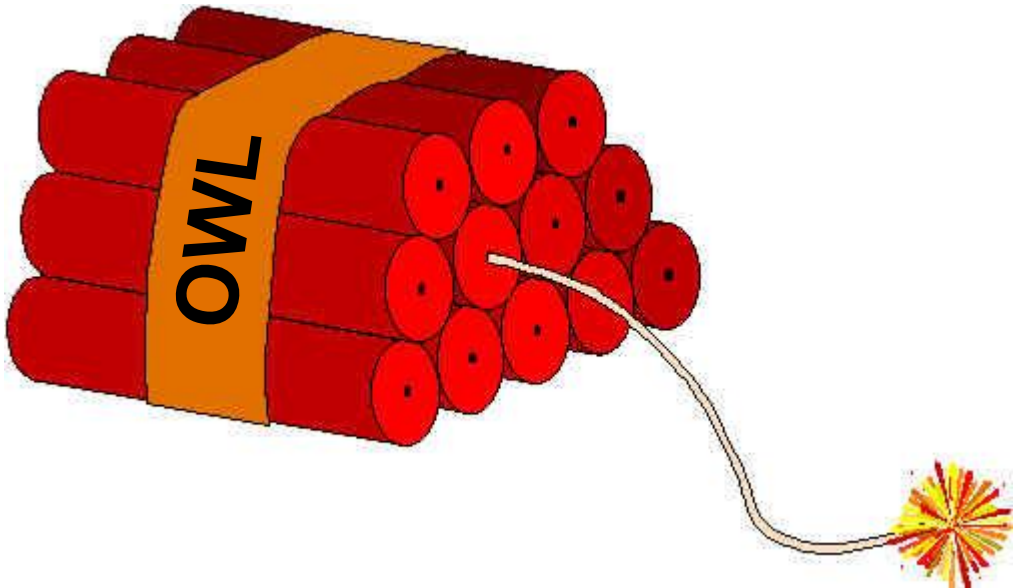
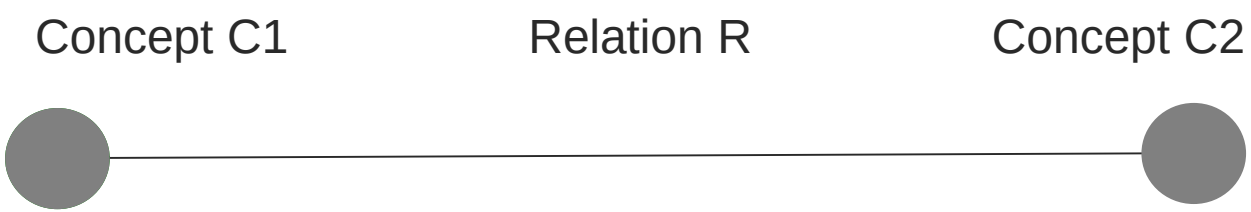
# Unintended models in OBO - OWL ontologies

Stefan Schulz
Freiburg University Medical Center, Germany



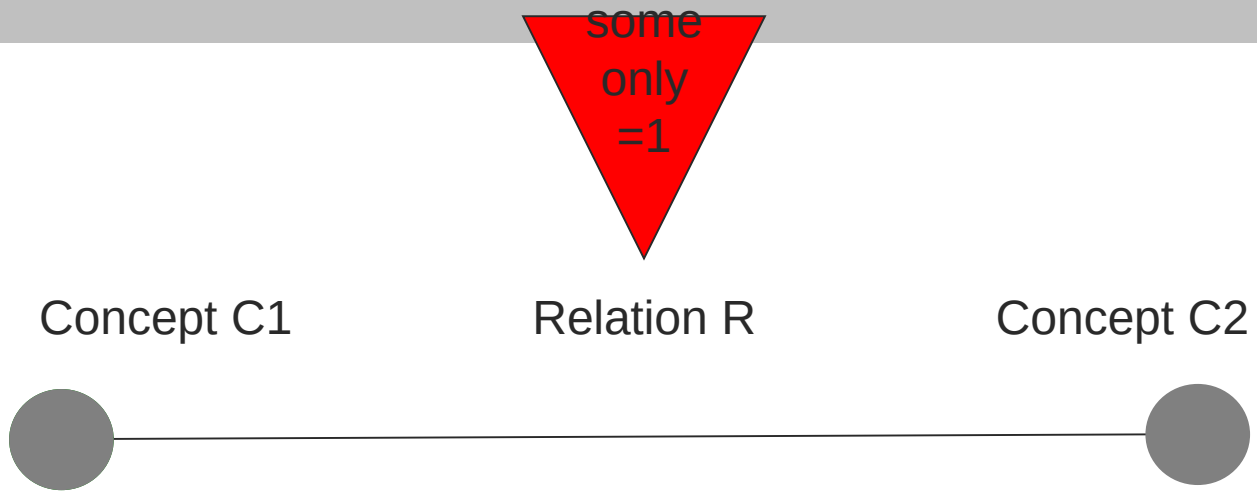
# Naïve, semantic network style approach to relations




| *Examples* | | |
|---|---|---|
| *Hepatitis* | *hasLocation* | *Liver* |
| *Hand* | *hasPart* | *Thumb* |
| *Aspirin* | *treats* | *Headache* |

