Some Interesting Phenomenon Occurring During Self-learning Process with Its Psychological Interpretation

Ryszard Tadeusiewicz

1 Introduction

This book is dedicated for the eminent scientist, former president of IEEE and candidate for 2018 IEEE President-Elect, wonderful man and—last but not least—my friend, professor Jacek Zurada. Professor Zurada is one of the best experts in (among other) computational intelligence [1], neural networks [2] and machine learning areas [3]. Therefore selecting the material for this chapter I must prefer scientific results related to quoted areas.

Neural networks are useful tools for solving many practical problem (e.g. [4–7]). But every of such solution is interesting for limited number of readers, working with similar problems and similar applications. Therefore we select more interesting observations, which are related to the phenomena observed during the neural network self-learning process. Because of same similarity to psychological processes [8], observed during natural activity in our own mind, we call such phenomena "artificial dreams" [9]. This name is similar to the title of Hamid Ekbia's book [10], but the meaning of this term in our works is slightly different. In Ekbia's book "artificial dreams" are presented as unrealized and unrealizable projects related to Artificial Intelligence. In our research we do observe "artificial dreams" as **spontaneous and unexpected processes, emerging automatically from the natural self-learning procedures**.

R. Tadeusiewicz (✉)
AGH University of Science and Technology, Krakow, Poland
e-mail: rtad@agh.edu.pl
URL: http://www.tadeusiewicz.pl

A.E. Gawęda et al. (eds.), *Advances in Data Analysis with Computational Intelligence Methods*, Studies in Computational Intelligence 738,
https://doi.org/10.1007/978-3-319-67946-4_4

The phenomena under consideration are very interesting and exciting, therefore can be mysterious, why so rare are reported by Artificial Intelligence or Computational Intelligence researchers? Yet so many people perform self-learning processes for many purposes—so why such phenomena are still not discovered and described?

The answer is simple. Most papers describing methods and results of the self-learning (even in neural networks, which are main tool considered in this work) are mainly **goal-oriented**. The researcher or practitioner are concentrate on the applications, not on the tool and its behavior. Authors of almost all papers first try to obtain the best result in terms of solving of specified problem (e.g. building of neural network based model of some process or finding the neural solution of the pattern recognition problem). Therefore the discussion of the self-learning results taking into account only the final result (e.g. quality of the model or correctness of classification), while the phenomena discussed in this paper occur when the self-learning system is not learned enough. In all works known to the author at this time nobody see on the details of network (or other self-learning system) behavior during the learning process. Meanwhile some phenomena observed during the self-learning process are really interesting, because totally unexpected.

2 Self-learning and Learning

In this paper we take into consideration **self-learning** process, instead of more known and more useful (from technical point of view) machine learning process. Let us describe the main difference between such two processes, because it will be important from the main thesis of this paper.

During the regular learning process we have the "teacher", who teach "pupil" (in fact it is machine) on the base of examples of properly solved tasks. In machine learning teacher is an algorithm, powered with examples database, but the main idea of teaching is based on simple scheme: get the knowledge from teacher and put it the pupil. After learning process "artificial pupil" can take an exam, where quality of learned knowledge can be evaluated and assessed. On Fig. 1 you can see how it works on the base of gender recognition problem.

In contrast to this scheme self-learning process is based on the knowledge discovery methods. The pupil (in fact it is still machine) can accept input data, but there are no teacher, who can explain, what the data means. Therefore self-learning must not only accumulate knowledge, but it must **discover** this knowledge without any external help. It is in general difficult task, but many successful applications prove this way effective. On Fig. 2 you can see how it works also on the base of gender recognition problem. Self-learning system after connecting with many input data can differentiate man from women, but off course cannot give the proper names to the genders. During the exam self-learning system can give classification for new person (sometimes proper, and sometimes not—as every artificial classification system) using symbols of classes instead names.

