

# A LINEAR-TIME SIMULATION OF DETERMINISTIC $d$ -LIMITED AUTOMATA

ALEXANDER RUBTSOV

HIGHER SCHOOL OF ECONOMICS, MOSCOW, RUSSIA

DEVELOPMENTS IN LANGUAGE THEORY  
AUGUST 18, 2021

# PARSING OF FORMAL LANGUAGES

- CFL's can be parsed in  $O(n^\omega)$  [L. Valiant'75]
- LR-grammars can be parsed in  $O(n)$  and generate DCFL's [D. Knuth'65]
- Parsing Expression Grammars (PEG's) can be parsed in  $O(n)$  [A. Birman and J. Ullman'70, B. Ford'o4]

The formal language's class description is an open question

# THEORETICAL BOUNDS

- 2DPDA-recognizable languages are linear-time recognizable  
[S. Cook'70]
- **Conditional lower bounds.** Improvement of Valiant's parser leads to fast algorithms for various problems, e.g., Boolean matrix multiplication,  $k$ -Clique problem  
[L. Lee'02, A. Abboud et al.'15]

# $d$ -LIMITED AUTOMATA

**$d$ -Limited Automaton** is a Turing Machine that can change only the input cells in first  $d$  visits [T. Hibbard'67]

- 1-LA's recognize regular languages
- For  $d \geq 2$   $d$ -NLA's recognize CFL's
- 2-DLA's recognize exactly DCFL's

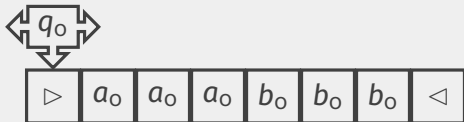
[T. Hibbard'67, G. Pighizzini and A. Pisoni'15]

