

Adding Logic to the Toolbox of Molecular Biology

Giovanni Boniolo · Marcello D’Agostino ·
Mario Piazza · Gabriele Pulcini

the date of receipt and acceptance should be inserted later

Abstract The aim of this paper is to argue that logic can play an important role in the “toolbox” of molecular biology. We show how biochemical pathways, i.e., transitions from a molecular aggregate to another molecular aggregate, can be viewed as deductive processes. In particular, our logical approach to molecular biology — developed in the form of a natural deduction system — is centered on the notion of Curry-Howard isomorphism, a cornerstone in nineteenth-century proof-theory.

1 Introduction

The last few decades have seen a strong wind of pluralism blowing across logic and philosophy of science ([Cook, 2010], [Beall and Restall, 2006], [Kellert *et al.*, 2006], [Giere, 2006]). In logic the shift from mathematics to computer science as the main source of problems and paradigmatic applications has brought about a dramatic change of attitude. The center stage is no longer occupied by classical logic, with its static and timeless declarative units (“propositions”), but a proliferation of logical systems — more attuned to the dynamic entities (“states”, “actions”) that dwell in artificial intelligence and programming — has come under the spotlight. In these areas a broader view of what constitutes

Giovanni Boniolo
Department of Health Sciences, University of Milan,
European Institute of Oncology (IEO)

Marcello D’Agostino
Department of Economics and Management, University of Ferrara

Mario Piazza
Department of Philosophy, University of Chieti-Pescara

Gabriele Pulcini
Centre for Logic, Epistemology and History of Science – State University of Campinas

a logical system [Gabbay, 1993; Gabbay, 1994] has offered more natural formalizations of human reasoning and has paved the way for a principled logical approach to the design and analysis of computer artefacts, over and beyond the tampering with classical logic typical of earlier efforts. On the other hand, in the philosophy of science the need has emerged for “new ways of writing about science that are less beholden than the old ways to the epistemological mores of theoretical physics, and more faithful to the actual practices not only of physics but of all of the natural sciences” [Keller, 2000, pp. S72–S73]. This need appears to be particularly urgent in certain fields, such as molecular biology, where “no division between theory and experiment has obtained, and where distinctions between representing and intervening, and more generally, between basic and applied science, are daily becoming more blurred” [Keller, 2000, p. S73].

Moreover, the growing practice of computer simulation in most research areas raises new challenging methodological issues. Computational models are being increasingly used to represent and analyse the implications of complex interactions and even when the general laws of such interactions are well understood (as may be the case in theoretical physics), the systems of equations that these laws determine in a specific application context are often intractable. This prompts for the necessity of simplifications and approximations in order to obtain tractable computational models. It has been argued that, under these circumstances, we should perhaps view the computational models themselves as first-class citizens, as genuine “attempts to provide — directly — representations of real systems, and not abstract models” [Winsberg, 2001, p. S450]. In this perspective, the challenges posed by the widespread use of computer simulations in the analysis of complex systems would be best addressed by a philosophy of science that places “less emphasis on the representational capacity of theories (and ascribes that capacity instead to models) and more emphasis on the role of theory in guiding, rather than determining, the construction of models” [Winsberg, 2001, p. S443].

This situation requires an approach that ascribes models a theoretical dignity independent of the blessing of a theory in the traditional sense. As Nancy Cartwright once put it: “I use the image of the tool-box of science to describe a kind of instrumentalism that I defend as a part of this movement to undermine the domination of theory” [Cartwright *et al.*, 1995, p. 138]. This is all the more appropriate when, as is the case of molecular biology, the scarcity of general strict laws¹ naturally suggests the view that a “theory” is “less a search for universal and abstract principles, than a more grounded theorizing, of a kind that is required to make sense of local experimental manipulations” [Keller, 2000, p. S73]. In theoretical physics, the rules governing state transitions in a model are somehow “deduced” from universally quantified statements via ad hoc simplifications and approximations that usually have a physical interpretation and are introduced for the sake of mathematical and computational

¹ Indeed, there is still a vivid debate on the existence of laws of nature in biology. For an introduction to this issue see [Hamilton, 2007]. For the last perspectives, see [Haufe, 2013; Raerinne, 2013].

tractability.² The transition rules may change depending on the complexity of the models and on how “close” they are to the real systems under investigation, but the universal laws from which they are ultimately derived are supposed to remain unchanged and their form is by no means dependent on the specific *context* to which they are applied. This is the picture given by the classical nomological-deductive model of explanation based on “covering laws”, for which classical first order logic may, *prima facie*, appear as a suitable representational tool. Yet, it is well-known that it is hard to find such covering-law explanations in life sciences.³ Here “most research has a different objective than subsuming phenomena under universal laws” and is instead “directed at revealing the particular processes at work in a given system (e.g., the particular substrates and enzymes involved in glycolysis or the particular operations performed in processing information)” [Bechtel, 1998, p. 307], so focusing on the system components and on their *interactions*. In this area being able to specify low-level, system-specific and object-oriented rules governing state transitions under on the basis of known interactions is as good as it gets.

Can logic play a role in the “toolbox” of molecular biology? The main aim of this paper is to suggest a positive answer to this question. We claim that logic can indeed play a role as a useful tool for systematizing and formalizing certain molecular mechanisms that can sensibly be described as “inferential”, more specifically those characterized by state transitions under controlled conditions. However, applying logic to a new area is never a one-directional process. As witnessed by its interaction with computer science, logic itself evolves as a result of pressure coming from its application areas and it is only likely that extant logical systems will need substantial modifications to cater for the specific needs of this new subject-matter.⁴ Indeed, we shall argue that such a modification is needed for representing the kind of state transitions that are typical of biomolecular models. Our contribution can be seen as part of the general movement to reconcile logic and philosophy of science recently witnessed by [Bentham, 2012].

The rest of the paper proceeds as follows. In the next section, we briefly discuss the failure of classical logic to formalize state transitions. On the other hand, even a more flexible logical system such as linear logic [Girard, 1987] cannot cope successfully with the kind of transitions with which the molecular biologist is typically concerned. Then, in section 3, we propose a possible way to rearrange the Curry-Howard isomorphism in order to fit a relevant class of biochemical types and express their interaction. In sections 4 and 5 we respectively introduce a context-sensitive consequence relation and a cluster of inference rules in the natural deduction style that hinge upon the notion

² On the role of models in theoretical physics see the classic [Redhead, 1980] and [Cartwright, 1983]. See also the more recent [Norton, 2012; Weisberg, 2013].

