Principles of the Design of D-Neuronal Networks I. Net Representation for Computer Simulation of a Melody Compositional Process *

Peter Fedor

Neurophysics Group, Institute of Animal Biochemistry and Genetics,
Slovak Academy of Sciences, 900 28 Ivanka pri Dunaji, Czechoslovakia

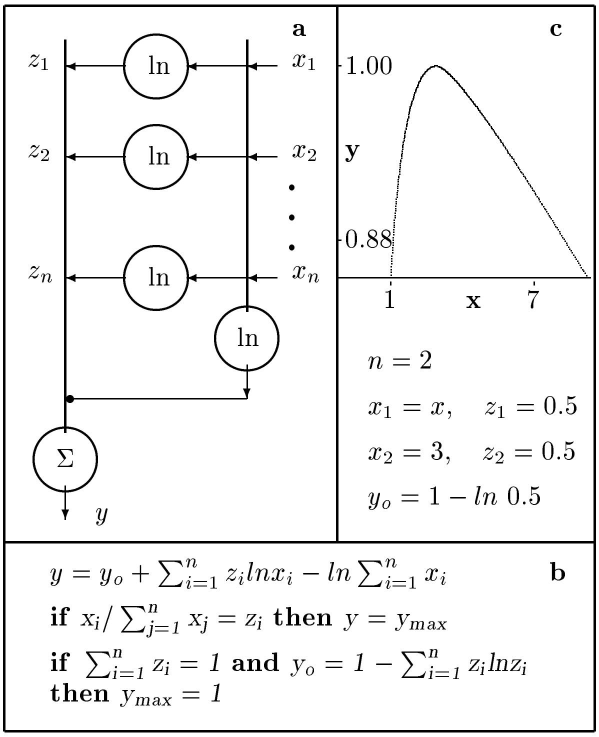
Abstract. We present a deterministic neural network model for the processing of temporal sequences of tones. Several types of model neural units are defined in a form suitable for the digital computer simulation of their nets. On the basis of the proposed units we set up prediction trees which serve for recognition and prediction of pitches that emerge in single-lined diatonic melodies. The prediction tree also evaluates the surprise evoked by a processed melody in regard to the pitch sequences currently memorized in this tree.

1 Introduction

The goal of this paper is to present a biologically plausible algorithm for the processing of pitch sequences being extracted from simple single-lined melodies.[1] We accept the opinion expressed recently by Hildreth and Koch [2] that algorithms used by the brain for processing sensory data "depend not only on the nature of the computation itself, but also on the properties and limitations of the hardware in which the algorithm is implemented". However, it is not easy to decide, even at the low levels of the brain organization (single neurons, elementary neuronal circuits), what models may be considered as a proper description of a "neural hardware". In a recent review on neural nets modelling Cowan and Sharp [3] have quoted: "neurons themselves are highly complicated electro-chemical nets, and can be activated only by appropriate space-time patterns of synaptic activation". It results from the works of Rall, [4] Butz and Cowan, [5] Koch, [6] and Koch and Poggio [7] which have elaborated the input-output description of a neuron with passive dendritic membrane whereas the neuron inputs have been defined in terms of input synaptic currents or even [6] in terms of transient conductance changes of the postsynaptic membrane in sites of synapses. For the neurons with the passive dendritic membrane an essential experimental and theoretical problem must be solved: how to relate quantitatively these conductance changes to space-time patterns of excitations taking place at the presynaptic nerve terminals. For many types of synapses a highly nonlinear transmission relation could be expected. A possible functional meaning of such a phenomenon is illustrated in Fig. 1. In addition, recent studies reveal many of different voltage-dependent conductances in mammalian neurons which endow them with autorhythmic electrical oscillatory properties. [8] Under these circumstances, the decision about implementability of proposed algorithms on neural hardware is a matter of the modeller's intuition. We hope that the presented formalism will be considered as a "provisional framework for organizing ways of thinking about the nervous system", as was required for the so-called simplified brain models by Sejnowsky *et al.* [9]

---

* *International Journal of Neural Systems* **3**(1), 65–73, 1992.



$$n = 2$$
$$x_1 = x, \quad z_1 = 0.5$$
$$x_2 = 3, \quad z_2 = 0.5$$
$$y_o = 1 - ln\ 0.5$$

$$y = y_o + \sum_{i=1}^{n} z_i ln x_i - ln \sum_{i=1}^{n} x_i$$
$$\textbf{if } x_i / \sum_{j=1}^{n} x_j = z_i \textbf{ then } y = y_{max}$$
$$\textbf{if } \sum_{i=1}^{n} z_i = 1 \textbf{ and } y_o = 1 - \sum_{i=1}^{n} z_i ln z_i \textbf{ then } y_{max} = 1$$

Fig. 1. *A discriminating neuron-like circuit. (a) Circuit scheme. (b) Input-output circuit relation. (c) Graph of the input-output relation for the chosen circuit parameters.*

The basic underlying idea of the formalism described in Sect. 2. consists in the assumption that there exists in the brain a representation of space-time features of the "useful part" of the acoustic world in the form of "prediction trees" which allow the brain to predict later acoustic events from those earlier. To each temporal sequence of features being permanently represented in a prediction tree a distribution of weights is assigned. This distribution is interpreted as the set of expectations of features which might follow. The expectation weights are the only tree variables that can be changed during the processing of a melody. An expectation weight of a feature increases if that feature repeated several times as the continuation of the same initial feature sequence. The weighted sum of all changes of all expectations evaluated at each time is considered to represent a variable called "the level of surprise". For the given prediction trees in which a simple "world of melodies" is represented, the melody compositional process is formulated as a search for that sequence of tones which maximize the mean level of surprise. We think that the presented view is in rough accordance with the Meyer's [10] verbal "implication-realization model" of the "embodied meaning" of music.

