Regular realizability problems and context-free languages

Alexander Rubtsov[2,3] Mikhail Vyalyi[1,2,3]

[1]Computing Centre of Russian Academy of Sciences

[2]Moscow Institute of Physics and Technology

[3]National Research University
Higher School of Economics

26 June 2015

Definition of the problems

Filter

We fix language $F \subseteq \Sigma^*$ called filter.

Definition of the problems

Filter

We fix language $F \subseteq \Sigma^*$ called filter.

$L(\mathcal{A}) \in \mathrm{REG}$ – input of the problem, where $\mathcal{A}$ is NFA.

Regular realizability problem

$$\mathrm{NRR}(F) = \{\mathcal{A} \mid \mathcal{A} \in \mathrm{NFA}, L(\mathcal{A}) \cap F \neq \varnothing\}$$

Examples

Periodic filters

- $\mathrm{Per}_1 = \{(1^k\#)^n \mid k, n \in \mathbb{N}\}$

$$111\#111\#\cdots\#111\# \in \mathrm{Per}_1$$

Examples

Periodic filters

- $\mathrm{Per}_1 = \{(1^k\#)^n \mid k, n \in \mathbb{N}\}$

$$111\#111\#\cdots\#111\# \in \mathrm{Per}_1$$

The problem $\mathrm{NRR}(\mathrm{Per}_1)$ is NP-complete

Examples

Periodic filters

- $\mathrm{Per}_1 = \{(1^k\#)^n \mid k, n \in \mathbb{N}\}$

$$111\#111\#\cdots\#111\# \in \mathrm{Per}_1$$

  The problem $\mathrm{NRR}(\mathrm{Per}_1)$ is NP-complete

- $\mathrm{Per}_2 = \{(w\#)^n \mid w \in \boldsymbol{\Sigma}^*, |\boldsymbol{\Sigma}| = 2, n \in \mathbb{N}\}$

$$w\,\#\,w\,\#\cdots\#\,w\,\# \in \mathrm{Per}_2$$

Examples

Periodic filters

- $\mathrm{Per}_1 = \{(1^k\#)^n \mid k, n \in \mathbb{N}\}$

$$111\#111\#\cdots\#111\# \in \mathrm{Per}_1$$

The problem $\mathrm{NRR}(\mathrm{Per}_1)$ is NP-complete

- $\mathrm{Per}_2 = \{(w\#)^n \mid w \in \boldsymbol{\Sigma}^*, |\boldsymbol{\Sigma}| = 2, n \in \mathbb{N}\}$

$$w\#w\#\cdots\#w\# \in \mathrm{Per}_2$$

The problem $\mathrm{NRR}(\mathrm{Per}_2)$ is PSPACE-complete

Examples

Periodic filters

- $\mathrm{Per}_1 = \{(1^k\#)^n \mid k, n \in \mathbb{N}\}$

$$111\#111\#\cdots\#111\# \in \mathrm{Per}_1$$

The problem $\mathrm{NRR}(\mathrm{Per}_1)$ is NP-complete

- $\mathrm{Per}_2 = \{(w\#)^n \mid w \in \Sigma^*, |\Sigma| = 2, n \in \mathbb{N}\}$

$$w\#w\#\cdots\#w\# \in \mathrm{Per}_2$$

The problem $\mathrm{NRR}(\mathrm{Per}_2)$ is PSPACE-complete

Anderson T., Loftus J., Rampersad N., Santean N., Shallit J. Detecting palindromes, patterns and borders in regular languages. Information and Computation. Vol. 207, 2009. P. 1096–1118.

M.V. Paper in Russian, 2009.

Rational dominance

Definition

A language $L \subseteq A^*$ is rationally dominated by $L' \subseteq B^*$ if there exists a rational relation R such that

$$L = \{u \in A^* \mid \exists v \in L' \ (v, u) \in R\}$$

$$L \leqslant_{rat} L'$$

Rational transductions

Definition

A finite state transducer (FST) $T(x) : \Delta^* \to \Gamma^*$ is a (nondeterministic) automaton with output tape $T = (\Delta, \Gamma, Q, q_0, \delta, F)$, where

- Δ – input alphabet;
- Γ – output alphabet;
- Q – set of states;
- $\delta : Q \times \Delta \cup \{\varepsilon\} \times \Gamma \cup \{\varepsilon\} \times Q$ – transitions relation;
- q_0 – initial state;
- F – set of accepting states.

If transducer T on input x has no path to accepting state, then $T(x) = \varnothing$.

Rational transductions

In other words

$$L \leqslant_{\mathrm{rat}} L' \Leftrightarrow \exists T \in \mathrm{FST} : L = T(L')$$

Reduction on filters

Proposition

$$F_1 \leqslant_{\mathrm{rat}} F_2 \Rightarrow \mathrm{NRR}(F_1) \leqslant_{\log} \mathrm{NRR}(F_2)$$

Rational cone

Definition

A rational cone is a class of languages closed under rational dominance. Denote by $\mathcal{T}(\mathrm{L})$ the least rational cone that includes language L and call it rational cone generated by L.

Rational cone

Definition

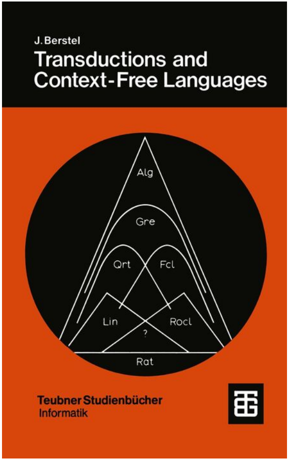
A rational cone is a class of languages closed under rational dominance. Denote by $\mathcal{T}(L)$ the least rational cone that includes language L and call it rational cone generated by L.

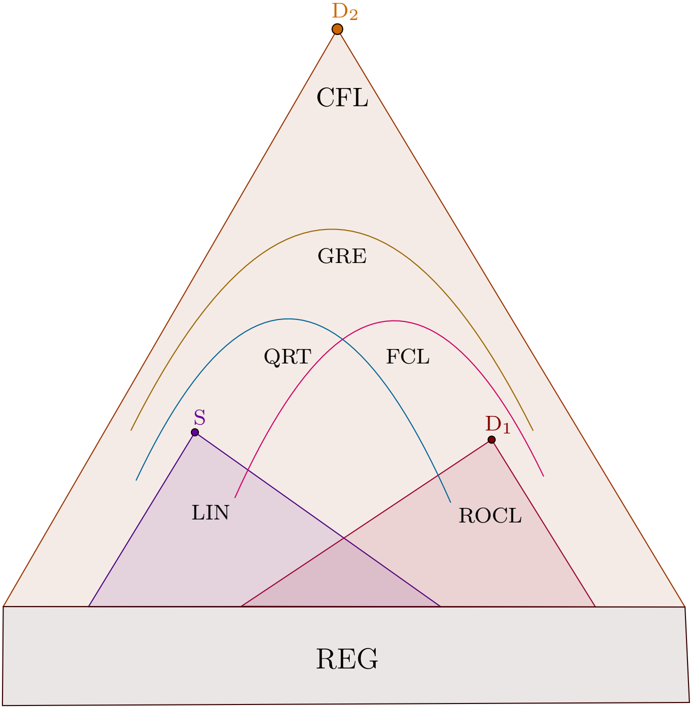
Theorem (Chomsky, Schützenberger)

$$\mathcal{T}(D_2) = \mathrm{CFL}$$

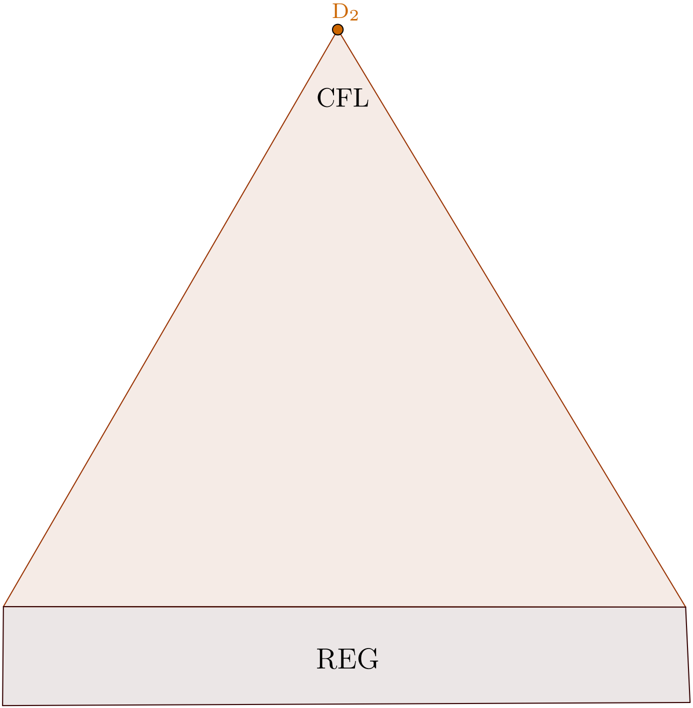
$$D_n = \langle S \to SS \mid a_1 S \bar{a}_1 \mid \cdots \mid a_n S \bar{a}_n \mid \varepsilon \rangle.$$

