ICANN19

Counting with Analog Neurons

Jiří Šíma

sima@cs.cas.cz

Institute of Computer Science
Czech Academy of Sciences

Proceedings of ICANN 2019, Part I, pp. 389–400

Manuscript Number **39**

Motivations: The Computational Power of NNs

(discrete-time recurrent NNs with the saturated-linear activation function)

depends on the information contents of weight parameters:

1. **integer** weights: finite automaton (Minsky, 1967)

2. **rational** weights: Turing machine (Siegelmann, Sontag, 1995)

   polynomial time $\equiv$ complexity class P

   polynomial time & increasing **Kolmogorov complexity** of real weights $\equiv$ a proper hierarchy of nonuniform complexity classes between P and P/poly

   (Balcázar, Gavaldà, Siegelmann, 1997)

3. arbitrary **real** weights: "super-Turing" computation (Siegelmann, Sontag, 1994)

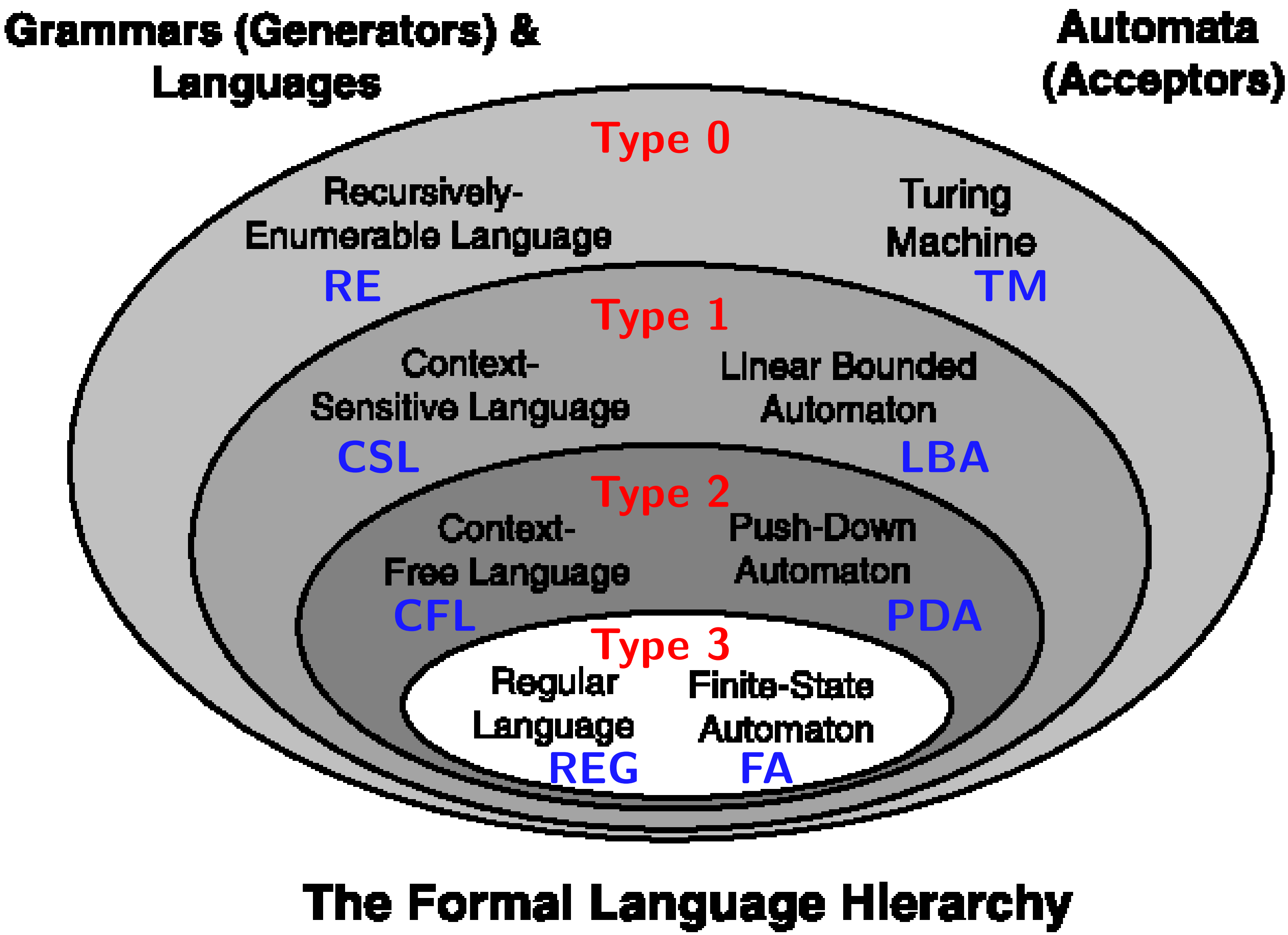
   polynomial time $\equiv$ nonuniform complexity class P/poly

