MERGING STATES IN DETERMINISTIC FUZZY FINITE TREE AUTOMATA BASED ON FUZZY SIMILARITY MEASURES

Somaye Moghari

Department of Mathematical Sciences
University of Shahrood
Shahrood
Iran
e-mail: s.moghari@shahroodut.ac.ir

Mohammad Mehdi Zahedi

Faculty of Mathematics and Computer
Shahid Bahonar University
Kerman
Iran
e-mail: zahedi_mm@modares.ac.ir

Reza Ameri

DepartmentSchool of Mathematics
Statistics and Computer Science
University of Tehran
Tehran
Iran
e-mail: ameri@umz.ac.ir

Abstract. This paper presents a contribution to the problem of measuring fuzzy similarity of states and merging them in a Deterministic Fuzzy Finite Tree Automaton (DFFTA). The main question is: how to merge some states of a complete and reduced DFFTA such that the languages of original automaton and minimized one be similar but not necessarily equal? In order to solving this problem, we generalize the concepts of distance and similarity measures between fuzzy sets to distance and similarity measures between states of DFFTA. Then, we define the notions of normal DFFTA and introduce an efficient algorithm (polynomial order of time complexity) for discovering similar state sets of a DFFTA.

Keywords: deterministic fuzzy tree automata, state reduction, fuzzy similarity measure.

1. Introduction

Fuzzy sets were introduced by Zadeh [29] as an extension of the classical notion of set. This theory can be used in a wide range of domains. One such domain is fuzzy automata theory first introduced by Wee [26]. A fuzzy automaton is a device which accepts a fuzzy set of words called fuzzy language. Automata have a long history both in theory and application [3], [2], [20], [19], [9], [1].

Finite Tree Automata (FTA) was introduced by Doner [7], [8] and Thatcher and Wright [24], [23]. Their goal was to prove the decidability of the weak second order theory of multiple successors. An FTA accepts a set of trees called recognizable tree language [5]. Trees appears in mathematics, computer science and other areas as formal terms, algebraic expressions, parse and derivation trees, computation trees, and generally as representations of hierarchically organized structures. A fuzzy set of trees, called recognizable fuzzy tree language if can be accepted by some Fuzzy FTA (FFTA) [18], [4], [10], [19]. The membership grade of each tree in language of FFTA is known as behavior of automaton. When the membership grade of trees takes values in a lattice (rather than in the unit interval of real numbers), the language called L-fuzzy tree language [12], [14], [13].

From a practical perspective, it is important that the considered automata are as small as possible (minimal). As well, some decision problems such as equivalence and intersection non-emptiness are closely related to minimization problem [5]. Current studies on minimizing DFFTA and deterministic weighted tree automata, focus on the two main strategies; language preserving minimization [12], [17], [15], [16], [20] and behavior preserving minimization [17].

Similarity plays an essential role in taxonomy, recognition, case based reasoning and many other fields [27], [25], [11]. We use the concept of similarity or approximate equality [22], [28] modeled on classes of fuzzy sets to define similarity between some states of DFFTA. We use this idea for similarity based merging states (state reduction) of DFFTA. We show how to find set of states similar to a given state and prove that merging a set of states which are pair wise similar, makes a DFFTA which its language is similar to original DFFTA.

In addition, we introduce the problem of similarity based state merging in DFFTA and show that this problem is not partitioning states by an equivalence relation, and it is different from language preserving and behavior preserving minimizing DFFTA.

This paper is organized as follows. Section 2 presents some mathematical preliminaries about L-Fuzzy sets and fuzzy finite tree automata. In Section 3, firstly, the concept of similarity and distance between L-fuzzy sets and its generalization to states of automata are presented. Then, we define the concept of normal DFFTA and introduce DFFTA normalizing algorithm. Also, we show how to find the set of similar states in a normal DFFTA.

2. Preliminaries

2.1. L-Fuzzy sets

We will present our results in the context of L-fuzzy sets, i.e., all the results presented below hold when membership takes values in a lattic. Basic concepts of the theory of ordered sets and lattices will be used as usual, see e.g., [6], [21]. Given a set L we can equip it with an order relationship $\leq$ and thus obtain a partially order set $(L, \leq)$. If, for every pair $x, y \in L$, $\inf(x, y)$ and $\sup(x, y)$ exist, we say that $(L, \leq)$ is a lattice. We denote $\inf(x, y)$ by $x \wedge y$ and $\sup(x, y)$ by $x \vee y$; then $\wedge, \vee$ are binary operations on L, and we say that $\ell = (L, \leq, \wedge, \vee)$ is a lattice. Also, denote by 0 and 1 the minimum and maximum elements of lattice L, respectively. A lattice $(L, \leq, \wedge, \vee)$ is complete if the least upper bound $\bigvee S$ and the greatest lower bound $\bigwedge S$ exist for all $S \subseteq L$. A completely distributive lattice is a complete lattice in which arbitrary meets $(\wedge)$ distribute over arbitrary joins $(\vee)$ and vice versa.

Definition 1. ([13],[21]) Given a nonempty set X and a lattice $\ell = (L, \leq, \wedge, \vee)$, an *L-fuzzy* set is characterized by its membership function $\mu_A : X \to L$, and $\mu_A(x)$ is interpreted as the grade of membership of element x in L-fuzzy set A for each $x \in X$. We denote by $\mathcal{F}(X, \ell)$ the set of all L-fuzzy sets on universal set X with membership grades in L.

Definition 2. ([13],[21]) Let X be a nonempty set and $\ell = (L, \leq, \wedge, \vee)$ be a lattice. For every $A, B \in \mathcal{F}(X, \ell)$ we have $A \subseteq B \Leftrightarrow \forall x \in X; \mu_A(x) \leq \mu_B(x)$.

2.2. Fuzzy finite tree automata

Our definitions in this section, are different from that in [4], [10], [17], [18], [20] only in some minor details.

The set of natural numbers is denoted by $\mathbb{N}$, and the set of finite strings over $\mathbb{N}$ is $\mathbb{N}^*$. The empty string is denoted by ε. A $\Sigma -$ *alphabet* is a finite and nonempty set of symbols. A *ranked alphabet* is a couple $(\Sigma, Arity : \Sigma \to \mathbb{N} \cup \{0\})$, which is the disjoint union of sets of $n - ary$ symbols $\Sigma_n = \{\sigma | Arity(\sigma) = n\}$ for all $n \geq 0$. The set $T_\Sigma(Q)$ of $\Sigma - trees$ indexed by Q is inductively defined to be the smallest set such that $Q \subseteq T_\Sigma(Q)$ and $\sigma(t_1, \ldots, t_n) \subseteq T_\Sigma(Q)$ for every $\sigma \in \Sigma_n$ and $t_1, \ldots, t_n \in T_\Sigma(Q)$. We write T_Σ for $T_\Sigma(\phi)$.

