Analyzing the Divide between FPGA Academic and Commercial Results

Presentation at FPT 2015,
Queenstown, New Zealand

Elias Vansteenkiste
Ghent University, Belgium
Elias.Vansteenkiste@ugent.be

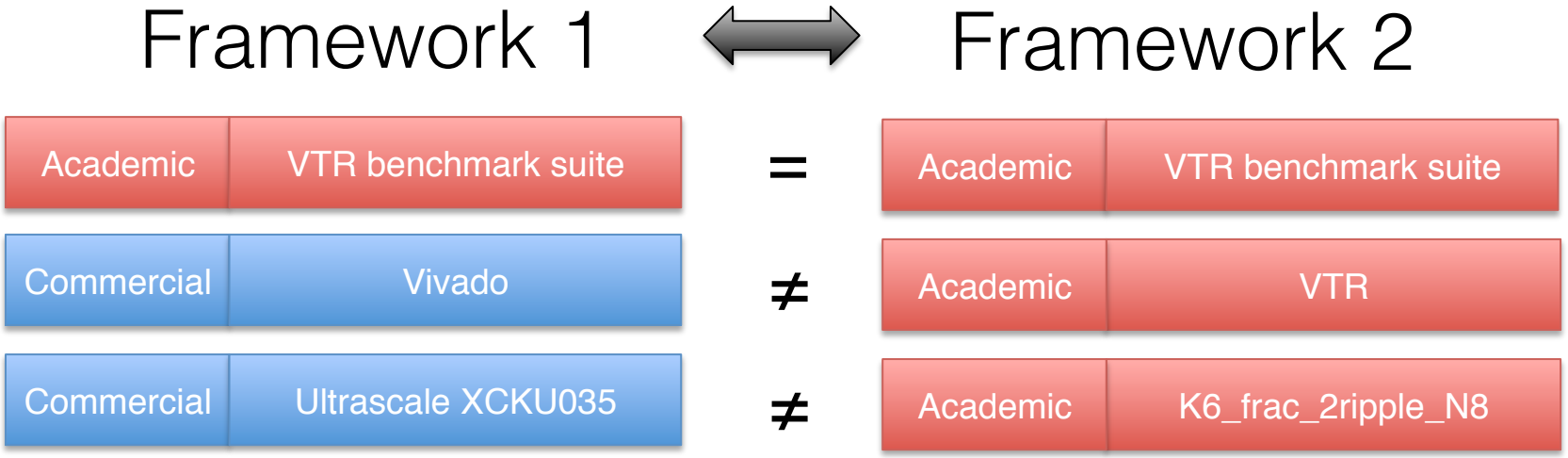
Alireza Kaviani and Henri Fraisse
Xilinx CTO Research Lab
{Alireza.Kaviani, Henri.Fraisse}@xilinx.com

Introduction

- Increasing capacity and performance of FPGAs
  - Process technology
  - Architecture
  - CAD tools
  - Contributions both from industry and academia
- Recent years: fewer architecture and tool ideas seeded from academic community!
- One possible reason:
  - Significant performance gap between the academic and commercial framework

Evaluation Frameworks

- Comparison with latest commercial tools and architectures



- Measures: Speed-performance, Area efficiency and Runtime

Quality of Results F1/F2

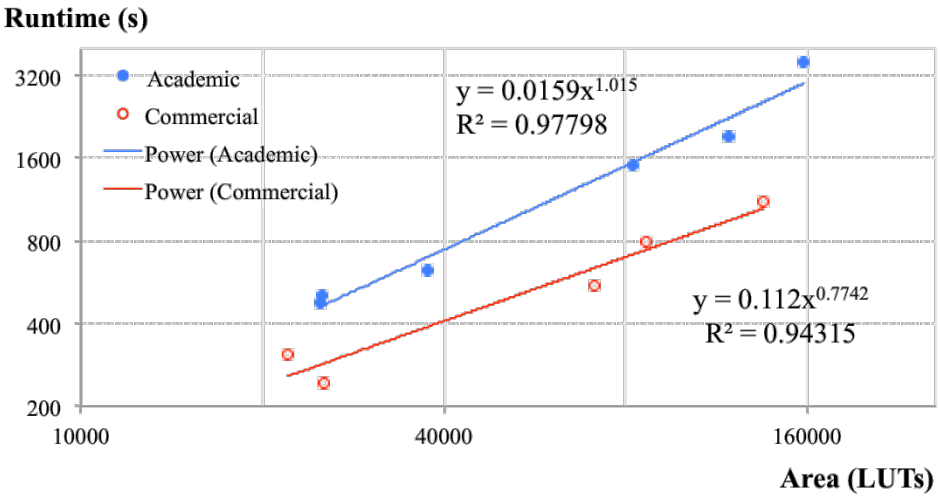
	Geomean	Geo. Std Dev
Speed-performance	2.24 x	1.45x
Area-efficiency	0.95 x	1.52x

