Chapter 12
Biological Information and Dormancy in Aquatic Population Modeling

Victor R. Alekseev

Abstract The importance of biological information in life activity of organisms and superorganisms at the overlying levels is not disputed by anyone, but it remains an area with scant investigation. There are several reasons, starting with the blurriness of the very concept of information and ending with methodological difficulties in its quantitative data tallying and measurement. Diapause is one of the most studied adaptations based on the detection of information signals about the state of the environment and its upcoming seasonal reconstructions and seems to be one of the most convenient bioinformational object for studying at this stage of development of this branch of science. The very significant progress of object-oriented modeling in recent decades provides a methodological basis for such studies. Different aspects of Shannon's theory of information in biological systems are also discussed.

Keywords Shannon's theory of information · Use of biological information · Effectiveness of biological functions · Shannon index, Informative links · Evaluation of biological information · Biological informative system

12.1 A Brief Review of the History of Bioinformation Research

The need to study the flow of information, along with the flow of energy and matter, was proclaimed at the beginning of the modern phase of the study of aquatic ecosystems (Khailov 1966). Major environmental programs have called for the study of theses flows but in most specific content remain empty, leaving biologists without firm foundations to follow-up research.

What is the reason for such long-lingering futility at one of the most important directions in the environment, assessing the state of organisms and its ability at reconstruction by neighboring activity? And what is the actual locus of information

V. R. Alekseev (✉)
Zoological Institute of Russian Academy of Sciences, St. Petersburg, Russia
e-mail: alekseev@zin.ru

links in superorganismic biosystems? In this paper, we propose one option for answering these questions.

The peak of interest in problems related to the assessment of the role of information in biological suborganismic systems occurred in Russia and abroad in the 1960s and 1970s (Shmalhausen 1968; Seravin 1973; Naumov 1973; Setrov 1975, etc.). By the beginning of the 1980s, the amount of work related to this topic decreased significantly, mainly because of limitations of starting theoretical assumptions. Turning to the origins of the bioinformation "boom," one should point to cybernetics (Wiener 1961), the systemic approach (von Bertalanffy 1969) and, of course, Shannon's mathematical theory of communication of optimal coding (1948). If the first two theories related to information would largely apply to descriptive constructs, pointing at the significance and the place of information links in the management of complex systems, Shannon's constructs allowed to quantify the quality of information for biology.

The famous Shannon index served as a quantitative basis for science—unprecedented in the breadth of application and the alleged effect of an explosive search for information links in highly divergent areas of science: from physics to linguistics. The Shannon index is extremely engaging in its simplicity for assessing information:

$$H = -\sum_{t=1}^{t=n} P_i * \log_2 P_i$$

where P_i is the probability of an event.

In addition to that, Shannon proposed several theorems describing the patterns of information transfer via communication channels from source to receiver; Shannon also considered the mathematical Boltzmann's H-theorem for thermodynamics and adapted it to the needs of information research. Among the most important contributions of Shannon, we find also strict descriptions of communication channels, sources of interference, and reliability criteria for communication.

However, having disregarded the meaning of the message being transmitted in the quantitative measure (H), Shannon posed a challenge to his apologists, one that has not been solved until now, as it appeared to be insoluble in principle. A wide recognition of that fact did not come immediately to the fore, including in the field of biology.

It should be said that the field biology is largely conservative in relation to constructs and models, which is also said of or even more so in applied sciences. Nevertheless, an indisputable impetus for biologists to search for and find meaningful information links in biosystems at various levels happened in 1953 with the decoding of the structure of the biological DNA code by Watson and Crick. This example, now part of students' textbooks, of the leading role of information in the implementation of such a determining function of the organism as it is reproduction,

made an extremely strong impression and gave rise to numerous followers in different biological disciplines. However, to the great disappointment of researchers on the role of information in biology, the situation with the genome was the exception rather than the rule. As the organization of biological systems becomes more complex, the probabilistic (balancing) principle of organizing information links comes to replace the rigid rules of the information bank (DNA). This is due to the fact that the internal environment of the organism, as well as its structure, turned out to be a more rigid system than the one that develops in superorganismic structures. Thus, it is very difficult to program and implement all possible variants of the relationship of the organism with the environment (as opposed to the stages of embryonic development, for instance). It became apparent that the question of the role of information in the biological systems at the superorganism level is decided differently than in the underlying systems.

Moreover, even among the followers of Shannon, voices began to be heard more and more loudly that although the theory of information is so broad that it can be applied to any subject, it is nevertheless advisable to apply it in cases where the decisive factor is the stream of communication rather than the meaning of the message being transmitted (Ichas 1960).

