Chapter 6
Linguistics, Semiotics, and Philosophy

6.1 Linguistics

6.1.1 The Biology-Linguistics Connection

The idea that language may provide a useful metaphor or analogy for describing and understanding the complexity inherent in living systems was expressed by Pattee (1968), Marcus (1974), and others more than four decades ago. The biology-linguistics connection was significantly strengthened by the uncovering of the isomorphism between cell and human languages (to be discussed in Sect. 6.1.2) (Ji 1997a, b, 1999b, 2001, 2002a). Unlike the familiar biology-physics connection, which is characterized by *determinism*, the biology-linguistic connection may be said to be characterized by *quasi-determinism* (Ji et al. 2009b), which is akin to the rule-governed creativity (RGC) in linguistics (see Sect. 6.1.4). RGC refers to the fact that humans are endowed with the ability to generate an indefinitely large number of *meaningful* sentences from a finite number of words and grammatical rules. RGC is also related to another linguistic phenomenon known as the *arbitrariness of signs*, that is, the *arbitrariness* of the relation between signs and their objects resulting from the lack of any physical laws mandating it (Sect. 6.1.4) (Lyons 1992, 1993; Culler 1991).

An indirect evidence for the quasi-deterministic nature of biology surfaced during the DIMACS (Discrete Mathematics and Computer Science) Workshop *on Bimolecular Networks: Topological Properties and Evolution*, held at Rutgers on May 11–13, 2005. At this meeting, Alfonso Valencia from the National Center of Biotechnology in Spain gave a lecture entitled “Biodegradation network, and all what we need for its study.” Based on his research experience in the field of the structure-function correlations in proteins, he came to the conclusion that protein folds and functions might not be predictable from amino acid sequence data. Valencia’s pessimism seems to go against the prevailing presupposition of

S. Ji, *Molecular Theory of the Living Cell: Concepts, Molecular Mechanisms, and Biomedical Applications*, DOI 10.1007/978-1-4614-2152-8_6,

biophysicists working in the field of protein folding that 3-dimensional folds of proteins should be ultimately predictable based on their amino acid sequence information alone, the view being referred to as the *Anfinsen's dogma* (Newman and Bhat 2007) in analogy to the *Central Dogma* in molecular biology (see Sect. 11.1). Anfinsen's dogma is based on Anfinsen's discovery in 1954 of the spontaneous refolding of ribonuclease A after denaturation. He found that the enzyme refolded into its native conformation if the environmental conditions employed were carefully controlled, that is, if the denaturant urea was removed before 2-mercaptoethanol, but the enzyme did not refold correctly if the order of removing the denaturants were revered. As will be discussed in Sect. 11.1, Anfinsen's dogma may not be fully supported by more recent experimental findings.

Valencia's pessimistic conclusion regarding protein structure-function correlation reminded me of a similar situation that transpired between the sixteenth and the mid-nineteenth century in the field of the *theory of algebraic equations* (Aleksandrov et al. 1984, pp. 261–278, Vol. I). The following is a list of the key events in the development of the theory of algebraic equations:

1. Ferrari (1522–1565) solved the general fourth-degree (i.e., quartic) polynomial equation of the type, $x^4 + ax^3 + bx^2 + cx + d = 0$ in the radical form (i.e., including the square root of n, where n is a positive number).
2. In 1824, Abel (1802–1829) proved that the fifth-degree (i.e., quintic) polynomial equations could not be solved in the radical form.
3. In "Memoir on the conditions of solvability of equations in radicals" published in 1846, Galois (1811–1832) explained why the quintic or higher-order polynomial equations cannot be solved in radicals. In the process, Galois was led to formulate a new mathematical theory, that is, the *group theory,* which has since been found to apply to a wide range of mathematical problems, providing a universal organizing principle in modern mathematics.

It is interesting to note that it took three centuries for mathematicians to realize that, although the fourth- and lower-order polynomial equations could be successfully solved in radical forms, the fifth- and higher-order ones could not be so solved. The reason for this was found to be that the coefficients of the quintic and higher-order equations belonged to a different *field* than the field to which the quartic and lower-order equations belonged, the former field being insoluble and the latter solvable (thus constituting the so-called *Galois group*) (http://en.wikipedia.prg/wiki/Galois_theory). Similarly, based on the experimental and theoretical evidences that have accumulated during the latter decades of the twentieth century and the first one of the twenty-first, I came to the conclusion between 2005 and 2009 that, even though the 2-D structures (i.e., α-helices and β-sheets) of proteins can be largely determined based on amino acid sequences alone, the 3-D and higher-order structures of proteins might not be so determined because the 3-D and higher-order structures of proteins are functions not only of their amino acid sequences but also of the time- and space-dependent microenvironmental conditions inside the cell under which proteins fold. A similar idea was proposed by Klonowski and Klonowska (1982). This idea may be alternatively expressed as follows:

> The 2-D structures of proteins are deterministic and predictable based on their amino acid sequences alone which are largely time-independent, but their 3-D structures are non-deterministic and unpredictable because i) proteins are sensitive to the space-and time-dependent microenvironmental conditions under which they fold, and ii) the information concerning their environment is largely lost to the past. (6.1)

For the convenience of discussion, we may refer to Statement 6.1 as the postulate of the *unpredictability of the 3-D protein folds* (U3PF), which is here suggested to be analogous to the *insolvability of the fifth-order polynomial equations* (I5PE) in mathematics. The main point of constructing Table 6.1 (see below) is to suggest that, just as the centuries-long attempt to solve the fifth-order polynomial equation (5PE) in mathematics had been instrumental in establishing the *group theory*, so the decades-long effort on the part of biologists and biophysicists to solve the 3-D protein folds (3PF) problem may lead to the development of a novel theory of life, the beginning of which is here suggested to be the *theory of gnergons* (Sects. 2.3.2 and 4.9). According to this theory, all self-organizing processes in living systems (including protein folding) are driven by gnergons, discrete units of gnergy defined as the complementary union of information (gn-) and energy (−ergy). We can express these ideas more simply in terms of the following two formal statements:

I5PE ----> Group Theory (6.2)

U3PF ==> GnergonTheory (6.3)

where the arrow "X ---- > Y" reads "X leads to Y via crisp logic" and "==>" reads "X leads to Y via fuzzy logic". In other words, the *group theory* was arrived at based on *crisp logic*, whereas the *gnergon theory* may involve uncertainties and fuzzy logic (Sect. 5.2.5).

The concept of gnergons may provide a theoretical framework for *the principle of rule-governed creativity* (Sect. 6.1.4), the *rule-governedness* reflecting the *energy* principle and the *creativity* reflecting the *information* that encodes the consequences of the historical contingencies associated with biological evolution. Rule-governedness is predictable and deterministic while *creativity* is unpredictable and nondeterministic. Thus, it may be concluded that the principle of *rule-governed creativity* embodies the principle of the complementarity between *determinism* and *nondeterminism* on the one hand and the *predictability* and *unpredictability* (*or creativity*) on the other. Alternatively, we may refer to rule-governed creativity as "freedom within constraints."

The possible analogy between the field of algebraic equations in mathematics and that of protein folding in biology is summarized in Table 6.1.

One possible reason why the protein structures and functions cannot be predicted based on their amino acid sequence data alone may be because biological systems in general (of which proteins are parts) obey the principle of complementarity between

Table 6.1 The postulated analogy between the insolvability of the fifth-degree (or quintic) polynomial equations (I5PE) and the *unpredictability of the 3-D folds* (U3PF) of proteins based on their amino acid sequence data

		Theory of	
		Algebraic equations	Protein folding
Solved	What is	Fourth-degree and lower-order polynomial equations	2-D folds (i.e., α-helices, β-sheets)
	By whom	Ancient mathematicians	Computational biologists of the twentieth century
Insolvable	What is	Fifth-degree and higher-order polynomial equations	3-D and higher-order folds based on amino acid sequence information alone
	Proved by	N.H. Abel in 1824	Experimental and theoretical results accumulated by 2010
	Insolvability explained	By E. Galois in 1846	Probably because the 3-D protein folds are the function of (1) amino acid sequence and (2) the time- and space-dependent intracellular micro-environmental conditions under which proteins fold, and the information concerning (2) is largely lost to the past
New theory	Emerged	Group theory	*A New theory of Life* based on the principle of *physical determinism* and *historical contingencies*, here identified with the *Gnergon theory*. Gnergons are defined as the discrete units of gnergy, the complementary union of information (*gn-*) and energy *(−ergy*) that are postulated to underlie all self-organizing processes in the Universe including life (Sects. 2.3.2, 6.1.2, and 11.1) (Ji 1991, pp. 152–156)

Table 6.2 A postulated relative importance of *laws* and *rules* in physics, biology, and linguistics

	Physics (law-based)	Biology (law-rule complementarity-based)	Linguistics (rule-based)
Laws (governing *matter*/*energy*)	+ + + + +	+ + +	+
Rules (governing symbols or signs carrying *information*)	+	+ + +	+ + + + +

the *predictable* or *determinism* (the domain of physics) and *the unpredictable* or *creativity* (the domain of evolutionary biology and linguistics; or the domain of rule-governed creativity [Ji 1997a; Lyons 1992]). It seems likely that (1) *rules* wrought by evolution or social conventions and (2) the physical *laws* of nature play equally fundamental roles in biology in agreement with Pattee (2008) and Barbieri (2003). We may refer to this idea as the *complementarity between determinism and nondeterminism* (CDN). CDN so defined may be unique to biology as indicated by the third column in Table 6.2. CDN is related to the concept of *matter-symbol complementarity* that has been advanced by H. Pattee (1982, 2001, 2008; Umerez 2001) over the past three decades, according to which all living systems embody two complementary aspects – the physical law-governed *energetic*/*material* aspect and the evolutionary rule-governed *symbolic* aspect. This idea was renamed as the *von Neumann-Pattee principle of matter-sign complementarity* in Ji (1999b) to reflect not only the history of the development of this important concept starting with von Neumann but also its affinity to the more general notion of information/energy *complementarity* embodied in the new biology-based philosophical framework known as *complementarism* (see Sect. 2.3, and Ji 1995). The theory of organic codes proposed by Barbieri (2003) may be viewed as another species of the biological theories based on the matter-symbol complementarity and the complementarity between determinism (matter) and nondeterminism (symbol, or codes) (CDN). Furthermore, it is suggested in Sect. 6.1.3 that CDN is related to the *arbitrariness of signs*, one of the 13 design features of human language, that may have evolved to maximize the ability of messages to transmit information (Ji 1997a, pp. 36–37). *Nondeterminism*, *arbitrariness,* and *creativity* may all reflect different aspects of the same essential feature of the message source of a communication system, that is, the freedom for a sender to choose different messages, which maximizes when all messages have an equal probability for selection and hence which message happens to be chosen is *arbitrary* (Ji 1997a, pp. 36–37).

If the content of Table 6.2 is correct, biology may be described as neither *physics* nor *linguistics* but *a combination of both*. This same idea may be expressed as follows:

Biology is a complementary union of physics and linguistics. (6.4)

Physics and linguistics are the complementary aspects of biology. (6.5)

Biology has two complementary aspects – physics and linguistics. (6.6)

Physics is law-based, linguistics is rule-based, and biology is based on both physical laws and evolutionary rules. (6.7)

“**The language is a system of signs that represent concept**.”

1) *Signs* => *Molecules*
2) *Systems* => *Self-Organizing Systems*
3) *Concepts* => *Gene-Directed Cell Processes*

“The cell language is a self-organizing system of molecules, some of which encode, act as signs for, or trigger, gene-directed cell processes.”

Fig. 6.1 The “formal” derivation of the definition of cell language from that of human language given by Saussure (Culler 1991; Ji 2002b)

To understand biology, it is necessary to understand both physics and linguistics. (6.8)

It is impossible to understand biology based on the laws of physics and chemistry alone. (6.9)

Since linguistics is an important branch of the more general theory of signs, namely, *semiotics*, it behooves us to inquire into the connection between biology and linguistics on the one hand (Sect. 6.1.2) and biology and semiotics on the other (see Sect. 6.2).

