Repairing $\mathcal{EL}$ ontologies using weakening and completing

Ying Li, Patrick Lambrix

LINKÖPING UNIVERSITY

ESWC23

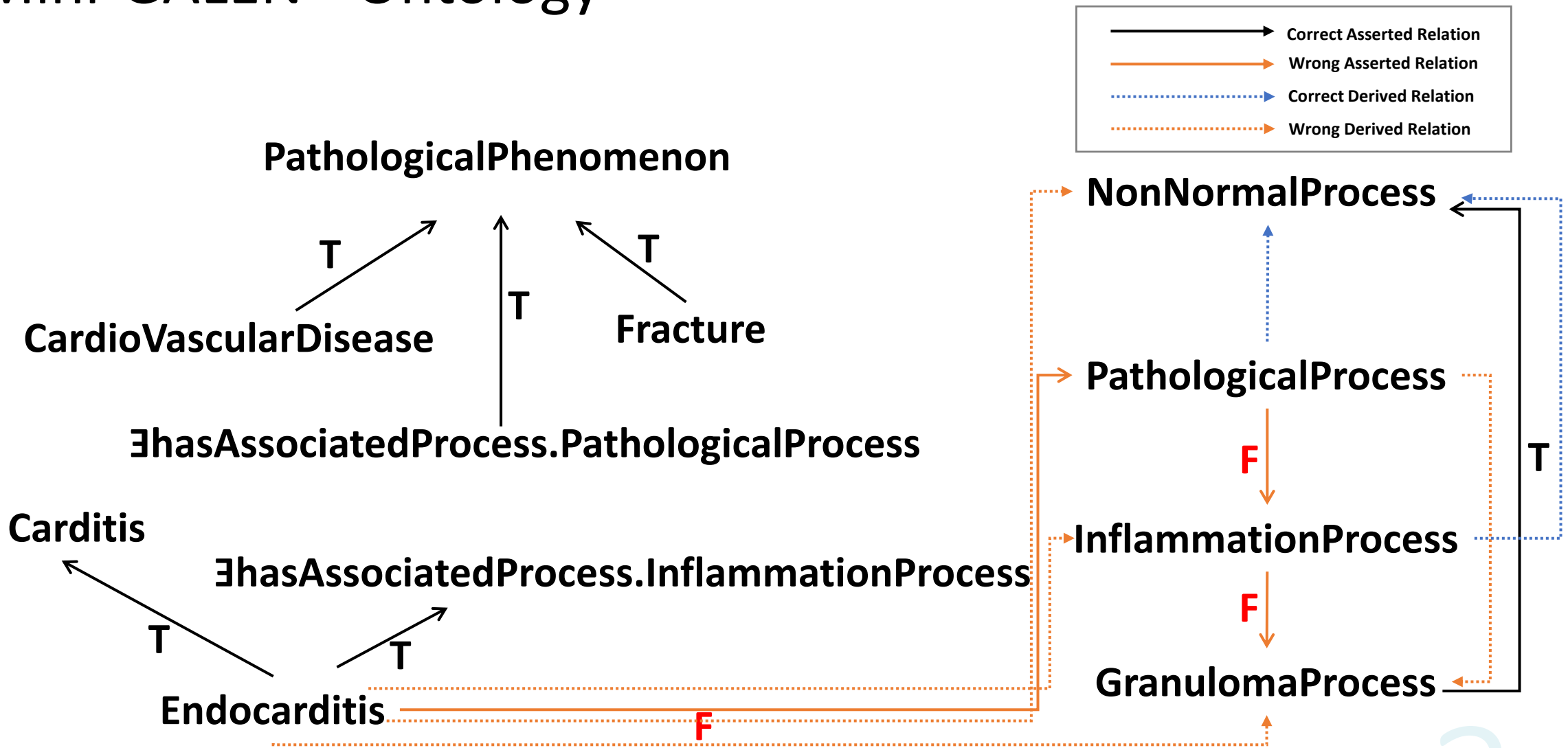
Background

The quality of ontologies in terms of correctness and completeness is crucial for developing high-quality ontology-based applications.

An ontology which contains wrong knowledge may lead to logical problems (i.e., incoherence or inconsistency) or statements that are not correct in the domain of the ontology (modeling defects).

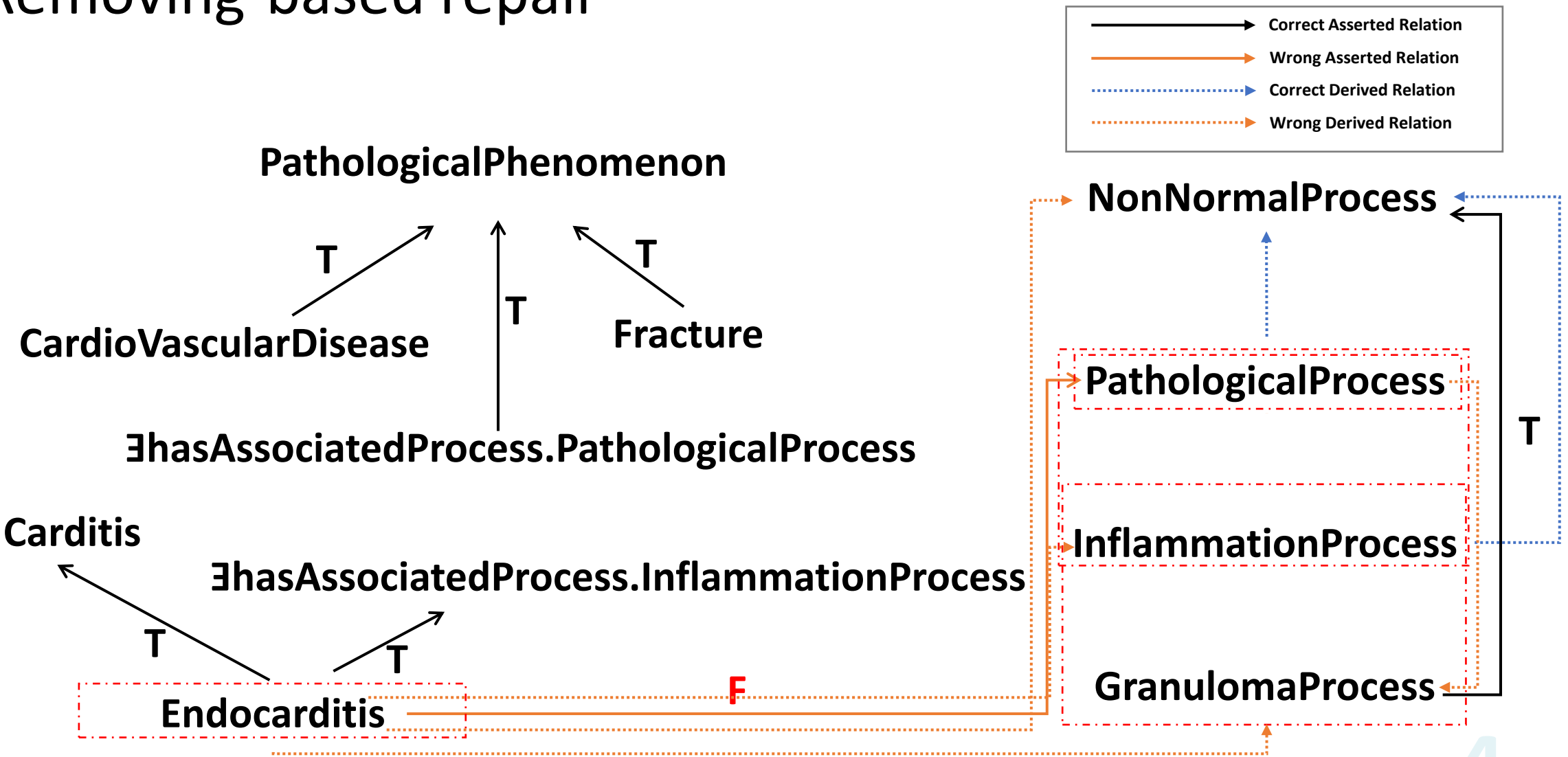
When repairing ontologies by removing unwanted axioms, some correct axioms may not be derivable anymore.

