

HUMBOLDT-UNIVERSITÄT ZU BERLIN



Automatically finding Metamorphic Relations in Computational Material Science Parsers

Sebastian Müller

Valentin Gogoll

Anh Duc Vu

Timo Kehr

Lars Grunske · 11 October 2022

Computational Material Science in a Nutshell

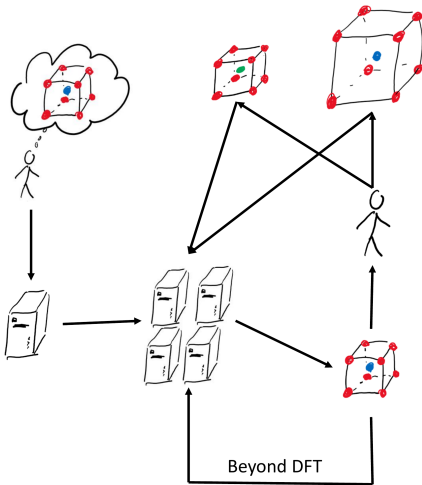
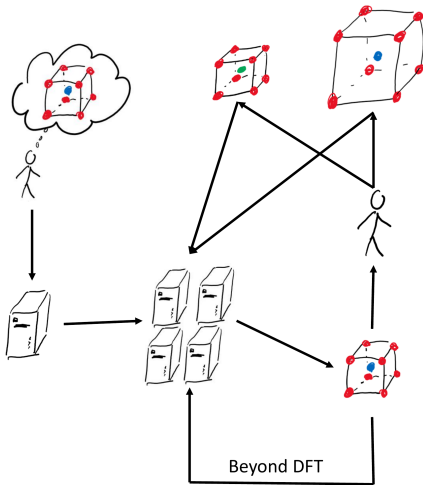


Figure: CMS Workflow. Adapted from Martin Kuban, HU Berlin

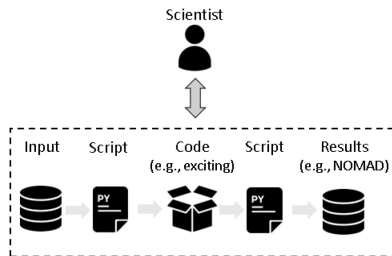
Computational Material Science in a Nutshell



In a nutshell:

- ▶ Human in the loop,
- ▶ Resource intensive,
- ▶ Monolithic, and
- ▶ Numerical approximations.

Figure: CMS Workflow. Adapted from Martin Kuban, HU Berlin



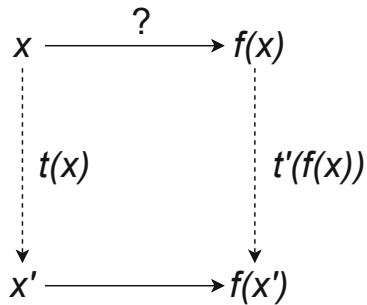
¹<https://nomad-lab.eu/>

Metamorphic Testing²

$$x \xrightarrow{\quad ? \quad} f(x)$$

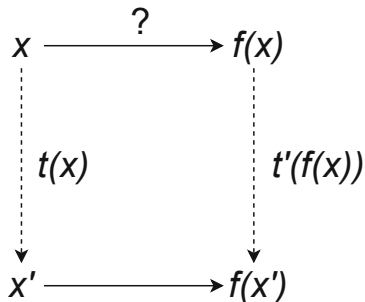
²T. Y. Chen et al.: "Metamorphic testing: A new approach for generating next test cases", 1998

Metamorphic Testing²



²T. Y. Chen et al.: "Metamorphic testing: A new approach for generating next test cases", 1998