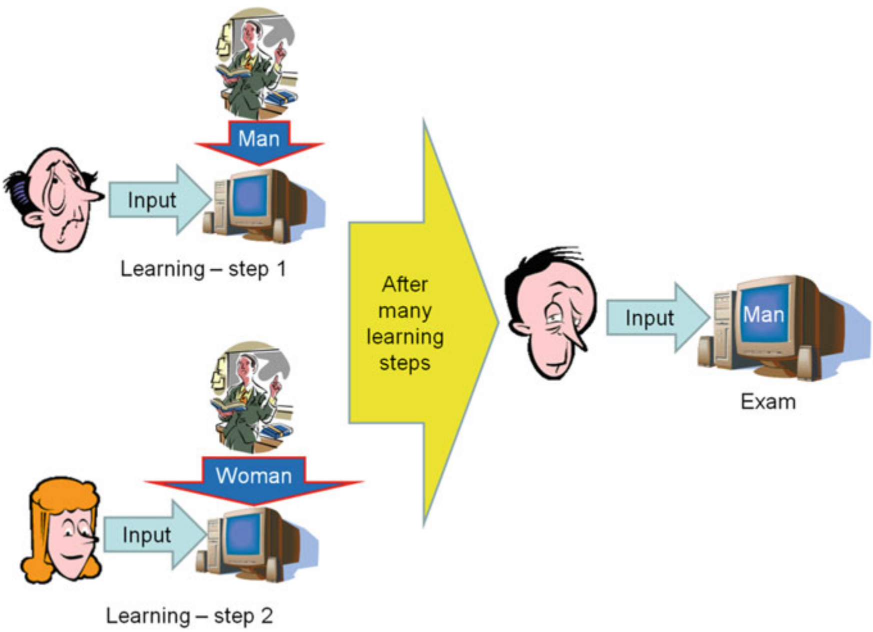


Fig. 1 Learning and exam in supervised learning

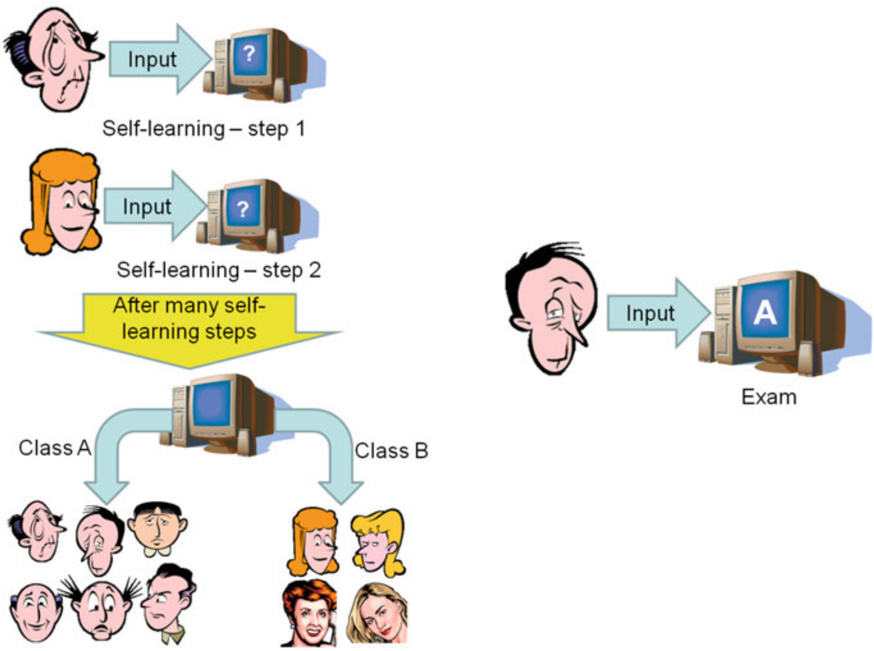


Fig. 2 Learning and exam in unsupervised learning (self-learning)

There are many learning and self-learning systems, but for purpose of this paper we selected neural networks as a tools, in which we can observe the “artificial dreams” phenomena, discussed in our works. Neural networks are in general known for almost everybody, but we try tell some words about simple (and interesting!) application of self-learning neural network, which will be base for further consideration.