The prediction trees are networks of the supposed "neural units" called the discrimination unit, the time-shifting unit, the prediction unit, the suprise-evaluating unit, the summation unit, and the negation unit. All these units are subsequently described in individual subsections of Sect. 2. The possible biological interpretation of the whole model is shortly discussed in Sect. 3. The results of the network computer simulation will be presented in Part II of this series.

Table I. Input pitch representation of all the used tones. Columns $\mathbf{p}^1$,$\mathbf{p}^2$,... ,$\mathbf{p}^8$ represent tones **c,d,... ,c'**, respectively.

	$\mathbf{p}^1$	$\mathbf{p}^2$	$\mathbf{p}^3$	$\mathbf{p}^4$	$\mathbf{p}^5$	$\mathbf{p}^6$	$\mathbf{p}^7$	$\mathbf{p}^8$
1	1	0	0	0	0	0	0	1
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	1	0	0	0	1	0	0	0
6	0	1	0	0	0	1	0	0
7	0	0	1	0	0	0	1	0
8	1	0	0	1	0	0	0	1
9	0	1	0	0	1	0	0	0
10	1	0	1	0	0	1	0	0
11	0	0	0	1	0	0	0	0
12	0	1	0	0	0	0	1	0
13	1	0	0	0	1	0	0	1
14	0	0	1	0	0	0	0	0
15	0	1	0	1	0	1	0	0
16	0	0	1	0	1	0	1	0
17	0	0	0	1	0	0	0	1
18	0	0	0	0	0	1	0	0
19	0	0	0	0	1	0	0	0
20	0	0	0	0	0	0	1	0
21	0	0	0	0	0	1	0	1
22	0	0	0	0	0	0	1	0

2 Pitch prediction trees

When a listener hears a single-lined melody he attributes to each tone several features: a pitch, a metrical stress, a loudness, a timbre, a duration, and a spacial location. It is a result of an intricated "categorization process" segmenting the acoustic signal along the temporal dimension into different components and describing them subsequently as discrete hierarchically organized features of the processed signal. The extracted temporal patterns of metrical stresses (onsets of tones with or without weak and strong accents) deliminate hierarchies of pitch events, [11,12] and so on. Taking into consideration the fact that the simplest music may be expressed as a sequence of tones taken from a diatonic musical scale, each tone with an appropriate accent and duration, but all with a uniform timbre and loudness, we define the input net representation $\mathbf{M}(d)$ of a tone sequence with the length d as

$$\mathbf{M}(d) = \mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(t), \ldots, \mathbf{x}(d), \qquad (1a)$$

where

$$\mathbf{x}(t) = m(t).\mathbf{p}(t) \ , \quad (t = 1, 2, \ldots, d). \qquad (1b)$$

Here: $m(t)$ $(m(t) \in \{0,1,2,3\})$ denotes the metrorhythmical event occurred at time t . The values 3, 2 or 1 of m represent the onset of a tone with a strong accent, with a weak accent or without accent, respectively. If $m(t) = 0$ then the tone being occurred at time t–1 is still persisting at time t; $\mathbf{p}(t)$ $(\mathbf{p}(t) \in \{\mathbf{p^1}, \mathbf{p^2}, \ldots, \mathbf{p^8}\})$ is that n-tuple from Table I which corresponds to the pitch $p(t)$ $(p(t) \in \{c, d, e, f, g, a, h, c'\})$ occurred at time t. The jth component of $\mathbf{p^i}$ will be denoted as p^i_j. The introduced tone sequence representation will be considered as the input of all the neural units and their networks proposed in the next text.

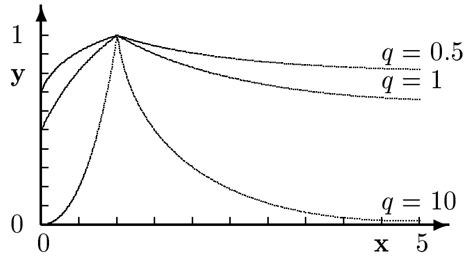


Fig. 2. *Graphs of the discrimination function D for three values of the quality of discrimination and* $\mathbf{x} = (x, 1)$, $\mathbf{z} = (1/2, 1/2)$ *and* $\Theta = 0$. *The function D is defined in text.*