³ And not even in certain areas of physics where models are all there is. On this point see [Boniolo *et al.*, 2002]. Concerning the debate on the difference between scientific laws and accidental generalizations, see [Boniolo, 2009].

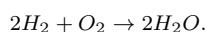
⁴ For interesting efforts in this direction, see [Danos and Laneve, 2004; Chauduri and Despeyroux, 2010].

of *control set* introduced in [Boniolo et al., 2013] and further developed in [D’Agostino et al., 2014]. Section 6 is devoted to more technical questions aiming at a refinement of the control set apparatus. In the final section, we briefly discuss some practical application of the calculus in molecular biology.

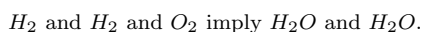
2 Which logic for state transitions?

As Jean-Yves Girard once observed, classical first order logic is particularly inconvenient for formalizing state transitions:

A typical consequence of the excessive focusing of logicians on mathematics is that the notion of state of a system has been overlooked. [...] We shall consider below the example of states in (summary!) chemistry, consisting of lists of molecules involved in a reaction [...] Observe that summary chemistry is modelled according to precise protocols, hence can be formalized: it can eventually be written in mathematics. But in all cases, one will have to introduce an extraneous temporal parameter, and the formalization will explain, in classical logic, how to pass from the state S (modelled as (S, t)) to a new one (modelled as $(S', t + 1)$). This is very awkward, and it would be preferable to ignore this ad hoc temporal parameter. In fact, one would like to represent states by formulae, and transitions by means of implications of states, in such a way that S' is accessible from S exactly when $S \rightarrow S'$ is provable from the transitions, taken as axioms. But here we meet the problem that, with usual logic, the phenomenon of updating cannot be represented. For instance take the chemical equation



A paraphrase of it in current language could be



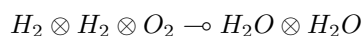
Common sense knows how to manipulate this as a logical inference; but this common sense knows that the sense of “and” here is not idempotent (because the proportions are crucial) and that once the starting state has been used to produce the final one, it cannot be reused. [Girard, 1995, p. 4]

Granted that “logic is the only way to rationalize bricolage” [Girard, 1987, p. 7], what is needed here are a type of conjunction and a type of implication that do not satisfy the traditional conditions of weakening and contraction:

$$A \wedge B \rightarrow A, \quad (\text{Weakening})$$

$$A \rightarrow A \wedge A, \quad (\text{Contraction})$$

for “contraction would say that the proportions do not matter, whereas weakening would enable us to add an atom of carbon to the left, that would not be present on the right” [Girard, 1995, p. 5]. These are the features of linear conjunction “ \otimes ” and linear implication “ \multimap ”. A correct representation would therefore be:



and “it turns out that if we take chemical equations written in this way as axioms, then the notion of linear consequence will correspond to the notion of accessible state from an initial one” [Girard, 1995, p. 5]. In the rest of this

section we shall take Girard’s suggestion seriously and outline a correspondence between linear deductions and (bio)chemical reactions that is akin to the well-known Curry-Howard correspondence, a cornerstone in nineteenth-century proof-theory. Next we shall explain why linear logic, as it stands, is not apt to represent conveniently the kind of state transitions that occur in molecular biology.

3 Correspondence between biochemical transitions and logical deductions

As shown by H.B. Curry in the Thirties [Curry, 1934] and W.A. Howard in the Sixties [Howard, 1980] of the last century, there is a strong correspondence between a well-defined class of computer programs (essentially those developed by means of functional programming languages) and logical proofs or, more generally, between logical systems and models of computation like typed λ -calculus and related variants. Such a correspondence clearly emerges when logical formulae are interpreted as *types of computational objects*. In this way, any theorem of the theory under focus represents the type of the program encoded by its corresponding proofs (see, [Howard, 1980], [Girard et al., 1989], [de Queiroz et al., 2011]). In a similar vein, we propose a correspondence in which logical formulae are interpreted as *types of molecular aggregates*, each proof as a reaction channel (i.e., a type of biochemical reactions) and each theorem as the type of the reaction channels encoded by its proofs.

By *molecular aggregate*, we mean a set of molecules satisfying some standard conditions for their potential interaction. Given two molecular aggregates x and y , their *composition* $x \circ y$ is the molecular aggregate consisting of all the molecules of x and all the molecules of y . We assume that aggregate composition is idempotent, associative and commutative. These properties correspond to the assumption that an “aggregate” is essentially a set of molecules in which *any* pair of molecules are in a position to interact with each other (whenever the conditions for interaction are met).

We aim at representing a biochemical pathway, namely the transition from a molecular aggregate to another molecular aggregate, as a deductive process. For this purpose we use a logical system that we call *Zsyntax*⁵ and is summarily described as follows:

- The *atomic formulae* of *Zsyntax* represent the elementary types, i.e., one-element aggregates consisting of a single item of a given molecular species, such as the tumor suppressor protein TP53; the caretaker gene Brest Cancer Type 1 (*BRAC1*), etc.
- The *interaction operator* \odot is an extra-logical operator and is used with the following meaning: we say that a molecule x is of type $A \odot B$ when there are molecules y of tpe A and z of type B , that interact to yield x . Notice

⁵ A preliminary version of this system and its possible role in the formalization of molecular biology were discussed in [Boniolo et al., 2010].

that Z-interaction is not associative: $(A \odot B) \odot C$ may be different from $A \odot (B \odot C)$. For, even if A interacts with B and the resulting product $A \odot B$ interacts with C , it does not follow that B interacts with C , or that $B \odot C$ or $A \odot (B \odot C)$ exists.⁶

