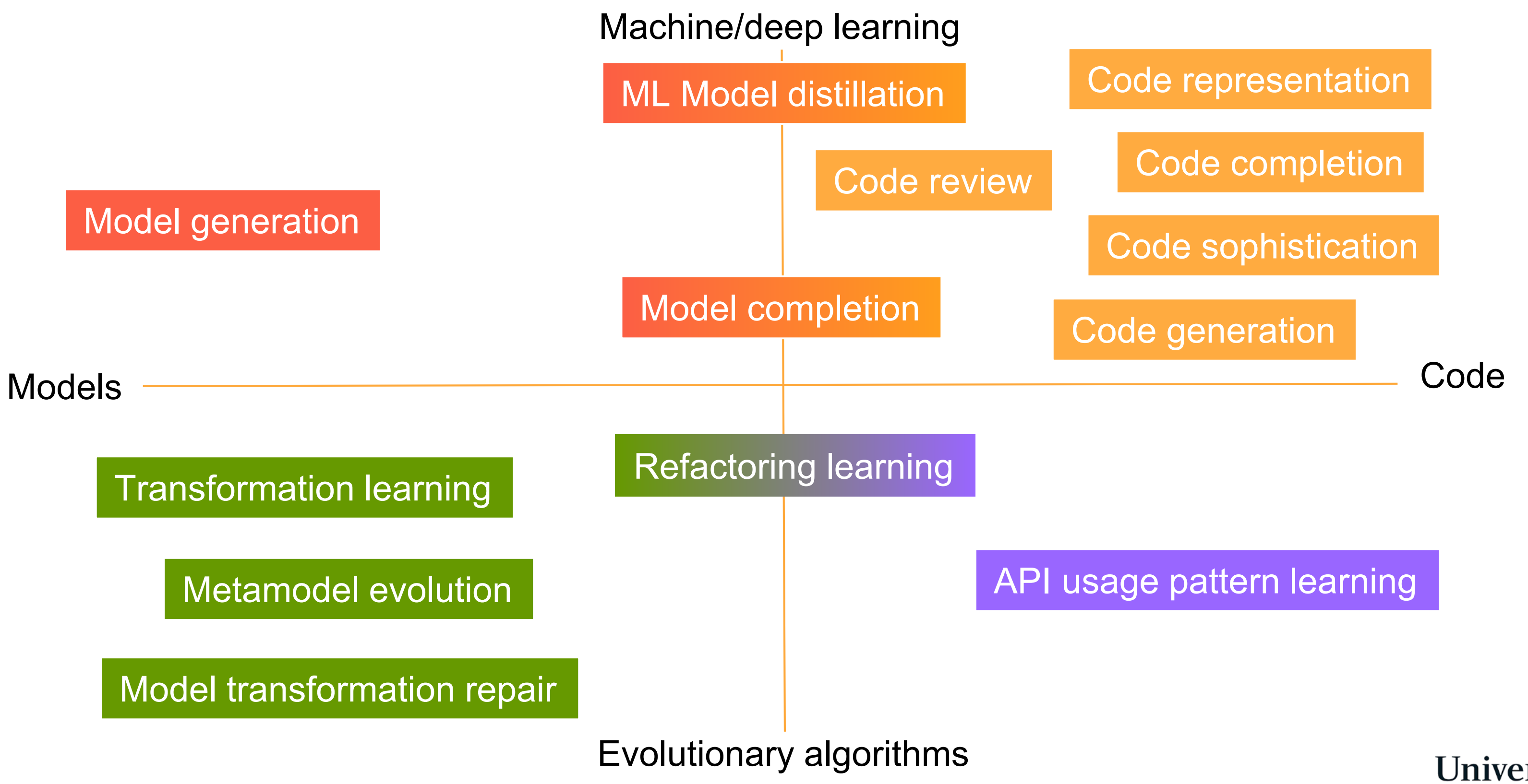
AI-enabled software-development automation

Houari Sahraoui

SEMTL 09/2022

Artificial Intelligence for Software Engineering



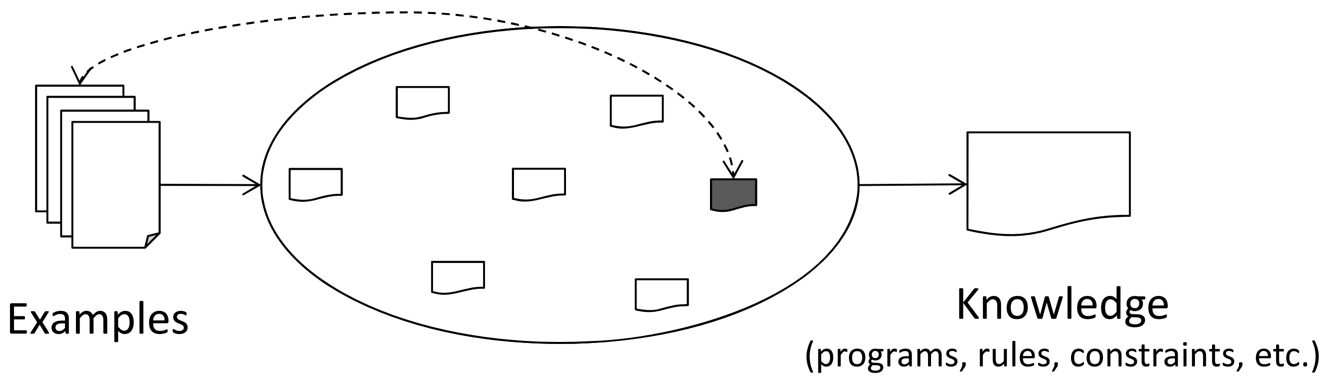
Software Engineering for Artificial Intelligence

