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

**Datasheet for the decision
of 14 June 2024**

Case Number:	T 1952/21 - 3.5.06
Application Number:	18174351.9
Publication Number:	3572982
IPC:	G06N3/00, G06N3/04
Language of the proceedings:	EN

Title of invention:
MACHINE LEARNING SYSTEM

Applicant:
Robert Bosch GmbH

Headword:
Reinforcement learning/BOSCH

Relevant legal provisions:
EPC Art. 52, 56
RPBA 2020 Art. 13(2)

Keyword:
Claims - clarity (yes)
Reinforcement learning a technical field (no)
Inventive step - all requests (no)

Decisions cited:

G 0001/19, T 1326/06, T 0702/20, T 1294/16

Catchword:

Case Number: **T 1952/21 - 3.5.06**

D E C I S I O N
of Technical Board of Appeal 3.5.06
of 14 June 2024

Appellant: Robert Bosch GmbH
                                Postfach 30 02 20
(Applicant)   70442 Stuttgart (DE)

Representative: Robert Bosch GmbH
                                C/IPE41
                                Postfach 30 02 20
                                70442 Stuttgart (DE)

Decision under appeal: **Decision of the Examining Division of the
                                European Patent Office posted on 15 July 2021
                                refusing European patent application No.
                                18174351.9 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chairman   M. Müller
Members: T. Alecu
                    K. Kerber-Zubrzycka

Summary of Facts and Submissions

I.   The appeal lies from the decision of the Examining
        Division to refuse the application. With the statement
        of grounds of appeal the Appellant requested that the
        decision of the Examining Division be set aside and
        that a patent be granted on the basis of the main
        request, identical to the main request underlying the
        decision under appeal, or on the basis of one of three
        auxiliary requests, all filed with the statement of
        grounds of appeal.

II.   The Examining Division refused the application for lack
        of clarity and for lack of inventive step in view of

        D1: Mnih V. et al.: "Asynchronous Methods for Deep
        Reinforcement Learning", Proceedings of the 33rd
        International Conference on Machine Learning, pages
        1928-1937, 2016.

III.   In a communication accompanying a summons to oral
        proceedings, the Board informed the Appellant of its
        provisional opinion that the main and the first two
        auxiliary requests were not allowable for lack of
        inventive step, that it tended not to admit the third
        auxiliary request, which added a new independent claim,
        but that in substance also this last request did not
        appear to be allowable.

IV.   With the letter of 15 May 2024, the Appellant filed a
        new third auxiliary request replacing the previous one.

V.   Claim 1 of the main request defines:

A machine learning system (10), comprising:
an input unit (20);
a processing unit (30);
and an output unit (40);
wherein, the input unit is configured to provide the
processing unit with input data;
wherein, the processing unit is configured to process
the input data to generate processing path input data;
wherein, the processing unit is configured to implement
a first processing path comprising a feed-forward
neural network to process the processing path input
data to generate first intermediate data;
wherein, the processing unit is configured to implement
a second processing path comprising a feed-forward
neural network to process the processing path input
data to generate second intermediate data, wherein said
feed-forward neural network comprises stochastic units;
wherein, the processing unit is configured to implement
a value output path comprising a feed-forward neural
network to process the first intermediate data and the
second intermediate data to generate value output data;
wherein, the processing unit is configured to implement
a policy output path comprising a feed-forward neural
network to process the first intermediate data and the
second intermediate data to generate policy output
data; and
wherein, the output unit is configured to output the
value output data and output the policy output data.

VI.   Claim 1 of the first auxiliary request differs from
        claim 1 of the main request by specifying that "*the*
        stochastic units comprise stochastic activations".