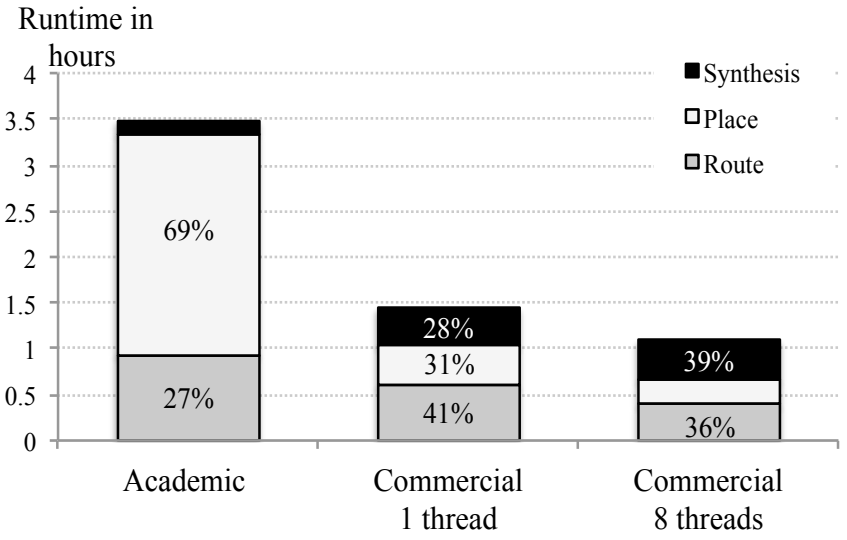
- Speed-performance gap is worrisome
- Area-efficiency gap is acceptable

Runtime

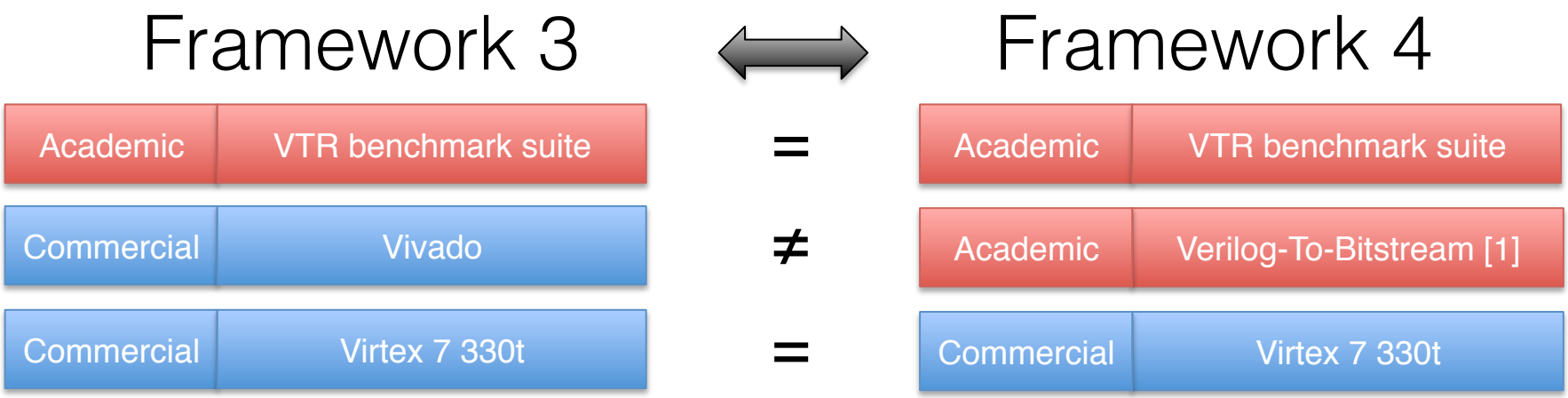
- Geomean: 2.2x
- Geo. Std. Dev.: 1.3x
- RT-size scalability:
  - RT gap x2 for each 60K LUT increase