- Supporting scientists in using ML/DL algorithms
  - From a scientific problem to a ML/DL implemention: a conceptual framework
  - Methods and tools to support non-computer scientists to use Ml/DL

SE and AI for Cyber-Bio-Physical Systems

- Digital shadows and digital twins
- AI to improve simulation
- Simulation to improve design space exploration
- Application to smart agriculture (vertical farming)

Learning with GP



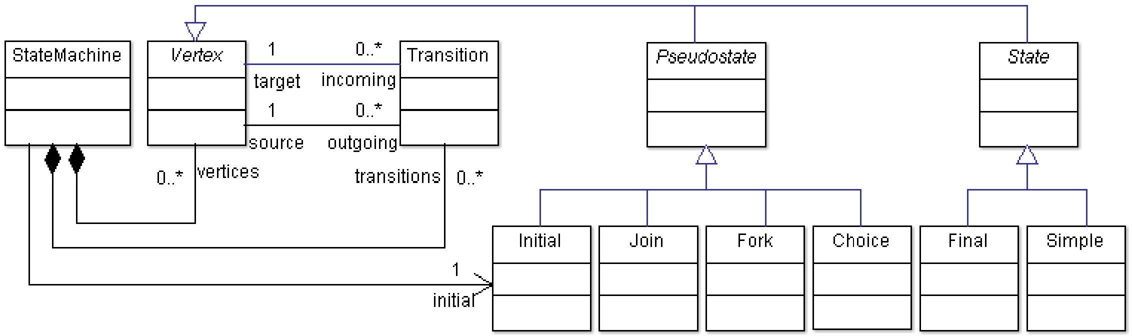
- Issues
  - Large (infinite) search space
  - Bloating
  - Single fitness peak

Genetic Programming vs ML/DL

- Genetic programming
  - 👍 Small amount of examples to learn from
  - 👍 Allow to learn complex executable artefacts
  - 👎 Low generalizability
- ML/DL
  - 👎 Large data sets to learn from or to fine-tune existing models
  - 👎 Work well only with small artefacts
  - 👍 High generalizability

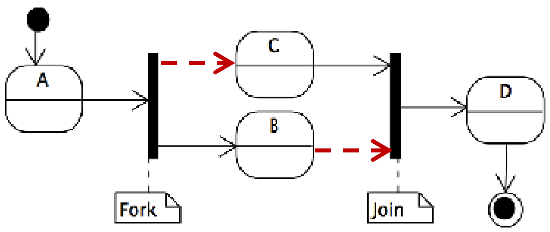
Learning Well-Formedness Rules with GP

- Why learning WFRs?



