

1. Understanding Nature Through the Symbiosis of Information Science, Bioinformatics, and Neuroinformatics

Nikola Kasabov

This chapter presents some background information, methods, and techniques of information science, bio- and neuroinformatics in their symbiosis. It explains the rationale, motivation, and structure of the Handbook that reflects on this symbiosis. For this chapter, some text and figures from [1.1] have been used. As the introductory chapter, it gives a brief overview of the topics covered in this *Springer Handbook of Bio-/Neuroinformatics* with emphasis on the symbiosis of the three areas of science concerned: information science (informatics) (IS), bioinformatics (BI), and neuroinformatics (NI). The topics presented and included in this Handbook provide a far from exhaustive coverage of these three areas, but they clearly show that we can better understand nature only if we utilize the methods of IS, BI, and NI, considering their integration and interaction.

1.1 Nature as the Ultimate Inspiration and Target for Science	1	1.2 Information Science (IS)	3
		1.2.1 The Scope of IS	3
		1.2.2 Probability, Entropy, and Information	3
		1.2.3 Statistical Information Methods	5
		1.2.4 Machine-Learning Methods	5
		1.2.5 Knowledge Representation	6
		1.3 Bioinformatics	7
		1.3.1 Biology Background	7
		1.3.2 Data Analysis and Modeling in Bioinformatics	8
		1.4 Neuroinformatics	9
		1.4.1 Human Brain and Nervous System ..	9
		1.4.2 Data Analysis and Modeling in Neuroinformatics	10
		1.5 About the Handbook	11
		1.5.1 Bioinformatics	11
		1.5.2 Neuroinformatics	11
		1.6 Conclusion	12
		References	12

1.1 Nature as the Ultimate Inspiration and Target for Science

Science aims at understanding nature. Scientific methods are inspired by principles from nature, too. The beauty of our world is that, along with the fascinating geographical formations, it has life as a variety of biological species. Most importantly, it has the highest level of life: the conscious brain. Nature and life have been the ultimate inspirations and targets for three important areas of science: information science (IS, also called informatics), bioinformatics (BI), and neuroinformatics (NI).

IS deals with *generic methods* for information processing. BI applies these methods to biological data and also develops its own *specific* information processing methods. NI applies the methods of IS to brain and

nervous system data and also develops its own *specific* methods.

Nature evolves in time. The most obvious example of an evolving process is life. Life is defined in the *Concise Oxford English Dictionary* as follows:

a state of functional activity and continual change peculiar to organized matter, and especially to the portion of it constituting an animal or plant before death; animate existence; being alive.

Continual change, along with certain stability, is what characterizes life. Modeling living systems requires that the continuous changes are represented in the model,

i. e., that the model adapts in a lifelong mode and at the same time preserves some features and principles that are characteristic to the process. This stability–plasticity dilemma is a well-known principle of life. In a living system, evolving processes are observed at different levels (Fig. 1.1).

At the quantum level, particles are in a complex evolving state all the time, being at several locations at the same time, which is defined by probabilities. General evolving rules are defined by several principles, such as entanglement, superposition, etc. [1.2–6]. The discovery of atomic structure by physics and chemistry revolutionized understanding of these quantum principles. Among the scientists who contributed to this is Ernest Rutherford (1871–1937) (Fig. 1.2).

At a molecular level, ribonucleic acid (RNA) and protein molecules, for example, evolve and interact in a continuous way based on the deoxyribonucleic acid (DNA) information and on the environment. The central dogma of molecular biology constitutes a general evolving rule, but what are the specific rules for different species and individuals? The area of science that deals with the information processing and data manipulation at this level is BI. At the cellular level (e.g., neuronal cells) all the metabolic processes, cell growth, cell division, etc., are evolving processes [1.7].

At the level of cell ensembles, or at a neural network level, an ensemble of cells (neurons) operates in concert, defining the function of the ensemble or the network through learning, for instance, perception of sound, perception of an image, or learning languages [1.8].

In the human brain, complex dynamic interactions between groups of neurons can be observed when certain cognitive functions are performed, e.g., speech and language learning, visual pattern recognition, reasoning, and decision making [1.9–11].

At the level of a population of individuals, species evolve through evolution (Charles Darwin). A bio-

logical system evolves its structure and functionality through both lifelong learning of an individual and the evolution of populations of many such individuals. In other words, an individual is a result of the evolution of many generations of populations, as well as a result of its own developmental lifelong learning processes [1.11, 12].

Processes at different levels from Fig. 1.1 show general characteristics, such as those listed below:

1. *Frequency*: Frequency, denoted as F , is defined as the number of signal/event repetitions over a period of time T (seconds, minutes, centuries). Some processes have stable frequencies, but others change their frequencies over time. Different processes from Fig. 1.1 are characterized by different frequencies defined by their physical parameters. Usually, a process is characterized by a spectrum of frequencies. Different frequency spectrums characterize brain oscillations (e.g., delta waves), speech signals, image signals, and quantum processes.
2. *Energy*: Energy is a major characteristic of any object or organism. Albert Einstein's most celebrated formula defines energy E as depending on the mass of the object m and the speed of light c as $E = mc^2$. Defining the energy in a living system is more complicated. The energy of a protein, for example, depends not only on the DNA sequence that is translated into this protein, but on the three-dimensional (3-D) shape of the protein and on external factors.
3. *Information*: Information is a characteristic that can be defined in different ways, as discussed in Sect. 1.2.
4. *Interaction*: There are many interactions within each of the six levels from Fig. 1.1 and across these levels. Interactions are what make a living organism complex. Understanding them is also a challenge for BI and NI; For example, there are complex interactions between genes in a genome, and between

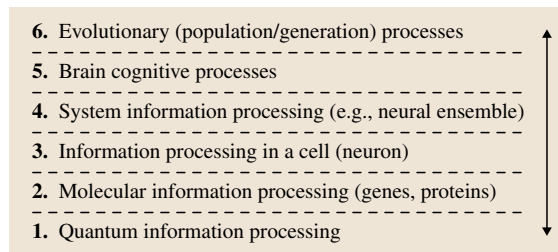


Fig. 1.1 Six levels of evolving processes in a higher-level living organism (after [1.1])



Fig. 1.2 Ernest Rutherford (1871–1937)

proteins and DNA. There are complex interactions between the genes and the functioning of each neuron, a neural network, and the whole brain [1.13].