- $d$ -DLA's recognizable languages ( $d$ -DCFL's) form a hierarchy

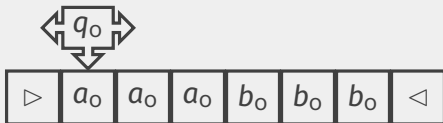
## Our Contribution

We provide a linear-time simulation algorithm for  $d$ -DLA's

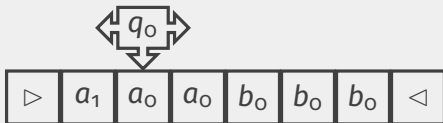
# EXAMPLE



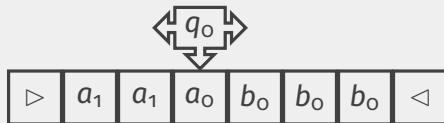
# EXAMPLE



# EXAMPLE

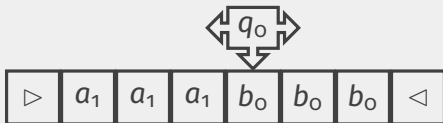


# EXAMPLE

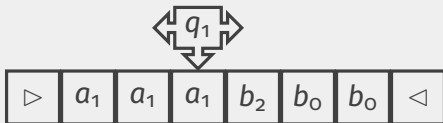




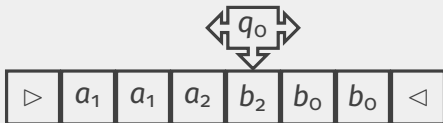
# EXAMPLE



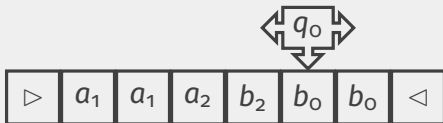
# EXAMPLE



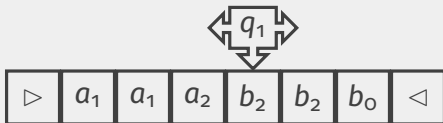
# EXAMPLE



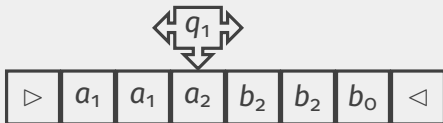
# EXAMPLE



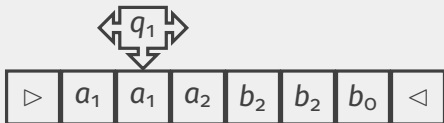
# EXAMPLE



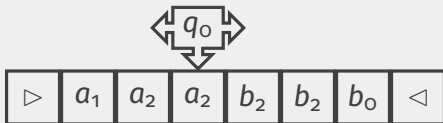
# EXAMPLE



# EXAMPLE

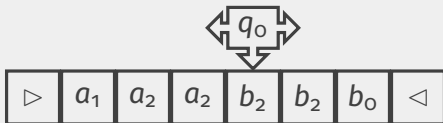


# EXAMPLE

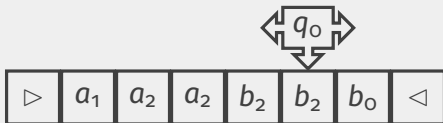




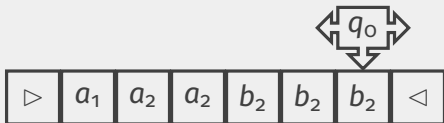
# EXAMPLE



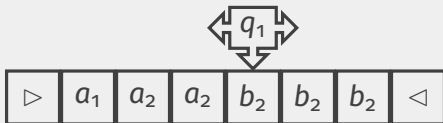
# EXAMPLE



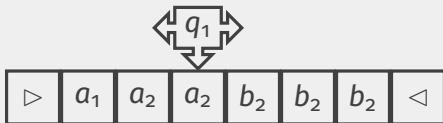
# EXAMPLE



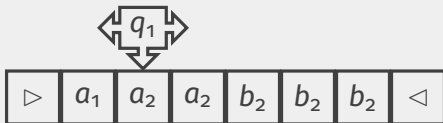
# EXAMPLE



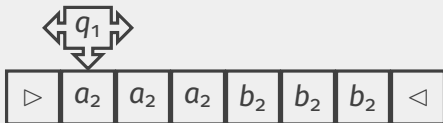
# EXAMPLE



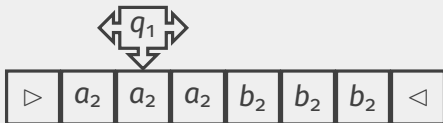
# EXAMPLE



# EXAMPLE

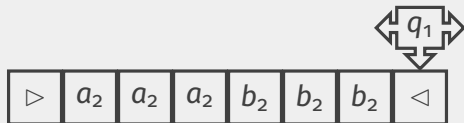


# EXAMPLE

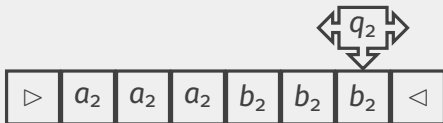




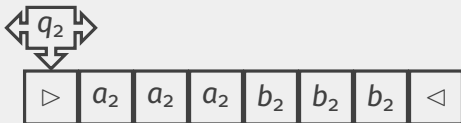
# EXAMPLE



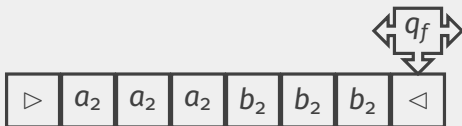
# EXAMPLE



# EXAMPLE



# EXAMPLE



# MORE INTUITION ON $d$ -DLA'S

Denote  $L_{\#} = \{\#a^n b^n c^m \mid n, m \geq 0\}$  and  $L_{\$} = \{\$a^m b^n c^n \mid n, m \geq 0\}$

- $L_{\#}$  and  $L_{\$}$  are DCFL's
- $(L_{\#} \cup L_{\$})^R$  is not a (2-)DCFL, but a 3-DCFL

# MORE INTUITION ON $d$ -DLA'S

Denote  $L_{\#} = \{\#a^n b^n c^m \mid n, m \geq 0\}$  and  $L_{\$} = \{\$a^m b^n c^n \mid n, m \geq 0\}$

- $L_{\#}$  and  $L_{\$}$  are DCFL's
- $(L_{\#} \cup L_{\$})^R$  is not a (2-)DCFL, but a 3-DCFL

$$L_{\#,\$} = \{a^n b^n c^m \mid n, m \geq 0\} \cup \{a^m b^n c^n \mid n, m \geq 0\}$$

- $L_{\#,\$}$  is an inherently ambiguous language
- $L_{\#,\$}$  is not a  $d$ -DCFL

**Theorem (T.Hibbard'67)**

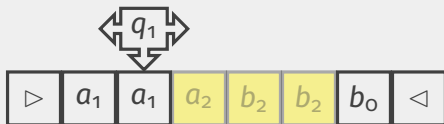
*Each  $d$ -DCFL is generated by an unambiguous CF-grammar*

# RELATED RESULTS

- A language recognizable in linear time by a (deterministic) TM is regular [F. Hennie'65]
- Each 1-DLA can be transformed into a linear-time TM [B. Guillon, L. Prigioniero'19]

# OUR CONSTRUCTION

- We modify Shepherdson's technique of simulation of 2DFA by 1DFA
- A segment of cells of rank  $d$  is replaced by a mapping that describes the behavior of  $d$ -DLA on the segment