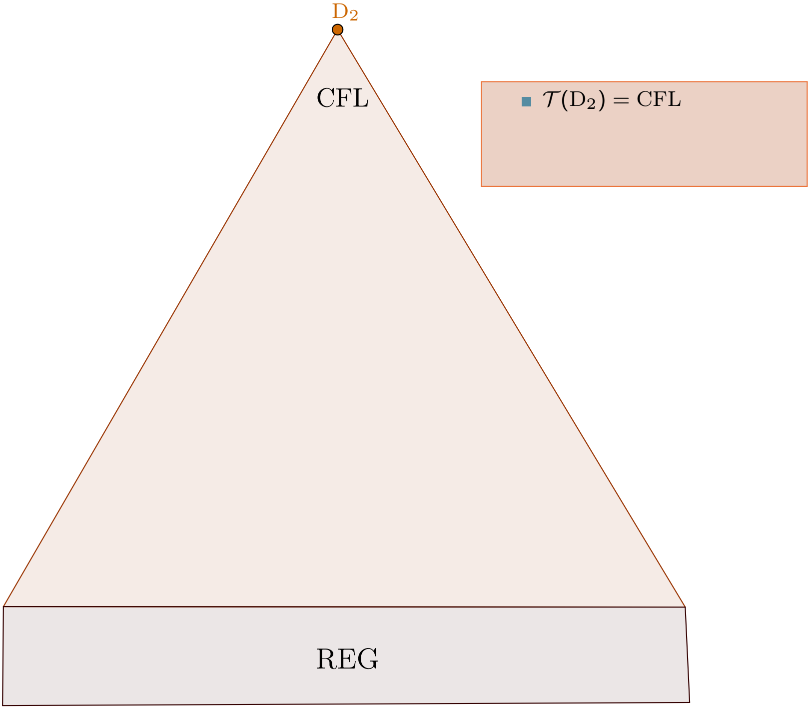
J. Berstel's book cover



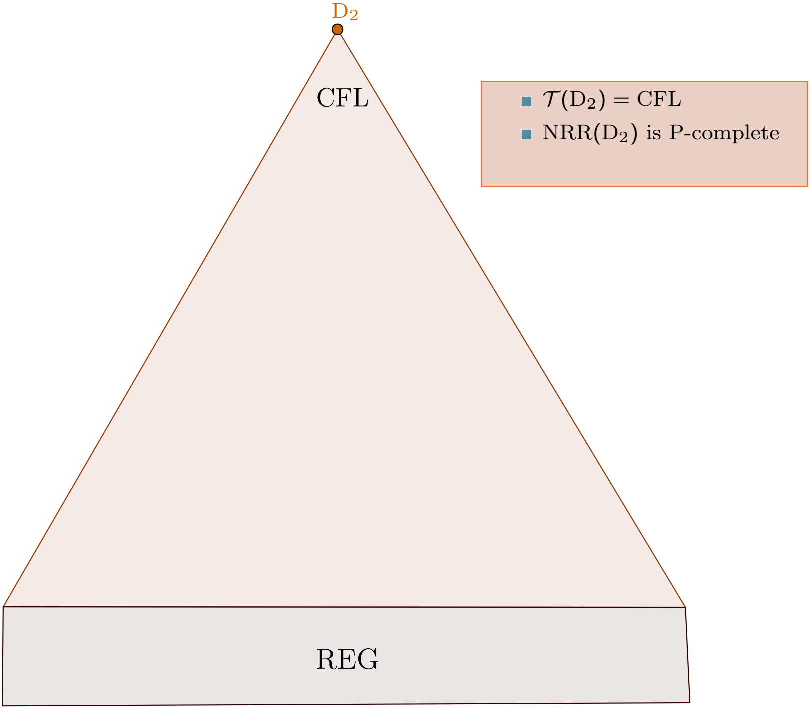


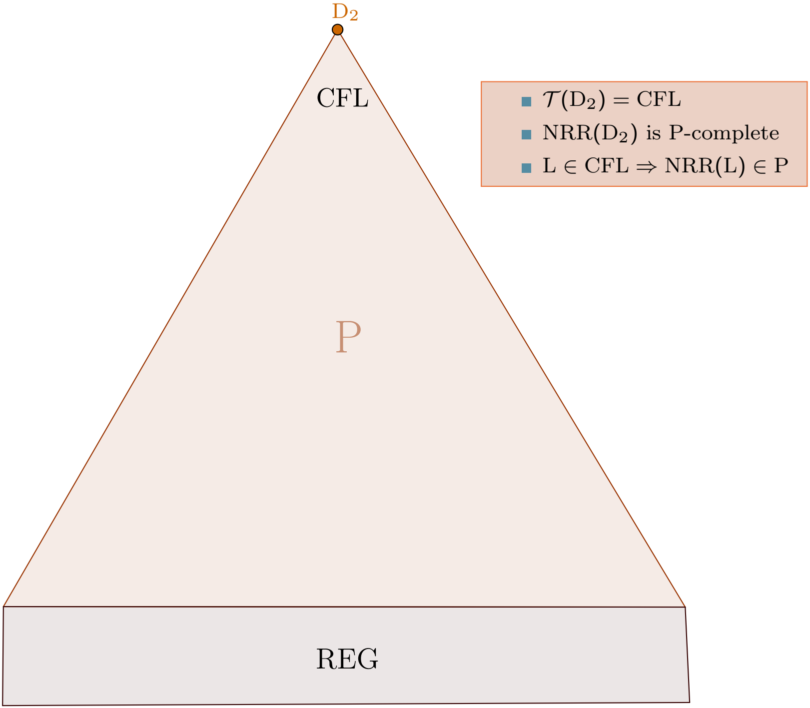
D_2
CFL
GRE
QRT
FCL
S
D_1
LIN
ROCL
REG


D_2
CFL
REG

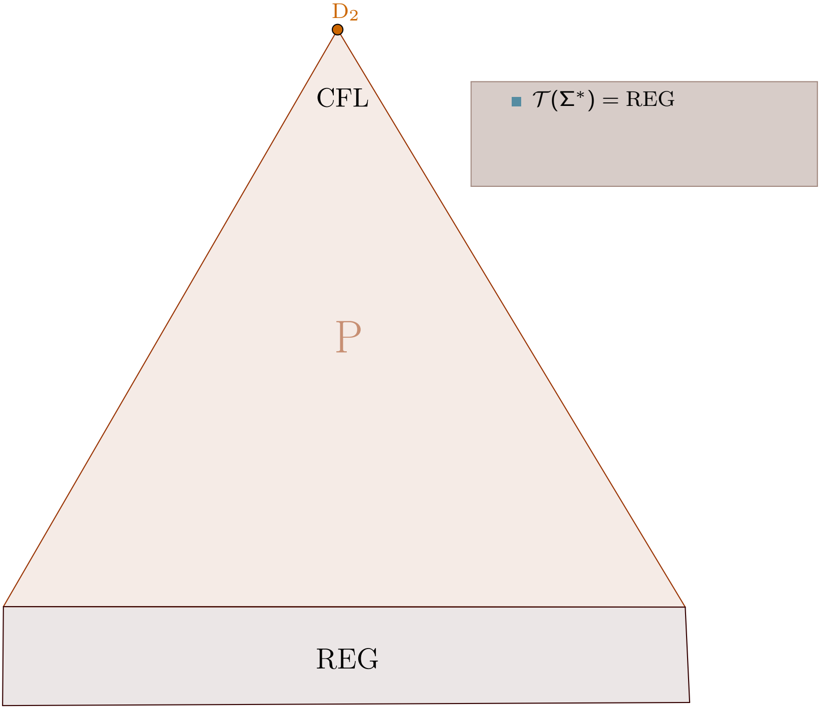


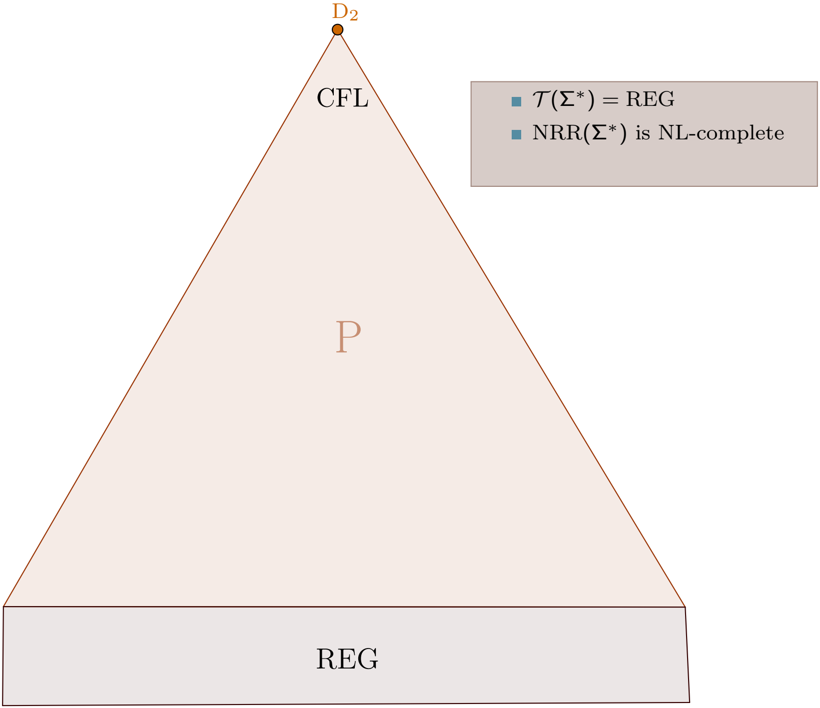
D_2
CFL