Abnormalities in some of these interactions are known to cause brain diseases, and many of them remain unknown at present [1.14].

1.2 Information Science (IS)

Information science is the area of science that develops generic methods and systems for information and knowledge processing, regardless of the domain specificity of this information.

1.2.1 The Scope of IS

IS incorporates the following subject areas:

- Data collection and data communication (sensors and networking)
- Information storage and retrieval (database systems)
- Methods for information processing (information theory)
- Creating computer programs and information systems (software engineering and system development)
- Acquisition, representing, and processing knowledge (knowledge engineering)
- Creating intelligent systems and machines (artificial intelligence).

Generally speaking, *data* are raw entities: numbers, symbols etc., e.g., 36.

Information is labeled, understood, interpreted data, e.g., the temperature of the human body is 36°C.

Knowledge is the understanding of a human, the way we do things, the interpretable in different situations, general information; e.g., IF the human temperature is between 36°C and 37°C degrees, THEN the human body is in a healthy state.

Some basic ways to represent data, information, and knowledge are presented in Sect. 1.2.2.

1.2.2 Probability, Entropy, and Information

The formal theory of probability relies on the following three axioms, where $p(E)$ is the probability of an event E to happen and $p(\neg E)$ is the probability of an event not to happen. E_1, E_2, \dots, E_k is a set of mutually exclusive events that form a universe U :

Axiom 1.1

$$0 \leq p(E) \leq 1.$$

Axiom 1.2

$$\sum p(E_i) = 1, E_1 \cup E_2 \cup \dots \cup E_k = U, U\text{-problem space.}$$

Corollary 1.1

$$p(E) + p(\neg E) = 1.$$

Axiom 1.3

$p(E_1 \vee E_2) = p(E_1) + p(E_2)$, where E_1 and E_2 are mutually exclusive events.

Probabilities are defined as:

- Theoretical: some rules are used to evaluate the probability of an event.
- Experimental: probabilities are learned from data and experiments – throw a die 1000 times and measure how many times the event “getting a 6” has happened.
- Subjective: probabilities are based on common-sense human knowledge, such as defining that the probability of getting a 6 after throwing a die is (1/6)th, without really throwing it many times.

A random variable x is characterized at any moment of time by its uncertainty in terms of what value this variable will take in the next moment – its *entropy*. A measure of uncertainty $h(x_i)$ can be associated with each random value x_i of a random variable x , and the total uncertainty $H(x)$, called the *entropy*, measures our lack of knowledge, the seeming disorder in the space of the variable x

$$H(X) = \sum_{i=1, \dots, n} p_i h(x_i), \quad (1.1)$$

where p_i is the probability of the variable x taking the value x_i .

The following axioms for the entropy $H(x)$ apply:

- Monotonicity: if $n > n'$ is the number of events (values) that a variable x can take, then $Hn(x) > Hn'(x)$; so, the more values x can take, the greater the entropy.

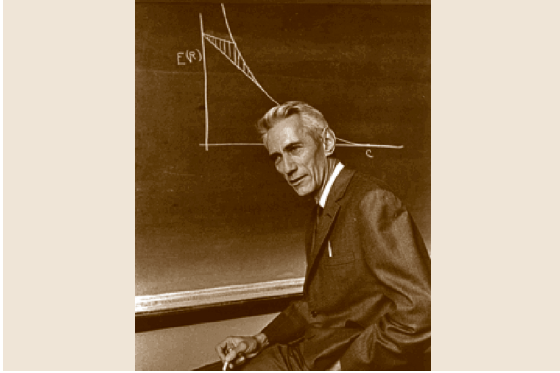


Fig. 1.3 Claude Shannon (1916–2011)

- **Additivity:** if x and y are independent random variables, then the joint entropy $H(x, y)$, meaning $H(x \text{ AND } y)$, is equal to the sum of $H(x)$ and $H(y)$.

The following log function satisfies these two axioms:

$$h(x_i) = \log(1/p_i) . \quad (1.2)$$

If the log has base 2, the uncertainty is measured in [bit], and if it is the natural logarithm \ln , then the uncertainty is measured in [nat],

$$H(x) = \sum_{i=1, \dots, n} [p_i h(x_i)] = -c \sum_{i=1, \dots, n} (p_i \log p_i) , \quad (1.3)$$

where c is a constant.

Based on *Shannon's* (Fig. 1.3) measure of uncertainty – *entropy* – we can calculate an overall probability for a successful prediction for all states of a random variable x , or the predictability of the variable as a whole

$$P(x) = 2^{-H(x)} . \quad (1.4)$$

The maximum entropy is calculated when all the n values of the random variable x are equiprobable, i.e., have the same probability $1/n$ – a uniform probability distribution

$$H(x) = - \sum_{i=1, \dots, n} p_i \log p_i \leq \log n . \quad (1.5)$$

The *joint entropy* between two random variables x and y (for example, an input and an output variable in a system) is defined by the formulas

$$H(x, y) = - \sum_{i=1, \dots, n} p(x_i \text{ AND } y_j) \times \log p(x_i \text{ AND } y_j) , \quad (1.6)$$

$$H(x, y) \leq H(x) + H(y) . \quad (1.7)$$

The *conditional entropy*, i.e., the uncertainty of a variable y (output variable) after observing the value of a variable x (input variable), is defined as

$$H(y|x) = - \sum_{i=1, \dots, n} p(x_i, y_j) \log p(y_j|x_i) , \quad (1.8)$$

$$0 \leq H(y|x) \leq H(y) . \quad (1.9)$$

Entropy can be used as a measure of the information associated with a random variable x , its uncertainty, and its predictability.