Definition 3. Let $\ell = (L, \leq, \wedge, \vee)$ be a completely distributive lattice. A *fuzzy finite tree automaton* over ℓ is a system $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$, where:

1. Σ is a finite set of ranked alphabets called input symbols.
2. Q is a finite set of symbols called states.
3. $\Gamma : Q \to L$ is an L-fuzzy set on Q and called the set of final states.
4. $\delta = \{\delta_\sigma : Q^n \times Q \times \Sigma_n \to L \,|\, \sigma \in \Sigma_n,\, n \geq 0\}$ is a finite set called transition rules.
5. $\rho : T_\Sigma(Q) \times Q \to L$ is called the run map of FFTA M, and defined by induction on structure of $t \in T_\Sigma(Q)$:
   (a) If $t = \sigma \in \Sigma_0$, then $\rho(t)(q) = \delta(q, \sigma)$, for all $q \in Q$.
   (b) If $t = \sigma(t_1, \ldots, t_n)$ for some $\sigma \in \Sigma_n$ and $t_1, \ldots, t_n \in T_\Sigma$, then
   $$\rho(t)(q) = \bigvee_{q_1,\ldots,q_n \in Q} \left(\delta(q_1, \ldots, q_n, q, \sigma) \wedge \bigwedge_{i=1}^{n} \rho(t_i)(q_i) \right).$$

6. $\beta : T_\Sigma \to L$ is an L-fuzzy set on a set of trees $t \in T_\Sigma$, called behavior of FFTA M, and defined by:
$$\beta(t) = \bigvee_{q\in Q} \rho(t)(q) \wedge \Gamma(q).$$

An FFTA $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ *accepts* a tree $t \in T_\Sigma$ iff $\beta(t) > 0$. Also, the set of all trees accepted by M is known as fuzzy tree language $L(M)$ *recognized* by FFTA M. As well, $\mu_{L(M)}(t) = \beta(t)$. In other words, a recognizable fuzzy tree language is a fuzzy tree language recognized by some FFTA.

An FFTA $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ is called *deterministic* if for every $\sigma \in \Sigma_n$ and $q_1, \dots, q_n \in Q$, where $n \geq 0$, there exist at most one $q \in Q$, such that $\delta(q_1, \dots q_n, q, \sigma) > 0$.

An FFTA $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ is called *complete* if for every $\sigma \in \Sigma_n$ and $q_1, \dots, q_n \in Q$, where $n \geq 0$, there exist at least one $q \in Q$, such that $\delta(q_1, \dots q_n, q, \sigma) > 0$.

An FFTA $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ is called *reduced* if for every $q \in Q$, there exists at least one $t \in T_\Sigma$ such that $\rho(t)(q) > 0$.

3. Similarity based merging states of DFFTA

3.1. Distance and similarity measure

Definition 4. ([21],[28]) Let X be a nonempty set and ℓ be a lattice. A real function $\mathcal{D} : \mathcal{F}(X, \ell)^2 \to [0, 1]$ is called a *distance measure* on $\mathcal{F}(X, \ell)$, if $\mathcal{D}$ satisfies the following properties:

1. $\mathcal{D}(A, B) = \mathcal{D}(B, A)$; $\forall A, B \in \mathcal{F}(X, \ell)$,
2. $\mathcal{D}(A, A) = 0$; $\forall A \in \mathcal{F}(X, \ell)$,
3. $\forall A, B, C \in \mathcal{F}(X, \ell)$, if $A \subseteq B \subseteq C$, then $\mathcal{D}(A, B) \leq \mathcal{D}(A, C)$ and $\mathcal{D}(B, C) \leq \mathcal{D}(A, C)$;

Definition 5. ([21],[28]) Let X be a nonempty set and ℓ be a lattice. A real function $\mathcal{S} : \mathcal{F}(X, \ell)^2 \to [0, 1]$ is called a *similarity measure* on $\mathcal{F}(X, \ell)$, if $\mathcal{S}$ satisfies the following properties:

1. $\mathcal{S}(A, B) = \mathcal{S}(B, A)$; $\forall A, B \in \mathcal{F}(X, \ell)$,
2. $\mathcal{S}(A, A) = \bigvee \{\mathcal{S}(B, C) \mid B, C \in \mathcal{F}(X, \ell)\}$; $\forall A \in \mathcal{F}(X, \ell)$,
3. $\forall A, B, C \in \mathcal{F}(X, \ell)$, if $A \subseteq B \subseteq C$, then $\mathcal{S}(A, C) \leq \mathcal{S}(A, B)$ and $\mathcal{S}(A, C) \leq \mathcal{S}(B, C)$.

Proposition 6. ([28]) *There exists a one-to-one correlation between all distance measures and all similarity measures, where a distance measure $\mathcal{D}$ and its corresponding similarity measure $\mathcal{S}$ satisfy $\mathcal{D} + \mathcal{S} = 1$.*

The similarity measure generated by the distance measure $\mathcal{D}$ is denoted by $\mathcal{S}\langle\mathcal{D}\rangle = 1 - \mathcal{D}$, and the distance measure generated by similarity measure $\mathcal{S}$ is denoted by $\mathcal{D}\langle\mathcal{S}\rangle = 1 - \mathcal{S}$.

Lemma 7. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA and $L^{\gamma}(q)$ be an L-fuzzy set with membership function $\mu_{L^{\gamma}(q)}(t) = \rho(t)(q) \wedge \gamma$, where $q \in Q$ and $\gamma \in \ell$. Then,*

$$L(M) = \bigcup_{q \in Q} L^{\Gamma(q)}(q).$$

Proof. Since M is a DFFTA, for every $t \in T_{\Sigma}$ with $\mu_{L(M)}(t) > 0$, so there exists exactly one $q \in Q$ such that $\rho(t)(q) \wedge \Gamma(q) > 0$. Therefore,

$$L(M) = \bigcup_{q \in Q} \{t \mid t \in T_{\Sigma}, \rho(t)(q) \wedge \Gamma(q) > 0\} = \bigcup_{q \in Q} L^{\Gamma(q)}(q). \qquad \blacksquare$$