In other words, Shannon's information theory is not applicable to assess the role of information in biological systems, but only to study and evaluate the effectiveness of its transmission in coded form. In this case, this theory has a methodological rather than a conceptual significance. A serious analysis of this aspect was made by L. Seravin (1973) and M. Setrov (1975).

12.2   Universal Language and Information Flows in the Ecosystem

Attempts to apply the theory of optimal coding for studying the role of information in biological systems, including complex ones such as biocenoses, did not stop. The idea of an ecological code—one for the ecosystem—was put forward and actively discussed (see Levich 1983 for review). It was assumed that there is a language (ecocode) that can be deciphered, in which information is exchanged by all participants in the ecosystem. The study of the ecocode was supposed to be similar to mastering language. The scheme of understanding the ecocode by the researcher, in principle internally consistent, was constructed with the involvement of the logic of formal linguistics.

Among the texts of the ecological code were included such indicators as abundance and species composition of planktonic organisms. Mandelbrot (1972) proposed to equate communities with texts, types with words, individuals of one species with word forms, etc. Levich (1983), using his version of analogies, equated indi-

viduals with letters, one-time communities with words, and succession of communities with text. All these analogies gave their authors formal grounds to draw as mathematical patterns of numbers obtained in the analysis of textual materials, in particular Zipf's law (Zipf 1949):

$$n(i) \sim \frac{1}{i}$$

where *i* is the rank of the word in the ranked series and *n(i)* is the occurrence of a word of rank *i*.

Despite the apparent dubiousness of such associations for the biologist, certain grounds for using regularities like Zipf's law (1949) have proven useful. Even before Zipf's law was published, Willis (1922) used a rank distribution for the analysis of the fauna of Ceylon and derived quite convincing material concerning the distribution of species in isolates. It is quite obvious that these studies are grounded on general statistics laws underlying the distribution of a wide range of objects that obey the law of large numbers in probability theory. It could be said that authors overlooked the notions of information significance of the rank-size distribution, and this would also be true for the inhabitants of a biocenosis. There is almost no doubt that organisms are completely deprived of both the ability and the biological equipment for analyzing the species structure of their ecosystem, as well as using information that could be obtained through them. But the fact that the ecocode should be understood not only by the researcher but also by the biota's organisms follows from the definition of ecocode and how it is perceived through collected or derived data.

Thus, we have come to another fairly common misconception regarding the evaluation of the role of information in biological systems inferred on the basis of probabilistic regularities of indices and indicators and their identification (including the usage of Shannon index). We also tend to draw conclusions from the analysis of indices purportedly to obtain biological information, but rather with the category of problems of studying the role of information in biological systems. In the best case, such studies allow us to build a mathematical apparatus developed in, and for, other fields of science detecting all kinds of diversity and heterogeneity in biological systems. These techniques are likely to be ultimately unsuitable to accommodate the role of biological information as regards to the structures and functions of biological systems.

The question arises: Which way is more productive? To answer this, it is necessary to determine what exactly gives information to the individual of the population and the biocenosis. First, the advantage in performing a particular function that will life expectancy, i.e., again. The question of the usefulness of information for the organism is deemed as an important one because, only in this case, it is possible to develop appropriate adaptations fixed either in a series of generations (e.g., be it instincts or biorhythms) or during the lifespan of an individual (conditioned reflexes, to mention one). So what is the measure of usefulness of information can be used for the organism? L. Seravin (1973) proposed two elementary indicators in this

capacity: the real gain $C_p = P_2 - P_1$ and the relative gain $C_o = P_2/P_1$, where P_2 and P_1 are the probabilities of the event, respectively, with and without the information. Enduring systems, including biological ones, are designed for maximum efficiency; therefore, the value of information increases for unlikely events and decreases for probable ones.

