

# L&PS – Logic and Philosophy of Science

Vol. VIII, No. 1, 2010

MARCO CARAPEZZA, MARCELLO D'AGOSTINO, <i>Logic and the Myth of the Perfect Language</i>	p. 1
MIRKO DI BERNARDO, <i>Ordine gratuito, morfogenesi autonoma e complessità semantica</i>	31
HECTOR FREYTES, ANTONIO LEDDA, GIUSEPPE SERGIOLI, <i>Continuous Functions as Quantum Operations: a Probabilistic Approximation</i>	83
MARCO GIUNTI, <i>Panorama e prospettive dell'approccio dinamico in scienza cognitiva</i>	101
MARTINVALDO KONIG, <i>Gödel Łukasiewicz Logic</i>	119
Information on the Journal	143



# Logic and the Myth of the Perfect Language\*

Marco Carapezza  
Department FIERI-AGLAIA  
University of Palermo  
e-mail: caramak@unipa.it

Marcello D’Agostino  
Department of Human Sciences  
University of Ferrara  
e-mail: marcello.dagostino@unife.it

1. The myth of the logically perfect language
2. Wittgenstein’s ‘adequate notation’ and the futility of logic
3. The perfect language and mathematical logic
4. ‘Unfeasible’ tautologies
5. Perfect language and computational complexity
6. Cook’s theorem and the inevitable imperfection of logical languages
7. Logic in a network of imperfect languages

ABSTRACT. We argue that the dream of a ‘perfect language’ – namely, a universal, unambiguous and semantically transparent medium of expression –, whose intriguing story has been told by Umberto Eco (1993), is deeply intertwined with the myth of *instant rationality*: the idea that a perfect language is one in which all logical relations become immediatly *visible*, so that the language itself “does the thinking for us” (Frege 1884). In the first part of this paper we trace this version of the dream in the works of Leibniz, Frege, Russell and Wittgenstein. In the second part we re-examine it in the light of more recent negative results in logic and theoretical computer science.

KEYWORDS: perfect language; instant rationality; logic; computational complexity.

---

\* A preliminary version of this paper was presented at the workshop “Linguaggio, verità e storia in matematica”, Mussomeli, 9 febbraio 2008. We wish to thank Gabriele Lolli, Franco Lo Piparo and Gianluigi Olivieri for interesting questions and suggestions.

## 1. The myth of the logically perfect language

“It [operating with figures mechanically] only became possible at all after the mathematical notation has, as a result of genuine thought, been so developed that it does the thinking for us so to speak” (Frege 1884, p. iv). These words by Gottlob Frege clearly and concisely express the myth of the ‘logically perfect language’ that constitutes the theme of the present essay.

Reflection on the perfect (or original) language has been a recurrent theme in scientific-philosophical thought beginning from Greek antiquity.<sup>1</sup> However, this phrase has been used to mean several things that are quite different from one another. So, let us immediately clarify that here we will not deal with projects based on the idea that names express the profound essence of things or give information on lexical semantics, thus proving to be ‘semantically transparent’. We will deal, instead, with the ‘logically perfect language’, meaning a language that can guarantee the correctness of processes of reasoning and one in which we can immediately recognize the relationships between propositions simply by means of *sensory perception*. Hence, this is a meaning of perfection in which semantic transparency concerns ‘logical words’ and not lexical terms. All that is required of the latter is referential univocity: to each conceptual content there must correspond only one term – a *real character*, to use the happy expression coined by Francis Bacon.

Like many other inventors of philosophical languages, Gottfried Wilhelm Leibniz believed that the possibility of a perfect language was founded on identification of the primitive notions that make up the knowable. A number should be attributed to every primitive notion and these numbers should combine together, producing all possible notions. A system of translation of these numbers into consonants and vowels would then allow us to assign to each notion a *character*<sup>2</sup> that refers univocally to the notion designated by it. Lastly, on the combinatory rules of these characters there should be founded the *ars iudicandi*:

I necessarily arrived at this remarkable thought, namely that a kind of alphabet of human thoughts can be worked out and that everything can

---

<sup>1</sup> Among the numerous works that could be mentioned, two that after many years have aroused interest in this theme are Couturat (1903a) and Eco (1993).

<sup>2</sup> For an explanation of how to construct the ‘characters’, see Leibniz (1679a). More in general, on pp. 42-92 there is the first systematic attempt by Leibniz to develop a logical calculus. For an English translation of many of his writings, see Leibniz (2000). On the similarity and differences between Leibniz’s ideas and those of other seventeenth-century authors engaged in projects for creating calculation languages, like George Dalgarno and John Wilkins, see Rossi (2000).

be discovered and judged by a comparison of the letters of this alphabet and an analysis of the words made from them (Leibniz 1679c, p. 222).

Although for Leibniz all knowledge was potentially translatable into the *Characteristica*, this instrument could also have worked *in specific domains*, and therefore could also have been used *before* the whole alphabet of human thought was identified. In this connection, the grammar of this language would be unique and indifferent to the sphere of application. Such grammar would not be different from logic, seen as “the art of using the understanding not only to judge proposed truth, but also to discover hidden truth” (Leibniz 1696, p. 475). Hence, it would be, at one and the same time, a method of discovery (*ars inveniendi*) and a method of decision (*ars iudicandi*). Thus, once the fundamental notions have been identified, there could be a *calculus ratiocinator* in the light of which all controversies become vain:

Obviously, once this is performed, every paralogism is nothing but a *calculation mistake*, and [...] every sophism, expressed in this kind of new writing, is nothing but a solecism or a barbarism, to be resolved easily through the laws of this philosophical grammar itself. Henceforth, when controversies arise, there would be no more need of disputation between two philosophers than between two accountants. For it would suffice to take their pencils in their hands, sit down with their counting-tables (having called in a friend, if they like) and say to each other: *let us calculate* (Leibniz 1684, p. 200, our translation).

Such a system would not be applicable to historical-natural language,<sup>3</sup> but to ideographic language, the *characteristica universalis*, in which there is a rigorous one-to-one correspondence between simple signs and simple ideas and, consequently, between compound signs and compound ideas. Only a similar artificial language would ensure this possibility, and not ordinary languages. The fact is that the latter:

although they serve for reasoning, nevertheless they are subject to innumerable misunderstandings, nor can they be employed for calculation, that is to say in such a way that the errors of reasoning can be discovered (Leibniz 1684, p. 205).

---

<sup>3</sup> Awareness of the historical character of languages is at the heart of Leibniz’s theoretical reflection, and does not contradict logico-combinatory works like *Dissertatio de arte combinatoria*. See Mugnai (1976) and Gensini (1991).

Hence *characteristica* will not make all men equally able to attain prodigious results in discovery, but it will make them all able not to commit errors or, possibly, to recognize errors by themselves, whether committed by them or by others. On one side, therefore, we have a system of discovery, linked to individual ability, and on the other a system of control that would become mechanical, and even *immediate*:

And this is the advantage of our method – we can judge *at once*, through numbers, whether proposed proposition are proved, and so we accomplish, solely with the guidance of characters and the use of a definite method which is truly analytic, what others have scarcely achieved with the greatest mental effort and by accident. And therefore we can succeed in presenting conclusion within our own century which would scarcely be provided for mortals in many thousands of years otherwise (Leibniz 1679b, p. 236, our emphasis).

Thus the *characteristica universalis* should make us *immediately* able to judge the correctness of reasoning and this immediateness would be made possible by the reduction of control to simple sensory perception. As for Leibniz, in later authors too this strong idea of *instant rationality* will characterize the projects of logically perfect languages.

“Leibniz’s dream”, as Giuseppe Peano was to call this grandiose programme of the German philosopher,<sup>4</sup> was taken up in the late nineteenth century and the early twentieth, interweaving with projects for refounding logic. Explicit references to Leibniz’s programme can be found in Boole and, even more markedly, in Frege, who at the beginning of his *Begriffsschrift* harks back precisely to Leibniz’s “*calculus philosophicus or ratiocinator*” (1879, p. 6). In this work, Frege, after indicating in the symbolic systems of arithmetic, geometry and chemistry partial realizations of Leibniz’s project, though relating to particular fields of knowledge, claims for his own ideography the merit of having added a new field and “indeed the central one, which borders on all the others” (1879, p. 7). In an article devoted to clarifying the difference between Peano’s ideography and his own, Frege actually writes: “it [the *Begriffsschrift*] is, to use an expression coined by Leibniz, a *lingua characterica*” (Frege 1897).<sup>5</sup>

---

<sup>4</sup> See Peano (1908). The expression is taken up by Bertrand Russell (1901). On this topic see also Mondadori (1986).

<sup>5</sup> An analogous concept is found in Frege (1880-81). This reference is important for understanding that what Frege really meant to obtain was not a calculus – and so a variant of Boole’s logic. Thanks to the possibilities afforded by “predicate letters, variables and quantifiers”, the

An important implication of Leibniz's programme is that if one succeeded in taking it to its conclusion, there would *be no more need for logic as such*: correct syntactic formulation would in itself constitute a guarantee of correct logical reasoning. That the 'end of logic' is the inevitable result of the fulfilment of Leibniz's dream is particularly clear in the words of Frege:

if we had a logically perfect language we would perhaps further need no logic, or we could read it off language. But we are at a vast distance from being in this condition (Frege 1915, p. 252).

## 2. Wittgenstein's 'adequate notation' and the futility of logic

Although in Wittgenstein's work there are no explicit references to Leibniz, and it is a very arduous undertaking to reconstruct his library, it appears very likely that the work of the German philosopher, possibly through the mediation of Frege and Russell, was very much present in the mind of the author of the *Tractatus*. In this connection, there are manifold (though little investigated) echoes of Leibniz in his writings.<sup>6</sup>

In the *Tractatus*,<sup>7</sup> Wittgenstein raises the question of an 'adequate notation', meaning a notation in which the grammatical structure and the logical structure of sentences coincide, and one through which each sentence *shows* its sense. For Wittgenstein, the sense of a proposition is to be identified with the possibility of its being true or false: "The sense of a proposition is its agreement and disagreement with the possibilities of the existence and non-existence of the states of affairs" (T. 4.2).<sup>8</sup> He distinguishes two types of propositions, elementary propositions [*Elementarsätze*] and propositions that have a

---

proposition becomes articulated and a meaning can be expressed by means of a language that can be considered a *lingua characteristic*, as already remarked by Van Heijenoort (1967b). A similar view is expressed in Hintikka (1997). Recently Korte has maintained that the decision to consider his *Begriffsschrift* a *lingua characteristic* "is related to his logistic program, which was to show that judgements of arithmetic are not synthetic, as Kant had claimed, but analytic" (Korte 2010).

<sup>6</sup> Among the few authors that relate the works of the first analytic philosophers with earlier artificial language projects, there are Soren Stenlund (2002) and Nicolay Milkov (2006).

<sup>7</sup> The reference is to Wittgenstein (1921). Here we quote the English translation, sometimes with slight modifications.

<sup>8</sup> Here we will not consider the consequences of an adequate notation in the framework of the distinction between senselessness [*Sinnlos*] of logical propositions, which at all events represent the framework of the world (cf. T. 6.124), and the nonsense [*Unsinn*] of non-logical propositions (cf. T. 4.003). See the now classic Diamond (1991), and Conant (2000).

complex structure, being formed by elementary propositions. While the truth of elementary propositions consists in the existence or non-existence of a certain fact about the world, the truth of the other propositions depends on the relations that link the elementary propositions contained in them: complex propositions are truth functions of the elementary propositions. As Wittgenstein writes: “A proposition is the expression of agreement and disagreement with the truth-possibilities of the elementary propositions” (T. 4.4).

Thus if the sense of a proposition consists in the conditions in which it is true or false, an adequate notation should be able to show these conditions explicitly. Wittgenstein has no doubt about the fact that this should be possible. In the *Tractatus* there is even a certain tension in relation to the idea that the proposition already shows its meaning in ordinary language – “a proposition *shows* its sense” (T. 4.022) – and not only in the presence of adequate symbolism. Nevertheless, Wittgenstein shares with Frege and Russell, however different their positions on common language may be, the idea that: “[in common language] it is humanly impossible to deduce the logic of language” (T. 4.002), because the grammatical structure does not mirror the logical structure of the sentence itself. The logic underlying linguistic utterances could instead be made evident by a more appropriate symbolism, one capable of making it immediately *visible*<sup>9</sup> without resorting to any ‘deductive process’. This seems also to be lurking in the back of Russell’s mind, when he writes:

In a logically perfect language, there will be one word and no more for every simple object, and anything that is not simple will be expressed by a combination of words. [...] A language of that sort will be *completely analytic*, and will show *at a glance* the logical structure of the facts asserted or denied (Russell 1918, p. 176, our emphasis).

For Wittgenstein the logically perfect language is the practicable result of his proposal of a new kind of symbolism in which recognition of tautologies can be *immediate*. Since the deducibility of  $q$  from  $p$  (where  $p$  indicates the conjunction of the premises and  $q$  the conclusion of a deductive process) is equivalent to the tautologyhood of  $p \rightarrow q$ , the correctness of the inference of  $q$  from  $p$  would prove, in a symbolism of the kind, to be immediately visible.

Frege and Russell had demonstrated that it was possible, although extremely complicated, to show in a purely formal way the tautological character of a proposition through the creation of a proof system founded on self-

---

<sup>9</sup> Wittgenstein made explicit use of visual metaphors probably more than any other philosopher and the centrality of vision in cognitive processes is one of the main themes of the *Tractatus*.



evident logical axioms (propositions whose tautological character is immediately recognizable) starting from which it was possible to derive all the other tautologies through valid rules of inference (which preserve tautology-hood).

We prove a logical proposition by creating it out of other logical propositions by applying in succession certain operations, which again generate tautologies out of the first. (And from a tautology only tautologies follow.) (T. 6.126).

In Wittgenstein's view, however, the systems of Frege and Russell were not as perspicuous as they should. In the first place, their approach made logical relationships opaque, causing propositions that are in fact identical to appear as distinct. For example,  $p \rightarrow q$  (if  $p$  then  $q$ ) and  $\sim p \vee q$  (not  $p$  or  $q$ ), though different signs, for Wittgenstein are *the same* proposition because both are true for all values of  $p$  and  $q$ , except in the case in which  $p$  is true (T) and  $q$  is false (F). Secondly, the privileged position of the tautologies that play the role of axioms is entirely arbitrary in such systems. Furthermore, recognizing that a proposition is a tautology, through a formal derivation, may generally be an extremely complicated process, and this appears in sharp contrast with the idea that "every tautology itself shows that it is a tautology" (T. 6.127).<sup>10</sup> Hence for Wittgenstein, the systems of Frege and Russell do not realize an 'adequate notation' for expressing logical connections. Indeed, in such a notation

That the truth of one proposition follows from the truth of other propositions, we perceive from the structure of the propositions (T. 5.13, our emphasis)

and

we can recognize in an adequate notation the formal properties of the propositions by mere inspection (T. 6.122).

Hence:

'Laws of inference', which – as in Frege and Russell – are to justify the conclusion are senseless and would be superfluous (T. 5.132).

---

<sup>10</sup> For a discussion of the effective decidability of the propositions of logic within the structure of the *Tractatus* see Marconi (2005).

Thus the outcome of this position is the thesis according to which in an *adequate notation logical deduction would be wholly superfluous!*<sup>11</sup>

It seems plausible that, according to Wittgenstein, an adequate notation could be provided by the famous method of the truth-tables introduced by himself in the *Tractatus*. The truth-table of a proposition explicitly shows its truth conditions in terms of the truth and falsehood of the elementary propositions contained in it. Further, for Wittgenstein every truth-table constitutes a propositional sign (T. 4.442):

For example the following is a propositional sign:

«

p	q	
T	T	T
F	T	T
T	F	
F	F	T

».

The table in the example illustrates the conditions of truth (and of falsehood) of the implication *if p then q* (but also of the disjunction *not-p or q*) in terms of the truth and falsehood of the elementary propositions *p* and *q*. It follows that *if p then q* is false only in the case in which *p* is true and *q* false, while it is instead true in all other cases. An abbreviation of this scheme could be  $(TT-T)(p, q)$  or, more explicitly,  $(TTFT)(p, q)$ . Other logical connections between the same elementary propositions would evidently give rise to other configurations.

Thus, the ideal of an adequate notation, a symbolic system in which propositions explicitly show their own truth conditions, seems to be fully accomplished. The theoretical move that allows for this result consists in considering the truth-table itself as a ‘propositional sign’, that is to say, a configuration of signs able to serve as a proposition. The difficulty about recognising tautologies that afflicted Frege’s and Russell’s systems seems to dissolve into a symbolism in which logical relations become immediately (and *instantly*) visible. In this vein, Wittgenstein writes: “if two propositions contradict one another, then their structure shows it; the same is true if one of them follows from the other. And so on” (T. 4.1211). Hence we could say that we *see* from the propositional signs  $(TTFT)(p, q)$  and  $(FFTF)(p, q)$  – that is to say from a comparison between the arrangements of the signs *T* and *F* appearing in them – that

---

<sup>11</sup> On this point see also D’Agostino and Floridi (2009).

the propositions in question contradict one another.<sup>12</sup> Therefore we do not need any logical demonstrations.

By way of example, remembering that “Every proposition of logic is a *modus ponens* present in signs” (T. 6.1264), consider this simple inference:

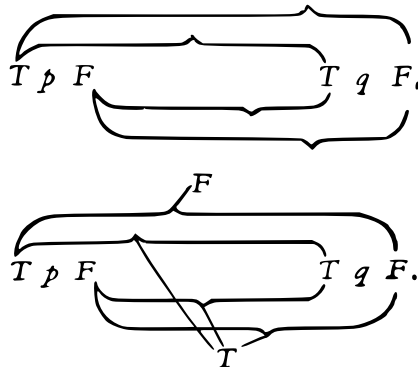
*If it rains I put on my hat*  
*It rains*  
*Then I put on my hat*

Let us now consider the truth-table that shows that this inference – which exemplifies *modus ponens* – is tautological (*p* and *q* stand for arbitrary propositions):

<i>p</i>	<i>q</i>	$p \rightarrow q$	$(p \rightarrow q) \wedge p$	$((p \rightarrow q) \wedge p) \rightarrow q$
T	T	T	T	T
F	T	T	F	T
T	F	F	F	T
F	F	T	F	T

Recognition of the tautological character of the proposition in the last column – which guarantees the correctness of the inference in question – does not here require any ‘logical demonstration’, but the simple inspection of the propositional sign representing the proposition in the tabular notation.

In the *Tractatus* (6.1203), Wittgenstein also provides another example of adequate notation. It is not clear whether this intended to be an alternative method to the tabular one, serving for visual recognition of the tautologies or, as it would seem, amounts to essentially the same method:



<sup>12</sup> On this point see Piana (1973, p. 54).

Let  $p$  and  $q$  be the connected elementary propositions in a complex proposition. Each proposition may have two truth values T and F that will be connected to those of the other proposition, yielding the same values as obtained with the truth-tables.<sup>13</sup>

As we will see more clearly in the next sections, the central problem in both methods is that the number of possible assignments of truth-values increases exponentially with the number of elementary propositions occurring in the complex proposition, making both the tabular notation and the graphic one outlined above entirely unfeasible. Wittgenstein does not consider this problem,<sup>14</sup> and does not provide examples of how the most complex tautologies should be written in one of these symbolisms.

We will argue, in the next sections, that this is by no means an empirical problem that falls outside the austere perspective of *Tractatus*, and that the unfeasibility of the symbolisms proposed by Wittgenstein is only a special case of a general logical problem that (with all probability) does not admit of any solution.

### 3. The perfect language and mathematical logic

Through the interpretation of the truth-table as a propositional sign, the myth of the logically perfect language seems to find partial realization in Wittgenstein's *Tractatus*, though in the narrow domain of propositional logic. In this connection, the truth-table of a proposition fully expresses the sense of the logical words that occur in it and it would therefore provide an example of 'adequate notation' that, according to Wittgenstein, would make the process of logical deduction entirely superfluous. It is true that Church-Turing's undecidability theorem (1936)<sup>15</sup> excludes the possibility of finding a similar perfect language for more powerful logics, such as the logic of quantifiers (with non-monadic predicates). But it is also true that this negative result does not imply that there cannot be, even for the logic of quantifiers, an 'almost-perfect' language, which works reasonably well in all practical contexts. After all, the impact of the undecidability theorem on the massive attempts set going in the

---

<sup>13</sup> Max Black (1964, comment on T 6.1203) defines this system "coarse and unusable", while claiming that its tabular equivalent is by far better.

<sup>14</sup> As Pasquale Frascolla puts it, the speaker able to see all meta-logical relations is not the empirical speaker, because he would be a logically omniscient being; and Frascolla observes that this raises a new problem: "in what sense, if any, can formal relations be said to exist even when the speaker is not in fact able to see them?" (2006, p. 180).

<sup>15</sup> For a historical reconstruction of the decision problem and of the negative results of Church and Turing, see Kneale and Kneale (1962).

1950s to realize Leibniz' dream through so-called 'automated deduction' was practically none. If anything, there prevailed the enormous impression aroused by the *positive* results, also made possible by theoretical contributions such as Herbrand's theorem, that provided methods of great practical interest for reducing quantificational reasoning to Boolean reasoning.

Analogous considerations also hold for other famous negative results of modern logic, first of all Gödel's theorems. These meta-mathematical exploits struck at the heart the hard-core of Hilbert's philosophical programme on the foundations of mathematics, that is to say, the assumption that an adequate formalization of mathematics is possible (Gödel's first theorem), and the assumption that it should be possible to show the consistency of this formalization with reliable means (Gödel's second theorem).<sup>16</sup> From the point of view of deductive *practice* and the possibility of a logically perfect language that would make it trivial or even superfluous, the first theorem is clearly the important one, but it is not at all obvious what its real scope is. As Solomon Feferman has observed:

A common complaint about this result is that it just uses the diagonal method to "cook up" an example of an undecidable statement. What one would really like to show undecidable by PA or some other formal system is a natural number-theoretical or combinatorial statement of prior interest. The situation is analogous to Cantor's use of the diagonal method to infer the existence of transcendental numbers from the denumerability of the set of algebraic numbers; however, that did not provide any natural example. The existence of transcendentals had previously been established by an explicit but artificial example by Liouville. Neither argument helped to show that  $e$  and  $\pi$ , among other reals, are transcendental, but they did at least show that questions of transcendence are non-vacuous. Similarly, Gödel's first incompleteness theorem shows that the question of decidability of sentences by PA or any one of its consistent extensions is non-vacuous. That suggests looking for natural arithmetical statements which have resisted attack so far to try to see whether that is because they are not decided by systems that formalize a significant part of mathematical practice (2006, pp. 436-437).

But so far this search has produced no results.<sup>17</sup> Indeed, one of the most accredited candidates, the statement of Fermat's conjecture, has recently been proved,

---

<sup>16</sup> For an entertaining discussion of Hilbert's programme within the framework of the search for the perfect language, see Chaitin (2009), inspired by Eco (1997).

<sup>17</sup> Paris and Harrington's result, according to which a certain modification of a famous theorem of Ramsey is independent of Peano arithmetic, in a sense constitutes an exception. Nev-

and according to some logicians its proof could be formalized in Peano arithmetic. For none of the other candidates that survive, including Goldbach's conjecture and Riemann's hypothesis, has anyone ever succeeded in showing their independence of Peano arithmetic or of any of its coherent extensions.

Hence the famous negative results of 20<sup>th</sup>-century meta-mathematics, obviously without denying their importance on the philosophical side, have not had such a practical impact as to establish once and for all that the very idea of a logically perfect language is void of content, that is to say, cannot have any real instantiation that corresponds, *more or less*, to what Leibniz, Frege and Wittgenstein had in mind. From this point of view, perhaps it is not the negative results that are surprising. What is truly surprising, as George Kreisel once remarked, is the discovery that certain branches of mathematics – such as pure logical reasoning on truth functions and quantifiers or elementary Euclidean geometry – are, after all, mechanizable.<sup>18</sup>

Therefore in this perspective the emphasis must be, if anything, on the numerous *positive* results, first of all the decidability of elementary Euclidean geometry, proved by Tarski in 1951,<sup>19</sup> and one may wonder what Leibniz's reaction would have been to a result of the kind! From this perspective, the dream of a logically perfect language would perhaps be impossible 'in principle' – that is to say, there would be no foundation to the grand theoretical pretension of finding a *universal* adequate notation, in which the solution of *all* logical problems can be 'read' off their statement itself – but it may well be realizable 'in practice', in the vast majority of the interesting cases that come to our attention and that, in the imperfect language of ordinary reasoning, can only be solved through tiresome (and therefore highly fallible) deductive processes. Leibniz's dream, and the myth of the perfect language that is indissolubly associated with it, would die as a *philosophical* research programme, but would survive as the metaphysical hard-core of a *scientific* research programme that, in spite of the negative results, appears to be highly progressive.<sup>20</sup>

---

ertheless, as Feferman himself observes, even in this case it was not a matter of an undemonstrable proposition whose truth had previously been the object of conjecture. For Ramsey's theorem and Paris and Harrington's result see Mangione and Bozzi (1993, in particular p. 516 and pp. 856-857).

<sup>18</sup> On this point see Kreisel and Krivine (1971, p. 165-166).

<sup>19</sup> Tarski (1951).

<sup>20</sup> As Gregory Chaitin puts it: "There's a wonderful intellectual tension between incompleteness and the fact that people still believe in formal proof and absolute truth. People still want to go ahead and carry out Hilbert's program and actually formalize everything, just as if Gödel and Turing had never happened!" (2009, p. 19).

However, this hope for the realizability of a ‘practically perfect’ logical language is also, with all probability, entirely unfounded. We will see that this pessimistic conclusion is the consequence of a profound result obtained in a field of research, that of computational complexity, whose origins are apparently very distant from typical philosophical problems, such as that of the perfect language, and are instead linked to a vast constellation of practical problems raised by the development of computer technology.

#### 4. ‘Unfeasible’ tautologies

In his *Games of Arithmetic and Interesting Problems*, published for the first time in 1925, Giuseppe Peano introduces a series of problems serving to make “arithmetic more pleasant and less boring.” Among the “captious problems”, in which “the true answer is not the one that first presents itself to the mind”, there was the following one:

A party of 7 travellers go to a hotel and ask for a bed for each traveller. The hotelier answers: “I only have six beds, distinguished with the letters A, B, C, D, E and F. But I will try to accommodate you anyway.” So he tells two travellers to sleep in bed A, then one in bed B, and that is three; then one in C, and that is four; then one in D, and that is five; then one in E, and that is six; then he takes one of those that were in A and transfers him to F, and that is seven. Thus the 7 travellers sleep in 6 beds, one per bed. How did he do this? Anyone who plays the game represents the beds with six cards, and proceeds fast, so that the listener does not realize that a traveller has been counted twice (Peano 1925).

The impossibility of an ‘honest’ solution to the problem of the hotel keeper is ensured by a fundamental combinatorial principle, also known as ‘pigeonhole principle’ or ‘Dirichlet’s principle’, according to which  $n + 1$  objects cannot be placed in  $n$  boxes unless one of the boxes contains more than one object. Despite its obviousness from the intuitive point of view, this principle is surprisingly useful for showing an enormous variety of mathematical facts, from simple curiosities, for example that in London there are at least two people that have the same number of hairs, to numerical matters that are anything but obvious, for example that among  $N$  integers chosen at random there are always two whose difference is divisible by  $N - 1$ , or that for any irrational number  $a$  there exist infinite rational numbers  $r = p/q$  such that  $|a - r| < q^{-2}$ .

It is well known that this important combinatorial principle can be expressed in a simple and natural way through a class of tautologies of Boolean logic. Indicating with  $p_{i,j}$  the sentence according to which the object  $i$  occupies box  $j$ , the proposition

$$(*) \quad (p_{1.1} \vee p_{1.2}) \wedge (p_{2.1} \vee p_{2.2}) \wedge (p_{3.1} \vee p_{3.2}) \wedge \neg(p_{1.1} \wedge p_{2.1}) \wedge \\ \wedge \neg(p_{1.2} \wedge p_{2.2}) \wedge \neg(p_{1.1} \wedge p_{3.1}) \wedge \neg(p_{1.2} \wedge p_{3.2}) \wedge \neg(p_{2.1} \wedge p_{3.1}) \wedge \\ \wedge \neg(p_{2.2} \wedge p_{3.2})$$

asserts, for example, that three objects can be placed in two boxes in such a way that each box contains at most one object. The impossibility of the situation described by this proposition is therefore a special case of the pigeonhole principle, with  $n = 2$ . The interesting thing is that this proposition is *logically* inconsistent (even simply at the level of Boolean logic) and therefore its negation is a *tautology*. Hence the pigeonhole principle is expressed by the class of all the tautologies obtained by negating the propositions constructed on the model of (\*) for each positive whole number  $n > 1$ .

Well, the tautology that expresses the pigeonhole principle for a given  $n$  contains  $f(n) = (n + 1) \times n$  distinct propositional letters and the length of the expression representing it in the still “imperfect” logical language of Frege and Russell contains (including the parentheses and counting each  $p_{i,j}$  as a single symbol) an overall number of symbols equal to:

$$g(n) = 7/2 n^3 + 11/2 n^2 + 4n + 1.$$

On the other hand, the number of lines contained in the complete truth-table for this expression, that is to say in its translation into the ‘adequate notation’, is equal to  $h(n) = 2^{f(n)}$ . The problem is that with an increase in  $n$ , the length of the translation into the ‘logically perfect’ language of the truth-tables grows much faster than the length of the proposition expressed in a standard Boolean language. While  $g(n)$  is a *polynomial* function of  $n$  (of degree 3),  $h(n)$  is an *exponential* function of  $n$ , and it is well known that the speed of growth of exponential functions very soon takes the length of the translation beyond the limits of feasibility. For example, returning to the ‘captious’ problem of Peano, to express the fact that 7 travellers cannot be put in 6 beds, a string containing 979 symbols is sufficient in an ordinary logical language, but its translation into the perfect language would require the construction of a truth-table containing 4 398 046 511 104 lines (without calculating the length of each line)! It can easily be verified that, even for relatively low values of  $n$ , the overall number of



symbols occurring in the truth-table required to ‘see immediately’ that the expression is a tautology would be higher than the number of atoms contained in the known universe, while the same tautology expressed in the imperfect standard language contains only a few thousand symbols. It follows that translating a proposition from the ordinary logical language of Frege and Russell into the logically perfect language of the truth-tables – in which it is possible to recognize that a certain expression is a tautology simply by examining the expression itself – is a task that is *practically impossible*. If a logically perfect language exists, and it must be possible to use it in practice, it cannot be that in which the propositional signs are constituted by the truth-tables.

This *reductio ad absurdum* of the idea that a logically perfect language could be that of the truth-tables is obviously not sufficient to render research in this direction entirely vain. There might be a more concise way to give a complete representation of the sense of a proposition – that is to say, to perform the same task as Wittgenstein assigned to the truth-tables – which however does not determine any combinatorial explosion. After all, in the truth-table method there are obvious inefficiencies that could easily be eliminated: (a) the same propositional letter can occur many times, so that a symbolic expression could be ‘compressed’ by resorting to a representation in the form of a graph, and (b) the enumeration of the truth conditions is manifestly redundant: for example, any ‘state of affairs’ that makes  $p$  true also makes true the disjunction  $p \vee q$ , and it is not necessary to consider *all* possible assignments to the atomic propositions that occur in  $q$ . It is legitimate to wonder whether the elimination of these inefficiencies would not in itself be sufficient to render truly perfect the language of the truth-tables; or, alternatively, whether it is possible to contrive some other representation of propositions, maybe wholly different from that proposed by Wittgenstein, that really is able to satisfy his request for an ‘adequate notation’ in which the tautological character of a proposition proves to be ‘immediately’ perceivable. At this precise point the theory of computational complexity enters the scene. Its impact on the scientific research programmes descending from Leibniz’ dream and from the myth of the perfect language, which survived the limitative theorems of the 1930s without big traumas, has been very important and has involved a profound change of perspective whose philosophical consequences are not yet fully understood.

## 5. Perfect language and computational complexity

The theory of computational complexity can be considered a refinement of the

traditional theory of computability taking into account the *resources* (time and space) used by algorithms. Its principal innovation consists in having replaced the concept of ‘effective procedure’ with that of ‘feasible procedure’.<sup>21</sup> An *effective procedure* or *algorithm* by and large consists in a ‘mechanical method’, i.e. one executable *in principle* by a machine, to solve a given class of problems. An effective procedure is *feasible* when it can also be carried out by a machine *in practice*, and not only in principle. The expression ‘in practice’ involves a certain degree of vagueness that computational complexity researchers have removed by agreeing to consider as feasible, executable in practice, algorithms that can be performed in *polynomial* space and time. To clarify the meaning of this convention, let us observe first of all that a *problem* is usually associated with a class of strings of symbols, in which every string identifies a particular instance of the problem. Thus the ‘tautology problem’ (abbreviated with TAUT) can be associated with the class of strings of symbols that represent ‘well-formed formulas’ in a standard logical language, let us say the language of the *Principia Mathematica* of Russell and Whitehead. A solution to a particular instance of this problem consists, in this case, in an answer of the ‘yes-or-no’ type: ‘yes’ (usually encoded with the symbol ‘1’) if the string in question is a tautology, ‘no’ (usually encoded with the symbol ‘0’) if it is not. The string of symbols in (\*), which we used above to encode a particular instance of the ‘pigeonhole problem’, represents a logically inconsistent formula, and since the negation of an inconsistent formula is a tautology, a correct algorithm solving TAUT has to produce the answer ‘yes’ whenever it receives as input the negation of (\*). A problem like TAUT, whose solution consists in an affirmative or negative answer, is referred to as a *decision problem*. Not all interesting problems are decision problems. Indeed, many of the problems that are met in practice are described as *determination problems* (‘what is the area of a circle of radius  $r$ ?’), and among them a major role is played by *optimization problems*, problems that require finding an optimal solution within a set of possible solutions (‘given a graph and a starting node  $s$ , find a minimum pathway from  $s$  to a given other node of the graph’). Nevertheless, these problems can be transformed into equivalent decision problems, so that it is possible to simplify the analysis by making reference only to decision problems and to the resources employed by the algorithms solving them.

The *running time*  $T(n)$  of an algorithm measures the (maximum) number of steps that the algorithm has to perform as a function of the complexity of the input, that is to say, of the length of the string of symbols that encodes a particular instance of the problem. Likewise, the *running space*  $S(n)$  can be defined

---

<sup>21</sup> For an excellent exposition, still valid after thirty years, see Garey and Johnson (1979).

*Size of largest problem instance solvable in one hour*

<i>Running time</i>	<i>With present computer</i>	<i>With computer 100 times faster</i>	<i>With computer 1000 times faster</i>
$n$	N1	100 N1	1000 N1
$n^2$	N2	10 N2	31.6 N2
$n^3$	N3	4.64 N3	10 N3
$n^5$	N4	2.5 N4	3.98 N4
$2^n$	N5	N5+6.64	N5+9.97
$3^n$	N6	N6+4.19	N6+6.29

FIG. 1: *Effect of improved technology on several polynomial and exponential time algorithms (from Garey and Johnson 1979).*

as the (maximum) number of units of memory that the algorithm has to use as a function of the complexity of the input. The problem is: how do these functions grow with an increase in the complexity  $n$  of the input? There is widespread agreement in defining as *feasible* (i.e. solvable in practice) problems that can be solved by algorithms working in polynomial time, i.e. ones for which a polynomial  $p$  exists such that, for any input of complexity  $n$ , the algorithm yields an answer in a number of steps  $\leq p(n)$ . (Solvability in polynomial time implies solvability in polynomial space, so that the time factor is privileged in the definition of the feasibility of a problem.) The underlying idea is that polynomial algorithms do not produce the kind of combinatorial explosion that we have observed in relation to truth-tables.

If, instead, the most efficient algorithm possible works in super-polynomial time, for example in *exponential* time (that is, described by a function like  $2^n$ ), the problem is considered *unfeasible*. Although certain polynomial time algorithms may well be highly inefficient and, for some inputs, even much more inefficient than exponential time algorithms that solve the same problem – for example when the running time is described by a polynomial of a very high degree, such as  $n^{100}$  – it must be observed that (a) their *asymptotic* behaviour is always enormously more efficient than that of exponential time algorithms; this also implies that, in practice, the actual time employed by these algorithms is highly sensitive to technological innovation, leading to the construction of faster and faster computers, while the latter is wholly irrelevant for exponential time algorithms (in this connection see Figure 1); (b) in actual fact the polynomial time algorithms that have emerged from attempts to solve natural problems arising in any research areas have a running time described by a polynomial of a low degree (usually not above three).

The class **P** is the class of all *tractable* decision problems, i.e. those that can be solved through an algorithm working in polynomial time. The bad news is that many interesting problems met in mathematical and technological research, for which a decision procedure is known, do *not* belong to **P**. Among these there stands out the problem of establishing whether a certain proposition is a theorem of elementary Euclidean geometry. Although, as we have already mentioned, this problem was ‘solved’ by Tarski in 1951, twenty-three years later Fischer and Rabin proved that it is in fact an intractable problem.<sup>22</sup> Other decidable problems whose intractability has been shown are:

- The arithmetic of the addition of natural numbers (Fischer and Rabin 1974; decidability shown by Pressburger in 1929).
- The arithmetic of the multiplication of natural numbers (Fischer and Rabin 1974).
- The theory of linear orders (which is obtained by the first-order axioms that express transitivity, totality and anti-symmetry; non-elementary lower bound demonstrated by A. R. Meyer in 1975; decidability shown by Rabin in 1969).<sup>23</sup>

What can one say about TAUT, our problem of recognizing tautologies? It is licit to require that a ‘logically perfect language’ should be a language  $L$  in which:

1. Inside  $L$  the problem of tautology should be solvable in polynomial time: if it has to be possible to recognize a tautology ‘immediately’, simply by examining the propositional sign expressing it, that is to say, if one must be in a condition of ‘seeing’ from the sign itself that a certain expression is a tautology, then a minimum requirement is that there exists a polynomial time algorithm that allows one to make such recognition;
2. The translation from ordinary logical language into  $L$  *should* be feasible, that is to say, it should be possible to express in  $L$  what can be expressed in the ordinary logical language without producing any combinatorial explosion that would make the logically perfect language  $L$  ‘perfectly’ useless for the purposes of deductive practice.

Well, Wittgenstein’s ‘logically perfect language’, in which the propositional signs coincide with the truth-tables, satisfies the first requirement – given any

---

<sup>22</sup> See Rabin (1977).

<sup>23</sup> For all these results, see Rabin (1977).

formal language in which it is possible to represent the truth-tables, it is possible to recognize in polynomial time whether a string of symbols in this language encodes the truth-table of a tautology<sup>24</sup> – but does not satisfy the second, since the length of the translation increases, as we have seen, exponentially in relation to the length of the proposition translated. It is to be observed that this does not only happen for artificial examples – for ‘logical monsters’ constructed *ad hoc* in order to obtain the desired result – but is a general phenomenon that manifests itself as the number of atomic propositions involved increases and therefore, as we have seen, also concerns tautologies expressing entirely natural logical principles, of great utility in both mathematical and ordinary reasoning. Thus the ‘adequate’ notation proposed by Wittgenstein in the *Tractatus* is not a logically perfect language. But does another formal language exist that satisfies both our requirements? We can call this question ‘the problem of the logically perfect language’.

## 6. Cook’s theorem and the inevitable imperfection of logical languages

In 1971 Stephen Cook proved a result from which it immediately follows that the problem in question is very probably unsolvable (Cook 1971). To outline the meaning of this result we have first to introduce a new concept: that of *non-deterministic algorithm*. Let us consider the complementary problem of TAUT, i.e. the *satisfiability* problem, which we can abbreviate as SAT: given any string of symbols expressing a proposition (a ‘well-formed formula’ again) in a standard Boolean language, the problem is to determine whether an assignment of truth values (true or false) to the atomic propositions that occur in it exists – or, as Leibniz would have put it, whether there is ‘a possible world’ – that renders this proposition true. In such a case the formula in question is said to be *satisfiable*. Since a tautology is a formula that is true for all assignments of this type (true in ‘all possible worlds’), and given that a formula is true if and only if its negation is false, it follows that a formula is satisfiable if and only if its negation is not a tautology, so that SAT and TAUT are complementary problems.

Let us now imagine an ‘algorithm’ attempting to determine whether a given formula is satisfiable using ‘guesswork’, that is to say by assigning at random a truth value to each atomic proposition and verifying whether the assignment thus obtained makes the given proposition true. An algorithm of the

---

<sup>24</sup> For this purpose it is sufficient to ensure that the truth values are assigned correctly on each line and that the final column only contains “1”.

kind is ‘non-deterministic’ in that there are no instructions that exactly specify how the truth values that are assigned to the elementary propositions are to be chosen. Hence, the computation is no longer expressed as a sequence of steps, but as a *tree* of possible choices. Well, the running time of such a non-deterministic algorithm is the (maximum) number  $T(n)$  of steps that must be performed for any input of length  $n$ , in the case in which the choices made are always the ‘luckiest’ possible, that is to say, those that lead to recognizing the yes-instances of the problem (in our case to an assignment, if any, that makes the input formula true). In the case of SAT, supposing that the luckiest choices of truth values are always made for the elementary propositions, it can be verified in linear time that the assignment thus obtained is correct, and therefore SAT can be ‘solved’ in polynomial time by a non-deterministic algorithm that simply consists in guessing an assignment and then verify that it is correct.

The class of problems analogous to SAT, i.e. the ones that can be solved in polynomial time by a similar ideal ‘algorithm’ always making the best choices, is named **NP** (in which the letter ‘N’ stands for ‘non-deterministic’ and the letter ‘P’ for ‘polynomial’). The class **NP** can also be characterized, as emerges from our discussion, as the class of problems for which it is possible to verify in polynomial time whether a putative proof that the answer is “yes” is indeed correct. For example, a given assignment to the atomic formulas that makes a complex formula true can be seen as a proof that the formula in question is satisfiable (that is, it belongs to SAT), and it is easy to verify that the proof, once found, is correct. It has emerged that the answer to many important problems is very difficult to find through ordinary deterministic algorithms, although the yes-instances admit of proofs that can be easily verified once they have been found. These problems are therefore in **NP**, but have eluded so far every attempt to show that they are also in **P**. One of these recalcitrant problems is precisely SAT.

It is legitimate to wonder whether luck really is so essential in solving problems in **NP**. Given a problem in **NP** is it perhaps always possible to find a deterministic algorithm that solves it in polynomial time? If the answer were affirmative then **P** would be equal to **NP**; in the opposite case the two classes would be different. Given its enormous scientific and technological impact, this is one of the seven ‘problems of the millennium’ for whose solution the Clay Mathematics Institute has put up a prize of seven million dollars (one million for each problem).<sup>25</sup> The point is that, among the problems in **NP**, there are some that are *really* hard: not only have they always eluded the efforts of

---

<sup>25</sup> See the site of the Clay Institute: <http://www.claymath.org/millennium>.

researchers to find deterministic algorithms solving them in polynomial time, but it is also possible to show that if one of these were solved feasibly, then *all* problems in **NP** would be. An example is the famous ‘travelling salesman problem’, abbreviated as TSP: given a network of towns connected by roads and a numerical bound  $B$ , the problem is finding a tour, if any, that visits all the towns exactly once and has length no more than  $B$ . It can be shown that, for any problem  $\Pi$  in **NP**, a feasible procedure exists (i.e. one with polynomial running time) for translating any instance  $I$  of  $\Pi$  into an instance  $I'$  of TSP in such a way that  $I'$  is a yes-instance of TSP if and only if  $I$  is a yes-instance of  $\Pi$ . Therefore, if TSP were in **P**, that is to say, if it were a tractable problem, then every problem in **NP** would also be tractable and therefore **P** would be equal to **NP**. Problems of this type are called **NP-complete**. Another example of an **NP-complete** problem is the ‘problem of the partial sums’: given a finite set of integers, the problem is to determine whether it includes a subset such that the sum of its elements is zero. It can quickly be verified whether or not a given subset yields a solution to the problem, but no method is known for finding a solution that is significantly more efficient than checking, one by one, all the subsets (which are exponential in number). The list of the **NP-complete** problems increases very fast<sup>26</sup> and covers hundreds of interesting issues, belonging to a variety of research fields, which have always eluded any attempt to find an efficient algorithmic solution. Recently the problem of *sudoku*, well known to puzzle-solvers, has joined the list.<sup>27</sup>

Since all problems in **NP** are polynomially reducible to each of the **NP-complete** problems, the latter are all equivalent to one another (that is to say, polynomially reducible to one another). Therefore, either they are all tractable and **P=NP**, or they are all intractable and **P≠NP**. The dominant conjecture among researchers is that **P≠NP** and, accordingly, that all **NP-complete** problems are intractable. This conjecture is used in numerous application areas (also ones that are critical from the security point of view, such as cryptography) *as if* it were demonstrated. From this point of view, its epistemological status is no different than that of a well corroborated hypothesis in a theory belonging to the empirical sciences: we behave as if it were true, though aware that one day or another it could prove false.

---

<sup>26</sup> For an ample overview check the latest edition of Garey and Johnson (1979).

<sup>27</sup> In its most general version the *sudoku* problem requires inserting numbers between 1 and  $n^2$  in a matrix of  $n^2 \times n^2$  elements subdivided into  $n^2$  sub-matrixes of dimension  $n^2$ , so that no number appears more than once in every line, column and sub-matrix. For a proof of its **NP-completeness** see Yato (2003).

