Integral To

GO Annotation Call 2011-11-22

Outline

- How to interpret a GAF line
  - BP
  - CC
- How integral_to affects meaning
- Inference and has_part

The meaning of an annotation

c1	c2	c3	c5	ev
MGI	MGI:98297	Shh	GO:0060173 (limb development)	IMP

what does it ***say****?*

SYNTAX

what does it ***mean****?*

SEMANTICS

The meaning of an annotation

c1	c2	c3	c5	ev
MGI	MGI:98297	Shh	GO:0060173 (limb development)	IMP

what does it ***say***?

default relation for BP; made explicit in GPAD

some *Shh* product actively participates in ***some*** limb development process

what does it ***mean***?

General rule: annotations near the ontology root are *non-specific*

How do we say…

- Shh products are required for **all types of limb development** in mouse
  - Approximation:
    - annotate to subclasses
      - **forelimb** development
      - **hindlimb** development
  - What if there are many subtypes?
    - E.g. gene product required for all types of neuron development

Proposed qualifier to strengthen meaning

c1	c2	c3	c4	c5	ev
MGI	MGI:98297	Shh	**integral_to**	GO:0060173 (limb development)	IMP

what does it ***say***?

all limb development process (in mouse)
has active participant some ***some*** *Shh* product

inverts direction and relation + strengthens quantifier

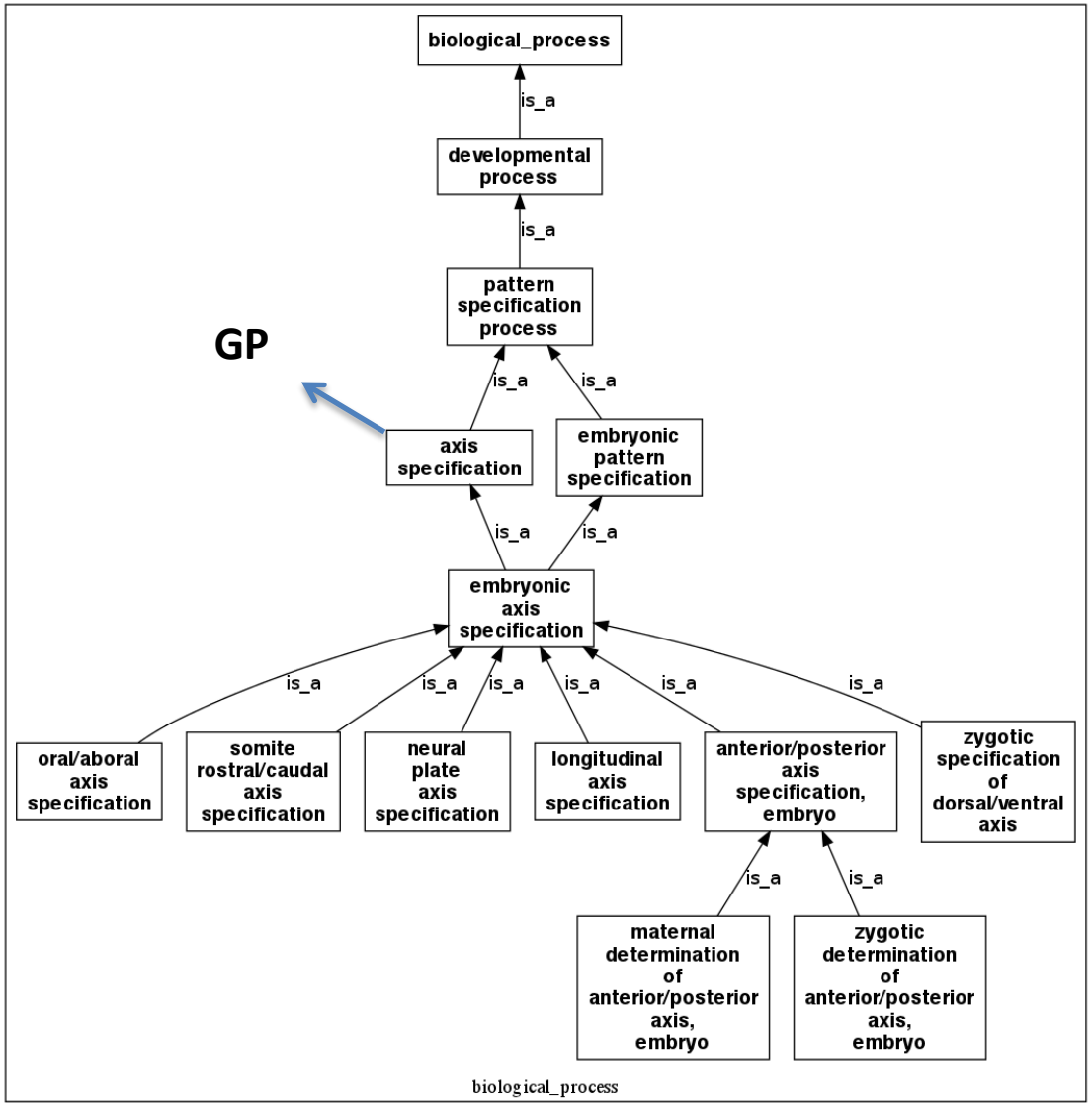
some *Shh* product actively participates in ***some*** hindlimb development process