Mini-GALEN[1] Ontology



[1] an example inspired by the GALEN ontology

Removing-based repair



Remove all the wrong axioms from the ontology

Motivation

Preserve correct axioms as much as possible when repairing ontologies to alleviate the negative effects of removing wrong axioms.

We propose an interactive approach combining removing, weakening and completing.

Preliminaries

EL ontologies

- N_C: atomic concepts
- N_R: atomic roles
- A TBox is a finite set of axioms which in EL are general concept inclusions (GCIs).

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
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	$P \sqsubseteq Q$	$P^{\mathcal{I}} \subseteq Q^{\mathcal{I}}$

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

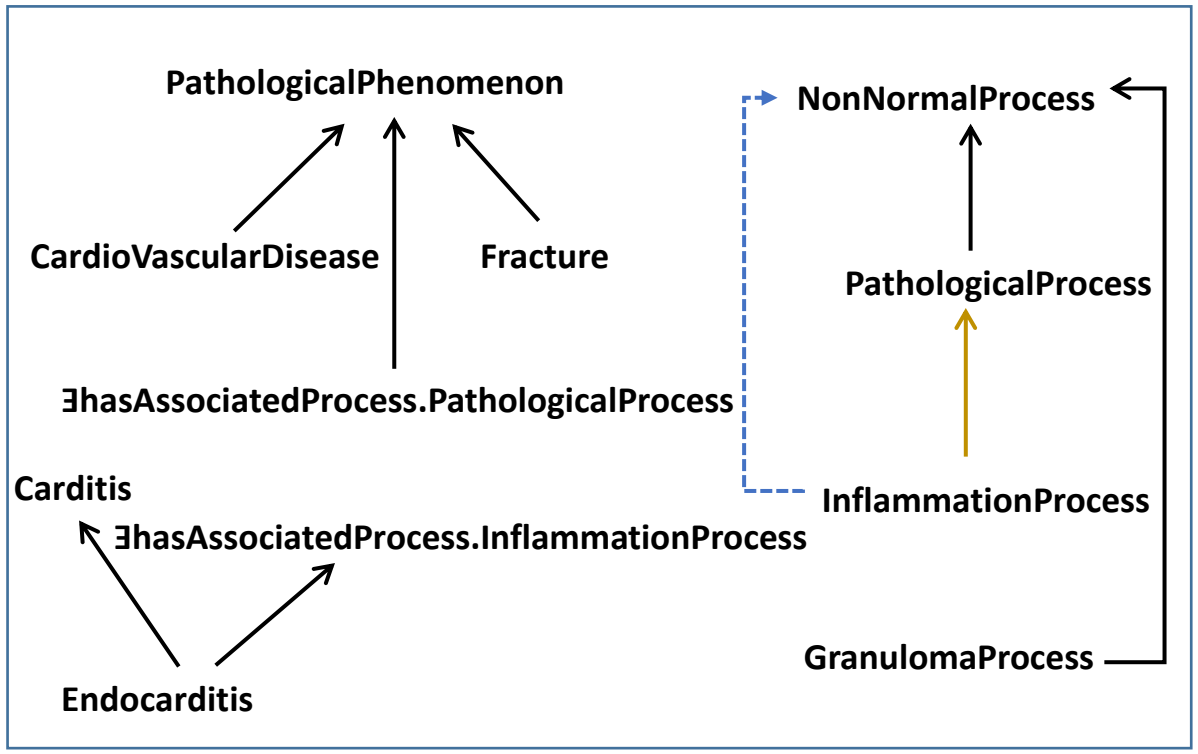
Preference relation

More complete – ontologies

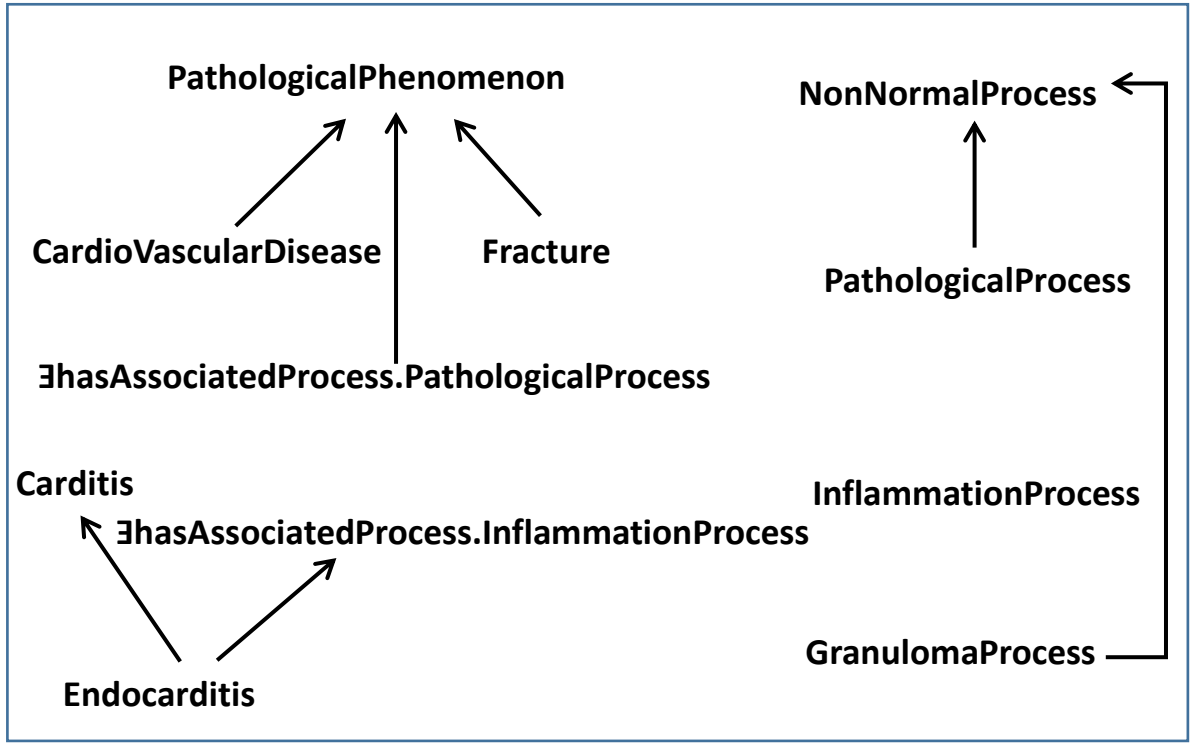
O_1 is more complete than O_2 iff

1) all correct knowledge in O_2 can also be derived in O_1.
2) there is correct knowledge in O_1 that cannot be derived in O_2.

Preference relation



O1



O2

Or(InflammationProcess ⊆ PathologicalProcess) = true
O_1 is more complete than O_2

Problem formulation

An ontology represented by TBox T

Domain expert

Definition 1. (Repair) Let $\mathcal{T}$ be a TBox. 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 finite set of TBox axioms A such that
(i) $\forall\ \psi \in A$: $Or(\psi) =$ true;
(ii) $\forall\ \psi \in W$: $(\mathcal{T} \cup A) \setminus W \not\models \psi$.

