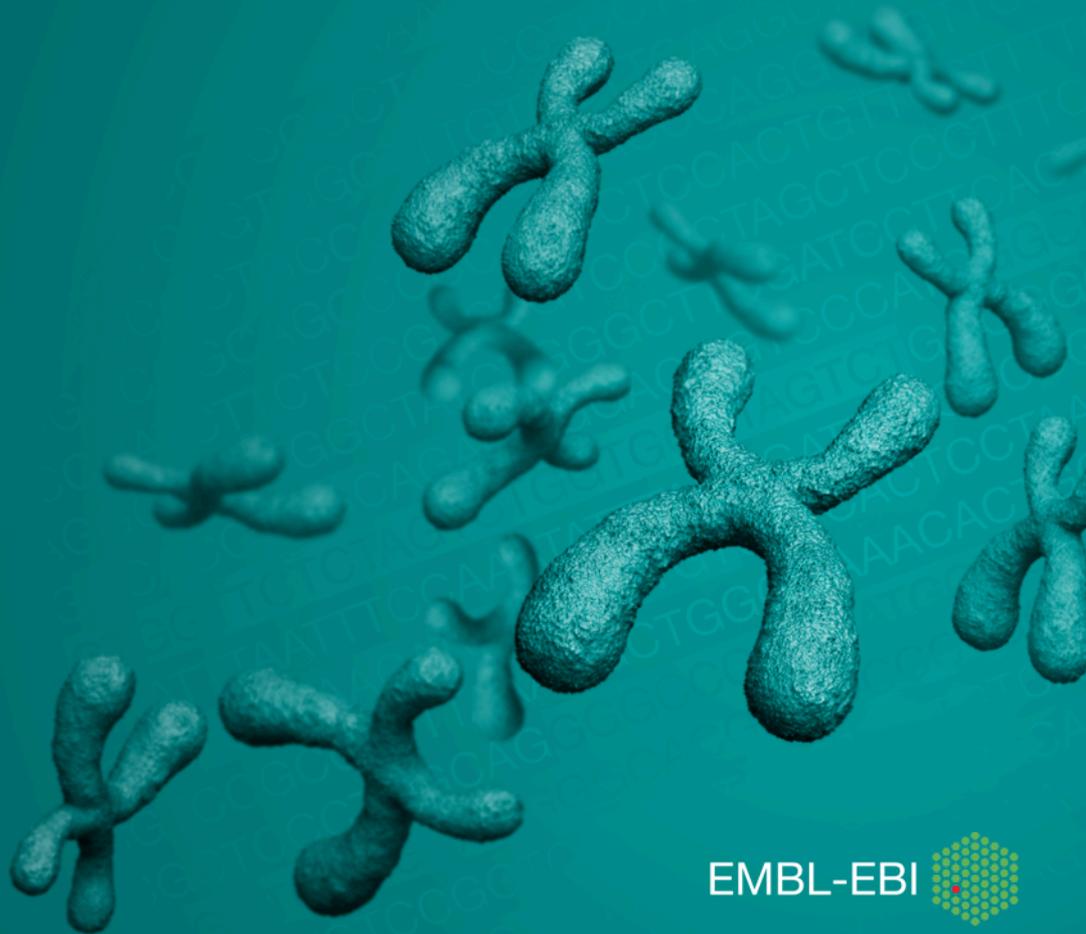


# OWL primer for GO users

David Osumi-Sutherland



# What is OWL?

- Web Ontology Language
  - Can express everything in OBO and more.



- Certified web standard



- Fast reasoning software allows:
  - automated ontology building;  
error checking; queries



# I'm an annotator, why should I care about OWL?

- OBO is OWL
  - OBO 1.4 spec defines OBO as a mapping to OWL.
- OWL reasoning is used by GO to:
  - Automate classification during ontology building
  - Check for errors (inconsistencies) in ontology and annotations
  - Drive TermGenie
  - **Tell you how annotation extensions fold**

