Improved Ramsey-Based Büchi Complementation

Stefan Breuers, Christof Löding, and Jörg Olschewski

RWTH Aachen University

Abstract. We consider complementing Büchi automata by applying the Ramsey-based approach, which is the original approach already used by Büchi and later improved by Sistla et al. We present several heuristics to reduce the state space of the resulting complement automaton and provide experimental data that shows that our improved construction can compete (in terms of finished complementation tasks) also in practice with alternative constructions like rank-based complementation. Furthermore, we show how our techniques can be used to improve the Ramsey-based complementation such that the asymptotic upper bound for the resulting complement automaton is $2^{\mathcal{O}(n \log n)}$ instead of $2^{\mathcal{O}(n^2)}$.

1 Introduction

The aim of this paper is to present several techniques to improve the Ramsey-based approach to complementation of Büchi automata, which was originally used by Büchi when he introduced this model of automata on infinite words to show the decidability of monadic second order logic over the successor structure of the natural numbers [2]. The method is called Ramsey-based because its correctness relies on a combinatorial result by Ramsey [10] to obtain a periodic decomposition of the possible behaviors of a Büchi automaton on an infinite word. Starting from a Büchi automaton with n states the construction yields a complement automaton with $2^{\mathcal{O}(n^2)}$ states [13]. A non-trivial lower bound of $n!$ for the complementation of Büchi automata was shown by Michel in [8]. The gap between the lower and the upper bound was made tighter by a determinization construction presented by Safra [11] from which a complementation construction with upper bound $2^{\mathcal{O}(n \log n)}$ could be derived. An elegant and simple complementation construction achieving this upper bound was presented by Klarlund in [7] using progress measures, now also referred to as ranking functions. Based on Klarlund's construction the gap between upper and lower bound has been tightened by Friedgut, Kupferman and Vardi [5] and later even further by Schewe [12], leaving only a polynomial gap compared to the improved lower bound that was obtained by Yan [16]. Besides the optimizations of the rank-based approach, a different construction has been developed by Kähler and Wilke in [6], usually called slice-based complementation, and the determinization construction of Safra has been optimized by Piterman [9].

Given all these different constructions, there has recently been an increased interest in experimental evaluations of the different complementation methods.

L. Birkedal (Ed.): FOSSACS 2012, LNCS 7213, pp. 150–164, 2012.

In experimental studies conducted by Tsai et al. [15], it turned out that the Ramsey-based approach was not only inferior to the more modern approaches (determinization-based, rank-based, and slice-based) in terms of the size of the resulting complement automata, but that in most cases its implementation did not even terminate within the imposed time and memory bounds.

Though performing inadequately for complementation, the Ramsey-based approach is used in two other fields, namely universality checking and inclusion checking [4,1]. For these two purposes specific optimization techniques have been developed.

Since the Ramsey-based approach to complementation has got an appealingly simple structure and is nice to teach, our aim is to improve it on both the practical level by using heuristics to reduce the size of the complement automata, and the theoretical level by using the ideas underlying our heuristics to obtain a general optimization of the method that meets the upper bound of $2^{\mathcal{O}(n \log n)}$.

We have implemented these heuristics [17] and conduct experiments on a large set of example automata, showing that the improved construction can compete with other methods for complementation.

The remainder of the paper is structured as follows. In Section 2 we start with some basic definitions and in Section 3 we repeat the original Ramsey-based complementation method. In Section 4 we present our heuristics that are used in the implementation, and we discuss our experimental results. Finally, in Section 5 we show how the ideas from Section 4 can be used to obtain an improvement of the Ramsey-based construction also on a theoretical level.

2 Preliminaries

The set of infinite words over an alphabet Σ is denoted by Σ^ω. For an infinite word $\alpha = a_0 a_1 \cdots$, with $\alpha[i]$ we denote the letter a_i at position i.

An *automaton* is a tuple $\mathcal{A} = \langle Q, \Sigma, q_0, \delta, F \rangle$ where Q is a finite set of states, Σ is a finite alphabet, $q_0 \in Q$ is the initial state, $\delta \colon Q \times \Sigma \to 2^Q$ is the transition function and $F \subseteq Q$ is the set of accepting states. The transition function δ can be extended to a function $\delta^* \colon 2^Q \times \Sigma^* \to 2^Q$ on subsets of the state set and on words in the natural way. By RSC denote the set of all subsets of states that are reachable by the subset construction, i.e. RSC $:= \{P \subseteq Q \mid \exists w \in \Sigma^* \colon \delta^*(\{q_0\}, w) = P\}$. $\mathcal{A}$ is called *deterministic* if $|\delta(q,a)| \leq 1$ for all $q \in Q$ and $a \in \Sigma$.

For automata we consider two different semantics: the usual NFA semantics, in which the automaton accepts a language of finite words, and the Büchi semantics, in which the automaton accepts a language of infinite words. This is made precise in the following definitions.

A *run* of $\mathcal{A}$ on a word $\alpha \in \Sigma^\omega$ is an infinite sequence of states $\rho = p_0, p_1, \ldots$ such that $p_0 = q_0$, and $p_{i+1} \in \delta(p_i, \alpha[i])$ for all $i \geq 0$. A run ρ is *Büchi-accepting* if $\rho(i) \in F$ for infinitely many $i \geq 0$. A *path* from p to q in $\mathcal{A}$ of a word $u = a_1 \cdots a_n \in \Sigma^*$ is a finite sequence of states $p_0, \ldots, p_n$ such that $p = p_0$, $q = p_n$, and $p_i \in \delta(p_{i-1}, a_i)$ for all $1 \leq i \leq n$. The path is *NFA-accepting* if $p_0 = q_0$ and $p_n \in F$.