   exponential time $\equiv$ any I/O mapping

filling the gap between integer and rational weights w.r.t. the Chomsky hierarchy

regular (Type 3) $\times$ recursively enumerable (Type 0) languages

The Traditional Chomsky Hierarchy



The Analog Neuron Hierarchy

α**ANN** $=$ a binary-state NN with **integer** weights
$+$ α **extra analog-state neurons** with **rational** weights

the computational power of NNs increases with the number α of extra analog neurons:

$$\text{(Type 3) FAs} \equiv \text{0ANNs} \subseteq \text{1ANNs} \subseteq \text{2ANNs} \subseteq \text{3ANNs} \subseteq \ldots \equiv \text{TMs (Type 0)}$$

$\uparrow$ **integer** weights — $\times$ **Chomsky hierarchy** **Type 1, 2 ?** — $\uparrow$ **rational** weights

Known Results:

- classifying **1ANNs** within the Chomsky hierarchy (Šíma, 2017):
  - upper bound: **1ANNs** $\subset$ **LBAs** $\equiv$ **CSLs (Type 1)**
  - lower bound: **1ANNs** $\not\subset$ **PDAs** $\equiv$ **CFLs (Type 2)**
    $\left(L_1 = \left\{x_1 \ldots x_n \in \{0,1\}^* \mid \sum_{k=1}^{n} x_{n-k+1} \left(\frac{3}{2}\right)^{-k} < 1\right\} \in \text{1ANNs} \setminus \text{CFLs}\right)$
  - **1ANNs** with **"quasi-periodic"** weights $\subseteq$ **FAs** $\equiv$ **REG (Type 3)**
- the analog neuron hierarchy collapses at **3ANNs** (Šíma, 2018):

$$\text{3ANNs} = \text{4ANNs} = \text{5ANNs} = \ldots \equiv \text{TMs} \equiv \text{RE (Type 0)}$$

The Main Result: Separating 2ANNs From 1ANNs

"counting" language $L_{\#} = \{0^n 1^n \mid n \geq 1\} \in$ **2ANNs** $\setminus$ **1ANNs**

$L_{\#}$ is a (non-regular) deterministic context-free language (DCFL)

accepted by a deterministic push-down automaton (DPDA)

1. $L_{\#} \notin$ **1ANNs**:

Theorem 1. *The deterministic context-free language $L_{\#}$ cannot be recognized by a neural network 1ANN with one extra analog unit having real weights.*

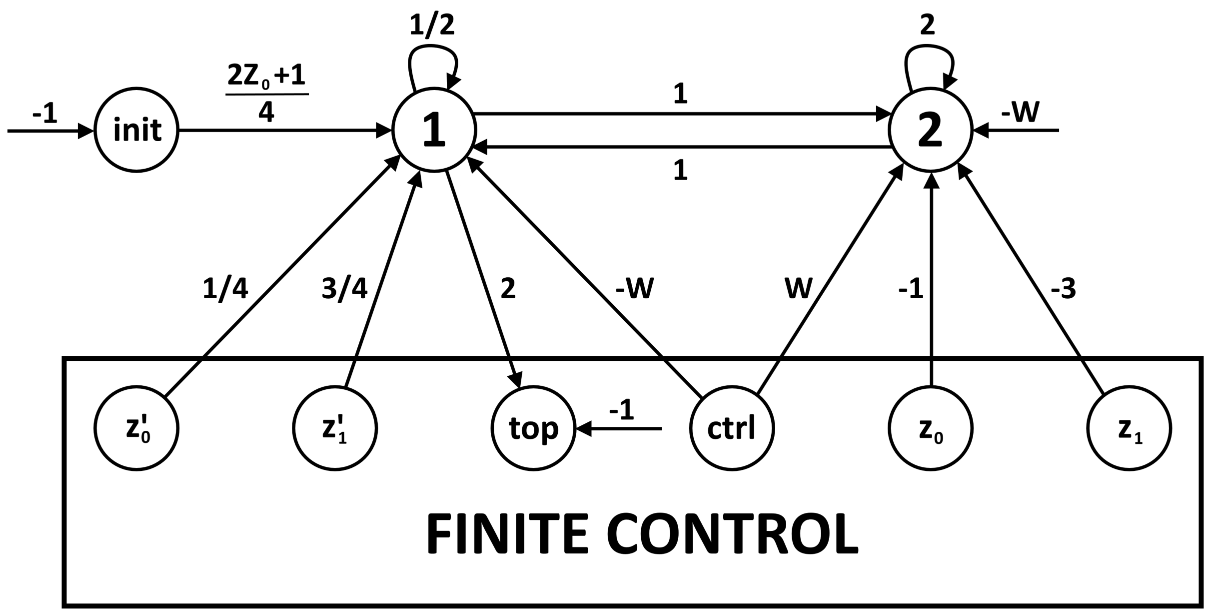
generalizes to (**DCFLs** $\setminus$ **REG**) $\cap$ **1ANNs** $= \emptyset$