6.1.2 The Isomorphism Between Cell and Human Languages: The Cell Language Theory

Human language can be defined as a system of *signs* obeying a set of rules that enables humans to communicate with one another. In other words, human language is a necessary condition for human communication. Similarly, there must be a language unique to living cells in multicellular (Ji 1997a, b) as well as unicellular (Stock et al. 2000) organisms, since cells must communicate among themselves in order to survive by carrying out their specialized biological activities in a coordinated manner. Such a language was named “cell language” in Ji (1997a). Cell language was defined as “a self-organizing system of molecules, some of which encode, act as signs for, or trigger, gene-directed cell processes” (Ji 1997a). This definition of cell language was inspired by the definition of human language given by Saussure (Culler 1991): “The language is a system of signs that represent concept.” The definition of cell language can be *formally* derived from that of human language given by Saussure by applying the following transformations: (1) replace “signs” with “molecules,” (2) replace “systems” with “self-organizing systems;” and (3) replace “concepts” with “gene-directed cell processes” (see Fig. 6.1).

Human and cell languages obey a common set of *linguistic* (or more generally *semiotic*) principles (Sect. 6.2), including *double articulation, arbitrainess of signs*

(Sect. 6.1.4), *rule-governed creativity*, the *energy requirement* of information transduction, storage, and transmission (Sect. 4.8) (Ji 1997a, 2001). Both human and cell languages can be treated as 6-tuples, $\{L, W, S, G, P, M\}$, where L is the alphabet, W is the lexicon or the set of words, S is a set of sentences, G is a set of rules governing the formation of sentences from words (called the *first articulation*) and the formation of words from letters (the *second articulation*), P is a set of physical mechanisms necessary and sufficient to implement a language, and finally M is a set of objects or processes, both symbolic and material, referred to by words, sentences, and their higher-order structures (e.g., texts). In Table 6.3, cell and human languages are compared with respect to the components of the linguistic 6-tuple. Table 6.3 contains two important concepts, *conformons* and *IDSs*, which play fundamental roles in the *Bhopalator model* of the living cell (Ji 1985a, b, 1991, 2002b), the user of cell language, as discussed in Chaps. 8 and 9. It is convenient to refer to cell language as *cellese* and human language as *humanese* (Ji 1999b), and the science of *cell biology* may be viewed as the translation of *cellese* to *humanese*. To the best of my knowledge, the first concrete application of the cellese concept was made by Aykan (2007) in formulating his so-called message-adjusted network (MAN) model of the gastro-enteropancreatic endocrine system.

Just as human language can be viewed as a *linear* network of letters forming words (i.e., *second articulation*), words forming sentences (i.e., *first articulation*), and sentences forming texts (i.e., *third articulation* [Ji 2005a, pp. 17–18]), so bionetworks (e.g., individual proteins or their networks known as metabolic networks) can be viewed as *multidimensional* generalizations of linguistic networks, where, for example, amino acids can be compared to letters, proteins to words, complexes of proteins to sentences, and network of complexes as texts (see Rows 7, 8, and 9 in Table 6.3). In addition to these structural or morphological similarities, there is a set of conventional/evolutionary rules and physical principles that is common to both human and cell languages, including the following:

1. The principle of self-organization (PSO) (6.10)

The phenomenon of self-organization was first observed in physical (e.g., Bernard instability [Kondepudi and Prigogine 1998; Kondepudi 2008]) and chemical systems (e.g., Belousov–Zhabotinsky reaction) as discussed in Sect. 3.1. Since the cell is an example of self-organized systems, it would follow that one of its functions, namely, communication with its environment including other cells (and hence cell language itself), must be self-organizing. Self-organization on the cellular level entails generating molecular forces from exergonic chemical reactions occurring internally. Also, since human communication is built upon (or presupposes) cell communication, it too must be an example of self-organizing processes. Therefore, it can be concluded that both cell and human languages are rooted in (or ultimately driven by) self-organizing chemical reaction–diffusion systems.

2. The minimum energy requirement for information transmission (6.11)

Both human and cell languages can be viewed as means of transmitting information in space and/or time. All information transmission requires dissipating free energy as

Table 6.3 A *formal* comparison between human and cell languages (Ji 1997a, 1999b)

	Human language (*humanese*)	Cell language (*cellese*)
1. Alphabet (L)	Letters	4 Nucleotides (or 20 amino acids)
2. Lexicon (W)	Words	Genes (or polypeptides)
3. Sentences (S)	Strings of words	Sets of genes (or polypeptides) expressed (or synthesized) coordinately in space and time dictated by DNA folds[a] (cell states)
4. Grammar (G)	Rules of sentence formation	The *physical laws* and *biological rules* mapping DNA sequences to folding patterns of DNA (polypeptides) under biological conditions[b]
5. Phonetics (P)	Physiological structures and processes underlying phonation, audition, and interpretation, etc.	Concentration and mechanical waves responsible for information and energy transfer and transduction driven by *conformons*[c] and *intracellular dissipative structures* (IDSs)[d]
6. Semantics (M)	Meaning of words and sentences	*Codes* mapping molecular signs to gene-directed cell processes
7. First articulation	Formation of sentences from words	Organization of gene expression events in space and time through *non-covalent interactions*[e] between DNA and proteins (or Space- and time-dependent non-covalent interactions among proteins, DNA, and RNA molecules). Thus, macromolecular complexes can be viewed as molecular analogs of sentences
8. Second articulation	Formation of words from letters	Organization of nucleotides (or amino acids) into genes (or polypeptides) through covalent interactions[f]
9. Third articulation	Formation of texts from sentences	Organization of chemical concentration gradients in space and time called *dissipative structures* (Babloyantz 1986; Kondepudi and Prigogine 1998) or *dissipatons* (see Sect. 3.1.5) in order to "reason" and "compute"[g]

[a] Just as verbal sentences (as written) are strings of words arranged linearly in the Euclidean space, so the cell-linguistic (or molecular) sentences are visualized as series of gene expression events arranged in time leading to dissipative structures or dissipatons (Chap. 9)

[b] Of all the folds of DNA and polypeptides allowed for by the laws of physics and chemistry, only small subsets have been selected by evolution (thereby giving rise to *biological information*) to constitute the gernome of a cell

[c] Sequence-specific conformational strains that carry both free energy (to do work) and genetic information (to control work) (Ji 1974a, 2000) (Chap. 8). Conformons are thought to provide immediate driving force (or serve as the force generators) for all nonrandom molecular processes inside the cell. Experimental evidence for conformons is discussed in Sect. 8.3

(continued)

Table 6.3 (continued)

[d] Space- and time-specific intracellular gradients of ions, biochemicals, and mechanical stresses (e.g., of the cytoskeletal system) that serve as the immediate driving forces for all cell functions on the microscopic level (see Chap. 9)

[e] Also called "conformational" interactions which involve neither breaking nor forming covalent bonds and depend only on the rotation around, or bending of, covalent bonds. Non-covalent interactions implicate smaller energy changes (typically around 1–3 kcal/mol) than covalent interactions which entail energy changes in the range of 30–100 kcal/mol

[f] Molecular interactions that involve changes in covalent bonds, that is, changes in valence electronic configurations around nuclei of atoms within a molecule

[g] This row is added to the original table published in (Ji 1997a,b). The *third articulation* (Ji 2005a) is a generalization and an extension of *second articulation*. Intercellular communication through chemical concentration gradients is well established in microbiology in the phenomenon of *quorum sensing* (Sect. 15.7) (Waters et al. 2008; Stock et al. 2000), whereby bacteria express a set of genes only if there are enough of them around so that they can combine and coordinate their efforts to accomplish a common task which is beyond the capability of individual bacteria. This phenomenon can be viewed as a form of *reasoning* and *computing* on the molecular level and the cell therefore can be viewed as *the smallest computational unit* (Ji 1999a), which may be referred to as *the computon,* a new term used here for the first time

mandated by Shannon's channel capacity equation (see Sect. 4.8). For artificial communication systems, the requisite energy is provided *externally* (e.g., a power station); for natural communication systems such as cells, the needed energy is generated from chemical reactions occurring *internally* utilizing chemicals provided by their environment. This difference in the sources of energy may have profound role in determining the global differences between artificial and living systems (e.g., macro vs micro sizes of system components).

3. The complementarity between determinism and non-determinism (6.12)

The process of communication can be viewed as a complementary union of *determinism* and *nondeterminism*. The deterministic aspect of communication reflects both the energy requirement (e.g., PSO, MERIT) and the syntactic rules (e.g., grammar) inherent in the language employed in communication, and the nondeterministic aspect (e.g., the principle of the arbitrariness of signs [PAS], the principle of rule-governed creativity [RGC], both described in Sect. 6.1.4) reflects the freedom of choice available to the sender of a message. Shannon's formula, Eq. 4.2, coupled with the definition of information given in Eq. (4.4), clearly indicates that, when there is no choice (i.e., no uncertainty), there is no information (Pattee 2008, p. 119), since "no choice" means "no selection," which in turn signifies "no reduction" in uncertainty.

To summarize, cell and human languages are *symmetric* with respect to at least five principles. Thus, to borrow the idioms of the group theory in mathematics, it may be stated that cell and human languages are the members of a *symmetry group* that has five "symmetry operators," here identified with (1) PSO, (2) MERIT, (3) CDN, (4) PAS, and (5) RGC, and hence may be designated as SG(5), where S and G stand for symmetry and group, respectively, and the Arabic numeral indicates the number of the principles that remain unchanged (or invariant, or symmetric) when

Table 6.4 An estimation of the average information content, I, or the complexity, H, of a linguistic text or a metabolic pathway based on the *cellese-humanese isomorphism* thesis and the simplified version of Shannon's formula, Eq. (4.3). The cellese is postulated to consist of two sub-languages – DNese and proteinese

Language	Letters in alphabet (a)	Letters in a word (b)	Words in a sentence (c)	Sentences in a text (d)	Complexity[a] of a text (H or I, in bits)
English	26	~10	~10	~10	~4.7 × 10^3
DNese	~60 (nucleotide triplets)	~100 (genes)	~10 (genes co-expressed)	~10 (genes working as a pathway)	~5.9 × 10^4
Proteinese	20 (amino acids)	~100 (polypeptide)	~10 (complexes/ metabolons)	~10 (metabolic pathways)	~4.3 × 10^4

[a]The complexity of a linguistic system (viewed from the perspective of the message source) is measured in terms of Shannon's entropy, H, that is, Eq. (4.3), which is equivalent to information, I, when viewed from the receiver's point of view (Seife 2006)

one language is replaced by the other. In other words, cell and human languages may be said to belong to a linguistic symmetry group with five symmetry operators, that is, the SG(5) group.

The set of the five rules common to cell and human languages may be divided into two complementary subsets – (1) *physical laws* (to be denoted as the P set) and (2) *linguistic* or *semiotic principles* (to be denoted as the L set) (See Sect. 6.2). It is clear that PSO and MERIT belong to the P set, and that the members of the L set include the principles of triple articulation as indicated in Table 6.3, the principles of the *arbitrariness of signs* and *rule-governed creativity* that are discussed next. These results agree with the matter-symbol complementarity thesis of Pattee (1969, 2008) and the basic tenets of the semantic biology advocated by Barbieri (2003, 2008a, b).

6.1.3 *The Complexities of the Cellese and the Humanese*

One of the most useful results that can be derived from the *cellese-humanese isomorphism* thesis is our ability to estimate the complexity (or the information content per symbol) of the cellese based on our experience with the humanese (see Table 6.4). The maximum *complexity* (viewed from the perspective of the message source) or the maximum information content (viewed from the receiver's perspective) (Seife 2006) of an English text can be estimated using the simplified version of Shannon's formula (see Eq. 43), that is,

$$\mathrm{I} = \mathrm{cbd}\ \log_2 \mathrm{a} \qquad (6.13)$$

where a is the number of letter in an alphabet, b is the number of letters in a word, c is the number of words in a sentence, and d is the number of sentences in a text.

In other words, Eq. 6.13 is based on the *principle of triple articulations* (PTA), denoted as 1, 2, and 3 as shown in Scheme (6.14):

$$\textbf{Letters} \xrightarrow{1} \textbf{Words} \xrightarrow{2} \textbf{Sentences} \xrightarrow{3} \textbf{Texts} \tag{6.14}$$

The *cellese hypothesis* (Ji 1997a, 1999b) assumes that PTA, Eq. 6.14, applies to the molecular processes occurring in the living cell and identifies the three levels of articulations of the cellese as shown in Scheme (6.15):

$$\textbf{Monomers} \xrightarrow{1} \textbf{Biopolymers} \xrightarrow{2} \textbf{Compexes} \xrightarrow{3} \textbf{Networks} \tag{6.15}$$

We will refer to Scheme (6.15) as the principle of the *triple articulations of the cellese* (TAC).

