LABORATOIRE BORDELAIS DE
RECHERCHE EN INFORMATIQUE

8 September, 2017

Report on the PhD Thesis of Szczepan Hummel Entitled: Topological Complexity of Sets Defined by Automata and Formulas

The thesis studies topological complexity of various versions of automata and variants of monadic second-order logic over infinite words and infinite trees. Since Buchi and Landweber in the late 60-ties the questions of topological complexity have gained increasing importance in language theory as they allowed to get new insights and obtain new separation results. For languages of infinite words and trees, topological complexity plays a similar role to computational complexity in theory of recursive languages. For example, one can obtain a separation result between two formalisms by showing that all languages definable by one formalism belong to a particular level of Borel hierarchy, and the other formalism can define a language outside this level.

For languages of infinite words, the thesis studies the power of BS-automata. These are finite automata extended with limited form of counters on which one can perform boundedness (B) and unboundedness (S) tests. The author establishes the topological complexity of languages of BS-automata, as well as its subclasses: B-automata and S-automata. Then he considers alternating BS-automata and shows that they can define languages much higher in the hierarchy than nondeterministic BS-automata. In consequence, unlike for standard finite automata, nondeterministic and alternating BS-automata have different expressive power. These developments serve as a basis for showing that the extension of MSO with unboundedness quantifier can express sets arbitrary high in the projective hierarchy. A striking consequence of this result is that it is not possible to capture the whole MSOL by a model of deterministic or even alternating automata with a winning condition on a fixed level of the projective hierarchy.

For languages of trees, the author concentrates on unambiguous languages. These are languages accepted by non-deterministic automata with the property that on every tree they have at most one accepting run. Unambiguous languages are quite a mysterious class. For example, it is not known if it is decidable if a given regular language of infinite trees in unambiguous. Before results of the author, the shared feeling was that unambiguous languages are close to deterministic languages. The results of the thesis show that unambiguous languages are topologically much more complex than deterministic languages.

Let me now give a more detailed description of the thesis presenting and commenting some notable results. The thesis is divided into three chapters: an introductory chapter followed by one chapter with results on word languages, and one chapter with results on tree languages.

The chapter on word languages starts with a presentation of classical results on regular languages of infinite words, automata, and monadic second-order logic (MSOL). It then introduces an extension of MSOL with unbounding quantifier, MSOL+U, and gives topological complexity of two most classic examples of non-regular languages definable in this logic. This allows to show that weak variant of MSO+U is strictly weaker than unrestricted MSO+U. The important part of the chapter is devoted to showing that on every level of projective hierarchy there is an MSO+U definable language. The proof is by a very elegant reduction from multi-branching trees that are standard examples of sets hard for the levels of this hierarchy. The remainder of the chapter is devoted to BS-automata. First, the author establishes topological complexity of nondeterministic B-automata and S-automata, as well as BS-automata. Next, he introduces a model of alternating BS automata and studies its complexity. The model is natural even though due to the results on MSOL+U mentioned above, it is clear that it cannot capture the whole logic, since languages of such automata belong necessarily to the second level of the projective hierarchy. The author shows that alternating BS-automata can define languages arbitrary high in Borel hierarchy. This implies that, unlike for standard automata, alternating model is stronger than nondeterministic model, or even Boolean combinations of nondeterministic BS-automata.

The chapter on trees gives lower bounds on the topological complexity of the class of languages recognized by unambiguous parity tree automata. It starts with an introduction to these automata, and a very good overview of known complexity results, in particular results concerning topological complexity. There are many interesting subclasses of regular tree languages. One hierarchy of subclasses is defined by the size of a parity condition need to accept a language. For example, Büchi languages are close to the bottom of this hierarchy as they need automata with (0,1) conditions. Different subclasses are obtained by considering automata's behavior: deterministic, nondeterministic, or alternating. Two more subtle classes are central for this chapter: unambiguous and bi-unambiguous languages. An automaton in unambiguous if on every tree it has at most one accepting run. A language is unambiguous if it is the language of an unambiguous automaton. A language is bi-unambiguous if it and its complement are unambiguous. These two classes have quite good closure properties, but important questions about them are still open. For example, we do not know if it is decidable whether a given regular language is unambiguous (or bi-unambiguous).

Probably one of the most useful results of the thesis is an example of an unambiguous language that is analytic-complete. The language is simple and even bi-unambiguous, since its complement is recognized by a deterministic automaton. Moreover, this language can be recognized by a Büchi automaton, but not by an unambiguous Büchi automaton. This way, a single language refutes many possible conjectures. The main part of the chapter is devoted to an operation producing from a bi-ambiguous language a topologically more complex one. For this to work the language should satisfy additionally an elegant condition called stretchability. This operation is then accompanied with limit operation, together they allow to obtain a sequence of length ω^2 of strictly more complex (in the sense of Wedge reducibility) bi-unambiguous languages. This result indicates that bi-unambiguous languages may have quite a rich structure from the topological point of view.

The organization of the thesis is very clear, the flow of arguments is logical, it is easy to find necessary definitions and lemmas. Mathematical writing is very good, formulations are precise, to the point, and readable at the same time. The thesis can very well serve as a starting point for somebody interested in topological complexity of regular languages of infinite words and trees.

In summary, the thesis brings new, very useful, results on topological complexity of languages of infinite words and trees. For infinite words, the results show a clear difference between automata models with boundedness/unboundedness tests and monadic second-order logic extended with unboundedness quantifier. For infinite trees, the thesis provides a very much needed results on unambiguous languages, confirming the richness and interest of this class.

In my opinion, the thesis satisfies all the requirements, and I declare that it merits to be accepted.

I. Walukiewicz

Igor Walukiewicz
Directeur de rechereche CNRS