i.e. **1ANNs** $\cap$ **DCFLs** $=$ **0ANNs** `(Šíma, Plátek, 2019)`

2. $L_{\#} \in$ **DCFLs** $\subset$ **2ANNs**:

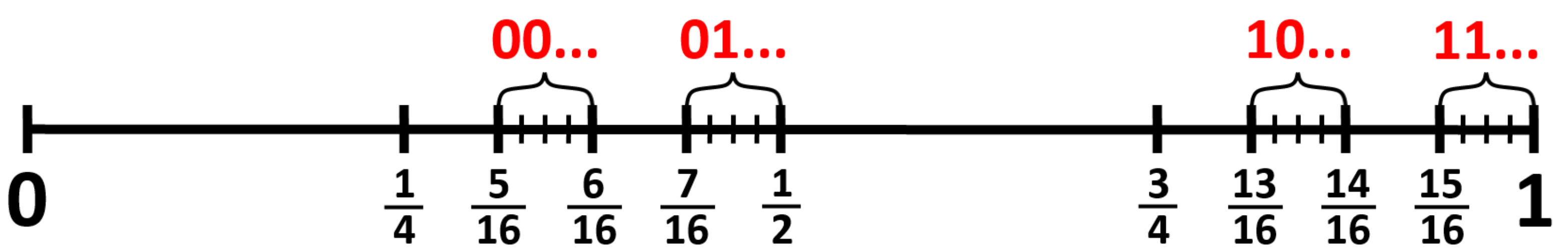
Theorem 2. *For any deterministic context-free language $L \subseteq \{0,1\}^*$, there is a neural network 2ANN with two extra analog units having rational weights, $\mathcal{N}$, which accepts $L = \mathcal{L}(\mathcal{N})$.*

A Schema of 2ANNs Simulating DPDAs



the stack contents $x_1 \ldots x_n \in \{0,1\}^*$ are encoded by states $y_1, y_2 \in [0,1]$ of analog neurons 1 (push), 2 (pop) using Cantor-like set (Siegelmann,Sontag,1995):

$$\mathtt{code}(x_1 \ldots x_n) = \sum_{i=1}^{n} \frac{2x_i + 1}{4^i} \in [0,1]$$