It is interesting to note that the complexities of linguistic and molecular texts (see the last column of Table 6.4) are the same within one order of magnitude. The *cellese* can be viewed as the *formal* aspect of the living cell whereas the set of physicochemical principles and laws embodied in "biocybernetics" (Ji 1991) represents the *physical* (i.e., energetic/material) aspect of the living cell. In other words, it may be stated that

> The *cell language theory* (Ji 1991, 1999b) and *biocybernetics* (Ji 1991) are the complementary aspects of the Bhopalator, the molecular model of the living cell. (6.16)

6.1.4 Double Articulation, Arbitrariness of Signs, and Rule-Governed Creativity

Of the 13 design features of human language described by Hockett (1960), three of them stand out in terms of their possible application to biology. These are (1) *double articulation* (extended to the triple articulation described in Table 6.3, (2) *arbitrariness of signs,* and (3) *rule-governed creativity* (see Table 6.6). It will be shown below that these features have molecular counterparts in cell language and may be necessary to maximize the channel capacity of biological communication systems (Ji 1997a), thereby facilitating biological evolution itself.

In Table 6.3, cell and human languages are compared from a formal (i.e., linguistic) point of view. In contrast, Table 6.5 compares cell and human languages from a physical point of view.

One of the design features of the human language, *arbitrariness of signs*, states that there is *no inevitable link* between the signifier (also called signs or representamen) (see Fig. 6.2) and the signified (object or referent) (Lyons 1993, p. 71). The arbitrary nature of signs in human language contributes to the flexibility and versatility of language, according to linguists. In addition, the author suggested that the arbitrariness of signs maximizes the amount of the information that can be

Table 6.5 A *physical (or material)* comparison between human and cell languages

	Human language (*Humanese*)	Cell language (*Cellese*)
1. Scale	Macroscopic	Microscopic
2. Signifier	Words	Molecules
3. Signified	Concepts	Gene-directed molecular processes
4. Rules wrought by	Social conventions	Biological evolution
5. Information transmission by	Sounds and light (i.e., sound and electromagnetic waves)	Conformons[a] & IDSs[b] (i.e., mechanical and concentration waves)
6. Maximum Information Principle made possible by	Arbitrariness of signs with respect to their objects or referents	Arbitrariness of molecular signs with respect to their target functions

[a] Conformational strains of biopolymers localized in sequence-specific sites (Chap. 8)
[b] Intracellular dissipative structures such as gradients of ions, metabolites, proteins, etc. inside the cell (Chap. 9)

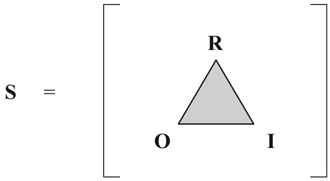

Fig. 6.2 A diagrammatic representation of the *Peircean sign triad*

transmitted by a sign, the idea being referred to as the *Maximum Information Principle* (Ji 1997a, pp. 36–37). Since cell language is isomorphic with human language, both belonging to the symmetry group, SG(5) (see Sect. 6.1.2), the *arbitrariness of signs* should apply to molecular signs in cell language, leading to the following inference:

> Just as the link between signs and their objects is arbitrary in human language, so the relation between molecular signs and their objects (or referents) are arbitrary, likely because such arbitrariness is necessary to maximize the amount of the information transmitted through or carried by molecular signs. (6.17)

For convenience, we will refer to Statement 6.17 as the *principle of the arbitrariness of molecular signs* (PAMS). Some experimental data supporting PAMS will be discussed in Sect. 12.10, where yeast RNAs are found to be divided into two distinct groups called the *cis*- and *trans-regulatory groups*, based on their genotypes, the former being less arbitrary (and thus carrying less genetic information) than the latter by a factor of about 3.

The *principle of arbitrariness of molecular signs* may be viewed as an aspect of the more general principle of *rule-governed creativity* (Ji 1997a). Both these principles appear to apply to multiple levels of biological organizations (as indicated in Table 6.6), from protein folding (Row 1a) to other processes on the molecular (Row 1b, 1c, and 1d) and cellular (Rows 2 and 3) levels.

Table 6.6 The principles of the arbitrariness of molecular signs, rule-governed creativity, and constrained freedom in action at various levels of living systems

Levels		Sign (rule, constraints)	Object/function (creativity, freedom)
1. Molecules	(a) Protein folding	Amino acid sequences	3-D shapes or folds
	(b) Catalysis	Protein shape	Chemical reaction catalyzed
	(c) Allostery	Allosteric ligand	Chemical reaction regulated
	(d) Binding	Transcription factor	Structural genes expressed
2. Cell–extracellular interactions		Intercellular messengers	Signal transduction pathways
3. Cell–intracellular interactions		Genome	Morphology, physiology

The arbitrary relation between amino acid sequence and the 3-D shape of a protein (see Row 1a in Table 6.6), which in turn determines its function, has already been pointed out in Tables 6.1 and 6.2 and is further discussed in Sect. 11.1. But protein folds are not entirely independent of amino acid sequences or completely dependent on them either, which may therefore be more accurately described as “quasi-deterministic” (Ji et al. 2009b). Although point mutations have been demonstrated to alter the shapes and functions of some proteins (but not all), it has also been found that an identical amino acid sequence can lead to more than one dominant conformations or folds, depending on the environmental conditions under which proteins fold. In fact, the Anfinsen’s classic experiments with ribonuclease A carried out in 1954 clearly demonstrate how sensitively dependent ribonuclease A conformations are on the environmental conditions under which it folded. The refolding of the denatured ribonuclease A induced by the removal of urea followed by the removal of 2-mercaptoethanol led to the native conformation of the enzyme with the 100% recovery of its enzymic activity but, when the refolding was induced by removing the denaturants in the reverse order, that is, removing 2-mercaptoethanol first followed by the removal of urea, the enzyme folded into nonnative conformations with only 1% of its enzymic activity recovered. Thus, the Anfinsen experiment of 1954 supports the notion that *conformations of proteins are the functions of both* (1) *amino acid sequences* and (2) *the environmental conditions under which proteins fold.* These dual conditions for protein folding constitute the core of the *unpredictability of the 3-D protein folds* (U3DPF) (see Statement 6.1). Thus the principle of arbitrariness of molecular signs (PAMS), Statement 6.17, may best regarded as reflecting an aspect of the molecular version of the principle of *rule-governed creativity* (*RGC*), another of the 13 design features of human language (Hockett 1960). RGC states that native speakers are able to produce an indefinitely large number of novel sentences based on finite sets of words and grammatical (or syntactic) rules and that these sentences can be understood by others in the linguistic community even though they never encountered them before (Lyons 1992, pp. 228–231; Harris 1993, pp. 57–58, 99–100). A *molecular version* of RGC may be stated as follows:

> Just as humans can produce an indefinitely large number of novel and meaningful sentences based on finite sets of words and grammatical rules, so living cells have evolved to produce an indefinitely large number of novel (i.e., unpredictable) functional molecular processes based on finite sets of molecules and physicochemical principles. (6.18)

Statement 6.18 may be referred to as the principle of rule-governed productivity, the principle of constrained freedom (PCF), or the principle of rule-governed molecular creativity. The principle of constrained freedom is symmetric or isomorphic with the principle of rule-governed creativity with respect to the following transformations.

1. Replacing "rule-governed" with "constrained"
2. Replacing "creativity" with "freedom"

These mutually replaceable elements in quotation marks may be considered to form a group comparable to the permutation group of Galois in his theory of polynomial equations (http://en.wikipedia.org/wiki/Galois_theory).

Just as it is impossible to predict the 3-D folds of a protein based on its amino acid sequence, so it is suggested in Row 1b in Table 6.6 that *it would be impossible to predict the nature of the chemical reaction that is catalyzed by an enzyme based solely on the 3-D shape (also called conformers, not to be confused with conformons of Chap.* 8*) of the enzymes alone, because the link between protein shape and the chemical reactions it catalyzes is not deterministic but arbitrary within physicochemical constraints (and hence quasi-deterministic),* reflecting the uncertainty about the environmental conditions under which biological evolution has selected the particular enzyme-catalyzed reaction.

The arbitrariness of the link between the shape of an allosteric ligand and the enzymic reaction it regulates (Row 1c) was pointed out by J. Monod (1971) who referred to it as "gratuity." Similarly, it is suggested in Row 1d that the link between the shape of a transcription factor and the nature of the structural gene whose expression it regulates is arbitrary within physicochemical constraints (i.e., *quasi-deterministic*), presumably to maximize the efficiency of the information transfer mediated by transcription factors (Ji 1997a).

Again in analogy to the unpredictability of the 3-D protein folds from amino acid sequences alone, so it is thought to be impossible to predict a priori the nature of the signal transduction pathways being activated based on the 3-D shape of intercellular messengers (Row 2) such as hormones, cytokines, and autoinducers.

Finally, Row 3 in Table 6.6 suggests that there may be no inevitable (i.e., deterministic) link between a genome and its phenotype, including the morphology and physiological processes of the organism involved. For example, human anatomy and physiology are arbitrarily related to and hence cannot be predicted from the human genome based on the laws of physics and chemistry alone. Again, to the extent that the link between a genome and its phenotype is arbitrary in the above sense, the genome has been optimized in order to transfer information from one generation to the next which entails information transfer in space and time. The identical twin studies of the human brain cognitive functions using functional magnetic resonance imaging (fMRI) technique (Koten Jr. et al. 2009) indicates that brain functions, such as memorizing and recognition, are partly gene-dependent and partly gene-independent, that is, quasi-deterministic with respect to genetic influence, consistent with the *principle of constrained freedom.*

6.2 Semiotics

Semiotics is the study of signs that dates back to ancient times when farmers predicted the weather from cloud patterns in the sky, or doctors diagnosed diseases based on the symptoms of patients. The American chemist-logician-philosopher Charles Sanders Peirce (1839–1914) has made a major contribution to establishing the field of modern semiotics which has been applied to a wide range of disciplines from linguistics to art, to philosophy, and to biology (Sebeok 1990; Emmeche 2002, 2003; Hoffmeyer 1996; Barbieri 2008a, b, c; Fernández 2008). Since signs can be divided into two types – macroscopic (e.g., stop signs) and microscopic (e.g., DNA) – based on their physical sizes, it would follow that semiotics itself can be divided into two branches – *macrosemiotics* and *microsemiotics* (Ji 2001, 2002a). Few biologists would deny that DNA molecules are *molecular signs,* since they encode (or refer to) RNA and protein molecules that are different from themselves. Likewise, few biologists would deny that the cell is the smallest physical system that can read and implement the genetic information/instructions encoded in DNA, leading to the following conclusions:

Molecular and cell biology constitute a part of *biosemiotics*, the study of living systems viewed as sign processors (Emmeche 2003), and since the cell is arguably the smallest DNA-based physical system that can process molecular information and perform molecular computation in the sense of Wolfram (2002) (Ji 1999a) and since the cell is the smallest unit of all living systems, *microsemiotics* constitutes the foundation of *biosemiotics*, just as *statistical mechanics* underlies *thermodynamics.*

6.2.1 The Peircean Theory of Signs

According to Peirce,

> A *sign*, . . ., is *something* which stands to *somebody* for *something* in some *respect* or *capacity*. (Buchler 1955, p. 99) (6.19)

Thus, "apple" is a sign referring to a juicy spherical fruit to someone, E, who speaks English. But "apple" is not a sign for a Korean, K, who does not understand English. For K, the sign, S, for the same object, O, is not "apple" but "sah-gwah."

