Completing the is_a tree for the cellular component ontology

Jane Lomax

The GO cellular and process ontologies do not currently have a complete is_a tree, that is, where every term has at least one path of is_a relationships back to the top node. We hope to create a full is_a tree for the cellular component ontology. There are several advantages to doing this: it would be ontologically correct, as logically everything is a type of something else; it would allow for more accurate queries and reasoning; it would work well with ontology tools other than OBO-Edit, which all assume is_a completeness in ontologies; and it would help simplify the visualisation of the tree by allowing alternation between is_a and part_of views. I present a version of the cellular component ontology with the full is_a paths, created by making several new high-level parent terms.

The ontology now, expanded at 'srb-mediator complex ; GO:0016592'.



Why do we need full is_a paths?

Ontologically correct

Logically, everything that exists is a sub-type of something else; this applies to all entities in GO.

More accurate queries and reasoning

At present, we are unable to easily make a query such as 'give me all the types of membrane' or 'give me all parts of a membrane'; adding these extra paths will allow us to do this. Or put another way, without these extra paths, GO is not complete.

Works well with other ontology tools

All ontology tools in the world, with the exception of DAG-Edit/OBO-Edit, assume that the is_a tree for an ontology is complete, so completing the is_a paths will allow GO to be used with other tools. It will also prevent external groups making incorrect assumptions about GO in order to make it work with other tools, such as replacing part_of relations with is_a.

Improved visualisation

Having complete is_a and part_of paths will allow us to make tools that allow you to move between is_a and part_of views. This is a much more intuitive way to view GO, and will avoid having to view the highly tangled, complicated mixed-relation view.





'intracellular organelle component' node expanded. Notice that most