Cook's theorem, which opened up the way to the theory of **NP**-completeness, consists precisely in the statement that in a standard Boolean language:

*SAT is NP-complete.*

It is therefore highly plausible (though unproven) that in a standard Boolean language like that of Frege and Russell, SAT is an intractable problem. Since a proposition is a tautology if and only if it is not satisfiable, any solution to SAT is also a solution to TAUT, the tautology problem; therefore, in a standard Boolean language, we must expect TAUT to be intractable too.

The probable intractability of the tautology problem in a standard Boolean language immediately implies that our two requirements for a logically perfect language cannot *simultaneously* be satisfied. If in a given language  $L$  tautologies can immediately be recognized through mere *inspection* of the symbols (that is to say in linear or, at worst, polynomial time), it is then highly implausible, via Cook's theorem and the related conjecture that  $\mathbf{P} \neq \mathbf{NP}$ , that there is a feasible translation from the ordinary logical language to  $L$ . For, its existence would also imply the existence of an efficient deterministic algorithm to solve the tautology problem (in the standard language), which instead, according to the currently accepted conjecture, does not exist. It therefore seems that the dream of a perfect language, in which the answer to a question is contained in its very 'clear and distinct' formulation, is unattainable in practice even in the case of apparently very simple problems for which it has long been known that it can be attained in principle. Hence there is no such thing as *instant rationality* (even for the most basic problems).

## 6. Logic in a network of imperfect languages

The conclusion that we reached in the previous section forces us to revisit the old idea of the presumed 'tautological' character of propositional logic. How is it possible for a logical truth 'to say nothing' (T. 6.11) if recognizing the fact that it 'says nothing' is such a hard problem that, with all probability, it does not admit any practical algorithmic solution? We should conclude that what a proposition 'says' cannot be fully understood or communicated, except in the simplest cases (in which only few atomic propositions occur), even with the help of the fastest computers that can be built compatibly with the laws of nature. But such a conclusion certainly appears paradoxical and constitutes a particularly strong version of what the philosopher Jaakko Hintikka has called 'the scandal of deduction':



C.D. Broad has called the unsolved problems concerning induction a scandal of philosophy. It seems to me that in addition to this scandal of induction there is an equally disquieting scandal of deduction. Its urgency can be brought home to each of us by any clever freshman who asks, upon being told that deductive reasoning is ‘tautological’ or ‘analytical’ and that logical truths have no ‘empirical content’ and cannot be used to make ‘factual assertions’: in what other sense, then, does deductive reasoning give us new information? Is it not perfectly obvious there is some such sense, for what point would there otherwise be to logic and mathematics? (Hintikka 1973, p. 222)

Hintikka’s thesis, in brief, is that the idea of the ‘tautological’ or ‘analytical’ character of deductive reasoning clashes both with intuition and with the discovery, by Church and Turing in the 1930s, that quantificational logic is undecidable.

Hintikka proposes to resolve the scandal through a revision of the traditional notion of semantic information (from which it would result that logical reasoning does not produce any new information), on whose basis it is possible to argue that the truths of quantificational logic are *not* tautological.<sup>28</sup> Nevertheless, this conclusion cannot be extended to propositional logic, given that for the latter a mechanical decision procedure exists. Thus, according to Hintikka, “the term ‘tautology’ does characterize very aptly the truth and inferences of propositional logic. One reason for its one-time appeal to philosophers was undoubtedly its success in this limited area” (Hintikka 1973, p. 154). It is precisely this conclusion of Hintikka’s that is challenged, according to our reconstruction, by the consequences of Cook’s result discussed in the previous sections.<sup>29</sup>

The thesis that deductive logic (in general) is ‘tautological’ has been taken up more recently by Carlo Cellucci (1998, 2000) in the context of his criticism of the ‘closed world conception’, that is to say the traditional point of view according to which mathematical theories are closed systems that cannot exchange information with the environment and are based on the axiomatic method. According to this orthodox view, mathematics is correctly represent-

---

<sup>28</sup> For a detailed criticism of Hintikka’s proposal, which however does not address the themes discussed in this paper, see Sequoiah-Grayson (2008).

<sup>29</sup> A thorough discussion of the consequences of Cook’s theorem on the very idea that propositional logic is tautological, inspired by some ideas in D’Agostino and Mondadori (2000), can be found in D’Agostino and Floridi (2009). On the related theme of ‘logical omniscience’ see also D’Agostino (2010).

ed by formal systems in which, according to a famous analogy by Gottlob Frege, theorems are contained in axioms “like plants in seeds” (Frege 1884, § 8), and its truths are established in a rigorously deductive way. But, since deduction cannot increase information (given that logic is ‘tautological’), and indeed generally reduces it (the conclusions are usually weaker than the premises), the closed world conception fails to explain the utility and creativeness of mathematical reasoning. According to Cellucci, this *reductio ad absurdum* of the closed world conception shows, (a) that mathematics cannot be adequately described by formal systems, and (b) that the methods of mathematics cannot be contained in the narrow confines of traditional deductive logic, but must be sought in alternative logics, able to represent reasoning processes that increase the information content. To the ‘closed world conception’ Cellucci opposes an ‘open world conception’, according to which mathematical knowledge is more similar to a ‘distributed environment’ of open systems, each of which affords a partial representation of the corresponding domain and is able to exchange information with the others.

The image of the growth of mathematical knowledge emerging from this analysis is undoubtedly more realistic and interesting than the traditional one. Nevertheless, one of the premises on which it is founded, that is to say the presumed ‘tautological’ character of deductive reasoning, appears to be difficult to justify in the light of our discussion. If it is true that mathematics is a network of open systems, this is largely independent of the thesis that deductive reasoning is tautological, which can hardly be defended even in the elementary domain of propositional logic. Nevertheless, the consequences of the theory of **NP**-completeness that we have discussed here lend further support to Cellucci’s open world conception. In this connection, Cook’s theorem can be exploited to maintain that the ‘closed world conception’ is untenable even within (traditional) deductive logic! It is not possible – or more exactly it is highly unlikely – that there exists a *single* formal system for classical propositional logic that can be used *in practice* to solve *all* problems in this domain. We have seen that one of these formal systems, the method of truth-tables, fails miserably in the practical attempt to recognize the tautological character of a fundamental combinational principle, the so-called ‘pigeonhole principle’. We chose this principle, by way of example, because it is an ‘intuitively obvious fact’ widely used in mathematical practice, not an artificial problem concocted for the sole purpose of challenging a particular formal system. Furthermore, the class of tautologies that express this principle is difficult even to *prove* for most complete formal systems used in automated deduction (in the technical sense that the shortest proof has exponential length). We can therefore main-

tain that all these formal systems are *practically incomplete* and that their incompleteness emerges in relation to classes of logical truths whose mathematical meaning is unquestionable, quite independently of the intractability proof.

On the other hand, it is not difficult to construct formal systems for propositional logic in which the pigeonhole problem *can* ‘easily’ be solved (in polynomial time).<sup>30</sup> Cook’s theorem implies, however, that all these formal systems must also, with all probability, be ‘practically incomplete’ too, in that they will not succeed in recognizing in polynomial time other (infinite) classes of tautologies. Therefore the only hope, even in this elementary domain, is to build up a ‘distributed environment’ of logical systems that are able to exchange information with one another, each of which yields a partial but feasible representation of a fragment of propositional logic. It follows from Cook’s theorem that this ‘network’ of systems cannot be finite,<sup>31</sup> so that the search for a solution to our logical problems will necessarily be, in accordance with the open world conception, a potentially infinite process. The myth of the logically perfect language, of a characteristic language for logic that can save us from the toil of deduction, has to give way to a pluralistic vision in which, instead of a *single* ‘perfect’ language, there is a (potentially infinite) *variety* of logical languages, each of which is inevitably ‘imperfect’ and can only attain perfection in relation to a partial domain, to a tiny fragment of the universe of logical relations. It is only from such a continual interaction of imperfect languages, from the full heuristic unfolding of the ‘perfection/imperfection’ opposition, that there can gradually (and never completely) emerge a practical solution to the tautology problem, as well as to other general problems that for centuries have challenged the ability of logicians and mathematicians.

## REFERENCES

- BLACK, Max (1964): *Companion to Wittgenstein’s Tractatus*, Cambridge: Cambridge University Press.
- CELLUCCI, Carlo (1998): *Le ragioni della logica*, Bari: Laterza.
- (2000): “The Growth of Mathematical Knowledge: An Open World View”, in E. Grosholz and H. Berger (eds.), *The Growth of Mathematical Knowledge*, Dordrecht: Kluwer, pp. 153-176.

---

<sup>30</sup> For an example, see Dunham and Wang (1976).

<sup>31</sup> Otherwise the combination of these systems would afford a polynomial decision procedure for the tautology problem.

- CHAITIN, Gregory (2009): *The Search for the Perfect Language*, lecture delivered to the Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, on September 21, 2009; published in Portuguese in *Dicta & Contradicta*, 4, 2009, <http://www.cs.auckland.ac.nz/~chaitin/pi.html>.
- CONANT, James (2000): "Elucidation and Nonsense in Frege and Early Wittgenstein" in A. Crary and R. Read (eds.), *The New Wittgenstein*, London: Routledge, pp. 174-217.
- COOK, Stephen (1971): "The Complexity of Theorem-Proving Procedures", in *Conference Record of Third Annual ACM Symposium on Theory of Computing*, New York: ACM, pp. 151-158.
- COUTURAT, Luis (1903a): *Histoire de la langue universelle*, with Léopold Leau, Paris: Hachette.
- (1903b): *Opuscles et fragments inédits de Leibniz*, Paris: Alcan.
- D'AGOSTINO, Marcello (2010): "Depth-Bounded Boolean Logic and the Problem of Logical Omniscience", in H. Hosni and F. Montagna (eds.), *Probability, Uncertainty and Rationality*, Berlin: Springer (SNS Series).
- D'AGOSTINO, Marcello and FLORIDI, Luciano (2009): "The Enduring Scandal of Deduction. Is Propositional Logic Really Uninformative?", *Synthese*, 167 (2), pp. 271-315.
- D'AGOSTINO, Marcello and MONDADORI, Marco (1999): "La logica è davvero analitica?", in *Studi in Onore di Franco Crispini*, Bollettino Filosofico dell'Università della Calabria, 15.
- DUNHAM, Bradford and WANG, Hao (1976): "Toward Feasible Solutions of the Tautology Problem", *Annals of Mathematical Logic*, 10, pp. 117-154.
- DIAMOND, Cora (1991): *The Realistic Spirit*, Cambridge (Mass): The MIT Press.
- ECO, Umberto (1993): *La ricerca della lingua perfetta*, Roma-Bari: Laterza (English translation: *The Search for the Perfect Language*, Oxford: Blackwell, 1997).
- FEFERMAN, Solomon (2006): "The Impact of the Incompleteness Theorems on Mathematics", *Notices of the AMS*, 53 (4), pp. 434-439.
- FRASCOLLA, Pasquale (2006): *Understanding Wittgenstein's Tractatus*, London: Routledge.
- FREGE, Gottlob (1879): *Begriffsschrift, eine der arithmetischen Nachgebildete Formelsprache des reinen Denkens* (English translation, by S. Bauer-Mengelberg: *Begriffsschrift, a Formula Language Modelled upon that of Arithmetic, for Pure Thought*, in Van Heijenoort 1967a, pp. 1-82).
- (1880-81) "Booles rechnende Logik und die Begriffsschrift" (English translation, by P. Long and R. M. White: "Boole's Logical Calculus and the Concept-script", in Frege 1991).

- (1884): *Die Grundlagen der Arithmetik* (English translation, by J. L. Austin: *The Foundations of Arithmetic*, Oxford: Blackwell, 1980).
- (1897): “Über die Begriffsschrift des Herrn Peano und meine eigene”, in *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Mathematisch-Physische Klasse*, 48, pp. 361-378 (English translation: “On Mr. Peano’s Conceptual Notation and My Own”, in McGuinness 1984, pp. 234-248).
- (1915): “Meine Grundlegenden logischen Einsichten” (English translation: “My Basic Logical Insights”, in Frege 1991, pp. 251-252).
- (1991): *Posthumous Writings*, edited by H. Hermes, F. Kambartel and F. Kaulbach, New York: Wiley & Sons.
- GAREY, Michael and JOHNSON, David S. (1979): *Computers and Intractability. A Guide to the Theory of NP-Completeness*, New York: W.H. Freeman.
- GENSINI, Stefano (1991): *Il naturale ed il simbolico*, Roma: Bulzoni.
- HINTIKKA, Jaakko (1973): *Logic, Language Games and Information. Kantian Themes in the Philosophy of Logic*, Oxford: Clarendon Press.
- (1997): “Lingua Universalis vs. Calculus Ratiocinator. An Ultimate Presupposition of Twentieth-Century Philosophy”, in *Selected Papers*, Dordrecht: Kluwer, vol. II, pp. IX-X.
- KNEALE, William and KNEALE, Martina (1962): *The Development of Logic*, Oxford: Oxford University Press.
- KORTE, Tapio (2010): “Frege’s Begriffsschrift as a *lingua characteristic*”, *Synthese*, 174, pp. 283-294.
- KREISEL, George and KRIVINE, Jean Louis (1971): *Elements of Mathematical Logic*, Amsterdam: North Holland.
- LEIBNIZ, Gottfried Wilhelm (1679a): “Elementa Characteristicae universalis”, in Couturat (1903b, pp. 42-49).
- (1679b): “Elementa Calculi”, in Couturat (1903b, pp. 49-57) (English translation in Leibniz 1969, pp. 235-239).
- (1679c): *Sine titulo*, in Leibniz (1890, pp. 184-189) (English translation: “On the general characteristic”, in Leibniz 1969, pp. 221-228).
- (1684): *Sine titulo* (“De Scientia universali seu calculo philosophico”), in Leibniz (1890, pp. 198-203) (English translation: “On the General Characteristic”, in Leibniz 1969, pp. 221-228).
- (1696): “Letter to Gabriel Wagner”, in Leibniz (1969, p. 475).
- (1969): *Philosophical Papers and Letters*, edited by L. L. Loemker, London: Kluwer.
- (1890): *Philosophische Schriften*, edited by C. I. Gerhardt, vol. VII, Berlin.
- (2000): *Logical Papers*, Oxford: Oxford University Press.

- MARCONI, Diego (2005): "Frascolla on Logic in the *Tractatus*", in *Dialectica*, 57, pp. 97-107.
- MANGIONE, Corrado and BOZZI, Silvio (1993): *Storia della Logica. Da Boole ai nostri giorni*, Milano: Garzanti.
- MCGUINNES, Brian (ed.) (1984): *Collected Papers on Mathematics, Logic, and Philosophy*, Oxford: Basil Blackwell.
- MILKOV, Nikolay (2006): "A New Interpretation of Leibniz's Concept of *characteristica universalis*", in H. Poser (ed.), *Einheit in der Vielheit. Proceedings of the 8<sup>th</sup> International Leibniz-Congress*, Hanover, 2006, pp. 606-614.
- MONDADORI, Marco (1986): "Il sogno di Leibniz"; reprinted in M. D'Agostino, G. Giorello and S. Veca (eds.), *Logica e politica. Per Marco Mondadori*, Milano: Fondazione Arnoldo e Alberto Mondadori, 2001, pp. 15-25.
- MUGNAI, Massimo (1976): *Astrazione e realtà*, Milano: Feltrinelli.
- PEANO, Giuseppe (1908): *Formulario Mathematico [1895-1908]*, Torino: Bocca.
- (1925): *Giochi di aritmetica e problemi interessanti*, Torino: Paravia.
- PIANA, Giovanni (1973): *Interpretazione del "Tractatus" di Wittgenstein*, Milano: Il Saggiatore.
- RABIN, Michael O. (1977): "Decidable theories", in J. Barwise (ed.), *Handbook of Mathematical Logic*, Amsterdam: North-Holland.
- ROSSI, Paolo (2000): *Logic and the Art of Memory. The Quest for a Universal Language*, London: Athlone Contemporary European Thinkers (English translation of *Clavis universalis*, 2<sup>nd</sup> edition, Bologna: il Mulino, 1983).
- RUSSELL, Bertrand (1901): "Mathematicians and the Metaphysicians", in *Mysticism and Logic and Other Essays*, London: G. Allen & Unwin, 1917.
- (1918): *The Collected Papers of Bertrand Russell. Vol. VIII: The Philosophy of Logical Atomism and Other Essays (1914-1919)*, London: Routledge, 1986, pp. 157-246.
- STENLUND, Soren (2002): "On the Linguistic Turn in Philosophy", in M. Gustafsson and L. Hertzberg (eds.), *The Practice of Language*, Dodrecht: Kluwer, pp. 11-51.
- TARSKI, Alfred (1951): *A Decision Method for Elementary Algebra and Geometry*, Berkeley: University of California Press.
- VAN HEIJENOORT, Jean (ed.) (1967a): *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Cambridge (Mass.): Harvard University Press.
- (1967b): "Logic as Calculus and Logic as Language", *Synthese*, 17, pp. 324-330.
- WITTGENSTEIN, Ludwig (1921): *Logisch-philosophische Abhandlung/Tractatus Logico-philosophicus*, edited by B. F. McGuinness and J. Schulte, Frankfurt:

Surkamp, 2001 (English translation by D. F. Pears and B. F. McGuinness, London: Routledge, 1974).

YATO, Takayuki (2003): *Complexity and Completeness of Finding Another Solution and its Application to Puzzles*, Master Thesis, Graduate School of Science, University of Tokyo, <http://www-imai.is.s.u-tokyo.ac.jp/~yato/data2/MasterThesis.pdf>





# Ordine gratuito, morfogenesi autonoma e complessità semantica

Mirko Di Bernardo  
Dipartimento di Ricerche Filosofiche,  
Università di Roma Tor Vergata  
e-mail: [diber.mirko@gmail.com](mailto:diber.mirko@gmail.com)

1. Gratuità dei processi cellulari e sistemi di regolazione genica
2. Reti geniche e modelli NK
3. Ordine dinamico e ontogenesi
4. *Meaningful Complexity* e selezione naturale

SOMMARIO. Questo lavoro costituisce un tentativo di rivisitare i sistemi monodiani di regolazione genica e la fondamentale nozione di gratuità dei processi cellulari, quei processi vale a dire da cui scaturisce la morfologia di ogni organismo vivente, alla luce dei nuovi scenari aperti dallo studio del significato in biologia. Una legge potenziale universale, infatti, potrebbe emergere come la chiave di volta interpretativa dei sistemi adattativi complessi. Con tale spirito, quindi, l'articolo mostra come Kauffman, utilizzando il linguaggio della Dinamica per interpretare i fenomeni biologici, inventa un modello matematico (plausibile da un punto di vista biologico) capace di inserire il mistero dell'ontogenesi in un quadro teorico più ampio, in cui la biologia si trova improvvisamente in "dialogo" con altri saperi come per esempio la matematica, la fisica, la teoria del caos, l'informatica e la teoria dei sistemi. Tali considerazioni ci permettono, inoltre, di inferire che l'ipotesi del biochimico americano secondo cui i tipi cellulari sono attrattori dinamici è ancora attendibile. Infine, a partire dalle investigazioni scientifiche e dalle esplorazioni metodologiche di Kauffman, l'articolo mette in luce il fatto che oggi la nuova sfida della teoria della complessità consiste nell'ambizioso progetto di elaborare una teoria dell'informazione a carattere semantico (in cui la complessità semantica svolge un ruolo selettivo) capace di interpretare il misterioso linguaggio della vita, nonché il

complesso rapporto che sussiste, in ambito evolutivo, fra i processi profondi di auto-organizzazione e la selezione naturale.

PAROLE CHIAVE: teoria della complessità, auto-organizzazione, omeostasi, morfogenesi autonoma, ordine dinamico, teoria dell'informazione semantica, selezione naturale.

## **1. Gratuità dei processi cellulari e sistemi di regolazione genica**

Tutti gli organismi vivi, dalle singole cellule ai pluricellulari, possono produrre i propri componenti, in qualche modo possono auto-costruirsi.

Lo sviluppo embrionale è un esempio sorprendente di questa potenzialità. È un processo continuo di crescita e differenziamento nel quale i singoli eventi sono disposti a cascata a partire dal primo evento fondante: la fecondazione. Con questo termine si indica l'ingresso dello spermatozoo nell'uovo seguito, nell'immediato, dall'attivazione citoplasmatica e, dopo alcune ore, dalla fusione/riorganizzazione dei due nuclei, quello maschile e quello femminile. La comparsa di una cellula totalmente nuova a seguito della fusione delle due cellule sessuali può essere considerata come un caso particolare di endo-simbiosi, il meccanismo cellulare grazie al quale, nel corso dell'evoluzione, sono comparse le novità più significative. A seguito della fusione dello spermatozoo e dell'uovo avviene una tale riorganizzazione molecolare che l'individuo cellulare che ne deriva (lo zigote) risulta totalmente nuovo.

Nei mammiferi, e quindi anche nell'uomo, la riorganizzazione è accompagnata anche dall'entrata in gioco di nuove molecole, ma vi sono molte specie in cui questo non accade e la novità nasce esclusivamente dal cambiamento di organizzazione di molecole già esistenti. A essere nuovo è l'assetto organizzato e non le singole molecole. È nuova la forma che compare e che, una volta completata dal patrimonio genetico, si manterrà identica in tutte le cellule dell'embrione fino alla morula, ovvero una piccola massa di cellule staminali. A differenza della fecondazione, però, in questo caso la forma non cambia, ma viene conservata e i costituenti materiali coinvolti non sono molecole ma cellule.

Se la morula viene divisa in due parti, ognuna di esse può proseguire normalmente il proprio sviluppo embrionale e dar luogo a un organismo completo (i gemelli monovulari). La forma della morula di partenza è una sferetta con cellule esterne e cellule interne. Ed è la stessa forma che si deve necessariamente conservare nelle due morule figlie perché possano proseguire la cresci-

ta; è invece ininfluyente il fatto che siano costituite da un numero dimezzato di cellule.

Dopo la fecondazione, quindi, come abbiamo accennato, lo zigote umano subisce una rapida divisione cellulare che conduce verso la formazione di 8-16 blastomeri. Tali cellule migrano nelle tube di Falloppio ed entrano nell'utero. Durante la migrazione, inoltre, la massa cellulare si incava e forma una sfera chiamata "blastocisti". Circa una settimana dopo la fecondazione, la blastocisti si spinge in profondità nella parete dell'utero, dando inizio così alla fase dell'annidamento. Durante tale fase, dunque, le cellule esterne che avvolgono la blastocisti si aprono la via formando un primo abbozzo di placenta.

Già a questo stadio primitivo, pertanto, assistiamo ai due processi fondamentali dello sviluppo: il differenziamento cellulare e la morfogenesi. Nelle prime fasi dell'ontogenesi le cellule embrionali sono totipotenti, poi man mano che lo sviluppo procede, alcuni geni si disattivano, mentre altri rimangono attivi: le cellule si differenziano e non riescono più a tornare indietro. Stando così le cose, quindi, lo zigote rappresenta contemporaneamente sia una cellula singola che una singola specie cellulare la quale, in un secondo tempo, durante il corso dei circa cinquanta eventi di divisione cellulare che separano lo zigote dal neonato, darà vita a una moltitudine di tipi cellulari differenti. Sulla base di recenti calcoli (si veda Alberts *et al.* 2002), sono stati individuati nel corpo umano almeno duecentosessantacinque tipi diversi di cellule, tutte specializzate per esplicare funzioni specifiche nei tessuti e negli organi. Il differenziamento inizia con la formazione prima di due, poi di tre strati distinti: finissimi strati di cellule sovrapposte chiamati "foglietti embrionali". Questi tre foglietti contengono in sé il futuro sviluppo di tutti gli organi. Il primo strato, l'ectoderma (dal greco *ectòs*, "fuori"), fabbricherà le parti più esterne come la pelle, i peli, le unghie, ma anche il sistema nervoso e gli organi di senso. Il secondo strato, il mesoderma (da *mésos*, "medio"), fornirà l'impalcatura di sostegno come le ossa e i muscoli, ma anche il sangue e l'apparato urogenitale. Il terzo strato, l'endoderma (da *éndon*, "dentro"), produrrà gli organi interni per l'alimentazione e la respirazione (stomaco, fegato, polmoni).

Stando così le cose, quindi, il differenziamento cellulare consiste nell'aumento della varietà di specie cellulari, mentre la sua coordinazione in tessuti organizzati e organi è definita "morfogenesi". Come può dunque una singola cellula dare luogo alla complessità di un essere umano? Quali sono i meccanismi che regolano questa auto-organizzazione misteriosa? L'ordine naturale di cui risulta possibile fare esperienza quotidianamente è il frutto di milioni di anni di inesplicabili e improbabili incidenti, oppure insieme alla selezione naturale vi è anche un'altra forza che contribuisce a creare le stupefacenti architet-

ture della biosfera? In altre parole, esiste il caso puro nella natura? E, nell'eventualità in cui la risposta fosse affermativa, quale sarebbe, allora, il ruolo giocato dalla *randomness* nei processi evolutivi? Infine, come è possibile delineare oggi il rapporto esistente fra ordine e caos?

Nelle ricerche del 1953 Monod aveva chiarito come gli esseri viventi siano strutture derivanti dall'interazione di composti chimici che costituiscono il risultato di principi connessi necessariamente alla struttura della materia. Dal punto di vista chimico, gli organismi viventi sono caratterizzati fondamentalmente da meccanismi di regolazione. Un meccanismo di regolazione particolare fu scoperto da Monod negli anni sessanta: si trattava del meccanismo relativo all'azione degli enzimi allosterici grazie ai quali, a partire da quel momento, sarebbe stato possibile spiegare non solo le inibizioni da prodotto finale delle sequenze delle reazioni di biosintesi nei batteri o negli altri organismi, ma anche i rapidi ed efficienti controlli delle attività enzimatiche nei punti cruciali del metabolismo in generale. Dietro a tutto ciò vi era l'idea secondo cui l'attività di un enzima potesse essere regolata da interazioni chimiche diverse da quelle dirette che intercorrono tra enzima e substrato. Monod, Changeux e Jacob, dunque, nel rivoluzionario articolo del 1963 dal titolo "Allosteric Proteins and Cellular Control System", così scrivono:

L'attività biologica di molte proteine è controllata da specifici metaboliti che non interagiscono direttamente con i substrati o con i prodotti delle reazioni. L'effetto di questi agenti regolatori sembra scaturire esclusivamente da un'alterazione conformazionale (transizione allosterica) indotta nella proteina quando quest'ultima si lega con l'agente. Si suppone che questo meccanismo giochi un ruolo essenziale nella regolazione dell'attività metabolica e forse anche nel controllo specifico della sintesi della proteina. [...] Sembrerebbe, in altre parole, che alcune proteine, agendo in determinati passaggi metabolici, siano elettivamente dotate di specifiche funzioni di regolazione e coordinazione; attraverso l'azione di queste proteine, una data reazione biochimica alla fine viene controllata da un metabolita che si comporta apparentemente come un segnale fisiologico piuttosto che come un componente chimicamente necessario della reazione stessa (Monod *et al.* 1963, p. 306).

Questa idea costituiva la generalizzazione di un meccanismo già riconosciuto da Jacob e Monod a livello degli acidi nucleici la cui struttura, come quella delle proteine, dà luogo a funzioni complesse. A livello del metabolismo cellulare, inoltre, se pure a ogni stadio l'enzima a esso deputato esegue il suo

compito in modo impeccabile, la somma di tutte le attività si concluderebbe inevitabilmente nel caos se esse non fossero subordinate le une alle altre per formare un sistema coerente. È noto che negli animali esistono sistemi capaci di garantire, su vasta scala, la coordinazione dell'attività dell'organismo come, per esempio, nel sistema endocrino e nel sistema nervoso, i quali assicurano la coordinazione tra cellule. Le operazioni cibernetiche elementari, dunque, sono garantite da proteine specializzate il cui compito è quello di integrare e individuare l'informazione chimica.

Si presuppone che queste proteine possiedano due, o almeno due, siti ricettori non sovrapponibili e stereospecificamente differenti. Uno di questi, il sito attivo, lega il substrato ed è responsabile dell'attività biologica della proteina. L'altro, o sito allosterico, è complementare alla struttura di un altro metabolita, l'effettore allosterico, che si lega specificamente e reversibilmente. La formazione del complesso effettore dell'enzima allosterico non attiva una reazione che coinvolge l'effettore stesso: si suppone che esso apporti una discreta alterazione reversibile della struttura molecolare della transizione allosterica o della proteina, che modifica le proprietà del sito attivo, cambiando uno o più dei parametri cinetici che caratterizzano l'attività biologica della proteina (ivi, p. 307).

Gli enzimi allosterici, dunque, secondo Monod, “rappresentano un'unità funzionale chimica e, insieme, un elemento mediatore di interazioni regolatrici” (1970, p. 69). Essi, infatti, costituiscono una classe particolare in virtù delle proprietà che li distinguono dagli enzimi classici: come questi ultimi, essi riconoscono, associandovisi, un substrato specifico e attivano la sua conversione in prodotti; tuttavia, in più, riconoscono elettivamente uno o molteplici altri composti la cui associazione con la proteina ha l'effetto di modificare (aumentare o inibire) la sua attività nei confronti del substrato.

Sulla base di un complesso di fatti sperimentali, dunque, il biologo francese postula che le interazioni allosteriche siano dovute a transizioni discrete della struttura molecolare della proteina stessa. Alcune proteine, quindi, presentano due o più stati strutturali; la transizione allosterica per cui la molecola passa reversibilmente da uno stato all'altro modifica le proprietà di riconoscimento stereospecifiche della proteina in virtù delle diverse strutture steriche dei due stati.

Gli enzimi allosterici non sono quasi mai soggetti a una sola modalità di regolazione poiché, in linea di massima, essi dipendono contemporaneamente da parecchi effettori allosterici che possono essere antagonisti oppure cooperati-

vi. In tal senso, allora, situazione frequente è una regolazione ternaria che comprende: *a*) l'attivazione da parte del substrato e l'inibizione da parte del prodotto finale della sequenza di reazioni; e *b*) l'attivazione in parallelo da parte di un metabolita appartenente alla stessa famiglia del prodotto finale.

L'enzima riconosce simultaneamente i tre effettori, misura le loro contrazioni relative e, in ogni istante, la sua attività rappresenta la somma di queste tre informazioni. Per esempio, nello stato *R* la proteina potrà associarsi a un ligando  $\alpha$ , ma non ad un ligando  $\beta$  che sarà riconosciuto (contrariamente ad  $\alpha$ ) dallo stato *T*. La presenza di uno dei due ligandi farà sì che uno dei due stati si stabilizzerà alle spese dell'altro; poiché, inoltre, le loro rispettive associazioni con la proteina si escludono a vicenda,  $\alpha$  e  $\beta$  saranno antagonisti. Si prenda ora un terzo ligando  $\gamma$  (per esempio il substrato) che si associa esclusivamente con la forma *R* in un sito della molecola diverso da quello in cui si fissa  $\alpha$ . Monod fa constatare che  $\alpha$  e  $\gamma$  coopereranno alla stabilizzazione della proteina nello stato attivo, mentre il ligando  $\beta$  agirà da inibitore.

Da questo schema proposto da Monod, dunque, si evince un fatto di fondamentale importanza per il prosieguo di tutta la presente trattazione: le interazioni cooperative o antagoniste dei tre ligandi avvengono del tutto indirettamente; tra essi, infatti, non esistono interazioni, che si verificano solo fra la proteina e ciascun ligando preso separatamente.

Un'ipotesi implicita assolutamente essenziale, sebbene negativa, in questa descrizione è che un effettore allosterico, dal momento che si lega in un sito del tutto distinto dal sito attivo e dal momento che esso non partecipa, a nessuno stadio, nella reazione attivata dalla proteina, non deve possedere nessuna particolare reazione chimica o metabolica con il substrato stesso. Si considera quindi che la specificità di qualsiasi effetto allosterico e la sua attuale manifestazione debba risultare esclusivamente dalla costrizione specifica della molecola della proteina stessa, permettendole di subire una particolare e discreta alterazione conformazionale reversibile, innescata dalla legatura dell'effettore allosterico. L'assenza di qualsiasi analogia chimica obbligatoria che sia inerente o di qualsiasi reattività tra substrato ed effettore allosterico sembra essere un fatto di estrema importanza biologica [...]. In aggiunta sono evidentemente essenziali, per dare una definizione propria di effettori allosterici e distinguerli dall'azione dei coenzimi, dei substrati secondari o analoghi di substrati i quali reagiscono tutti con il substrato o con il sostituto e quindi devono possedere qualche relazione strutturale con esso o qualche reattività chimica verso di esso. Detto ciò, non si dovrebbe escludere la possibilità che l'azione di certi coenzimi o di altri effettori dell'enzima

possano coinvolgere degli effettori allosterici in aggiunta al loro ruolo classico come reattori transitori o trasportatori (*ibidem*).

Dal modello che Monod viene a tracciare, pertanto, una nozione di fondamentale importanza viene ben presto a emergere: la nozione di *gratuità*, cioè quella relazione arbitraria (di struttura e di reattività) dal punto di vista chimico tra il substrato di un enzima allosterico e i ligandi che attivano o inibiscono la sua attività. Già nell'articolo del 1961 dal titolo "General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation", Jacob e Monod, esaminando i meccanismi di regolazione cellulare, avevano messo in luce il fatto che nella regolazione per mediazione di una proteina allosterica, sorprendentemente, tutto è possibile:

Questa discussione ci suggerisce una possibilità particolarmente interessante. Cioè che, non essendoci ancora una volta nessuna correlazione obbligatoria tra substrati specifici e inibitori di enzimi allosterici, l'effetto non deve essere limitato all'inibizione del prodotto finale. Questa, infatti, è la ragione principale per evitare il termine *inibizione del prodotto finale* in una discussione generale di questo meccanismo (Monod e Jacob 1961, p. 391).

Stando così le cose, quindi, è soprattutto l'individuazione del concetto di *gratuità*, ovvero di "indipendenza chimica tra la funzione stessa e la natura dei segnali chimici a cui è subordinata", che caratterizza l'impostazione concettuale della ricerca monodiana; le istruzioni biologiche, infatti, costituiscono essenzialmente un messaggio, non sono cioè legate a un rapporto di necessità chimica con le molecole che le "interpretano". Con gli stessi componenti elementari del DNA si possono comporre i messaggi più diversi e proteine simili possono compiere azioni molto diverse negli opportuni contesti. I sistemi di regolazione dirigono l'attività delle cellule facendone così un'unità funzionale:

Tra il substrato di un enzima allosterico e i ligandi che attivano o inibiscono la sua attività non esiste alcuna relazione, chimicamente necessaria, di struttura o di reattività. La specificità delle interazioni è indipendente, in definitiva, dalla struttura dei ligandi: essa è determinata per intero da quella della proteina nei suoi diversi stati, struttura a sua volta liberamente e arbitrariamente dettata da quella di un gene. Ne deriva – ed è questo il punto fondamentale – che in fatto di regolazione per mediazione di una proteina allosterica tutto è possibile. [...] Il principio operativo delle interazioni allosteriche autorizza dunque una completa li-

bertà nella scelta delle subordinazioni che, sfuggendo a qualsiasi vincolo chimico, saranno in grado di obbedire meglio ai soli vincoli fisiologici, grazie ai quali verranno selezionati secondo la maggior coerenza e efficacia che conferiranno alla cellula o all'organismo. In definitiva è la gratuità stessa di questi sistemi che, aprendo all'evoluzione molecolare un campo di esplorazioni e di esperimenti praticamente infinito, le ha permesso di costruire l'immensa rete di interconnessioni cibernetiche che fanno di un organismo un'unità funzionale autonoma, le cui prestazioni sembrano trascendere le leggi della chimica, se non addirittura sfuggire ad esse (Monod 1970, p. 74).

Finora sono state analizzate le trasformazioni delle molecole di piccole dimensioni e la mobilitazione del potenziale chimico; quando tali prestazioni si analizzano in scala molecolare, quindi, appaiono, agli occhi di Monod, completamente interpretabili in termini di interazioni chimiche specifiche liberamente scelte e organizzate da proteine regolatrici: "È proprio nella struttura di queste molecole che si deve individuare la fonte ultima dell'autonomia o, più esattamente, dell'autodeterminazione che caratterizza gli esseri viventi nelle loro prestazioni" (*ibidem*). L'analisi delle interazioni allosteriche dimostra, dunque, come una sola molecola proteica si rivela già capace non solo di attivare elettivamente una reazione, ma di regolare la propria attività in funzione di molteplici informazioni chimiche. Risulta evidente allora che le prestazioni teleonomiche non sono proprietà esclusive dei sistemi complessi il cui ordine trova il proprio fondamento ultimo nella stupefacente armonia di interazioni chimiche libere:

La struttura e la reattività *sui generis* di un effettore allosterico sono importanti per l'interpretazione dei suoi effetti. Non vi rimane nessun reale paradosso chimico, una volta ammesso che un effetto allosterico è indiretto, che viene mediato interamente dalla proteina e che è dovuto a una specifica transizione della sua struttura. Comunque la gratuità, chimicamente parlando, di certi effetti allosterici appare quasi scioccante a prima vista, ma è questa stessa arbitrarietà che conferisce a tali effetti un significato fisiologico unico e l'interpretazione biologica dell'apparente paradosso diviene ovvia. La struttura specifica di ogni proteina dell'enzima, ovviamente, è un mero prodotto di selezione, necessariamente limitato, tuttavia, dalla struttura e dalle proprietà chimiche dei veri reagenti. [...] Una proteina allosterica regolatrice quindi deve essere considerata come un prodotto specializzato dell'ingegneria selettiva, che permette a una interazione indiretta, positiva o negativa, di avere luogo tra metaboliti che altrimenti non interagirebbero o addirittura non po-



trebbero reagire in alcun modo, eventualmente conducendo una particolare reazione sotto il controllo di un composto estraneo o indifferente. In questo modo è possibile comprendere come si possa essere stabilita, attraverso la selezione di adeguate strutture di proteine allosteriche, qualsiasi connessione di controllo fisiologicamente utile tra qualsiasi percorso in una cellula o qualsiasi tessuto in un organismo. È quasi superfluo precisare che il funzionamento chimico integrato di una cellula richiede l'esistenza di simili sistemi di controllo. Il punto importante della presente discussione è che questi circuiti di controllo non avrebbero potuto operare, vale a dire non sarebbero potuti evolvere, se i loro meccanismi elementari fossero stati limitati a interazioni chimiche dirette [...] tra percorsi differenti. Utilizzando alcune proteine non solo come catalizzatori o trasportatori ma come ricevitori molecolari e traduttori di segnali chimici, si ottiene la libertà da legami chimici che altrimenti sarebbero insuperabili, permettendo alla selezione di sviluppare e interconnettere il circuito immensamente complesso degli organismi viventi. È in questo senso che le reazioni allosteriche devono essere riconosciute come i componenti più caratteristici ed essenziali dei sistemi di controllo cellulare (Monod *et al.* 1963, p. 325).

Alla luce di tutto ciò, dunque, appare chiaro come la chimica cellulare comprenda anche un altro livello di sintesi: quello della sintesi delle macromolecole, ovvero degli acidi nucleici e delle proteine (le quali includono gli stessi enzimi). Anche a questo livello, infatti, esistono alcuni sistemi regolatori il cui studio però, afferma Monod, è molto più arduo di quello relativo agli enzimi allosterici.

Sulla base della fondamentale nozione di gratuità dei processi cellulari, quindi, lo studioso francese porta avanti le fondamentali ricerche concernenti i geni regolatori per le quali, nel 1965, gli viene assegnato il Premio Nobel per la medicina e la fisiologia insieme a François Jacob e André Lwoff. Tali ricerche forniscono, ancora oggi, le basi concettuali per interpretare il differenziamento cellulare e l'ontogenesi. La principale scoperta che portò alla comprensione delle procedure che presiedono, a livello biologico, al controllo genico, avvenne però negli anni cinquanta, con l'analisi genetica dettagliata operata da Jacob e Monod dell'azione degli enzimi del metabolismo del lattosio nel batterio *Escherichia coli*; tale scoperta condusse quindi, nel 1961, alla formulazione al livello del modello dell'operone per la regolazione genica.

L'informazione strutturale scritta come una sequenza di deossiribonucleotidi in un gene viene prima trascritta in una sequenza ribonucleotidica: il messaggero. Il messaggero si aggancia a un ribosoma, dove ha

luogo la trascrizione in una sequenza polipeptidica, gli aminoacidi vengono trasferiti dall'aminoacido S-RNA e posizionati lungo la sequenza da un appropriato accoppiamento di base tra messaggero e S-RNA. Questo sistema è controllato al livello della sintesi del messaggero da agenti specifici, i repressori, che sono in grado di riconoscere e legare loci genetici elettivamente certi chiamati "operatori", i quali fungono apparentemente da punti di iniziazione esclusivi per la prima trascrizione. Il segmento di DNA, la cui trascrizione è dunque coordinata da un dato operatore, può coinvolgere uno o più geni (cistroni); esso costituisce un'unità di espressione genetica chiamata "operone". La sintesi della proteina (o delle proteine) è pertanto governata dal repressore omologo che, a sua volta, viene sintetizzato sotto il controllo di uno speciale gene regolatore. Nella maggior parte dei casi, se non in tutti, l'attività del repressore, e cioè presumibilmente la sua abilità di legare il corrispondente operatore, è controllata da piccoli composti molecolari specifici che si comportano o da effettori positivi (attivando il repressore e bloccando in questo modo il messaggero e la sintesi della proteina) o da effettori negativi (inibendo il repressore e inducendo così la sintesi del messaggero e della proteina/proteine). Gli effettori positivi del repressore sono chiamati "co-repressori". Gli effettori negativi del repressore sono chiamati "induttori" (ivi, p. 325).

Il biologo francese, dunque, fornisce il primo esempio di un sistema di regolazione della trascrizione genetica attraverso lo studio concernente il controllo di espressione del gene dell'operone Lac (operone del lattosio). Come abbiamo dinanzi mostrato attraverso le parole di Jacob e Monod, l'operone è un'unità genetica del DNA che controlla la sintesi dello mRNA, nella sua composizione rientrano uno o più geni strutturali adiacenti che codificano la produzione di proteine e un gene operatore. Anche se il sistema sperimentale usato dai due scienziati era il batterio comune dell'*Escherichia coli*, i concetti di fondo della regolazione cellulare da loro scoperti sono validi per tutti gli organismi. Il "sistema lattosio" regola la sintesi di tre proteine nel batterio *Escherichia coli*: la galattosidopermeasi, che consente ai galattosidi di penetrare e di accumularsi all'interno delle cellule la cui membrana, in assenza di questa proteina, risulta essere impermeabile a tali zuccheri; una seconda proteina che idrolizza i  $\beta$ -galattosidi; ed infine una terza proteina che ha un'importanza minore. I componenti del sistema regolatore, come abbiamo accennato, sono il gene regolatore ( $G_R$ ), la proteina repressore ( $R$ ), il segmento operatore del DNA ( $O$ ), il segmento promotore del DNA ( $p$ ) e una molecola di galattoside induttore ( $\beta$ -g). Il gene regolatore coordina la sintesi della proteina repressore, la quale riconosce il segmento operatore e vi si associa formando così un composto stabile.

In questo stato la sintesi dell'mRNA viene bloccata a causa dell'impedimento molecolare. Il repressore, pur riconoscendo anche i  $\beta$ -galattosidi, vi si associa in modo stabile solo "allo stato libero"; pertanto, in presenza di  $\beta$ -galattosidi, "il complesso operatore-repressore si dissocia consentendo la sintesi del messaggero e, quindi, delle proteine" (Monod 1970, p. 71). Le due interazioni del repressore sono non covalenti e reversibili e l'induttore non è modificato dall'associazione con il repressore.

La logica di questo sistema è pertanto di un'estrema semplicità: il repressore inattiva la trascrizione ed è, a sua volta, inattivato dall'induttore. Da questa duplice negazione deriva un effetto positivo, cioè un'affermazione. La logica di questa negazione della negazione non è dialettica: non sfocia in una proposizione nuova ma nella reiterazione di quella originale, inscritta nella struttura del DNA in conformità al codice genetico. La logica dei sistemi biologici di regolazione dunque non obbedisce alla logica di Hegel ma all'algebra di Boole, alla stessa stregua di quella dei calcolatori (ivi, p. 73).

Secondo Monod, quindi, risulta possibile fornire un modello cibernetico del processo di strutturazione di una proteina globulare a partire dalla formazione stocastica delle sequenze polipeptidiche, un processo, cioè, la cui sorgente (emissione) è chiaramente casuale, in particolare markoviana, e di tipo sostanzialmente stazionario. La misura dell'incertezza relativa alla sorgente, dunque, è una misura di incertezza essenziale indipendente dalla nostra ignoranza; esistono, cioè, eventi per cui, sulla base dei dati presenti, è logicamente impossibile determinare una formula nella quale siano implicate le descrizioni di stato future.