So, it is evident that the definition of a sign, S, must include, in addition to O, a third element that Peirce referred to as *interpretant*, I, which is well characterized in the following paragraph quoted in (Houser et al. 1998):

> A *sign* is a thing which serves to convey knowledge of some other thing, which it is said to *stand for* or *represent*. This thing is called the *object* of the sign; the idea in the mind that the sign excites, which is a mental sign of the same object, is called an *interpretant* of the sign. (6.20)

Thus, the interpretant is the effect that S has on the mind of its interpreter or as the mechanisms or processes by which the interpreter or the processor of S is made to connect O and S. That is, in order for a sign process to occur successfully, there must be interactions among three elements, S, O, and I, within the sign processor. It was Peirce who first recognized the necessity of invoking these three elements in the definition of a sign and their actions (which he called "semiosis"). In other words, a sign, according to Peirce, is an irreducible triad of S, O, and I, which idea is often referred to as the "irreducibility of the sign triad" or the "triadicity of a sign." It is important to note that, in this definition of a sign, the term "sign" has dual roles – as a *component* of the sign triad and as the *sign triad itself*. To distinguish between these two roles, Peirce coined the term "representamen" to refer to the narrower sense of the term sign (Buchler 1955, p. 121). Thus, we may represent the Peircean definition of a sign diagrammatically as follows:

S = sign, **R** = representamen (also often called a *sign or a sign vehicle*), **O** = object, and **I** = interpretant. Unless pointed out otherwise, sign usually means **R**, a component of the irreducible sign triad. Also, it is important to note that the interpreter of **R** or the material system that process **R**, thereby implementing semiosis, is not explicitly discussed in semiotics literature but is assumed to be present. We may use the triangle itself to represent this interpreter, thus graphically distinguishing between *interpretant* (one of the three apexes or nodes) and *interpreter* (the triangle itself). It is important to note that the bracket symbolizes the *inrreducibility* of Peircean sign triad, that is, none of the three elements can be replaced by any other.

Although the study of signs can be traced back to the beginning of the human history as already pointed out, the investigation of signs as a fundamental science did not begin until the Portuguese monk John Poinsot (1589–1644) and C. S. Peirce (apparently independently of Poinsot) undertook their comprehensive and systematic studies of signs (Deely 2001).

The definition of signs that Peirce formulated can be extended to molecular biology, although Peirce probably did not know that such a possibility existed because he died about four decades before Watson and Crick discovered the DNA double helix that ushered in the era of molecular biology. Genes encoded in DNA fit the definition of the Peircean sign because they encode and stand for their complementary transcripts, RNA molecules and their functions, which are evidently distinct from the molecular structure of DNA. One plausible candidate for the *interpretant* for DNA viewed as a molecular sign is the *state of the cell*, since whether a given gene encoded in DNA is transcribed to RNA or not depends on the

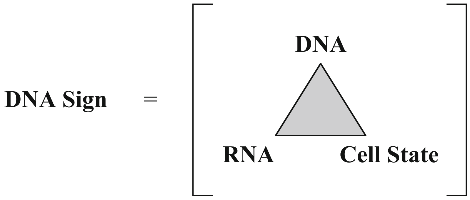

Fig. 6.3 Genes encoded in DNA as an example of Peircean signs at the molecular level. The role of interpretant is suggested to be fulfilled by *cell states*, and the interpreter of DNA is postulated to be the *cell* itself represented by the triangle. This definition seems to be consistent with the finding that only a select set of genes are expressed in cells at any given time and under any given environmental condition depending on the internal state of the cell (Nishikawa et al. 2008)

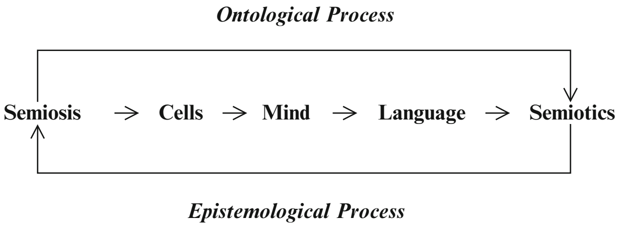

Fig. 6.4 The cyclical, or reversible, relation between *semiosis* and *semiotics*. The expression "A –> B" should be read as "B presupposes A" or "B cannot exist without A." The *upper arrow* from *left* to *right* indicates the *ontological* process in the Universe known as *evolution*, while the *lower arrow* from *right* to *left* signifies the *epistemological* causal relation resulting from the inferential activities of the human mind. It is assumed that ontological processes are independent of the human mind but epistemological processes are dependent on it. This figure is consistent with the principle of closure discussed in Sect. 6.3.2

state the cell is in, leading to the following diagrammatic representation of DNA as a sign (Ji 2002a) (Fig. 6.3).

Peirce distinguished between *semiotics* and *semiosis*. Semiotics is the systematic knowledge that human mind has created about semiosis based on empirical data, while semiosis refers to the totality of the natural and artificial processes whose occurrence requires the mediating role of signs. Thus, we may logically conclude that, *although semiotics depends on human mind, semiosis does not.* The causal relation between *semiotics* and *semiosis* may be represented diagrammatically as shown in Fig. 6.4.

6.2.2 *The Principle of Irreducible Triadicity: The Metaphysics of Peirce*

According to the metaphysics of Peirce, all phenomena, material or mental, living or nonliving, comprise three basic elements or aspects – *Firstness* (e.g., quality,

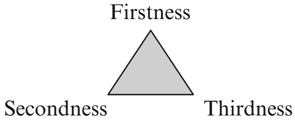

Fig. 6.5 A diagrammatic representation of the principle of irreducible triadicity of Firstness, Secondness, and Thirdness of Peirce (Goudge 1969; Hausman 1997; de Waal 2001; Sheriff 1994; Feibleman 1946)

feeling, possibilities), *Secondness* (e.g., facts, actualities, reaction, interaction, brute force), and *Thirdness* (e.g., generality, laws, habit-taking, representation, reasoning) (Fig. 6.5). For example, in logic, there are three kinds of relations; C = monadic, A = dyadic, and B = triadic relation. We may represent this principle diagrammatically as follows:

The *Threeness* plays a fundamental role in the metaphysics of Peirce, metaphysics being the study of the most general traits of reality. *Reality is the object of the conclusions one cannot help drawing*. As pointed out by Pierce, "When a mathematical demonstration is clearly apprehended, we are forced to admit the conclusion. It is evident; and we cannot think otherwise." (Goudge 1969). Metaphysics studies "the kinds of phenomena with which every man's experience is so saturated that he usually pays no particular attention to them." One way to get a feel of the three metaphysical categories of Peirce is through some of the examples that Peirce gave of these categories throughout his career. These are collected in Table 6.7, which was adopted from (Debrock 1998). It is evident that the examples are not logically tight, and, indeed, they are "vague" or "fuzzy" (Sect. 5.2.5), and even contradictory in some cases, having some overlaps here and there and missing some examples as well. Nevertheless, it is possible to recognize (1) the unmistakable family resemblances among most of the items listed within each category (i.e., within each column) and (2) distinct family characteristics present among the three categories (i.e., within each row).

6.2.3 *Peircean Signs as Gnergons*

One corollary of Fig. 6.4 is that the elucidation of the connection between *semiotics* and *life* would be tantamount to elucidating the principles underlying *semiosis* itself (in agreement with Sebeok 1990), and this is because life (as exemplified by cells and mind) presupposes semiosis. Based on the information-energy complementarity principle discussed in Sect. 2.3.2, we can conclude that, like all fundamental processes in nature, *semiosis* must have two complementary aspects – the *energetic/material* (e.g., computer hardware, or ATP in cells) and the *informational* (e.g., computer software or genetic information encoded in DNA). Of these two aspects,

Table 6.7 The evolution of Peirce's nomenclature of categories (Reproduced from Debrock 1998 except items 8 and 9)

	Year (*Peirce's age*)	Firstness	Secondness	Thirdness
1	1867 (*28*)	Quality	Relation	Representation
2	1891 (*52*)	First	Second	Third
3		Spontaneity	Dependence	Mediation
4		Mind	Matter	Evolution
5		Chance	Law	Tendency to take habits
6		Sporting	Heredity	Fixation of character
7		Feeling	Reaction	Mediation
8	1894 (*55*)	–	–	Learning
9		–	–	Government
10	1896 (*57*)	Quality	Fact	Law
11	1897 (*58*)	Ideas of feelings	Acts of reaction	Habits
12		Quality	Shock/vividness	–
13		Feeling	Reaction	Thought
14	1898 (*59*)	Quality	Reaction	Mediation
15		First qualities/ideas	Existence/reaction	Potential/continuity

the traditional semiotics as formulated by Peirce has emphasized primarily the *informational* aspect of semiosis, apparently ignoring the equally fundamental e*nergetic/material* aspect. It was only with the advances made in both *experimental* and *theoretical* branches of molecular and cell biology during the past several decades that *the essentiality of the energy/material aspect of semiosis has come to light* (Ji 1974a, b, 1985a, b, 1988, 1991, 1997a, b, 1999b, 2000, 2002a, b, 2004a, b). Thus it has been postulated that all self-organizing processes in the Universe, including semiosis, are driven by a complementary union *of information* and *energy,* that is, *gnergy* (Sects. 2.3.2 and 4.13) and (Ji 1991, 1995). Since information can be alternatively called "gnon" (from the Greek root *gnosis* meaning knowledge) and energy "ergon" (from Greek root *ergon* meaning work or energy), the *gnergon*, the discrete unit of *gnergy*, can be viewed as the complementary union of the *gnon* and the *ergon*:

$$\text{Gnergon} = \text{Gnon}^\wedge\text{Ergon} \tag{6.21}$$

where the symbol "^" denotes a *generalized complementarity relation as* defined in Sect. 2.3.3 (Ji 1991, 1995). That is, "C = A^B" reads as "A and B are complementary aspects of C," or "C is a complementary union of A and B." Since it has been postulated that Gnergy serves as the universal driving force for all self-organizing processes in this Universe (see Fig. 4.8), including molecular processes in the living cell (Ji 1991), we can interpret Fig. 6.4 as implying the following general statement:

Life results from *semiosis* driven by *gnergy*. (6.22)

Those not familiar with Peirce's (1839–1914) semiotics may think of signs as synonymous with "symbols" like stop signs and written words on printed pages.

Table 6.8 The classification of signs based on the dual trichotomies – (1) the ontological/material trichotomy (OT) (*first row*), and (2) the phenomenological/formal (PT) trichotomy (*first column*) (Ji 2002c)

	OT		
PT	Firstness (Potentiality)	Secondness (Facts)	Thirdness (Law)
Firstness (Sign)	*Qualisign*	*Sinsign*	*Legisign*
Secondness (Object)	*Icon*	*Index*	*Symbol*
Thirdness (Interpretant)	*Rheme*	*Dicent Sign*	*Argument*

Such a view is frequently referred to as "glossocentric" or "language-centered." But the concept of signs according to Peirce is much more general and includes not only linguistic symbols, but also icons (e.g., portraits, statutes, maps, electronic circuit diagrams), and indexes (e.g., smokes, laughter, fever, weathervane). The generality of signs is, in part, due to the fact that we think in signs. As someone said: *Think of an elephant; do you have an elephant in your head?* The neuronal firing patterns associated with our thoughts are signs representing their objects, whatever they may be, because neuronal firing patterns are not identical with the objects that they stand for. Peirce divides signs into a total of nine classes (Buchler 1955):

> Signs are divisible by three trichotomies; first, according as the sign itself is a mere quality ["qualisign"], is an actual existent ["sinsign"], or is a general law ["legisign"; secondly, according as the relation of the sign to its object consists in the sign's having some character in itself ('icon'), or in some existential relation to the object ["index"], or in its relation to an interpretant ["symbol"]; thirdly, according as its interpretant represents it as a sign of possibility ["rheme"] or as a sign of fact ["dicent sign"] or a sign of reason ["argument"]. (6.23)

The term "interpretant" here can be understood as the effect that a sign has on the mind of an interpreter, or as "meaning," "significance," or "more advanced sign." The above classification of signs by Peirce is summarized in Table 6.8.

Each of the nine types of signs appearing in the *interior* of Table 6.8 has dual aspects (reminiscent of the wave/particle duality of light) – (1) the *ontological* (or *material*) aspect and (2) the *phenomenological* (or *formal*) aspects, which appear on the *margins* of the table. The ontological/material aspect of a sign can be identified with *energy*/*matter* properties, while the phenomenological/formal aspect with *informational* properties. It is for this reason that the Peircean signs located in the interior of Table 6.8 can be viewed as examples of *gnergons*, the discrete units of gnergy postulated to be the ultimate cause of, or ground for, all self-organizing (or pattern-forming) processes in the Universe (Ji 1991, 1995). Since all sign processes (semiosis) can be viewed as species of self-organizing processes, ultimately driven by the free energy of exergonic chemical reactions (e.g., ATP hydrolysis or oxidation of NADH) or physical processes (e.g., heat flow, solar radiation, the Big Bang, etc.), it would follow that gnergons are the ultimate causes of semiosis (Ji 1995, 2002c) consistent with Fig. 4.8.

