

The Modular Structure of an Ontology: Atomic Decomposition

Chiara Del Vescovo, Bijan Parsia,
Uli Sattler, & Thomas Schneider

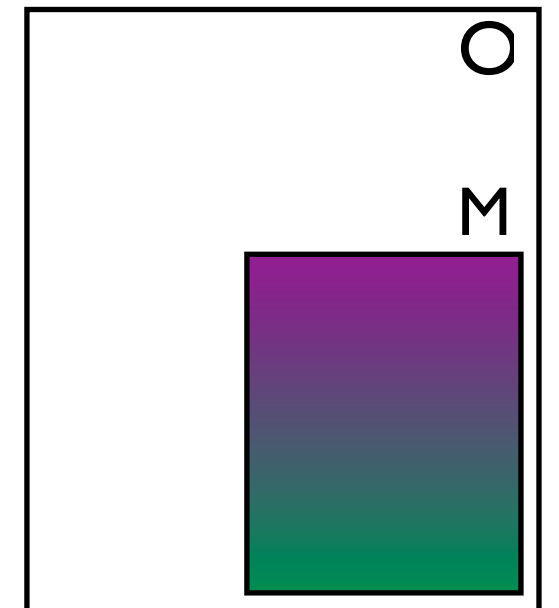
The University of Manchester, UK
Universitaet Bremen, Germany

Ontologies & Modules

- ▶ An **ontology** is a finite set of axioms in a (description) logic
- ▶ A **module** $M(\Sigma, O) \subseteq O$ encapsulates knowledge w.r.t a signature Σ :

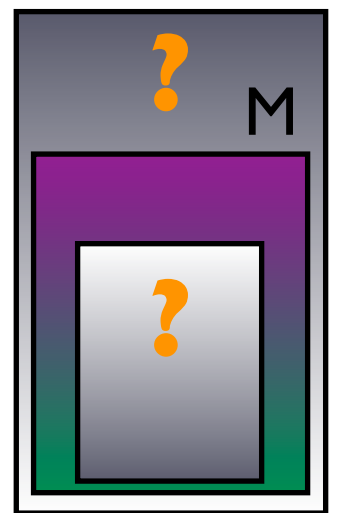
for all α with $\text{sig}(\alpha) \subseteq \Sigma$: $O \models \alpha$ iff $M(\Sigma, O) \models \alpha$

$M(\{\text{part}\}, \text{Mereology.owl}) = \{\text{Trans: part, part InverseOf: PartOf, Trans: partOf}\}$

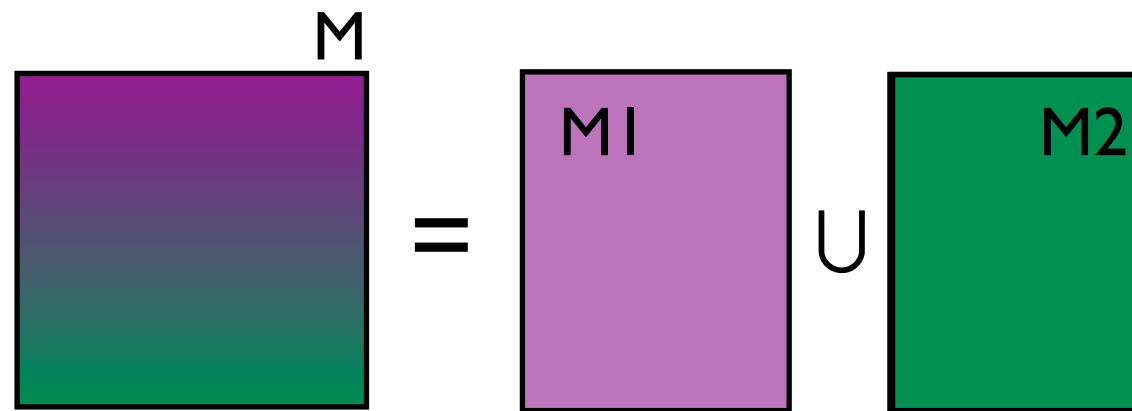


Modular *Structure*

- ▶ 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 dependence between axioms