# I'm an annotator, why should I care about OWL?

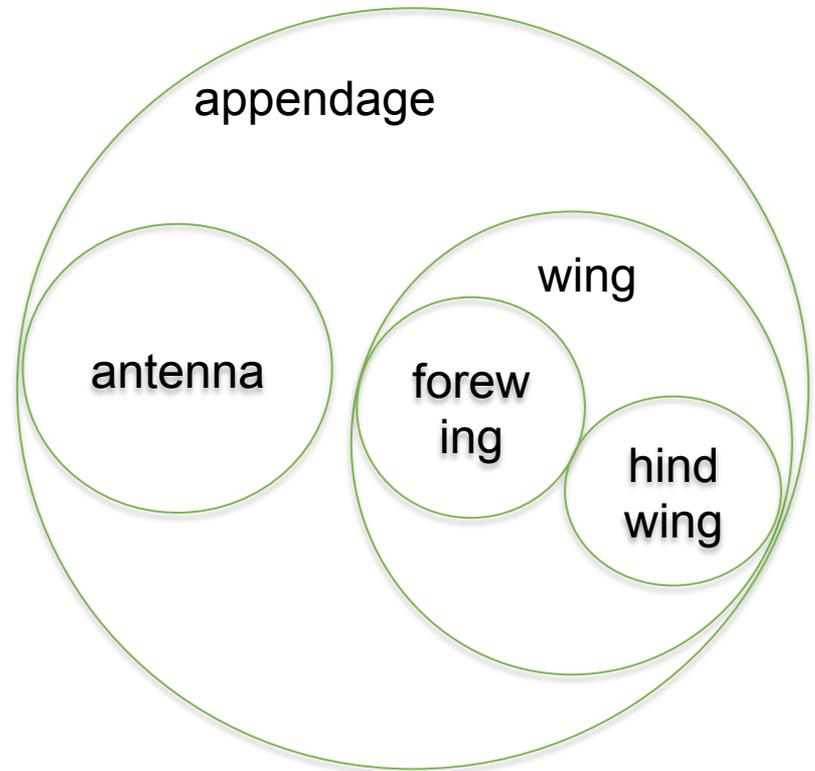
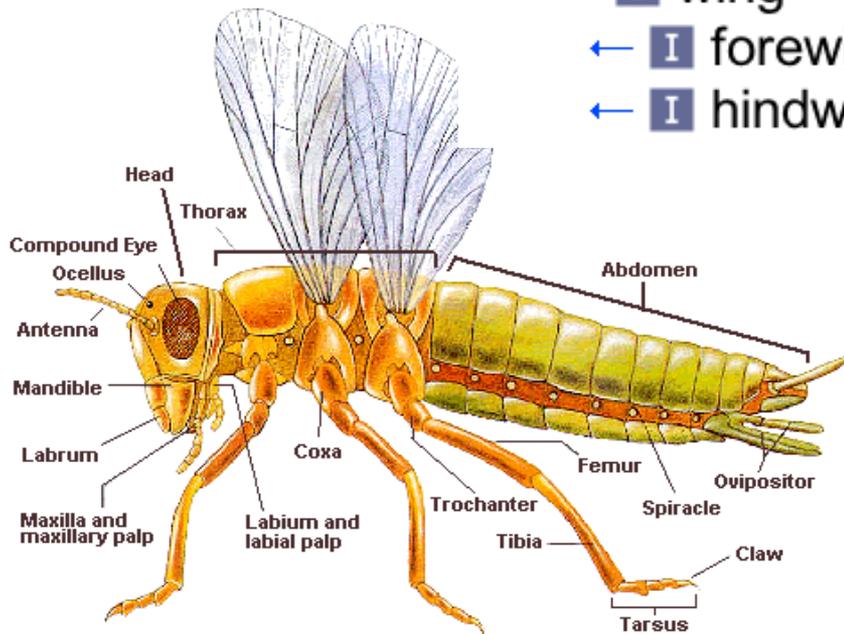
- OBO is OWL
  - OBO 1.4 spec defines OBO as a mapping to OWL.
- OWL reasoning is used by GO to:
  - Automate classification during ontology building
  - Check for errors (inconsistencies) in ontology and annotations
  - Drive TermGenie
  - **Tell you how annotation extensions fold**

**YOU CAN'T KNOW HOW AN ANNOTATION EXTENSION  
WILL FOLD WITHOUT ASKING A REASONER!**

# What is an ontology ?

A classification

- ▣ Classes
  - ▣ — appendage
    - ← I antenna
  - ▣ — leg
    - ← I leg
  - ▣ — wing
    - ← I forewing
    - ← I hindwing



# OBO-OWL cheat sheet: classification

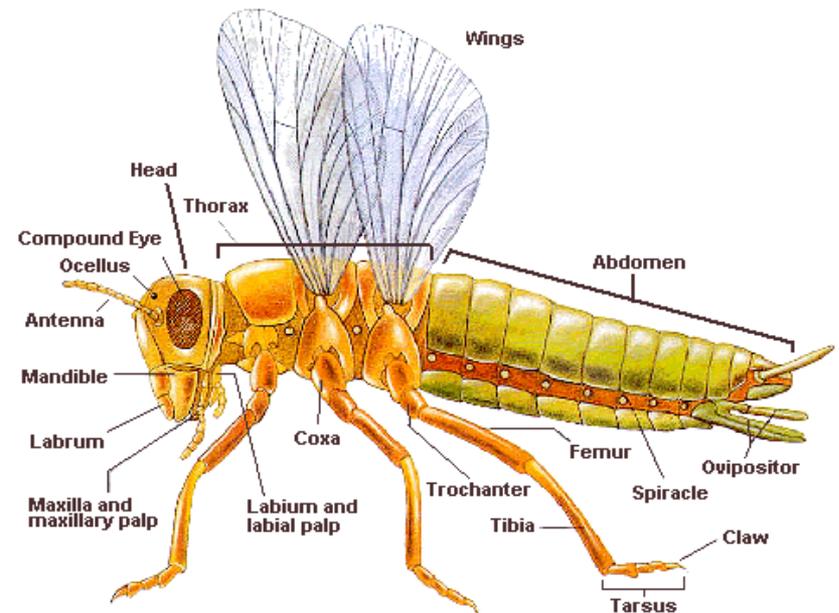
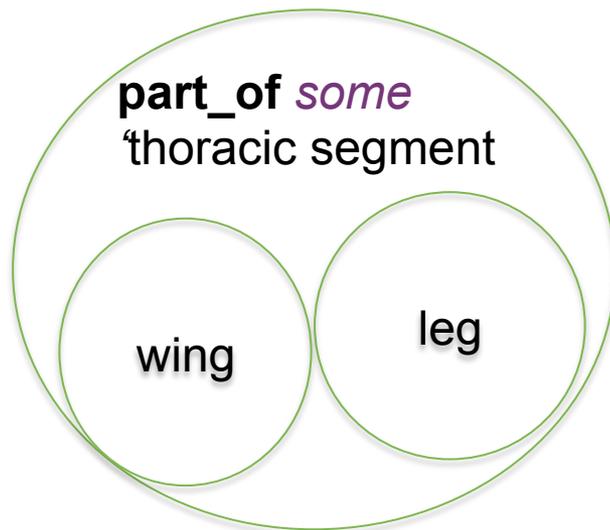
OWL: antenna `SubClassOf` appendage