$\mathcal{T}(D_2) = CFL$
REG


D_2
CFL
$\mathcal{T}(D_2) = \mathrm{CFL}$
$\mathrm{NRR}(D_2)$ is P-complete
REG

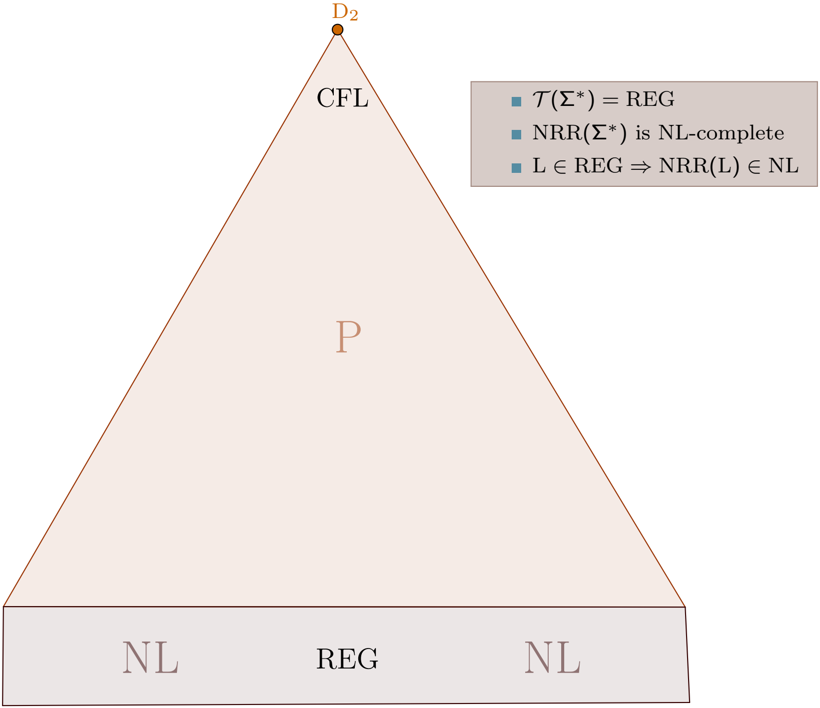


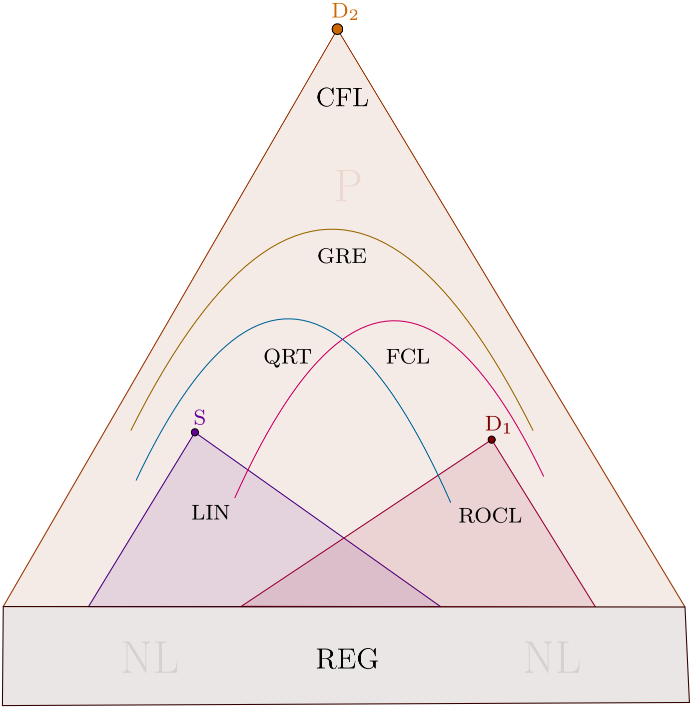
D_2
CFL
P
REG
$\mathcal{T}(D_2) = CFL$
$NRR(D_2)$ is P-complete
$L \in CFL \Rightarrow NRR(L) \in P$