VII.   Claim 1 of the second auxiliary request differs from claim 1 of the main request by defining

"*A machine learning system (10) <u>for reinforcement learning</u> ..*"

VIII.   Claim 1 of the third auxiliary request differs from claim 1 of the main request by defining

"*A machine learning system (10) <u>for reinforcement learning of control policy output data configured for controlling an engine, or valve, electrical circuit, or heating system, or robotic arm, or aerial drone, or humanoid robot, or self-driving vehicle</u> ..*" and by further defining in the last two features the policy output data as "*<u>control</u> policy output data*".

Reasons for the Decision

The application

Background and prior art

1.   The application relates to reinforcement learning. In reinforcement learning (see e.g. the published application, paragraph 2), an agent explores the environment according to a policy, determining which action the agent takes (e.g. move right) at every juncture as a function of its current state (e.g. its position in the environment). The agent receives rewards, positive or negative. In this way it can "learn" the value of the various actions and states. The goal of training is to maximize a value function which reflects the expected sum of rewards given a certain action.

2.   The application (see paragraph 4) builds upon the
        method of D1, called A3C (asynchronous advantage actor-
        critic). That method separately approximates the policy
        and value models as neural networks. The raw input
        (describing the environment) is preprocessed in
        sequence by a feed forward network (CNN - for spatial
        input description) and a recurrent neural network (LSTM
        - for time dependencies). The result is fed to the
        value and the policy networks (see the published
        application, figure 1).

2.1   Developments of that method, termed NoisyNet A3C in the
        current application, inject randomness into the
        training by using stochastic weights (e.g. by adding
        random noise or using stochastic models) in the policy
        and value networks. This allows for further exploration
        of the parameter space (see the published application,
        paragraph 5 and figure 2).

Contribution

3.   According to the application (paragraph 6), such
        stochasticity in the dynamics of the controlled system,
        and also a lack of training data, can lead to imperfect
        decisions. The application therefore proposes (para-
        graphs 10 to 13, 95 to 98, and figure 5) the use of a
        feed-forward intermediate layer between the LSTM layer
        of A3C and the policy and value networks, comprising a
        (standard) deterministic CNN and a CNN with stochastic
        units, exemplified as neurons with stochastic activa-
        tion functions, working in parallel. The outputs of the
        two networks are concatenated and fed to the policy and
        value networks. Alternatively, the full intermediate
        layer may be stochastic. This intermediate layer is

said to provide for better exploration, faster
convergence, and better policies (paragraphs 15 to 17).

Main request: clarity

4.   The Examining Division decided (reasons 11) that
        claim 1 lacked clarity as the term "*stochastic unit*"
        did not have a generally accepted meaning in the art
        and the application did not provide a clear definition
        either, in particular as to where the stochasticity
        "originated from".

5.   In the Board's view, the skilled person would
        understand that a "*neural network compris[ing]
        stochastic units*" was one comprising neurons the output
        of which is partly determined by a stochastic element.
        There is no need to specify the exact origin of
        stochasticity (e.g. stochastic weight models or added
        random noise - see points 2 and 3 above) for the scope
        of the claim to be clear - it covers any possible
        "origin". Therefore, the Board does not follow the
        objection of the Examining Division.

Main request: inventive step

6.   The Examining Division acknowledged the differences to
        D1 as being those related to the intermediate layer
        containing stochastic units (see the decision, reasons
        12.2, but also point 3 above). However, it reasoned
        (reasons 12.3) that they were "*limited in their effect
        to the way how a mathematical model in the form of a
        neural network internally processes abstract data*", so
        that they could not contribute to the technical
        character of the claimed invention. The Examining

Division also stated that the claim did not "*serve a specific technical purpose*".

6.1 In response to the Appellant's arguments, it considered (reasons 12.4) that reinforcement learning was not actually claimed, and anyway it did not "*refer to a technical field but to a machine learning approach*". Also, the various alleged advantages (see e.g. point 3 above) were not derivable from the claimed matter (decision, reasons 12.6 and 12.8).

6.2 Further, the Examining Division was not convinced that the case law related to simulations or cryptography, esp. RSA, applied to the present case (see the decision, reasons 12.5 and 12.9, respectively 12.7).

