

**Internal distribution code:**

- (A) [ - ] Publication in OJ
- (B) [ - ] To Chairmen and Members
- (C) [ - ] To Chairmen
- (D) [ X ] No distribution

**Datasheet for the decision  
of 7 November 2022**

**Case Number:** T 0702/20 - 3.5.06

**Application Number:** 14882049.1

**Publication Number:** 3089081

**IPC:** G06N3/04

**Language of the proceedings:** EN

**Title of invention:**

HIERARCHICAL NEURAL NETWORK DEVICE, LEARNING METHOD FOR  
DETERMINATION DEVICE, AND DETERMINATION METHOD

**Applicant:**

Mitsubishi Electric Corporation

**Headword:**

Sparsely connected neural network/MITSUBISHI

**Relevant legal provisions:**

EPC Art. 56

RPBA 2020 Art. 13(2)

**Keyword:**

Amendment after summons - taken into account (yes)

Inventive step - (no)

**Decisions cited:**

G 0001/19, T 0641/00, T 0154/04, T 1326/06, T 1294/16,  
T 1924/17

**Catchword:**

A neural network defines a class of mathematical functions which, as such, is excluded matter. As for other "non-technical" matter, it can therefore only be considered for the assessment of inventive step when used to solve a technical problem, e.g. when trained with specific data for a specific technical task.



**Beschwerdekammern**  
**Boards of Appeal**  
**Chambres de recours**

Boards of Appeal of the  
European Patent Office  
Richard-Reitzner-Allee 8  
85540 Haar  
GERMANY  
Tel. +49 (0)89 2399-0  
Fax +49 (0)89 2399-4465

Case Number: T 0702/20 - 3.5.06

**D E C I S I O N**  
**of Technical Board of Appeal 3.5.06**  
**of 7 November 2022**

**Appellant:**  
(Applicant)

Mitsubishi Electric Corporation  
7-3, Marunouchi 2-chome  
Chiyoda-ku  
Tokyo 100-8310 (JP)

**Representative:**

Pfenning, Meinig & Partner mbB  
Patent- und Rechtsanwälte  
Theresienhöhe 11a  
80339 München (DE)

**Decision under appeal:**

**Decision of the Examining Division of the  
European Patent Office posted on 17 December  
2019 refusing European patent application No.  
14882049.1 pursuant to Article 97(2) EPC.**

**Composition of the Board:**

**Chairman** M. Müller  
**Members:** T. Alecu  
A. Jimenez

## Summary of Facts and Submissions

I. The appeal is against the decision of the Examining Division to refuse the application. The Appellant requests that the decision of the Examining Division be set aside, and that a patent be granted on the basis of a main request corresponding to the first auxiliary request underlying the contested decision, or on the basis of a single auxiliary request filed during the oral proceedings before the Board.

II. The application was refused for a lack of inventive step (Article 56 EPC) starting from document

D1: JOYDEEP GHOSH ET AL: "Structural adaptation and generalization in supervised feed-forward networks", JOURNAL OF ARTIFICIAL NEURAL NETWORKS, vol. 1, no. 4, January 1994, pages 431-458.

III. In the grounds of appeal, the Appellant referred to documents (numbering by the Board)

DA1: Mitsubishi Electric develops Compact Hardware AI for Implementation on Small-scale FPGAs", <https://www.mitsubishielectric.com/news/2018/0214-g.html>, press release from 14 February 2018, and

DA2: Mocanu et al., "Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science", Nature Communications (2018).

IV. Claim 1 of the main request defines:

*A hierarchical neural network apparatus (1) implemented on a computer comprising*

a weight learning unit (20) to learn weights between a plurality of nodes in a hierarchical neural network, the hierarchical neural network being formed by loose couplings between the nodes in accordance with a sparse parity-check matrix of an error correcting code, wherein the error correcting code is a LDPC code, spatially-coupled code or pseudo-cyclic code, and comprising an input layer, intermediate layer and output layer, each of the layers comprising nodes; and

a discriminating processor (21) to solve a classification problem or a regression problem using the hierarchical neural network whose weights between the nodes coupled are updated by weight values learned by the weight learning unit (20)

