An Approach for Repairing Ontologies and Ontology networks

Ying Li

Main Supervisor: Patrick Lambrix

Co-Supervisor: Olaf Hartig

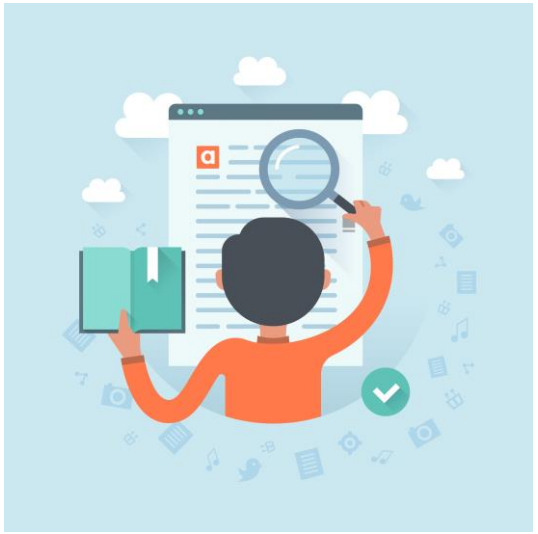
LINKÖPING UNIVERSITY

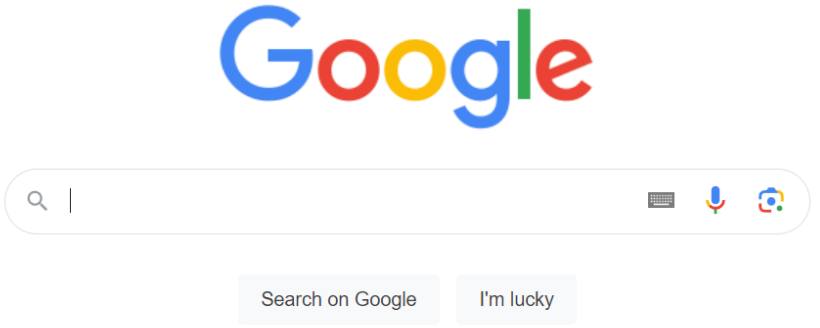
Agenda

- Background and Motivation
- Preliminaries
- Problem Formulation
- Repairing Approach
- Conclusion and Future work

Background and Motivation

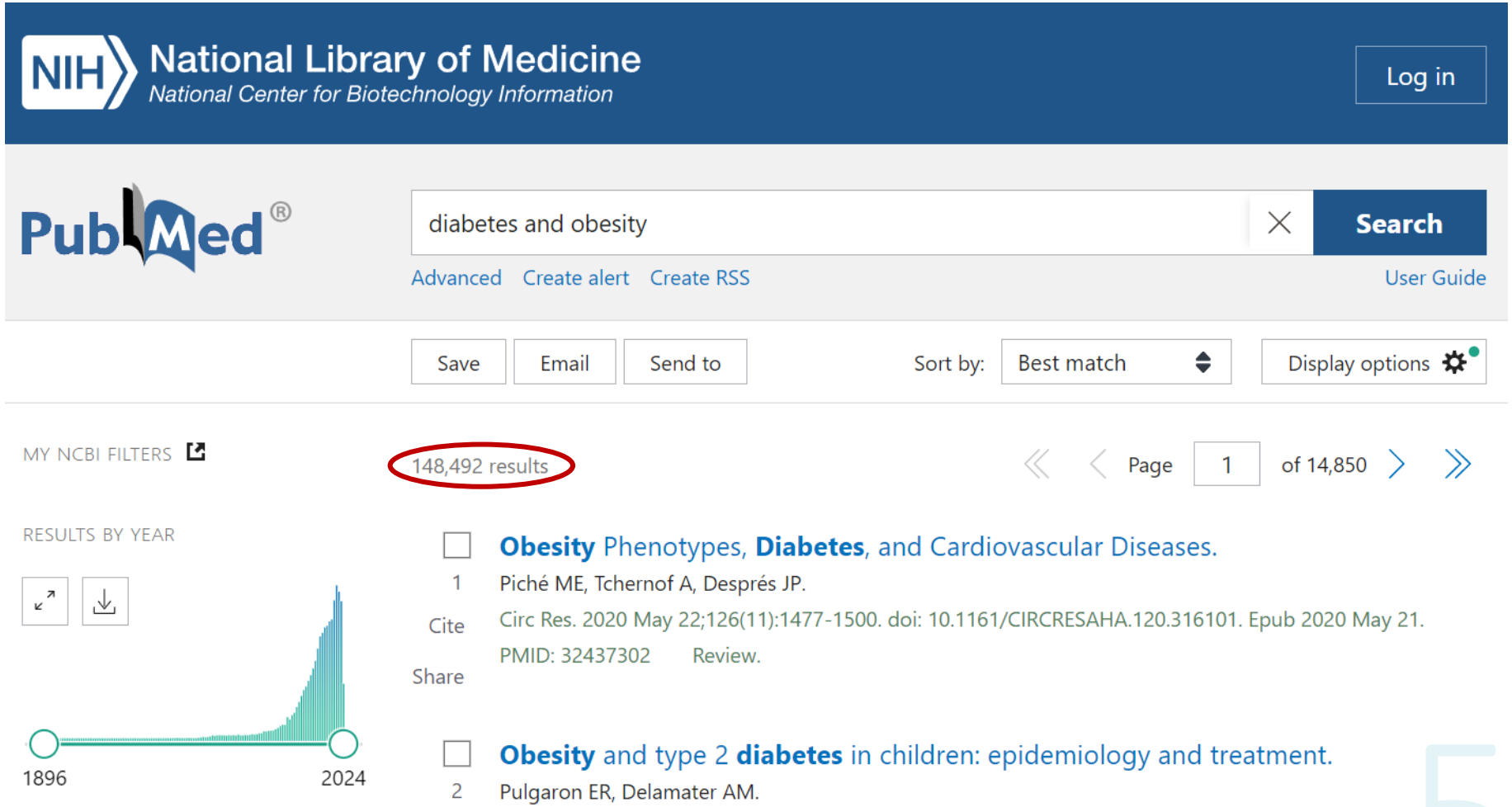
A researcher is conducting a literature review on the relationship between diabetes and obesity to identify potential areas for further research……





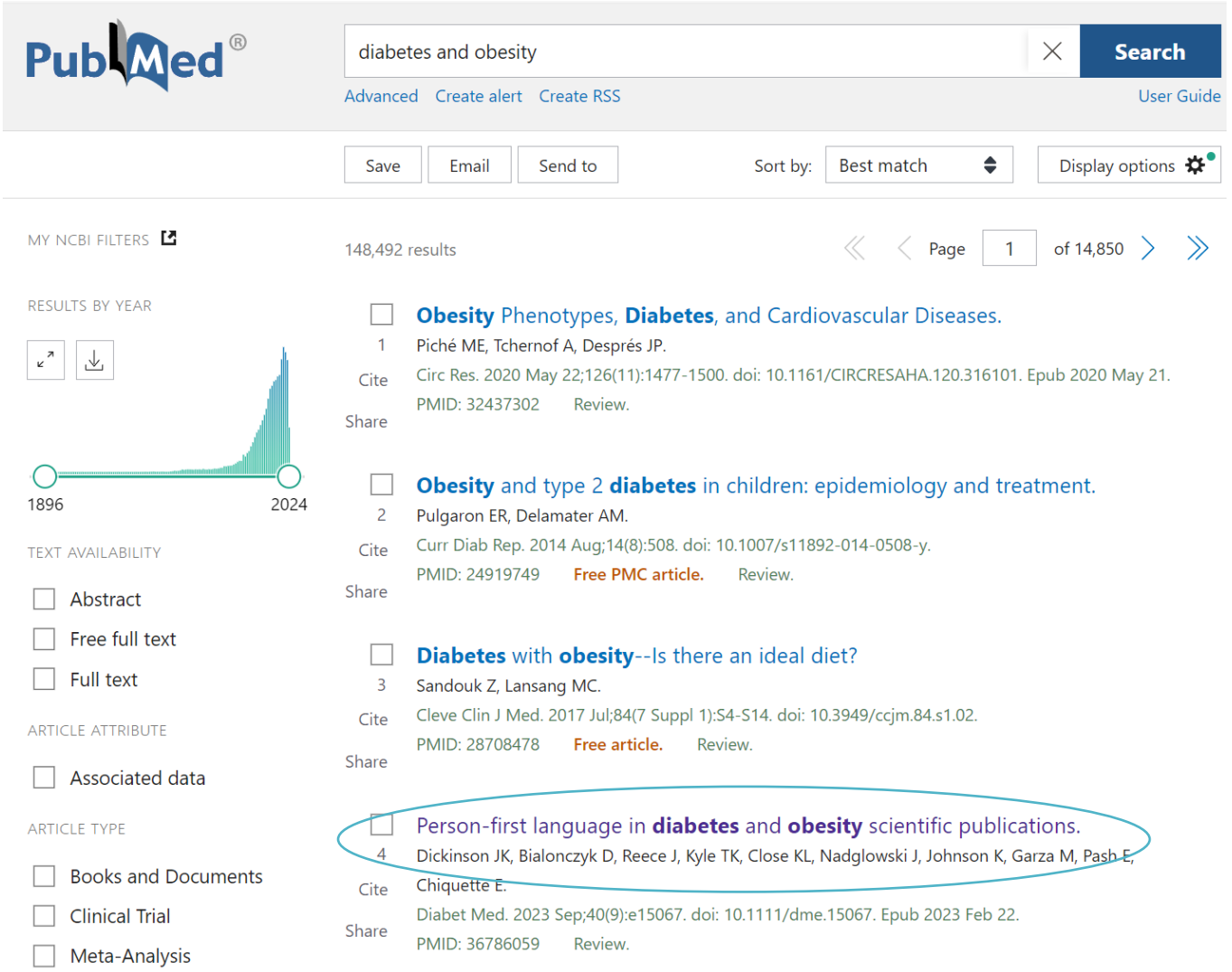
Keyword-based search

overwhelming number of results



Keyword-based search

- ❑Keyword Ambiguity
- ❑Synonymy
- ❑Lack of Context



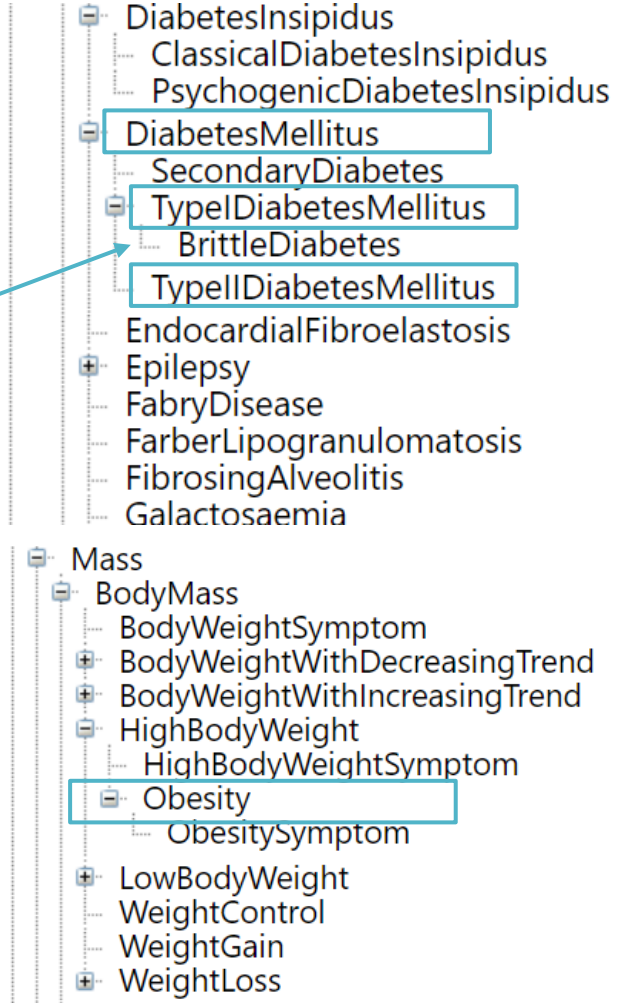
Ontology-based search

- Standardized Terminology

➢ A biomedical ontology contains concepts related to diabetes, obesity, and their interrelationship.