Metamorphic Testing²



For any $x \in \mathbb{R}$ the following relation must hold:

$$\sin(x) = \sin(x + 2\pi)$$

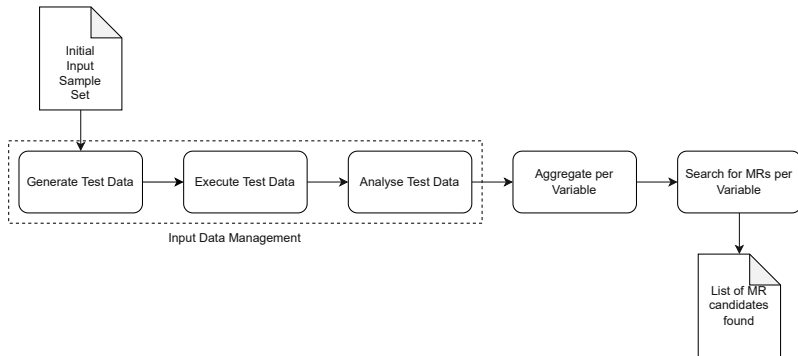
That is: If $t(x) := x + 2\pi$, then $f(x)$ must be equal to $f(t(x))$.

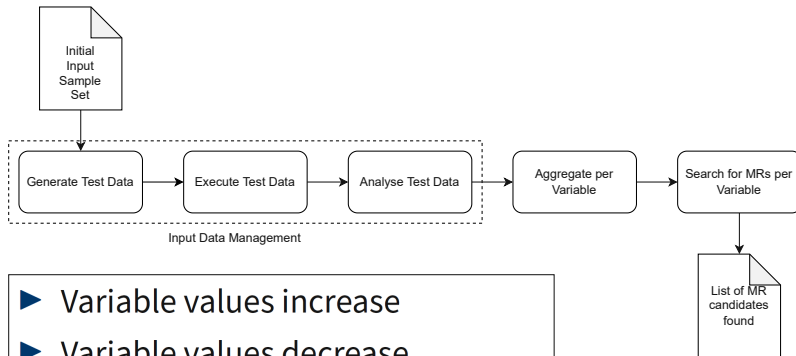
²T. Y. Chen et al.: "Metamorphic testing: A new approach for generating next test cases", 1998

Research Questions

- RQ1:** Is it feasible to develop a tool to automatically find MRs for the Density Functional Theory parser for the `exciting` software package?
- RQ2:** Can such a tool find MRs and what is the reached quantity and quality of the found MRs?

Tool





- ▶ Variable values increase
- ▶ Variable values decrease
- ▶ Variable values approach ∞
- ▶ Variable values approach $-\infty$
- ▶ Variable values are negated
- ▶ Variable values are multiplied by 0

Datasets

Number	Material	Program Version	Entry ID
G1_F1	Gallium(III)Oxide	NITROGEN	znPclmrcW6PbToktDWG5MbMdo8_e
G1_F2	Carbon	OXYGEN	fL0MnyQ45rsaT8nxWRddgRiFvsFD
G1_F3	Boron(III)Nitride	NITROGEN-14	q5eFRvvLzpcNpXkEfp7neO8qr5ev
G1_F4	Boron(III)Nitride	NITROGEN-14	iGiyXyo4z2QzWVFpwJdGAi0QgCY3
G2_F1	Gallium(III)Oxide	NITROGEN	wDjCpU9CklzHMFHNIw1uYpOUGq8U
G2_F2	Carbon	OXYGEN	G2-ES2n20BlomnDt55yPVtb4L2x-
G2_F3	Boron(III)Nitride	NITROGEN-14	oJCWC5WHvKwwUgsPxxfxzn5x0CzK
G2_F4	Boron(III)Nitride	NITROGEN-14	k06u-7wXbzg1tcClFVGNLBtNgnUB

Table: Test Data Set and Validation Set as sampled from the NOMAD repository.

Verifiable MR Candidates

Numerics Identity	Increasing	Decreasing	Negation	Addition	Multiplication
28	36	0	78	9	9
Lists Identity	Length	Vectors Identity	Strings Identity	Spaces	Data type
2	4	1	1	1	1

Table: Number of verifiable MR candidates. For better readability we aggregated all MR candidates of a type per data type of the variables under consideration.

Answering RQ1 – Feasibility

Yes, it is not only possible to create such a tool, it also seems sensible to do so, when considering the amount of MRs found by just this simple prototype tool.

Answering RQ2 – Quantity and Quality

We are able to find 170 MR candidates that are verifiable across both our data sets. Albeit, these found candidates are not very complex MRs. This is due to just using the simple prototype – further search strategies, as well as MR templates may further improve this result.

Your Questions?

