

Internal distribution code:

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

**Datasheet for the decision
of 22 May 2023**

Case Number: T 0761/20 - 3.5.06

Application Number: 11183934.6

Publication Number: 2439684

IPC: G06N99/00

Language of the proceedings: EN

Title of invention:

Automated assessment of examination scripts

Applicant:

The Chancellor, Masters and Scholars of the
University of Cambridge

Headword:

Automated script grading/UNIVERSITY OF CAMBRIDGE

Relevant legal provisions:

EPC Art. 52(1), 52(2), 56, 112(1)(a)
RPBA 2020 Art. 13

Keyword:

Patentable invention - field of technology - computer
implemented invention - mathematical method
Inventive step - (no)
Referral to the Enlarged Board of Appeal - (no)

Decisions cited:

G 0001/19, T 0641/00, T 1784/06, T 0489/14, T 0755/18,
T 0702/20

Catchword:

According to G 1/19, a *direct* link with physical reality is not required for a technical effect to exist. However an at least *indirect* link to physical reality, internal or external to the computer, is required. The link can be mediated by the intended use or purpose of the invention ("when executed" or when put to its "implied technical use"). (see point 20)



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 0761/20 - 3.5.06

D E C I S I O N
of Technical Board of Appeal 3.5.06
of 22 May 2023

Appellant:
(Applicant)
The Chancellor, Masters and Scholars of the
University of Cambridge
Cambridge ESOL Examinations
Syndicate Buildings
1 Hills Road
Cambridge
Cambridgeshire CB1 2EU (GB)

Representative:
MacDougall, Alan John Shaw
Mathys & Squire
The Shard
32 London Bridge Street
London SE1 9SG (GB)

Decision under appeal: **Decision of the Examining Division of the
European Patent Office posted on 14 November
2019 refusing European patent application No.
11183934.6 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 sole request underlying the contested decision was rejected for a lack of inventive step over document
- D2: Briscoe et al., "Text processing tools and services from iLexIR Ltd", Proceedings of the LangTech Conference 2008, pages 145-148
- II. 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 claims subject to the decision under appeal, which were also re-filed with the statement of grounds of appeal.
- III. In a communication accompanying a summons to oral proceedings, the Board indicated its provisional opinion that claim 1 lacked an inventive step. In particular, it stated that it did *"not see that the claimed method, as a whole, brought a technical contribution to any field of technology, the only provided results relating to the task of script grading which is not technical in nature"*. The Board also raised objections under Article 84 EPC, lack of support.
- IV. With a letter of 14 November 2022, the Appellant filed a new main request and five auxiliary requests. The amendments were intended as a response to the objections under Article 84 EPC raised by the Board.

V. As to inventive step, the Appellant provided arguments that the claimed invention provided a contribution to the field of "educational technology". Should the Board not accept that, the Appellant requested that the following questions be referred to the Enlarged Board of Appeal under Article 112(1) (a) EPC:

(a) What are the characteristics of a field of human activity that make it fall within the definition of being "a field of technology" under the EPC and how do these characteristics differ from the characteristics of a field of human activity that is not considered to be within a field of technology under the EPC?

(b) Is educational technology a "field of technology" under the EPC?

VI. On 12 December 2022 the Appellant indicated that they would not attend the oral proceedings and asked for a decision based on the state of the file. The oral proceedings were subsequently cancelled by the Board.

VII. Claim 1 of the main request defines:

A computer-implemented method of grading scripts (145) comprising text, the method comprising:

training an automated computerized text assessment system to grade text of scripts, the training including, by a computer device (900):

receiving (210) a plurality of training linguistic vectors ($x_1, x_2, x_3, \dots, x_n$) each training linguistic vector comprising a plurality of numerical values representing linguistic features of text within a training script (105);

receiving, for each of a plurality of pairs of said training linguistic vectors, ranking data ($x_1 <_r x_2$) that defines which one of the pair of training linguistic vectors is representative of a better training script (105);

generating a plurality of difference training vectors ($x_j - x_i$) each difference training vector being calculated as a difference between a pair of said training linguistic vectors ranked by said ranking data; and

performing an iterative process (220 - 270) to adapt a weight vector to a trained model weight vector by:

- i) calculating a dot product between a current weight vector and each difference training vector to generate a respective scalar value for each difference training vector;
- ii) determining (230), for each difference training vector, if the current weight vector misclassified the difference training vector in dependence upon a comparison result obtained by comparing the scalar value for the difference training vector with a threshold;
- iii) generating (240) an aggregate vector, \tilde{a} , by summing the difference training vectors that said determining determines are misclassified and normalizing with a current timing factor;
- iv) updating (250) the current weight vector by summing the current weight vector with the generated aggregate vector;
- v) reducing (260) the timing factor; and
- vi) repeating steps i) through v) until the timing factor reaches a predetermined value, whereupon the then current weight vector becomes (280) said trained model weight vector;

generating a linguistic vector comprising a plurality of numerical values representing linguistic

features of text of an input script (145) that is to be graded;

calculating, a dot product between the trained model weight vector and the linguistic vector for the text of the input script that is to be graded to generate a scalar value for the input script; and

outputting a grade for the input script using the scalar value generated for the input script (145).