OBO: antenna `is_a` appendage

# Relationships record necessary conditions for class membership

Being part of a thoracic segment is a necessary condition of being in the class leg

'leg' *SubClassOf* **part\_of** *some* thoracic segment

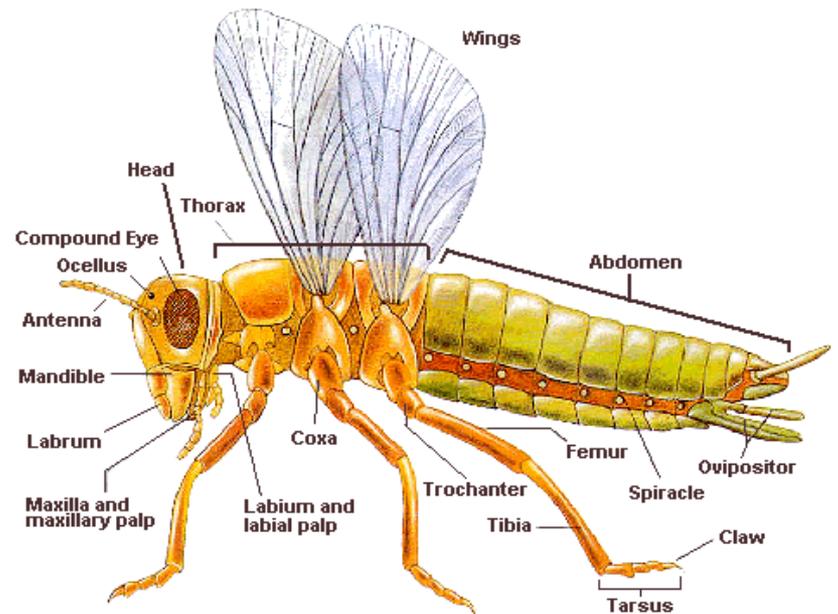


# class – class relationships are quantified

- Class:Class relationships are many to many
  - Does the relation apply to all or just some of the class ?
  - we specify this with quantifiers:
    - *some*  $\exists$  : there exists,
    - $\forall$  : for all, *all, only, every*

# relationships between classes use quantifiers

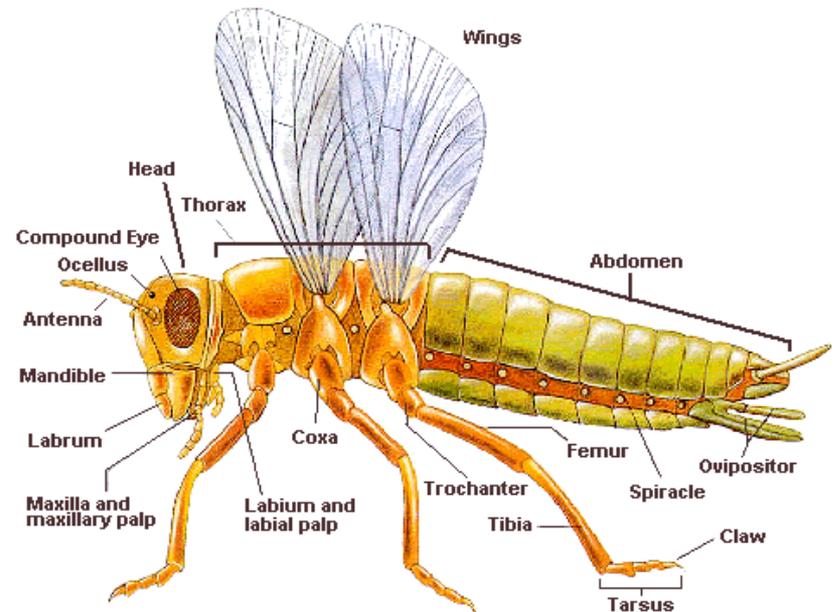
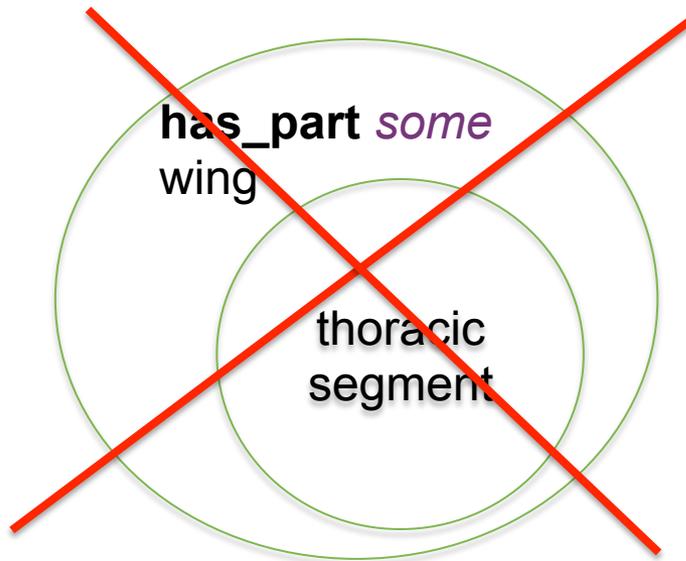
- OBO (**quantifiers hidden**)
  - name: leg
  - relationship: part\_of thoracic segment
- OWL (MS):
  - leg *SubClassOf* part\_of *some* 'thoracic segment'