Complementarism, a scientific metaphysics rooted in both contemporary biology and Bohr's complementarity (Sect. 2.3.4), states that the ultimate reality consists in a complementary union of *information* and *energy*, that is, *gnergy*. Since signs are species of gnergons, it would follow that Peirce's semiotics falls within the domain of complementarism. This assertion may be supported by the following arguments:

1. Peirce's semiotics deals mainly with macroscopic signs, that is, signs with macroscopic dimensions "perfusing" the Universe; Peirce dealt mainly with *macrosemiotics*. This is not surprising because Peirce died in 1914, about four decades before the discovery of DNA double helix that ushered in the age of molecular biology and *microsemiotics* (Ji 2001, 2002a).
2. Complementarism can be applied not only to Peirce's semiotics (as suggested above) but also to molecular and cell biology, as evident in the formulation of the theory of "microsemiotics" based on the gnergy concept (Ji 2002a, c). *Microsemiotics* can be regarded as synonymous with the twin theories of the living systems known as *biocybernetics* (Ji 1991) and *cell language theory* (Ji 1997a). Thus the following relation suggests itself:

$$\begin{aligned}\text{Complementarism} &= \textit{Macrosemiotics} + \textit{Microsemiotics}\\ &= \text{Peirce's semiotics} + \text{Biocybernetics}/\\ &\quad\ \text{Cell Language Theory}\end{aligned} \tag{6.24}$$

Consistent with Peirce's triadic ontology, the principle of *complementarity* may itself be manifested in the Universe in three distinct modes:

Firstness = **Complementarity in metaphysics** (e.g., Yin and Yang as complementary aspects of the Tao of Lao-tze; Extension and Thought as the complementary aspects of Substance of Spinoza; Body and Mind as the complementary aspects of the Flesh of Merleau-Ponty [Dillon 1997])

Secondness = **Complementarity in physics** (e.g., the wave-particle duality of light)

Thirdness = **Complementarity in life sciences** (e.g., hysterical anesthesia of William James [Stephenson 1986]), physiology (i.e., the left-right hemispheric specialization [Cook 1986]), and molecular and cell biology (e.g., the *information-energy complementarity* of *gnergy* [Ji 1991, 1995])

These ideas are schematically represented in Fig. 6.6.

If the ideas expressed in Fig. 6.6 are correct, the separation and divergence of physics and metaphysics that are widely believed to have begun with Galileo's experiments with falling bodies in the seventeenth century may be expected to be reversed through the mediating role of the life sciences in the twenty-first century. In other words, the principle of information/energy complementarity manifested in *biology* (Ji 1991, 1995) may provide the theoretical framework for integrating *metaphysics* and *physics*.

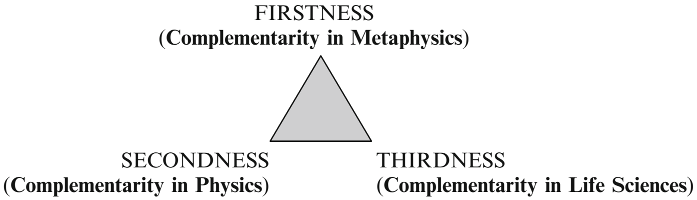

Fig. 6.6 The three modes of being of the generalized complementarity (Ji 1995). This diagram suggests the possibility that *life sciences* as Thirdness may serve as the mediator between *metaphysics* and *physics*. *Life science* may be viewed as synonymous with *cognitive sciences, since all organisms are cognizant of and interact with their environment*. The three nodes of the triangle may also be interpreted diachronically (Sect. 4.5): Firstness gave rise to Secondness, which in turn gave rise to Thirdness

6.2.4 *Macrosemiotics versus Microsemiotics: The Sebeok Doctrine of Signs*

As indicated in Sect. 6.2, we can divide semiotics into two branches – the *macrosemiotics* dealing with macroscopic signs such as written words and texts, and the *microsemiotics* concerned with molecular signs such as DNA, cytokines, and neurotransmitters, etc. Peirce did not have access to the empirical evidence that came to light only in the mid-twentieth century, that semiotic processes are not confined to the macroscopic world (*macrosemiosis*) but also occur on the molecular level (*microsemiosis*). The possibility of extending Peircean semiotics from macroscale to microscale was clearly foreseen by Sebeok in 1968 when he wrote (as cited in Deely 2001):

> . . .the genetic code must be regarded as the most fundamental of all semiotic networks and therefore as the prototype for all other signaling systems used by animals, including man. From this point of view, molecules that are quantum systems, acting as stable physical information carriers, zoosemiotic systems, and, finally, cultural systems, comprehending language, constitute a natural sequel of stages of ever more complex energy levels in a single universal evolution. It is possible, therefore, to describe language as well as living systems from unified cybernetic standpoint . . . A mutual appreciation of genetics, animal communication studies, and linguistics may lead to a full understanding of the dynamics of semiotics, and this may, in the last analysis, turn out to be no less than the definition of life. (6.25)

Elsewhere (Ji 2001), it was suggested that Statement 6.25 be referred to as the *Sebeok doctrine of signs* for convenience of reference.

The first full-length paper on microsemiotics was published in (Ji 2002a). Despite the enormous difference in the sizes of the sign processors involved in macro- and microsemiosis (see Table 6.9 below), it is surprising that there exists a

Table 6.9 A comparison between the physical dimensions of the *macrosemiotic* and *microsemiotic agents*. Notice that the linear dimension of the human body is about five orders of magnitude greater than that of the cell (Adapted from Ji 2001)

Parameters	Macrosemiotics	Microsemiotics
1. *Sign processor or agent*	Human body	Cell
2. *Size*	Macroscopic	Microscopic
Linear size (m)	~1	$\sim 10^{-5}$
Volume (m^3)	~1	$\sim 10^{-15}$
3. *Number of cells involved*	$\sim 10^{13}$	1
4. *Signs used for communication*	Words and sentences	Molecules
Linear size (m)	$\sim 10^{-3}$	$\sim 10^{-8}$
Volume (m^3)	$\sim 10^{-9}$	$\sim 10^{-24}$
5. *Mechanics obeyed*	Classical	Classical and quantum
6. *Thermal stability at ~25°C*	Yes (i.e., rigid)	No (i.e., thermally fluctuating)
7. *Powered (or driven) by*	Chemical reactions	Chemical reactions

set of principles that is common to the semiotic processes on both these levels as evidenced by the isomorphism found between human and cell languages (see Table 6.3) (Ji 1997a, b, 1999b, 2001, 2002a). This unexpected finding may be rationalized if we can assume that semiosis, the process of handling *information*, is scale-free, just as the process of handling *energy* are scale-free as evidenced by the universal applicability of the laws of energy and entropy to all structures and processes in the Universe from the microscopic to the cosmological, another evidence supporting *the information-energy complementarity* principle discussed in Sect. 2.3.2.

6.2.5 Three Aspects of Molecular Signs: Iconic, Indexical, and Symbolic

If *macrosemiotics* and *microsemiotics* are isomorphic as asserted by the cell language theory (Ji 1997a, 2001), it may be inferred that the triadic aspects of macrosigns (i.e., signs with macroscopic sizes, Table 6.9), namely, the iconic, indexical, and symbolic aspects (Table 6.8), may also be found in microsigns (or molecular signs). As already indicated in Sects. 6.2.1 and 6.2.3, (1) a *sign* stands for something (called *object or signified*) to someone (interpreter, receiver, or sign processor) in some context (environmental contingencies), and (2) there are three kinds of signs – *iconic* signs (e.g., a statute) related to their objects by *similarity*, *indexical* signs (e.g., smoke) related to their objects by *causality*, and *symbolic* signs (e.g., words) related to their objects by *convention, rules,* and *codes* which are *arbitrary* from the standpoint of the laws of physics and chemistry.

Applying these concepts and definitions to the molecular information processing systems in the living cell, it may be conjectured (1) that DNA serves as the sign for RNA to cells during the transcription step catalyzed by *transcriptosomes*, RNA in turn serving as the sign for proteins during the translation step catalyzed by *ribosomes*, (2) that the relation between DNA and RNA during transcription is primarily iconic (due to Watson-Crick base paring) and indexical (requiring the mechanical energy stored in DNA as *conformons* (Ji 2000) to power orderly molecular motions), and (3) the relation between mRNA and protein synthesized during translation is *iconic* (owing to the complementary shapes of codons and anticodons), *indexical* (requiring *conformons* in the ribosome to drive the orderly movement, or *translation*, of aminoacyl tRNA molecules along the mRNA track), and *symbolic* (due to the *arbitrariness of the relation* between the codons of mRNA and the corresponding amino acids carried by tRNA, i.e., the arbitrariness of the genetic code) (Barbieri 2003, 2008c).

If these conjectures prove to be correct in principle, it would be logical to conclude that biological information processing in the cell cannot be completely characterized in terms of the *laws* of physics and chemistry alone but requires, in addition, the *rules* (e.g., genetic codes) engendered by biological evolution, thus supporting the *von Neumann–Pattee principle of matter-sign complementarity* as applied to biological systems (Pattee 2001, 2008; Ji 1999a, b). In other words, biology is best viewed not as an autonomous science separate from physics and chemistry as some evolutionary biologists assert but a *triadic* science based on *physics, chemistry,* and *semiotics* on equal footings.

6.2.6 Human and Cell Languages as Manifestations of Cosmolanguage

The proposition that the cell possesses its own language, "the cell language," seems almost tautological in view of the fact that cells communicate, since *no communication would be possible without a language*. The natural question that then arises concerns the relation between human language and cell languages. There may be three possibilities:

1. Human language has evolved from cell language.
2. Both cell and human languages are different manifestations of a third language that exists independent of, and serves as the source of, them.
3. Possibilities (1) and (2) are not mutually exclusive but represent the *diachronic* and the *synchronic* manifestations, respectively, of the fundamental characteristics of the Universe we inhabit, namely, that *the final cause of our Universe is to know itself through Homo sapiens*. (Such a Universe was named *the Self-Knowing Universe* or *Universum sapiens* in Ji [1991].)

The author is inclined to accept the third possibility. If this view is true, we are living in the *Self-Knowing Universe* where both cell and human languages exist as diachronic manifestations of a third language which may be referred to as the Cosmological language (or *Cosmolanguage*, for short). By invoking the existence

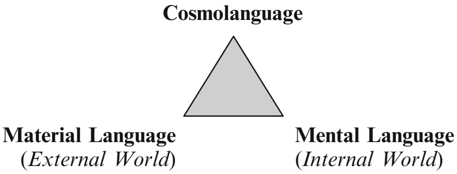

Fig. 6.7 The postulate that the cosmolanguage is manifest in two ways – externally as material language (including cell language) and internally as mental language (exclusive to *Homo sapiens* ?)

of the cosmolanguage, I am in effect postulating that the *language principle* (or more generally *semiotic principles*) applies to all phenomena in the Universe. In Ji (2002a), I expressed the same conclusion as follows:

> *... the principles of language (and associated semiotic principles of Peirce, including rule-governed creativity and double articulation) are manifested* at two levels – at the material level in the external world as well as at the mental level in the internal world. We may refer to this phenomenon as the "principle of the dual manifestations of language or semiosic principles", or the "language duality" for short. Like the wave/particle duality in physics, this matter/mind duality may be a reflection of a deep-lying complementarity which may be identified with the following triad...: (6.26)

Figure 6.7 can be read in two ways – diachronically (or ontologically) as indicating the evolution of the mental and material languages from the cosmolanguage, and synchronically (or epistemologically) as indicating that the material and mental languages are complementary aspects of the cosmolanguage. Both these interpretations are consistent with the model of the Universe called the Shillongator proposed in (Ji 1991). Figure 6.7 may be consistent with Wolfram's *Principle of Computational Equivalence* (Sect. 5.2.1) if we view language, communication, and computation as fundamentally related.

6.2.7 Semiotics and Life Sciences

S*emiotics* and the *science of life* (i.e., biology, agricultural science, and medicine) have had a long and venerable history of interactions (e.g., ancient physicians in both East and West diagnosed the diseases of patients based on symptoms; farmers used cloud patterns to predict weather, etc.), but the connection between *semiotics* and *life sciences* in general may have undergone a significant weakening when the reductionist scientific methodologies were imported into life sciences from physics and chemistry around the nineteenth century. The reductionist trend in physics began with the birth of the mathematically oriented physics following the successful experiments with falling bodies performed by Galileo in the seventeenth century. After over three centuries of domination of physical and biological sciences by reductionism, a new trend seems to be emerging in physics and life sciences that emphasizes *integration* and *holism*, without necessarily denying the fundamental importance of *reductionism* (Elsasser 1998; von Baeyer 2004; Emmeche 2002; Hoffmeyer 1996, 2008; Fernández 2008). As a concrete example of such a new trend, we may cite the isomorphism found between the cell language and the human language (see Table 6.3). One of the major goals of this book is to reveal the deep connection that exists between *life* and *semiosis*, thereby laying the foundation for a *semiotic theory of life*, or organisms viewed as systems of *molecular signs* and *sign processes* (Hoffmeyer 1996).