or comprising

a weight pre-learning unit (22) to learn weights between a plurality of nodes in a deep neural network, the deep neural network being formed by loose couplings between the nodes in accordance with a sparse parity-check matrix of an error correcting code, wherein the error correcting code is a LDPC code, spatially-coupled code or pseudo-cyclic code, and comprising an input layer, a plurality of intermediate layers and an output layer, each of the layers comprising nodes; and

a discriminating processor (21) to solve a classification problem or a regression problem using the deep neural network whose weights between the nodes coupled are updated by weight values learned by the weight pre-learning unit (22)

and

a weight adjuster (23) to perform supervised learning to adjust the weights learned by the weight pre-learning unit (22) by supervised learning; and wherein the weights are learned by the weight pre-learning unit (22) by performing unsupervised learning; and

*the weights between the nodes coupled are updated by weight values adjusted by the weight adjuster (23).*

- V. Claim 1 of the auxiliary request differs therefrom only in that it defines the hierarchical neural network apparatus to be implemented on a microcomputer.

## **Reasons for the Decision**

### *The application*

1. The application relates to a neural network apparatus, a method of classifier learning and a discrimination method (paragraph 1). It starts from the observations (paragraphs 2 and 3) that a standard fully connected neural network requires a large amount of computations and may lead to overfitting, i.e. a classifier that learns the training data too well and is not able to generalise.
- 1.1 It is thus proposed to reduce the number of connections between the nodes; the application talks about "loose coupling" in this context. Unlike the prior art cited in paragraph 3 of the application, the connections are established prior to the training, independently of the learning data (paragraph 6), according to the check matrix of an error correcting code (figure 4, paragraphs 7, 19-23).
- 1.2 According to the application at paragraph 27: "*making the loose couplings between the nodes based on the check matrix of an error correcting code enables the classifier learning and discriminating processing to be performed at high speed while maintaining the discrimination performance*".

*The decision under appeal*

2. There is agreement between the Examining Division and the Appellant (grounds of appeal page 2), and indeed the Board, that the difference between the claimed invention and D1 resides in that the different layers of the neural network are connected in accordance with an error code check matrix.
3. The Examining Division argued (12.1.3) that these distinguishing features *"do not serve a technical purpose, and they are not related to a specific technical implementation either. They merely pertain to the initial, fixed structural definition of an abstract mathematical neural network-like model with unknown, possibly abstract data in- and outputs by means of a binary-valued matrix prior to the model's further simulation and manipulation by means of a computer"*.

*The Appellant's arguments*

4. The Appellant argued in the grounds of appeal that the whole system claimed served a technical purpose (grounds of appeal, bottom of page 2 and first full paragraph on page 3).
- 4.1 The claims related to machine learning *"which serves a technical purpose by solving a well defined technical problem by mathematical means"*. This argument was supported by analogy to T 1326/06 (issued by this Board in a different composition), in which it had been recognized, in the Appellant's words, *"that methods relating to data encoding and/or decoding can serve a technical purpose even though they are almost entirely*

*based on mathematical algorithms and used for encrypting and decoding abstract data".*

- 4.2 The system was implemented by a computer, so a specific technical implementation was present. T 697/17 stated in section 3.5 of its reasons that *"describing a technical feature at a high level of abstraction does not necessarily take away the feature's technical character"*. The Appellant argued that *"By analogy, the possibility that the neural network apparatus may process unknown, possibly abstract data in- and outputs should not necessarily take away the technical character of the distinguishing feature"* (page 3, paragraph 3).
- 4.3 The distinguishing feature solved the problem of improving the learning capability and efficiency of a machine (page 4, paragraph 3) by reducing the required computational resources and preventing overfitting (page 3, paragraph 4). This *"paved the way for the development of compact hardware artificial intelligence"* as shown by DA1. These technical effects were demonstrated in scientific papers and *"the specific design of sparse neural networks has become a major research trend in the field of machine learning recently"* as shown by DA2 (page 4, paragraphs 1 and 2).
5. In its preliminary opinion, the Board tended to agree with the Examining Division and, in particular, did not consider that machine learning in general solved a technical problem or constituted a field of technology within the meaning of Article 52(1) EPC. In response, the Appellant provided the following arguments during the oral proceedings before the Board.
6. The claim was directed to a a neural network apparatus. An artificial neural network was a mathematical algo-



rithm meant to mimic the human brain, by replicating biological optimization. It was implemented and trained in hardware, on a computer; the application itself also referred to a microcomputer. It allowed the automation of complex tasks, so that the computer could perform them instead of a human; automation was generally recognized by the case law as a technical problem. A neural network was thus not an abstract mathematical method, but it used mathematics to solve a technical problem, as was the case in cryptography. Artificial neural networks were therefore to be considered as defining a field of technology.