3 Self-learning Neural Network

The phenomena described in this paper can be discovered, as mentioned above, in almost all types of neural networks and for almost all methods of learning (both supervised and unsupervised). In this paper we decided take into account such situation:

Let we have one-layer linear neural network. It means as the input to the network we consider n-dimensional vectors $\boldsymbol{X} = < x_1, x_2, \ldots, x_n >$, the knowledge of the network is represented by collection of weight vectors $\boldsymbol{W_j} = < w_{1j}, w_{2j}, \ldots, w_{nj} >$ for all neurons ($j = 1, 2, \ldots, \mathrm{L}$), which outputs can be obtained by mans of simplest and very known equation:

$$y_j = \sum_{i=1}^{n} w_{ij} x_i \tag{1}$$

The network learns on the base of simple hebbian rule: If on step p we obtain the input vector $\boldsymbol{X_p} = < x_{1p}, x_{2p}, \ldots, x_{np} >$ than the correction of the weight vector $\Delta \boldsymbol{W_j}$ (p) depends on the output value y_{jp} calculated by the j-th neuron for $\boldsymbol{X_p}$ according to the Eq. (1), and on the value of input vector $\boldsymbol{X_p}$ according to the formula:

$$\Delta W_j(p) = \eta\, y_{jp} X_p \tag{2}$$

where η is the learning rate coefficient ($\eta < 1$).

Of course new value of weight vector $\boldsymbol{W_j}$ at the next step ($p + 1$) of the self-learning process can be calculated by means of formula:

$$W_j(p+1) = W_j(p) + \Delta W_j(p) = W_j(p) + \eta\, y_{jp} X_p \tag{3}$$

which must be applied for all neurons (for all $j = 1, 2, \ldots, \mathrm{L}$). It is easy to find out, that the result of such calculations are different for neurons with positive output y_{jp} calculated as the answer for input signal $\boldsymbol{X_p}$, and different for neurons with negative output. In first case the weight vector of the neuron $\boldsymbol{W_j}$ (p) is changed toward to the position of actual input signal $\boldsymbol{X_p}$ (attraction), in second case the weight vector of the neuron $\boldsymbol{W_j}$ (p) is changed backward to the position of actual input signal $\boldsymbol{X_p}$

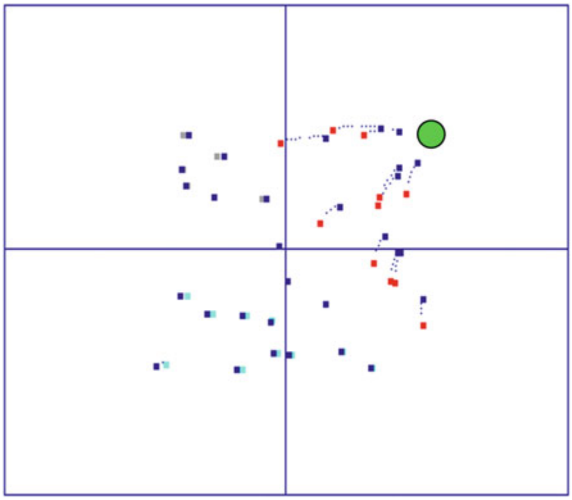


Fig. 3 Migration of the weight vectors during one step at the self-learning process

(repulsion). This process is presented on Fig. 3, where big ring denotes position of input signal $\boldsymbol{X_p}$, and the small squares denotes positions of weight vectors of the neurons. The "migration" of the weight vectors can be observed on this plot—one are attracted toward the input signal, where the other are pushed in opposite direction.

The same process performed by big populations of self-learned neurons is presented on Fig. 4.

