Endmarkers Can Make a Difference

Moshe Y. Vardi*
IBM Almaden Research Center

Abstract

We show that the nonemptiness problem for two-way automata with only one endmarker over unary alphabets is complete for nondeterministic logarithmic space. This should be contrasted with the corresponding problem for two-way automata with two endmarkers, which is known to be NP-complete.

1 Introduction

Rabin and Scott [RS59] introduced two-way finite automata, which are allowed to move in both directions along their input tape. One may ask whether adding endmarkers to input tapes will cause an increase in expressive power of two-way automata, because of the added ability of the automata to go to the beginning or the end of the tape checking for a certain property. As was observed by Shepherdson [Sh59], this is not the case, since two-way automata are equivalent to one-way automata, and the latter can simply guess the beginning and end of the input tape.

We show here that while endmarkers do not add any expressive power, they can cause an increase in computational complexity. We consider the nonemptiness problem for two-way automata over unary alphabets. It is known that for two-way automata with two endmarkers the problem is NP-complete [Ga76]. We show here that with only one endmarker the problem is complete for nondeterministic logarithmic space.

Interestingly, the new result was motivated by a practical application related to optimization of database logic programs. In that application, automata over unary alphabets (with only one endmarker) are used to detect certain mappings between database queries. See [Va88] for details.

*Part of the research reported here was done while the author was visiting the Weizmann Institute of Science, Rehovot, Israel. Address: IBM Almaden Research K53-802, 650 Harry Rd., San Jose, CA 95120-6099.

2 Basic Definitions

A (nondeterministic) *two-way automaton* $A = (\Sigma, S, S_0, \rho, F)$ consists of an alphabet Σ, a finite set of states S, a set of initial states $S_0 \subseteq S$, a transition function $\rho : S \times \Sigma \to 2^{S\times\{-1,0,1\}}$, and a set of accepting states $F \subseteq S$. Intuitively, a transition indicates not only the new state of the automaton, but also whether the head should move left, right, or stay in place.

A *configuration* of A is a member of $S \times I\!N$, i.e., a pair consisting of a state and a "position". A *run* is a sequence of configurations, i.e., an element of $(S \times I\!N)^\star$. The run $(s_0, j_0), \ldots, (s_m, j_m)$ is a run of A on a word $w = a_0 \ldots a_{n-1}$ in $\Sigma^\star$ if $0 \leq j_m \leq n$, and for all i, $0 \leq i < m$, we have that $0 \leq j_i < n$, and there is some $(t, k) \in \rho(s_i, a_{j_i})$ such that $s_{i+1} = t$ and $j_{i+1} = j_i + k$. This run is *accepting* if $s_0 \in S_0$, $j_0 = 0$, $j_m = n$, and $s_m \in F$. A *accepts* w if it has an accepting run on w.

The set of words accepted by A is denoted by $L(A)$. The *nonemptiness problem* for two-way automata is to determine, given a two-way automaton A, whether $L(A)$ is nonempty. It is known that the nonemptiness problem for two-way automata is PSPACE-complete [Ga76].

Let b and e be fixed symbols, designating left and right endmarkers, respectively. A *two-way automaton with endmarkers* $A = (\Sigma, S, S_0, \rho, F)$ consists of an alphabet Σ that does not contain b or e, a finite set of states S, a set of initial states $S_0 \subseteq S$, a transition function $\rho : S \times (\Sigma \cup \{b, e\}) \to 2^{S\times\{-1,0,1\}}$, and a set of accepting states $F \subseteq S$. A accepts a word $w \in \Sigma^*$ if it has an accepting run over bwe (i.e., w with b concatenated from the left and e concatenated from the right). A two-way automaton with a *left* (resp., *right*) endmarker has only the endmarker b (resp., e). It accepts a word $w \in \Sigma^*$ if it has an accepting run over bw (resp., we).

