

Multi-language Diversity Benefits Autoformalization

Albert Q. Jiang, Wenda Li, Mateja Jamnik

qj213@cam.ac.uk

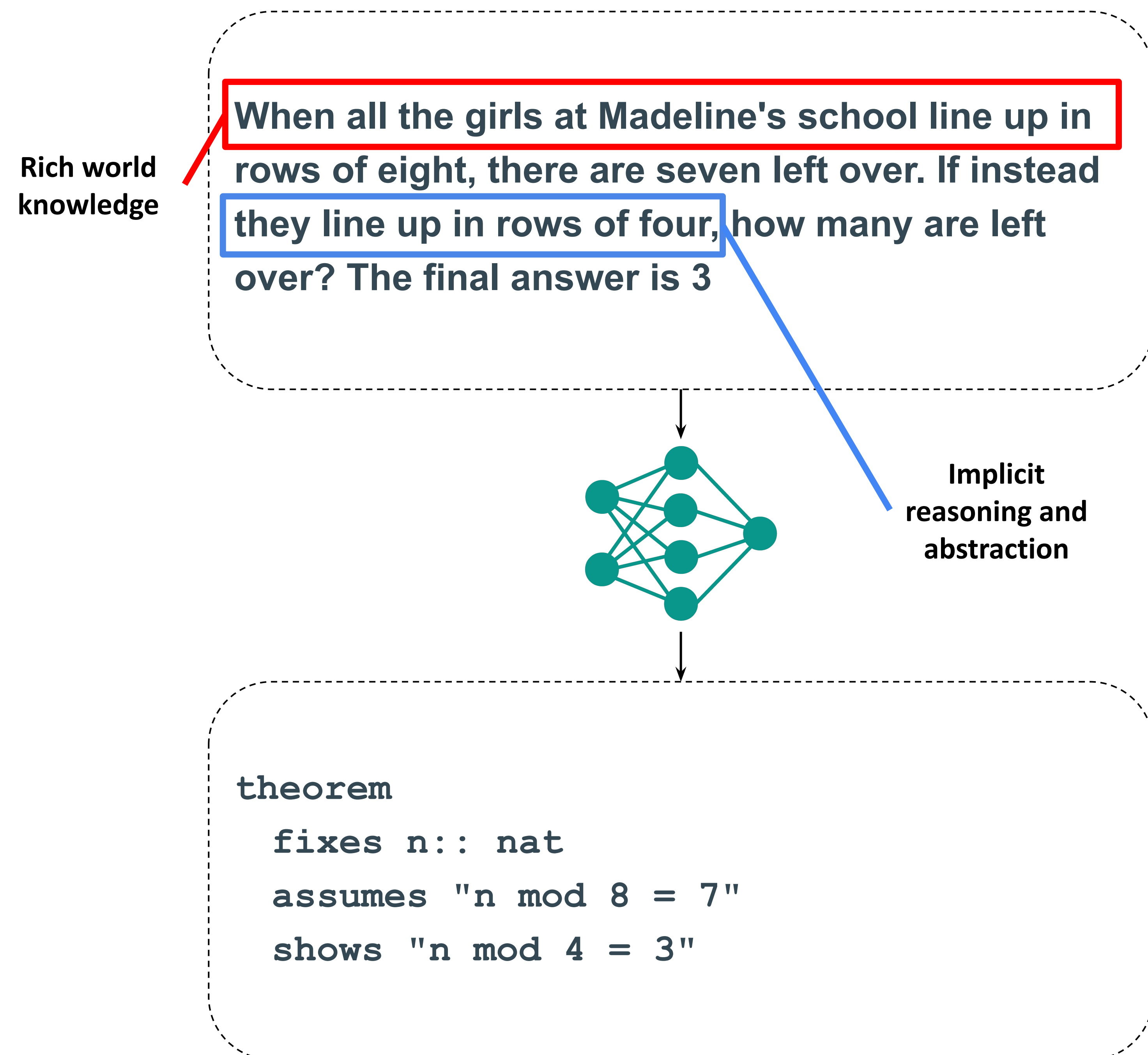


UNIVERSITY OF