- RT breakdown:
  - Placement: relative large runtime consumer
  - Small benchmark designs, easy to route





Architecture gap versus Tool gap

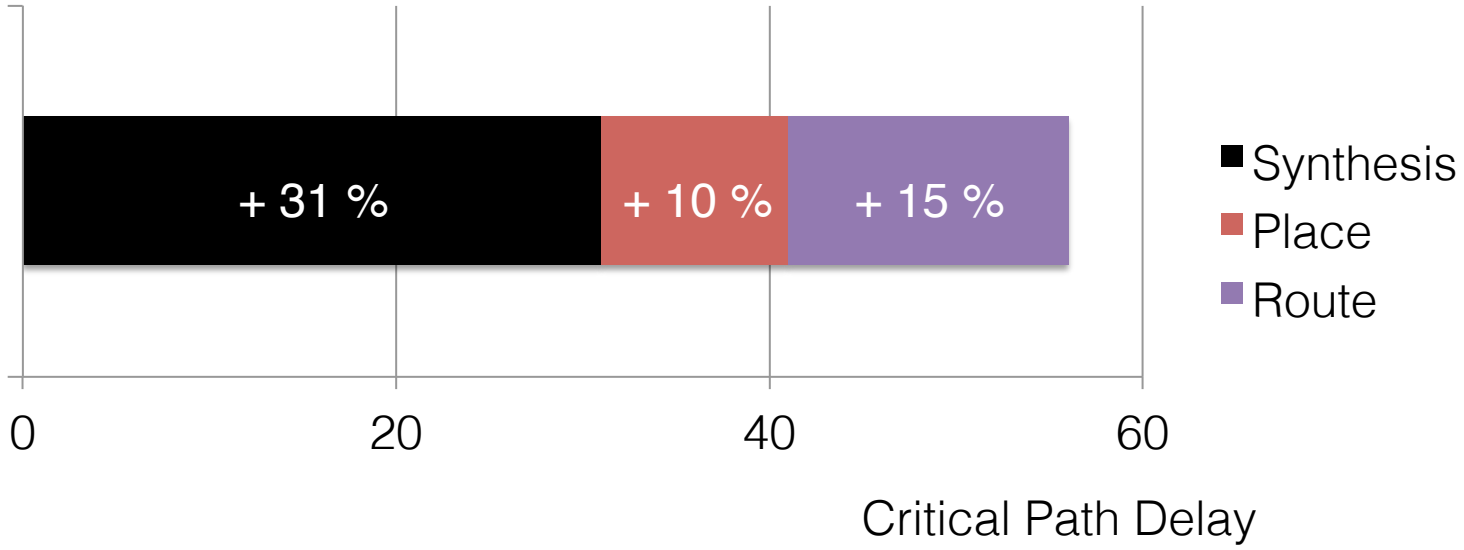


	F1 vs F2	F3 vs F4
Speed-performance	2.2x	2.1x
Area efficiency	-5%	-25%
Runtime	2.2x faster	5.5x faster

Architecture gap?
VTR not optimized to exploit commercial architectural features

[1] E.Hung, "**Mind The (Synthesis) Gap: Examining Where Academic FPGA Tools Lag Behind Industry**," in FPL 2015