The nonemptiness problem for two-way automata with endmarkers is PSPACE-complete; if restricted to unary alphabets (i.e., alphabets consisting of a single symbol), then the problem is NP-complete [Ga76].

3 Automata with No Endmarker

We first consider automata with no endmarker.

Proposition 3.1: *Let A be a k-state two-way automaton over a unary alphabet $\Sigma = \{a\}$. If $L(A)$ is nonempty, then $a^n \in L(A)$ for some $n < 2k$.*

Proof: Let $A = (\Sigma, S, S_0, \rho, F)$. Let $(s_0, j_0), \ldots, (s_m, j_m)$ be an accepting run of A of shortest length. Let $j_m = n$ (i.e., the length of the input word is n). Define $H(i) = \min\{j_p \,|\, i \leq p \leq m\}$ for $0 \leq i \leq m$. Intuitively, $H(i)$ is the leftmost position reached at or after the configuarion (s_i, j_i).

Claim 1: $j_i - H(i) \leq k$ for $0 \leq i \leq m$. Suppose to the contrary that there exist some i such that $j_i - H(i) > k$. There there are p, q such that $i \leq p < q \leq m$, $s_p = s_q$, and

$j_p > j_q$. But then, for $d = j_p - j_q > 0$, we have that $(s_0, j_0), \ldots, (s_p, j_p), (s_{q+1}, j_{q+1} + d), \ldots, (s_m, j_m + d)$ is an accepting run of A of length $m - d < m$ (on a word of length $n + d > n$) – contradicting our assumption on the minimality of m.

Claim 2: $n < 2k$. Suppose to the contrary that $n \geq 2k$. Define $I(j) = \min\{p \mid j_p = k+j\}$ for $0 \leq j \leq k$. Intuitively, $I(j)$ is a pointer to the first configuarion to reach the position $k + j$. Clearly, $I(j) < I(j + 1)$ for $0 \leq j < k$. Thus, there are p, q such that $0 \leq p < q \leq k$ and $s_{I(p)} = s_{I(q)}$. Let $d = j_{I(q)} - j_{I(p)} = q - p > 0$. Then we have that $(s_0, j_0), \ldots, (s_p, j_p), (s_{q+1}, j_{q+1} - d), \ldots, (s_m, j_m - d)$ is an accepting run of A of length $m - d < m$ (on a word of length $n - d < n$). This run is legal, since $H(I(q)) - d \geq j_{I(q)} - k - d = k + q - k - d = p \geq 0$, by Claim 1. We have contradicted our assumption on the minimality of m, so we must have $n < 2k$. ∎

4 Automata with Right Endmarkers

Let $A = (\Sigma, S, S_0, \rho, F)$ be a two-way automaton with a right endmarker over the unary alphabet $\Sigma = \{a\}$. Let $R_n \subseteq S^2$, $n > 0$, be the set of state pairs s, t such that there is a run $(s_0, j_0), \ldots, (s_m, j_m)$ of A on a^n, where $j_0 = j_m = n - 1$, $s_0 = s$, and $s_m = t$. That is, $(s, t) \in R_n$ if A has a run on a^n starting on the right in state s and ending on the right in state t.

We now prove several properties of the relations R_n, $n \geq 0$.

Lemma 4.1: *$(s, t) \in R_n$ if and only if there is a sequence $s_0, \ldots, s_{m-1}$ such that $s = s_0$, $t = s_{m-1}$, and for all i, $0 \leq i < m - 1$ we have that either*

- *$(s_{i+1}, 0) \in \rho(s_i, a)$, or*
- *there are states s'_i and s''_i such that $(s'_i, -1) \in \rho(s_i, a)$, $(s'_i, s''_i) \in R_{n-1}$, and $(s_{i+1}, 1) \in \rho(s''_i, a)$.*