Agli occhi del biologo francese, pertanto, ciò può avvenire a due livelli che sono quello della sorgente or ora accennata e quello delle mutazioni. Nel primo caso l'incertezza essenziale va intesa in senso relativo: abbiamo a che fare, infatti, con una funzione di probabilità  $M$  che assegna, sul piano logico, valori equiprobabili a tutte le descrizioni di stato e che corrisponde, sotto certi aspetti, a un automa 0-Limitato (si veda Carsetti 1972). A livello delle mutazioni, invece, abbiamo a che fare con un'incertezza intesa in senso assoluto attraverso l'intervento dello schema teorico della fisica quantistica che non possiamo approfondire in questa sede. I passi successivi della costruzione di Monod, inoltre, consistono nell'individuazione delle funzioni catalitiche, delle funzioni regolatrici e delle funzioni epigenetiche; lo studio di tali funzioni, infatti, permette allo studioso di passare dal livello delle proteine a quello della cellula: le funzioni catalitiche corrispondono molto bene alle funzioni selettive (funzioni conosciute in cibernetica come CPM), le funzioni regolatrici a tut-

ti i processi classici di retroazione e controllo, e infine le funzioni epigenetiche all'esecuzione del programma in un dato contesto ambientale.

Ciò che, a nostro giudizio, è importante rilevare in questa sede consiste nel fatto che tutte queste funzioni vanno inquadrare nell'ambito delle sorgenti stazionarie, per cui la formulazione di un modello integrato non è più lasciata all'inventiva del singolo, bensì viene determinata strettamente per passi successivi a partire dalle strutture iniziali. È questo il caso del modello del programma genetico ideato da Jacob e Monod per "interpretare" le dinamiche dello sviluppo.<sup>1</sup> Si può, pertanto, affermare, secondo i due scienziati, che immediatamente dopo il concepimento esiste in essenza il programma completo di sviluppo di un nuovo essere vivente e che questo programma è singolare e discriminante per ogni organismo. In ogni divisione cellulare si produce una replica esatta del genoma e degli altri componenti cellulari. Tutte le cellule degli esseri pluricellulari, quindi anche quelle dell'uomo, ricevono una copia identica dell'insieme di geni che c'è nella cellula originale unica. Alla luce di tutto ciò, dunque, Monod così scrive: "Nella misura in cui tutte le strutture e prestazioni dell'organismo sono la risultante delle strutture e delle attività delle proteine che lo costituiscono, si deve considerare tale organismo l'espressione epigenetica ultima del messaggio genetico stesso" (ivi, p. 202).

Se è chiaro quanto detto finora, allora non ha senso discutere su quando ha inizio la vita. Dal punto di vista genetico si constata che questo accade nel momento stesso della fecondazione: l'unico momento nel quale c'è un prima e un dopo rispetto all'identità genetica (Jacob e Monod 1961a, 1961b). Nel concepimento si costituisce un programma completo, diverso da quello dei genitori e capace di rivelarsi nel corso dello sviluppo dell'organismo pluricellulare in questione.<sup>2</sup> In questo spirito, quindi, proseguiamo la nostra disamina facendo innanzitutto riferimento alla proposta teorica di Kauffman sui processi di auto-organizzazione e di complessità emergente, espressa prima nel testo fondativo *The Origins of Order* (1993) e poi nel più "poetico" *A casa nell'universo* (1995).

## 2. Reti genetiche e modelli NK

Fin dagli anni settanta Kauffman era convinto del fatto che la fonte della biodiversità non poteva essere esclusivamente la selezione naturale: doveva es-

---

<sup>1</sup> Per un approfondimento di questo tema si veda Di Bernardo (2007a).

<sup>2</sup> Per approfondire la questione relativa ai limiti della nozione monodiana di programma genetico si vedano Atlan (1985) e Di Bernardo (2007b).

serci, infatti, qualcos'altro alla base del meraviglioso ordine della biosfera: "In qualche modo volevo riuscire a dimostrare che l'ordine è presente fin dall'inizio, senza dover essere costruito, né prodotto dall'evoluzione. Speravo di riuscire a dimostrare che l'ordine in un sistema regolatore di geni è naturale, quasi inevitabile. Doveva essere gratuito e spontaneo" (Kauffman 1995, pp. 40-41). Se tale ragionamento era corretto, allora la misteriosa auto-organizzazione della vita doveva essere l'altra faccia della selezione naturale. Stando così le cose, quindi, se da un lato i dettagli genetici specifici di ogni determinato organismo erano il prodotto di mutazioni casuali e della selezione naturale in perfetto accordo con la teoria di Darwin, dall'altro l'organizzarsi della vita, quindi l'ordine, doveva essere qualcosa di più profondo ed essenziale. In quegli anni quindi, agli occhi del giovane studioso, l'ordine costituiva una risposta al mistero dell'esistenza umana, ovverossia era in grado di spiegare la nostra condizione di creature viventi e pensanti in un universo apparentemente governato dalla casualità, dal caos e dalla cieca legge naturale.

A giudizio di Kauffman, dunque, l'ordine ci insegna che noi potremmo essere anche molto di più di un prodotto accidentale della natura. In effetti, secondo lui, Darwin aveva ragione nel sostenere che gli uomini e tutti gli altri sono senza dubbio eredi di quattro miliardi di anni di casualità nelle mutazioni, nelle catastrofi e nelle lotte per la sopravvivenza e non certo frutto di un intervento divino. Comunque, ci teneva a precisare che la selezione darwiniana non spiegava tutto. Darwin non sapeva nulla dell'auto-organizzazione, dei continui tentativi della materia di disporsi in strutture sempre più complesse, anche in presenza delle incessanti forze di dissoluzione descritte dal secondo principio della termodinamica; né tanto meno sapeva che le energie dell'ordine e dell'auto-organizzazione si riscontrano tanto nel crearsi dei sistemi viventi quanto nel formarsi dei fiocchi di neve o delle celle di convezione in una pentola di brodo bollente. Alla luce di tutto ciò, dunque, la storia della vita, a giudizio di Kauffman, era certamente la storia di eventi accidentali e casuali, ma anche quella dell'ordine: un tipo di creatività profonda, interna, intessuta della trama stessa della natura. Come era possibile, dunque, coniugare in seno alla teoria dell'evoluzione le mutazioni casuali e la selezione naturale così come pensate da Darwin con l'emergenza dell'ordine e l'auto-organizzazione che affondano le loro radici in un olismo irriducibile, nato non dal misticismo, ma dalla necessità matematica?

Per comprendere in che modo l'auto-organizzazione possa essere un motore dell'evoluzione, appare necessario soffermarsi sui sistemi complessi. Negli ultimi venticinque anni in tutti i settori delle scienze naturali e sociali si è manifestato un enorme interesse per tali sistemi, tuttavia questi studi sono anco-

ra così recenti che non esiste neppure una definizione di complessità che abbia validità generale. Ciò nonostante, alcune proprietà dei sistemi complessi stanno cominciando a chiarirsi. Il caos deterministico ha la capacità di generare aleatorietà: per effetto del caos, infatti, alcuni sistemi dinamici non lineari inizialmente ordinati possono con il tempo divenire affatto disorganizzati. Condizioni iniziali molto simili possono condurre a esiti diversissimi. In meteorologia, per esempio, il caos è esemplificato dal cosiddetto “effetto farfalla”: una farfalla che batte le ali a Rio de Janeiro può alterare il tempo a Chicago.

Tuttavia, pur essendo molto affascinante, il caos costituisce solo un aspetto del comportamento dei sistemi complessi. Esiste, infatti, anche un concetto per nulla intuitivo che Kauffman (1991), in un articolo dove vengono anticipate alcune nozioni fondamentali presenti in *The Origins of Order*, allora ancora in stampa, definisce con il termine *anticaos*, ovvero quel fenomeno secondo cui alcuni sistemi molto disordinati cristallizzano spontaneamente in uno stato altamente ordinato. Ogni sistema complesso presenta caratteristiche che si possono chiamare “locali” perché descrivono il modo in cui i singoli elementi del sistema sono collegati tra loro e il modo in cui possono interagire reciprocamente. Pertanto, dato un insieme qualsiasi di caratteristiche locali, si può costruire un ampio gruppo (o famiglia) costituito da tutti i sistemi complessi compatibili con quelle caratteristiche. Mediante una “meccanica statistica di nuovo tipo”, quindi, secondo Kauffman, si possono identificare le caratteristiche medie dei diversi sistemi della famiglia: i singoli sistemi della famiglia possono essere diversissimi, ciò nonostante le strutture e i comportamenti statisticamente tipici costituiscono la base migliore per prevedere le proprietà di ogni dato sistema.

Il primo passo compiuto in tal senso, allora, consiste nell’idealizzare il comportamento di ciascun elemento del sistema (di ciascun gene nel caso del genoma), assimilandolo a una semplice variabile binaria (attiva o inattiva). Per studiare il comportamento di migliaia di elementi fra loro accoppiati, quindi, Kauffman ricorre alle reti booleane stocastiche, ovvero al modello matematico di reti in cui ogni elemento di un insieme  $N$  di elementi ha due possibili stati e riceve  $K^3$  segnali dagli altri elementi. Il *comportamento dinamico* di ciascuna variabile, che essa sia attiva (1) o inattiva (0) nel momento successivo, è governato da una *regola di commutazione logica* o *funzione booleana*<sup>4</sup> che specifica l’attività della variabile regolata nell’istante seguente. Poiché ciascun

<sup>3</sup> Dove  $K$  sta per il numero di variabili di segnali che regolano un dato elemento binario.

<sup>4</sup> Nell’algebra di Boole tra gli elementi si possono compiere le operazioni binarie di congiunzione (E), di disgiunzione debole (O) e di disgiunzione esclusiva (O ESCLUSIVA). Se

elemento può essere attivo o inattivo, il numero di combinazioni dei  $K$  segnali è semplicemente  $2^K$ . Per ciascuna di queste combinazioni una determinata funzione booleana deve specificare se l'elemento che è stato regolato è attivo o inattivo. Dal momento che esistono due opzioni per ogni combinazione di stati di  $K$  segnali, il numero complessivo delle funzioni booleane  $F$  di  $K$  segnali è  $F = 2^{2^K}$ . Il numero di possibili funzioni booleane, quindi, aumenta con la stessa rapidità con cui aumenta  $K$  (Kauffman 1993, p. 188). Se una rete non riceve segnali dall'esterno del sistema viene considerata autonoma e il suo comportamento dipende unicamente da se stessa. Tale rete viene individuata scegliendo, per ogni elemento binario, quali elementi  $K$  fungeranno come suoi segnali regolatori e assegnando a ogni elemento binario una delle possibili funzioni booleane di  $K$  segnali. I modelli matematici dei sistemi biologici che Kauffman esamina in *The Origins of Order*, dunque, vengono definiti “reti booleane NK autonome stocastiche”. Queste ultime contengono  $N$  elementi, ciascuno dei quali ha  $K$  ingressi; a ogni elemento vengono assegnati a caso sia gli ingressi, sia una delle possibili funzioni booleane. Assegnando valori a  $N$  e a  $K$ , dunque, è possibile definire una famiglia di reti aventi le medesime caratteristiche locali; una rete stocastica è una rete scelta a caso in questa famiglia. La figura 1a alla pagina successiva mostra una rete booleana autonoma formata da tre elementi. Ognuno riceve segnali dagli altri due. L'elemento 1 è regolato dalla funzione E mentre gli elementi 2 e 3 sono regolati dalla funzione O. Il grande biochimico ci tiene a sottolineare come la classe più semplice di reti booleane sia simultanea, e ciò significa che tutti gli elementi aggiornano le loro attività a ogni istante. Così, ogni elemento esamina le attività dei suoi segnali  $K$ , consulta la sua funzione booleana e assume il successivo stato di attività prescritto. Tutto questo è riassunto da Kauffman nella figura 1b (alla pagina successiva) dove vengono riscritte le regole booleane. Ognuna delle  $2^3$  possibili combinazioni di attività dei tre elementi corrisponde a uno stato dell'intera rete: “Ogni stato ad un certo punto fa sì che tutti gli elementi valutino i valori dei propri segnali regolatori e che in un preciso momento assumano la giusta attività successiva. Su una successione di momenti, il sistema

---

si opera tra due elementi  $A$  e  $B$  entrambi veri, il valore della funzione di congiunzione tra di essi sarà anch'esso vero (affinché la funzione di congiunzione sia vera, tutti gli oggetti della funzione devono essere veri). Nell'operazione di disgiunzione debole, se  $A$  è vero e  $B$  è falso, allora la funzione è vera (almeno uno degli elementi deve essere vero). Nel caso dell'operazione di disgiunzione esclusiva abbiamo che la funzione assume valore vero se uno degli elementi  $A$  e  $B$  è vero, mentre assume valore falso se  $A$  e  $B$  sono entrambi veri o falsi. Kauffman utilizza le regole booleane sostituendo però alla variabile binaria “vero”/“falso” quella di “attivo”/“inattivo”.

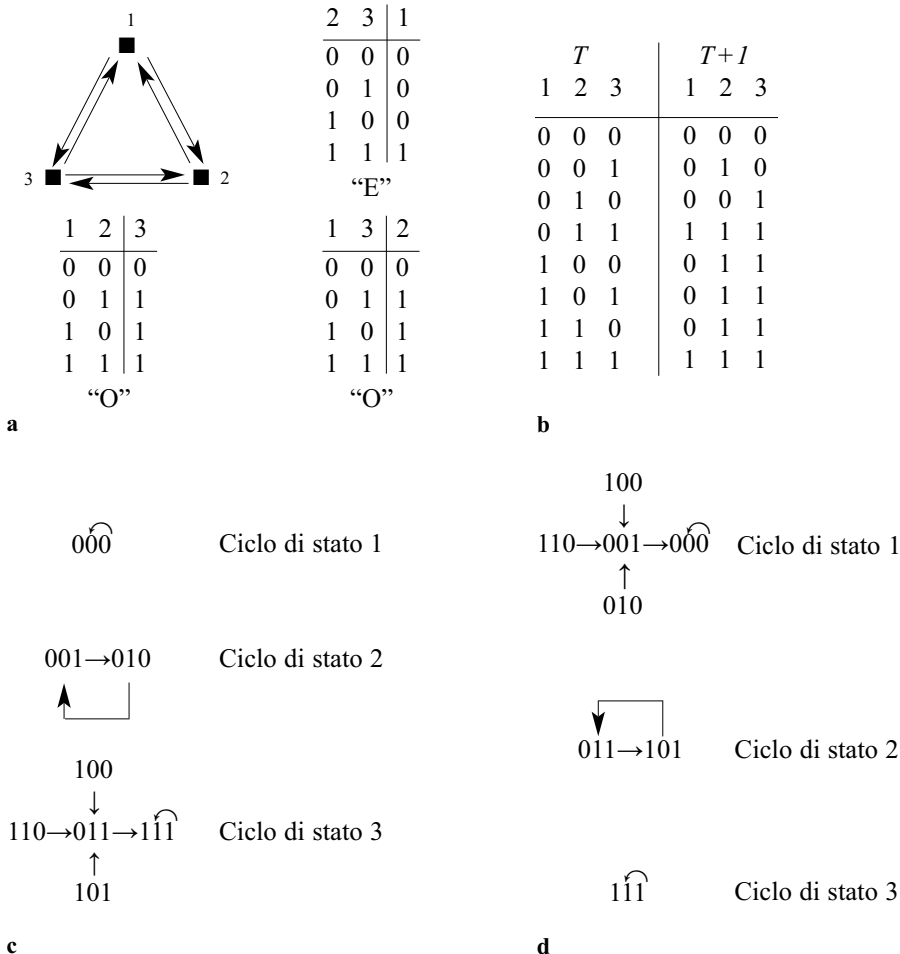


FIG. 1: Un diagramma booleano. (a) Diagramma di collegamento in una rete booleana con tre elementi binari, ognuno dei quali invia segnali agli altri due. (b) Regole booleane di (a) riscritte per mostrare, per tutti i  $2^3 = 8$  stati al tempo T, le attività assunte da ogni elemento al tempo successivo, T+1. Leggendo da sinistra a destra, questa figura mostra lo stato successivo per ogni stato di partenza. (c) Il grafico dello stato di transizione, o “campo di comportamento”, dell’automa rete booleana di (a) e (b), ottenuto mostrando stati di transizione verso stati successivi connessi da frecce. (d) Gli effetti di un cambiamento della regola dell’elemento 2 da O a E (tratto da Kauffman 1995).



passa attraverso una successione di stati chiama traiettoria” (ivi, p. 190; si veda la figura 1c). La caratteristica fondamentale di una rete booleana stocastica è che essa possiede un numero finito di stati, pertanto il sistema prima o poi deve tornare in uno stato che ha già assunto. Ebbene, dato il suo comportamento deterministico, il sistema ripeterà la stessa successione di stati che ha percorso in precedenza, continuando così a ripassare indefinitamente per lo stesso ciclo di stati. Questi cicli vengono definiti “attrattori dinamici” della rete:

Traiettorie differenti possono convergere verso un singolo stato che non cambia nel tempo – ovvero uno stato stabile – raggiungendolo in un limite di tempo infinito. Inoltre lo stato stabile è un attrattore zero-dimensionale, o puntiforme, e il suo bacino d’attrazione è un intero volume di stati che risiedono sulle traiettorie che fluttuano verso quell’attrattore. [...] L’idea dei bacini d’attrazione e degli attrattori di stato stabile puntiformi è essenzialmente paragonabile ad una regione montuosa con colline, rilievi, valli, laghi, ed un sistema di drenaggio delle acque. I laghi corrispondono ad attrattori puntiformi; i bacini di drenaggio, ai bacini d’attrazione. Così come una regione montuosa può avere molti laghi e bacini di drenaggio, allo stesso modo un sistema dinamico può avere molti attrattori, ciascuno dei quali drena il proprio bacino. Pertanto viene naturale concepire lo spazio di stato come ripartito in bacini di attrazione disgiunti. Quando viene rilasciato da uno stato iniziale, il sistema dinamico si trova su una traiettoria che giace su di un solo bacino, ed il sistema fluttua verso l’attrattore di quel bacino (ivi, p. 176).

Questa restrizione sta a significare che ogni bacino disgiunto conduce a un unico attrattore e quindi che i diversi attrattori costituiscono il numero totale dei comportamenti a lungo termine e alternativi del sistema. Generalmente un attrattore è definito come una serie di punti o stati nello spazio di stato in cui delle traiettorie, entro un certo volume di spazio di stato, convergono asintoticamente nel tempo (ivi, p. 178). Tuttavia, oltre a semplici stati stabili, i sistemi dinamici continui possono ammettere anche attrattori più complessi. Tra le diverse classi di attrattori complessi individuate da Kauffman quella degli attrattori strani o caotici, scoperta da Lorenz nel 1963 e successivamente rivisitata da Ruelle (1979) e da Mayer-Kress (1986) costituisce, senza dubbio, il cuore del modello matematico di reti autonome stocastiche che stiamo descrivendo.

Tali attrattori si differenziano dai cicli limite e dagli stati stabili essenzialmente per due proprietà. La prima riguarda la divergenza delle traiettorie sull’attrattore. Se, infatti, il sistema viene rilasciato da due punti sull’attrattore, che sono arbitrariamente vicini l’uno all’altro, le traiettorie che susseguono ri-

mangono sulla superficie dell'attrattore, ma divergono l'una dall'altra. Dopo aver fluttuato per un tempo sufficiente sull'attrattore, le due traiettorie possono poi trovarsi in una posizione arbitrariamente molto distante su di esso. La seconda proprietà concerne la bassa dimensionalità che può essere assunta anche in uno spazio di stato alto-dimensionale. Pertanto, un sistema può avere cento variabili, ma il flusso può essere ristretto a un attrattore strano a due dimensioni, una superficie ripiegata che si richiude su se stessa in quello spazio 100-dimensionale. Dal punto di vista dell'intero spazio di stato l'attrattore è in effetti un oggetto molto piccolo, il sistema è "in scatolato" in uno spazio di stato ridotto, nonostante il suo comportamento all'interno di quel piccolo volume sia caotico (nel senso di elevata sensibilità alle condizioni iniziali) (Kauffman 1993, p. 179).

Questo, agli occhi di Kauffman, costituisce un punto molto importante da rimarcare poiché il comportamento di un tale sistema che presenta un caos basso-dimensionale è molto più ordinato rispetto al comportamento di un sistema che vaga attraverso grandi tratti di spazio di stato su attrattori alto-dimensionali.<sup>5</sup> Nei testi successivi, però, lo studioso americano mostrerà come neanche i frattali siano modellazioni teoriche adeguate in grado di "interpretare" matematicamente la realtà profonda di quello strano fenomeno emergente che d'abitudine indichiamo con il termine "vita" (Kauffman 2000, pp. 170-183; 2008, p. 111; si veda anche Di Bernardo, di prossima pubblicazione). In accordo con l'ultimo Kauffman, dunque, a nostro giudizio, l'essenza della vita implica una complessità che resta mistero; non essendo tuttavia possibile approfondire ora queste considerazioni, appare opportuno ritornare ad analizzare le reti booleane stocastiche.

I cicli di stato delle reti hanno, come abbiamo già mostrato, lunghezza variabile fra 1 e  $2^N$ . Nel caso di lunghezza massima del ciclo, il sistema passerà per tutte le posizioni possibili, fino a ritornare al punto di partenza, ovvero alla condizione iniziale. Così, lasciata a se stessa, una rete finisce prima o poi in uno dei suoi attrattori e vi rimane; tuttavia, se viene di nuovo perturbata, la sua traiettoria può cambiare. La semplice rete booleana descritta nella figura 1a ha tre attrattori (Kauffman 1993, p. 190; si veda anche la figura 1c). Kauffman discute due tipi di perturbazioni: le perturbazioni minime e le perturbazioni strutturali. Una perturbazione minima in una rete booleana consiste nel capo-

---

<sup>5</sup> La dimensionalità di un attrattore strano spesso non è un intero. La definizione di dimensione dipende da come cambia la densità di punti sull'attrattore con un raggio in tutte le cento dimensioni, che sia lontano da qualsiasi punto arbitrario sull'attrattore stesso. Per un ulteriore approfondimento di questo tema si veda Mandelbrot (1977).

volgimento transitorio dell'attività di un elemento allo stato opposto (figura 1c). Il capovolgimento transitorio di qualsiasi elemento allo stato attivo fa in modo che il sistema si sposti in uno dei due bacini d'attrazione rimanenti. Ne consegue che questo attrattore è stabile a tutte le possibili perturbazioni minime. A volte, però, basta alterare l'attività anche di un solo elemento per scatenare nei regimi di comportamento una valanga di cambiamenti. Tali cambiamenti (mutazioni) sono danni che si possono propagare nella rete in varia misura. Una perturbazione strutturale, invece, secondo Kauffman, costituisce un'alterazione permanente dei collegamenti o delle funzioni booleane della rete. Queste perturbazioni includono lo scambio degli ingressi di due elementi o la trasformazione della funzione O di un elemento in funzione E (figura 1d). Come le perturbazioni minime, anche quelle strutturali possono causare danni e le reti hanno nei loro confronti gradi diversi di stabilità.

In *The Origins of Order*, capolavoro del 1993, Kauffman, integrando alcuni risultati di ricerche portate avanti a partire dagli anni settanta con le geniali intuizioni relative alla transizione di fase e al margine del caos risalenti alla metà degli anni ottanta, offre alla biologia teorica un nuovo paradigma in grado di coniugare in seno alla teoria dell'evoluzione le mutazioni casuali e la selezione naturale con l'emergenza dell'ordine e l'auto-organizzazione che affondano le loro radici in un olismo irriducibile nato da una precisa necessità a carattere matematico. Nel suo volume, infatti, presenta i risultati di circa trent'anni di lavoro. Egli esamina specificatamente i seguenti casi:  $K = N$ ,  $K = 1$  e  $K = 2$ . Nel primo caso ogni elemento riceve i segnali da tutti gli altri elementi, incluso se stesso. Poiché una rete stocastica  $K = N$  ha il massimo disordine, il successore di ciascuno stato è del tutto casuale: la rete, infatti, assume un comportamento caotico. Un indice del disordine di questi sistemi è che al crescere del numero degli elementi la lunghezza dei cicli di stato aumenta con legge esponenziale. Anche se ogni transizione di stato avvenisse in un microsecondo, ci vorrebbe un tempo miliardi di volte più lungo della vita dell'universo affinché questa rete percorra tutto il suo attrattore. Le reti  $K = N$  manifestano inoltre una sensibilità massima alle condizioni iniziali. Pertanto, variazioni anche minime di solito causano danni (alterazioni negli schemi di attività) cospicui e quasi sempre immediati. Questi aspetti caotici del comportamento e della struttura non riguardano solo le reti  $K = N$ , ma continuano a manifestarsi anche quando  $K$  decresce fino al valore di tre circa. Nel secondo caso si tratta di reti casuali in cui ogni elemento riceve solo un input. In tali reti non succede niente di rilevante poiché esse cadono velocemente in cicli di stato molto brevi al punto che spesso consistono in un solo stato. Nel terzo caso, invece, lo scenario è completamente diverso: nelle reti  $K = 2$ , infatti, tanto

il numero di cicli diversi quanto la loro lunghezza media si riducono fino a un valore circa pari alla radice quadrata del numero di elementi. I cicli di questi sistemi, dunque, sono stabili rispetto a quasi tutte le perturbazioni minime, mentre le perturbazioni strutturali alterano solo di poco il loro comportamento dinamico. Pertanto, nelle reti stocastiche con due soli ingressi per elemento, ciascun attrattore è stabile rispetto a quasi tutte le perturbazioni minime e la maggior parte delle mutazioni altera solo di poco gli attrattori. Il regime di rete ordinata è quindi caratterizzato da un comportamento omeostatico: dopo una perturbazione le reti tornano di norma al loro attrattore di partenza. Alla luce di quanto detto sinora, dunque, Kauffman scrive: “L’ordine spontaneo che abbiamo appena scoperto nelle reti  $K = 2$  e le loro generalizzazioni descritte sotto sottolineano una speranza concreta di spiegare buona parte dell’ordine che si è visto nel comportamento coordinato e ordinato dei sistemi regolatori genetici che sottolineano a loro volta l’ontogenesi in assenza di selezione” (1993, p. 202).

È, quindi, possibile domandarsi quali leggi permettono lo stupefacente ordine gratuito che si osserva nelle reti  $K = 2$ . La risposta, agli occhi di Kauffman, sembra risiedere nel fatto che tali reti sviluppano una maglia connessa, o nucleo congelato, di elementi. ognuno dei quali congela o nello stato 1 o nello stato 0. Il nucleo congelato crea dei muri di attraversamento, o filtrazione, che rompono il sistema in isole funzionalmente isolate di elementi non congelati i quali vengono separati dai muri di elementi congelati che impediscono ai primi di influenzarsi reciprocamente. La formazione di tali isole da un nucleo congelato filtrante sembra essere una condizione di ordine sufficiente nelle reti booleane. Il regime di confine in cui un nucleo congelato inizia a filtrare e in cui la regione non congelata comincia a rompersi in isole non congelate è la fase di transizione fra ordine e caos. Tutti i risultati ottenuti sinora indicano che una misteriosa fase limite separa le reti che si trovano in un regime ordinato (dove si forma un raggruppamento gigante congelato di elementi la cui attività è 1 o 0) da quelle che mostrano dinamiche caotiche in cui gli elementi fissi nell’attività 1 o 0 si riducono a piccolissime isole distanti tra loro. Proprio in mezzo vicino alla transizione di fase, proprio ai confini del caos, possono verificarsi i comportamenti più complessi, comportamenti, vale a dire, abbastanza ordinati da assicurare una stabilità, ma al contempo pieni di flessibilità e sorprese. Questo è veramente ciò che Kauffman intende per *complessità*:

L’idea centrale è che, se una rete è profonda nella fase di congelamento, allora non potrà verificarsi una grande computazione al suo interno. Nella fase caotica la dinamica è troppo disordinata per essere utile. Pic-

coli cambiamenti in un punto qualsiasi causano un danno alla maggior parte degli altri elementi del sistema. Al confine tra ordine e caos, invece, il regime congelato si scioglie e le isole non congelate, isolate funzionalmente, sono in lieve contatto mutevole l'una con l'altra. Sembra plausibile che il comportamento più complesso possa avvenire in questa regione di confine (ivi, p. 219).

Ai margini del caos emergono comportamenti dinamici interessanti. In corrispondenza di questa transizione di fase esisterebbero isole non congelate grandi e piccole. Le perturbazioni minime causano numerose valanghe piccole e poche valanghe grandi. I vari siti della rete possono pertanto comunicare tra loro (influenzare vicendevolmente il proprio comportamento) secondo una distribuzione che manifesta un andamento a potenza: i siti vicini comunicano spesso tramite molte piccole valanghe di danni, mentre i siti lontani comunicano più di rado tramite poche grandi valanghe. Inoltre, i sistemi in bilico tra ordine e caos possono ricoprire un ruolo fondamentale anche per quanto riguarda l'evoluzione poiché sembrano avere la migliore capacità di evolversi. Le mutazioni e la selezione naturale, così come divise da Darwin, possono migliorare un sistema biologico tramite l'accumulo di successive modificazioni secondarie nello stesso modo in cui con i rabberciamenti si può migliorare una tecnologia.

### 3. Ordine dinamico e ontogenesi

Riassumendo brevemente il modello del circuito genetico così come delineato da Jacob e Monod, abbiamo che la proteina repressore si lega al sito dell'operatore usando un sito specifico sulla proteina repressore. La molecola di allolattosio (derivato metabolico del lattosio) si lega a un secondo sito sulla proteina repressore. Il legame della molecola di allolattosio cambia la forma della proteina repressore, modificando la forma del suo primo sito, e quindi riduce l'affinità della proteina repressore con il segmento operatore del DNA. A questo punto l'allolattosio, legandosi a un secondo sito sul repressore, lo estrae dall'operatore permettendo altresì la sintesi della  $\beta$ -galattosidasi (gene che metabolizza il lattosio). Tuttavia, poiché l'allolattosio agisce mediante il sito allosterico della proteina repressore, ciò implica che la forma della molecola di allolattosio non abbia alcuna connessione con le conseguenze finali delle sue azioni, ovvero con la capacità di regolare l'attività dei geni. In questo caso, dunque, il substrato e l'inibitore competitivo hanno caratteristiche molecolari simili; nonostante ciò, l'allolattosio, poiché agisce sul sito alloste-

rico del repressore, può essere usato altrettanto bene come segnale per controllare la trascrizione di un gene che codifica, per esempio, l'actina, la miosina ecc.:

La forma della molecola controllore non deve avere necessariamente un legame con il prodotto finale del processo così controllato. Come entrambi gli autori hanno fatto notare, l'azione che si esplica mediante altri siti implica una libertà profonda da un punto di vista molecolare per creare circuiti genetici di logica arbitraria e complessa (Kauffman 1995, p. 136).

Jacob e Monod, quindi, hanno mostrato che i geni possono attivarsi e disattivarsi a vicenda e che i circuiti genetici possono avere configurazioni alternative di attività genica le quali costituiscono i diversi tipi di cellule di un organismo. Quale è dunque la struttura di una simile rete genetica? Quali regole misteriose governano il comportamento dei geni e dei loro prodotti nelle reti che controllano l'ontogenesi?

Abbiamo visto nel paragrafo precedente come Kauffman risponda a tali domande con i modelli di reti booleane. Ora lo stesso concetto è applicabile ai circuiti di regolazione genetica elaborati da Jacob e Monod. Kauffman, infatti, in *A casa nell'universo. Le leggi del caos e della complessità* "interpreta" i sistemi genetici di regolazione come reti booleane:

Possiamo immaginare il gene strutturale per la beta-galattosidasi attivo o inattivo, trascritto o non trascritto. Possiamo immaginare la proteina repressore legata o no al sito dell'operatore, cioè accesa o spenta. Possiamo immaginare l'allolattosio legato o no al secondo sito della proteina repressore. Sebbene questa sia sicuramente una semplificazione, possiamo comunque estenderla a reti di geni e di loro prodotti che interagiscono reciprocamente in enormi reti di circuiti di regolazione. [...] In questo contesto la proteina repressore è un segnale molecolare diretto verso l'operatore, mentre lo stesso operatore è un segnale di regolazione diretto verso l'attività del gene della beta-galattosidasi. Le funzioni, o regole, booleane [...] indicano le combinazioni di segnali molecolari che attivano o inibiscono l'attività di un dato gene. Per esempio, l'operatore è controllato sia dal repressore che dall'allolattosio. L'operatore è legato dal repressore, a meno che l'allolattosio si leghi al repressore ed estrometta il repressore dall'operatore. Così l'operatore è controllato dalla funzione booleana NON SE. Il gene che produce la betagalattosidasi è inattivo, ma solo se l'allolattosio non è presente (ivi, p. 140; si veda la figura 2).

Come abbiamo visto in precedenza, i modelli di reti booleane con  $N$  geni possono trovarsi in uno dei  $2^N$  stati in cui una combinazione differente di geni viene attivata o disattivata. Mentre i geni seguono le regole booleane e si attivano o disattivano a seconda delle indicazioni dei segnali molecolari che ricevono, tale rete segue una traiettoria nel suo spazio di stato, una traiettoria che alla fine converge verso un attrattore di ciclo di stato attorno al quale il sistema girerà da quel momento in avanti. Stando così le cose, dunque, Kauffman sostiene che grandi reti di geni possono mostrare spontaneamente l'ordine necessario per l'ontogenesi. Egli, infatti, individua nei piccoli attrattori da cui scaturisce un ordine immenso la fonte ultima dell'ordine dell'ontogenesi, un ordine gratuito che viene regolato dalle leggi dell'auto-organizzazione e che solo in un secondo tempo subisce l'intervento dalla selezione naturale:

Jacob e Monod hanno suggerito che i comportamenti alternativi stabili del loro circuito genetico – gene 1 attivo e gene 2 inattivo; gene 1 inattivo e gene 2 attivo – potevano essere i differenti tipi di cellule di una rete genomica. Io suggerisco che ogni buco nero attrattore di un ciclo di stato nell'immenso spazio di stato di una rete genomica sia un tipo cellulare differente. Una rete che consiste anche solo di un modesto numero di geni può esplorare potenzialmente un enorme spazio di stato. Ma se ho ragione, una manciata di attrattori la spinge in poche direzioni. A seconda di quale sia il ciclo di stato attorno a cui la rete sta orbitando, verranno attivati o disattivati differenti geni, e quindi verranno prodotte proteine diverse. La rete genomica si comporterà come un tipo differente di cellula (ivi, p. 142).

A partire da questa ipotesi, dunque, secondo lo studioso statunitense, si possono fare molte previsioni, poiché in questo nuovo quadro concettuale diverse caratteristiche dei sistemi genomici e dell'ontogenesi sembrano trovare la giusta collocazione. Dall'uomo all'idra i geni insieme all'RNA e ai prodotti proteici risultano essere collegati in reti interconnesse di interazioni regolatrici, reti di regolazione genomica, il cui comportamento collettivo coordina lo sviluppo dallo zigote all'adulto. Noi possediamo circa duecentosessantacinque tipi di cellule diverse, un moscerino della frutta ha circa quindicimila geni strutturali e forse sessanta tipi di cellule e così via. Quali sono, dunque, i principi misteriosi che stanno alla base dell'ordine spontaneo dell'ontogenesi?

Kauffman, esaminando i circuiti genetici dei batteri e degli organismi superiori, mette in luce tre proprietà che caratterizzano le reti genomiche. La prima si riferisce ai meccanismi di regolazione cellulare, ovvero ogni gene o altra variabile molecolare viene regolato direttamente da un numero piuttosto

piccolo di segnali molecolari. Per esempio, l'operatore del lattosio viene regolato da due segnali molecolari: l'allolattosio e la proteina repressore. La seconda riguarda le regole booleane che descrivono le attività dei geni. In particolare abbiamo che regole che descrivono le attività di geni differenti differiscono l'una dall'altra. Per esempio, l'operone del lattosio viene "de-represso" dal lattosio secondo la regola booleana del NON SE. Altri geni vengono attivati da segnali molecolari secondo le regole booleane O o E, o seguono regole più complesse. Infine la terza proprietà concerne un particolare sottogruppo di funzioni booleane: le funzioni di canalizzazione. Considerando i geni come variabili binarie (attivi o inattivi per quanto riguarda la trascrizione e presenti assenti per quanto concerne i loro segnali di controllo) e utilizzando la semplificazione delle funzioni booleane introdotta nel paragrafo precedente, geni noti vengono regolati da funzioni di canalizzazione, ovvero regole booleane che hanno la semplice proprietà per cui almeno uno dei segnali molecolari ha un valore, che potrebbe essere 1 o 0, che da solo può determinare completamente la risposta del gene regolato.

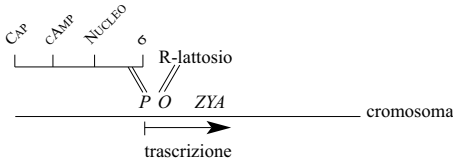
La funzione O è un esempio di funzione canalizzante. Un elemento regolato da questa funzione è attivo al momento successivo se il suo primo o il suo secondo input o entrambi sono attivi nell'attimo precedente. Se il primo input è attivo, allora siamo sicuri che l'elemento regolato sarà attivo nell'istante successivo indipendentemente dall'attività del secondo segnale. Quasi tutti i geni regolati che si trovano nei virus, nei batteri e negli organismi superiori di cui abbiamo conoscenza sono governati, nella semplificazione booleana, da funzioni booleane canalizzanti. Nel caso dell'operone Lac, per esempio, secondo Kauffman, il sito operatore è governato dalla funzione booleana canalizzante NOT IF<sup>6</sup>. Se il repressore è assente, allora il sito operatore è sicuramente libero, a prescindere dalla presenza o dall'assenza dell'allolattosio. Anche se l'allolattosio è presente, l'operatore è sicuramente libero, indipendentemente dalla presenza o dall'assenza della proteina repressore. Questo accade perché l'allolattosio si lega al sito allosterico sul repressore, estromettendolo dal sito operatore (ivi, p. 145). Il biochimico americano sottolinea come, sorprendentemente, quasi tutte le reti di regolazione genomica che possiedono queste tre proprietà esibiscano, sotto certi aspetti, la maggior parte dell'ordine presente nella biosfera.

Nel paragrafo precedente abbiamo visto che le reti booleane costruite casualmente si comportano seguendo tre regimi: caotico, ordinato e complesso

---

<sup>6</sup> La funzione booleana NON SE specifica l'attività dell'operatore nel momento successivo, dati ognuno dei quattro possibili stati dei segnali di regolazione (si veda la figura 2).





ALLOLATTOSIO	REPRESSORE	OPERATORE
0	0	0
0	1	1
1	0	0
1	1	0

“NON SE”

CAP	cAMP	NUCLEO	SIGMA	PROMOTORE
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

“E”

P	Op	ZYA
0	0	0
0	1	0
1	0	1
1	1	0

**b**

**a**

FIG. 2: Circuiti genetici. (a) La tavola superiore rappresenta l'operone del lattosio nel batterio E. coli. Z, Y e A sono geni strutturali; O è il sito operator, P è il sito promotore ed R la proteina repressore che si lega all'operatore e blocca la trascrizione, a meno che essa stessa non sia legata dal lattosio o dall'allolattosio. Il promotore è regolato da quattro fattori molecolari: l'AMP ciclico, il nucleo dello RNA polimerasi, il fattore sigma e la CAP. Tutti e quattro i fattori devono essere presenti perché la trascrizione abbia inizio. La seconda tavola mostra una funzione booleana che descrive la regolazione dell'operatore mediante il repressore e l'allolattosio. Per il sito dell'operatore, 0 = libero, 1 = legato. Per il repressore e il lattosio, 0 = assente, 1 = presente. La funzione booleana NON SE specifica l'attività dell'operatore nel momento successivo, dati ognuno dei quattro possibili stati dei segnali di regolazione. La terza tavola mostra la regolazione del promotore in base ai suoi quattro segnali molecolari. La funzione booleana è la funzione E di quattro variabili. (b) L'ultima tavola mostra la funzione booleana per la trascrizione dei geni strutturali nell'operone del lattosio. La trascrizione richiede che il promotore sia legato e attivo (1) e che l'operatore sia libero (0); questa è la funzione booleana NON SE (tratto da Kauffman 1993).

al confine con il caos. Per garantire che la maggior parte dei sistemi si trovino nel regime ordinato bastano due costrizioni, ovvero occorre che ogni elemento binario riceva  $K = 2$  o meno segnali. Oppure, se la rete ha meno di  $K = 2$  segnali per nodo, certe tendenze delle regole booleane espresse dal parametro  $P$  possono essere modificate per ottenere l'ordine richiesto. La maggior parte delle funzioni booleane con un gran numero di segnali non sono canalizzate. Kauffman fa l'esempio della funzione della disgiunzione esclusiva con due input: l'attività dei segnali 1 o 0 da soli non è sufficiente per garantire l'attività del gene regolato:

Probabilmente non è casuale che i geni regolati, e la maggior parte degli altri processi biochimici, sembrano governati da funzioni canalizzanti, perché le funzioni canalizzanti sono rare tra le possibili funzioni booleane, e diventano sempre più rare all'aumentare del numero dei segnali,  $K$ . Ma sono semplici da costruire chimicamente. Perciò l'abbondanza di funzioni canalizzanti riflette sia la selezione di una rara specie di regola booleana che una semplicità chimica. In entrambi i casi l'abbondanza di funzioni canalizzanti sembra essere molto importante per il comportamento ordinato dei sistemi genomici regolatori (ivi, p. 146).

Il numero di funzioni booleane possibili con un numero  $K$  di segnali differenti, come abbiamo visto, è  $2^{2^K}$ . Se consideriamo il caso  $K = 2$  input abbiamo che, a parte la funzione della disgiunzione esclusiva e della sua complementare SE E SOLO SE, quattordici delle sedici funzioni booleane risultano essere canalizzanti. Diversamente, per  $K = 4$  input solo il 5 % è canalizzante. Da ciò, dunque, si evince facilmente che il numero dei segnali  $K$  e quello delle corrispondenti funzioni canalizzanti sono inversamente proporzionali. Le funzioni canalizzanti sono semplici da costruire da un punto di vista molecolare. Si consideri, per esempio, un enzima con due segnali, esso risulterà attivato se uno dei due input sarà attivo. Per realizzare un enzima che rappresenti la funzione canalizzante O, quindi, basta creare un enzima con un solo sito allosterico: il legame di uno dei due input molecolari al sito altera la conformazione dell'enzima e lo attiva. Per realizzare un enzima che rappresenti la funzione non canalizzante della disgiunzione esclusiva, invece, servirebbero due distinti siti allosterici. Il legame del suo input molecolare (effettore) a uno solo dei due siti, dovrebbe alterare l'enzima per permettere l'attività. Il legame simultaneo a entrambi i siti allosterici, o a nessuno dei due, infatti, non dovrebbe consentire la medesima modificazione che attiva l'enzima. In linea di principio, dunque, secondo Kauffman, questa "macchina chimica" è certamente possibile da ottenere, tuttavia senza ombra di dubbio appare molto più complicata della funzione O; pertan-

to, risulta molto più semplice “costruire” una “macchina molecolare” che realizza le funzioni canalizzanti piuttosto che quelle non canalizzanti:

L'importanza dell'evidente semplicità chimica delle funzioni canalizzanti è la seguente: grandi reti di elementi binari governate principalmente da funzioni canalizzanti si collocano spontaneamente in un regime ordinato. Un'enorme quantità di ordine gratuito è disponibile per l'ulteriore scrematura della selezione. Se le funzioni canalizzanti nelle cellule sono abbondanti perché sono chimicamente semplici, allora la stessa semplicità chimica è sufficiente per produrre un ordine spontaneo su larga scala (ivi, p. 147).

Ebbene, in questo ambito di discorso, ciò che interessa rilevare è, però, che quest'ordine spontaneo risulta cruciale per comprendere il comportamento del genoma: “Il dogma centrale della biologia dello sviluppo afferma solo che specie cellulari differenti rappresentano differenti configurazioni di attività dello stesso sistema genomico” (*ibidem*). Tali considerazioni, tuttavia, non ci aiutano molto, poiché il genoma umano presenta un numero di combinazioni di attività genica talmente grande che, pur utilizzando la semplificazione del modello di reti binarie (attivo-inattivo) il quale non prende in considerazione né i diversi gradi di espressione dei geni né i differenti livelli di attività degli enzimi, le cellule non avrebbero la possibilità di esplorare un simile campo di configurazioni nell'intera esistenza di tutte le creature possibili. A questo punto, dunque, secondo Kauffman, si impongono alcune considerazioni di ordine generale e metodologico:

Il mistero incomincia a diradarsi se consideriamo la possibilità che, a causa della maniera in cui sono costruite, le reti genomiche si posizionino nel regime ordinato. Invece di vagabondare per tutta la mappa, queste reti vengono attratte da una manciata di attrattori, buchi neri nello spazio di stato genomico. Una cellula nell'orbita di un particolare attrattore esprimerà certi geni e proteine che la faranno comportare in un certo modo. La stessa cellula, orbitante attorno ad un attrattore differente, esprimerà altri geni e proteine. Così la nostra ipotesi afferma che le diverse specie cellulari sono attrattori nel repertorio della rete genomica. In questo quadro molte delle proprietà conosciute dell'ontogenesi trovano un loro posto. Ogni tipo cellulare deve essere confinato in una frazione infinitesimale delle configurazioni possibili dell'attività genica. Proprio questo comportamento sorge spontaneo nel regime ordinato (ivi, p. 148).

