Internal distribution code:

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

Datasheet for the decision of 5 November 2018

Case Number: T 0815/17 - 3.5.06

Application Number: 11733755.0

Publication Number: 2591449

IPC: G06N3/063

Language of the proceedings: EN

Title of invention:
METHODS AND SYSTEMS FOR REPLACEABLE SYNAPTIC WEIGHT STORAGE IN NEURO-PROCESSORS

Applicant:
Qualcomm Incorporated

Headword:
Replaceable synaptic weight storage/QUALCOMM

Relevant legal provisions:
EPC Art. 56, 111(1)
RPBA Art. 13(1), 13(3)

Keyword:
Inventive step (no)
Late-filed requests - admitted (yes)
Right to two instances (no)
Remittal to the department of first instance (no)

Decisions cited:
G 0010/93

Catchword:

Europäisches Patentamt
European Patent Office
Office européen des brevets

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 0815/17 - 3.5.06

D E C I S I O N
of Technical Board of Appeal 3.5.06
of 5 November 2018

Appellant: (Applicant)	Qualcomm Incorporated 5775 Morehouse Drive San Diego, CA 92121 (US)
Representative:	Wegner, Hans Bardehle Pagenberg Partnerschaft mbB Patentanwälte, Rechtsanwälte Prinzregentenplatz 7 81675 München (DE)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 26 October 2016 refusing European patent application No. 11733755.0 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chairman	W. Sekretaruk
Members:	M. Müller
	S. Krischer

Summary of Facts and Submissions

I. The appeal is against the decision of the examining division to refuse the European patent application No. 11 733 755 for lack of inventive step over the document

D1: US 5 448 682 A.

The decision also mentioned other documents without, however, relying upon them in its reasons. Of these, the board will make reference to the document

D2: Misra J *et al.*, "Artificial neural networks in hardware: A survey of two decades of progress", Neurocomputing, vol. 74, no. 1-3, Elsevier Science Publishers, pp. 239-255, online available since 5 May 2010.

II. Notice of appeal was filed on 4 January 2017, the appeal fee being paid on the same day. A statement of grounds of appeal was received on 2 March 2017. The appellant requested that the decision be set aside and a patent be granted on the basis of claims according to a main request or one of two auxiliary requests as filed with the grounds of appeal.

III. In an annex to a summons to oral proceedings, the board introduced three documents of its own motion, in particular

D6: US2006017578 A1,

and informed the appellant of its preliminary opinion that claim 1 of all requests, properly construed, lacked inventive step over D1 in view of common

knowledge or D2, and over other prior art documents such as D6. Clarity objections were also raised.

IV. In response to the summons, with a letter dated 8 October 2018, the appellant filed amended claims according to a main request and three auxiliary requests and requested the grant of a patent on this basis. It also requested that the case be remitted to the examining division for further prosecution so as not to deprive it of an instance in view of the fact that the board had introduced new documents. For justification of this request it referred to Enlarged board of appeal decision G 10/93, item VIII. With its response, the appellant withdrew its request for oral proceedings and requested that a decision based on the state of the file be taken.

V. The oral proceedings were then cancelled.

VI. Claim 1 of the main request reads as follows

"An electrical circuit, comprising:

a neuro-processor chip (202) with a plurality of neuron circuits (102, 106) and a plurality of synapses (104), wherein each synapse (104) connects a pair of neuron circuits (102, 106) of the plurality of neuron circuits (102, 106); and

a removable random access memory (206) connected to the neuro-processor chip (202) storing weights ($\omega_{1..P}$) of the at least one synapse (104), wherein the weights ($\omega_{1..P}$) define, at least in part, a function of the neuro-processor chip (202), wherein the removable random access memory (206) is external to the neuro-processor chip (202) and connected to the neuro-processor chip (202) via an interface circuit (204) and

wherein the removable random access memory (206) comprises a non-volatile memory device."

In addition to claim 1 of of the main request, that of auxiliary request 1 requires, just before the specification of what the weights define,

"... wherein the synapse weights ($\omega_{1..P}$) are trained within the neuro-processor chip (202) ..."