CAMBRIDGE

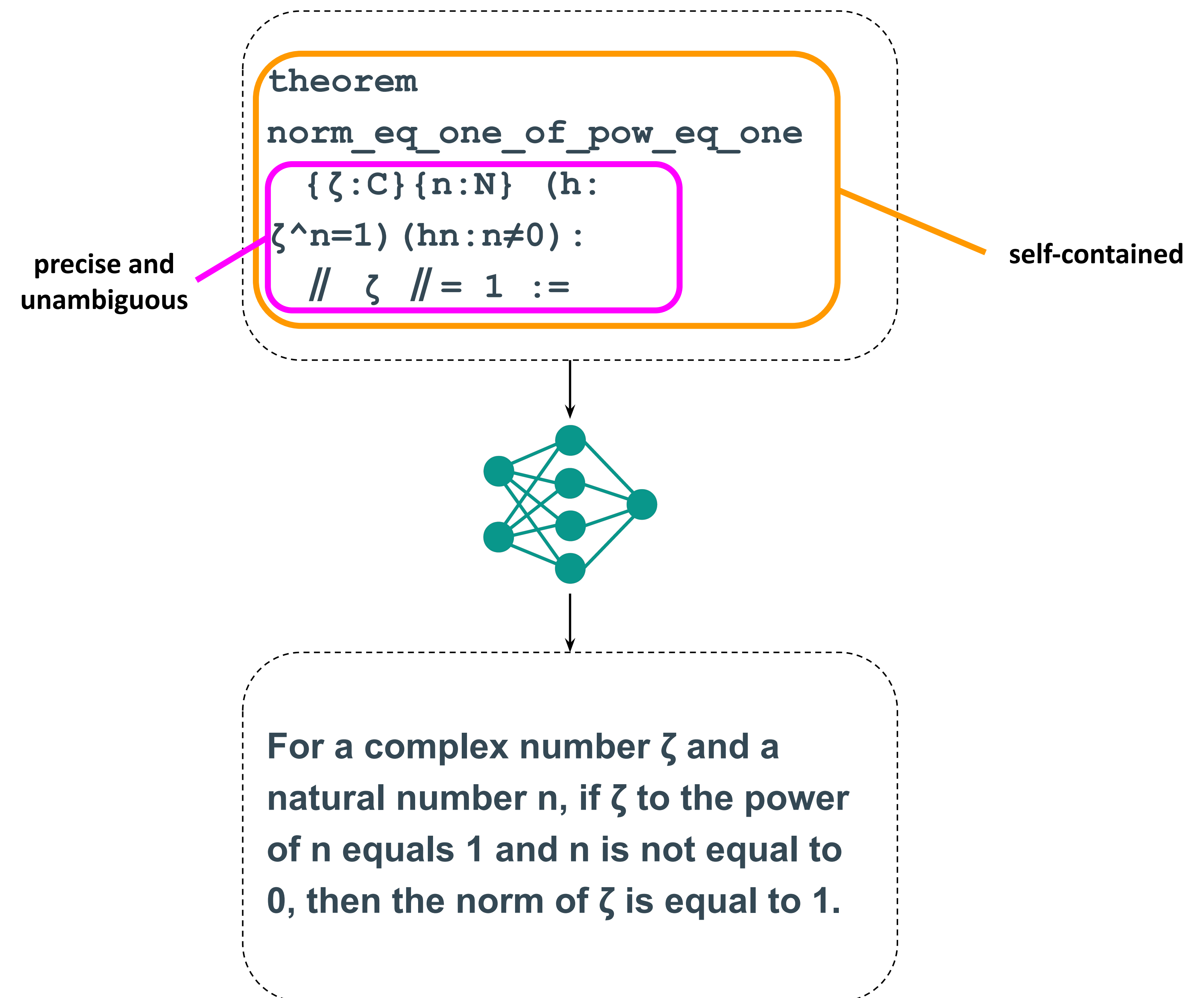


THE UNIVERSITY
of EDINBURGH