Everybody know, what results after many steps of such self-learning process, performed by the network connected with a real data stream. If the data are not

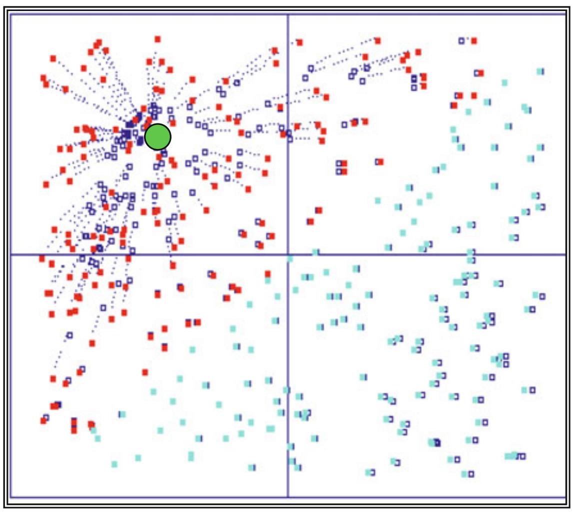


Fig. 4 Migration of the weight vectors in the biggest self-learned network

uniformly distributed, the neurons are divided (spontaneously!) onto groups, when every group is dedicated to the one cluster of the input data. Moreover the values of the weights vectors of the neurons belonging to each group are more or less precisely located in the center of selected cluster of the data. It means, that after the self-learning process inside the neural network we have neurons, which can be used as detectors (or sentinels) for every cluster (group of similar signals), present in observed data stream and automatically discovered by the network.

This process described above is not ideal, because as everybody know, spontaneous migration of the weight vector for every independent neuron leads to many pathologies: every attractor have many neurons as the detectors (over-representation), and sometimes some important attractors can be omitted (no one neuron decide to point out this region of input space). Everybody know also, how to solve this problem: the much better solution is to use Kohonen network and methodology of self-organizing maps.

Yes, but in this work we do not try to made the best self-organized representation of the data. Our goal is definitely other: we are searching for very simple model of the learning of neural network, because on the base of this model we try to show, how (and why) the learned network sometimes presents behavior, which can be interpreted as "artificial dreams".

4 How and Where Artificial Dreams Phenomena Can Be Discovered?

Let assume we must design spacecraft for discovery mysterious world of distant star and planetary robot, which will be send on the ground of totally unknown planet, inhabited by some species of alien monsters. Our robot must collect as many information about aliens as is possible without any a'priori knowledge (Fig. 5). The ideal form of the main computer installed on the robot desk is self-learning neural network, which can collect and systematize information about all creatures found on the exotic planet. After return the spacecraft to the Earth we can obtain from the robot main computer self-learned memory information about number of species of aliens and about their properties, thanks to similar kind classification like shown on Fig. 2.

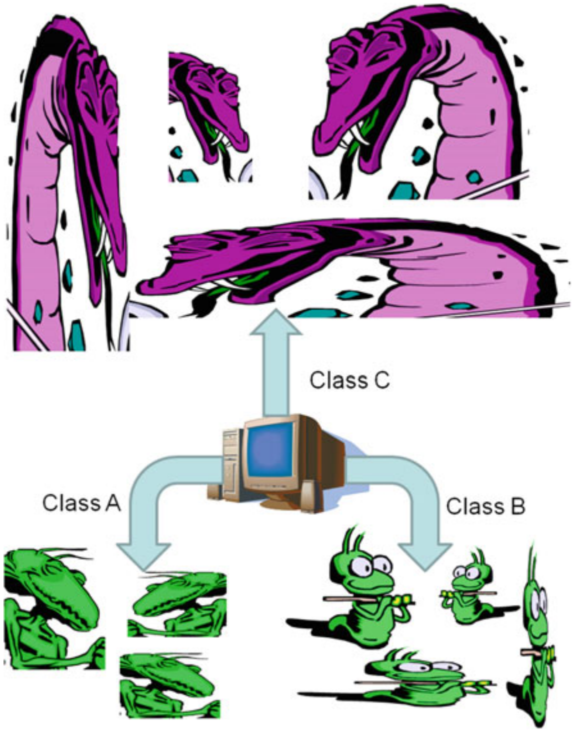
For most researchers only interesting result of computer memory investigation is like shown on Fig. 6. The way, how this classification was obtained by the self-learning process is out of area of interest of most researchers.