Definition 8. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. An operator $\Pi : \mathcal{P}(\ell) \to \ell$ is called centroid operation on δ, iff for every $\overline{\delta} \subseteq \delta$, $M' = (\Sigma, Q, \Gamma, \delta', \ell, \rho', \beta')$, $M'' = (\Sigma, Q, \Gamma, \delta'', \ell, \rho'', \beta'')$ and $\gamma, \gamma', \gamma'' \in \ell$ with

$$\forall q_1, \ldots, q_n, q \in Q, \sigma \in \Sigma_n, n \geq 0;$$

$$\gamma = \delta(q_1, \ldots, q_n, q, \sigma),$$

$$\gamma'' = \prod_{\substack{n \geq 0,\ \sigma \in \Sigma_n, \\ q_1, \ldots, q_n,\ q \in Q, \\ \overline{\delta}(q_1, \ldots, q_n, q, \sigma) > 0}} \overline{\delta}(q_1, \ldots, q_n, q, \sigma),$$

$$\overline{\delta}(q_1, \ldots, q_n, q, \sigma) > 0 \Rightarrow \begin{cases} \delta'(q_1, \ldots, q_n, q, \sigma) = \gamma', \\ \delta''(q_1, \ldots, q_n, q, \sigma) = \gamma'', \end{cases}$$

$$\overline{\delta}(q_1, \ldots, q_n, q, \sigma) = 0 \Rightarrow \begin{cases} \delta'(q_1, \ldots, q_n, q, \sigma) = \gamma, \\ \delta''(q_1, \ldots, q_n, q, \sigma) = \gamma, \end{cases}$$

where, ρ', β' are corresponding to δ', and ρ'', β'' are corresponding to δ''; it holds $\mathcal{D}(L(M), L(M'')) \leq \mathcal{D}(L(M), L(M'))$.

Definition 9. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, Π be a centroid operation on δ, $q, q' \in Q$ and $t \in T_{\Sigma}(Q)$. Then, $\rho(t)\left(q \xrightarrow{\Pi} q'\right)$ is defined by:

1. $\rho(t)\left(q \xrightarrow{\Pi} q\right) = \rho(t)(q)$,
2. $\rho(t)\left(q \xrightarrow{\Pi} q'\right)$

$$= \bigvee_{q_1, \ldots, q_n \in Q} \left(\left(\prod_{\substack{q_i \notin \{q, q'\} \Rightarrow q'_i = q_i, \\ q_i \in \{q, q'\} \Rightarrow q'_i \in \{q, q'\}, \\ i \in \{1, \ldots, n\},\ p \in \{q, q'\}}} \delta(q'_1, \ldots, q'_n, p, \sigma) \right) \wedge \left(\bigvee_{\substack{q_i \notin \{q, q'\} \Rightarrow q'_i = q_i, \\ q_i \in \{q, q'\} \Rightarrow q'_i \in \{q, q'\}, \\ i \in \{1, \ldots, n\},\ p \in \{q, q'\}}} \bigwedge_{i=1}^{n} \rho(t_i)\left(q'_i \xrightarrow{\Pi} q_i\right) \right) \right)$$

Definition 10. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA, $\mathcal{D}$ be a distance measure on $L(M)$, $\mu_{L^{\gamma}(q)} = \rho(t)(q) \wedge \gamma$ and $\mu_{L^{\gamma}(q \xrightarrow{\Pi} q')} = \rho(t)\left(q \xrightarrow{\Pi} q'\right) \wedge \gamma$; where Π is a centroid operation on δ, $\gamma \in \ell$, $q, q' \in Q$ and $t \in T_{\Sigma}(Q)$. Then, we develope the distance measure $\mathcal{D}$ on Q by:

$$\mathcal{D}^{\Pi}(q, q') = \bigwedge_{\gamma \in \ell} \mathcal{D}\left(L^{\Gamma(q)}(q) \cup L^{\Gamma(q')}(q'), L^{\gamma}\left(q \xrightarrow{\Pi} q'\right)\right); \ \forall q, q' \in Q.$$

Theorem 11. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, Π be a centroid operation on δ, $\mathcal{D}$ be a generalized distance measures on $L(M)$ and Q, $x \in [0,1]$ and Q' be a subset of Q such that for every $q, q' \in Q'$ it holds $\mathcal{D}^{\Pi}(q, q') \leq x$. Then, there exists a $\gamma \in \ell$ such that for every $q \in Q'$ it holds:*

$$\mathcal{D}\left(\bigcup_{q' \in Q'} L^{\Gamma(q')}(q'), \ \bigcup_{q' \in Q'} L^{\gamma}\left(q' \xrightarrow{\Pi} q\right)\right) \leq x.$$

Proof. Let

$$\gamma_{q,q'} = \bigwedge\left\{\gamma \in \ell \mid \mathcal{D}^{\Pi}(q, q') = \mathcal{D}\left(L^{\Gamma(q)}(q) \bigcup L^{\Gamma(q')}(q'), \ L^{\gamma}\left(q \xrightarrow{\Pi} q'\right)\right)\right\},$$

$$M_q = \bigvee_{q' \in Q'} \gamma_{q,q'} \ , \ m_q = \bigwedge_{q' \in Q'} \gamma_{q,q'} \ ; \ \forall q \in Q',$$

$$M = \bigwedge_{q \in Q'} M_q \ \textit{and} \ m = \bigvee_{q \in Q'} m_q.$$

Now, we prove that the above inequality holds for all $\gamma \in [m, M]$. To show this, on the contrary, let there exists a $\gamma' \in [m, M]$ such that

$$\mathcal{D}\left(\bigcup_{q' \in Q'} L^{\Gamma(q')}(q'), \ \bigcup_{q' \in Q'} L^{\gamma'}\left(q' \xrightarrow{\Pi} q\right)\right) > x.$$

Since M is deterministic and from Lemma 7, $L^{\Gamma(q)}(q) \cap L^{\Gamma(q')}(q') = \phi$ for every $q, q' \in Q$. So, there exists a $q \in Q'$ with $\mathcal{D}\left(L^{\Gamma(q)}(q), L^{\gamma'}(q)\right) > x$. Furthermore, we have $[m, M] \subseteq [m_q, M_q]$ which implies that $L^{m_q}(q) \subseteq L^{\gamma'}(q) \subseteq L^{M_q}(q)$. Let $M_q, m_q \in Q'$ be two states such that $M_q = \gamma_{q,q_M}$ and $m_q = \gamma_{q,q_m}$, where $q_m, q_M \in Q'$. Now, if $\Gamma(q) \leq \gamma'$, then $L^{\Gamma(q)}(q) \subseteq L^{\gamma'}(q) \subseteq L^{M_q}(q)$, and $\mathcal{D}^{\Pi}(q, q_M) > x$, which is a contradiction. On the other hand, if $\gamma' \leq \Gamma(q)$, and $L^{m_q}(q) \subseteq L^{\gamma'}(q) \subseteq L^{\Gamma(q)}(q)$, thus $\mathcal{D}^{\Pi}(q, q_m) > x$, which also is a contradiction. ■