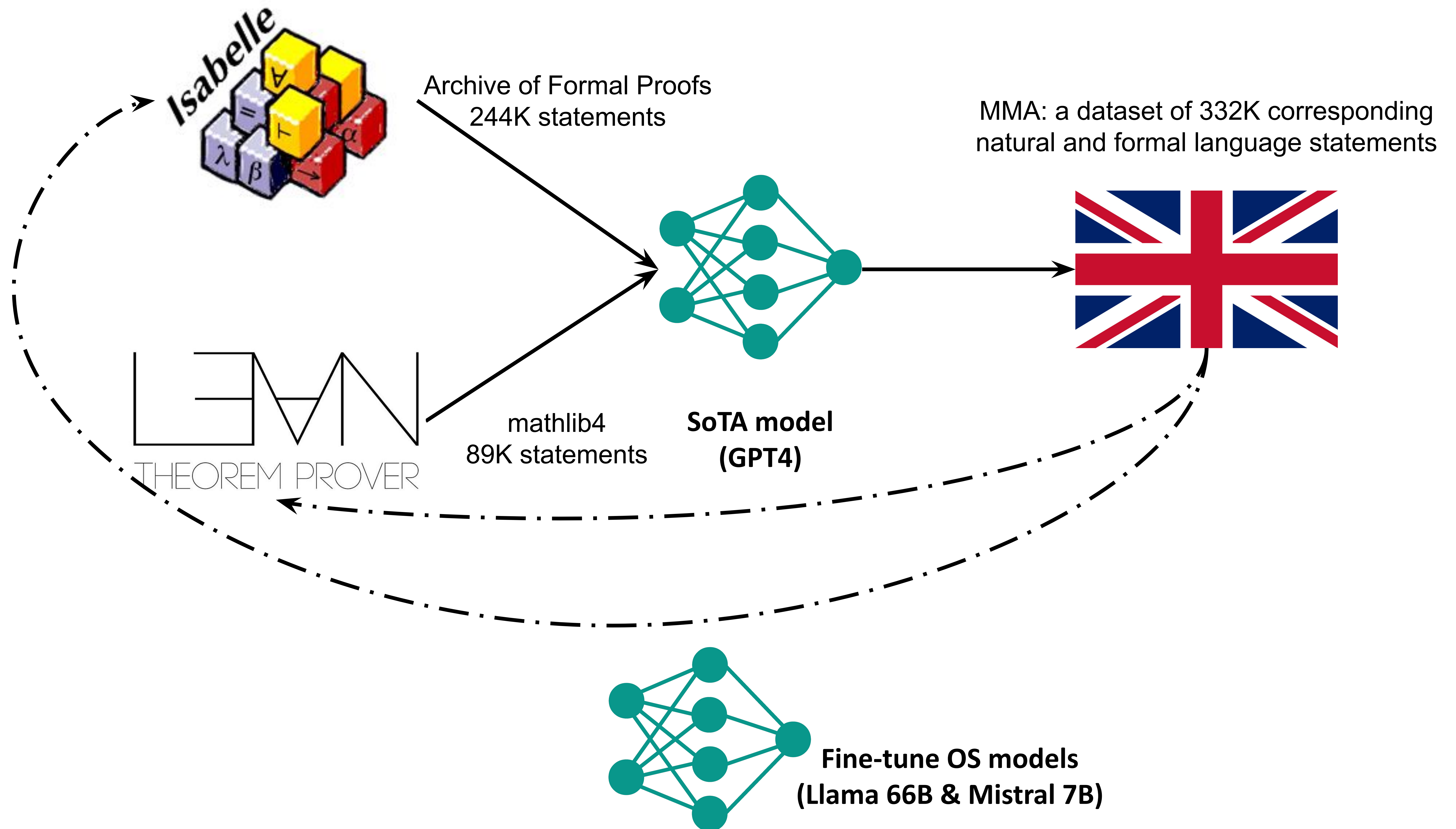
Autoformalizing mathematics is hard



