FoNeCo: Analytical Foundations of Neurocomputing

Grant Project

Provider: The Czech Science Foundation (GA ČR)

Evaluation Panel: P103 Cybernetics and Information Processing

Type & Code: Standard grant project No. 19-05704S

Beneficiary: Institute of Computer Science, AS CR

Duration: 3 years (2019-2021)

Total Budget: approx. 5 million CZK

Team: 6 key researchers & 2 students

Researchers: Jérémie Cabessa (University of Paris), Jan Kalina, Věra Kůrková, Martin Plátek (Charles University), Petr Savický, Jiří Šíma (principal investigator)

Student Assistants (programmers): Tomáš Jurica (Nicole Tobišková), Jan Tichavský

Project Research

Motivations:

- successful softcomputing methods such as neurocomputing (e.g. deep learning) are of heuristic or statistical nature
- prevailing empirical research based on computer simulations using benchmark or practical training dataset
- demands for theoretical analysis and justification of used models which can even help in proposing more rigorous/efficient methods

Methodology:

(artificial) neural networks employed for brain modeling and engineering applications are formalized in mathematical definitions as idealized abstract machines (e.g. analog numerical parameters are considered to be true real numbers)

these formal models of NNs are investigated using the theoretical **tools:**
functional analysis, formal languages & automata theory, complexity theory, robust statistics, algorithmic and computational learning theory etc.

Research Directions

- classifying subrecursive NNs between finite automata (integer weights) and Turing machines (rational weights) within the Chomsky hierarchy
- bio-inspired NN model of Synfire Rings: Turing universality, learning algorithms
- approximation theory of NNs: estimating model complexity, suboptimal solutions of learning tasks
- robust fitting of NNs: robust estimators avoiding outliers in MLPs, RBFs, convolutional networks
- computational complexity of deep learning: modified loading problem with pre-trained subnetworks as an external oracle, analysis for new types of units (e.g. rectified linear unit ReLU)
- software & numerical experiments: implementation and testing of proposed robust and bio-inspired training algorithms

The Computational Power of NNs

depends on the information contents of weight parameters:

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

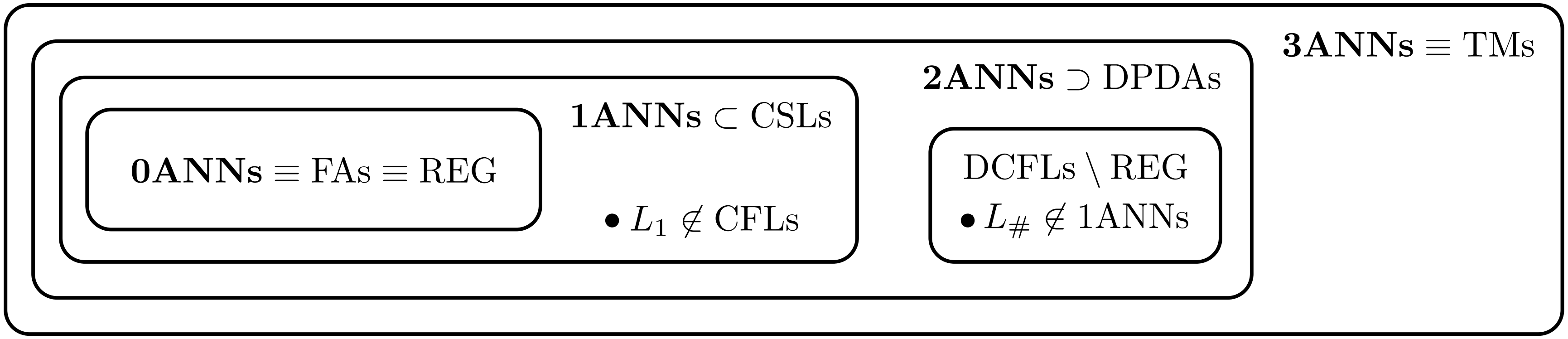
2. rational weights: Turing machine (Siegelmann, Sontag, 1995)
   polynomial time $\equiv$ complexity class P

   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

Analog Neuron Hierarchy

α**ANN** is a binary-state NN with α extra analog-state neurons and rational weights



$$\text{FAs} \equiv \text{0ANNs} \subsetneqq \text{1ANNs} \subsetneqq \text{2ANNs} \subseteq \text{3ANNs} = \text{4ANNs} = \ldots \equiv \text{TMs}$$

Classifying αANNs within the **Chomsky hierarchy** (Šíma, 2019):

integer-weight 0ANNs ≡ "quasi-periodic" 1ANNs ≡ FAs ≡ regular languages REG **(Type-3)**