context Fork inv :
self.outgoing −> size() > 1

context Join inv :
self.incoming −> size() > 1

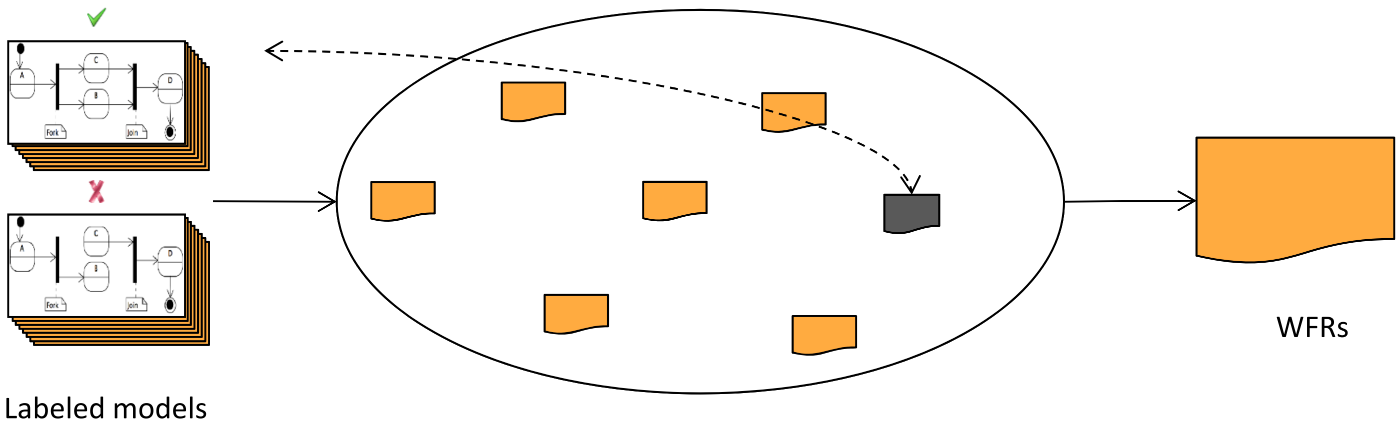


Université de Montréal

geodes

Learning Well-Formedness Rules

- Learning principle

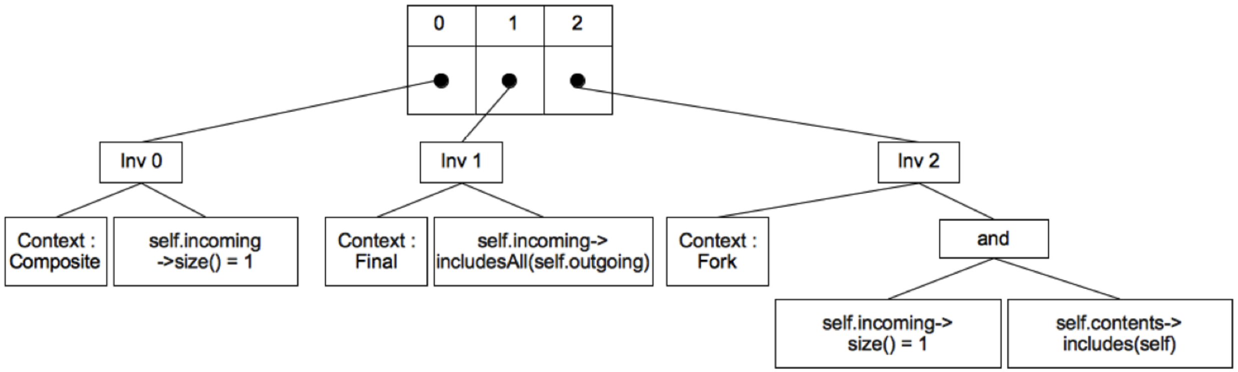


Faunes, M., Cadavid, J., Baudry, B., Sahraoui, H., & Combemale, B. Automatically searching for metamodel well-formedness rules in examples and counter-examples. MODELS2013

Université de Montréal

Learning Well-Formedness Rules

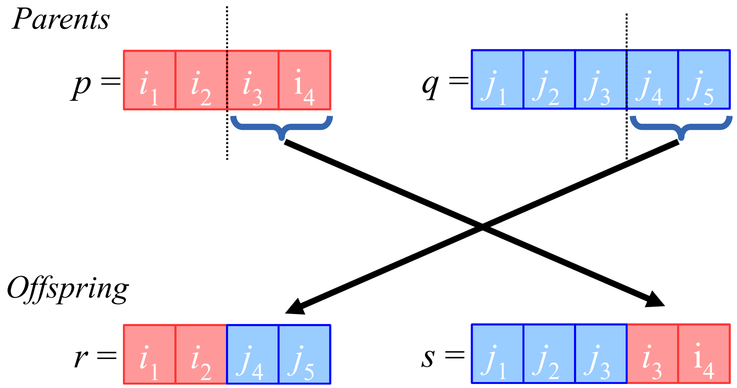
- Candidate solution
  - Set of OCL invariants



Learning Well-Formedness Rules

- Genetic operators

Crossover



Mutation

- At invariant set level
  - Add a new invariant
  - Delete an invariant
  - Combine two invariants with “implies”
- At invariant level
  - Change a comparison operator
  - Change a logical operator
  - Incrementing/decrementing a numerical constant
  - Replace an attribute or a reference (same type and context).
  - Replacing a sub-tree.
  - Negating an invariant