Unfortunately!

The example with spacecraft and aliens was rather fantastic and science-fiction based (in fact it was only the joke!), nevertheless the problem under consideration is real and serious. Self-learning system are used often, eagerly and for many purposes. But in fact everybody who use self-learning systems is interested only on final result in terms of classification ability or data clustering, when the way of

Fig. 5 Hypothetical spacecraft robot powered by self-learning neuro-computer

Fig. 6 Content of the memory of self-learning spacecraft after discovery alien planet



learning process is disregarded. Meanwhile we try show in this paper, that the unstable and transitory phenomena, observed in neural networks **during** the self-learning process, are also very interesting, impressive and inspiring.

Such phenomena can be discovered long time after the start of learning, when the network knows nothing because of random values assigned to all it weights. Self-learning process goes then automatically, so typical researcher starts performing another job or goes home. At the same time moment, when such unusual phenomena can be observed, occur long time before final point of learning process, when the network knows (almost) everything and can be exploited according to the plan. Such phenomena can be classified as errors of not matured enough self-learning neural network and therefore can be disregarded. But try give them some psychological interpretations.

5 How Manifest Artificial Dreams?

Observed phenomena can be all disregarded as learning imperfections, but some of them can also be interpreted as "artificial dreams" performed by the artificial neural networks. It can give us new interpretation of the human ability to the imagination, fantasy and also poetry. It can be presented even on the base of the very simple neural network models, but of course the most interesting results can be investigated by means of the networks deployed with high level of similarity to the real brain structures what means big level of complication of the neural structure and also complicated forms of observed phenomena. Before we show and discuss considered phenomena we must shot description of the example problem, in which "artificial dreams" can be very easy encountered.

Let us assume now, that we take into account very simple example problem, which must be solved by the neural network during the self-learning process. In this exemplary problem we assume, that we have four clusters in the input data. Let assume for clear and easy graphical presentation of the results, that the attractors preset in the data (most typical examples) are localized exactly at the centers of four subparts (quarter) of the input space (Fig. 7). The base of this space is defined by two parameters: *body form* and *body shape* (whatever it means). In such space we will observe process of differentiation of four various groups living beings (women, birds, fishes, snakes) shown (one example for every class) on Fig. 7.

In this case self-learning process in simulated neural network after some thousands of learning steps leads to the situation, when almost every neuron become member of one from the four separate groups, located (in sense of localization of weight vectors) at the points corresponding with the centers of the clusters discovered in the input data stream. Three snapshots from the learning process are presented on the Fig. 8.

Typical user of the neural network takes into account mainly last snapshot, presenting, how many neurons are located in proper positions after the learning process and how precisely the real values of attractors coordinates are reproduced

