Decision problems for parameterized top-down tree transducers

T. Tang*

This paper presents the results of a study on the complexity of certain decision problems—membership, equivalence, unification, and functionality—for top-down tree transducers (TDTs) with output parameters (PTDTs). Membership problem is L-complete for deterministic models DTDTs and PDTDTs, belongs to LOGCFL for TDTs, and it is NP-complete for PTDTs. Equivalence checking of complete DTDTs and PDTDTs can be performed in cubic time. The same time $O(n^3)$ is required for computing the most general unifier of a pair of PDTDTs as well as for computing the functional instance of a PTDT. Minimization of complete PDTDTs is NP-complete in contrast with polynomial-time complexity of minimization problem for DTDTs.

Keywords: tree transducers, decision problems, output parameters, computational complexity.

1. Introduction

Finite trees are hierarchical data structures, examples of which include abstract syntax trees or structured documents. Top-down tree transducers model the data-processing workflow: they construct an output tree by traversing an input tree, starting from its root. At the design stage, the transformation rules may be known, while the output labels—such as document tags or operations for a target platform—remain undefined; these must be specified at a later stage or during the system's installation by the user. Thus, this approach gives rise to two main problems: 1) how to verify whether a given system template satisfies the specified requirements for any parameter values, and 2) how to select parameter values such that the system meets the given requirements. Corresponding decision problems have been studied for finite-state parameterized word transducers [1]. In this paper, the same problems are investigated for top-down parameterized tree transducers.

* *Tang Tianxiang* — PhD student, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Shenzhen MSU-BIT University, e-mail: tangtx@cs.msu.ru, ORCID: 0009-0009-6040-8108

Published under CC-BY 4.0

2. Top-down tree transducers

For a ranked alphabet Γ, the set $\mathcal{T}(\Gamma)$ of finite ground terms is generated by

$$t ::= c \mid f(t_1, \ldots, t_m), \qquad c \in \Gamma^{(0)}, \quad f \in \Gamma^{(m)}$$

where $\Gamma^{(m)}$ denotes the symbols of arity m. Each term represents a rooted ordered tree: $f(t_1, \ldots, t_m)$ has a root labelled by f and ordered subtrees represented by $t_1, \ldots, t_m$, while a nullary symbol c represents a leaf.

A *parameterized top-down tree transducer* (PTDT) is a tuple $\Pi = (\Sigma, \Delta, \mathcal{P}, \mathcal{S}, S_0, R)$, where Σ and Δ are finite ranked input and output alphabets containing nullary symbols, $\mathcal{P}$ is a finite ranked alphabet of parameters disjoint from $\Sigma \cup \Delta$, $\mathcal{S}$ is a finite set of states, $S_0 \in \mathcal{S}$ is the initial state, and R is a finite set of rules

$$S(f(x_1, \ldots, x_k)) \longrightarrow \beta, \qquad S \in \mathcal{S}, \quad f \in \Sigma^{(k)}$$

Here $x_1, \ldots, x_k$ are distinct placeholders for the immediate input subtrees. The right-hand side is generated by

$$\beta ::= Q(x_i) \mid a(\beta_1, \ldots, \beta_m)$$

where $Q \in \mathcal{S}$, $1 \le i \le k$, $a \in (\Delta \cup \mathcal{P})^{(m)}$, and $m \ge 0$; for $m = 0$, the second alternative is simply a. A call $Q(x_i)$ recursively processes the i-th input subtree in state Q. Input variables occur only in such calls and may be repeated or omitted, allowing copying and deletion.

An ordinary *top-down tree transducer* (TDT), denoted by π, is the case $\mathcal{P} = \emptyset$. A transducer is *deterministic* if it has at most one rule per pair (S, f), and *complete* if it has at least one. The deterministic models are abbreviated DTDT and PDTDT.

Parameter substitutions are arity-preserving maps $\eta\colon \mathcal{P} \to \mathcal{P} \cup \Delta$; they are *ground* if $\eta(\mathcal{P}) \subseteq \Delta$. Their sets are denoted by $\mathrm{Subst}(\mathcal{P}, \Delta)$ and $\mathrm{GSubst}(\mathcal{P}, \Delta)$, respectively. We assume $\Delta^{(m)} \neq \emptyset$ whenever $\mathcal{P}^{(m)} \neq \emptyset$. The instance $\Pi\eta$ is obtained by simultaneously replacing every occurrence of each parameter p in all rules by $\eta(p)$, retaining its arguments. Ground instances are ordinary TDTs. Different parameters may share a value; parameter values remain fixed throughout a computation.