- The *bonding language* \mathcal{L}_\odot consists of all the formulae that can be built out of the atomic formulae by means of the interaction operator \odot .
- The *aggregative Z-conjunction* \otimes is a logical operator for representing molecular aggregates. We say that an aggregate x is of type $A \otimes B$ to mean that there are two disjoint aggregates y of type A and z of type B , such that $x = y \circ z$. (By “disjoint aggregates” we mean aggregates that do not share any molecule.) Like multiplicative conjunction in linear logic, aggregative Z-conjunction is not idempotent, that is, $A \otimes A \neq A$. This feature is necessary to account for stoichiometry. So, an aggregate of type $ATP \otimes ATP \otimes ATP$ is an aggregate consisting of three molecules of ATP, where *ATP* is adenosine triphosphate. It follows from its definition that \otimes is both associative and commutative because of the corresponding properties of aggregate composition. The associativity of \otimes allows us to use expressions of the form $A_1 \otimes \dots \otimes A_n$ with no need for parentheses.
- The *Z-conditional* \rightarrow is the key logical operator for representing state-transitions. We say that x is of type $A \rightarrow B$ if and only if there exists a reaction that, for any y of type A (and disjoint from x), transforms $x \circ y$ (the composition of x and y) into an aggregate z of type B . For example, we know that there is a reaction that transforms any aggregate $x \circ y$, where x is of type *MDM2* and y is of type TP53, into an aggregate z of type MDM2, where *MDM2* is the gene and MDM2 the coded protein. Then, according to our definition, every molecule x of type *MDM2* is *ipso facto* of type $TP53 \rightarrow MDM2$. Another example is the enzymatic reaction that takes place when an enzyme E catalyzes a substrate S into a product P . In this case, we say that each molecule of type E is also of type $S \rightarrow (E \otimes P)$.
- The *selective Z-conjunction* \wedge is useful to represent multiple pathway reactions, i.e., when there are multiple reaction channels that may transform a given aggregate either into an aggregate of type A or into an aggregate of type B at our choice, i.e., depending on some unspecified, albeit controlled, conditions (e.g., temperature). In these circumstances we can say that the product of such a reaction is of type $A \wedge B$. For example, we may have that a given initial aggregate begets a product A at room temperature, but the product B if temperature is increased. In this case we say that the product is of type $A \wedge B$.
- The *Z-disjunction* \vee is useful to represent non-deterministic reaction, i.e., reactions that transform a given aggregate x either into an aggregate of

⁶ For example, consider the *Trp Operon* of *Escherichia coli* that does not bind the Trp-repressor unless the latter has not previously bound to Tryptophan. Then:

$$(\text{Tryptophan} \odot \text{Trp-repressor}) \odot \text{Trp Operon} \neq \text{Tryptophan} \odot (\text{Trp-repressor} \odot \text{Trp Operon})$$

and the formula on the right-hand side does not correspond to any molecular type.

formula	type
$A \odot B$ (with $A, B \in \mathcal{L}_\odot$)	one-element aggregate that results from the interaction of a molecule of type A with a molecule of type B , if such molecules interact, nothing otherwise
$A \rightarrow B$	aggregate x such that, for every disjoint aggregate y of type A , the composition $x \circ y$ triggers a biochemical reaction that delivers an aggregate of type B
$A \otimes B$	composition of an aggregate of type A with a <i>disjoint</i> aggregate of type B
$A \wedge B$	an aggregate that may be of type A or of type B at our choice, depending on known and controlled conditions.
$A \vee B$	aggregate that is either of type A or of type B , depending on some unknown or uncontrolled conditions.
\top	empty aggregate.

Table 1 Correspondence between formulae and types of molecular objects.

type A or into an aggregate of type B depending on unknown or unspecified conditions. For example, suppose that an initial aggregate begets a product of type A or a product of type B , depending on conditions that are beyond our control. In these circumstances, we may regard the reaction as a non-deterministic process and say that the product is of type $A \vee B$.

- The *unit formula* \top is a fixed formula representing the *unit type*, i.e., the one which is inhabited only by the “empty aggregate” denoted by \emptyset . Clearly, the composition of \emptyset with any aggregate x is equal to x . Thus, both $\top \otimes A$ and $A \otimes \top$ denote the same type denoted by A .
- The *full language* \mathcal{L} of Zsyntax consists of all the formulae that can be built out of the formulae of the bonding language \mathcal{L}_\odot by means of the logical operators $\otimes, \rightarrow, \wedge, \vee, \top$.

Given the meaning of the Zsyntax operators, the formulae of \mathcal{L} correspond to suitable types of molecular aggregates in accordance with the scheme in Table 1. To simplify our notation, we shall write $x : A$ to mean that an aggregate x is of type A .

4 A context-sensitive consequence relation

In the sequel we shall use the upper case greek letters Γ, Δ, A as metalinguistic variables ranging over finite sequences of \mathcal{L} -formulae and represent a sequence by simply listing its elements separated by commas (with no delimiters).⁷

⁷ Observe that, under these conventions, “ A ” represents both the formula A and the sequence consisting of one occurrence of A .

We let a finite sequence $\Gamma = A_1, \dots, A_n$ ($n \geq 1$) of formulae stand for an arbitrary aggregate x such that (i) $x = y_1 \circ \dots \circ y_n$ for some disjoint y_1, \dots, y_n , and (ii) y_i is of type A_i for all $i = 1, \dots, n$. When these conditions are met, we say that x is an *instance* of Γ and write $x : \Gamma$. Thus, each finite sequence of formulae acts like the description of a molecular state and we call it a *Z-state*. We say that a Z-state is *elementary* if all the formulae in it belong to the bonding language \mathcal{L}_\odot . We use the symbol \boxplus to denote the empty Z-state.

Let \vDash be a relation between Z-states defined as follows: $\Gamma \vDash \Delta$ if there exists a (possibly null) reaction — i.e., a possibly empty sequence of state transitions — that transforms any instance of Γ into an instance of Δ . We shall view \vDash as the analog of a logical consequence relation.

Notice that, given the meaning of \otimes , if $x : A_1, \dots, A_n$, then $x : A_1 \otimes \dots \otimes A_n$. Hence, in our “consequence relation” \vDash , unlike in Gentzen-style sequents, the commas within a sequence of formulae have the same meaning on both sides of the consequence sign, corresponding to the meaning of \otimes (“+” in the chemist’s shorthand for describing reactions):

$$A_1, \dots, A_m \vDash B_1, \dots, B_n \iff A_1 \otimes \dots \otimes A_m \vDash B_1 \otimes \dots \otimes B_n. \quad (1)$$

It immediately follows from the above definition that \vDash satisfies:

$$\Gamma \vDash \Gamma \quad (\text{Reflexivity})$$

and

$$\Gamma \vDash \Delta \text{ and } \Delta \vDash A \implies \Gamma \vDash A. \quad (\text{Transitivity})$$

Given the commutativity of aggregate composition, \vDash satisfies also:

$$\Gamma, A, B, \Delta \vDash A \implies \Gamma, B, A, \Delta \vDash A. \quad (\text{Exchange})$$

Moreover, it follows from the meaning of \otimes and \rightarrow that:

$$\Gamma, A \vDash \Delta \iff \Gamma \vDash A \rightarrow \otimes \Delta, \quad (2)$$

which provides the analog of a deduction theorem in our setting.