Applicando il modello NK al genoma umano, quindi, Kauffman mostra come la lunghezza degli attrattori dei cicli di stato cresca con la radice quadrata del

numero dei geni. Il grande studioso statunitense, dunque, tenta di spiegare perché i centomila geni dell'*Homo sapiens*<sup>7</sup> dirigono il differenziamento di soli duecentocinquantesi tipi cellulari quando, a livello teorico, sono possibili  $10^{30000}$  diversi tipi di attività. Variando il valore di  $K$  fino a farlo diventare uguale a 2 (ogni elemento della rete è connesso con altri due elementi), infatti, il numero dei cicli di stato è uguale alla radice quadrata di  $N$ , ovvero a 317, che è approssimativamente il numero dei tipi cellulari degli organismi umani. I piccoli attrattori del regime ordinato, dunque, costituiscono l'ordine gratuito.

Il tempo previsto necessario perché cellule compiano un ciclo attorno ai propri attrattori è completamente plausibile da un punto di vista biologico. Ci vuole un tempo nell'ordine da uno a dieci minuti per attivare o disattivare un gene. Di conseguenza il tempo per compiere un'orbita attorno al ciclo di stato dovrebbe variare da 317 a 3170 minuti, o da circa cinque a cinquanta ore – proprio all'interno del campo di comportamenti della cellula! In effetti il ciclo più ovvio cui una cellula va incontro è il ciclo di divisione cellulare. Nei batteri in fase di crescita rapida, il ciclo cellulare impiega circa venti minuti. Vi sono cellule che rivestono l'intestino in un'area chiamata cripte di Leberkuhn che compiono un ciclo ogni otto ore; altre cellule del nostro corpo compiono un ciclo in circa cinquanta ore. Così se i tipi di cellule corrispondono ad attrattori di cicli di stato, allora un ciclo cellulare può essere visto come una cellula che attraversa la lunghezza del suo ciclo di stato. E la scala di tempo per attraversare l'attrattore è la scala di tempo reale del ciclo cellulare. Una rete genetica con due segnali per gene ( $K = 2$ ), o una rete piena di funzioni canalizzanti, non solo esibisce un ordine spontaneo, ma un ordine simile a quello riscontrato nelle cellule reali. [...] Perfino reti con  $K = 4$  o  $K = 5$ , che si trovano già ampiamente in un regime caotico, hanno dei cicli di stato estremamente lunghi. Qui consideriamo delle reti genomiche costruite in maniera completamente casuale, soggette solo a costrizioni note presenti nelle reti genomiche reali, e troviamo dei tempi di ciclo approssimativamente esatti da un punto di vista biologico (ivi, p. 149).

Kauffman nel suo volume mostra come, in organismi che vanno dai batteri al lievito fino agli esseri umani, la lunghezza dei cicli cellulari aumenta con la radice quadrata  $\sqrt{N}$  del numero dei geni: la durata media del ciclo cellulare, infat-

---

<sup>7</sup> I dati forniti da Kauffman nel testo del 1995 sono imprecisi: oggi, infatti, sappiamo che il genoma umano è costituito da circa trentamila geni e che i tipi cellulari umani attualmente conosciuti sono duecentosessantacinque.

ti, varia all'incirca come la radice quadrata del numero dei geni presenti nell'organismo. Tale grafo, inoltre, permette al biochimico di inferire che,

come predetto anche dal nostro modello, la distribuzione attorno alla media è fortemente asimmetrica poiché la maggior parte delle cellule, a diversi gradi di complessità genomica, possiede un breve ciclo cellulare, mentre poche di esse esibiscono un ciclo lungo. La stessa distribuzione asimmetrica viene riscontrata nel modello delle reti genomiche con  $K = 2$  input per gene. [...] Anche se è stata fatta una grande quantità di lavoro sul ciclo cellulare, non ne sappiamo comunque abbastanza per riscontrare qualcosa di più di una notevole corrispondenza statistica tra la teoria e le osservazioni (ivi, p. 150).

Kauffman, inoltre, mette in evidenza un'altra corrispondenza significativa. Si tratta del rapporto fra il logaritmo del numero dei tipi cellulari in organismi di molti *phyla* diversi e il logaritmo del DNA contenuto in ogni cellula. Sulla base di risultati accumulati in diversi anni di lavoro si evince la linearità di tale rapporto: il numero dei tipi cellulari, infatti, aumenta secondo la radice quadrata del DNA contenuto nella cellula. Pertanto, assumendo che il numero di geni strutturali e regolatori sia proporzionale al DNA contenuto in ogni cellula, lo studioso statunitense conclude che il numero di tipi di cellule aumenterà in funzione della radice quadrata del numero dei geni (secondo il grafo, inoltre, il numero dei geni aumenta dai batteri agli esseri umani). Se, come sostiene Kauffman, un tipo cellulare è un attrattore di ciclo di stato, allora dovremmo essere in grado di prevedere il numero di tipi cellulari come una funzione del numero dei geni presenti nell'organismo. Con  $K = 2$  input per gene, e più generalmente con delle reti canalizzanti, il numero medio di attrattori di ciclo di stato è solo approssimativamente la radice quadrata del numero dei geni. Se questa teoria è corretta, dovremmo essere in grado di prevedere la relazione fra il numero di geni e il numero di tipi cellulari che dovrebbero aumentare proporzionalmente alla radice quadrata del primo valore.

Con il numero di geni sull'asse delle  $x$  e il numero di tipi cellulari sull'asse delle  $y$ , i dati mostrati da Kauffman in diversi grafi sembrano confermare questa previsione: il numero dei tipi cellulari aumenta, in effetti, all'incirca in funzione della radice quadrata del numero dei geni. Nel 2004, tuttavia, l'International Human Genome Sequencing Consortium, il consorzio pubblico che tutt'ora sta lavorando sulla mappatura del genoma umano, ha rimarcato il fatto che nell'*Homo sapiens* il numero dei geni codificanti per le proteine è approssimativamente compreso fra venticinque e trentamila. Tali ricerche hanno messo in luce, inoltre, come nella scala evolutiva la complessità biologica de-

gli organismi non sia proporzionale al numero dei geni. Tuttavia, è bene precisare il fatto che, nonostante gli erronei dati numerici, il modello delle reti geniche elaborato da Kauffman, a seguito di adeguati accorgimenti, rimane comunque valido. Per continuare a “interpretare” il differenziamento cellulare attraverso il linguaggio dei sistemi dinamici, ovvero analizzando lo spazio delle fasi, si potrebbe pensare, infatti, di modificare il modello NK riducendo il valore di  $N$  ( $25000 < N < 30000$ ) e aumentando il valore  $K$  per poter osservare, altresì, nel sistema, un comportamento che dia luogo a un numero di attrattori che approssimi il numero di tipi cellulari umani. Questo dato, comunque, mostra che, considerando il comportamento del sistema genomico nello spazio delle fasi, quando ogni gene è connesso con più di due altri geni ( $K > 2$ ), il genoma umano non si pone più nella transizione di fase, ai confini con il caos, ma si stabilizza nella zona caotica dello spazio in cui cicli di stato diventano estremamente lunghi.

Come abbiamo già accennato nei paragrafi precedenti, per cogliere il comportamento ordinato nelle reti in cui  $K > 2$ , Kauffman introduce la variabile  $P$ , che denota l’omogeneità interna della rete booleana. Con tale termine si indica la disposizione degli elementi della rete ad assumere lo stesso valore (per esempio, in una rete di quattro segnali in cui l’elemento considerato assume, su sedici possibili stati, quindici valori identici, il valore diventa:  $P = 15/16 = 0,9375$ ). Quando il valore  $P$  tende ad aumentare, allora la rete, anche se  $K > 2$ , può mostrare comportamenti ordinati. A ogni modo, come lo studioso statunitense ha mostrato in alcuni articoli recenti (Kauffman 2004a; Ribeiro e Kauffman 2007), anche quando nella rete  $K \geq 5$  il sistema mostra un’elevata sensibilità alle condizioni iniziali, nel senso che delle piccole differenze iniziali crescono con il tempo nell’attività del modello, la sensibilità degli attrattori alle perturbazioni appare comunque modesta. Nonostante tali accorgimenti, però, i dati forniti dallo Human Genome Sequencing Consortium sembrano comunque smentire le leggi di elevamento a potenza da lui ipotizzate con tanto entusiasmo in particolare nei testi del 1993 e del 1995. Egli, tuttavia, in uno dei suoi più rilevanti articoli pubblicati di recente, così si esprime:

Le nostre intuizioni a proposito dei requisiti per l’ordine si sono rivelate semplicemente errate. I risultati che mettono a confronto questa legge di scala della radice quadrata con il comportamento periodico delle cellule, nonché le durate dei cicli cellulari negli organismi come funzione del loro DNA per cellula, appaiono interessanti. Il periodo del ciclo cellulare medio è all’incirca una funzione della radice quadrata del DNA totale. L’avvertimento naturalmente è che molto di quel DNA è DNA spazzatura. Pertanto, la legge di scala, rispetto alle regioni codificanti e

ai siti *cis*, è più ripida, ma probabilmente non come quella esponenziale. Ciò indica che le cellule sono, appunto, nel regime ordinato. Come è noto, è stato recentemente mostrato (Gibbs 2003) che parte del DNA spazzatura può codificare per l'RNA il quale gioca un ruolo regolatore, così il numero complessivo dei geni può essere di gran lunga maggiore rispetto al numero dei geni strutturali. La vera legge di scala della durata del ciclo cellulare come funzione del numero complessivo dei geni, fin qui, può essere di gran lunga più lenta rispetto al caso in cui si utilizzassero geni strutturali. Un insieme proprio dovrebbe prevedere la vera legge di potenza, ed anche la distribuzione del ciclo cellulare attorno alla mediana. [...] Questa caratteristica attualmente non è del tutto testabile poiché non conosciamo il numero complessivo dei geni nelle cellule (2004a, p. 588).

Analizzando il genoma umano, dunque, ci si è accorti che rispetto ai geni codificanti (geni strutturali) condivisi quasi totalmente con lo scimpanzé, nell'uomo aumentano i geni funzionali e in particolare il DNA spazzatura, ovvero il numero di regioni non codificanti (introni). Alcuni studiosi (Gibbs 2003) hanno mostrato che il numero elevato di proteine prodotte nelle cellule umane e la maggiore complessità che caratterizza i nostri sistemi dipendono proprio dal DNA spazzatura che nella nuova visione sistemica diviene fondamentale. Gli esoni, infatti, costituiscono meno del 2% del nostro genoma, mentre gli introni ne rappresentano circa il 25%. Fin dal momento della loro scoperta, nel 1978, gli introni sono stati considerati DNA spazzatura (insieme alla restante parte di DNA che non contiene geni). Recentemente, invece, il completamento delle sequenze genomiche di altri organismi (cane, scimpanzé, topo) ha consentito di confrontarle con quella umana (genomica comparativa). Le sequenze che hanno una funzione vengono conservate, ovvero si modificano poco tra i diversi organismi durante il processo evolutivo. Le sequenze degli esoni, che servono a codificare proteine, sono risultate infatti molto simili nelle diverse specie. Le sequenze degli introni non rappresentano spazzatura, ma contengono informazioni importanti per il funzionamento dei nostri 30000 geni. Non è infatti tanto il numero di geni, quanto il modo in cui il loro funzionamento è regolato a rendere l'uomo uomo, il cane cane e lo scimpanzé scimpanzé.

Se, per esempio, confrontiamo il genoma di *Drosophila melanogaster* con quello di *Caenorhabditis elegans*, notiamo che il primo (14000 esoni) ha meno geni del secondo (19000 esoni), pur essendo apparentemente più complesso; tuttavia, l'apparente paradosso è facilmente risolvibile nel momento in cui esaminiamo il DNA non codificante: il moscerino della frutta infatti ha di media 10000 introni, contro i 5000 del verme. Stando così le cose, quindi, Alberts

*et al.* (2002) sostengono che, sebbene il kit di costruzione molecolare abbia poche tipologie di parti, le istruzioni di assemblaggio, come specificato dalla sequenza regolatrice nel DNA non codificante, sembrano comunque essere più voluminose. In questo contesto, allora, le leggi di elevamento a potenza ipotizzate da Kauffman potrebbero ancora essere valide qualora venissero inseriti nella variabile  $N$  geni funzionali, introni e altri metaboliti. In continuità con tale prospettiva, infatti, anche la corrispondenza tra il logaritmo del numero di tipi cellulari in organismi di molti *phyla* diversi e il logaritmo del DNA contenuto in ogni cellula può essere rivisitata mantenendo in vita, altresì, l'ipotesi secondo cui i tipi cellulari corrispondono ad attrattori alternativi di reti booleane costruite casualmente:

Esiste una legge di scala per il numero degli attrattori di ciclo di stato come funzione di  $N$  che sia in grado di prevedere correttamente le relazioni esponenziali tra il numero di tipi cellulari di un organismo e il numero dei suoi geni (magari includendo la possibilità di RNA spazzatura che giochi un ruolo regolatore)? Nel mio lavoro precedente, ho trovato l'evidenza numerica che, per una rete  $K = 2$ , il numero di attrattori aumenta in funzione della radice quadrata del numero,  $N$ , di geni. Questo risultato si è recentemente dimostrato errato. Ho preso a campione spazi di stato molto ampi e ho lasciato piccoli bacini d'attrazione. Per le reti piccole esiste l'evidenza numerica che la legge di scala è lineare per reti  $K = 2$ , che come è risaputo, si trovano nella fase di transizione tra ordine e caos. Più recentemente, Socolar ed io abbiamo trovato la prova che per le reti  $K = 2$  la proporzionalità esponenziale è più veloce che lineare, ma è più lenta che profonda nel regime ordinato. Questo fatto, se riscontrato, sta a significare che non possiamo prevedere la legge di scala per il numero di tipi di cellule come una funzione del numero dei geni senza sapere dove, e se, si trovino le cellule nel regime ordinato e quale sia il numero di geni. D'altra parte, l'insieme corretto, le reti booleane casuali, le reti ad invarianza di scala, le reti del piccolo mondo, le reti medusa, o altre, dovrebbero essere in grado di rispettare la legge di scala appena osservata, se la posizione nel regime ordinato viene misurata [...] e se si conosce il numero di geni (2004b, p. 585).

Alla luce di tutto ciò, dunque, queste considerazioni ci permettono di inferire che l'ipotesi secondo cui i tipi cellulari sono attrattori dinamici è ancora attendibile; Kauffman, quindi, utilizzando il linguaggio della Dinamica per interpretare i fenomeni biologici, inventa un modello matematico (plausibile da un punto di vista biologico) capace di inserire il mistero dell'ontogenesi in un quadro teorico più ampio, in cui la biologia si trova improvvisamente in "dia-



logo” con altri saperi, come per esempio la matematica, la fisica, la teoria del caos, l’informatica e la teoria dei sistemi. Ebbene, in questo contesto, attraverso la teoria delle reti booleane costruite casualmente, lo studioso statunitense risponde in modo definitivo all’esigenza originaria del Theoretical Biology Club (rappresentata dalla linea di ricerca iniziata da Waddington con gli studi relativi alla canalizzazione ed all’assimilazione genetica) di costruire un nuovo paradigma organicistico non vitalista dello sviluppo in cui la biologia si dotasse di quel potere di spiegazione logica e matematica che avevano sempre avuto le scienze fisiche.<sup>8</sup> Con tale spirito, quindi, Kauffman, donando un codice matematico e un nuovo metodo di approccio alla biologia, offre, nello stesso tempo, un modello efficace anche per quanto concerne la spiegazione dell’omeostasi:

L’omeostasi, la tendenza dei tipi cellulari a rimanere uguali nonostante le perturbazioni, è essenziale per la vita. Prendiamo una rete booleana con migliaia di variabili e lasciamo che si stabilizzi attorno ad un attrattore di ciclo di stato. Di quando in quando cambiamo l’attività di un singolo gene nel modello. Dopo quasi tutte queste perturbazioni, il sistema ritorna al ciclo di stato dal quale è stato spostato. Questa è precisamente l’omeostasi, ed è gratuita nel regime ordinato. Ma l’omeostasi non può essere completa. Se lo zigote si differenzia per vie divergenti in tipi cellulari intermedi, che a loro volta si differenziano nei tipi cellulari definitivi del neonato o dell’adulto, allora di quando in quando una perturbazione dovrà spingere una cellula in un nuovo bacino di attrazione che scorre verso un nuovo attrattore – cioè in un nuovo percorso di sviluppo che la porterà a diventare un nuovo tipo cellulare. Allo stadio iniziale dell’embrione, per esempio, è risaputo che le cellule dello stato germinativo dell’ectoderma, che stanno per differenziarsi in cellule della pelle, possono essere stimolate da un messaggio molecolare che le fa saltare in un nuovo schema di differenziazione, in cui diventeranno cellule nervose. Allo stesso tempo, sappiamo del lavoro di circa cent’anni che ogni tipo cellulare può essere spinto a cambiare solo lungo percorsi adiacenti. Negli embrioni allo stato iniziale le cellule dell’ectoderma possono trasformarsi da cellule epidermiche a neuroni, ma non possono diventare le cellule che foderano lo stomaco e che secernono acido cloridrico (1995, p. 152).

Lo zigote di tutti gli organismi pluricellulari si divide e avvia le cellule figlie verso percorsi divergenti con alcune possibilità a ogni bivio (rappresentabile

---

<sup>8</sup> Per un ulteriore approfondimento si veda Continenza (1987).

in un grafico come una biforcazione) per creare la diversità finale dei tipi cellulari. In linea di principio, dunque, il modello di sistemi genomici nel regime ordinato proposto da Kauffman, effettivamente, esibisce in modo gratuito le proprietà or ora accennate:

Per la maggior parte delle perturbazioni, un sistema genomico influenzato da un attrattore esibirà un ritorno omeostatico allo stesso attrattore. I tipi cellulari sono fondamentalmente stabili. Ma in seguito ad alcune perturbazioni, il sistema viene catturato da un attrattore differente, così la differenziazione avviene in maniera naturale. E l'altra proprietà critica è questa: a partire da qualsiasi singolo attrattore, è possibile subire delle transizioni solo verso alcuni attrattori vicini, e da questi altre perturbazioni conducono il sistema ad ulteriori attrattori. Ogni lago è vicino solo a pochi altri laghi. Una cellula dell'ectoderma può essere spinta facilmente in un attrattore in cui dovrà formare una cellula della retina, ma le cellule dell'ectoderma non possono essere spinte facilmente nel sistema in cui si formano le cellule intestinali. Senza un'ulteriore selezione, le reti genomiche nel regime ordinato mostrano spontaneamente una proprietà fondamentale che ha caratterizzato l'ontogenesi per forse un miliardo di anni: le cellule si differenziano lungo percorsi divergenti che partono dallo zigote, per produrre i molti tipi cellulari dell'adulto (ivi, p. 153).

Ci si trova così di fronte a un bivio: dato cioè che l'omeostasi è la caratteristica fondamentale dell'ontogenesi, allora o ammettiamo che la selezione ha "lottato" per un miliardo di anni al fine di mantenere intatta tale caratteristica, che quindi viene condivisa da tutti gli organismi pluricellulari in virtù della discendenza comune, oppure dobbiamo ammettere con Kauffman che il differenziamento lungo percorsi divergenti è una peculiarità inserita così profondamente nelle reti genomiche canalizzanti che questo aspetto profondo dell'ontogenesi "appare come espressione dell'ordine gratuito indipendentemente dal successivo vaglio della selezione". Se quest'ultima ipotesi è giusta, allora "la selezione non è l'unica fonte di ordine nell'ontogenesi". A tale riguardo, inoltre, lo studioso statunitense fornisce ulteriori prove:

Ricordiamo che nel regime ordinato, un componente rosso di geni congelato, sia nello stato attivo che inattivo, forma un raggruppamento gigante che si propaga attraverso la rete genomica, lasciandosi dietro delle isole verdi di geni funzionalmente isolati, che si accendono e si spengono in configurazioni complesse. Se il sistema genomico è davvero in un regime ordinato, allora dovrebbero comparire questo componente

congelato e le isole lampeggianti. Se così fosse, allora una grande percentuale dei geni dovrebbe trovarsi nello stesso stato di attività in tutti i tipi cellulari del corpo. Essi dovrebbero corrispondere al componente congelato che investe tutta la rete. In effetti, si ritiene che circa il 70% dei geni costituiscano un nucleo comune, attivi simultaneamente in tutti i tipi cellulari dell'organismo di un mammifero. Numeri simili valgono anche per le piante. Questo nucleo comune fa parte di un componente congelato (ivi, p. 154).

Tale discorso implica, quindi, che solo una frazione dei geni determina le differenze esistenti fra le cellule. Per esempio, in una pianta con ventimila geni le differenze nell'espressione genica fra tipi cellulari diversi è, sulla base di recenti calcoli, circa il cinque per cento. Questo valore, secondo Kauffman, è molto vicino alla frazione prevista sulla base del numero di geni presunti nelle isole lampeggianti non congelate in uno stato fisso. Infine, nel regime ordinato la perturbazione dell'attività di un singolo gene si propaga solo a una piccola frazione dei geni dell'intera rete: nessun gene, infatti, dovrebbe essere in grado di innescare una valanga di alterazioni (effetto farfalla) che si propagano a cascata fra gli altri geni.

Le considerazioni dinanzi accennate, dunque, agli occhi dello studioso statunitense, rappresentano degli indizi che avvalorano la tesi secondo cui la selezione non è l'unica fonte di ordine nell'ontogenesi; tesi, quest'ultima, che egli ha sostenuto, con sfumature diverse, fin da quando era uno studente di medicina. All'inizio degli anni sessanta, quando cominciò a occuparsi degli insiemi auto-catalitici e dei suoi modelli di reti del genoma, infatti, il giovane Kauffman credeva davvero che la vita fosse plasmata quasi del tutto dall'auto-organizzazione e che la selezione naturale fosse solo una manifestazione secondaria: nulla lo dimostrava meglio dello sviluppo dell'embrione, dove i geni interagenti si organizzavano in diverse configurazioni corrispondenti a vari tipi di cellule e dove le cellule interagenti si organizzavano a formare i differenti tessuti. Così il grande studioso si concentrò sulla possibilità di conciliare l'auto-organizzazione e la selezione naturale. All'inizio pensò al problema in termini di una "lotta": i due fattori, infatti, agivano da antagonisti finché non raggiungevano uno stato di equilibrio in cui la selezione non poteva più modificare gli organismi. Quest'immagine accompagnò Kauffman fino alla metà degli anni ottanta, quando arrivò al Santa Fe Institute e, dopo aver conosciuto Langton, cominciò a studiare il margine del caos. Langton fece capire al nuovo collega che il margine del caos era una regione speciale a sé, ovvero sia il luogo in cui si potevano trovare sistemi con comportamenti complessi simili a quelli della vita (Langton 1986). Da quel momento Kauffman, studiando fe-

nomeni complessi come, per esempio, semplici molecole lipidiche che vanno alla deriva nell'acqua e si raggruppano poi in bolle simili a cellule, l'ordine gratuito in reti che collegano decine e decine di migliaia di variabili, oppure il potenziale della strutturazione della vita in gruppi di molecole reagenti, si accorse che dietro a tali fenomeni si nascondeva in realtà una misteriosa combinazione di selezione naturale e auto-organizzazione, un connubio che presto divenne il cuore teorico di riferimento della linea di ricerca trentennale volta all'individuazione, in seno alla biologia teorica, di leggi in grado di governare il caos e la complessità.<sup>9</sup>

Egli, infatti, capì che l'auto-organizzazione non è la forza più potente in biologia poiché i sistemi viventi, come abbiamo visto, non sono profondamente radicati nel regime ordinato, bensì sono molto vicini alla transizione di fase al margine del caos dove i fenomeni sono più sciolti e fluidi. A questo punto, dunque, agli occhi dello studioso, il problema concernente il rapporto tra selezione e auto-organizzazione assunse una meravigliosa chiarezza: la selezione naturale non è l'antagonista dell'auto-organizzazione poiché è più simile a una legge del moto, ovvero a una forza che spinge di continuo sistemi emergenti (auto-organizzanti) verso il margine del caos. In seguito Kauffman, attraverso la simulazione informatica di reti di lampadine interconnesse, cercò di verificare se, nelle reti genetiche, l'evoluzione conduca tipi reali di cellule verso il margine del caos. Un piccolo numero di simulazioni riuscite non dimostra niente di per sé, ma bastò a convincere lo studioso del fatto che stesse camminando verso la direzione giusta. Così, tale idea divenne il fulcro-teorico di tutta la sua attività di ricerca presso il Santa Fe Institute, un cammino i cui risultati, come abbiamo mostrato nei paragrafi precedenti, sarebbero stati descritti in modo sistematico solo a metà degli anni novanta, in particolare nel testo del 1995 dove lo studioso, ormai famoso, non può esimersi, verso la fine del capitolo dedicato al mistero dell'ontogenesi, dal ricordare al lettore la sua convinzione iniziale, con la quale, a questo punto, ci pare opportuno concludere il presente paragrafo:

Nel 1964, quando iniziai i miei studi di medicina, avevo un sogno di cui non conosco le origini. A quel tempo mi ero da poco accostato alla biologia, ed ero poco addentro alle sue meraviglie in cui si intrecciano contingenza storica, selezione, progetto, casualità e prodigio. Penso che come giovane scienziato non potevo ancora iniziare a capire la potenza

---

<sup>9</sup> Il primo a chiedersi se la selezione possa spingere le reti booleane per l'elaborazione parallela ai margini del caos è stato Packard; in tal senso, di fondamentale importanza fu per Kauffman la lettura di Packard (1988).

della selezione naturale, la cui sottigliezza mi è apparsa sempre più straordinaria nel corso dei tre decenni successivi. Ma il sogno, scaturito da chissà quale origine sconosciuta, c'è sempre. Se i biologi hanno ignorato l'auto-organizzazione, non è perché l'auto-organizzazione non fosse profonda e pervasiva. Ma è perché noi biologi dobbiamo ancora capire come ragionare sui sistemi governati simultaneamente da due sorgenti d'ordine. E tuttavia chi [...] potrebbe non essere attraversato da un pensiero cruciale? Se mai riusciremo a ottenere una teoria finale nella biologia, dovremo di sicuro capire la mescolanza dell'auto-organizzazione e della selezione. Dovremo affrontare il fatto di essere espressioni naturali di un ordine più profondo. E alla fine scopriremo nel nostro mito della creazione che dopo tutto eravamo attesi (1995, p. 154).

#### **4. *Meaningful Complexity* e selezione naturale**

Alla luce di quanto detto sin qui, la storia della vita, a giudizio di Kauffman, è certamente la storia di eventi accidentali e casuali, ma anche quella dell'ordine: di qui la necessità di postulare un tipo di creatività profonda, interna, intessuta della trama stessa della natura. Lo studio delle reti booleane mostra che tali reti sono modelli logico-matematici (anche se limitati) di una ampia classe di sistemi dinamici non lineari. Gli attrattori di queste reti possono simulare il naturale oggetto di interesse. Da un punto di vista biologico, come abbiamo mostrato in precedenza, è possibile ipotizzare che questi attrattori corrispondano ai tipi cellulari, mentre da un punto di vista conoscitivo risulta possibile interpretare tali attrattori come la naturale classificazione che la rete fa del mondo esterno. Queste scoperte rappresentano un prudente allargamento dei risultati ottenuti nel campo della termodinamica di non-equilibrio.

In particolare è importante precisare, a questo proposito, che tale allargamento riguarda, prima di tutto, la natura e la dinamica dei processi di differenziamento, il collegamento, in prospettiva, che esiste tra questi ultimi processi e la successiva formazione di particolari bacini di attrazione. Come abbiamo visto nel paragrafo precedente, nonostante gli erronei dati numerici relativi al numero dei geni umani, l'ipotesi di Kauffman secondo cui i tipi cellulari sono attrattori dinamici è tutt'ora valida. Kauffman, pur avendo segnato (mediante l'applicazione delle reti booleane stocastiche alla biologia) il cammino delle ricerche nell'ambito della complessità biologica dagli anni settanta fino a oggi, e nonostante i numerosi e originali sentieri esplorati nell'ultimo decennio (la teoria dell'agente autonomo, il concetto di cicli di lavoro termodinamico, la rivisitazione del concetto di significato secondo Peirce e la teoria

dell'informazione istruttiva)<sup>10</sup>, a differenza di alcuni scienziati che hanno posto l'accento dei loro studi sulla possibilità di costruire una teoria semantica dell'informazione biologica, ovvero una rinnovata teoria algoritmica dell'informazione basata su di una logica intensionale e sul riferimento a strumenti matematici innovativi, rimane, per alcuni aspetti, ancorato a un modello matematico (basato su una logica estensionale), che permane, a livello formale, quello presentato con tanta cura e lungimiranza nell'articolo del 1970 e a cui, ancora oggi, si fa continuo riferimento in tanti centri di ricerca nel mondo (si veda Kauffman 1970).

Questo lavoro, dunque, attraverso le esplorazioni scientifiche e le investigazioni metodologiche di Kauffman, con nessuna pretesa di esaustività, costituisce comunque una sorta di cammino lungo i sentieri evolutivi della complessità biologica, un cammino, vale a dire, che si propone di esaminare i vari modelli interpretativi, via via elaborati da diversi studiosi, al fine di sondare le presunte “colonne d’Ercole della biologia contemporanea”, ovvero le misteriose procedure dell’auto-organizzazione. Ecco allora che le modellazioni matematiche (i processi di Markov e l’algebra di Boole) che consentivano al primo Kauffman di interpretare l’evoluzione dei sistemi dinamici e i processi stocastici dell’espressione genica non sono più sufficienti:

I sistemi dinamici di cui stiamo parlando possono essere definiti, sotto certe condizioni, in uno spazio discreto in cui l’evoluzione si mostra come un processo markoviano. In realtà lo spazio di stato di sistemi che danno luogo a una dinamica caotica è essenzialmente continuo. A questo livello le deviazioni dalla media non possono essere viste come eventi localizzati. Ciò nonostante, in molti casi la dinamica caotica può essere presentata nella forma di una catena di Markov utilizzando strumenti ben fondati offerti dalla classica teoria dell’informazione. Ciò comporta un passaggio dall’entropia fisica all’entropia informazionale. Inoltre, dobbiamo sottolineare che la considerazione dei processi di Markov e degli automi di Markov riguarda anche la descrizione del comportamento dinamico di sistemi caratterizzati dalla presenza di un reticolo di componenti booleani identici, che interagiscono localmente nello spazio e alla presenza di specifici fenomeni di soglia. A livello di tali reti, alcuni pattern computazionali globalmente emergenti possono venire configurandosi come una funzione del tempo ecc. Questo fatto ben noto si basa sull’esistenza di particolari attrattori della rete dinamica dissipativa. A seconda dello stimolo iniziale il sistema finirà in diffe-

---

<sup>10</sup> Per un approfondimento di tali tematiche si vedano Kauffman (2000) e (2008).

renti bacini d'attrazione. I processi di Markov, in questo senso, giocano un ruolo importante nell'interfaccia tra la fisica del non-equilibrio e i modelli teorici della rete genetica, vale a dire, nell'interfaccia tra ordine fisico e ordine cognitivo come sostengono gli studi classici nel campo della cibernetica e dell'intelligenza artificiale. [...] Queste proprietà compaiono in almeno quattro modi distinti: auto-catalisi, auto-assemblaggio, auto-organizzazione, auto-rappresentazione. Determinante in ogni manifestazione di questo tipo è il flusso di energia, il flusso di energia intrinseca (secondo la distinzione di Maxwell tra *intrinsic energy* e *exergy*). Una delle caratteristiche principali che distinguono queste tipologie di sistemi è il fatto che per loro non esiste alcuna soluzione in forma chiusa. Il loro comportamento si può prevedere quantitativamente solo se si simula la dinamica. Se ora spostiamo la nostra attenzione sull'ordine biologico e prendiamo in considerazione il funzionamento del DNA, costatiamo facilmente che, per quanto riguarda l'ordine fisico, le attività regolatrici e i processi di differenziazione giocano un ruolo cruciale in questo particolare tipo di sistema. Non siamo più di fronte a un semplice automa stocastico di Markov, bensì ci troviamo dinanzi a un automa cellulare complesso che è in grado di canalizzare il flusso di energia in modo tale da permettere alle potenzialità nascoste nello stesso flusso di essere rivelate progressivamente in una maniera differenziata (Carsetti 1993, pp. 112-114).

Da Prigogine a Kauffman, quindi, a nostro giudizio, è possibile scorgere una linea di ricerca coerente basata sull'individuazione di principi che caratterizzano la dinamica caotica e, da un punto di vista più generale, la natura dello stato intermedio (il cristallo aperiodico di cui parlava Schrödinger). Questi principi fanno riferimento essenzialmente, in accordo con lo schema del neodarwinismo, all'esistenza di una precisa dialettica tra mutazione, selezione e differenziamento. Tali studiosi, infatti, con le loro ricerche, senza dubbio offrono una prima caratterizzazione di questo tipo di dialettica utilizzando, in modo creativo, gli strumenti messi a disposizione dalla contemporanea teoria della complessità. A questo punto, però, risulta possibile chiedersi: anche se questo schema è parzialmente corretto, è anche vero? In altre parole, è possibile spiegare l'intera complessità dei processi di auto-organizzazione all'interno di un contesto generale markoviano anche se allargato in virtù delle considerazioni relative ai ruoli giocati dalla selezione naturale e dai processi di differenziamento cellulare?

Come osserva correttamente Atlan, in un sistema naturale che si auto-organizza (un sistema biologico) la finalità non viene stabilita dall'e-

sterno. Ciò che si auto-organizza è la funzione stessa con il suo significato. L'origine del significato nell'organizzazione del sistema è una proprietà emergente. Inoltre, l'origine del significato è strettamente connessa a precise opzioni di osservazione. Se progettiamo di costruire una rete cellulare complessa per stimolare le attività di un cervello dobbiamo tenere conto della capacità, per conto di questa rete, di osservare il significato nel mondo e il "significato" che scaturisce in se stesso. Il comportamento della rete, in altre parole, possiede un significato non solo nel senso che essa risulterà autonoma, ma anche in quello secondo cui essa risulterà osservata e intenzionalmente connessa a una produzione continua di possibili nuove interpretazioni. Così, per fare in modo che una fonte di informazione possa mostrare un comportamento autonomo di auto-organizzazione, dobbiamo aggiungere ai processi di mutazione, di selezione e di differenziamento, or ora accennati, particolari capacità di osservazione, di auto-osservazione, di simulazione e di interpretazione. Il problema può essere spiegato come segue: adesso siamo finalmente in grado di modellare il comportamento di reti cellulari capaci di evolvere e di differenziarsi in un modo semplice, utilizzando gli strumenti offerti dalla contemporanea teoria della complessità. A quali strumenti concettuali dobbiamo ricorrere per essere in grado di modellare il processo di auto-generazione del significato? [...] Da un punto di vista oggettivo dobbiamo osservare, prima di tutto, che la frontiera tra ordine e caos sembra poter offrire strumenti molto più sofisticati alla selezione. In particolare essa è in grado di offrire, invece di mutazioni puntiformi, una variabilità articolata capace di guidare l'ambiente a rivelare se stesso e a manifestare potenzialità nascoste che vivono progressivamente a un livello profondo. [...] In riferimento a questo particolare "paesaggio" i vincoli imposti dalle pressioni selettive a livello della dinamica dell'automa cellulare dissipativo possono permettere, in realtà, una canalizzazione più complessa del flusso informazionale in input. [...] Le pressioni selettive, che plasmano la complessità variata, la quale permette l'emergenza di nuove correlazioni stabili e di nuovi vincoli, costituiscono la condizione necessaria per l'emergere di nuove forme di ordine. Nuovi principi generativi saranno canalizzati. Ci sarà, di conseguenza, la possibilità di incapsulare questi principi nella materia. Questo fatto, a sua volta, permetterà il generarsi di nuove forme di invarianza e di dissipazione. [...] In quest'ottica, da un punto di vista biologico, il genoma, non più considerato come un puro sistema di replicazione bensì come un sistema che si auto-organizza volto all'espressione e al rinnovamento del tipo di replicazione creativa della fonte, sembra articolare la propria attività di costruzione seguendo stadi di sviluppo precisi e differenziati. La sua produzione di complessità variata (come viene ottenuta lungo la realizzazione del processo di *embodi-*



*ment* al livello fenotipico, ovvero lungo la realizzazione di specifiche strutture dissipative che possiedono un carattere non replicativo) viene plasmata e vincolata, successivamente, da pressioni selettive secondo procedure specifiche e secondo particolari disposizioni di reti di vincoli (ivi, pp. 116-118).

Mediante lo sviluppo del processo di incarnazione or ora delineato il genoma determina, pertanto, la costruzione progressiva di uno specifico canale per un'ulteriore espressione parziale del contenuto informativo profondo e della rivelazione di nuove forme di incomprimibilità. Così, il nuovo ordine che possibilmente sorge in conseguenza di questo tipo di rivelazione si iscriverà all'interno e al di sopra della vecchia struttura, ovvero sia all'interno e al di sopra delle "fibre" biochimiche dell'ultimo ordine rivelato. In tal senso, allora, secondo Carsetti, risulta lecito inferire che, in base alle considerazioni or ora richiamate, è impossibile, in linea di principio, ricreare attraverso la semplice simulazione informatica la vita reale:

L'uomo può solamente fare di se stesso uno strumento al fine di determinare, utilizzando modelli di simulazione adatti, un'espressione del *bios* più profonda e differente. L'informazione profonda di cui stiamo parlando non è una struttura particolare o un ordine o un insieme di correlazioni. È il luogo (il *locus* d'azione) dei principi generativi, dei flussi informativi profondi. Questi flussi diventano un ordine solo quando riescono a esprimersi funzionalmente come uno specifico atto di sintesi all'interno della cornice rappresentata dai vincoli determinati a loro volta dalla realizzazione progressiva di un adeguato canale di auto-organizzazione. In altre parole, questo tipo di sintesi può aver luogo solamente in riferimento, prima di tutto, alla dialettica riguardante il collegamento che esiste tra la produzione di complessità variata e l'attività di selezione. La reale manifestazione di questa dialettica, di questa dinamica interattiva, si mostra come un processo di generazione di informazione, un processo che si articola attraverso la successiva apparizione di forme specifiche di ordine e coerenza e di specifici attrattori (ivi, p. 119).

Ebbene, lo studioso italiano sottolinea come sia precisamente la capacità autonoma, a livello del canale (del processo di incarnazione), di articolare secondo queste forme nel modo corretto, sfruttando, passo dopo passo, i sentieri adeguati, e di riflettere lo schema generale di questi stessi sentieri (per costruire anche un adeguato programma di simulazione), che può, finalmente, aprire la strada alla rivelazione dei proto-programmi viventi inespressi nei

flussi di informazione e alla trasformazione di principi generativi in specifiche proprietà in grado di generare l'informazione intrinseca ai sistemi dinamici:

In questo modo la fonte di informazione profonda, una volta incapsulata nell'iscrizione (una volta incapsulata, cioè, nella materia come una struttura biologica replicativa), continua a "dettare" la rappresentazione successiva dei pattern e dei vincoli che caratterizzano questa stessa iscrizione durante la realizzazione progressiva, a livello di superficie, delle attività dissipative. Dopo il completamento del processo di incarnazione (e durante il suo sviluppo), il sistema, grazie a specifiche procedure riflessive (e anche grazie ad altri strumenti di analisi), procede, dunque, a delineare in termini schematici un recupero dell'intero processo. Se questo tipo di recupero riesce a estrarre e a generare il "significato" riguardante l'iscrizione, ovvero a ricostruire la "memoria" nelle forme adatte, avremo come risultato un potenziale innesco per tutti quei processi funzionali che condurranno forse alla chiusura definitiva del processo di incapsulamento e, allo stesso tempo, al possibile dispiegamento di nuovi programmi di sintesi e di azione che vivono a livello profondo (*ibidem*).

Attualmente, dunque, la sorgente dell'informazione, per giungere a una forma di espressione stabile, deve incapsulare se stessa in specifiche proprietà generative. Tali proprietà, afferma Carsetti, devono risultare iscritte nella materia fisica del sistema in modo da dare luogo alla possibilità di generare, in maniera ripetitiva e autonoma, la complessità variata, prodotto delle proprietà stesse.

Ciò risulterà possibile solo nella misura in cui i principi generativi verranno ad essere storizzati nella materia per quanto concerne l'arco formale della loro realizzazione effettiva. Sarà, in questo senso, necessario che l'azione esterna di tali principi venga letta e riconosciuta dal sistema così come essa è venuta realizzandosi al suo interno (nella metamorfosi ad esso relativa), in modo tale che il sistema stesso possa, infine, delineare, per così dire, in negativo una rappresentazione, sul piano formale, della azione operata dai principi stessi. [...] Tale rappresentazione riprodurrà all'interno del sistema i moduli della canalizzazione realizzatasi e potrà, quindi, essere applicata direttamente, tramite trasformazioni specifiche operate entro i confini del sistema, afflussi informativi puri o misti, non necessariamente cioè contraddistinti dalla presenza articolata e funzionale dei moduli suddetti. Si realizza, in tal modo, un processo di assimilazione mediata ed indiretta attraverso la

quale, tramite la ricostruzione operata dell'azione dei nuclei informativi esterni, il sistema giunge a storizzare, in maniera oggettiva, l'informazione necessaria per rendere stabile ed autonomo il meccanismo di formazione della complessità variata. Per modulare, in una parola, l'autonomia delle proprietà generatrici secondo condizioni di oggettività. [...] Trasformando, processi in proprietà, operando una assimilazione precisa di principi generativi esterni, il sistema ha la possibilità non solo di offrire, in modo stabile ed autonomo, complessità variata ai flussi informativi profondi, ma ha anche la possibilità di generare dall'interno, tramite processi di riorganizzazione continua, una esplorazione sempre più ampia dei percorsi della variabilità possibile; di iniziare, cioè, a delineare quella che, a livelli di massima complessità, si rivelerà come una precisa e mirata azione di coagulo (ottenuta per mezzo della costruzione di rappresentazioni astratte) per il manifestarsi di sempre nuovi aspetti della informazione di profondità (Carsetti 1989, pp. 94-95).

Alla luce di tutto ciò, dunque, in questo quadro teorico, appare chiaramente come per realizzare l'antico progetto del Theoretical Biology Club legato alla costruzione di una biologia teorica indipendente dalla fisica e dalla chimica, non bastino gli accorgimenti, le rimodulazioni e il disegno astratto relativi a una meccanica statistica a carattere rinnovato così come individuati e perseguiti da Kauffman in *The Origins of Order* e più volte rivisitati nei suoi testi successivi. Per costruire a livello biologico una meccanica statistica concernente geni e macromolecole (in azione) occorre fare i conti sino in fondo con l'informazione profonda, un'informazione, vale a dire, non misurabile tramite il ricorso agli strumenti offerti dalla tradizionale teoria dell'informazione di Shannon basata su di una matematica troppo semplice e quindi "incompatibile" con la complessità dei fenomeni vitali. Occorre, in altre parole, definire, come abbiamo accennato sopra, i principi di una nuova teoria dell'informazione algoritmica (cioè di una nuova teoria della complessità), non esclusivamente ancorata a una base proposizionale, bensì articolata al livello di una dimensione logica a carattere predicativo e stratificato. Una tale teoria della complessità dovrebbe essere in grado, fra l'altro, di mostrarci come sia possibile parlare, senza contraddizione alcuna, di non esistenza di algoritmi finiti in relazione a problemi che pure risultano ben posti in termini di unicità e di esistenza (la non esistenza è un dato di partenza ineliminabile così come, sul versante fisico, in accordo con Prigogine, è un dato primitivo l'esistenza di una *randomness* che trova il suo fondamento nella dinamica). Ebbene, tutto ciò implica anche l'elaborazione di una semantica intensionale ed iperintensionale per i processi ri-

correnti di auto-organizzazione, nonché la costruzione di modelli di simulazione di automi dotati di basi intensionali e di funzioni riflessive e interpretative. In tal senso, la prima tappa di un progetto così vasto dovrebbe, a nostro giudizio, far riferimento, almeno da un punto di vista astratto, ai tentativi in atto di delineare nuovi principi concettuali atti a definire un *background* logico adeguato per una corretta semantica dei processi.

Possiamo facilmente intuire che i suddetti tentativi finora si sono concentrati sulla sperimentale definizione di (almeno) due nuovi concetti centrali: il concetto di verità considerata non come invarianza, ma come emergenza, e il concetto di modello che si auto-organizza. Per quanto riguarda il primo concetto non abbiamo più a che fare con una nozione di verità considerata come una semplice forma di propagazione invariante della verità all'interno di un *setting* logico monotonic. La verità ora sembra essere definibile solo in riferimento alle procedure non-monotoniche, alle strutture del secondo ordine, alle azioni di sistemi accoppiati, all'esistenza di funzioni intensionali specifiche. In questo senso, l'emergenza della verità sembra essere collegata in modo specifico ai processi di rivelazione e di articolazione della fonte d'informazione originaria. Per ciò che concerne questo particolare contesto non sembra possibile, per esempio, parlare della verità in modo corretto, senza effettuare preliminarmente una distinzione tra informazione di superficie e informazione profonda. Tuttavia, ciò che meglio caratterizza, in poche parole, quella particolare rivoluzione di paradigma nell'ambito della semantica scientifica del nostro secolo rappresentata da un vero delinarsi di una adeguata semantica dei processi, è forse (se facciamo un riferimento essenziale agli aspetti dinamici della nozione di significato) il concetto generale di modello che si auto-organizza. Da un punto di vista semantico, la prima metà del Novecento è stata caratterizzata dalla costruzione teorica, da Russel a Tarski e Henkin, di una particolare nozione di modello (semantico): una nozione secondo la quale un modello appare essenzialmente vincolato a relazioni atomiche individuali, insiemi e forme logiche invarianti. Una proprietà unaria in questo contesto è rappresentata da una classe di individui; una relazione binaria, invece, è rappresentata da tutte le possibili coppie ordinate di individui atomici che soddisfino quella particolare relazione e così via. In questo modo, parlando in termini generali, un modello  $M = \{ S, V \}$  per una logica del primo ordine (senza uguaglianza) consiste in un dominio non-vuoto  $S$  (il dominio di interpretazione) insieme a una funzione  $V$  che assegna a ogni simbolo di funzione di posto  $n$  una funzione da  $S^n$  a  $S$  e a ogni simbolo di predicato di posto  $n$  una relazione di posto  $n$  su  $S$  (Carsetti 1996, pp. 103-104).