The *mutual information* between two random variables, also simply called the *information*, can be measured as

$$I(y; x) = H(y) - H(y|x) . \quad (1.10)$$

The process of online information entropy evaluation is important, as in a time series of events; after each event has happened, the entropy changes and its value needs to be reevaluated.

Information models based on probability include:

- Bayesian classifiers
- Hidden Markov models (HMM).

A *Bayesian classifier* uses a conditional probability estimated to predict a class for a new datum (1.11), which is represented as the conditional probability between two events C and A , known as the Bayes formula (Tamas Bayes, 18th century)

$$p(A|C) = p(C|A)p(A)/p(C) . \quad (1.11)$$

It follows from (1.11) that

$$\begin{aligned} p(A \wedge C) &= p(C \wedge A) = p(A|C)p(C) \\ &= p(C|A)p(A) . \end{aligned} \quad (1.12)$$

Problems with the Bayesian learning models relate to unknown prior probabilities and the requirement of a large amount of data for more accurate probability calculation. This is especially true for a chain of events A, B, C, \dots , where the probabilities $p(C|A, B), \dots$, etc. need to be evaluated. The latter problem is addressed in techniques called hidden Markov models (HMM).

HMM [1.15] is a technique for modeling the temporal structure of a time series or sequence of events. It is a probabilistic pattern-matching approach which models a sequence of patterns as the output of a random process. The HMM consists of an underlying Markov chain

$$\begin{aligned} P[q(t+1)|q(t), q(t-1), q(t-2), \dots, q(t-n)] \\ \approx P[q(t+1)|q(t)] , \end{aligned} \quad (1.13)$$

where $q(t)$ is state q sampled at time t .

1.2.3 Statistical Information Methods

Correlation coefficients represent possible relationships between variables. For every variable x_i ($i = 1, 2, \dots, d_1$), its correlation coefficients $\text{Corr}(x_i, y_j)$ with all other variables y_j ($j = 1, 2, \dots, d_2$) are calculated. Equation (1.14) is used to calculate the Pearson correlation between two variables x and y based on n values for each of them

$$\text{Corr} = \text{SUM}_i \frac{(x_i - Mx)(y_i - My)}{(n-1)\text{Std}x\text{Std}y}, \quad (1.14)$$

where Mx and My are the mean values of the two variables x and y , and $\text{Std}x$ and $\text{Std}y$ are their respective standard deviations.

The t -test and the signal-to-noise ratio (SNR) evaluate how important a variable is to discriminate samples belonging to different classes. For the case of a two-class problem, the SNR ranking coefficient for a variable x is calculated as an absolute difference between the mean value $M1x$ of the variable for class 1 and the mean $M2x$ of this variable for class 2, divided by the sum of the respective standard deviations

$$\text{SNR}_x = \text{abs} \frac{M1x - M2x}{\text{Std}1x + \text{Std}2x}. \quad (1.15)$$

A similar formula is used for the t -test:

$$t\text{-test}_x = \text{abs} \frac{M1x - M2x}{\text{Std}1x^2/N_1 + \text{Std}2x^2/N_2}, \quad (1.16)$$

where N_1 and N_2 are the numbers of samples in class 1 and class 2, respectively.

Principal component analysis (PCA) aims at finding a representation of a problem space X defined by its variables $X = \{x_1, x_2, \dots, x_n\}$ into another orthogonal space having a smaller number of dimensions defined by another set of variables $Z = \{z_1, z_2, \dots, z_m\}$, such that every data vector x from the original space is projected onto a vector z of the new space, so that the distance between different vectors in the original space X is maximally preserved after their projection into the new space Z .

Linear discriminant analysis (LDA) is a transformation of classification data from the original space into a new space of LDA coefficients that has an objective function to preserve the distance between the samples, using also the class label to make them more distinguishable between the classes.

Multiple Linear Regression Methods (MLR)

The purpose of MLR is to establish a quantitative relationship between a group of p independent variables

(X) and a response y

$$y = XA + b, \quad (1.17)$$

where p is the number of independent variables, y is an $n \times 1$ vector of observations, X is an $n \times p$ matrix of regressors, A is a $p \times 1$ vector of parameters, and b is an $n \times 1$ vector of random disturbances. The solution to the problem is a vector, A' , which estimates the unknown vector of parameters.

The *least-squares* solution is used so that the linear regression formula or another model approximates the data with the least root-mean-square error (RMSE) as

$$\text{RMSE} = \text{SQRT} \left\{ \text{SUM}_{i=1,2,\dots,n} \frac{(y_i - y'_i)^2}{n} \right\}, \quad (1.18)$$

where y_i is the desired value from the dataset corresponding to an input vector x_i , y'_i is the value obtained through the regression formula for the same input vector x_i , and n is the number of samples (vectors) in the dataset.

Another error measure is also used to evaluate the performance of the regression and other models – the nondimensional error index (NDEI) – the RMSE divided by the standard deviation of the dataset:

$$\text{NDEI} = \text{RMSE}/\text{Std}. \quad (1.19)$$

1.2.4 Machine-Learning Methods

Machine learning is an area of IS concerned with the creation of information models from data, the representation of knowledge, and the elucidation of information and knowledge from processes and objects. Machine learning includes methods for feature selection, model creation, model validation, and knowledge extraction.

One of the widely used machine-learning method is artificial neural networks (ANNs) [1.9, 11, 16–22].

