# Modularity in Ontologies: Atomic decomposition

*Thomas Schneider*[1] Dirk Walther[2]

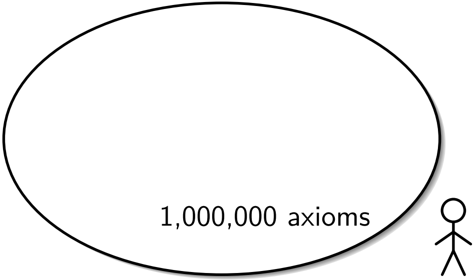
[1]Department of Computer Science, University of Bremen, Germany

[2]Center for Advancing Electronics Dresden, TU Dresden, Germany

ESSLLI, 15 August 2013

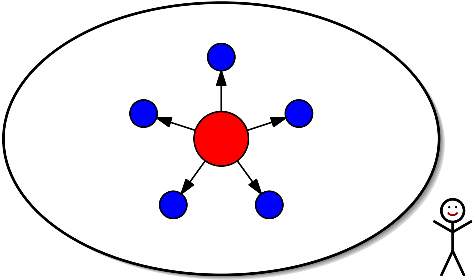
# What is my ontology about?

We can't inspect all its axioms.



# What is my ontology about?

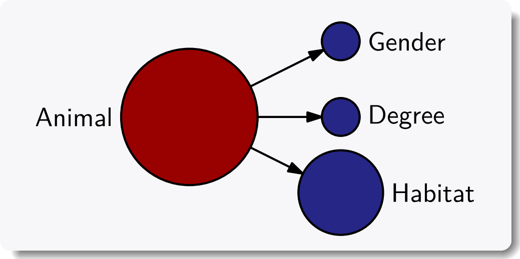
We can inspect its modular structure, obtained *a posteriori*.



# We bet Robert Stevens

- Ontology about periodic table of the chemical elements
- Logical structure ≈ intended modelling?
  - What are its main parts?
  - How do they logically interact with each other?
- Challenge: *automatic* partition into meaningful modules

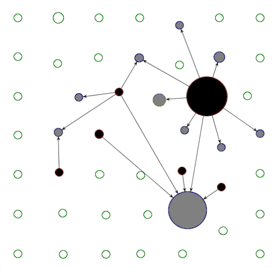
Partition of *Koala* via E-connections in Swoop




- importing part
- imported but non-importing part
- isolated part

→ "imports vocabulary from"

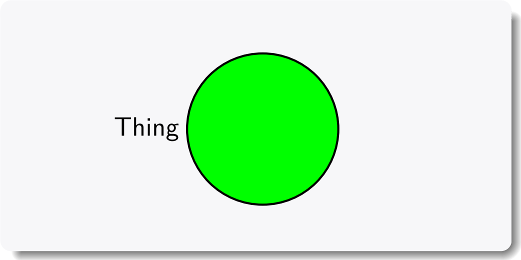
# Partition for ontology *SWEET*



- importing part
- imported but non-importing part
- isolated part

→ "imports vocabulary from"

# Partition for ontology *Periodic*




- importing part
- imported but non-importing part
- isolated part

→ "imports vocabulary from"

# Modular structure via LBMs – goals

- Draw conclusions on characteristics of an ontology:
  - Which topics does $\mathcal{O}$ cover?
  - How do they interact with each other?
  - How strongly are certain terms connected in $\mathcal{O}$?
  - Does $\mathcal{O}$ have superfluous parts?
  - Agreement between logical and intended intuitive modelling?
- Guide users in choosing the right signatures/modules

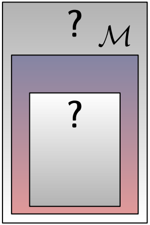
# Modular *structure*

Remember: Ontology $\mathcal{O}$ is a set of axioms; module $\mathcal{M} \subseteq \mathcal{O}$

Modules are great: if you know your (seed) signature ...
and for "module local" tasks such as reuse

*Single* module extraction does *not* help if you

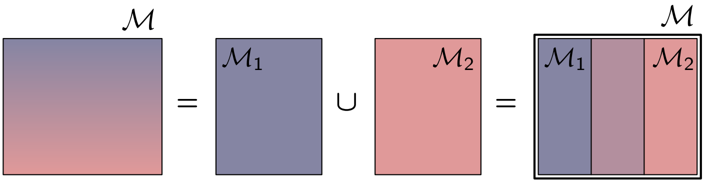
- do *not* know the *right* seed signature
- want to understand *other* modules
- want to understand *axiom dependency structure*