The Appellant's arguments

7. The Appellant argued that the distinguishing features contributed to the technical character of the invention for the following reasons.

8. First, the system design was motivated by technical considerations of the internal functioning of the computer. A computer was a deterministic system and therefore limited to deterministic operations, and the claimed stochastic units overcame that limitation by creating the desired stochastic property within the deterministic computer (statement of grounds of appeal, pages 11 and 12; and the Appellant's letter of 14 May 2024, section 1.A.a).

9. Secondly, the claim implicitly defined reinforcement learning: from the required value and policy output paths the skilled person would understand that the

claim relates to reinforcement learning (see e.g. the statement of grounds of appeal, pages 8 and 9).

10. The claimed approach brought advantages in this field due to the claimed layer comprising stochastic units. There should not be a general requirement, neither in the field of reinforcement learning nor, more generally, in the field of artificial intelligence (AI), to provide experiments as evidence for technical effects. The established standard for establishing alleged advantages required only *sufficient evidence* (in support, the Appellant referred to the case law book of the Board of Appeals, 10th Edition, Chapter I.D-4.2; see letter of 14 May 2024, section 1.B.i). The Appellant argued that logical reasoning alone could constitute the required sufficient evidence.

10.1 Accordingly, the Appellant offered theoretical considerations for the present case (letter of 14 May 2024, section 1.B.a). The Appellant stated that stochasticity during training allowed for a wider exploration of the field of possible actions. It also allowed the optimization algorithm to escape local optima and find the global optimum. The obtained solution balanced deterministic and probabilistic signals, and by gaining a probabilistic perspective enabled it to assess ambiguous scenarios more effectively and to generalize better.

10.2 The corresponding scientific publication of the inventors:

Shang W., van der Wal D., van Hoof H., Welling M. (2020) Stochastic Activation Actor Critic Methods

showed (abstract, table 2 and figure 2) that these
advantages could indeed be obtained, using benchmarks
typically used in the art (statement of grounds of
appeal, middle of page 12).

10.3   Although these results were related to video games, the
person skilled in the art was able to obtain these
advantages for any technical field; he or she could
start with the hyperparameter sets of the prior art and
modify them by trial and error, until a working
configuration was obtained. This was within the skills
of the skilled person, who must be considered to have
experience in parametrizing neural networks.

10.4   For example, the invention was grounded in a technical
project, namely that of an ABS breaking system.
Applying the principles of the application, the
inventors were able to reduce the breaking distance
significantly.

11.   These advantages were to be taken into account as
potential technical effects in the sense of G 1/19, as
they occurred when the system was used as intended,
namely for reinforcement learning (statement of grounds
of appeal, page 10).

12.   The field of machine learning, and in particular
reinforcement learning, was technical (see e.g. the
statement of grounds of appeal, pages 12 to 14).

12.1   The Appellant argued this first by comparison with
cryptography (RSA), which according to case law was a
technical application. In particular T 1326/06,
reasons 6.4, stated the following: "*RSA was a
breakthrough in the development of cryptography: RSA is
regarded as the first practicable, concretely*

implementable asymmetric cryptosystem and is now a central component in numerous cryptographic security systems. The mathematics underlying RSA thus serves directly to solve a concrete technical problem".

12.2   In the Appellant's view, reinforcement learning was a similar breakthrough for autonomous systems, where it is "*the only practicable and concretely implementable solution*". It was therefore incorrect to require a limitation to a specific application. In fact "*reinforcement learning has crossed much further the border between technical and non-technical than RSA*" – i.e. is more technical than the latter and more remote from mere mathematics – "*as it uses many more technical aspects to achieve its purpose. [...] both an agent and an environment of the agent are required and the agent control is learned as well as improved, while RSA only requires data in terms of public-/private keys and electronic messages*". These parallels should lead to the conclusion that also reinforcement learning is a technical field, even if the use of reinforcement learning is not the same as that of RSA.