Nella logica classica, come tutti sanno, i modelli sono normalmente invarianti (modelli standard). Tuttavia, come è evidente, il mondo può cambiare. Quando si pensa a un cambiamento in un modello in genere si pensa a un cambiamento negli attributi degli individui del modello oppure a un cambiamento nel suo dominio. Ciò malgrado, un modello può mutare anche in conseguenza dell'intervento dell'osservatore sulla struttura ultima del modello, nonché del pieno sviluppo di alcuni e ben precisi meccanismi di auto-organizzazione interna. Ecco, dunque, il delinarsi del fondamentale passaggio dai modelli standard a quelli non-standard.

Questo passaggio rappresenta davvero uno dei possibili percorsi che possono condurci a una prima idea di un concetto ancora più generale: il concetto di modello che si auto-organizza. Quando entriamo nel regno di un modello che si auto-organizza non siamo più legati solo all'esistenza di individui e di forme logiche invarianti: siamo, al contrario, legati in prima istanza, all'articolazione di generatori specifici e di attributi complessi, all'esistenza di attrattori particolari, chiusure operazionali, flussi ricorrenti di informazione e, in generale, a una architettura a più strati di schemi che si auto-organizzano. Inoltre, un modello che si auto-organizza sembra essere necessariamente confinato all'interno dei limiti di un sistema accoppiato: strettamente connesso, cioè, all'effettiva articolazione di ben definiti processi auto-referenziali e (in prospettiva) di osservazione. Al livello di un modello che si auto-organizza siamo in presenza di processi specifici di partizione funzionale e di contrazione vincolata, che governano l'apparenza, a livello di superficie, di attributi complessi, individui e mondi. In questo nuovo contesto teorico, i concetti tradizionali di soddisfazione, verità invariante e conseguenza semantica saranno necessariamente soggetti a precise trasformazioni secondo specifiche tecniche logiche non-standard (ivi, pp. 104-105).

Il nucleo teorico di questo nuovo concetto è rappresentato dall'intuizione secondo cui il terreno della semantica non è costituito da domini di individui atomici, bensì da flussi di informazione, ovvero da processi generativi e ricorrenti. Questi processi determinano (secondo una dimensione verticale) una precisa gerarchia di livelli caratterizzata in modo intensionale.

Il passaggio da un livello a un altro è soggetto a trasformazioni algebriche dinamiche governate dall'azione di vincoli specifici che dipendono dalle differenti tipologie di livello; è il coinvolgimento di questi processi che costituisce, per esempio, al primo livello, quelle particolari unità funzionali rappresentate dagli attributi complessi. Ed è, precisamente,

con riferimento a questi particolari tipi di costituzione logica che gli individui-oggetti di dominio possono poi iniziare ad articolare la propria esistenza. I flussi di informazione, in questo senso, stanno alla base della “vita” reale degli individui. Essi rappresentano le vere “fibre” della struttura del modello dinamico. Pertanto, gli individui-oggetti non sembrano essere entità atomiche o primarie (a livello logico), bensì il risultato finale dell’effettiva articolazione di alcuni processi specifici, in particolare di un processo di generazione interiore e di un processo reale di partizione (funzionale). In questo modo possiamo vedere che al livello dei processi semantici entrano in gioco il Tempo e il flusso di informazione. In particolare, da un punto di vista tecnico, a entrare in gioco sono delle specifiche determinazioni del Tempo (ivi, p. 105).

Eccoci, dunque, ancora una volta di fronte ad alcune delle geniali intuizioni di Prigogine. Secondo lo studioso russo, infatti, per spiegare l’irreversibilità (e stocasticità) si devono considerare gli stati con una rottura di simmetria temporale propagata mediante leggi che sono esse stesse dovute a una rottura di simmetria. La rottura di simmetria temporale, quindi, in questo contesto, rappresenta uno strumento essenziale per sviluppare un nuovo livello di comprensione nel quale la razionalità non viene più identificata con l’idea di certezza. Allo stesso modo, dunque, risulta possibile affermare che nella semantica dei processi si assiste alla progressiva introduzione di concetti correlati a particolari condizioni di rotture di simmetria che si danno a livello logico e informativo: di qui la necessità, per esempio, del ricorso al concetto di partizione, al concetto di modello auto-organizzantesi e così via. Ebbene, mediante questa introduzione è attualmente possibile penetrare in territori nuovi della semantica, in un terreno, vale a dire, che appare strettamente determinato dal progressivo ampliamento di processi evolutivi.

Ma come abbiamo appena sottolineato, al livello della semantica dei processi così come a livello di una teoria dell’informazione multidimensionale, non siamo più limitati dentro i confini di semplici strutture markoviane. Questo fatto costituisce una vera linea di demarcazione rispetto all’apparato formale sviluppato da Nicolis e da Prigogine a livello della loro esplorazione delle basi teoriche della teoria della complessità. Nel quadro di riferimento della semantica dei processi dobbiamo, in realtà, ricorrere al delineamento di nuovi e più complessi spazi di informazione, di nuove misure di complessità in grado di articolare non solo a livello proposizionale (come, per esempio, l’entropia di Shannon), ma anche a livello del primo ordine e a livello del secondo ordine. È solo in riferimento a questi spazi di informazione più complessi che i

suddetti processi di partizione funzionale e di auto-organizzazione dinamica risulteranno definibili secondo un preciso *setting* matematico (ivi, p. 106).

In quest'ottica, allora, in accordo con Prigogine e Carsetti, non abbiamo a che fare solo con due diverse concezioni del tempo, vale a dire, il tempo come ripetizione (invarianza) e il tempo come disgregazione (dissoluzione), bensì ci troviamo dinanzi a un terzo concetto di tempo in grado, cioè, di superare questo dualismo: il tempo come costruzione, una costruzione che appare ai nostri occhi, simultaneamente, come creazione e come ritrovamento, anche se questa stessa costruzione passa attraverso stati specifici di invarianza e di degradazione.

Il tempo appare, dunque, strettamente legato all'intensionalità e allo sviluppo di un tessuto gerarchico di processi ricorrenti. Questo tessuto si presenta, al contempo, come creazione e rivelazione. Come la creazione continua di nuove forme di autonomia e, contemporaneamente, come la rivelazione continua di nuovi livelli di potere generativo. Ciò appare a noi come una emergenza incessante di significati sempre nuovi, una emergenza che plasma, in modo consecutivo e successivo, le determinazioni del Tempo. Quelle particolari determinazioni (o schemi, in termini kantiani) che forgiavano, secondo precisi modelli matematici, la variegata e vincolata espressione del linguaggio della Vita (ivi, p. 106).

Va sottolineato, inoltre, come a livello evolutivo non ci si trovi normalmente dinanzi a fenomeni stocastici puri, bensì, come si è detto, a fenomeni legati ai meccanismi del caos deterministico. La complessità variata, in altre parole, appare essere il risultato dell'azione di proprietà produttive dell'informazione semantica intrinseche ai sistemi dinamici. Queste proprietà possono, tuttavia, giungere a realizzarsi e a esprimersi, in modo compiuto, solo attraverso un collegamento costante con l'azione selettiva propria di principi generativi, di flussi (o nuclei) informativi specifici, che vivono nella realtà esterna al sistema (una realtà con la quale il sistema stesso risulta accoppiato), nonché attraverso il manifestarsi pieno, all'interno dello stesso sistema, di "un processo di corporificazione in atto degli spazi originari della virtualità pura". In altri termini, secondo Carsetti l'articolazione piena dei nuclei informativi esterni risulterà possibile solo nella misura in cui il sistema si rivelerà in grado di offrire a tali nuclei l'opportunità reale di declinarsi secondo specifici labirinti del possibile, sia secondo una dimensione naturale diretta, sia secondo una articolazione guidata sulla base della utilizzazione di particolari trasformazioni. Nel-

la misura in cui, in altre parole, da un punto di vista astratto, le configurazioni offerte risulteranno capaci di agire come esca e come coagulo adatti per specifici processi di canalizzazione dell'informazione proveniente dall'ambiente esterno (si veda Carsetti 2005, pp. 91-92):

Rispetto ad un quadro teorico siffatto, il problema reale non è quello di individuare di volta in volta i differenti possibili compromessi tra fluttuazioni, da un lato, e processi di stabilizzazione, dall'altro; il problema reale è quello di “seguire” i contorni reali dell'evoluzione propria della dinamica generativa e produttiva in atto in modo da poter innescare, sul piano effettivo, un reale dispiegamento della informazione di profondità. In tal senso, la costruzione di nuovi algoritmi atti a delineare l'invenzione di ulteriori modelli cognitivi si porrà come uno dei passaggi essenziali per il realizzarsi di una intelligibilità migliore della complessità propria dei sistemi viventi e, al tempo stesso, per una più ricca articolazione delle potenzialità evolutive loro proprie. I processi di auto-organizzazione dovrebbero essere considerati, innanzitutto, come una sorta di “arco” a carattere teleonomico in grado di comporre insieme *meaningful complexity*, da un lato e selezione per eliminazione, dall'altro. L'invarianza vera (come vita e come generazione) può esistere solo nel quadro di una morfogenesi autonoma in atto e viceversa. È l'opera di invenzione (e di generazione nella invarianza, in una invarianza che non è semplice ripetizione e che si costituisce ogni volta tramite un rinnovamento ed un ritrovamento che si danno a livello profondo) congiunta al radicamento del significato che determina l'evoluzione, i salti e gli equilibri punteggiati, le condizioni relative al delinarsi successivo di sempre nuove modalità di espressione da parte del vivente. In questo quadro, la selezione da parte del significato può mostrarsi autonoma nella misura, innanzitutto, in cui si rivela capace di “modulare” l'irruzione continua di nuovi apparati generativi. Essa non è, allora, legata al solo gioco dei possibili, al farsi canale per il caso puro, bensì al farsi canale per l'iscrizione del *file* nell'*humus* del significato in vista di un rinnovamento e di un ritrovamento della creatività originaria. È il rivelarsi autonomo, nell'invenzione, di nuova incomprimibilità possibile che, in effetti, determina l'emergere di una creatività semantica nuova, tramite anche il racconto relativo alla sua opera-costruzione. È in questo senso che le procedure del riferimento, se *successful*, possono giungere a modulare la canalizzazione e a porsi, così, come base per il nascere stesso di nuova invarianza (tramite il percorso della morfogenesi). Non si tratta di scoprire nuovi “territori”, ma di farsi matrice ed arco per il loro scaturire autonomo. Non c'è un processo casuale autonomo già in essere (cosa in sé) ed una attività di selezione e sintesi tramite “ritaglio” possi-



bile, tramite, in particolare, il ricorso a procedure del riferimento intese come modalità di semplice irreggimentazione. Le procedure, in realtà, sono funzionali alla costruzione di nuova incomprimibilità: il significato dà la possibilità di realizzare un ancoraggio olistico; è, esattamente, ciò che permette al categoriale di affiorare come “arborizzazione”. È in questo modo che si può assicurare l’esistenza di un tempo della invenzione ma non della semplice ripetizione, di un tempo, segnatamente, del rinnovamento e del ritrovamento che si rivela continuamente come possibile a misura dell’opera. Perché ciò possa darsi le procedure del riferimento debbono porsi, pertanto, come arco tra due ben differenti attività selettive, tra invarianza (a cui corrisponde il massimo della variabilità potenziale) da un lato, e morfogenesi autonoma, dall’altro. Esse potranno, tuttavia, nutrire l’incomprimibilità nuova solo ove si dia un reale processo di annidamento dello spazio originario della virtualità pura. Non è il ritaglio, allora, che conta, ma il racconto ed il raccordo ben riusciti. È l’iscrizione operante che permette lo scaturire di nuova incomprimibilità che necessariamente mi bypassa. Potrò, quindi, giungere a vedere per principi una incomprimibilità nuova che perviene a manifestarsi come fusione in atto di nuclei emergenti della creatività nell’unità del significato operante. La nuova invenzione che nasce dà, in effetti, luogo al sorgere ed all’aprirsi degli occhi di una mente: posso vedere in quanto mente perché nuovo significato giunge ad articolarsi e ad annidarsi mio tramite (e solo nella misura in cui ciò avviene). Ecco una Vita che si realizza come attività cognitiva, che giunge ad articolarsi come molteplicità di monadi-menti capaci di porsi come osservatori nel mondo e del mondo, come molteplicità, vale a dire, di “oggetti strani” in grado, appunto in quanto tali, di osservare un mondo (*Natura Naturata*) che risulta a sua volta popolato di osservatori nonché di eventi che risultano ad essi necessariamente correlati (ivi, pp. 102-104).

Questa nuova visione sistemica inscindibilmente legata allo studio del significato, dunque, costituisce, tra l’altro, uno degli attuali tentativi di rispondere alla domanda di Schrödinger, una domanda, che ha inquietato ed entusiasmato al tempo stesso l’essere umano fin dall’alba della nostra civiltà: basti pensare alle ricerche scientifiche di Anassimandro. Lungo il corso di questo lavoro abbiamo avuto la possibilità di mostrare, sulla scorta dell’opinione di illustri studiosi, come per fare, oggi, i conti seriamente con la domanda “Che cos’è la vita?”, si sia costretti a non rinchiudersi nella pura biologia e ad aprirsi, quindi, ad altri saperi, alcuni classici come la fisica, la chimica e la matematica, e altri, invece, più recenti come, per esempio, la scienza dell’informazione, la bioinformatica e la scienza del caos deterministico. Nonostante i passi da gi-

gante compiuti dalla scienza da trent'anni a questa parte siamo, tuttavia, ancora molto lontani da una risposta adeguata a questa domanda: a tutt'oggi, infatti, non esiste una risposta a carattere compiuto, esistono al contrario fasci correlati di ipotesi, agglomerati di ardite congetture in grado di trasformare lo studio scientifico della vita in un'interrogazione. Alla luce di tutto ciò, dunque, a nostro giudizio, nonostante il raggiungimento di sorprendenti risultati e l'apertura di nuovi orizzonti di ricerca prima impensabili, la scienza contemporanea si trova davanti all'impossibilità concreta di poter rispondere in maniera definitiva alla domanda riguardante l'essenza ultima della vita: la complessità che la pervade resta, pertanto, intrisa di mistero. Lungi dal farci cadere in una condizione di sconforto, quest'affermazione, tuttavia apre nuovi orizzonti al nostro sguardo e ci invita a raccogliere le nostre forze per un balzo in avanti ancora più articolato e profondo. Per concludere, quindi, possiamo affermare che tramite la nascita del nuovo paradigma scientifico legato alla scienza della complessità, oggi l'affascinante questione concernente l'essenza della vita ci addita la necessità di fare riferimento a una circolarità nuova in grado di avvicinare una filosofia che scopre sempre più in profondità il ruolo di analisi per quel che riguarda le grandi domande relative al senso della vita a una scienza che a sua volta gradatamente giunge a divenire consapevole della necessità di non smarrire mai, lungo il corso della ricerca, i profondi interrogativi concernenti il significato<sup>11</sup>.

## RIFERIMENTI BIBLIOGRAFICI

- ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. e WALTER, P. (2002): *Molecular Biology of the Cell*, IV ed., New York: Garland Science.
- ATLAN, H. (1985): "Complessità, disordine e auto-creazione del significato", trad. it. in G. Bocchi e M. Ceruti (a cura di), *La sfida della complessità*, Milano: Mondadori, 2007, pp. 134-154.
- CARSETTI, A. (1972): "Modelli cibernetici e biologia molecolare", in *Problemi filosofici della biologia*, Bologna: CNR, pp. 5-16.
- (1989): "Teoria della complessità e modelli della conoscenza", *La Nuova Critica*, 9-10, pp. 61-103.

---

<sup>11</sup> Per un approfondimento circa l'esame effettivo delle nuove misure d'informazione relative al "codice semantico" dei sistemi biologici e in particolare della mente umana, quali esse sono state divise, nel corso di questi ultimissimi anni, si vedano: Carsetti (di prossima pubblicazione) e Carsetti (a cura di) (2010).

- (1993): “Meaning and Complexity: the Role of Non-Standard Models”, *La Nuova Critica*, 22, pp. 57-86.
- (1996): “Chaos, Natural Order and Molecular Semantics”, *La Nuova Critica*, 27-28, pp. 83-107.
- (2005): “Entropia, morfogenesi autonoma e sistemi cognitivi”, *La Nuova Critica*, 45-46, pp. 59-105.
- *Epistemic Complexity and Knowledge Construction*, Berlin: Springer, di prossima pubblicazione.
- (2010) (a cura di): *Causality, Meaningful Complexity and Embodied Cognition*, Berlin: Springer, 2010.
- CONTINENZA, B. (1987): “Il Theoretical Biology Club e la fallacia del riduzionismo”, *Metamorfosi*, 6, pp. 169-210.
- DI BERNARDO, M. (2007a): *Per una rivisitazione della dottrina monodiana della morfogenesi autonoma alla luce dei nuovi scenari aperti dalla post-genomica*, Roma: Aracne.
- (2007b): “La Genomica funzionale: la crisi della concezione del DNA come programma genetico”, *Dialegesthai*, 9, <http://mondodamani.org/dialegesthai>.
- (2010): “Che cos'è la vita? Complessità e significato in biologia”, *Información Filosófica*, 13, pp. 1-12.
- *I sentieri evolutivi della complessità biologica nell'opera di S. A. Kauffman*, Milano: Mimesis, di prossima pubblicazione.
- GIBBS, W. W. (2003): “Unseen Genome: Gems among the Junk”, *Scientific American*, 289, 5, pp. 46-53.
- JACOB, F., e MONOD, J. (1961a): “On the Regulation of Gene Activity”, *Cold Spring Harbor Symposium on Quantitative Biology*, 26, pp. 193-211.
- (1961b): “Genetic Regulatory Mechanisms in the Synthesis of Proteins”, *Journal of Molecular Biology*, 3, pp. 318-356.
- KAUFFMAN, S. A. (1970): “Behaviour of Randomly Constructed Genetic Net: Binary Element Nets”, in C. H. Waddington (a cura di), *Towards a Theoretical Biology*, Chicago: Aldine Publishing, pp. 18-36.
- (1991): “Antichaos and Adaptation”, *Scientific American*, 265, pp. 64-70 (trad. it. “Anticaos ed evoluzione biologica”, *Le Scienze*, 47, 1991, pp. 82-91).
- (1993): *The Origins of Order. Self-Organization and Selection in Evolution*, New York: Oxford University Press.
- (1995): *At Home in the Universe. The Search of the Laws of Self-Organization and Complexity*, New York: Oxford University Press (trad. it. *A casa nell'universo. Le leggi del caos e della complessità*, Roma: Editori Riuniti, 2001).
- (2000): *Investigations*, New York: Oxford University Press (trad. it. *Esplorazioni evolutive*, Torino: Einaudi, 2005).

- (2004a): “A Proposal for Using the Ensemble Approach to Understand Genetic Regulatory Networks”, *Theoretical Biology*, 230, 4, pp. 581-590.
- (2004b): “The Ensemble Approach to Understand Genetic Regulatory Networks”, *Physica, A* 340, pp. 733-740.
- (2008): *Reinventing the Sacred. A New View of Science, Reason and Religion*, New York: Basic Books (trad.it. *Reinventare il sacro. Una nuova concezione della scienza, della ragione e della religione*, Torino: Codice Edizioni, 2010).
- LANGTON, C. G. (1986): “Studying Artificial Life with Cellular Automata”, *Physica*, 22D, pp. 120-149
- MAYER-KRESS, G. (a cura di) (1986): *Dimensions and Entropies in Chaotic Systems: Quantification of Complex Behavior*, Berlin: Springer.
- MANDELBROT, B. (1977): *The Fractal Geometry of Nature*, San Francisco: Freeman.
- MONOD, J. (1970): *Le hasard et la necessite: essai sur la philosophie naturelle de la biologie moderne*, Paris: Seuil (trad. it. *Il caso e la necessità*, Milano: Mondadori).
- MONOD, J., e JACOB, F. (1961): “General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth, and Differentiation”, *Cold Spring Harbor Symposium on Quantitative Biology*, 26, pp. 306-329.
- MONOD, J., CHANGEUX, J., e JACOB, F. (1963): “Allosteric Proteins and Cellular Control System”, *Journal of Molecular Biology*, 6, pp. 306-329.
- PACKARD, N. (1988): “Adaptation to the Edge of Chaos”, in J. A. S. Kelso, A. J. Mandell e M. F. Schlesinger (a cura di), *Dynamic Patterns in Complex Systems*, Singapore: World Scientific, pp. 293-301.
- RIBEIRO, A., e KAUFFMAN, S.A. (2007): “Noisy Attractors and Ergodic Sets in Models of Gene Regulatory Networks”, *Journal of Theoretical Biology*, 247, 4, pp. 743-755.
- RUELLE, D. (1979): “Sensitive Dependence on Initial Condition and Turbulent Behavior of Dynamical Systems”, *Annals of the New York Academy of Sciences*, 316, pp. 408-416.

# Continuous Functions as Quantum Operations: a Probabilistic Approximation\*

Hector Freytes  
Instituto Argentino de Matemática (CONICET)  
e-mail: hfreytes@gmail.com

Antonio Ledda  
University of Cagliari  
e-mail: antonio.ledda78@gmail.com

Giuseppe Sergioli  
University of Cagliari  
e-mail: giuseppe.sergioli@gmail.com

- 1 Introduction
- 2 Basic notions
- 3 A probabilistic Stone-Weierstrass type theorem
- 4 Representing the standard PMV-operations
- 5 Conclusions and open problems

ABSTRACT. In this note we propose a version of the classical Stone-Weierstrass theorem in the context of quantum operations, by introducing a particular class of quantum operations, dubbed *polynomial quantum operations*. This result permits to interpret from a probabilistic point of view, and up to a certain approximation, any continuous function from the real cube  $[0, 1]^n$  to the real interval  $[0, 1]$  as a quantum operation.

KEYWORDS: quantum operations; PMV-algebras; quantum computation.

---

\*A. Ledda and G. Sergioli were supported by Regione Autonoma della Sardegna, POR Sardegna FSE-M.S. 2007-2013 L.R. 7/2007.

## 1. Introduction

Since the classical work of Birkhoff and von Neumann [4], logical and algebraic perspectives of several aspects of quantum theory have been proposed. Leading examples are orthomodular lattices [21], and effect algebras, appeared independently under several names (e.g. D-posets [23]) as a generalization of orthomodular posets [13, 16]. Moreover, effect algebras play a fundamental role in various studies on fuzzy probability theory [3, 17] also.

In recent times, quantum computation itself stirred increasing attention, and an array of related algebraic structures arose [5, 6, 18, 15]. Those structures stem from an abstract description of circuits obtained by combinations of quantum gates [7]. Let us mention as examples *quantum MV-algebras* [12], *quasi MV-algebras*,  $\sqrt{\cdot}$ *quasi MV-algebras* and *product MV-algebras* [25, 14, 10, 28].

Even if those structures are plainly related to quantum computing, some of the functions in their types are algebraic abstractions of irreversible transformations, e.g. the truncated disjunction “ $\oplus$ ” [7].

These observations provides a general framework for a probabilistic-type representation of continuous functions in the real interval  $[0, 1]$  as *quantum operations* in the sense of [24]. Therefore, in the present paper we show that all those algebraic structures mentioned so far are fully settled into the general model of quantum computing, based on quantum operations acting on density operators.

The irreversible quantum operational approach has plenty of advantages in the implementation of quantum computational devices: as Aharonov, Kitaev and Nisan discussed [1], there are several relevant problems to deal with in the usual unitary model of quantum computation. Those problems (such as measurements, or noise and decoherence ) disappear in the non-unitary (i.e. non-reversible) model. In fact, although quantum computational processes permit measurements in the middle of the computation, however, the state of the computation after a measurement is a mixed state. Moreover, to implement quantum computers, noise and, in particular, decoherence are important obstacles. The main problem in this interface between quantum physics and quantum computation models is that quantum noise and decoherence are non-unitary operations that cause a pure state to become a mixed state.

In this note, we propose a simple construction on density matrices (dubbed *polynomial operations*) that permits to resort, in terms of probability distribu-

tions, to a Stone-Weierstrass type theorem. That result implies that any continuous function can be regarded, from a probabilistic point of view, and up to a certain approximation, as a quantum operation, Theorems 3.3 and 3.4. The results are organized as follows: in Section 2 we provide all the basic notions, in Section 3 we show an overview of a probabilistic version of the Stone-Weierstrass theorem in the framework of quantum operations (further details are showed in [11]), in Section 4 some applications of our main result to the case of product MV-algebra [28] are given, and lastly in Section 5 some possible future investigation issues are illustrated and a few conclusive remarks drawn.

## 2. Basic notions

A quantum system in a pure state is described by a unit vector in a Hilbert space. In the Dirac notation a pure state is denoted by  $|\psi\rangle$  ( $\langle\psi|$ ). A *quantum bit* or *qubit*, the fundamental concept of quantum computation, is a pure state in the Hilbert space  $\mathbb{C}^2$ . The standard orthonormal basis  $\{|0\rangle, |1\rangle\}$  of  $\mathbb{C}^2$  is called the *logical basis*. Thus a qubit  $|\psi\rangle$  may be written as a linear superposition of the basis vectors with complex coefficients  $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$  with  $|c_0|^2 + |c_1|^2 = 1$ .

Quantum mechanics reads out the information content of a pure state via the Born rule, according to which the probability value assigned to a qubit is defined as follows:

$$p(|\psi\rangle) = |c_1|^2.$$

The states of interest for quantum computation lie in the tensor product  $\otimes^n \mathbb{C}^2 = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2$ , where  $\otimes^n \mathbb{C}^2 = \mathbb{C}^2$  if  $n = 1$ . The space  $\otimes^n \mathbb{C}^2$  is a  $2^n$ -dimensional complex Hilbert space. The  $2^n$ -*computational basis* consists of the  $2^n$  orthogonal states  $|t\rangle$  ( $0 \leq t \leq 2^n$ ), where  $t$  is in binary representation and it can be seen as the tensor product of the states  $|t_1\rangle \otimes |t_2\rangle \otimes \dots \otimes |t_n\rangle$  where  $t_j \in \{0, 1\}$ . A pure state  $|\psi\rangle \in \otimes^n \mathbb{C}^2$  is generally a superposition of the basis vectors:  $|\psi\rangle = \sum_{t=1}^{2^n} c_t |t\rangle$  with  $\sum_{t=1}^{2^n} |c_t|^2 = 1$ .

In general, a quantum system is not in a pure state. This might be because the system is coupled with an environment, it is subject to a measurement process etc. In those cases, the state-evolution is no longer reversible and the system is said to be in a *mixed state*. A convenient mathematical description of a

mixed state is given by the notion of *density operator*, i.e. an Hermitian positive operator  $\rho$  on a  $2^n$ -dimensional complex Hilbert space with trace  $\text{tr}(\rho) = 1$ .

A pure state  $|\psi\rangle$  can be represented as a limit case of mixed state in the form  $\rho = |\psi\rangle\langle\psi|$ . In particular, each vector of the logical basis of  $\mathbb{C}^2$  can be associated to a density operator  $P_0 := |0\rangle\langle 0|$  or  $P_1 := |1\rangle\langle 1|$  that represents the *falsity-property* and the *truth-property*, respectively. One can represent an arbitrary density matrix  $\rho$  in terms of a tensor products of the Pauli matrices:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

in the following way:

$$\rho = \frac{1}{2^n} \sum_{\mu_1 \dots \mu_n} P_{\mu_1 \dots \mu_n} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n},$$

where  $\mu_i \in \{0, x, y, z\}$  for each  $i \in \{1, \dots, n\}$ . The real expansion coefficients  $P_{\mu_1 \dots \mu_n}$  are given by  $P_{\mu_1 \dots \mu_n} = \text{tr}(\sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_n} \rho)$ . Since the eigenvalues of the Pauli matrices are  $\pm 1$ , the expansion coefficients satisfy the inequality  $|P_{\mu_1 \dots \mu_n}| \leq 1$ . In what follows, for sake of simplicity, we will use without distinction  $I$  or  $\sigma_0$ . We denote by  $\mathcal{D}(\otimes^n \mathbb{C}^2)$  the set of all density operators of  $\otimes^n \mathbb{C}^2$ ; hence the set  $\mathcal{D} = \bigcup_{i \in N} \mathcal{D}(\otimes^n \mathbb{C}^2)$  will be the set of all possible density operators. Moreover, we can identify in each space  $\mathcal{D}(\otimes^n \mathbb{C}^2)$  two special operators  $P_0^{(n)} = \frac{1}{2^n} I^{n-1} \otimes P_0$  and  $P_1^{(n)} = \frac{1}{2^n} I^{n-1} \otimes P_1$  that represent, in this framework, the falsity-property and the truth-property, respectively. The *probability of truth*  $p$  of a density operator  $\rho$  is dictated by the Born rule and equals

$$p(\rho) = \text{tr}(P_1^{(n)} \rho).$$

In case  $\rho = |\psi\rangle\langle\psi|$ , where  $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ , then  $p(\rho) = |c_1|^2$ .

Let  $\rho \in \mathcal{D}(\mathbb{C}^2)$ . Then  $\rho$  can be represented as a linear superposition  $\rho = \frac{1}{2}(I + r_x \sigma_x + r_y \sigma_y + r_z \sigma_z)$ , where  $r_x, r_y, r_z$  are real numbers such that  $r_x^2 + r_y^2 + r_z^2 \leq 1$ . Therefore, every density operator  $\rho$  in  $\mathcal{D}(\mathbb{C}^2)$  has the matrix representation:

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix} = \begin{pmatrix} 1 - \alpha & \beta \\ \beta^* & \alpha \end{pmatrix} \quad (1)$$



Furthermore, any real number  $\lambda$  ( $0 \leq \lambda \leq 1$ ) uniquely determines a density operator as follows:

$$\rho_\lambda = (1 - \lambda)P_0 + \lambda P_1 = \frac{1}{2}(I + (1 - 2\lambda)\sigma_z) = \begin{pmatrix} 1 - \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad (2)$$

In virtue of (1) and (2), one may verify that, whenever  $\rho \in \mathcal{D}(\mathbb{C}^2)$ , then  $p(\rho) = \frac{1 - r_z}{2}$  and  $p(\rho_\lambda) = \lambda$ . Thus each density operator  $\rho$  in  $\mathcal{D}(\mathbb{C}^2)$  can be written as

$$\rho = \begin{pmatrix} 1 - p(\rho) & a \\ a^* & p(\rho) \end{pmatrix} \quad (3)$$

In the usual model of quantum computation the state of a system is pure and the operations (*quantum gates*) are represented by unitary operators. Nevertheless, in case a system is not completely isolated from the environment its evolution is, in general, irreversible. A model of quantum computing that relates to that phenomenon is mathematically described by means of *quantum operations* (as quantum gates) acting on density operators (as information quantities).

Given a finite dimensional complex Hilbert space  $H$ , we will denote by  $\mathcal{L}(H)$  the vector space of all linear operators on  $H$ . Let  $H_1, H_2$  be two finite dimensional complex Hilbert spaces. A *super operator* is a linear operator  $\mathcal{E} : \mathcal{L}(H_1) \rightarrow \mathcal{L}(H_2)$  sending density operators to density operators [2]. This is equivalent to say that  $\mathcal{E}$  is trace-preserving and positive, i.e. sends positive semi-definite Hermitian operators to positive semi-definite Hermitian operators. A super operator  $\mathcal{E}$  is said to be a *quantum operation* iff the super operator  $\mathcal{E} \otimes I_H$  is positive, where  $I_H$  is the identity super operator on an arbitrary finite dimensional complex Hilbert space  $H$ . In this case  $\mathcal{E}$  is also called *completely positive*. The following theorem, due to K. Kraus [24], provide an equivalent definition of quantum operations:

**Theorem 2.1.** A linear operator  $\mathcal{E} : \mathcal{L}(H_1) \rightarrow \mathcal{L}(H_2)$  is a quantum operation iff  $\forall \rho \in \mathcal{L}(H_1)$ :

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger$$

for some set of operators  $\{A_i\}$  such that  $\sum_i A_i^\dagger A_i = I$ .

### 3. A probabilistic Stone-Weierstrass type theorem

The aim of the present section is to propose a representation, in probabilistic terms, of a particular class of polynomials via quantum operations. Such a result will be expedient to prove a probabilistic Stone-Weierstrass type theorem. First of all, let us introduce some notations and preliminary definitions. The term *multi-index* denotes an ordered  $n$ -tuple  $\alpha = (\alpha_1, \dots, \alpha_n)$  of non negative integers  $\alpha_i$ . The *order* of  $\alpha$  is given by  $|\alpha| = \alpha_1 + \dots + \alpha_n$ . If  $\mathbf{x} = (x_1, \dots, x_n)$  is an  $n$ -tuple of variables and  $\alpha = (\alpha_1, \dots, \alpha_n)$  a multi-index, the monomial  $\mathbf{x}^\alpha$  is defined by  $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ . In this language a real polynomial of order  $k$  is a function  $P(\mathbf{x}) = \sum_{|\alpha| \leq k} a_\alpha \mathbf{x}^\alpha$  such that  $a_\alpha \in \mathbb{R}$ .

Let  $\mathbf{x} = (x_1, \dots, x_n)$  and  $k$  be a natural number. If we define the set  $D_k(\mathbf{x})$  as follows:

$$D_k(\mathbf{x}) = \{(1-x_1)^{\alpha_1} x_1^{\beta_1} \dots (1-x_n)^{\alpha_n} x_n^{\beta_n} : \alpha_i + \beta_i = k, i \in \{1, \dots, n\}\} \quad (4)$$

then we obtain the following useful lemmas:

**Lemma 3.1.** Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be a family of matrices such that

$$\mathbf{X}_i = \begin{pmatrix} 1-x_i & b_i \\ b_i^* & x_i \end{pmatrix}$$

and let  $\mathbf{X} = (\otimes^k \mathbf{X}_1) \otimes (\otimes^k \mathbf{X}_2) \otimes \dots \otimes (\otimes^k \mathbf{X}_n)$ . Then

$$Diag(\mathbf{X}) = D_k(x_1, \dots, x_n).$$

*Proof.* It can be verified that  $\otimes^k \mathbf{X}_i = \{h_1 h_2 \dots h_k : h_j \in \{(1-x_i), x_i\}, 1 \leq j \leq k\} = \{(1-x_1)^{\alpha_1} x_1^{\beta_1} : \alpha + \beta = k\}$ . Thus,  $(\otimes^k \mathbf{X}_1) \otimes (\otimes^k \mathbf{X}_2) \otimes \dots \otimes (\otimes^k \mathbf{X}_n) = \{(1-x_1)^{\alpha_1} x_1^{\beta_1} \dots (1-x_n)^{\alpha_n} x_n^{\beta_n} : \alpha_i + \beta_i = k, i \in \{1, \dots, n\}\}$ . Whence our claim follows.  $\square$

**Lemma 3.2.** Let  $\mathbf{x} = (x_1, \dots, x_n)$  and  $k$  be a natural number. For any monomial  $\mathbf{x}^\alpha$ , such that  $|\alpha| \leq k$ , the following conditions hold:

1.  $\mathbf{x}^\alpha = \sum_{\mathbf{y} \in D_k(\mathbf{x})} \delta_{\mathbf{y}} \mathbf{y}$ ;
  2.  $1 - \mathbf{x}^\alpha = \sum_{\mathbf{y} \in D_k(\mathbf{x})} \gamma_{\mathbf{y}} \mathbf{y}$ ;
- where  $\delta_{\mathbf{y}}$  and  $\gamma_{\mathbf{y}}$  are in  $\{0, 1\}$ .

*Proof.* First, let us define, for any  $i \in \{1, \dots, n\}$ , a matrix  $\mathbf{X}_i$  as follows

$$\mathbf{X}_i = \begin{pmatrix} 1 - x_i & 0 \\ 0 & x_i \end{pmatrix}$$

- 1) Let  $\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$  such that  $|\alpha| \leq k$ . Thus, there are  $s_1, \dots, s_n$  such that  $\alpha_i + s_i = k$ . Let  $\mathbf{W} = (\otimes^{s_1} \mathbf{X}_1) \otimes (\otimes^{s_2} \mathbf{X}_2) \otimes \dots \otimes (\otimes^{s_n} \mathbf{X}_n)$  and consider the matrix  $\mathbf{W}\mathbf{x}^\alpha$ . In view of Lemma 3.1,  $\text{Diag}(\mathbf{W}\mathbf{x}^\alpha) \subseteq D_k(x_1, \dots, x_n)$  since every element in  $\text{Diag}(\mathbf{W}\mathbf{x}^\alpha)$  is a monomial of order  $nk$ . Further, since  $\text{tr}(\mathbf{W}\mathbf{x}^\alpha) = (\text{tr} \mathbf{W})\mathbf{x}^\alpha = 1\mathbf{x}^\alpha = \mathbf{x}^\alpha$ , we obtain  $\mathbf{x}^\alpha = \text{tr}(\mathbf{W}\mathbf{x}^\alpha)$ , i.e. the required polynomial expansion.
- 2) Let  $\mathbf{X} = (\otimes^k \mathbf{X}_1) \otimes (\otimes^k \mathbf{X}_2) \otimes \dots \otimes (\otimes^k \mathbf{X}_n)$ . By Lemma 3.1,  $\text{Diag}(\mathbf{X}) = D_k(x_1, \dots, x_n)$  and  $\text{tr}(\mathbf{X}) = 1$ . Upon recalling that  $\mathbf{x}^\alpha = \sum_{\mathbf{y} \in D_k(\mathbf{x})} \delta_{\mathbf{y}} \mathbf{y}$ , we define  $\gamma_{\mathbf{y}} = 1$  if  $\delta_{\mathbf{y}} = 0$  and  $\gamma_{\mathbf{y}} = 0$  if  $\delta_{\mathbf{y}} = 1$ . Therefore,

$$\begin{aligned} 1 &= \text{tr}(\mathbf{X}) \\ &= \sum_{\mathbf{y} \in D_k(\mathbf{x})} \delta_{\mathbf{y}} \mathbf{y} + \sum_{\mathbf{y} \in D_k(\mathbf{x})} \gamma_{\mathbf{y}} \mathbf{y} \\ &= \mathbf{x}^\alpha + \sum_{\mathbf{y} \in D_k(\mathbf{x})} \gamma_{\mathbf{y}} \mathbf{y} \end{aligned}$$

$$\text{and } 1 - \mathbf{x}^\alpha = \sum_{\mathbf{y} \in D_k(\mathbf{x})} \gamma_{\mathbf{y}} \mathbf{y}.$$

□

In virtue of the previous claims, we can prove a technical but rather important theorem:

**Theorem 3.3.** Let  $\mathbf{x} = (x_1, \dots, x_n)$  be an  $n$ -tuple of variables, and let  $P(\mathbf{x}) = \sum_{\mathbf{y} \in D_k(\mathbf{x})} a_{\mathbf{y}} \mathbf{y}$  be a polynomial such that  $\mathbf{y} \in D_k(\mathbf{x})$ ,  $0 \leq a_{\mathbf{y}} \leq 1$  and the restriction  $P(\mathbf{x}) \upharpoonright_{[0,1]^n}$  be such that  $0 \leq P(\mathbf{x}) \upharpoonright_{[0,1]^n} \leq 1$ . Then there exists a polynomial quantum operation  $\mathcal{P} : \mathcal{L}(\otimes^{nk} \mathbb{C}^2) \rightarrow \mathcal{L}(\otimes^{nk} \mathbb{C}^2)$  such that, for any  $n$ -tuple  $\sigma = (\sigma_1, \dots, \sigma_n)$  in  $\mathcal{D}(\mathbb{C}^2)$ ,

$$\mathfrak{p}(\mathcal{P}((\otimes^k \sigma_1) \otimes \dots \otimes (\otimes^k \sigma_n))) = P(\mathfrak{p}(\sigma_1), \dots, \mathfrak{p}(\sigma_n)).$$

Moreover,

$$\mathcal{P}((\otimes^k \sigma_1) \otimes \dots \otimes (\otimes^k \sigma_n)) = \left( \frac{1}{2^{nk-1}} \otimes^{nk-1} I \right) \otimes \rho_{P(\mathfrak{p}(\sigma_1), \dots, \mathfrak{p}(\sigma_n))}.$$

*Proof.* Let  $\sigma_1, \dots, \sigma_n$  be density operators on  $\mathbb{C}^2$ . Assume that for any  $\sigma_i$

$$\sigma_i = \begin{pmatrix} 1 - x_i & b_i \\ b_i^* & x_i \end{pmatrix}$$

Hence,  $p(\sigma_i) = x_i$ . Evidently,  $\sigma = (\otimes^k \sigma_1) \otimes \dots \otimes (\otimes^k \sigma_n)$  is a  $2^{nk} \times 2^{nk}$  matrix and, by Lemma 3.1,  $\text{Diag}(\sigma) = D_k(x_1, \dots, x_n)$ . Thus, each  $\mathbf{y} \in D_k(\mathbf{x})$  can be seen as the  $(i, i)$ -th entry of  $\text{Diag}(\sigma)$ . Further, the polynomial  $P(\mathbf{x}) = \sum_{\mathbf{y} \in D_k(\mathbf{x})} a_{\mathbf{y}} \mathbf{y} = \sum_{j=1}^{2^{nk}} a_j \mathbf{y}_j$  is such that every  $\mathbf{y}_j$  is the  $(j, j)$ -th entry of  $\text{Diag}(\sigma)$ . Let, now,  $\mathbf{y}_{j_0} \in \text{Diag}(\sigma)$ .

a) We want to place the elements of the form  $a_{j_0} \mathbf{y}_{j_0}$  in the  $(2s, 2s)$ -th entries of a  $2^{nk} \times 2^{nk}$  matrix. Let us consider the  $2^{nk} \times 2^{nk}$  matrix  $A_{j_0}^{2s} = \sqrt{\frac{a_{j_0}}{2^{nk-1}}} D_{j_0}^{2s}$  such that  $D_{j_0}^{2s}$  has 1 just in the  $(2s, j_0)$ -th entry and 0 in any other entry. One may verify that  $A_{j_0}^{2s} \sigma (A_{j_0}^{2s})^\dagger$  is the required matrix. Moreover:

$$\sum_{2s} A_{j_0}^{2s} \sigma (A_{j_0}^{2s})^\dagger = \frac{1}{2^{nk-1}} \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & a_{j_0} \mathbf{y}_{j_0} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & a_{j_0} \mathbf{y}_{j_0} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

b) Let us recall that  $1 = \sum_{\mathbf{y} \in D_k(\mathbf{x})} \mathbf{y} = \sum_{j=1}^{2^{nk}} \mathbf{y}_j$ . Then:

$$1 - \sum_{j=1}^{2^{nk}} a_j \mathbf{y}_j = \sum_{j=1}^{2^{nk}} \mathbf{y}_j - \sum_{j=1}^{2^{nk}} a_j \mathbf{y}_j = \sum_{j=1}^{2^{nk}} (1 - a_j) \mathbf{y}_j.$$

We now want to stick the elements of the form  $(1 - a_{j_0}) \mathbf{y}_{j_0}$  into the  $(2s - 1, 2s - 1)$ -th entries of a  $2^{nk} \times 2^{nk}$  matrix. Let us consider the  $2^{nk} \times 2^{nk}$  matrix  $A_{j_0}^{2s-1} = \sqrt{\frac{1 - a_{j_0}}{2^{nk-1}}} D_{j_0}^{2s-1}$  such that  $D_{j_0}^{2s-1}$  have 1 just in the  $(2s - 1, j_0)$ -th entry and 0 in any other entry. Again, one may verify that  $A_{j_0}^{2s-1} \sigma (A_{j_0}^{2s-1})^\dagger$  is the required matrix. Furthermore:

$$\sum_{2s-1} A_{j_0}^{2s-1} \sigma (A_{j_0}^{2s-1})^\dagger = \frac{1}{2^{nk-1}} \begin{pmatrix} (1 - a_{j_0}) \mathbf{y}_{j_0} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & (1 - a_{j_0}) \mathbf{y}_{j_0} & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Thus, some calculations show that:

$$\begin{aligned} \mathcal{P} &= \sum_{j_0} \sum_{2s} A_{j_0}^{2s} \sigma(A_{j_0}^{2s})^\dagger + \sum_{j_0} \sum_{2s-1} A_{j_0}^{2s-1} \sigma(A_{j_0}^{2s-1})^\dagger \\ &= \left( \frac{1}{2^{kn-1}} \otimes^{nk-1} I \right) \otimes \begin{pmatrix} 1 - \sum_{j=1}^{2^{nk}} a_j \mathbf{y}_j & 0 \\ 0 & \sum_{j=1}^{2^{nk}} a_j \mathbf{y}_j \end{pmatrix} \end{aligned}$$

Now, set  $A = \sum_{j_0} \sum_{2s} (A_{j_0}^{2s})^\dagger A_{j_0}^{2s} + \sum_{j_0} \sum_{2s+1} (A_{j_0}^{2s+1})^\dagger A_{j_0}^{2s+1}$ . Our task is to verify that  $A = I$ .