Theorem 12. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, Π be a centroid operation on δ, and $\mathcal{D}$ be a generalized distance measures on $L(M)$ and Q; $x \in [0,1]$, and Q' be a subset of Q. If there exists a $\gamma \in \ell$ such that for every $q \in Q'$ it holds:*

$$\mathcal{D}\left(\bigcup_{q' \in Q'} L^{\Gamma(q')}(q'), \ \bigcup_{q' \in Q'} L^{\gamma}\left(q \xrightarrow{\Pi} q'\right)\right) \leq x.$$

Then, for any FFTA $\tilde{M} = (\Sigma, \tilde{Q}, \tilde{\Gamma}, \tilde{\delta}, \ell, \tilde{\rho}, \tilde{\beta})$ with

1. $\tilde{Q} = \{\tilde{q}\} \cup Q \backslash Q'$, *where* $\tilde{q} \notin Q$,
2. $\tilde{\Gamma}(\tilde{q}) = \gamma$ *and* $\tilde{\Gamma}(q) = \Gamma(q); \forall q \in \tilde{Q} \backslash \{\tilde{q}\}$,
3. $\forall \tilde{q}_1, \ldots, \tilde{q}_n \in \tilde{Q}, \forall q \in \tilde{Q} \backslash \{\tilde{q}\}$;

$$\tilde{\delta}(\tilde{q}_1, \ldots, \tilde{q}_n, q, \sigma) = \prod_{\substack{\tilde{q}_i = \tilde{q} \Rightarrow q'_i \in Q', \\ \tilde{q}_i \neq \tilde{q} \Rightarrow q'_i = \tilde{q}_i, \\ 1 \leq i \leq n}} \delta(q'_1, \ldots, q'_n, q, \sigma),$$

$$\tilde{\delta}(\tilde{q}_1, \ldots, \tilde{q}_n, \tilde{q}, \sigma) = \prod_{\substack{\tilde{q}_i = \tilde{q} \Rightarrow q'_i \in Q', \\ \tilde{q}_i \neq \tilde{q} \Rightarrow q'_i = \tilde{q}_i, \\ 1 \leq i \leq n, q' \in Q'}} \delta(q'_1, \ldots, q'_n, q', \sigma),$$

4. $\tilde{\rho}$ *and* $\tilde{\beta}$ *are corresponding to* $\tilde{\delta}$,

it holds $\mathcal{D}\left(L(M), L(\tilde{M})\right) \leq x$.

Proof. Let

$$L_1 = \bigcup_{q \in Q'} L^{\Gamma(q)}(q) \quad \text{and} \quad L_2 = \bigcup_{q \in Q \backslash Q'} L^{\Gamma(q)}(q).$$

Then, from Lemma 7, we have

a) $L(M) = \bigcup_{q \in Q} L^{\Gamma(q)}(q) = L_1 \cup L_2$.

b) $L(\tilde{M}) = \bigcup_{q \in \tilde{Q}} L^{\tilde{\Gamma}(q)}(q) = L^{\gamma}(\tilde{q}) \cup L_2$.

Now, from (a) and (b), and by the assumption of the theorem,

$$\mathcal{D}\left(L(M), L(\tilde{M})\right) = \mathcal{D}\left(L_1, L^{\gamma}(\tilde{q})\right) \leq x. \qquad \blacksquare$$

3.2. Normalizing DFFTA

Remark 13. In this manuscript, without lose of generality, we assume that Σ and Q are ordered sets.

Definition 14. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. An ordering on set δ is defined as follows:

1. If $\sigma < \sigma'$ then $\delta(q_1, ..., q_n, q, \sigma) < \delta(q'_1, ..., q'_m, q', \sigma')$.
2. If there exists $i \in \{1, ..., n\}$ such that $q_i < q'_i$, and $q_j = q'_j$ for $j \in \{1, ..., i-1\}$, then $\delta(q_1, ..., q_n, q, \sigma) < \delta(q'_1, ..., q'_n, q', \sigma)$.
3. If $q < q'$ then $\delta(q_1, ..., q_n, q, \sigma) < \delta(q_1, ..., q_n, q', \sigma)$.

where $q_1, ..., q_n, q, q'_1, ..., q'_m, q' \in Q$, $0 \leq n \leq m$, $\sigma \in \Sigma_n$ and $\sigma' \in \Sigma_m$.

Definition 15. Let S be an ordered set. We define the function $f^* : S \to \mathbb{N}$ called offset of $x \in S$ by

$$f^*(x) = |\{x' \in S | x' < x\}|.$$

Lemma 16. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a complete and reduced DFFTA. Then, for every $r, r' \in \delta$ with*

$$r : \delta(q_1, ..., q_i', ..., q_n, q', \sigma),$$
$$r' : \delta(q_1, ..., q_i, ..., q_n, q, \sigma),$$

it holds

$$f^*(r) = f^*(r') + (f^*(q_i') - f^*(q_i)) \times |Q|^{n-i},$$

where, $q_1, ..., q_i, ..., q_n, q, q', q_i' \in Q$, $\sigma \in \Sigma_n$, $q_i < q_i'$ and $1 \leq i \leq n$.

Proof. Since M is complete, reduced and deterministic; order of rules $r \in \delta$ related to each $\sigma \in \Sigma_n$ is like the sequence of $(n+1)$–*digit* numbers in base $|Q|$, where the value of each digit $q \in Q$ is $f^*(q)$. Now, the proof is straightforward. ∎

Corollary 17. *Let M be a complete and reduced DFFTA. The time complexity of accessing the membership grade of each fuzzy transition rule is $O(l)$, where l is the maximum rank of Σ–alphabet.*

Example 18. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a complete and reduced DFFTA defined by