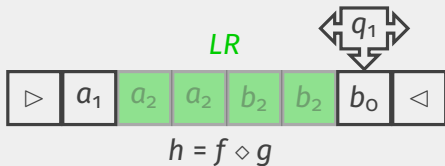
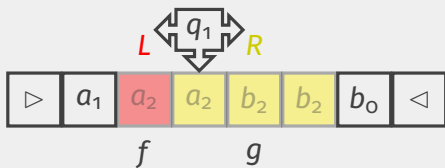


$$g : \vec{q}_0 \mapsto \vec{q}_0, \overleftarrow{q}_1 \mapsto \overleftarrow{q}_1, \overleftarrow{q}_2 \mapsto \overleftarrow{q}_2$$

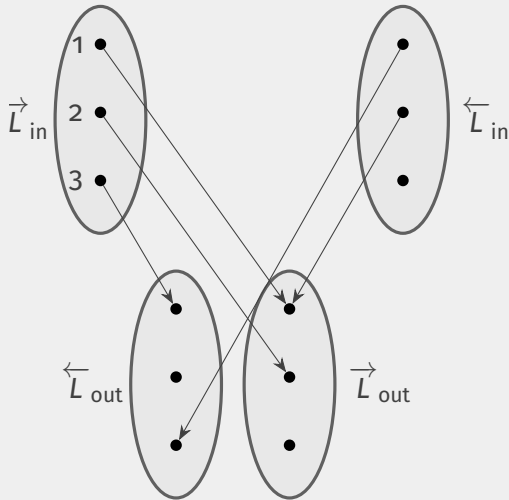
- We modify the  $d$ -DLA model: the tape of TM is replaced by the doubly-linked list



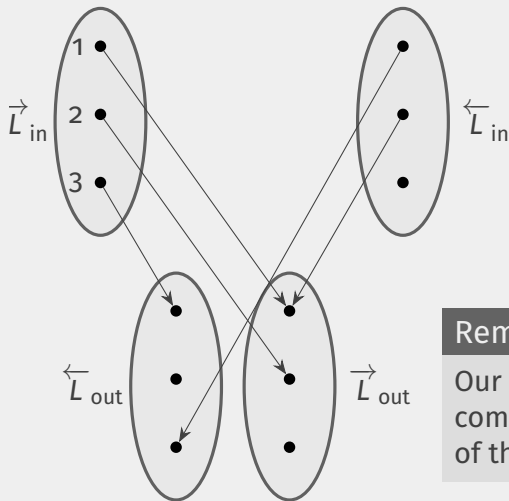
# “COMPOSITION” OF MAPPINGS



# MAPPING REPRESENTATION



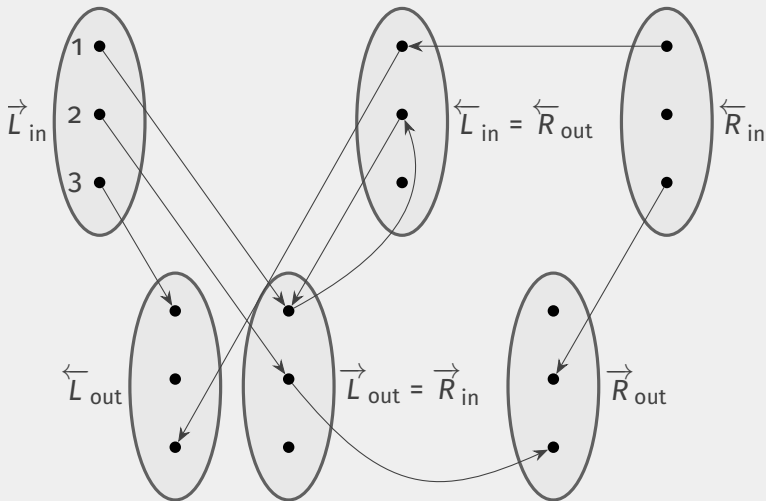
# MAPPING REPRESENTATION



## Remark

Our algorithm of “composition” computing is an effective variant of the Birget’s construction

# LINEAR-TIME "COMPOSITION" COMPUTING



# AMORTIZED ANALYSIS



# CONCLUSION

Denote by  $m$  the length of the  $d$ -DLA  $M$ 's description and by  $n$  the length of the input  $w$

## Our Results

- $O(mn)$  algorithm in RAM for the membership problem
- It is a linear-time simulation algorithm if  $M$  is fixed.
- A more careful analysis leads to the bound  $O(kn + m)$ , where  $k$  is the number of  $M$ 's states.

**Open Question:** Do PEG's generate all  $d$ -DCFL's?

**Semi-open Question:** Prove that  $d$ -DCFL's do not include the language of palindromes