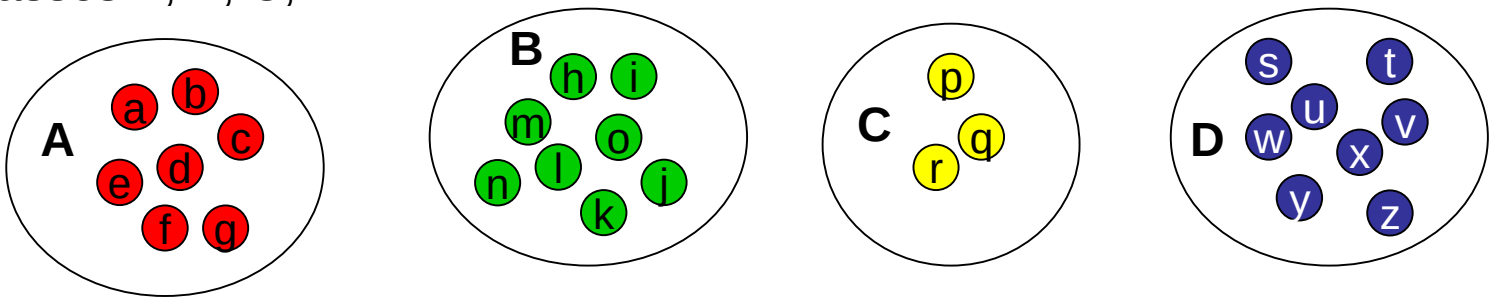
This was mainly the starting point of the OBO format

# OWL-DL approach to relations: requires quantification

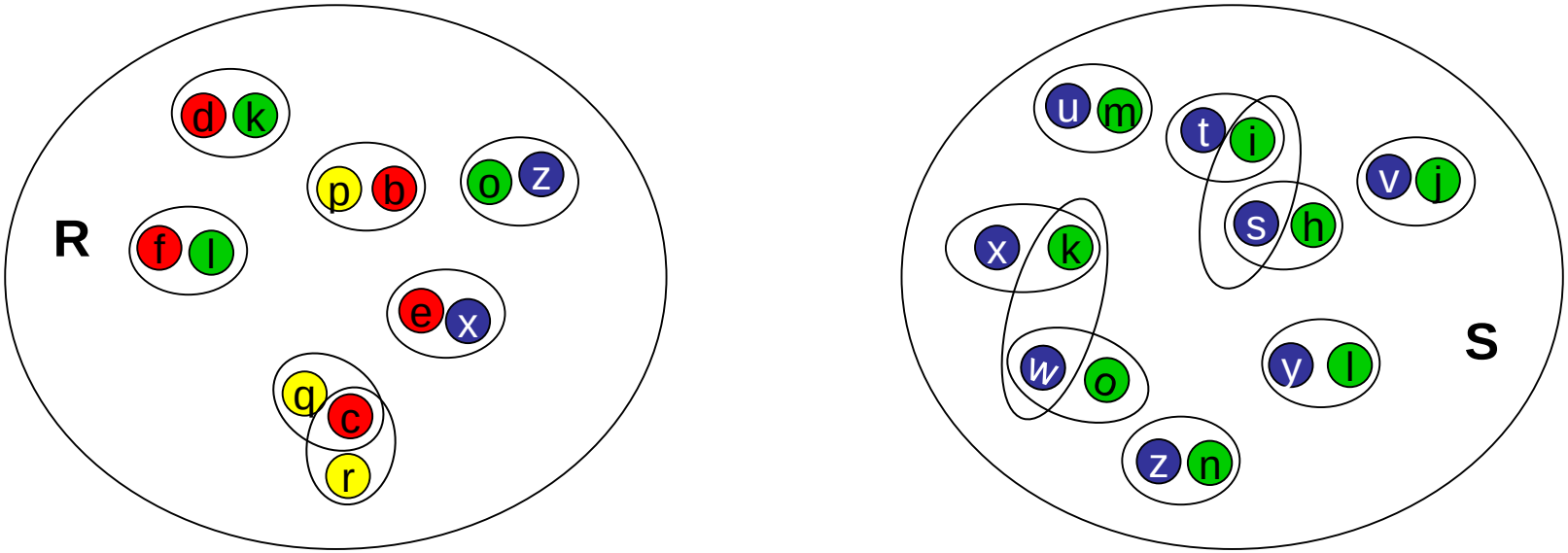



| *Examples* | | |
|---|---|---|
| *Hepatitis* | *hasLocation* | *Liver* |
| *Hand* | *hasPart* | *Thumb* |
| *Aspirin* | *treats* | *Headache* |