some *Shh* product actively participates in ***some*** forelimb development process

what does it ***mean***?

What does this buy us?

- Annotations higher up the subclass (is_a) hierarchy are **stronger**
- Is this a big deal?
  - Can we not just make separate annotations to the subclasses?
    - hard sometimes
      - e.g. neuron development
      - axis specification



if a gene product is integral to axis specification, it is integral to all subclasses

CC example

c1	c2	c3	c5	ev
SGD	S000000119	MCM2	GO:0042555 (MCM complex)	IDA

what does it ***say***?

some *MCM2* product <u>is part of</u> ***some*** MCM complex

what does it ***mean***?

Stronger

c1	c2	c3	c4	c5	ev
SGD	S000000119	MCM2	**integral_to**	GO:0042555 (MCM complex)	?

what does it ***say***?

every MCM complex (in S cer) has part ***some*** *MCM2* product

some *MCM2* product is part of ***some*** MCM complex

what does it ***mean***?

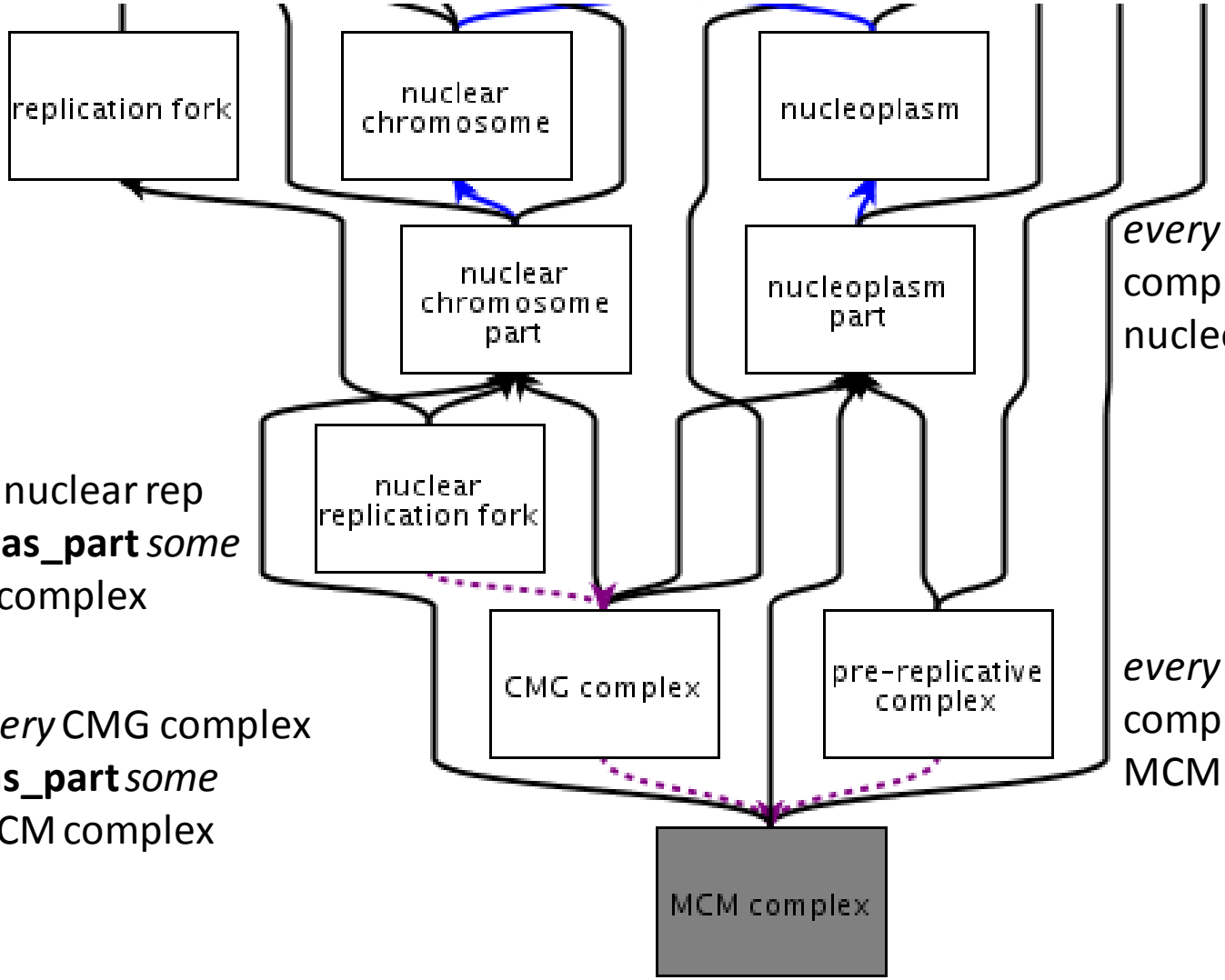
Why would we bother to say this?

- Answer:
  - it gives us more inferences because of has_part
    - eh? why?
- Conceptual leap:
  - think of annotations as being *part of the ontology*
  - think of the ontology graph *as being sentences*

Inference

- Input:
  - **every** CMG complex has_part **some** MCM complex [ontology]
  - **every** (Scer) MCM complex has_part **some** MCM2 product [annotation]
  - has_part is **transitive**
- Output:
  - **every** CMG complex has_part **some** MCM2 product

Is this a useful inference?