- 6.1 Though implemented by way of a computer program, a neural network was not a conventional computer program in that its functioning was not determined by the programmer but by the data used for the training. The programmer could not predict how the neural network would work. If its execution was stopped, the programmer would not understand the significance of the values of its mathematical parameters; also in that it was similar to cryptography. A neural network implemented on a computer set-up that computer to function like an artificial brain.
- 6.2 The current application contributed to the domain of neural networks. As already explained in the grounds of appeal, the new network structure with sparse connections allowed for a more efficient implementation by reducing the computing and storage requirements, so that networks could be placed on smaller devices.
- 6.3 The fact that data remained abstract did not mean that a technical problem could not be acknowledged (T 697/17, reasons 3.5, as already submitted in the grounds of appeal). What was a technical field remained

an open question. The Enlarged Board considered in G 1/19 (reference being made to points 67 and 85 of the reasons) that it was "never possible to give an exhaustive list of (positive or negative, alternative or cumulative) criteria for assessing whether a computer-implemented process solves a technical problem", and that technical effects, such as "better use of storage", could occur "within the computer".

*The Board's opinion*

*Technical background: neural networks*

7. A neural network is composed of nodes, called "neurons", linked to each other by edges transmitting the output of one neuron to the input of another. Each neuron implements a parameterized mathematical function, typically a weighted addition of its inputs followed by a nonlinear operator (e.g. a threshold, a sigmoid function etc.); the parameters are called weights.
- 7.1 The structure of the network, i.e. the neuron types (the functions they implement) and the way in which they are connected differs from one network to another, but there are always neurons receiving the input data, and others producing the output. For instance, in a classical feed-forward network, the neurons are organized in layers, each inner layer receiving inputs from the preceding layer and outputting data to the following layer, the first layer receiving the input data, and the last layer outputting the result.
- 7.2 In principle, it is possible, if cumbersome, to replace the inputs to each neuron by the mathematical functions implemented by the nodes of the previous layer and write down the mathematical function that the network

implements as a whole, i.e. the output as a function of the input. The function definition so obtained is determined by the structure of the network, which gives it its general form, and by its weights.

- 7.3 Each set of weights defines a different such function. Thus neural networks with a particular structure constitute a class of mathematical functions, and each member of the class is defined by its weights.
  
- 8. The network is used to "learn" a relationship between pairs of input and output data using known such pairs (training data) so that, when presented with new input data, it can output the "correct" result. The learning process proceeds by changing the values of the weights until the relationship is learned to a level deemed satisfactory, by minimising a loss function depending on the errors made on the training set and their cost.
  - 8.1 The learning process "fixes" the weights in a network (although they may also be changed by re-training at a later time). It may be said that the learning produces a member of the class of functions which is suitable to replicate the input-output relationship expressed in the training data.
  
  - 8.2 The capability of a neural network to learn that relationship, and thus to fulfil its task of providing a correct result on unseen data, is based essentially on "brute force". A large amount of configurable parameters (the weights) are provided so that the functions represented by the neural networks can approximate a large set of input-output distributions, given sufficient data. Nonetheless, the structure of the neural network determines and constrains the class of functions it can represent and thus the set of input-

output distributions it can model. The accuracy of the trained neural network also depends on the adequacy of the loss function and of the training data quality (e.g. data coverage and accuracy).

*Legal background: exclusion and technicality*

9. Article 52(1) EPC provides that:

*European patents shall be granted for any inventions, in all fields of technology, provided that they are new, involve an inventive step and are susceptible of industrial application.*

Article 52(2) EPC provides a list of things that, in particular, *shall not be considered as inventions within the meaning of Article 52(1) EPC, inter alia mathematical methods* (paragraph (a) and *programs for computers* (paragraph (c)).