6.2.8 Semiotics and Information Theory

The study of information may not be successfully carried out without the aid of *semiotics*. This is because information is carried by *signs* (without signs, no information can be generated, transformed, stored, or transmitted) and the study of signs in general is the domain of *semiotics*. Nauta (1972) states a similar view in greater details:

> ...Much work has been done in the field of pure information theory, but the problems concerning the meaning (i.e., semantics vis-à-vis syntactic; my addition) and application (i.e., pragmatics: my addition) of information have largely been neglected. In our opinion, these important problems can be tackled only from a semiotic point of view. The key to these problems will be the analysis of signals, signs and symbols. (Nauta 1972, p. 29) (6.27)

> Semiotics, divided into transmission theory, syntactics, semantics and pragmatics, and subdivided into pure, descriptive, and applied semiotics, offers a general framework for the study of information processes and for the development of a universal theory of information. In its generalized form, semiotics encompasses the following fields: Logistics (artificial symbols) Linguistics (symbols) Semiotics in a narrower sense (signs) Automatics, the study of automatic processes and pre-coded representations and mechanisms (signals). (Nauta 1972, pp. 61–62) (6.28)

Nauta distinguishes three information carriers – “signals,” “signs,” and “symbols” (Table 6.10). He defines signals as carriers of *form* but not *meaning* or *function*; signs as carriers of form and meaning but not of function; symbols as carriers of form, meaning, and functions. This contrasts with Peirce’s division of signs into “iconic signs,” “indexical signs,” and “symbolic signs,” each of which can have form, meaning, and function (Table 6.10).

It is not clear to me why Nauta invoked his triad of information carriers rather than using Peirce’s original sign triad, but it may be possible to represent Nauta’s information carriers as linear combinations of Peirce’s triadic signs. Writing Nauta’s information carriers with capital letters and Peirce’s signs with lower-case letters, we may construct a set of algebraic equations as shown below, where

Table 6.10 Definition of signals, signs, and symbols according to Nauta (1972, p. 159)

	Form	Meaning	Function
Signals	+	–	–
Signs	+	+	–
Symbols	+	+	+

doubly indexed coefficients, a_{ij}, indicate the degree of contribution of Peircean signs to a given information carrier (IC) of Nauta:

$$\begin{aligned} \text{Signal} = \text{IC}_1 &= a_{11}\ \text{icon} + a_{12}\ \text{index} + a_{13}\ \text{symbol} \\ \text{Sign} = \text{IC}_2 &= a_{21}\ \text{icon} + a_{22}\ \text{index} + a_{23}\ \text{symbol} \\ \text{Symbol} = \text{IC}_3 &= a_{31}\ \text{icon} + a_{32}\ \text{index} + a_{33}\ \text{symbol} \end{aligned} \tag{6.29}$$

In general, we may write:

$$\mathbf{A}x = b \tag{6.30}$$

with

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad x = \begin{bmatrix} \text{icon} \\ \text{index} \\ \text{symbol} \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} \text{IC}_1 \\ \text{IC}_2 \\ \text{IC}_3 \end{bmatrix}$$

Equation 6.30 may be viewed as an algebraic expression for the relation between *information theory* (as represented by b) and semiotics (as represented by x) and **A** as the rule of transforming the Peircean semiotics to the information theory according to Nauta (1972).

More recently, Debrock (1998, pp. 79–89) proposed a novel theory of information viewing information as *events* rather than as *entities* and suggested that such a dynamic approach to information may be consistent with the Peirce's theory of signs. Debrock's suggestion seems consistent with the postulate that Peircean signs are gnergons, the source of energy and information to drive all self-organizing processes, including informed events (see Sect. 6.2.3).

6.2.9 *The Cell as the Atom of Semiosis*

The following statement is often made as a useful metaphor:

The cell is the atom of life. (6.31)

In addition, it is asserted here that :

The cell is the atom of semiosis. (6.32)

The term "semiosis" is defined as any physicochemical processes that are mediated by *signs* such as *communication*, *computation,* and DNA-directed *construction.* This triad of processes was referred to as *the C-triad* in (Ji and Ciobanu 2003).

One consequence of combining Statements 6.31 and 6.32 is the corollary that the *cell provides the physical basis and mechanisms for both living processes and semiosis.* A theoretical model of the cell, capable of achieving both these functions, was first proposed in 1983 in an international conference on the Living State held in Bhopal, India, and hence was named the *Bhopalator* (Fig. 2.11) (Ji 1985a, b, 2002b). One of the basic principles underlying the Bhopalator is that of *information-energy complementarity* as manifested in two ways – as *conformons* (conformational strains of biopolymers harboring mechanical energy in sequence-specific sites; see Chap. 8) and as *IDSs* (intracellular dissipative structures such as cytosolic calcium ion gradient; see Chap. 9).

6.2.10 The Origin of Information Suggested by Peircean Metaphysics

In this section, the general problem of the origin of information (including biological and nonbiological) is discussed based on Peirce's metaphysics (Sect. 6.2.2). As is evident in the following quotations, Peirce made a clear distinction between *possibility*, Firstness, and actuality, Secondness (see Table 6.7):

> Possibility implies a relation to what exists. (Hartshorne and Weiss 1931–1935, paragraph #531)
>
> . . .a possibility remains possible when it is not actual (Hartshorne and Weiss 1932, paragraph #42)
>
> . . .possibility evolves the actuality (Hartshorne and Weiss 1932, paragraph #453)
>
> In order to represent to our minds the relation between the universe of possibilities and the universe of actual existent facts, if we are going to think of the latter as a surface, we must think of the former as three-dimensional space in which any surface would represent all the facts that might exist in one existential universe. (Hartshorne and Weiss 1933, paragraph #514)

Feibleman (1946) summarized the essence of Peirce's' distinction between *possibility* and *actuality* as follows:

> Not all *possibles* can exist: *actualit*y is a selection of them.

When I read this statement, especially the term "selection," it occurred to me that Peirce's metaphysics might provide a philosophical foundation for the *origin of information* in this Universe, since information can be broadly defined as resulting from the *selection* of a set of objects, events, or entities from a larger set of them. The formalism is very simple. Let us designate the number of all possibilities (or *possibles* of Peirce) out of which this Universe originated as p, and the number of actual existents (which may be called "actuals") as a. Then the primordial information associated with (or imparted on) this Universe, to be designated as $\mathrm{I_C}$, where C means "cosmological," may be expressed simply as the binary logarithm of the ratio between these two numbers (assuming for simplicity that all *possibles* have equal probabilities of being actualized):

$$\mathrm{I_C} = \log_2 (p/a) \text{ bits} \tag{6.33}$$

Although it is almost impossible to measure or determine p and a (and hence $\mathrm{I_C}$), the mere fact that we can write down a mathematical expression relating these two quantities to the information content of the Universe may be significant.

Equation 6.33 describes only the informational aspect of the origin of the Universe. The energy aspect of the origin of the Universe appears adequately described by the Big Bang theory in physics. That is, the energy requirement for the selection process implicated in Eq. 6.33 is met by the dissipation of free energy (or entropy production in this case, since the Universe is isolated) attending the expansion of the Universe:

$$p \xrightarrow{\text{Entropy Production}} a \tag{6.34}$$

where the arrow indicates that a actuals have been selected out of p possibles (i.e., $p > a$). In Ji (1991), it was suggested that p might be identified with (all possible) *superstrings*, and hence a may now be identified with a subset of p reified into elementary particles constituting all the material entities extant in this Universe. The total number of particles in this Universe has been estimated to be approximately 10^{80}, which is known as the Eddington number (Barrow and Tipler 1986, p. 225). These a actuals are thought to possess sufficient *information* and *energy* (i.e., *gnergy*) to evolve higher-order structures such as atoms and molecules, stars, planets, galaxies, the biosphere, and organisms including humans, under appropriate conditions emergent at specific epochs in the history of the Universe (see Fig. 15.12). It is interesting to note that a similar view was recently put forward by a group of cosmologists (Kane et al. 2000). The biological information encoded in living systems may be viewed as ultimately derived from the Cosmological Information, $\mathrm{I_C}$, through a series of information *transductions*, similar to the well-studied phenomenon of signal transductions occurring in the living cell (Sect. 12.16). If this view of the origin of information is correct, a set of interesting inferences could be made:

1. What happens in this Universe cannot be completely random, including biological evolution. That is, biological evolution may be constrained (or directed) by the cosmological information, $\mathrm{I_C}$, encoded in nonliving material entities (i.e., abiotic matter).
2. All information associated with this Universe may be continuous with (or traced back to) the origin of the cosmological information at the time of and prior to the Big Bang.

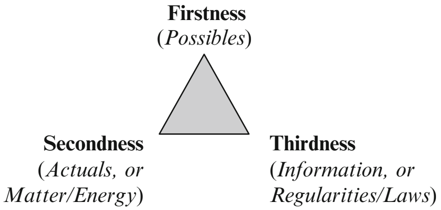

Fig. 6.8 A postulated evolution (or reification) of *possibles* into *actuals* and associated *information* (*and laws*). The nodes are read in the counter clock-wise direction starting from the top node

3. *Possibles*, *Actual*s, and *Information* may reflect the ontological triad of Peirce.

The similarity between Figs. 6.8 and 4.5 may be significant. The similarity may be transformed into an identity simply by equating the Gnergy with the Possibles of Peircean metaphysics, leading to the following conclusions:

Gnergy is the source of possibles out of which all actuals in the Universe are derived. (6.35)

6.2.11 The Triadic Model of Function

The notion of the *structure-function correlation* is widely discussed in biology. In fact, biology may be defined as the scientific study of the correlations between *structure* and *function* of living systems at multiple levels of organization, from molecules to the human body and brain (Polanyi 1968; Bernstein 1967; Kelso and Zanone 2002). The concept of *functio*n is not dichotomous or dyadic as the familiar phrase "structure-function correlation" may suggest but is here postulated to be *triadic* in the sense that a function involves three essential elements – *structure*, *processes*, and *mechanisms*, *all organized within an appropriate boundary or an environmental condition that constrains the processes to perform a function.* M. Polanyi (1891–1976) clearly realized the fundamental role played by boundary conditions in effectuating living processes at the molecular, cellular, and higher levels (Polanyi 1968). A similar idea was expressed by N. Bernstein (1967) at the level of human body movement. Polanyi's and Bernstein's ideas may be expressed in the language of information theory:

$$I_X = \log_2(w_0/w_x)\ \text{bits} \tag{6.36}$$

where I_X is the Shannon information (Sect. 4.3) associated with Function X, w_0 is the number of all possible processes allowed for by the laws of physics and chemistry, and w_x is the number of processes actually selected by the boundary conditions to perform Function X. Equation 6.36 quantitatively expresses the idea that functions are processes selected (or constrained) by appropriate boundary

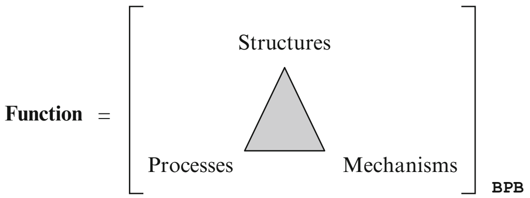

Fig. 6.9 A diagrammatic representation of the triadic conception of function in biology. This diagram presents function as an irreducible triad *of structures*, *processes*, and *mechanisms*. BPB stands for Bernstein-Polanyi boundaries. The boundary-sensitive mechanisms are thought to select only those dissipative structures that perform a desired function out of all possible processes permitted by the laws of physics and chemistry

conditions to perform Function X at a given level of biological organization. For the convenience of discussion, it is suggested that the boundary conditions that constrain and enable Function X to appear from the processes allowed for by the laws of physics and chemistry be referred to as the *Bernstein-Polanyi boundaries* and the information, I_X, embodied in (or needed to specify) such boundaries be referred to as the *Bernstein-Polanyi information*. The Bernstein-Polanyi boundaries (BPBs) reduce the degree of freedom of the components of the system so that they have no choice or freedom but to perform the motions or movements that constitute a function at a given level of organization. Thus, boundaries, constraints, and reduced degrees of freedom are all synonymous terms referring to a function (Polanyi 1968; Bernstein 1967). The triadic conception of function can then be diagrammatically represented as shown in Fig. 6.9:

One advantage of Fig. 6.9 is that it provides a geometric template to organize the four terms that are obviously related with one another, that is, *function*, *structure*, *process,* and *mechanism.* It may be significant that the triadic definition of a function given in Fig. 6.9 is isomorphic with the triadic definition of a sign given by Peirce (1839–1914) (see Fig. 6.2) and consistent with his metaphysics that all phenomena comprise three basic elements (Sect. 6.2). Table 6.11 lists various examples of functions in biology and their triadic components.