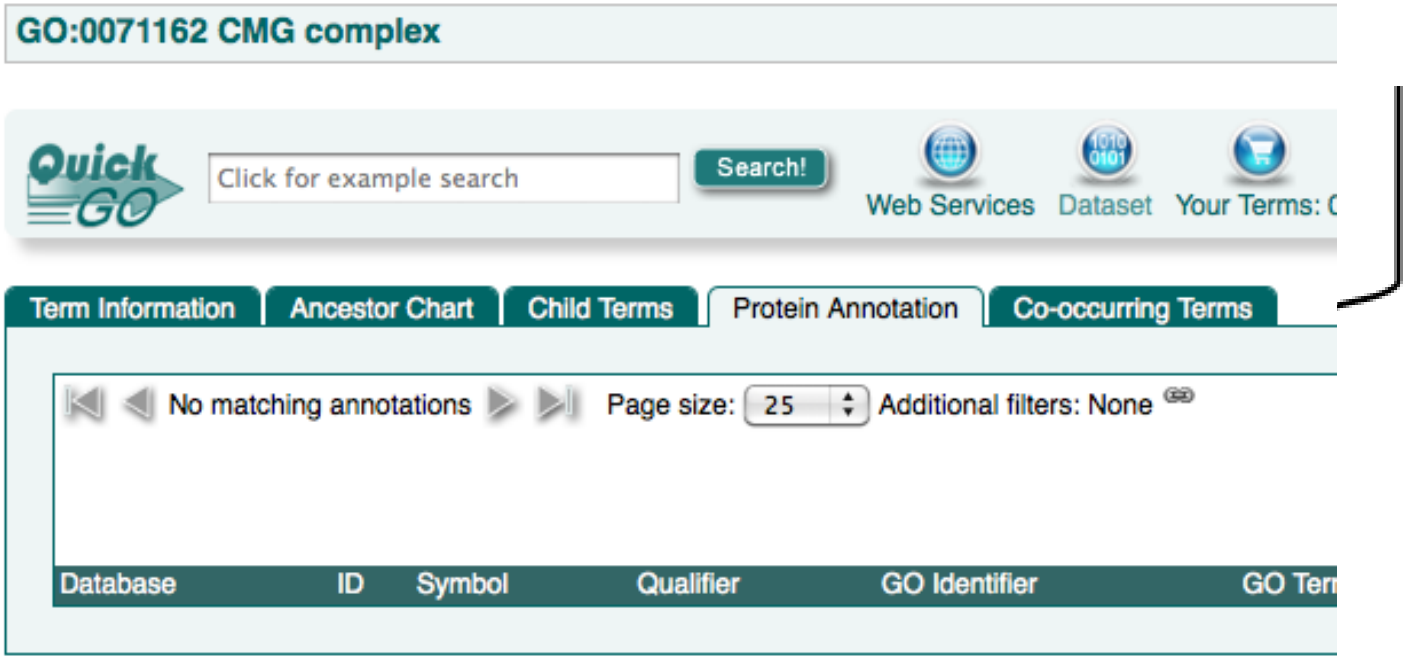


every pre-replicative complex **is part of** *some* nucleoplasm

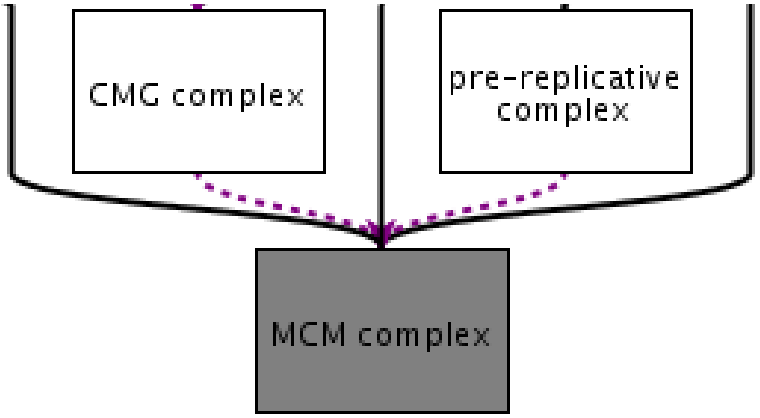
every nuclear rep fork **has_part** *some* CMG complex