2.1 Pitch discrimination performed by D-units

Each theory of pitch perception assumes that the inner ear performs a type of a spectral analysis of input sounds. [13] To explain the finding that overtones alone can yield the perception of pitch corresponding to a frequency that is not actually present, many models of feature-extracting mechanisms operating on the spectrum has been proposed (for a short review of these models see Ref. 14). Goldstein [15] postulated the existence of a "central pitch processor" searching for the fundamental frequency f_o the harmonics $k_i.f_o$ of which agree best with the frequences occuring in the spectrum of the perceived tone. The improvement of this model performed by Duifhuis *et al.* [16] comprises, as a part of the central pitch processor, the so-called harmonic sieve which selects harmonics for the proper comparison process. Since the definition of the melody representation (1) supposes that such a selection has already been done, for the estimation of the pitch p belonging to an input n-tuple $\mathbf{x}$ it is sufficient to use the discrimination function D defined as

$$D(\mathbf{x}; \mathbf{z}, q, \Theta) = \begin{cases} \left(\sum_{i=1}^{n} min(\kappa_i, z_i)\right)^q & \text{if } \left(\sum_{i=1}^{n} min(\kappa_i, z_i)\right)^q \geq \Theta \ and \ \sum_{j=1}^{n} x_j \geq \epsilon \\ 0, & \text{otherwise} \end{cases},$$

where

$$\kappa_i = \frac{x_i}{\sum_{j=1}^{n} x_j}. \tag{2}$$

Here $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ $(x_i > 0)$; $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ $(z_i \geq 0, \sum_{j=1}^{n} z_j = 1)$; $0 \leq \Theta \leq 1$; $0 < q < q_{max}$; . The quantities q, Θ, ϵ $(0 < \epsilon \ll 1)$ are called the quality of discrimination, the discrimination threshold, and the noise threshold, respectively. It was shown in Fedor and Majerník [17] that it holds:

i) if $\sum_{i=1}^{n} x_i \geq \epsilon$ then $D(\mathbf{x}; \mathbf{z}, q, \Theta) = D(k\mathbf{x}; \mathbf{z}, q, \Theta)$ for any $k > 1$;

ii) $0 \leq D \leq 1$;

iii) $D = 1$ if and only if it holds $\sum_{i=1}^{n} x_i \geq \epsilon$ and $\mathbf{x} = k\mathbf{z}$, where $k > 0$;

iv) if $\Theta = 0$ and $\sum_{i=1}^{n} x_i \geq \epsilon$ then $D = 0$ if and only if it holds $\mathbf{x}.\mathbf{z} = \mathbf{0}$.

The graphs of this function for $\mathbf{x} = (x, 1), \mathbf{z} = (0.5, 0.5), \Theta = 0$ are depicted in Fig. 2 for three various values of q. The function

$$y(t) = c(t).D(\mathbf{x}(t); \mathbf{z}, q, \Theta), \tag{3}$$

represents the input-output desription of a "neural unit" called here the discrimination unit (D-unit). $c(t)$ $(c(t) \in \{0,1\})$ is its boolean controling input at time t, $\mathbf{x}(t)$ its "specific" input at time t, $\mathbf{z}$ its permanent inner or memory state, and $y(t)$ its output at time t. The D-unit is *opened* at time t if it holds $c(t) = 1$. The network graphical notation of a D-unit is depicted in Fig. 3a. The output of an opened D-unit with the inner state $\mathbf{z^i} = \mathbf{p^i}/\sum_{k=1}^{n} p_k^i$ and the input $\mathbf{p^j}$ defines the *measure of consonancy* of the inputs $\mathbf{p^i}$ and $\mathbf{p^j}$ by given values of the discrimination threshold Θ and quality q. The D-unit is *immature* if it is possible to find at least two different inputs $\mathbf{p^i}$ and $\mathbf{p^j}$ which are consonant in regard to it, i.e. they can evoke in it nonzero outputs. Otherwise, the D-unit is *mature.*

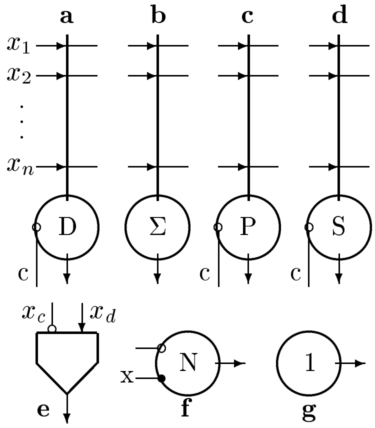


Fig. 3. *List of the network graphical notations of the proposed neural units. (a) D-unit with two excitatory controlling inputs. (b) Summation unit with an excitatory and an inhibitory controlling input. (c) Prediction unit with an excitatory controlling input. (d) Surprise-evaluating unit with an excitatory controlling input. (e) Time-shifting unit. (f) Negation unit. (g) Source of value* 1.