ANNs are computational models that mimic the nervous system in its main function of adaptive learning. An ANN consists of small processing units – artificial neurons – connected with each other. It has two major functions: learning, which is the process of presenting examples to the ANN and changing the connection weights, and recall, which is the process of presenting new examples to the trained ANN and examining its reaction (output). The connection between the neurons are analogized to the synaptic weights in the nervous system. Most of the known ANNs learning algorithms are influenced by a concept introduced by *Hebb* [1.23]. He proposed a model for unsupervised learning in which

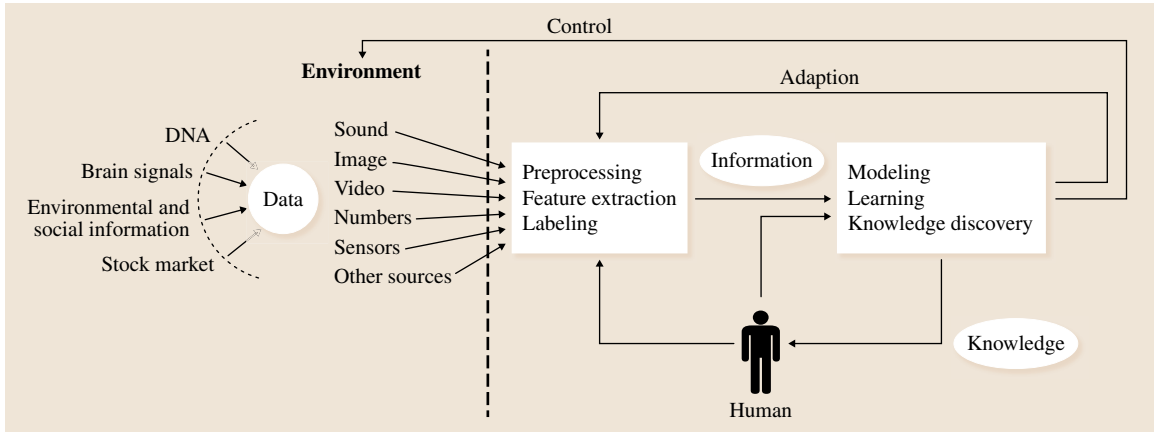


Fig. 1.4 From data to information and knowledge, and then back to information modeling (after [1.1])

the synaptic strength (weight) is increased if both the source and the destination neurons become simultaneously activated. It is expressed as

$$w_{ij}(t+1) = w_{ij}(t) + c o_i o_j, \quad (1.20)$$

where $w_{ij}(t)$ is the weight of the connection between the i th and j th neurons at moment t , and o_i and o_j are the output signals of neurons i and j at the same moment t . The weight $w_{ij}(t+1)$ is the adjusted weight at the next time moment $(t+1)$.

In general terms, a learning system $\{S, W, P, F, L, J\}$ is defined by its structure S , its parameter set P , its variable (e.g., connections) weights W , its function F , its goal function J , and a learning procedure L . The system learns if the system optimizes its structure and its function F when observing events z_1, z_2, z_3, \dots from a problem space Z . Through a learning process, the system improves its reaction to the observed events and captures useful information that may be later represented as knowledge.

Another class of machine-learning methods are inspired by evolution of biology in nature, being called *evolutionary computation* (EC) [1.24, 25]. Here, learning is concerned with the performance not only of an individual system but of a population of systems that improve their performance through generations. The best individual system is expected to emerge and evolve from such populations. EC methods, such as genetic algorithms (GA), utilize ideas from Darwinism.

Another popular machine-learning method is called *support vector machines* (SVM). It was first proposed by Vapnik and his group at AT&T Bell Laboratories [1.26]. For a typical learning task defined as probability estimation of output values y depending on

input vectors x

$$P(x, y) = P(y|x)P(x), \quad (1.21)$$

a SVM classifier is used to build a decision function

$$f_L : x \rightarrow \{-1, +1\} \quad (1.22)$$

based on a training set

$$f_L = L(S_{\text{train}}), \quad (1.23)$$

where $S_{\text{train}} = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

In SVM, the primary concern is to determine an optimal separating hyperplane that gives a low generalization error. Usually, the classification decision function in the linearly separable problem is represented by

$$f_{w,b} = \text{sign}(w \cdot x + b). \quad (1.24)$$

In SVM, this optimal separating hyperplane is determined by giving the largest margin of separation between vectors that belong to different classes. It bisects the shortest line between the convex hulls of the two classes.

1.2.5 Knowledge Representation

The ultimate goal of information processing is the creation of knowledge. The process of knowledge acquisition from nature is a continuous process that will never end. This knowledge is then used to understand nature, to preserve it, to model it, and to predict events. From data to information and knowledge and then back: this is what information science is concerned with (Fig. 1.4).

Different types of knowledge can be used in machine-learning systems, some of them being [1.1, 21]:

- Propositional logic rules (Aristotle)
- First-order logic rules
- Fuzzy logic rules [1.27]
- Semantic maps
- Schemata
- Metarules
- Finite automata
- Higher-order logic.

Fuzzy logic is an extension of propositional logic. It was first introduced by *Zadeh* in 1965 [1.27]. It deals

with fuzzy propositions that can have any truth value between true (1) and false (0). Fuzzy propositions are used to represent fuzzy rules that better represent human knowledge [1.21, 28], e.g., IF the gene expression of gene G is High AND the age of the person is Old THEN the risk of cancer is High, where fuzzy membership functions (rather than Yes or No values) are used to represent the three propositions.

Modeling and knowledge can be [1.26, 29]:

- Global: valid for the whole population of data
- Local: valid for clusters of data [1.1, 30]
- Personalized: valid only for an individual [1.1, 26, 31].

1.3 Bioinformatics

1.3.1 Biology Background

Bioinformatics brings together several disciplines – molecular biology, genetics, microbiology, mathematics, chemistry and biochemistry, physics, and of course informatics, with the aim of understanding life. The *theory of evolution through natural selection* (Charles Darwin, 1809–1882, Fig. 1.5) was a significant step towards understanding species and life.

With the completion of the first draft of the human genome and the genomes of some other species, the task is now to be able to process this vast amount of ever-growing dynamic information and to discover new knowledge.