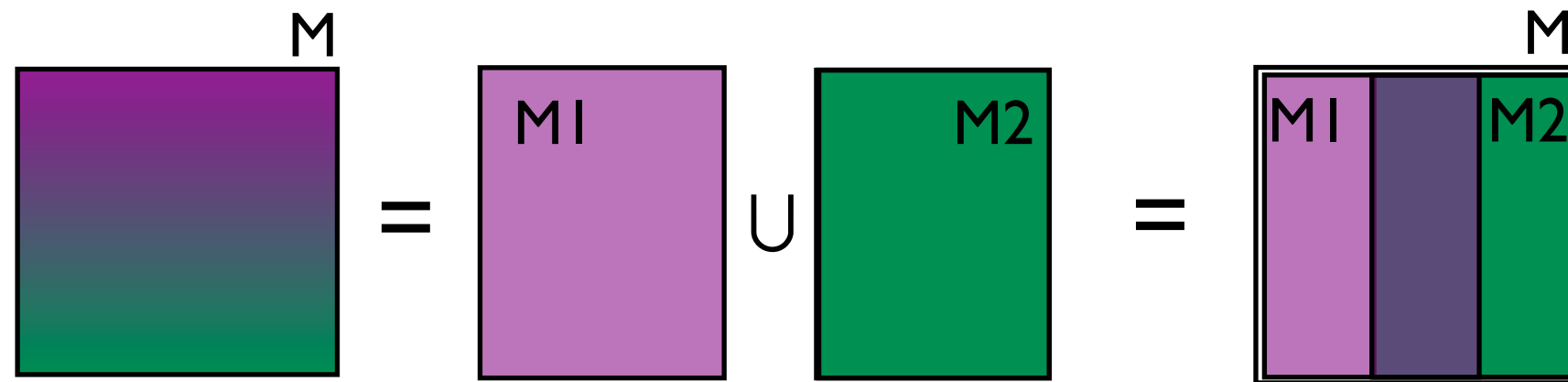


Are all modules significant?



- ▶ To understand M, one must
 - ▶ understand the dependancy structure of M1
 - ▶ understand the dependancy structure of M2
 - ▶ **nothing** else: M1 and M2 have no further dependancies
- ➔ M is **not** significant: it is a **fake** module
 - ▶ Thus, M1 and M2 may be “significant”
 - ▶ knowing that M is “only” a union is important

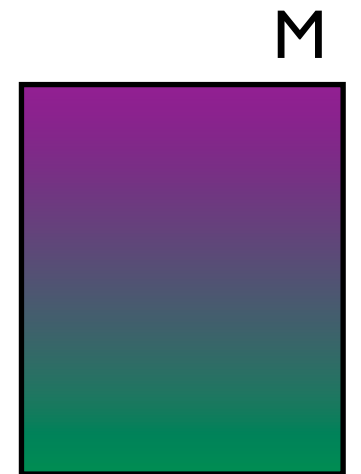
Are all modules significant?



- ▶ To understand M , one must
 - ▶ understand the dependancy structure of $M1$
 - ▶ understand the dependancy structure of $M2$
 - ▶ **nothing** else: $M1$ and $M2$ have no further dependancies
- ➔ M is **not** significant: it is a **fake** module
 - ▶ Thus, $M1$ and $M2$ may be “significant”
 - ▶ knowing that M is “only” a union is important

Are all modules significant?