2.2 Prediction and surprise-evaluating units

The input-output relation of a prediction unit is given by (3) in which we put

$$q = 1, \tag{4a}$$

$$\Theta = 0. \tag{4b}$$

In contradiction to the D-unit inner state $\mathbf{z}$, the prediction (and also the surprise-evaluating unit) inner or memory state $\mathbf{z}(t)$, is not constant. It is described by the system of the following difference equations:

$$z_i(t+1) = \frac{z_i(t) + A(t)x_i(t)}{\sum_{j=1}^{n} z_j(t) + A(t)\sum_{j=1}^{n} x_j(t)}, \quad (i = 1, 2, \ldots, n), \tag{4c}$$

with the initial conditions

$$z_i(0) = \frac{1}{n}, \quad (i = 1, 2, \ldots, n). \tag{4d}$$

Here: $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ $(x_i(t) > 0)$ is the input n-tuple of the prediction unit,

$$A(t) = A_o c(t), \tag{4e}$$

is the adaptivity of the prediction unit at time t whereby A_o is a positive constant and $c(t)$ the boolean controlling input.

The summation of (4c) yields an important property

$$\sum_{i=1}^{n} z_i(t) = 1 \quad \text{for any time t.} \tag{5a}$$

The further evident property of (4c) consists in the validity of the inequalities

$$0 \le z_i(t) \le 1 \quad \text{for any i and t.} \tag{5b}$$

The value $z_i(t)$ represents the expectation weight or the prediction that a signal arrives at the ith input at time t.

The input-output relation of a surprise-evaluating unit (SE-unit) is given by

$$y(t) = c(t)[1 - D(\mathbf{x}(t); \mathbf{z}(t), 1, 0)] = \frac{1}{2} c(t) \sum_{i=1}^{n} \left| \frac{x_i(t)}{\sum_{j=1}^{n} x_j(t)} - z_i(t) \right| . \tag{6}$$

The evolution of the memory state $\mathbf{z}(t)$ is described by (4c). It follows from (6) that the output of a SE-unit may be considered as the measure of the distance between the n-tuple $\mathbf{z}$ of expectation weights and the current normalized input $\kappa(t)$. The greater distance the greater surprise evoked by the input $\kappa(t)$. It is clear that the output of a SE-unit is a "negation" of the output of the corresponding prediction unit. For this reason, it is not necessary to work with both units. In what follows the SE-units and "negation units" will be used. The so-called negation unit is defined by means of the formula

$$y(t) = \begin{cases} 1 & \text{if } c(t).[1 - x(t)] \ge \Theta_n \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $c(t)$ is its boolean controlling input, $x(t)$ its "specific" input at time t, and Θ_n its threshold.

The network graphical notations of the defined units are depicted in Fig. 3c, d,f.

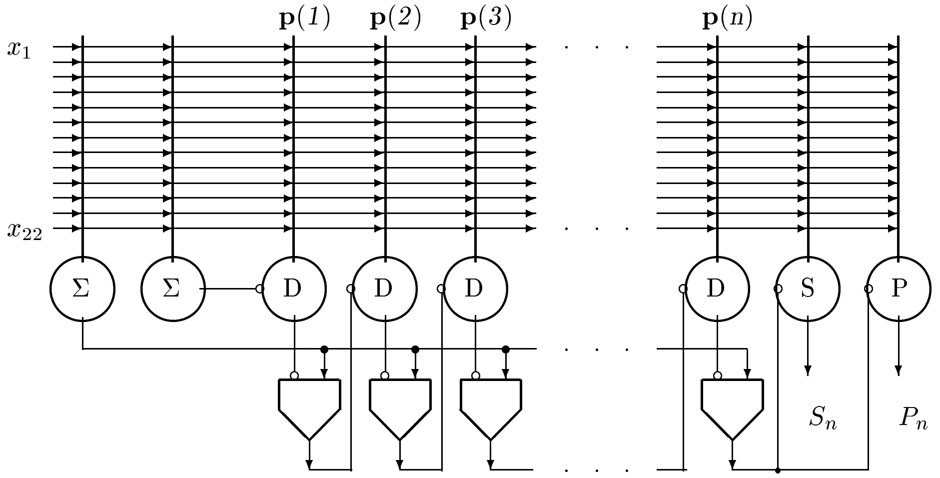


Fig. 4. *Structure of a network recognizing a class of tone sequences. The inner state of the ith net D-unit is adjusted to recognize the pitch tuple $\mathbf{p}(i)$. A sequence belongs to the recognized class if each its terms is "consonant" with the corresponding D-unit inner state. The net prediction unit and the surprise-evaluation unit compute, respectively, the prediction P_n and the surprise S_n of the pitch event $\mathbf{p}(n+1)$.*