National Center for Biomedical Ontology (NCBO) BioPortal

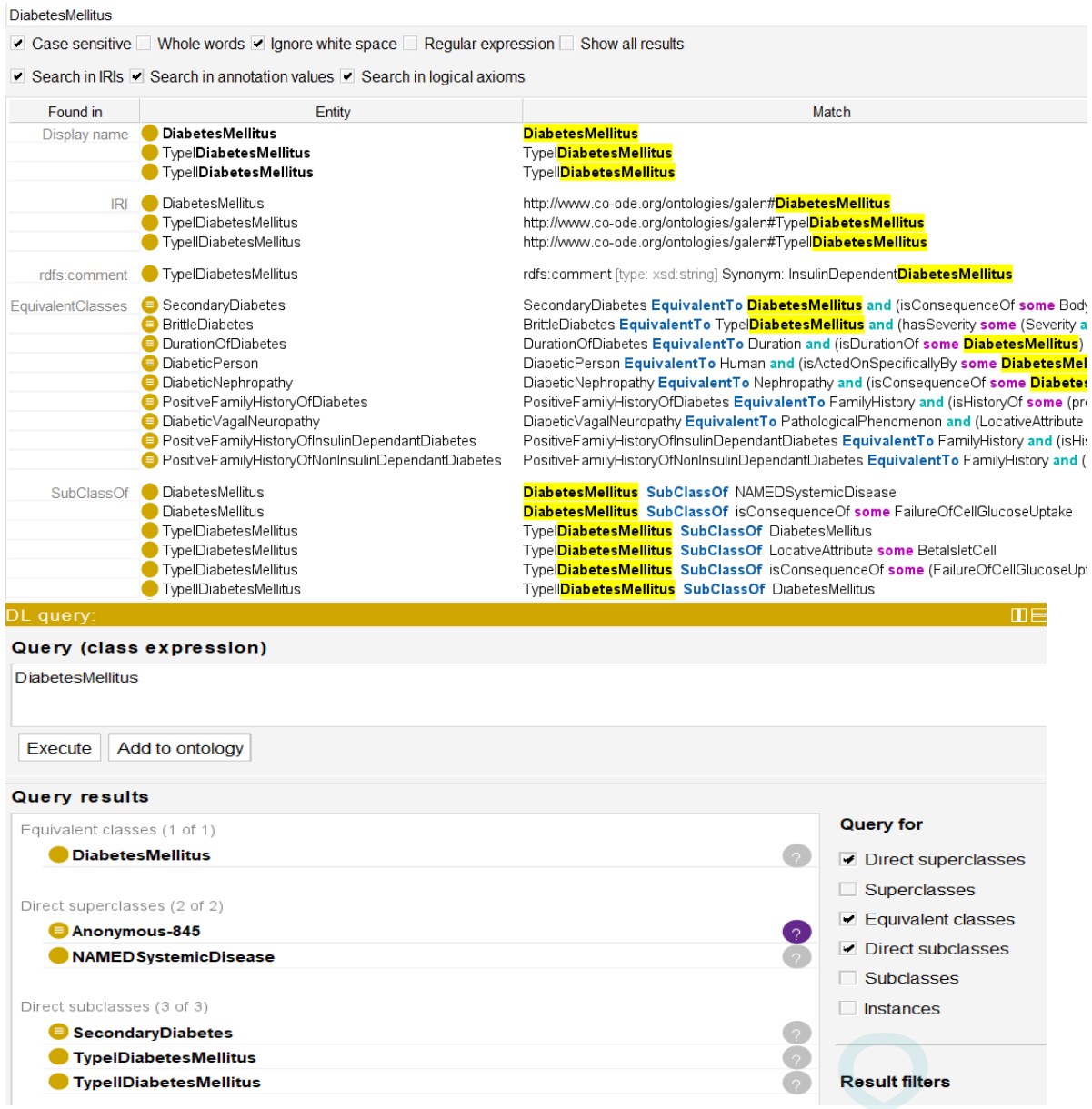
e.g., Concepts: DiabetesMellitus and Obesity



Ontology-based search

- Semantic Search
  - ➢specify the query in terms of synonyms, specific subtypes or aspects of diabetes…….

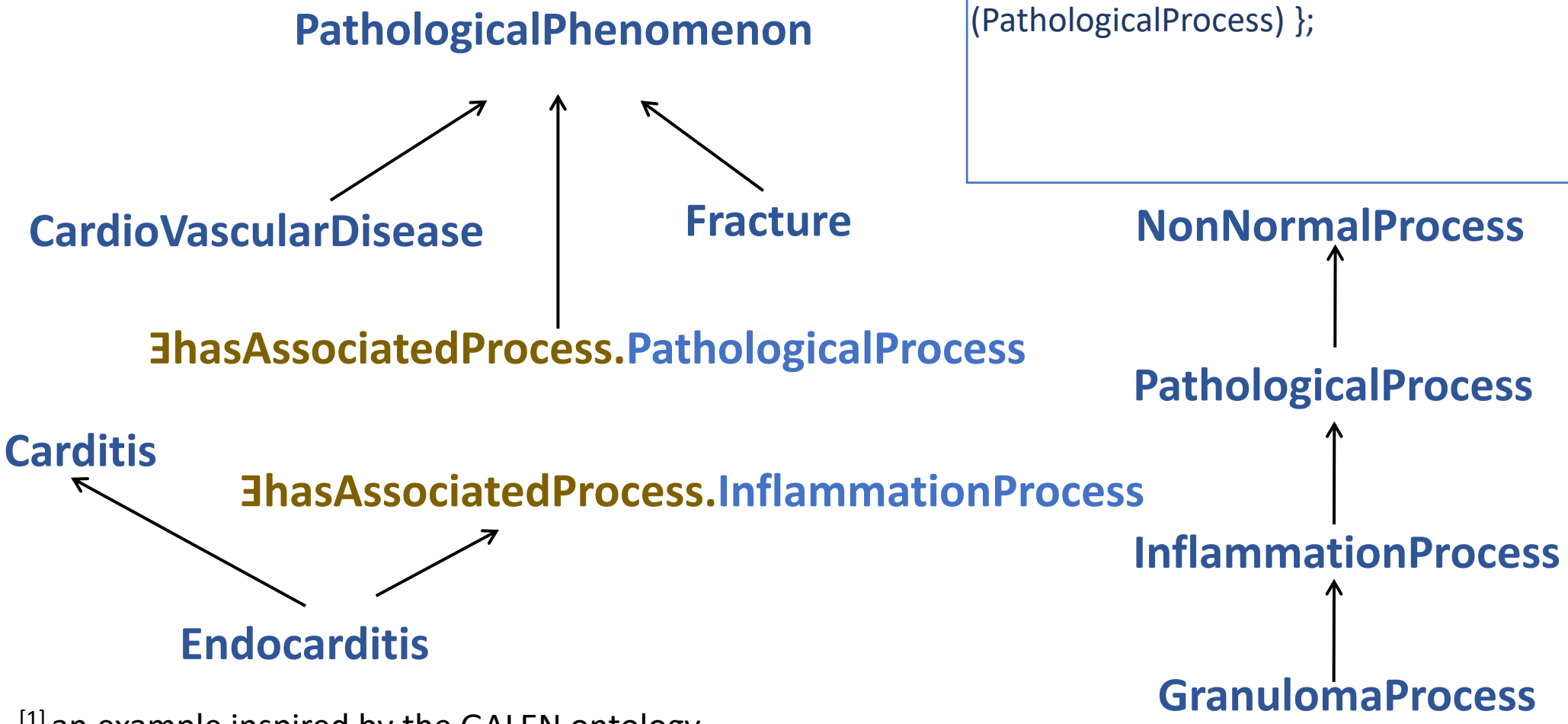
- Interoperability and data integration (Query Expansion)
  - ➢expand search queries to include synonyms, related terms…..



Example: Mini-GALEN[1] Ontology

Concepts

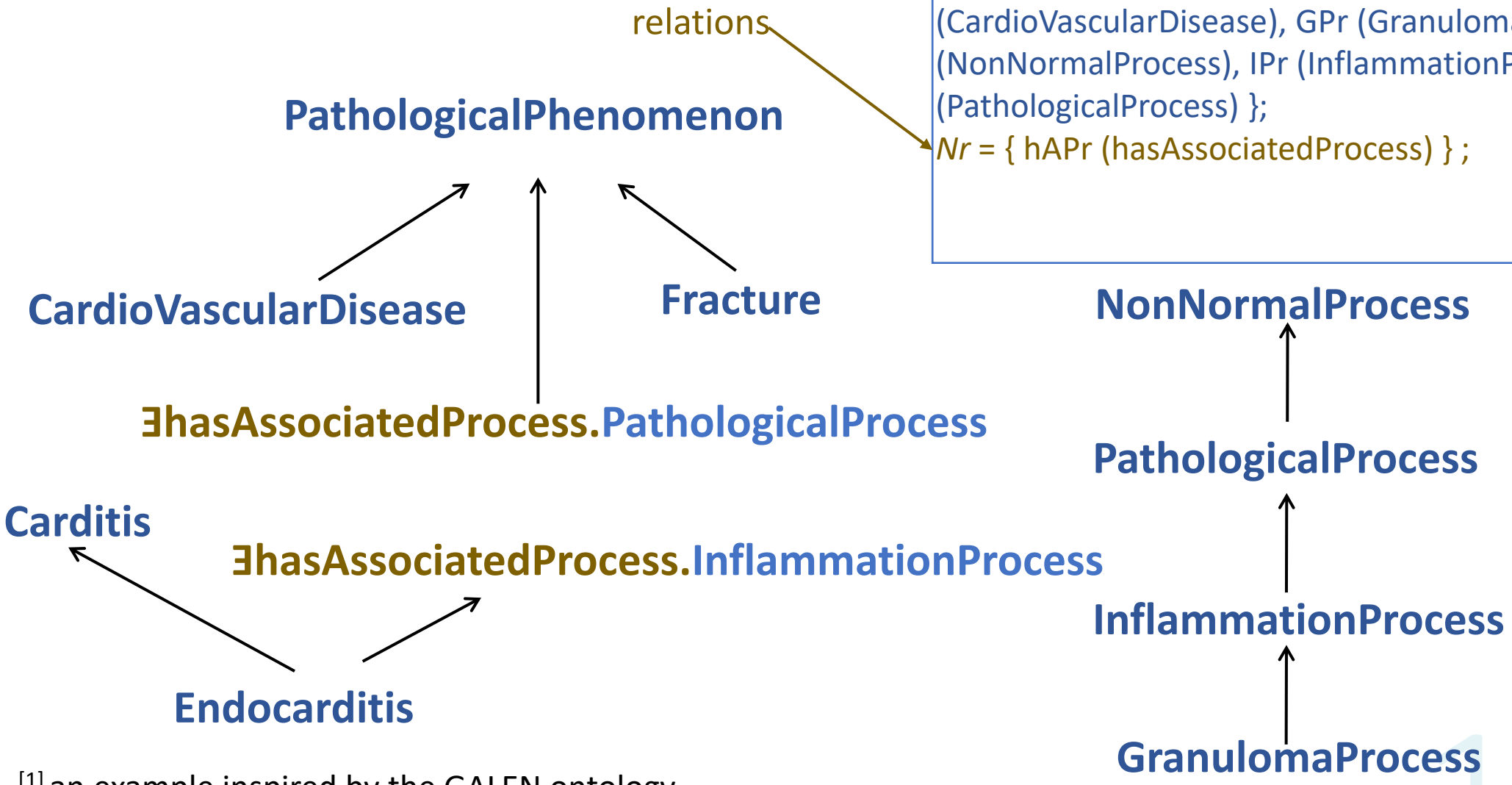
Nc = {PPh (PathologicalPhenomenon), F(Fracture), E (Endocarditis), C (Carditis), CVD (CardioVascularDisease), GPr (GranulomaProcess), NPr (NonNormalProcess), IPr (InflammationProcess), PPr (PathologicalProcess) };