A repair A is a set of correct axioms that when added to the TBox where the axioms in W are removed will not allow deriving the axioms in W.

W is the set of the wrong asserted axioms that we want to remove from the ontology.

Repairing – basic operations

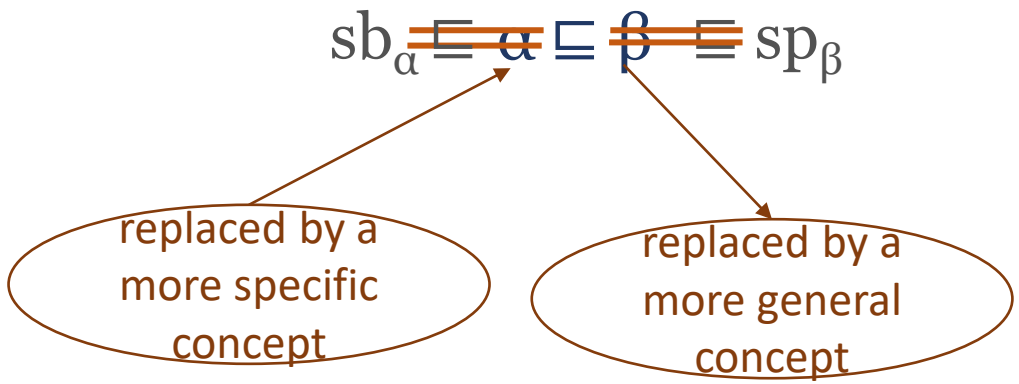
- Removing: removing the given asserted wrong axioms from the ontology.

- Weakening: finding correct axioms that are weaker than a given axiom; applied to the wrong axioms.

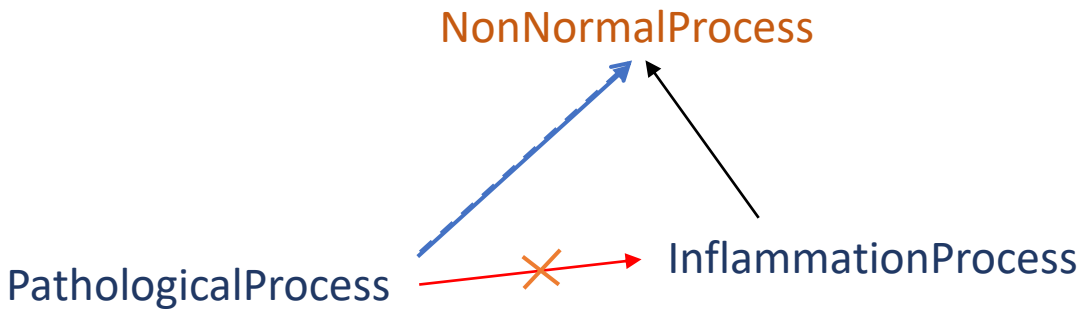
Dual operations where the former finds weaker axioms and the latter stronger axioms

- Completing: finding additional correct axioms that are not derivable from the ontology yet and that would make a given axiom derivable; applied to the weakened axioms.

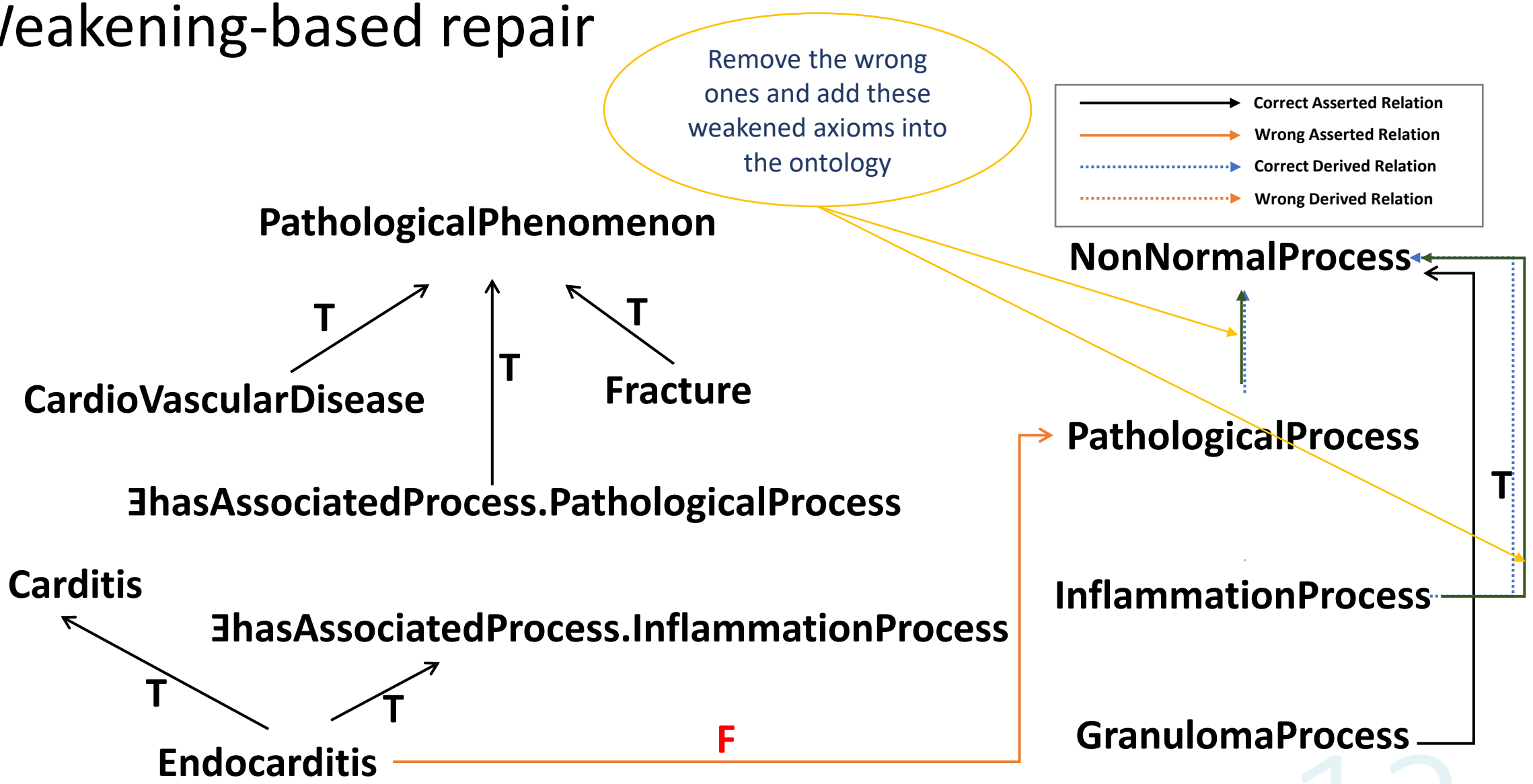
Weakening - finding correct weaker axioms