Claim 1 of auxiliary request 2 reads as follows:

"A system, comprising:
a first neuro-processor chip (202) with a first plurality of neuron circuits (102, 106) and a first plurality of synapses (104), wherein each of the first synapses (104) connects a pair of neuron circuits (102, 106) of the first plurality of neuron circuits (102, 106); and a second neuro-processor chip (202) with a second plurality of neuron circuits (102, 106) and a second plurality of synapses (104), wherein each of the second synapses (104) connects a pair of neuron circuits (102, 106) of the second plurality of neuron circuits (102, 106);
a removable random access memory (206) connected to the first neuro-processor chip (202) storing weights ($\omega_{1..P}$) of the at least one synapse (104) of the first synapses, wherein the synapse weights ($\omega_{1..P}$) are trained within the first neuro-processor chip (202), wherein the weights ($\omega_{1..P}$) define, at least in part, a function of the first neuro-processor chip (202), wherein the removable random access memory (206) is external to the first neuro-processor chip (202) and connected to the first neuro-processor chip (202) via a first interface circuit (204) transferring the weights ($\omega_{1..P}$) from the frist neuro-processor chip (202) to the

removable random access memory (206) and wherein the removable random access memory (206) comprises a non-volatile memory device which when connected to a second neuro-processor chip, causes the second neuro-processor chip to execute the function of the first neuro-processor chip (202) based at least in part on the values of the weights ($\omega_{1..P}$) transferred from the removable random access memory (206) to the second neuro-processor chip, via a second interface circuit (204) of the second neuro-processor chip."

Claim 1 of auxiliary request 3 differs from the of auxiliary request 2 in that the "removable random access memory" is now also specified to be "replaceable" and wherein the "training" feature is amended so that it now reads as:

"... wherein the synapse weights ($\omega_{1..P}$) are trained within the first neuro-processor chip (202) for the plurality of neuron circuits (102, 106) and are stored in the replaceable removable random access memory (206) ..."

All request also contain independent method and computer program product claims which correspond in substance to that of the circuit or system claim 1.

Reasons for the Decision

Admission of the amended sets of claims
Article 13(1,3) RPBA

1. It is true that the amended sets of claims were filed "timely" with respect to the time limit within which "Any amendments to [the appellant's] case should be

submitted" according to the board's summons to oral proceedings (see the appellant's letter of 8 October 2018, page 2, penultimate paragraph).

1.1 Nonetheless, they constitute amendments to the appellant's case within the meaning of Article 13(1) RPBA which "may be admitted and considered at the Board's discretion [...] exercised in view of inter alia the complexity of the new subject matter submitted, the current state of the proceedings and the need for procedural economy". While the timeliness of the parties' response to the board's summons may have a bearing on how the board exercises its discretion in view of "the current state of the proceedings and the need for procedural economy", so does the substantive complexity of the new subject matter. Article 13(3) RPBA even states that amendments "shall not be admitted" if they raise new and inappropriately complex issues. Therefore, in contrast to what the appellant seems to assume (*loc. cit.*), there is no unconditional obligation on the board to consider the amendments as to their merits. Also the board's preliminary opinion has not stated or implied that.

1.2 In fact, however, the board judges that the new submissions do not raise new and inappropriately complex issues, in particular none which the board could not deal with in writing, and thus exercises its discretion under Article 13(1) RPBA and admits the newly filed requests.

Remittal for further prosecution

2. The appellant considered it "likely that [it would be] deprived of an instance" by the board introducing newly cited prior art documents unless the case was remitted

to the first instance" and that this would be contrary to the Enlarged Board of Appeal decision G 10/91.

2.1 The board agrees that the appellant may be "deprived of an instance" if a board of appeal was the first and only instance to assess newly introduced prior art documents. However, Article 111(1) EPC gives the board discretion to exercise any power within the competence of the department which was responsible for the decision appealed, irrespective of whether the appellant might be "deprived of an instance". As many board of appeal decisions have stated clearly, the EPC does not provide a "right to two instances".

2.2 Moreover, G 10/93 is consistent with this finding. The passage referred to by the appellant (item VIII) contains the submission of a party in that case rather than the Enlarged Board's position. In fact, the Enlarged Board found rather clearly that "If the board holds that the application is not patentable, it can confirm the decision. There is no obligation to remit the decision to the first instance" (see the very last two sentence of the reasons).

The invention

3. The application relates to neuro-processors equipped with replaceable and removable non-volatile storage for holding the synaptic weights (see for instance paragraphs 1, 6, 12, 21; figures 1 and 2; original claim 14; all references being to the application as originally filed). It is disclosed that the training of the weights can be performed within the neuro-processor (see paragraph 25). After training, the weights may be replicated to another weight memory which may then be used in combination with another neuro-processor.

Thereby, the time- and power-consuming training need not be repeated (paragraph 26). It is also disclosed that the removable memory connected to a neuro-processor may be replaced with another one so that the neuro-processor performs the function defined by the weights stored in this other memory (see for instance original claim 6).

The prior art

4. D1 discloses a neuro-processor with on-chip random access memory, which stores synaptic weights once they have been learnt (see esp. figure 1B, no. 10, and column 1, lines 23-32; column 2, lines 26-29). The weights stored in RAM 10 are transmitted to "each synapse" (see column 1, lines 66-68, and column 2, line 2) in order to be stored in a "D latch" (see figure 1B, no. 80, and column 2, lines 30-33). Thereby, they define the "function of the neuro-processor chip". D1 also discloses the desire for the neuro-chip to carry out several functions (see column 1, lines 13-16 and 19-22).