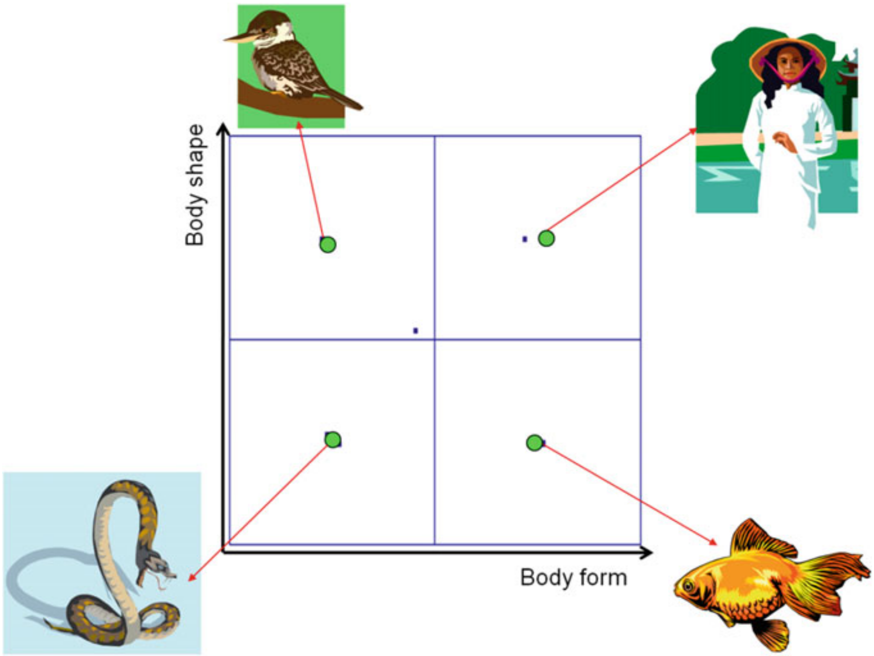


Fig. 7 The example problem. Detail description in the text

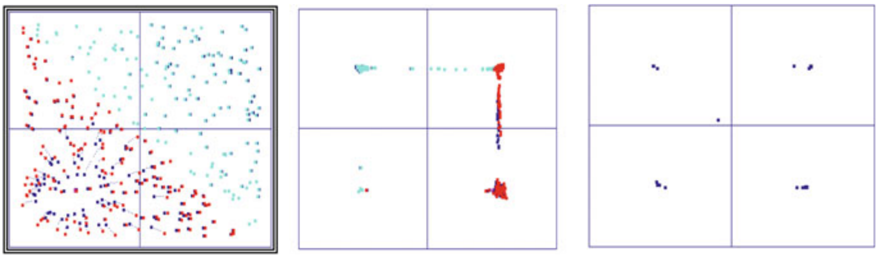


Fig. 8 Three stages of the self-learning process

by the neurons parameters. For our consideration the medium snapshot will be most interesting, because in presents something strange: situation, when knowledge of the network is definitely not complete, but also the initial chaos was partially removed. This stage of learning process is usually skipped by neural network researchers, because apparently man cannot find anything interesting in this plots: the learning process is not ready yet, it's all.

Apparently.

In fact what we see on the central plot on Fig. 8 is registration of "artificial dream". We must only think in terms of special interpretation…

6 Special Interpretation of the Intermediate Stages of Learning Process

In all goal oriented investigations when using neural networks researchers are interested in final result of learning process, which must useful and accurate. Almost nobody takes into consideration intermediate stages shown on Fig. 8. But when we try to understand, what can means in fact form of plotting, repeated on Fig. 9—we must find out, that although it is not real **dream**, it can be interpreted as very exciting **model of artificial dream**. In fact on the plotting presented on Fig. 9 we can point out the localizations of the neurons, which can recognize some (named) objects from real world. After learning all neurons will be attributed to the real world objects, like girls, fishes and birds. But when we have very early stage of learning process, we can find in the population of neurons both real-world related detectors and fantasy-world related detectors. On the line connecting points representing for example girls with the point representing fishes we can find neurons, which are ready to recognize objects, which parameters (features) are partially similar to the girls shapes, and partially include features taken from the other real