- VIII. The auxiliary requests define steps v) and vi) of the update procedure in more detail. Their exact wording is not pertinent to the current decision (see the corresponding part of the reasons below).

Reasons for the Decision

The application

1. The application relates to automated assessment of scripts written in examination, in particular English for Speakers of Other Languages (ESOL) examinations (paragraphs 1 to 3).
- 1.1 The system comprises a feature analysis module, denoted as RASP (robust accurate statistical parsing) which extracts and numerically quantifies linguistic features of text (paragraphs 52 to 56) to form a feature vector.
- 1.2 This feature vector is used to grade scripts on the basis of discriminative models, such as SVM or large margin perceptrons, including a variant, said to be new, called the Timed Aggregate Perceptron (TAP, see paragraph 28). In the TAP training procedure, unlike in standard perceptron training, a timing parameter reduces the update rate as a function of how far the process has progressed, of the magnitude of the increase

in empirical loss, and of the balance of the training distributions (paragraph 36). This has the role of providing an approximate solution that prevents overfitting (by early stopping).

- 1.3 The application describes embodiments with binary outputs based on SVM or TAP, useful for pass/fail grading systems (paragraphs 24 to 39), and an embodiment denoted as a modification of the TAP using preference ranking (paragraphs 41 to 49).
- 1.4 In the latter embodiment, the perceptron's success is measured by its ability to correctly rank pairs of training samples on the basis of its scalar output; this scheme is conceptually aimed at reducing errors in relative grading (i.e. the decision which test to assign a higher score) as opposed to errors in absolute grading. The output of a perceptron is in essence the result of a dot product between the learned weight vector and the incoming sample. In the standard perceptron, to reduce the ranking errors, the weight vector is updated in the direction of the misclassified samples; in the proposed variant, the update direction is provided by the sum of the difference vectors between the samples of the misclassified pairs. This variant can be used both for binary fail/pass grading and for non-binary grading.
- 1.5 The application discloses performance assessments of the described methods (paragraphs 62 to 73) based on how well its results correlate with those of prior art systems and of human markers (or examiners/raters) (see Tables 4 and 5, paragraphs 63 and 71). According to those results, the preference ranking TAP model outputs grades that correlate with those provided by human markers almost as well as the human markers' grades

correlate with one other. Also, the preference ranking TAP outperforms TAP on a binary task, while binary TAP and SVM outperform prior art systems.

- 1.6 The requests on file are all based on the TAP preference ranking model.

Main request: admittance (Article 13 RPBA 2020)

2. The only amendment carried out in the present main request was to replace, in independent claims 1 and 10, the term "combining" by the term "summing". This addresses and overcomes the lack of support objection raised by the Board (for the first time) in its preliminary opinion at point 2.1. In view of this, the Board decides to admit the amended main request.

Inventive step

The decision under appeal

3. The Examining Division has started its inventive step analysis from document D2 and acknowledged (decision, reasons 2.2) that a number of features were not disclosed by D2. These features are those defining the preference ranking variant of the TAP, the Examining Division considering that D2 disclosed an automated grading system using TAP.

- 3.1 The Examining Division then argued that

"2.3 The distinguishing features above are merely representing mathematical or linguistic operations and entities, implemented on a general-purpose computer. Said features are not directed to a specific technical implementation going beyond the common use of a

general-purpose computer, and their implementation would be, therefore, straightforward for the person skilled in computer science.

2.4 Furthermore, the above differences are not limited to a technical purpose, since it is not specified how the input and the output of the sequence of mathematical or linguistic steps of this difference relate to a technical purpose, so that said difference would be causally linked to a technical effect. In particular, it is noted that grading text scripts is not considered as serving a technical purpose, in the first place."

- 3.2 Thus, it considered that neither the features themselves nor their claimed purpose were technical, so that they did not contribute to a technical effect, and that their implementation on a computer was straightforward.

