# Efficient Formal Verification of Neural Networks

## Efficient Formal Verification of Neural Networks

**Keywords**: Formal Verification, Neural Network, Why3, CAISAR, Yices

### Institution

The French Alternative Energies and Atomic Energy Commission (CEA) is a key player in research, development, and innovation. Drawing on the widely acknowledged expertise gained by its 16,000 staff spanned over 9 research centers with a budget of 4.1 billion Euros, CEA actively participates in more than 400 European collaborative projects with numerous academic (notably as a member of Paris-Saclay University) and industrial partners. Within the CEA Technological Research Division, the CEA List institute addresses the challenges coming from smart digital systems.

Among other activities, CEA List's Software Safety and Security Laboratory (LSL) research teams design and implement automated analysis in order to make software systems more trustworthy, to exhaustively detect their vulnerabilities, to guarantee conformity to their specifications, and to accelerate their certification. The lab recently extended its activities on the topic of AI trustworthiness and gave birth to a new research group: AISER (Artificial Intelligence Safety, Explainability and Robustness).

### Scientific context

Formal verification of neural network is a very active field, using techniques ranging from abstract interpretation (Singh et al. 2019), interval bound propagation (Wang et al. 2021), Mixed Integer Linear Programming (Wong and Kolter 2017) or Satisfaction Modulo Theory (SMT) calculus (Katz et al. 2019). SMT-based approach are of particular interest, because neural networks have a conceptually simple structure: no loops and almost linear arithmetic. The main obstacle is computing non-linear activation functions, which has exponential complexity. (Katz et al. 2019) approach uses temporary bound relaxations to reduce the search space. Preliminary experiments in our lab showed another promising venue: using Fourier-Motzkin variable elimination within the MCSAT framework, along with activation pattern learning.

## Internship

The aim of this internship is to study the applicability of variable substitutions to accelerate SMT solvers on neural network formal verification. To do so, the intern will use the CAISAR open-source platform to manipulate the neural networks control flow as logical formulaes. A suggested SMT solver to test along CAISAR is Yices, as it has a native OCaml API, which is incidentally the language CAISAR is written in.

The broad internship goals are:

- familiarization with the state-of-the-art on neural network verification
- implementation of a prototype variable elimination heuristic within the CAISAR platform
- benchmark on selected datasets and verification problems

## Qualifications

The candidate will work at the crossroads of formal verification and artificial intelligence. As it is not realistic to be expert in both fields, we encourage candidates that do not meet the full qualification requirements to apply nonetheless. We strive to provide an inclusive and enjoyable workplace. We are aware of discriminations based on gender (especially prevalent on our fields), race or disability, we are doing our best to fight them.

- **Minimal**
  - Master student or equivalent (2nd/3rd engineering school year) in computer science
  - knowledge of OCaml
  - knowledge of SMT solving, First Order Logic
  - ability to work in a team, some knowledge of version control
- **Preferred**
  - notions of AI and neural networks
  - knowledge of Why3

## Characteristics

The candidate will be monitored by two research engineers of the team.

- **Duration:** 5 to 6 months from early 2023
- **Location:** CEA Nano-INNOV, Paris-Saclay Campus, France
- **Compensation:**
  - €700 to €1300 monthly stipend (determined by CEA compensation grids)

  - maximum €229 housing and travel expense monthly allowance (in case a relocation is needed)
  - CEA buses in Paris region and 75% refund of transit pass
  - subsidized lunches
  - 3 days of remote work

## Application

If you are interested in this internship, please send to the contact persons an application containing:

- your resume;
- a cover letter indicating how your curriculum and experience match the qualifications expected and how you would plan to contribute to the project;
- your bachelor and master 1 transcripts;
- the contact details of two persons (at least one academic) who can be contacted to provide references.

Applications are welcomed until the position is filled. Please note that the administrative processing may take up to 3 months.

## Contact persons

For further information or details about the internship before applying, please contact:

- Julien Girard-Satabin (julien.girard2@cea.fr) (also available on LinkedIn)
- François Bobot (francois.bobot@cea.fr)

## References

Katz, Guy, Derek A. Huang, Duligur Ibeling, Kyle Julian, Christopher Lazarus, Rachel Lim, Parth Shah, et al. 2019. “The Marabou Framework for Verification and Analysis of Deep Neural Networks.” In *Computer Aided Verification*, edited by Isil Dillig and Serdar Tasiran, 11561:443–52. Cham: Springer International Publishing.

Singh, Gagandeep, Timon Gehr, Markus Püschel, and Martin Vechev. 2019. “An Abstract Domain for Certifying Neural Networks.” *Proceedings of the ACM on Programming Languages* 3 (POPL): 1–30.

Wang, Shiqi, Huan Zhang, Kaidi Xu, Xue Lin, Suman Jana, Cho-Jui Hsieh, and J. Zico Kolter. 2021. “Beta-CROWN: Efficient Bound Propagation with Per-neuron Split Constraints for Complete and Incomplete Neural Network Robustness Verification.” October 31, 2021. http://arxiv.org/abs/2103.06624.

Wong, Eric, and J. Zico Kolter. 2017. “Provable Defenses Against Adversarial Examples via the Convex Outer Adversarial Polytope.” *Proceedings of the 35th International Conference on Machine Learning*, November. https://arxiv.org/abs/1711.00851v3.