To analyse the modular structure of the ontology:

- significant modules
- significant relations between modules
- ... which reveals logical dependency between axioms

# Are all modules significant?




To understand $\mathcal{M}$, one must understand

- the dependency structure of $\mathcal{M}_1$
- the dependency structure of $\mathcal{M}_2$
- nothing else: $\mathcal{M}_1$ and $\mathcal{M}_2$ have no further dependencies

$\rightsquigarrow$ $\mathcal{M}$ is not significant: it is a fake module

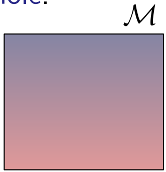
- $\mathcal{M}_1$ and $\mathcal{M}_2$ may be "significant"
- Knowing that $\mathcal{M}$ is "only" a union is important

# Are all modules significant?

Consider a module $\mathcal{M}$ that is not fake.

To understand $\mathcal{M}$, one has to understand $\mathcal{M}$ as a whole.

- all axioms in $\mathcal{M}$ logically interact
- in different ways – but interact




"Not fake" implies significant: genuine

# How many (fake, genuine) modules are there?

The number of *all* modules can and does grow exponentially in $|\mathcal{O}|$
[Del Vescovo, Parsia, Sattler, Schneider 2010]

Given a set of genuine modules,

- Unions lead to fake modules

$\rightsquigarrow$ The space of fake modules is exponential

- But not *every* union of genuine modules is a module

**Question 1**

Is module growth primarily due to trivial combinations?
I.e., are most modules fake?

# Yes!

**Theorem 1**

Each genuine module is the smallest module that contains $\alpha$, for some axiom $\alpha \in \mathcal{O}$.

[Del Vescovo, Parsia, Sattler, Schneider 2011]

$\rightsquigarrow$ The family of genuine modules is linear in $|\mathcal{O}|$.

Most modules are fake!

Proof exploits properties of modules

- uniqueness, monotonicity, self-containedness, ...
- which are satisfied by all locality-based modules

# Wrap-up so far

- Ontology $\mathcal{O}$ can have exponentially many modules
- **Genuine modules** are of particular interest: $\mathcal{M} \neq \mathcal{M}_1 \cup \mathcal{M}_2$
- Central theorem:

$$\mathcal{M} \text{ is genuine} \quad \text{iff} \quad \mathcal{M} = \ldots\text{-mod}(\text{sig}(\alpha), \mathcal{O})$$

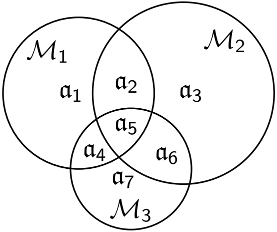
(works for different module notions, e.g., LBMs, MEX)

$\rightsquigarrow$ There are **linearly many** genuine modules!

# Atoms

**Atom**

$=$ atomic part of a module in the context of all other modules



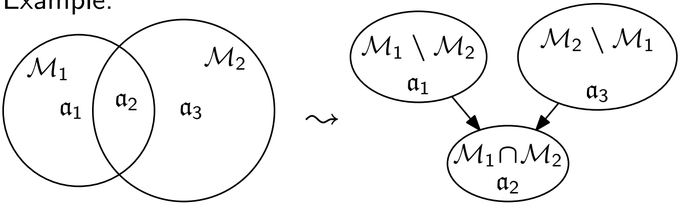

- Example:
- Easy to see: $\forall \mathfrak{a} \forall \mathcal{M} : \mathfrak{a} \subseteq \mathcal{M}$ or $\mathfrak{a} \cap \mathcal{M} = \emptyset$

**Important result:**

All atoms can be computed using only the genuine modules
$\rightsquigarrow$ linearly many!

# Dependency between atoms

- Example:




- Axioms in $\mathfrak{a}_1$ depend on those in $\mathfrak{a}_2$ – but not vice versa
- In general: $\mathfrak{a}$ depends on $\mathfrak{b}$ if
  iff $\mathfrak{a}$ needs $\mathfrak{b}$ to form a module:
  iff $\forall \mathcal{M} : \mathfrak{a} \subseteq \mathcal{M} \Rightarrow \mathfrak{b} \subseteq \mathcal{M}$

**Important result:**
All dependencies can be computed using only the genuine modules
$\rightsquigarrow$ in polynomial time!

