How to Distinguish Parthood from Location in Bio-Ontologies

Stefan Schulz[a,b], Philipp Daumke[a],
Barry Smith[c,d], Udo Hahn[e]

[a]Department of Medical Informatics, Freiburg University Hospital, Germany

[b]Health Informatics Laboratory, Paraná Catholic University, Brazil

[c]Department of Philosophy, The New York State Center of Excellence in Bioinformatics and Life Sciences, University at Buffalo, NY, USA.

[d]IFOMIS, Saarland University, Germany

[e]Jena University Language and Information Engineering (JULIE) Lab, Germany

Ontologies of Biological Structure ("Anatomies")

- Foundational Model of Anatomy (FMA)
- Human Anatomy portions in OpenGalen, SNOMED CT, NCI ontology,…
- Cell Component branch in Gene Ontology
- Open Biological Ontologies (OBO):
  - Human development
  - Mouse (adult / embryo), Zebrafish, Drosophila, C. elegans,…
  - General plant, maize, cereal plant,…
  - Increasing repository of biological structure descriptions

Example: **Adult Mouse Anatomy**

Term Detail

MA term: **metatarsal bone digit 1**
MA id: **MA:0001369**
Number of paths to term: **9**

Ⓘdenotes an 'is-a' relationship
Ⓟdenotes a 'part-of' relationship

mouse anatomy
- Ⓟadult mouse
  - Ⓟanatomic region
    - Ⓘlimb
      - Ⓘhindlimb
        - Ⓟfoot
          - Ⓟfoot bone
            - Ⓘmetatarsal bone
              - Ⓘmetatarsal bone digit 1 [MA:0001369]
              - Ⓘmetatarsal bone digit 2
              - Ⓘmetatarsal bone digit 3
              - Ⓘmetatarsal bone digit 4
              - Ⓘmetatarsal bone digit 5

mouse anatomy
- Ⓟadult mouse
  - Ⓟanatomic region
    - Ⓘlimb
      - Ⓘhindlimb
        - Ⓟfoot
          - Ⓟmetatarsus
            - Ⓟmetatarsal bone
              - Ⓘmetatarsal bone digit 1 [MA:0001369]
              - Ⓘmetatarsal bone digit 2
              - Ⓘmetatarsal bone digit 3

- Orthogonal *Part-of* and *Is-a* hierarchies are backbones of bio-ontologies
- *Part-of* and *Is-a* are generally considered "foundational relations"
- Recent standardization of the semantics of *Is-a* and *Part-of* as asserted between classes

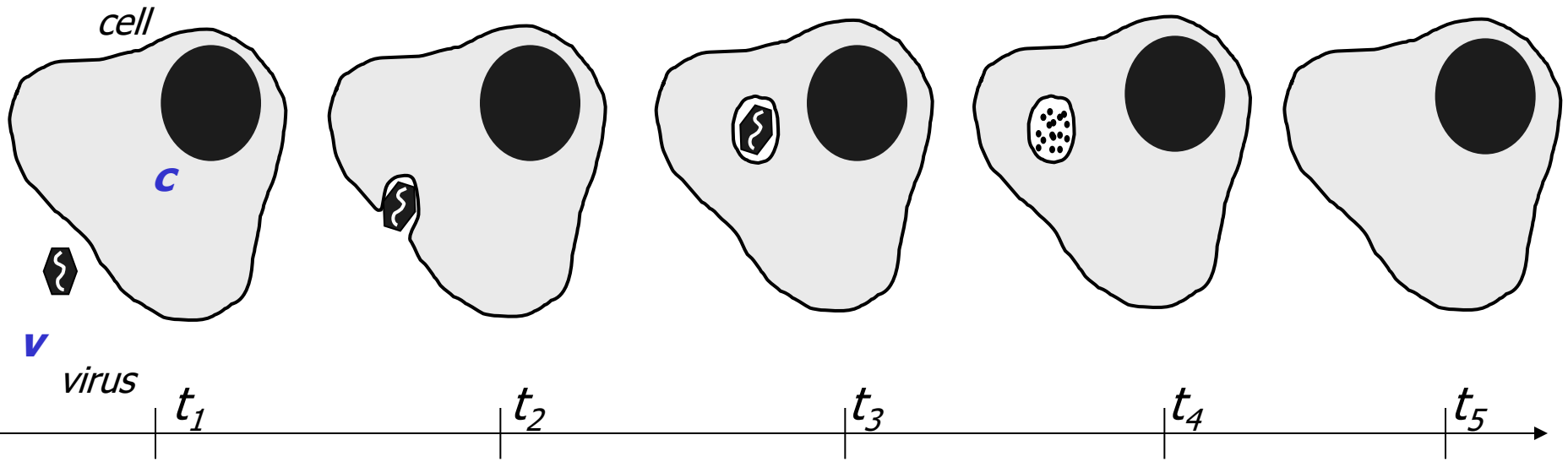
Smith et al.: Relations in Biomedical Ontologies. Genome Biology, 2005, 6 (5)