Deoxyribonucleic acid (DNA) is a chemical chain, present in the nucleus of each cell of an organism; it consists of ordered double-helix pairs of small chemical molecules (bases), adenine (A), cytosine (C), guanine (G), and thymine (T), linked together by a sugar phosphate nucleic acid backbone (Fig. 1.6).

The *central dogma of molecular biology* states that DNA is transcribed into RNA, which is translated into proteins.

DNA contains millions of base pairs, but only 5% or so is used for production of proteins, and these are the segments of the DNA that contain genes. Each gene is a sequence of base pairs that is used in the cell to produce RNA and/or proteins. Genes have a length of hundreds to thousands of bases.

Ribonucleic acid (RNA) has a similar structure to DNA, but here thymine (T) is substituted by uracil (U). In pre-RNA, only segments that contain genes are ex-

tracted from the DNA. Each gene consists of two types of segments: exons, which are segments translated into proteins, and introns, which are segments that are considered redundant and do not take part in protein production. Removing the introns and ordering only the exon parts of the genes in a sequence is called splicing, and this process results in the production of a messenger RNA (mRNA) sequences.

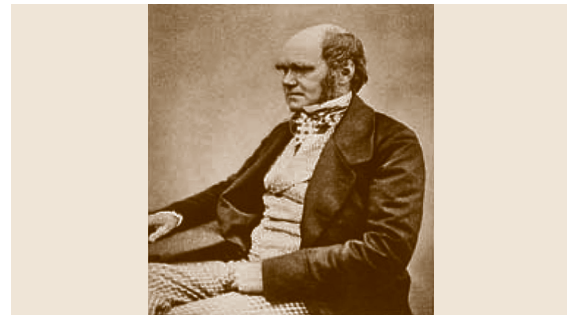


Fig. 1.5 Charles Darwin (1809–1882)



Fig. 1.6 DNA is organized as a double helix

mRNAs are directly translated into proteins. Each protein consists of a sequence of amino acids, each of them defined by a base triplet, called a codon. From one DNA sequence, many copies of mRNA produced, the presence of a certain gene in all of them defining the level of gene expression in the cell and indicating what and how much of the corresponding protein will be produced in the cell.

Genes are complex chemical structures that cause dynamic transformation of one substance into another during the whole life of an individual, as well as the life of the human population over many generations. When genes are *in action*, the dynamics of the processes in which a single gene is involved are very complex, as this gene interacts with many other genes and proteins, and the process is influenced by many environmental and developmental factors.

Modeling these interactions, learning about them, and extracting knowledge are major goals for bioinformatics.

Bioinformatics is concerned with the application of the methods of information science for the analysis, modeling, and knowledge discovery of biological processes in living organisms.

1.3.2 Data Analysis and Modeling in Bioinformatics

There are five main phases of information processing and problem solving in most bioinformatic systems:

1. Data collection, e.g., collecting biological samples and processing them
2. Feature analysis and feature extraction – defining which features are more relevant and therefore should be used when creating a model for a particular problem (e.g., classification, prediction, decision making)
3. Modeling the problem, which consists of defining the inputs, outputs, and type of the model (e.g., probabilistic, rule-based, connectionist), training the model, and statistical verification
4. Knowledge discovery in silico, in which new knowledge is gained through analysis of the modeling results and the model itself

5. Verifying the discovered knowledge in vitro and in vivo – biological experiments in both laboratory and real life to confirm the discovered knowledge.

When creating models of complex processes in molecular biology, the following issues must be considered:

- How to model complex interactions between genes and proteins, between the genome and the environment.
- Both stability and repetitiveness are features that need to be modeled, because genes are relatively stable carriers of information.
- Dealing with uncertainty; For example, when modeling gene expressions, there are many sources of uncertainty, e.g., alternative splicing (a splicing process of the same RNAs resulting in different mRNAs); mutation in genes caused by ionizing radiation (e.g., x-rays); chemical contamination, replication errors, viruses that insert genes into host cells, etc. Mutated genes express differently and cause the production of different proteins.

There are many problems in bioinformatics that require solution through data analysis and modeling. Typical problems are:

- Discovering patterns from DNA and RNA sequences (e.g., promoters, binding sites, splice junctions)
- Analysis of gene expression data and gene profile creation
- Protein discovery and protein function analysis
- Modeling interaction networks between genes, proteins, and other molecules
- System biology approach to modeling the whole dynamic process of biological functioning
- Creating biological and medical prognostic and decision-support systems

All the above tasks require different information methods, both generic (taken from IS) and specific, being created for the particular analysis and modeling of a specific problem and type of biological data.

1.4 Neuroinformatics

1.4.1 Human Brain and Nervous System

The human brain can be viewed as a dynamic, evolving information-processing system, probably the most complex one. Processing and analysis of information recorded from brain and nervous system activity, and modeling of perception, brain functions, and cognitive processes, aim at understanding the brain and creating brain-like intelligent systems. This is a subject of *neuroinformatics*.

The brain evolves initially from stem cells. It evolves its structure and functionality from an embryo to a sophisticated biological information processing system (Fig. 1.7).

In an embryo, the brain grows and develops based on genetic information and nutritional environment. The brain evolves its functional modules for vision, speech and language, music and logic, and many other cognitive tasks.

There are *predefined* areas of the brain that are *allocated* for language and visual information processing, for example, but these areas may change during the neuronal evolving processes. The paths of the signals traveling in, and the information processes of, the brain are complex and different for different types of information. Even at the age of 3 months, some functional areas are already formed, but identical embryos with the same genetic information can develop in different ways to reach the state of an adult brain. This is because of the environment in which the brain evolves. Both the genetic information (nature) and the environment (nurture) are crucial factors. They determine the *evolving rules* for the brain. The challenge is how to reveal these rules and eventually use them in brain models. Are they the same for every individual?

A significant step in understanding the brain and the nervous system was the discovery of the structure of the neural system by Ramón y Cajal (1852–1934, Figs. 1.8, 1.9).