# Directionality and quantifiers

'wing' *SubClassOf* **part\_of** *some* thoracic segment ✓

'thoracic segment' *SubClassOf* **has\_part** *some* 'wing' ✗



# Defining necessary and sufficient conditions for class membership

- English

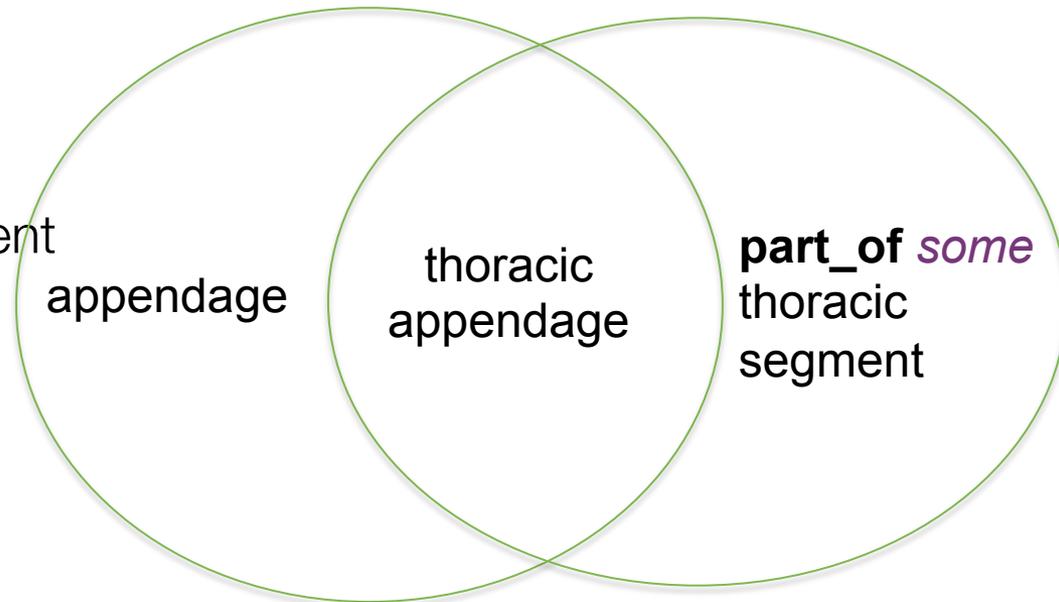
- Any appendage that is part of some thoracic segment is a thoracic appendage

- OWL

- thoracic appendage

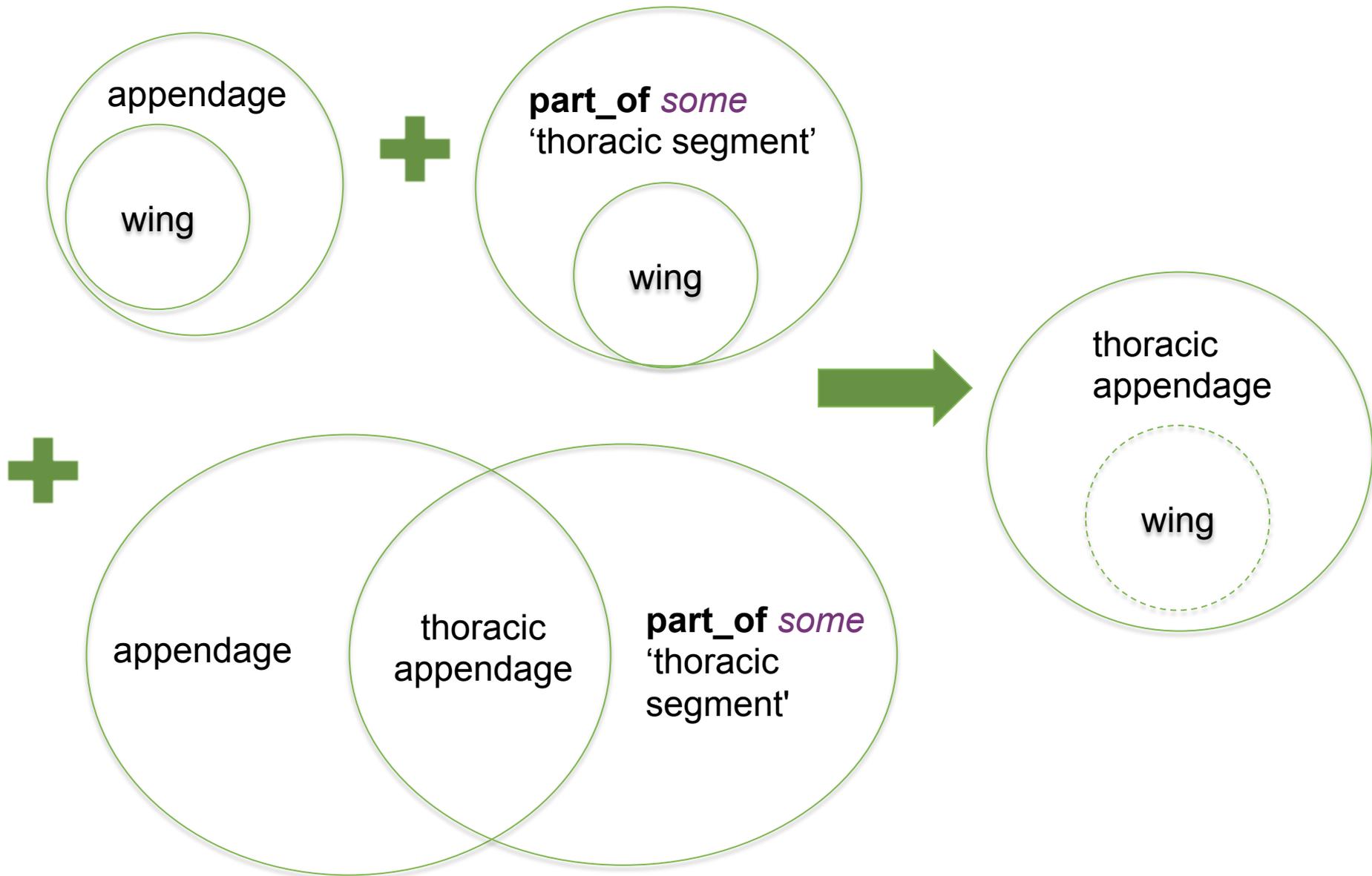
EquivalentTo 'appendage'

