Modularity of Ontologies in an Arbitrary Institution

Yazmin Angelica Ibañez[1] Till Mossakowski[2] Don Sannella[3]
Andrzej Tarlecki[4]

[1]University of Bremen

[2]Faculty of Computer Science, Otto-von-Guericke University of Magdeburg

[3]Laboratory for Foundations of Computer Science, University of Edinburgh

[4]Institute of Informatics, University of Warsaw

February 18, 2015

Motivation

Size of ontologies SNOWMED CT[1] or GALEN[2] is huge
$\Rightarrow$ reuse only those parts that cover all the knowledge about that subset of relevant terms.
This leads to the module extraction problem:

> *given a subset Σ of the signature of an ontology $\mathcal{O}$, find a (minimal) subset of that ontology that is "relevant" for the terms in Σ.*

[1] `http://ihtsdo.org/snomed-ct/`
[2] http://www.opengalen.org/

Example

$$
\begin{array}{lll|lll}
\text{Male} & \equiv & \text{Human} \sqcap \neg\text{Female}, & \text{Father} & \sqsubseteq & \text{Human}, \\
\text{Human} & \sqsubseteq & \forall\text{has\_child}.\text{Human}, & \text{Father} & \equiv & \text{Male} \sqcap \exists\text{has\_child}.\top
\end{array}
$$

Terms of interest: $\Sigma = \{\text{Male}, \text{Human}, \text{Female}, \text{has\_child}\}$.
Let $\mathcal{M}$ = grey shaded axioms. Then

$\boxed{\mathcal{M} \textit{ is a } \Sigma\textit{-module of } \mathcal{O}\textit{, i.e. } \mathcal{O} \textit{ has the same } \Sigma\textit{-consequences as } \mathcal{M}.}$

E.g., $\text{Male} \sqcap \exists\text{has\_child}.\top \sqsubseteq \text{Human}$ follows from $\mathcal{O}$, but also from $\mathcal{M}$.

The same in first-order logic

$$
\begin{array}{rl}
\forall x.\text{Male}(x) & \leftrightarrow \text{Human}(x) \wedge \neg\text{Female}(x), \\
\forall x.\text{Human}(x) & \rightarrow \forall y.\text{has\_child}(x, y) \rightarrow \text{Human}(y) \\
\forall x.\text{Father}(x) & \rightarrow \text{Human}(x), \\
\forall x.\text{Father}(x) & \leftrightarrow \text{Male}(x) \wedge \exists y.\text{has\_child}(x, y)
\end{array}
$$

Goal of this work

- Generalise the notion of module (extraction) to an arbitrary logical system
- Provide a semantics for module extraction in DOL

Institutions (Goguen/Burstall 1984)

Definition

An institution consists of

- a category **Sign** of *signatures*,
- a sentence functor $\mathbf{Sen}\colon \mathbf{Sign} \longrightarrow \mathbb{S}et$
  - for $\sigma\colon \Sigma \longrightarrow \Sigma'$, we have $\mathbf{Sen}(\sigma)\colon \mathbf{Sen}(\Sigma) \longrightarrow \mathbf{Sen}(\Sigma')$,
- a model functor $\mathbf{Mod}\colon \mathbf{Sign}^{op} \longrightarrow \mathbb{C}at$
  - for $\sigma\colon \Sigma \longrightarrow \Sigma'$, we have $\mathbf{Mod}(\sigma)\colon \mathbf{Mod}(\Sigma') \longrightarrow$,
- a satisfaction relation $\models_\Sigma \subseteq |\mathbf{Mod}(\Sigma)| \times \mathbf{Sen}(\Sigma)$,

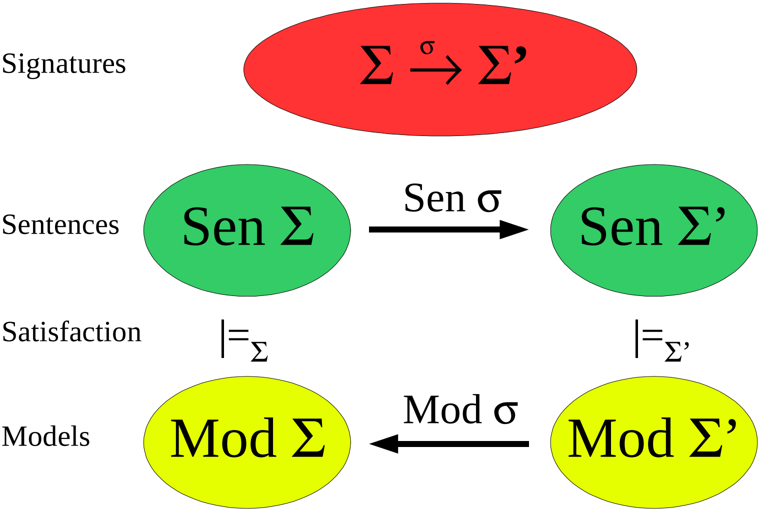
such that the following satisfaction condition holds:

$$M' \models_{\Sigma'} \mathbf{Sen}(\sigma)(\varphi) \text{ if and only if } \mathbf{Mod}(\sigma)(M)' \models_\Sigma \varphi$$

or shortly

$$M' \models_{\Sigma'} \sigma(\varphi) \text{ if and only if } M'|_\sigma \models_\Sigma \varphi$$

Institutions: formalisation of notion of logical system



Sample institutions

- propositional logic
- description logics, OWL
- first-order, higher-order logic, polymorphic logics
- logics of partial functions
- modal logic, epistemic logic, deontic logic, logics of knowledge and belief, agent logics
- μ-calculus, dynamic logic
- spatial logics, temporal logics, process logics, object logics
- intuitionistic logic
- linear logic, non-monotone logics, fuzzy logics
- paraconsistent logic, database query languages

Inclusive Categories

Definition

An inclusive category is a category with a broad subcategory[a] which is a partially ordered class with a least element (denoted $\emptyset$), non-empty products (denoted $\cap$) and finite coproducts (denoted $\cup$), such that for each pair of objects A, B, the following is a pushout in the category:

$$\begin{tikzcd} A \cap B \arrow[r, hook] \arrow[d, hook] & A \arrow[d, hook] \\ B \arrow[r, hook] & A \cup B \end{tikzcd}$$

[a] That is, with the same objects as the original category.

For any objects A and B of an inclusive category, we write $A \subseteq B$ if there is an inclusion from A to B; the unique such inclusion will then be denoted by $\iota_{A \subseteq B} \colon A \hookrightarrow B$, or simply $A \hookrightarrow B$.

Definition

An institution $I = (\mathbf{Sign}, \mathbf{Sen}, \mathbf{Mod}, \models)$ is inclusive if

- **Sign** is an inclusive category,
- **Sen** is inclusive and preserves intersections,[a] and
- each model category is inclusive, and reduct functors are inclusive.[b]

Moreover, we asume that reducts w.r.t. signature inclusions are surjective on objects.

[a] That is, for any family of signatures $\mathbb{S} \subseteq |\mathbf{Sign}|$, $\mathbf{Sen}(\bigcap \mathbb{S}) = \bigcap_{\Sigma \in \mathbb{S}} \mathbf{Sen}(\Sigma)$.

[b] That is, we have a model functor $\mathbf{Mod} \colon \mathbf{Sign}^{op} \to \mathbb{IC}at$, where $\mathbb{IC}at$ is the (quasi)category of inclusive categories and inclusive functors.

Notation

In inclusive institutions, if $\Sigma_1 \subseteq \Sigma_2$ via an inclusion $\iota\colon \Sigma_1 \hookrightarrow \Sigma_2$ and $M \in \mathbf{Mod}(\Sigma_2)$, we write $M|_{\Sigma_1}$ for $M|_\iota$.

$\mathbf{Sen}(\iota)\colon \mathbf{Sen}(\Sigma_1) \to \mathbf{Sen}(\Sigma_2)$ is the usual set-theoretic inclusion, hence its application may be omitted.

(Weakly) Union-exact Institutions

Definition

An inclusive institution I is called (weakly) union-exact, if all intersection-union signature pushouts in **Sign** are (weakly) amalgamable. More specifically, the latter means that for any pushout

$$\begin{tikzcd} \Sigma_1 \cap \Sigma_2 \arrow[r] \arrow[d] & \Sigma_1 \arrow[d] \\ \Sigma_2 \arrow[r] & \Sigma_1 \cup \Sigma_2 \end{tikzcd}$$

in **Sign**, any pair $(M_1, M_2) \in \mathbf{Mod}(\Sigma_1) \times \mathbf{Mod}(\Sigma_2)$ that is compatible in the sense that M_1 and M_2 reduce to the same $(\Sigma_1 \cap \Sigma_2)$-model can be amalgamated to a unique (or weakly amalgamated to a not necessarily unique) $(\Sigma_1 \cup \Sigma_2)$-model: there exists a (unique) $M \in \mathbf{Mod}(\Sigma_1 \cup \Sigma_2)$ that reduces to M_1 and M_2, respectively.