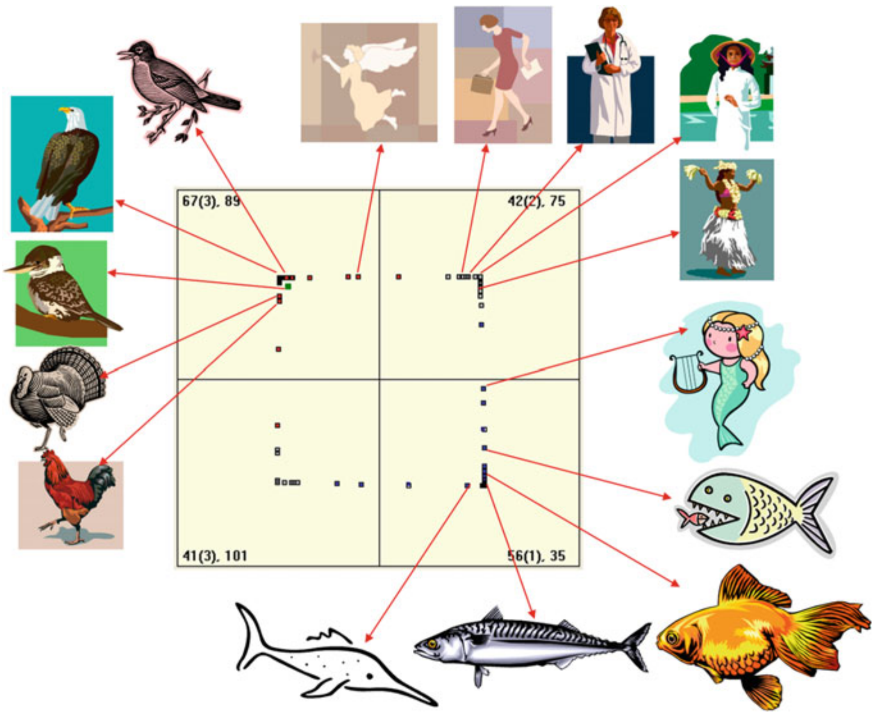


Fig. 9 Parameters of self-learning neural network shows after encoding, that some of neurons spontaneously produce imaginations of non-existing beings. There are the “artificial dreams”!

objects, for example fishes (e.g. tiles). Another hybrid imagination is creature having features taken from girls and from birds. Perhaps it can be angel?

Isn't it something known in the plots shown on Fig. 9? Obviously in real world object like some of plotted here cannot exist. The objects of such properties cannot also be elements of learning data stream, because input information for the network is every time taken from the real world examples. Nevertheless in neural network structure learning process forms neurons, which want to observe and recognize such not real objects.

Isn't it some kind of "artificial dreams"?

Very interesting is fact, that the fantasy-oriented objects, like presented on Fig. 9, encountered during the learning process, never are unrestricted or simply random. We can find only such neurons, which are able to recognize some hybrids, fantastic, but build from the real elements. Isn't it analogy to the tell-stories or myths?

Limited volume of this presentation not allows us to present many other examples of the "artificial dreams" encountered during the learning processes in neural networks. But one more example can be also interesting, because it shows another kind of fantasy identified in neural network behavior. This form of fantasy can be called "making giants". Example of such behavior of the learned network is presented on the Fig. 10. When the network is learned by means of examples of real world object—in the neural structures the prototypes of these objects are formed and enhanced. This process goes over the big population of neurons and leads to the forming of internal representation (in neural structures) of particular real objects. Neurons belonging to these representation can recognize every real object of the type under consideration. It is very known and regular process.

But sometimes in contrast to this regular pattern we can observe single neurons, which parameters are formed in such way, which leads to the surprise after interpretation. Let us assume, that real objects on the base of which the network was learned during the experiment illustrated on Fig. 10, was lion. The network can "see" many lions (of course as a collections of parameters, representing selected data about lions—e.g. how toll is lion, how long and sharp is lion's tooth and so on). After some learning period inside the network we have some imagination of real lion. This imagination, given as collection of parameters (neurons weights), enable us to recognize every real lion. But some neurons have parameters, which enable to recognize surreal lion, much bigger than real one, with biggest tooth and with much more dangerous claws. The relations and proportions between parameters are the same, as for real lions (see on Fig. 10 relations between parameters of real objects and relations between parameters of the imprinted in weights of refugee neuron imagination of the "giant"—both belonging to the same line, coming from the root of coordination system), but so big lion cannot exist. Nevertheless we can find neuron ready for recognition of this giant, although it not exists!