[1] an example inspired by the GALEN ontology

Example: Mini-GALEN[1] Ontology

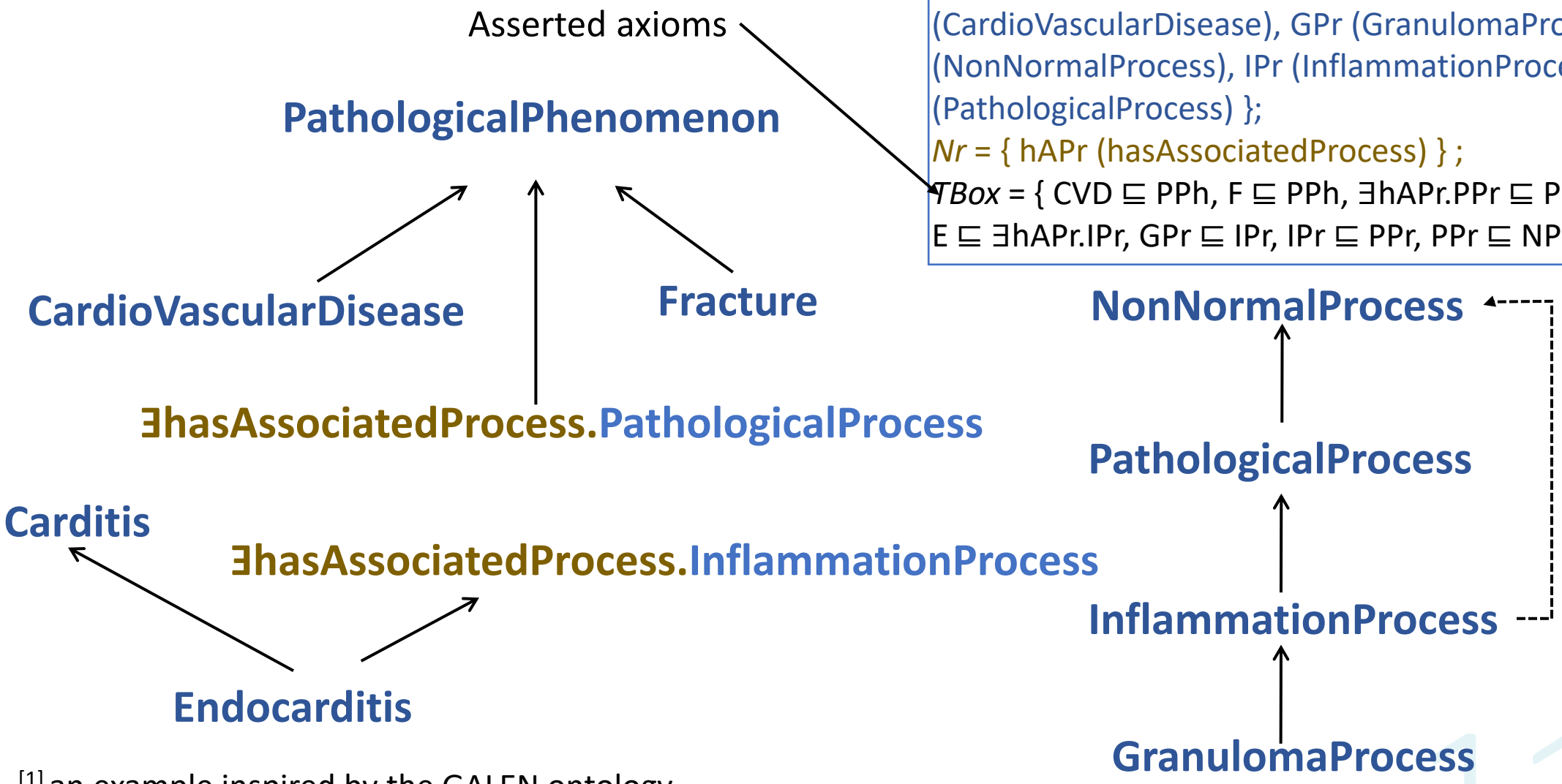
Nc = {PPh (PathologicalPhenomenon), F(Fracture), E (Endocarditis), C (Carditis), CVD (CardioVascularDisease), GPr (GranulomaProcess), NPr (NonNormalProcess), IPr (InflammationProcess), PPr (PathologicalProcess) };
Nr = { hAPr (hasAssociatedProcess) } ;



[1] an example inspired by the GALEN ontology

Example: Mini-GALEN[1] Ontology

Nc = {PPh (PathologicalPhenomenon), F(Fracture), E (Endocarditis), C (Carditis), CVD (CardioVascularDisease), GPr (GranulomaProcess), NPr (NonNormalProcess), IPr (InflammationProcess), PPr (PathologicalProcess) };
Nr = { hAPr (hasAssociatedProcess) } ;
TBox = { CVD ⊑ PPh, F ⊑ PPh, ∃hAPr.PPr ⊑ PPh, E ⊑ C, E ⊑ ∃hAPr.IPr, GPr ⊑ IPr, IPr ⊑ PPr, PPr ⊑ NPr }



[1] an example inspired by the GALEN ontology

Ontology Alignment

- Defining the relations between the terms in different ontologies
- An alignment: a set of mappings between ontologies

➢ generated by tools and domain experts

i.e., dead(Galen) ⊑ Dead(SNOMED)

Dead(SNOMED) ⊑ dead(Galen)

GALEN ONTOLOGY	SNOMED CT
Cefaclor	Cefaclor
SpecificGravity	Specific gravity
DoxazosinMesylate	Doxazosin mesylate
HighDensityLipoprotein	High density lipoprotein
dead	Dead
NephroticSyndrome	Nephrotic syndrome
Lysosome	Lysosome
Thyroiditis	Thyroiditis
Ethylenediamine	Ethylenediamine
Vanadium	Vanadium
Formication	Formication
Neisseria	Neisseria
TransOralApproach	Transoral approach

❖ An equivalence mapping P is equivalent to Q => two subsumption mappings P ⊑ Q and Q ⊑ P

Defects in Ontologies and Alignments

- Axioms in ontologies/mappings between ontologies are not correct.
  - It may lead to logical problems (i.e., incoherence or inconsistency) or statements that are not correct in the domain of the ontology

e.g., TAMBIS ontology contained 144 unsatisfiable concepts

- Ontologies/Ontology networks are not complete
  - missing relations

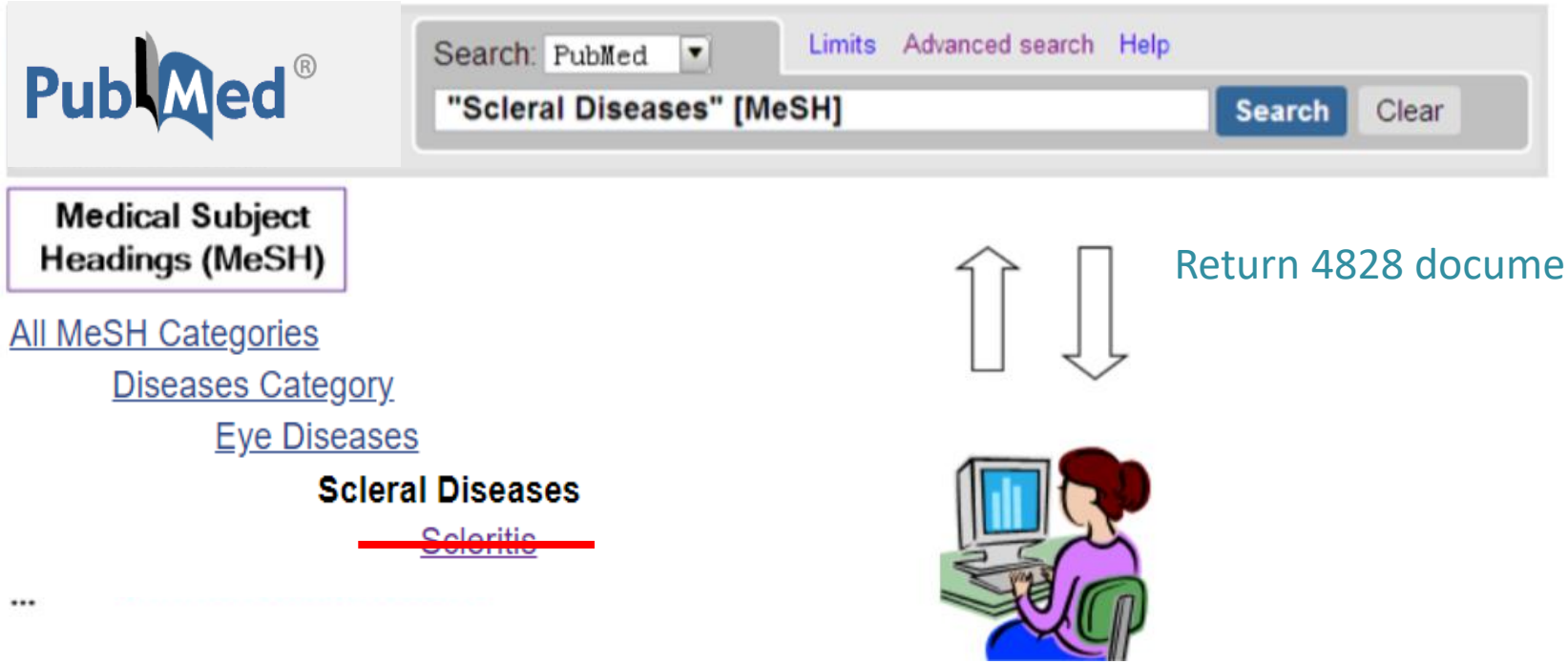
e.g., MA and NCI (Ontology Alignment Evaluation Initiative, Anatomy track), 121 and 83, is-a missing relations

Influence of Defects

- Incomplete results from the ontology based-query

Influence of Defects

- Incomplete results from the ontology based-query



Ontology debugging

- Semantically-enabled applications require high-quality ontologies and alignments.

➢Detection

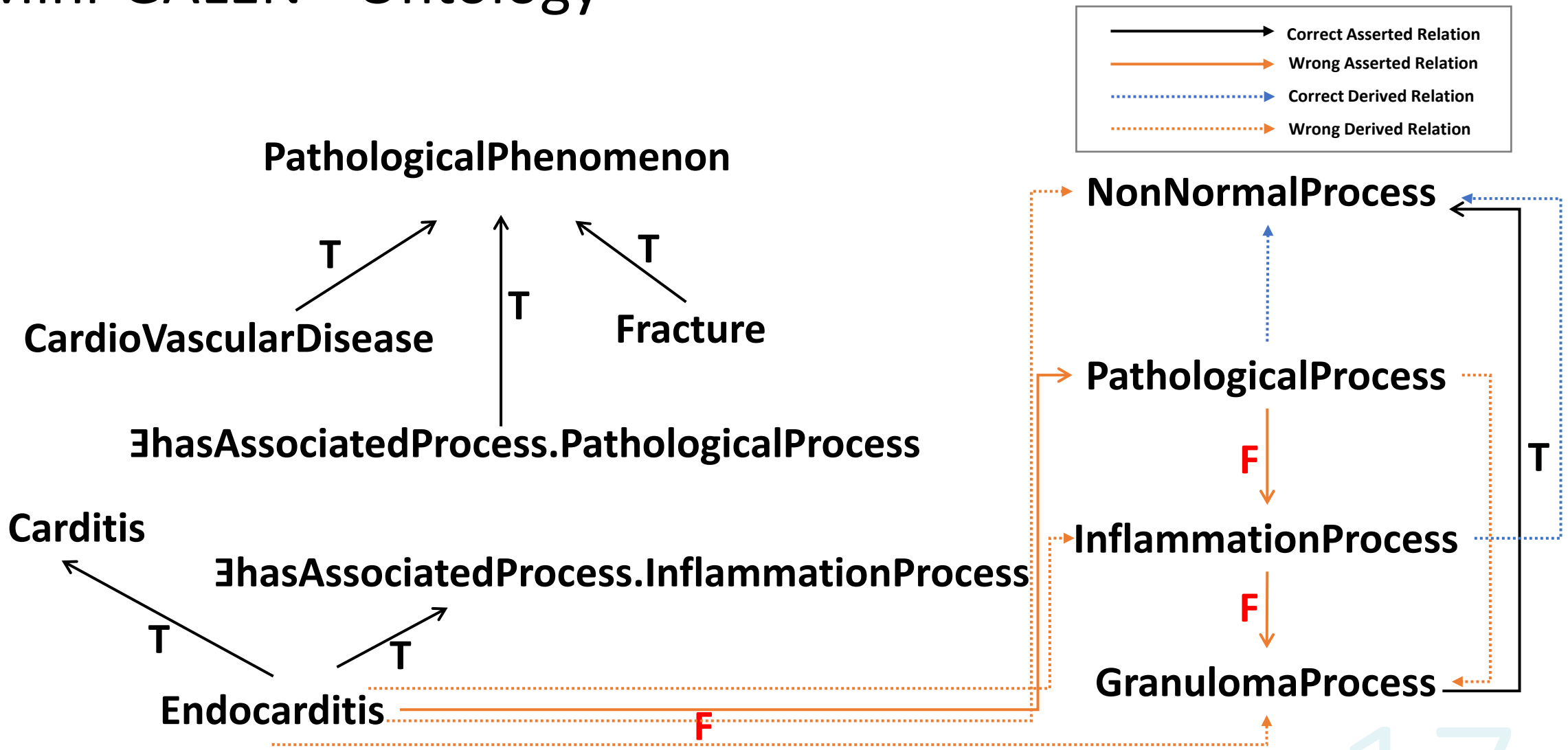
➢Repair

Removing (wrong relations)

