

Limited Automata and Descriptive Complexity

Giovanni Pighizzini

(joint papers with Andrea Pisoni – DCFS 2013, NCMA 2013)

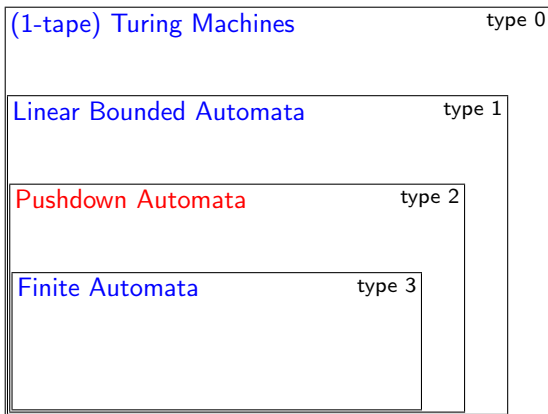
Dipartimento di Informatica
Università degli Studi di Milano, Italy

Riunione PRIN 2013

Milano, 28–29 novembre 2013



The Chomsky Hierarchy



Limited Automata [Hibbard'67]

One-tape Turing machines with restricted rewritings

Definition

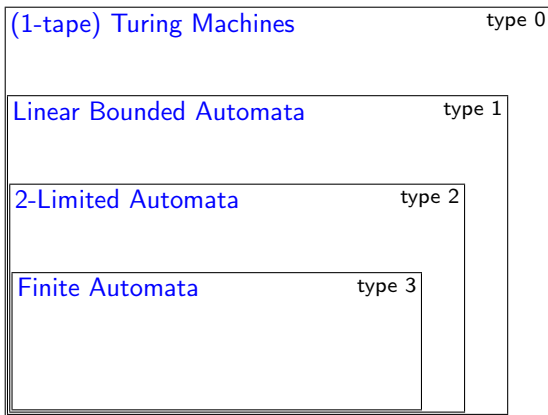
Fixed an integer $d \geq 1$, a *d-limited automaton* is

- ▶ a one-tape Turing machine
- ▶ which is allowed to rewrite the content of each tape cell *only in the first d visits*

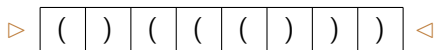
Computational power

- ▶ For each $d \geq 2$, *d-limited automata* characterize context-free languages [Hibbard'67]
- ▶ 1-limited automata characterize regular languages [Wagner&Wechsung'86]

The Chomsky Hierarchy



Example: Balanced Parentheses



- (i) Move to the right to search a closed parenthesis
- (ii) Rewrite it by #
- (iii) Move to the left to search an open parenthesis
- (iv) Rewrite it by #
- (v) Repeat from the beginning

Special cases:

- (i') If in (i) the right end of the tape is reached then scan all the tape and *accept* iff all tape cells contain #
- (iii') If in (iii) the left end of the tape is reached then *reject*

Each cell is rewritten only in the first 2 visits!

2-Limited Automata \rightarrow Pushdown Automata

Problem

How much it costs, in the description size, the simulation of 2-LAs by PDAs?

Our result

Exponential cost!
(optimal)

Deterministic case

- ▶ Determinism is preserved by the simulation *provided that the input of the PDA is right end-marked*
- ▶ Without end-marker: *double exponential* size for the simulation of D2-LAs by DPDA
- ▶ Conjecture: this cost cannot be reduced

CFLs \rightarrow 2-Limited Automata

New transformation based on:

Theorem ([Chomsky&Schützenberger'63])

Every context-free language $L \subseteq \Sigma^$ can be expressed as*

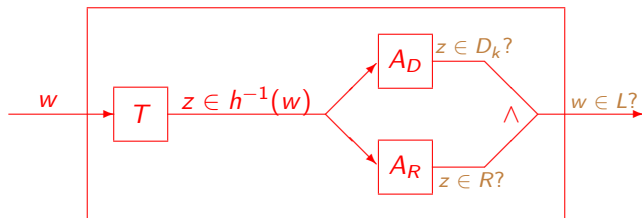
$$L = h(D_k \cap R)$$

where, for $\Omega_k = \{(1,)_1, (2,)_2, \dots, (k,)_k\}$:

- ▶ $D_k \subseteq \Omega_k^*$ is a Dyck language
- ▶ $R \subseteq \Omega_k^*$ is a regular language
- ▶ $h : \Omega_k \rightarrow \Sigma^*$ is an homomorphism

Furthermore, it is possible to restrict to *non-erasing* homomorphisms [Okhotin'12]