Proof: It is easy to see that the condition is sufficient. For the converse, assume that $(s, t) \in R_n$. Then there is a run $(s_0, j_0), \ldots, (s_m, j_m)$ of A on a^n, where $j_0 = j_m = n - 1$, $s_0 = s$, and $t_0 = t$. Let k, l be such that $k < l$, $j_k = j_l = n - 1$, and $j_{k'} < n - 1$ for $k < k' < l$. If $l = k + 1$, then we must have $(s_l, 0) \in \rho(s_k, a)$. Otherwise, $l > k + 1$, in which case $(s_{k+1}, -1) \in \rho(s_k, a)$, $(s_l, 1) \in \rho(s_{l-1}, a)$, and $(s_{k+1}, j_{k+1}), \ldots, (s_{l-1}, j_{l-1})$ is a run of A on a^{n-1}, where $j_{k+1} = j_{l-1} = n - 2$. Thus, $(s_{k+1}, s_{l-1}) \in R_{n-1}$. The claim follows. ∎

Lemma 4.2: *$R_n \subseteq R_{n+1}$ for $n \geq 0$.*

Proof: Let $(s, t) \in R_n$. Then there is a run $(s_0, j_0), \ldots, (s_m, j_m)$ of A on a^n, where $j_0 = j_m = n - 1$, $s_0 = s$, and $t_0 = t$. Clearly, $(s_0, j_0 + 1), \ldots, (s_m, j_m + 1)$ is a run of A on a^{n+1}, where $j_0 = j_m = n$, $s_0 = s$, and $t_0 = t$. The claim follows. ∎

Lemma 4.3: *If $R_n = R_{n+1}$, then $R_{n+1} = R_{n+2}$.*

Proof: By Lemma 4.2, we have that $R_{n+1} \subseteq R_{n+2}$. It remains to prove the inverse containment.

Let $(s, t) \in R_{n+2}$. By Lemma 4.1, there is a sequence $s_0, \ldots, s_{m-1}$ such that $s = s_0$, $t = s_{m-1}$, and for all i, $0 \leq i < m - 1$ we have that either

- $(s_{i+1}, 0) \in \rho(s_i, a)$, or
- there are states s'_i and s'_{i+1} such that $(s'_i, -1) \in \rho(s_i, a)$, $(s'_i, s'_{i+1}) \in R_{n+1}$, and $(s_{i+1}, 1) \in \rho(s'_{i+1}, a)$.

Since $R_n = R_{n+1}$, we can replace R_{n+1} by R_n in the condition. It follows that $(s, t) \in R_{n+1}$. ∎

From Lemma 4.2 and 4.3 we get:

Lemma 4.4: *If $|S| = k$, then $R_n = R_{k^2}$ for all $n \geq k^2$.*

The R_n's can be used to give a "one-way" characterization of acceptance (cf. [Sh59]).

Lemma 4.5: *A accepts a^n if and only if there is a sequence $t_0, \ldots, t_m$, $m > n$, such that*

1. *$t_0 \in S_0$ and $t_m \in F$,*
2. *for $0 \leq i \leq n - 1$, there is a state t'_i such that $(t_i, t'_i) \in R_{i+1}$ and $(t_{i+1}, 1) \in \rho(t'_i, a)$,*
3. *for $n \leq i < m - 1$, either $(t_{i+1}, 0) \in \rho(t_i, e)$, or there are states t'_i and t''_i such that $(t'_i, -1) \in \rho(t_i, e)$, $(t'_i, t''_i) \in R_n$, and $(t_{i+1}, 1) \in \rho(t''_i, a)$,*
4. *$(t_m, 1) \in \rho(t_{m-1}, e)$.*