1ANNs $\not\subset$ PDAs ≡ context-free languages CFLs **(Type-2)**

$\left(L_1 = \left\{x_1 \ldots x_n \in \{0,1\}^* \,\middle|\, \sum_{k=1}^{n} x_{n-k+1} \left(\frac{3}{2}\right)^{-k} < 1\right\} \in \text{1ANNs} \setminus \text{CFLs}\right)$

2ANNs ⊃ DPDAs ≡ deterministic context-free languages DCFLs

1ANNs ∩ DCFLs = 0ANNs $\quad \left(L_\# = \left\{0^n 1^n \,\middle|\, n \geq 1\right\} \in \text{DCFLs} \setminus \text{1ANNs}\right)$

1ANNs ⊂ LBAs ≡ context-sensitive languages CSLs **(Type-1)**

rational-weight 3ANNs ≡ TMs ≡ recursively enumerable languages **(Type-0)**

Contributions to General Theory

a **quasi-periodic number** characterizing 1ANNs that recognize regular languages:

for a fixed real base (radix) β ($|\beta| > 1$) and a finite set $A \neq \emptyset$ of real digits, every β-expansion

$$x = (0 \,.\, a_1\, a_2\, a_3 \ldots)_\beta = \sum_{k=1}^{\infty} a_k\, \beta^{-k} \qquad \text{where} \quad a_k \in A$$

is eventually quasi-periodic:

$$\Big(0 \,.\, \underbrace{a_1 \ldots a_{m_1}}_{\substack{\text{preperiodic}\\ \text{part}}}\ \underbrace{a_{m_1+1} \ldots a_{m_2}}_{\text{quasi-repetend}}\ \underbrace{a_{m_2+1} \ldots a_{m_3}}_{\text{quasi-repetend}}\ \underbrace{a_{m_3+1} \ldots a_{m_4}}_{\text{quasi-repetend}} \ldots\Big)_\beta$$

such that

$$(0 \,.\, \overline{a_{m_1+1} \ldots a_{m_2}})_\beta = (0 \,.\, \overline{a_{m_2+1} \ldots a_{m_3}})_\beta = (0 \,.\, \overline{a_{m_3+1} \ldots a_{m_4}})_\beta = \cdots$$

Example: the plastic $\beta \approx 1.324718$ ($\beta^3 - \beta - 1 = 0$), $A = \{0, 1\}$

$$1 = (0 \,.\, 0\ \underbrace{100}\ \underbrace{0\,011\,011\,1}\ \underbrace{0\,011\,1}\ \underbrace{100} \ldots)_\beta$$

with quasi-repetends: $(0 \,.\, \overline{100})_\beta = (0 \,.\, \overline{0(011)^i 1})_\beta = \beta$ for every $i \geq 1$

Contributions to General Theory (continued)

the **simplest non-regular deterministic context-free language**:

$$L_\# = \{0^n 1^n \mid n \geq 1\}$$

can be reduced to any language in the class DCFLs $\setminus$ REG by a finite automaton

a counterpart to the hardest problems in the complexity classes such as NP-complete problems

Publications Already Dedicated to FoNeCo Project:

Journals:

1. M. Marozzi, A. Mukherjee, J. Kalina: Interpoint distance tests for high-dimensional comparison studies. *Journal of Applied Statistics*, 2019. (in print)
2. J. Šíma: Subrecursive neural networks. *Neural Networks*, 116:208–223, 2019.
3. J. Šíma: Analog neuron hierarchy. 47p., 2019. (submitted)

Conferences:

1. J. Cabessa, J. Šíma: Robust optimal-size implementation of finite state automata with synfire ring-based neural networks. Proceedings of ICANN 2019, LNCS 11727, Springer, 2019.
2. J. Cabessa, A. Villa: A memory-based STDP rule for stable attractor dynamics in Boolean recurrent neural networks. Proceedings of IJCNN 2019, IEEE, 2019.
3. J. Kalina, N. Tobišková, J. Tichavský: A nonparametric bootstrap comparison of variances of robust regression estimators. Proceedings of MME 2019, MatfyzPress, 2019.
4. J. Kalina, P. Vidnerová: Implicitly weighted robust estimation of quantiles in linear regression. Proceedings of MME 2019, MatfyzPress, 2019.
5. J. Kalina, P. Vidnerová: Robust training of radial basis function networks. Proceedings of ICAISC 2019, LNAI 11508, pp. 113–124, Springer, 2019.
6. V. Kůrková: Probabilistic bounds for approximation by neural networks. Proceedings of ICANN 2019, LNCS 11727, Springer, 2019.
7. J. Šíma: Counting with analog neurons. Proceedings of ICANN 2019, LNCS 11727, Springer, 2019.
8. J. Šíma, M. Plátek: One analog neuron cannot recognize deterministic context-free languages. 2019. (submitted)