Individuals a, b, c, …z
Classes A, B, C, D




Relations R, S

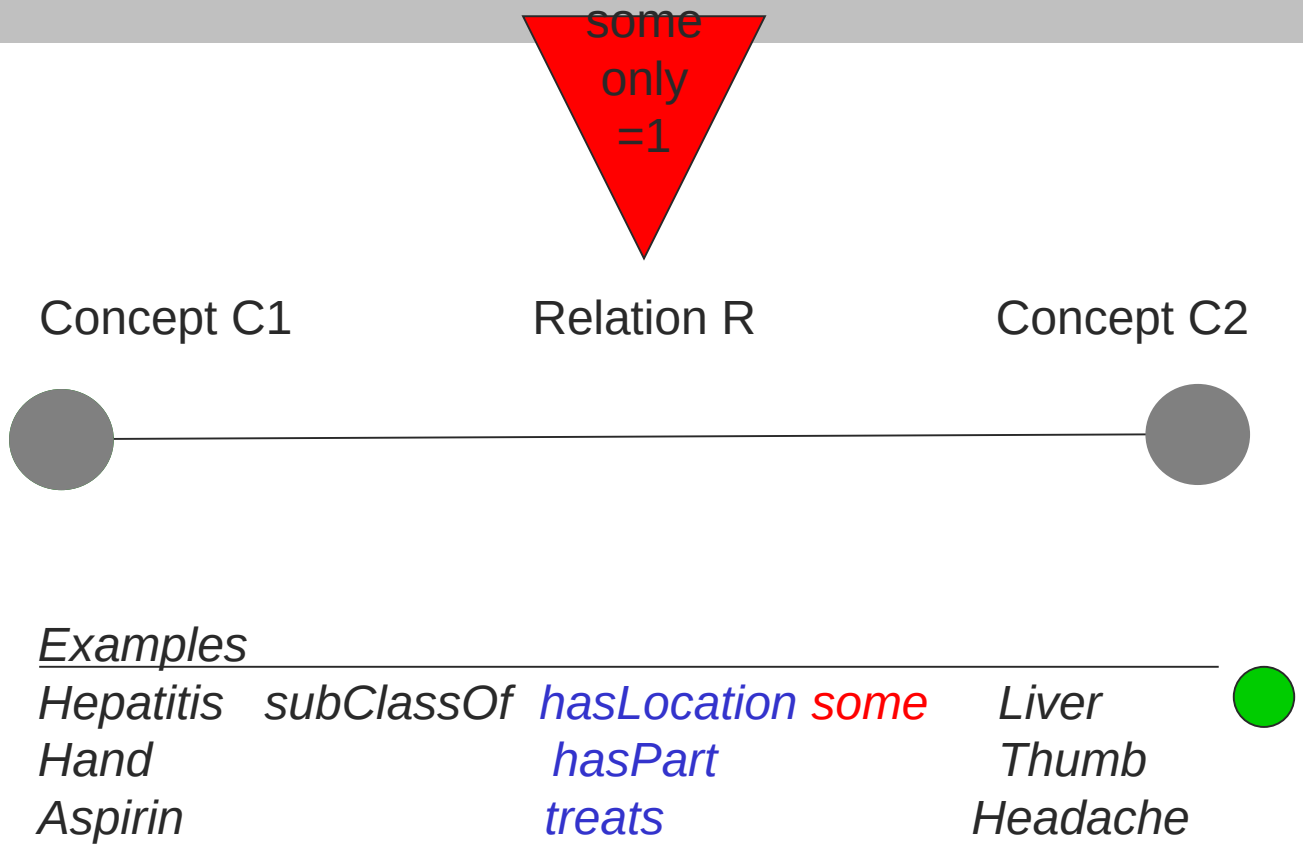



*C subclassOf R some A*

*D subclassOf S some B*
*B subclassOf inv_S some D*

What can we say about A and B?

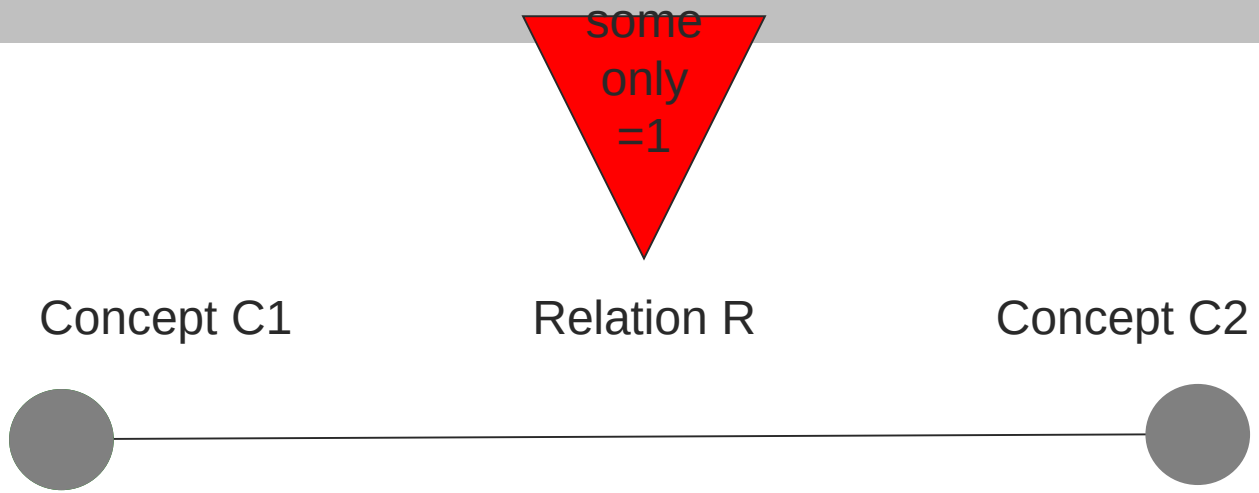
# OWL-DL approach to relations: requires quantification




