Limitations of Non-deterministic Finite Automata Imposed by One Letter Input Alphabet

Laura Mančinska, Māris Ozols,
Renāte Praude and Agnese Zalcmane

violeta@navigator.lv, marozols@yahoo.com,
renka@inbox.lv, agnese.zalcmane@gmail.com
Department of Computer Science, University of Latvia,
Raiņa bulvāris 29, Rīga, LV-1459, Latvia

Conference: FCS'05
Keywords: finite automata, NFA, one letter input alphabet, limitations of NFA, unreachable subsets

Abstract

NFA usually requires significantly less states than DFA to recognize the same language. NFAs in one letter input alphabet are more restricted and the gap between NFAs and DFAs decreases, because the power of NFA is in its ability to reach many subsets of its state set. We discuss limitations of DFAs in one letter input alphabet and show that approximately 1/4 of all subsets are unreachable and for every fixed $k \in \{2,\ldots,number\_of\_states-2\}$ at least one subset of size k is unreachable.

1. Introduction

It is known that Non-deterministic Finite Automata (NFA) recognize the same languages as Deterministic Finite Automata (DFA). For every n-state NFA (n-NFA) there exists equivalent DFA with at most 2^n states. It is known, that for 3 or more letter input alphabet it is possible to construct n-NFA for which equivalent DFA cannot be built with less than 2^n states. If only 2 letter input alphabet is used, it is possible to construct n-NFA for witch equivalent DFA requires at least 2^n-1 states. There are no published results for one letter input alphabet. Currently Ansis Rosmanis is researching this problem and has unpublished result that states: for each n there exists n-NFA with one letter input alphabet (n-NFA1) for which equivalent DFA has at least polynomial (with respect to n) size, but does not exist n-NFA1 for which equivalent DFA requires exponential (with respect to n) number of states.

In this paper we will discuss only NFAs with one letter input alphabet (NFA1). In section 3 we propose a brief combinatorial illustration for the fact that for each n-NFA1 approximately 1/4 of its configurations is unreachable (thus equivalent DFA can be built with at most $3 \cdot 2^{n-2}$ states). In section 4 we prove that for each $k \in \{2,\ldots,n-2\}$ every NFA1 cannot reach at least one of it's subset of size k (if k does not belongs to interval mentioned, then for fixed $k \in \{0,1,n-1,n\}$ it is trivial to construct such an automaton). It is important to note, that estimation of number of all reachable configurations done by Ansis Rosmanis does not say anything about subsets of fixed size. Thus our proof shows the weakness of NFA1s even for fixed size of its subsets.

This research was supported by Grant No. 05.1528 from the Latvian Council of Science and by the European Commission, Contract IST-1999-11234 (QAIP) phone: +371-7224363, fax: +371-7820153.

2. Terms and Notations

NFA1 – Non-deterministic Finite Automaton (NFA) whose input alphabet consists of one letter

n-NFA1 – NFA1 that has n states

q – a state of NFA1

Q_n – a set that contains all states of n-NFA1 ($|Q_n|=n$)

$q_i\text{--->}q_j$ – represents a transition (due to reading one input letter) from q_i to q_j ($q_i,q_j\in Q_n$)

$q_i\text{-}\varepsilon\text{->}q_j$ - represents an ε-transition from q_i to q_j($q_i,q_j\in Q_n$)

configuration (or subset) – a subset of Q_n

Conf(t) – the configuration of n-NFA1 at the moment t (after reading in a word of length t). This subset includes those and only those states which are reachable from some initial state by reading word of length t in ($Conf(t)\subseteq Q_n$).