Is *part-of* a Foundational Relation ?

- Foundational relations are supposed to be robust with regard to individual interpretations.
- Observation: many assertions of parthood are tied to human perception and belief

Is *part-of* a Foundational Relation ?

- Foundational relations are supposed to be robust with regard to individual interpretations.
- Many assertions of parthood are tied to human perception and belief



Parthood assertions are controversial

Instances of …

Part ?	**Whole**	**Part ?**	**Whole**
Transplant	Organism	Thyroxin Molecule	Thyroid Gland
Mitochondrium	Cell	Alanin Molecule	Collagen Fiber
E.Coli bacterium	Intestine	Bolus of Food	Stomach
H_20 molecule	Cytoplasm	Transfused Blood	Body
Glioblastoma	Brain	Zygote	Uterus
Brain metastasis	Brain	Artificial Head	Femur

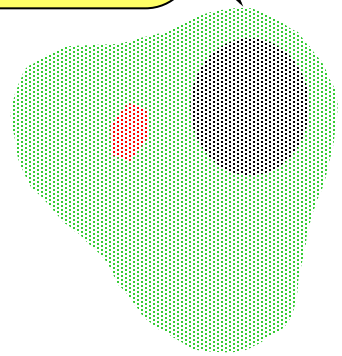
Beyond controversy: *Located-in (region-contained-in)*

$$\textit{located-in}\ (x, y, t) =_{def} \textit{part-of}\ (r\ (x, t), r\ (y, t))$$

Beyond controversy: *Located-in (region-c…in)*

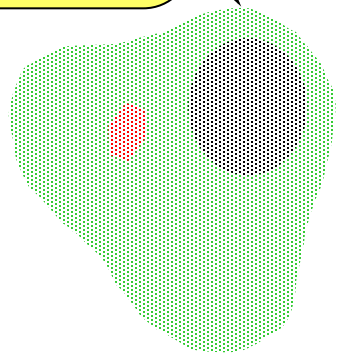
Parthood between regions = point set inclusion