- ▶ Consider a module M that is **not fake**
- ▶ To understand M , one has to understand M **as a whole**
 - ▶ all axioms in M logically interact
 - ▶ in different ways - but interact
- ▶ Not fake implies significant: **genuine**



Ratio of Fake to Genuine

- ▶ Given a set of genuine modules
 - ▶ unions lead to fake modules,
 - the space of fake modules is large (exponential)
 - ▶ but not every union of genuine modules is a module
- ▶ The cardinality of the set of all modules can and does grow exponentially in the size of O
 - ▶ See Parsia&Schneider, KR2010 & WOMO 2010
- ▶ *Is module growth primarily due to trivial combinations?*
 - ▶ are most modules **fake**?

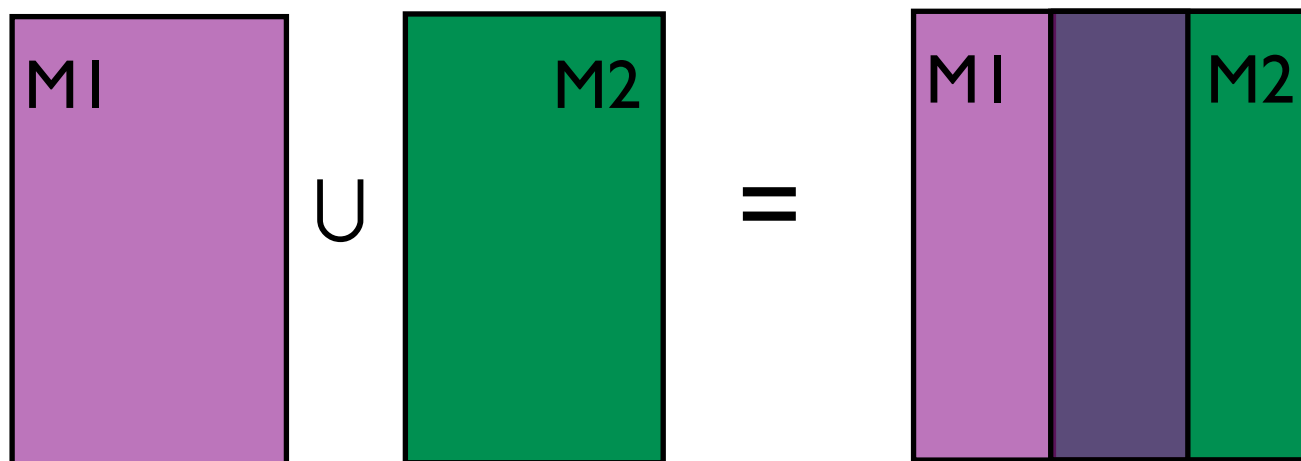
Yes!

Theorem 1: Each genuine module is the smallest module for some axiom $\alpha \in \mathcal{O}$.