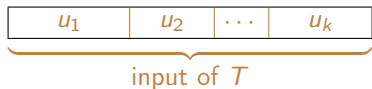
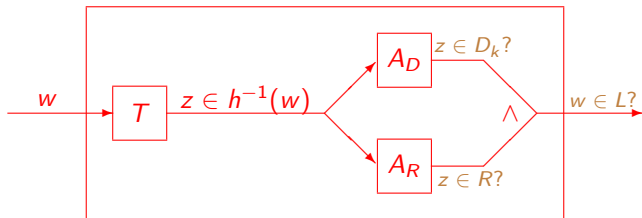
CFLs \rightarrow 2-Limited Automata



L context-free language, with $L = h(D_k \cap R)$

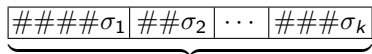
- ▶ T nondeterministic transducer computing h^{-1}
- ▶ A_D 2-LA accepting the Dyck language D_k
- ▶ A_R finite automaton accepting R

CFLs \rightarrow 2-Limited Automata



$$z = \sigma_1 \sigma_2 \dots \sigma_k \in h^{-1}(w)$$

$$h(\sigma_i) = u_i$$



(padded) input of A_D and A_R

Non erasing homomorphism!

Not stored into the tape!

Each σ_i is produced "on the fly" and replaced by γ_i , its first rewriting by A_D



Pushdown Automata \rightarrow 2-Limited Automata

PDA_s \rightarrow 2-LA_s

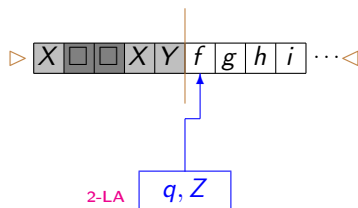
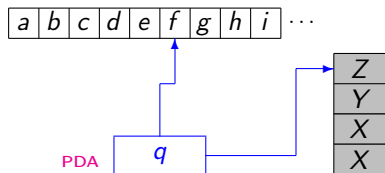
Polynomial cost!

DPDA_s \rightarrow D2-LA_s

Polynomial cost!

(in the description size)

Pushdown Automata \rightarrow 2-Limited Automata



Normal form for (D)PDAs:

- ▶ at each step, the stack height increases at most by 1
- ▶ ϵ -moves cannot push on the stack

Each (D)PDA can be simulated by an equivalent (D)2-LA of polynomial size

2-Limited Automata \equiv Pushdown Automata

Summing up...

- ▶ Descriptive complexity point of view

2-LAs \rightarrow PDAs

Exponential gap

PDAs \rightarrow 2-LAs

Polynomial upper bound

- ▶ Determinism vs Nondeterminism

Deterministic Context-Free Languages \equiv Deterministic 2-LAs

On the other hand:

$$L = \{a^n b^n c \mid n \geq 0\} \cup \{a^n b^{2^n} d \mid n \geq 0\} \in \text{det-3-LA} - \text{DCFL}$$

Infinite hierarchy [Hibbard'67]:

$$\text{det-}d\text{-LA} \supset \text{det-}(d-1)\text{-LA}, \text{ for each } d \geq 2$$

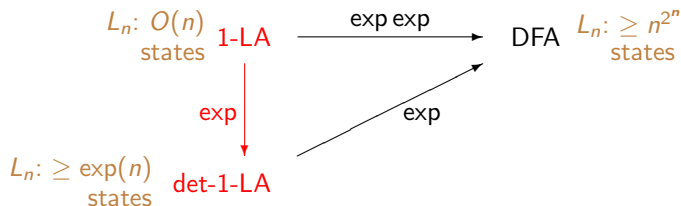
1-Limited Automata \rightarrow Finite Automata

Costs in states of the optimal simulations of n -state 1-LAs by finite automata:

	DFA	NFA
nondet. 1-LA	$2^{n \cdot 2^{n^2}}$	$n \cdot 2^{n^2}$
det. 1-LA	$n \cdot (n + 1)^n$	$n \cdot (n + 1)^n$

These upper bounds do not depend on the alphabet size of M !

Nondeterminism vs Determinism in 1-LAs



Corollary

Removing nondeterminism from 1-LAs requires exponentially many states

Cfr. Sakoda and Sipser question [Sakoda&Sipser'78]:

How much it costs in states to remove nondeterminism from two-way finite automata?

Futher Investigations

- ▶ Descriptive complexity aspects for $d > 2$

We conjecture that for $d > 2$ the size gap from d -limited automata to PDAs remains exponential

- ▶ Descriptive complexity aspects in the unary case

- Unary context-free languages are regular [Ginsburg&Rice'62]

- Ex: $L_n = (a^{2^n})^*$

	size
2-LA	$O(n)$
DPDA	$O(n)$
minimal DFA	2^n
minimal 2NFA	2^n