A neuron, which receives signals (spikes) through its *dendrites* and emits output signals through its *axon*, is connected to thousands other neurons through *synapses*. The synaptic connections are subject to *adaptation and learning* and represent the *long-term memory* of the brain.

Neurons can be of different types according to their main functionality [1.11]. There are, for example: sensory neurons, motor neurons, local interneurons, projection interneurons, and neuroendocrine cells.

It is through the organization of neurons into ensembles that functional compartments emerge. Neurosciences provide a very detailed picture of the organization of the neural units in the functional compartments (functional systems). Each functional system is formed by various brain regions that are responsible for processing of different types of information. It is shown that the paths which link different components of a functional system are hierarchically organized.

It is mainly in the *cerebral cortex* where the cognition functions take place. Anatomically, the cerebral cortex is a thin, outer layer of the cerebral hemisphere with thickness of around 2–4 mm. The cerebral cortex is divided into four lobes: frontal, parietal, temporal, and occipital.

Several principles of the evolving structure, functions, and cognition of the brain are listed below:

- Redundancy; i. e., there are many redundant neurons allocated to a single stimulus or a task; e.g., when a word is heard, there are hundreds of thousands of neurons that are immediately activated.
- Memory-based learning; i. e., the brain stores exemplars of facts that can be recalled at a later stage. Some studies suggest that all human actions, including learning and physical actions, are based on memorized patterns.
- Learning is achieved through interaction of an individual with the environment and with other individuals.
- Inner processes take place, e.g., information consolidation through sleep learning.

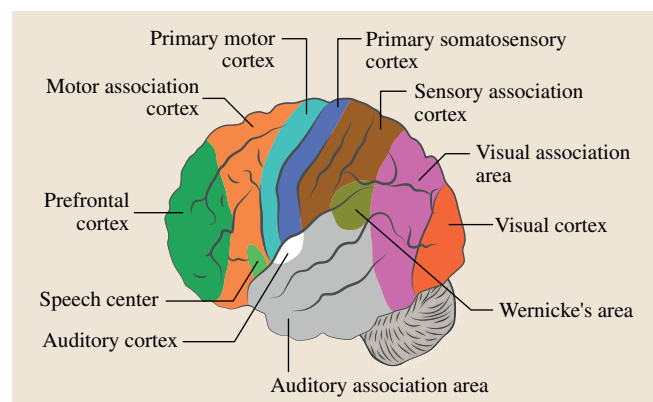


Fig. 1.7 The brain evolves its structure and functionality through genetic information and developmental learning (after [1.32])



Fig. 1.8 Santiago Ramón y Cajal (1852–1934) (after [1.33])



Fig. 1.9 A drawing by Ramón y Cajal of a neuronal circuitry (after [1.33])

- The learning process is continuous, evolving, and lifelong.
- Learning and memory are of three main types: short term (within hundreds of milliseconds), which is manifested in the synapses, the neuronal membrane potentials, and the spiking activity of neurons; long term (within minutes or days), which is manifested in the synaptic weights; and genetic (within months, years, or millions of years), which is manifested in the genes and their expressions). These three types of memory and learning interact in a dynamic way in the brain.
- Through the process of evolving brain structures (neurons, connections), higher-level concepts emerge; these are embodied in the structure and represent a level of abstraction.

1.4.2 Data Analysis and Modeling in Neuroinformatics

The brain is the most complex information processing machine. It processes data, information, and knowledge at different levels. Modeling the brain as an informa-

tion processing machine has different results depending on the goals of the models and the detail with which the models represent the genetic, biological, chemical, physical, physiological, and psychological rules and the laws that govern the functioning and behavior of the brain.

Several levels of brain data analysis and modeling can be distinguished.

Molecular/Genetic Level

At the genetic level, the genome constitutes the input information, while the phenotype constitutes the output result, which causes:

1. Changes in the neuronal synapses (learning)
2. Changes in DNA and its gene expression [1.12].

Neurons from different parts of the brain, associated with different functions, such as memory, learning, control, hearing, and vision, function in a similar way, and their functioning is genetically defined. This principle can be used as a unified approach to building different neuronal models to perform different functions, such as speech recognition, vision, learning, and evolving. The genes relevant to particular functions can be represented as a set of parameters of a neuron. These parameters define the way the neuron functions and can be modified through feedback from the output of the neuron.

Single Neuronal Level

There are many information models of neurons that have been explored in neural network theory (for a review, see [1.11]). Among them are:

1. Analytical models. An example is the Hodgkin–Huxley model (1952) [1.34].
2. McCulloch and Pitts-type models (1943) [1.35].
3. Spiking neuronal models (Maas, Gerstner, Kistler, Izhikevich, Thorpe, Wysocki et al.) [1.29, 36–45].
4. Neurogenetic models, where a gene regulatory network is part of the neuronal model ([1.8, 32, 38, 40]).

Neural Network (Ensemble) Level

Information is processed in ensembles of neurons that form a functionally defined area, such as sensory modalities [1.46]. The human brain deals mainly with five sensory modalities: vision, hearing, touch, taste, and smell. Each modality has different sensory receptors. After the receptors perform the stimulus transduction, the information is encoded through the excitation of neural action potentials. The information is encoded us-

ing the average of pulses or the time interval between pulses. This process seems to follow a common pattern for all sensory modalities, however there are still many unanswered questions regarding the way the information is encoded in the brain.

Cognitive Level

Information is processed in the whole brain through many interacting modules. Many neuronal network modules are connected together to model a complex brain structure and learning algorithms [1.47, 48]. To date, the most effective means available for brain activ-

ity measurements are electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI). Once the data from these measurement protocols has been transformed into an appropriate state-space representation, an attempt to model different dynamic brain functions can be made.

Modeling the entire brain is far from having been achieved, and it will take many years to achieve this goal, but each step in this direction is useful towards understanding the brain and towards the creation of intelligent machines that will help people [1.49].

