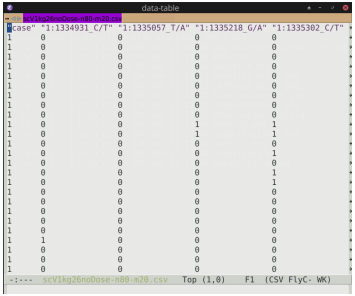
# Combining learning and reasoning for Bio-AI

Nil Geisweiller

SingularityNET & OpenCog Foundations

# Why?

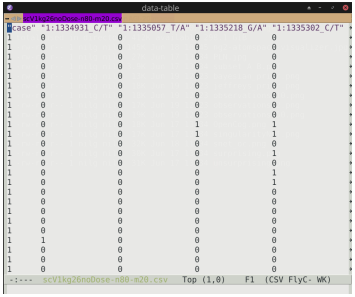
Why combining machine learning and reasoning?



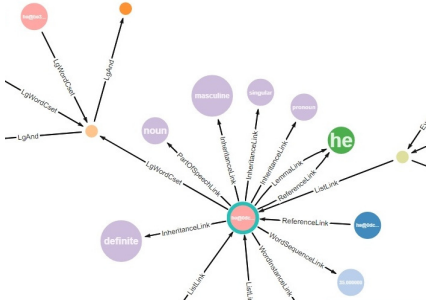

⊗

# Why?

Why combining machine learning and reasoning?



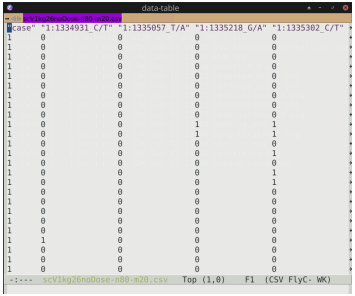

⊗



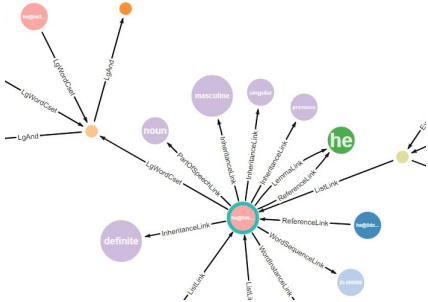

Massive amount of indirect evidence

# Why?

Why combining machine learning and reasoning?




⊗




Massive amount of indirect evidence

Ultimate answer to overfitting

# Why?

What about simulation?

Impractical without abstractions

Reasoning
⇓
Abstractions

# Why?

Help learning (and reasoning)

- Reasoning for meta-learning
  - Filter relevant features
  - Guide optimization
- Learning for meta-reasoning
  - Discover inference control patterns
  - Create contextual Hebbian links

# Learning & reasoning over the Bio-AtomSpace

- Learning:
  - MOSES (program evolution)
    ⇒ Predictive models
  - Pattern Miner (frequent pattern mining)
    ⇒ Discover abstractions
- Reasoning:
  - Pattern Miner
  - PLN (Probabilistic Logic Networks)
    ⇒ Use existing and discovered background knowledge

# Example

1. From table of Gene expression of individuals:

```
(Implication (stv 0.6 0.02)
  (And
    (Expression Gene-1)
    (Expression Gene-2))
  Longevity)
```

2. From GO (Gene Ontology)

```
(Member (stv 1 1) Gene-2 GO-1)
```

3. From mining GO and SMP (Small Molecule Pathway)

```
(Similarity (stv 0.7 0.2) GO-1 SMP-2)
```

4. From SMP

```
(Member (stv 1 1) Gene-3 SMP-2)
```

5. Inferred conclusion

```
(Implication (stv 0.65 0.01)
  (And
    (Expression Gene-1)
    (Expression Gene-2)
    (Expression Gene-3))
  Longevity)
```

# Status

- Discovered simple patterns
  - Pattern size: 2 conjuncts
  - GO + SMP dataset: 1M atoms
  - Time: couple hours
- Inferred short trails
  - Trail size: about 8 steps
  - GO dataset: 650K atoms
  - Time: couple hours
- Focused on longevity

# Difficulties

- Porting data into the atomspace
- Finding good queries
- Very resource hungry (millions of atoms)
  - CPU: 1 single step can take 20+ minutes
  - RAM: 1 single step can take 64GB+
  - Need ECAN!
- Advanced forms of reasoning
  - Stress-test on PLN

# To do

- Experiment with other domains, COVID-19, Cancer
- Complete Multi-threaded Rule Engine
- Integrate ECAN
- Integrate spatio-temporal reasoning
- Experiment with inference control meta-learning