For an ordinary TDT π, a rewriting step, denoted by $\longrightarrow_\pi$, replaces any occurrence of $S(f(t_1, \ldots, t_k))$ by $\beta\{x_1/t_1, \ldots, x_k/t_k\}$, where $t_i \in \mathcal{T}(\Sigma)$. Rewriting may take place below output symbols; a successful computation ends with a ground output term containing no state calls. The transduction relation is defined by

$$\mathrm{TR}(\pi, S) = \{(t, T) \in \mathcal{T}(\Sigma) \times \mathcal{T}(\Delta) \mid S(t) \longrightarrow_\pi^* T\}$$

A DTDT defines a partial function $[\![S]\!]_\pi$ with $[\![S]\!]_\pi(t) = T$ if and only if $(t, T) \in \mathrm{TR}(\pi, S)$. This function is total for a complete DTDT.

3. Membership problem

The membership problem consists of determining, for a given TDT π and ground terms t and T (represented explicitly as prefix strings), whether the condition $(t, T) \in \mathrm{TR}(\pi, S_0)$ holds. In the case where Π is a parameterized transducer, the problem is to determine whether there exists a substitution $\eta \in \mathrm{GSubst}(\mathcal{P}, \Delta)$ such that $(t, T) \in \mathrm{TR}(\Pi\eta, S_0)$

Theorem 1. *Membership problem is* L*-complete for DTDT and PDTDT, it is in* LOGCFL *for TDT, and it is* NP*-complete for PTDT.*

4. Equivalence and unification problems

Equivalence of two states of a complete TDT means that $\mathrm{TR}(\pi, S_1) = \mathrm{TR}(\pi, S_2)$. Equivalence problem consists of determining, given a TDT π and a pair of its states S_1, S_2 whether these states are equivalent. Since this problem is undecidable for TDTs in general [4], we consider it only for complete DTDTs. Unification problem for a pair of PTDTs Π_1, Π_2 requires to compute such a substitution (unifier) $\theta \in \mathrm{Subst}(\mathcal{P}, \Delta)$ that TDTs $\Pi_1\theta\eta$ and $\Pi_2\theta\eta$ are equivalent for every $\eta \in \mathrm{GSubst}(\mathcal{P}, \Delta)$. A unifier θ is called the most general one if every ground unifier θ' is of the form $\theta\rho$, $\rho \in \mathrm{GSubst}(\mathcal{P}, \Delta)$.

Theorem 2. *Equivalence problem for complete DTDTs can be decided in $O(n^3)$ time, where n is the total size of the transducer descriptions.*

Theorem 3. *There exists an $O(n^3)$-time algorithm which for any pair of complete PDTDTs either computes their most general unifier, or reports that these PDTDTs are not unifiable.*

The approach of [2] is refined to enable its application to the functionality problem.

5. Functionality problem

A TDT π is functional if the relation $\mathrm{TR}(\pi, S_0)$ is function. Functionality problem consists of determining, given a TDT π whether it is functional.

Theorem 4. *Functionality problem for complete TDTs can be decided in $O(n^3)$ time.*

Theorem 5. *There exists an $O(n^3)$-time algorithm which for any complete PTDT Π either computes a ground substitution η such that the instance $\Pi\eta$ is functional, or reports that all ground instances of Π are not functional.*

6. Minimization

In [3] it was shown that minimization of complete DTDTs can be performed in polynomial time. Minimization problem for PDTDTs is set up as follows. Given a complete PDTDT Π and integer k, one needs to determine if there exists $\eta \in \mathrm{GSubst}(\mathcal{P}, \Delta)$ such that the instance $\Pi\eta$ is equivalent to a DTDT M with $|\mathcal{S}_M| \leq k$.

Theorem 6. *Minimization problem for complete PDTDT is* NP*-complete.*

7. Projection problem

Projection problem is that of checking, given a TDT π and a ground term T, if there exists such a ground term t that $(t, T) \in \mathrm{TR}(\pi, S_0)$.

Theorem 7. *Projection problem for DTDTs is* PSPACE*-complete.*

References

[1] T. Tang, V. A. Zakharov, "On the complexity of decision problems for parameterized finite state synchronous transducers", *Implementation and Application of Automata (CIAA 2024)*, LNCS, **15015**, Springer, Cham, 2024, 332–346. DOI: 10.1007/978-3-031-71112-1_24.

[2] Z. Deng, T. Tang, V. A. Zakharov, "An algebraic approach to the equivalence checking of deterministic top-down tree transducers", *Implementation and Application of Automata (CIAA 2025)*, LNCS, **15981**, Springer, Cham, 2026, 86–98. DOI: 10.1007/978-3-032-02602-6_7.

[3] J. Engelfriet, S. Maneth, H. Seidl, "Deciding equivalence of top-down XML transformations in polynomial time", *Journal of Computer and System Sciences*, **75**:5 (2009), 271–286. DOI: 10.1016/j.jcss.2009.01.001.

[4] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, C. Löding, S. Tison, M. Tommasi, *Tree Automata Techniques and Applications*, INRIA, 2008.

Received on October 20, 2026