*Examples*

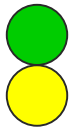
| | | | |
|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation some* | *Liver* |
| *Hand* | | *hasPart* | *Thumb* |
| *Aspirin* | | *treats* | *Headache* |

# OWL-DL approach to relations: requires quantification

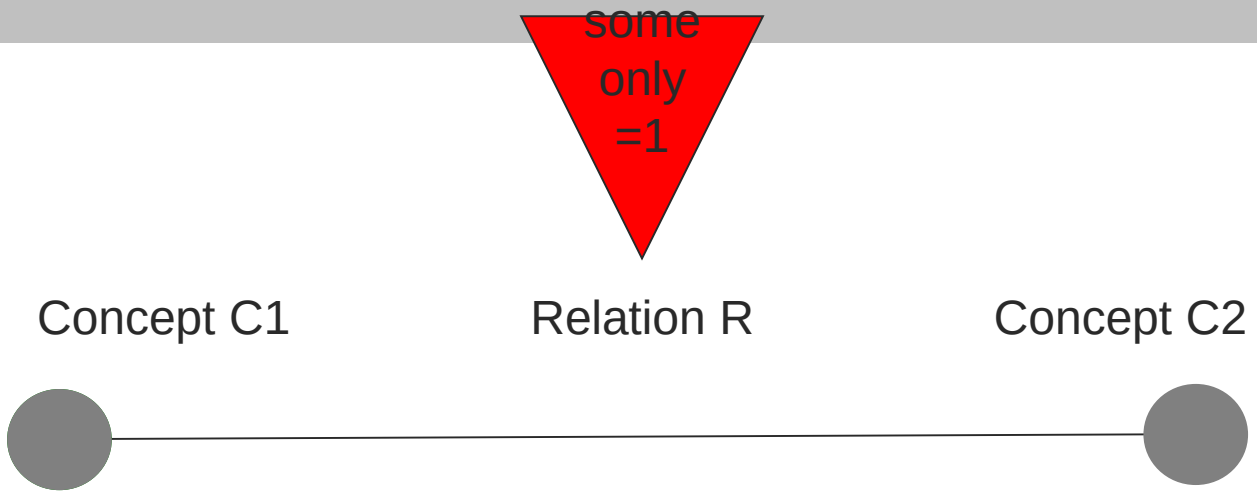



*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | | *treats* | | *Headache* |

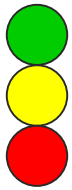


# OWL-DL approach to relations: requires quantification

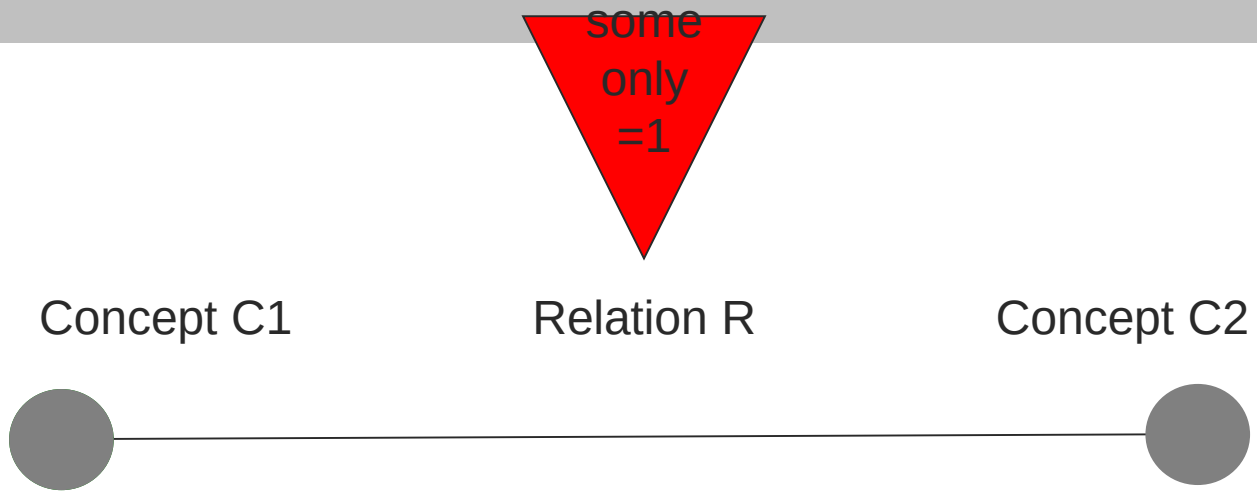



*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | *subClassOf* | *treats* | *some* | *Headache* |