Proof: It is easy to see that the condition is sufficient. For the converse, assume that A accepts a^n. Then there is an accepting run $(s_0, j_0), \ldots, (s_p, j_p)$ of A on $a^n e$. For each i, $0 \leq i < n$, let k_i be minimal such that $j_{k_i} = i$, and let t_i be s_{k_i}. Also, let $t_n, \ldots, t_{m-1}$ be the longest sequence $s_{l_1}, \ldots, s_{l_{m-n}}$, where $j_{l_i} = n$ for $1 \leq i \leq m - n$. Finally, let t_m be s_p. Clearly, $t_0 \in S_0$, $(t_m, 1) \in \rho(t_{m-1}, e)$, and $t_m \in F$. For $0 \leq i < n$, let t'_i be $s_{k_{i+1}-1}$. Clearly, $(t_i, t'_i) \in R_{i+1}$ and $(t_{i+1}, 1) \in \rho(t'_i, a)$. Finally, for $1 \leq i < m - n$, either $l_{i+1} = l_i + 1$, in which case $(t_{i+n}, 0) \in \rho(t_{i+n-1}, e)$, or $(s_{l_i+1}, -1) \in \rho(s_{l_i}, e)$, $(s_{l_i+1}, s_{l_{i+1}-1}) \in R_n$, and $(s_{l_{i+1}}, 1) \in \rho(s_{l_{i+1}-1}, a)$. ∎

From Lemmas 4.5 and 4.4 it follows that there are "short witnesses" for nonemptiness.

Lemma 4.6: *Suppose A has k states. If $L(A)$ is nonempty, then $a^n \in L(A)$ for some $n \leq k^2 + k$.*

Proof: Let n be minimal such that $a^n \in L(A)$. Consider the sequence $t_0, \ldots, t_m$ guaranteed to exist by Lemma 4.5. Suppose $n > k^2 + k$. Then there are p, q such that $k^2 + 1 \leq p < q \leq k^2 + k + 1$ and $t_p = t_q$. Since $R_p = R_q$ by Lemma 4.4, the sequence $t_0, \ldots, t_{p-1}, t_q, \ldots, t_m$ satisfies the condition of Lemma 4.5 for a^{n+p-q}. It follows that $a^{n+p-q} \in L(A)$ - contradicting the minimality of n. ∎

5 Automata with Left Endmarkers

Let $A = (\Sigma, S, S_0, \rho, F)$ be a two-way automaton with a left endmarker over the unary alphabet $\Sigma = \{a\}$. Let $\bar{R}_n \subseteq S^2$, $n > 0$, be the set of state pairs s, t such that there is a run $(s_0, j_0), \ldots, (s_m, j_m)$ of A on a^n, where $j_0 = j_m = 0$, $s_0 = s$, and $s_m = t$. That is, $(s, t) \in \bar{R}_n$ if A has a run on a^n starting on the left in state s and ending on the left in state t.

The relations $\bar{R}_n$, $n \geq 0$, are analogous to the relations R_n, $n \geq 0$, of the previous section. As in the previous section, we can prove the following property.

Lemma 5.1: *If $|S| = k$, then $\bar{R}_n = \bar{R}_{k^2}$ for all $n \geq k^2$.*

As in the previous section, the $\bar{R}_n$'s can be used to give a "one-way" characterization of acceptance.

Lemma 5.2: *A accepts a^n if and only if there is a sequence $t_0, \ldots, t_m$, $m > n$, such that*

1. *$t_0 \in S_0$ and $t_m \in F$,*
2. *for $0 \leq i < m - n - 1$, either $(t_{i+1}, 0) \in \rho(t_i, b)$, or there are states t'_i and t''_i such that $(t'_i, 1) \in \rho(t_i, b)$, $(t'_i, t''_i) \in \bar{R}_n$, and $(t_{i+1}, -1) \in \rho(t''_i, a)$,*
3. *there is a state t'_{m-n-1} such that $(t'_{m-n-1}, 1) \in \rho(t_{m-n-1}, b)$, and $(t'_{m-n-1}, t_{m-n}) \in \bar{R}_n$.*
4. *for $m - n \leq i < m - 1$, there is a state t'_i such that $(t'_i, 1) \in \rho(t_i, a)$, and $(t'_i, t_{i+1}) \in \bar{R}_{m-i-1}$.*
5. *$(t_m, 1) \in \rho(t_{m-1}, a)$.*

