

Universität  
Rostock



Traditio et Innovatio

# Lifted Bayesian Filtering in Multi-Entity Systems

## Dissertation

to obtain the academic degree of  
**Doktor-Ingenieur (Dr.-Ing.)**  
of the Faculty of Computer Science and Electrical Engineering  
at the University of Rostock

submitted by

**Stefan Lüdtke**

born on August 5, 1991 in Rostock

November 23, 2020

## Abstract

*Bayesian filtering* (BF) is a general probabilistic framework for estimating the state of a dynamic system that can be observed only indirectly through noisy measurements. This thesis focuses on systems that consist of multiple, interacting *entites* (e.g. agents or objects), for which the system dynamics can be specified naturally by *multiset rewriting systems* (MRSs). Unfortunately, BF in MRSs is computationally challenging due to the combinatorial explosion in the state space size.

Therefore, we investigate efficient BF algorithms for such multi-entity systems. The main insight is that the state space that is underlying an MRS exhibits a certain *symmetry*, which can be exploited to increase inference efficiency.

This thesis provides five main contributions. First, we show how distributions over multisets can be decomposed into two factors: A distribution over the structures and multiplicities of entities, and a distribution over values of the entities' properties. This representation allows to group together entities with identical structure, thus achieving a substantial reduction in representation complexity. As this representation bears some similarity to other concepts from lifted probabilistic inference, we call it a *lifted* representation.

Secondly, we introduce a BF algorithm that works directly on this lifted representation, which is able to achieve a factorial reduction in space and time complexity, compared to conventional, ground filtering. When observations or system dynamics break symmetry, the algorithm automatically adapts by *splitting*.

When a maximally parallel action execution semantics is used – when all entities can act in parallel – exact BF can become intractable due to the large number of parallel actions. To alleviate this problem, our third contribution is a Markov chain Monte Carlo algorithm that *samples* parallel actions instead of performing full enumeration.

Fourth, we address the problem that due to symmetry breaks, the algorithm must perform splitting, so that the model can become completely propositional over time and inference becomes intractable. This is done by introducing inverse *merging* operations for a number of practically relevant special cases.

Finally, we empirically evaluate the lifted BF algorithm on real-world human activity recognition domains, and show that the algorithm can be more efficient than propositional BF. To the best of our knowledge, this is the first attempt to provide BF for systems with MRS dynamics and the first attempt that allows to perform prediction and update directly on the lifted representation.

## Zusammenfassung

Bayes'sches Filtern (BF) ist ein probabilistischer Ansatz zur Zustandsschätzung dynamischer Systeme, die nur über verrauschte Sensordaten beobachtet werden können. In dieser Arbeit werden Systeme betrachtet, die aus mehreren, interagierenden Entitäten (z.B. Agenten oder Objekten) bestehen. Die Systemdynamik solcher Systeme kann auf natürliche Art durch *Multiset Rewriting Systems* (MRS) spezifiziert werden. BF in MRS ist allerdings durch die kombinatorische Explosion in der Größe des Zustandsraums solcher Systeme rechnerisch aufwendig.

Aus diesem Grund betrachten wir in dieser Arbeit effiziente Algorithmen für BF in Systemen mit MRS-Dynamik. Die wesentliche Erkenntnis ist, dass der Zustandsraum solcher Systeme *Symmetrien* aufweist, die ausgenutzt werden können, um die Effizienz der Inferenz zu erhöhen.

Diese Arbeit leistet fünf wesentliche Beiträge. Als Erstes zeigen wir, wie Verteilungen über Multimengen in zwei Faktoren aufgeteilt werden können: Eine Verteilung über die Struktur und die Multiplizität der Entitäten und eine Verteilung über die Werte der Entitäten. Diese Repräsentation der Verteilung erlaubt es, Entitäten mit identischer Struktur zu gruppieren, sodass die Repräsentationskomplexität beträchtlich reduziert wird. Da diese Repräsentation Zusammenhänge zu anderen Konzepten aus dem Forschungsbereich der *Lifted Probabilistic Inference* aufweist, bezeichnen wir diese als *geliftete* Repräsentation.

Zweitens führen wir einen BF-Algorithmus ein, der direkt auf der gelifteten Repräsentation arbeitet, wodurch sich die Zeit- und Platzkomplexität gegenüber dem grundierten Algorithmus um einen faktoriellen Faktor reduziert. Wenn die Beobachtungen oder die Systemdynamik zu einem Symmetriebruch führen, kann der Algorithmus die Repräsentation automatisch durch *Splitting* anpassen.

Wenn eine maximal-parallele Aktionsausführungs-Semantik genutzt wird, d.h. wenn alle Entitäten gleichzeitig agieren können, ist exaktes BF aufgrund der großen Anzahl paralleler Aktionen nicht mehr durchführbar. Um dieses Problem zu lösen, entwickeln wir einen *Markov Chain Monte Carlo*-Algorithmus, der parallele Aktionen sampelt anstatt diese vollständig aufzuzählen.

Viertens behandeln wir das Problem, dass der Algorithmus durch Symmetriebrüche wiederholt splitten muss, sodass das Modell im Laufe der Zeit vollständig grundiert wird, wodurch Inferenz nicht mehr durchführbar ist. Hierzu führen wir für eine Reihe von praktisch relevanten Fällen *Merging*-Algorithmen ein, die sich invers zum Splitting verhalten.

Schließlich evaluieren wir den gelifteten BF-Algorithmus empirisch mit realen Aktivitätserkennungsszenarien und zeigen, dass der Algorithmus effizienter als grundiertes BF sein kann. Der in dieser Arbeit eingeführte BF-Algorithmus erlaubt erstmals BF in Systemen mit MRS-Dynamik. Es handelt sich um den ersten Ansatz, um BF direkt auf der gelifteten Repräsentation durchzuführen.