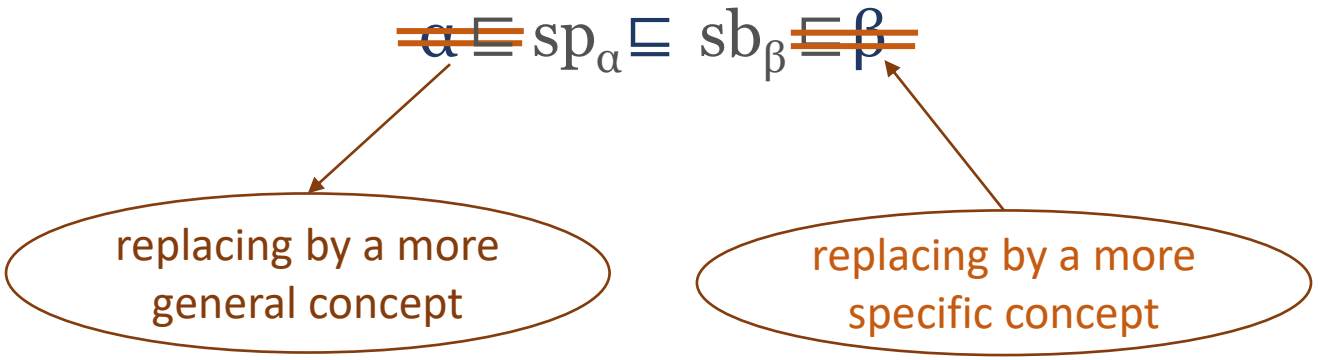
For the repairing, a wrong axiom $\alpha \sqsubseteq \beta$ can be replaced by a correct weaker axiom $sb_\alpha \sqsubseteq sp_\beta$, thereby mitigating the effect of removing the wrong axiom.



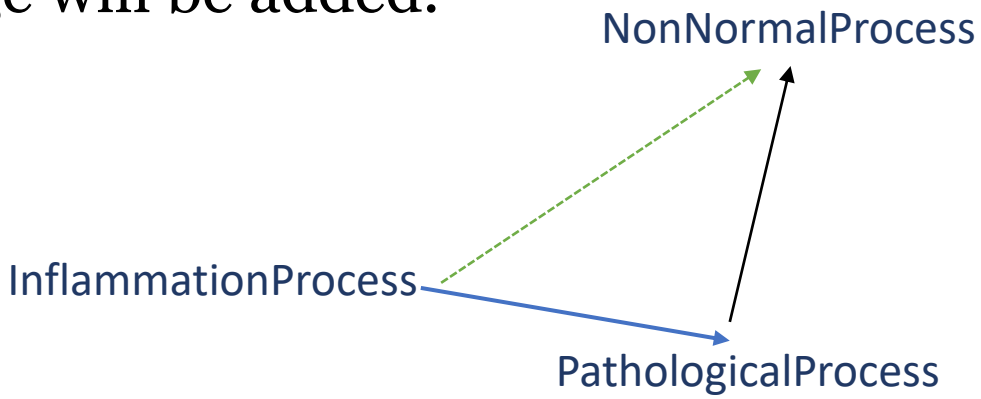
Weakening-based repair



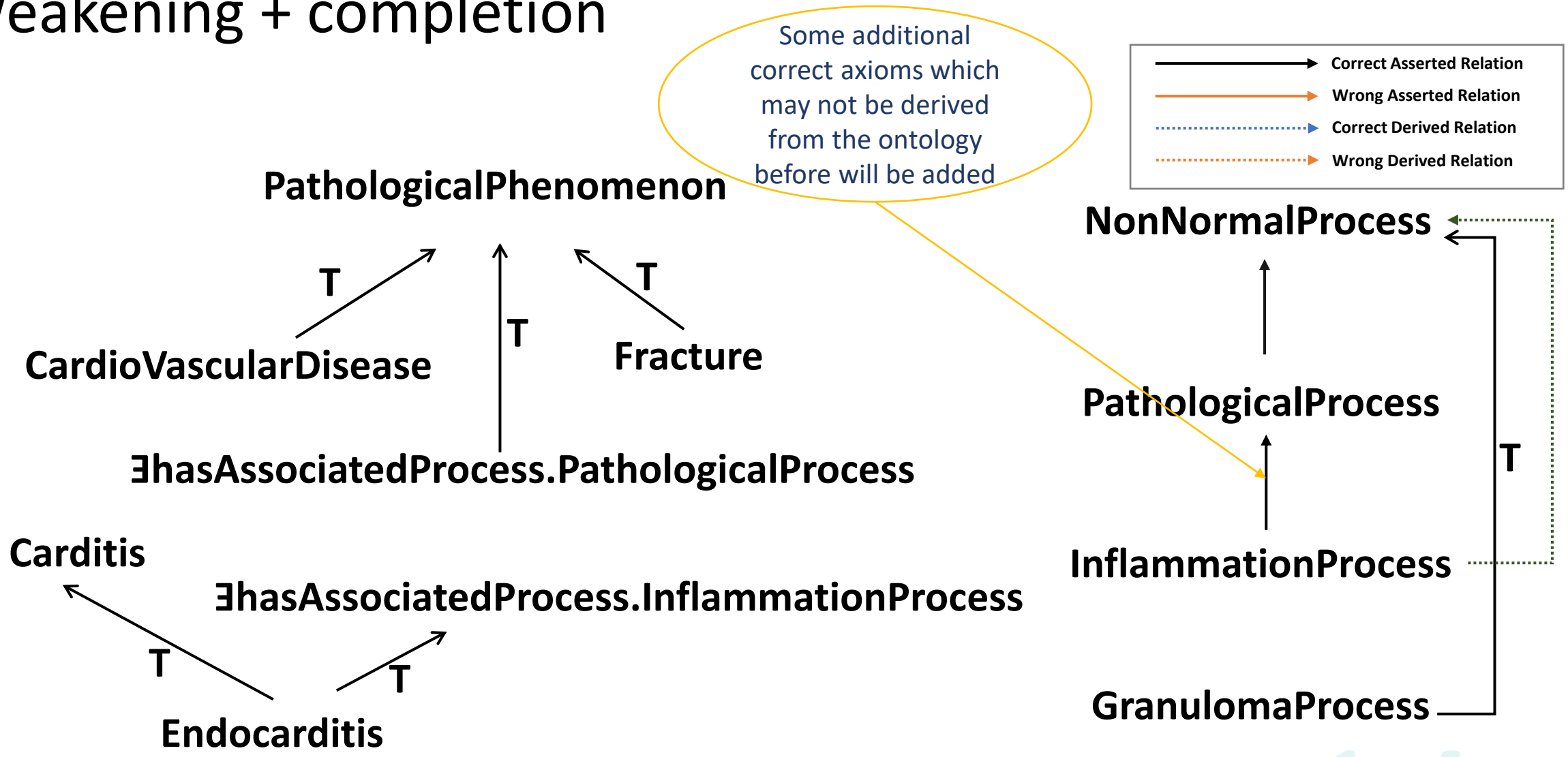
Completion - finding additional correct axioms



When adding $sp_\alpha \sqsubseteq sb_\beta$ into the ontology, then $\alpha \sqsubseteq \beta$ can be derivable and more correct knowledge will be added.