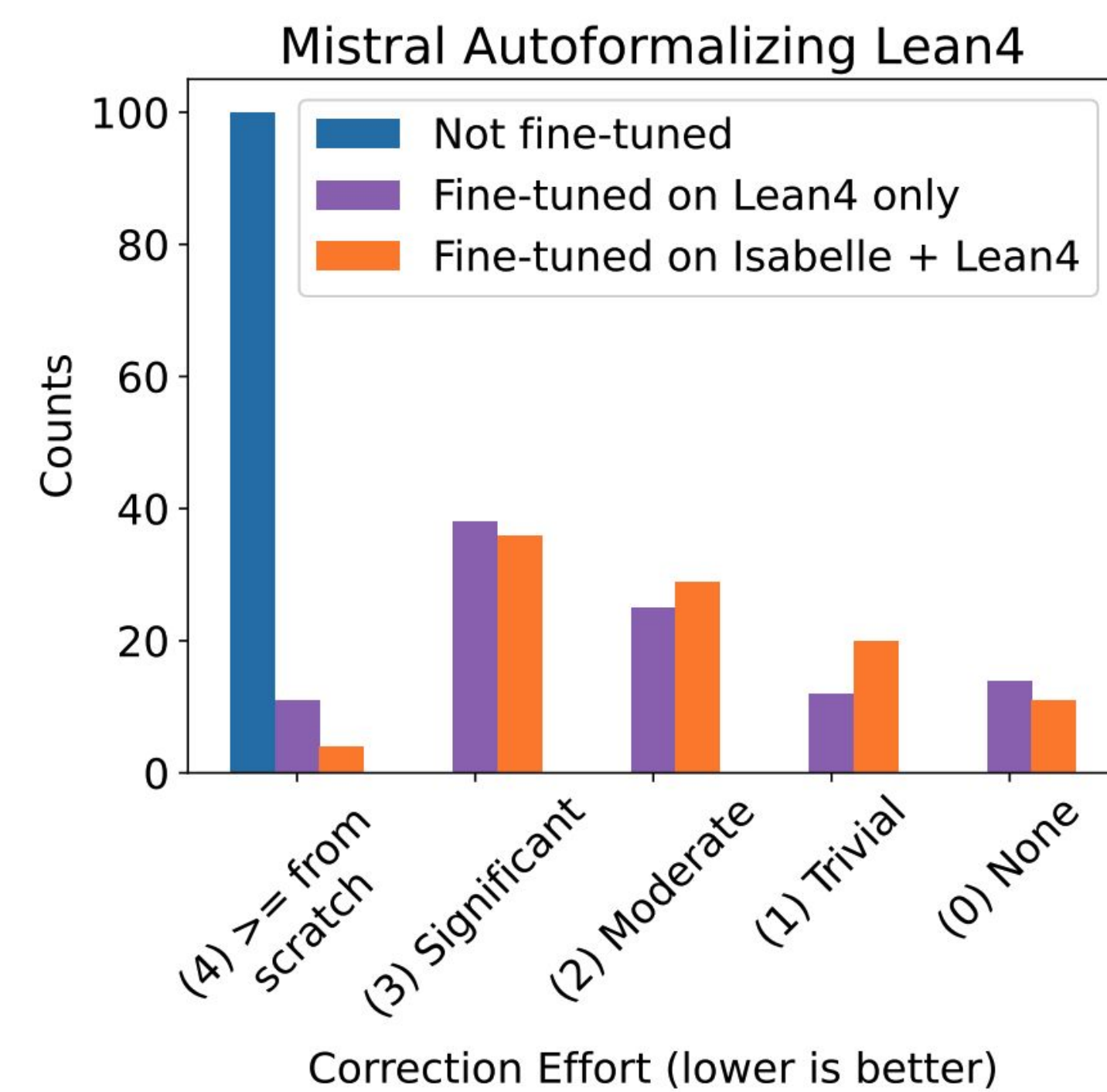
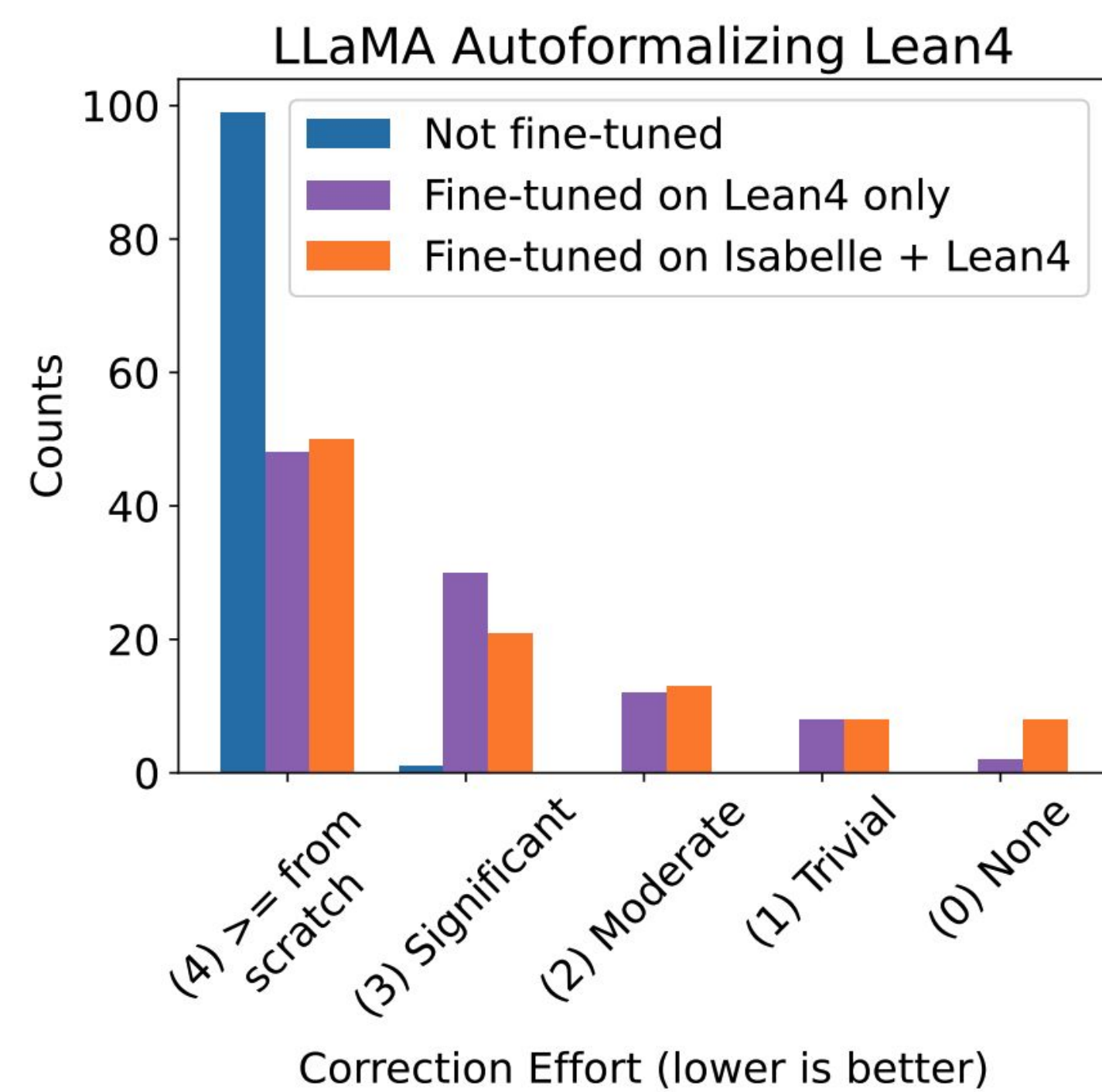
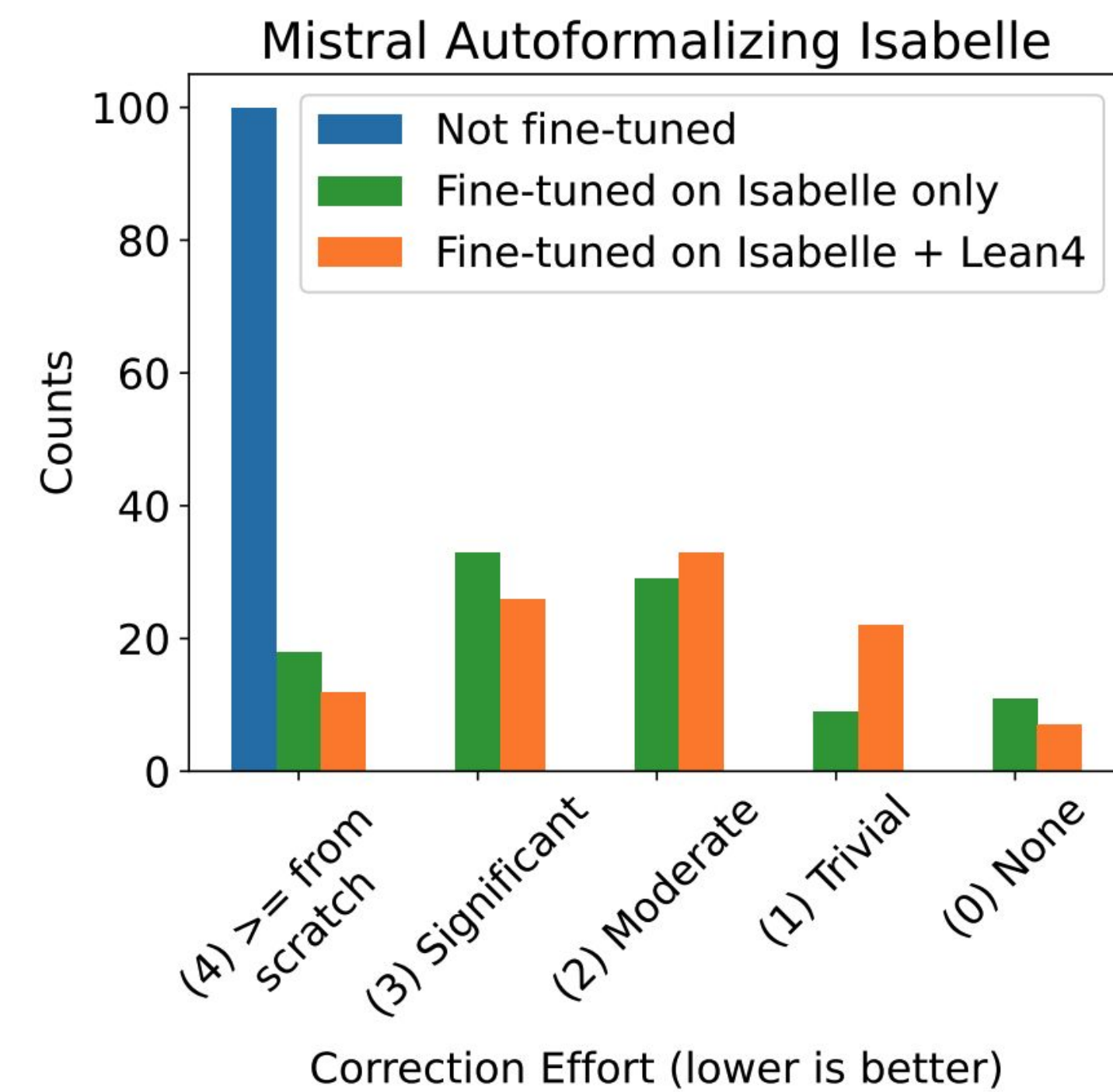