6.2.12 The Principle of Prescinding

The term “prescinding” refers to our mental capacity (or tendency) to focus on some aspect of a phenomenon, process, or structure in exclusion of other aspects for the convenience of thought. For example, according to the diagram shown in Fig. 6.10, “function” comprises three irreducible aspects, “structure,” “process,” and “mechanism.” However, it is often convenient to focus on one of these aspects of function and discuss the structure-function, the process-function, or the mechanism-function correlation for the purpose of convenient analysis and communication. Which of these

Table 6.11 Examples of various functions and their elements in biology

Function	Structure	Process	Mechanism
1. Transcription	DNA template	RNA polymerization	RNA polymerase driven by conformons[a]
2. Translation	mRNA, tRNA, rRNA	Peptidyl transfer reaction	Directed movement of the ribosome components driven by conformons
3. Amino acyl tRNA synthesis	tRNA anticodons	Amino acylation of tRNA	Allosteric control of amino acylation by tRNA anticodon
4. Protein folding	Amino acid sequence	Rate of translation	Environment-sensitive protein conformation
5. Enzymic catalysis	Protein folds	Chemical reactions	Conformon-driven regulation of the activation energy barrier
6. Semiosis	Representamen (or signifier, sign vehicle)	Object (or signified, referent)	Interpretant (or codemaking, mapping, habit-forming, evolution)

[a] Conformons are the mechanical energy stored in sequence-specific sites within biopolymers that are generated from exergonic chemical reactions and drive all orderly molecular motions inside the cell including enzymic catalysis, molecular motors, pumps, rotors, and chromatin remodeling (see Sect. 8.1)

Epistemology

‖

[**Innate Knowledge** encoded in **Cell Language**]

↓

[**External Knowledge** expressed in **Human Language**]

Fig. 6.10 Epistemology as the study of the rules governing the conversion of cell-linguistic texts into human-linguistic texts, and as a cell language-based interpretation of Socrates' doctrine of *anamnesis* and of the theory of *pre-reflective experience* of Merleau-Ponty (Dillon 1997, pp. 1–2)

correlations is chosen for discussion would depend on the context of the discourse at hand and on the perspectives of the discussants involved. Peirce (1868) views "prescision" as a form of "abstraction" (Colapietro 1993) and characterizes it in connection with the related terms "discrimination" and "dissociation" as follows:

> The terms "prescision" and "abstraction," which were formerly applied to every kind of separation, are now limited, not merely to mental separation, but to that which arises from attention to one element and neglect of the other. Exclusive attention consists in a definite conception or supposition of one part of an object, without any supposition of the other. Abstraction or prescision ought to be carefully distinguished from two other modes of mental separation, which may be termed discrimination and dissociation. Discrimination has to do merely with the senses of terms, and only draws a distinction in meaning. Dissociation is that separation which, in the absence of a constant association, is permitted by the law of association of images. It is the consciousness of one thing, without the necessary simultaneous consciousness of the other. Abstraction or prescision, therefore,

> supposes a greater separation than discrimination, but a less separation than dissociation. Thus I can discriminate red from blue, space from color, and color from space, but not red from color. I can prescind red from blue, and space from color (as is manifest from the fact that I actually believe there is an uncolored space between my face and the wall); but I cannot prescind color from space, nor red from color. I can dissociate red from blue, but not space from color, color from space, nor red from color.

6.3 Philosophy

6.3.1 The "Five Causes Doctrine"

The *Four Causes Doctrine* of Aristotle (384–322 B.C.) consists of the first four items listed below but lacks the last cause which is here thought to be essential to be taken into account in view of the fundamental discoveries made in evolutionary and developmental biology since Darwin (Waddington 1957; Gerhart and Kirschner 1997; West-Eberhard 2003; Carroll 2006; Jablonka 2006, 2009). I elected to refer to this new addition to Aristotle's list of the four causes as the "original" cause but it could as well be referred to as a "historical" cause:

1. The *material cause* (What is it made out of?)
2. The *efficient cause* (How does it work?)
3. The *formal cause* (What is it?)
4. The *final cause* (What is it for?)
5. The *original cause* (Where does it come from? How did it originate?)

The five causes doctrine asserts that no structure, process, or phenomenon can be said to be completely known or understood until and unless the associated five causes have been determined and described. In other words, no theory, especially biological ones, would be deemed complete unless the five causes described above are fully detailed.

It may be objected that it is not necessary to invoke the fifth cause because it can be viewed as a part of the efficient cause. I would agree to such an opinion under one condition – recognizing two kinds of the efficient causes – (1) rapid and (2) slow. The *rapid efficient cause* takes effect in times shorter than the lifetime (or cycling time) of the system under consideration, and the *slow efficient cause* act over times much longer than the lifetimes of individual systems. The former may be referred to as the *synchronic* efficient cause and the latter as the *diachronic* efficient cause (see related discussions in Sects. 4.5 and 6.3.2). Alternatively, we can adopt the five causes and associate the efficient cause with the synchronicity (and developmental biology) and the fifth cause with diachronicity (and evolutionary biology).

The Four Causes Doctrine of Aristotle was formulated over 2,000 years ago based on abstractions from everyday human experiences in the macroscopic world. Since then two major developments have taken place in human knowledge – (1) the development of the theory of biological evolution formulated by Darwin in 1859 and (2) the development of quantum mechanics that began with the discovery of quanta by M. Planck in 1900. Therefore, it would not be too surprising if it is found that the Four Causes Doctrine cannot be extended to the contemporary biological sciences and physics without updating. The addition of the fifth cause to the *Five Causes Doctrine* may be viewed as a natural consequence of taking into account of the modern theory of biological evolution, and a further modification may be necessary in order to take cognizance of the quantum revolution of the twentieth century. In what form this predicted modification of the Four Causes Doctrine should take is not yet clear.

6.3.2 The Principle of Closure

When two entities, A and B, need each other for their own existence, so that without A, B cannot exist or function, or vice versa, A and B can be said to exhibit or embody the *principle of closure*.

Semantic Closure. The principle of closure defined above was inspired by, and is a generalization of, the concept of "semantic closure" or "semiotic closure" formulated by H. Pattee (1995, 2001) who characterized *semantic closure* as follows:

> …this complex interrelationship of strong and weak bonds… that allows the realization of von Neumann's quiescent *symbolic description* and *dynamic material construction*.

The Principle of Ontological and Epistemic Closure. Before the cell can read DNA, an *epistemic* act, the cell must have been endowed with such a capability through evolution, an *ontological* process. Before the human brain can know anything, an *epistemic* act, it must have been endowed with the knowing capability through biological evolution, an *ontological* act. These statements are consistent with the pre-fit hypothesis of ligand–protein interactions, including enzymic catalysis (Sect. 7.1.3), which in fact may provide the molecular rationale for their validity. As will be discussed in Sect. 7.1.3, the pre-fit hypothesis is rooted in the generalized Franck–Condon principle imported from quantum physics and well supported by recent experimental findings in molecular biology Kurakin (2009). It seems to me that there is a real possibility that the pre-fit hypothesis can rationalize, in molecular terms, Kant's *Copernican Revolution in Philosophy (CRP)*, namely, the claim that objects conform to our knowledge rather than the other way around (http://en.wikipedia.org/wiki/Copernican_Revolution_(metaphor). That is, our *knowledge* is pre-fit to *objects* just as *enzyme active sites* are pre-fit to their *substrates and products* (see Fig. 7.2). We may refer to this idea as the *Copernican Revolution (CRB) in Biology* and suggest that *CRB and CRP belong to the same category of paradigm shifts* and further that *CRB underlies CRP*.

The Diachronic and Synchronic Closure. Although humans can use a language without knowing its past history which is related to the *synchronicity* emphasized by Saussure (Culler 1991), a language cannot be effective as a means of communication among members of a community without its long history of development and evolution (*diachronicity*). Biologists can describe all the physics, chemistry, and biochemistry of the living cell (*synchronicity*), but it would be impossible for them to understand the workings of the cell without taking into account the long history of biological evolution (*diachronicity*). *This is because the synchronic properties of the cell* (e.g., ligand–receptor interactions, cell cycle, chemotaxis, etc.) *are needed for its evolution and the evolutionary process is needed for the emergence of such synchronic properties endowed with genetic information.*

The Closure Relation Between Boundary Conditions and the Dynamics of Physical Systems. It is clear that no physical laws can describe any observable properties without *there* being specific boundary conditions. In other words, the equation of motion describing a physical system, which embodies laws of physics, cannot be solved without the initial and boundary conditions applicable to that system (Pattee 1995).

6.3.3 The Anthropic Principle

Cosmologists have found that the numerical values of the fundamental physical constants such as c (speed of light), G (gravitational constant), h (Planck constant), e (electronic charge), m_e (electron mass), and m_n (neutron mass) must be precisely what they are in order for our Universe to evolve to contain those elements (e.g., carbon, nitrogen, iron, etc.) that are essential for life to exist on this planet (Barrow and Tipler 1986; Kane et al. 2000). Deviations by even a few percent from these values have been found to lead to alternate model universes devoid of carbon atoms, for example. Therefore, it is clear that there is a *closure relation* between the existence of life (A) in this Universe and the numerical values of the key physical constants (B) that characterize the structure of this Universe: that is, Without B, no A; or A presupposes B. In this view, the anthropic principle is a species of the principle of closure (Sect. 6.3.2).

6.3.4 The Table Theory

It appears that we acquire our knowledge about an unfamiliar object only in terms of what we already know, reminiscent of the biological principle that organisms derive from preceding organisms (except when life first originated), and ligands bind only to those receptors that are pre-fit to their shapes (Sect. 7.1.3). We may state this idea as follows:

> Our knowledge about an unknown object can be increased only in terms of the properties of an already familiar object. (6.37)

Statement 6.37 may be related to Socrates' (ca. 470–399 B.C.) *doctrine of recollection*, or *anamnesis*, according to which knowledge can only come from recollection. A similar idea was advanced by the French phenomenologist Maurice Merleau-Ponty (1907–1961) in his theory of "pre-reflective experience": (Dillon 1997, pp. 1–2). An intriguing possibility to account for the phenomenon of

Table 6.12 A formalization of the "table theory" or an analogical inference. *Internal* and *external relations* may also be referred to as "intra-" and "inter-system" relations

Parameters		F (familiar)		U (unfamiliar)
1. Components: $f_1, f_2, f_3, \ldots, f_n$			***AR***	
$u_1, u_2, u_3, \ldots, u_n$		$\mathbf{f_1}$	<——>	$\mathbf{u_1}$
2. Relations:	***IR***	↕		↕
↕: Internal (or intra-system) Relation (***IR***)		$\mathbf{f_2}$	<——>	$\mathbf{u_2}$
<——>: Analogical (or inter-system)		↕		↕ (?)
Relation (***AR***)		$\mathbf{f_3}$	<——>	$\mathbf{u_3}$
3. The "Table Symmetry Principle":		↕		↕
(1) If **F** and **U** are *isomorphic*,		.		.
(2) if ***IR***(f_i, f_{i+1}) is known,		.		.
(3) if ***AR***(f_i, u_i) is known, and		.		.
(4) if ***AR***(f_{i+1}, u_{i+1}) is known, then		↕		↕
(5) ***IR***(u_i, u_{i+1}) = ***IR***(f_i, f_{i+1})		$\mathbf{f_n}$	<——>	$\mathbf{u_n}$

anamnesis may be opened up by the postulated isomorphism between cell and human languages (Table 6.3). Because we are made up of cells which are in turn made up of material entities originating in nature, we may already *know* how cells and nature work by virtue of the communication mediated by cell language between the human brain and its constituent cells, although we may not be able to express this knowledge for the purpose of communication among humans because it is not encoded in human language. To do so, we must convert our *innate* (or *internal*) knowledge encoded in cell language (which may be identified with "pre-reflective experience" of Merleau-Ponty) into what may be called the *external* or *objective* knowledge expressed in human language, and this postulated process of *cell-human language transduction* (or *translation*) may constitute the heart of epistemology. The language mediating the communication between cells (C) and humans (H) may be referred to as *the CH language*, distinct from human language (which may be called *the HH language*) and cell language (*the CC language*). Through CH language, humans may be able to communicate with the Universe itself, since cells are the embodiment of the laws of nature and the historical record of the Universe. We may represent this series of ideas diagrammatically as shown in Fig. 6.10.

