# Automata learning: from probabilistic to quantum

Chu, W.

# Chapter 8

# Conclusions

In this thesis, we developed various algorithms for learning automata, varying in the learning methods as well as in the type of automata. After preliminary concepts on probabilistic automata and hidden Markov models, we studied several passive learning probabilistic automata algorithms in Chapters 3 and 4. Our initial research question was if we could develop an effective passive learning algorithm tailored for deterministic probabilistic regular distributions, achieving a separation between structural information and probabilistic characteristics. To answer this question, in Chapter 3 we presented a novel passive learning algorithm that keeps separate structural from probabilistic information. The construction not only enhances the clarity of the model but makes the model learning more modular and flexible, and it makes it easier to identify and address issues related to, e.g. the language accepted by a complex large system, without any obfuscation from the probabilistic information. More importantly, it allowed us to reuse algorithms that are better suited to handle structural aspects without unnecessary entanglement with probability computation. With a few experiments, we have shown the advantage of this method over a classical state merging algorithm. On the negative side, our technique is bound to learn only deterministic probabilistic systems and would be sub-optimal when learning models that have inherent uncertainty.

For this reason, we moved to our second research question asking whether separating structure and probability would help in learning not-necessarily deterministic regular distributions. We give a partial answer in Chapter 4, where we successfully combine the learning of a non-deterministic structure from positive and negative samples from a regular distribution and the learning of the probabilistic structure via different precise and approximation methods. The answer is partial because we know there are non-deterministic regular distributions that cannot be learned precisely with our new method, even if we gradually increase the sample size. Current existing techniques suffer the same problem, but again, our technique separating structure from probabilities proved to be superior in a small experimental benchmark.

Our approach to learning probabilistic automata by first identifying the structure and then optimizing the probabilities allows for a clear separation of concerns, where the structure learning phase focuses on establishing the correct topology of the automaton, followed by a more specialized optimization of transition probabilities. In addition to more precise and robust parameter learning, this two-phase process also improves interpretability, providing better insights into the underlying Markov process being modeled and making the probabilities within a known structure more meaningful. In terms of application areas, even if not treated in this thesis, our method could be particularly beneficial across various domains such as natural language processing, biological sequence analysis, and robotic control systems, where domain-specific structures can be customized for better performance.

Because of the success of our general separation of concerns methodology, our last research question was about the possibility of applying it directly to other systems that exhibit probabilistic behavior, such as quantum systems: simplifying their inherently complex nature by a clear separation between the design of the automaton's topology and the fine-tuning of its transition amplitudes. To answer this question we moved from the context of passive to active learning and took inspiration from learning weighted automata in learning the structure of quantum automata via their Hankel matrices. In Chapter 6 we devise a novel approach to learning measure-once one-way quantum finite automata combining active learning with genetic and evolutionary optimization algorithms. This approach could be used in learning models from the realistic setting of linear optics. Conversely, we show as proof of concept, that 2 measure-once one-way quantum finite automata can be implemented in terms of linear optics. We leave open the research question about the precise correspondence between the two.

Our results pave the way for future research in a broader context. In fact, passive (and to some extent active) automata learning presents a powerful methodology that can be used for deriving concise abstract models of recurrent neural networks (RNNs) [54], thereby facilitating verification and analysis. In this context, automata learning combined with abstraction techniques may help represent essential behavioral patterns and dependencies within the network. The probabilistic automata learned from the RNN's dynamics will enable a more tractable approach to verification tasks using formal methods, such as model checking, to assess properties like correctness and convergence, contributing to the development of more robust and interpretable neural network models.

Even if not directly related to the focus of this thesis as organized by the three research questions above, it is worthwhile mentioning an additional result that may be helpful in a more general context. In Chapter 2 we developed a novel method for effectively calculating the Euclidean distance between regular distributions represented by probabilistic automata.

This algorithm allows us to assess how well the model captures the desired distribution. Additionally, the Euclidean distance is a fundamental metric used in various machine learning algorithms. For an example outside the context of this thesis, our algorithm can be used in optimization tasks for parameter tuning for probabilistic models, where it is important to have a metric that quantifies the difference between the model's output and the target distribution.