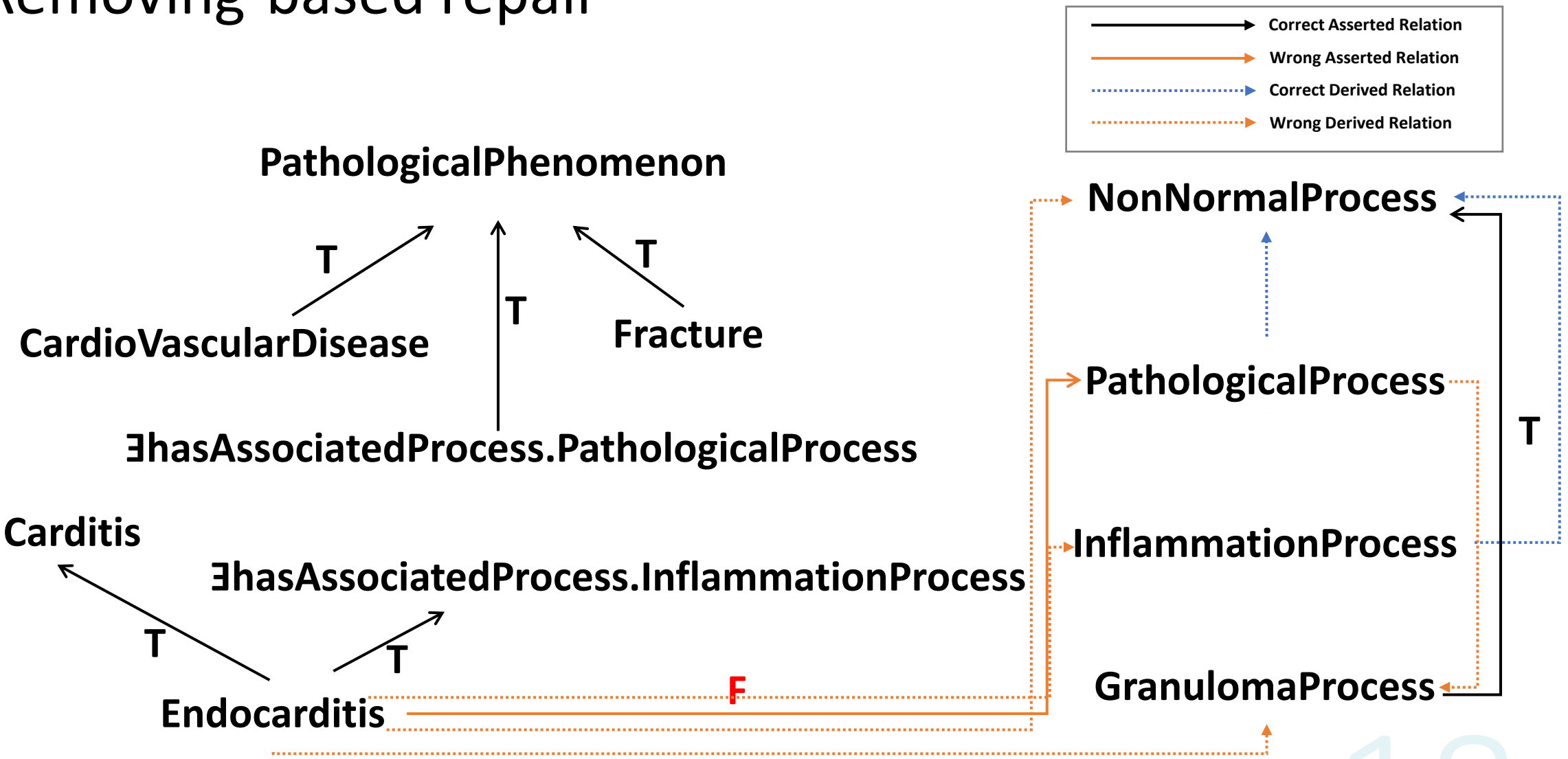
Completing (missing relations)

- Traditional debugging techniques repair ontologies/ontology networks with wrong knowledge by removing unwanted axioms and mappings, but may thereby remove consequences that are correct in the domain.

Mini-GALEN[1] Ontology



Removing-based repair



Remove all the wrong axioms from the ontology

Motivation

Preserve correct knowledge as much as possible when removing wrong axioms from the ontology/ontology network.

A domain expert (Oracle)
will validate the results

We proposed an interactive repairing approach to mitigate these effects of removing wrong axioms by, in addition to removing those axioms, also adding correct knowledge.

Completeness:
Weakening
Completion

Correctness:
Debugging
Removing

Preliminaries

$\mathcal{EL}_\bot$ Ontologies

- N_C: atomic concepts
- N_R: atomic roles
- A TBox is a finite set of axioms which in $\mathcal{EL}_\bot$ are general concept inclusions (GCIs).

Table 1. $\mathcal{EL}_\bot$ syntax and semantics.

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
bottom	$\bot$	$\emptyset$
conjunction	$P \sqcap Q$	$P^{\mathcal{I}} \cap Q^{\mathcal{I}}$
existential restriction	$\exists r.P$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x,y) \in r^{\mathcal{I}} \wedge y \in P^{\mathcal{I}}\}$
GCI	$P \sqsubseteq Q$	$P^{\mathcal{I}} \subseteq Q^{\mathcal{I}}$

❖ We base our work and examples on $\mathcal{EL}_\bot$, the discussions hold for ontologies represented by DLs in general.

Ontology network

Ontologies are represented using DL TBoxes.

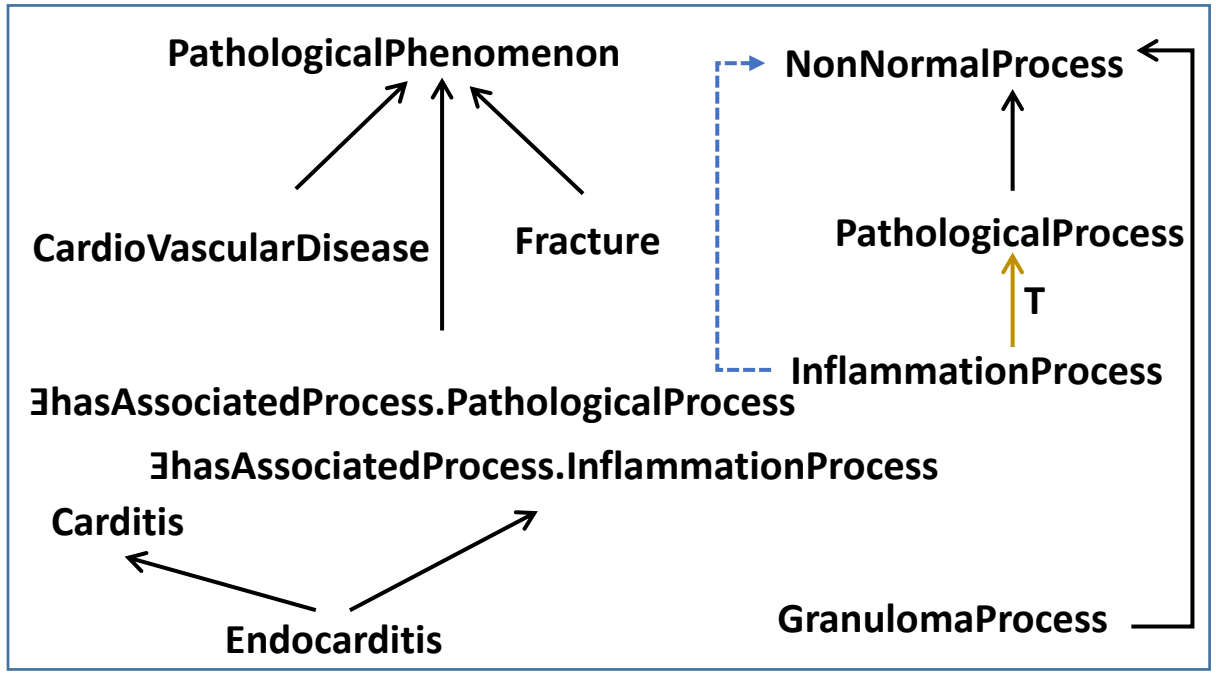
An alignment between two ontologies is a set of mappings between the ontologies.

Definition 1. Let $\mathcal{T}_1, ..., \mathcal{T}_n$ be TBoxes representing ontologies $\mathcal{O}_1, ..., \mathcal{O}_n$, respectively. For $i, j \in [1..n]$ with $i < j$, let $\mathcal{A}_{ij}$ be an alignment between ontology $\mathcal{O}_i$ and $\mathcal{O}_j$. The network of the ontologies and their alignments is then represented by TBox $\mathcal{T} = (\bigcup_{i=1..n} \mathcal{T}_i) \cup (\bigcup_{i,j=1..n,i<j} \mathcal{A}_{ij})$.

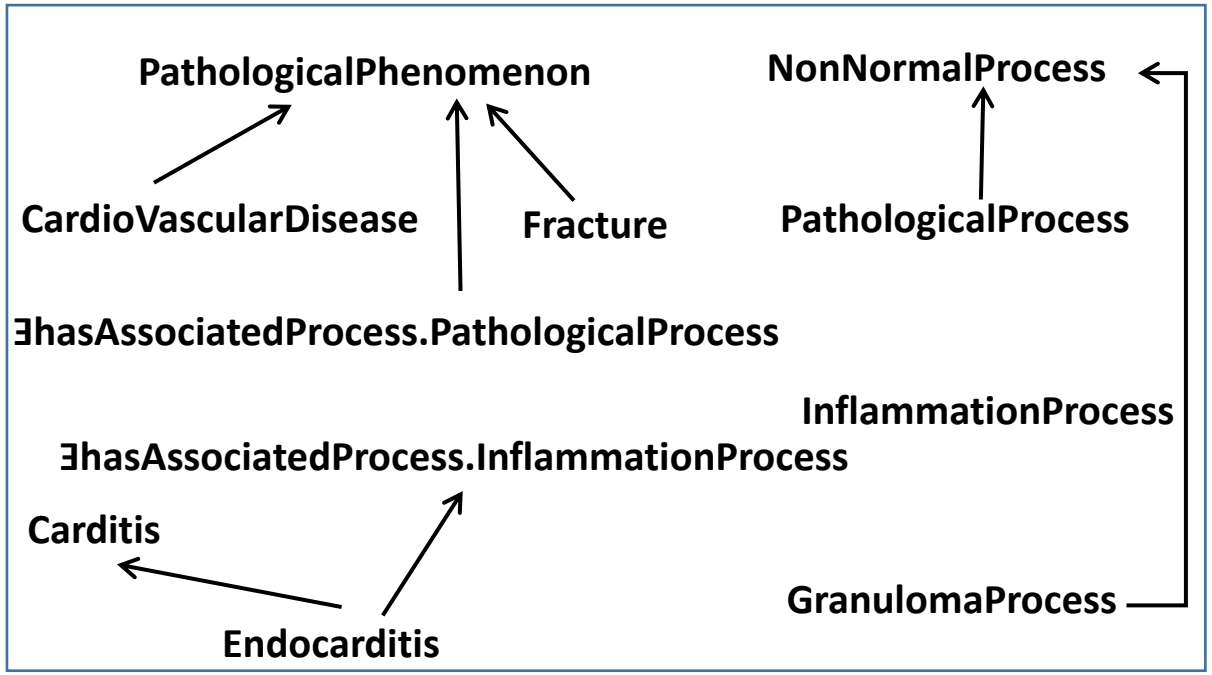
More complete TBoxes

Definition 2. TBox T_1 is more complete than *TBox T_2* iff

1) all correct knowledge in T_2 can also be derived in T_1.
2) there is correct knowledge in T_1 that cannot be derived in T_2.



T1



T2

Or(InflammationProcess ⊑ PathologicalProcess) = true

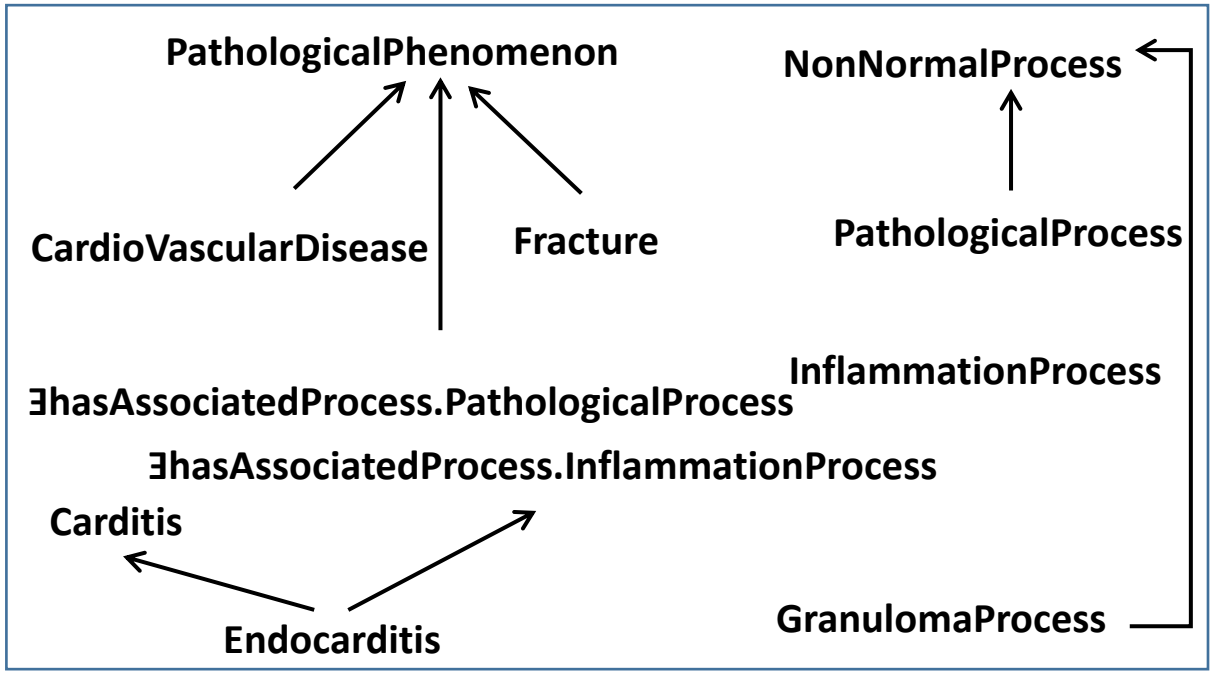
T_1 is more complete than T_2

Or: an oracle, when given an axiom, can answer whether this axiom is correct or wrong in the domain of interest of the ontology.