and part\_of some thoracic segment



- OBO

- name: thoracic appendage
- intersection\_of: appendage
- intersection\_of: part\_of thoracic segment



# How automated classification helps ontology building

- Record
  - one classification (is\_a)
  - relationships (necessary conditions for class membership)
- Reason:
  - To find what other classifications are applicable

+ve: It is very hard to find all relevant classifications by hand.

-ve: Only works if relations well defined and understood

# We don't need to make a class to express a concept in OWL

Query (class expression)

appendage **and** part\_of **some** 'thoracic segment'

Execute Add to ontology

Query results

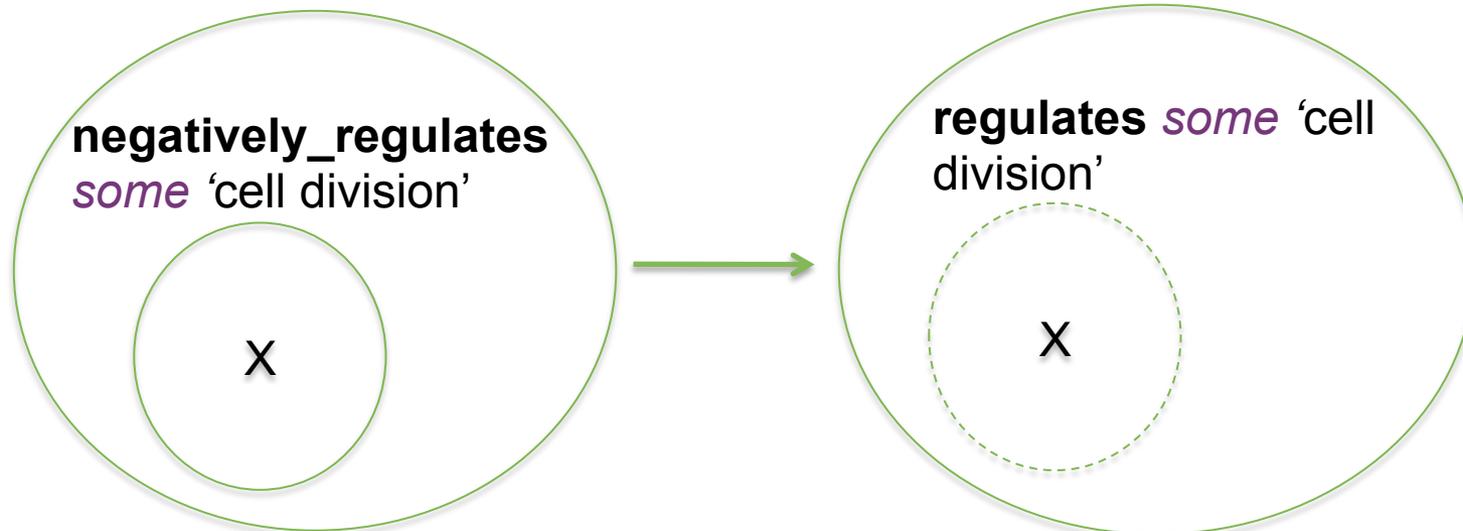
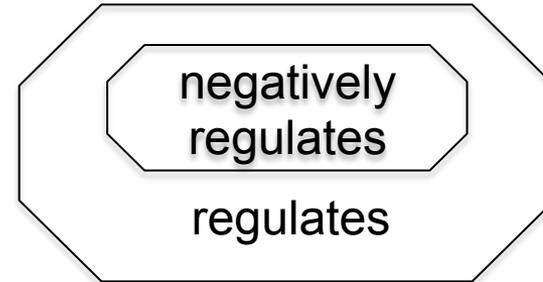
Descendant classes (6)

- **'mesothoracic leg'**
- **'metathoracic leg'**
- **'prothoracic leg'**
- **haltere**
- **leg**
- **wing**

- Just as in annotation extensions

# Some relations entail others

- regulates
  - 'negatively regulates'
  - 'positively regulates'