In Ji (1991), the essence of the above ideas was formalized under the rubric of "table theory." The term "table" is employed here, because the theory utilizes a 2-D table as an essential graphical tool for comparing the properties of a familiar (F) object with those of an unknown or unfamiliar (U) object. The *table theory* has three key elements – (1) sets of components or nodes for F or U; (2) two kinds of relations, internal relation (IR) and analogical relation (AR) (IR and AR may also be referred to as "intrasystem" and "intersystem" relations, respectively); and (3) the principle of *table symmetry* stating that, if F and U are isomorphic (i.e., obey a common set of principles), IR of U can be inferred from the IR of F given that AR exists between the components of F and U. These ideas are summarized in Table 6.12. The main objective of comparing two objects, F and U, is to discover the *relations* among the components of U (i.e., the vertical arrows among the u's;

see the question mark next to the vertical double arrow by using the relations among the components of F based on the assumption that a set of similarity relations (see the horizontal arrows) holds between components of F and the corresponding components of U.

The cell language theory (see Table 6.3) formulated on the basis of the similarity between the properties of living cells and those of human language may be viewed as one of the most detailed examples of the application of the "table theory."

6.3.5 The Principle of Mӧ̋bius Relations

The Möbius strip is "a one-sided surface that is constructed from rectangle by holding one end fixed, rotating the opposite end through 180°, and applying it to the first end" [*Webster's Ninth New College Dictionary*]. The essential geometric properties of the Möbius strip may be characterized in terms of the following two propositions:

The Möbius strip consists of two opposite surfaces, A and B, when viewed locally. (6.38)

Surfaces, A and B, merge into one another when viewed globally. (6.39)

Statements 6.38 and 6.39 may be combined into one:

Locally A **or** B; globally A **and** B. (6.40)

Statement 6.40 may be viewed as an alternative expression of what is referred to as the *global–local complementarity* (or the *forest-tree complementarity*). In all these statements, the terms "local" and "global" may be replaced with "synchronic" and "diachronic," respectively. For the definitions of "synchronicity" and "diachronicity," see Sects. 4.5 and 6.3.2.

6.3.6 The Pragmatic Maxim of Peirce

According to Peirce, the meaning of a word or a concept can be equated with the totality of the practical effects or consequences that the word has in life:

> In order to ascertain the meaning of an intellectual conception one should consider what practical consequences might conceivably result by necessity from the truth of that conception; and the sum of these consequences will constitute the entire meaning of the conception. (Goudge 1969, p. 153) (6.41)

This maxim will become useful in defining the meaning of life in Sect. 16.1.

6.3.7 A New Architectonics Based on the Principle of Information-Energy Complementarity

Architectonics is the science of systematizing all knowledge. We can recognize three distinct types of sciences. By "science" is meant a system of human knowledge organized according to some rules. Also I am including matter within the concept of energy, since energy and matter are interconnected and interconvertible through $E = mc^2$ (Shadowitz 1968).

1. "Energy science" or *energetics*, the study of energy transformation, transmission, and storage in the Universe, for example, physics, chemistry, quantum mechanics, thermodynamics, chemical kinetics, Newtonian mechanics, statistical mechanics, etc.,
2. "Information science" or *informatics*, the study of information transduction, transfer, and storage, for example, linguistics, computer science, informatics, logic, mathematics, etc., and
3. "Energy-Information science" or *gnergetics*, the study of goal-directed or teleonomic processes driven by gnergy, for example, cosmology, biology, cognitive science, philosophy, and religion.

According to complementarism (see Sect. 2.3.4), all irreconcilable opposites can be viewed as complementary aspects of a third which transcends the level where the opposites are recognized or have meanings. Thus, if we view information and energy as irreconcilable opposites, then there must be a third for which energy and information are complementary aspects. This third entity was given the name "gnergy" in the mid-1980s. We can express the same ideas algebraically thus:

$$\text{Ontology} : \text{Gnergy} = \text{Gnon}^{\wedge}\text{Ergon} \tag{6.42}$$

$$\text{Epistemology} : \text{Gnergetics} = \text{Gnonics}^{\wedge}\text{Ergonics} \tag{6.43}$$

where the symbol, ^, denotes the *complementary relation* of Bohr, generalized in complementarism in terms of three complementarian logical criteria (see Sect. 2.3.3).

Equations 6.42 and 6.43 provide the logical basis for equating "Energy-Information science" with gnergetics, or gnergy science, in (3) above. We can summarize all these ideas in the form of a diagram (Fig. 6.11):

The tree types of sciences described in (1) through (3) above and those described in Fig. 6.11 are related as follows:

Energy Science = Energetics (or Ergonics)
Information science = Informatics (or Gnonics)
Energy-Information Science = Gnergetics (or Gnergonics)

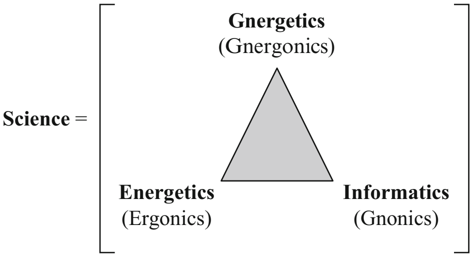

Fig. 6.11 The trichotomization of sciences based on the metaphysics of complementarism (Sect. 2.3.4). Gnegons, ergons, and gnons are discrete units of gnergy, energy, and information, respectively

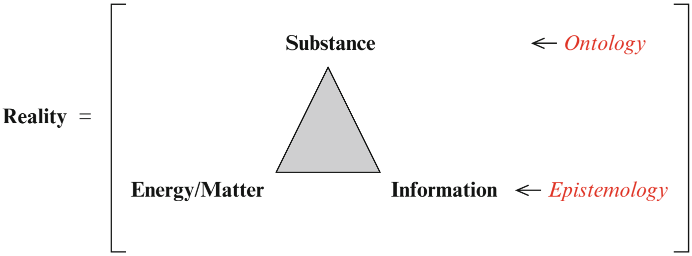

Fig. 6.12 The "triadic theory of reality" (TTR). TTR organizes seven fundamental concepts in physics and metaphysics (i.e., *energy, matter, information, epistemology,* and *ontology, Substance,* and *Reality*) within a coherent conceptual framework or network rooted in the Peircean theory of signs as depicted in Fig. 6.2 (see Sect. 6.2.1)

This is the systematization (or the architectonics) of knowledge suggested by the information/energy complementarity principle (Sect. 2.3.2).

6.3.8 The Triadic Theory of Reality

Wheeler (1998) presupposes the existence of the questioner or questioners who ask a series of binary questions, the answers to which are thought to constitute knowledge or science. Since the questioner came after the Universe and there must have been something to question about, it seems logical to infer that ontology precedes epistemology. Both Wheeler (1998) and Bohr (1933, 1958; Murdoch 1987; Pais 1991) seem to be concerned primarily with *epistemology*, taking ontology for granted.

If this analysis is right, it may be necessary to go beyond Wheeler and Bohr by extending the Bohr's principle of complementarity to *ontology*, and one way to do this may be to integrate all the major concepts discussed in philosophy into a logically coherent system using the triadic diagram discussed in Sect. 6.2.1.

TTR shown in Fig. 6.12 may be regarded as the periodic table of philosophy and a geometrical version of complementarism, according to which the ultimate reality is the complementary union of irreconcilable opposites. It should be noted that there are three kinds of complementarities embedded in Fig. 6.12.

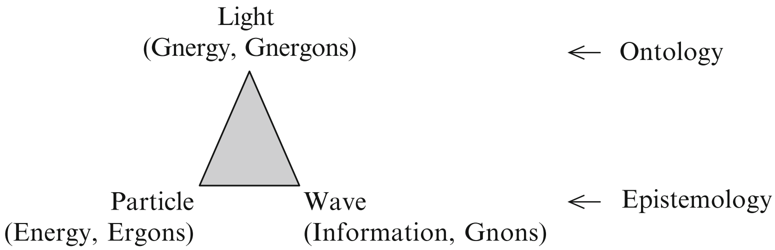

Fig. 6.13 Light as gnergy or the complementary union of information (gn-) and energy (−ergy). Gnergons are defined as discrete units of gnergy; ergons are discrete units of energy; and gnons are discrete units of information. The double labeling of each node is necessitated by the continuity-discontinuity complementarity

1. The "horizontal complementarity" = Energy/Matter and Information (or Life-Information, Liformation; see Table 2.6) as the complementary aspects of Substance.
2. The "vertical complementarity" = Ontology and Epistemology as complementary aspects of Reality.
3. The "complementarity of complementarities" = The "horizontal" and "vertical" complementarities as complementary aspects of Reality.

TTR described in Fig. 6.12 suggests that Wheeler is primarily concerned with what is here called the *horizontal complementarity* without acknowledging the third term, *Substance*, and Bohr may have glimpsed both the *horizontal* and *vertical* complementarities when he stated that "Contraries are complementary," since this statement can be shown to be recursive on the concept of complementarity (Ji 2008b).

TTR may shed a new light on the long-lasting debate between Bohr and Einstein on the interpretation of quantum mechanics (Murdoch 1987; Petruccioli 1993; Cushing 1994; Johansson 2007). For example, the reason that we observe the wave property (interference patterns) and particle property (photoelectric effects) of light (and other quantum entities generally referred to as quons [Herbert 1987]) depending on the measuring apparatus used may be because

Light is BOTH waves and particles even before it is measured. (6.44)

In my lexicon, light as directly observed by the human eye (before measurement) corresponds to gnergy which appears as either waves or particles depending on which measuring apparatus light has gone through. This situation can be described diagrammatically as shown in Fig. 6.13:

Two observations seem to support the triadic interpretation of light:

1. A stream of electrons arriving at a screen past a hole one at a time produces *the Airy pattern* characterized by a set of dots that form concentric circles, the dots

reflecting the particle property and the circular patterns indicating the wave property of the electron wave (Herbert 1987, p. 62)

2. The de Broglie equation quantitatively relates the wave property (wavelength) and particle property (momentum) of a quon:

$$(\text{Wavelength}) = (\text{Planck constant})/(\text{Momentum}) \tag{6.45}$$

One simple interpretation of both the Airy pattern and de Broglie equation is to assume that light (or quons) is both waves and particles at the same time, as Bohm has been advocating throughout his career following de Broglie's idea.

If the above analysis is correct, *the Bohr-Einstein* and *Bohr-Bohm* debates may have a possible solution:

> Both Bohr and his opponents are right in one sense and wrong in another, because their arguments are on two separate levels of reality – Bohr on the epistemological level and his opponents on the ontological level. To the extent that Bohr (his opponents) confined his (their) argument(s) to the epistemological (ontological) level, Bohr's (his opponents') argument(s) is (are) valid. To the extent that Bohr (his opponents) asserted the validity of his (their) argument(s) beyond the epistemological (ontological) level, his (their) argument (s) loses legitimacy. (6.46)

6.3.9 The Type-Token Distinction

Philosophers distinguish between *types* and *tokens*. A type is an abstract object that does not exist anywhere but tokens are concrete realizations of a type that exist physically somewhere at some time. For example, the living cell is a type but the cells constituting my body are tokens of the living cell. At least two examples illustrating the type-token distinction appear in this book. *Gnergy* is a type and *gnergons* are its tokens (Sect. 2.3.2). *MAPK* is a type and its tokens include *JNK, ERK1/2,* and *RK* proteins in the MAPK signaling cascade discussed in Sect. 12.16. Seger and Kreebs (1995) refer to MAPK as a "generic name" while JNK, ERK ½, and RK as "specific names" (see the margins of the table shown in Fig. 12.35).