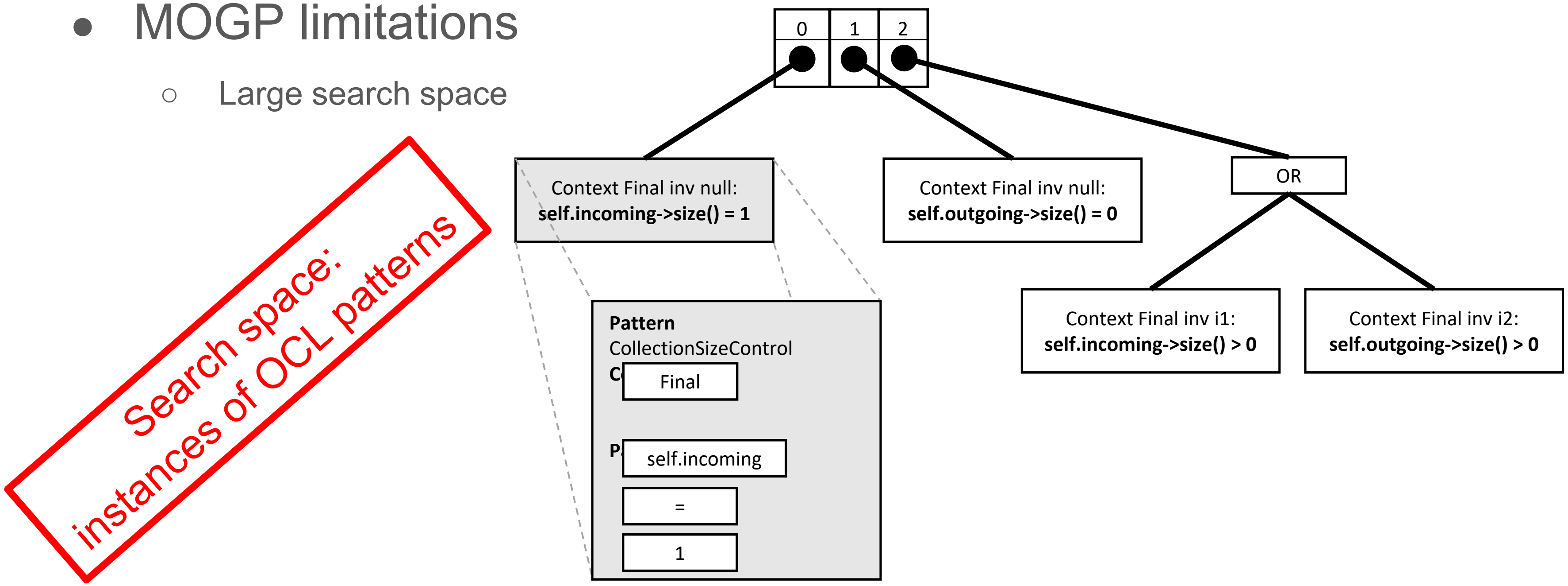
geodes

Learning Well-formedness rules

- Objectives
  - Classification accuracy of valid and invalid examples
  - Size of the WFR set
    - # leafs in the tree of WFR set, normalized between 0 and 1

Learning Well-formedness rules

- MOGP limitations
  - Large search space



Gómez, J. J. C., Baudry, B., & Sahraoui, Searching the boundaries of a modeling space to test metamodels. ICST *2012*

Learning Well-formedness rules

- MOGP limitations
  - Single fitness peak

Problem

Examples

	a	b	c
Ex1	2	3	7
Ex2	2	4	8
Ex3	2	5	9
Ex4	2	1	5

Solution

c = 2a + b

Fitness of a solution = % of correct examples

***fit*(c = 2a + b) =**

100%

Learning Well-formedness rules

- MOGP limitations
  - Single fitness peak

Problem

Examples

	a	b	c
Ex1	2	3	7
Ex2	2	4	8
Ex3	2	5	9
Ex4	2	1	5
Ex5	3	7	16

Solution

~~**c = 2a + b**~~

$c = a^2 + b$

Fitness of a solution = % of correct examples

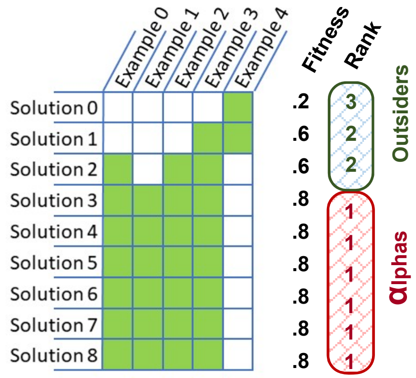
***fit*(c = 2a + b) =** 

80%

Université de Montréal

Learning Well-formedness rules

- MOGP limitations
  - Single fitness peak



~~$c = a^2 + b$~~

$c = 2a + b$

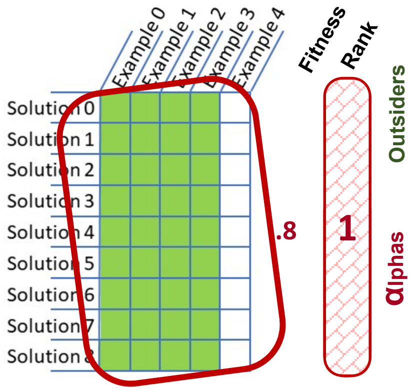
Some examples are **more frequently solved in general..**

..and **rare cases** might be neglected !

- Diversity of solutions decreases

Learning Well-formedness rules

- MOGP limitations
  - Single fitness peak



~~$c = a^2 + b$~~

$c = 2a + b$

- Diversity of solutions decreases

Some examples are **more frequently solved in general..**

..and **rare cases** might be neglected !