D_2
CFL
$\mathcal{T}(\Sigma^*) = \mathrm{REG}$
P
REG

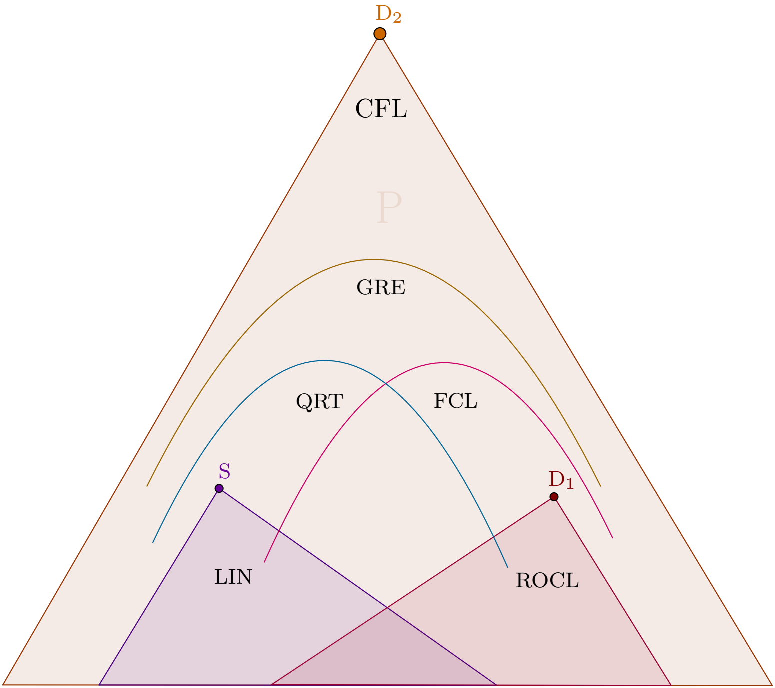


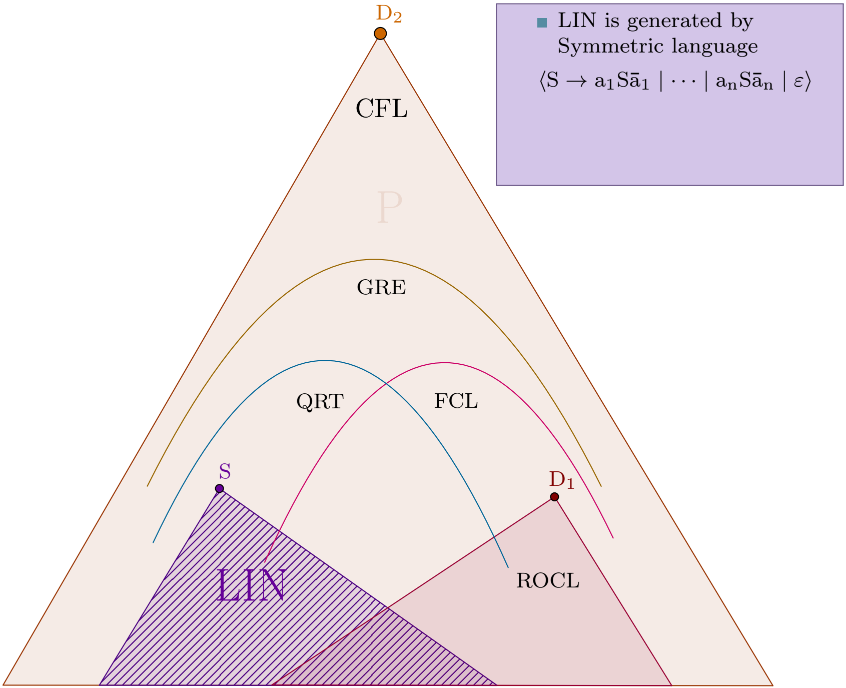
D_2
CFL
P
REG
$\mathcal{T}(\mathbf{\Sigma}^*) = \mathrm{REG}$
$\mathrm{NRR}(\mathbf{\Sigma}^*)$ is NL-complete

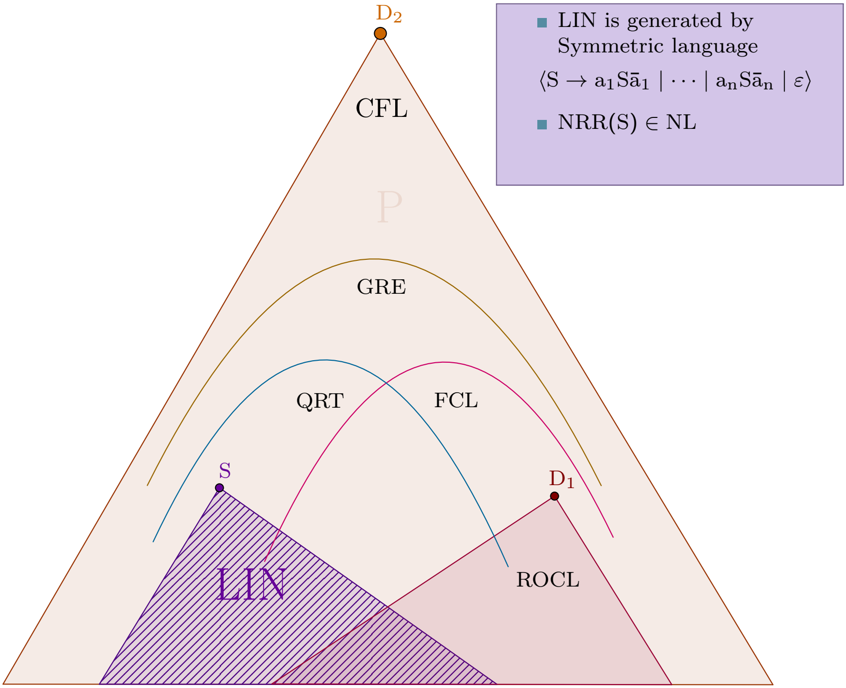

D_2
CFL
P
NL
REG
NL
$\mathcal{T}(\boldsymbol{\Sigma}^*) = \mathrm{REG}$
NRR($\boldsymbol{\Sigma}^*$) is NL-complete
L $\in$ REG $\Rightarrow$ NRR(L) $\in$ NL



D2
CFL
P
GRE
QRT
FCL
S
D1
LIN
ROCL
NL
REG
NL


D2
CFL
P
GRE
QRT
FCL
S
D1
LIN
ROCL


LIN is generated by
Symmetric language
⟨S → a₁Sā₁ | ⋯ | aₙSāₙ | ε⟩
D2
CFL
P
GRE
QRT
FCL
S
D1
LIN
ROCL



LIN is generated by Symmetric language
$\langle S \to a_1 S \bar{a}_1 \mid \cdots \mid a_n S \bar{a}_n \mid \varepsilon \rangle$
$NRR(S) \in NL$
D_2
CFL
P
GRE
QRT
FCL
S
D_1
LIN
ROCL

Th. (Anderson, Loftus, Rampersad, Santean, Shallit, 2009)

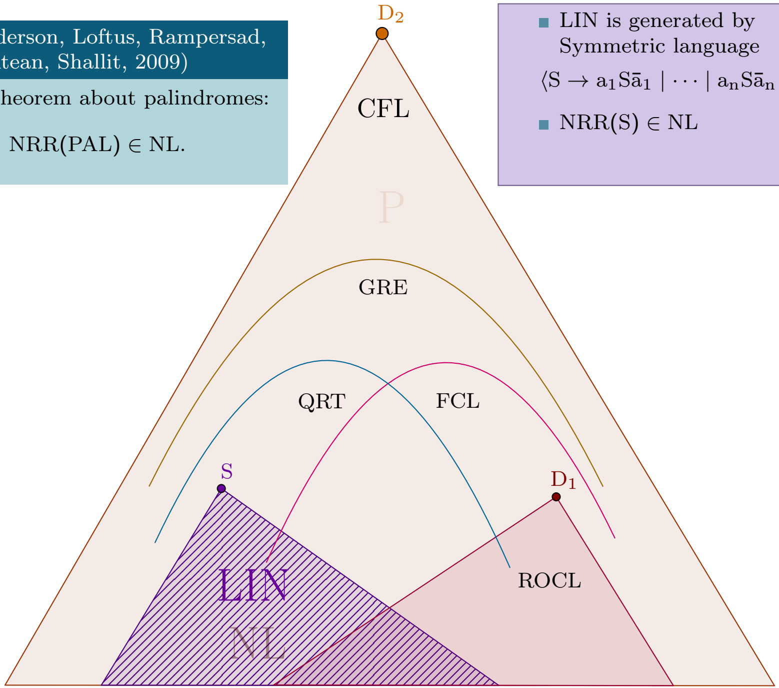
Similar theorem about palindromes:

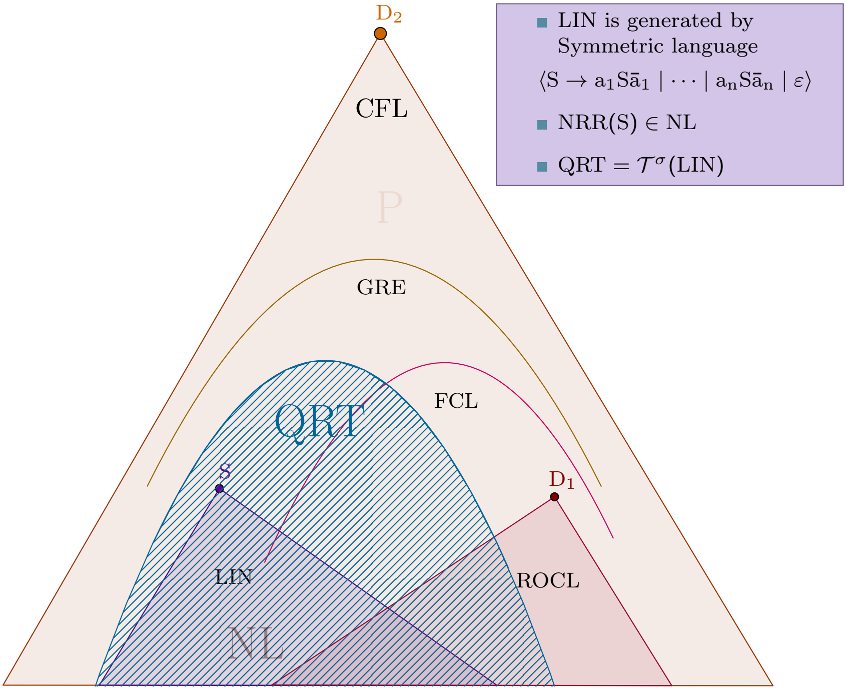
$$\mathrm{NRR(PAL)} \in \mathrm{NL}.$$

- LIN is generated by Symmetric language

$$\langle S \to a_1 S \bar{a}_1 \mid \cdots \mid a_n S \bar{a}_n \mid \varepsilon \rangle$$

- $\mathrm{NRR(S)} \in \mathrm{NL}$





LIN is generated by Symmetric language
$\langle S \to a_1 S \bar{a}_1 \mid \cdots \mid a_n S \bar{a}_n \mid \varepsilon \rangle$
NRR(S) $\in$ NL
QRT $= \mathcal{T}^\sigma$(LIN)
D_2
CFL
P
GRE
FCL
QRT
S
D_1
LIN
ROCL
NL

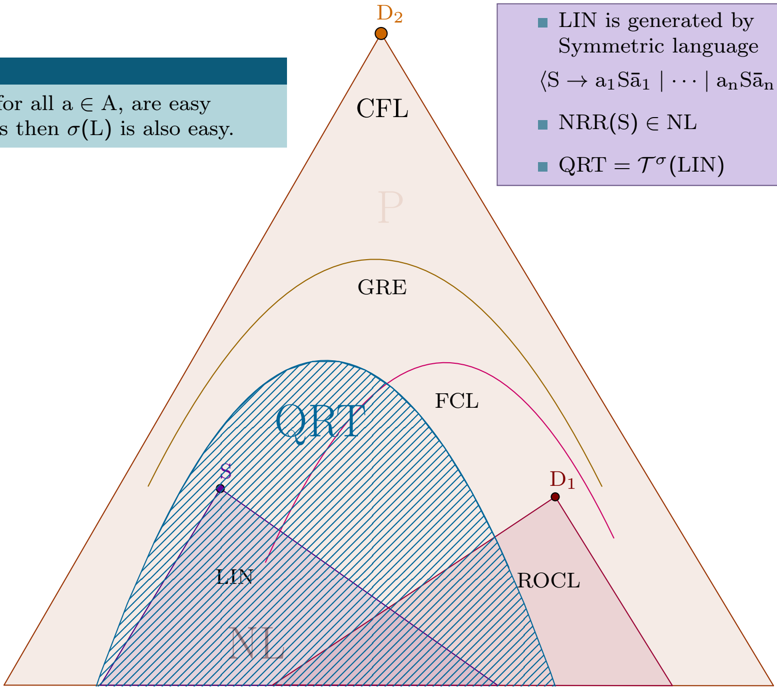
Lemma

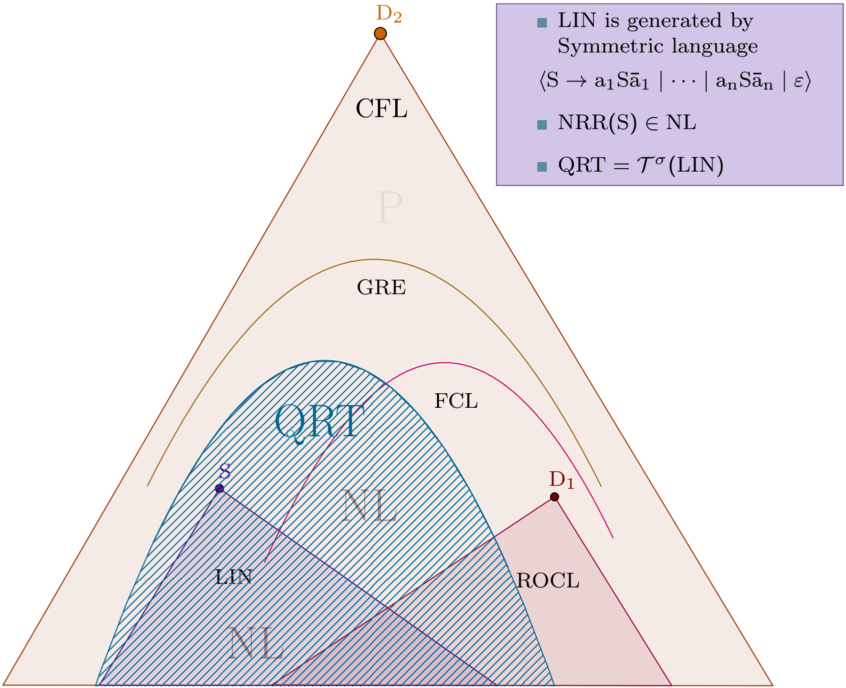
If L, L_a for all $a \in A$, are easy languages then $\sigma(L)$ is also easy.

- LIN is generated by Symmetric language

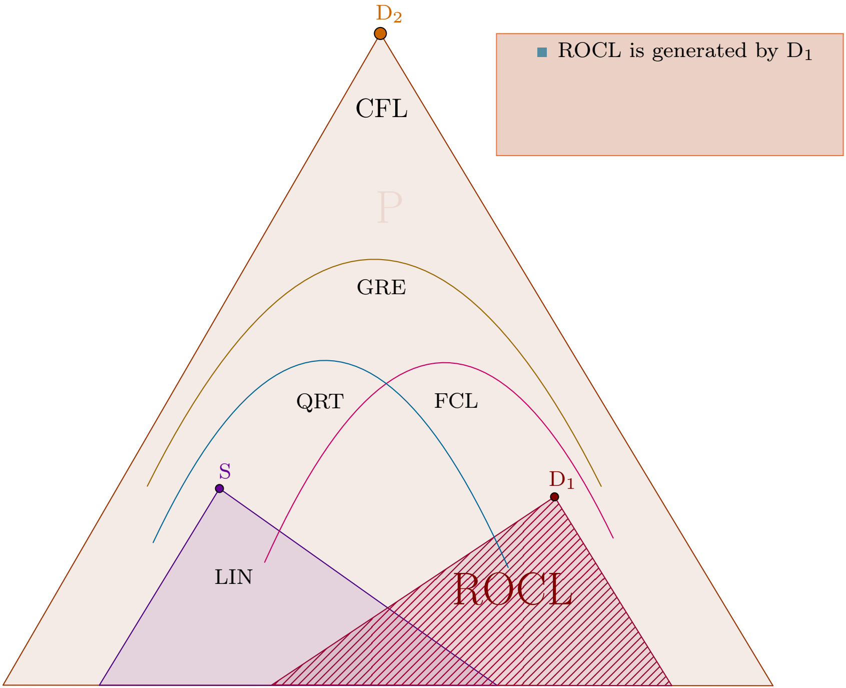
$\langle S \to a_1 S \bar{a}_1 \mid \cdots \mid a_n S \bar{a}_n \mid \varepsilon \rangle$

- $NRR(S) \in NL$
- $QRT = \mathcal{T}^{\sigma}(LIN)$





D2
CFL
P
GRE
FCL
QRT
S
D1
NL
LIN
ROCL
NL
LIN is generated by Symmetric language
⟨S → a1Sā1 | · · · | anSān | ε⟩
NRR(S) ∈ NL
QRT = 𝒯σ(LIN)