- c) First of all, notice that the matrix  $(A_{j_0}^{2s})^\dagger A_{j_0}^{2s}$  has the value  $\frac{a_{j_0}}{2^{nk-1}}$  just in the  $(j_0, j_0)$ -th entry, while any other entry is 0. Thus, the matrix  $\sum_{2s} (A_{j_0}^{2s})^\dagger A_{j_0}^{2s}$  has the value  $\frac{2^{nk-1} a_{j_0}}{2^{nk-1}} = a_{j_0}$  in the  $(j_0, j_0)$ -th entry and all the other entries equal 0. Hence:

$$\sum_{j_0} \sum_{2s} (A_{j_0}^{2s})^\dagger A_{j_0}^{2s} = \begin{pmatrix} a_1 & 0 & 0 & 0 & \dots \\ 0 & a_2 & 0 & 0 & \dots \\ 0 & 0 & a_3 & 0 & \dots \\ 0 & 0 & 0 & a_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- d) On the other hand the matrix  $(A_{j_0}^{2s-1})^\dagger A_{j_0}^{2s-1}$  has the value  $\frac{1-a_{j_0}}{2^{nk-1}}$  just in the  $(j_0, j_0)$ -th entry and 0 in any other. Thus, the matrix  $\sum_{2s-1} (A_{j_0}^{2s-1})^\dagger A_{j_0}^{2s-1}$  has the value  $\frac{2^{nk-1}(1-a_{j_0})}{2^{nk-1}} = 1 - a_{j_0}$  in the  $(j_0, j_0)$ -th entry and any other is 0. Hence:

$$\sum_{j_0} \sum_{2s-1} (A_{j_0}^{2s-1})^\dagger A_{j_0}^{2s-1} = \begin{pmatrix} 1 - a_1 & 0 & 0 & 0 & \dots \\ 0 & 1 - a_2 & 0 & 0 & \dots \\ 0 & 0 & 1 - a_3 & 0 & \dots \\ 0 & 0 & 0 & 1 - a_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Thus,  $\sum_{j_0} \sum_{2s} (A_{j_0}^{2s})^\dagger A_{j_0}^{2s-1} + \sum_{j_0} \sum_{2s-1} (A_{j_0}^{2s-1})^\dagger A_{j_0}^{2s-1} = I$ . Whence, by Theorem 2.1,  $\mathcal{P}$  is a quantum operation.

□

In virtue of Theorem 3.3, we can now establish a probabilistic version of the classical Stone-Weierstrass theorem.

**Theorem 3.4.** Let  $\mathbf{x} = (x_1, \dots, x_n)$  be an  $n$ -tuple of variables and  $f : [0, 1]^n \rightarrow (0, 1)$  be a continuous function. Then, for each  $\varepsilon > 0$  there exists a quantum operation  $\mathcal{P} : \mathcal{L}(\otimes^{nk}\mathbb{C}^2) \rightarrow \mathcal{L}(\otimes^{nk}\mathbb{C}^2)$  and a constant  $M \geq 1$  such that, for any density matrix  $\sigma = (\otimes^k\sigma_1) \otimes \dots \otimes (\otimes^k\sigma_n)$ , the following inequality holds:

$$|\mathfrak{p}(\mathcal{P}(\sigma)) - \frac{1}{M}f(\mathfrak{p}(\sigma_1), \dots, \mathfrak{p}(\sigma_n))| \leq \varepsilon.$$

*Proof.* Let  $f : [0, 1]^n \rightarrow (0, 1)$  be a continuous function. By the classical Stone-Weierstrass theorem, there exists a polynomial  $P_0(x_1, \dots, x_n) = \sum_{|\alpha| \leq k} a_\alpha \mathbf{x}^\alpha$  such that for each  $\varepsilon > 0$ ,  $|P_0 - f| \leq \frac{\varepsilon}{2}$ . Let  $a_{\alpha_1}, \dots, a_{\alpha_n}$  be positive coefficients and  $a_{\beta_1}, \dots, a_{\beta_s}$  be negative coefficients in the polynomial  $P_0(x_1, \dots, x_n)$ . Let  $M$  be a positive real number such that  $\frac{\sum_{i=1}^s |a_{\beta_i}|}{M} \leq \frac{\varepsilon}{2}$ . Let us define a polynomial  $P$  by  $P(x_1, \dots, x_n) = \sum_{i=1}^n \frac{a_{\alpha_i}}{M} \mathbf{x}^{\alpha_i} + \sum_{j=1}^s \frac{|a_{\beta_j}|}{M} (1 - \mathbf{x}^{\beta_j})$ . Then, we obtain that  $P(x_1, \dots, x_n) = \frac{1}{M}P_0(x_1, \dots, x_n) + \frac{\sum_{i=1}^s |a_{\beta_i}|}{M}$ . Therefore, in  $[0, 1]^n$ :

$$\begin{aligned} |P - \frac{1}{M}f| &= |P - \frac{1}{M}P_0 + \frac{1}{M}P_0 - \frac{1}{M}f| \\ &\leq |P - \frac{1}{M}P_0| + |\frac{1}{M}P_0 - \frac{1}{M}f| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{M} \\ &\leq \varepsilon \end{aligned}$$

So, by Lemma 3.2, we obtain that  $P(\mathbf{x}) = \sum_{\mathbf{y} \in D_k(\mathbf{x})} a_{\mathbf{y}} \mathbf{y}$  is such that  $a_{\mathbf{y}} \geq 0$  and  $0 \leq P(\mathbf{x}) \leq 1$  in  $[0, 1]^n$ . Whence, by Theorem 3.3, there exists a quantum operation  $\mathcal{P} : \mathcal{L}(\otimes^{nk}\mathbb{C}^2) \rightarrow \mathcal{L}(\otimes^{nk}\mathbb{C}^2)$  associated to  $P$  such that for each  $n$ -tuple  $\sigma = (\sigma_1, \dots, \sigma_n)$ , with  $\sigma_i$  in  $\mathcal{D}(\mathbb{C}^2)$ ,  $\mathfrak{p}(\mathcal{P}((\otimes^k\sigma_1) \otimes \dots \otimes (\otimes^k\sigma_n))) = P(x_1/\mathfrak{p}(\sigma_1), \dots, x_n/\mathfrak{p}(\sigma_n))$ .<sup>1</sup> Thus

$$|\mathfrak{p}(\mathcal{P}((\otimes^k\sigma_1) \otimes \dots \otimes (\otimes^k\sigma_n))) - \frac{1}{M}f(\mathfrak{p}(\sigma_1), \dots, \mathfrak{p}(\sigma_n))| \leq \varepsilon.$$

□

<sup>1</sup> By  $x_i/\mathfrak{p}(\sigma_i)$  we mean the attribution of the value  $\mathfrak{p}(\sigma_i)$  to the variable  $x_i$ .

#### 4. Representing the standard PMV-operations

In this section we apply the results obtained to two functions (namely, the product  $t$ -norm  $\bullet$ , and the Łukasiewicz conorm  $\oplus$ ) of definite importance in fuzzy logic. Let us recall some notions first.

The *standard PMV-algebra* [10, 28] is the algebra

$$[0, 1]_{PMV} = \langle [0, 1], \oplus, \bullet, \neg, 0, 1 \rangle,$$

where  $[0, 1]$  is the real unit segment,  $x \oplus y = \min(1, x + y)$ ,  $\bullet$  is the real product, and  $\neg x = 1 - x$ . This structure plays a notable role in quantum computing, in that it describes, in a probabilistic way, a relevant system of quantum gates named *Poincarè irreversible quantum computational algebra* [5, 8]. The connection between quantum computational logic with mixed states and fuzzy logic, comes from the election of a system of quantum gates such that, when interpreted under probabilistic semantics, they turn out in some kind of operation in the real interval  $[0, 1]$ . The above-mentioned *PMV – algebra* is a structure that represents algebraic counterpart of the probabilistic semantics conceived from the continuous  $t$ -norm. On the other hand, the use of fuzzy logics (and infinite-valued Łukasiewicz logic in particular) in game theory and theoretical physics was pioneered in [26, 27], linking the mentioned structures with Ulam games and  $AF - C^*$ -algebras, respectively. We will pay special attention to the study of the Łukasiewicz  $t$ -norm, due to its relation with Ulam games and its possible applications to error-correction codes in the context of quantum computation.

Evidently,  $\neg$  can be expressed as a polynomial in the generator system  $D_1(x)$ ; whence by Theorem 3.3, it is representable as a polynomial quantum operation. A possible representation can be the following:  $NOT(\rho) = \sigma_x \rho \sigma_x^\dagger$ . In fact,  $p(NOT(\rho)) = 1 - p(\rho)$ .

Furthermore,  $\bullet$  can be represented by a polynomial in the generator system  $D_2(x, y)$ . According with the construction presented in Theorem 3.3, the following representation obtains. Let us consider the following matrices:

$$G_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad G_2 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad G_3 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{aligned}
 G_4 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & G_5 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & G_6 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 G_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} & G_8 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

One may verify that  $\sum_{i=1}^8 G_i(\tau \otimes \sigma)G_i^\dagger = \frac{1}{2}I \otimes \rho_{p(\tau)p(\sigma)}$  where  $\sigma, \tau \in \mathcal{D}(\mathbb{C}^2)$ . Thus, by Kraus representation Theorem [24],  $\sum_{i=1}^8 G_i(\tau \otimes \sigma)G_i^\dagger$  is a quantum operation, and  $p(\sum_{i=1}^8 G_i(\tau \otimes \sigma)G_i^\dagger) = p(\tau) \bullet p(\sigma)$ . That quantum operation represents the well known quantum gate *IAND* modulo a tensor power [5, 29]. As regards the Łukasiewicz conorm  $\oplus$ , it can be seen that it is not a polynomial, see Figure 1.

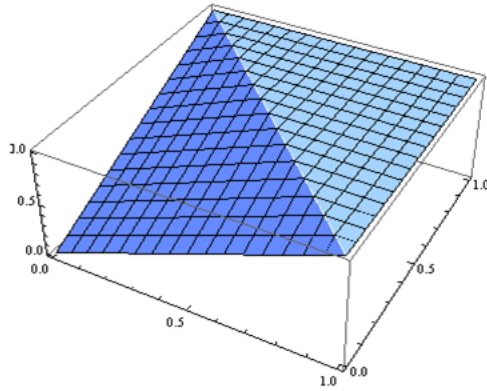


FIG. 1: *The Łukasiewicz conorm*

Therefore, our idea is to obtain a polynomial  $P(x, y)$  in some generator system  $D_k(x, y)$ , such that  $P(x, y)$  can approximate the Łukasiewicz sum.



By using numerical methods, we get the following approximating polynomial of  $\oplus$  in  $[0, 1]$ :

$$P(x, y) = \frac{5}{12}x(1-x) + \frac{5}{12}y(1-x) + \frac{5}{12}x(1-y) + \frac{5}{12}y(1-y) + \frac{1}{2}x + \frac{1}{2}y,$$

whose graph is depicted in Figure 2.

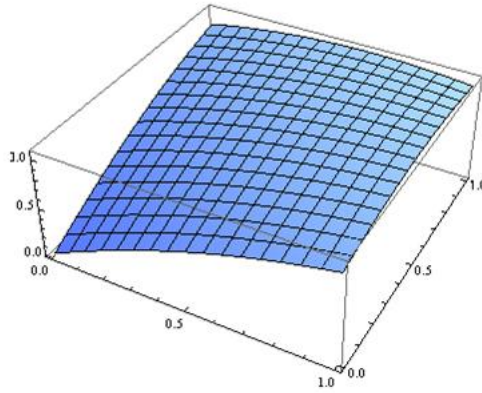


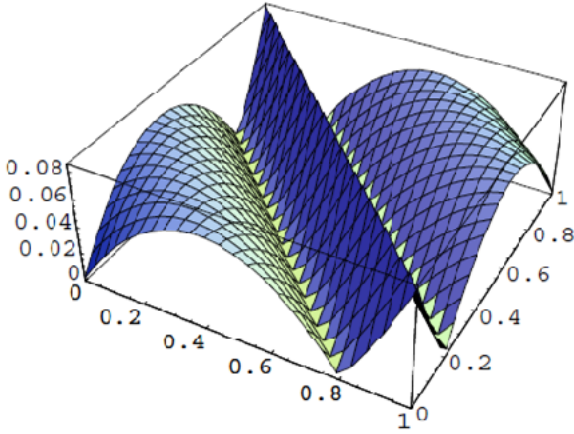
FIG. 2:  $P(x, y)$

Let us remark that  $0 \leq P(x, y) \leq x \oplus y$ . Then,  $e = \max_{[0,1]} \{(x \oplus y) - P(x, y)\} \leq 0.08$ , as Figure 3 shows. Furthermore, one readily realizes that  $P(x, y)$  is a polynomial obtained from the generator system  $D_2(x, y)$ , that also satisfies the hypothesis of Theorem 3.3. Thus,  $P(x, y)$  is representable as a polynomial quantum operation  $\mathcal{P}_\oplus$ , where

$$p(\mathcal{P}_\oplus(\tau \otimes \sigma)) = (p(\tau) \oplus p(\sigma)) \pm 0.08.$$

## 5. Conclusions and open problems

In virtue of the results in Section 4, it turns out that the approximation obtained in the case of PMV-algebras is definitely accurate. Further, in [11], authors

FIG. 3:  $(x \oplus y) - P(x, y)$ 

show a convergence theorem that allows to achieve every degree of accuracy; the price to pay is the increasing of the degree of the approximating polynomial. In our opinion, that is an interesting achievement, since it provides a (quantum) computational motivation for the investigation of algebraic structures equipped with the Łukasiewicz sum and, to a certain extent, it relates “classical” fuzzy logic to quantum computational logics.

Nonetheless, some general remarks on the whole construction are in order as well.

1. If one wants to apply our results to the models of (quantum) computing, efficiency is of central importance. Unfortunately, since the number of copies required in our construction corresponds to the degree of the approximating polynomial, it is impossible to generally specify the dimension of the Hilbert space required to achieve, given a certain  $\varepsilon$ , the approximating polynomial.
2. Since the work of Ekert and other scholars [22, 20, 19, 9], a direct study of estimations of linear and non-linear functionals of (quantum) states using quantum networks has been proposed. This approach has the advantage that it bypasses quantum tomography, providing more direct estimations

of both linear and non-linear functionals of a state. It could be of interest, in our opinion, to investigate if, and in what cases, our construction can be carried out by a quantum network.

## REFERENCES

- [1] D. AHARANOV, A. KITAEV and N. NISAN (1997), “Quantum circuits with mixed states”, *Proc. 13th Annual ACM Symp. on Theory of Computation, STOC*, pp. 20-30.
- [2] G. ALBER, T. BETH, M. HORODECKI, P. HORODECKI, R. HORODECKI, M. RÖTTLER, H. WEINFURTER, R. F. WERNER and A. ZEILINGER (2001), *Quantum information: an introduction to basic theoretical concepts and experiments*, Springer, Dordrecht.
- [3] S. BUGAJSKI, K. E. HELLWIG and W. STULPE (1998), “On fuzzy random variables and statistical maps”, *Rep. Math. Phys.*, 41, pp. 1-11.
- [4] G. BIRKHOFF and J. VON NEUMANN (1936), “The logic of quantum mechanics”, *Ann. Math.*, 37, pp. 823-843.
- [5] G. CATTANEO, M. L. DALLA CHIARA, R. GIUNTINI and R. LEPORINI (2004), “Quantum computational structures”, *Math. Slov.*, 54, pp. 87-108.
- [6] M. L. DALLA CHIARA, R. GIUNTINI and R. GREECHIE (2004), *Reasoning in Quantum Theory, Sharp and Unsharp Quantum Logics*, Kluwer, Dordrecht-Boston-London.
- [7] M. L. DALLA CHIARA, R. GIUNTINI and R. LEPORINI (2005), “Logics from Quantum Computation”, *Int. J. Quant. Inf.*, 3, 2, pp. 293-337.
- [8] G. DOMENECH and H. FREYTES, “Fuzzy propositional logic associated with quantum computational gates” (2006), *Int. J. Theor. Phys.*, 34, pp. 228-261.
- [9] A. EKERT, C. M. ALVES, D. K. L. OI, P. HORODECKI, M. HORODECKI and L. C. KWEK (2002), “Direct estimations of linear and non-linear functionals of a quantum state”, *Phys. Rev. Lett.*, 88, 21, pp. 217901-217905.

- [10] F. ESTEVA, L. GODO and F. MONTAGNA (2001), “The  $L\Pi$  and  $L\Pi_{\frac{1}{2}}$ : two complete fuzzy systems joining Łukasiewicz and Product Logic”, *Arch. Math. Log.*, 40, pp. 39-67.
- [11] H. FREYTES, G. SERGIOLI and A. ARICÒ (2010), “Representing continuous T-norms in quantum computation with mixed states”, *Journal of Physics A, Math. Theor.*, 43, <http://iopscience.iop.org/1751-8121/43/46/465306/>. Forthcoming.
- [12] R. GIUNTINI (1996), “Quantum MV-algebras”, *St. Log.*, 56, pp. 393-417.
- [13] R. GIUNTINI and H. GREULING (1989), “Toward a formal language for unsharp properties”, *Found. Phys.*, 19, pp. 931-945.
- [14] R. GIUNTINI, A. LEDDA and F. PAOLI (2007), “Expanding quasi-MV algebras by a quantum operator”, *St. Log.*, 87, 1, pp. 99-128.
- [15] R. GIUNTINI, A. LEDDA, G. SERGIOLI and F. PAOLI, “Some generalizations of fuzzy structures in quantum computational logic”, *International Journal of General Systems*, forthcoming.
- [16] S. GUDDER (1996), “Examples, problems, and results in effect algebras”, *Int J. Theor. Phys.*, 35, pp. 2365-2376.
- [17] S. GUDDER (1998), “Fuzzy probability theory”, *Dem. Math.*, 31, pp. 235-254.
- [18] S. GUDDER (2003), “Quantum computational logic”, *Int. J. Theor. Phys.*, 42, pp. 39-47.
- [19] P. HORODECKI (2003), “Measuring quantum entanglement without prior state reconstruction”, *Phys. Rev. Lett.*, 90, 16, pp. 167901-167905.
- [20] P. HORODECKI and A. EKERT (2003), “Direct detection of quantum entanglement”, *Quant. Computers and Computing*, 4, 2, pp. 111-120.
- [21] G. KALMBACH (1983), *Orthomodular lattices*, Academic Press, London and New York.
- [22] KEYL M. and WERNER R. F. (2001), “Estimating the spectrum of a density operator”, *Physical Review A*, 64, 5, pp. 052311-052316.

- [23] F. KÔPKA and F. CHOVANEC (1994), “D-posets”, *Math. Slov.*, 44, pp. 21-34.
- [24] K. KRAUS (1983), *States, effects and operations*, Springer-Verlag, Berlin.
- [25] A. LEDDA, M. KONIG, F. PAOLI and R. GIUNTINI (2006), “MV algebras and quantum computation”, *St. Log.*, 82, 2, pp. 245-270.
- [26] D. MUNDICI (1986), “Interpretation of AF  $C^*$ -algebras in the Łukasiewicz sentential calculus”, *Funct. Anal.*, 65, pp. 15-63.
- [27] D. MUNDICI (1993), “Ulam games, Łukasiewicz logic and AF  $C^*$ -algebras”, *Fundam. Inform.*, 18, pp. 151-161.
- [28] D. MUNDICI and B. RIEČÁN (2002), “Probability on MV-algebras”, in E. PAP (ed.) *Handbook of Measure Theory*, North Holland, Amsterdam, pp. 869-909.
- [29] M. A. NIELSEN and I. L. CHUANG (2000), *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge.



# Panorama e prospettive dell'approccio dinamico in scienza cognitiva

Marco Giunti

Alophis – Applied Logic, Philosophy and History of Science  
Dipartimento di Scienze Pedagogiche e Filosofiche, Università di Cagliari  
e-mail: giunti@unica.it

1. Introduzione: genesi dell'approccio dinamico
2. Diverse tendenze all'interno dell'approccio dinamico
3. L'unificazione mancata di Van Gelder
4. Dinamicismo vs computazionalismo puro: approcci dinamico, simbolico e connessionista
5. Il simulazionismo come dogma metodologico della scienza cognitiva
6. L'alternativa dinamica al simulazionismo: la metodologia Galileiana
7. La prospettiva di una scienza cognitiva programmaticamente Galileiana

**Appendice.** Sistemi dinamici, fenomeni e modelli Galileiani

SOMMARIO. In questo articolo ripercorreremo prima di tutto alcune tappe significative della genesi e del successivo sviluppo dell'approccio dinamico alla scienza cognitiva, a partire dall'inizio degli anni novanta. In particolare, cercheremo di mettere in luce come, fin dall'inizio, questo approccio abbia in realtà incluso diverse tendenze, soltanto parzialmente unificate dall'uso di modelli, metodi e concetti che, in senso lato, possono ricondursi alla teoria matematica dei sistemi dinamici. Cercheremo poi di chiarire il rapporto fra l'approccio dinamico e i due approcci alternativi più importanti, simbolico e connessionista. La tesi che sosterremo è che gran parte delle tendenze che ricadono sotto l'etichetta di approccio dinamico si differenzia sia dall'approccio simbolico che da quello connessionista da due punti di vista: (i) per il riconoscimento dell'*analisi dinamica dei modelli* come condizione necessaria delle spiegazioni di molti fenomeni cognitivi e (ii) per il conseguente abbandono del computa-

zionalismo come quadro di riferimento sufficiente per tali spiegazioni. Tuttavia, metteremo anche in luce un aspetto di sostanziale *continuità metodologica* fra i tre approcci, ovvero l'assunzione del *simulazionismo* come metodologia fondamentale. Nella parte finale del saggio verrà analizzata criticamente proprio questa assunzione, mettendo in luce come tale scelta metodologica implichi una sostanziale debolezza epistemologica della scienza cognitiva. Concluderemo infine l'articolo sostenendo che l'assunzione programmatica di una metodologia alternativa, comunemente utilizzata nella modellizzazione dinamica di molti fenomeni (fisici, biologici, economici ecc.), può aprire per la scienza cognitiva la possibilità di un superamento di tale debolezza epistemologica, se non in generale, almeno per quanto riguarda la descrizione e spiegazione di *alcuni tipi* di fenomeni cognitivi.

PAROLE CHIAVE: modelli dinamici; modello Galileiano; dinamicismo; computazionalismo; simulazionismo.

## 1. Introduzione: genesi dell'approccio dinamico

Le origini dell'approccio dinamico in scienza cognitiva possono essere ricondotte ad almeno cinque diverse linee di ricerca che risalgono fino alla metà del secolo scorso, ma che soltanto all'inizio degli anni novanta hanno cominciato a trovare una parziale unificazione metodologica e concettuale.

L'evento che, molto verosimilmente, può essere identificato con la vera e propria nascita di ciò che oggi è comunemente noto come "approccio dinamico" è la Indiana University Conference on Dynamic Representation in Cognition, che si svolse a Bloomington dal 14 al 17 novembre 1991 per iniziativa di un gruppo di filosofi, linguisti e psicologi dell'Indiana University, fra i quali Tim Van Gelder, Robert Port, Esther Thelen e James Townsend.<sup>1</sup>

Come sottolineato da Van Gelder e Port (1995), le radici più remote dell'approccio dinamico possono rintracciarsi nella *cibernetica* della fine degli anni quaranta e degli anni cinquanta. In effetti, "l'idea che la dinamica potesse fornire un quadro generale per una scienza unitaria della cognizione fu la

---

<sup>1</sup> In tutto questo saggio, quando il nome di un autore è citato senza uno specifico riferimento bibliografico, si intende che lavori esemplificativi delle ricerche di tale autore possono essere agevolmente reperiti nei seguenti volumi: Smith e Thelen (1993), Thelen e Smith (1994), Van Gelder e Port (1995), Kelso (1995), Tschacher e Dauwalder (2003).



base di uno dei più importanti libri di quel periodo, *Design for a Brain* (Ashby 1952)” (Van Gelder e Port 1995, p. 37).

Una seconda fonte che ha largamente contribuito alla disseminazione di idee dinamiche nel campo della ricerca cognitiva è la *teoria delle reti neurali*, i cui primi sviluppi risalgono agli anni cinquanta-sessanta. Successivamente, risultò particolarmente importante “il lavoro del gruppo di Stephen Grossberg che, in un contesto connessionista, applicò idee dinamiche a un ampio spettro di diverse funzioni cognitive” (*ibidem*). A questo riguardo, Van Gelder e Port hanno espressamente riconosciuto:

siccome le reti connessioniste sono sistemi dinamici, fu inevitabile che strumenti dinamici divenissero importanti per la comprensione del loro comportamento e, quindi, della natura delle funzioni cognitive. La recente rapida affermazione dell'approccio dinamico è perciò dovuta, in larga misura, a questa rinascita del connessionismo e al suo sviluppo in una direzione dinamica (*ibidem*).

Una terza fonte è costituita da una linea di ricerca che ha cominciato ad applicare allo studio del comportamento di sistemi biologici, e in particolare della coordinazione motoria, concetti e principi dinamici originariamente sviluppati nell'ambito della fisica. Tali studi rientrano a tutti gli effetti nel campo della scienza cognitiva in quanto comprendono anche varie abilità motorie e percettive degli esseri umani. Fra gli esponenti più noti di questa linea di ricerca possiamo ricordare Michael T. Turvey, Peter N. Kugler e J. A. Scott Kelso.

Una quarta fonte deriva direttamente da alcuni sviluppi della teoria matematica dei sistemi dinamici e, in particolare, dalle applicazioni della teoria delle biforcazioni o della teoria delle catastrofi di René Thom allo studio del linguaggio e, più in generale, delle strutture cognitive. Uno dei principali esponenti di questa tradizione di ricerca è Jean Petitot.

Infine, varie ricerche di tipo dinamico si sono sviluppate anche a partire dall'approccio ecologico di James J. Gibson alla psicologia della percezione. Secondo Gibson, il fattore chiave per un'adeguata spiegazione della percezione è la struttura dell'informazione già presente nello stimolo, e non tanto l'elaborazione di esso da parte di meccanismi interni. Queste intuizioni di Gibson sono state riprese e sviluppate utilizzando metodi dinamici, che permettono di descrivere formalmente la struttura dell'informazione sottostante a vari processi percettivi e motori. Fra gli esponenti di questa tendenza possiamo per esempio ricordare Geoffrey P. Bingham.

## 2. Diverse tendenze all'interno dell'approccio dinamico

Come abbiamo visto nel paragrafo precedente, l'approccio dinamico si costituì all'inizio degli anni novanta alla confluenza di diverse tradizioni di ricerca sulla cognizione, le quali utilizzavano modelli, metodi e concetti in senso lato riconducibili alla teoria matematica dei sistemi dinamici. Dando ora uno sguardo di insieme allo sviluppo successivo, dagli anni novanta fino a oggi, possiamo distinguere in esso almeno quattro distinte tendenze che, in modi diversi e interrelati, si ricollegano alle cinque tradizioni del periodo originario.

La prima tendenza può essere convenientemente indicata con l'etichetta di *connessionismo dinamico*. Secondo questa posizione i modelli che meglio descrivono i fenomeni cognitivi sono le reti neurali, ma il comportamento di tali reti non può essere compreso se non analizzando la loro *dinamica*, anche attraverso adeguati strumenti matematici. Fra l'altro, ciò significa che i processi cognitivi non possono essere visti, semplicemente, come *computazioni*. Al contrario, il concetto centrale nell'analisi dei modelli connessionisti e, conseguentemente, nella spiegazione dei fenomeni cognitivi da essi descritti, diviene quello di *evoluzione dinamica*. Ciò è tanto più vero quanto più le reti utilizzate si discostano dalla semplice architettura di un *feed-forward network*. Infatti, sebbene tutte le reti neurali siano sistemi dinamici, il funzionamento di un *feed-forward network* si presta naturalmente anche a un'analisi in termini computazionali; questo tipo di analisi, però, diviene molto meno informativo già per semplici *recurrent network* a tempo discreto, per non parlare poi delle reti interattive completamente continue (cioè, continue sia riguardo al tempo che ai livelli di attivazione delle singole unità). "On the Proper Treatment of Connectionism" (Smolensky 1988) può essere considerato una sorta di manifesto del connessionismo dinamico, ma in questa tendenza rientrano a pieno titolo anche altri autorevoli esponenti dell'approccio connessionista, fra i quali possiamo per esempio ricordare Grossberg, Elman e Pollack.

La seconda tendenza, che possiamo denominare *modellizzazione dinamica a poche dimensioni*, utilizza come modelli vari tipi di sistemi dinamici (continui, semicontinui, discreti, deterministici, probabilistici) caratterizzati però da alcuni tratti comuni. In primo luogo, gran parte di questi modelli sono specificati o mediante sistemi di equazioni differenziali ordinarie o mediante sistemi di equazioni alle differenze finite, a seconda che l'evoluzione temporale sia supposta continua o discreta.

A prima vista, anche le reti neurali sembrerebbero quindi rientrare in questo tipo di modello, perché la variazione del livello di attivazione di una *generica* unità della rete è anch'essa specificata da un'equazione (differenziale o al-

le differenze finite) e, conseguentemente, la dinamica di *tutta* la rete è descritta da un sistema di equazioni (i) tutte della stessa forma e (ii) in numero pari al numero delle unità della rete. Tuttavia, nei modelli a poche dimensioni la singola equazione *non rappresenta* la variazione del livello di attivazione di una generica unità di una rete e quindi le condizioni (i) e (ii) non sono di solito rispettate. Inoltre, proprio a causa di (ii), lo spazio degli stati di una rete neurale ha naturalmente un numero molto elevato di dimensioni, mentre il numero delle dimensioni in questo tipo di modelli è spesso molto basso, nell'ordine di poche unità, o perfino di una sola.

Secondo, ciascuna dimensione (o *componente*) dello spazio degli stati ha tipicamente una specifica *interpretazione*, che la fa corrispondere a una ben precisa proprietà, o *grandezza*, del fenomeno cognitivo o, più in generale, del sistema reale (naturale o artificiale) descritto dal modello. Non in tutti i modelli a poche dimensioni le grandezze corrispondenti alle componenti sono misurabili, ma in alcuni modelli (si veda per esempio Van Geert 1995) lo sono certamente – se non tutte, almeno alcune. Se alcune grandezze sono misurabili, la corrispondenza fra il modello e il fenomeno *non è una simulazione*, più o meno fedele, di alcuni aspetti (quantitativi o qualitativi) del fenomeno stesso da parte di aspetti simili del modello stesso, ma è invece una puntuale corrispondenza quantitativa fra le serie temporali misurate di una o più grandezze e le serie temporali previste dal modello.

Ciò che è importante notare qui non è tanto il carattere quantitativo della corrispondenza (anche le simulazioni sono spesso in grado di fornire ottime corrispondenze quantitative con i dati sperimentali) ma, piuttosto, che questa corrispondenza quantitativa si situa al livello delle *componenti di base* o, *costitutive*, della dinamica del modello stesso, mediante le *grandezze* che interpretano ciascuna dimensione dello spazio degli stati. Si contrasti questo tipo di relazione con il modo in cui un modello connessionista simula un fenomeno cognitivo. Anche supposto che la simulazione sia molto accurata, non c'è alcuna corrispondenza fra le componenti di base del modello connessionista e il fenomeno stesso. Infatti, ciascuna di tali componenti rappresenta il livello di attivazione di una singola unità (neurone artificiale) del modello ma, ovviamente, tale livello di attivazione non ha nessuna interpretazione sul fenomeno descritto. Piuttosto, l'interpretazione del modello avviene a un livello *più elevato* o *superficiale* della sua dinamica, per esempio a quello di un particolare attrattore, che magari rappresenta il raggiungimento di un certo livello di prestazione in un compito assegnato (come il riconoscimento o la classificazione di un dato input).

Come detto, non in tutti i modelli a poche dimensioni le grandezze che interpretano le componenti sono misurabili. A seconda di quante lo siano e dei

metodi più o meno diretti di misurazione, tali modelli risultano più o meno nettamente distinguibili dai *modelli di simulazione* tipici sia dell'approccio connessionista che di quello simbolico. A titolo di esempio, riportiamo una nostra valutazione (in effetti molto grossolana e aperta a revisioni) di quanto alcuni gruppi di modelli a poche dimensioni siano più o meno nettamente distinguibili dai modelli di simulazione. Mettendo per primi i modelli più nettamente distinguibili, otteniamo il seguente ordinamento: (1) i modelli della crescita cognitiva di Van Geert; (2) vari modelli relativi alla coordinazione, al controllo motorio e alle abilità percettive elaborati da Turvey, Carello, Kelso, Saltzman e altri; (3) i modelli dei processi di decisione (*Decision Field Theory*) di Townsend e Busemeyer; (4) i modelli degli agenti autonomi di Beer.

La terza tendenza, che possiamo denominare *descrizione dinamica di dati senza modelli*, è quella esemplificata da una serie di studi di Esther Thelen, Linda Smith e altri (Smith e Thelen 1993; Thelen e Smith 1994) sullo sviluppo cognitivo e motorio. Spesso, nello studio di fenomeni cognitivi piuttosto complessi, i dati a disposizione non permettono la costruzione di modelli adeguati. Tuttavia, un uso attento di concetti e metodi propri della teoria matematica dei sistemi dinamici può permettere di estrarre importanti aspetti qualitativi (ma a volte anche quantitativi) dai dati disponibili. Per esempio, negli studi condotti da Esther Thelen sullo sviluppo della capacità di raggiungere e afferrare oggetti, i movimenti adeguati sono concettualizzati come particolari attrattori che, a un certo stadio di sviluppo, caratterizzano lo spazio dei possibili movimenti. Questi attrattori emergono via via che la dinamica si modifica sotto l'influenza di parametri quali la pratica delle azioni stesse e la crescita corporea.

Possiamo infine denominare la quarta tendenza *interpretazione cognitiva di strutture dinamiche*. Come nella tendenza precedente, anche in questa non si tratta di costruire *modelli* dinamici, ma di utilizzare, più o meno direttamente, la *teoria matematica* dei sistemi dinamici per la descrizione o spiegazione di particolari aspetti della cognizione. Tuttavia, mentre nella tendenza precedente, l'utilizzo della teoria è strettamente vincolato dai dati disponibili – e quindi si può parlare di una direzione della ricerca che va dai dati alla teoria –, in questa tendenza la direzione è rovesciata, in quanto si parte direttamente dalla teoria matematica, che permette di costruire una serie di strutture dinamiche, e si dà poi a queste strutture una particolare interpretazione di tipo cognitivo, anche in mancanza di dati empirici che supportino più o meno direttamente tale interpretazione. Questa tendenza è la più speculativa delle quattro prese qui in considerazione ed è esemplificata dalle ricerche di Jean Petitot sulla costruzione di strutture linguistiche (sia sintattiche che semantiche) mediante l'approccio morfodinamico di René Thom.

### 3. L'unificazione mancata di Van Gelder

Abbiamo visto nel paragrafo precedente come l'approccio dinamico sia in realtà composto da diverse tendenze, soltanto parzialmente unificate dall'uso di modelli, metodi e concetti che, in senso lato, possono ricondursi alla teoria matematica dei sistemi dinamici. Tuttavia, negli anni novanta, Tim Van Gelder tentò di unificare queste diverse tendenze avanzando la proposta dell'*ipotesi dinamica* (*Dynamical Hypothesis – DH*). Nelle sue intenzioni, questa ipotesi avrebbe dovuto assumere lo stesso ruolo fondante che l'*ipotesi del sistema simbolico fisico* di Newell e Simon (Newell e Simon 1972, 1976; Newell 1980) aveva avuto per l'approccio simbolico. L'ipotesi dinamica fu esposta da Van Gelder nel 1998 in un noto "articolo bersaglio" (*target article*) di *Behavioral and Brain Sciences*. Contro le aspettative dell'autore, però, la proposta non riuscì a suscitare il consenso sperato, e anche alcuni degli stessi sostenitori di posizioni dinamiche, specialmente fra i connessionisti, non risparmiarono a Van Gelder critiche piuttosto corrosive e tutto sommato ingenerose. Senza addentrarci troppo in una ricostruzione dei vari motivi che scatenarono questa sorta di fuoco amico contro l'ipotesi dinamica, ci preme però rilevare come essa in effetti presentasse diversi punti deboli, che possiamo così sintetizzare.

Con l'ipotesi dinamica Van Gelder tentò di definire un particolare *tipo di modello*, da lui denominato *quantitativo*, che avrebbe dovuto caratterizzare l'approccio dinamico e risultare ben distinto sia dai modelli simbolici dell'approccio classico che da quelli connessionisti non dinamici. Tuttavia, come messo in evidenza da diversi commentatori, fra i quali anche Randy Beer, la definizione di Van Gelder risulta troppo poco elaborata, ed è perciò sostanzialmente incapace di delineare un tipo di modello che rappresenti una chiara alternativa ai modelli simbolici o a quelli connessionisti. Inoltre, l'ipotesi dinamica non chiarisce sufficientemente quali siano i vantaggi che potrebbero derivare dall'utilizzo di un modello dinamico quantitativo invece di uno simbolico o connessionista.

Abbiamo visto nel paragrafo 2 che, sul piano metodologico del rapporto fra oggetto di indagine e costruito teorico, l'approccio dinamico oscilla fra almeno quattro diverse posizioni. Per i connessionisti dinamici, tale rapporto è quello tipico della modellizzazione con le reti neurali, ovvero, il modello permette di *simulare*, in modo più o meno fedele, alcuni aspetti ritenuti significativi del fenomeno indagato. Per i sostenitori dei modelli a poche dimensioni, invece, il rapporto di elezione è quello tipico della modellizzazione dinamica, ovvero, una corrispondenza fra serie temporali (misurate e previste) di valori

di grandezze, dove ciascuna di tali grandezze costituisce l'interpretazione di una delle componenti del modello. Per la descrizione dinamica di dati si tratta invece di far vedere come i dati a disposizione possano essere inseriti in quadri concettuali tratti dalla teoria matematica dei sistemi dinamici. Infine, per l'interpretazione di strutture dinamiche, il rapporto fra oggetto di indagine e costruito teorico non comporta una vera e propria considerazione di dati empirici, ma soltanto la possibilità di costruire plausibili e significative interpretazioni cognitive di strutture matematiche astratte.

Quale di queste quattro metodologie è quella preferita o, per così dire, "ufficiale" dell'approccio dinamico? O nessuna di queste è sufficiente e una nuova metodologia unificata, una specie di sintesi di tutte e quattro, dovrebbe essere adottata? O forse dovrebbe essere sviluppata una via intermedia? L'ipotesi dinamica di Van Gelder non dà alcuna risposta a queste domande, che però sono cruciali se l'approccio dinamico vuole trasformarsi in un programma di ricerca realmente unitario e non rimanere una somma piuttosto estrinseca di tendenze simili, ma sostanzialmente indipendenti.

#### **4. Dinamicismo vs computazionalismo puro: approcci dinamico, simbolico e connessionista**

Che cosa c'è di nuovo e *più caratteristico* nell'approccio dinamico? È possibile cristallizzarne un nucleo metodologico che possa fornirgli un'identità forte, paragonabile a quella dei due approcci principali della scienza cognitiva, simbolico e connessionista? Da quanto abbiamo visto nei paragrafi precedenti, se tale nucleo esiste, non può certamente aspirare a includere *tutti* i diversi aspetti o tendenze dell'approccio dinamico; il tentativo di Van Gelder è infatti fallito proprio perché rivolto a conciliare *troppi* elementi eterogenei, peraltro senza indicare una vera e propria base metodologica comune.

Lasciando per il momento aperta la risposta a questa domanda, è però possibile enucleare un'ipotesi metodologica molto generale, ma non per questo meno significativa, che è certamente sottoscrivibile (ed è di fatto, seppur solo implicitamente, sottoscritta) da gran parte delle diverse componenti dell'approccio dinamico. Questa ipotesi, che denominiamo *dinamicismo*, può essere così formulata:

[D] Molti fenomeni cognitivi non sono adeguatamente spiegabili se non si analizza la *dinamica* di opportuni modelli di tali fenomeni.

È abbastanza evidente che il dinamicismo sia accettabile dai connessionisti dinamici e anche dai sostenitori della modellizzazione a poche dimensioni. Infatti, da un lato, il connessionismo dinamico [CD] non è nient'altro che una specificazione di [D]; tale specificazione consiste semplicemente nel sostituire “modelli” con “modelli connessionisti”:

[CD] Molti fenomeni cognitivi non sono adeguatamente spiegabili se non si analizza la *dinamica* di opportuni modelli *connessionisti* di tali fenomeni.

Dall'altro, per quanto riguarda la modellizzazione a poche dimensioni, l'analisi dinamica è certamente un ingrediente necessario per gran parte delle spiegazioni che si basano su questo tipo di modelli.

Per quanto riguarda invece la descrizione dinamica di dati e l'interpretazione di strutture dinamiche, non è altrettanto evidente che per queste tendenze [D] sia accettabile. O, per lo meno, [D] non è certamente esemplificato nella loro pratica di ricerca perché, come abbiamo visto, essa non fa ricorso a modelli. Tuttavia, non possiamo escludere che anche molti dei sostenitori di queste due tendenze non sottoscrivano [D], se non come principio metodologico concreto, almeno come ideale normativo.

È importante notare che il dinamicismo è inconsistente con la seguente formulazione del computazionalismo, che denominiamo *computazionalismo puro* [CP]:

[CP] Ogni fenomeno cognitivo è adeguatamente spiegabile se si analizzano *soltanto* le *computazioni* che si verificano in opportuni modelli di tale fenomeno.

D'altra parte, il computazionalismo puro è certamente implicato dall'ipotesi del sistema simbolico fisico o, più in generale, da una qualsiasi formulazione standard dei capisaldi teorici dell'approccio simbolico. Per quanto riguarda invece l'approccio connessionista, bisogna distinguere fra connessionismo *computazionale* e *dinamico*. L'interpretazione computazionale del connessionismo è certamente molto diffusa e non è azzardato dire che essa sottenda molte delle formulazioni più influenti di tali posizioni, per esempio quelle ormai classiche di Rumelhart e McClelland (1986). Il connessionismo computazionale [CC] può essere espresso nel seguente modo:

[CC] Ogni fenomeno cognitivo è adeguatamente spiegabile se si analizzano *soltanto* le *computazioni* parallele e distribuite che si verificano in opportuni modelli connessionisti di tale fenomeno.

È quindi chiaro che [CP] segue logicamente da [CC], essendo il secondo una specificazione del primo. Possiamo perciò concludere che il dinamicismo [D], essendo inconsistente con il computazionalismo puro [CP], è sufficiente a distinguere gran parte dell'approccio dinamico sia dall'approccio simbolico che da quello connessionista computazionale.

Si noti infine che, sebbene il dinamicismo sia inconsistente con il computazionalismo puro, è però ovviamente compatibile con forme più deboli, o impure, di computazionalismo. Per esempio, quella che si ottiene da [CP] sostituendo "ogni" con "qualche", o un'altra, ancora più debole, che afferma che l'analisi computazionale dei modelli è necessaria per la spiegazione di alcuni fenomeni cognitivi.

## 5. Il simulazionismo come dogma metodologico della scienza cognitiva

Abbiamo visto che gran parte dell'approccio dinamico risulta distinto dall'approccio simbolico e dal connessionismo computazionale in quanto accetta il dinamicismo [D] e, conseguentemente, rifiuta il computazionalismo puro [CP]; da un altro punto di vista, però, c'è una forte omogeneità metodologica nella pratica della modellizzazione di tutti e tre questi approcci. Tale omogeneità scaturisce dal fatto che essi condividono tacitamente un'ipotesi metodologica molto generale; questa ipotesi, che denominiamo *simulazionismo*, può essere così formulata:

[S] Ogni fenomeno cognitivo è adeguatamente spiegabile sulla base di opportuni *modelli di simulazione* del fenomeno stesso.

Notiamo prima di tutto che, a differenza del dinamicismo [D] o del computazionalismo puro [CP], il simulazionismo non prescrive un tipo di *analisi dei modelli* a cui le spiegazioni dei fenomeni cognitivi dovrebbero conformarsi ma, piuttosto, un tipo di *modello* su cui tali analisi dovrebbero poi essere condotte. Essendo il tipo di modello proposto molto generale (*modello di simulazione* – si veda la definizione sotto), esso include sia modelli simbolici che connessionisti e anche, almeno in linea di principio, modelli dinamici a poche dimensioni. Proprio per questa ragione il simulazionismo, anche se originariamente è stato uno dei principi metodologici fondanti dell'approccio simbolico, è oggi un principio trasversale, comune a tutti i più importanti approcci della scienza cognitiva. Spesso esso è tanto più radicato quanto più la sua accettazione non è esplicita e cosciente, ma tacita e scontata, tanto che a volte anche modelli che



di fatto *non sono* modelli di simulazione vengono però presentati dai loro stessi autori come se lo fossero (un esempio abbastanza chiaro dell'operare di questa "falsa coscienza metodologica" può essere trovato in Van Geert 1995).