a subset S is reached at the moment t – S=Conf(t)

a subset S is reachable – there exists t such as Conf(t)=S

a subset S is not reachable – there does not exist t such as Conf(t)=S (in fact, it is sufficient to show that the automaton has not reached subset S before reaching some subset twice)

|S| - the number of elements in the set S

k-subset – a subset of Q_n which contains exactly k states

cycle – a set of states $\{q_0,q_1,...,q_{c-1}\}$ such as $\forall i\in\{0,...,c-1\}:q_i\text{--->}q_{i\oplus 1}$, where ⊕ denotes addition modulus c (c is called the length of cycle)

chain – a set of states $\{q_0,q_1,...,q_{c-1}\}$ such as $\forall i\in\{0,...,c-2\}:q_i\text{--->}q_{i+1}$ (c is called the length of chain, q_{c-1} is called the end of the chain)

3. Unreachable Configurations

It is known that each n-NFA can be transformed without changing the number of states and the amount of reachable configurations so that it does not contain ε-transitions. Thus in this section only n-NFA's without ε-transitions will be discussed.

Theorem T1 For each n-NFA1 at least ~1/4 of its configurations is unreachable.

Lemma T1[L1] If NFA1 contains subgraph depicted in Fig. 1 ($q_1\neq q_2$) then at least ~1/4 of its configurations is unreachable.

Proof T1[L1] Let us denote the set of all configurations that contains q_0 by N_0 ($|N_0|=2^{n-1}$) and the set of configurations that contains q_1 and q_2 by $N_{1\&2}$ ($|N_{1\&2}|=2^{n-2}$). If $Conf(t)\in N_0$ then $Conf(t+1)\in N_{1\&2}$ (it also concerns configurations that are included both in N_0 and $N_{1\&2}$). If the same configuration is reached twice then configurations that have not been reached until that moment will not be reached at all. Thus only $|N_{1\&2}|+1$ configurations from N_0 are reachable. The other $2^{n-2}-1$ will be unreachable. The amount of unreachable configurations forms approximately 1/4 of all (2^n) configurations.

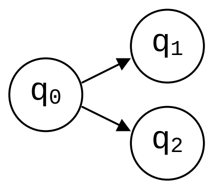

Fig. 1

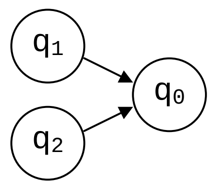

Fig. 2

Lemma T1[L2] If NFA1 contains subgraph depicted in Fig. 2 ($q_1\neq q_2$) then at least ~1/4 of its configurations is unreachable.

Proof T1[L2] Let us denote the set of all configurations that contains q_0 by N_0 ($|N_0|=2^{n-1}$) and the set of configurations that contains q_1 or q_2 by N_{1v2} ($|N_{1v2}|=2\cdot 2^{n-1}-2^{n-2}=3\cdot 2^{n-2}$). If $Conf(t)\in N_{1v2}$ then $Conf(t+1)\in N_0$ (it also concerns configurations that are included both in N_0 and N_{1v2}). If the same configuration is reached twice then configurations that have not been reached until that moment will not be reached at all. Thus only $|N_0|+1$ configurations from N_{1v2} are reachable. The other $2^{n-2}-1$ will be unreachable. The amount of unreachable configurations forms approximately 1/4 of all (2^n) configurations.

Lemma T1[L3] If NFA1 contains neither subgraph depicted in Fig. 1 nor subgraph depicted in Fig. 2. ($q_1 \neq q_2$) then at least ~1/2 of its configurations is unreachable.

Proof T1[L3] In this case the number of both incoming and outcoming arrows for each state is 0 or 1. Thus automaton consists of separate parts and each part is either a cycle or a chain. Let us perform the following transformations that do not influence the amount of reachable configurations. Remove all cycles that do not contain any initial state or contain only initial states. Thus all cycles of length 1 will be removed. Leave only one state in each cycle as initial. Find the chain with the most distant initial state viewed from its end and leave this state as only initial state in the chain. Remove all other chains. Now we have gained automaton that consists of cycles (with one initial state in each) of length greater then one and at most one chain with one initial state. Thus the number of initial states cannot be greater than $\lceil n/2 \rceil$. The amount of states contained in Conf(t) is not increasing in time. It means that automaton cannot reach its subsets containing more than $\lceil n/2 \rceil$ states. The number of k-subsets (where $k \in \{0,...,\lceil n/2 \rceil\}$) is approximately 2^{n-1} or 1/2 of all n-NFA1's subsets.