2.3 Discrimination of pitch sequences

In order to propose a "neural mechanism" which ought to elucidate how temporal patterns of metrical stresses deliminate in a melody the hierarchies of pitch features which are memorized and recognized invariantly to tempo and tone lengthes, the so-called time-shifting elements (T-shifters) have been introduced in Fedor. [1] Since we let time be discrete this supposed "neural unit" may be described by the following simple relations:

$$x_1(t) = \begin{cases} 1 & \text{if } x_c(t) \geq \Theta_T \\ 0 & \text{otherwise} \end{cases}, \tag{8a}$$

$$x_2(t) = \begin{cases} 1 & \text{if } x_d(t) \geq \Theta_T \\ 0 & \text{otherwise} \end{cases} \tag{8b}$$

$$z(t+1) = z(t) + x_1(t)[1 - z(t)] - x_2(t)[1 - x_1(t)]z(t), \tag{8c}$$

$$y(t) = z(t).x_2(t). \tag{8d}$$

Here: $x_c(t)$ $(x_c(t) > 0)$ and $x_d(t)$ $(x_d(t) > 0)$ are the so-called charging and discharging inputs of a T-shifter, respectively; $z(t)$ $(z(t) \in \{0,1\})$ denotes the inner state of the T-shifter being called in the case $z = 1$ or $z = 0$ the charged or the discharged state, respectively; $y(t)$ $(y(t) \in \{0,1\})$ is the output; Θ_T $(0 < \Theta_T < 1)$ is the T-shifter threshold. The T-shifter network graphical notation is depicted in Fig. 3e.

The T-shifter "charging-discharging equation" (8c) secures that (i) a T-shifter being in the discharged state at time t will be charged at the next time t+1 only if its charging input was activated at time t; (ii) a T-shifter being in the charged state at time t will be discharged at next time t+1 only if its discharging input was activated at time t and its charging input not. It means that a charged T-shifter is "waiting" in the charged state till the arrival of a signal at its discharging input. When a T-shifter charging input is connected with the output of a D-unit we set $\Theta_T = \Theta$.

The network consisting of the D-units which are linearly interconnected by means of T-shifters (see Fig. 4) represents a device for the recognition of a defined class of pitch sequences which are present in the processed melody $\mathbf{M}(d)$. This recognition process is obviously invariant in regard to the tone lengths and to tempo of the processed melody since the discharging inputs of all the T-shifters in Fig. 4 are connected with the output of the so-called summation unit detecting the onset of the input tuples $\mathbf{x}$ the pitch of which ougth to be discriminated by some network D-units. The ith term $\mathbf{x}(i)$ of the input sequence $\mathbf{M}(d)$ will be discriminated by the ith network D-unit if this one was opened by the output of the corresponding T-shifter. However, this event occurs only if the corresponding T-shifter was charged by the (i-1)th D-unit at a previous time moment.

The output of the other network summation unit opens the network first D-unit, i.e. it signals the onset of all the subsequences of $\mathbf{M}(d)$ which ougth to be processed by the network.

The input-output relation of a summation unit is given by

$$y(t) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i(t) \geq \Theta_s \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where $x_i(t)$ $(x_i(t) \geq 0)$ is the ith input at time t. The network graphical notations of a summation unit is shown in Fig. 3b. In the network depicted in Fig. 4 the threshold Θ_s $(\Theta_s > 0)$ is set up at the value at which the output of the summation unit signals the onset of a tone, e.g. the tone with the strong or/and weak accent. If we replace the network summation units by SE-units (or by a combination of both the units) we may reach that the net will discriminate only tones with a new pitch.

2.4 Pitch prediction trees

A network which ought to predict possible continuations of a temporal pitch sequence $\mathbf{p}(1), \mathbf{p}(2), \ldots, \mathbf{p}(d)$ and to evaluate the surprise evoked by the pitch event $\mathbf{p}(n+1)$ may contain as its part a network the organization of which is similar to that depicted in Fig. 4. A prediction tree for pitch prediction and surprise evaluation is then a tree-like structure the each tree-path of which is destinated to recognize a class of pitch sequences. If a tree-path had the same organization as that in Fig. 4 and if it contained *immature D-units* the following undesirable phenomena would appear (i) several tone sequences would activate the same tree path and (ii) a tone sequence would activate several tree pathes.

The prediction tree containing immature D-units we shall call *the immature tree.* Otherwise, we call it *the mature tree.* To prevent the immature trees from the above-mentioned phenomena (or at least to diminish their occurence) the D-units must be controlled not only by outputs of T-shifters but also by inhibitory outputs of SE-units. In this manner a mechanism arises which secures that a prediction tree recognizes and memorizes only sequences which are not too surprising.