The Appellant's arguments

4. The Appellant disagreed both with the assessment of the differences in view of D2, submitting that more features distinguished the claimed invention over D2, (statement of grounds of appeal, points 9 and 22 to 26), and with the analysis regarding technicality reproduced above (statement of grounds of appeal, points 12 to 21).

- 4.1 Regarding the former, the Appellant indicated that D2, while using the RASP engine, did not provide for the extraction of linguistic vectors (points 22-23), and that while D2 taught the use of TAP for text classification, it did not teach using TAP for script grading

(points 24-25), nor did it teach ranking between different scripts (point 26).

- 4.2 Regarding the latter the Appellant submitted (statement of grounds of appeal, points 12 to 15) that the problem addressed is not that of grading scripts per se, acknowledging that "*the manual process of grading scripts by a human marker may not be technical*" (see letter of 14 November 2022 point 11), but that of "*providing a computer system that can automatically grade text scripts [and provide grades] that correlate well with the grades provided by human markers*". Also, the distinguishing features reflected "*further technical considerations*", for which case it stated that G 3/08 "*guaranteed*" that a technical character is present (presumably referring to point 13.5.1 of the reasons).
5. After receiving the Board's provisional opinion, with the letter of 14 November 2022 the Appellant submitted (point 9) that the question at the heart of the Board's opinion seemed to be "*what is a technical field?*" and argued that the invention provided a technical contribution in the field of "*educational technology*", defined as "*the combined use of computer hardware, software, and educational theory and practice to facilitate learning*". This field drew from "*practical education experience*" as well as from "*theoretical knowledge from various disciplines such as communication, education, psychology, sociology, artificial intelligence, and computer science*" and "*encompass[ed] several domains including learning theory, computer-based training, online learning, and m-learning, where mobile technologies are used*".

The Appellant also stated (point 10) that there was "a long tradition going back to the 1940s of getting machines to grade multiple-choice questions - i.e. using OMR sheets and readers. This mechanical approach is taken for granted nowadays. Inventions of the kind the applicant has come up with can handle much more complex responses including scripts. This is a rapidly expanding area of Educational Technology that deploys novel uses of AI. The applicant's invention contributes to this field by providing a system that can automatically grade a script in a manner that correlates better with the grades provided by a human marker in comparison to the prior art techniques. [...] with the automatic and accurate grading provided by the present invention, the student can be provided with near instantaneous feedback which improves their learning of the subject".

The Board's opinion

Differences and technical problem

6. As the Board understands the argument of the Examining Division, it does not depend on whether the differences to D2 also comprise the ones advanced by the Appellant, as the claim as a whole can be said to only define "mathematical or linguistic steps" used for "grading text scripts". This means that, if the argument of the Examining Division is correct, the claim as a whole is not "causally linked to a technical effect".
7. Also the Appellant, challenging the finding of the Examining Division, refers to the claim as a whole when it states an alleged contribution to the art and the corresponding technical problem solved. This is appropriate, as the specific effects of any distinguishing

features over D2 are only relevant for inventive step if it can be acknowledged at all that a technical problem is solved. If that is the case, the differences themselves might give rise, for instance, to an argument that the results according to the invention correlate *better* with those of human markers than the prior art methods (instead of merely "well").

8. The Board shares this view and will therefore also address the claim in its entirety to assess whether a combination of features solving a technical problem can be identified.
9. The claim defines a method of automated script grading using machine learning, which is effectively a computer implemented process. Such processes may have technical effects - and thus be deemed to solve a technical problem - at their input or output, but also by way of their execution (see G 1/19, reasons 85). A technical effect may also be acknowledged in view of their purpose, i.e. an (implied) technical use of their output (see G 1/19, reasons 137).

Technical effects "within the computer"

10. The claimed method contains steps for extracting numerical "linguistic" vectors from scripts (for all considered samples, training scripts and scripts to be graded), a step of training a perceptron, and a step of using the perceptron to grade the scripts.
 - 10.1 The extraction of linguistic vectors, which is the step providing the input to the grading perceptron, is not detailed in the claim. According to the description (see paragraph 52), they are defined and selected to capture sufficient information for evaluating the de-

gree of linguistic competence; they can be said to provide a "mathematical" summary of a script. Since the claim provides no detail as to the contents of the vector, this step cannot be considered to provide any contribution on its own, be it related to the script acquisition (e.g. scanning or OCR) or modelling, or to any optimization within the computer.