5. D2 is a survey of hardware support for artificial neural networks (ANNs). It discloses *inter alia* the use and advantage of non-volatile storage ("SRAM" or "static RAM") for synaptic weights (see page 241, right column, paragraph 3; page 242, left column, last paragraph; page 245, left column, paragraph 5; right column, paragraph 3, last sentence), and both on-chip and off-chip learning (page 241, left column, paragraph 3; page 244, right column, paragraph 1, last sentence). In section 4.2 (see page 248, left column, last paragraph), it discusses a chip called ETANN

("Electrically Trainable Analog Neural Network") which uses the services of a host computer for training.

6. Document D6 discloses a neuro-processor with an external and non-volatile weight memory (see D6, figure 1, nos. 22 and 85; paragraphs 19, 33 and 37) and states that the "ANN training can be performed" on an external computer and "loaded onto the" neuro-chip (see paragraph 52).

Claim construction

7. A central feature of the independent claims of all request is that the synaptic weight memory be "removable" - or even "replaceable removable" (auxiliary request 3) - from the neuro-processor chip.

7.1 The board considers it evident that a component removable from a chip cannot be part of the same integrated circuit (IC) as the neuro-processor chip. On the other hand, any component which is not part of an IC is "removable" from it in the sense that it can be separated from the chip without destroying or damaging the latter.

7.2 The appellant challenges this view by referring to paragraph 22 of the application, which discloses that "A synaptic memory 206 may be implemented as a separate and external removable memory, which may be connected to a neuro-processor 2020 through an interface circuit 204", and to figure 2, which depicts the interface circuit labelled with "Data Read Write Address", and by stating that a removable memory needs to be "intended to be removed" and to be "configured to be removable, for example by taking care of the memory not being

damaged when the memory is removed" (see the appellant's letter dated 8 October 2018).

7.3 The board however notes that neither the claims nor, in fact, the application itself, specifies any structural feature (or "configuration") of the synaptic memory that would express (or correspond to) the "intent" that it be removable - or replaceable, for that matter. The board appreciates that the skilled person would understand a component only to be "removable" without destroying any other component. The claims do not allow any conclusion, however, as to how difficult the removal or replacement might be or much time if might take.

7.4 The board therefore sticks to its preliminary opinion (see point 8 of the summons) that any component which is not integrated in the neuro-chip is removable and replaceable within the meaning of the terms as claimed.

Inventive step, Article 56 EPC

8. In its decision, the examining division found that D1 differed from the claimed invention by the provision of an external, removable memory (see point 1.2 of the reasons) and considered this as an obvious solution to the problem of replacing the weights in an neuro-processor (see point 1.4 of the reasons) in view of the fact that D1 suggested the desire to "enable different functionality of neural networks" (see point 1.5 of the reasons).

9. The appellant stressed that the invention implemented the idea of "on-chip learning and off-chip storage",

i.e. of training a set of weights on the chip and storing the trained weights in the removable, non-volatile storage for them to be reusable on other neuro-chips (see grounds of appeal, page 3, paragraph 3, and the paragraph bridging pages 7 and 8). It argued that the weight memory in D1 was an "integral part of the chip" (see page 6, paragraph 7) and that D1 disclosed only on-chip training and gave no indication that weights, once learnt, could be reused on a different chip; on a different chip, the learning had to be repeated (see page 6, last two paragraphs, and page 7, paragraph 1). The skilled person could learn from D2 to perform off-chip training on a host computer and download the training results to the chip (page 7, paragraphs 2-4). However, neither D1 nor D2, separately or in combination, disclosed or suggested a removable RAM for the synaptic weights as claimed (page 7, penultimate paragraph).

10. The board's position is as follows.

Main request

11. It is uncontroversial that D1 is a suitable starting point for the assessment of inventive step. It is also uncontroversial that D1 discloses all the features of claim 1 of the main request except, at best, that the random access memory for storing synaptic weights (see D1, figure 1A, no. 10) is

(a) non-volatile
(b) removable, and
(c) external to the neuro-processor chip.

12. As regards feature (a), D1 does not disclose whether the weight memory is volatile or non-volatile.

12.1 Contradicting this observation, the appellant referred to the "D latch" (figure 1B, no. 80) which was volatile storage and stated that D1 did not contain any prompt for the skilled person to change the D latch to a non-volatile memory (see the appellant's letter of 8 October 2018, page 8, paragraph 2). However, the D latches depicted in figure 1B are part of the circuitry of a single synapse and thus do not store more than one synapse weight. It is RAM no. 10 (see D1, fig. 1A) which corresponds to the claimed random access memory storing the - i.e. all - synaptic weights. Apparently, whether the D latches are volatile is immaterial for the question of whether RAM 10 is. Hence, the appellant's argument is not pertinent.