$$\begin{aligned}
&\Sigma_0 = \{\alpha\}, \Sigma_1 = \{\lambda\}, \Sigma_2 = \{\sigma\}, \Sigma = \{\alpha, \lambda, \sigma\}, \ell = ([0,1], \geq, \wedge, \vee),\\
&Q = \{q_1, q_2, q_3, q_4\}, \Gamma = \{(q_1, 0.7), (q_2, 0.5), (q_3, 0.5)\} \text{ and}\\
&\delta = \{r_1 : \delta(q_1, \alpha) = 0.9, r_2 : \delta(q_1, q_4, \lambda) = 0.8, r_3 : \delta(q_2, q_4, \lambda) = 0.8,\\
&\quad r_4 : \delta(q_3, q_4, \lambda) = 0.8, r_5 : \delta(q_4, q_2, \lambda) = 0.8, r_6 : \delta(q_1, q_1, q_3, \sigma) = 0.7,\\
&\quad r_7 : \delta(q_1, q_2, q_3, \sigma) = 0.7, r_8 : \delta(q_1, q_3, q_2, \sigma) = 0.6, r_9 : \delta(q_1, q_4, q_4, \sigma) = 0.6,\\
&\quad r_{10} : \delta(q_2, q_1, q_3, \sigma) = 0.6, r_{11} : \delta(q_2, q_2, q_3, \sigma) = 0.6, r_{12} : \delta(q_2, q_3, q_2, \sigma) = 0.6,\\
&\quad r_{13} : \delta(q_2, q_4, q_4, \sigma) = 0.2, r_{14} : \delta(q_3, q_1, q_2, \sigma) = 0.3, r_{15} : \delta(q_3, q_2, q_2, \sigma) = 0.3,\\
&\quad r_{16} : \delta(q_3, q_3, q_3, \sigma) = 0.3, r_{17} : \delta(q_3, q_4, q_4, \sigma) = 0.2, r_{18} : \delta(q_4, q_1, q_4, \sigma) = 0.1,\\
&\quad r_{19} : \delta(q_4, q_2, q_4, \sigma) = 0.1, r_{20} : \delta(q_4, q_3, q_4, \sigma) = 0.1, r_{21} : \delta(q_4, q_4, q_4, \sigma) = 1\}.
\end{aligned}$$

Now, it holds

$$f^*(r_{16}) = f^*(r_{14}) + (f^*(q_3) - f^*(q_1))|Q|^{Arity(\sigma)-2} = 13 + (2-0) \times 4^{2-2} = 15.$$

Definition 19. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. For any $q \in Q$ we define the *accessibility grade set* by

$$\mathcal{A}(q) = \{\rho(q)(t) | t \in T_\Sigma, \rho(q)(t) > 0\}.$$

Lemma 20. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. Then, the set of maximum accessibility grade ($\vee \mathcal{A}(q)$) of all $q \in Q$ can be calculated in $O(l|Q|^l)$, where l is the maximum rank of $\Sigma - alphabet$.*

Proof. We define the Algorithm 1 for computing $\bigvee \mathcal{A}(q)$ for all $q \in Q$:

Algorithm 1. Computing maximum accessibility grade of all states.

---

0 Input: $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$
1 $\forall q \in Q; V(q) = 0$
2 $\forall \sigma \in \Sigma_0, q \in Q; V(q) = V(q) \vee \delta(q, \sigma)$
3 $A_{\max} = \bigvee_{q \in Q} V(q)$
4 $Q_{acc} = \phi$
5 Repeat
6    $Q_{new} = \{q | q \in Q,\ V(q) = A_{\max}\}$
7    $\forall q_1, \ldots, q_n \in Q_{acc} \cup Q_{new}, \{q_1, \ldots, q_n\} \cap Q_{new} \neq \phi, q \in Q,\ \sigma \in \Sigma_{n>0};$
\- $V(q) = V(q) \vee (\delta_\sigma(q_1, \ldots, q_n, q, \sigma) \wedge A_{\max})$
8    $A_{\max} = \bigvee_{q \in Q, V(q) < A_{\max}} V(q)$
9    $Q_{acc} = Q_{acc} \cup Q_{new}$
10 Until $Q = Q_{acc}$
11 Output: $V(q); \forall q \in Q.$

Lines 5 to 10 consist a loop that in each iteration, processes at least one transition rule. Since each transition rule will be processed only one time, the number of repetitions of this loop is $O(|\delta|)$. We note that the number of repeating lines 7 to 9 is not more than $|\delta|$ times (see $\{q_1, \ldots, q_n\} \cap Q_{new} \neq \phi$). Furthermore, making each transition rule requires combining l states. In the other hand, according to Lemma 7 and Corollary 17 accessing the membership grade of each rule is $O(l)$, which can be merged with the process of transition rule making. Therefore this loop can be done in $O(l|\delta|)$. ∎

Definition 21. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. For any $q \in Q$ we define the *behavior grade set* by

$$\mathcal{B}(q) = \{\beta(t) | t \in T_\Sigma, \rho(q)(t) \wedge \Gamma(q) > 0\}.$$

Lemma 22. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, for every $q \in Q$, it holds*

$$\mathcal{B}(q) = \{\gamma \wedge \Gamma(q) | \gamma \in \mathcal{A}(q)\}.$$

Proof. Since M is deterministic, the condition $\rho(q)(t) \wedge \Gamma(q) > 0$ implies that $\beta(t) = \rho(q)(t) \wedge \Gamma(q)$. Hence,

$$\mathcal{B}(q) = \{\rho(t)(q) \wedge \Gamma(q) | t \in T_\Sigma, q \in Q, \rho(q)(t) > 0\} = \{\gamma \wedge \Gamma(q) | \gamma \in \mathcal{A}(q)\}.$$ ∎

Corollary 23. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, for all $q \in Q$, it holds*

$$\vee\mathcal{B}(q) = \vee\mathcal{A}(q) \wedge \Gamma(q).$$

Proof. Using Lemma 22 we have

$$\mathcal{B}(q) = \{\gamma \wedge \Gamma(q)|\gamma \in \mathcal{A}(q)\} = \vee\{\gamma|\gamma \in \mathcal{A}(q)\} \wedge \Gamma(q) = \vee\mathcal{A}(q) \wedge \Gamma(q). \quad \blacksquare$$

Corollary 24. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, the $\vee\mathcal{B}(q)$ for all $q \in Q$ can be calculated in $O(l|Q|^l)$, where l is the maximum rank of $\Sigma -$ alphabet.*

Proof. It is and immediate consequence of the Lemma 16 and Corollary 23. ■

Definition 25. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. The *next* set of every $q \in Q$ is defined by:

$$next(q) = \{q'|1 \leq i \leq n, \sigma \in \Sigma_n, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, q' \in Q, \\ \delta(q_1, \ldots, q_{i-1}, q, q_{i+1}, \ldots, q_n, q', \sigma) \geq 0\}$$

Lemma 26. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, the order of time complexity for calculating next set of all $q \in Q$ is $O(l|Q|^l)$, where l is the maximum rank of $\Sigma -$ alphabet.*

Proof. It is sufficient that for all $\delta(q_1, \ldots, q_n, q, \sigma) \geq 0$, with $q_1, \ldots, q_n, q \in Q$, $1 \leq n$ and $\sigma \in \Sigma_n$, add q to sets $next(q_1)$,..., $next(q_n)$. ■

Definition 27. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be an FFTA. The *follow* set of each $q \in Q$ is the smallest set with the following properties:

1. $next(q) \subseteq follow(q)$,
2. If $q' \in follow(q)$ then, $next(q') \subseteq follow(q)$.