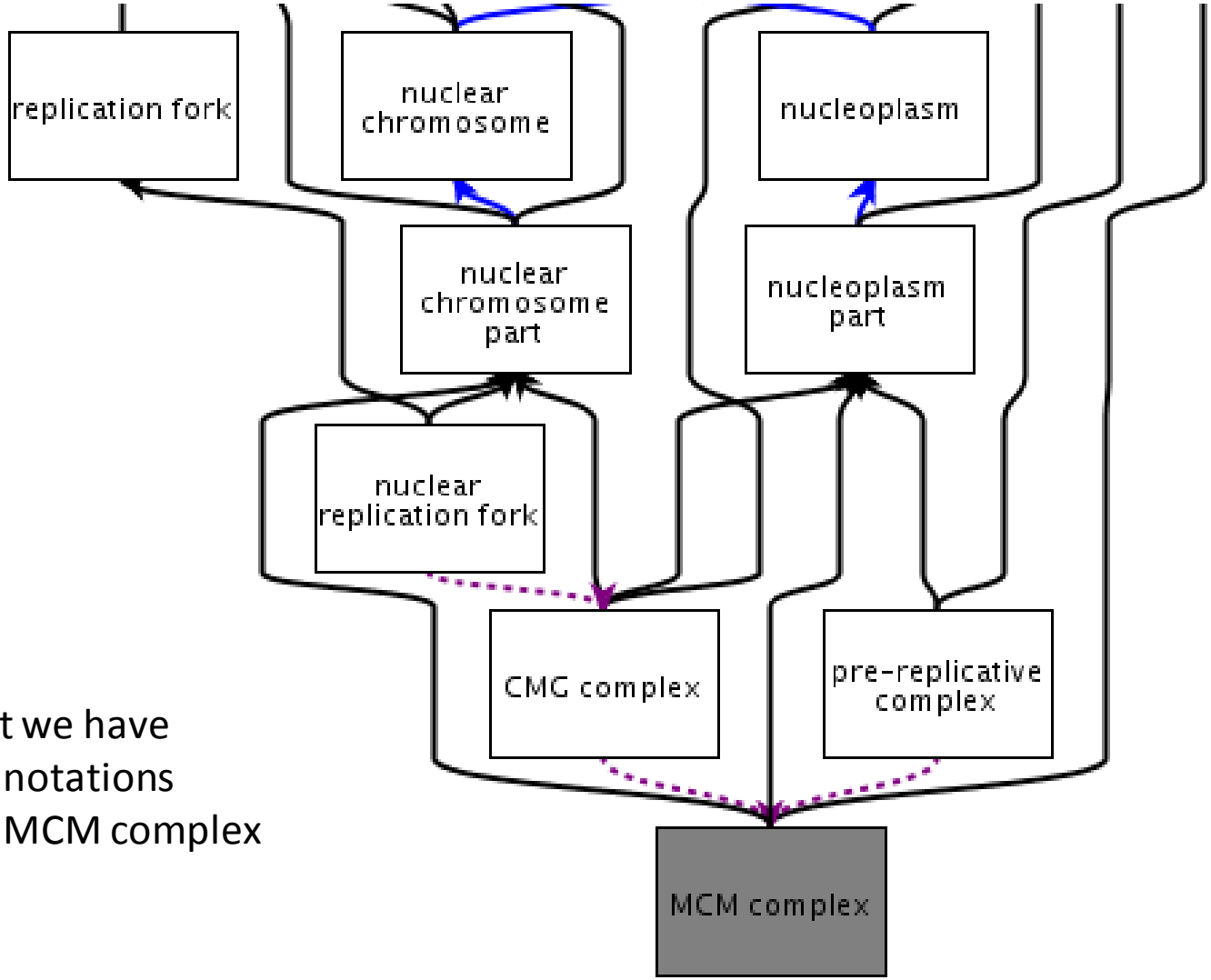
every pre-replicative complex **has_part** *some* MCM complex