- ▶ 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

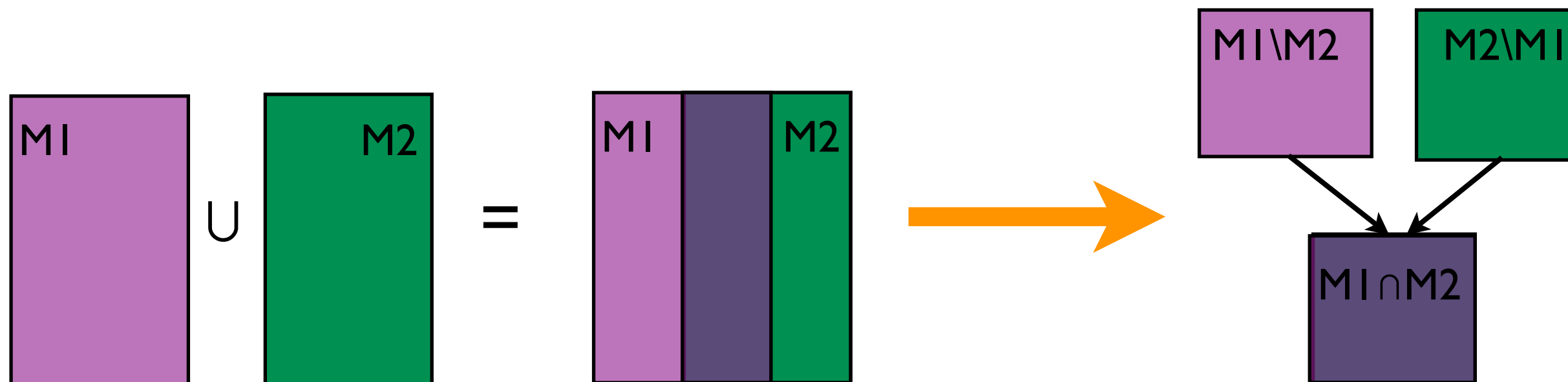
Relations between Modules

- ▶ Genuine modules may overlap
- ▶ This exposes significant logical dependence between axioms:
 - ▶ axioms in $M1 \setminus M2$ depend on axioms in $M1 \cap M2$

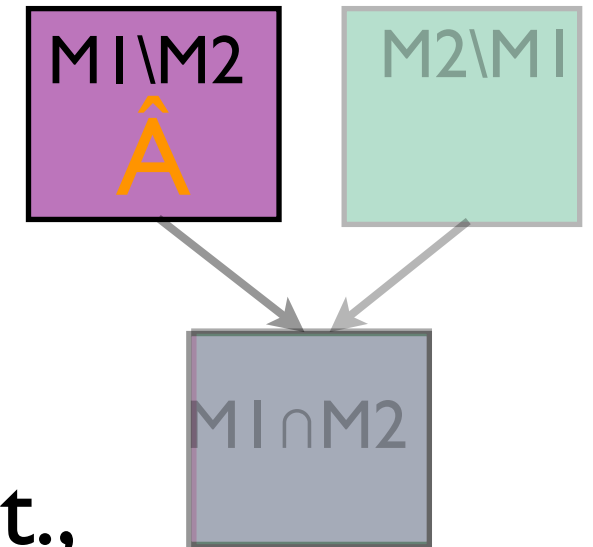


Relations between Modules

- ▶ Genuine modules may overlap
- ▶ This exposes significant logical dependence between axioms:
 - ▶ axioms in $M1 \setminus M2$ depend on axioms in $M1 \cap M2$

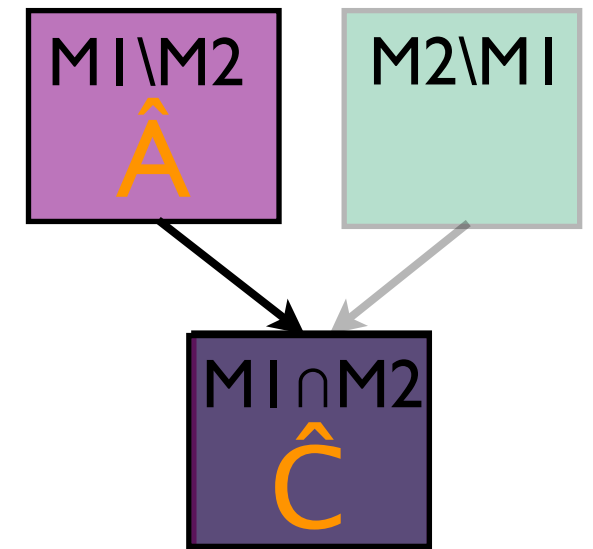


Atoms



- ▶ $\hat{A} \subseteq O$ is an **atom** if it is a maximal set s.t., for each module M , either $\hat{A} \subseteq M$ or $\hat{A} \cap M = \emptyset$.
 - ➔ The smallest module for an axiom α contains the whole atom to which α belongs!
 - ➔ Axioms in an atom are logically interdependent
 - ➔ Any two atoms are disjoint
 - ➔ The family of atoms is a partition of the ontology
 - ▶ Only linearly many atoms
 - ➔ Each GM is a disjoint union of atoms

Proposition: There is a 1-1 correspondence between genuine modules and atoms.