Mind The Synthesis Gap

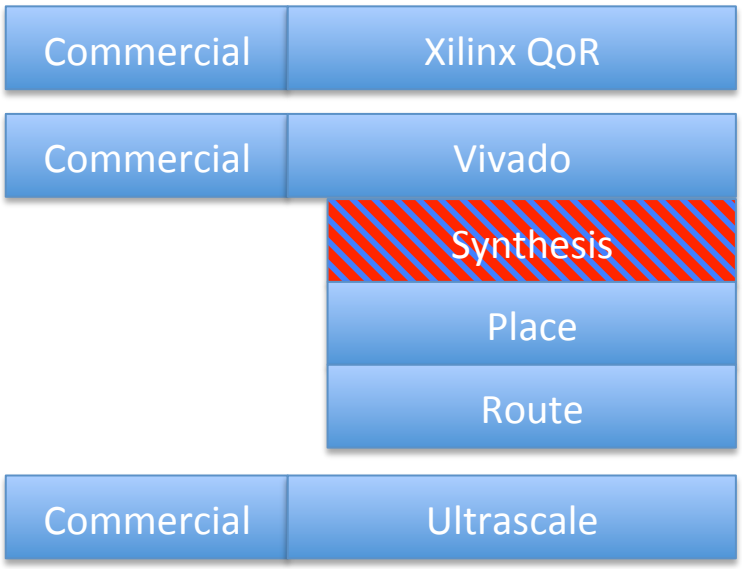


Reproduced from [1]

[1] E.Hung, "**Mind The (Synthesis) Gap: Examining Where Academic FPGA Tools Lag Behind Industry**," in FPL 2015

Hybrid Flow: ABC-Vivado

- Combine merits of academic tools with the credibility of commercial tools.



ABC is used for logic optimization and technology mapping

ABC-Vivado: Motivations

- Evaluation of
  - Quality of logic optimization and technology mapping
  - Architecture and tool optimization ideas, for example [2], [3], [4], [5], …

[2] S. Ray, A. Mishchenko, N. Een, R. Brayton, S. Jang, and C. Chen, "**Mapping into LUT structures**," in *DATE 2012*, pp. 1579-1584.

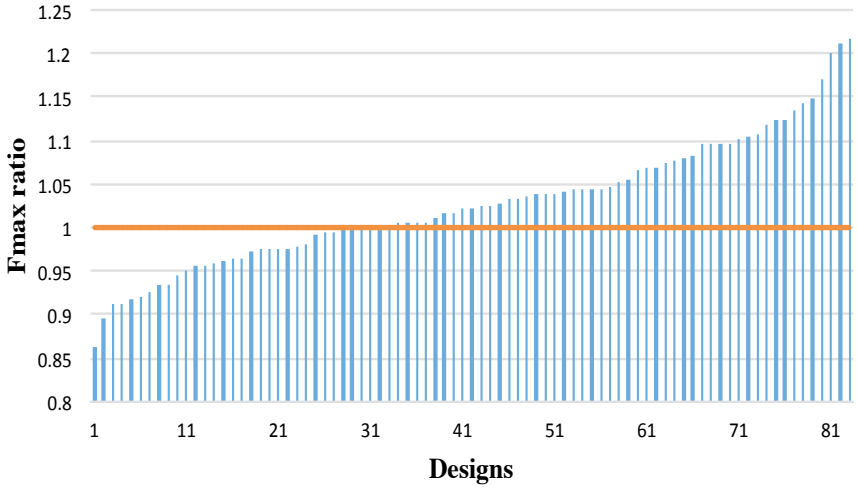
[3] A. Mishchenko, "**Enumeration of irredundant circuit structures**," in *IWLS* 2014.

[4] H. Parandeh-Afshar, H. Benbihi, D. Novo, and P. Ienne, "**Rethinking FPGAs: Elude the Flexibility Excess of LUTs with And-Inverter Con**es," in *FPGA 2012*

[5] A. Mishchenko, R. Brayton, S. Jang and V. Kravitz, "**Delay Optimization using SOP balancing**," in ICCAD *2011*

Hybrid Flow Results

- Initial results: basic script
  - All metrics within 5% except for runtime
  - Logic opt. and tech. mapping NOT the reason for the divide
- Best ABC script
  - Consists of multiple iterations of
    - SOP balancing [5]
    - Mapping to 6 LUTs
  - Retaining the best result
  - Fmax +2.8%
  - Area (CLBs) -1.8%
- At the expense of runtime