- 10.2 The claimed perceptron model is a linear mathematical function mapping the input numerical vectors to output grades. Specific details are only claimed with regard to its training procedure, which is optimized to preserve the ranking of grades, as opposed to minimizing the absolute error in output grades (see point 1.4 above). The model is not based on technical considerations relating to the internal functioning of a computer (e.g. targeting specific hardware or satisfying certain computational requirements), and the preference ranking is chosen merely according to its educational purpose, which does not relate to any effects within the computer either.
- 10.3 Also the final step of using the perceptron to grade the scripts provides no effects within the computer.
11. In principle, the claimed training procedure might constitute a technical contribution to the state of the art (see e.g. G1/19, reasons 33). Taken alone, however, this is a mathematical method, so this contribution is in the - excluded - field of mathematical methods (see T 0702/20 and T 0755/18, catchwords) and is therefore not a patentable contribution.
12. Thus the Board cannot identify any technical problem solved be it at the input, or in generating the output grade output, or by execution of the claimed process.

Technical effect via "implied technical use"

13. What remains as a potentially patentable contribution is the purpose of the claimed system to provide an automated tool for script grading. This corresponds to the problem formulated by the Appellant, namely *"providing a computer system that can automatically grade text scripts [and provide grades] that correlate well with the grades provided by human markers"*. The questions to be answered are (i) whether this problem is, or implies, a technical one, and (ii) whether it is actually solved (T 641/00, reasons 5 and 6).
14. Turning first to question (ii), the Board remarks that the human grading process is a cognitive task in which the marker evaluates the content of the script (e.g. language richness and grammatical correctness) to assign a grade.
 - 14.1 The assigned grade depends on the content of the script itself, but is also at least partly subjective: the marker will have preferences as to style and language, and will be influenced by experience and grades assigned to scripts in the past.
 - 14.2 The Board thus doubts that the problem of automating script grading is defined well enough that one can properly assess whether it has been solved, i.e. in the sense that it provides a system that can actually replace different human markers and provide "correct" grades.
 - 14.3 The Appellant has captured this in the problem formulation by the qualifier *"correlate well"*. Given the results in the application, showing that the claimed system provides results that agree with the ground truth

on the same level as the markers agree with each other, the Board is satisfied that the system can produce outputs that "*correlate well*" with the training data from human markers. The Board has no occasion to challenge that the invention may for instance be useful, as the Appellant submitted, for the (self-)evaluation of linguistic competences by students.

15. In its communication, the Board questioned under Article 84 EPC whether the claims of the main request comprised all the features necessary to produce this result. However, given that the Appellant was willing to amend the claims to overcome this objection, the Board leaves this question open and proceeds on the assumption that the problem, as qualified by the Appellant, is solved.
16. Under this assumption, there is a first argument that any automation of human tasks, irrespective of the task, is sufficient to conclude that a technical problem is solved, as it reduces human labor.
 - 16.1 This argument, however, contradicts the requirement of G 1/19 that there must be a technical purpose. Though G 1/19 was related to computer-implemented simulations, its reasons apply to computer-implemented methods other than simulations as well.
 - 16.2 The Enlarged Board stated that "*information which may reflect properties possibly occurring in the real world [...] may be used in many different ways*", that "*a claim concerning the calculation of technical information with no limitation to specific technical uses would therefore routinely raise concerns with respect to the principle that the claimed subject-matter has to be a technical invention*" (reasons 98),

and that "*[i]f the claimed process results in a set of numerical values, it depends on the further use of such data (which use can happen as a result of human intervention or automatically within a wider technical process) whether a resulting technical effect can be considered in that assessment*" (reasons 124), and concluded that "*such further [technical] use has to be at least implicitly specified in the claim*" (reasons 137).

- 16.3 Therefore, the argument that a technical problem is already solved by the mere provision of any automated tool cannot succeed.
17. As stated above, the Board assumes that the claimed invention serves the purpose of supporting its users in evaluating linguistic competences, as the Appellant argued. The Board also cannot see any other implied purposes. The question remains whether the assessment of linguistic competences, or maybe merely providing a grade, is a technical purpose.

What is technical?

18. The Appellant considers that automated grading makes a technical contribution in the field of "*educational technology*" and, if the Board disagrees, asks the question "*what is a technical field?*" or "*a field of technology?*".
19. The Board understands these two questions to be equivalent. The express reference to "*fields of technology*" in Article 52(1) EPC, introduced with the EPC 2000 in order to bring Article 52 EPC in line with Article 27(1) TRIPS, was not intended to change the established understanding that patent protection is "*reserved for creations in a technical field*", i.e. involving a

"*technical teaching [...] as to how to solve a particular technical problem*" (see OJ EPO Special edition 4/2007, 48, but also G 1/19, reasons 24, and T 1784/06, reasons 2.4).