13.   The Board suggested in its preliminary opinion (see also below) that the appropriate starting point for the assessment of the technical character and potential technical contributions of reinforcement learning was the decision G 1/19, in which the Enlarged Board of Appeal had made several observations on the examination of computer-related inventions in general, in particular that a specific technical purpose may be needed to establish a technical effect.

14.   In response, the Appellant argued the following in its letter of 14 May 2024 (section 1.A.b).

14.1   It might be correct that G 1/19 required a specific
        technical purpose to establish a technical contribu-
        tion. But this requirement, strictly applied, would
        mean that some earlier case law finding certain fields
        (such as RSA) to be "patent-eligible" would no longer
        be applicable. The Appellant questioned whether this
        development was "*aligned with a teleological interpre-
        tation of the EPC*", in particular when (see page 3 of
        that letter) "*the understanding of society of the term
        technical or technology becomes broader over time due
        to the exponential technologic[al] advancement (also
        referred to as technological evolution)*" and, at the
        same time, "*the case-law of the EPC steadily* narrows
        its understanding" of technology. The Appellant also
        asked: "*Is this in line with the original intention of
        the EPC, or is it contrary to the idea of the EPC?*"

14.2   In the oral proceedings the Appellant argued that the
        EPC was old and written with traditional, for instance
        mechanical, inventions in mind. It was understood that
        such inventions existed to make people's lives easier,
        for instance by supporting or taking over manual tasks.
        Nowadays software implementing artificial intelligence
        (AI) often has the same purpose, albeit emulating a
        different class of human capabilities, and this trend
        would intensify in the future. Although AI methods
        indeed relied heavily on applied mathematics and (big)
        data processing, they were applicable in many technical
        fields and thus of independent value.

14.3   AI inventions therefore deserved patent protection,
        which was also desirable in order not to discourage
        their publication, which was beneficial for the public.

14.4   Moreover, it was a question of fairness how narrow a
        technical application or purpose as required by G 1/19
        would have to be. Limiting the protection granted to an
        AI invention to a very specific technical application
        did not provide fair protection, if it relied on ideas
        which are broadly applicable. This was the case here
        because the invention was an improvement of reinforce-
        ment learning which was generally applicable, e.g. to
        cars and robotics.

14.5   In summary, the Appellant asserted a disconnect between
        the patent system and the real world. It argued that
        "everyone in the real world" would acknowledge AI or
        machine learning as "technical" and that the case law
        needed to recognise this evolution of technology (see
        also the statement of grounds of appeal, page 15,
        fourth paragraph).

15.   The distinguishing features therefore had to be accep-
        ted as solving a technical problem. They were also not
        disclosed or rendered obvious by the cited prior art.
        Hence the claimed invention involved an inventive step.

The Board's opinion

16.   The Appellant's allegation that stochastic units over-
        come the limitations of the "*deterministic*" computer,
        goes beyond reinforcement learning and relates to a
        computer in general.

17.   The Board remarks that pseudorandom number generators
        were known to the person skilled in the art. Their use,
        in general or in the more specific context of
        "*stochastic units*" (which the Appellant acknowledged to
        be known in the art, see the statement of grounds of
        appeal, page 4), does not change in substance the

computer, which remains as "*deterministic*" as any
conventional computer. So the Board cannot see a
contribution on this level.

18.   On the more narrow level of reinforcement learning, the
       non-deterministic behaviour of the claimed system is
       considered below.

19.   The Board agrees with the Appellant that the skilled
       person would understand the claimed system to be one
       "*for", i.e. meant to be used in, "reinforcement
       learning*". The Appellant submission is, in a nutshell,
       that this field is technical and that the claimed
       invention makes improvements in this field.