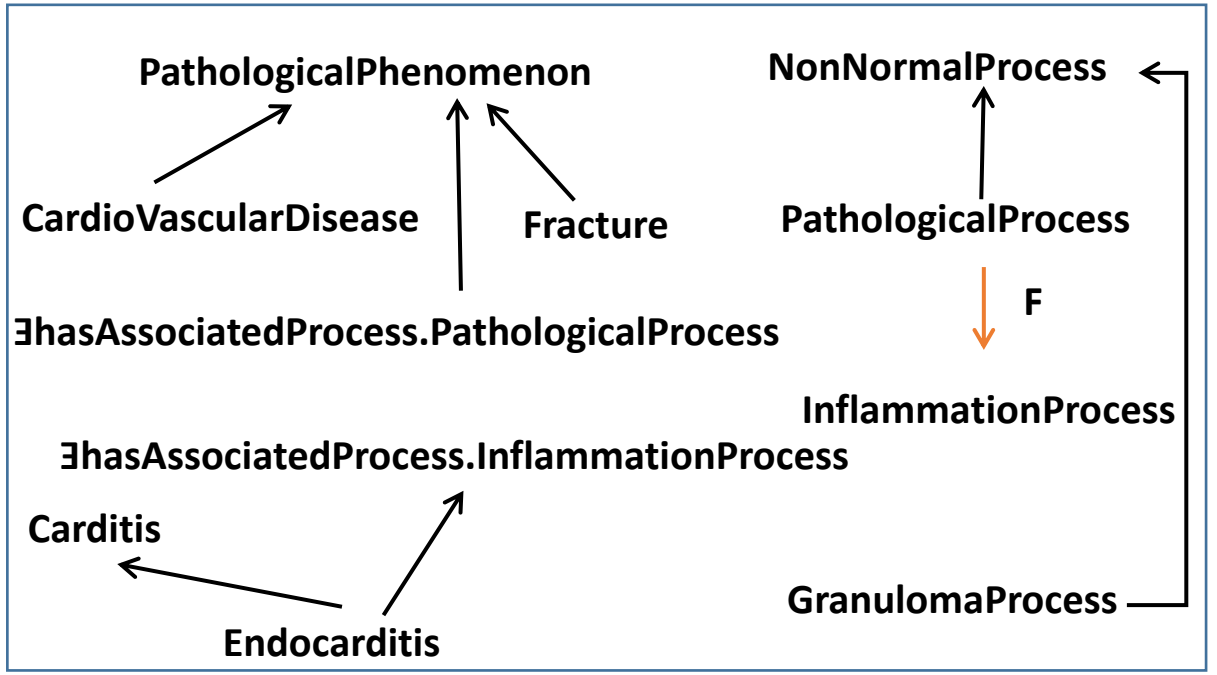
More complete TBoxes

Definition 3. TBox T_1 is less incorrect than *TBox T_2* iff

1) all incorrect knowledge in T_1 can also be derived in T_2.
2) there is incorrect knowledge in T_2 that cannot be derived in T_1.



T1



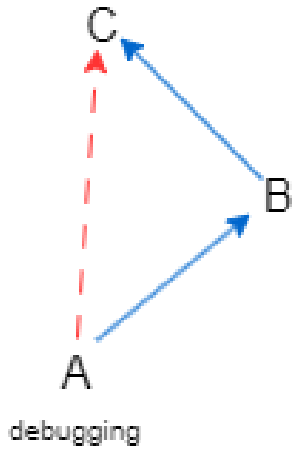
T2

Or(PathologicalProcess ⊑ InflammationProcess) = false

T_1 is less incorrect than T_2

- Or: an oracle, when given an axiom, can answer whether this axiom is correct or wrong in the domain of interest of the ontology.

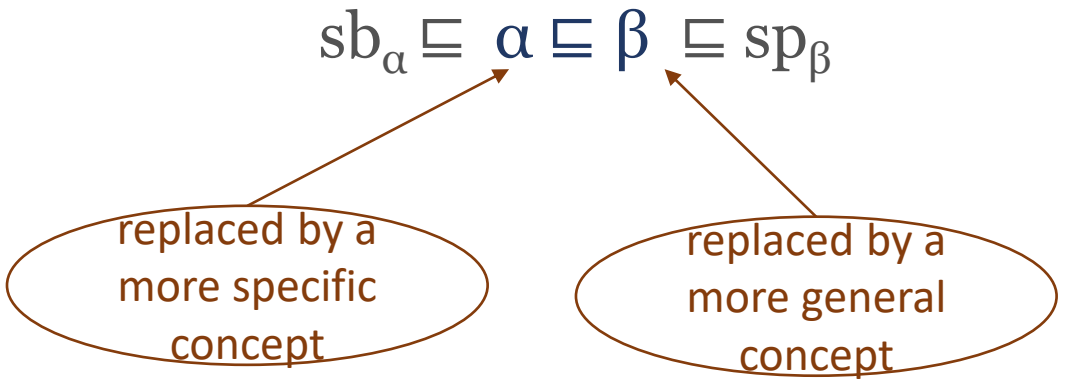
Debugging and Removing- finding wrong axioms

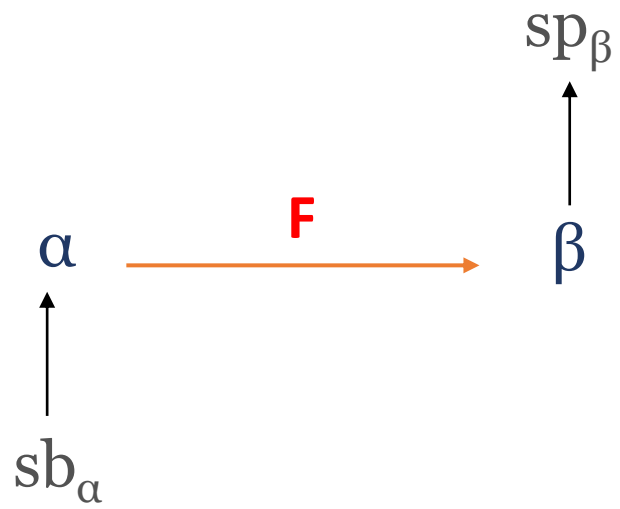


A *Justification* for an axiom α in T is a set of axioms $T' \subseteq T$ such that $T' \models \alpha$ and $\forall\, T'' \subsetneq T' : T'' \not\models \alpha$

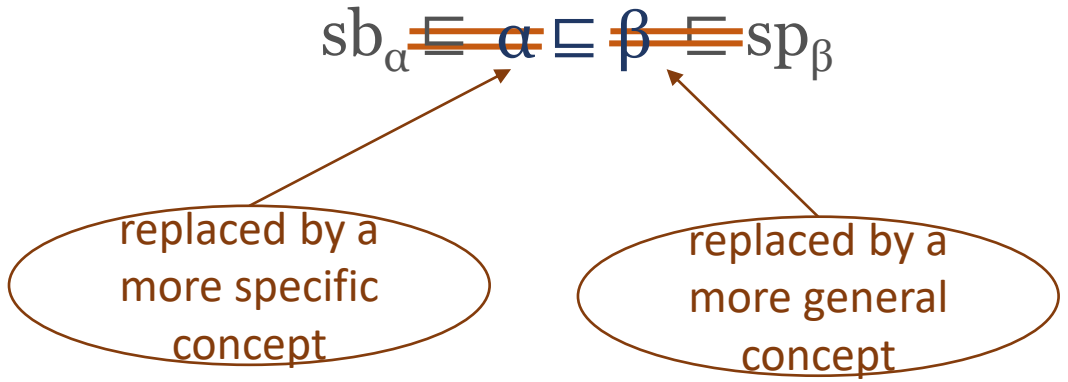
Jus(A ⊑ C)={ A ⊑ B, B ⊑ C}

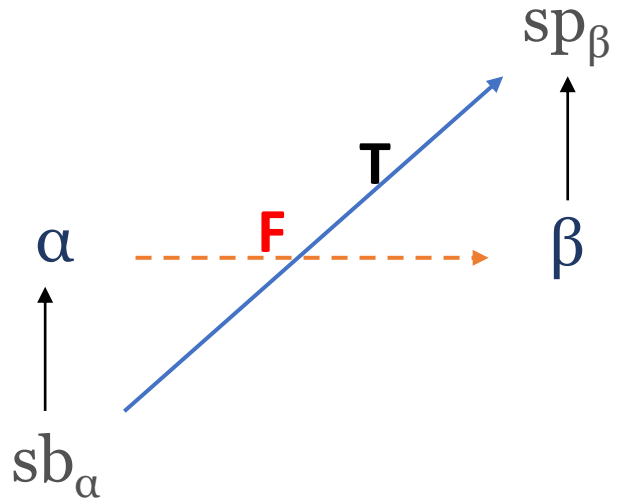
Weakening - finding correct weaker axioms





Weakening - finding correct weaker axioms

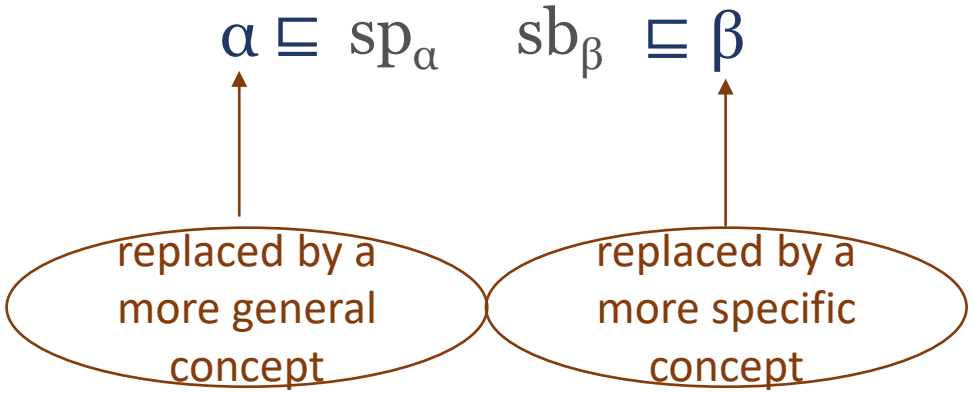


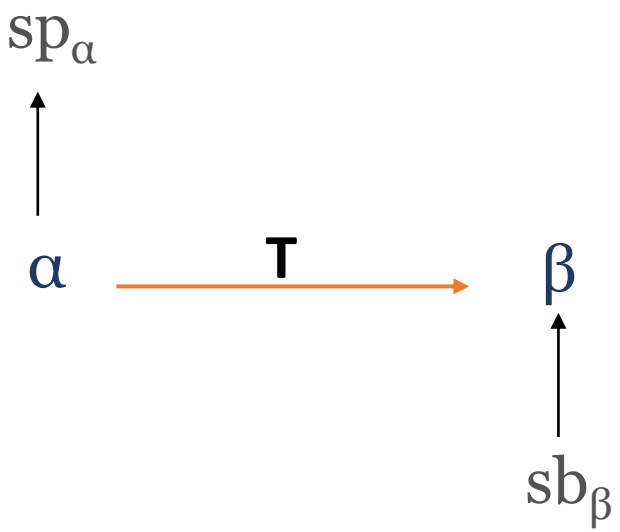


Replace wrong $\alpha \sqsubseteq \beta$

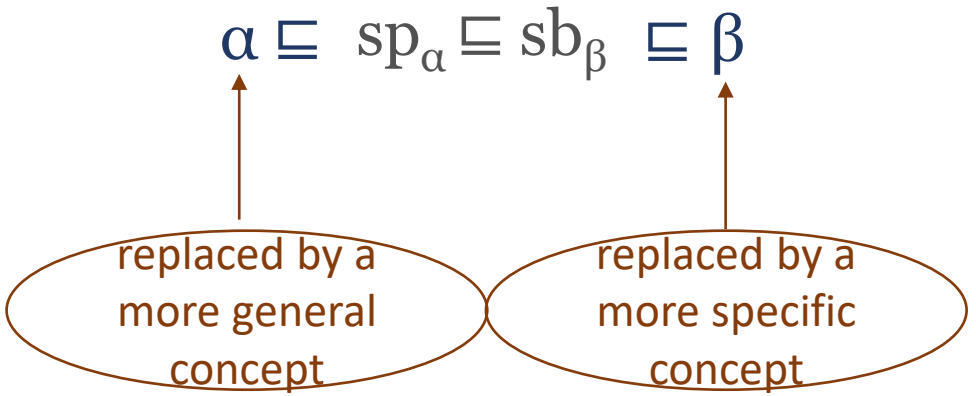
with correct $sb_\alpha \sqsubseteq sp_\beta$