Université de Montréal

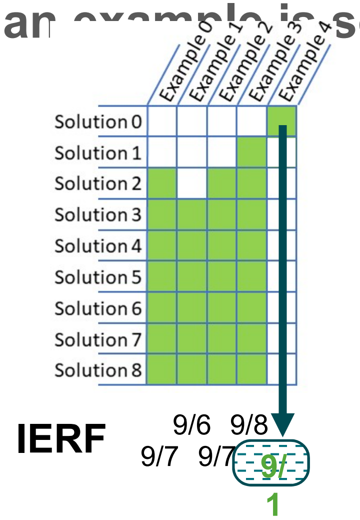
geodes

Social Diversity

- Fitness of a solution
  - Proportional to the **number of examples solved**
  - Offset by the **frequency of which an example is solved** by the population's individuals (IERF)

An example is **less frequently solved..?**

..it becomes **more valuable !**

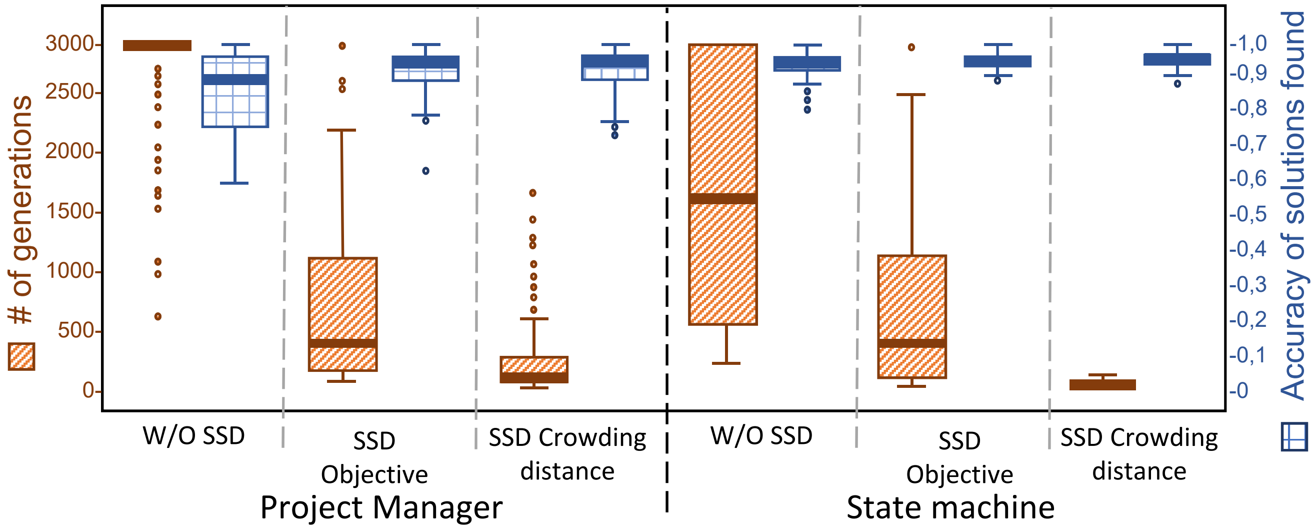


Batot, E., & Sahraoui, H. Promoting social diversity for the automated learning of complex MDE artifacts. *Software and Systems Modeling*, 2022.