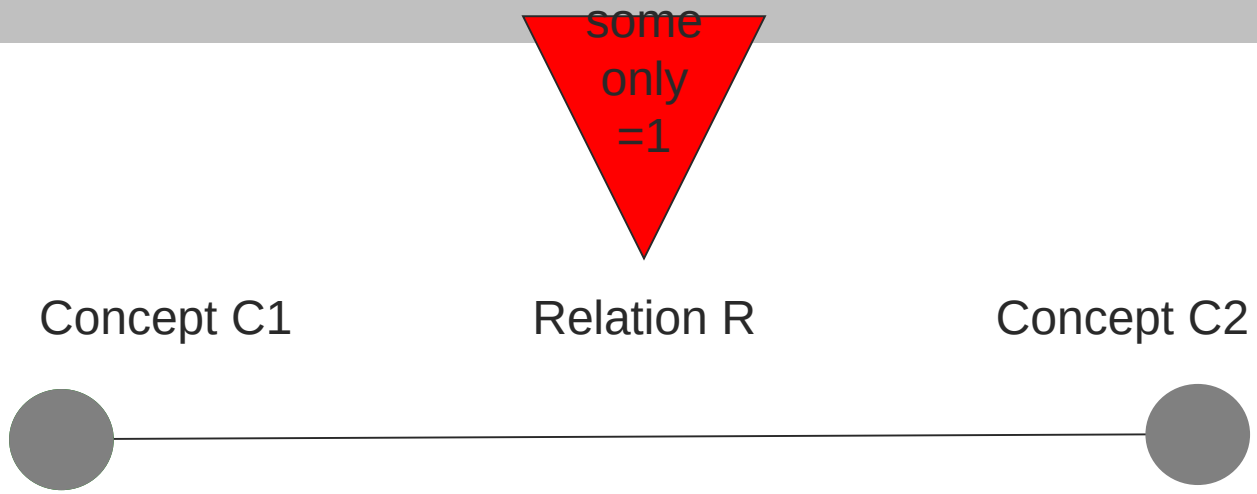
# OWL-DL approach to relations: requires quantification




*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | *subClassOf* | *treats* | *some* | *Headache* |
| *Hand* | *subClassOf* | *hasPart* | *only* | *Thumb* |

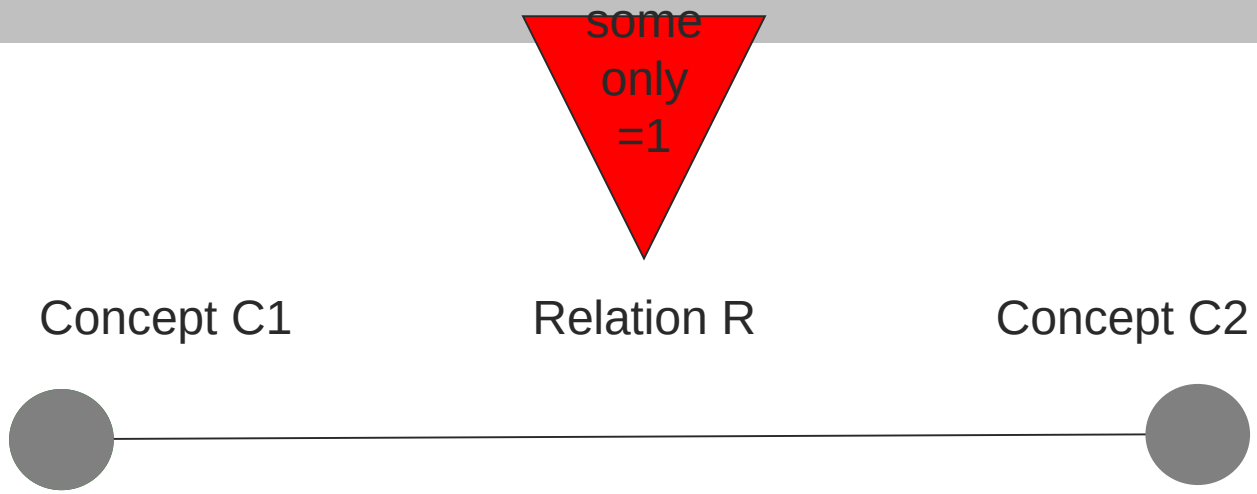
# OWL-DL approach to relations: requires quantification




*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | *subClassOf* | *treats* | *some* | *Headache* |
| *Hand* | *subClassOf* | *hasPart* | *only* | *Thumb* |
| *Thumb* | *subClassOf* | *partOf* | *some* | *Hand* |

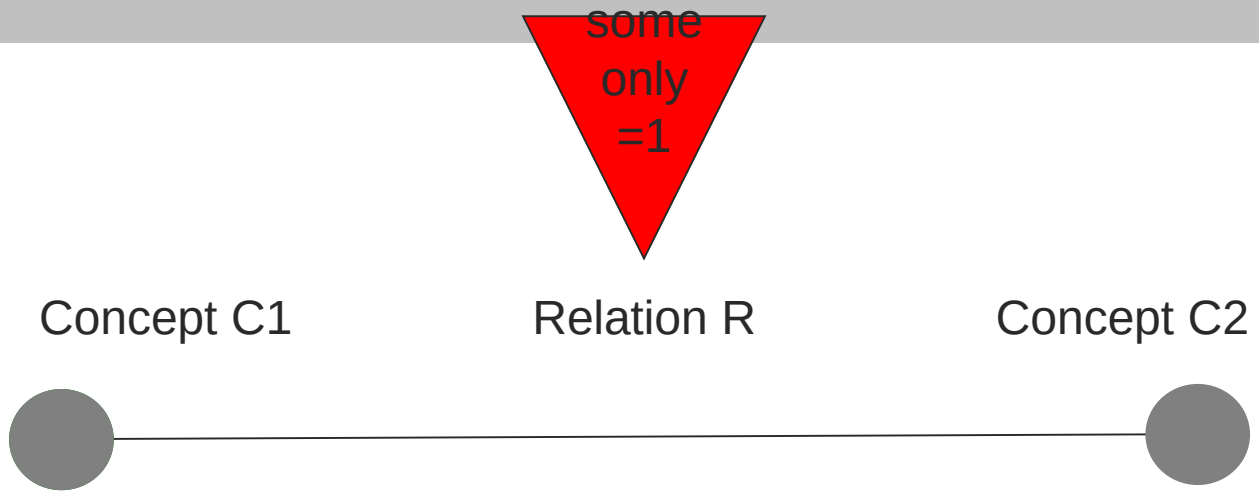
# OWL-DL approach to relations: requires quantification