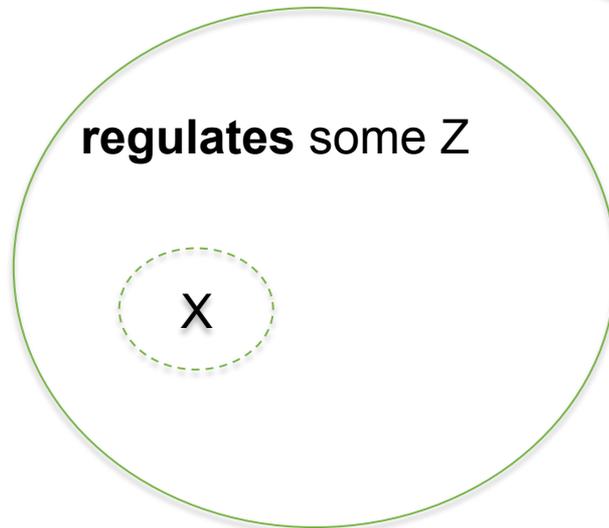
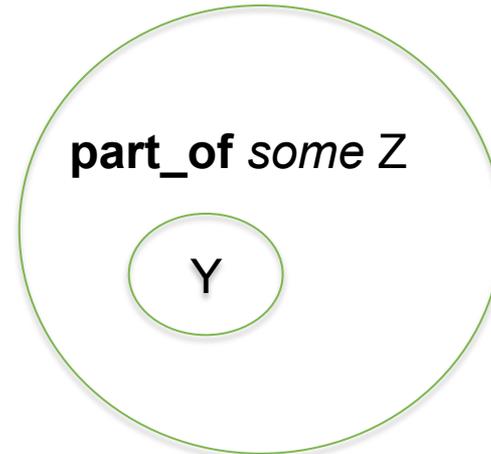
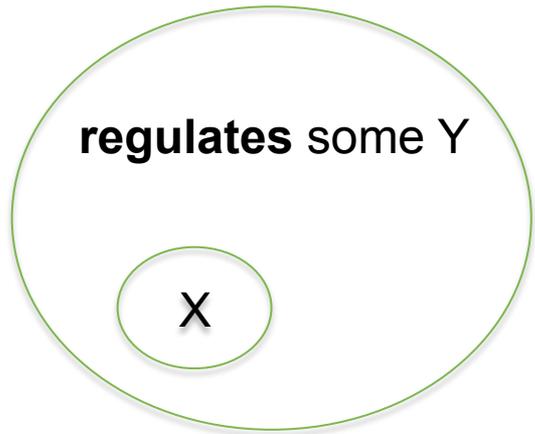
# Rules

*IF X regulates Y  
AND Y part\_of Z*  **X regulates Z**

Property chains 

**regulates**  **'part of'**  **SubPropertyOf** **regulates**

*IF X regulates Y  
AND Y part\_of Z*  $\longrightarrow$  **X regulates Z**



# Annotation extension -> OWL

Gene Product	GO term (c5)	C16
ABC1	apoptotic process	occurs in photoreceptor

Query (class expression)

'apoptotic process' and 'occurs in' some 'photoreceptor cell'

Execute Add to ontology

Query results

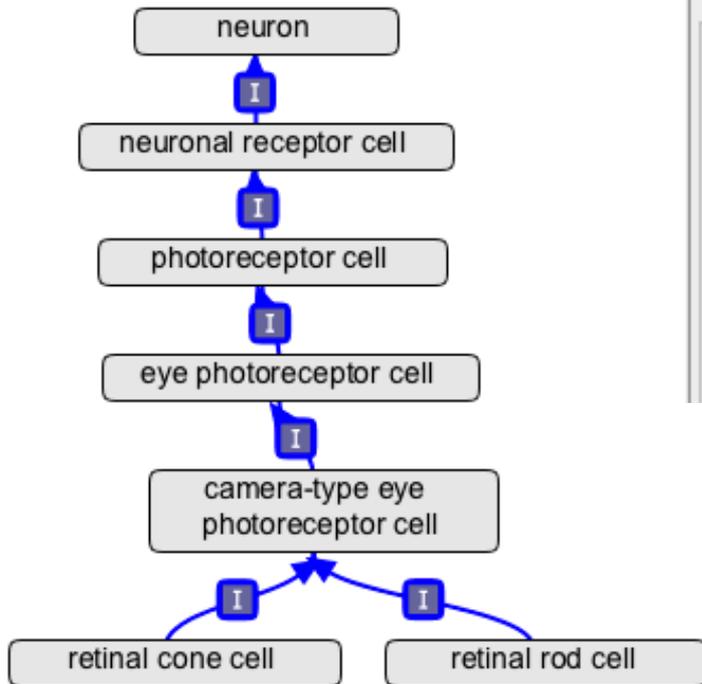
Super classes (1)

- 'neuron apoptotic process'

Descendant classes (2)

- 'retinal cone cell apoptotic process'
- 'retinal rod cell apoptotic process'

# Folding



Query (class expression)

'apoptotic process' and 'occurs in' some 'photoreceptor cell'

Execute Add to ontology

Query results

Super classes (1)

- 'neuron apoptotic process'

Descendant classes (2)

- 'retinal cone cell apoptotic process'
- 'retinal rod cell apoptotic process'

folding

- ▼ 'neuron apoptotic process'
  - ▼ 'apoptotic process and occurs in some photoreceptor cell'
    - 'retinal cone cell apoptotic process'
    - 'retinal rod cell apoptotic process'

# Unfolding

Gene Product	GO term (c5)	C16
ABC1	B cell apoptotic process	

Description: 'B cell apoptotic process'

Equivalent To  **'apoptotic process' and 'occurs in' some 'B cell'**

Gene Product	GO term (c5)	C16
ABC1	apoptotic process	occurs in B cell

# OWL translation of C16 with multiple clauses

Gene Product	GO term (c5)	C16
CASQ2	sequestering of calcium ion	occurs_in sarcoplasmic reticulum, occurs_in cardiac muscle cell

The screenshot shows a web interface for a DL query. At the top, there is a yellow bar labeled "DL query:". Below it, a section titled "Query (class expression)" contains the following text: "'sequestering of calcium ion' and 'occurs in' some 'sarcoplasmic reticulum' and 'occurs in' some 'cardiac muscle cell'". Below the query text are two buttons: "Execute" and "Add to ontology". Below the query section is a section titled "Query results" which shows "Super classes (2)":

- 'maintenance of location in cell' (with a question mark icon)
- 'sequestering of calcium ion' (with a question mark icon)

← Note – not nested

← Sometimes reasoning will give classifications you might not expect.