12.2 D1 discloses that the weights stored in RAM 10 are obtained from learning, which is well-known to be a time-consuming operation. Hence, the skilled person would obviously want to protect the computed weights against loss. As a solution, it would also have been obvious to choose RAM 10 as a non-volatile storage. In the board's judgment, the skilled person would have found this obvious from first principles, from common general knowledge as described in D2, which discloses non-volatile storage for synaptic weights (see the references above), or from D6 which discloses an "external flash memory", i.e. non-volatile storage, for holding synaptic connections weights (see paragraph 37).

13. With regard to feature (b), it is worth noting that D1 discloses the possibility that the "configuration" of figures 1A and 1B is "integrated on a chip", but it does not require integration (see column 2, lines 21-23).

13.1 If "integrated", the RAM is not removable. The goal of D1 to have a "programmable multilayer neural network with which a user can perform multiple functions" is, in the board's view, achieved by the fact that the neuro-processor of D1 has a RAM from which the weights are loaded into the neural network circuit (see column 1, lines 23-26 and 55-58; column 2, lines 30-44; and figure 1A) so that the neural network executes a different function depending on the weight values stored in the RAM. The board thus considers that this goal is insufficient motivation for the skilled person to make the memory removable. In this, the board disagrees with the decision under appeal (see point 1.5 of the reasons).

13.2 However, when the configuration of D1 is not "integrated", which option D1 does not exclude, any of its components will be "removable", at least in a broad sense (see point 7.4 above), RAM 10 included. The board thus opines that "removability" of the weight memory is a necessary consequence of the obvious option to implement the configuration of D1 in a non-integrated manner.

14. With regard to feature (c), the board considers that any memory "removable" from a chip must be construed as "external" to that chip, at least when removed. Moreover, the memory would have to be "connectable" to the chip via suitable wiring which qualifies as the claimed "interface circuit".

15. In view of the above, the board comes to the conclusion that claim 1 of the main request lacks inventive step over D1, Article 56 EPC.

Auxiliary request 1

16. Claim 1 of auxiliary request 1 adds the feature that the synapse weights are trained "on chip".

16.1 The appellant explained that, in its view, D1 and D2 disclosed "on chip learning and on chip storage" and "off-chip learning and off-chip storage" but neither D1 nor D2 - nor any other cited prior art document - disclosed "on-chip learning and off-chip storage" as claimed (see the letter of 8 October 2018, table on top of page 13; with respect to D6, see page 16, paragraph 4; with respect to D1, see also the grounds of appeal, page 6, penultimate paragraph, to page 7, paragraph 1).

16.2 Even though the board has no occasion to doubt the appellant's analysis of D1, D2 or D6, it does not accept that the questions where the synaptic weights are "learned" and where they are stored are intimately related. More specifically, the board disagrees with the suggestion that off-chip storage as claimed would become non-obvious if the learning was carried out on-chip. In particular the preceding argument why the skilled person would find it obvious in D1 to practice RAM 10 as a removable memory does not depend on whether the synaptic weights were generated on- or off-chip.

16.3 Therefore - and on the understanding that D1 itself discloses on-chip learning so that the added feature does not represent a further difference over D1 - the board considers that the above assessment of claim 1 of the main request carries over directly to claim 1 of auxiliary request which, hence, also lacks inventive step, Article 56 EPC.

Auxiliary request 2

17. The weights stored in RAM 10 of D1 define, in the words of claim 1, "at least in part, a function of the [...] neuro-processor chip". Moreover, D1 discloses that the weights "obtained from learning" are stored in RAM 10 (see column 2, lines 26-29). Assuming on-chip learning, this means that the weights are transferred from the [...] neuro-processor chip [...] to the removable random access memory". The board has found above that it would have been obvious in the context of D1 to use a removable and non-volatile RAM. The board takes the view that a component which is "removable" from one processor can, by the same token, also be connected to another one. Although claim 1 does not specify that the two neuro-processor chips are equivalent, this is clearly intended and certainly not excluded. When the memory component is non-volatile, and absent any further limitations on the circuitry in question, it follows that the re-connected memory component will "cause[] the second neuro-processor chip to execute the function of the first neuro-processor chip [...]". Therefore, also claim 1 of auxiliary request 2 lacks inventive step, Article 56 EPC.

Auxiliary request 3

18. Claim 1 of auxiliary request 3 differs from that of auxiliary request 2 only in that the random access memory is not only "removable" but also "replaceable". As explained above, the board takes the view that this term does not, on its own, imply any structural limitations of the claimed circuitry that could change the preceding inventive step assessment, Article 56 EPC.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chairman:

B. Atienza Vivancos W. Sekretaruk

Decision electronically authenticated