Presentations

Definition

A presentation in an institution $I = (\mathbf{Sign}, \mathbf{Sen}, \mathbf{Mod}, \models)$ is a pair

$$P = (\Sigma, \Phi),$$

where $\Sigma \in |\mathbf{Sign}|$ is a signature and $\Phi \subseteq \mathbf{Sen}(\Sigma)$ is a set of Σ-sentences.

Σ is also denoted as $\mathrm{Sig}(P)$, Φ as $\mathrm{Ax}(P)$.
We extend the model functor to presentations and write $\mathbf{Mod}(\Sigma, \Phi)$ (or sometimes $\mathbf{Mod}(\Phi)$ if the signature is clear) for the full subcategory of $\mathbf{Mod}(\Sigma)$ that consists of the models of (Σ, Φ), i.e., $|\mathbf{Mod}(\Sigma, \Phi)| = \{M \in |\mathbf{Mod}(\Sigma)| \mid M \models_\Sigma \Phi\}$.

Ontologies

Definition

An ontology $\mathcal{O}$ in a logic given as the institution I is just a set of sentences

For each ontology $\mathcal{O}$, its signature $\mathrm{Sig}(\mathcal{O})$ is the least signature over which all the sentences in $\mathcal{O}$.

Note: the standard institutional concept to consider ontologies as presentations does not work for the definition of module below.

Conservative extensions

Definition

Consider ontologies $\mathcal{O}' \subseteq \mathcal{O}$ and a signature $\Sigma \in |\mathbf{Sign}|$.

1. $\mathcal{O}$ is a model Σ-conservative extension (Σ-mCE) of $\mathcal{O}'$, if

   for every $(\mathrm{Sig}(\mathcal{O}') \cup \Sigma)$*-model* $\mathcal{I}'$ *of* $\mathcal{O}'$*, there exists a* $(\mathrm{Sig}(\mathcal{O}) \cup \Sigma)$*-model* $\mathcal{I}$ *of* $\mathcal{O}$ *such that* $\mathcal{I}'|_\Sigma = \mathcal{I}|_\Sigma$.

2. $\mathcal{O}$ is a consequence Σ-conservative extension (Σ-cCE) of $\mathcal{O}'$, if for every Σ-sentence α, we have

$$\mathcal{O} \models \alpha \text{ iff } \mathcal{O}' \models \alpha.$$

For an ontology $\mathcal{O}$ and a signature Σ, we define

$$\mathcal{O}{\uparrow}\Sigma = (\mathsf{Sig}(\mathcal{O}) \cup \Sigma, \mathsf{Ax}(\mathcal{O})).$$

Definition

Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be ontologies and Σ a signature. Then $\mathcal{O}_1$ and $\mathcal{O}_2$ are model Σ-inseparable, written $\mathcal{O}_1 \equiv^m_\Sigma \mathcal{O}_2$ if,

$$\{\mathcal{I}|_\Sigma \mid \mathcal{I} \in |\mathbf{Mod}(\mathcal{O}_1{\uparrow}\Sigma)|\} = \{\mathcal{I}|_\Sigma \mid \mathcal{I} \in |\mathbf{Mod}(\mathcal{O}_2{\uparrow}\Sigma)|\}$$

Note that in the literature, a simpler condition is usually used:

$$\{\mathcal{I}|_\Sigma \mid \mathcal{I} \models \mathcal{O}_1\} = \{\mathcal{I}|_\Sigma \mid \mathcal{I} \models \mathcal{O}_2\}$$

However, this is wrong! Consider:

$$\{C \sqsubseteq C\} \equiv^m_{\{C,C'\}} \{C' \sqsubseteq C'\}$$

Consequence Inseparability

Definition

$\mathcal{O}_1$ and $\mathcal{O}_2$ are consequence Σ-inseparable, written $\mathcal{O}_1 \equiv^s_\Sigma \mathcal{O}_2$, if for all Σ-sentences φ

$$\mathcal{O}_1 \models \varphi \text{ iff } \mathcal{O}_2 \models \varphi$$

Proposition

Model-theoretic inseparability implies consequence inseparability, but not vice versa.