*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | *subClassOf* | *treats* | *some* | *Headache* |
| *Hand* | *subClassOf* | *hasPart* | *only* | *Thumb* |
| *Thumb* | *subClassOf* | *partOf* | *some* | *Hand* |
| *Aspirin* | *subClassOf* | *treats* | *only* | *Headache* |

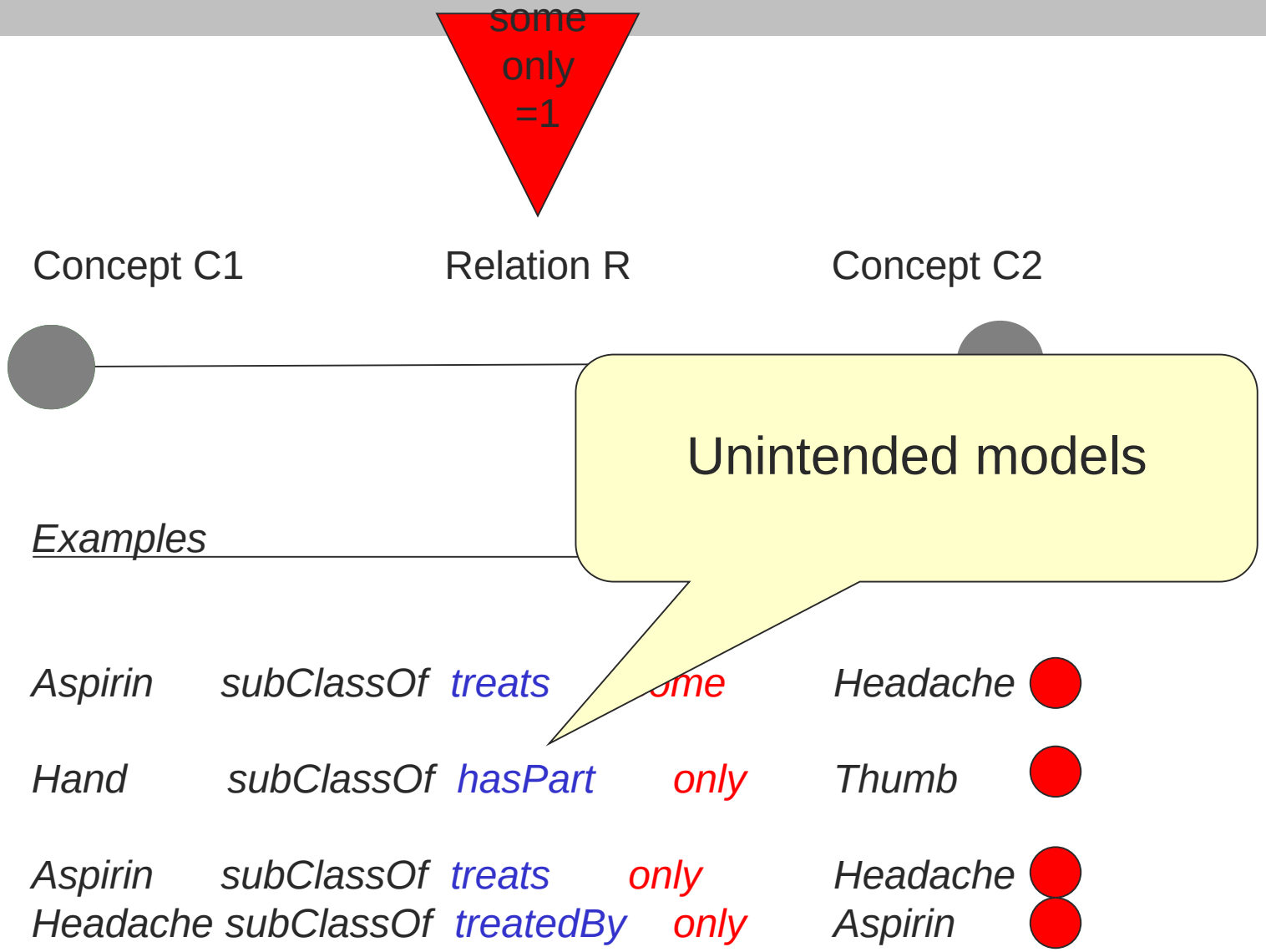
# OWL-DL approach to relations: requires quantification




*Examples*

| | | | | |
|---|---|---|---|---|
| *Hepatitis* | *subClassOf* | *hasLocation* | *some* | *Liver* |
| *Hand* | *subClassOf* | *hasPart* | *some* | *Thumb* |
| *Aspirin* | *subClassOf* | *treats* | *some* | *Headache* |
| *Hand* | *subClassOf* | *hasPart* | *only* | *Thumb* |
| *Thumb* | *subClassOf* | *partOf* | *some* | *Hand* |
| *Aspirin* | *subClassOf* | *treats* | *only* | *Headache* |
| *Headache* | *subClassOf* | *treatedBy* | *only* | *Aspirin* |