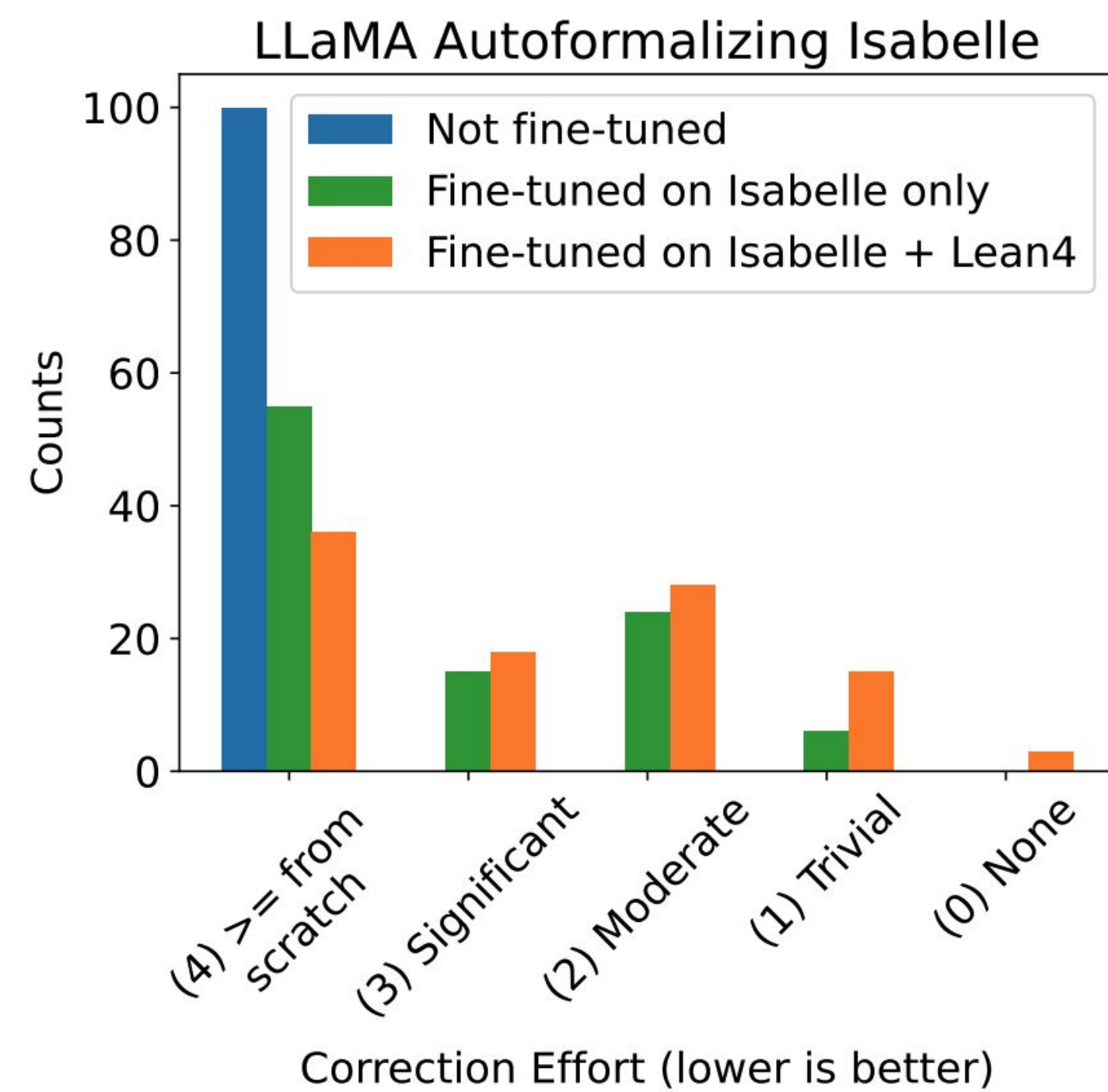
Informalisation is easier