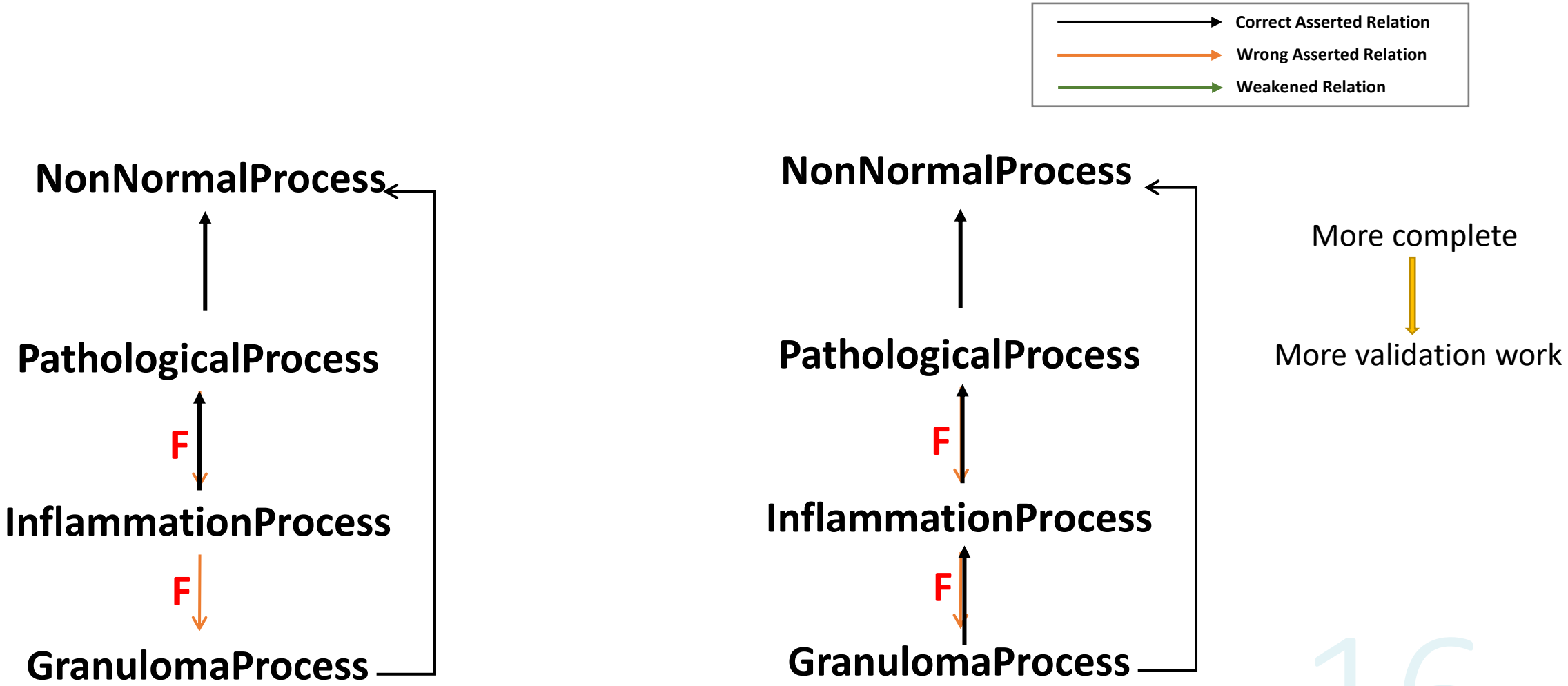
Weakening + completion



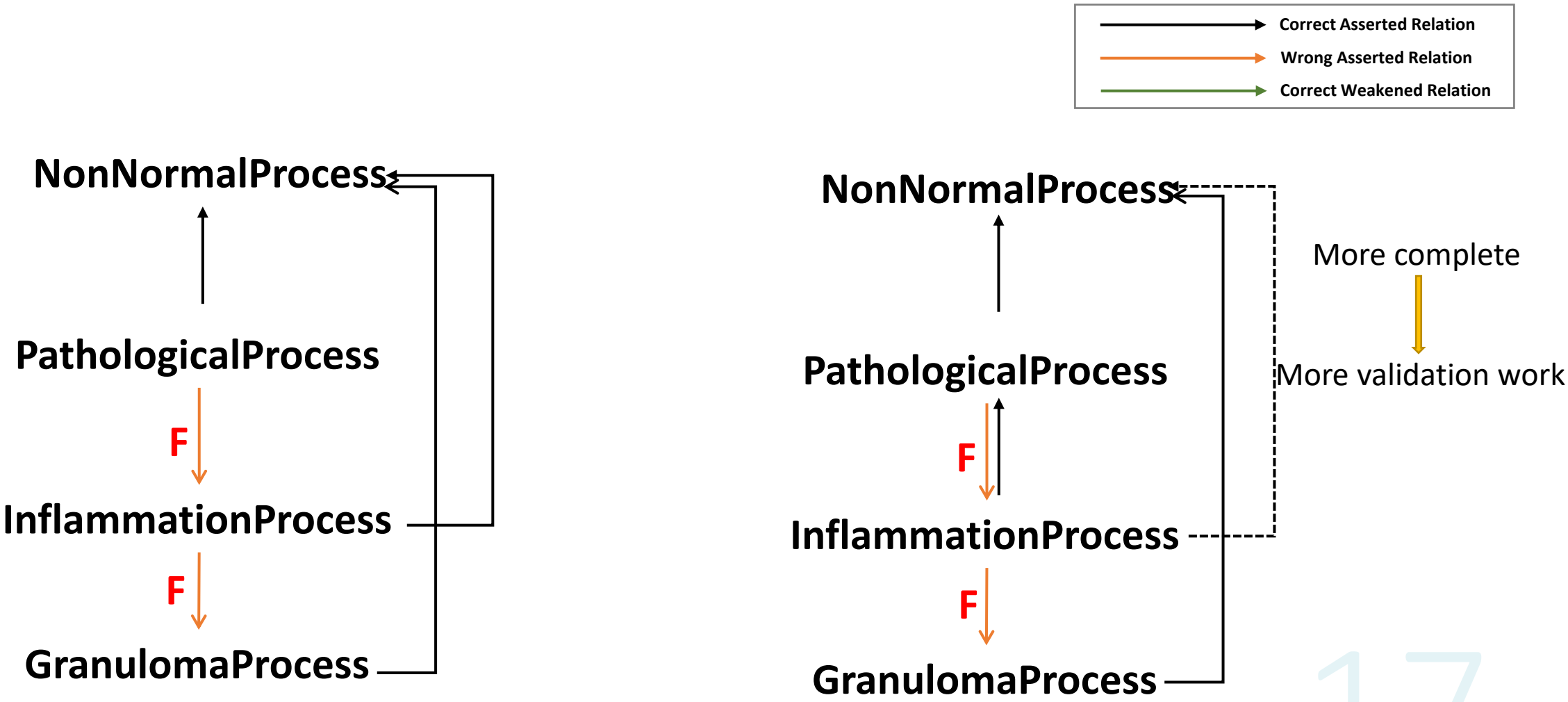
Combinations (choices, order)

- Removing before / after weakening / completing
- Removing/weakening/completing all at once / one at a time
- The order of removing 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

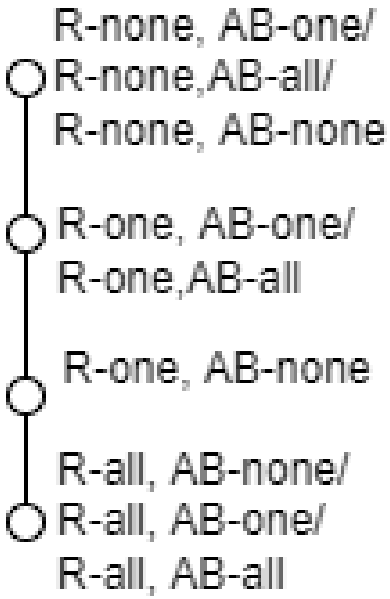
Removing before/after completion(Algorithm C9/C8)