Their patentability shall be excluded (Article 52(3) EPC) *only to the extent to which a European patent application or European patent relates to such subject-matter or activities as such.*

10. A neural network relates to both programs for computers and to mathematical methods. The question to be answered is whether it relates only to such subject-matter "as such" or whether it relates to something more, and, in particular, to something that can fulfil the patentability conditions of the EPC.
11. Article 52(1) EPC is understood as setting out four requirements to be fulfilled by a patentable invention: there must be an invention, and if there is an invention, it must satisfy the requirements of novelty,

inventive step, and industrial applicability (see, e.g., G 1/19, reasons 30 (A)).

- 11.1 Established case law for computer-implemented inventions defines a corresponding two-step approach (also known as the "two hurdle approach") (see, e.g., G 1/19, reasons 37-39). In the first step, it is assessed whether there is an invention in the meaning of Article 52(1), in view of the exclusions in Article 52(2) EPC, by the so-called "any hardware" approach. It is considered that an invention is something that possesses technical character, and that this condition is fulfilled the moment any technical means, e.g. a computer, is claimed (cf. G 1/19, reasons 28 and 29).
- 11.2 In the second step, following the so-called Comvik approach (T 641/00), it is made sure that only features contributing to the technical character of the invention are considered for the assessment of, in particular, inventive step (cf. G 1/19, reasons 30 (F), cited from T 154/04). In particular, "non-technical" features, understood in this context as features which, on their own, would fall within a field excluded from patentability Article 52(2) EPC (see, e.g., T 1294/16, reasons 35), can only be considered for this assessment if they contribute to solving a technical problem (see also T 1924/17, reasons 15 to 19). Even technical features may be ignored with regard to inventive step if they do not contribute towards solving a technical problem (see G 1/19, reasons 33).
- 11.3 Accordingly, whether a claimed invention is patentable or not can often be decided by focusing on the technical problems it solves, and by means of which combination of features, be they technical or not, and by answering the question of whether this combination of

features is obvious. As a consequence, the Boards of Appeal often limit their objections to ones under inventive step even if it might be possible to raise other objections, too.

*The case at hand*

*Main request*

12. The claimed neural network apparatus may have, as argued by the Appellant, a new and non-obvious structure. The proposed network structure, however, only defines a class of mathematical functions (see above points 7 and 8), which, as such, is excluded matter. As for other "non-technical" matter, it can therefore only be considered for the assessment of inventive step when used to solve a technical problem (see above point 11).

12.1 The Appellant has argued that the claimed neural network solved a technical problem by providing effects within the computer related to the implementation of neural networks (storage requirements), and that neural networks generally solve technical problems by automating human tasks. Though the Appellant has not argued this, the Board remarks that a technical problem may also be solved if the outputs of the system have an implied further technical use (G 1/19, reasons 137).

*Effects "within the computer"*

13. The Appellant has emphasised that the claim is to a neural network apparatus implemented on a computer.

13.1 This makes the application pass the first hurdle with the "any hardware" approach. However, in view of the question which technical problem might be solved, the

Board notes that the implementation does not require any adaptation of the computer. This might be why the Examining Division referred to the lack of a "specific" technical implementation. The *compact hardware* referred to in DA1 is neither part of the present claims, nor of the application, for that matter.

14. The Appellant argued that the proposed modification in the neural network structure, in comparison with standard fully-connected networks, would reduce the amount of resources required, in particular storage, and that this should be recognized as a technical effect, following G 1/19, reasons 85.

14.1 The Board notes that, while the storage and computational requirements are indeed reduced in comparison with the fully-connected network, this does not in and by itself translate to a technical effect, for the simple reason that the modified network is different and will not learn in the same way. So it requires less storage, but it does not do the same thing. For instance, a one-neuron neural network requires the least storage, but it will not be able to learn any complex data relationship. The proposed comparison is therefore incomplete, as it only focuses on the computational requirements, and insufficient to establish a technical effect.

*The neural network apparatus as an automation tool*

15. The appellant has also argued that neural networks apparatuses are artificial brains, and that artificial brains solve an automation problem, because they can carry out various complex tasks, instead of the human, without being programmed specifically for one task or another. In its argument, the appellant stressed that the neural networks mimic the human brain and that

their behaviour cannot be predicted or understood by their programmer.