Lemma 28. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, the order of time complexity for calculating the follow set of all $q \in Q$ is $O(l|Q|^{\max(l,2)})$, where l is the maximum rank of $\Sigma -$ alphabet.*

Proof. According to Lemma 26, the time complexity of calculating *next* set of all $q \in Q$ is $O(l|Q|^l)$. Now, according to properties of *follow* set, computing it for each state can be done by a simple recursive process with the time complexity $O(|Q|^2)$. We mention that the total time complexity, when $l = 1$ is $O(|Q|^2)$ and otherwise, is $O(l|Q|^l)$. ■

Definition 29. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. The *maximum follow grade* of each $q \in Q$ is defined by:

$$\mathcal{G}(q) = \bigvee_{q' \in follow(q)} \mathcal{B}(q').$$

Lemma 30. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. The time complexity for calculating $\mathcal{G}(q)$ for all $q \in Q$ is $O(l|Q|^{\max(l,2)})$, where l is the maximum rank of $\Sigma -$ alphabet.*

Proof. It is similar to Lemma 28. ∎

Definition 31. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, $\sigma \in \Sigma_n$, $q_1, \ldots, q_n, q \in Q$ and $\gamma \in \ell$. A transition rule $\delta(q_1, \ldots, q_n, q, \sigma) = \gamma$ is called *normal* iff

$$\gamma = \Big(\Gamma(q) \vee \mathcal{G}(q)\Big) \wedge \bigwedge_{i=1}^{n} \big(\vee \mathcal{A}(q_i)\big).$$

Also, M is a *normal FFTA* iff all transition rules in δ be *normal* and for each $q \in Q$ it holds $\Gamma(q) \leq \vee\mathcal{A}(q)$.

Theorem 32. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA. Then, normalizing M can be done in order of time complexity $O(l|Q|^l)$, where l is the maximum rank of $\Sigma -$ alphabet.*

Proof. We define Algorithm 2 for *normalizing M*:

Algorithm 2. Normalizing a DFFTA.

---

0 Input: $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$
1 $\forall q \in Q; V(q) = \vee\mathcal{A}(q)$
2 $A_{\max} = \bigvee_{q \in Q} V(q)$
3 $Q_{acc} = \phi$
4 $\delta^{\Delta} = \phi$
5 Repeat
6 $\quad Q_{new} = \{q | q \in Q, V(q) = A_{\max}\}$
7 $\quad \forall q \in Q, \sigma \in \Sigma_n, n \geq 0, q_1, \ldots, q_n \in Q_{acc} \cup Q_{new}$,
- $\quad\quad \{q_1, \ldots, q_n\} \cap Q_{new} = \phi \Leftrightarrow n = 0$;
- $\quad\quad\quad \delta^{\Delta}(q_1, \ldots, q_n, q, \sigma) = \delta(q_1, \ldots, q_n, q, \sigma) \wedge A_{\max} \wedge \Big(\Gamma(q) \vee \mathcal{G}(q)\Big)$
8 $\quad A_{\max} = \bigvee_{q \in Q, V(q) < A_{\max}} V(q)$
9 $\quad Q_{acc} = Q_{acc} \cup Q_{new}$
10 Until $Q = Q_{acc}$
11 $\forall q \in Q; \Gamma^{\Delta}(q) = \Gamma(q) \wedge V(q)$
12 Output: $M = (\Sigma, Q, \Gamma^{\Delta}, \delta^{\Delta}, \ell, \rho, \beta)$.

According to Lemma 30, calculating $\mathcal{G}(q)$ for all $q \in Q$ can be done in $O(l|\delta|)$. Lines 5 to 10 are a loop that in each repetition, adds at least one transition rule to δ^{Δ}. Thus making all transition rules must be repeats $|\delta|$ times and making each rule, involves combining n states. Therefore, this loop can be done in $O(n|\delta|)$. Calculating the complexity of other lines is straightforward. ∎

3.3. Merging states of DFFTA

Definition 33. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, Π be a centroid operation on δ, $\mathcal{S}\langle\mathcal{D}\rangle$ be a similarity measure corresponding to distance measure $\mathcal{D}$ on $L(M)$ and Q, and $\tau \in [0,1]$. The similarity relation $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ on Q, is defined by:

$$q \overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow} q' \Longleftrightarrow 1 - \mathcal{D}^{\Pi}(q,q') \geq \tau \ ; \ \forall q, q' \in Q.$$

Lemma 34. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a normal DFFTA, Π be a centroid operation on $\delta, \mathcal{S}$ be a similarity measures on $L(M)$ and Q, $\tau \in [0,1]$, and $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ be a similarity relation on Q. Also, let Q' be a subset of Q such that for every $q, q' \in Q'$ we have $q \overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow} q'$, $\delta_{Q'} = \{r : \delta(q_1, ..., q_n, q, \sigma) = c | c > 0,\ \sigma \in \Sigma_n, q_1, ..., q_n \in Q'\}$, and let $d = \prod_{r \in \delta_{Q'}} r$ and $\delta'_{Q'} = \{r : \delta(q_1, ..., q_n, q, \sigma) = d \mid \sigma \in \Sigma_n, q_1, ..., q_n \in Q'\}$. Then, it holds $\mathcal{S}(\delta_{Q'}, \delta'_{Q'}) \geq \tau$.*

Proof. Let $x = 1 - \tau$. for every $q, q' \in Q'$ it holds $q \overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow} q'$; thus, $\mathcal{D}^{\Pi}(q,q') \leq x$. Then, from Theorem 11, there exists a $\gamma \in \ell$ such that for every $q \in Q'$ it holds:

$$\mathcal{D}\left(\bigcup_{q' \in Q'} L^{\Gamma(q')}(q'), \bigcup_{q' \in Q'} L^{\gamma}\left(q' \overset{\Pi}{\longrightarrow} q\right)\right) \leq x.$$

From Theorem 12 and since M is normal, we have $\mathcal{D}(\delta_{Q'}, \delta'_{Q'}) \leq x$. ■

Theorem 35. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a complete and reduced DFFTA and $\mathcal{S}$ be a similarity measure corresponding to distance measures $\mathcal{D}$ on $L(M)$ and Q. Then, the similarity relation $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ is not transitive for some $\tau \in [0,1]$ and centroid operation Π on δ.*