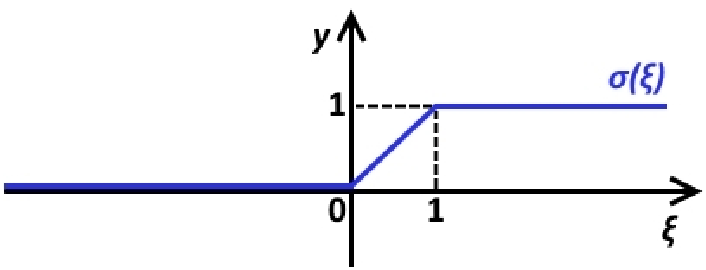


Implementing the Stack Operations

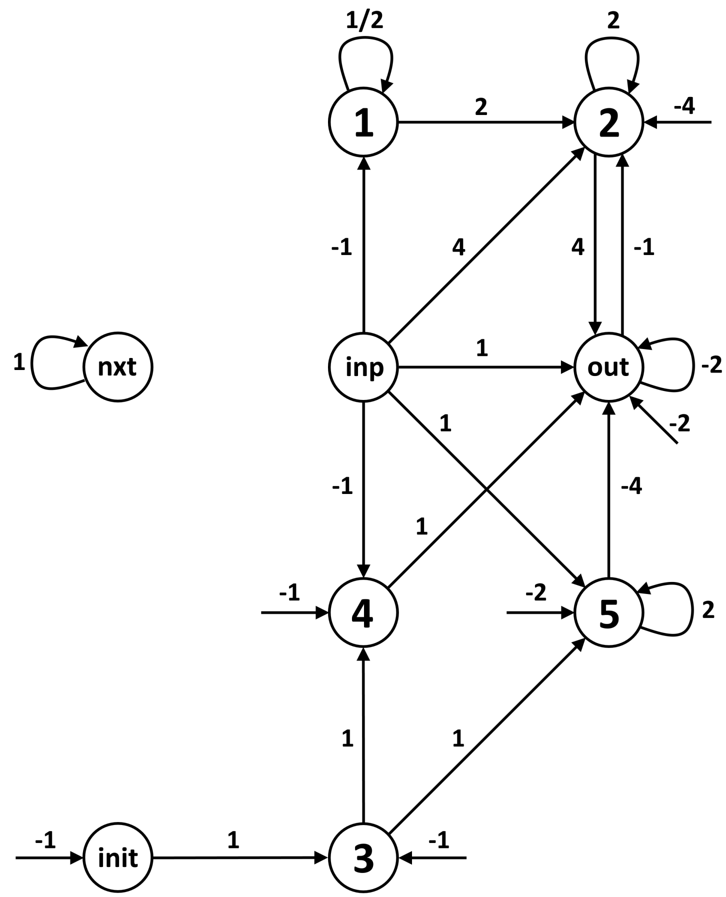
$$\mathtt{top}(y_1) = H(2y_1 - 1) = \begin{cases} 1 & \text{if } y_1 \geq \frac{1}{2} \text{ i.e. } y_1 = \mathsf{code}(1\, x_2 \ldots x_n) \\ 0 & \text{if } y_1 < \frac{1}{2} \text{ i.e. } y_1 = \mathsf{code}(0\, x_2 \ldots x_n) \end{cases}$$

$$\mathtt{pop}(y_2) = \sigma(4y_2 - 2\,\mathtt{top} - 1) = \mathsf{code}(x_2 \ldots x_n)$$

$$\mathtt{push}(y_1, b) = \sigma\left(\tfrac{1}{4}\, y_1 + \tfrac{2b+1}{4}\right) = \mathsf{code}(b\, x_1 \ldots x_n) \quad \text{for } b \in \{0, 1\}$$



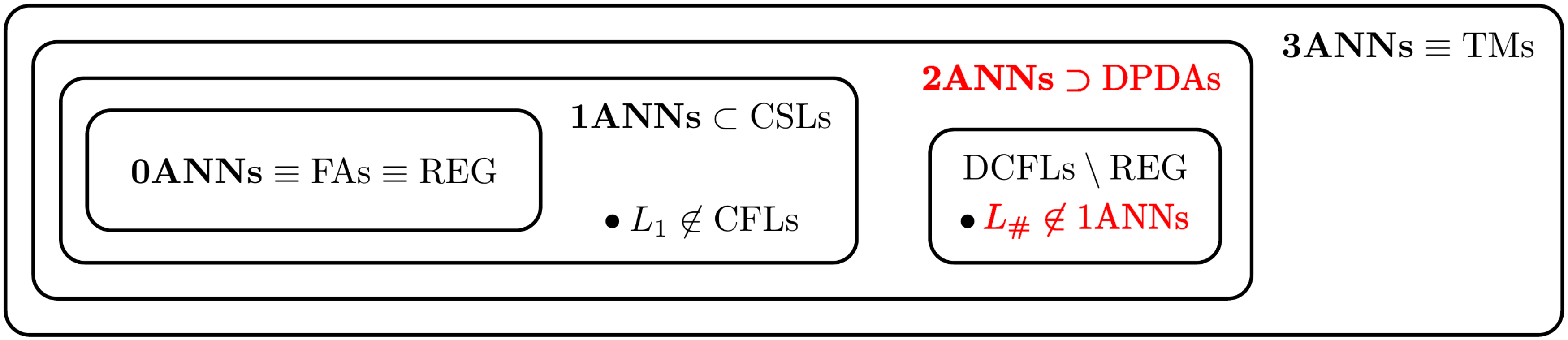
+ synchronizing the swaps between y_1 and y_2



Example of a 2ANN recognizing $L_{\#}$

A Summary of the Analog Neuron Hierarchy

$$\text{FAs} \equiv \text{0ANNs} \subsetneqq \text{1ANNs} \subsetneqq \text{2ANNs} \subseteq \text{3ANNs} \equiv \text{TMs}$$



Open Problems:

- the separation of the 3rd level: **2ANNs** $\subsetneqq$ **3ANNs** **?**
- strengthening the 2nd level separation to the nondeterministic CFLs:

$$(\textbf{CFLs} \setminus \textbf{REG}) \cap \textbf{1ANNs} = \emptyset \quad \textbf{?}$$

- a proper "natural" hierarchy of NNs between integer and rational weights which can be mapped to known infinite hierarchies of REG/CFLs **?**