Definition (Kontchakov/Wolter/Zakharyaschev 2011)

An inseparability relation $\mathcal{S} = \langle \equiv^{\mathcal{S}}_{\Sigma} \rangle_{\Sigma \in |\mathbf{Sign}|}$ is a family of equivalence relations. It is monotone if

1. for any signatures $\Sigma' \subseteq \Sigma$, $\equiv^{\mathcal{S}}_{\Sigma} \subseteq \equiv^{\mathcal{S}}_{\Sigma'}$
   Intuition:
   the inseparability relation gets finer when the signature gets larger
2. if $\mathcal{O}_1 \subseteq \mathcal{O}_2 \subseteq \mathcal{O}_3$ and $\mathcal{O}_1 \equiv^{\mathcal{S}}_{\Sigma} \mathcal{O}_3$ then $\mathcal{O}_1 \equiv^{\mathcal{S}}_{\Sigma} \mathcal{O}_2$ and $\mathcal{O}_2 \equiv^{\mathcal{S}}_{\Sigma} \mathcal{O}_3$
   Intuition:
   since larger ontologies capture more of "the knowledge of interest", we also require that any ontology squeezed between an ontology and its inseparable extension is inseparable from both of them.

Proposition

$\langle \equiv^{m}_{\Sigma} \rangle_{\Sigma \in |\mathbf{Sign}|}$ *and* $\langle \equiv^{s}_{\Sigma} \rangle_{\Sigma \in |\mathbf{Sign}|}$ *are monotone.*

Robustness

Definition

An inseparability relation $\mathcal{S} = \langle \equiv^{\mathcal{S}}_{\Sigma} \rangle_{\Sigma \in |\mathbf{Sign}|}$ is

- robust under signature extensions, if for all ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$ and all signatures Σ, Σ' with $\Sigma' \cap (\mathrm{Sig}(\mathcal{O}_1) \cup \mathrm{Sig}(\mathcal{O}_2)) \subseteq \Sigma$

$$\mathcal{O}_1 \equiv_{\Sigma} \mathcal{O}_2 \text{ implies } \mathcal{O}_1 \equiv_{\Sigma'} \mathcal{O}_2$$

- robust under replacement if for all ontologies $\mathcal{O}$, $\mathcal{O}_1$ and $\mathcal{O}_2$ and all signatures Σ with $\mathrm{Sig}(\mathcal{O}) \subseteq \Sigma$, we have

$$\mathcal{O}_1 \equiv_{\Sigma} \mathcal{O}_2 \text{ implies } \mathcal{O}_1 \cup \mathcal{O} \equiv_{\Sigma} \mathcal{O}_2 \cup \mathcal{O}$$

- robust under joins, if for all ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$ and all signatures Σ with $\mathrm{Sig}(\mathcal{O}_1) \cap \mathrm{Sig}(\mathcal{O}_2) \subseteq \Sigma$, we have for $i = 1, 2$

$$\mathcal{O}_1 \equiv_{\Sigma} \mathcal{O}_2 \text{ implies } \mathcal{O}_i \equiv_{\Sigma} \mathcal{O}_1 \cup \mathcal{O}_2$$

Robustness for Arbitrary Institutions

Theorem

Model inseparability is robust under replacement.

In a union-exact inclusive institution, model inseparability is also robust under signature extensions and joins.

Ontology Modules

Definition (Kontchakov/Wolter/Zakharyaschev 2011)

Let $\mathcal{O}$ be an ontology, $\mathcal{M} \subseteq \mathcal{O}$ and Σ a signature. We call $\mathcal{M}$

- a (plain) Σ-S-module of $\mathcal{O}$ induced by $\mathcal{S}$ if $\mathcal{M} \equiv_{\Sigma}^{\mathcal{S}} \mathcal{O}$;
- a self-contained Σ-S-module of $\mathcal{O}$ induced by $\mathcal{S}$ if $\mathcal{M} \equiv_{\Sigma \cup \mathsf{Sig}(\mathcal{M})}^{\mathcal{S}} \mathcal{O}$;
- a depleting Σ-S-module of $\mathcal{O}$ induced by $\mathcal{S}$ if $\mathcal{O} \setminus \mathcal{M} \equiv_{\Sigma \cup \mathsf{Sig}(\mathcal{M})}^{\mathcal{S}} \emptyset$.

Proposition

For any ontology $\mathcal{O}$, $\mathcal{M} \subseteq \mathcal{O}$ and signature Σ, $\mathcal{M}$ is a Σ-m-module of $\mathcal{O}$ if and only if $\mathcal{O}$ is a model Σ-conservative extension of $\mathcal{M}$.

Proposition

$\mathcal{M}$ self-contained Σ-m-module of $\mathcal{O}$ $\Rightarrow$ $\mathcal{M}$ (plain) Σ-m-module of $\mathcal{O}$.

$\mathcal{M}$ depleting Σ-m-module of $\mathcal{O}$ $\Rightarrow$ $\mathcal{M}$ self-contained Σ-m-module of $\mathcal{O}$.