Université de Montréal

Social Diversity

- Evaluation: accuracy and convergence

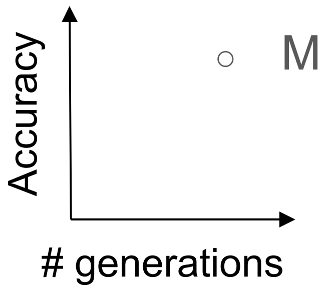


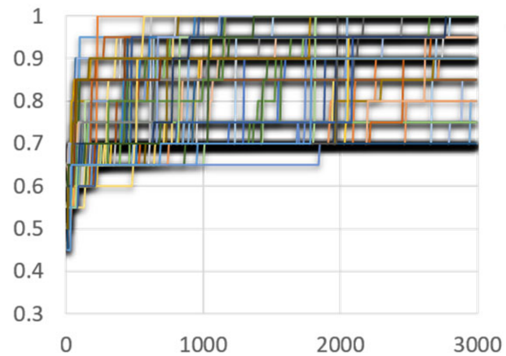
Université de Montréal

geodes

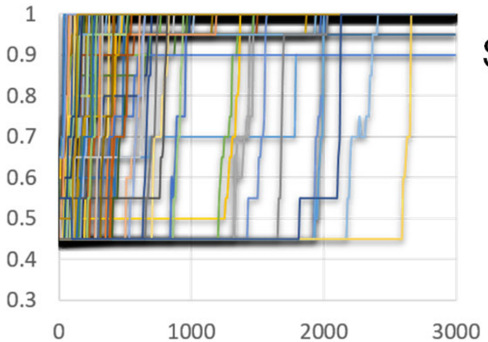
Social Diversity

- Evaluation: sensitivity
  - Multiple runs

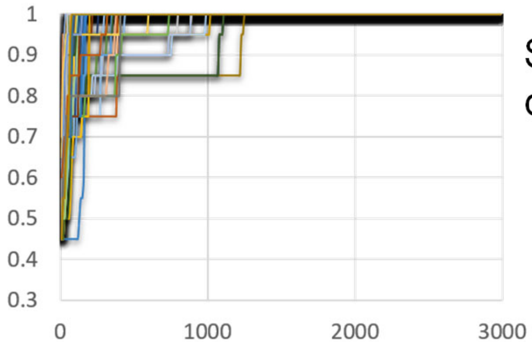




w/o SSD



SSD objective

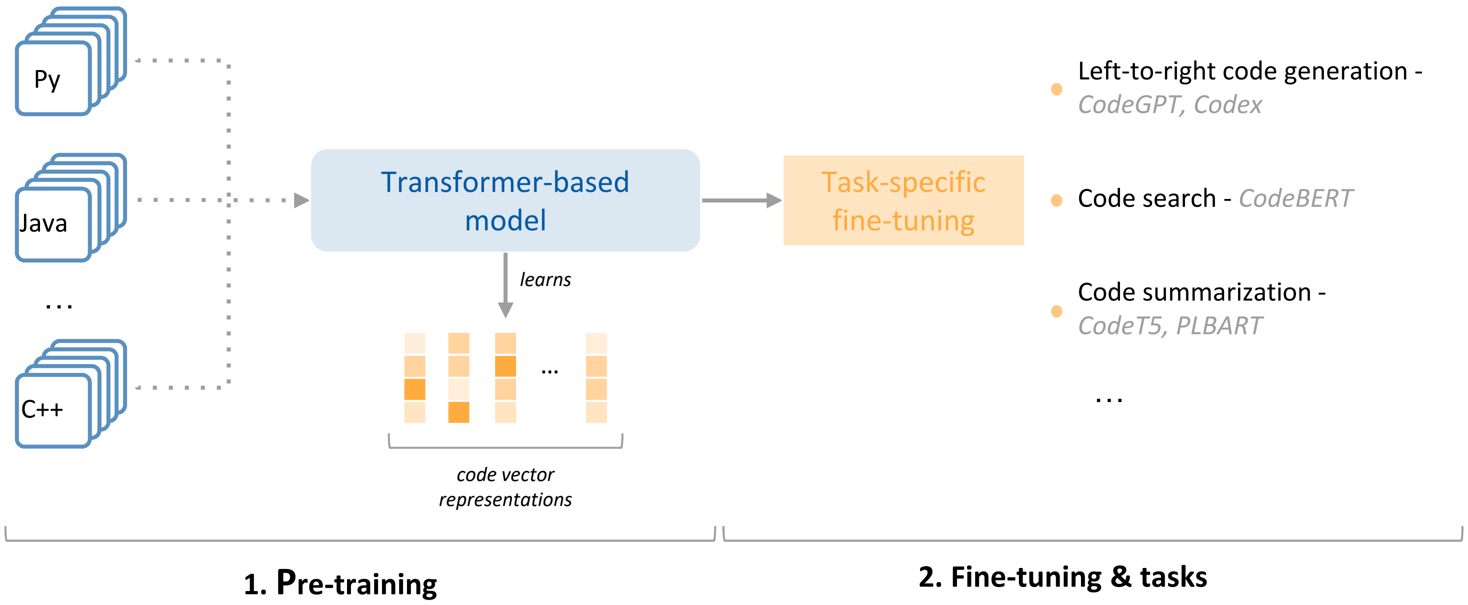


SSD crowding distance

Université de Montréal

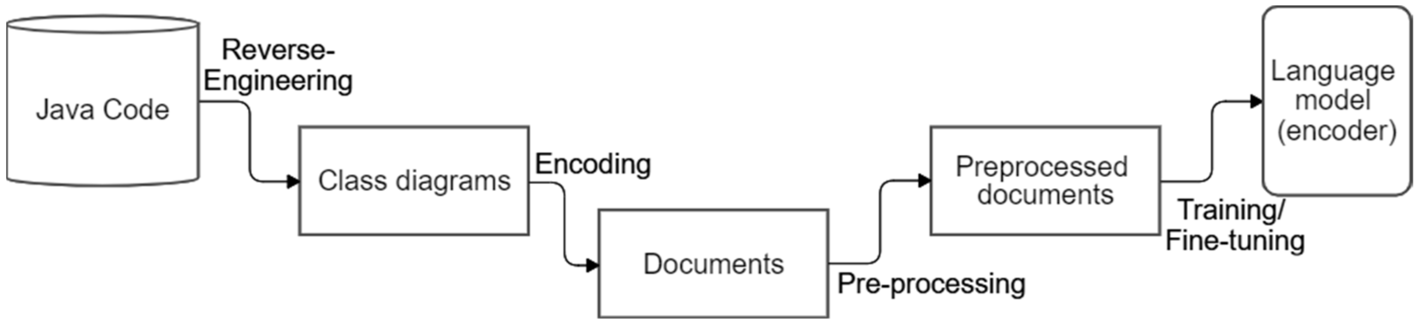
(Deep) Learning Software Artefacts