Proof. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be DFFTA defined by:

$$\begin{aligned}
&\Sigma_0 = \{\alpha\},\ \Sigma_2 = \{\sigma\},\ \Sigma = \{\alpha, \sigma\},\ Q = \{q_1, q_2, q_3\},\\
&\Gamma = \{(q_1, 0.2), (q_2, 0.3), (q_3, 0.6)\}, \text{ and}\\
&\delta = \{r_1 : \delta(q_1, \alpha) = 0.6, r_2 : \delta(q_1, q_1, q_2, \sigma) = 0.6,\\
&\qquad r_3 : \delta(q_1, q_2, q_3, \sigma) = 0.6, r_4 : \delta(q_1, q_3, q_3, \sigma) = 0.6,\\
&\qquad r_5 : \delta(q_2, q_1, q_1, \sigma) = 0.6, r_6 : \delta(q_2, q_2, q_1, \sigma) = 0.6,\\
&\qquad r_7 : \delta(q_2, q_3, q_2, \sigma) = 0.6, r_8 : \delta(q_3, q_1, q_2, \sigma) = 0.6,\\
&\qquad r_9 : \delta(q_3, q_1, q_1, \sigma) = 0.6, r_{10} : \delta(q_3, q_2, q_1, \sigma) = 0.6\}.
\end{aligned}$$

Also, let $\ell = ([0,1], \geq, \wedge, \vee)$ be a lattice and $\overset{\mathcal{S},\text{mid},0.8}{\longleftrightarrow}$ be a similarity relation, where $\mathcal{S}(A,B) = 1 - \bigvee_{a \in A, b \in B} |a - b|; \forall A, B \in \mathcal{F}(T_{\Sigma}, \ell)$ and $mid(C) = \frac{\vee C + \wedge C}{2}$; $\forall C \subseteq [0,1]$. (In this example, $[0,1]$ is the unit interval of real numbers.) Now, we have $q_1 \overset{\mathcal{S},\text{mid},0.8}{\longleftrightarrow} q_2$ and $q_2 \overset{\mathcal{S},\text{mid},0.8}{\longleftrightarrow} q_3$, where $q_1 \overset{\mathcal{S},\text{mid},0.8}{\longleftrightarrow} q_3$ does not hold. ■

Corollary 36. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, $\prod$ be a centroid operation on δ, $\mathcal{S}$ be a similarity measure on $L(M)$ and on Q, $\tau \in [0, 1]$, and $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ be a similarity relation on Q. Then, for some $P = \bigcup_{q \in Q} \{q' \overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow} q\}$, there exists $p, p' \in P$ such that $p \cap p' \neq \phi$.*

Definition 37. Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a reduced and complete DFFTA. A *Merging Dependency Graph* (MDG) $G = (V, E)$ on M is a directed graph with the following properties:

1. $V = \Big\{\{q, q'\} |\, q, q' \in Q\Big\}$,
2. $E \subseteq V^2$,
3. $\big(\{p, p'\}, \{q, q'\}\big) \in E$ if and only if there exist $\sigma \in \Sigma_n$, $1 \leq i \leq n$ and $q_1, ..., q_n, q, q', p, p' \in Q$ such that $\delta(q_1, ..., q_{i-1}, q, q_{i+1}, ..., q_n, p, \sigma) > 0$ and $\delta(q_1, ..., q_{i-1}, q', q_{i+1}, ..., q_n, p', \sigma) > 0$.

Lemma 38. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a complete and reduced DFFTA. The order of time complexity of making MDG $G = (V, E)$ for M is $O(l\,|Q|^{l+1})$, where l is the maximum arity of $\Sigma -$ alphabet.*

Proof. All rules in δ must be processed for constructing set E. As well, the set of rules related to each $\sigma \in \Sigma_n$, denoted by δ_σ, compare with each other. Therefore, it is sufficient to prove that the order of time complexity for processing all rules in δ_σ is $O(n|Q|^{n+1})$. Now, we introduce Algorithm 3 to process all rules in δ_σ for constructing the edges in set E:

Algorithm 3. Making MDG for a DFFTA.

---

0 Input: $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$
1 $E = \phi$
2 $\forall r : \delta(q_1, \ldots, q_n, q, \sigma) > 0;$
3 $\quad \forall i \in \{1, \ldots, n\};$
4 $\quad\quad \forall r' : \delta(q_1, \ldots, q'_i, \ldots, q_n, q', \sigma) > 0, q'_i \in Q, q'_i > q_i;$
5 $\quad\quad\quad E = E \cup \Big\{\big(\{q, q'\}, \{q_i, q'_i\}\big)\Big\}$
6 Output: E.

The proof of correctness of this algorithm is straightforward; thus, we analyze its order of time complexity. Lines 2 to 5 are three nested loops which cause lines 4 and 5 process $\frac{n|Q|^{n+1}}{2}$ times. According to Lemma 16, the order of time complexity for finding offset of r' by r is $O(1)$. Furthermore, we assume that a two dimension array (adjacency matrix) is used for holding set E. Thus, the order of time complexity for adding an edge to E is $O(1)$. Therefore, the order of time complexity for processing all rules in $\delta_\sigma$is $O(n|Q|^{n+1})$. ∎

Lemma 39. *Let $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$ be a DFFTA, Π be a centroid operation on δ, $\mathcal{S}$ be a similarity measures on $L(M)$ and Q, and $\tau \in [0, 1]$. Then, making graph $G = (V, E)$ such that $V = Q$ and E is corresponding to similarity relation $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ on Q can be done in order of complexity $O(l|Q|^{\max(2l,4)})$, where l is the maximum arity of $\Sigma-$ alphabet.*

Proof. Without lose of generality, we can assume that M is normal. Now, we introduce Algorithm 4 for processing all rules in δ_σ and constructing graph G.

Algorithm 4. Making MDG corresponding to similarity relation $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$.

---

0 Input: $M = (\Sigma, Q, \Gamma, \delta, \ell, \rho, \beta)$
1 $E = \left\{(q, q')|q, q' \in Q, \mathcal{S}^{\frac{1}{2}}\big(\mathcal{B}(q), \mathcal{B}(q')\big) \geq \tau\right\}$
2 $\forall\ r : \delta(q_1, \ldots, q_n, q, \sigma) > 0\ ,\ r' : \delta(q'_1, \ldots, q'_n, q', \sigma) > 0;$
3 $\quad R_{rr'} = \phi$
4 $\quad \mathcal{S}^{\frac{1}{2}}(r, r') < \tau\ ,\ \forall i \in \{1, \ldots, n\}; (q_i, q'_i) \in E\ \Rightarrow$
5 $\qquad R_{rr'} = \left\{(q_i, q'_i)|i \in \{1, \ldots, n\}\right\}$
6 $R = \bigcup_{r,r' \in \delta} R_{rr'}$
7 Repeat
8 $\quad (q, q') = \bigvee R$
9 $\quad E = E - \{(q, q')\}$
10 $\quad \forall r, r' \in \delta; (q, q') \in R_{rr'} \Rightarrow R_{rr'} = \phi$
11 $\quad R = \bigcup_{r,r' \in \delta} R_{rr'}$
12 Until $R = \phi$
13 Output: $G = (Q, E)$.