Proof: It is easy to see that the condition is sufficient. For the converse, assume that A accepts a^n. Then there is an accepting run $(s_0, j_0), \ldots, (s_p, j_p)$ of A on ba^n. Let $t_0, \ldots, t_{m-n-1}$ be the longest sequence $s_{l_1}, \ldots, s_{l_{m-n}}$, where $j_{l_i} = 0$ for $1 \leq i \leq m - n$. Also, for each i, $m - n \leq i \leq m - 1$, let k_i be maximal such that $j_{k_i} = i - m + n + 1$, and let t_i be s_{k_i}. Finally, let t_m be s_p. The reader can verify that the sequence $t_0, \ldots, t_m$ satisfies the required conditions. ∎

From Lemmas 5.2 and 5.1 it follows that there are "short witnesses" for nonemptiness.

Lemma 5.3: *Suppose A has k states. If $L(A)$ is nonempty, then $a^n \in L(A)$ for some $n \leq k^2 + k$.*

Proof: Let n be minimal such that $a^n \in L(A)$. Consider the sequence $t_0, \ldots, t_m$ guaranteed to exist by Lemma 5.2. Suppose $n > k^2 + k$. Then there are p, q such that $0 \leq p < q \leq k$ and $t_{m-1-k^2-p} = t_{m-1-k^2-q}$. Since $\bar{R}_n = \bar{R}_{n+p-q}$ and $\bar{R}_{k^2+p} = \bar{R}_{k^2+q}$ by Lemma 5.1, the sequence $t_0, \ldots, t_{p-1}, t_q, \ldots, t_m$ satisfies the condition of Lemma 5.2 for a^{n+p-q}. It follows that $a^{n+p-q} \in L(A)$ - contradicting the minimality of n. ∎

6 Complexity

From Lemmas 4.6 and 5.3 we get:

Proposition 6.1: *Let A be a k-state two-way automaton with only one endmarker over a unary alphabet $\Sigma = \{a\}$. If $L(A)$ is nonempty, then $a^n \in L(A)$ for some $n \leq k^2 + k$.*

Our main result follows from Proposition 6.1.

Theorem 6.2: *The nonemptiness problem for two-way automata with one endmarker over unary alphabets is complete for nondeterministic logarithmic space.*

Proof: Hardness follows from the same result for one-way automata [Jo75]. It remains to show that the problem is solvable in nondeterministic logarithmic space.

Let A be a two-way automaton with k states over a unary alphabet $\Sigma = \{a\}$. By Proposition 6.1, if $L(A)$ is nonempty, then $a^n \in L(A)$ for some $n \leq k^2 + k$. To check whether $L(A)$ is nonempty, we guess a number $n \leq k^2 + k$, and then simulate A over ba^n or $a^n e$. At any point of the simulation we just have to remember n and some configuration (s_i, j_i). This requires only logarithmic space. ∎

Acknowledgements. I am grateful to the two referees for their helpful suggestions. In particular, both referees suggested Proposition 3.1 and its proof.

References

[Ga76] Galil, Z.: Hierarchies of complete problems. *Acta Informatica* 6(1976), pp. 77–88.

[Jo75] Jones, N.D.: Space-bounded reducibility among combinatorial problems. *J. Computer and System Sciences* 11(1975), pp. 68–85.

[RS59] Rabin, M.O., Scott, D.: Finite automata and their decision problems. *IBM J. Research and Development*, 3(1959), pp. 114–125.

[Sh59] Shepherdson, J.C.: The reduction of two-way automata to one-way automata. *IBM J. Research and Development*, 3(1959), pp. 199–201.

[Va88] Vardi, M.Y.: Decidability and undecidability results for boundedness of linear recursive queries. *Proc. 7th ACM Symp. on Principles of Database Systems*, 1988, pp. 341–351.