Robustness for Modules

Robustness under signature restrictions.
A module of an ontology w.r.t. a signature Σ is also a module of this ontology w.r.t. any subsignature of Σ.
Intuition: We do not need to import a different module when we restrict the set of terms that we are interested in.
Robustness under signature extensions.
A module of an ontology $\mathcal{O}$ w.r.t. a signature Σ is also a module of $\mathcal{O}$ w.r.t. any $\Sigma' \supseteq \Sigma$, if $\Sigma' \cap \mathrm{Sig}(\mathcal{O}) \subseteq \Sigma$.
Intuition: we do not need to import a different module when extending the set of relevant terms with terms not from $\mathcal{O}$.
Robustness under replacement.
If $\mathcal{M}$ is a module of $\mathcal{O}$ w.r.t. Σ, then the result of importing $\mathcal{M}$ into another ontology $\mathcal{O}'$ is a module of the result of importing $\mathcal{O}$ into $\mathcal{O}'$:

$$\mathcal{M} \text{ is } \Sigma\text{-module of } \mathcal{O} \Rightarrow \mathcal{O}' \cup \mathcal{M} \text{ is } \Sigma\text{-module of } \mathcal{O}' \cup \mathcal{O}$$

This is called module coverage in the literature: importing a module does not affect its property of being a module.

Properties	**Module Notions**		
	plain	self-contained	depleting
inseparability	$\mathcal{O} \equiv^m_\Sigma \mathcal{M}$	$\mathcal{O} \equiv^m_{\Sigma \cup \mathsf{Sig}(\mathcal{M})} \mathcal{M}$	$\mathcal{O} \setminus \mathcal{M} \equiv^m_{\Sigma \cup \mathsf{Sig}(\mathcal{M})} \emptyset$
mCE (dCE)	✓	✓	✓
self-contained	✗	✓	✓
depleting	✗	✗	✓
robustness under signature restrictions	✓	✓	✓
robustness under signature extensions	$\Sigma' \cap \mathsf{Sig}(\mathcal{O}) \subseteq \Sigma$ plus weak union-exactness	$\Sigma' \cap \mathsf{Sig}(\mathcal{O}) \subseteq \Sigma$ plus weak union-exactness	$\Sigma' \cap \mathsf{Sig}(\mathcal{O}) \subseteq \Sigma$ plus weak union-exactness
robustness under replacement	$\mathsf{Sig}(\mathcal{O}') \subseteq \Sigma$	$\mathsf{Sig}(\mathcal{O}') \cap \mathsf{Sig}(\mathcal{O}) \subseteq \Sigma \cup \mathsf{Sig}(\mathcal{M})$	$\mathsf{Sig}(\mathcal{O}') \cap \mathsf{Sig}(\mathcal{O}) \subseteq \Sigma \cup \mathsf{Sig}(\mathcal{M})$

Minimum Modules

Theorem (Kontchakov/Wolter/Zakharyaschev 2011)

Let $\mathcal{O}$ be an ontology and Σ be a signature. Then there is a minimum depleting Σ-m-module of $\mathcal{O}$ (indeed, a minimum depleting Σ-S-module for any monotone inseparability relation S robust under replacement.)

$$
\begin{array}{l}
\texttt{input } \mathcal{T}, \Sigma \\
\mathcal{M} := \emptyset; \\
\mathcal{W} := \emptyset; \\
\texttt{while } (\mathcal{T} \setminus \mathcal{M}) \neq \mathcal{W} \texttt{ do} \\
\quad \texttt{choose } \alpha \in (\mathcal{T} \setminus \mathcal{M}) \setminus \mathcal{W} \\
\quad \mathcal{W} := \mathcal{W} \cup \{\alpha\}; \\
\quad \texttt{if } \mathcal{W} \not\equiv_{\Sigma \cup \mathsf{sig}(\mathcal{M})} \emptyset \texttt{ then} \\
\quad\quad \mathcal{M} := \mathcal{M} \cup \{\alpha\}; \\
\quad\quad \mathcal{W} := \emptyset \\
\quad \texttt{endif} \\
\texttt{end while} \\
\texttt{output } \mathcal{M}
\end{array}
$$

By contrast, minimum plain or self-contained modules not always exist.
$\Rightarrow$ use minimum depleting Σ-module for DOL semantics.

Conclusions and Future work

Conclusions

- Generalised various module notions and theorems to an arbitrary institution
- corrected a small error in the definition of model inseparability
- found a semantics for module extraction in DOL: the minimum depleting Σ-module

Future work

- efficient computability of modules
- generalise various notions of locality to an arbitrary institution