every CMG complex **has_part** *some* MCM complex



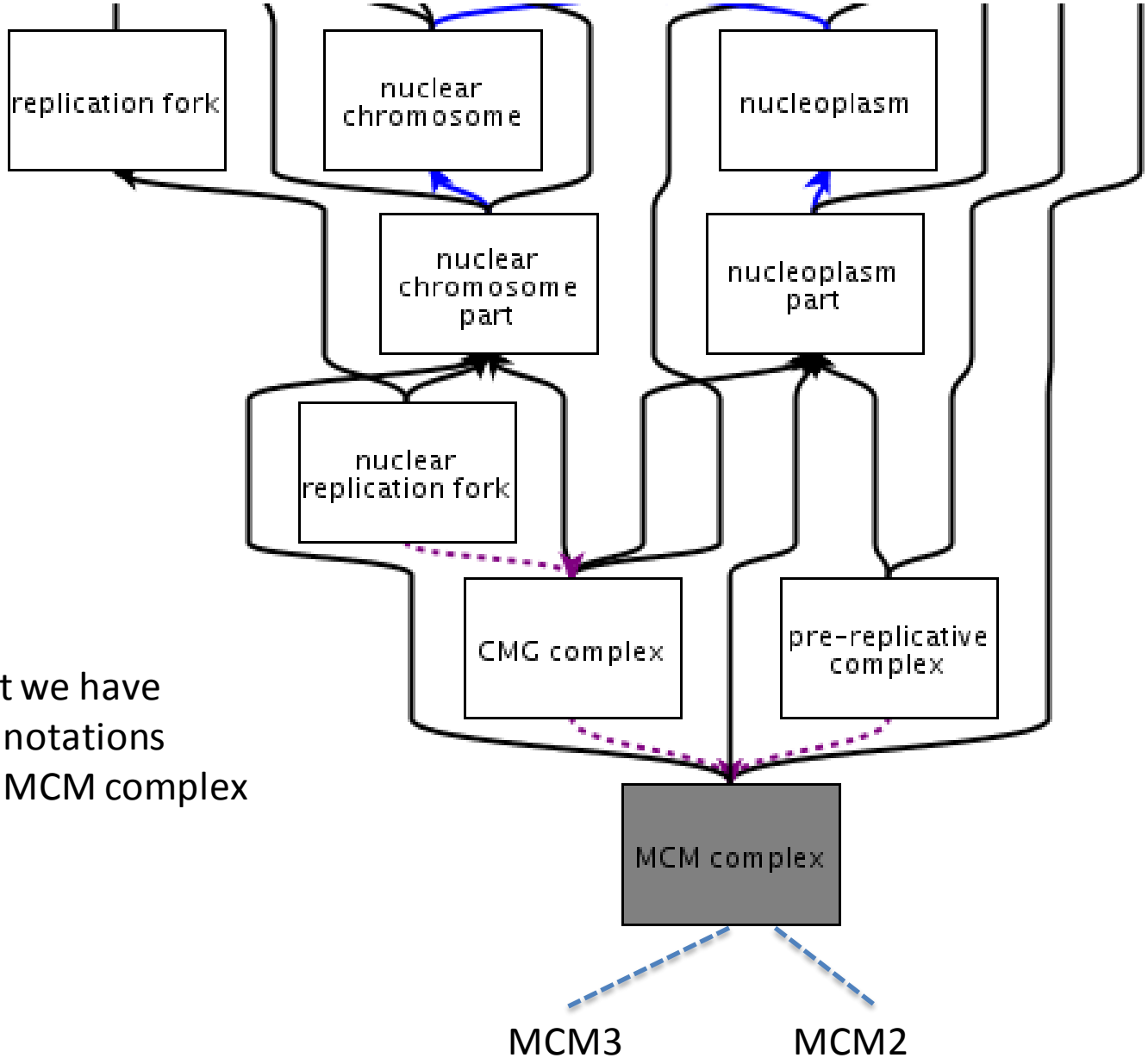
CMG complex
has no annotations visible



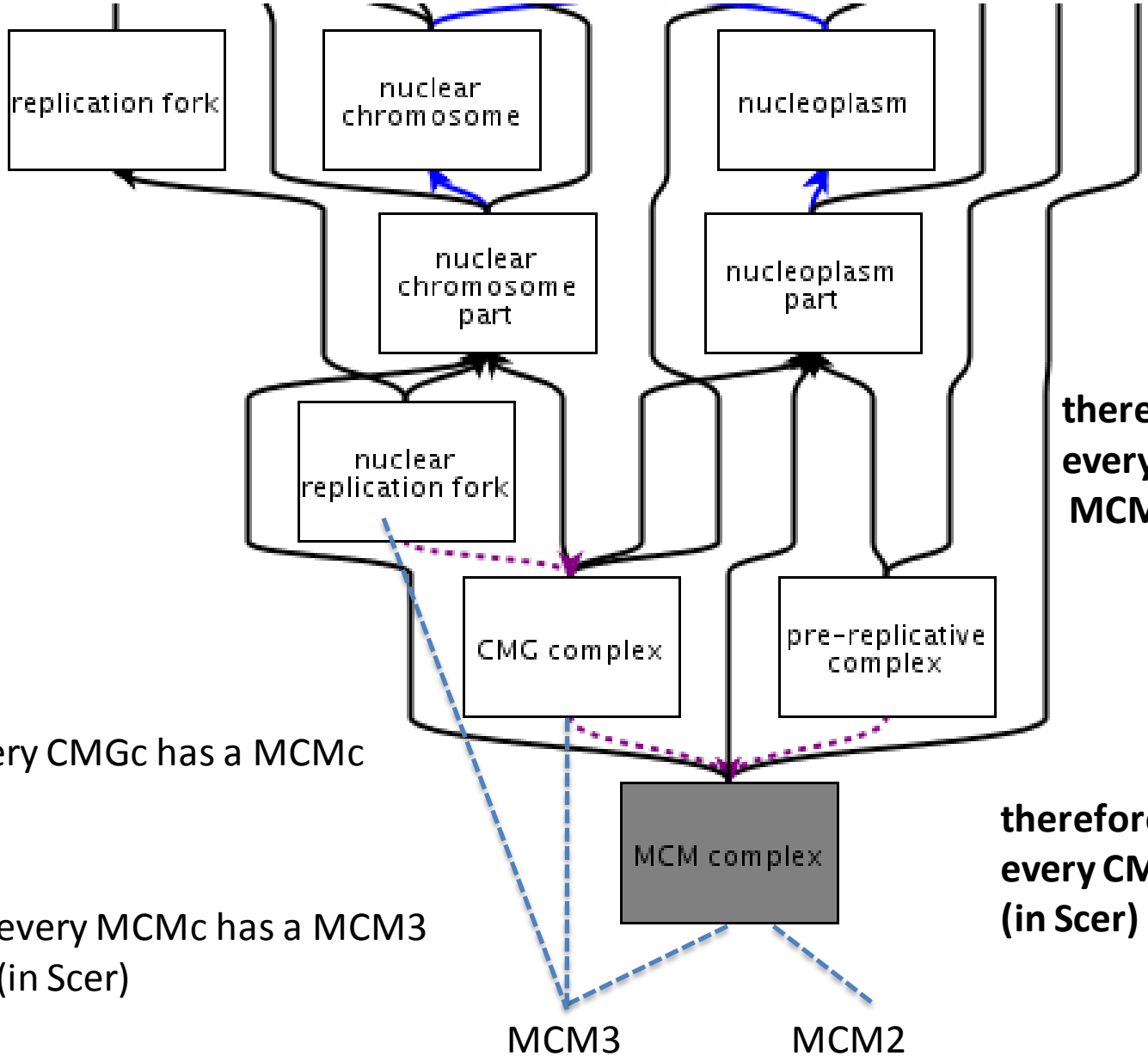


yet we have
annotations
to MCM complex

MCM3 MCM2



yet we have
annotations
to MCM complex



therefore:
every NRF has a MCM3 (in Scer)

every CMGc has a MCMc

therefore:
every CMGc has a MCM3 (in Scer)

every MCMc has a MCM3 (in Scer)

This works for BP too

- Relation chain:
  - if A has_part B
  - and B has_active_participant C
  - then A has_active_participant C

Seeding annotations

- Going back to strengthen old annotations is a lot of work
  - hard
    - how do you know for sure the process requires the gene product
    - easier for complexes – 'true by definition'
- For complexes, we can use other sources
  - E.g. PRO
- For processes, pathway databases
- Use qualifier moving forward

Conclusions

- Current annotations have some-some semantics
- Integral_to makes a strong statement about the complex/process, for that species
- This allows composition with has_part

OWL formalism

- Normal annotation:
  - <Ann> is_about some (<GP> and <R> some <X>)
- Annotation with integral_to qualifier
  - <Ann> is_about some (<GP> and <R> some <X>)
  - (<X> and in_taxon <T>) SubClassOf inverseOf(<R>) some <GP>