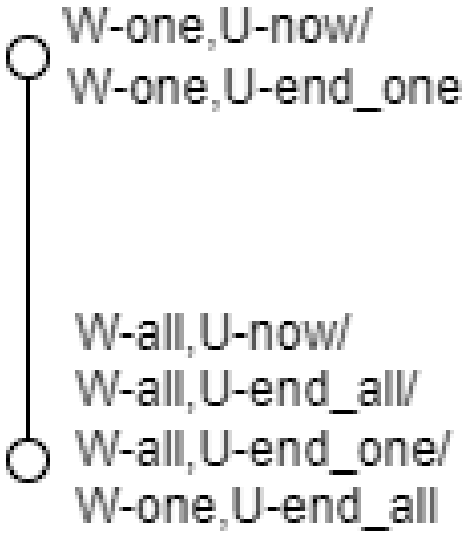
Update all/one during completion(Algorithm C9/C10)



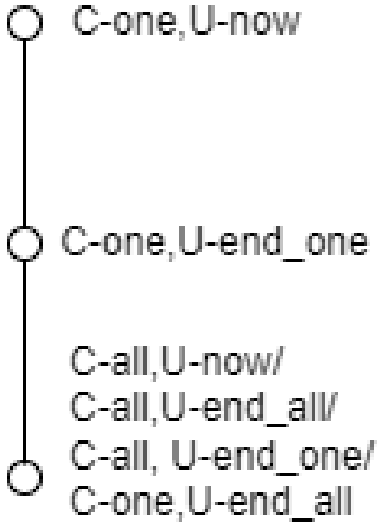
Compare algorithms using the Hasse diagrams



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



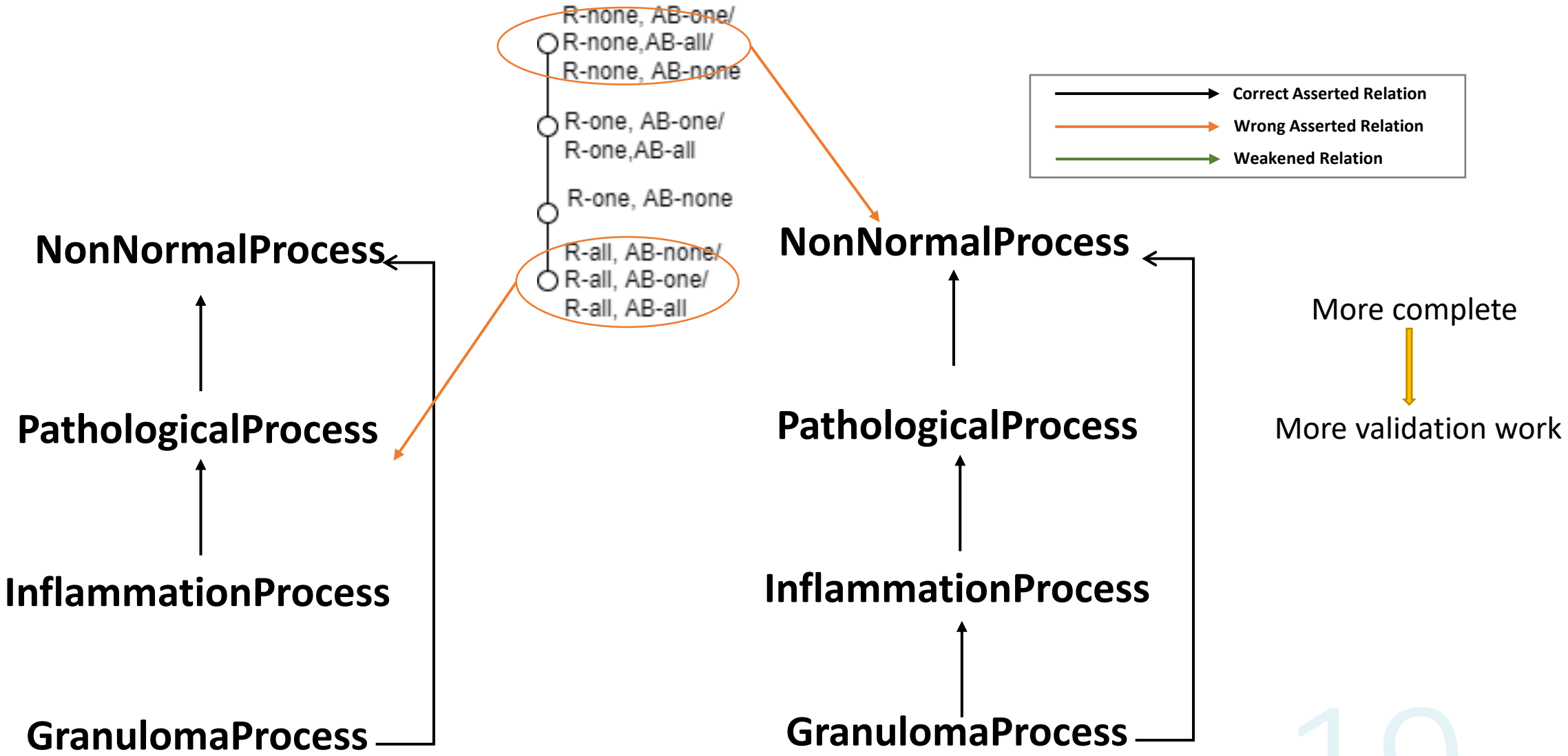
W: weakening (one at a time/all at once)
U: update the ontology



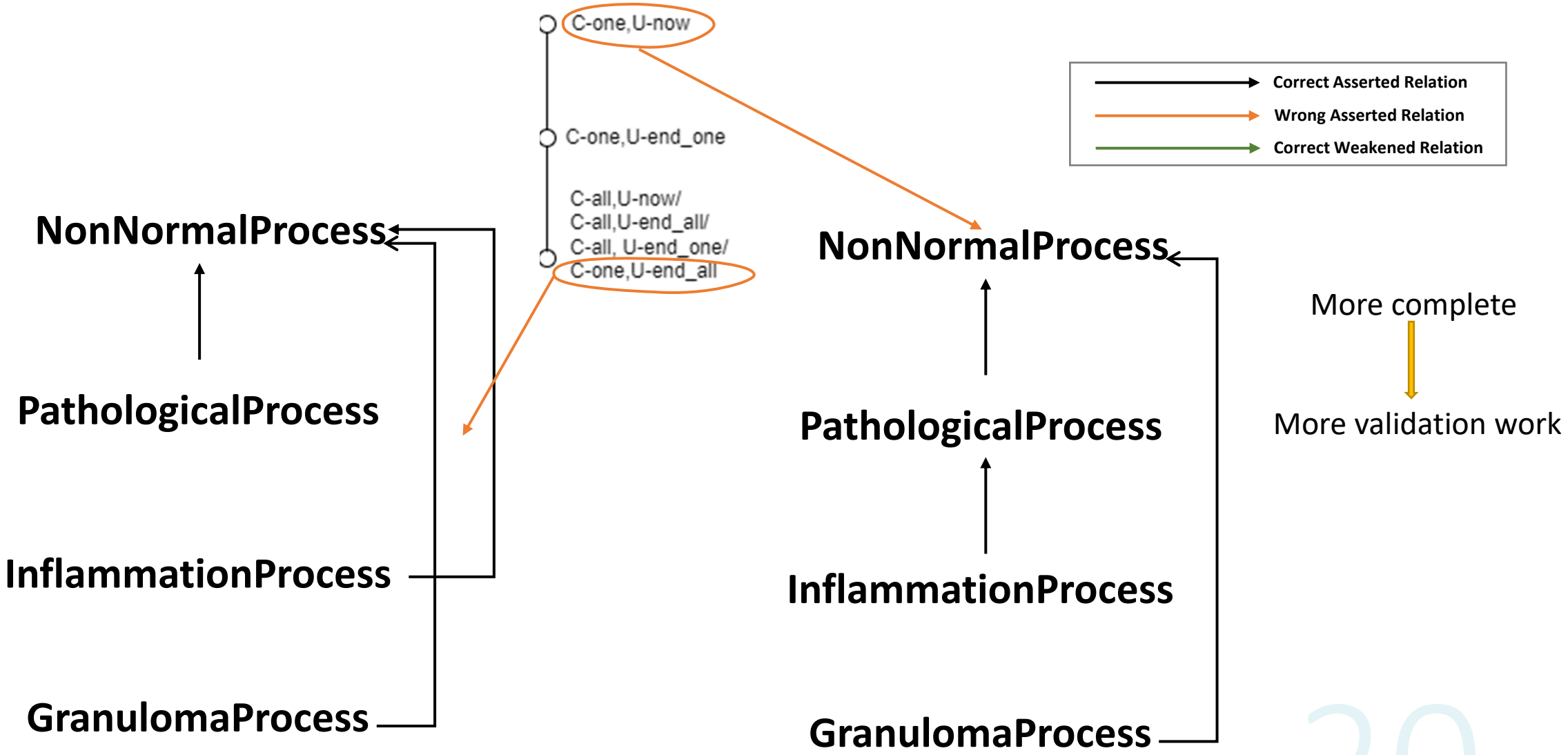
C: completing (one at a time/all at once)
U: update the ontology