Notice that, in accordance with (2), when Γ is empty and Δ is a single formula B , we have that $A \vDash B$ if and only if $\top \vDash A \rightarrow B$. Here, the formula \top plays a special role in that it gives us a formal tool to express a piece of biochemical knowledge (the existence of a reaction) by simply saying that the empty aggregate is of a certain type expressed by a Z-conditional. For example, the fact that two molecules of types A and B interact can be expressed as follows:

$$\top \vDash A \otimes B \rightarrow A \odot B.$$

According to our conventions, this means that the composition of the empty aggregate \emptyset with any aggregate x of type $A \otimes B$ (which is equal to x itself) triggers a reaction that delivers a product of type $A \odot B$. In general, the fact that there exists a reaction that transforms an aggregate of type A into an aggregate of type B is expressed by the statement $\top \vDash A \rightarrow B$.

Moreover, our definitions imply that:

$$A_1 \rightarrow (A_2 \rightarrow \cdots (A_{n-1} \rightarrow A_n) \cdots) \vDash (A_1 \otimes \cdots \otimes A_{n-1}) \rightarrow A_n \quad (3)$$

and

$$(A_1 \otimes \cdots \otimes A_{n-1}) \rightarrow A_n \vDash A_1 \rightarrow (A_2 \rightarrow \cdots (A_{n-1} \rightarrow A_n) \cdots). \quad (4)$$

Let us say that a formula A is *valid* if $\top \vDash A$. It then follows from (2), (3) and (4) that

$$\Gamma \vDash \Delta \text{ if and only if } \otimes \Gamma \rightarrow \otimes \Delta \text{ is valid.} \quad (5)$$

Our “consequence relation” \vDash shares some important features with the consequence relation of linear logic. In particular, the Z-correspondence outlined above implies that the following properties of classical (and intuitionistic) consequence *do not hold*:

$$\Gamma \vDash A \implies \Gamma, \Delta \vDash A \quad (\text{Weakening})$$

and

$$\Gamma, B, B \vDash A \implies \Gamma, B \vDash A. \quad (\text{Contraction})$$

This amounts to saying that the following Z-conditionals are both *invalid*:

$$A \otimes B \rightarrow A \quad (6)$$

$$A \rightarrow A \otimes A. \quad (7)$$

However, in linear logic the multiplicative conjunction \otimes is monotonic with respect to \vDash (and \rightarrow), that is:

$$A \vDash B \implies C \otimes A \vDash C \otimes B \quad (8)$$

or, equivalently,

$$A \rightarrow B \text{ is valid } \implies C \otimes A \rightarrow C \otimes B \text{ is valid.} \quad (9)$$

This monotonicity of \otimes (that is typical of any conjunction operator in most known logical systems) is at odds with the fact that many important state transitions that are of interest in molecular biology are *context-sensitive*: a reaction may not take place depending on the molecular context in which it occurs. Consider, for example, the well-known effect called *enzyme inhibition*. An enzyme inhibitor, say I , is a molecule that binds to an enzyme, say E , and decreases its activity. If no inhibitor is present, we can assume that an enzyme E always links to its substrate S , that is: $E \otimes S \rightarrow E \odot S$ is a valid Z-conditional and, given a generic aggregate of type C ,

$$C \otimes E \otimes S \rightarrow C \otimes (E \odot S)$$

is also valid. However, if C is equal to the inhibitor I , we have that the above conditional is not valid and

$$I \otimes E \otimes S \rightarrow (E \odot I) \otimes S$$

is valid instead. This means that the transition from $C \otimes E \otimes S$ to $C \otimes E \odot S$ is not licensed for *all* possible substitutions of the metalinguistic variable C , but only for those that do not give rise to known exceptions.

Hence, while being clearly a close relative of the consequence relation of linear logic, our consequence relation \vDash does not intuitively satisfy (8) and (9) and so departs from linear logic for dropping the unrestricted monotonicity of \otimes with respect to \vDash and \rightarrow . Hence, to pursue our project of encoding biochemical reactions as proofs in a logical system, we need a refinement of linear logic that allows for expressing this controlled monotonicity of the aggregative conjunction \otimes that is crucial to express the kind of context-sensitive state transitions that are ubiquitous in molecular biology.

5 Context-sensitive inference rules

In this section we present a set of inference rules that form a system of natural deduction for context-sensitive state transitions.⁸

Amongst inference rules, we distinguish between *empirical* rules and *formal* rules. Empirical rules represent elementary transitions (with no intermediate products). The general format of an empirical rule T is:

$$\frac{\Gamma, A_1, \dots, A_m : \Gamma \notin \mathfrak{S}_T}{\Gamma, B_1, \dots, B_n}, \quad (10)$$

where

- Γ is a metalinguistic variable ranging over Z-states,
- $A_1, \dots, A_m, B_1, \dots, B_n$ are fixed formulae of the bonding language \mathcal{L}_\odot ,
- the elementary state transition

$$\frac{A_1, \dots, A_m}{B_1, \dots, B_n},$$

is empirically corroborated, and

- \mathfrak{S}_T is the set of all Z-states Δ for which it is *known* that the transition

$$\frac{\Delta, A_1, \dots, A_m}{\Delta, B_1, \dots, B_n},$$

does *not* take place.

The rule says that the transition from any aggregate described by the elementary Z-state A_1, \dots, A_m to an aggregate described by the elementary Z-state B_1, \dots, B_n takes place in every molecular context Γ that satisfies the *control*

⁸ We have chosen the natural deduction format for reasons of perspicuity. In fact the context-sensitive approach via control sets proposed here could be adapted to other formats that are more appropriate for automated reasoning purposes, e.g., by adapting the tableau method for linear logic presented in [D'Agostino *et al.*, 1999].

condition specified after the colon. So, the control condition restricts the allowed substitutions for the context variable Γ to the Z -states that do not give rise to known exceptions. Similar control conditions can be specified for every empirical rule, depending on the current state of knowledge concerning the types of molecules mentioned in the rule. The set \mathfrak{S}_T specified in the control condition of an empirical rule is called the *control set* associated with the rule. Notice that, for each control set \mathfrak{S} , by definition $\perp \notin \mathfrak{S}$ and $\top \notin \mathfrak{S}$.