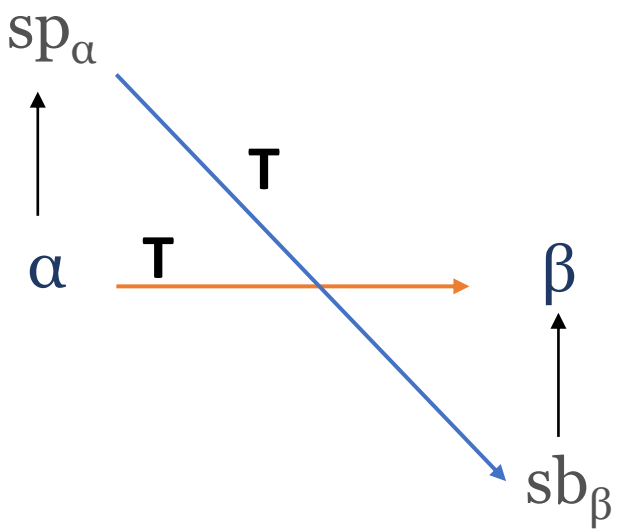
Completing - finding correct stronger axioms





Completing - finding correct stronger axioms





Replace correct $\alpha \sqsubseteq \beta$

with correct $sp_\alpha \sqsubseteq sb_\beta$

Problem Formulation

Problem Formulation

Domain expert

An ontology/ontology network represented by TBox T

Definition 4. (Repair) Let TBox $\mathcal{T} = (\bigcup_{i=1..n} \mathcal{T}_i) \cup (\bigcup_{i,j=1..n,i<j} \mathcal{A}_{ij})$ represent a network of ontologies $\mathcal{O}_i$ represented by TBoxes $\mathcal{T}_i$, and their alignments $\mathcal{A}_{ij}$. Let Or be an oracle that given a TBox axiom returns true or false. Let W be a finite set of TBox axioms in $\mathcal{T}$ such that $\forall\, \psi \in W$: $Or(\psi) =$ false. Then, a repair for Debug-Problem $\mathrm{DP}(\mathcal{T}, Or, W)$ is a tuple (A, D) where A and D are finite sets of TBox axioms such that

(i) $\forall\, \psi \in A$: $Or(\psi) =$ true;
(ii) D is a finite set of *asserted* axioms in $\mathcal{T}$;
(iii) $\forall\, \psi \in D$: $Or(\psi) =$ false;
(iv) $\forall\, \psi \in W$: $(\mathcal{T} \cup A) \setminus D \not\models \psi$.

W is a set of the wrong axioms to remove from the ontology.

A repair (A,D) is a tuple containing two sets:
A: a set of correct axioms to add to the TBox
D: a set of wrong asserted axioms to remove from the Tbox.

When the axioms in D are removed and the axioms in A are added, the wrong axioms in W cannot be derived anymore.

Repairing Approach

Basic framework

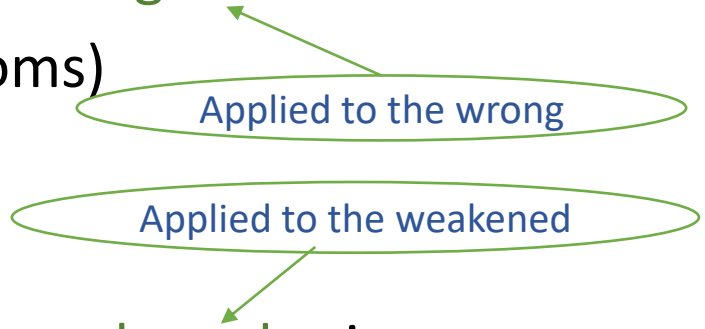
- Debugging: find the wrong asserted axioms in the ontology
- Removing: remove the wrong asserted axioms from the ontology

Remove wrong

- Weakening: Find correct weakened axioms of the wrong axioms

(Mitigate the negative effect of removing wrong axioms)

Dual operations

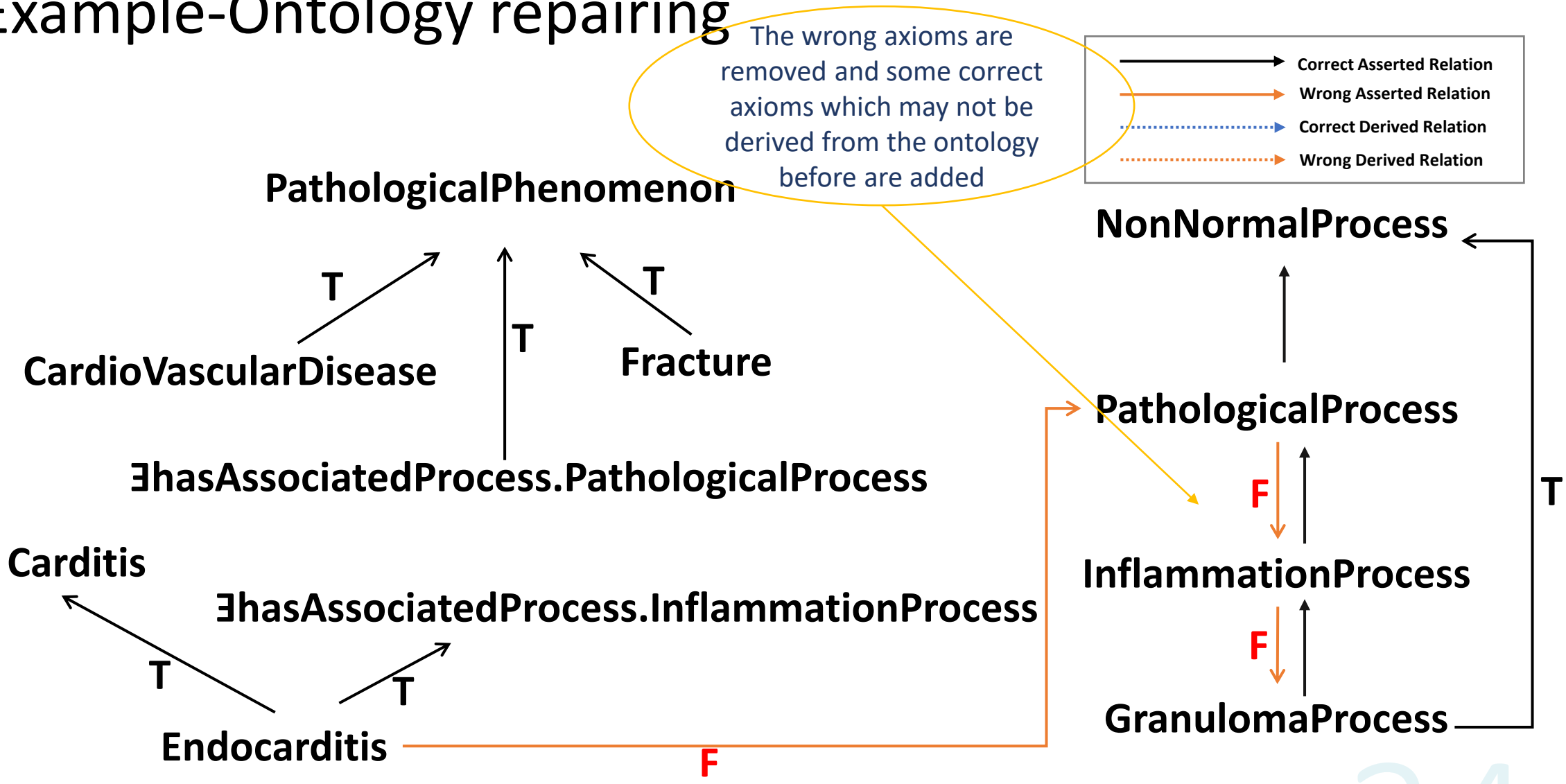


- Completing: Find correct completed axioms of the weakened axioms

(Make the ontology more complete)

Add correct

Example-Ontology repairing



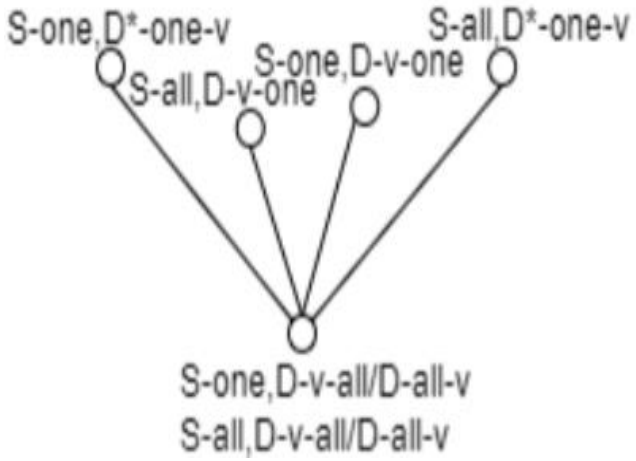
Combinations of basic operations

Choices regarding in which order to perform the operations, and when to update the ontology, i.e.,

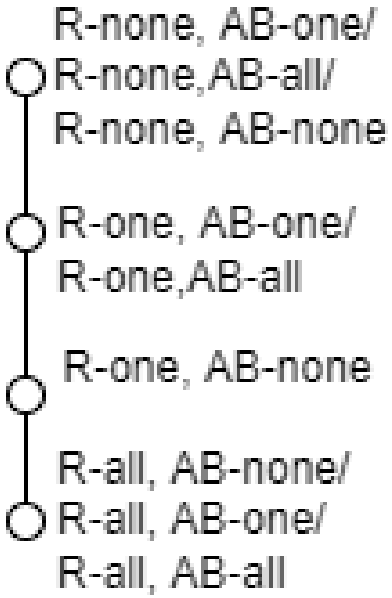
- Removing wrong axioms before/after weakening/completing
- Removing/weakening/completing all at once/one at a time
- The order of removing wrong axioms (which axiom will be removed first, second.....)
- Completing one at a time and adding new correct axioms as soon as they are found /wait until the end

In general, using as much knowledge as possible may lead to more complete ontologies, but also more validation effort.

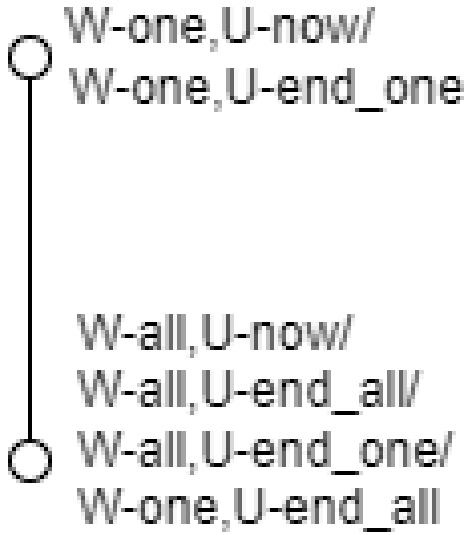
Compare algorithms using the Hasse diagrams



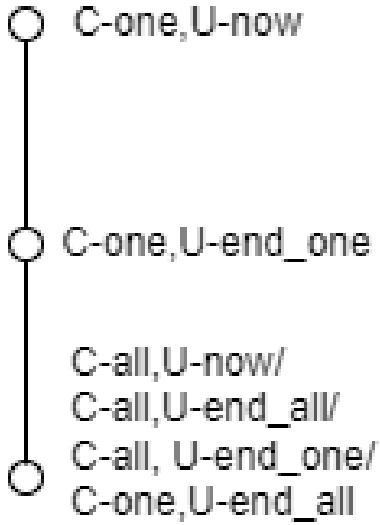
S: Compute the justifications for (one at a time/ all at once)
D: Validate the axioms (all asserted axioms/one hitting set)



R: remove wrong axioms (one at a time/ all at once/none)
AB: add back wrong axioms (none/one/all)