1.5 About the Handbook

This Springer Handbook includes 12 parts, 6 of them covering topics from BI and 6 from NI. Each part includes chapters, and each chapter introduces topics that integrate BI and IS, or NI and IS, or BI, NI, and IS together.

1.5.1 Bioinformatics

Part A is about *understanding information processes in biological systems*. It includes chapters that reveal the information processing at cellular level, genomics level, proteomics level, and evolutionary molecular biology point of view.

Part B covers the *methods of molecular biology*, including: analysis of DNA sequences, analysis and discovery of microRNA signatures, discovery of regulatory elements in RNA, protein data modeling, and protein structure discovery.

Part C presents different *machine-learning methods for analysis, modeling, and knowledge discovery from bioinformatics data*. It includes chapters that review the applications of different methods, such as Bayesian classifiers and support vector machines (SVM); case-based reasoning; hybrid clustering; fuzzy logic; and phylogenetic cladograms.

Part D presents more sophisticated methods for *integrated, system biology analysis* and modeling in BI, including chapters on: inferring interaction network from Omics data; inferring gene transcription networks; analysis of transcriptional regulations, inferring genetic networks using differential evolution; pattern discovery in protein-protein networks; visual representation of molecular networks; and a pipeline model for identi-

fying somatic mutations with examples from leukemia and colon cancer.

Part E presents *databases and ontologies* that contain structured bioinformatics data to enable worldwide research and study in bioinformatics. It includes chapters on bioinformatics databases and bioinformatics ontology systems.

Part F is about *applications of bioinformatics in medicine, health, and ecology*. It includes chapters on modeling cancer stem formation, epigenetics, immune system control, nutrigenomics, nanomedicine, personalized medicine, health informatics, and ecological informatics.

1.5.2 Neuroinformatics

Part G is about *understanding information processes in the brain and the nervous system*. It includes chapters on information processes at a lower, synaptic level, spiking neural networks that represent and model information processes at a neuronal ensemble level; brain connectivity study based on fMRI data; and information processes at the level of the whole brain.

Part H introduces *advanced signal processing methods for brain signal analysis and modeling*. The methods are applicable to study spatiotemporal spiking activities of single neurons and neuronal ensembles along with spiking activity of the whole cortex. This part includes chapters on adaptive filtering in kernel spaces for spike train analysis, analysis and visualization of multiple spike trains, and the multivariate empirical mode decomposition method for time-frequency analysis of EEG signals.

Part I is concerned with modeling *perception, sensation, and cognition*. It includes chapters on modeling vision, modeling the gustatory system, perception and motor control modeling based on EEG with application for brain–computer interfaces, spiking neural network and neurogenetic systems for spatio- and spectrotemporal brain data analysis and modeling, and models of natural language.

Part J presents *neuroinformatics databases and systems* to help brain data analysis and modeling. It includes chapters on brain-gene ontology systems, neuroinformatics databases, and worldwide organizations.

Applications of neuroinformatics methods for understanding and curing of brain diseases is presented in Part K. It contains chapters on Alzheimer disease

genetic regulatory networks, integrating data and prior knowledge for understanding Alzheimer disease, a system biology approach to modeling and understanding Parkinson and Alzheimer disease, modeling gene dynamics in epilepsy, predicting outcome of stroke, and surface electromyography methods for nerve–muscle system rehabilitation using the case study of stroke.

Nature-inspired integrated information technologies, presented in the last Part L, combine different principles from the biology, brain, and quantum levels of information processing (Fig. 1.1). It includes chapters on brain-like robotics, interactive, developmental multimodal robotic systems, quantum and biocomputing integration, and integrated brain-, gene-, and quantum-inspired computational intelligence.

1.6 Conclusion

This chapter presents a brief overview of the topics covered in this *Springer Handbook of Bio-/Neuroinformatics*, with emphasis on the symbiosis of the three areas of science concerned: information science (informatics), bioinformatics, and neuroinformatics. The

topics presented and included in the Handbook provide a far from exhaustive coverage of these three areas, but they show clearly that we can better understand nature only if we utilize the methods of IS, BI, and NI, considering their integration and interaction.

References

- 1.1 N. Kasabov: *Evolving Connectionist Systems: The Knowledge Engineering Approach* (Springer, London 2007)
- 1.2 R.P. Feynman, R.B. Leighton, M. Sands: *The Feynman Lectures on Physics* (Addison-Wesley, Redding 1965)
- 1.3 R. Penrose: *The Emperor's New Mind* (Oxford Univ. Press, Oxford 1989)
- 1.4 R. Penrose: *Shadows of the Mind. A Search for the Missing Science of Consciousness* (Oxford Univ. Press, Oxford 1994)
- 1.5 C.P. Williams, S.H. Clearwater: *Explorations in Quantum Computing* (Springer, Berlin 1998)
- 1.6 M. Brooks: *Quantum Computing and Communications* (Springer, Berlin, Heidelberg 1999)
- 1.7 D.S. Dimitrov, I.A. Sidorov, N. Kasabov: Computational biology. In: *Handbook of Theoretical and Computational Nanotechnology*, Vol. 1, ed. by M. Rieth, W. Sommers (American Scientific Publisher, New York 2004), Chap. 21
- 1.8 N. Kasabov, L. Benuskova: Computational neurogenetics, *Int. J. Theor. Comput. Nanosci.* **1**(1), 47–61 (2004)
- 1.9 F. Rosenblatt: *Principles of Neurodynamics* (Spartan Books, New York 1962)
- 1.10 W. Freeman: *Neurodynamics* (Springer, London 2000)
- 1.11 M. Arbib (Ed.): *The Handbook of Brain Theory and Neural Networks* (MIT, Cambridge 2003)
- 1.12 H. Chin, S. Moldin (Eds.): *Methods in Genomic Neuroscience* (CRC, Boca Raton 2001)
- 1.13 J.J. Hopfield: Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982)
- 1.14 National Center for Biotechnology Information (US): *Genes and Disease [Internet]* (NCBI, Bethesda 1998), available online at <http://www.ncbi.nlm.nih.gov/books/NBK22183/>
- 1.15 L.R. Rabiner: A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE* **77**(2), 257–285 (1989)
- 1.16 S. Grossberg: On learning and energy – Entropy dependence in recurrent and nonrecurrent signed networks, *J. Stat. Phys.* **1**, 319–350 (1969)
- 1.17 D.E. Rumelhart, G.E. Hinton, R.J. Williams (Eds.): *Learning internal representations by error prop-*