Control sets are empirically determined, i.e., they result from empirical knowledge obtained in the laboratory. The immediate consequence is that their content typically changes over time and the resulting logical system is *open*: theorems may lose their status depending on modifications in our empirical knowledge base. Moreover, in general we require that \mathfrak{S}_T is a *computable* set. If tractability is an important consideration, one may specifically require that \mathfrak{S}_T is polynomial-time computable⁹.

In the enzyme inhibition example we have the following two empirical rules:

$$\frac{\Gamma, E, S : \Gamma \notin \mathfrak{S}_{T_1}}{\Gamma, E \odot S} \quad T_1 \quad (11)$$

with $I \in \mathfrak{S}_{T_1}$, and

$$\frac{\Gamma, I, S, E : \Gamma \notin \mathfrak{S}_{T_2}}{\Gamma, E \odot I, S.} \quad T_2 \quad (12)$$

Notice that: (i) the actual content of each control set is not fixed once and for all, but depends on the current state of empirical knowledge; (ii) the context-sensitive rule is equivalent to an infinite list of state transitions, one for each allowed substitution of the context variable Γ .

Formal rules are entirely expressed in terms of metalinguistic variables and can be distinguished into *structural* and *logical* rules. The structural rules are assumptions on the notion of Z -state and do not involve any operator of \mathcal{L} . Our only structural rule is the following *Exchange* rule:

$$\frac{\Gamma, A_1, A_2, \Delta}{\Gamma, A_2, A_1, \Delta.} \quad (\text{Ex})$$

The rule says that the ordering of formula-occurrences in a Z -state is immaterial.¹⁰ We could do without such a structural rule by stipulating that Z -states are *multisets*, rather than sequences, of formulae. So, $(A, A, B) \neq (A, B)$ and $(A, A, B) = (B, A, A) = (A, B, A)$. (This is essentially the same as assuming that (Ex) is always applied implicitly so that the actual order of the elements of a Z -state – but not their repetitions – can be ignored).

⁹ In some applications one may require that \mathfrak{S}_T is closed under given conditions. For example, a condition implicitly used in [Boniole *et al.*, 2013] and [D'Agostino *et al.*, 2014] is that if Γ (intended as multiset) is included in Δ , and Γ is in \mathfrak{S}_T , then Δ is also in \mathfrak{S}_T . This condition may be reasonable in specific application contexts but not in others, where extending a context may defuse its power of blocking a derivation.

¹⁰ For a proof-theoretical analysis of exchange rules, see [Piazza, 2001].

The logical rules license the transitions from a Z-state to another that are justified by the very meaning of the logical operators of \mathcal{L} . It is not difficult to see that the meaning of \rightarrow , as explained in Table 1, immediately justifies the following rule of \rightarrow -*introduction*:

$$\frac{\Gamma, \Delta \quad \boxed{\begin{array}{c} \Gamma, A \\ \vdots \\ B \end{array}}}{A \rightarrow B, \Delta} \quad (\rightarrow \mathcal{I})$$

This rule says that any instance of the Z-state Γ, Δ is *ipso facto* an instance of the Z-state $A \rightarrow B, \Delta$ (and so the latter is trivially “obtained” from the former by the empty sequence of state transitions), whenever it can be shown that any instance of Γ is of type $A \rightarrow B$. In accordance with the meaning of \rightarrow this amounts to showing that there is a sequence of state transitions leading from any instance of Γ, A to an instance of B , which is done in the subordinate proof enclosed in the box. For example, the following is a proof of $E \rightarrow E \odot S$ from S :

$$\frac{S \quad \boxed{\begin{array}{c} S, E \\ \hline E \odot S \end{array}}}{E \rightarrow E \odot S}$$

where the inference in the box is licensed by the empirical rule (11) with Γ equal to the empty Z-state. Observe that, by definition, the latter does not belong to the control set \mathfrak{S}_{T_1} , and so the control condition is satisfied.

The very meaning of \rightarrow justifies also the following rule of \rightarrow -*elimination*:

$$\frac{A \rightarrow B, A}{B}. \quad (13)$$

However, owing to the context-sensitivity of state transitions, such a rule cannot be generalized by including a context-variable Γ , unless a suitable control condition is specified. This leads to the following rule of *controlled* \rightarrow -*elimination*.

$$\frac{\Gamma, A \rightarrow B, A : \Gamma \notin \mathfrak{S}_{A \rightarrow B}}{\Gamma, B}. \quad (\rightarrow \mathcal{E})$$

where $\mathfrak{S}_{A \rightarrow B}$ is the set of Z-states Γ for which it is known that $\Gamma, A \rightarrow B, A \neq \Gamma, B$ when A and B are replaced by the specific \mathcal{L} -formulae involved in the rule application. Thus, the control set $\mathfrak{S}_{A \rightarrow B}$ depends on the application of the rule.

To stick to the enzyme inhibition example, the rule licenses the transition

$$\frac{\Gamma, E \rightarrow E \odot S, E}{\Gamma, E \odot S}$$

provided that the context Γ does not belong to the control set \mathfrak{S}_{T_1} specified in the empirical rule (11). In more complex cases, determining the control set associated with the specific instance of $A \rightarrow B$ involved in a given application of $\rightarrow \mathcal{E}$ depends on the empirical rules used in the known derivations of $A \rightarrow B$ from elementary Z-states.¹¹ Observe that the weaker version of \rightarrow -elimination in (13) is a special case of $\rightarrow \mathcal{E}$ that obtains when Γ is the empty Z-state (which, by definition, does not belong to any control set). Observe also that the uncontrolled version of $\rightarrow \mathcal{E}$ corresponds to the special case in which the control set $\mathfrak{S}_{A \rightarrow B}$ is empty.

Similarly, the meaning of \otimes justifies the following rules of \otimes -introduction and \otimes -elimination :

$$\frac{\Gamma, A, B,}{\Gamma, A \otimes B} \quad (\otimes \mathcal{I})$$

and

$$\frac{\Gamma, A \otimes B,}{\Gamma, A, B} \quad (\otimes \mathcal{E})$$

As for \wedge , we have one introduction rule:

$$\frac{\Gamma}{\begin{array}{|c|c|} \hline \Gamma & \Gamma \\ \hline \vdots & \vdots \\ \hline A, \Delta & B, \Delta \\ \hline \end{array}}{A \wedge B, \Delta} \quad (\wedge \mathcal{I})$$

and two elimination rules:

$$\frac{\Gamma, A \wedge B}{\Gamma, A} \quad (\wedge \mathcal{E}1)$$

$$\frac{\Gamma, A \wedge B}{\Gamma, B} \quad (\wedge \mathcal{E}2)$$

Symmetrically, we have two introduction rules for \vee :

$$\frac{\Gamma, A}{\Gamma, A \vee B} \quad (\vee \mathcal{I}1)$$

and

$$\frac{\Gamma, B}{\Gamma, A \vee B} \quad (\vee \mathcal{I}2)$$

¹¹ See Section 6 below.