This algorithm clusters transition rules that cannot merge together (because of their grade of transition) by clustering the set Q. Also, line 1 makes an initial clustering based on the similarity of states. According to Corollary 24, the time complexity of line 1 is $O(l|Q|^l)$, where l is the maximum arity of Σ. Lines 2 to 5 is a loop that repeats $|Q|^{2n}$ times, where n is the arity of σ. The order of lines 4 and 5 is $O(l)$; therefore, the order of time complexity for lines 2 to 5 is $O(l|Q|^{2l})$. As well, the order of time complexity of line 6 is $O(l|Q|^{2l})$. Lines 7 to 12 is a loop that repeats $O(|Q|^2)$ times and its order of time complexity is $O(l|Q|^{\max(2l,4)})$. Therefore, the total order of time complexity of algorithm $O(l|Q|^{\max(2l,4)})$. ∎

4. Conclusion

We contribute the problem of similarity based merging states of DFFTA. Firstly, the concept of similarity and distance measure of fuzzy sets is generalized for states of FFTA and a similarity relation $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$ is defined on Q. We prove that this relation is not transitive; therefore, minimizing DFFTA by $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$, is not similar to traditional minimization algorithms. As well, the concept of normalizing DFFTA with a polynomial time algorithm is introduced. Then, normal DFFTA is used for obtaining MDG and defining an ordering on Q^2. Furthermore, we present Algorithm 4 and show that the order of time complexity for making similarity relation graph on Q corresponding to $\overset{\mathcal{S},\Pi,\tau}{\longleftrightarrow}$, is $O(l|Q|^{max(4,2l)})$, where l is the maximum arity of Σ-*alphabet*.

References

[1] Arbib, M.A., *Towards automata theory to brain theory*, International Journal of Man-Machine Studies, 7 (3) (1975), 279–295.

[2] Bozapalidis, S., Bozapalidoy, O.L., *On the recognizability of fuzzy languages*. I. Fuzzy Sets and Systems, 157 (17) (2006), 2394–2402.

[3] Bozapalidis, S., Bozapalidoy, O.L., *On the recognizability of fuzzy languages*. II. Fuzzy Sets and Systems, 159 (1) (2008), 2394–2402.

[4] Bozapalidis, S., Bozapalidoy, O.L., *Fuzzy tree language recognizability*, Fuzzy Sets and Systems, 161 (5) (2010), 716–734.

[5] Comon, H., Dauchet, M., Gilleron, R., Jacquemard, F., Lugiez, D., Loding, C., Tison, S., Tommasi, M., *Tree automata: techniques and applications*, Online, 2007. Available: http://tata.gforge.inria.fr.

[6] Davey, B.A., Priestley, H.A., *Introduction to lattices and order*, 2nd ed. Cambridge University Press, New York, 2002.

[7] Doner, J.E., *Decidability of the weak second-order theory of two successors*, Notices American Mathematic Society, 12 (1) (1965), 365–368.

[8] Doner, J.E., *Tree acceptors and some of their applications*, Journal of Computer and System Sciences, 4 (5) (1970), 406–451.

[9] Dubois, D., Prade, H., *Fuzzy sets and systems: theory and applications*, Academic Press, New York, 1980.

[10] Esik, Z., Guangwu, L., *Fuzzy tree automata*, Fuzzy Sets and Systems, 158 (2007), 1450–1460.

[11] Fallah, M.K., Moghari, S., Nazemi, E., Zahedi, M.M., *Fuzzy ontology based document feature vector modification using fuzzy tree transducer*. In *Proceedings of the 2008 IEEE International Conference on Signal Image Technology and Internet Based Systems* (2008), 38–44.

[12] Ghorani, M., Zahedi, M.M., *Characterizations of complete residuated lattice-valued finite tree automata*, Fuzzy Sets and Systems, 199 (2012), 28–46.

[13] Goguen, J.A., *L-fuzzy sets*, Journal of Mathematical Analysis and Applications, 18 (1) (1967), 145–174.

[14] Gratzer, G.A., *General lattice theory*, 2nd ed. Birkhauser, 1998.

[15] Maletti, A., *Minimizing deterministic weighted tree automata*. In *Proceedings of the Second International Conference on Language and Automata: Theory and Applications, LNCS*, vol. 5196, Springer, Berlin (2008), 357–372.

[16] Maletti, A., *Minimizing deterministic weighted tree automata*, Information and Computation, 207 (11) (2009), 1284–1299.

[17] Moghari, S., Zahedi, M.M., Ameri, R., *Minimization of fuzzy finite tree automata*. In *Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Applications* (ICMSA2010) Universiti Tunku Abdul Rahman, Kuala Lumpur, Malaysia (2010), 1044–1052.

[18] Moghari, S., Zahedi, M.M., Ameri, R., *New direction in fuzzy tree automata*, Iranian Journal of Fuzzy Systems, 8 (5) (2011), 59–68.

[19] Mordeson, J., Malik, D.S., *Fuzzy discrete structures*, 1st ed., Physica-Verlag, New York, 2000.

[20] Mordeson, J., and Malik, D.S., *Fuzzy automata and languages: theory and applications*, Chapman & Hall, London, 2002.

[21] Nguyen, H.T., and Walker, E.A., *A first course on fuzzy logic*, CRC Press, Boca Raton, 1997.

[22] Pappis, C.P., *Value approximation of fuzzy systems variables,* Fuzzy Sets and Systems, 39 (1991), 111–115.

[23] Thatcher, J.W., Wright, J.B., *Generalized finite automata*, Notices American Mathematic Society, 12 (1965), 820.

[24] Thatcher, J.W., Wright, J.B., *Generalized finite automata with an application to a decision problem of second-order logic*, Mathematical System Theory, 2 (1) (1968), 57–82.

[25] Wang, D.G., Meng, Y.P., Li, H.X., *A fuzzy similarity inference method for fuzzy reasoning*, Computers & Mathematics with Applications, 56 (10) (2008), 2445–2454.

[26] Wee, W.G., *On generalization of adaptive algorithm and application of the fuzzy sets concept to pattern classification*, PhD thesis, Purdue University, Lafayette, IN, 1967.

[27] Xiong, N., *Fuzzy rule-based similarity model enables learning from small case bases*, Applied Soft Computing, 13 (4) (2013), 2057–2064.

[28] Xuecheng, L., *Entropy, distance measure and similarity measure of fuzzy sets and their relations*, Fuzzy Sets and Systems, 52 (1992), 305–308.

[29] Zadeh, L.A., *Fuzzy sets*, Information and Control, 8 (3) (1965), 338–353.

Accepted: 28.04.2014