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

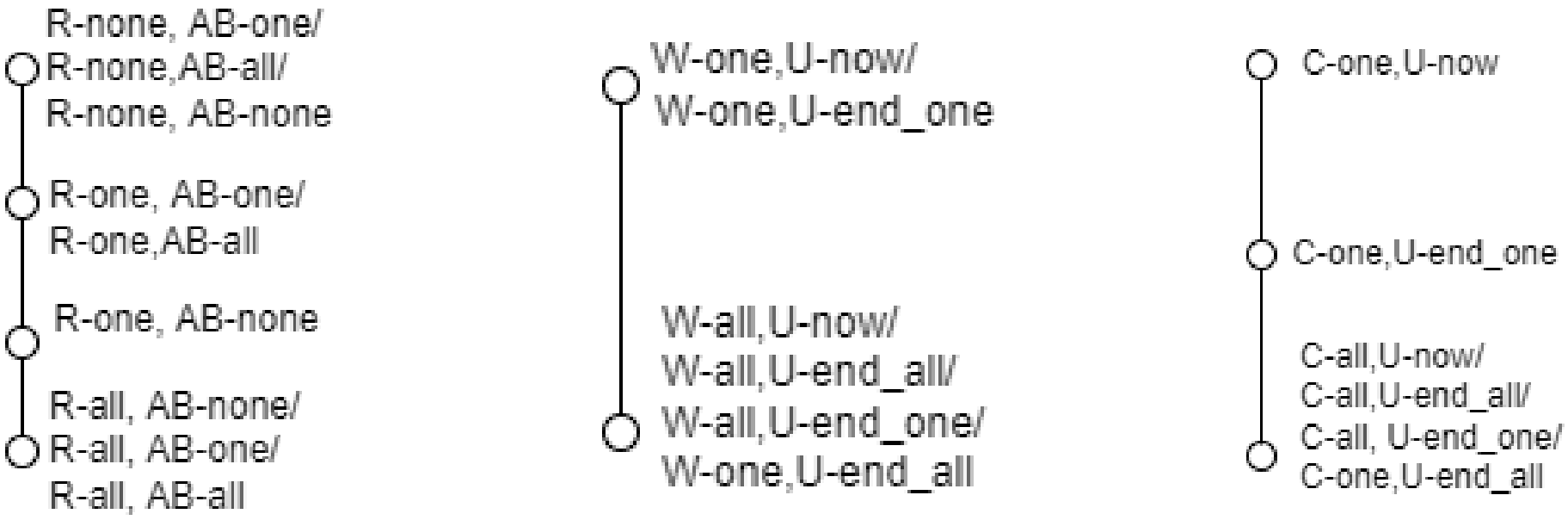
Removing before/after completion(Algorithm C9/C8)



Update all/one during completion(Algorithm C9/C10)