# OWL-DL approach to relations: requires quantification



# Unintended models in medical ontologies

*SNOMED CT:*

*Tonsillectomy planned SubClassOf*

*rg some (associatedProcedure some Tonsillectomy) and…*

*SNOMED CT:*

*Congenital absence of bile duct SubClassOf*

*findingSite some BileDuctStructure …*

*NCI Ontology:*

*Skin_Squamous_Cell_Carcinoma_in_situ SubClassOf*

*(diseaseMayHaveFinding some Erythema) and…*

*GALEN:*

*Vomitus subclassOf contains some carrot*

# Survey of existentially quantified relations in OBO-OWL ontologies

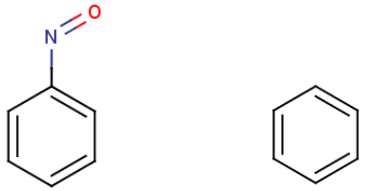
| | |
|---|---|
| part_of | 51650 |
| regional_part_of | 19665 |
| constitutional_part_of | 12392 |
| branch_of | 7167 |
| has_functional_parent | 3978 |
| has_role | 2662 |
| systemic_part_of | 2658 |
| develops_from | 2392 |
| start | 2320 |
| end | 2317 |
| DESCENDENTOF | 1868 |
| regulates | 1497 |
| is_conjugate_base_of | 1266 |
| is_conjugate_acid_of | 1266 |
| negatively_regulates | 1240 |
| positively_regulates | 1224 |
| is_enantiomer_of | 1156 |
| **has**_part | 959 |
| is_substituent_group_from | 712 |
| bounds | 660 |
| has_parent_hydride | 615 |
| lacks_modification | 538 |
| DESCINMALE | 514 |
| is_tautomer_of | 396 |
| DESCINHERM | 380 |
| attaches_to | 192 |
| has_modification | 179 |
| start_stage | 179 |
| has_quality | 160 |
| overlaps | 158 |
| starts_at_end_of | 158 |
| fasciculates_with | 134 |
| achieves_planned_objective | 130 |
| has_units | 128 |
| has_specified_output | 84 |
| unit_of | 80 |
| increased_in_magnitude_relative_to | 76 |
| end_stage | 74 |
| decreased_in_magnitude_relative_to | 74 |
| preceded_by | 68 |
| has_function | 68 |
| connected_to | 65 |
| has_specified_input | 58 |
| derives_from | 47 |
| is_specified_output_of | 28 |
| has_origin | 25 |
| agent_in | 17 |
| has_regexp | 14 |
| realizes | 13 |
| process is result of | 13 |
| surface_of | 12 |
| has Boundary | 11 |
| has Molecular Constituent | 10 |
| role_of | 10 |
| has Regional Part | 8 |
| starts_axis | 7 |
| is composed of | 7 |
| transcribed_to | 7 |
| finishes_axis | 7 |
| has_axis | 6 |
| continuous With | 6 |
| reciprocal_of | 6 |
| towards | 6 |
| adjacent_to | 6 |
| is_realized_by | 5 |
| is part of | 5 |
| has_cross_section | 5 |
| follows_axis | 4 |
| non_functional_homolog_of | 4 |
| is_concretization_of | 4 |
| opposite_to | 4 |
| innervated_by | 3 |
| member_of | 3 |
| has grain | 3 |
| variant_of | 3 |
| has Morphological Type | 2 |
| has Aggregate Part | 2 |
| bearer_of | 2 |
| is_manufactured_by | 2 |
| has Location Of | 2 |
| hasRole | 2 |
| anterior_to | 1 |
| sibling | 1 |
| singly_occurring_form_of | 1 |
| boundary Of | 1 |
| bound_to | 1 |
| associated Cell Component | 1 |
| associated Cellular Element | 1 |
| approximately_perpendicular_to | 1 |
| has Component | 1 |
| is dose in | 1 |
| is frequency in | 1 |
| hasParent | 1 |
| hasChild | 1 |
| left_of | 1 |
| function_of | 1 |
| has Neurotransmitter | 1 |
| deep_to | 1 |
| has been genotyped as | 1 |
| is enrolled in | 1 |
| guided_by | 1 |
| orthogonal_to | 1 |
| passes_through | 1 |
| dorsal_to | 1 |
| distal_to | 1 |
| posterior_to | 1 |
| objective_achieved_by | 1 |

# Example 1: Protein Ontology

- 'chordin isoform 1 unmodified form' *subclassOf*
  'chordin isoform 1 *and*
  lacks_modification
  *some* 'post-translational protein modification‘
- Problem: existence of instance of 'post-translational protein modification is expressed by logic but the intended meaning of „lacks“ is the contrary
- Relevance:
  530 lacks_modification in Protein Ontology

# Example 2: ChEBI

- *nitrosobenzene subclassOf*
  *has_parent_hydride some benzene*




- Problem: if chemicals are interpreted as extending to real chemical molecules, the axiom states that for each nitrosobenzene molecule there must be at least one benzene molecule.
- Relevance:
  9389 *is_conjugate_base_of, is_enantiomer_of, has_functional_parent, has_parent_hydride, is_conjugate_acid_of, is_tautomer_of, is_substituent_group_from*
- only in ChEBI