16. The Board sees no evidence that a neural network functions like a human brain. While its structure is inspired by that of the human brain, this does not imply that they can actually function like one.

16.1 Moreover, whilst the functioning of a neural network may not be foreseeable prior to training and the programmer may not understand the significance of its individual parameters, as the Appellant argued, the neural network still operates according to the programming of its structure and learning scheme. Its parameters and provided results are fully determined, given the training data and the training procedure: at its core, as explained above, a neural network is a mathematical approximation function, which can be simple and understandable if the network is small (e.g. an approximating line going through a set of 2D points for a single neuron perceptron). It is only the sheer complexity of a larger neural network that makes it appear unpredictable. That a learning system is complex is not sufficient to conclude that it replicates the functioning of a brain. The Board also notes that the claims do not determine any specific degree of complexity.

16.2 The Appellant thus has not convinced the Board that neural networks in general function like a human brain or can replace the human in performing complex tasks. Even less so has the Appellant established that the claimed neural network solves the "brain" automation problem in general.



17. The claims do not further specify any particular task, i.e. type of relationship to be learned, for the neural network. Hence, it cannot be said either that the claimed neural network solves any specific automation problem.

*Implied "further technical use"*

18. The claimed learning and use of the network "to solve a classification problem or a regression problem" (where classification is merely regression with discrete outputs corresponding to the classes), can use any data. The outputs of the neural network do not have therefore any implied "further technical use"; they may, for instance, be related to forecasting stock market evolution. In cryptography, the example provided by the Appellant, the situation is different: the encryption of digital messages was found to address the technical problem of increasing system security by preventing data access to parties not in possession of the decryption key (T 1326/06 reasons 6 and 7).

*Conclusion*

19. The claim as a whole specifies abstract computer-implemented mathematical operations on unspecified data, namely that of defining a class of approximating functions (the network with its structure), solving a (complex) system of (non-linear) equations to obtain the parameters of the functions (the learning of the weights), and using it to compute outputs for new inputs. Its subject matter cannot be said to solve any technical problem, and thus it does not go beyond a mathematical method, in the sense of Article 52(2) EPC, implemented on a computer.

- 19.1 Under the "any hardware" definition of the first hurdle it is to be concluded that the claimed matter is not excluded from patentability but does not involve an inventive step in the sense of Article 56 EPC.

*Further remarks*

20. The Board stresses that there can be no reasonable doubt that neural networks can provide technical tools useful for automating human tasks or solving technical problems. In most cases, however, this requires them to be sufficiently specified, in particular as regards the training data and the technical task addressed. What specificity is required will regularly depend on the problem being considered, as it must be established that the trained neural network solves a technical problem in the claimed generality.
21. For the sake of completeness, the Board also notes the following: even if, as the Appellant argued, general methods for machine learning, and neural networks in particular, were to be considered as matter not excluded under Article 52(2) EPC, it would remain questionable whether the proposed loose connectivity scheme actually provides a benefit beyond the mere reduction of storage requirements, for instance a "good" trade-off between computational requirements and learning capability.
- 21.1 In terms of learning, the Appellant asserted that the new structure avoided overfitting, but did not justify this assertion. As explained above, the performance of a given neural network structure, including whether overfitting occurs, generally depends on the data characteristics. Here, however, data characteristics are not considered when the network connectivity is

determined (see 1.1 above). The Board notes that the prior art cited in the application, as well as DA2 (see the algorithm in Box 1, bottom of the second page), relate to data-driven sparsity, i.e. the connectivity scheme is learned from the task at hand, based on the training data distribution.

- 21.2 Hence the Board cannot see in this particular case, considering the content of the application, for which type of learning tasks the proposed structure may be of benefit, and to what extent.

*Auxiliary request*

22. This request differs from the main request only by specifying a microcomputer instead of a computer, which amendment was meant to emphasise the limited resources and therefore the relevance of a small network size. However, as this matter has already been part of the discussion for the main request, the Board admits the auxiliary request and finds it unallowable for lack of inventive step, too (see also T 1294/16, reasons 18.2 to 18.4).

**Order**

**For these reasons it is decided that:**

The appeal is dismissed.

The Registrar:

The Chairman:



L. Stridde

Martin Müller

Decision electronically authenticated