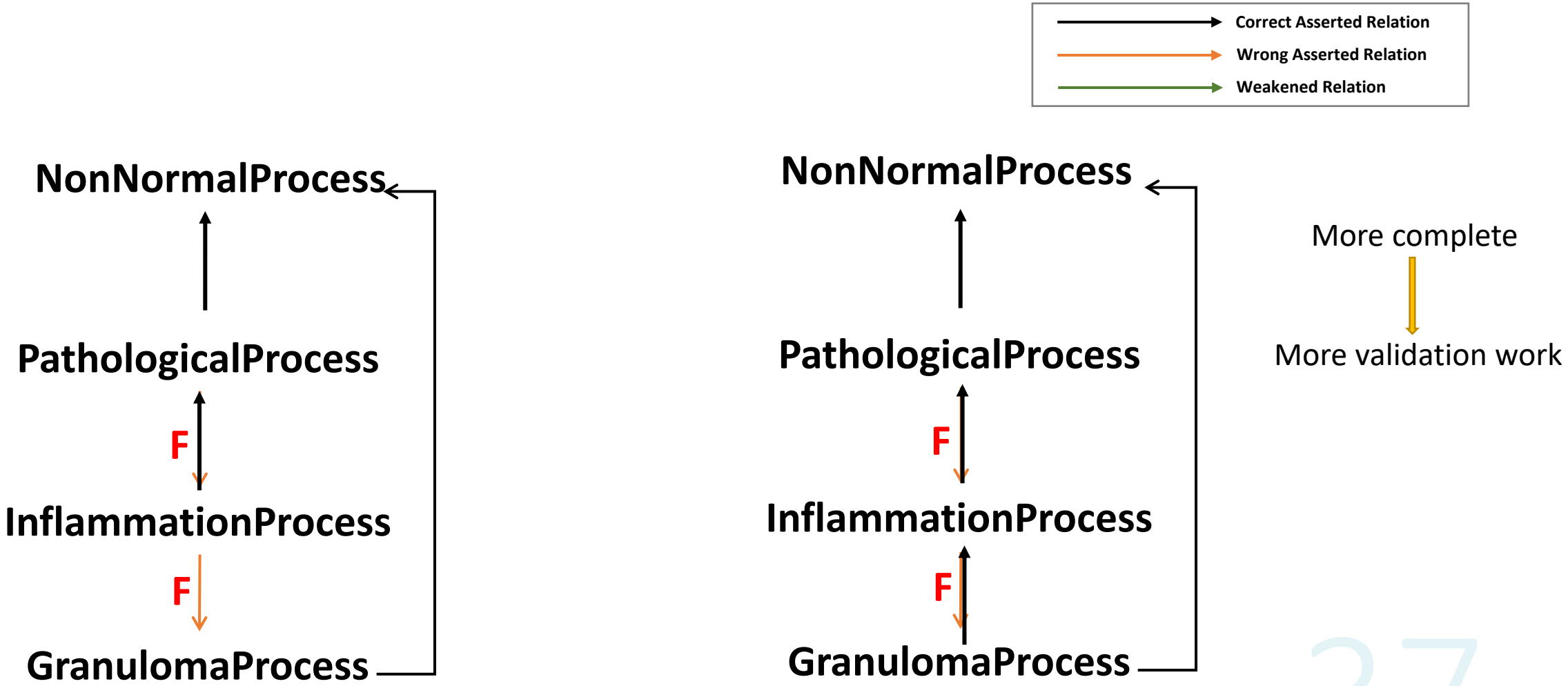
W: weaken (one at a time/all at once)
U: update the ontology



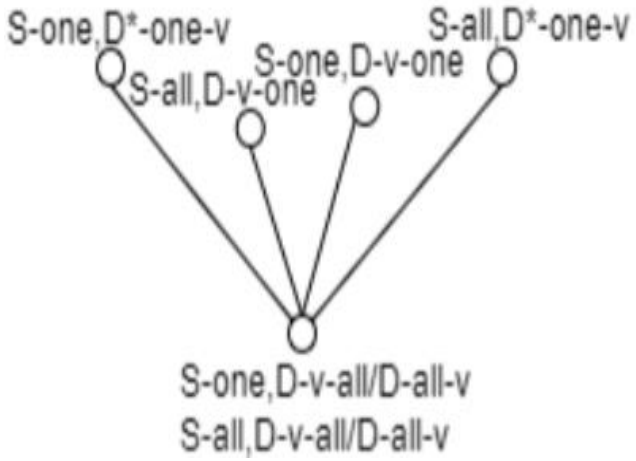
C: complete (one at a time/all at once)
U: update the ontology

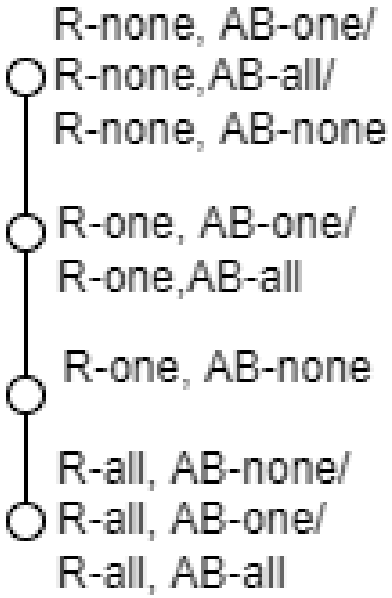
Using operators higher up in the diagrams leads to more complete ontologies and more validation work.

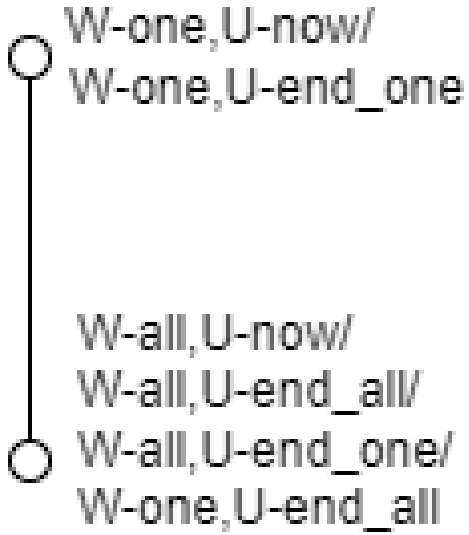
Removing before/after completion(Algorithm C9/C8)

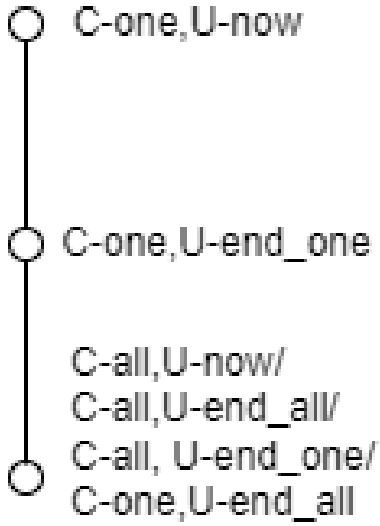


Combinations of removing, weakening and completing



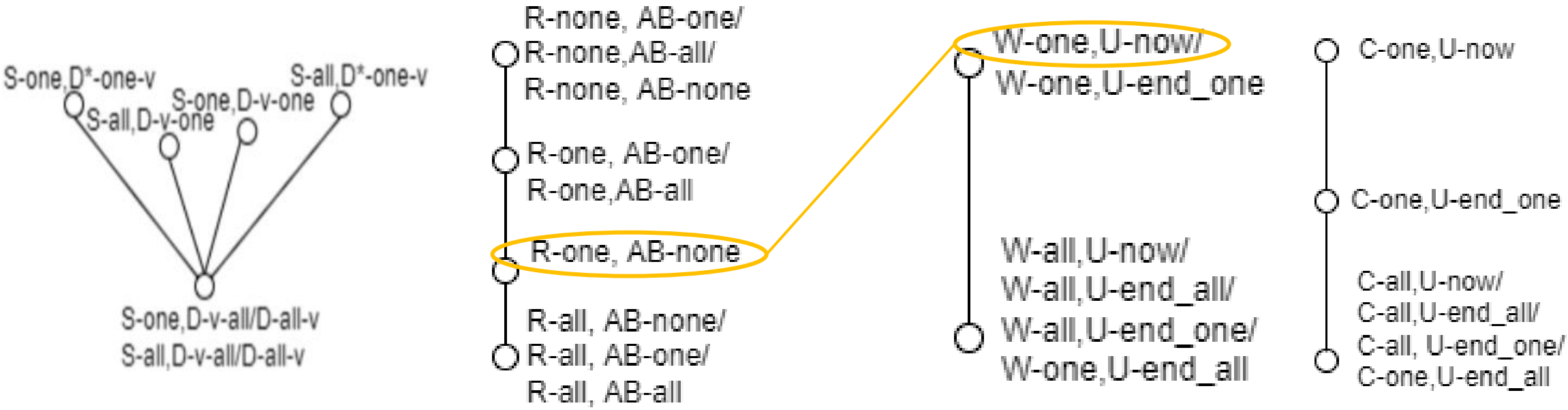




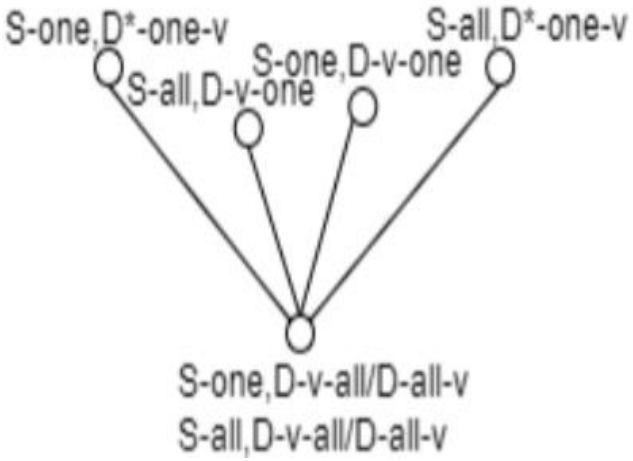


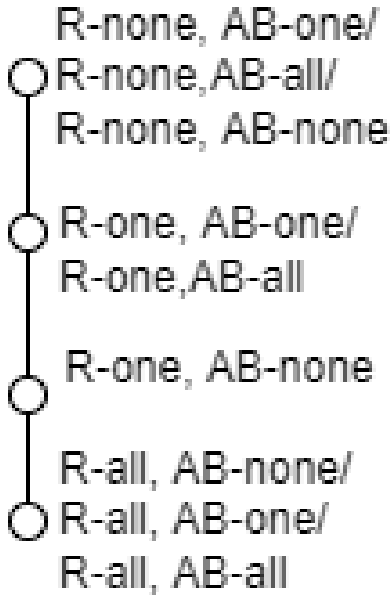
No previous work on combinations

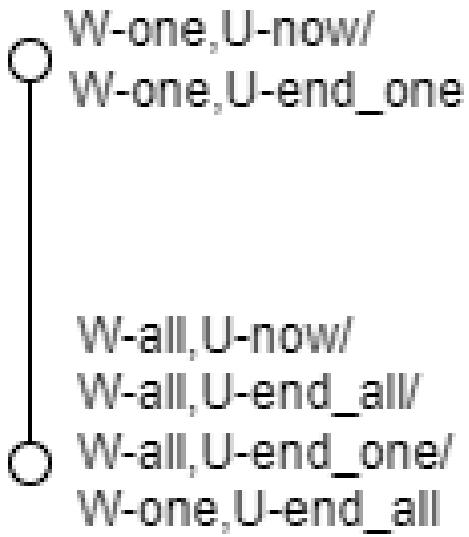
Previous work on removing and weakening

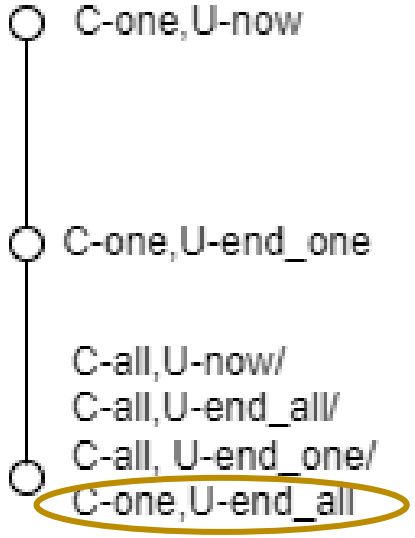


Previous work on completing









Ontology vs. Ontology network

There are also choices regarding the autonomy level of the ontologies and alignments in the ontology network, which reflects the policies of the ontology and alignment owners regarding updating and computing for their ontologies and alignments.

- O (ontology) / M(mappings)
- MO (materialized ontology) / MM(materialized mappings)
- ON (ontology network)

Ontology vs. Ontology network

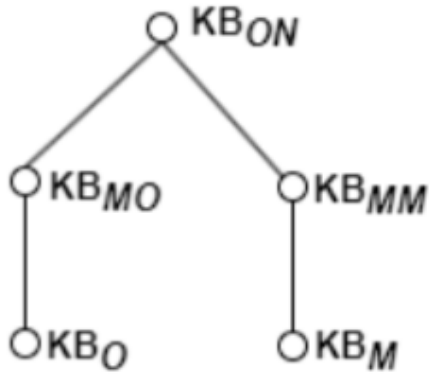
There are also choices regarding the autonomy level of the ontologies and alignments in the ontology network, which reflects the policies of the ontology and alignment owners regarding updating and computing for their ontologies and alignments.