- State-of-the-art learning paradigm



(Deep) Learning Software Artefacts

- Dealing with data scarcity for software modeling tasks

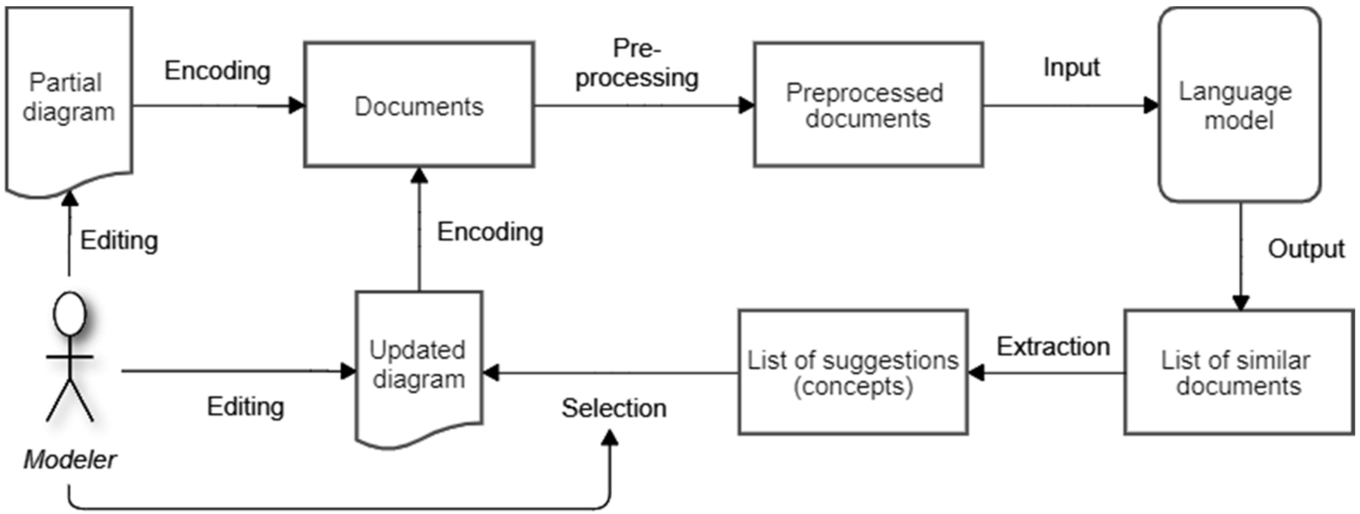


Capuano, T., Sahraoui, H., Freney, B., Vanderose, B., Learning from Code Repositories to Recommend Model Classes, ECMFA/JOT 2022

geodes

(Deep) Learning Software Artefacts

- Model completion using code repositories

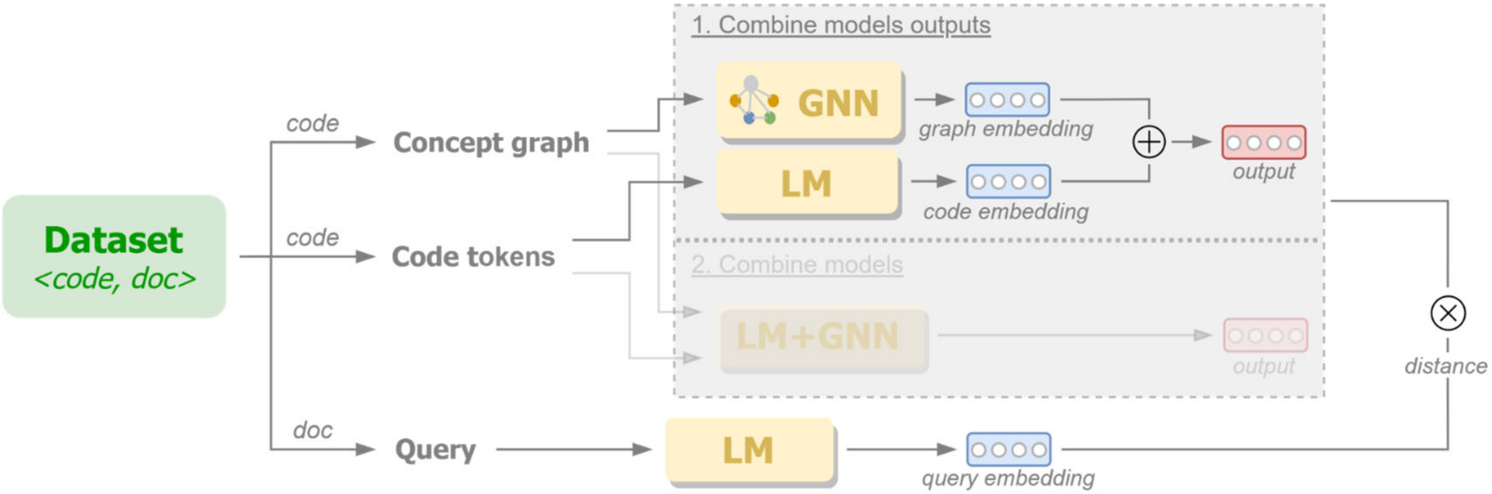


Capuano, T., Sahraoui, H., Freney, B., Vanderose, B., Learning from Code Repositories to Recommend Model Classes, ECMFA/JOT 2022

