OWL Representation of Drug Activities on Biological Systems

Author	**Samuel Croset**
Supervisors	Dietrich Rebholz-Schuhmann
Studies/Stage	PhD student, second year
Affiliation	European Bioinformatics Institute
E-Mail	croset@ebi.ac.uk

Aims and Objectives of the Research

The pharmaceutical industry is currently facing a decrease in productivity. Despite the increasing investments made in research and development, the number of newly approved drugs on the market is not growing. To overcome this issue, a strategy, drug-repurposing, aims at re-assigning already approved drugs towards new disease indications. This approach has witnessed numerous successes such as the safe rehabilitation of withdrawn drugs such as Duloxetine for a totally different use than the original one [1]. Repurposing opportunities can arise from fortunate situations or advantageous side-effects, but researchers are developing numerous methodologies to identify and predict potentially relevant cases in a systematic way. We are working on such a methodology, based on description logic. Our method aims at integrating knowledge from the molecular to the phenotypic level, with data coming from biological pathways repositories and Gene Ontology (GO) annotations. The method would offer a formal framework on which drug re-purposing hypotheses could be tested and explained via a reasoner running on the top of the integrated knowledge.

Justification for the Research Topic

Motivations behind the Description Logic approach: The vast majority of drug re-profiling approaches are based on previous knowledge stored in public databases. From this information, statistical methods can process the data and identify the significant re-purposing opportunities based on features association. It is however difficult to provide a pragmatic explanation behind the predictions made in order to support further improvements for the development of the drug. Moreover, the outcomes strongly depend on the choice of the features used [2]. A different approach, developed within the systems biology field, models the dynamics of a set of molecules known to be involved in a disease, in order to predict the action of a drug. Such techniques provide solid arguments to explain the mechanism of action of the compound, but are sensitive to missing knowledge or parameters. Moreover it is difficult to integrate additional data from the phenotypic level, as the modelling is often built upon chemical reactions only [2]. Our approach is complementary yet different to the ones previously described. It uses description logic to capture the rational thinking behind the design of a drug towards a disease. From a set of biological facts and based on a series of deductions, our framework could explain the logical reasons why a drug would influence a phenotype. Facts can be extracted from ontologies and large public repositories and inked together in order to create axioms. The axioms can describe any kind of level of abstraction; therefore the methodology can scale across various biological levels, from the molecule to the phenotypic trait. Open Biomedical Ontologies have a Web Ontology Language (OWL) representation, which enables the use of tools and reasoning engines coming from the semantic web community. An OWL reasoner running over the axioms would provide a proof of concept for the new effect of the drug on a particular phenotype, in an automated way.

Research Questions

What is the minimum and necessary set of axioms that enables the representation of the biological system?

Are there enough quality facts in biomedical databases in order to make useful deductions? What are the advantages brought by a reasoning engine?

How can we represent the effect of a drug in description logic? Can we prove the correctness of an indication for an active compound?

How can it be evaluated? Can we suggest new indications for already approved drugs?

Research Methodology

Activity abstraction layer: The medical value of a drug is tightly bound to the subsequent activity that it has on a biological body. In order to ease the integration of the various biological layers, we argue for a representation at a common semantic level, scaling up from chemicals to phenotypes. Considering the activity of the members of a system enables such a representation. In other words, our methodology does not represent a molecule for its physical structure, but considers rather the biological function carried by this molecule. We refer to the activity born by a compound by adding some curly brackets around the compound name. For instance the activity of a compound A is written {A} (Figure 1). The GO provides definitions and annotations which we re-use, in order to describe the abstract activity born by a biological entity.

Knowledge integration: The common level of representation of the actors involved in the system as activities eases the integration ask. Indeed, abstract molecular activities can be linked together using the positive/negative regulation relationships provided by the Relation Ontology (RO), in order to create triples. It is also possible to furthermore link phenotypic processes as described in the GO using these same relations. We argue that these two relations are intuitive and accurate enough to capture dependencies mong he molecular activities of a biological system. The triples, which we refer as axioms or facts, are generated from the content of pathway databases: We convert the chemical reactions representation into activity dependencies relationships (Figure 1). The connection of facts creates an activity map, here the nodes are activities/processes and the edges are relations from the RO.
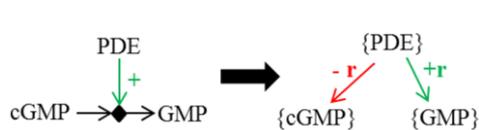
Figure 1: Conversion of an enzymatic reaction as represented in pathway databases into activity dependencies axioms, sing he positive/negative regulation from the RO.

Reasoning: The activity map is represented in OWL in order to re-use the reasoning frameworks already developed. The reasoner provides a mean to check the consistencies of the integrated facts that are building the activity map. It also computes the causal effects of drugs acting on the map: In biological systems, a drug disturbs first the activity of a molecular target, and then a perturbation wave follows, influencing the activities and processes linked to the original target. The reasoner renders this effect by changing the states of activity of the nodes of the activity map that are directly or indirectly connected to the drug target. The nodes are either becoming positively or negatively perturbed, according to the state of the previous nodes to which they are linked to, and to the type of edge chaining the two nodes (positive or negative regulation). From this series of logical steps, the reasoner can demonstrate and explain the causal influence an active compound has on a large set of integrated molecular activities and phenotypic processes.

References

1.   Thor, K. B. & Katofiasc, M. A. (1995), Effects of duloxetine, a combined serotonin and norepinephrine reuptake inhibitor, on central neural control of lower urinary tract function in the chloralose-anesthetized female cat, in: J. Pharmacol. Exp. Thera. 274.
2.   Sanseau P, Koehler J. (2011), Editorial: computational methods for drug repurposing, in: Brief Bioinform. 12(4)