# OWL translation of C16 finding equivalent GO terms

Gene Product	GO term (c5)	C16
xyz10	apoptotic process	occurs_in some B cell

Query (class expression)

'apoptotic process' and 'occurs in' some 'B cell'

Execute Add to ontology

Query results

Equivalent classes (1)

- 'B cell apoptotic process'

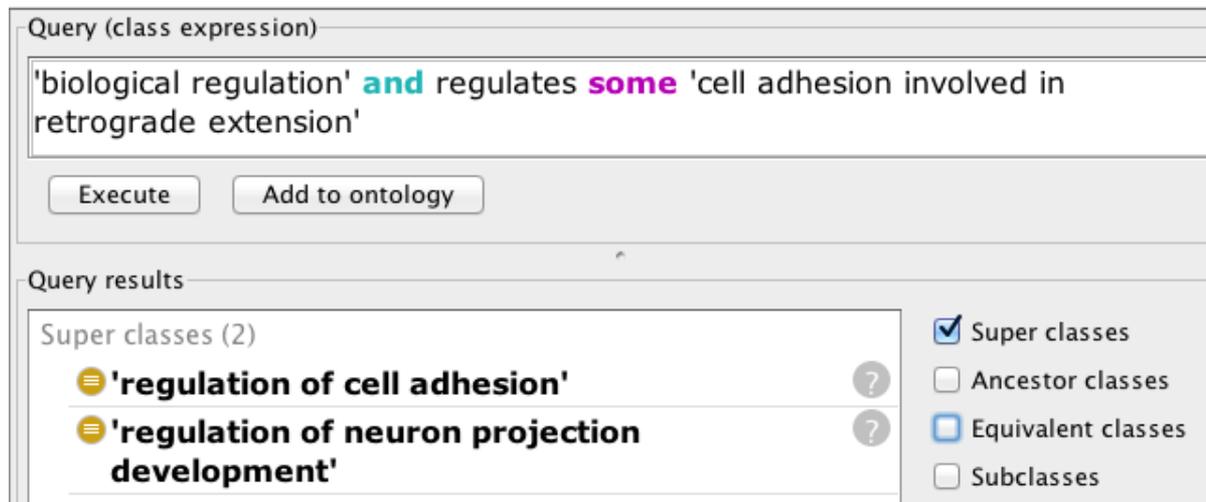
Super classes (1)

- 'lymphocyte apoptotic process'

← AE equivalent to existing term

# TermGenie examples

TG: regulation of cell adhesion involved in retrograde extension



The screenshot shows the TermGenie query interface. The query input field contains the text: "'biological regulation' and regulates some 'cell adhesion involved in retrograde extension'". Below the input field are two buttons: "Execute" and "Add to ontology". The "Query results" section displays two super classes: "'regulation of cell adhesion'" and "'regulation of neuron projection development'". To the right of the results are four checkboxes: "Super classes" (checked), "Ancestor classes", "Equivalent classes", and "Subclasses".

is\_a: GO:0010975 {is\_inferred="true"} ! regulation of neuron projection development

is\_a: GO:0030155 {is\_inferred="true"} ! regulation of cell adhesion

# The reasons for inferred classification can be complicated

Query (class expression)

'biological regulation' **and** regulates **some** 'cell adhesion involved in retrograde extension'

Execute    Add to ontology

Query results

Super classes (2)

- 'regulation of cell adhesion'
- 'regulation of neuron projection development'

Super classes  
 Ancestor classes  
 Equivalent classes  
 Subclasses

Explanation for: 'biological regulation' and (regulates some 'cell adhesion involved in retrograde extension') SubClassOf 'regulation of neuron projection development'

- 1) regulates o 'part of' **SubPropertyOf** regulates In ALL other justifications ?
- 2) 'retrograde extension' **SubClassOf** 'neuron projection development' In ALL other justifications ?
- 3) 'cell adhesion involved in retrograde extension' **SubClassOf** 'part of' **some** 'retrograde extension' In NO other justifications ?
- 4) 'regulation of neuron projection development' **EquivalentTo** 'biological regulation' **and** (regulates **some** 'neuron projection development') In ALL other justifications ?

# Using Protégé to find usage of relations

ary.org/obo/go.owl) : [/repos/go\_local\_ed\_git/workinggo.owl]

Search: acts\_on\_populatio

Case sensitive  Whole words  Ignore white space  Regular expression  Show all results

Search in IRIs  Search in annotation values  Search in logical axioms

Found in	Entity	Match
Display name	acts_on_population_of	acts_on_population_of

Usage: acts\_on\_population\_of

Show:  this  disjoints

- ▼ **'alpha-beta T cell proliferation'**
  - 'alpha-beta T cell proliferation' **EquivalentTo** 'cell proliferation' **and** (acts\_on\_population\_of some 'alpha-beta T cell proliferation')
- ▼ **'B cell homeostasis'**
  - 'B cell homeostasis' **EquivalentTo** 'homeostasis of number of cells' **and** (acts\_on\_population\_of some 'B cell')
- ▼ **'B cell homeostatic proliferation'**
  - 'B cell homeostatic proliferation' **EquivalentTo** 'cell proliferation' **and** ('part of' some 'homeostatic process') **and** (acts\_on\_population\_of some 'B cell')
- ▼ **'B cell proliferation'**
  - 'B cell proliferation' **EquivalentTo** 'cell proliferation' **and** (acts\_on\_population\_of some 'B cell')
- ▼ **'B-1 B cell homeostasis'**
  - 'B-1 B cell homeostasis' **EquivalentTo** 'homeostasis of number of cells' **and** (acts\_on\_population\_of some 'B-1 B cell')

# Take home messages

- If you understand the relations you use then the classification should look after itself.
- Except in the most trivial cases, you can't work out how an annotation extension will fold.
- Tooling support may be necessary to help understand the implications of annotation extensions

# Tooling support for Annotation Extensions

- A web tool to allow annotators to find how extensions will fold, check for equivalence, etc?

