Centre Nacional de Microelectrònica IMB

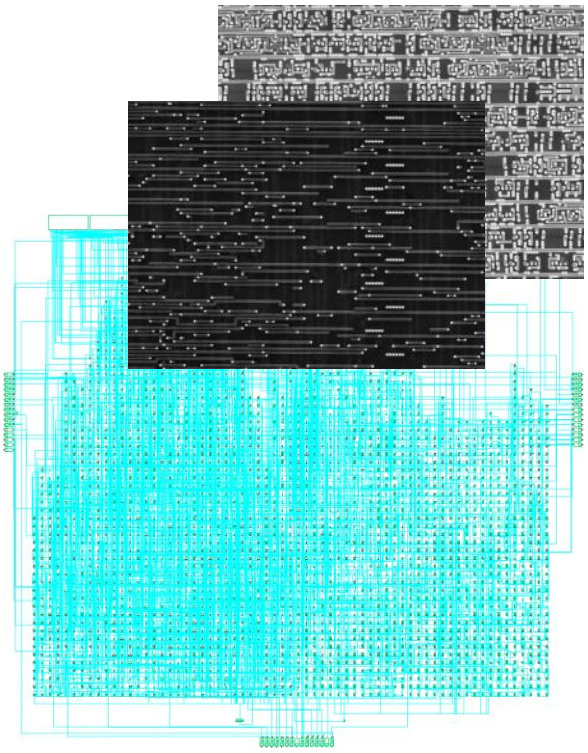
CSIC
Consejo Superior de Investigaciones Científicas

Master Thesis (with possibility of Engineering position)
at the Reverse Engineering research group of IMB-CNM(CSIC)

Automatic Dataset Generator for AI training in Integrated Circuit Architecture Extraction

Description

The reverse engineering of large-scale integrated circuits (ICs) is highly demanded in fields like competitive analysis, patent protection or digital forensics, among others. With chips exceeding one billion transistor count, the complexity of digital ICs is leading to the automation of most of the reverse engineering tasks in order to achieve reliable results in a reasonable time. In particular, the extraction of the chip architecture at register-transfer logic (RTL) level from the flat sea-of-gates netlist obtained through physical imaging still remains a very time-consuming semi-manual process, which is prone to introduce bugs in the architecture analysis. New AI-based methods may solve this problem, but they will surely require a huge dataset of IC netlists properly labeled to train their machine learning models. This project aims to create a new tool to automatically generate a large dataset of digital synthesis cases for some relevant functional blocks defined with hardware description languages (HDLs). Indeed, dataset growing techniques will be based on variability from synthesis algorithms, target CMOS technologies, block functional parameters and optimization figures.



Background and skills

- Electronic engineering or any similar curriculum covering the following topics: CMOS technology basics, digital circuit design and semi-custom IC design methodology.
- Knowledge of EDA tools and HDLs for semi-custom IC design.
- Experience in Python programming language.
- Capability of working as a team.
- Good spoken and written English.

Tasks

The student will develop a Python-based tool for the automation of logic synthesis through Cadence EDA tools. The input will consist on digital functional models described in Verilog, while the output will be a dataset containing flat netlists of logic gates properly labeled for their use in the training of graph machine learning models. Indeed, the tool will be capable of managing standard-cell logic libraries of CMOS technology nodes from 350 nm to 22 nm in order to artificially create subset libraries for the purpose of dataset growing. Furthermore, the tool interface will allow the control of the synthesis variability through optimization figures like timing, power and area. The student will be also in charge of testing the new tool with some digital functions of interest. All the above tasks will be performed in the IMB-CNM lab facilities at the UAB Bellaterra Campus. The success of the proposed Master Thesis could lead to a paid continuation of the work.

Contact Jofre Pallarès Cuxart
jofre.pallares@imb-cnm.csic.es
Dr. Francesc Serra Graells
paco.serra@imb-cnm.csic.es