[5] A. Mishchenko, R. Brayton, S. Jang and V. Kravitz, "**Delay Optimization using SOP balancing**," in ICCAD *2011*

Academic Benchmark Suites

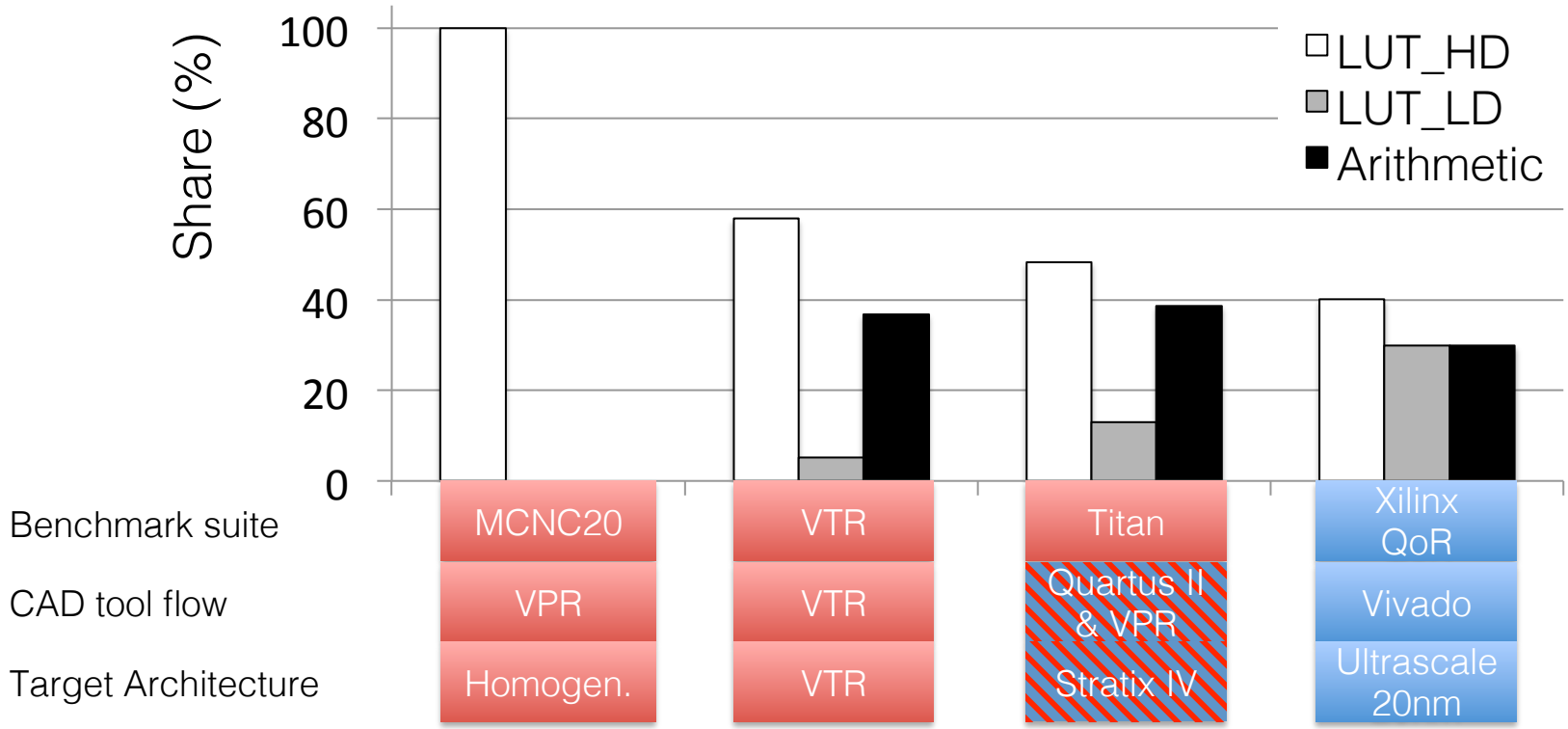
- Can't separate benchmark designs from their framework
- Overview of academic frameworks
  - 20 MCNC – VPR - homogeneous LUT-only architecture (48nm) [6][7][8]
  - VTR - k6_frac_2ripple_N8_22nm
  - Titan – Quartus II & VPR - Stratix-IV blocks and academic routing architecture