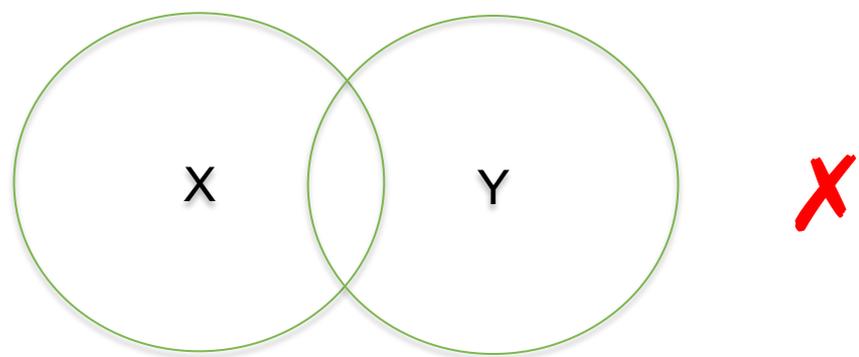
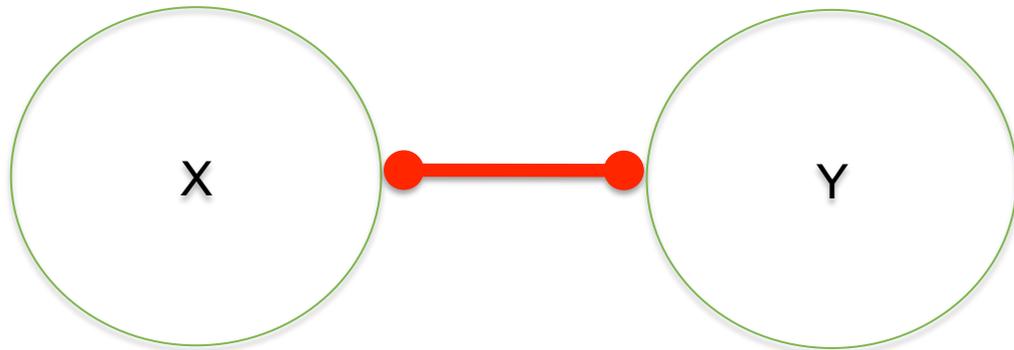
PROBABLY WON'T USE SLIDES FROM  
HERE ON

# ERROR MESSAGES ARE YOUR FRIENDS!

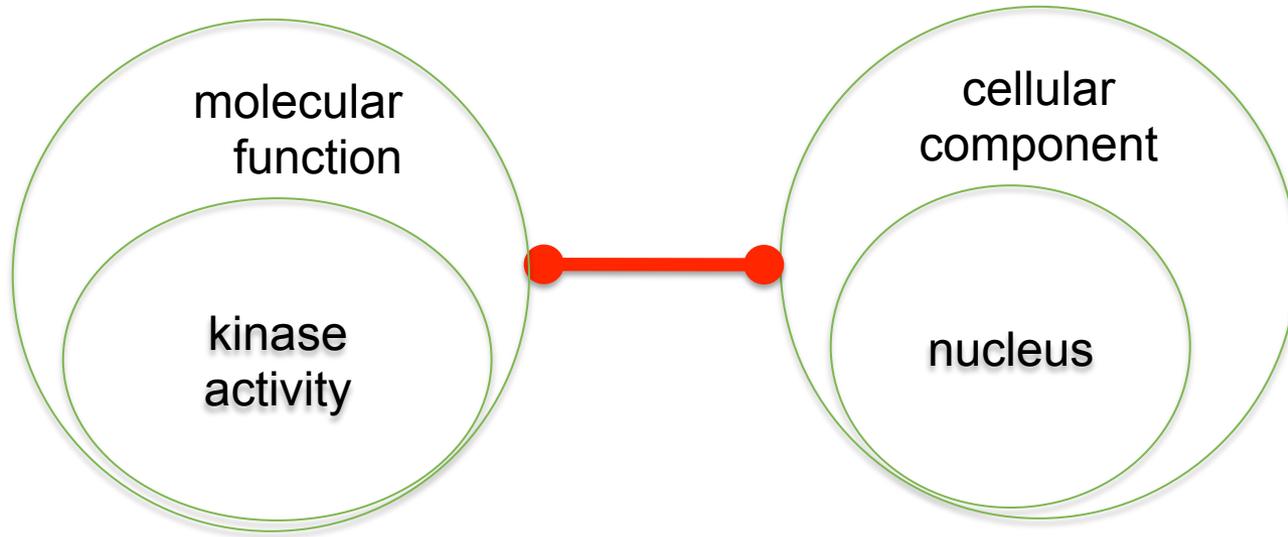
- They tell you you've screwed up before you get embarrassing emails complaining that you've screwed up

# Some classes don't intersect

OWL `DisjointWith` OBO: `disjoint_from` 



# Some classes don't intersect



nucleus *is\_a* kinase activity

nucleus *subClassOf* kinase activity