Atomic Decomposition

- ▶ Dependence between atoms:
 - ▶ $\hat{A} \succcurlyeq \hat{C}$ if, for each M : $\hat{A} \subseteq M$ implies $\hat{C} \subseteq M$
 - ▶ The smallest module containing $\hat{A}$ also contains $\hat{C}$!
 - ▶ Axioms in $\hat{A}$ logically depend on axioms in $\hat{C}$

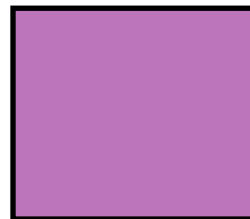
Theorem 2: The relation $\succcurlyeq$ is reflexive, antisymmetric, and transitive.

- ▶ a Hasse diagram exposes 2 logical dependencies amongst axioms in atoms & between atoms

Mereology Ontology

42 axioms

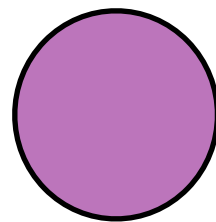
1952 modules



Mereology Ontology

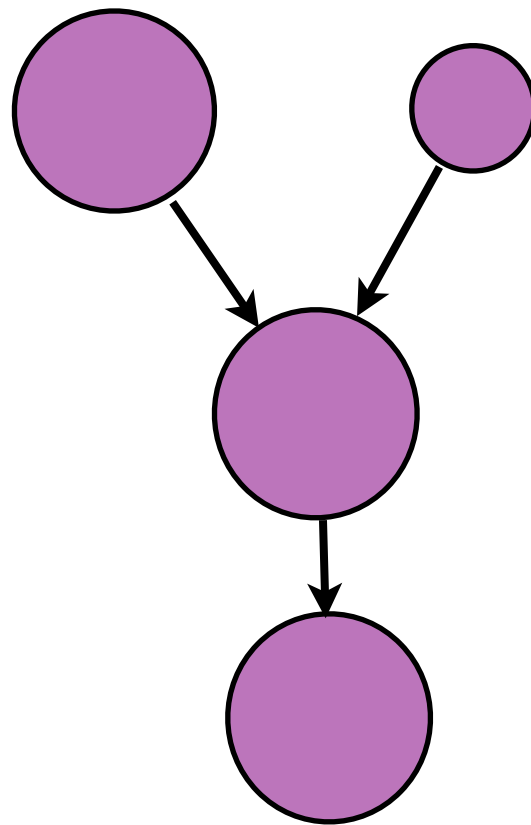
42 axioms

1952 modules



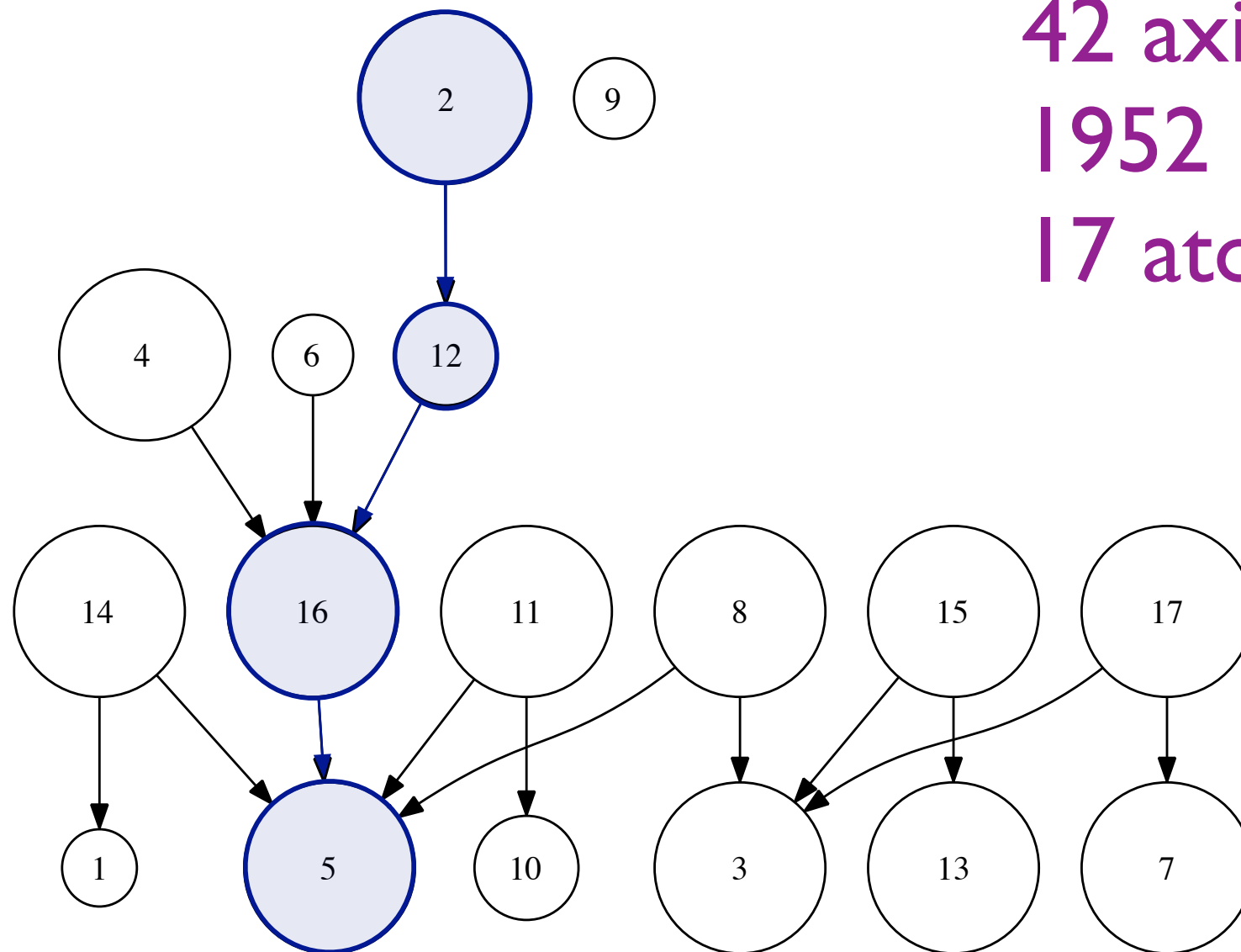
Mereology Ontology

42 axioms
1952 modules



Mereology Ontology

42 axioms
1952 modules
17 atoms/GMs



- ▶ Can we compute all genuine modules?
 - ▶ and all atoms
 - ▶ with their dependencies?
 - ▶ ...without computing all modules?!

Yes!

- ▶ Remember:

Theorem 1: Each genuine module is the smallest module for some axiom $\alpha \in O$.

- ▶ extract $M(\text{sig}(\alpha), O)$
 - ▶ $\leq$ linearly many module extractions
- ▶ AD induced by the comparison of GMs
 - ▶ quadratic procedure

In Reality?

- ▶ We have decomposed 181 OWL ontologies in Bioportal
- ▶ Decomposability: average
 - ▶ nr. axioms/atom: 1.73
 - ▶ max nr. axioms/atom: 86
 - ▶ nr. axioms/GM: 66
 - ▶ max nr. axioms/GM: 143

Future Work

- ▶ More on dependency of axioms
 - ▶ between atoms and **sets** of atoms
- ▶ Labels for atoms
 - ▶ different labels for different tasks
- ▶ Applications
 - ▶ All Module Count: see WoMO 2011
 - ▶ Fast Module Extraction
 - ▶ Topicality for Ontology Comprehension: see ICCS 2011
 - ▶ ...