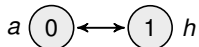


# Human vs. AI interpretations



Statement: *Abelard and Heloise are in love.*



Human agent:

AI (after NLP): *Abelard is in love and Heloise is in love.*

AI (after NLP2FOL):  $\exists x, \text{love}(\text{abelard}, x) \wedge \exists x, \text{love}(\text{heloise}, x)$

## How many models (e.g. MACE4)?

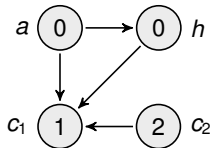
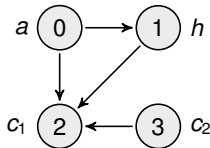
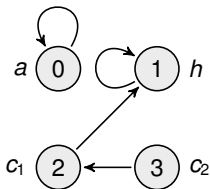
```
assign(max_models, -1).
assign(domain_size, 4).
formulas(assumptions).
    exists x love(abelard,x).
    exists x love(heloise,x).
end_of_list.
```

```
===== STATISTICS =====
For domain size 4.
Current CPU time: 0.00 seconds (total CPU time: 5.66 seconds)
Ground clauses: seen=2, kept=2.
Selections=278522, assignments=557049, propagations=18, current
Rewrite_terms=23, rewrite_bools=20, indexes=18.
Rules_from_neg_clauses=0, cross_offs=0.

===== end of statistics =====
User_CPU=5.66, System_CPU=10.18, Wall_clock=25.
Exiting with 278528 models.
----- process 4061 exit (all_models) -----
Process 4061 exit (all_models) Sun Jul 28 11:51:49 2019
```



## Reducing 278,528 models



- 1 UNA: *abelard*  $\neq$  *heloise* 163,840 models
- 2 Assume love is not narcissistic:  $\forall x, \neg \text{love}(x, x)$ . (5,120)
- 3 Assume someone can love only one person at a time:  $\text{love}(x, y) \wedge \text{love}(x, z) \rightarrow y = z$ . (80)
- 4 Remove isomorphic interpretations `isofilter` (74)
- 5 2 Skolem consts: assume no interest in love relations between them (17)

### Remarks

- 1 order of reductions is computationally relevant (186,976 models,  $\sim 2h$ )
- 2 which domain knowledge to add is subject to interpretation