Magic Numbers and Subset Construction

Samik Datta
Sayantan Mahinder

Magic Number, α

- Iwama, Matsuura, Paterson defined a magic number as an integer α between n and 2^n (both inclusive) such that there is no minimal NFA of n states which require exactly α states in the minimal equivalent DFA.

- We know that n and 2^{n-1} are not magic numbers.
- Why? The division automaton, the DFA for the (n-1)th symbol from the RHS is 0.

- We will investigate the question, whether 2^n ,in particular, is a magic number? More optimistically … are there any magic numbers at all?

Fooling Set $\Im$ for language L

$$\Im = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$$

- Fooling set (pair of strings) satisfies the following 2 conditions
- 1. For all k, $x_k y_k \in L$
- 2.For all different i, j, at least one of the followings are satisfied (cross-over terms).

  $x_i y_j \notin L \qquad x_j y_i \notin L$

- Example: for $L = 1^k$ a fooling set is

$$\{(\varepsilon, 1^k), (1^1, 1^{k-1}), \ldots, (1^k, \varepsilon)\}$$

Minimality of NFA

- Lemma: If L is a regular language, them the # of states in a NFA accepting L is $\geq |\mathfrak{I}|$

- Proof outline: All the intermediate states reached after reading the first string (of the pair) in the fooling set are different. Prove using contradiction.

- Corollary: To prove that a given NFA for L with n states is minimal, we can demonstrate a fooling set of cardinality n.

A skeleton NFA

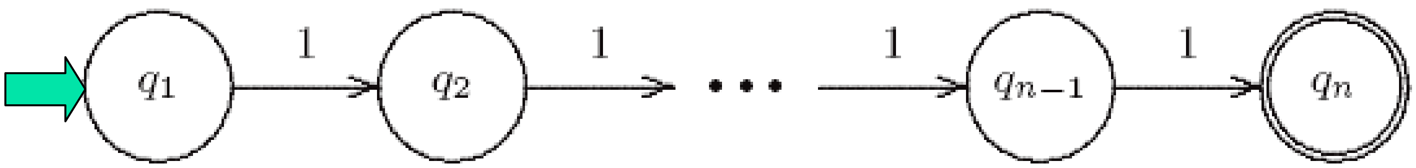

Fig. 1. Transitions on reading 1 of the NFA M

- The NFA M has n states and have only 2 restrictions on it's transition
- $1.\delta(q_i, 1) = \{q_{i+1}\}$ for all i=1,2, … ,n-1

  $2.\delta(q_n, 1) = \Phi$
- All the other transitions for the rest of the alphabet (except 1) are arbitrary.
- q_1 is the only start state, q_n is the only accepting state.

An useful theorem

- Theorem: For any NFA M satisfying 1 and 2, the following 2 facts hold good …
- 1. M is minimal among the NFA s accepting L(M).
- 2. The DFA consisting of the reachable states after the subset construction is minimal, too.

- Proof outline:
- (1) Show a fooling set of cardinality n. $\{(\varepsilon, 1^{n-1}), (1^1, 1^{n-2}), \ldots, (1^{n-1}, \varepsilon)\}$
- (2) The string 1^{n-i} leads a state of the DFA (obtained by subset construction) with q_i as an element to a final state and another state without q_i as an element to the non-final state. Therefore, there are no 2 equivalent states in the power set of Q which are reachable from the start state.

The bound is tight! A_k

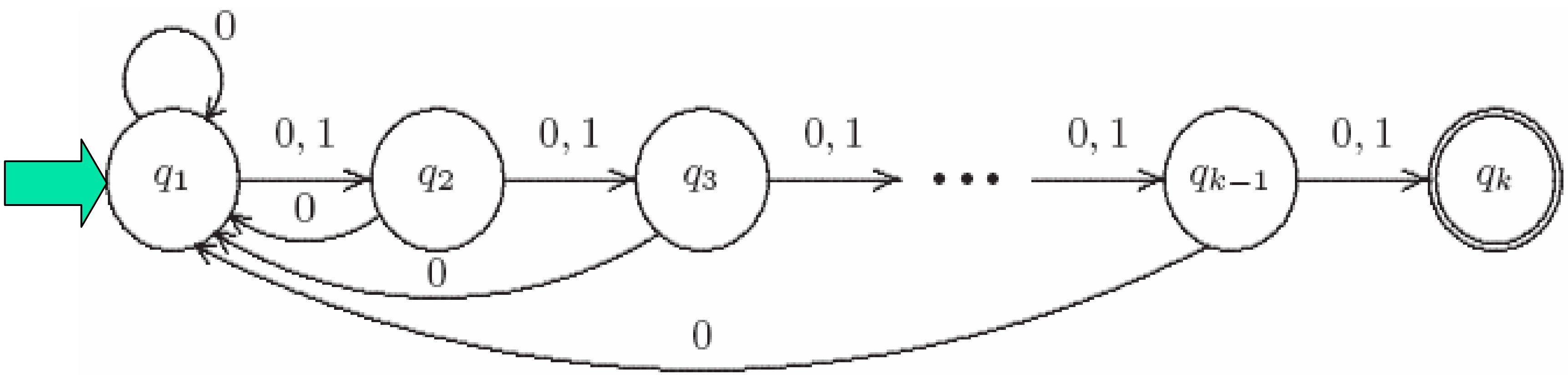

Fig. 3. The nondeterministic finite automaton A_k

- It is a variation of the "skeleton NFA" we considered in the last slide , having {0,1} as the alphabet, and the transitions on 0 defined as in the figure.
- Please note the back arrow, forward arrow labeled 0 from each state. What demands them to be present? …

But why $\Delta(A_k, k) = 2^k$?

