

Substance concentrations as conditions for the realization of dispositions

Janna Hastings^{1*}, Christoph Steinbeck¹, Ludger Jansen² and Stefan Schulz³

¹ European Bioinformatics Institute, Hinxton, UK

² Department of Philosophy, University of Rostock, Germany

³ Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria

Abstract. Ontologies aim to represent what is general, by means of universal statements. In contrast, dispositional predications capture knowledge about what is likely to happen if a certain set of circumstances obtain, which is crucial in investigative research such as in drug discovery and systems biology, where entities which are constitutionally dissimilar can nevertheless have similar behavior in a biological context. While such dispositional properties are increasingly included in biomedical ontologies, the circumstances under which the dispositions are realized are seldom explicitly modeled, and doing so is problematic due to the necessary restriction to binary relations in OWL ontologies. In this paper we address this shortcoming, focusing on the bioactivity of small molecules at varying levels of concentration within a living organism as our problem domain, although our approach is generalizable to other problems. We discuss the ontological nature and representation of dispositions and their realization; consider the nature of concentrations and their representation; and finally we detail an approach to linking dispositions to the conditions for their realization which regards conditions as triggers for the process in which the disposition is realized.

1. Introduction

A fundamental tenet of ontologies in general and biomedical ontologies in particular is to make statements that are universally true. These statements are often considered to be statements about universals that, in turn, imply universally quantified statements about the instances of the universals involved (Smith, 2006) that is, they describe categorical properties. On the other hand, dispositional or functional properties are often contrasted with categorical properties because they specify what will occur *if* the correct circumstances obtain, which is a hypothetical property (Arp, 2008).

Dispositional or functional views on biomedical information are crucial in investigative research such as in drug discovery and systems biology, where entities which are constitutionally dissimilar can nevertheless have surprisingly converging behavior in a biological context; and *vice versa*. Thus we seem to face a dilemma: On the one hand, dispositional statements seem to be hypothetical rather than categorical, but on the other hand they are essential to a proper description of the biomedical domain, thus are increasingly being included in biomedical ontologies, examples of which are the ChEBI 'roles' (de Matos, 2010) and the Gene Ontology molecular functions (GO Consortium, 2000).

This dilemma is dissolved, we will argue, by recognizing that what is hypothetical in a disposition ascription is not the ascription of the disposition itself, but the expectation of its realization, which is conditional. The hard problem that remains is how to represent the realization conditions of a biomedical disposition, which are seldom explicitly included in biomedical ontologies. One reason for this may be the technical difficulty in adequately capturing the required nuances within a formalism allowing for only binary relations, such as OWL (Schulz, 2009).

In this paper we address this shortcoming, focusing on the bioactivity of small molecules at varying levels of concentration within a living organism as our problem domain, although our approach is generalizable to other problems. First, we discuss the ontological nature and representation of dispositions and their realization; then we consider the nature of concentrations and their representation. Finally, we explicitly link dispositions to the conditions for their realization.

2. Background

We will shortly present the biochemical and ontological background needed for our discussion.

2.1. Biochemical background

Small molecule bioactivity: Small molecules such as drugs or metabolites are essential ingredients in all the processes of life. The presence or absence of varying quantities of specific kinds of molecules can mean the difference between life and death. The biochemical mechanisms underlying the bioactivity are extremely varied and complex, although the basic mechanism is the binding (usually involving several non-covalent chemical interactions) of the small molecule to some organic macromolecular target. These mechanisms are regulated by the surrounding environmental conditions. For example, the transport of oxygen in the human bloodstream from the lungs to the cells where it is consumed is allosterically regulated, displaying sigmoidal behavior as a function of the concentration of the substrate (oxygen) (Berg 2002). The partial pressure of oxygen in the lungs is 100 torr, while that in the tissues is 20 torr. The change in oxygen pressure results in a change in the binding affinity of the oxygen, which results in a release of around 66% of the carried oxygen. The key is the *change in binding affinity* which depends on the *concentration* of the substrate.

2.2. Ontological background

Dispositions: In the aftermath of the verificationism of the logical positivist movement, dispositions have long been regarded as dubious or superfluous. Meanwhile, however, their importance for science has been rediscovered (Cartwright, 1989) and dispositions are again being discussed in Ontology (Mumford, 1998; Molnar, 2003). We will here follow the Basic Formal Ontology (BFO) (IFOMIS, 2010) and treat dispositions as dependent continuants. Continuants are entities that have no temporal parts, but exist as a whole at every moment of their existence. They are dependent because they need a bearer in order to exist (as all properties and relations do) (Arp, 2008; Jansen 2008). Dispositions are special realizables, that is, they are related to processes which are their realizations: dissolving in water is the realization of water solubility, and conducting electricity the realization of conductivity.

