 9 & 9 & 198 \\ 0 & 2 & 2 & 12 \end{pmatrix}$$

Question: Did the algorithm forget α in the linear independent check? In the algorithm of [30], the authors state “let k be maximum such that $\{v_1, \dots, v_k\}$ are linearly independent”. Shouldn’t it be “let k be maximum such that $\{\alpha, v_1, \dots, v_k\}$ are linearly independent”?

Applying the forward reduction using Schützenberger construction as described in [30] Eq. 4.2:

$$\begin{aligned} \vec{F} \mu(a) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 9 & 9 & 198 \\ 0 & 2 & 2 & 12 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \vec{\mu}(a) \vec{F} \\ &= \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 9 & 9 & 198 \\ 0 & 2 & 2 & 12 \end{pmatrix} = \begin{pmatrix} m_{11} & 9m_{12} + 2m_{13} & 9m_{12} + 2m_{13} & 198m_{12} + 12m_{13} \\ m_{21} & 9m_{22} + 2m_{23} & 9m_{22} + 2m_{23} & 198m_{22} + 12m_{23} \\ m_{31} & 9m_{32} + 2m_{33} & 9m_{32} + 2m_{33} & 198m_{32} + 12m_{33} \end{pmatrix} \end{aligned}$$

More compactly represented in order to set up the system of linear equations to be solved:

$$\begin{pmatrix} m_{11} & 9m_{12} + 2m_{13} & 9m_{12} + 2m_{13} & 198m_{12} + 12m_{13} \\ m_{21} & 9m_{22} + 2m_{23} & 9m_{22} + 2m_{23} & 198m_{22} + 12m_{23} \\ m_{31} & 9m_{32} + 2m_{33} & 9m_{32} + 2m_{33} & 198m_{32} + 12m_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

From that it immediately follows, that $m_{11} = m_{21} = m_{31} = 0$. For m_{12} and m_{13} we get

$$\left(\begin{array}{cc|c} 9 & 2 & 1 \\ 198 & 12 & 0 \end{array} \right) \xrightarrow{II=II-6I} \left(\begin{array}{cc|c} 9 & 2 & 1 \\ 144 & 0 & -6 \end{array} \right) \Rightarrow m_{12} = \frac{-1}{24}, \quad m_{13} = \frac{1}{2} \left(1 - 9 \cdot \frac{-1}{24} \right) = \frac{33}{48}$$

As m_{22}, m_{23} and m_{32}, m_{33} are defined by the same system of equation, their value is equal. That is:

$$\left(\begin{array}{cc|c} 9 & 2 & 0 \\ 198 & 12 & 0 \end{array} \right) \xrightarrow{II=II-6I} \left(\begin{array}{cc|c} 9 & 2 & 0 \\ 144 & 0 & 0 \end{array} \right) \Rightarrow m_{22} = m_{23} = 0 = m_{32} = m_{33}$$

Thus

$$\vec{\mu}(a) = \begin{pmatrix} 0 & \frac{-1}{24} & \frac{33}{48} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Regarding $\vec{\mu}(b)$ we get

$$\begin{aligned} \vec{F}_{\mu(b)} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 9 & 9 & 198 \\ 0 & 2 & 2 & 12 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 18 \\ 0 & 0 & 0 & 4 \end{pmatrix} = \vec{\mu}(a) \vec{F} \\ \Rightarrow & \begin{pmatrix} h_{11} & 9h_{12} + 2h_{13} & 9h_{12} + 2h_{13} & 198h_{12} + 12h_{13} \\ h_{21} & 9h_{22} + 2h_{23} & 9h_{22} + 2h_{23} & 198h_{22} + 12h_{23} \\ h_{31} & 9h_{32} + 2h_{33} & 9h_{32} + 2h_{33} & 198h_{32} + 12h_{33} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 18 \\ 0 & 0 & 0 & 4 \end{pmatrix} \end{aligned}$$

From that it immediately follows, that $h_{11} = h_{21} = h_{31} = 0$. For h_{12} and h_{13} we get

$$\left(\begin{array}{cc|c} 9 & 2 & 0 \\ 198 & 12 & 0 \end{array} \right) \xrightarrow{II=II-6I} \left(\begin{array}{cc|c} 9 & 2 & 0 \\ 144 & 0 & 0 \end{array} \right) \Rightarrow h_{12} = h_{13} = 0$$

$$\left(\begin{array}{cc|c} 9 & 2 & 0 \\ 198 & 12 & 18 \end{array} \right) \xrightarrow{II=II-6I} \left(\begin{array}{cc|c} 9 & 2 & 0 \\ 144 & 0 & 18 \end{array} \right) \Rightarrow h_{22} = \frac{1}{8}, \quad h_{23} = \frac{1}{2}(-9 \cdot \frac{1}{8}) = \frac{-9}{8}$$

$$\left(\begin{array}{cc|c} 9 & 2 & 0 \\ 198 & 12 & 4 \end{array} \right) \xrightarrow{II=II-6I} \left(\begin{array}{cc|c} 9 & 2 & 0 \\ 144 & 0 & 4 \end{array} \right) \Rightarrow m_{12} = \frac{1}{36}, \quad m_{13} = \frac{1}{2}(-9 \cdot \frac{1}{36}) = \frac{-1}{8}$$