$$located\text{-}in\ (x, y, t) =_{def} part\text{-}of\ (r\ (x, t), r\ (y, t))$$



Relation Hierarchy

Parthood between regions = point set inclusion



located-in $(x, y, t) =_{def}$
part-of $(r(x, t), r(y, t))$

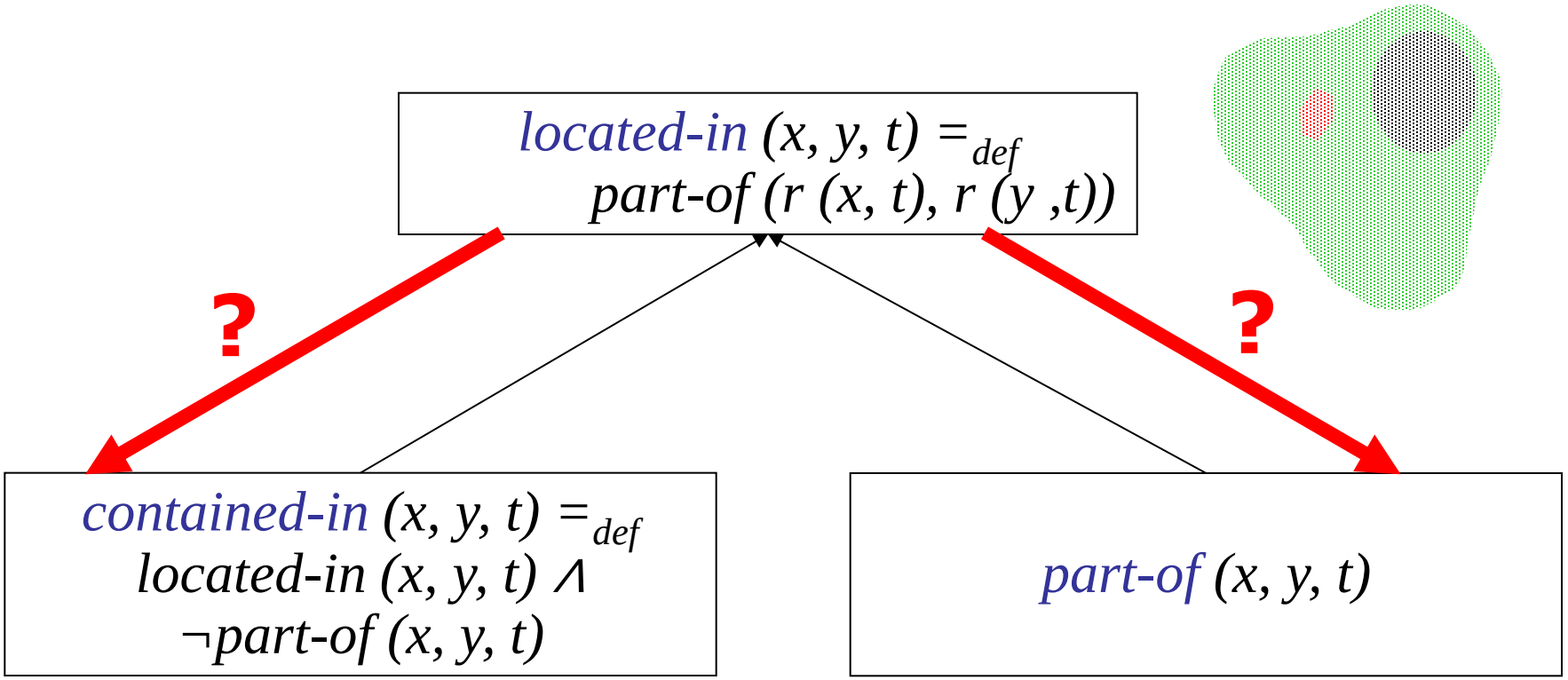
contained-in $(x, y, t) =_{def}$
located-in $(x, y, t) \wedge$
$\neg$*part-of* (x, y, t)

part-of (x, y, t)

Problem Statement

- Parthood always implies spatial location, but spatial location does not always imply parthood
- Under which circumstances can we infer parthood from spatial location ? When does inclusion without parthood obtain ?

Relation Hierarchy



Proposal: Four criteria for inferring parthood

1. Sortality
2. Genetic identity
3. Life Cycle
4. Function / Integrity

Inferring part from spatial inclusion: 1. Sortality

Rules out objects of certain sort as parts:

- x is material, y is immaterial:

$$Solid(x) \wedge Hole \rightarrow (y) \wedge located\text{-}in(x, y) \rightarrow \neg\, part\text{-}of(x, y)$$

$$located\text{-}in(myBrain, myCranialCavity) \rightarrow \neg\, part\text{-}of(myBrain, myCranialCavity)$$

- x is an non-biological artifact:

$$located\text{-}in(myPacemaker, myBody) \rightarrow \neg\, part\text{-}of(myPacemaker, myBody)$$

$$located\text{-}in(myInlay, myTooth) \rightarrow \neg\, part\text{-}of(myInlay, myTooth)$$

Inferring part from spatial inclusion: 2. Genetic Identity

Rules out objects of different genetic origin:

- Symbionts:

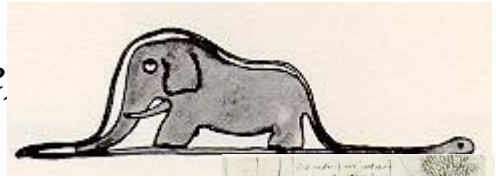
  located-in (anEcoliBacterium , myIntestine) $\rightarrow$
  $\neg$ *part-of (anEcoliBacterium , myIntestine)*

- Parasites:

  located-in (anEchinococcus, myLiver) $\rightarrow$
  $\neg$ *part-of (anEchinococcus, myLiver)*

- Preys:

  located-in (anElephant, aSnake) $\rightarrow$
  $\neg$ *part-of (anElephant, aSnake)*