Concentrations: Concentrations are system properties, i.e. they are properties of a complex bearer, a mixture, and a concentration ascription describes how much of one ingredient (or fraction) is contained in the mixture. Like dispositions, concentrations cannot exist without a bearer; they, too, are dependent continuants. Concentrations are relational properties: a concentration is always a concentration of something in something, e.g., of alcohol in an alcohol-water mixture. Often, as in this case, the mixture in question is a solution, and the concentration in question is the concentration of the solute in the solvent. Concentrations relate two amounts of matter. In this aspect the ontological notion of concentration is stricter than the common sense concept of concentration where we say "100% alcohol" but refer to a pure substance and not a mixture, or where we characterize a non-alcoholic beverage by "0% alcohol concentration" but mean the complete absence of alcohol in the mixture.

3. Models

3.1. Dispositions

Dispositional properties can be viewed in two ways. On the one hand, dispositions point to their realization in the future, which is only hypothetical. For example, considering the disposition of aspirin to treat pain, we find that there are many molecules of aspirin which never treat any pain and many instances of pain which are never treated by aspirin. On the other hand, dispositions are part of the present state of the things they are ascribed to. And in so far as they are present properties of their bearers they are neither hypothetical nor a matter of probability only. It is not the disposition, but its realization which is hypothetical. We can formalize this dispositional property of aspirin in terms of its realization along the following lines:

$$\text{PortionOfAspirin} \sqsubseteq \exists \text{ bearerOf.}(\text{Disposition} \sqcap \\ \forall \text{ hasRealization.}(\text{Treating} \sqcap \exists \text{ hasParticipant.Pain}))$$

Using this pattern we express that for each particular portion of aspirin there is a particular disposition of a certain kind. This disposition is then described by constraining the kind of process by which it can be realized. But, a portion of a dozen aspirin molecules is not sufficient to treat pain in any organism. Furthermore, a

hundred 300mg tablets ingested at the same time may cause complications such as severe bleeding and intoxication before ever treating pain. Yet in the formula above, these conditions are not described. We will address this point later, as we now turn to a model of concentrations.

3.2. Concentrations

Let us take a simple example referring to instances of substance portions, e.g. a particular mixture of 10g of water with 10g of glucose. In discussing this example, we will use lower case letters for particulars and initial capitals for universals. We have three entities of interest:

(i) the water/glucose mixture wg_{mix} , (ii) the water fraction w_{coll} , i.e. the collection of all water molecules, and (iii) the glucose fraction g_{coll} , i.e. the collection of all glucose molecules. Following Schulz (2006), we distinguish between molecule collections (homogeneous pluralities of molecules of the same kind) and compounds (entities that are defined as sums of non-overlapping sortally distinct parts). Mixtures are a special case of compounds, and the fractions are their non-overlapping but maximally mixed components.

w_{coll} and g_{coll} are components of wg_{mix} ; wg_{mix} is the mereological sum of w_{coll} and g_{coll} .

These three particulars bear the following qualities:

- They have a mass (wg_{mix} 20g; w_{coll} and g_{coll} 10g each);
- They have a defined number of molecules (cardinality);
- Only wg_{mix} has a defined volume, as w_{coll} and g_{coll} are scattered objects.

We turn to the question: what are the concentrations and which entities are the bearers? Note that there are different kinds of concentration, the most important in biology and medicine being:

- mass percentage: mass of g_{coll} / mass of wg_{mix}
- mole fraction: number of molecules in g_{coll} / number of molecules in wg_{mix}
- mass/volume percentage: mass of g_{coll} / volume of wg_{mix}

In all cases we have

- a portion of a mixture of a kind (here wg_{mix}), which bears qualities like mass, volume, temperature;
- fractions which are component of this mixture (here g_{coll} and w_{coll}), but which are not mixture themselves, and which bear qualities like mass (but not volume);
- concentration of fractions in mixtures, e.g. g_{coll} in wg_{mix} .

To formalize this ontologically, we observe that all particular collections of glucose molecules instantiate the class G_{coll} , and have granular (repeated multitudinously) parts which instantiate G :

$$\begin{aligned}
G &\sqsubseteq \text{EntireMolecule} \\
G_{coll} &\sqsubseteq \text{HomogeneousCollection} \\
G_{coll} &\equiv \exists \text{ hasGranularPart}.G \sqcap \forall \text{ hasGranularPart}.G
\end{aligned}$$

For a mixture with several components, its "fractions":

$$\begin{aligned}
WG_{mix} &\sqsubseteq \text{Mixture} \\
WG_{mix} &\sqsubseteq = 1 \text{ hasComponent}.G_{coll} \sqcap = 1 \text{ hasComponent}.W_{coll}
\end{aligned}$$

where each component is a distinct fraction of the mixture. A concentration can be ascribed to a particular homogeneous collection *iff* this collection is a component of a mixture, as expressed by the axiom:

$\exists \text{ bearerOf}.Concentration \equiv \text{Homogeneous collection} \sqcap \exists \text{ componentOf}.Mixture$
with

$Concentration \sqsubseteq \exists \text{ inheresIn}.(HomogeneousCollection \sqcap \exists \text{ componentOf}.Mixture)$

The class *Concentration* can then further be specified in terms of the kind of concentration as explained above, e.g.

$$\begin{aligned}
VolumeConcentration &\sqsubseteq Concentration \\
MassConcentration &\sqsubseteq Concentration
\end{aligned}$$

as well as in terms of the participating substance portions:

$$\begin{aligned}
BloodGlucoseVolumeConcentration &\equiv VolumeConcentration \sqcap \\
&\exists \text{ inheresIn}.(PortionOfGlucose \sqcap \exists \text{ componentOf}.PortionOfBlood)
\end{aligned}$$

This states that wherever there is a blood glucose volume concentration there must be a portion of glucose and a portion of blood. In contradistinction to the glucose/water example we here have blood as an overly complex mixture with probably tens of thousands of fractions. The example demonstrates, however, that the exact composition of the mixture does not need to be specified. Quantitative measures of concentrations, as referred to in common discourse, are attributes of instances of *Concentration*.

3.3. Conditional realization of dispositions

We now have a formalism for defining dispositions, in terms of the process in which they are realized; and concentrations, in terms of the substances from which they are composed. We here attempt to relate the latter as a precondition for the realization of the former. We need to create a relationship between a disposition, a condition (concentration), and a realization in a biological process. This cannot be straightforwardly represented in OWL as the required relationship is ternary rather than binary. Indeed, if the probability of the realization of the disposition is also made explicit, the resulting relationship is quaternary: see the *Has_realization_under_conditions_with_probability* relation introduced in Schulz & Jansen (2009). The challenge at this point is to do justice to the ontology of

dispositions, within the restrictions of OWL. In section 3.1, we suggested the following example of a disposition ascription:

$$\begin{aligned} \text{PortionOfAspirin} \sqsubseteq \exists \text{ bearerOf.}(\text{Disposition} \sqcap \\ \forall \text{ hasRealization.}(\text{Treating} \sqcap \exists \text{ hasParticipant.Pain})) \end{aligned}$$

In section 3.2, we were able to define the volume concentration of aspirin in the blood along the following line:

$$\begin{aligned} \text{BloodAspirinVolumeConcentration} \equiv \text{VolumeConcentration} \sqcap \exists \text{ inheresIn.} \\ (\text{PortionOfAspirin} \sqcap \exists \text{ componentOf.PortionOfBlood}) \end{aligned}$$

We know that a certain concentration of aspirin in the blood is necessary in order to have the pain relieving disposition of aspirin realized. We can revise our suggestion from section 3.1 in order to incorporate this fact by using our new way to express concentrations plus a new relation *hasTrigger*. While a full discussion of this relation would require more space than we have available here, for the purpose of this study it is sufficient to interpret a trigger rather simplistically as a circumstance without which a process cannot occur. Combining these tools, we get:

$$\begin{aligned} \text{PortionOfAspirin} \sqsubseteq \exists \text{ bearerOf.}(\text{Disposition} \sqcap \forall \text{ hasRealization.}(\text{Treating} \sqcap \\ \exists \text{ hasParticipant.Pain} \sqcap \exists \text{ hasTrigger. SufficientConcentration})) \end{aligned}$$

where, of course, *SufficientConcentration* \sqsubseteq *BloodAspirinVolumeConcentration*.

4. Discussion

The problem of the ontology of dispositional properties is not a new one, although its relevance to biomedical informatics is recent. Many representatives of the formal ontology community defend the perspective that representations of non-categorical properties lie at the borderline or outside the realm of ontology (Rector, 2008; Schulz, 2009) and emphasize that the current representational formalisms such as OWL are not well suited to express modal or probabilistic knowledge and lead to unintended models if used to represent, say, the knowledge that a disease *X* *may* have the symptom *Y*, or that a molecule *A* *tends* to interact with a molecule *B* (Schulz, 2010). Others advocate the inclusion of dispositions and tendencies in their ontologies (Schulz & Jansen, 2009; Jansen, 2007), and our approach aligns with the latter. Key to our approach is the analysis of dispositional properties as being *necessarily* realized if the correct circumstances obtain, thus allowing us to reformulate the circumstances as a trigger for the process of realization.

Dispositional properties are closely related to probabilistic knowledge representation. The expression of such probabilistic knowledge within OWL ontologies has been discussed by Rector et al. (2008), who propose several workarounds including in particular the introduction of an explicit construct for ‘may’ into the language syntax and semantics to accommodate relations of possibility, although it is still not clear how the *magnitude* of a probability should be captured in their proposed formalism.