So that

$$\vec{\mu}(b) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{-9}{8} \\ 0 & \frac{1}{36} & \frac{-1}{8} \end{pmatrix}$$

and

$$\vec{\eta} = \vec{F}\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 9 & 9 & 198 \\ 0 & 2 & 2 & 12 \end{pmatrix} \cdot (0, 0, 0, 1)^T = (0, 198, 12)^T$$

And finally $\vec{\mathcal{A}} = (3, \{a, b\}, \vec{\mu}, (1, 0, 0), (0, 198, 12)^T)$

Question: α is defined as e_1 in the paper. I just assumed that it's the first column vector of \vec{F} . Is that correct?

As the automaton is minimal with 3 states, the backwards reduction is omitted for now (will be added when correctness of the test case is affirmed or rejected). Further another paper by the same author [28] states with reference to [48] that $\mathcal{A} \equiv \vec{\mathcal{A}} \equiv \overleftarrow{\mathcal{A}}$, also providing a proof of this in the Appendix A.1

To assure that this is still the same automaton, compute the following examples with the new automaton:

- $\vec{\alpha} \vec{\mu}(a) = (0, \frac{-1}{24}, \frac{33}{48})$; $\vec{\alpha} \vec{\mu}(b) = (0, 0, 0)$
- $\vec{\mu}(a) \vec{\eta} = \begin{pmatrix} 0 & \frac{-1}{24} & \frac{33}{48} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} (0, 198, 12)^T = (198 \frac{-1}{24} + 12 \frac{33}{48}, 0, 0)^T = (0, 0, 0)^T$
- $\vec{\mu}(b) \vec{\eta} = (0, \frac{45}{4}, 4)$
- $\vec{\alpha} \vec{\mu}(a) \vec{\eta} = 0 + \frac{198}{1536} - \frac{132}{1024} = 0$

- $\vec{\mu}(a)^2 = 0$, but $\vec{\mu}(b)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{-1}{64} & 0 \\ 0 & 0 & \frac{11}{64} \end{pmatrix}$
- $\vec{\alpha} \vec{\mu}(a) \vec{\mu}(b) = (1, 0, 0)^T \begin{pmatrix} 0 & \frac{-1}{24} & \frac{33}{48} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{-9}{8} \\ 0 & \frac{1}{36} & \frac{-1}{8} \end{pmatrix} = (0, \frac{1}{72}, \frac{-5}{128})$
- $\vec{\alpha} \vec{\mu}(a) \vec{\mu}(b) \vec{\eta} = (0, \frac{1}{72}, \frac{-5}{128}) \cdot (0, 198, 12)^T = \frac{73}{32}$
- $\vec{\alpha} \vec{\mu}(a) \vec{\mu}(b) \vec{\mu}(a) \vec{\eta} = (0, \frac{1}{72}, \frac{-5}{128}) \cdot (0, 0, 0)^T = 0$
- $\vec{\alpha} \vec{\mu}(a) \vec{\mu}(b) \vec{\mu}(b) \vec{\eta} = (0, \frac{1}{72}, \frac{-5}{128}) \cdot (0, \frac{45}{4}, 4) = 0$
- **This should not be non-zero, something is wrong!**
- $\vec{\alpha} \vec{\mu}(a) \vec{\mu}(b) \vec{\mu}(b) \vec{\mu}(b) \vec{\eta} = (0, \frac{1}{72}, \frac{-5}{128}) \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{-9}{8} \\ 0 & \frac{1}{36} & \frac{-1}{8} \end{pmatrix} \cdot (0, \frac{45}{4}, 4) = \frac{-73}{2048}.$

With the addition of one additional b the weight is again zero, but for two more bs it's again non-zero. That is the "minimization" added an alternating mistake.

There are four obvious reasons:

1. I miscalculated something or there is a typo
2. I did not understand the algorithm and thus applied it wrongly
3. It is necessary to also apply the backwards reduction to correct the mistake introduced by the forward reduction. Thus [28] is incorrect
4. the algorithm itself is incorrect