Back-translate to improve autoformalization



Training on multi-language MMA improves autoformalization



Informal statement:

Prove that there is no rational number whose square is 12.

Ground truth in Isabelle:

```
theorem exercise_1_2: assumes "∃ x ∈ ℚ. x^2 = 12" shows False
```

Base LLaMA model:

Theorem square_of_12_is_not_rational:

```
"(1/x)^2 = 12"
```

```
"x = 12"
```

```
"x = 1/12"
```

```
"x = 1/12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of_one)
```

```
"x = 12" by (simp add: square_of_one_divide_by_square_of"
```

Fine-tuned on Isabelle only:

```
lemma
```

```
"⊢ {N = 12} {N = 24}"
```

```
:= 2 * {N = 24}"
```

Fine-tuned on Isabelle + Lean4:

```
lemma "¬ (∃ x. is_rat x ∧ x^2 = 12)"
```

Conclusion

- We construct MMA, a dataset of 332K corresponding informal-formal mathematical statement pairs in both Isabelle and Lean4, by informalising formal repositories.
- We show that fine-tuning models on MMA significantly improves their autoformalization abilities, especially when using both formal languages.
- Iterative back-translation in the spirit of our paper might enable large-scale autoformalization and much stronger neural theorem provers.

Poster: Wednesday 11AM - 2PM