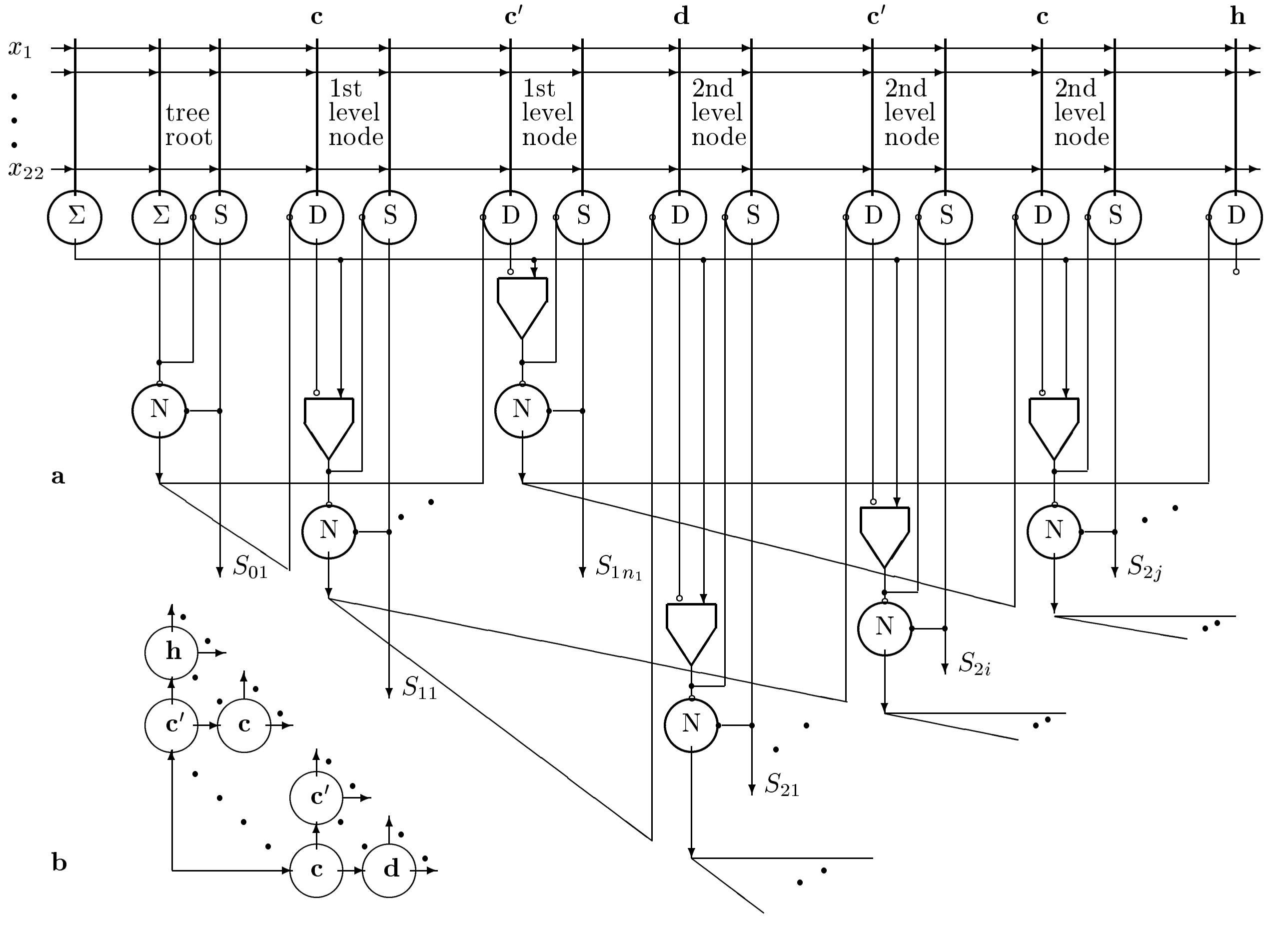


Fig. 5. *Prediction tree. (a) Structure. (b) The simplified schema. The tree nodes are depicted as circles in which the inner states* $\mathbf{z}$ *of their D-units are inscribed.*

The organization of a prediction tree is schematically given in Fig. 5. Each tree-node contains a D-unit, a SE-unit, a T-shifter, and a negation unit. Output branches of a node negation unit are connected with the controlling inputs of D-units located in other tree-nodes. These connections (called in what follows the c-connections) create a true tree structure. The greater the threshold Θ_n of a tree negation unit the smaller the output of the SE-unit must be in order to open the D-unit controlled by this negation unit. For this reason the value $1 - \Theta_n$ will be called the *tree surprise threshold.* For mature trees we set the surprise threshold at its highest value 1.0 for immature trees at a lower value, e.g. at 0.8 if the D-unit threshold and the quality of discrimination were set at 0.39 and 1.0, respectively. The root-node of a prediction tree differs in its organization from the other tree-nodes. It does not contain a T-shifter. The tree root summation unit output signales the onset of tone sequences to be processed by the tree. The network input tones for which the output of the summation unit lying in the root is nonzero we shall call the *tree-starting tones.*For example, the tones with $m \geq 2$ may represent the tree starting tones. In this case the tree will process all the subsequences of the input sequence $\mathbf{M}(d)$ which begin with a tone bearing the weak or strong accent. On the other side, the nonzero outputs of the other summation unit depicted in Fig. 5 on the left of the tree root, signal the onsets of the tones the pitches of which ougth to be processed by the prediction tree, e.g. the occurence of such inputs $\mathbf{x}(t)$ for which it holds $m(t) > 0$. The tree input tones eliciting these nonzero outputs will be called the *tree-shifting tones.*As it was mentioned in Sect. 2.3, replacing the network summation units by SE-units (or by a combination of both the units) we can reach that only tones which bear a new pitch and an appropriate accent can function as the tree-starting and shifting tones.