Proof T1
Lemmas T1[L1], T1[L2] and T1[L3] cover all cases and in each case at least ~1/4 of all configurations remains unreachable. Thus T1 has been proved.

4. Unreachable Configurations of Fixed Size

Theorem T2 It is not possible to construct n-NFA1 ($n \geq 4$) that could reach all its k-subsets for arbitrary chosen $k \in \{2,...,n-2\}$.

Statement T2[S1] For each n-NFA1 ($n \geq 4$) and for each k-subset ($k \in \{2,...,n-2\}$) two states $p_1,p_2 \in Q_n$ that belong to this subset and two states $r_1,r_2 \in Q_n$ that does not belong to this subset can be found. This is because $k \geq 2$ and $k \leq n-2$.

Statement T2[S2] For each n-NFA1 ($n \geq 4$), each $k \in \{2,...,n-2\}$ and each quadruple of states $p_1,p_2,r_1,r_2 \in Q_n$ k-subset $W \subseteq Q_n$ (such as $p_1,p_2 \in W$, but $r_1,r_2 \notin W$) can be found (follows from T2[S1]).

Lemma T2[L1] If n-NFA1 ($n \geq 4$) contains ε-transition then for each $k \in \{2,...,n-2\}$ unreachable k-subset can be found.

Proof T2[L1] If n-NFA1 contains ε-transition then a pair of states $q_p,q_r \in Q_n$ can be found such as q_p-ε->q_r. It means that a subset that includes q_p, but does not include q_r will be unreachable. According to T2[S2] for each $k \in \{2,...,n-2\}$ such k-subset can be found (for instance, by taking $p_1=q_p$ and $r_1=q_r$).

From now on only n-NFA's without ε-transitions will be discussed. Theorem T2 will be proved by using *reductio ad absurdum*. We assume that it is possible to construct n-NFA1 required in T2 and then examine the various properties this n-NFA1 should have and derive a contradiction. The proof will be divided into two parts. In the first part (T2a) the assumption that the n-NFA1 contains a cycle of length greater than three will be made. In the second part (T2b) n-NFA1s that does not contain a cycle of length greater than three will be examined.

Theorem T2a It is not possible to construct n-NFA1 ($n \geq 4$) containing a cycle of length $c \geq 4$ that could reach all its k-subsets for arbitrary chosen $k \in \{2,...,n-2\}$.

Statement T2a[S1] For each n-NFA1 ($n \geq 4$) containing a cycle C of length $|C| \geq 4$, for each $k \in \{2,...,n-2\}$ and each quadruple of states $p_1,p_2,r_1,r_2 \in C$ k-subset W (such as $p_1,p_2 \in W$, but $r_1,r_2 \notin W$) can be found (follows from T2[S2]).

Lemma T2a[L1] The amount of simultaneously reachable states in each NFA1's cycle cannot decrease in time.

Proof T2a[L1] For each m different states $q_{a1},q_{a2},...,q_{am} \in C$ one can find m different states $q_{b1},q_{b2},...,q_{bm} \in C$ such as $\forall j \in \{1,...,m\}: q_{aj} \dashrightarrow q_{bj}$ where $m \in \{1,...,c\}$. It can be done by choosing $bj=aj \oplus 1$. If $|Conf(t) \cap C|=m$ then

$|Conf(t+1) \cap C| \geq m$ thus the amount of reachable states in a cycle cannot decrease with time.

Lemma T2a[L2] If NFA1 contains a cycle C then for each pair of reachable subsets $S_1=Conf(t_1)$ and $S_2=Conf(t_2)$, where $t_2>t_1$ and $|S_1 \cap C|=|S_2 \cap C|$, some d can be found such as $S_2 \cap C$ can be obtained by rotating $S_1 \cap C$ in the direction of cycle's arrows by d units.