# In reality?

**Next:**

- Example
- Decomposition of some existing ontologies
  (Thanks to Chiara Del Vescovo and Nicolas Matentzoglu for pictures.)

# Example

$$\exists \text{hasChild}.\top \sqsubseteq \text{Animal} \quad (\alpha_1)$$
$$\top \sqsubseteq \forall \text{hasChild}.\text{Animal} \quad (\alpha_2)$$
$$\text{Mother} \equiv \exists \text{hasChild}.\top \sqcap \exists \text{hasGender}.\{\text{female}\} \quad (\alpha_3)$$
$$\exists \text{hasGender}.\top \sqsubseteq \text{Animal} \quad (\alpha_4)$$
$$\top \sqsubseteq \forall \text{hasGender}.\text{Gender} \quad (\alpha_5)$$

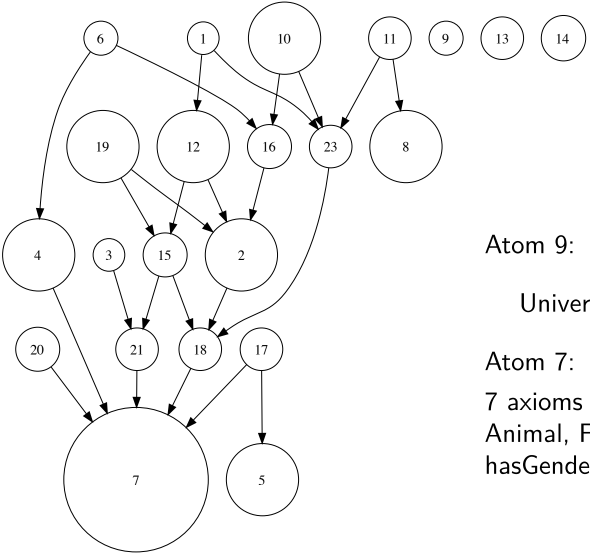
$\rightsquigarrow$ **AD via $\bot$-modules:**

$$\bot\text{-mod}(\text{sig}(\alpha_1), \mathcal{O}) = \{\alpha_1, \alpha_2\}$$
$$\bot\text{-mod}(\text{sig}(\alpha_2), \mathcal{O}) = \{\alpha_1, \alpha_2\}$$
$$\bot\text{-mod}(\text{sig}(\alpha_3), \mathcal{O}) = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$$
$$\bot\text{-mod}(\text{sig}(\alpha_4), \mathcal{O}) = \{\alpha_4, \alpha_5\}$$
$$\bot\text{-mod}(\text{sig}(\alpha_5), \mathcal{O}) = \{\alpha_4, \alpha_5\}$$

$\rightsquigarrow$ **Atoms and dependencies:**

$\{\alpha_1, \alpha_2\}$, $\{\alpha_3\}$, $\{\alpha_4, \alpha_5\}$ — $\{\alpha_3\}$ depends on the other two

# Decomposition of Koala (⊤⊥*)



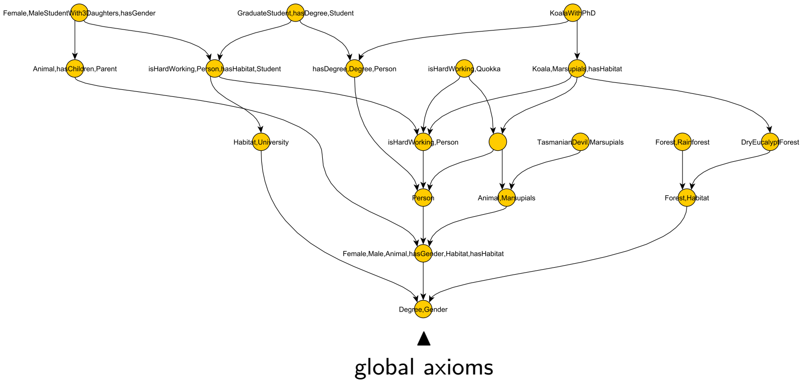

Atom 9:

University $\sqsubseteq$ Habitat

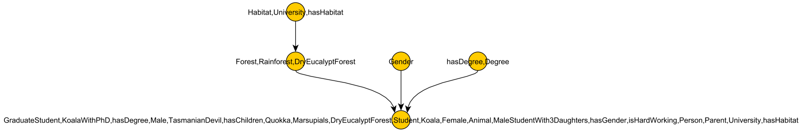
Atom 7:

7 axioms about Animal, Female, Male, hasGender, hasHabitat

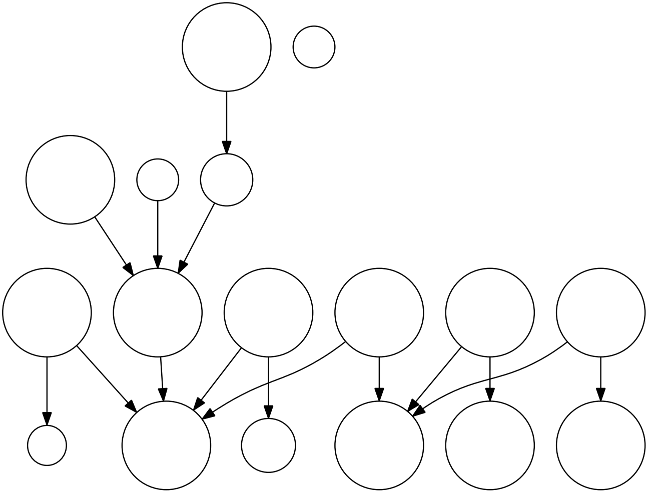
# Decomposition of Koala (⊥)



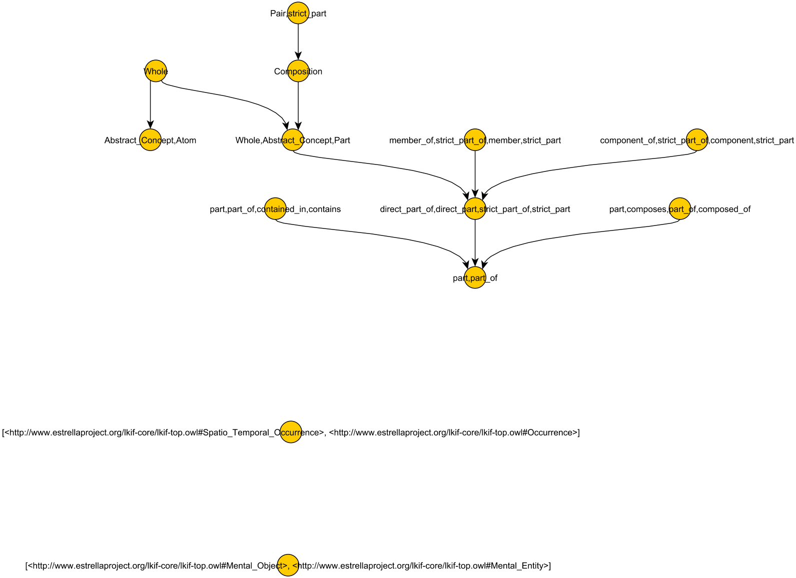
# Decomposition of Koala (⊤)



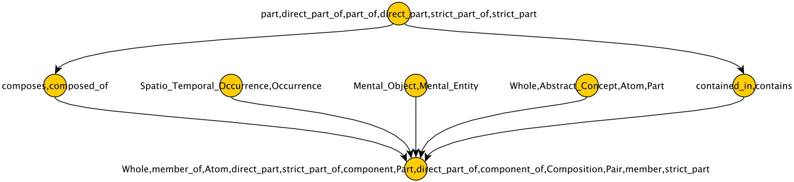
# Decomposition of Mereology (⊤⊥*)



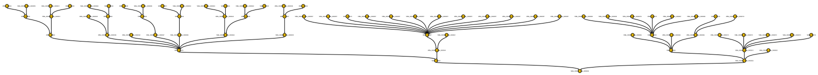
# Decomposition of Mereology (⊥)



# Decomposition of Mereology ($\top$)



# Decomposition of c-elegans (⊥)



# Summary

The atomic decomposition (AD) . . .

- is a linear representation of the potentially exponential set of all modules
- can be computed using a linear number of module extractions
- exposes 2 types of logical dependencies between axioms
- is implemented in the OWL API Tools (but not yet released) `http://sourceforge.net/projects/owlapitools`

**Thank you.**