and one elimination rule:

$$\frac{\Gamma, A \vee B}{\boxed{\begin{array}{c|c} \Gamma, A & \Gamma, B \\ \vdots & \vdots \\ \Delta & \Delta \end{array}}} \Delta \quad (\vee\mathcal{E})$$

Finally the rules for the unit formula \top are, trivially, the following:

$$\frac{\Gamma}{\Gamma, \top} \quad (\top\mathcal{I})$$

and

$$\frac{\Gamma, \top}{\Gamma} \quad (\top\mathcal{E})$$

We allow for the *simultaneous application* of several inference rules. Suppose, for example, that we are given two empirical rules T_1 and T_2 of the form:

$$\frac{\Gamma, A, B : \Gamma \notin \mathfrak{S}_{T_1}}{\Gamma, E} \quad T_1 \quad \frac{\Gamma, C, D : \Gamma \notin \mathfrak{S}_{T_2}}{\Gamma, F} \quad T_2 \quad (14)$$

with (i) $C, D \notin \mathfrak{S}_{T_1}$, (ii) $A, B \notin \mathfrak{S}_{T_2}$, (iii) $F \notin \mathfrak{S}_{T_1}$, (iv) $E \in \mathfrak{S}_{T_2}$. Then, the following transition is allowed by the simultaneous application of both rules:

$$\frac{A, B, C, D}{E, F}$$

since the control conditions are simultaneously satisfied. On the other hand, the following sequence of transitions is *not* allowed:

$$\frac{A, B, C, D}{\frac{E, C, D}{E, F}}$$

since the second transition is forbidden by the control condition of T_2 , while the following one:

$$\frac{A, B, C, D}{\frac{A, B, F}{E, F}}$$

is allowed, because no control condition is violated. Thus, in general, inference rules are *not permutable* and our formalism accounts for the temporal dimension of a transition process in a natural way.

A *Z-sequence based on Γ* is any finite sequence of Z-states such that its first element is Γ and each subsequent element is obtained from the preceding one by means of one or more simultaneous applications of inference rules. A *Z-proof of Δ from Γ* is a Z-sequence based on Γ ending with Δ . We shall write $\Gamma \vdash \Delta$ to mean that there exists a Z-proof of Δ from Γ .

Remark 1 If we ignore the empirical rules and the control condition of $\rightarrow \mathcal{E}$, we have that $\Gamma \vdash A$ if and only if A is provable from Γ in Girard's linear logic.

Finally, observe that each empirical rule T of the form (10) corresponds to the valid conditional:

$$A_1 \otimes \cdots \otimes A_m \rightarrow B_1 \otimes \cdots \otimes B_n. \quad (15)$$

and that the control set associated with each application of $\rightarrow \mathcal{E}$ to this conditional is the same as the one associated with T :

$$\frac{\Gamma, A_1 \otimes \cdots \otimes A_m \rightarrow B_1 \otimes \cdots \otimes B_n, A_1 \otimes \cdots \otimes A_m : \Gamma \notin \mathfrak{S}_T}{\Gamma, B_1 \otimes \cdots \otimes B_n.}$$

Thus, we can replace each empirical rule with a corresponding axiom. It is not difficult to show that the following is a derived rule of our system:

$$\frac{\Gamma}{\Gamma, A \rightarrow B} \quad \text{for every valid conditional } A \rightarrow B \quad (16)$$

which is also intuitively sound, since the empty aggregate is of type $A \rightarrow B$ whenever $A \rightarrow B$ is valid. Hence, axioms can be freely introduced *ad libitum* in any Z-state.

6 Elementary bases and elementary control sets

As they stand, control sets are highly redundant objects. We now show how their contents can be restricted to formulae of the bonding language \mathcal{L}_\odot . Let \mathbb{Z}_\odot be the set of all elementary Z-states, i.e., those whose formulae belong all to \mathcal{L}_\odot . An *elementary base* for an \mathcal{L} -formula A is defined inductively as follows:

1. for all $A \in \mathcal{L}_\odot$, Γ is an elementary base for A if and only if $\Gamma = A$;
2. Γ is an elementary base for $A \rightarrow B$ if and only if $\Gamma \in \mathbb{Z}_\odot$ and it is known that $\Gamma, A \vDash B$;
3. Γ is an elementary base for $A \otimes B$ if and only if there are Δ, Λ such that (i) $\Gamma = \Delta, \Lambda$, (ii) Δ is an elementary base for A and (iii) Λ is an elementary base for B ;
4. Γ is an elementary base for $A \wedge B$ if and only if Γ is an elementary base for A and Γ is an elementary base for B ;
5. Γ is an elementary base for $A \vee B$ if and only if either Γ is an elementary base for A or Γ is an elementary base for B .

Notice that an elementary base for A is a Z-state Γ such that (i) all the formulae of Γ belong to the bonding language \mathcal{L}_\odot and (ii) it can be shown that $\Gamma \vdash A$ with no application of the empirical rules. Let us denote by A^* the set of all elementary bases for the formula A . For every Z-state $\Gamma = A_1, \dots, A_n$, let Γ^* be the set of all Z-states $\Lambda = \Delta_1, \dots, \Delta_n$ such that, for all $i = 1, \dots, n$, Δ_i is in A_i^* .

Notice that the control sets associated with each empirical rule and each application of $\rightarrow \mathcal{E}$ are defined in such a way that if one of them contains Γ , then it contains also some $\Delta \in \Gamma^*$. We can make control sets less redundant by defining them as sets of elementary Z-states, i.e., Z-states in \mathbb{Z}_\odot , and replacing the preconditions for the empirical rules and for the controlled \rightarrow -elimination rule accordingly. The *elementary control set* associated with an empirical rule T of the form

$$\frac{\Gamma, A_1, \dots, A_m : \Gamma \notin \mathfrak{S}_T}{\Gamma, B_1, \dots, B_n}$$

is the set \mathfrak{S}_T^* defined as follows:

$$\mathfrak{S}_T^* = \{\Delta \mid \Delta \in \mathbb{Z}_\odot \text{ and it is known that } \Delta, A_1, \dots, A_m \not\ll \Delta, B_1, \dots, B_n\}.$$

Then, T can be replaced by:

$$\frac{\Gamma, A_1, \dots, A_m : \Gamma^* \cap \mathfrak{S}_T^* = \emptyset}{\Gamma, B_1, \dots, B_n}$$

The control condition of the rule says that none of the elementary bases for Γ belongs to the elementary control set \mathfrak{S}_T^* .