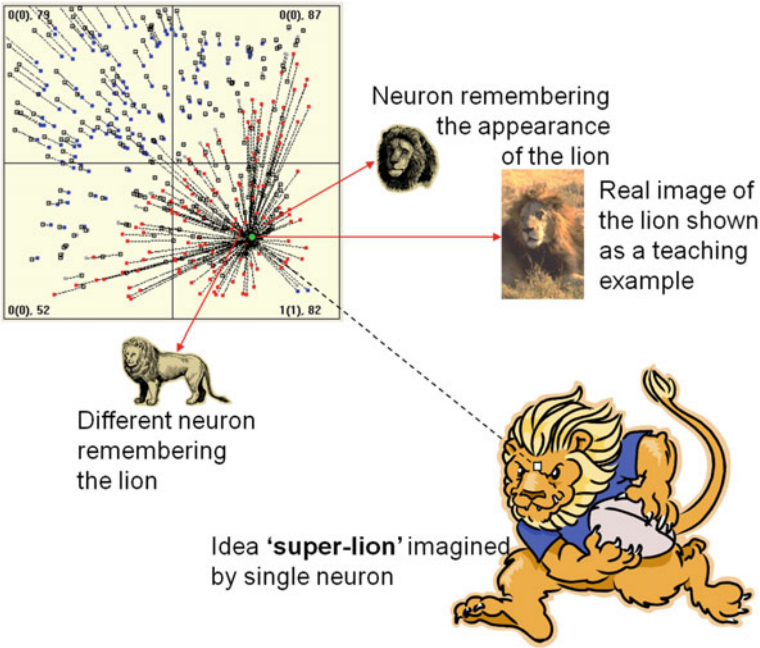


Fig. 10 Another form of "artificial dream". Description in the text

7 Concluding Remarks

Facts and comments presented in this paper definitely aren't very important from the scientific point of view and also are not applicable to the practical problem solving using neural networks. But as long as we use neural networks as the artificial systems very similar to the structures discovered in human brain—we still thinking about analogies between processes in our psychic and in neurocomputers. Results of simulations presented in this paper gives us new point to such considerations and we hope can be interesting for many neural network researchers bored with new learning paradigms, new network structures and new neurocomputing applications and searching for something absolutely different from the serious and boring standards. This paper is something for him!

References

1. Zurada, J.M., Marks, R.J., Robinson, C.J. (eds.): Computational Intelligence: Imitating Life. IEEE Press, New York (1994)
2. Zurada, J.M.: Introduction to Artificial Neural Systems. West Publishing Company, St. Paul, Minnesota (1992)

3. Cloete, I., Zurada, J.M. (eds.): Knowledge-Based Neurocomputing. MIT Press, Cambridge, Massachusetts (2000)
4. Sasiada, M., Fraczek-Szczypta, A., Tadeusiewicz, R.: Efficiency testing of artificial neural networks in predicting the properties of carbon nanomaterials as potential systems for nervous tissue stimulation and regeneration. Bio-Algorithms and Med-Systems (2017). doi:10.1515/bams-2016-0025
5. Mazurkiewicz, E., Tomecka-Suchoń, S., Tadeusiewicz, R.: Application of neural network enhanced ground penetrating radar to localization of burial sites. Appl. Artif. Intell. **30**(9), 844–860 (2016). doi:10.1080/08839514.2016.1274250
6. Smyczyńska, J., Hilczer, M., Smyczyńska, U., Stawerska, R., Tadeusiewicz, R., Lewiński, A.: Artificial neural models—a novel tool for predictying the efficacy of growth hormone (GH) therapy in children with short stature. Neuroendocrinol. Lett. **36**(4), 348–353 (2015). ISSN 0172-780X; ISSN-L 0172-780X
7. Tadeusiewicz, R.: Neural networks in mining sciences—general overview and some representative examples. Arch. Min. Sci. **60**(4), 971–984 (2015). doi:10.1515/amsc-2015-0064
8. Tadeusiewicz, R.: Using neural networks for simplified discovery of some psychological phenomena. In: Rutkowski, L. et al. (eds.) Artificial Intelligence and Soft Computing, LNAI 6114, pp. 104–123. Springer, Berlin, Heidelberg, New York (2010)
9. Tadeusiewicz, R., Izworski, A.: Learning in neural network—unusual effects of “Artificial Dreams”. In: King et al. (eds.) Neural Information Processing, Lecture Notes in Computer Science, Part I, vol. 4232, pp. 211–218, Springer, Berlin, Heidelberg, New York (2006)
10. Ekbia, H.: Artificial Dreams: The Quest for Non-Biological Intelligence, Cambridge University Press (2008)