- $\Delta(M, n)$ denotes the # of states in the minimal DFA equivalent to minimal NFA M with n states.
- Proof outline: To show that all the states in P(Q) are reachable in the subset construction, use induction on the cardinality of the set concerned.
- Basis: Cardinality 0,1: All k+1 such states are reachable.
- Hypothesis: All states with cardinality <= l-1 are reachable.
- Induction: To show that all the states with cardinality l are reachable, note that

$$\delta(\{q_{i_2-i_1}, q_{i_3-i_1}, \dots, q_{i_l-i_1}\}, 01^{i_1-1}) = \{q_{i_1}, q_{i_2}, \dots, q_{i_l}\}$$

Another family, B_k

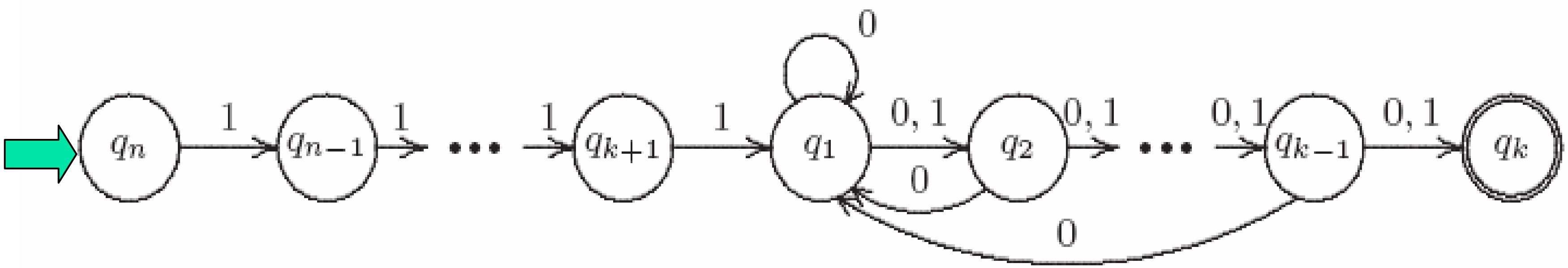

Fig. 4. The nondeterministic finite automaton B_k

$1 \leq k \leq n-1$

$$\Delta(B_k, n) = 2^k + n - k$$

Proof outline: Use the previous result … be careful to prove that no other state is reachable. Minimality follows from our good old lemma.

Yet another family! $M_{k,j}$

$$\Delta(M_{k,j},n) = 2^k + n - k + j \qquad 1 \le j \le 2^k - 1$$

- Trick: Take your alphabet to be big enough, consisting of $2^{n-1}+1$ letter, including 0,1.

Construction: start with B_k
Add transitions from the accepting state on letters $a_1, a_2, \dots, a_j$ (none 0,1) to the states $S_1, S_2, \dots, S_j$ where each such state is of the form $\wp(\{q_1, q_2, \dots, q_k\} \circ \{q_{k+1}\}$ except the $\{q_{k+1}\}$

Proof outline: It is easy to see all the newly added j states are reachable, but we have to be careful to show that no other state is reachable.

Magic Number is a Myth!

- The case, α = n is trivial
- Else α satisfies $2^k + n - k \leq \alpha < 2^{k+1} + n - (k+1)$

- In case when α is the left limit, consider B_k
- Else consider $M_{k,j}$

- If α is 2 power n, consider A_n

Wait a minute … what happens in small alphabet ?

- In the paper, Galina Jiraskova was able to give the proof of no magic number using 2n sized alphabet (unlike we did here for exponential order). But, it is not a construction, but an existence.

- In case of {0,1}, the same author proved that there is no magic number of $O(n^2)$

- But the question whether there are some magic numbers of $\omega(n^2)$ is still open.
- In case of {1}, Chrobak proved that no minimal NFA with n states needs $\omega(e^{\sqrt{n \ln n}})$ states in the minimal equivalent DFA.

The question whether there exist some magic # less than that is still open …

Any practical implications ?

- Yes, these bounds are necessary to analyze the algorithms involving the finite automata
- There is a field called the *state complexity theory*, which gives lower bound for minimum number of states needed to recognize certain regular languages, and other regular languages obtained by applying various operations like Reversal, Shuffle, Quotient, Prefix etc. on those regular languages.

References

- 0. Galina Jiraskova: Note on Minimal Finite Automata. Mathematical Institute, Slovak academy of sciences. Slovakia.
- 1. M. Chrobak: Finite automata and unary languages. Theoretical Computer Science

  47(1986), 149-158
- 2. J. Hromkoviˇc: Communication Complexity and Parallel Computing. Springer 1997
- K. Iwama. A. Matsuura and M. Paterson: A family of NFA's which need $2n - a$ deterministic states. Proc. MFCS'00, Lecture Notes in Computer Science 1893, Springer-Verlag 2000, pp.436-445
- F. Moore: On the bounds for state-set size in proofs of equivalence between deterministic, nondeterministic and two-way finite automata. IEEE Trans. Comput.

  C-20, pp.1211-1214, 1971

Thank you!

Questions