On the class of Church-Rosser languages

Gundula Niemann and **Friedrich Otto**

Fachbereich Mathematik/Informatik, Universität Kassel, D–34109 Kassel

{*niemann,otto*}*@theory.informatik.uni-kassel.de*

September 9, 1997

For a finite, length-reducing, and confluent string-rewriting system the word problem can be solved in linear time by the so-called normal form algorithm [Boo82]. Motivated by this fact McNaughton et al [MNO88] introduced the *Church-Rosser languages.* A Church-Rosser language $L \subset \Sigma^*$ is given through a finite, length-reducing, and confluent string-rewriting system R on some alphabet Γ properly containing Σ, two irreducible strings $t_1, t_2 \in (\Gamma \smallsetminus \Sigma)^*$, and an irreducible letter $Y \in \Gamma \smallsetminus \Sigma$ satisfying the following condition for all strings $w \in \Sigma^*$:

$$w \in L \text{ if and only if } t_1 w t_2 \rightarrow_R^* Y.$$

Hence, the membership problem for a Church-Rosser language is decidable in linear time. It follows immediately that the class CRL of Church-Rosser languages is contained in the class CSL of context-sensitive languages.

It remained open, however, whether the class CRL is closed under complement. Accordingly, McNaughton et al introduced the class of *Church-Rosser decidable languages* CRDL. This subclass of the class CRL contains the class DCFL of deterministic context-free languages, and it is closed under complement. Also it remained open at the time whether or not every context-free language is a Church-Rosser language, although it was conjectured that not even the linear language $L_0 := \{ww^\sim | w \in \{a,b\}^*\}$ is a Church-Rosser language. Here $w^\sim$ denotes the reversal of the string w.

These questions remained open until another, seemingly unrelated development had taken place. Dahlhaus and Warmuth [DaWa86] considered the class GCSL of *growing context-sensitive languages.* These languages are generated by context-sensitive grammars each production rule of which is strictly length-increasing. They proved that these languages have membership problems that are decidable in polynomial time. Although it might appear from the definition that GCSL is not an interesting class of languages, Buntrock and Loryś showed that GCSL is an *abstract family of languages* [BuLo92]. Exploiting these closure properties they characterized the class GCSL through various other classes of grammars that are less restricted [BuLo92, BuLo94].

Using these grammars Buntrock and Otto [BuOt95] obtained a characterization of the class GCSL by a nondeterministic machine model, the so-called *shrinking pushdown automaton with two pushdown stores* (sTPDA). The input for such a machine is provided as the intial contents of one of the pushdown stores, and it accepts either by final state or (equivalently) by empty pushdown stores. A positive weight is assigned to each tape symbol and each internal state symbol of the machine. By adding up the weights this gives a weight for each configuration. Now it is required that the weight of the actual configuration decreases with each step of the machine. It is with respect to these weights that the two-pushdown automaton is called *shrinking*.

Since the sTPDA is a nondeterministic device, it was only natural to consider the class of languages that are accepted by the deterministic variant of it. As it turned out the deterministic sTPDA accept exactly the so-called *generalized Church-Rosser languages*, which are obtained from the Church-Rosser languages by admitting finite, *weight-reducing*, and confluent string-rewriting systems in the definition [BuOt95]. Thus, the class GCRL of generalized Church-Rosser languages coincides with the class of 'deterministic growing context-sensitive languages.' In particular, it follows that this class is closed under complement. Further, Buntrock and Otto concluded from this result that the language classes CFL and GCRL, and therewith the classes CFL and CRL, are indeed incomparable under set inclusion. Finally, this yields the following chain of inclusions:

$$\mathsf{DCFL} \subset \mathsf{CRDL} \subseteq \mathsf{CRL} \subseteq \mathsf{GCRL} \subset \mathsf{GCSL} \subset \mathsf{CSL},$$

where it was left open whether or not the two inclusions $\mathsf{CRDL} \subseteq \mathsf{CRL} \subseteq \mathsf{GCRL}$ are proper or not.

Here we show that none of these two inclusions is actually a proper one, that is, the three language classes CRDL, CRL, and GCRL all coincide. Our proof makes use of the above-mentioned characterization of the generalized Church-Rosser languages through the deterministic sTPDA. We will prove that each language that is accepted by some deterministic sTPDA is actually a Church-Rosser decidable language. Hence, $\mathsf{GCRL} \subseteq \mathsf{CRDL}$ implying that the three classes above actually coincide. Hence, the class of Church-Rosser languages can be interpreted as the class of deterministic growing context-sensitive languages.

References

[Boo82] R.V. Book. Confluent and other types of Thue systems. *J. Association Computing Machinery*, 29:171–182, 1982.

[BuLo92] G. Buntrock and K. Loryś. On growing context-sensitive languages. In W. Kuich, editor, *Proc. of 19th ICALP*, Lecture Notes in Computer Science 623, pages 77–88. Springer-Verlag, Berlin, 1992.

[BuLo94] G. Buntrock and K. Loryś. The variable membership problem: Succinctness versus complexity. In P. Enjalbert, E.W. Mayr, and K.W. Wagner, editors, *Proc. of STACS 94*, Lecture Notes in Computer Science 775, pages 595–606. Springer-Verlag, Berlin, 1994.

[BuOt95] G. Buntrock and F. Otto. Growing context-sensitive languages and Church-Rosser languages. In E.W. Mayr and C. Puech, editors, *Proc. of STACS 95*, Lecture Notes in Computer Science 900, pages 313–324. Springer-Verlag, Berlin, 1995.

[DaWa86] E. Dahlhaus and M. Warmuth. Membership for growing context-sensitive grammars is polynomial. *J. Computer System Sciences*, 33:456–472, 1986.

[MNO88] R. McNaughton, P. Narendran, and F. Otto. Church-Rosser Thue systems and formal languages. *Journal Association Computing Machinery*, 35:324–344, 1988.