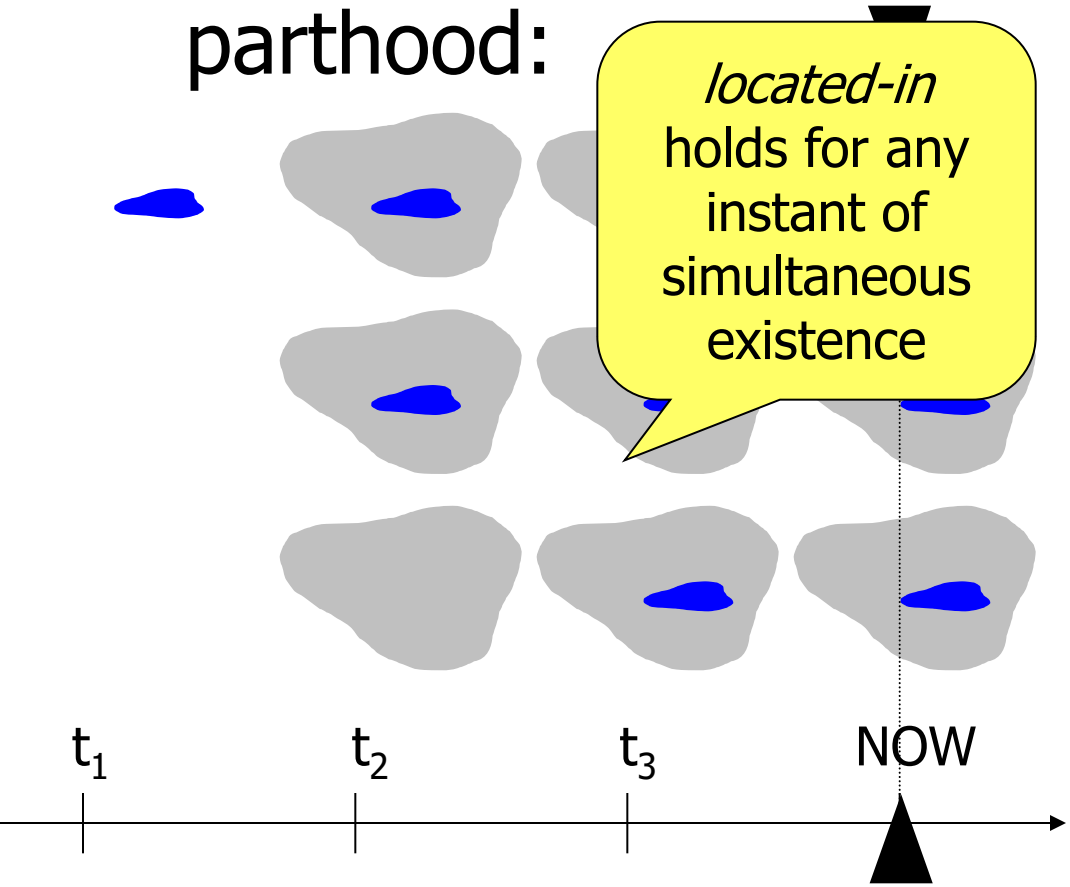


- Zygotes, Embryos, Fetuses:

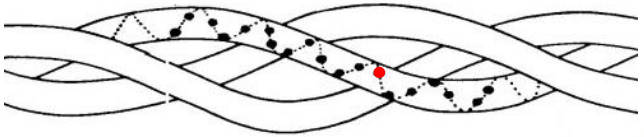
  located-in (Leonardo, Caterina) $\rightarrow$
  $\neg$ *p (Leonardo, Caterina)*

Inferring part from spatial inclusion: 3. Life Cycle

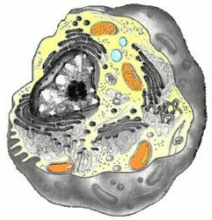
3. Life Cycle patterns which allow to assert parthood:



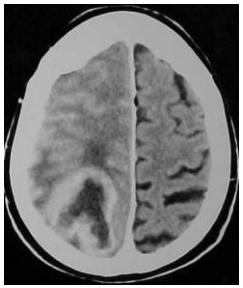
aGlycinMolecule, aCollagenFiber



aCytoplasm, aCell

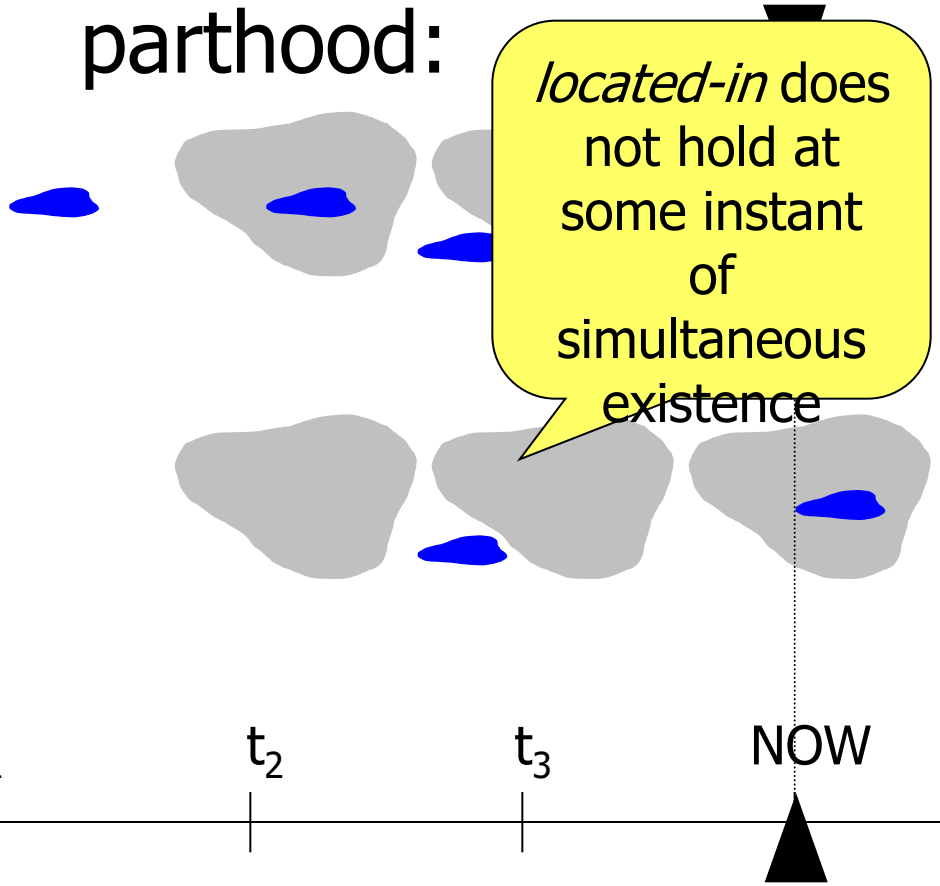


aGlioblastoma, aBrain

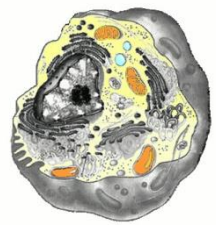


Inferring part from spatial inclusion: 3. Life Cycle

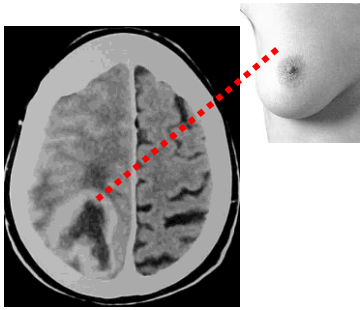
3. Life Cycle patterns which allow to rule out parthood:



aWaterMolecule, aCell



aBrainMetastasis, aBrain

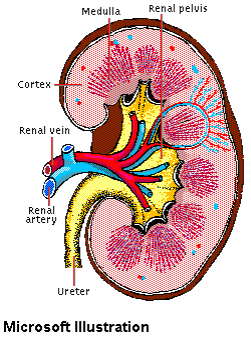


Inferring part from spatial inclusion: 4. Function / Integrity

4. Related to function or integrity

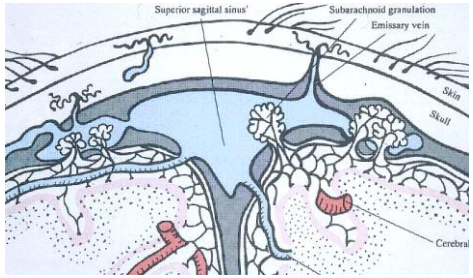
- Transplants

functionally_related (aTransplant, anOrganism)
∧ located-in (aTransplant, anOrganism) →
part-of (aTransplant, anOrganism)



- Body Substances:

functionally_related (myCSF, myCNS)
∧ located-in (myCSF, myCNS) →
part-of (myCSF, myCNS)