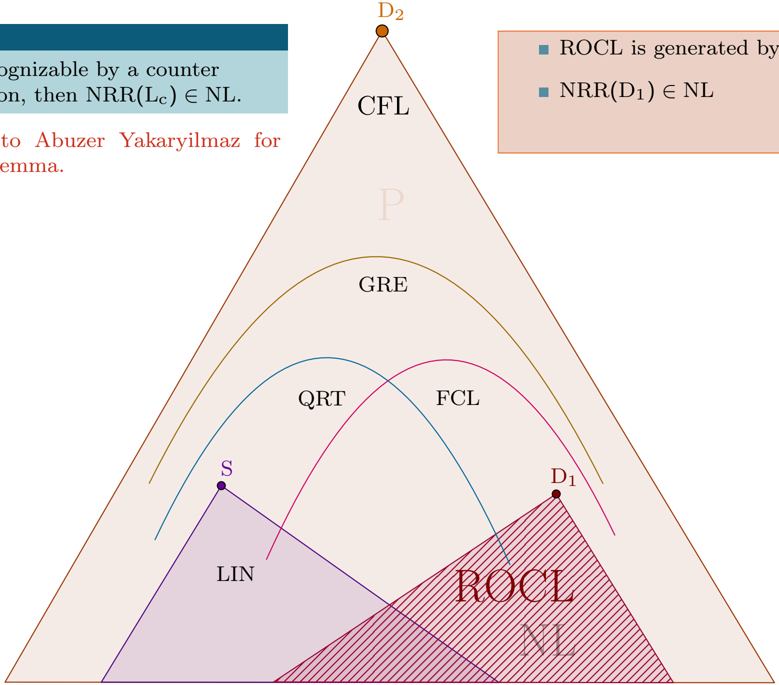


D2
ROCL is generated by D1
CFL
P
GRE
QRT
FCL
S
D1
LIN
ROCL

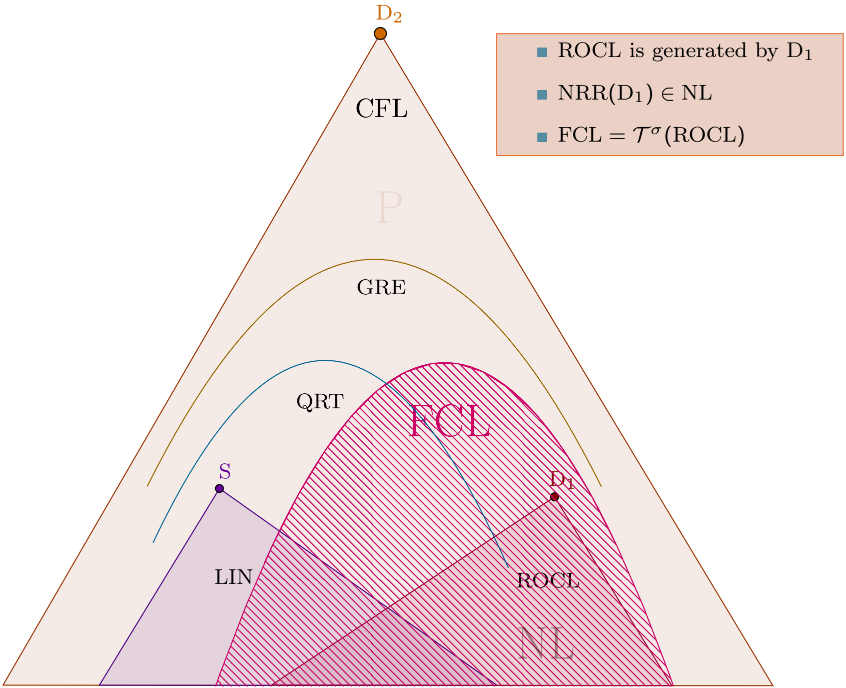
Lemma

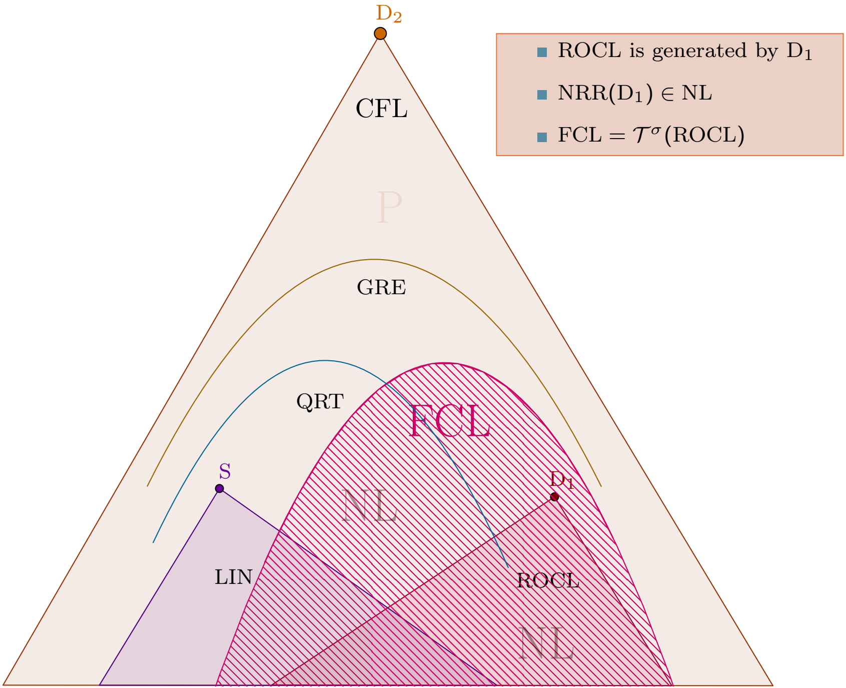
If L_c recognizable by a counter automaton, then $NRR(L_c) \in NL$.

Thanks to Abuzer Yakaryilmaz for the key-lemma.

