The equivalence of finite valued transducers
(on HDTOL languages) is decidable

Karel Culik II
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada
N2L 3G1

Juhani Karhumäki 
Department of Mathematics
University of Turku
Turku, Finland

Short title: Finite valued transducers.

ABSTRACT

We show a generalization of the Ehrenfeucht Conjecture: For every language there
exists a (finite) test set with respect to normalized k-valued finite transducers with bounded
number of states. Further, we show that for each HDTOL language such a test set can be
found effectively. As a corollary we solve an open problem by Gurari and Ibarra: The
equivalence problem for finite valued finite transducers is decidable. This is the first time the
equivalence problem is shown to be decidable for a larger class of multivalued transducers.

1. Introduction

Let $L$ be an arbitrary language over $\Sigma$. We say that a finite subset $F$ of $L$ is a test set (with respect to
morphisms) for $L$ if whenever two morphisms agree on $F$ they agree on $L$ as well. The Ehrenfeucht conjecture, cf. [CS2] and [K], states that each language possesses a test set. It has been shown valid recently byAlbert and Lawrence [AL1], and independently by Guba [Gub]. In [McN,P,Sa,St] several variations of the latter
proof have been given, while in [AL2] the conjecture has been generalized for $k$-bounded substitutions, i.e., for
substitutions satisfying that the cardinalities of the images of letters are bounded by a fixed constant $k$.

We show here that the Ehrenfeucht conjecture holds also for much more general types of mappings,
namely for those defined by normalized finite transducers with bounded number of states and bounded degree
of nondeterminism on inputs of length one. (The term “normalized” refers to the fact that the transducer reads in
a single step either the empty word or a letter.) In particular, it follows that the conjecture holds for normalized
$k$-valued transducers with bounded number of states. Moreover, we show that for this family a test set can be
effectively found for each HDTOL language. Our main motivation is a corollary of this result, the decidability of
the equivalence problem for $k$-valued finite transducers. In order to put our theorem into perspective we
next list the history of main results on the equivalence of finite transducers.

* This work was supported by the Natural Sciences and Engineering Research Council of Canada under Grant
A-7403.
† This work has been done during the second author’s visit at the University of Waterloo.
Equivalence problems for various types of finite transducers (finite automata with outputs) have been extensively studied since the beginning of automata theory. In his famous paper [Mo] Moore showed that the equivalence problem is decidable for length preserving deterministic sequential machines. This decidability result was then gradually extended as follows: For deterministic gem's Jones and Laaser [JL], cf. also Blattner and Head [BH2], for single-valued transducers Schützenberger [Sc] and independently Blattner and Head [BH1], for deterministic two-way transducers Gurari [Gur1] and [Gur2], and finally for single-valued two-way transducers the authors [CK3]. A strictly larger class than that of deterministic gem's, but incomparable with the other classes above, is the class of deterministic two-tape acceptors, for which the decidability of the equivalence problem was proved by Bird [Bi].

On the other hand, Fischer and Rosenberg [FR] showed the undecidability of the equivalence problem for nondeterministic finite transducers. At the same time Griffiths [Gr] generalized this undecidability for c-free nondeterministic gem's, and finally Ibarra [I1] proved it for c-free gem's with unary input (or output) alphabet. Here, we further narrow the gap between the decidable and undecidable equivalence problems for finite transducers. We show that the problem is decidable for finite valued finite transducers. We give a detailed proof for the case of one-way finite transducers, but using the techniques from our previous paper [CK3] the result can be straightforwardly extended to finite valued two-way finite transducers, too.

Actually, we prove a considerably stronger result, namely that for every HDTOL language and two natural members n and k there effectively exists a test set with respect to normalized k-valued finite transducers with at most n states. Clearly, this result implies that given an HDTOL language L and two finite valued transducers we can test whether they are equivalent on L.

The decidability of testing the equivalence of mappings of certain types on languages from a family L has been considered by many authors. Most relevant results from the point of view of this paper can be listed as follows.

Testing the equivalence of mappings which are realized by finite transducers on a regular language is a special case of the equivalence problem for finite transducers, since a restriction of a finite transduction to a regular set is again a finite transduction. In [KK] it was shown that this problem is decidable for (multivalued) mappings of the form "morphism followed by inverse morphism", while it is undecidable for the mappings of the form "inverse morphism followed by morphism". This latter undecidability result was generalized in [Ma] for inverses of finite substitutions. A lot of attention was given to the problem of deciding the equivalence of two morphisms on a given language. Indeed, in order to prove the decidability of the DOL sequence equivalence problem in [CF] it was shown that the equivalence of two morphisms can be tested on the DOL language generated by one of them. Subsequently the following cases were shown to be decidable. Morphisms on a context-free language in [CS1], morphisms on an HDTOL language in [CK2], and as generalizations of [CS1] and [CK2] single-valued finite transducers on a context-free language in [C], and single-valued two-way finite transducers on an HDTOL language in [CK3]. For any family L of languages which is effectively closed under inverse morphisms and intersections with regular sets, and such that languages in L have effectively constructible semilinear Parikh maps, another generalization of [CS1] was shown in [12]; It is decidable whether two deterministic two-way finite transducers are equivalent on a language from L.

Some of the above decidability results are based, or can be based, on the effective existence of test sets for some families of languages, which is known to hold in the following cases: For regular languages [CS2], for context-free languages [ACK], for supports of k-rational formal power series, with k a field, [RR], for DOL languages [Ru], and for HDTOL languages [CK2]. The last result in this list was generalized in [CK3], where it was shown that each HDTOL language possesses a test set even with respect to single-valued two-way (sequential) transducers with bounded number of states. We shall see here that this result can be further extended to the corresponding family of k-valued transducers, but that it does not hold in general, or even for gem's with bounded number of states.

When considering a restricted class of transducers it is certainly important whether this class is effectively given, that is whether we can test the membership in the class for an arbitrary transducer. To test whether a given finite transducer is a (deterministic) gem is trivial. The decidability of single-valuedness for one-way finite transducers was shown in [Sc], and for two-way finite transducers in [CK9]. It should be noted that in both these cases the decidability of the single-valuedness actually easily implies the decidability of the equivalence problem. In [GI] it has been shown decidable whether a given finite transducer is k-valued for a given k. As will be outlined in Section 6, our earlier result to decide the single-valuedness for a two-way finite transducer can be extended to decide the k-valuedness as well.

The rest of this paper is organized as follows. In Section 2 we review some basic notions in order to fix our terminology and establish an important generalization of the Ehrenfeucht conjecture in terms of systems of equations. In the next section we introduce transducer schemata as an auxiliary device in proving the existence of a test set with respect to a large family of transducers. It is shown that such a test set exists (noneffectively) for each language. In Section 4 we show that for each HDTOL language a test set with respect to normalized k-valued transducers with bounded number of states exists effectively. Since regular sets are a subfamily...