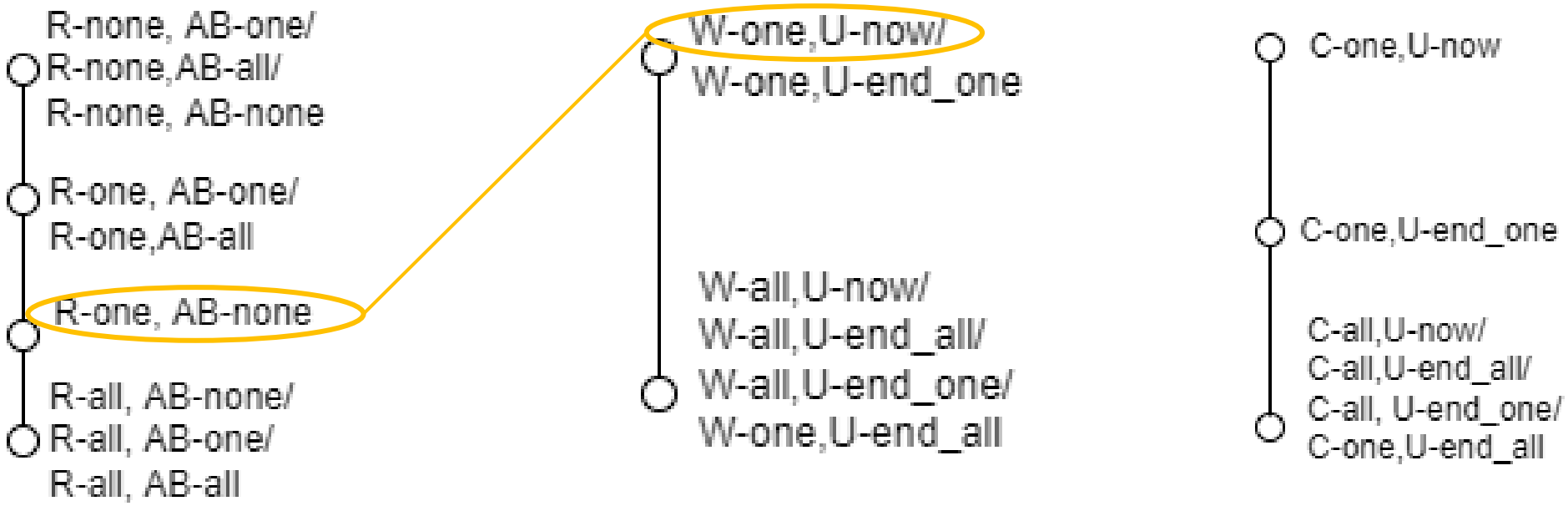


Combinations of removing, weakening and completing

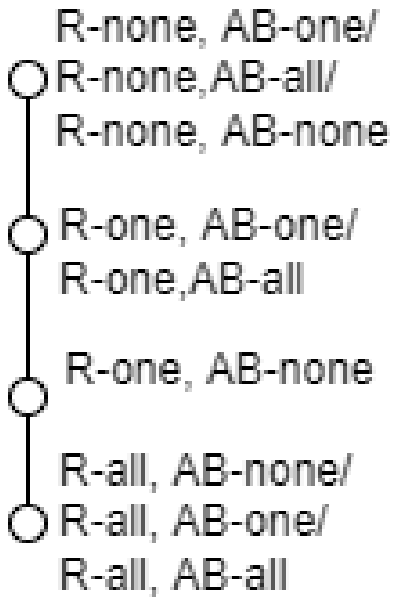


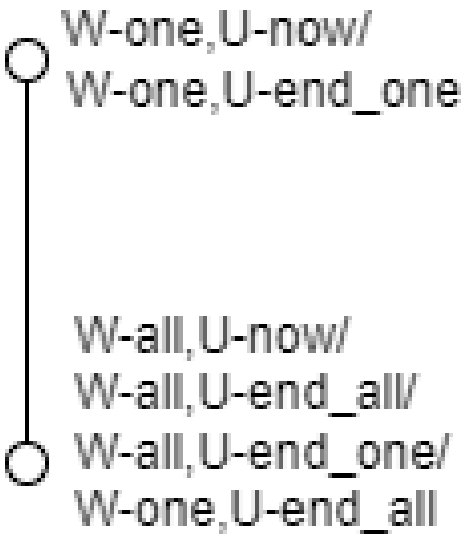
No previous work

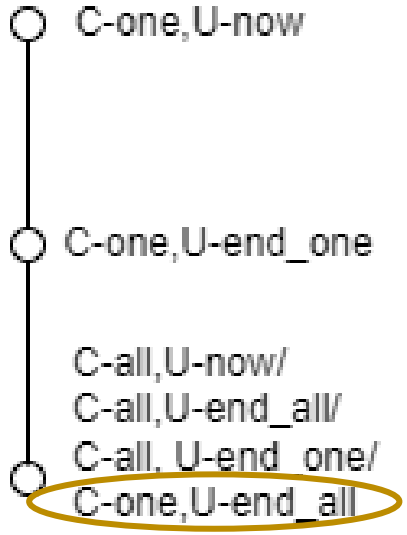
Previous work on removing and weakening



Previous work on completing







Experiments

	Mini-GALEN	Pizza	EKAW	OFSMR	PACO	NCI
Concepts	9	74	100	159	224	3304
Roles	1	33	8	2	23	1
Axioms	20	341	801	1517	1153	30364

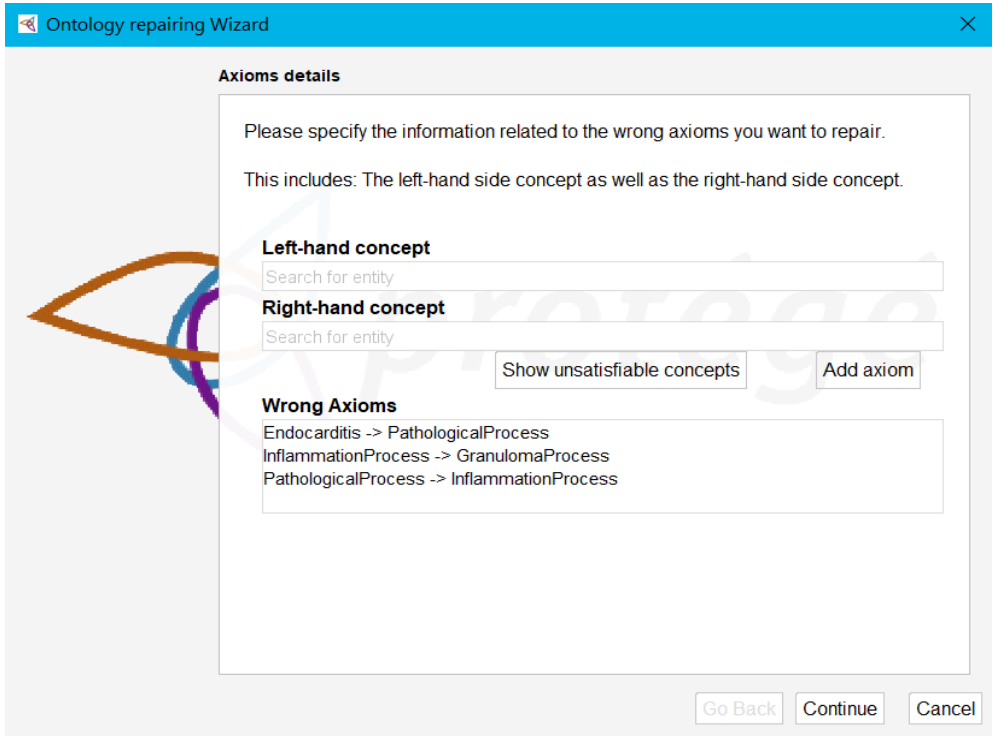
- Each ontology contains 3-5 wrong axioms
- For the Mini-GALEN ontology, the highest validation work is 8 times more than the least one.

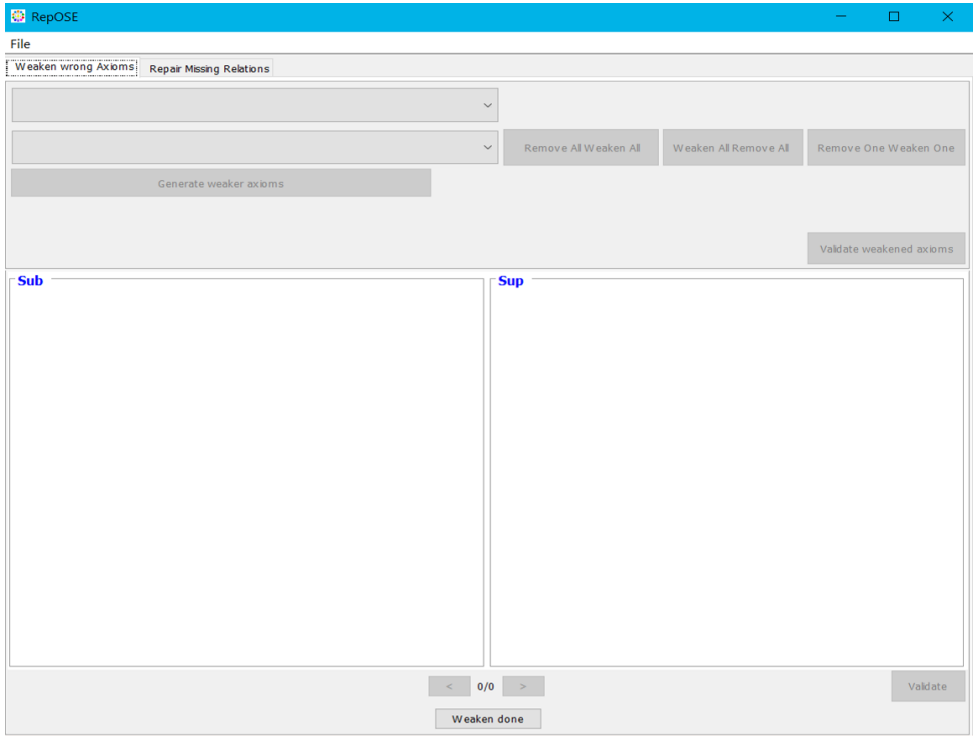
Full results can be found in the supplementary material

Systems

We implemented two systems:

- A Protégé plugin for repairing based on Algorithm C9
- The EL version of the RepOSE system





The systems are available on https://www.ida.liu.se/~patla00/publications/ESWC2023/

Conclusion

An approach combining removing, weakening, completing and expert validation.

Previous work only considered one particular combination.

A trade-off between completeness and the amount of domain expert validation.

Hasse diagrams - using operators higher up in the diagrams leads to more complete ontologies and more validation work.