Thus, the main reason (and therefore the criterion for researchers) for information links can be considered limited or uneven spatial-temporal distribution of any of the factors (resources) affecting the success of the vital functions. Previously, a number of authors have addressed some striking examples of the participation of information in the implementation of such functions (Seravin 1973; Zelikman 1977; Levich 1980, 1983) which enables us to infer an approximate structure in information links of the organism as an element of the population and an element of the superorganismic system (Fig. 12.1). Almost all the communication channels known for biological objects can act as carriers of information in providing each of the functions in question: chemical (such as metabolites, attractants, etc.), optical (such as visual image, photoperiod), and mechanical vibrations of the environment, electromagnetic waves (Zelikman 1977). Interaction between individuals of one population can be considered as a sort of multinodal information network covering the habitat of the population and increasing the probability of survival of any of the members of the network. In the diagram, this is indicated by the selection of the plane of intrapopulation interactions.
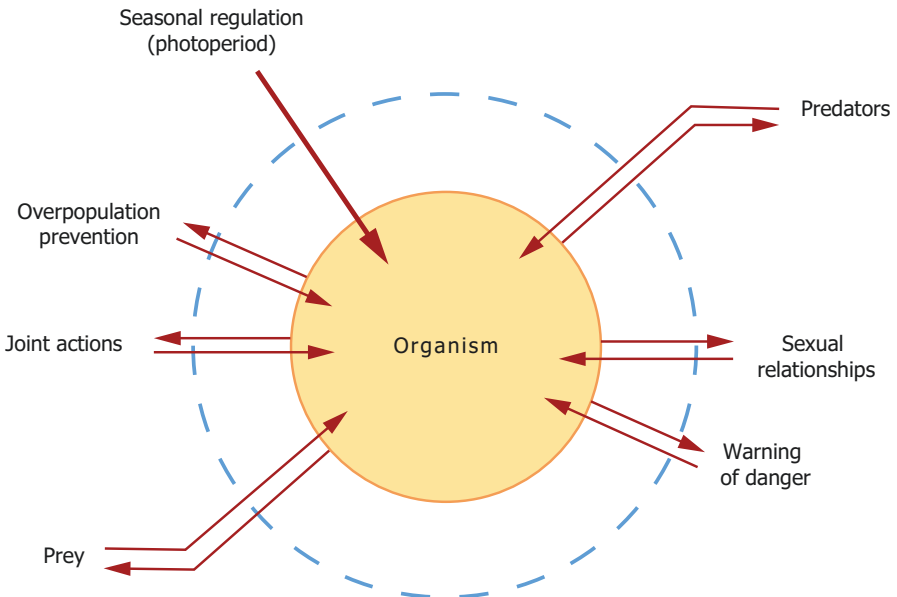

Fig. 12.1 Schematic diagram of vertical and horizontal information links of the organism. Dashed line shows the plane of intrapopulation links

The relationship between the higher and lower levels should be considered as an element of the "pyramid of information links," which in some ways resembles a trophic system. Since one type food feeding and control by the same predator is very rare in nature, vertical information interaction is carried out in a multidimensional space. In our opinion, vertical and horizontal information communications in many respects give concrete outlines and certain stability to superorganismal structures.

Returning to the question of the biological language of ecosystems (i.e., Levich's ecological code), it should be said that it was not possible to detect it due to the fact that the original premises did not correspond to reality. In real-world ecosystems, instead of just one universal language, the "Babylonian confusion" of extremely simple dialect functions, which is understandable to the circle of organisms, is limited by information links. Each organism in search of food and trying to evade a predator is constantly forced, as a first task, to collect information about these factors in order to increase the effectiveness of actions. During periods of reproduction or collaborative interactions (hunting, collective protection in a pack, etc.), individuals send and perceive information that is understandable only to those to whom it is addressed, that is, it does not go beyond members of the same population co-inhabiting in the same biotope. Exceptions are rare and refer to species with similar niches (Alekseev 1990).

Of the few languages that can claim to be universal because of direct connection with such a generally necessary adaptation, as the seasonal adjustment of communities, the photoperiod and temperature can be considered very simple codes. The perception of specific photoperiod values in the binary system (day/night) is largely species-specific. Nevertheless, the photoperiod and partly the temperature, passing and acting approximately unambiguously, causing the completion of the active part of the life cycle and leading to the formation of various kinds of biological dormancy among ecosystem members, can claim the role of flows of biologically significant information penetrating through a dense multidimensional multi-member information network of organisms of the whole ecosystem. In this sense, the factors of transition to biological dormancy can be considered as flows of bioinformation.

12.3 Features of the Study of the Role of Bioinformation in the Aquatic Environment

Specific study of information links should be based on the structure of the information system (Fig. 12.2). Problems arising in the study of information links can be grouped in accordance with the elements of the information system: (a)

Fig. 12.2 Schematic diagram of the information system. (Modified from Shannon 1948)

identification of the information source and destination; (b) establishment of the type of communication channel (chemical, optical, etc.); (c) identification of the code (chemical formula, frequency of oscillations, photoperiod); (d) evaluation of the upper and lower thresholds of signal perception; (e) determination of the range and noise immunity of the communication system; (f) clarification of the form of response to the signal; and (g) study of the mechanism of the transmitter and receiver of signals.

The works of the S. Shvarts school, on amphibian larvae (Schwarz et al. 1976), and the Leningrad school of entomology (Danilevsky 1961), on terrestrial insects, can serve as good research examples of this kind of information links, even though they reached there from their own path each from others. In addition, some environmental studies of energy and substance flows, in particular those studying the efficacy of predator nutrition, do not always include an explicit evaluation of information parameters (Ivlev 1961; Krylov 1989). Such indicators, which are of interest from the point of view of studying information, may include the radius of the predator reaction (assessment of the ability of nontactile detection), the proportion of successful predator attacks, and the assessment of the ability of the victim to detect and avoid the predator by analyzing chemosignals—as exemplified by kairomones (Lampert 2011).