Per comprendere più in dettaglio il significato di [S] è opportuno tenere sempre presente il seguente schema della *metodologia simulazionista*. Dato un certo fenomeno cognitivo  $H$ , esso determina un ambiente con un relativo compito  $C_H$  (*task-environment*  $C_H$ ); si tratta quindi di (i) elaborare un modello  $M$  che, una volta implementato su un normale calcolatore digitale, permetta al calcolatore di eseguire il compito  $C_H$  in modo rilevantemente simile a come lo eseguirebbe un essere umano o, più in generale, a come lo eseguirebbe un reale agente cognitivo  $A$ ; (ii) formulare l'ipotesi che alcuni dei processi caratteristici del modello  $M$  corrispondano in misura adeguata ai reali processi cognitivi che stanno alla base della prestazione (performance) dell'agente  $A$  relativamente al compito  $C_H$ ; (iii) ricercare appropriati dati sperimentali che permettano di confermare o smentire tale ipotesi; (iv) accettare o rigettare l'ipotesi a seconda dei dati conseguiti.

Infine, si tenga presente che un *modello di simulazione di un fenomeno cognitivo*  $H$  può essere definito come un qualunque modello  $M$  che soddisfi le ipotesi (i) e (ii). Conseguentemente, l'esito positivo/negativo del metodo della simulazione consiste nell'accettare/rifutare che  $M$  sia un modello di simulazione di  $H$ .

## **6. L'alternativa dinamica al simulazionismo: la metodologia Galileiana**

I modelli di simulazione costituiscono il paradigma dominante nella scienza cognitiva. Tuttavia, a causa della loro costituzione, i modelli di simulazione hanno forti limitazioni epistemologiche, sia al livello della *descrizione* dei dati relativi a uno specifico fenomeno, che a quello della loro *spiegazione*.

Il limite descrittivo riguarda la corrispondenza fra dati simulati e dati reali, che non è diretta e intrinseca al modello ma, al più, indiretta ed estrinseca. Infatti, un modello di simulazione non incorpora proprietà misurabili (grandezze) del fenomeno reale fra le sue componenti di base; al contrario la corrispondenza è ottenuta a un livello più superficiale e globale, ovvero fra i dati simulati relativi ad alcuni processi tipici del modello e i dati reali relativi al fenomeno in questione.

Il limite esplicativo riguarda la qualità delle spiegazioni basate su modelli di simulazione. Tipicamente, esse non sono né compiute né realistiche, ma so-

no invece spiegazioni in linea di principio che fanno spesso ricorso anche a entità o processi chiaramente fittizi. Questo secondo limite, come d'altra parte anche il primo, dipende dal fatto che le componenti di base di un modello di simulazione non corrispondono direttamente ad aspetti reali del fenomeno indagato.<sup>2</sup> Di conseguenza, qualsiasi spiegazione basata sull'analisi di un tale modello è irrimediabilmente destinata a introdurre una serie di elementi fittizi che non hanno alcun corrispettivo nel fenomeno reale.

Questi due limiti epistemologici tipici dei modelli di simulazione sono agevolmente superati da un tipo di modello dinamico (*modello Galileiano*) comunemente utilizzato nella modellizzazione di molti fenomeni naturali e sociali (fisici, biologici, economici ecc.). I modelli Galileiani, così chiamati (Giunti 1995, 1997) perché fu Galileo che per primo li utilizzò sistematicamente per render conto dell'oscillazione di un pendolo, della caduta libera di un grave, del moto su un piano inclinato e del lancio di un proiettile, sono sistemi dinamici (discreti o continui, reversibili o irreversibili) il cui spazio degli stati ha  $n$  ( $1 \leq n$ ) dimensioni o componenti. Il punto fondamentale è che ciascuna componente ha un'interpretazione precisa e definita, in quanto essa è identificata con l'insieme dei possibili valori di una grandezza del fenomeno reale che il modello descrive. Alcune di queste  $n$  grandezze (ma non necessariamente tutte) sono *misurabili* e, rispetto a queste, il modello risulta *empiricamente corretto*, nel senso che tutte le misurazioni di tali grandezze corrispondono (entro i limiti di precisione delle misurazioni stesse) ai valori previsti dal modello.<sup>3</sup>

È chiaro che, a causa della loro costituzione, i modelli Galileiani sono in grado di superare i limiti descrittivi ed esplicativi tipici dei modelli di simulazione. Infatti, in primo luogo, la descrizione dei dati è diretta e intrinseca, in quanto ciascuna componente del modello determina i valori di una specifica grandezza del fenomeno descritto. Secondo, le spiegazioni supportate da questo tipo

---

<sup>2</sup> Si potrebbe obiettare che, almeno per quanto riguarda i modelli simbolici, i processi computazionali e le strutture di dati che li caratterizzano non sono affatto arbitrari, ma hanno invece una chiara e puntuale interpretazione relativamente ai fenomeni cognitivi descritti da tali modelli. Possiamo concordare con questa affermazione, ma essa non dimostra affatto che le *componenti di base* di tali modelli corrispondano ad aspetti reali dei fenomeni indagati. Al contrario, siccome un qualsiasi modello simbolico non è nient'altro che il *sistema dinamico* implicitamente specificato da un programma, ciò significa proprio che l'interpretazione del modello, in quanto si limita soltanto a *parti del programma*, non ha niente a che vedere con le componenti di base del modello stesso, ovvero, con le componenti dello spazio degli stati di tale sistema.

<sup>3</sup> Una definizione più rigorosa del concetto di *modello Galileiano* è possibile, ma essa presuppone un chiarimento preliminare delle due nozioni di *sistema dinamico* e *fenomeno* (dinamico). Per questo si rimanda all'Appendice.

di modello sono compiute e realistiche; infatti, siccome ciascuna componente del modello corrisponde a una ben precisa grandezza del fenomeno reale, qualsiasi spiegazione basata sull'analisi di tale modello non può introdurre alcun elemento fittizio o arbitrario. Per queste ragioni, chiunque sia interessato al miglioramento dei risultati della scienza cognitiva, sia al livello descrittivo che a quello esplicativo, dovrebbe prendere in seria considerazione la prospettiva di costruire modelli Galileiani di fenomeni cognitivi (Giunti 1995, 1997, 2005).

## **7. La prospettiva di una scienza cognitiva programmaticamente Galileiana**

Abbiamo visto nel paragrafo precedente che la metodologia Galileiana è, almeno in linea di principio, in grado di fornire alla scienza cognitiva l'opportunità di superare la debolezza epistemologica insita nella metodologia simulazionista. La domanda che sorge spontanea è allora se, a oggi, questa sia soltanto una possibilità teorica o se non sia in qualche modo già supportata da una pratica di ricerca consolidata. La risposta a questa domanda deve prendere in considerazione diversi aspetti.

Come già notato nel paragrafo 2, la modellizzazione dinamica a poche dimensioni utilizza molto raramente veri e propri modelli di simulazione, e alcuni dei modelli proposti all'interno di questa tendenza sono certamente modelli Galileiani. Di fatto, la metodologia Galileiana è quella che meglio si accorda con il tipo di indagine portato avanti da questa tendenza, ma ciò non è stato ancora adeguatamente sottolineato o riconosciuto esplicitamente, probabilmente per due ordini di motivi.

Da un lato, come già notato, a causa della pervasività e persistenza del dogma simulazionista, anche alcuni modelli che non sono di simulazione, ma a tutti gli effetti Galileiani, si presentano invece travestiti in guisa simulazionista, rendendo in effetti difficile il riconoscimento della novità e del vero valore di queste ricerche per il miglioramento della scienza cognitiva.

Dall'altro, i modelli Galileiani, come al momento definiti, non riescono a includere *tutti* i tipi di modelli utilizzati nella modellizzazione a poche dimensioni. Per esempio, alcuni di questi modelli sono di tipo probabilistico e quindi non rientrano fra i modelli Galileiani, perché questi ultimi sono tutti sistemi dinamici deterministici. Dunque, per poter includere tutti o almeno gran parte dei modelli utilizzati da questa tendenza, il concetto di modello Galileiano dovrebbe essere generalizzato, in modo da sussumere almeno alcuni fra i tipi più diffusi di modelli probabilistici.

Tuttavia, anche limitandoci a modelli deterministici, i modelli Galileiani si sono già dimostrati utili per la descrizione accurata e la spiegazione di *alcuni tipi* di fenomeni cognitivi. Fra essi, possiamo ricordare i fenomeni riguardanti la *crescita cognitiva*, studiati da Van Geert, o i fenomeni riguardanti la *coordinazione motoria* studiati, fra gli altri, da Turvey e Kelso. Un altro tipo di fenomeno cognitivo che è molto probabilmente descrivibile in modo accurato mediante modelli di tipo Galileiano è quello dei fenomeni di *computazione umana*, ovvero, tutti quei fenomeni cognitivi che riguardano l'esecuzione di calcoli con carta e penna (Giunti 2009).<sup>4</sup>

## Appendice

### Sistemi dinamici, fenomeni e modelli Galileiani

#### *Sistemi dinamici*

Un sistema dinamico è un tipo di modello matematico che esprime formalmente l'idea di un sistema deterministico arbitrario, reversibile o irreversibile, con tempo o spazio degli stati discreto o continuo. Siano  $Z$  gli interi,  $Z^+$  gli interi non negativi,  $R$  i reali e  $R^+$  i reali non negativi; quella sotto è la definizione esatta di un sistema dinamico.

- [1] *DS* è un sistema dinamico sse *DS* è una coppia  $(M, (g^t)_{t \in T})$  tale che:
1.  $M$  è un insieme non vuoto;  $M$  rappresenta tutti i possibili stati del sistema ed esso è chiamato lo *spazio degli stati*;
  2.  $T = Z, Z^+, R$  o  $R^+$ ;  $T$  rappresenta il tempo del sistema ed è chiamato l'*insieme tempo*; ogni  $t \in T$  è detta una *durata* del sistema;
  3.  $(g^t)_{t \in T}$  è una famiglia di funzioni da  $M$  a  $M$ ; per ogni  $t \in T$ , la funzione  $g^t$  è detta la *transizione di stato di durata t*, o il *t-avanzamento*, del sistema;
  4. per ogni  $v, t \in T$ , per ogni  $x \in M$ ,  $g^0(x) = x$  e  $g^{v+t}(x) = g^v(g^t(x))$ .
- [2] Un *sistema dinamico discreto* è un sistema dinamico il cui spazio degli stati è finito o numerabile e il cui insieme tempo è  $Z$  o  $Z^+$ ; esempi di si-

---

<sup>4</sup> Per ampliare ulteriormente il panorama delle diverse tendenze dell'approccio dinamico, si vedano Ward (2002) e Schöner (2008). Per la prospettiva di un approccio Galileiano in scienza cognitiva e il rapporto con le posizioni computazionaliste si vedano Giunti (1992, 1996, 1998) e Giunti e Giuntini (2007). Per un'aggiornata sintesi dei diversi tipi di modellizzazione in scienza cognitiva si veda McClelland (2009).

stemi dinamici discreti sono le macchine di Turing e gli automi cellulari (con un numero finito di stati non quiescenti).

- [3] Un *sistema dinamico continuo* è un sistema dinamico che non è discreto; esempi di sistemi dinamici continui sono le iterazioni di funzioni su  $R$  e i sistemi specificati da equazioni differenziali ordinarie.

### *Fenomeni*

In generale, un *fenomeno*  $H$  può essere pensato come una coppia  $(F, B_F)$  composta di due elementi distinti. Il primo elemento  $F$  è una *descrizione funzionale* di (i) un tipo astratto di sistema reale  $AS_F$  e (ii) uno schema spazio-temporale generale  $CS_F$  delle sue interazioni causali; in particolare, la descrizione funzionale del sistema astratto  $AS_F$  specifica i suoi elementi strutturali (o parti funzionali), la loro organizzazione e le relazioni reciproche, mentre la descrizione funzionale dello schema causale  $CS_F$  specifica le condizioni *iniziali* dell'evoluzione di  $AS_F$ , le condizioni ulteriori durante tutta l'evoluzione successiva (condizioni *al contorno*) e, eventualmente, le condizioni sotto le quali l'evoluzione di  $AS_F$  termina (condizioni *finali*). Il secondo elemento  $B_F$  è l'insieme di tutti i sistemi concreti di tipo  $AS_F$  che soddisfano anche lo schema di interazioni causali  $CS_F$  o, il che è lo stesso,  $B_F$  è l'insieme di tutti i sistemi reali che soddisfano la descrizione funzionale  $F$ ;  $B_F$  è detto il *dominio di applicazione*<sup>5</sup> del fenomeno  $H$ .

Per esempio, sia  $H_{e\phi} = (F_{e\phi}, B_{Fe\phi})$  il fenomeno della caduta libera di un corpo di media grandezza in prossimità della terra, dove  $\phi \in [0, \psi]$  è un parametro reale non negativo il cui significato è spiegato sotto (d'ora in poi chiameremo  $H_{e\phi}$  *il fenomeno della caduta libera*). In questo caso, la descrizione funzionale  $F_{e\phi}$  è la seguente. Il tipo astratto di sistema reale  $AS_{Fe\phi}$  ha un solo elemento strutturale, ovvero, un corpo di media grandezza in prossimità della terra. Lo schema di interazioni causali  $CS_{Fe\phi}$  consiste nel (i) rilasciare il corpo a un istante arbitrario e con una velocità e posizione *puramente verticali* (relativamente alla superficie terrestre), in modo tale che il corpo colpisca la superficie terrestre a un qualche istante successivo e l'altezza massima raggiunta dal corpo non sia superiore a  $\phi$ ; (ii) durante tutto il moto del corpo, l'unica forza agente su di esso è il suo peso; (iii) il moto termina esattamente quando il corpo colpisce la super-

---

<sup>5</sup> Siccome la descrizione funzionale  $F$  di norma contiene un certo numero di idealizzazioni, nessun sistema reale o concreto  $RS$  soddisfa *esattamente*  $F$  ma, piuttosto, la soddisfa soltanto fino a un certo grado. Perciò, da un punto di vista formale, il dominio di applicazione  $B_F$  di un fenomeno  $H = (F, B_F)$  sarebbe meglio descritto come un insieme fuzzy.

ficie terrestre. Infine,  $B_{F_{e\phi}}$  è l'insieme di tutti i corpi reali di media grandezza in prossimità della terra che soddisfano il dato schema di interazioni causali. Un qualsiasi corpo appartenente a  $B_{F_{e\phi}}$  è detto un *grave (in caduta libera)*.

### Modelli Galileiani

Per ogni  $i$  ( $1 \leq i \leq n$ ), sia  $X_i$  un insieme non vuoto e  $DS = (M, (g^t)_{t \in T})$  sia un sistema dinamico il cui spazio degli stati  $M$  sia incluso nel prodotto cartesiano  $X_1 \times \dots \times X_n$ ; per ogni  $i$ , l'insieme  $C_i = \{x_i; \text{ per qualche } n\text{-upla } x \in M, x_i \text{ è l}'i\text{-esimo elemento di } x\}$  è detto *l' $i$ -esima componente di  $M$* . Un'interpretazione  $I_H$  di  $DS$  su un fenomeno  $H$  consiste nell'identificare ciascuna componente  $C_i$  con l'insieme di tutti i possibili valori di una grandezza  $M_i$  del fenomeno  $H$ , e l'insieme tempo  $T$  con l'insieme dei possibili valori del tempo  $T$  dello stesso  $H$ . Un'interpretazione  $I_H$  di  $DS$  su  $H$  è *empirica* se il tempo  $T$  e alcune delle grandezze  $M_i$  sono proprietà misurabili del fenomeno  $H$ . Una coppia  $(DS, I_H)$ , dove  $DS$  è un sistema dinamico con  $n$  componenti e  $I_H$  è un'interpretazione di  $DS$  su  $H$ , è detta un *modello del fenomeno  $H$* . Se l'interpretazione  $I_H$  è empirica, allora  $(DS, I_H)$  è un *modello empirico di  $H$* . Tale modello è detto *empiricamente corretto* se, per ogni grandezza misurabile  $M_i$ , tutte le misurazioni di  $M_i$  sono consistenti con i corrispondenti valori  $x_i$  determinati da  $DS$ . Un modello empiricamente corretto di  $H$  è anche chiamato un *modello Galileiano di  $H$*  (Giunti 1995; 1997, cap. 3). Infine, un *modello Galileiano* è un qualsiasi modello empiricamente corretto di qualche fenomeno.

A titolo di esempio, consideriamo il seguente sistema di due equazioni differenziali ordinarie  $\langle dy(t)/dt = \dot{y}(t), d\dot{y}(t)/dt = -g \rangle$ , dove  $g$  è una costante reale positiva. Le soluzioni di tali equazioni determinano univocamente il sistema dinamico  $DS_e = (Y \times \dot{Y}, (g^t)_{t \in T})$ , dove  $Y = \dot{Y} = T = R$  (i numeri reali) e, per ogni  $t, y, \dot{y} \in R$ ,  $g^t(y, \dot{y}) = (-gt^2/2 + \dot{y}t + y, -gt + \dot{y})$ . È immediato verificare che  $DS_e$  soddisfa la Definizione [1].

Consideriamo adesso nuovamente il fenomeno della caduta libera  $H_{e\phi}$  e interpretiamo la prima componente  $Y$  dello spazio degli stati di  $DS_e$  come l'insieme di tutti i possibili valori della *posizione verticale* di un grave arbitrario, la seconda componente  $\dot{Y}$  come l'insieme di tutti i possibili valori della *velocità verticale* del grave<sup>6</sup> e l'insieme tempo  $T$  di  $DS_e$  come l'insieme di tutti i possibili valori del *tempo fisico*. Sia  $I_{H_{e\phi}}$  questa interpretazione. Siccome tutte

<sup>6</sup> Per ogni grave  $a$ , se  $p_a$  è il punto dove  $a$  è inizialmente rilasciato, la posizione e la velocità verticali di  $a$  si intendono riferite a un asse con origine nel centro della terra che passa per  $p_a$ ; il verso positivo di tale asse è quello che va dal centro della terra al punto  $p_a$ .

e tre le grandezze considerate sono proprietà misurabili del fenomeno della caduta libera  $H_{e\phi}$ , in base alle definizioni precedenti,  $I_{H_{e\phi}}$  è un'interpretazione empirica di  $DS_e$  su  $H_{e\phi}$  e  $(DS_e, I_{H_{e\phi}}) = DS_{e\phi}$  è un modello empirico di  $H_{e\phi}$ .  $DS_{e\phi}$  è detto *il modello della caduta libera*. Se  $\phi$  è sufficientemente piccolo e per un valore appropriato della costante  $g$ , tale modello risulta anche empiricamente corretto<sup>7</sup> ed è quindi un esempio di modello Galileiano.

## RIFERIMENTI BIBLIOGRAFICI

- ASHBY, William Ross (1952): *Design for a Brain*, London: Chapman & Hall.
- GIUNTI, Marco (1992): *Computers, Dynamical Systems, Phenomena and the Mind*, Ph.D. dissertation, Bloomington (Ind.): Indiana University; pubblicata da University Microfilms Inc., Ann Arbor (Mich.): numero d'ordine UMI: 9301444.
- (1995): “Dynamical Models of Cognition”, in Port e Van Gelder (a cura di) (1995, pp. 549-571).
- (1996): “Beyond Computationalism”, in Garrison W. Cottrel (a cura di), *Proceedings of the 18<sup>th</sup> Annual Conference of the Cognitive Science Society* Mahwah (N.J.): L. Erlbaum Associates, pp. 71-75.
- (1997): *Computation, Dynamics, and Cognition*, New York: Oxford University Press.
- (1998): “Is Computationalism the Hard Core of Cognitive Science?”, in Vito M. Abrusci, Carlo Cellucci, Roberto Cordeschi e Vincenzo Fano (a cura di), *Prospettive della logica e della filosofia della scienza: Atti del convegno triennale della Società Italiana di Logica e Filosofia delle Scienze, Roma, 3-5 gennaio 1996*, Pisa: Edizioni ETS, pp. 255-267.
- (2005): “Dal simulazionismo al paradigma Galileiano”, in *Atti del XIX Congresso Nazionale dell'Associazione Italiana di Psicologia, sez. di Psicologia Sperimentale, Cagliari, 18-20 settembre 2005*, Cagliari: AIP.
- (2009): “Bidimensional Turing Machines as Galilean Models of Human Computation”, in Gianfranco Minati, Mario Abram e Eliano Pessa (a cura di), *Processes of Emergence of Systems and Systemic Properties: Towards a General Theory of Emergence*, Singapore: World Scientific, pp. 383-423.

---

<sup>7</sup> Ovviamente, se  $g = \text{accelerazione di gravità standard}$  ( $g = 9.80665 \text{ m/s}^2$ ), il modello della caduta libera  $DS_{e\phi} = (DS_e, I_{H_{e\phi}})$  risulta empiricamente corretto entro limiti di precisione sufficienti per molti scopi pratici, supposto che il parametro  $\phi$  sia sufficientemente piccolo.

- GIUNTI, Marco e GIUNTINI, Roberto (2007): “Macchine, calcolo e pensiero”, in Stefano Mancini (a cura di), *Sguardi sulla scienza dal giardino dei pensieri*, Milano: Mimesis, pp. 39-67.
- KELSO, J. A. Scott (1995): *Dynamic Patterns: The Self-Organization of Brain and Behavior*, Cambridge (Mass.): The MIT Press.
- MCCLELLAND, James L. (2009): “The Place of Modeling in Cognitive Science”, *Topics in Cognitive Science*, 1, pp. 11-38.
- NEWELL, Allen (1980): “Physical Symbol Systems”, *Cognitive Science*, 4, pp. 135-183.
- NEWELL, Allen e SIMON, Herbert (1972): *Human Problem Solving*, Englewood Cliffs (N.J.): Prentice Hall.
- (1976): “Computer Science as Empirical Enquiry: Symbols and Search”, *Communications of the Association for Computing Machinery*, 19, pp. 113-126.
- PORT, Robert F. e VAN GELDER, Timothy J. (a cura di) (1995): *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge (Mass.): The MIT Press.
- RUMELHART, David E. e MCCLELLAND, James L. (a cura di) (1986): *Parallel Distributed Processing*, 2 voll., Cambridge (Mass.): The MIT Press.
- SCHÖNER, G. (2008): “Dynamical Systems Approaches to Cognition”, in R. Sun (a cura di), *Cambridge Handbook of Computational Psychology*, New York: Cambridge University Press, pp. 101-126.
- SMITH, Linda B. e THELEN, Esther (a cura di) (1993): *A Dynamic Systems Approach to Development*, Cambridge (Mass.): The MIT Press.
- SMOLENSKY, Paul (1988): “On the Proper Treatment of Connectionism”, *Behavioral and Brain Sciences*, 11, pp. 1-74.
- THELEN, Esther e SMITH, Linda B. (1994): *A Dynamic Systems Approach to the Development of Cognition and Action*, Cambridge (Mass.): The MIT Press.
- TSCHACHER, Wolfgang e DAUWALDER, Jean-Pierre (a cura di) (2003): *The Dynamical Systems Approach to Cognition*, Singapore: World Scientific.
- VAN GEERT, Paul (1995): “Growth Dynamics in Development”, in Port e Van Gelder (a cura di) (1995, pp. 313-337).
- VAN GELDER, Timothy J. (1998): “The Dynamical Hypothesis in Cognitive Science”, *Behavioral and Brain Sciences*, 21, pp. 615-665.
- VAN GELDER, Timothy J. e PORT, Robert F. (1995): “It’s about Time: An Overview of the Dynamical Approach to Cognition”, in Port e Van Gelder (a cura di) (1995, pp. 549-571).
- WARD, Lawrence M. (2002): *Dynamical Cognitive Science*, Cambridge (Mass.): The MIT Press.



# Gödel Łukasiewicz Logic

Martinvaldo König  
University of Cagliari  
e-mail: martinvk@iol.it

- 1 The logical system
- 2 Basic algebraic notions
- 3 Algebraic properties
- 4 Logical results

**ABSTRACT.** This paper studies many-valued logic endowed with two different kinds of implications: Łukasiewicz implication and Gödel implication. This research focuses on the class of algebras containing the algebraic counterpart of this new logic: the class of Heyting Wajsberg algebras. We prove that this variety is a discriminator variety. We show Gödel Łukasiewicz Logic to be regularly algebraizable, strongly complete, decidable and to have the Deduction-Detachment Theorem.

**KEYWORDS:** many valued logics, algebraizable logics, bounded distributive lattices, MV-algebras, HW-algebras, discriminator variety, equationally definable principal congruences, strong completeness, subdirectly irreducible algebras.

## Introduction

Many valued logic arose by the study of Jan Łukasiewicz who first, in the early twenties of the last century, understood the importance of employing an infinite set of truth values in the semantics of a deductive system. Once interpreted this set of values as the unit interval of real numbers, he proposed the following interpretation for the implicational connective:

$$x \rightarrow_L y := \begin{cases} 1 & \text{if } x \leq y \\ 1 - x + y & \text{otherwise} \end{cases}$$

Later C. C. Chang [14] in order to prove the completeness of the axioms of Łukasiewicz  $\mathfrak{K}_0$ -valued propositional calculus proved the Lindenbaum-Tarski algebra of this deductive system to be an MV-algebra. Many other interpretations of the implicational connective have been introduced in the scientific literature.

After Lotfi-A. Zadeh conceived fuzzy set theory [25] a renewed interest pushed philosophers and mathematicians to investigate the possible ways to enlarge the previous studies in many valued logic. A long debate on which connective is the most appropriate to interpret intersection in fuzzy sets took place during several years and then the scientific community seemed to agree to define the intersection operator (and its dual union) in an axiomatic way as a class: the set of triangular norms (and its dual conorms) [26]. A triangular norm is any continuous function  $t : [0, 1]^2 \mapsto [0, 1]$  that satisfies the following four properties: it has to be commutative, associative, monotonic and to have 1 as neutral element. Moreover by the residuation law  $x \wedge y \leq z \Leftrightarrow x \leq y \rightarrow z$  any triangular norm is associated to an implication (i.e. its residuum). The class of the triangular norms is a range of functions limited at the top by the minimum (i.e. the maximal triangular norm) and at the bottom by the Łukasiewicz triangular norm  $t_L$  (i.e. the minimal triangular norm) defined for any  $x, y \in [0, 1]$ ,  $xt_Ly := \min\{0, x + y - 1\}$ . The residuation law connects Łukasiewicz triangular norm with Łukasiewicz implication and minimum with the following implicational connective:

$$x \rightarrow_G y := \begin{cases} 1 & \text{if } x \leq y \\ y & \text{otherwise} \end{cases}$$

This implication has been introduced by Kurt Gödel in [18] and then is usually known as Gödel implication. Hence, Gödel implication and Łukasiewicz implication are the two residua delimiting the whole range of triangular norms. For this reason we want to study many-valued logic endowed with these two different implications as primitive operators.

## 1. The logical system

Let  $\Lambda := \langle \rightarrow_G, \rightarrow_L, \mathbf{0} \rangle$  be a language of type  $\langle 2, 2, 0 \rangle$ . Gödel Łukasiewicz Logic, in symbols  $GLL := \langle \Lambda, \vdash_{GLL} \rangle$  is the deductive system presented by the following collection of axioms (Ax1-8) and inference rule (MP1). First we

recall, for the sake of readability, the non-primitive connectives definitions:

$$\begin{aligned}
 \neg\alpha &:= \alpha \rightarrow_L \mathbf{0} \\
 \sim\alpha &:= \alpha \rightarrow_G \mathbf{0} \\
 \alpha \wedge \beta &:= \neg((\neg\alpha \rightarrow_L \neg\beta) \rightarrow_L \neg\beta) \\
 \alpha \vee \beta &:= (\beta \rightarrow_L \alpha) \rightarrow_L \alpha = \neg(\neg\alpha \wedge \neg\beta) \\
 \alpha \leftrightarrow \beta &:= (\alpha \rightarrow_L \beta) \wedge (\beta \rightarrow_L \alpha)
 \end{aligned}$$

$$\begin{aligned}
 (\text{Ax1}) \quad &\alpha \rightarrow_G \alpha \\
 (\text{Ax2}) \quad &(\alpha \rightarrow_G (\beta \wedge \gamma)) \leftrightarrow ((\alpha \rightarrow_G \gamma) \wedge (\beta \rightarrow_G \gamma)) \\
 (\text{Ax3}) \quad &(\alpha \wedge (\alpha \rightarrow_G \beta)) \leftrightarrow (\alpha \wedge \beta) \\
 (\text{Ax4}) \quad &((\alpha \vee \beta) \rightarrow_G \gamma) \leftrightarrow ((\alpha \rightarrow_G \gamma) \wedge (\beta \rightarrow_G \gamma)) \\
 (\text{Ax5}) \quad &((\alpha \rightarrow_G \alpha) \rightarrow_L \alpha) \leftrightarrow \alpha \\
 (\text{Ax6}) \quad &(\alpha \rightarrow_L (\beta \rightarrow_L \gamma)) \leftrightarrow (\neg(\alpha \rightarrow_L \gamma) \rightarrow_L \neg\beta) \\
 (\text{Ax7}) \quad &\neg\sim\alpha \rightarrow_L \sim\sim\alpha \\
 (\text{Ax8}) \quad &(\alpha \rightarrow_G \beta) \rightarrow_L (\alpha \rightarrow_L \beta) \\
 (\text{MP1}) \quad &\frac{\alpha, \alpha \rightarrow_L \beta}{\beta}
 \end{aligned}$$

**Definition 1.1.**  $\text{MP2} := \frac{\alpha, \alpha \rightarrow_G \beta}{\beta}$ .

**Lemma 1.1.**  $\text{MP1} \Rightarrow \text{MP2}$ .

*Proof.* By (Ax8) and two applications of MP1. □

**Definition 1.2.** Let us consider the set of well formed formulas of  $GLL$  defined by induction in the traditional way, in symbols  $Fm(\Lambda)$ . An evaluation on  $Fm(\Lambda)$  is a mapping  $e : Fm(\Lambda) \mapsto [0, 1]$  such that:

$$e(\alpha \rightarrow_L \beta) = \min\{1, 1 - e(\alpha) + e(\beta)\}$$

$$e(\alpha \rightarrow_G \beta) = \begin{cases} 1 & \text{if } e(\alpha) \leq e(\beta) \\ e(\beta) & \text{otherwise} \end{cases}$$

**Definition 1.3.** A formula  $v \in Fm(\Lambda)$  is a 1-tautology iff  $e(v) = 1$  for any evaluation  $e$ .

## 2. Basic algebraic notions

Heyting algebras are the algebraic counterpart (i.e. the class of algebraic structures which verify exactly the provable formulae) of intuitionistic propositional logic (pp. 380-383 [15]). In the same way Wajsberg algebras are the algebraic counterpart of the  $\aleph_0$ -valued Łukasiewicz propositional calculus and Wajsberg algebras are termwise definitionally equivalent to MV-algebras [13].

The class of Heyting Wajsberg algebras has been first introduced by Giampiero Cattaneo and Davide Ciucci [9]. They explained that by the composition of the two primitive operators with the  $0$ -element it is possible to define the modal-like operators of necessity and possibility (and their duals). Moreover the pairs of these unary operators generate rough approximation spaces of Boolean elements and Heyting Wajsberg algebras define an abstract environment linking fuzzy and rough sets.

In the sequel we are going to show that the set of 1-tautologies of Gödel Łukasiewicz logic (i.e. the logic whose the algebraic equivalent semantics is the class of Heyting Wajsberg algebras) contains both the one of intuitionistic propositional logic and the one of Łukasiewicz many valued propositional logic. In order to do that we will prove that the same relationships among the corresponding equational theories hold. First we have to report some well known definitions.

**Definition 2.1.** A Heyting algebra is a structure  $\mathcal{A} = \langle A, \wedge, \vee, \rightarrow, 0 \rangle$  of type  $\langle 2, 2, 2, 0 \rangle$  in which the following axioms are satisfied:

- (H1)  $x \rightarrow x = y \rightarrow y$
- (H2)  $(x \rightarrow y) \wedge y = y$
- (H3)  $x \rightarrow (y \wedge z) = (x \rightarrow z) \wedge (x \rightarrow y)$
- (H4)  $x \wedge (x \rightarrow y) = x \wedge y$
- (H5)  $(x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z)$
- (H6)  $0 \wedge x = 0$

In [23] it is proved that Heyting algebras are equivalent to residuated lattices where the multiplication operator coincides with the lattice meet operator, i.e.,  $\forall x, y \in A : x \wedge y = x * y$ . Moreover the unary operator  $\sim x := x \rightarrow 0$  is a Brouwer negation, i.e., it satisfies the three following properties:

- (B1)  $x \leq \sim \sim x$
- (B2)  $\sim (x \vee y) = \sim x \wedge \sim y$

$$(B3) \quad x \wedge \sim x = 0$$

**Definition 2.2.** A Wajsberg algebra is a structure  $\mathcal{A} = \langle A, \rightarrow, \neg, 1 \rangle$  of type  $\langle 2, 1, 0 \rangle$  in which the following axioms are satisfied:

$$(W1) \quad 1 \rightarrow x = x$$

$$(W2) \quad (x \rightarrow y) \rightarrow ((y \rightarrow z) \rightarrow (x \rightarrow z)) = 1$$

$$(W3) \quad (x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x$$

$$(W4) \quad (\neg x \rightarrow \neg y) \rightarrow (y \rightarrow x) = 1$$

Let us remark that in any Wajsberg algebra it is possible to define an order relation as

$$a \leq b \quad \text{iff} \quad a \rightarrow b = 1$$

With respect to this order relation, 1 is the maximum element and  $0 := \neg 1$  the minimum element and once defined the meet and join operators in the usual way:  $x \leq y$  iff  $x \wedge y = x$  iff  $x \vee y = y$ , the structure  $\langle A, \wedge, \vee, 0, 1 \rangle$  is a distributive lattice.

**Definition 2.3.** An *MV-algebra* is a structure  $\mathcal{A} = \langle A, \oplus, \neg, \mathbf{0} \rangle$  of type  $\langle 2, 1, 0 \rangle$  such that, upon defining for any  $x, y \in A$  :  $x \vee y := \neg(\neg x \oplus y) \oplus y$ , the following conditions are required:

$$(MV1) \quad (x \oplus y) \oplus z = (y \oplus z) \oplus x$$

$$(MV2) \quad x \oplus \mathbf{0} = x$$

$$(MV3) \quad x \oplus \neg \mathbf{0} = \neg \mathbf{0}$$

$$(MV4) \quad x \vee y = y \vee x$$

$$(MV5) \quad \neg \neg x = x$$

It is useful to define also the dual concepts:  $\mathbf{1} := \neg \mathbf{0}$ ,  $x \odot y := \neg(\neg x \oplus \neg y)$ , and  $x \wedge y := \neg(\neg x \odot y) \odot y$ . We observe that the relation  $x \leq y \Leftrightarrow x \vee y = y$  induces in every MV-algebra a distributive lattice order. In what follows we denote by  $\mathcal{A}_{[0,1]}$  the standard MV-algebra whose support is the unit real interval and by  $\mathcal{A}_{[0,1] \cap \mathcal{Q}}$  the algebra whose support is the unit rational interval. Notice that in both these algebras for any  $x, y$  in their support,  $\neg x := 1 - x$ ,  $x \oplus y := \min\{1, x + y\}$  and  $\mathbf{0} := 0$ .

**Definition 2.4.** An MV-algebra  $\mathcal{A}$  is a *Stonean MV-algebra* if and only if for any  $x \in A$ , there exists a Boolean (i.e. idempotent) element  $z$  such that  $z = \bigvee \{y \mid y \wedge x = \mathbf{0}\}$ . This property is equivalent to the possibility to define a Stonean

negation  $\sim$ , for any  $x \in A, \sim x := z$ . Moreover a Stonean MV-algebra is an algebra  $\mathcal{A} = \langle A, \oplus, \neg, \sim, \mathbf{0} \rangle$  of type  $\langle 2, 1, 1, 0 \rangle$ .

**Definition 2.5.** Let  $\mathcal{A} = \langle A, \rightarrow_L, \rightarrow_G, \mathbf{0} \rangle$  be an algebraic structure of type  $\langle 2, 2, 0 \rangle$ .  $\mathcal{A}$  is a *Heyting Wajsberg algebra* if for any  $x, y, z \in A$ , once defined

$$\begin{aligned} \neg x &:= x \rightarrow_L \mathbf{0} \\ \sim x &:= x \rightarrow_G \mathbf{0} \\ x \wedge y &:= \neg((\neg x \rightarrow_L \neg y) \rightarrow_L \neg y) \\ x \vee y &:= (x \rightarrow_L y) \rightarrow_L y \\ \mathbf{1} &:= \neg \mathbf{0} \end{aligned}$$

the following identities are satisfied:

- (HW1)  $x \rightarrow_G x = \mathbf{1}$
- (HW2)  $x \rightarrow_G (y \wedge z) = (x \rightarrow_G z) \wedge (x \rightarrow_G y)$
- (HW3)  $x \wedge (x \rightarrow_G y) = x \wedge y$
- (HW4)  $(x \vee y) \rightarrow_G z = (x \rightarrow_G z) \wedge (y \rightarrow_G z)$
- (HW5)  $\mathbf{1} \rightarrow_L x = x$
- (HW6)  $x \rightarrow_L (y \rightarrow_L z) = \neg(x \rightarrow_L z) \rightarrow_L \neg y$
- (HW7)  $\neg \sim x \rightarrow_L \sim \sim x = \mathbf{1}$
- (HW8)  $(x \rightarrow_G y) \rightarrow_L (x \rightarrow_L y) = \mathbf{1}$

Let us introduce the following conventions:

$$\begin{aligned} x \oplus y &:= \neg x \rightarrow_L y \\ x \odot y &:= \neg(\neg x \oplus \neg y) \end{aligned}$$

Any HW-algebra  $\mathcal{A} = \langle A, \rightarrow_L, \rightarrow_G, \mathbf{0} \rangle$  has the MV-algebra  $\mathcal{A}^* = \langle A, \oplus, \neg, \mathbf{0} \rangle$  as term reduct and any HW-algebra  $\mathcal{A} = \langle A, \rightarrow_L, \rightarrow_G, \mathbf{0} \rangle$  has the bounded distributive lattice  $\mathcal{A}^{**} = \langle A, \wedge, \vee, \mathbf{0}, \mathbf{1} \rangle$  as term reduct ([10], [11]).

It is also shown in [11] (proposition 1.1) that the natural partial order  $\leq$  defined by  $\wedge$  or  $\vee$  (*i.e.*  $x \leq y := x \wedge y = x$  or  $x \leq y := x \vee y = y$ ) has the following property:

$$(P) \quad x \leq y \Leftrightarrow x \rightarrow_L y = \mathbf{1} \Leftrightarrow x \rightarrow_G y = \mathbf{1}$$

**Proposition 2.1.** In any linear MV-algebra a Stonean negation can be defined in a natural way in order to have a Heyting Wajsberg algebra term reduct.

*Proof.* By [11] any HW-algebra is termwise definitionally equivalent to a Stonean MV-algebra. An MV-algebra is Stonean when there can be defined a Stonean negation (see also [12]). Any linear MV-algebra is trivially Stonean once defined the Stonean negation  $\sim_0$ :

$$\sim_0 x := \begin{cases} \mathbf{0} & \text{if } x \neq \mathbf{0} \\ \mathbf{1} & \text{if } x = \mathbf{0} \end{cases}$$

Then any linear MV-algebra enriched in such a way has a HW-algebra term reduct.  $\square$

A HW-algebra  $\mathcal{A}$  is *linear* (or *totally ordered*) iff for any pair of elements  $x, y \in A$ , either  $x \leq y$  or  $y \leq x$ . A linear algebra can be also called for short a *chain*.

Now we introduce the most important example of Heyting Wajsberg algebra, the model we will prove to generate the whole set of Heyting Wajsberg algebras as quasi-variety.

**Example 2.1** (Standard HW-algebra). The algebra whose support represents the set of truth-values of *GLL* is

$$\mathcal{A}_{[0,1]} = \langle [0, 1], \rightarrow_L, \rightarrow_G, \mathbf{0} \rangle$$

where  $[0, 1] \subset \mathbb{R}$ , and the two binary operators are defined

$$a \rightarrow_L b := \min\{1, 1 - a + b\}$$

$$x \rightarrow_G y := \begin{cases} 1 & \text{if } x \leq y \\ y & \text{otherwise} \end{cases}$$

Another important example of Heyting Wajsberg algebra in order to study *GLL* follows.

**Example 2.2** (Lindenbaum-Tarski algebra of *GLL*). Let the binary relation  $\equiv$  on  $Fm(\Lambda)$  be defined by  $\alpha \equiv \beta$  iff  $\vdash_{GLL} \alpha \rightarrow_L \beta$  and  $\vdash_{GLL} \beta \rightarrow_L \alpha$ . Then  $\equiv$  is a congruence relation and the quotient set  $Fm(\Lambda)_{\equiv}$  becomes a HW-algebra with the operation  $\rightarrow_L, \rightarrow_G$  and the constant  $\perp$  defined by

$$[\alpha]_{\equiv} \rightarrow_L [\beta]_{\equiv} := [\alpha \rightarrow_L \beta]_{\equiv}$$

$$[\alpha]_{\equiv} \rightarrow_G [\beta]_{\equiv} := [\alpha \rightarrow_G \beta]_{\equiv}$$

$$\perp := \{\gamma \mid \vdash_{GLL} \beta \text{ and } (\beta \rightarrow_L \mathbf{0}) \equiv \gamma\}$$

The reader can find a self contained proof of the standard completeness of Heyting Wajsberg algebras in [21]. However this result could be also inferred in an indirect way by the connection between Heyting Wajsberg algebras and Stonean MV-algebras [11] because in [20] it is proved that the variety of Stonean MV-algebras is generated by the model whose support is the unitary real interval (i.e. the standard Stonean MV-algebra). We report this important result.

**Theorem 2.3.**  $\mathbf{HW} = \mathbf{HSP}(\mathcal{A}_{[0,1]})$ .

In the sequel we will adopt the following notation. Given a HW-algebra  $\mathcal{A}$ ,  $\forall x \in A$  and  $\forall n \in \mathbb{N}$ :

$$nx = \begin{cases} \mathbf{0} & \text{if } n = 0 \\ x & \text{if } n = 1 \\ \underbrace{x \oplus \dots \oplus x}_{n\text{-times}}, & \text{if } 2 \leq n \in \mathbb{N} \end{cases}$$

and

$$x^n = \begin{cases} \mathbf{1} & \text{if } n = 0 \\ x & \text{if } n = 1 \\ \underbrace{x \odot \dots \odot x}_{n\text{-times}}, & \text{if } 2 \leq n \in \mathbb{N} \end{cases}$$

We recall below an important basic result that will be useful in the sequel.

**Lemma 2.1.** Let  $\mathcal{A} = \langle A, \rightarrow, \neg, \mathbf{0} \rangle$  be a Wajsberg algebra. Then for any  $x, y, z \in A$  the following properties hold:

1.  $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$
2.  $x \rightarrow y = \neg y \rightarrow \neg x$

*Proof.* See Lemma 4.2.3 and Lemma 4.2.4 in [13]. □

In the next lemma and corollary Heyting, Wajberg and Heyting Wajsberg algebras are dealt in signatures that are not the shortest (i.e. with the smallest number of symbols) in which they had been defined previously but these signatures are however their extensions. This can be found often in the algebraic literature and we decide to adopt it for the sake of a clear wide comprehension. We advise the reader to consult also Lemma 16 in [16] because of its strict connection with the following lemma.



**Lemma 2.2.** An algebra  $\mathcal{A} = \langle A, \wedge, \vee, \rightarrow_L, \neg, \rightarrow_G, \sim, \mathbf{0}, \mathbf{1} \rangle$  of type  $\langle 2, 2, 2, 1, 2, 1, 0, 0 \rangle$  has a Heyting Wajsberg term reduct if and only if the following conditions are satisfied:

1.  $\langle A, \rightarrow_L, \neg, \mathbf{0}, \mathbf{1} \rangle$  is a Wajsberg algebra.
2.  $\langle A, \wedge, \vee, \rightarrow_G, \sim, \mathbf{0}, \mathbf{1} \rangle$  is a Heyting algebra.
3. The Wajsberg algebra partial order  $\leq^{\rightarrow_L}$  ( $a \leq^{\rightarrow_L} b := a \rightarrow_L b = \mathbf{1}$ ) and the Heyting algebra partial order  $\leq^{\rightarrow_G}$  ( $a \leq^{\rightarrow_G} b := a \rightarrow_G b = \mathbf{1}$ ) coincide.

*Proof.* We prove separately the two implications.

( $\Rightarrow$ ) For the left-to-right direction, let  $\mathcal{A}$  be a HW-algebra.

- (1) In [11] it is shown that Heyting Wajsberg algebras are termwise definitionally equivalent to Stonean MV-algebras that are trivially MV-algebras. MV-algebras are proved to be termwise definitionally equivalent to Wajsberg algebras [13]. Thus any Heyting Wajsberg algebra satisfies the axioms of Wajsberg algebras.
- (2) We have to prove that the Heyting Wajsberg algebra  $\mathcal{A}$  satisfies the axioms (H1)-(H6). (H1) follows easily from (HW1). (H3), (H4) and (H5) coincide with (HW2), (HW3) and (HW4). Since by Theorem 2.3 we have that  $\mathbf{HW} = \mathbf{HSP}(\mathcal{A}_{[0,1]})$ , (H2) can be just verified in  $\mathcal{A}_{[0,1]}$  where either  $x \leq y$  or  $y < x$  and where we remind that  $\rightarrow_G$  is defined

$$x \rightarrow_G y := \begin{cases} 1 & \text{if } x \leq y \\ y & \text{otherwise} \end{cases}$$

Since any Heyting Wajsberg algebra  $\mathcal{A}$  has a bounded distributive lattice  $\mathcal{A}^{**}$  term reduct (H6) is satisfied.