We say that the pitch sequence $\mathbf{P}(d) = \mathbf{p}(1), \mathbf{p}(2), \ldots, \mathbf{p}(d)$ is *represented* in a prediction tree, if this tree contains a tree-path the D-units of which are adjusted to discriminate subsequently the corresponding sequence terms, i.e. if it holds for the inner state $\mathbf{z}(i)$ of the ith path D-unit

$$\mathbf{z}(i) = \frac{\mathbf{p}(i)}{\sum_{j=1}^{k} p_j(i)} \ , \quad (i = 1, 2, \ldots, d). \tag{10}$$

However, a sequence that is represented in a tree need not to be memorized in it since for the inner states of the path SE-units need not hold relations analogical to (10). The pitch sequence $\mathbf{P}(d)$ is *memorized* in the tree if the d-th D-unit of one of its pathes gives nonzero output at time step d of the sequence processing being started at time 1. A prediction tree is *initiated* if the components of the inner states of all its SE-units are equal $1/n$. Otherwise, the prediction tree is *educated.*

A *hierarchy* of pitch prediction trees arises when the outputs of all the D-units of a pitch prediction tree create the input-tuple of another tree and/or the two trees differ in the definition of their starting and shifting tones. For example, a hierarchically lower tree is activated through its shifting input at time moments for which $m(t) \geq 1$ but the higher one always at time t when it holds $m(t) \geq 2$.

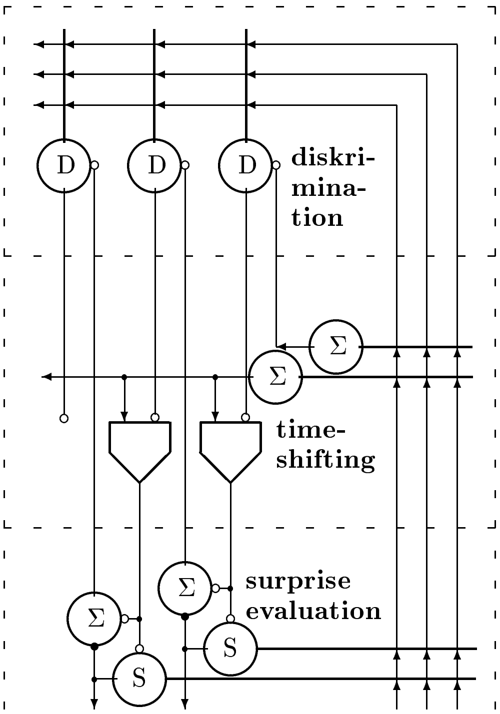


Fig. 6. *A schema of a prediction tree in which different types of neural units lie in different space compartments.*

3 Conclusions of Part I

We have proposed an innovated neural network model for the prediction of pitch sequences during the brain processing of single-lined melodies. The model is defined in the form of difference equations which are convenient for simulation on a digital computer. The results of such a simulation will be presented in Part II of this series. For this reason the general discussion of the model will be left to this part. Here we will briefly discuss possible biological interpretations of the ”memory” processes taking place in the proposed prediction trees and also developmental processes (not included in the presented model) which may underlie the tree genesis and maturation. In order to facilitate the discussion, a prediction tree is depicted in Fig. 6 in such a manner that different types of neural units are distributed in different space compartments.

The tree developmental processes – the tree genesis and the tree maturation – occuring in the upper compartment could be interpreted as remodelling of neural connections by environmental influences during a critical period of the brain postnatal development. Such changes of neural connections influenced by visual experience has been observed in the visual cortex of cats and monkeys. [18] The molecular and cellular mechanisms by means of which the experience is expressed in the structure of neural circuits are not known. However, it seems that the first protein whose biochemical characteristics change with the onset of the critical period has been identified. [19]

Let us shortly discuss the tree development from the viewpoint of its required function – to make it possible to memorize in a tree a limited number of sequences with a limited length repeating one by one many times. The tree maturation process is being considered as a long-term memory process resulting in the increase of the quality of discrimination (or threshold) of all the tree D-units. There are many possibilities of tree genesis and tree maturation.

In order to illustrate one possible tree developmental process let us suppose that the conections between D-units and T-shifters (c-connections) are genetically determined and the experience-dependent modification of connections between tree D-units and tree input lines occurs by means of the learning process running in all the tree D-units. This process is described by (4c) where we put instead of (4e) the following equation:

$$A(t) = A_o c(t) y(t) (q_{max} - q(t)). \tag{11}$$

Here: A_o denotes a positive constant called the constant of the D-unit adaptivity; $c(t)$ is the boolean controlling input; $y(t)$ is the D-unit output; $q(t)$ and q_{max} is the quality of discrimination at time t and its maximal admissible value, respectively.

Now we must describe the tree-maturation process, i.e. the evolution of the quality q of discrimination. Let us suppose that this process is described by the equation

$$q(t+1) = \begin{cases} q_{max} & \text{if } q(t) > \Theta_q \\ q(t) + B_o[y(t) - \frac{1}{2}](q_{max} - q(t)) & \text{otherwise} \end{cases}, \tag{12}$$

where $B_o \quad (0 \leq B_o \leq 1/q_{max})$ is a constant and Θ_q is a threshold value.