Probabilistic logics such as those described in Lucasiewicz (2008) provide a much more expressive formalism for this type of knowledge, but at the expense of additional complexity which may not be acceptable to the ordinary domain scientists who create and/or make use of biomedical ontologies.

Our analysis has been in the context of molecular bioactivity and concentrations, but could easily be extended to dispositional properties and conditions in general. More challenging will be the extension to a general treatment of conditions for ontological assertions, as the truths of most domain descriptions in biomedical science can be regarded as contextual truths which apply under certain circumstances. Such contextual truths include the composition and arrangement of bodily organs in organisms (the circumstances here are "normality" or "canonicity", and additionally "health" (Schulz & Hahn, 2007)), and the shape of molecular entities (where the circumstances include temperature, pressure, and environment).

5. Conclusion

Much recent work in biomedical ontology has focused on clarifying the top-level distinctions between kinds of entities in ontologies. Our work focuses on one particularly problematic kind of entity, *viz.* dispositional properties, which require a particular set of circumstances to obtain in order to be realized, regarding the relevant circumstances as a necessary trigger for realization. We provide an ontological analysis of concentrations as one kind of circumstance.

We see this work as a contribution to the analysis of dispositions and in particular to the explicit formalization of the conditions under which dispositions are realized. Future work will explore the representation of conditional properties of biomedical objects beyond dispositional properties, and extend our strategies to other triggering circumstances, like temperature, (blood) pressure or infections.

Acknowledgements

This work was supported by (i) the BBSRC, grant agreement number BB/G022747/1 within the "Bioinformatics and biological resources" fund; and (ii) by the DFG, grant agreement number JA 1904/2-1, SCHU 2515/1-1 GoodOD (Good Ontology Design).

References

- Arp, R., Smith, B. (2008) Function, Role and Disposition in Basic Formal Ontology. In *Proc. of the 11th Annual Bio-Ontologies SIG meeting, July 20, 2008, Colocated with ISMB 2008*, Toronto, Canada.
- Berg, J.M., Tymoczko, J.L., Stryer, L. (2002) *Biochemistry, fifth edition*. W.H Freeman and Company, New York 10010.
- Cartwright, N. (1989) *Nature's capacities and their measurement*, Oxford: Clarendon Press.
- de Matos, P., Alcántara, R., Dekker, A., Ennis, M., Hastings, J., Haug, K., Spiteri, I., Turner, S., and Steinbeck, C. (2010). Chemical entities of biological interest: an update. *Nucleic Acids Research* **38** (Database issue):D249-D254.

- Gene Ontology Consortium (2000) Gene ontology: tool for the unification of biology. *Nature Genetics* **25** (1):25-29.
- IFOMIS (2010) <http://www.ifomis.org/bfo>.
- Jansen, L. (2008) Classification, in: Munn, K., Smith, B. (eds.), *Applied Ontology*, Frankfurt/Lancaster: Ontos 2008, 159-172.
- Jansen, L. (2007) Tendencies and other Realizables in Medical Information Sciences. *The Monist* **90**, 534-555.
- Lukasiewicz, T. (2008) Expressive probabilistic description logics. *Artificial Intelligence* **172** (6-7): 852-883.
- Mumford, S. (1998) *Dispositions, 2nd ed.*, Oxford: Clarendon Press.
- Molnar, G. (2003) *Powers. A Study in Metaphysics*, Oxford: Oxford University Press.
- Rector, A., Stevens, R., Drummond, N. (2008) What Causes Pneumonia? The Case for a Standard Semantics for “may” in OWL. OWLED 2008.
- Schulz, S., Beisswanger, E., Hahn, U., Wermter, J., Kumar, A., Stenzhorn, H. (2006) From GENIA to BIOTOP - Towards a Top-Level Ontology for Biology. FOIS 2006: 103-114
- Schulz S., Jansen, L. (2009) Molecular interactions: On the ambiguity of ordinary statements in biomedical literature. *Applied Ontology* **4** (1): 21-34.
- Schulz, S., Stenzhorn, H., Boeker, M., Smith, B. (2009) Strengths and limitations of formal ontologies in the biomedical domain. *RECIIS* **3** (1): 31-45.
- Schulz, S., Schober, D., Tudose, I., Stenzhorn, H. (2010) The Pitfalls of Thesaurus Ontologization – the Case of the NCI Thesaurus. Submitted to *AMIA 2010*.
- Schulz, S., Hahn, U. (2007) Towards the ontological foundations of symbolic biological theories. *Artificial Intelligence in Medicine* **39** (3): 237-250.
- Smith, B., Kusnierczyk, W., Schober, D., Ceusters, W. (2006) Towards a reference terminology for ontology research and development in the biomedical domain. *Proc. of KR-MED 2006*, Baltimore, USA, 57-66.