Proof T2a[L2] Let us denote the elements of the set $Conf(t_1) \cap C$ by $q_{a1},q_{a2},...,q_{am} \in C$ ($m \leq |C|$). As in the proof of the T2a[L1], for each w>0 m different states $q_{b1},q_{b2},...,q_{bm} \in C$ such as $\{q_{b1},q_{b2},...,q_{bm}\} \subseteq Conf(t_1+w)$ can be found by choosing $bj=aj \oplus w$. S_2 can be obtained by rotating S_1 by $d=t_2-t_1$ in the direction of cycle's arrows. This is because the amount of reachable states in the cycle is growing with respect to time (T2a[L1]) and $|S_1 \cap C|=|S_2 \cap C|$.

Proof T2a Let us choose two k-subsets S_1 and S_2. So that states belonging to the set $S_1 \cap C$ are placed together and the four states mentioned in the statement T2[S2] are placed together in the following sequence: $p_1p_2r_1r_2$ (see Fig. 3). But states that states belong to the set $S_2 \cap C$ are not placed together as four mentioned states are sequence $p_1r_1p_2r_2$ (Fig. 4).

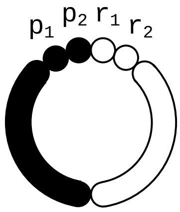

Fig. 3

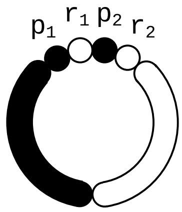

Fig. 4

Here two arrangements in the cycle has been gained that cannot be gained one from another by rotation. Thus these both k-subsets are mutually exclusive – if one of these k-subsets can be reached then other cannot and vice versa (follows from T2a[L2]). It means that for each n-NFA1 ($n \geq 4$) containing cycle of length greater than three for each $k \in \{2,...,n-2\}$ there can be found at least two mutually exclusive k-subsets. In other words, there does not exist k from interval $\{2,...,n-2\}$ such as any NFA1 could reach all its k-subsets.

Now NFA1s that does not contain cycle of length greater than three will be examined.

Theorem T2b It is not possible to construct n-NFA1 ($n \geq 4$) that does not contain a cycle of length $c \geq 4$ and could reach all its k-subsets for arbitrary chosen $k \in \{2,...,n-2\}$.

Lemma T2b[L1] If n-NFA1 does not contain any cycle one can find such state $r \in Q_n$ which cannot be reached more than once.

Proof T2b[L1] At first we will prove that there is a state in n-NFA1 that does not have any incoming arrow. Let us assume the opposite – each state has at least one incoming arrow but automaton does not contain any cycle. For arbitrary chosen state $q_1 \in Q_n$ one can find $q_2 \in Q_n$ such as q_2--->q_1 and $q_2 \neq q_1$ (if $q_2=q_1$ then there would be a cycle in automaton). Similarly $q_3 \in Q_n$ can be found such as q_3--->q_2 where q_3 is some state we have not dealt before i.e. $q_3 \neq q_2$ and $q_3 \neq q_1$. This is because there is not any cycle in the automaton examined. Proceeding in similar manner we will finally come to $q_n \in Q_n$. Yet for q_n it will not be possible to find previously unencoutered $q_x \in Q_n$ such as q_x--->q_n. Here the contradiction is derived, as there is some state r, which does not have incoming arrow. If the state r is initial then it can be reached only once, otherwise it cannot be reached at all.

Statement T2b[S1] If n-NFA1 ($n \geq 4$) does not contain any cycle then for each $k \in \{2,...,n-2\}$ unreachable k-subset can be found (this is because according to T2[S2] for each $k \in \{2,...,n-2\}$ at least two k-subsets containing state r mentioned in lemma T2b[L1] can be found).

Lemma T2b[L2] If n-NFA1 ($n \geq 4$) contains a cycle of length 1 (a state q pointing to itself) then for each $k \in \{2,...,n-2\}$ unreachable k-subset can be found.