[6] A. Petkovska, D. Novo, A. Mishchenko and P. Ienne, "**Constrained interpolation for guided logic synthesis**," in *ICCAD 2014*.
[7] DeHon, André, and Nikil Mehta**.** "**Exploiting partially defective LUTs: Why you don't need perfect fabrication,**" in *ICFPT 2013*
[8] P. E. Gaillardon, X. Tang, G. Kim and G. De Micheli (2015). "**A Novel FPGA Architecture Based on Ultrafine Grain Reconfigurable Logic Cells.**" in TVLSI 2015

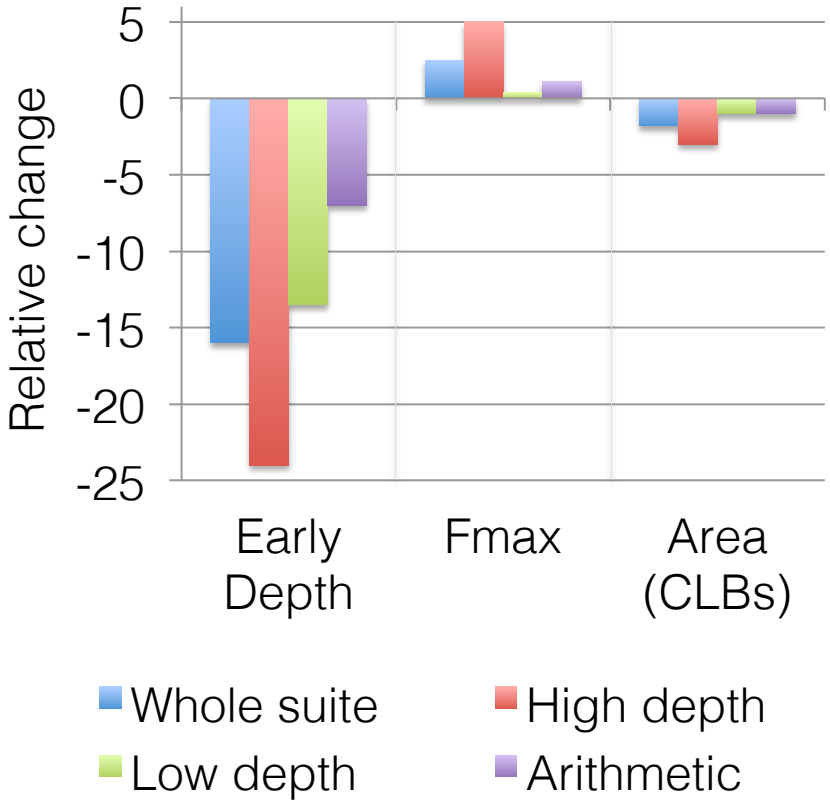
Commercial Design Suite
Xilinx QoR – Vivado - Ultrascale

- >80 designs
  - Mostly customer designs
- 100K<size<3M blocks
- Noticeable difference:
  - depth profile of the critical zone

Depth Profile Comparison



Results of the ABC-Vivado Flow



	Early Depth	Fmax	Area
Whole suite	-16	**+2,5**	-1,8
High depth	-24	**+5**	-3
Low depth	-13,5	**+0,4**	-1
Arithmetic	-7	**+1,1**	-1

- Less emphasis on early depth reduction
- Uncertainty about post routing critical path

Conclusions (1)

- Has the divide grown beyond acceptance?
  - Speed-performance 2.2x
  - Hard to assess the merits of academic ideas
- Solutions
  - Address the divide by leveraging the commercial frameworks as much as possible
  - Hybrid frameworks lead to credible results
    - The Vivado–ABC flow is able to improve QoR with 5% for HD designs
    - requires joint cooperation of academic and commercial interested parties

Conclusions(2)

- Benchmark design suites
  - Trend towards low depth, highly pipelined designs
  - Focus research on retiming instead of early depth reduction.
- Evaluation framework is available online
  - github.com/EliasVansteenkiste/EvaluationFramework

Questions