# Example 3: ChEBI

- *anisotropine methylbromide*

  *has_role some anti-ulcer drug*
- Problem: For each *anisotropine methylbromide* molecule there is such a role instance, regardless of it participates in a anti-ulcer treatment process
- Relevance:
  2639 *has_role*
- only in ChEBI

# Example 4: Mass spectrometry

- *ion_reaction subclassOf*

  *part_of some ion*
- Problem: For each instance of *ion_reaction* there is some ion it is part of. Existential implication OK, but wrong relation
- Relevance:

  113 *part_of,* most of them obviously with this problem
- only in MS

# Example 5: SOPHARM

- *patient subclassOf*

  *person and hasRole some patient_role and*

  *(is_part_of some clinical_trial_panel) or*

  *(is_enrolled_in some clinical trial)*

- Problem: nobody is a patient unless participating in a clinical trial. Not really a problem of logic, rather a problem of naming

# Example 5: SOPHARM

- *(inferred)*
  *calcium channel complex subclassOf*
  *part_of some cell*
- Problem: calcium channels must be parts of cells (unless part of means part at some instant in time). With this restriction, however, part-of is no longer transitive.

| | |
|---|---|
| part_of | 51650 |
| regional_part_of | 19665 |
| constitutional_part_of | 12392 |
| branch_of | 7167 |
| has_functional_parent | 3978 |
| has_role | 2662 |
| systemic_part_of | 2658 |
| develops_from | 2392 |
| Start | 2320 |
| End | 2317 |
| DESCENDENTOF | 1868 |
| Regulates | 1497 |
| is_conjugate_base_of | 1266 |
| is_conjugate_acid_of | 1266 |
| negatively_regulates | 1240 |
| positively_regulates | 1224 |
| is_enantiomer_of | 1156 |
| **has**_part | 959 |
| is_substituent_group_from | 712 |
| Bounds | 660 |
| has_parent_hydride | 615 |
| lacks_modification | 538 |
| DESCINMALE | 514 |
| is_tautomer_of | 396 |
| DESCINHERM | 380 |
| attaches_to | 192 |
| has_modification | 179 |
| start_stage | 179 |
| has_quality | 160 |
| overlaps | 158 |
| starts_at_end_of | 158 |
| fasciculates_with | 134 |
| achieves_planned_objective | 130 |
| has_units | 128 |
| has_specified_output | 84 |
| unit_of | 80 |

| | |
|---|---|
| increased_in_magnitude_relative_to | 76 |
| end_stage | 74 |
| decreased_in_magnitude_relative_to | 74 |
| preceded_by | 68 |
| has_function | 68 |
| connected_to | 65 |
| has_specified_input | 58 |
| derives_from | 47 |
| is_specified_output_of | 28 |
| has_origin | 25 |
| agent_in | 17 |
| has_regexp | 14 |
| realizes | 13 |
| process is result of | 13 |
| surface_of | 12 |
| has Boundary | 11 |
| has Molecular Constituent | 10 |
| role_of | 10 |
| has Regional Part | 8 |
| starts_axis | 7 |
| is composed of | 7 |
| transcribed_to | 7 |
| finishes_axis | 7 |
| has_axis | 6 |
| continuous With | 6 |
| reciprocal_of | 6 |
| towards | 6 |
| adjacent_to | 6 |
| is_realized_by | 5 |
| is part of | 5 |
| has_cross_section | 5 |
| follows_axis | 4 |
| non_functional_homolog_of | 4 |
| is_concretization_of | 4 |
| opposite_to | 4 |
| innervated_by | 3 |

| | |
|---|---|
| member_of | 3 |
| has grain | 3 |
| variant_of | 3 |
| has Morphological Type | 2 |
| has Aggregate Part | 2 |
| bearer_of | 2 |
| is_manufactured_by | 2 |
| has Location Of | 2 |
| hasRole | 2 |
| anterior_to | 1 |
| sibling | 1 |
| singly_occurring_form_of | 1 |
| boundary Of | 1 |
| bound_to | 1 |
| associated Cell Component | 1 |
| associated Cellular Element | 1 |
| approximately_perpendicular_to | 1 |
| has Component | 1 |
| is dose in | 1 |
| is frequency in | 1 |
| hasParent | 1 |
| hasChild | 1 |
| left_of | 1 |
| function_of | 1 |
| has Neurotransmitter | 1 |
| deep_to | 1 |
| has been genotyped as | 1 |
| is enrolled in | 1 |
| guided_by | 1 |
| orthogonal_to | 1 |
| passes_through | 1 |
| dorsal_to | 1 |
| distal_to | 1 |
| posterior_to | 1 |
| objective_achieved_by | 1 |

# Conclusion

- The use of OWL requires a precise ontological commitment
  - *is a hand without a thumb still a hand*
  - *what about a severed thumb*
- Many important statements cannot be adequately represented
  - OWL semantic enforces statements of the type „for all… some“ or „for all… only“
  - No way to express what is mostly or normally true
- Work in progress:
  - systematic analysis
  - definition of error classes
  - suggestion of alternative models