20.   The system for reinforcement learning as claimed is a
       neural network, comprising various sub-networks,
       implemented on a computer. The network, as a whole,
       defines a mathematical function mapping inputs into
       outputs. Effectively, the claim is to a mathematical
       method implemented on a computer.

21.   Considering this, the Board holds that the Enlarged
       Board decision G 1/19, addressing the patentability of
       computer-implemented mathematical models for simula-
       tion, should be the starting point when assessing the
       technical character of reinforcement learning. It is
       commonly accepted (also by the Appellant, see statement
       of grounds of appeal page 10, bottom) that a large part
       of the findings in G 1/19 apply to any computer
       implemented inventions.

22.   In G 1/19, the Enlarged Board of Appeal stated (reasons
       137) that (simulation) models by themselves are not
       technical but that "*they may contribute to technicality
       if, for example, they are a reason for adapting the*

computer or its functioning, or if they form the basis for a further technical use of the outcomes of the simulation". However, "*such further use has to be at least implicitly specified in the claim*".

23.   The implied use of the system in reinforcement learning requires, as the Appellant argued, an agent acting in an environment (see point 12.1 above). However, the agent and its environment need not exist in the real world, and can be completely virtual, e.g. part of a simulation model (a simulated agent acting within a simulated environment) or even a completely imaginary video game. The Board notes that both the prior art (see D1, section 5.1) and the scientific paper corresponding to the application referred to by the Appellant present results on video games. The concept of reinforcement learning in general does not imply a technical context.

24.   The Board has already explained above that the functioning of the computer, or the computer itself, are not adapted. A further technical use is also not implied by the claim. So, even if the advantages in reinforcement learning brought forward by the Appellant were to be acknowledged (which is not the case, see below from point 32 on), the Board must conclude, on the basis of G 1/19, that the claimed system does not solve a technical problem.

24.1   This conclusion is consistent with that in the case T 702/20, which is in many ways similar to the present one, where this Board (in a different composition) decided, also following G 1/19, that a trained machine learning model, namely a neural network, can "*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" (T 702/20, Catchword; see also reasons 12 and 17 to 19).

25.   The Appellant also argued that reinforcement learning was technical based on an analogy with the case law regarding cryptography, in particular RSA (see 12 above).

25.1   The Board notes that, as the Appellant also acknowledged, notwithstanding certain similarities, RSA and reinforcement learning are different and serve different purposes. In particular, RSA and other cryptographic methods have a specific, and at least implied, purpose, namely data security. This is not the case for reinforcement learning. So the findings regarding RSA cannot directly be transferred to reinforcement learning.

25.2   It is therefore immaterial for the present decision whether individual Board of Appeal decisions relating to RSA are still applicable after G 1/19 or whether, as the Appellant seemed to imply, they are now wrong, i.e. "bad law".

26.   The Appellant's opinion that decision G 1/19 has narrowed the scope of patentable subject matter and that this is in conflict with the evolution of technology and with a teleological interpretation of the EPC is noted. However, before the Board can deviate from the interpretations or explanations of the EPC given in G 1/19 it has to refer a question to the Enlarged Board (Article 21 RPBA). The Appellant did not propose a question to be referred, nor did it request that a suitable question be referred.

27.   The Board itself sees no reason to deviate from G 1/19
        in the present case.

27.1   The Appellant's argument that it should be possible to
        patent mostly abstract, mathematical inventions without
        a limitation to a specific technical application if
        they are generally applicable and have practical
        utility for a wide range of new products, may, from a
        business perspective, be a legitimate one. Although it
        may be assumed that the Appellant would find
        substantially less desirable an equally broad patent
        when held by a competitor.

27.2   But it was the lawmaker's choice to exclude from
        patentability, albeit only "as such", mathematical
        methods and programs for computers (see Articles 52(2)
        and (3) EPC).