- agation. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT/Bradford, Cambridge 1986)
- 1.18 T. Kohonen: *Self-Organizing Maps* (Springer, Berlin, Heidelberg 1997)
 - 1.19 S. Haykin: *Neural Networks – A Comprehensive Foundation* (Prentice Hall, Engelwood Cliffs, 1994)
 - 1.20 C. Bishop: *Neural Networks for Pattern Recognition* (Oxford Univ. Press, Oxford 1995)
 - 1.21 N. Kasabov: *Foundations of Neural Networks, Fuzzy Systems and Knowledge Engineering* (MIT, Cambridge 1996)
 - 1.22 S. Amari, N. Kasabov: *Brain-like Computing and Intelligent Information Systems* (Springer, New York 1998)
 - 1.23 D. Hebb: *The Organization of Behavior* (Wiley, New York 1949)
 - 1.24 X. Yao: Evolutionary artificial neural networks, *Int. J. Neural Syst.* **4**(3), 203–222 (1993)
 - 1.25 D.B. Fogel: *Evolutionary Computation – Toward a New Philosophy of Machine Intelligence* (IEEE, New York 1995)
 - 1.26 V. Vapnik: *Statistical Learning Theory* (Wiley, New York 1998)
 - 1.27 Z.A. Zadeh: Fuzzy sets, *Inf. Control* **8**, 338–353 (1965)
 - 1.28 T. Yamakawa, H. Kusanagi, E. Uchino, T. Miki: A new effective algorithm for neo fuzzy neuron model, *Proc. Fifth IFSA World Congress (IFSA, 1993)* pp. 1017–1020
 - 1.29 N. Kasabov: Global, local and personalized modeling and profile discovery in Bioinformatics: An integrated approach, *Pattern Recognit. Lett.* **28**(6), 673–685 (2007)
 - 1.30 M. Watts: A decade of Kasabov's evolving connectionist systems: A review, *IEEE Trans. Syst. Man Cybern. C* **39**(3), 253–269 (2009)
 - 1.31 Q. Song, N. Kasabov: TWNFI – Transductive neural-fuzzy inference system with weighted data normalization and its application in medicine, *IEEE Trans. Fuzzy Syst.* **19**(10), 1591–1596 (2006)
 - 1.32 L. Benuskova, N. Kasabov: *Computational Neuro-Genetic Modeling* (Springer, New York 2007)
 - 1.33 <http://www.wikipedia.org> (last accessed April 4 2012)
 - 1.34 A.L. Hodgkin, A.F. Huxley: A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* **117**, 500–544 (1952)
 - 1.35 W. McCulloch, W. Pitts: A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* **5**, 115–133 (1943)
 - 1.36 W. Gerstner: Time structure of the activity of neural network models, *Phys. Rev.* **51**, 738–758 (1995)
 - 1.37 E. Izhikevich: Simple model of spiking neurons, *IEEE Trans. Neural Netw.* **14**(6), 1569–1572 (2003)
 - 1.38 N. Kasabov, L. Benuskova, S. Wysoski: A Computational Neurogenetic Model of a Spiking Neuron, *IJCNN 2005 Conf. Proc.*, Vol.1 (IEEE, New York 2005) pp. 446–451
 - 1.39 N. Kasabov, R. Schliebs, H. Kojima: Probabilistic computational neurogenetic framework: From modeling cognitive systems to Alzheimer's disease, *IEEE Trans. Auton. Ment. Dev.* **3**(4), 1–12 (2011)
 - 1.40 N. Kasabov: To spike or not to spike: A probabilistic spiking neuron model, *Neural Netw.* **23**(1), 16–19 (2010)
 - 1.41 G. Kistler, W. Gerstner: *Spiking Neuron Models – Single Neurons, Populations, Plasticity* (Cambridge Univ. Press, Cambridge 2002)
 - 1.42 W. Maass, C.M. Bishop (Eds.): *Pulsed Neural Networks* (MIT, Cambridge 1999)
 - 1.43 S. Thorpe, A. Delorme, R. Van Rullen: Spike-based strategies for rapid processing, *Neural Netw.* **14**(6/7), 715–725 (2001)
 - 1.44 S. Wysoski, L. Benuskova, N. Kasabov: Evolving spiking neural networks for audiovisual information processing, *Neural Netw.* **23**(7), 819–835 (2010)
 - 1.45 S. Guen, S. Rotter (Eds.): *Analysis of Parallel Spike Trains* (Springer, New York 2010)
 - 1.46 E. Rolls, A. Treves: *Neural Networks and Brain Function* (Oxford Univ. Press, Oxford 1998)
 - 1.47 J.G. Taylor: *The Race for Consciousness* (MIT, Cambridge 1999)
 - 1.48 R. Koetter (Ed.): *Neuroscience Databases: A Practical Guide* (Springer, Berlin, Heidelberg 2003)
 - 1.49 D. Tan, A. Nijholt (Eds.): *Brain-Computer Interfaces* (Springer, London 2010)

<http://www.springer.com/978-3-642-30573-3>

Springer Handbook of Bio-/Neuro-Informatics

Kasabov, N.K. (Ed.)

2014, LVII, 1229 p. 490 illus., 200 illus. in color.,

Hardcover

ISBN: 978-3-642-30573-3