Proof T2b[L2] If q is initial state then it is impossible to reach a k-subset to which q does not belong (the existence of such k-subset follows from statement T2[S2]). If q is not initial state then there should exist $p \in Q_n$ such as p--->q as otherwise none k-subset containing q would be reachable. In this case it is not possible to reach more than one k-subset containing p, but not q. Yet there will be at least two such k-subsets (follows from T2[S2]: $p_1=p$, $r_1=q$, but p2 and r2 can be chosen in at least two different ways as there are at least 4 states in automaton)

Lemma T2b[L3] If none of n-NFA's ($n \geq 4$) initial states belongs to some cycle then for each $k \in \{2,...,n-2\}$ unreachable k-subset can be found.

Proof T2b[L3] Let us construct a new NFA1 A' that contains all initial states of given n-NFA1 A. There will be a transition from state q to state p in A' iff state p can be reached from q in the given automaton A. There will not be any cycles in A' not having been already in A. Thus there will be a state r in A' which cannot be reached more than once (follows from T2b[L1]). It can be seen that also the corresponding state in A will not be reachable more than once. In T2[S2] we concluded that for each $k \in \{2,...,n-2\}$ and each state r there is more than one k-subset that contains r. Thus for each k at least one k-subset will be unreachable.

Lemma T2b[L4] If there exists initial state which belongs to some cycle C_2 of length 2 in n-NFA1 ($n \geq 4$) then for each $k \in \{2,...,n-2\}$ at least one unreachable k-subset can be found.

Proof T2b[L4] According to T2a[L1] the amount of simultaneously reachable states in a cycle cannot decrease. Thus for each $k \in \{2,...,n-2\}$ k-subset that does not contain any of C_2 states will be unreachable. The existence of such k-subset follows from T2[S2].

Lemma T2b[L5] If there is more than one initial state in some n-NFA1's ($n \geq 4$) cycle C_3 of length three then for each $k \in \{2,...,n-2\}$ unreachable k-subset can be found.

Proof T2b[L5] By repeating similar arguments as in T2b[L4] it can be seen that for each $k \in \{2,...,n-2\}$ those k-subsets that does not contain two C_3 states will be unreachable.

Lemma T2b[L6] If there is initial state in some n-NFA1's ($n \geq 5$) cycle C_3 of length three then for each $k \in \{2,...,n-3\}$ ($k \neq n-2$) unreachable k-subset can be found.

Proof T2b[L6] In this case for arbitrary chosen three states $r_1,r_2,r_3 \in Q_n$ and for each $k \in \{2,...,n-3\}$ at least one k-subset that does not contain any of states r_1,r_2,r_3 can be found. It can be easily concluded that for each $k \in \{2,...,n-3\}$ k-subset that does not contain any of three states belonging to C_3 cannot be reached (here r_1,r_2,r_3 are chosen from cycle C_3).

Proof T2b According to T2b[S1] automaton contains at least one cycle. T2a does not allow cycles with more than 4 states. Cycles consisting of only one state are denied by T2b[L2]. Thus automaton contains at least one cycle of length two or three. According to T2b[L3] at least one of initial states must be in a cycle. Let us denote this cycle by C. According to T2b[L4] the length of C cannot be 2. Thus C consists of 3 states. There can be only one initial state in cycle C (follows from T2b[L5]). If there exists k, such as all k-subsets can be reached then according to T2b[L6] $k \notin \{2,...,n-3\}$ thus k could be only n-2. It means that all states that do not belong to cycle C have to be initial. Three different cases can be distinguished:

a) Cycle C shares two states with cycle C_2 of length 2 (see Fig. 5). There cannot be other kind of cycles of length 2 in the automaton otherwise there would be initial state in a cycle of length 2, but it is denied by T2b[L4].

b) Cycle C shares two states with at least one cycle of length 3 (see Fig. 6). Let us denote this cycle by C_3. There cannot be other kind of cycles of length 3 in the automaton. Otherwise there would be two initial states in a cycle of length 3, but it is denied by T2b[L5]. There cannot be cycle of length 3, which contains the same states as the cycle C, but with arrows pointing in opposite direction. Then a cycle of length 2 containing initial state would be formed.

c) Cycle C does not share any of its elements with other cycles (see Fig. 7).