A special place in the study of the role of biological information belongs to studies of seasonal adaptations and, foremost among them, the diapause or biological dormancy. This evolutionary adaptation comprises a complex mechanism of detection and analysis of heterogeneous information (including photoperiod, temperature, chemosignals), and simultaneously it materializes this information in the form of biological dormancy. Furthermore, the reactivation and subsequent induction of diapause coincide with the active part of the life cycle of organisms, and that affects the development of other species, primarily foraging elements and predators of diapause-using organisms. In other words, it optimizes the seasonal cycle of population development, combining the whole ecosystem into a single whole. Hence, diapause has a biorhythmical function, which, perhaps, is even more important than the protective function (Alekseev 1990).

Vertical information links in biological systems appear primarily in the experimental work of a physiological and biochemical kind, and it is usually well endowed from the methodological side—this is confirmed by results already obtained (Schwarz et al. 1976; Danilevsky 1961; Lampert 2011). The greatest difficulties for the formation of a full-fledged theory about the role and place of information in biological systems at the superorganism level lie in the need to perform an extremely large amount of work comparable only with the decoding of the human genetic code. Significant advances are achieved through model studies on the role of information in the functioning of both populations and more complex biosystems at different trophic levels (Alekseev and Fiks 1989; Alekseev and Umnov 2002; Alekseev and Kazantseva 2007). Research in this direction will shed light on the problem of a more accurate model description of the relationship of the species to the environment and assess the role of bioinformation not only qualitatively but also quantitatively (Alekseev and Kazantseva 2015).

References

Alekseev VR (1990) Diapause in crustaceans: ecological and physiological aspects. Nauka Academic Publishers, Moscow, 144 pp (in Russian)

Alekseev VR, Fiks BV (1989) The population dynamics model of the planktonic Cyclopid Mesocyclops leuckarti (Claus), based on Photoperiodism. Transof Zool Inst USSR Acad sci 205:96–107 (in Russian)

Alekseev VR, Kazantseva TI (2007) The usage of individual-oriented model in studies on the role of the maternal effect in the reproductive switch in Cladocera. Zhurnal obshchei biologii 68(3):242–251

Alekseev VR, Kazantseva TI (2015) The role of chemoreception in forming of Daphnia longispina sustainable population (the simulation experiments). Zhurnal obshchei biologii 76(5):377–389

Alekseev VR, Umnov AA (2002) The role of trophic and signal factors in diapause induction and seasonal dynamics in a Daphnia longispina population: a simulation model. Verh Int Verein Limnol 27:3927–3932

Danilevsky AS (1961) Photoperiodism and seasonal development of insects. Leningrad University Press (in Russian)

Ichas M (1960) Protein text. Theory of information in biology, Moscow, 399 p (in Russian)

Ivlev VS (1961) Experimental ecology of the feeding of fish. Yale University Press, New York, p 302

Khailov KM (1966) The problem of links and organization in life systems' evolution. Russ J Philos:4

Krylov PI (1989) Freshwater predative zooplankton feeding. Hydrobiology, Moscow, 146 p (in Russian)

Lampert W (2011) Daphnia: development of a model organism in ecology and evolution. Excell Ecol 21:1–275

Levich AP (1980) Structure of ecological communities. MGU Publication, pp 3–181 (in Russian)

Levich AP (1983) Biological system languages. Human and biosphere 8:68–77 (in Russian)

Mandelbrot B (1972) Theory of information and psycholinguistic. Mathematical methods in social sciences

Naumov NP (1973) Signal biological fields and their significance to animals. J Commun Biol 74:P808–P817 (in Russian, English summary)

Schwarz SS, Piaskolova OA, Dobrinskaya AA, Runkova GG (1976) Group effect in aquatic animals population and chemical ecology. Nauka Academic Publishers, 151 p (in Russian)

Seravin LN (1973) Theory of information from a point of view of biologist. Nauka Academic Publishers, 160 p (in Russian)

Setrov MI (1975) Informative processes in biological systems. Nauka Academic Publishers, 155 p (in Russian)

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423, 623–656

Shmalhausen II (1968) Cybernetic questions in biology. Nauka Academic Publishers, 224 p

von Bertalanffy L (1969) General system theory. George Braziller, New York, p 136

Wiener N (1961) Cybernetics or control and communication in the animal and the machine. M.I.T. Press, New York, 212p

Willis JC (1922) Age and area, Cambridge. https://doi.org/10.5962/bhl.title.30741

Zelikman EA (1977) Non-trophic regulatory interactions among marine invertebrates. Ocean Biol 2 (in Russian)

Zipf GK (1949) Human behavior and the principle of least effort, Cambridge