27.3   Mathematical methods have always been generally appli-
        cable (e.g. Pythagoras' theorem used to calculate dis-
        tances) and been applied in many new - and undoubtedly
        technical - inventions. This did not prevent the
        legislator to list mathematical methods amongst the
        things which, as such, are not to be considered
        inventions. The fundamental nature of mathematical
        methods and their wide applicability may in fact have
        been a reason *for* excluding them from patentability.

27.4   The Board accepts that the use of the term *technical* in
        the case law of the Boards of Appeal may differ from
        its use elsewhere in society, especially from its
        colloquial use. However, this does not mean that the
        Boards of Appeal interpret the law incorrectly: it is
        common place that the legal interpretation of a term
        may differ from its colloquial meaning. In particular,
        the Boards use the term "non-technical" to denote

matter excluded under Article 52(2) and (3) EPC. Any alternative interpretation of the terms "technical" and "non-technical" can only be used to justify the patentability of subject-matter to the extent that it does not contradict the law, in particular the exclusion of mathematical methods.

First and second auxiliary requests

28.   The amendments in the first and second auxiliary requests cannot change the conclusion on inventive step as their substance has already been considered (see points 4, 5 and 19 above).

Third auxiliary request

Admittance

29.   With the statement of grounds of appeal the Appellant had filed a third auxiliary request containing two independent system claims, of which the first one was similar to the one in the corresponding request underlying the decision. In its preliminary opinion, the Board indicated that it saw no reason for the presence of the second independent claim, which also caused issues with Rule 43(2)b) EPC, and was inclined not to admit this request.

30.   With the letter of 14 May 2024, the Appellant filed the current third auxiliary request, which is the same as the previous one, but with the second independent claim deleted. This removes the only reason the Board had advanced for non-admitting the previous request, without introducing any new issues. For that reason,

the Board admits this request (Article 13(2) RPBA, see also T 1294/16).

Inventive step

31.   In substance, the claim now specifies the learned "*control policy output*" to be "*configured for controlling*" one of several technical systems (e.g. "*an engine*"), but gives no details about either the specific technical system or its model.

32.   The Appellant argues that since a technical use is specified, a technical effect must be acknowledged. That is because the claimed approach has advantages for any technical field (see point 10 above).

33.   The Board disagrees that any advantage of the claimed invention is established over the whole breadth of the claim or even for any of the broadly defined claimed technical systems.

33.1   The theoretical arguments advanced only make it credible that in some cases, depending on the considered scenario, i.e. the environment and the task at hand, on the configuration of the processing paths, stochastic and deterministic, of the output paths, and on the manner in which stochasticity is implemented, these advantages may be achieved. This is because the arguments themselves are based on assumptions which are not developed, such as the structure of the optimization space, the type of stochasticity, the basic deterministic algorithm, the extra effort that may be required for stochastic exploration in view of the extra rewards obtained etc.

33.2   Also, it is well known in machine learning that no
        optimization algorithm is better than another one
        across all possible instances, or uses (a statement of
        this type being generally referred as a "no free lunch"
        theorem). This can also be seen in the practical advice
        given in the corresponding paper of the inventors (see
        Section 7): "*Stochastic activation is a general
        approach to improve A3C but not the panacea to every
        environment and task*".

33.3   Notably, the authors found it worth saying this word of
        caution even in the context of controlled video games
        environments. Real-world scenarios may present further,
        unexpected challenges. The Board notes that, during the
        oral proceedings, the Appellant did refer to results in
        real world scenarios (ABS braking), but did not give -
        let alone submit - any further details about them, or
        any explanation as to how they would relate to the
        disclosure of the current application.

34.   So the Board does not see sufficient evidence, be it by
        theoretical considerations, or by experiments, to con-
        clude that a technical effect is present over the full
        breadth of the claim. In fact, neither the application
        nor the corresponding paper of the inventors provide
        sufficient evidence for a technical effect in any
        specific technical field.

Conclusion

35.   None of the requests are allowable for lack of
        inventive step.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chairman:

L. Stridde   Martin Müller

Decision electronically authenticated