- O (ontology) / M(mappings)
- MO (materialized ontology) / MM(materialized mappings)
- ON (ontology network)

Ontology network

During the repairing process different levels of autonomy can be used at different stages:

More complete



Choices regarding computation

Choices regarding final answer

Different levels of autonomy

➢O (ontology) / M (mappings) --- Ontologies/Mappings are completely autonomous

- The set of wrong axioms W contains only axioms in the ontology (O) /mappings (M).
- Only the axioms within the ontology/mappings can be used for the computation of repairs. ⟶ KB_O / KB_M
- Solutions only include axioms in the ontology/mappings. ⟶ AS_O / AS_M

Different levels of autonomy

➢MO (materialized ontology) / MM (materialized mappings) --- computes derived axioms/mappings for the ontology/alignment, but then acts autonomously

- The set of wrong axioms W contains only axioms in the ontology/mappings.
- Only the axioms within the materialized ontology/mappings can be used for the computation of repairs. ⟶ KB_{MO} / KB_{MM}
- Solutions only include axioms in the ontology/mappings. ⟶ AS_O / AS_M

Different levels of autonomy

➢ON (ontology network) --- ontologies and alignments as integral parts of the network

- The set of wrong axioms W contain ontology axioms and mappings.
- The whole network is used for the computation of repairs. ⟶ KB_{OM}
- Solutions contain axioms and mappings. ⟶ AS_{ON}

Debugging

Asserted axioms in Ontologies O_1 and O_2

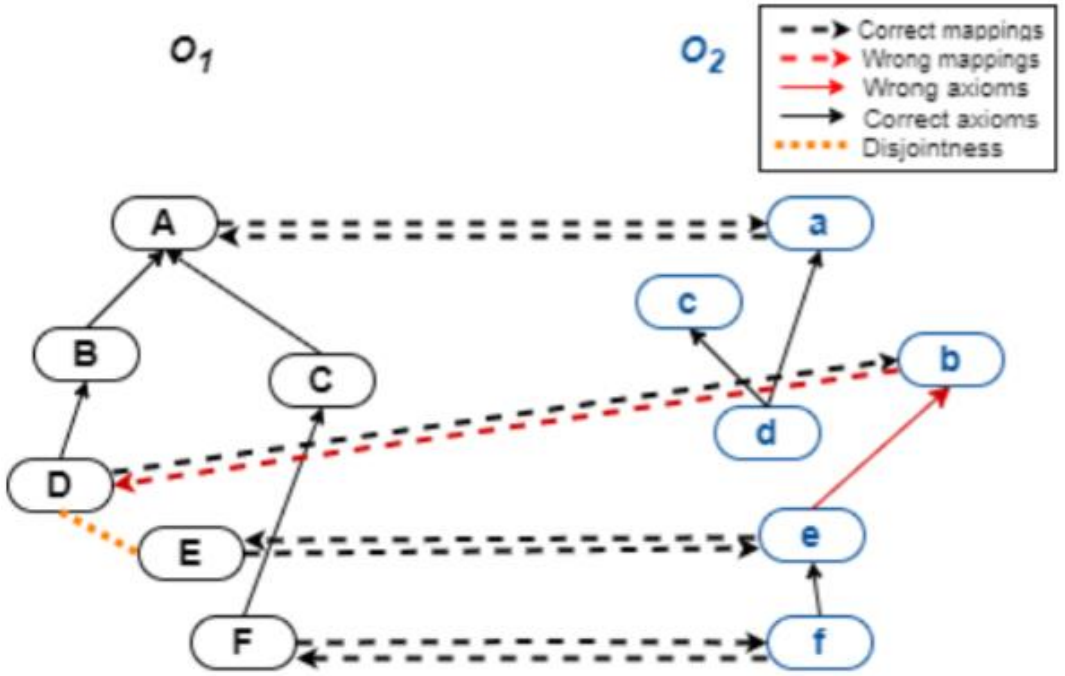
O_1: B ⊑ A, D ⊑B, C ⊑ A, D ⊑ B, F ⊑ C

O_2: d ⊑ c, d ⊑ a, e ⊑ b, f ⊑ e

Mappings between O_1 and O_2

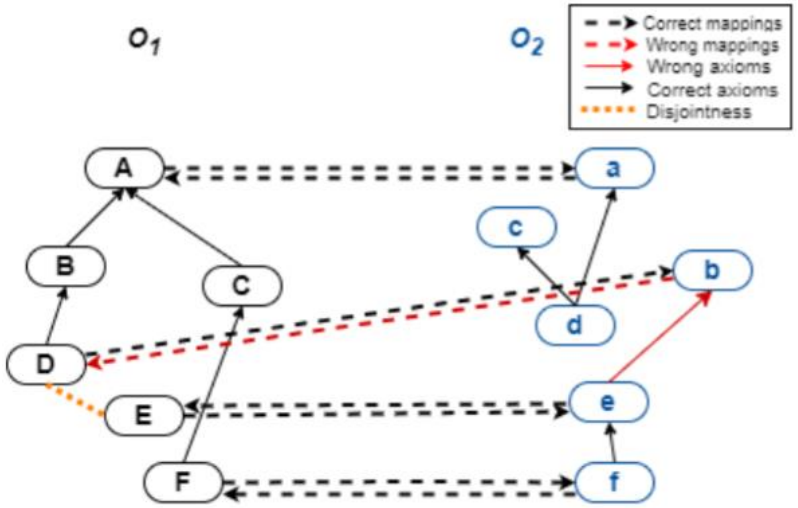
a ⊑ A, D ⊑ b, E ⊑ e, F ⊑ f

A ⊑ a, b ⊑ D, e ⊑ E, f ⊑ F



We have used 'ON' in the debugging step where we validate all axioms in the justifications. This results in a wrong axiom e ⊑ b and a wrong mapping b ⊑ D.

Weakening

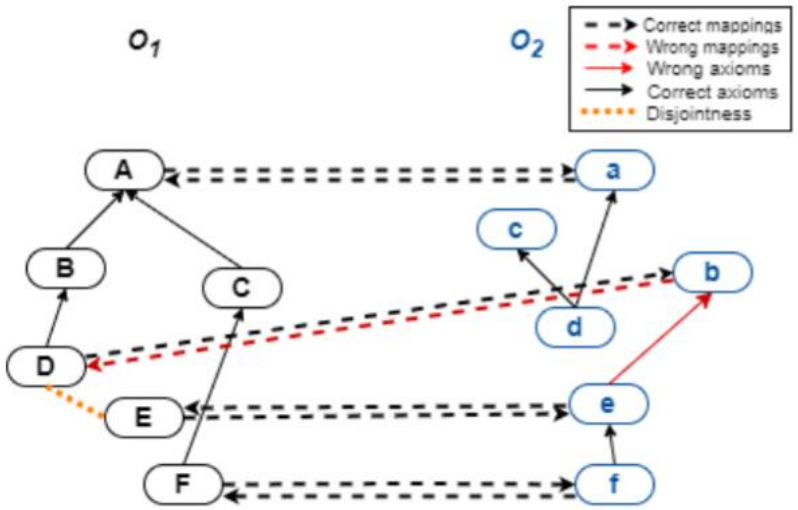


2, 4, 20 candidates for ‘O’, ’MO’ and ’ON’

Background knowledge base	KB_O $e \sqsubseteq b$	KB_{MO} $e \sqsubseteq b$	KB_{ON} $e \sqsubseteq b$		KB_M $b \sqsubseteq D$	KB_{MM} $b \sqsubseteq D$	KB_{ON} $b \sqsubseteq D$
$\lvert Sub(\alpha,\mathcal{T})\rvert$	2	2	4		2	6	6
$\lvert Sup(\beta,\mathcal{T})\rvert$	1	2	5		2	5	5
Weakened	$f \sqsubseteq b$	$f \sqsubseteq b$, $e \sqsubseteq a$	$f \sqsubseteq b$, $e \sqsubseteq a$, $E \sqsubseteq A$			$b \sqsubseteq B$	$b \sqsubseteq B$, $e \sqsubseteq a$

’ON’ leads to the most complete network

Completing



Background knowledge base	KB_O e ⊑ b	KB_{MO} e ⊑ b	KB_{ON} e ⊑ b		KB_M b ⊑ D	KB_{MM} b ⊑ D	KB_{ON} b ⊑ D
$\|Sub(\alpha,\mathcal{T})\|$ $\|Sup(\beta,\mathcal{T})\|$	2 1	2 2	4 5		2 2	6 5	6 5
Weakened	f ⊑ b	f ⊑ b, e ⊑ a	f ⊑ b, e ⊑ a, E ⊑ A			b ⊑ B	b ⊑ B, e ⊑ a
$\|Sup(\alpha,\mathcal{T})\|$ $\|Sub(\beta,\mathcal{T})\|$	2 1	4 3 3 5	10 7 7 6 11 11			5 7	5 7 7 11
Completed	f ⊑ b	f ⊑ b, e ⊑ d	f ⊑ b, e ⊑ d, E ⊑ C			b ⊑ B, B ⊑ b	b ⊑ B, B ⊑ b

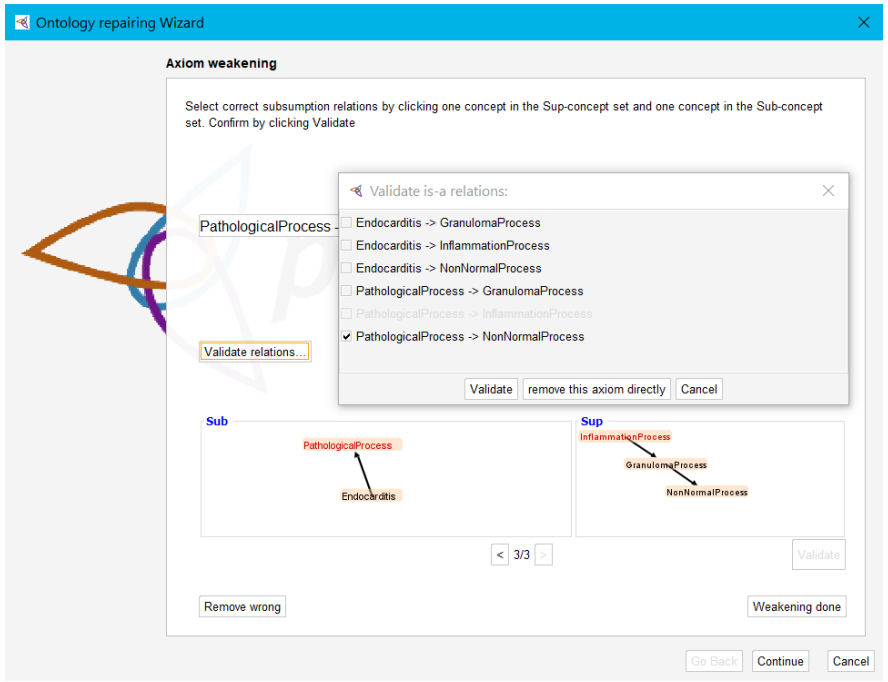
’ON’ leads to the most complete network

The network is repaired by removing the wrong axioms and adding the completed axioms.

RepOSE-A tool for repairing wrong axioms in ontologies/ontology networks.

Two versions:

- A Protégé plugin for ontology repairing based on Algorithm C9 in [1]
- The $\mathcal{EL}_\bot$ version of the RepOSE system for both ontology repairing and network repairing.



Conclusion and Future work

Conclusion

- We proposed an interactive approach to mitigate the negative effects of removing unwanted axioms from the ontology/ontology network.

- We defined combination operators that reflect choices in how and when to use the basic operations, as well as choices regarding the autonomy level of the ontologies and alignments in the ontology network.

Conclusion

- We showed the trade-offs for different combination strategies involving correctness, completeness and validation work.

- By using our framework together with existing algorithms for debugging weakening and completing, we essentially provide a blueprint for extending previous work and systems.

Future work

- TBox repairing -> ABox repairing
- Ontology -> Knowledge graph

References

[1] Ying Li, Patrick Lambrix (2023) Repairing $\mathcal{EL}$ Ontologies Using Weakening and Completing, The Semantic Web. ESWC 2023., p. 298-315
[2]Ying Li, Patrick Lambrix (2023) A System for Repairing EL Ontologies Using Weakening and Completing, The Semantic Web: ESWC 2023 Satellite Events, Hersonissos, Crete, Greece.
[3] Ying Li, Patrick Lambrix (2023) Repairing $\mathcal{EL}_\bot$ Ontologies using Debugging, Weakening and Completing (Extended abstract), 36th International Workshop on Description Logics, Rhodes, Greece, 2023.
[4] Ying Li, Patrick Lambrix (2023) Repairing Networks of Ontologies using Weakening and Completing, 18th International Workshop on Ontology Matching, Athens, Greece.
[5] Ying Li, Patrick Lambrix (2024) Repairing $\mathcal{EL}_\bot$ Ontologies Using Debugging, Weakening and Completing. (Submitted)
[6] Ying Li, Patrick Lambrix (2024) Repairing Networks of $\mathcal{EL}_\bot$ Ontologies Using Weakening and Completing. (Submitted)

LINKÖPING UNIVERSITY