Similarly, the *elementary control set* associated with the *conditional* $A \rightarrow B$ can be defined as follows:

$$\mathfrak{S}_{A \rightarrow B}^* = \{\Gamma \mid \Gamma \in \mathbb{Z}_\odot, \text{ and for some } \Delta \in (A \rightarrow B)^* \text{ it is known that } \Gamma, \Delta \not\ll \Gamma, B\}.$$

Then the controlled \rightarrow -elimination rule becomes the following:

$$\frac{\Gamma, A \rightarrow B, A : \Gamma^* \cap \mathfrak{S}_{A \rightarrow B}^* = \emptyset}{\Gamma, B}$$

7 Conclusion: why another formalism for biology?

That is the final question to which we seek a plausible answer. One of the main advantages of our logical approach to molecular mechanisms lies in the ‘openness’ of the proposed calculus. It makes good sense that this calculus is open on two distinct levels:

1. It envisages an *evolving* set of proper axioms formally encoding empirical information;
2. It is non monotonic in character, so that adding *new* information to a knowledge base may invalidate previous conclusions.

Thinking realistically about molecular mechanism requires substantial departure from classical logic (transition rules are context-dependent rather than universal laws), but also from more flexible logical systems such as linear logic that cannot accommodate the kind of transitions with which the molecular biologist is usually concerned. In our example about the concurrent enzyme inhibition, we have – as the terminology itself suggests – a phenomenon of

concurrency involving two processes: (i) the reaction delivering the compound $E \odot S$ and (ii) the reaction leading to the alternative compound $E \odot I$. The second reaction has *priority* over the first, so that, when three elements E , S and I are synchronously introduced, the final result will be $E \odot I$ with S playing the role of the residual component. If we look at the concurrent enzyme inhibition from our logical point of view, then it becomes obvious that this inhibition corresponds to a typical non-monotonicity phenomenon. This is to say, the treatment of non-monotonicity that we have proposed is basically a way to syntactically represent concurrency phenomena. Generalizing from this example, then, we may suppose the existence of a deep conceptual relation between the notion of concurrency and that of non-monotonicity, a relation which would deserve an independent and more detailed analysis.

Other significant benefits are intrinsic to the intuitive natural deduction formulation of the calculus. When logicians and philosophers have considered biochemical formalisms, indeed, they have generally done so from the perspective of Hilbert-style proof systems: these consist in sets of formulas, called *axioms*, and a few inference rules — standardly only one, the *modus ponens* — for combining them. Yet, Hilbert systems are cumbersome to use for proving theorems and say very little about the dynamics of the processes described. Natural deduction calculi are instead explicitly conceived to do justice to the scientific practice one intends to formally grasp. ‘Natural’ essentially means that the calculus accounts for the reasoning in a field, by promoting a two-way traffic between logic and specific phenomena that occur in that field. This is the reason why a natural deduction formalism like the one we have illustrated seems to be more suited to represent biochemical reactions as they are produced in accordance with laboratory protocols.

More realistic logical models may require to shift from standard formulas to labelled formulas as basic units of the contexts. By ‘labelled formulas’ we mean expressions of the form $A : \alpha, B : \beta, C : \gamma$ where α, β, γ are labelling terms, specifying the values of suitable parameters expressing any kind of additional information concerning the entities to which the formulas refer.¹² For example, in such a labelled deductive system, an application of *modus ponens* may look like the following (discarding the control part):

$$\frac{A \rightarrow B : \alpha, A : \beta}{B : f(\alpha, \beta)}.$$

where f is a suitable function depending on A and B .

Moreover, our logical setting may allow the user to design and implement heuristic algorithms. This is the case, in particular, when one needs to ‘fill gaps’ in formal biochemical reactions, i.e. when we know the output compounds, whilst some initials or intermediate reactants are left unspecified. This kind of ‘inverse deduction’ is known as *abduction*:

¹² A similar idea has been investigated in [Chauduri and Despeyroux, 2010] using Hybrid Linear Logic.

$$\frac{A \rightarrow B \quad B}{A}$$

Although such an inference pattern is logically unsound, corresponding to the fallacy of affirming the consequent, its spectrum is nonetheless very broad in empirical sciences. Indeed, abduction can be assimilated to the inference to the best explanation [Lipton, 2004], and offers a simple way to characterize the process of scientific discovery [Aliseda, 2006]. So, one of the most challenging tasks when trying to define formalisms for empirical sciences is that of permitting the logic at hand to perform abductive inferences. The intuitive idea is that the control sets device should allow us to reproduce meaningful forms of abductive reasoning by admitting ‘controlled’ (i.e. constrained by control sets) versions of the abduction rule.

Another main advantage of our formalism is that natural deduction proofs, as already pointed out, are capable of expressing an intrinsic computational meaning throughout the Curry-Howard correspondence. The *practical* importance of this correspondence cannot be overestimated. Computer proof assistants exploits it: in order to check whether a formal proof validates a certain statement, the software has to safely ‘execute’ the corresponding proof term. In general, proofs assistants interactively help users in writing formal proofs and then in translating them into the corresponding proof terms. Proof terms are expressed in the language of some suitable version of typed λ -calculus so that the choice of the type system defines the logical framework upon which the proof assistant is developed. For instance, the proof assistant *Coq* stresses the logical framework provided by the Calculus of Inductive Constructions ([Bertot and Castéran, 2004]). Hence there is another important reason for thinking our proposal relevant to molecular biology. Indeed, we presented a system of biochemical types together with a natural deduction system able to represent biochemical reactions as formal proofs; to this extent, the computational content of proofs can be preserved by designing a specific type system in which reductions are context-sensitive, i.e. elementary computational steps are performed under some global constraint on the whole structure of the proof term. In other words, the present proposal lends support to the broader project to design a sort of ‘biochemical proof assistant’ able to formally certify reactions once they are specified in logical terms.

Acknowledgements

G.P. acknowledges the support from FAPESP Post-Doc Grant 2013/22371-0, São Paulo State, Brazil.

References

- [Aliseda, 2006] A. Aliseda. *Abductive Reasoning*, volume 330. Springer Netherlands, 2006.

