

Abstract

Quantitative Automata and Logic for Pictures and Data Words

Parvaneh Babari

(Dissertation, 2017)

Mathematical logic and automata theory are two scientific disciplines with a close relationship that is not only fundamental for many theoretical results but also forms the basis of a coherent methodology for the verification and synthesis of computing systems. This connection goes back to a much longer history in the 1960s, through the fundamental work of Büchi-Elgot-Trakhtenbrot, which shows the expressive equivalence of automata and logical systems such as monadic second-order logic on finite and infinite words. This allowed the handling of specifications (where global system properties are stated), and implementations (which involve the definition of the local steps in order to satisfy the global goals laid out in the specifications) in a single framework. This connection has been extended to and well-investigated for many other structures such as trees, finite pictures, timed words and data words. For many computer science applications, however, quantitative phenomena need to be modelled, as well. Examples are vagueness and uncertainty of a statement, length of time periods, spatial information, and resource consumption. Weighted automata, introduced by Schützenberger, are prominent models for quantitative aspects of systems. The framework of weighted monadic second-order logic over words was first introduced by Droste and Gastin. They gave a characterization of quantitative behavior of weighted finite automata, as semantics of monadic second-order sentences within their logic. Meanwhile, the idea of weighted logics was also applied to devices recognizing more general structures such as weighted tree automata, weighted automata on infinite words or traces. The main goal of this thesis is to give logical characterizations for weighted automata models on pictures and data words as well as for Büchi-tiling systems in the spirit of the classical Büchi-Elgot theorem. As the second goal, we deal with synchronizing problem for data words. Below, we briefly summarize the contents of this thesis. Informally, a two-dimensional string is called a picture and is defined as a rectangular array of symbols taken from a finite alphabet. A two-dimensional language (or picture language) is a set of pictures. Picture languages have been intensively investigated by several research groups. In Chapter 1, we define weighted two-dimensional on-line tessellation automata (W2OTA) taking weights from a new weight structure called picture valuation monoid. This new weighted picture automaton model can be used to model several applications, e.g. the average density of a picture. Such aspects could not be modelled by semiring weighted picture automaton model. The behavior of this automaton model is a picture series mapping pictures over an alphabet to elements of a picture valuation monoid. As one of our main results, we prove a Nivat theorem for W2OTA. It shows that recognizable picture series can be obtained precisely as projections of particularly simple unambiguously recognizable series restricted to unambiguous recognizable picture languages. In addition, we introduce a weighted monadic second-order logic (WMSO) which can model average density of pictures. As the other main result, we show that W2OTA and a suitable fragment of our weighted MSO logic are expressively equivalent. In Chapter 2, we generalize the notion of finite pictures to $+\omega$ -pictures, i.e., pictures which have finite number of rows and infinite number of columns. We extend conventional tiling systems with a Büchi acceptance condition in order to define the class of Büchi-tiling recognizable $+\omega$ -picture languages. The class of recognizable $+\omega$ -picture languages is indeed, a natural generalization of ω -regular languages. We show that the class of all Büchi-tiling recognizable $+\omega$ -picture languages has the similar closure properties as the class of tiling recognizable languages of finite pictures: it is closed under projection, union, and intersection, but not under complementation. While for languages of finite pictures, tiling recognizability and EMSO-definability coincide, the situation is quite different for languages of $+\omega$ -pictures. In this setting, the notion of tiling recognizability does not even cover the language of all $+\omega$ -pictures over $\Sigma = \{a, b\}$ in which the letter a occurs at least once – a picture language that can easily be defined in first-order logic. As a consequence, EMSO is too strong for being captured by

the class of tiling recognizable $+\omega$ -picture languages. On the other hand, EMSO is too weak for being captured by the class of all Büchi-tiling recognizable $+\omega$ -picture languages. To obtain a logical characterization of this class, we introduce the logic EMSO $^\infty$, which extends EMSO with existential quantification of infinite sets. Additionally, using combinatorial arguments, we show that the Büchi characterization theorem for ω -regular languages does not carry over to the Büchi-tiling recognizable $+\omega$ -picture languages. In Chapter 3, we consider the connection between weighted register automata and weighted logic on data words. Data words are sequences of pairs where the first element is taken from a finite alphabet (as in classical words) and the second element is taken from an infinite data domain. Register automata, introduced by Francez and Kaminski, provide a widely studied model for reasoning on data words. These automata can be considered as classical nondeterministic finite automata equipped with a finite set of registers which are used to store data in order to compare them with some data in the future. In this chapter, for quantitative reasoning on data words, we introduce weighted register automata over commutative data semirings equipped with a collection of binary data functions in the spirit of the classical theory of weighted automata. Whereas in the models of register automata known from the literature data are usually compared with respect to equality or a linear order, here we allow data comparison by means of an arbitrary collection of binary data relations. This approach permits easily to incorporate timed automata and weighted timed automata into our framework. Motivated by the seminal Büchi-Elgot-Trakhtenbrot theorem about the expressive equivalence of finite automata and monadic second-order (MSO) logic and by the weighted MSO logic of Droste and Gastin, we introduce weighted MSO logic on data words and give a logical characterization of weighted register automata. In Chapter 4, we study the concept of synchronizing data words in register automata. The synchronizing problem for data words asks whether there exists a data word that sends all states of the register automaton to a single state. The class of register automata that we consider here has a decidable non-emptiness problem, and the subclass of nondeterministic register automata with a single register has a decidable non-universality problem. We provide the complexity bounds of the synchronizing problem in the family of deterministic register automata with k registers (k -DRA), and in the family of nondeterministic register automata with single register (1-NRA), and in general undecidability of the problem in the family of k -NRA. To this end, we prove that, for k -DRA, inputting data words with only $2k + 1$ distinct data values, from the infinite data domain, is sufficient to synchronize. Then, we show that the synchronizing problem for k -DRA is in general PSPACE-complete, and it is in NLOGSPACE for 1-DRA. For nondeterministic register automata (NRA), we show that Ackermann(n) distinct data, where n is the number of states of the register automaton, might be necessary to synchronize. Then, by means of a construction, proving that the synchronizing problem and the non-universality problem in 1-NRA are interreducible, we show the Ackermann-completeness of the problem for 1-NRA. However, for k -NRA, in general, we prove that this problem is undecidable due to the unbounded length of synchronizing data words.