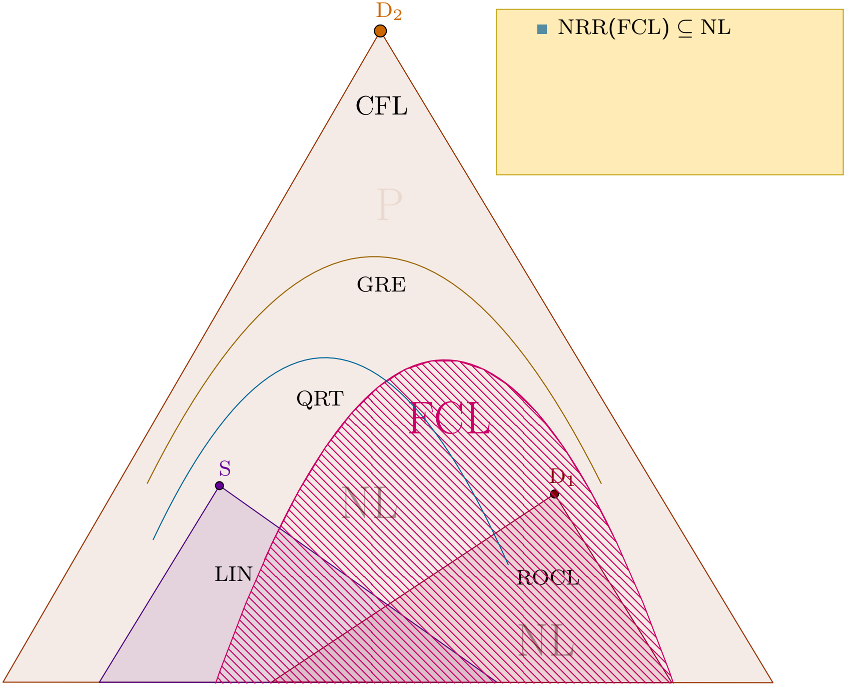


- ROCL is generated by D_1
- $NRR(D_1) \in NL$

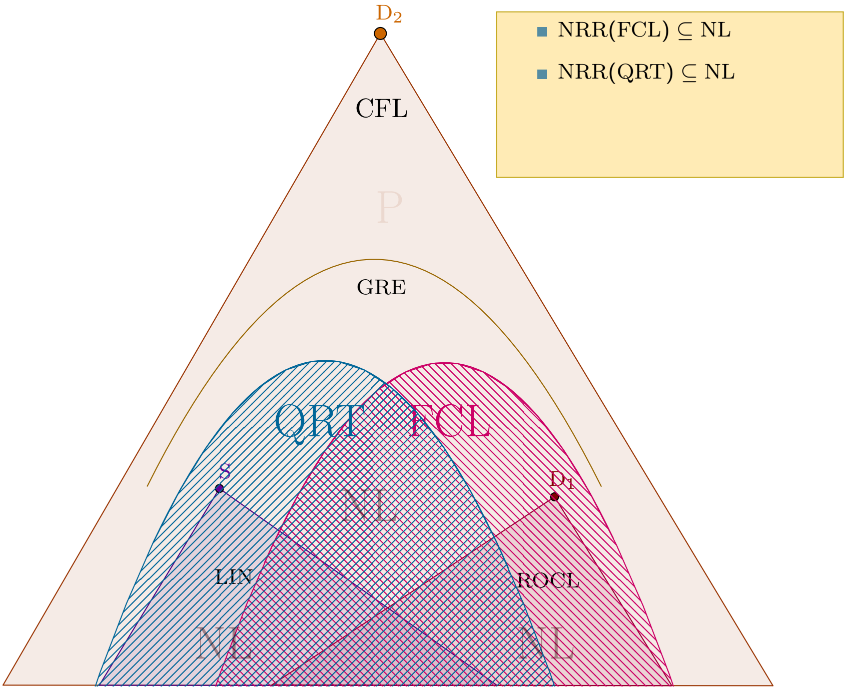

ROCL is generated by D_1
$NRR(D_1) \in NL$
$FCL = \mathcal{T}^{\sigma}(ROCL)$
D_2
CFL
P
GRE
QRT
FCL
S
D_1
LIN
ROCL
NL



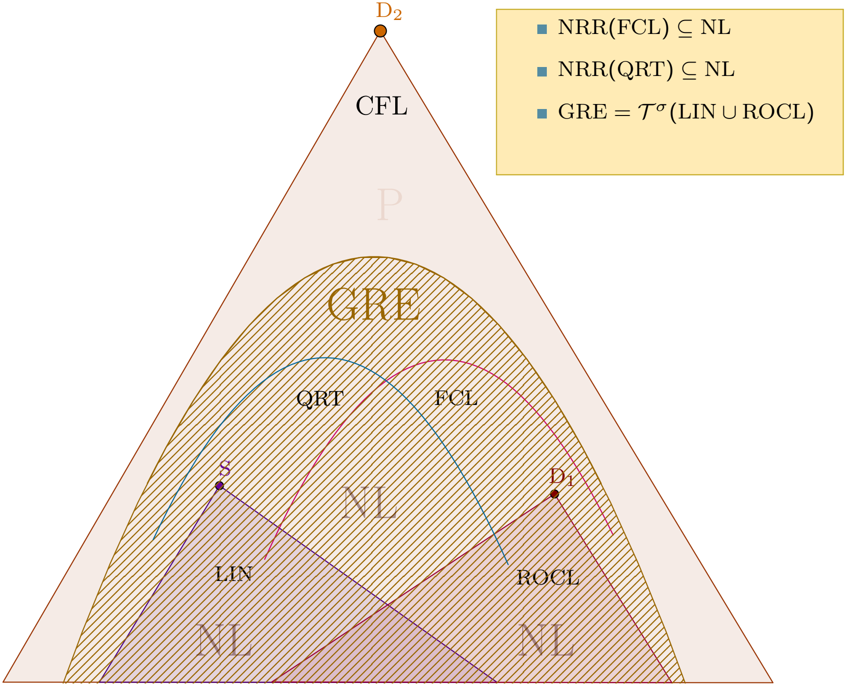
D2
ROCL is generated by D1
NRR(D1) ∈ NL
FCL = 𝒯^σ(ROCL)
CFL
P
GRE
QRT
FCL
S
D1
NL
LIN
ROCL
NL



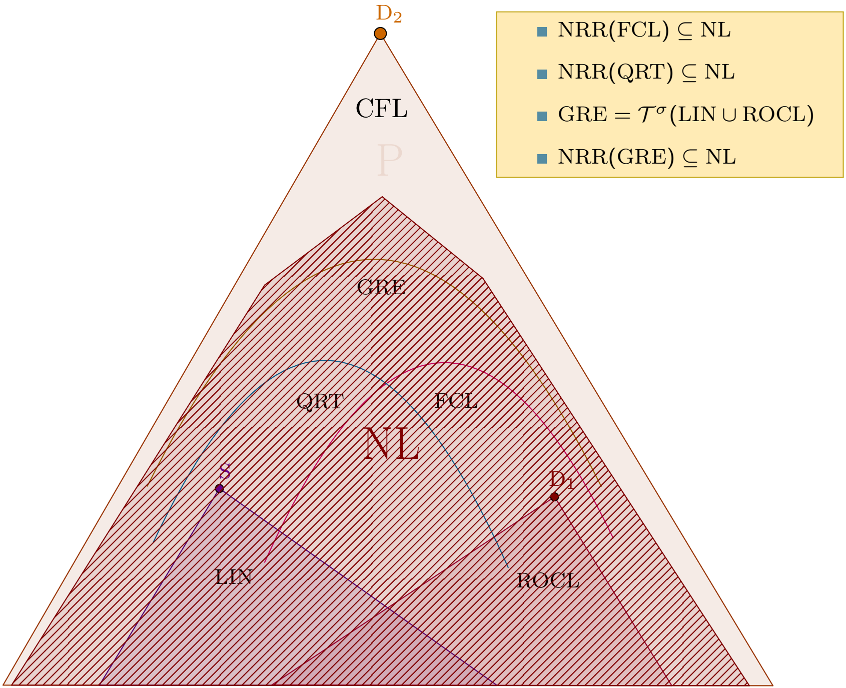
D_2
NRR(FCL) ⊆ NL
CFL
P
GRE
QRT
FCL
S
D_1
NL
LIN
ROCL
NL

