New Techniques for Proving the Decidability of Equivalence Problems *

Karel Culik II
Department of Computer Science
University of South Carolina
Columbia, S. C. 29208

Abstract. We will discuss several recently developed techniques for proving the decidability of the equivalence problem for devices defining languages and relations. Most of these techniques come from the development originated in the proof of the decidability of the DOL equivalence problem. Of particular importance is the recently shown validity of the Ehrenfeucht conjecture and its effective variants.

1 Introduction

We will discuss three recently developed techniques for proving the decidability of equivalence problems. We are mainly interested in the equivalence problems for devices defining languages or translations (relations). The literature on this subject is extensive. A selection of results on language equivalence as well as relevant auxiliary results are listed chronologically in Table 1, similarly for relational equivalence in Table 2.

At the previous ICALP meeting held in Finland the author presented a solution of the DOL equivalence problem obtained with I. Fris in [11]. The decidability proof was complex, however, it introduced new techniques which, together with newly encountered problems, turned out to be very fruitful. Indeed, our first new technique presented in Section 3 evolved from a new proof of the DOL equivalence problem (see [13]). This technique is based on the solution of the important Ehrenfeucht conjecture, (see [4,26]) that also originated in work on the DOL problem. The solution of the DOL problem involves the testing of equivalence of two morphisms on every string of the language generated by one of them. This leads to the problem of testing the morphic equivalence on a language, introduced in [19], and further to the investigation of different kinds of equivalence problems for mappings restricted to certain domains, see e.g. [2,37] and [44]. These problems become especially fruitful when the positive solution of Ehrenfeucht conjecture made possible to use our result from [12], stating that the existence of a test set for each HDTOL language implies its effective existence. The new proof of the DOL problem and its generalizations are based on this result. Further development involved the consideration of test sets with respect to certain classes of finite transducers rather than just morphisms. This lead to the proof of the decidability of the equivalence problem for single-valued two-way transducers (possibly restricted to HDTOL languages defined by nonerasing morphisms) [15]. Even without the restriction on the domains this result is an extension of the decidability of the equivalence problem for deterministic two-way transducers, a "classical" problem that was open for many years and finally solved by Gurari [27]. However, our result also generalizes the problem of morphic equivalence on HDOL and on DTOL languages, two

*Reported research was supported by the National Sciences Foundation under Grant No. CCR-8702752 and by the Natural Sciences and Engineering Research Council of Canada under Grant No. A-7403
examples of longstanding open problems where, in contrast to the above, the mappings are simple but the restricted domains are complex.

In [16] we further extended the above result; we have shown that it is decidable whether two finite valued transducers agree on an HDTOL language. This implies the decidability of the equivalence problem for finite valued transducers and demonstrates that unbounded ambiguity with respect to outputs is the crucial property that makes the equivalence problem for finite transducers undecidable. Indeed, it has been shown in [30] that without the assumption of finite valuedness, the problem is undecidable even for a unary input (or output) alphabet [30]. The finite valuedness result, the details of which are given in Section 3, can be extended to two-way transducers, provided that the HDTOL language is generated by nonerasing morphisms.

The second technique is based on synchronizability (i.e., indirect synchronization) of two equivalent devices. Devices like deterministic pushdown automata (DPDAs) or multitape deterministic finite automata (MDFAs) are difficult to test for equivalence since they might be equivalent but not synchronized as far as the movement of the stacks is concerned in the former and the reading from different input tapes in the latter. The main idea is to prove that for any two equivalent devices $P$ and $Q$ (from certain class) there exists a chain of devices $P = M_0, M_1, \ldots, M_n = Q$, such that $M_i$ and $M_{i+1}$ are equivalent and synchronized for $i = 0, 1, \ldots, n - 1$. That is, in order to test their equivalence, $P$ and $Q$ do not need to be synchronized directly, but rather in several steps. This motivated the following development in [17]. First, we have defined, in quite a general way, the notion of synchronized DPDAs and have shown that two synchronized DPDAs can be tested for equivalence. Further, if there are DPDAs $P = M_0, M_1, \ldots, M_n = Q$, such that $M_i$ and $M_{i+1}$ are synchronized, for $i = 0, 1, \ldots, n - 1$, then we call $P$ and $Q$ synchronous. We have shown that the equivalence of two synchronized DPDAs can be tested. Consequently, the equivalence problem is decidable for every class of DPDAs for which equivalence implies synchronizability. Although it is probably difficult to show that every two equivalent DPDAs are synchronizable even for a subclass of DPDAs we do conjecture that it is true for all DPDAs.

We say that two deterministic pushdown transducers are synchronized if their underlying DPDAs are synchronized. In [17] it has been shown that the equivalence problem for synchronized deterministic pushdown transducers is decidable by reducing it to the morphic equivalence for context-free languages [2]. An alternate proof of a slightly weaker result is in [31].

Our third new, yet unpublished technique is based on one particular result mentioned above, namely, the decidability of the morphic equivalence for DTOL languages (or sequences). It turns out that frequently computations of two equivalent but not synchronized devices, e.g. n-tape deterministic finite automata, can be "HDTOL-matched", which means that there is a HDTOL language $L$ each string of which encodes a pair of matching computations. Consequently, the testing of equivalence of the two given devices can be reduced to testing whether two morphisms agree on language $L$. Note, that we are using techniques originated in the DOL problem and some other results on L-systems, in attacking difficult decidability problems in "classical" automata theory. For example the equivalence problem for n-tape DFAs was explicitly stated already in 1968 ([23]). Its decidability for two tapes has been shown in [7], for single cycles in [41]. The corresponding inclusion problem has been shown undecidable in [23] and decidable for single cycles in [38].

Actually, synchronization of MDFAs (or DPDAs) can be defined as the existence of an HDTOL-matching, hence the last two techniques could be possibly combined.

After some preliminary definitions we discuss our new techniques for testing of equivalence in Sections 3 — 5. Note that the other topics motivated by the DOL problem mentioned above were thoroughly discussed in last year invited ICALP-lecture by J.Karhumäki [35] and also in his recent column on Formal languages in EATCS Bulletin [36].