- (3) This follows from (P).

( $\Leftarrow$ ) For the right-to-left direction, suppose  $\mathcal{A}$  satisfies (1)-(3). By (W2),  $(\mathbf{1} \rightarrow_L \mathbf{1}) \rightarrow_L ((\mathbf{1} \rightarrow_L x) \rightarrow (\mathbf{1} \rightarrow_L x)) = \mathbf{1}$  and then by (W1),  $x \rightarrow_L x = \mathbf{1} \rightarrow_L (x \rightarrow_L x) = \mathbf{1}$ . By (P) we have (HW1). It can be easily observed that (HW2), (HW3), (HW4), (HW5) correspond to (H3), (H4), (H5) and (W1). By combining the two statements in Lemma 2.1 (HW6) holds in any Wajsberg algebra.

One of the basic property in an MV-algebra is that for all  $x, y \in A$   $x \odot y \leq x \wedge y$ . By (H4)/(HW3)  $y \geq x \wedge y = x \wedge (x \rightarrow_G y) \geq x \odot (x \rightarrow_G y)$ . The Łukasiewicz implication  $\rightarrow_L$  is a residuum with respect to  $\odot$ , then by residuation law we have  $(x \rightarrow_G y) \odot x \leq y \Leftrightarrow (x \rightarrow_G y) \rightarrow_L (x \rightarrow_L y) = \mathbf{1}$  that is (HW8).

We need a remark in order to prove (HW7). Any Wajsberg algebra is termwise definitionally equivalent to an MV-algebra that is isomorphic to a subdirect product of a family of linear MV-algebras  $\{\mathcal{A}_i \mid i \in I\}$  [14]. Suppose  $\sim x = y$  and that in the subdirect representation each element  $z \in A$ , is expressed componentwise  $z = \langle z_1, \dots, z_n, \dots \rangle$  where  $1 \leq n \leq i \in I$ . We distinguish two cases:

1. By  $x \wedge y = \mathbf{0}$  (B3) if  $x_i \neq \mathbf{0}_i$  then  $y_i = \mathbf{0}_i$ .
2. Let  $x_i = \mathbf{0}_i$ .
  - 2.1. Suppose  $y_i \neq \mathbf{0}_i$ . Let us consider the set of element  $\{z^j \mid j \in J\}$  such that for any  $j \in J$ ,  $x \wedge z^j = \mathbf{0}$  and  $z_i^j \neq \mathbf{0}_i$ . This set is not empty because  $y$  belongs to it.  
 By (B3) for any  $j \in J$ ,  $\sim z_i^j = b_i = \mathbf{0}_i = x_i$ .  
 By (B1) for any  $j \in J$ ,  $z_i^j \leq \sim \sim z_i^j = \sim b_i = \sim x_i = y_i$  and hence  $y_i$  has to be the maximal element of  $\{z_i^j \mid j \in J\}$ .  
 Moreover  $y_i$  has to be  $\oplus$ -idempotent because otherwise  $k = 2y > y$  and  $y_i < k_i \in \{z_i^j \mid j \in J\}$  against maximality. Since a linear MV-algebra has the only two idempotent elements  $\mathbf{0}$  and  $\mathbf{1}$ ,  $y_i = \mathbf{1}_i$ .
  - 2.2.  $y_i = \mathbf{0}_i$ .

This means that in any case  $y$  is an idempotent element, i.e.  $\sim x \oplus \sim x = \sim x$ . By a well known MV-property (Theorem 1.5.3 in [13]) if  $y \oplus y = y$  then  $\neg y \wedge y = \mathbf{0}$  and then we obtain  $\sim x \wedge \neg \sim x \leq \mathbf{0}$  that by residuation law is equivalent to  $\neg \sim x \leq \sim \sim x$ . By hypothesis 3. we have (HW7). □

**Corollary 2.1.** Heyting-Wajsberg algebras are axiomatized in the signature  $\langle \wedge, \vee, \rightarrow_L, \neg, \rightarrow_G, \sim, \mathbf{0}, \mathbf{1} \rangle$  by:

1. A set of identities axiomatizing Heyting algebras in the signature  $\langle \wedge, \vee, \rightarrow_G, \sim, \mathbf{0}, \mathbf{1} \rangle$ .
2. A set of identities axiomatizing Wajsberg algebras in the signature  $\langle \rightarrow_L, \neg, \mathbf{1} \rangle$ .
3. The identity  $x \vee y = (x \rightarrow_L y) \rightarrow_L y$ .

From the previous lemma we conclude that Heyting Wajsberg algebras satisfy the following useful identities, which will be needed in the sequel:

$$x \rightarrow_L x = \mathbf{1}$$

$$x \rightarrow_L \mathbf{1} = \mathbf{1}$$

### 3. Algebraic properties

We recall the fundamental notion of discriminator variety on which a wide literature exists (see for example [4]).

**Definition 3.1.** A *discriminator term* on a set  $A$  is a function  $t : A^3 \mapsto A$  such that, for any  $a, b, c \in A$ :

$$t(a, b, c) = \begin{cases} c & \text{if } a = b \\ a & \text{otherwise} \end{cases}$$

**Definition 3.2.** An algebra  $\mathcal{A}$  is *subdirectly irreducible* if for every subdirect embedding  $f : \mathcal{A} \mapsto \prod_{i \in I} \mathcal{A}_i$  there is an  $i \in I$  such that  $\pi_i \circ f : \mathcal{A} \mapsto \mathcal{A}_i$  is an isomorphism.

**Definition 3.3.** A variety  $\mathbf{V}$  is said to be a *discriminator variety* if there exists a ternary term  $t$  such that  $t$  is a discriminator term on each subdirectly irreducible member of  $\mathbf{V}$ . A *pointed discriminator variety* is a discriminator variety with a constant term.

Now we can state a theorem that opens the path to many important other results:

**Theorem 3.1.** **HW** is a discriminator variety.

*Proof.* Since by Theorem 2.3  $\mathbf{HW} = HSP(\mathcal{A}_{[0,1]})$  in order to prove **HW** to be a discriminator variety we have just to find a ternary term that is a discriminator term on the standard HW-algebra  $\mathcal{A}_{[0,1]}$ . First we define:

$$\tau(a, b) := \neg \sim \neg((a \rightarrow_G b) \wedge (b \rightarrow_G a))$$

It can be easily verified that

$$\tau(a, b) = \begin{cases} 1 & \text{if } a \neq b \\ 0 & \text{otherwise} \end{cases}$$

Let us consider the following ternary term

$$\sigma(a, b, c) := (\tau(a, b) \wedge a) \vee (\sim \tau(a, b) \wedge c)$$

If  $a = b$ ,  $\tau(a, b) = 0$  and  $\sigma(a, b, c) = (0 \wedge a) \vee (1 \wedge c) = c$ . If  $a \neq b$ ,  $\tau(a, b) = 1$  and  $\sigma(a, b, c) = (1 \wedge a) \vee (0 \wedge c) = a$ . We have proved  $\sigma$  to be a discriminator term on  $\mathcal{A}_{[0,1]}$ , then on any member of **HW** and hence **HW** is a discriminator variety.  $\square$

A different discriminator term for HW-algebras has been provided in [2] but it has never been published. The property to be a discriminator variety opens the path to many related algebraic consequences. When such a variety has equationally definable principal congruences, if it is a variety generated by the real unit interval model so it is as quasi-variety: quasi-equations satisfied in the prototypical model are satisfied in any member of the variety. This result is known as strong completeness theorem.

In MV-algebras there is a one-one correspondence between ideals and congruences. To have this relationship in HW-algebras, in [21] a new adequate definition of filter (the dual concept of ideal) has been introduced. We are going to recall some basic definitions to introduce the concept of equationally definable principal congruences (i.e.EDPC) in order to prove the strong completeness theorem for HW-algebras with respect to the standard unit interval HW-algebra. On EDPC the four works of W.Blok and D.Pigozzi quoted in the references can give to the reader a deep and exhaustive insight.

**Definition 3.4.** A variety  $\mathbf{V}$  is congruence distributive (modular) if and only if for any  $\mathcal{A} \in \mathbf{V}$ , the lattice of congruences (i.e.  $\text{Con}(\mathcal{A})$ ) is a distributive (modular) lattice.

**Definition 3.5.** An algebra  $\mathcal{A}$  has the *congruence extension property* (CEP) if for every subalgebra of the same class  $\mathcal{B}$  and any congruence  $\theta \in \text{Con}(\mathcal{B})$  there is a congruence  $\phi \in \text{Con}(\mathcal{A})$  such that  $\theta = \phi \cap B^2$  where  $B^2$  is the set of all the 2-tuples of elements from  $B$ . A class  $\mathbf{K}$  of algebras has the CEP if every algebra in the class has the CEP.

**Definition 3.6.** Let  $\mathcal{A}$  be an algebra and  $a_1, \dots, a_n \in A$ , let  $\Theta(a_1, \dots, a_n)$  be a congruence generated by  $\{\langle a_i, a_j \rangle \mid 1 \leq i, j \leq n\}$ . i.e. the smallest congruence such that  $a_1, \dots, a_n$  are in the same class. The congruence  $\Theta(a_1, a_2)$  is called *principal congruence*.

**Definition 3.7.** A class of algebras  $\mathbf{K}$  is said to have *equationally definable principal congruences* (briefly EDPC) if there exists a finite set of quaternary terms  $p_i(x, y, z, w)$ ,  $q_i(x, y, z, w)$  of  $\mathbf{K}$  such that for every algebra  $\mathcal{A} \in \mathbf{K}$  and all  $a, b, c, d \in A$ ,

$$c \equiv d \pmod{\Theta^A(a, b)} \text{ if and only if } p_i^A(a, b, c, d) = q_i^A(a, b, c, d), \text{ for } i = 1, \dots, n.$$

Block and Pigozzi (Theorem 3.8 in [5]) proved every discriminator variety to have EDPC. We give an explicit proof for Heyting Wajsberg algebras for the sake of readability.

**Theorem 3.2.** The variety of Heyting Wajsberg algebras has EDPC.

*Proof.* In the previous theorem we have proved **HW** to be a discriminator variety. Then for any  $\mathcal{A} \in \mathbf{HW}$  there is a discriminator ternary term, such that for any  $a, b, c \in A$ :

$$t(a, b, c) = \begin{cases} c & \text{if } a = b \\ a & \text{otherwise} \end{cases}$$

We can define a couple of quaternary terms  $p_1$  and  $q_1$ :

$$p_1(a, b, c, d) := (d \rightarrow_G d) \rightarrow_G t(a, b, c)$$

$$q_1(a, b, c, d) := (c \rightarrow_G c) \rightarrow_G t(a, b, d)$$

Then **HW** has EDPC. □

**Corollary 3.1.** **HW** is congruence distributive (and thus congruence modular) and has the congruence extension property.

*Proof.* See Theorem 1.2 in [8]. □

The following theorem is part of the folklore of universal algebra. A clear presentation of this result can also be found in [1].

**Theorem 3.3.** Let  $K$  be a class of simple algebras such that the variety generated by  $K$  (we denote it  $V(K)$ ) has EDPC.  $V(K)$  coincides with the quasi-variety generated by  $K$ . In symbols,  $\mathbf{HSP}(K) = \mathbf{ISPPu}(K)$ .

**Theorem 3.4** (Strong completeness).  $\mathbf{HW} = \mathbf{ISPPu}(\mathcal{A}_{[0,1]})$ .

*Proof.* By Theorem 2.3  $\mathbf{HW} = \mathbf{HSP}(\mathcal{A}_{[0,1]})$  and in Theorem 3.2 it is shown **HW** to have EDPC.  $\mathcal{A}_{[0,1]}$  is trivially simple, then by Theorem 3.3 the standard **HW**-algebra generates the quasi-variety **HW**. □

Now we are going to show the finite model property for Heyting Wajsberg algebras. The following lemma and theorems provide the necessary steps.

It is important to remind that, by the classical construction, any real number can be identified by a sequence of Cauchy of rational numbers. I recall this definition:

**Definition 3.8.** A sequence  $a : n \mapsto a_n$  (we write  $\{a_n\}$ ) is a Sequence of Cauchy if and only if  $\forall \varepsilon > 0, \exists v, \forall m, n > v, d(a_m, a_n) < \varepsilon$ .

I am going to show that sequences of Cauchy are closed under the sum.

**Lemma 3.1.** Let  $a : n \mapsto a_n$  and  $b : n \mapsto b_n$  two sequences of Cauchy. Then  $a + b : n \mapsto a_n + b_n$  is a sequence of Cauchy.

*Proof.* By definition we have directly that  $\forall \varepsilon > 0, \exists v_1, \forall m, n > v_1, d(a_m, a_n) < \frac{\varepsilon}{2}$  and  $\forall \varepsilon > 0, \exists v_2, \forall m, n > v_2, d(b_m, b_n) < \frac{\varepsilon}{2}$ . Then  $\forall \varepsilon > 0, \exists v_i = \max\{v_1, v_2\}, \forall m, n > v_i, d(a_m + b_m, a_n + b_n) \leq d(a_m, a_n) + d(b_m, b_n) < \varepsilon$ .  $\square$

It can be easily observed that whether a sum is truncated (for instance,  $\oplus$  in  $\mathcal{A}_{[0,1] \cap Q}$ ) the property expressed by the previous lemma is not affected and it still holds. Moreover  $\{0_n\}$  and  $\{1_n\}$  are trivially sequences of Cauchy. Then  $\neg(n : n \mapsto a_n)$  is  $\{1_n\} + ((-n : n \mapsto a_n)) := (n : n \mapsto -a_n)$  and sequences of Cauchy are closed under  $\oplus$  and  $\neg$  componentwise.

The following theorem is part of the MV folklore but an explicit proof cannot be found in the literature. Then we have decided to prove and present it.

**Theorem 3.5.** Let  $\mathcal{A}$  be an MV-algebra,  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q}) = \text{HSP}(\mathcal{A}_{[0,1]})$ .

*Proof.* Since  $\mathcal{A}_{[0,1] \cap Q}$  is an MV-algebra and  $\mathbf{MV} = \text{HSP}(\mathcal{A}_{[0,1]})$  [13], we have  $\mathcal{A}_{[0,1] \cap Q} \in \text{HSP}(\mathcal{A}_{[0,1]})$  and then  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q}) \subseteq \text{HSP}(\mathcal{A}_{[0,1]})$ . On the other hand  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q})$  is closed under direct products, then an MV-algebra  $\mathcal{A}_{[0,1] \cap Q}^\omega$  whose support is made of infinite copies of  $\mathcal{A}_{[0,1] \cap Q}$  and whose operators are defined componentwise, belongs to  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q})$ . This variety is also closed under subalgebras. The set of the sequences of Cauchy in  $[0, 1] \cap Q$  endowed with componentwise-defined truncated sum and involutive negation  $\mathcal{A}_{[0,1] \cap Q}^{\{a_n\}}$  is an MV-subalgebra of  $\mathcal{A}_{[0,1] \cap Q}^\omega$ . Then  $\mathcal{A}_{[0,1] \cap Q}^{\{a_n\}} \in \text{HSP}(\mathcal{A}_{[0,1] \cap Q})$ . Since real numbers are identified by sequences of Cauchy there is trivially a homomorphism from  $\mathcal{A}_{[0,1] \cap Q}^{\{a_n\}}$  to  $\mathcal{A}_{[0,1]}$ .  $\mathcal{A}_{[0,1]}$  is the quotient of  $\mathcal{A}_{[0,1] \cap Q}^{\{a_n\}}$  modulo the congruence  $R$  defined by  $aRb := a$  and  $b$  have the same limit. Thus  $\mathcal{A}_{[0,1]} \in \text{HSP}(\mathcal{A}_{[0,1] \cap Q})$  and  $\text{HSP}(\mathcal{A}_{[0,1]}) \subseteq \text{HSP}(\mathcal{A}_{[0,1] \cap Q})$ . Then  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q}) = \text{HSP}(\mathcal{A}_{[0,1]})$ .  $\square$

**Theorem 3.6.** Let  $\mathcal{A}$  be a Stonean MV-algebra,  $\text{HSP}(\mathcal{A}_{[0,1] \cap Q}) = \text{HSP}(\mathcal{A}_{[0,1]})$ .

*Proof.* Since in [20] it is proved the variety of Stonean MV-algebras  $\mathbf{SMV} = \mathbf{HSP}(\mathcal{A}_{[0,1]})$  we have  $\mathbf{HSP}(\mathcal{A}_{[0,1] \cap \mathcal{Q}}) \subseteq \mathbf{HSP}(\mathcal{A}_{[0,1]})$ . We have to prove that  $\mathbf{HSP}(\mathcal{A}_{[0,1]}) \subseteq \mathbf{HSP}(\mathcal{A}_{[0,1] \cap \mathcal{Q}})$  that is, if  $\mathcal{A}_{[0,1] \cap \mathcal{Q}} \models t(x_1, \dots, x_n) = s(x_1, \dots, x_m)$  then  $\mathcal{A}_{[0,1]} \models t(x_1, \dots, x_n) = s(x_1, \dots, x_m)$ . We'll do it by induction on the number of occurrences of  $\sim$  in  $t \cup s$  (we denote it  $\sim\text{-occ}(t \cup s)$ ). If  $\sim\text{-occ}(t \cup s) = 0$  then  $t(x_1, \dots, x_n) = s(x_1, \dots, x_m)$  is an MV-equation and by Theorem 3.5, if  $\mathcal{A}_{[0,1] \cap \mathcal{Q}} \models t(x_1, \dots, x_n) = s(x_1, \dots, x_m)$  then  $\mathcal{A}_{[0,1]} \models t(x_1, \dots, x_n) = s(x_1, \dots, x_m)$ . By hypothesis of induction we suppose that this implication holds for each  $j \leq k = \sim\text{-occ}(t \cup s)$ . If  $\sim\text{-occ}(t \cup s) = k + 1$ , for some  $1 \leq i \leq n$ , without loss of generality we have that  $\mathcal{A}_{[0,1] \cap \mathcal{Q}} \models t(x_1, \dots, x_n, \sim t'(x_1, \dots, x_n)) = s(x_1, \dots, x_m)$  and  $\sim\text{-occ}(t' \cup s) \leq k$ . Since  $\sim$  in both  $\mathcal{A}_{[0,1] \cap \mathcal{Q}}$  and  $\mathcal{A}_{[0,1]}$  is  $\sim_0$  and

$$\sim_0(a) = \begin{cases} 0 & \text{if } a \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

there will be a term  $t^*$  for which, by induction hypothesis, if  $\mathcal{A}_{[0,1] \cap \mathcal{Q}} \models t^*(x_1, \dots, x_n, 0) = s(x_1, \dots, x_m)$  then  $\mathcal{A}_{[0,1]} \models t^*(x_1, \dots, x_n, 0) = s(x_1, \dots, x_m)$  and if  $\mathcal{A}_{[0,1] \cap \mathcal{Q}} \models t^*(x_1, \dots, x_n, 1) = s(x_1, \dots, x_m)$  then  $\mathcal{A}_{[0,1]} \models t^*(x_1, \dots, x_n, 1) = s(x_1, \dots, x_m)$ . It follows  $\mathcal{A}_{[0,1]} \models t(x_1, \dots, x_n, \sim t'(x_1, \dots, x_n)) = s(x_1, \dots, x_m)$ .  $\square$

An important feature in the interconnection between logic and algebra is the following: any logical finitely axiomatized propositional calculus is decidable if its Lindenbaum-Tarski algebra belongs to a variety that has the *finite model property* (FMP). We introduce this definition.

**Definition 3.9.** We say that a variety has the finite model property (FMP) if every identity that fails to hold in the class can be refuted in a finite member of the class. Varieties with FMP are said to be *generated by their finite members*.

Throughout this section, for each  $n = 1, 2, 3, \dots$  we shall use the notation:

$$\mathbf{Z}_{\frac{1}{n \cap [0,1]}} := \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\}.$$

Moreover we denote with  $\mathcal{A}_{n+1}$  the subalgebra of  $\mathcal{A}_{[0,1] \cap \mathcal{Q}}$  whose support is  $\mathbf{Z}_{\frac{1}{n \cap [0,1]}}$ .

**Theorem 3.7.** The variety of Stonean MV-algebras (**SMV**) has FMP.

*Proof.* By Theorem 3.6 and since  $\mathbf{SMV} = \mathbf{HSP}(\mathcal{A}_{[0,1]})$  [20], if an equation  $\alpha : t = s$  in the variables  $x_1, \dots, x_n$  fails in some Stonean MV-algebra  $\mathcal{A}$  then it fails in  $\mathcal{A}_{[0,1] \cap \mathcal{Q}}$ . Hence for some  $c_1, \dots, c_n \in [0, 1] \cap \mathcal{Q}$ ,  $t^{A_{[0,1] \cap \mathcal{Q}}}(c_1, \dots, c_n) \neq s^{A_{[0,1] \cap \mathcal{Q}}}(c_1, \dots, c_n)$ . So if  $c_1 \in \mathcal{A}_{m+1}, c_2 \in \mathcal{A}_{n+1}, \dots, c_n \in \mathcal{A}_{v+1}$  and  $\mathcal{A}_{ij}$  is the subalgebra of  $\mathcal{A}_{[0,1] \cap \mathcal{Q}}$  whose support is  $\mathbf{Z}_{\frac{1}{m \cdot n \cdot \dots \cdot v \cap [0,1]}}$  then  $t^{A_{ij}}(c_1, \dots, c_n) \neq s^{A_{ij}}(c_1, \dots, c_n)$ .  $\square$

**Corollary 3.2.** The variety of Heyting Wajsberg algebras (**HW**) has FMP.

*Proof.* By Theorem 3.7 and termwise equivalence between Stonean MV-algebras and Heyting Wajsberg algebras.  $\square$

**Theorem 3.8.** An algebra  $\mathcal{A} \in \mathbf{HW}$  is subdirectly irreducible if and only if it is a chain.

*Proof.* We prove separately the two implications.

( $\Rightarrow$ ) For the left-to-right direction, by definition for every subdirect embedding  $f : \mathcal{A} \mapsto \prod_{i \in I} \mathcal{A}_i$  there is an  $i \in I$  such that  $\pi_i \circ f : \mathcal{A} \mapsto \mathcal{A}_i$  is an isomorphism. Either by the subdirect representation Theorem for HW-algebras [21] or by the subdirect representation Theorem for Stonean MV-algebras [20] and termwise equivalence between Stonean MV-algebras and Heyting Wajsberg algebras for every HW-algebra  $\mathcal{A}$  there is a subdirect embedding  $h : \mathcal{A} \mapsto \prod_{i \in I} \mathcal{A}_i$  and for every  $i \in I$ ,  $\pi_i \circ h : \mathcal{A} \mapsto \mathcal{A}_i$  is an homomorphism and each  $\mathcal{A}_i$  is a chain. Then  $\mathcal{A}$  is isomorphic to a chain.

( $\Leftarrow$ ) For the right-to-left direction, suppose  $\mathcal{A}$  is a chain. By termwise equivalence between Stonean MV-algebras and Heyting Wajsberg algebras,  $\mathcal{A}$  is termwise definitionally equivalent to a linear Stonean MV-algebra  $\mathcal{A}'$ . Since by linearity in  $\mathcal{A}'$  we have  $\sim = \sim_0$ , there can be only two congruences-ideals:  $\{\mathbf{0}\}$  and the whole  $\mathcal{A}'$ , that is to say  $\mathcal{A}'$  is simple and hence so it is  $\mathcal{A}$ . By Theorem 8.1(II) in [19],  $\mathcal{A}$  is simple if and only if  $\mathcal{A}$  is subdirectly irreducible.  $\square$

**Lemma 3.2.** For each  $n, 0 \neq n \in \mathbf{N}$ , any two n-element Heyting Wajsberg chains are isomorphic.

*Proof.* Trivial, by termwise equivalence with MV-chains that are Stonean by  $\sim_0$ .  $\square$



Combining Theorem 3.8 and Lemma 3.2 the following theorem yields:

**Theorem 3.9.** Up to isomorphism, there is precisely one subdirectly irreducible HW-algebra for each finite cardinality.

The lattice of subvarieties of most discriminator varieties is totally ordered. Nevertheless I am going to show this is not the case of Heyting Wajsberg algebras. Let us notice  $\mathcal{A}_n$  the HW-algebra (or MV-algebra) of cardinality  $n$ .

**Theorem 3.10.** The lattice of subvarieties of Heyting Wajsberg algebras is not totally ordered. Moreover for any  $n, 2 < n \in \mathbb{N}$ ,  $\text{HSP}(\mathcal{A}_n) \not\subseteq \text{HSP}(\mathcal{A}_{n+1})$  and  $\text{HSP}(\mathcal{A}_{n+1}) \not\subseteq \text{HSP}(\mathcal{A}_n)$ .

*Proof.* The lattice of subvarieties of MV-algebras is not totally ordered (see [17]). In the same book it is shown that in the case of MV-algebras for any couple of prime natural numbers  $p_i, p_j$  we have that  $\text{HSP}(\mathcal{A}_{p_i+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_j+1})$  and  $\text{HSP}(\mathcal{A}_{p_j+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_i+1})$ . In [22] and [24] it is proved that any finite MV-algebra is isomorphic to a direct product of a family of linear MV-algebras. Any direct product of linear MV-algebras is a Stonean MV-algebra by the definition of Stonean negation  $\sim_0$  componentwise. Then any finite MV-algebra is Stonean and termwise definitionally equivalent to a HW-algebra. Since the Stonean operator can't be defined as a combination of the MV-operators  $\oplus$  and  $\neg$ , the class of theorems of a Stonean MV-algebra is the set-union of the set of theorems where the Stonean operator occurs and the set of theorems where it does not occur. These two sets are disjoint. Trivially an equation in the language of MV-algebras that is not satisfied in an MV-algebra continues not to be satisfied in the language of Stonean MV-algebra whereas the MV-algebra is Stonean. It follows that in the lattice of subvarieties of Stonean MV-algebras  $\text{HSP}(\mathcal{A}_{p_i+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_j+1})$  and  $\text{HSP}(\mathcal{A}_{p_j+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_i+1})$ . By termwise equivalence in the lattice of subvarieties of Heyting Wajsberg algebras  $\text{HSP}(\mathcal{A}_{p_i+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_j+1})$  and  $\text{HSP}(\mathcal{A}_{p_j+1}) \not\subseteq \text{HSP}(\mathcal{A}_{p_i+1})$ . For the sake of completeness we present a pair of equations in order to provide an infinite set of counterexamples to linearity. We leave to the reader the task to verify that for any integer  $n > 1$ ,

$$\begin{aligned} \mathcal{A}_{n+2} &\models (nx)^2 = (n-1)x \\ \mathcal{A}_{n+1} &\not\models (nx)^2 = (n-1)x \end{aligned}$$

and for any integer  $n \geq 0$ ,

$$\begin{aligned} \mathcal{A}_{n+1} &\models nx = (n+1)x \\ \mathcal{A}_{n+2} &\not\models nx = (n+1)x \end{aligned}$$

□

#### 4. Logical results

**Theorem 4.1.** Let  $\alpha$  be a formula of  $Fm(\Lambda)$ . Then,  $\vdash_{GLL} \alpha$  iff  $\alpha$  is a 1-tautology

*Proof.* We prove separately the two implications.

( $\Rightarrow$ ) For the left-to-right direction, it can be easily verified that axioms (Ax1)-(Ax9) are 1-tautologies and that Modus Ponens (MP1), the only deduction rule of  $GLL$ , cannot decrease the evaluation of an inferred formula.

( $\Leftarrow$ ) For the left-to-right direction, if  $\alpha$  is a 1-tautology then the algebraic term related to  $\alpha$ , meant in the natural traditional way (see p.21 in [13]),  $t_\alpha$  implies  $\mathcal{A}_{[0,1]} \models t_\alpha = 1$

By the standard algebraic completeness expressed in Theorem 2.3, if  $t_\alpha = 1$  is satisfied in  $[0, 1]$  it is satisfied in any model of **HW** and thus  $[\alpha]_{\equiv}$  is the top element of the Lindenbaum Tarski algebra of  $GLL$ . Hence  $\vdash_{GLL} \alpha$ . □

We have thus proved that  $GLL$  is the **1**-assertional logic of the variety **HW**, in Symbols  $GLL = S(\mathbf{HW}, 1)$ . Since **HW** is a **1**-regular variety, its **1**-assertional logic is regularly algebraisable with **HW** as equivalent algebraic semantics. A system of equivalence formulas is given by  $\{\alpha \rightarrow_L \beta, \beta \rightarrow_L \alpha\}$ .

Let  $\Phi := \langle \rightarrow_L, \neg \rangle$  be a language of type  $\langle 2, 1 \rangle$ .  $\aleph_0$ -valued Łukasiewicz logic, in symbols  $\mathbb{L} := \langle \Phi, \vdash_{\mathbb{L}} \rangle$  is the deductive system presented by the following collection of axioms (Ł1-4) and inference rule (MP):

- (Ł1)  $\alpha \rightarrow_L (\beta \rightarrow_L \alpha)$
- (Ł2)  $(\alpha \rightarrow_L \beta) \rightarrow_L ((\beta \rightarrow_L \gamma) \rightarrow_L (\alpha \rightarrow_L \gamma))$
- (Ł3)  $((\alpha \rightarrow_L \beta) \rightarrow_L \beta) \rightarrow_L ((\beta \rightarrow_L \alpha) \rightarrow_L \alpha)$
- (Ł4)  $(\neg \beta \rightarrow_L \neg \alpha) \rightarrow_L (\alpha \rightarrow_L \beta)$
- (MP)  $\frac{\alpha, \alpha \rightarrow_L \beta}{\beta}$

It can be easily observed that the language  $\Phi$  can be defined into the language  $\Lambda$  and thus the induced set of formulas  $Fm(\Lambda)$  is an extension of  $Fm(\Phi)$ .

**Theorem 4.2.** For any formula  $\psi$  in the set of formulas defined by induction on the traditional way in  $\Phi$  (i.e.  $Fm(\Phi)$ ):

$$\vdash_{\mathbb{L}} \psi \Rightarrow \vdash_{GLL} \psi$$

*Proof.* If  $\vdash_{\mathbb{L}} \psi$  by the semantical completeness theorem for the  $\aleph_0$ -valued propositional calculus proved by C. C. Chang (see [13]) and termwise equivalence between MV-algebras and Wajsberg algebras (also reported in [13]), let  $\mathscr{W}_{[0,1]}$  be the standard real unit interval Wajsberg algebra,  $\mathscr{W}_{[0,1]} \models \psi$ . By Lemma 2.2 the Wajsberg algebra  $\mathscr{W}_{[0,1]}$  is a term reduct of the standard real unit interval Heyting Wajsberg algebra  $\mathscr{A}_{[0,1]}$ , thus  $\mathscr{A}_{[0,1]} \models \psi$ . By Theorem 4.1,  $\vdash_{GLL} \psi$ .  $\square$

We introduce the intuitionistic propositional calculus IPC. Let  $\Omega := \langle \wedge, \vee, \rightarrow_G, \mathbf{0} \rangle$  be a language of type  $\langle 2, 2, 2, 0 \rangle$ . The intuitionistic propositional calculus, in symbols  $IPC := \langle \Omega, \vdash_{IPC} \rangle$  is the deductive system presented by the following collection of axioms (I1-9) and inference rule MP:

- (I1)  $\alpha \rightarrow_G (\beta \rightarrow_G \alpha)$
- (I2)  $\alpha \rightarrow_G (\beta \rightarrow_G (\alpha \wedge \beta))$
- (I3)  $(\alpha \wedge \beta) \rightarrow_G \alpha$
- (I4)  $(\alpha \wedge \beta) \rightarrow_G \beta$
- (I5)  $\alpha \rightarrow_G (\alpha \vee \beta)$
- (I6)  $\beta \rightarrow_G (\alpha \vee \beta)$
- (I7)  $(\alpha \vee \beta) \rightarrow_G ((\alpha \rightarrow_G \gamma) \rightarrow_G ((\beta \rightarrow_G \gamma) \rightarrow_G \gamma))$
- (I8)  $(\alpha \rightarrow_G \beta) \rightarrow_G ((\alpha \rightarrow_G (\beta \rightarrow_G \gamma)) \rightarrow_G (\alpha \rightarrow_G \gamma))$
- (MP)  $\frac{\alpha, \alpha \rightarrow_G \beta}{\beta}$

It can be easily observed that the language  $\Omega$  can be defined into the language  $\Lambda$ . Then the induced set of formulas  $Fm(\Lambda)$  is an extension of  $Fm(\Omega)$ .

**Theorem 4.3.** For any formula  $\rho$  in the set of formulas defined by induction on the traditional way in  $\Omega$  (i.e.  $Fm(\Omega)$ ):

$$\vdash_{IPC} \rho \Rightarrow \vdash_{GLL} \rho$$

*Proof.* If  $\vdash_{IPC} \rho$  by the completeness theorem of IPC respect to the class of Heyting algebras [15], let  $\mathscr{H}_{[0,1]}$  be the standard real unit interval Heyting algebra,  $\mathscr{H}_{[0,1]} \models \rho$ . By Lemma 2.2 the Heyting algebra  $\mathscr{H}_{[0,1]}$  is a term reduct of the standard real unit interval Heyting Wajsberg algebra  $\mathscr{A}_{[0,1]}$ , thus  $\mathscr{A}_{[0,1]} \models \rho$ . By Theorem 4.1,  $\vdash_{GLL} \rho$ .  $\square$

**Definition 4.1.** A deductive system  $S$  over a pointed language type with distinguished constant symbol  $\mathbf{1}$  is said to be a *pointed discriminator logic* if it arises as the  $\mathbf{1}$ -assertional logic of a pointed discriminator variety.

**Corollary 4.1.** Gödel Łukasiewicz Logic is a pointed discriminator logic.

**Definition 4.2.** Let  $S$  be a deductive system over a language type  $\Lambda$ . A non-empty set  $\{\sigma_i \mid i \in I\}$  of binary  $\Lambda$ -formulas is said to be a *deduction-detachment system* for  $S$  if, for all  $\Gamma \cup \{\alpha, \beta\} \subseteq Fm(\Lambda)$  and for all  $i \in I$ :

$$\Gamma, \alpha \vdash_S \beta \Leftrightarrow \Gamma \vdash_S \sigma_i(\alpha, \beta).$$

$S$  is said to have the (*uniterm*) *deduction-detachment theorem* (DDT) if it has a (unitary) deduction-detachment system.

DDT has been extensively studied in abstract algebraic logic. For a survey, see the tutorial paper [7].

**Theorem 4.4.** Gödel Łukasiewicz Logic has DDT.

*Proof.* We have proved the variety of Heyting Wajsberg algebras to be a discriminator variety and to have EDPC. W. Block and D. Pigozzi in [5] have proved that in such conditions the related arising  $\mathbf{1}$ -assertional logic (i.e. Gödel Łukasiewicz Logic) has DDT.  $\square$

**Theorem 4.5.** Gödel Łukasiewicz Logic is decidable.

*Proof.* In Corollary 3.2 we have proved the variety of Heyting Wajsberg algebras to have the finite model property (FMP). It is well known in abstract algebraic logics that if a finitely axiomatized variety has FMP then its  $\mathbf{1}$ -assertional logic is decidable. Thus  $GLL$  is decidable.  $\square$

**Theorem 4.6.** Gödel Łukasiewicz Logic is strongly complete.

*Proof.* In Theorem 3.4 we have proved the variety of Heyting Wajsberg algebras to be complete with respect to the standard unitary real interval model as quasi-variety (i.e. with respect to quasi-equations). Then its  $\mathbf{1}$ -assertional logic (i.e.  $GLL$ ) is strongly complete.  $\square$

## REFERENCES

- [1] VAN ALTEN, C. J. (1998), *An Algebraic Study of Residuated Ordered Monoids and Logics without Exchange and Contraction*, PhD thesis, University of Natal, Durban.
- [2] BIGNALL, R. J., VEROFF, R. and SPINKS, M. (2003), "On the assertional logics of the generic pointed discriminator and generic fixedpoint discriminator variety", preprint.
- [3] BLOK, W. J. and PIGOZZI, D. (1982), "On the Structure of Varieties with Equationally Definable Principal Congruences I", *Algebra Universalis*, 15, pp. 195-227.
- [4] BLOK, W. J., KÖLHER, P. and PIGOZZI, D. (1984), "On the Structure of Varieties with Equationally Definable Principal Congruences II", *Algebra Universalis*, 18, pp. 334-379.
- [5] BLOK, W. J. and PIGOZZI, D. (1994), "On the Structure of Varieties with Equationally Definable Principal Congruences III", *Algebra Universalis*, 31, pp. 1-35.
- [6] BLOK, W. J. and PIGOZZI, D. (1994), "On the Structure of Varieties with Equationally Definable Principal Congruences IV", *Algebra Universalis*, 32, pp. 545-608.
- [7] BLOK, W. J. and PIGOZZI, D. (2001), *Abstract Algebraic Logic and the Deduction Theorem*, Preprint.
- [8] BURRIS, S. and SANKAPPANAVAR, H.P. (1981), *A Course in Universal Algebra*, New York: Springer-Verlag.
- [9] CATTANEO, G. and CIUCCI, D. (2002), "Heyting Wajsberg algebras as an abstract environment linking fuzzy and rough sets", *Lecture Notes in Artificial Intelligence*, 2475, pp. 77-84.
- [10] CATTANEO, G., CIUCCI, D., GIUNTINI, R. and KONIG, M. (2004), "Algebraic Structures Related to Many Valued Logical Systems Part I: Heyting Wajsberg Algebras", *Fundamenta Informaticae*, 63, pp. 331-355.

- [11] CATTANEO, G., CIUCCI, D., GIUNTINI, R. and KONIG, M. (2004), “Algebraic Structures Related to Many Valued Logical Systems Part II: Equivalence Among Widespread Structures”, *Fundamenta Informaticae*, 63, pp. 357-373.
- [12] CATTANEO, G., GIUNTINI, R. and PILLA, R. (1998), “BZMV<sup>dm</sup> algebras and Stonean MV-algebras (Applications to fuzzy sets and rough approximations)”, *Fuzzy Sets and Systems*, 108, pp. 201-222.
- [13] CIGNOLI, R., D’OTTAVIANO, I. and MUNDICI, D. (2000), *Algebraic Foundation of Many-valued Reasoning*. Dordrecht: Kluwer.
- [14] CHANG, C. C. (1958), “An Algebraic Analysis of many valued logics”, *Transactions of the American Mathematical Society*, 88, pp. 476-490.
- [15] DUNN, J. M. and HARDEGREE, G. M. (2001), *Algebraic Methods in Philosophical Logic*. Oxford: Clarendon Press.
- [16] GIUNTINI, R., FREYTES, H., LEDDA, A. and PAOLI, F. (2009), “A discriminator variety of Gödel algebras with operators arising in quantum computation”, *Fuzzy Sets and Systems*, 160, pp. 1082-1098.
- [17] GALATOS, N., JIPSEN, P., KOWALSKI, T. and ONO, H. (2007), *Residuated Lattices: an Algebraic Glimpse at Substructural Logic*. Amsterdam: Elsevier.
- [18] GÖDEL, K. (1933), “Zum intuitionistischen Aussagenkalkül”, *Anzeiger Akademie der Wissenschaften Wien, Math.-Naturwissensch, Klasse*, 69, pp. 65-69.
- [19] JONSSON, B. (1995), “Congruence Distributive Varieties”, *Mathematica Japonica*, 42, pp. 353-402.
- [20] KONIG, M. (2005), “Many-valued Enriched with a Stonean Negation: a Direct Proof of Representation and Completeness”, *Logic and Philosophy of Science: an Electronic Journal*, 3, 1.
- [21] KONIG, M. (2011), “A Self Contained Proof of Standard Completeness for Heyting Wajsberg Algebras”, submitted.

- [22] MANGANI, P. (1973), “ Su certe algebre connesse con logiche a più valori”, *Bollettino dell’Unione Matematica Italiana*, 8, pp. 68-78.
- [23] MONTEIRO, A. (1955), “Axiomes Independants pour les Algebres de Brouwer”, *Revista de la Union Matematica Argentina y de la Asociacion Fisica Argentina*, 17, pp. 149-160.
- [24] TRACZYK, T. (1979), “On the Variety of Bounded BCK-algebras ”, *Mathematica Japonica*, 24, pp. 238-292.
- [25] ZADEH, L. A. (1965) “Fuzzy Sets”, *Information and Control* 8, pp. 338-353.
- [26] ZIMMERMANN, H. J. (1985), *Fuzzy set Theory and Its Applications*. Boston, Kluwer.





# **L&PS – Logic & Philosophy of Science**

## **Information on the Journal**

### **AIMS AND CONTENTS**

L&PS – Logic and Philosophy of Science is an on-line philosophical journal sponsored by the Department of Philosophy of the University of Trieste (Italy). The journal promotes both theoretical and historical research in the philosophy of science and logic, without excluding any particular cultural perspective.

Topics welcomed by the journal include:

- the theory of scientific knowledge and the analysis of the general methodological problems of science (such as scientific discovery, causation, scientific inference, induction and probability, the structure of scientific theories and their relations with empirical data);
- the methodological and foundational problems of the different sciences, from the natural, to the biomedical, to the social sciences;
- the problems related to the historical development of logic, in all its branches, and to the role of logical methods both in the general methodology of science and in the foundations of empirical and mathematical sciences;
- the philosophical problems raised by the development of the cognitive sciences and the philosophy of mind, with particular attention to those results that are relevant for the analysis of scientific practice;
- the epistemological problems related to Artificial Intelligence, robotics, virtual reality, and artificial life;
- the problems in the sociology and the history of science that are relevant to the philosophical investigation of science;
- the problems related to the ethics of science;
- the questions related to the historical and conceptual development of the philosophy of science and logic;
- the problems of the philosophy of language, with particular attention to those results that are relevant for logic and philosophy of science.

The journal will appear twice a year (in March and in September), and one of these issues will be usually devoted to a special topic.

## INFORMATION FOR THE AUTHORS

Papers submitted to the journal must be written either in Italian or in English, and must be accompanied by a short summary in English (and also in Italian for the articles written in Italian). All papers will be evaluated by anonymous referees.

In order to promote critical discussion and exchange among scholars, the journal is willing to publish reports on work in progress, to be submitted and evaluated according to the criteria already mentioned above.

The copyright is left to the authors, provided that any reprint of the paper explicitly mentions the version previously published in L&PS.

## EDITORIAL BOARD

Gilberto Corbellini (Roma) gilberto.corbellini@uniroma1.it

Mauro Dorato (Roma) dorato@uniroma3.it

Roberto Festa (Trieste) festa@units.it

Marco Giunti (Cagliari) giunti@unica.it

Roberto Giuntini (Cagliari) giuntini@unica.it

Simone Gozzano (L'Aquila) simone.gozzano@cc.univaq.it

Federico Laudisa (Milano) federico.laudisa@unimib.it

Francesco Paoli (Cagliari) paoli@unica.it

Mario Piazza (Chieti) m.piazza@unich.it

Guglielmo Tamburrini (Pisa) gugt@fls.unipi.it

### EDITORS IN CHIEF

Mauro Dorato

Roberto Festa

Roberto Giuntini

### ASSISTANT EDITORS

Marco Giunti

Francesco Paoli

## EDITORIAL ADISORY BOARD

Vito Michele Abrusci, *Roma*; Dario Antiseri, *Roma*; Giovanni Boniolo, *Pado-  
va*; Andrea Cantini, *Firenze*; Mirella Capozzi, *Roma*; Martin Carrier, *Bielefeld*;  
Arturo Carsetti, *Roma*; Ettore Casari, *Pisa*; Carlo Cellucci, *Roma*; Roberto  
Cordeschi, *Salerno*; Giorgio De Rossi, *Trieste*; Giuliano Di Bernardo, *Trento*;  
Rosaria Egidi, *Roma*; Maurizio Ferriani, *Bologna*; Maria Carla Galavotti, *Bo-  
logna*; Sergio Galvan, *Milano*; Pierdaniele Giaretta, *Padova*; Gurol Irzik,  
*Istanbul*; Theo A. F. Kuipers, *Groningen*; Diego Marconi, *Vercelli*; Enrico Mo-  
riconi, *Pisa*; Ilkka Niiniluoto, *Helsinki*; Francesco Orilia, *Macerata*; Paolo Par-  
rini, *Firenze*; Angelo Maria Petroni, *Bologna*; Huw Price, *Sydney*; Giorgio San-  
dri, *Bologna*; Marina Sbisà, *Trieste*; Silvano Tagliagambe, *Sassari*; Nicla Vas-  
sallo, *Genova*; Achille C. Varzi, *New York*; Alberto Voltolini, *Vercelli*; Gereon  
Wolters, *Konstanz*; Giancarlo Zanier, *Trieste*.

## TYPESETTING, GRAPHICAL ADVICE AND TECHNICAL ASSISTANCE

Gustavo Cevolani (g.cevolani@gmail.com)

Luca Tambolo (l\_tambolo@hotmail.com)

## WEBMASTER

Gustavo Cevolani (g.cevolani@gmail.com)