… but not: *part-of (thisVolumeOfUrine, myBladder),* because not essential for function

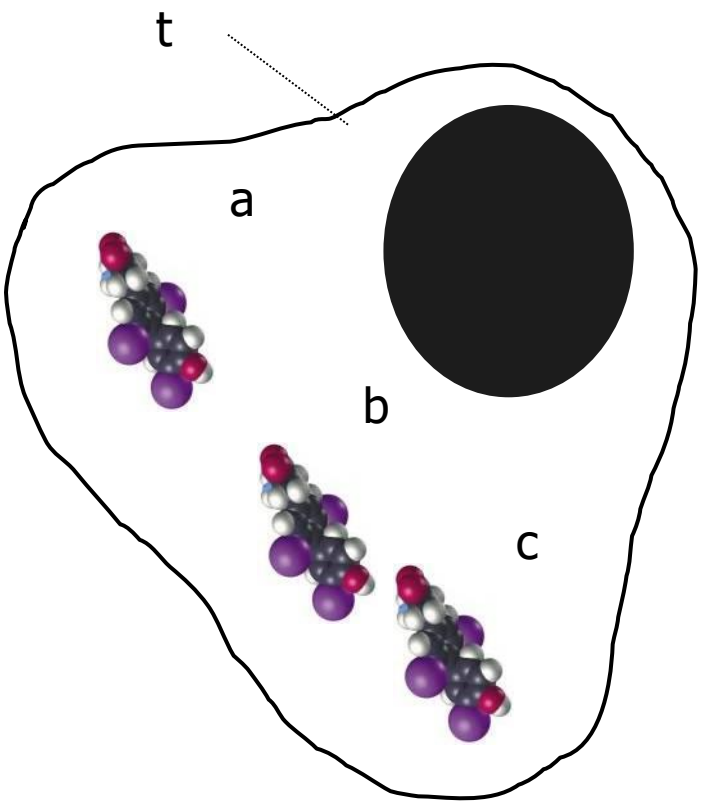
Inferring part from spatial inclusion: Decision algorithm

1 If **located-in** (*c*, *d*, *t*)
1 If **Artifact**(*c*) then
contained-in(*c*, *d*, *t*)
Else
4 If **function-integrity-relevant** (*c*, *d* , *t*) then
part-of (*c*, *d*, *t*)
Else
4 If not **same-genetic-origin** (*c*, *d*, *t*) or
1 (**instance-of** (*c*, *Material*) and
instance-of (*d*, *Immaterial*)) then
contained-in (*c*, *d*, *t*)
Else
3 If **hitherto-located-in** (*c*, *d*, *t*) or
(**hitherto-located-in** (*c*, *m*, *t*) and
part-of (*m*, *d*, *t*)) then
part-of (*c*, *d*, *t*)
Else
contained-in (*c*, *d*, *t*)
End If
End If
End If
End If
End If

Borderline cases

- Fuzzy notion of “artifact”: engineered tissue, genetically modified cells
- Unclear identity: e.g., tumors, metastases (where does their existence begin ?)
- “Sameness” of masses defined by their containers
  (air in the lung, blood in the heart, urine in the bladder)

Counter-intuitive consequences



- a: Thyroxine molecule synthesized by c -> *part-of (a, t)*
- b: Thyroxine molecule synthesized by other cell -> *contained-in (b, t)*
- c: Thyroxine molecule ingested as drug -> *contained-in (c, t)*

Acknowledgement: Anand Kumar

Conclusion

- Spatial location (topological) inclusion: non-controversial foundational relation for bio-ontologies
- *part-of* more useful exhibits human-dependent semantic bias
- Algorithmic approach for specializing location to either parthood or containment
- Problems persist: borderline cases, unintuitive cases, ill-defined notion of functionality / integrity

How to Distinguish Parthood from Location in Bio-Ontologies

Stefan Schulz[a,b], Philipp Daumke[a],
Barry Smith[c,d], Udo Hahn[e]

[a]Department of Medical Informatics, Freiburg University Hospital, Germany

[b]Health Informatics Laboratory, Paraná Catholic University, Brazil

[c]Department of Philosophy, The New York State Center of Excellence in Bioinformatics and Life Sciences, University at Buffalo, NY, USA.

[d]IFOMIS, Saarland University, Germany

[e]Jena University Language and Information Engineering (JULIE) Lab, Germany