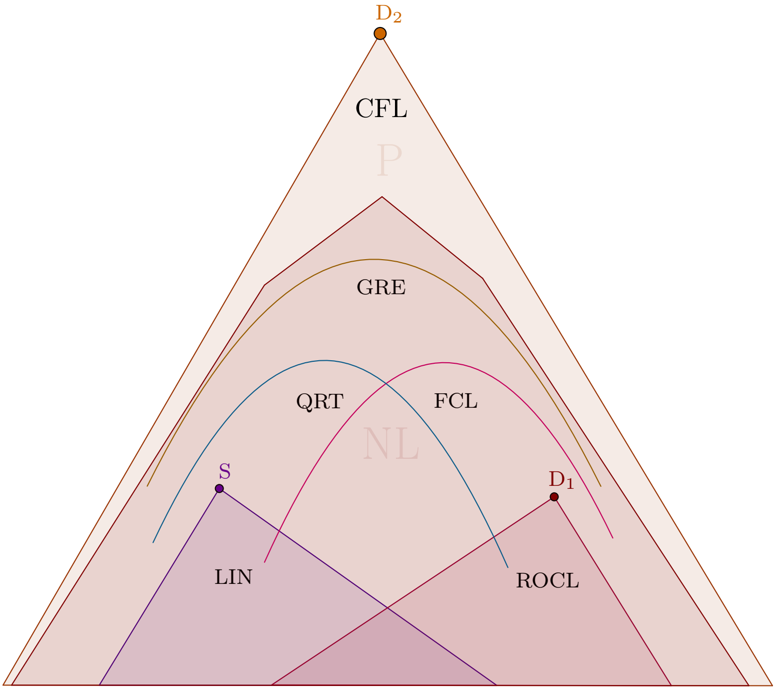


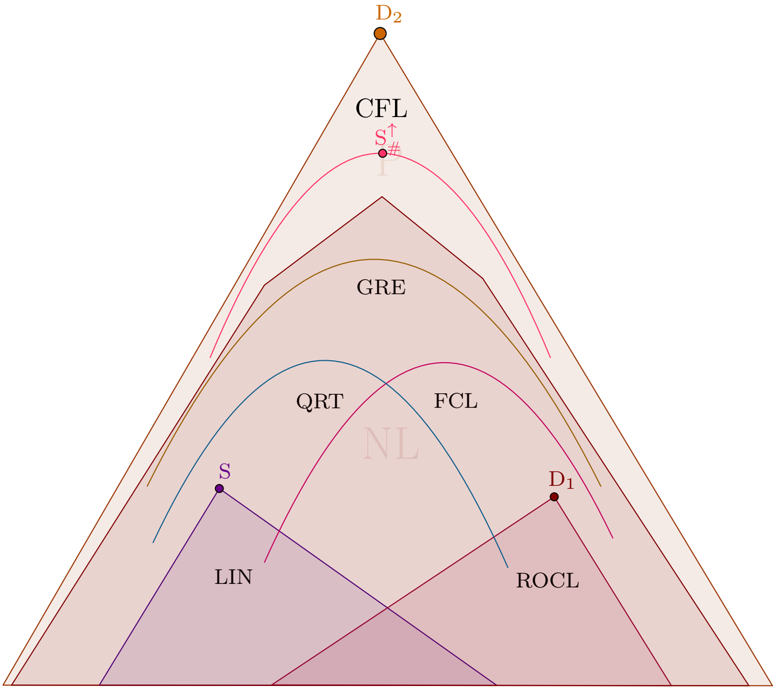
D2
NRR(FCL) ⊆ NL
NRR(QRT) ⊆ NL
CFL
P
GRE
QRT
FCL
S
D1
NL
LIN
ROCL
NL
NL



D2
NRR(FCL) ⊆ NL
NRR(QRT) ⊆ NL
GRE = $\mathcal{T}^\sigma$(LIN ∪ ROCL)
CFL
P
GRE
QRT
FCL
S
D1
NL
LIN
ROCL
NL
NL


D2
NRR(FCL) ⊆ NL
NRR(QRT) ⊆ NL
GRE = $\mathcal{T}^\sigma$(LIN ∪ ROCL)
NRR(GRE) ⊆ NL
CFL
P
GRE
QRT
FCL
NL
S
D1
LIN
ROCL

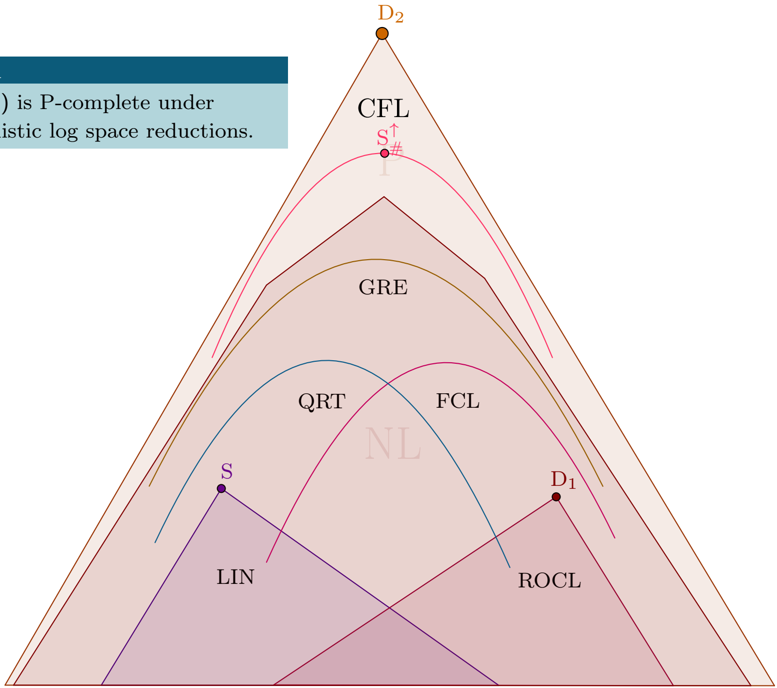

D$_2$
CFL
P
GRE
QRT
FCL
NL
S
D$_1$
LIN
ROCL



D2
CFL
S↑#
P
GRE
QRT
FCL
NL
S
D1
LIN
ROCL

Theorem

NRR($S^{\uparrow}_{\#}$) is P-complete under deterministic log space reductions.



Rational index

Definition

The rational index $\rho_L(n)$ of a language L is a function that returns the maximum length of the shortest word from the intersection of the language L and a language $L(\mathcal{A})$ recognizing by an automaton $\mathcal{A}$ with n states provided $L(\mathcal{A}) \cap L \neq \varnothing$:

$$\rho_L(n) = \max_{\mathcal{A}:|Q_{\mathcal{A}}|=n} (\min_{w}\{|w| : w \in L(\mathcal{A}) \cap L \neq \varnothing\})$$

Rational index

Definition

The rational index $\rho_L(n)$ of a language L is a function that returns the maximum length of the shortest word from the intersection of the language L and a language $L(\mathcal{A})$ recognizing by an automaton $\mathcal{A}$ with n states provided $L(\mathcal{A}) \cap L \neq \varnothing$:

$$\rho_L(n) = \max_{\mathcal{A}:|Q_{\mathcal{A}}|=n} (\min_{w}\{|w| : w \in L(\mathcal{A}) \cap L \neq \varnothing\})$$

Theorem (Boasson, Courcelle, Nivat, 1981)

If $L' \leqslant_{rat} L$ then there exists a constant c such that

$$\rho_{L'}(n) \leqslant cn(\rho_L(cn) + 1).$$

Properties of rational index

Proposition

Rational index of an arbitrary context-free language is bounded from below by a linear function.

Properties of rational index

Proposition

Rational index of an arbitrary context-free language is bounded from below by a linear function.

Theorem (Pierre 1992)

The rational index of any generator of the rational cone of CFL belongs to $\exp(\Theta(n^2/\log n))$.

Properties of rational index

Proposition

Rational index of an arbitrary context-free language is bounded from below by a linear function.

Theorem (Pierre 1992)

The rational index of any generator of the rational cone of CFL belongs to $\exp(\Theta(n^2/\log n))$.

Theorem (Pierre, Farrinone, 1990)

For a positive algebraic number $\gamma > 1$ there exists a context-free language with the rational index $\Theta(n^\gamma)$.

Complexity of RR problems

Theorem

Let F be a context-free filter with polynomially bounded rational index, then the problem NRR(F) belongs to NSPACE($\log^2 n$).

Complexity of RR problems

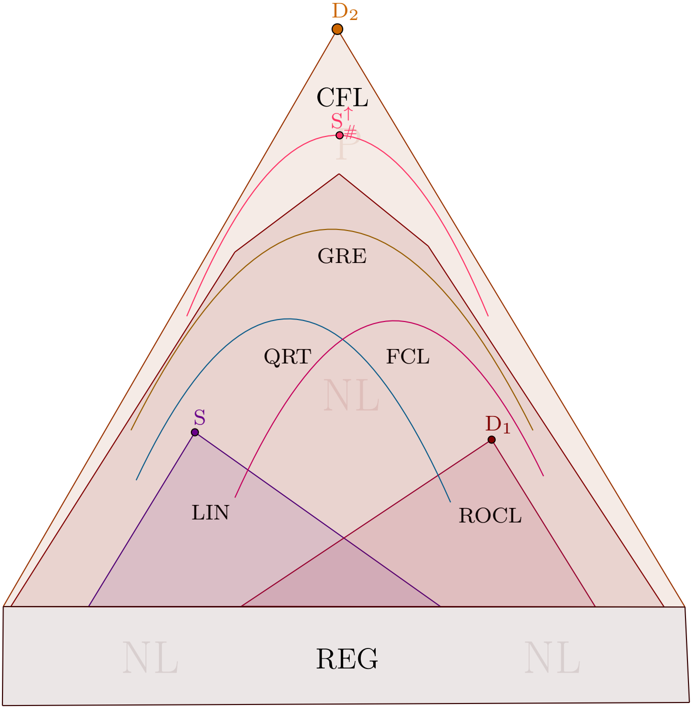
Theorem

Let F be a context-free filter with polynomially bounded rational index, then the problem NRR(F) belongs to $\mathrm{NSPACE}(\log^2 n)$.

Conjecture

Let F be a context-free filter with polynomially bounded rational index, then the problem NRR(F) belongs to NL.

Thank You!



D2
CFL
S↑#
P
GRE
QRT
FCL
NL
S
D1
LIN
ROCL
NL
REG
NL