- [Beall and Restall, 2006] J.C. Beall and G. Restall. *Logical Pluralism*. Clarendon Press, Oxford, 2006.
- [Bechtel, 1998] W. Bechtel. Representations and cognitive explanations: Assessing the dynamicist’s challenge in cognitive science. *Cognitive Science*, 22:295–318, 1998.
- [Bentham, 2012] J. Van Bentham. The logic of empirical theories revisited. *Synthese*, 186:775–792, 2012.
- [Bertot and Castéran, 2004] Y. Bertot and P. Castéran. *Interactive Theorem Proving and Program Development, Coq’Art: the Calculus of Inductive Constructions*. Springer-Verlag, Berlin, 2004.
- [Boniolo *et al.*, 2002] G. Boniolo, C. Petrovich, and G. Pisent. On the philosophical status of nuclear physics. *Foundations of Science*, 7:425–452, 2002.
- [Boniolo *et al.*, 2010] G. Boniolo, M. D’Agostino, and P. Di Fiore. Zsyntax: A formal language for molecular biology with projected applications in text mining and biological prediction. *PLoS ONE*, 5(3):e9511, 2010.
- [Boniolo *et al.*, 2013] G. Boniolo, M. D’Agostino, M. Piazza, and G. Pulcini. A logic of non-monotonic interactions. *Journal of Applied Logic*, 11:52–62, 2013.
- [Boniolo, 2009] G. Boniolo. Laws of nature: the kantian approach. In M. Bitbol, P. Kerszberg, and J. Petitot, editors, *Constituting Objectivity: Transcendental Perspectives on Modern Physics*, Western Ontario Series in the Philosophy of Science, pages 183–201. Springer-Verlag, 2009.
- [Cartwright *et al.*, 1995] N. Cartwright, T. Shomar, and M. Suárez. The toolbox of science. Tools for the building of models with a superconductivity example. *Poznań Studies in the Philosophy of the Sciences and the Humanities*, 44:137–149, 1995.
- [Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie*. Oxford University Press, Oxford, 1983.
- [Chauduri and Despeyroux, 2010] K. Chauduri and J. Despeyroux. A hybrid linear logic for constrained transition systems with applications to molecular biology. Technical report, INRIA-HAL, 2010. Available at http://www-sop.inria.fr/members/Joelle.Despeyroux/papers/hyll_report.pdf.
- [Cook, 2010] R.T. Cook. Let a thousand flowers bloom: a tour of logical pluralism. *Philosophical Compass*, 5(6):492–504, 2010.
- [Curry, 1934] H. Curry. Functionality in combinatory logic. *Proceedings of the National Academy of Science*, 20:584–590, 1934.
- [D’Agostino *et al.*, 1999] Marcello D’Agostino, Dov M. Gabbay, and Krysia Broda. Tableau methods for substructural logics. In M. D’Agostino, D.M. Gabbay, R. Hähnle, and J. Posegga, editors, *Handbook of Tableaux Methods*, pages 397–468. Kluwer Academic Publishers, 1999.
- [D’Agostino *et al.*, 2014] M. D’Agostino, M. Piazza, and G. Pulcini. A logical calculus for controlled monotonicity. *J. Applied Logic*, 12(4):558–569, 2014.
- [Danos and Laneve, 2004] V. Danos and C. Laneve. Formal molecular biology. *Theoretical Computer Science*, 325:69–110, 2004.

- [de Queiroz *et al.*, 2011] R.J.G.B. de Queiroz, A.G. de Oliveira, and D.M. Gabbay. *The functional interpretation of logical deduction*. Imperial College Press/World Scientific, London, 2011.
- [Gabbay, 1993] D.M. Gabbay. How to construct a logic for your application. In H.J. Ohlbach, editor, *GWAI '92: Advances in Artificial Intelligence*, volume 671 of *Lecture Notes in Computer Science*, pages 1–30, Berlin, 1993. Springer-Verlag.
- [Gabbay, 1994] D.M. Gabbay. What is a logical system? In D.M. Gabbay, editor, *What is a logical system?*, pages 179–216. Clarendon Press, Oxford, 1994.
- [Giere, 2006] R.N. Giere. *Scientific Pluralism*. University of Minnesota Press, 2006.
- [Girard *et al.*, 1989] J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*. Cambridge University Press, Cambridge, 1989.
- [Girard, 1987] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–101, 1987.
- [Girard, 1995] J.-Y. Girard. Linear logic: its syntax and semantics. In J.-Y. Girard, Y. Lafont, and L. Regnier, editors, *Advances in Linear Logic*, volume 222 of *London Mathematical Society Lecture Notes Series*, pages 1–42. Cambridge University Press, 1995.
- [Hamilton, 2007] A. Hamilton. Laws of biology, laws of nature: Problems and (dis)solutions. *Philosophy Compass*, 2(3):592–610, 2007.
- [Haufe, 2013] C. Haufe. From necessary chances to biological laws. *British Journal for the Philosophy of Science*, 64(2):279–295, 2013.
- [Howard, 1980] W. A. Howard. The formulae-as-types notion of construction. In J.P. Seldin and J.R. Hindley, editors, *To H. B. Curry: essays on combinatory logic, lambda calculus and formalism*, pages 479–490. Academic Press, London, New York, 1980.
- [Keller, 2000] E.F. Keller. Models of and models for. Theory and practice in contemporary biology. *Philosophy of Science*, 67:S72–S86, 2000.
- [Kellert *et al.*, 2006] S.H. Kellert, H.E. Longino, and C.K. Waters, editors. *Scientific Pluralism*. University of Minnesota Press, 2006.
- [Lipton, 2004] P. Lipton. *Inference to the Best Explanation*. Routledge, London, 2 edition, 2004.
- [Norton, 2012] J. Norton. Approximations and idealizations: Why the difference matters. *Philosophy of Science*, 79:207–232, 2012.
- [Piazza, 2001] M. Piazza. Exchange rules. *Journal of Symbolic Logic*, 66:509–516, 2001.
- [Raerinne, 2013] J. Raerinne. Stability and lawlikeness. *Biology and Philosophy*, 28(5):833–851, 2013.
- [Redhead, 1980] M. Redhead. Models in physics. *The British Journal for the Philosophy of Science*, 31:145–163, 1980.

-
- [Weisberg, 2013] M. Weisberg. *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press, Oxford, 2013.
- [Winsberg, 2001] E. Winsberg. Simulations, models and theories: complex physical systems and their representation. *Philosophy of Science*, 68:S442–S454, 2001.