Université de Montréal

(Deep) Learning Software Artefacts

- Multimodal learning with code concept graphs

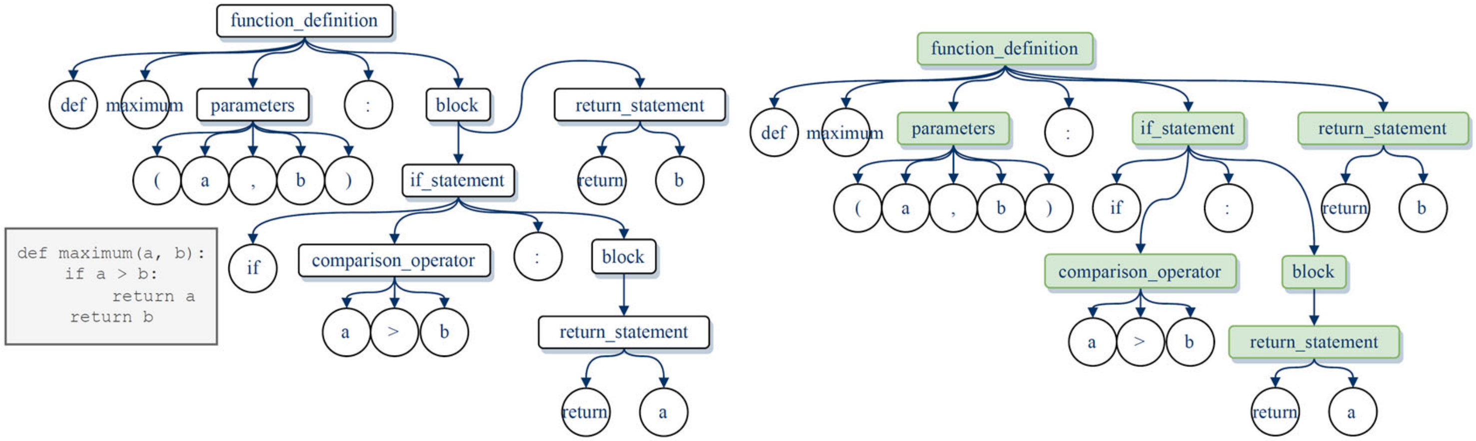


Weyssow, M.; Sahraoui, H.; Liu, B.; Better Modeling the Programming World with Code Concept Graphs-augmented Multi-modal Learning, ICSE-NIER2022,

geodes

(Deep) Learning Software Artefacts

- Recovering abstract syntax trees from hidden representations of pre-trained language models



(a) Ground-truth (b) GraphCodeBERT

López, J.A., Weyssow, M.; Cuadrado, J. S.; Sahraoui, H.; AST-Probe: Recovering abstract syntax trees from hidden representations of pre-trained language models, ASE2022,

Key Takeawys

Learning problems are often more difficult than they appear at first glance

“Low-hanging fruit” solutions are only initial steps towards solving the problems

Deep learning paradigm offers a world of possibilities for solving complex problems

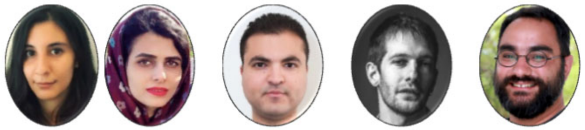
Important to choose the right the problem to solve with DL. “Choose the nail and then look for the hammer.”

geodes

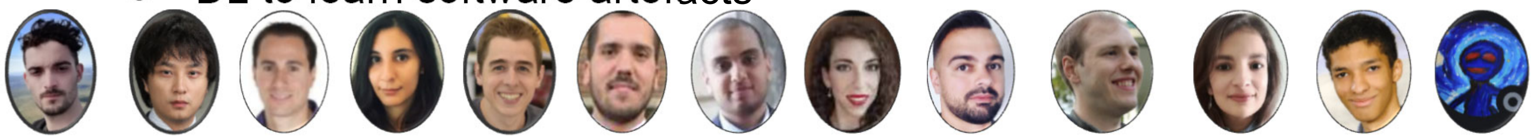
Acknowledgement

- Collaborators
    - MOGP to learn software artefacts

    - MOGP to repair software artefacts



    - DL to learn software artefacts



Université de Montréal