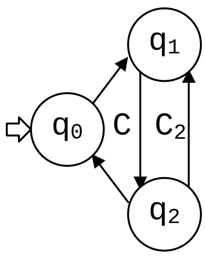

Fig. 5

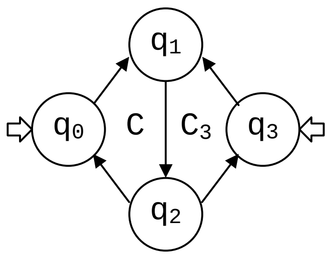

Fig. 6

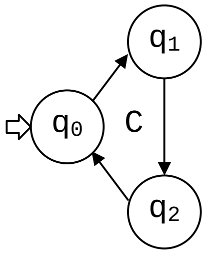

Fig. 7

For all three cases all states of automaton other than q_1 and q_2 are initial.

a) Let us look at two (n-2)-subsets S_1 and S_2 such as $S_1 \cap C = q_2$ and $S_2 \cap C = \{q_0, q_2\}$. These two subsets are mutually exclusive. S_1 cannot be reached after reaching S_2, because $|S_1 \cap C| < |S_2 \cap C|$ (according to T2a[L1] the amount of reachable states in each cycle cannot decrease). Yet also S_2 cannot be reached after reaching S_1 (if $Conf(t) \cap C = S_1$ then $(Conf(t+1) \cap C) \supseteq \{q_0, q_1\}$, $(Conf(t+2) \cap C) \supseteq \{q_1, q_2\}$ and $\forall d \geq 3: (Conf(t+d) \cap C) \supseteq \{q_0, q_1, q_2\}$).

b) In this case let us look at two (n-2)-subsets S_1 and S_2. S_1 contains q_1 and q_0, but does not contain q_2 and q_3. S_2 contains q_2 and q_3, but does not contain q_0 and q_1. These two subsets are mutually exclusive. This is because $|S_1 \cap C| = 2$ and $|S_1 \cap C_3| = 1$, but $|S_2 \cap C| = 1$ and $|S_2 \cap C_3| = 2$. According to T2a[L1] the amount of simultaneously reachable states in each cycles cannot decrease. This condition cannot be met no matter in what order S_1 and S_2 are reached.

c) In this case C is the only cycle in automaton. Let us construct new automaton using the same principle as in the proof of lemma T2b[L3] (to avoid cycles we will neglect transition q_0--->q_0). There will be no cycles in the automaton gained otherwise original automaton besides cycle C would contain at least one other cycle. Thus there will exist a state r which will not be reachable more than once in automaton constructed (follows from T2b[L1]) and also in original automaton. If $r = q_0$ then cycle C is not reachable from states which does not belong to it. Thus the number of simultaneously reachable states in C cannot increase, thus subsets containing more than one state of C will not be reachable. If $r \neq q_0$ then r cannot be reached more than once (but according to T2[S2] it belongs to at least two different (n-2)-subsets).

Proof T2 Theorems T2a and T2b together form T2. Thus by proving T2a and T2b, we have also proved T2.

5. References

[1] Andris Ambainis, Rusins Freivalds: 1-way quantum finite automata: strengths, weaknesses and generalizations CoRR quant-ph/9802062: (1998)
[2] Andris Ambainis, Richard F. Bonner, Rusins Freivalds, Marats Golovkins, Marek Karpinski: Quantum Finite Multitape Automata. SOFSEM 1999: 340-348
[3] Wilfried Brauer: On Minimizing Finite Automata. Bulletin of the EATCS 35: 113-116 (1988)
[4] Rusins Freivalds: Probabilistic Machines Can Use Less Running Time. IFIP Congress 1977: 839-842
[5] Rusins Freivalds: An answer to an open problem. Bulletin of the EATCS 23: 31-32 (1984)
[6] Eitan Gurari. An Introduction to the Theory of Computation. Computer Science Press, 1989.