- 19.1 The Board further notes that the field of "*educational technology*" as defined by the Appellant (see point 5 above) is a rather inhomogeneous one, covering insights from - and presumably contributions to - a wide range of "fields", technical ones and non-technical ones. It appears questionable, therefore, that this field can be considered a technical one as a whole. However, this question is not decisive.

- 19.2 What is decisive, according to established case law of the Boards of appeal, is whether the invention makes a contribution which may be qualified as technical in that it provides a solution to a technical problem. If this is the case, a contribution to a field of technology may be said to also be present. It is noted that the "field" of this contribution may be different from the one to which the patent more generally relates: for instance, inventions within the broad field of "*educational technology*" may make contributions in the field of computer science.

20. In G 1/19, the Enlarged Board followed its earlier case law and "*refrain[ed] from putting forward a definition for 'technical'*", because this term must remain open (section E.I.a, especially reasons 75 and 76; see also OJ EPO Special Edition 4/2007, 48). Nonetheless, the Enlarged Board provided considerations as to what may be considered technical.

- 20.1 The referring Board had suggested that a technical effect might require a "*direct link with physical*

reality, such as a change in or a measurement of a physical entity" (see T 489/14, reasons 11).

- 20.2 The Enlarged Board accepted that such a "*direct link with physical reality [...] is in most cases sufficient to establish technicality*" (reasons 88) and, in this context, that "*[i]t is generally acknowledged that measurements have technical character since they are based on an interaction with physical reality at the outset of the measurement method*" (reasons 99). It also stressed that an effect could also be "*within the computer system or network*" (i.e. internal rather than "*(external) physical reality*", see G 1/19, reasons 51 and 88).
- 20.3 It recalled that potential technical effects might also be sufficient (see also reasons E.I.e), i.e. "*effects which, for example when a computer program [...] is put to its intended use, necessarily become real technical effects*" (reasons 97).
- 20.4 And it also considered that calculated data, while "*routinely raising concerns with respect to the principle that the claimed subject-matter has to be a technical invention over substantially the whole scope of the claims*" might contribute to a technical effect by way of an implied technical use (reasons 98 and 137), "*e.g. a use having an impact on physical reality*" (reasons 137).
- 20.5 While the Enlarged Board of Appeal has thus found that a *direct* link with physical reality may not be required for a technical effect to exist, it has, in this Board's view, confirmed that an at least *indirect* link to physical reality, internal or external to the computer, is indeed required. The link can be mediated

by the intended use or purpose of the invention ("when executed" or when put to its "implied technical use").

21. Returning to the case at hand, the Board finds that automated script grading, by itself or via its intended use for evaluating linguistic competences, does not have an implied use or purpose which would be technical via any direct or indirect link with physical reality.

Conclusion

22. The claimed computer-implemented method of automated script grading does not provide a contribution to any technical and non-excluded field, be it by way of how the automation is carried out, or by way of its use; an inventive step according to Article 56 EPC can therefore not be acknowledged.

Auxiliary requests

23. The auxiliary requests are amended with respect to the main request in view of the Article 84 EPC objection raised by the Board. The amendments only concern details of the preference ranking method used for training the perceptron, with a view of defining a method that can be said to solve the problem stated by the Appellant. Since the Board has already assumed that (see point 14 above), these amendments have no impact on the assessment of inventive step carried out above.

Referral to the Enlarged Board of Appeal

24. According to Article 112(1) (a) EPC, the Board of Appeal shall refer any question to the Enlarged Board of Appeal if it considers that a decision is required in

order to ensure uniform application of the law, or if a point of law of fundamental importance arises.

- 24.1 As regards the first question, the Board considers that the case law of the Boards of Appeal on the question of what is "technical" or a "field of technology" is sufficiently uniform (see in particular G 1/19) so that a referral to the Enlarged of Appeal is not required.
- 24.2 As regards the second question, the Board notes the following. First, the term "educational technology" is too vague to be relevant for deciding the present case (cf. above, point 19.1). And secondly, even within a field of technology a patentable invention must be shown to solve a technical problem (see above, point 19.2). In the present case, the Board was unable to identify a specific technical problem solved by the invention. Therefore, a referral to the Enlarged Board is also not required for the second question.
25. The request to refer questions to the Enlarged Board of Appeal is therefore rejected.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar:

The Chairman:



L. Stridde

Martin Müller

Decision electronically authenticated