It follows from (11) and (12) that the quantity q denotes not only the quality of discrimination but, at the same time, the level of the D-unit maturation. If it holds $q = q_{max}$ the processes described by (11) and (12) are stopped. It is clear that the validity of (11) and (12) provides for an existence of a special educational strategy which enables to represent in an immature tree all before defined input sequences: it is necessary to repeat each sequence sufficiently many times.

The middle compartment in Fig. 6 contains the time-shifting units. Such a hypothetical bistabile neural unit could be arranged in the actual neural tissue by means of an oscilatory loop or by a single pacemaker cell the oscilatory activity of which is controlled (started and stopped) by the corresponding (on- and off-) inputs and the output of which is processed by an off-neuron. The fact that (i) oscilatory neurons and loops has been observed in thalamus [20] and (ii) there exists a corticofugal pathway on sensory thalamic nuclei supports the hypothesis that the cortical discrimination units are connected one to another through the supposed T-shifters located in thalamus. The observed sensitivity of single neurons in auditory cortex to several features of tonal sequences [21] may be considered as an experimental support for the existence of T-shifters in brain.

The bottom compartment in Fig. 6 contains the SE-units in which rapid reversible changes in "synaptic efficiency" z_i occur (see (6c)). At present, it is difficult to decide definitely what biochemical and/or structural changes may be interpreted as events underlying the short-term memory states persisting in range from seconds to several minutes. Maybe the short-term and the long-term potentiation of synaptic transmission observed in variety of brain areas and species [22,23] could be mentioned in this context.

References

1. P. Fedor, "Principles of the design of D-neuronal networks I. A neural model for pragmatic analysis of simple melodies," *Biol. Cybern.* **27**, 129–146 (1977).

2. E. C. Hildreth and C. Koch, "The analysis of visual motion: from computational theory to neuronal mechanisms," *Ann. Rev. Neurosci.* **10**, 477–533 (1987).

3. J. D. Cowan and D. H. Sharp, "Neural nets," *Quart. Rev. Biophys.* **21**, 365–427 (1988).

4. W. Rall, "Core conductor theory and cable properties of neurons," in *Handbook of physiology - The nervous system I: Cellular biology of neurons*, ed. E. R. Kandel and S. Geiger **S.3**, 39–97 (1977).

5. E. Butz and J. D. Cowan, "Transient potentials in systems of arbitrary dendritic geometry," *Biophys. J.* **14**, 661–689 (1974).

6. C. Koch, "Nonlinear information processing in dendritic trees of arbitrary geometries," Ph.D. thesis, University of Tübingen, 1982.

7. C. Koch and T. Poggio, "A theoretical analysis of electrical properties of spines," *Proc. R. Soc. Lond. B* **218**, 455–477 (1983).

8. R. R. Llinás, "The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function," *Science* **242**, 1654–1664 (1988).

9. T. J. Sejnowsky, C. Koch, and P. S. Churchland, "Computational neuroscience," *Science* **241**, 1299–1306 (1988).

10. L. B. Meyer, *Explaining music: essays and explorations* (University of California Press, Berkeley, 1973).

11. F. Lerdahl and R. Jackendoff, "Toward a formal theory of tonal music," *J. Music. Theory* **21**, 111–171 (1977).

12. D. Deutsch, "Organizational processes in music," in *Music, Mind, and Brain*, ed. M. Clynes (Plenum, New York, 1982).

13. L. Deng, C. D. Geisler, and S. Greenberg, "Responses of auditory-nerve fibers to multiple-tone complexes," *J. Acoust. Soc. Am.* **82**, 1989–2000 (1987).

14. E. de Boer and W. A. Dreschler, "Auditory psychophysics: spectrotemporal representation of signals," *Ann. Rev. Psychol.* **38**, 181–202 (1987).

15. J. L. Goldstein, "An optimum processor theory for the central formation of the pitch of complex tones," *J. Acoust. Soc. Am.* **54**, 1496–1516 (1973).

16. H. Duifhuis, L. F. Willems, and R. J. Sluyter, "Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception," *J. Acoust. Soc. Am.* **71**, 1568–1580 (1982).

17. P. Fedor and V. Majerník, "A neuron model as a universal element of self-learning networks for pattern recognition," *Biol. Cybern.* **26**, 25–33 (1977).

18. T. N. Wiesel, "Postnatal development of the visual cortex and the influence of environment," *Nature* **299**, 583–592 (1982).

19. C. Aoki and P. Siekewitz, "Plasticity in brain development," *Sci. Amer.* **259**, 34–42 (1988).

20. M. Steriade and R. Llinás, "The functional states of the thalamus and the associated neuronal interplay," *Physiol. Rev.* **68**, 649–742 (1988).

21. N. M. Weinberger and T. M. McKenna, "Sensitivity of single neurons in auditory cortex to contour: toward a neurophysiology of music perception," *Music Perception* **5**, 355–390 (1988).

22. R. S. Zucker, "Short-term synaptic plasticity," *Ann. Rev. Neurosci.* **12**, 13–31 (1988).

23. T. J. Teyler and P. DiScenna, "Long-term potentiation," *Ann. Rev. Neurosci.* **10**, 131–161 (1987).