



PATENTS ACT 1977

APPLICANT Institute for Information Industry

ISSUE Whether patent application GB2018211.9 complies
with section 1(2) of the Patents Act 1977

HEARING OFFICER Phil Thorpe

DECISION

Introduction

- 1 Patent application GB2018211.9 was filed on 19 November 2020 claiming a priority date of 30 September 2020 from Taiwanese application TW 109134256. The application was published as GB 2599739 A on 13 April 2022.
- 2 Despite several rounds of correspondence between the examiner and the applicant's attorney, the applicant has been unable to satisfy the examiner that the application meets the requirements of the Act. In particular, the examiner remains of the opinion that the claimed invention is excluded from patentability as a program for a computer as such. The application has not been searched.
- 3 In their letter dated 19 April 2022 the examiner invited the applicant to request a hearing, noting that a further response may result in the application being passed to a senior officer. The agent responded on 17 June 2022 with further arguments and did not request a hearing. The examiner remained unconvinced by the arguments, and the case was passed to me for a decision on the papers.

The invention

- 4 The invention relates to a federated machine learning system in which a host device 110 stores and updates a host model based on training results received over multiple training rounds from a number of client devices 1201-120k. Each of the client devices receives the host model, trains the host model based on their locally stored parameter sets 1211-121k, and sends the results to the host device. The host can then adjust the model based on the results received. As noted, the model can be trained over several training rounds. The invention provides that the host device can update the host model every time the host device receives the training result transmitted by any of the client devices or alternatively only to update the host model after receiving a predetermined number of the training results.

- 5 Due to factors such as the number or size of the parameter sets and hardware equipment, the training speed of the client devices may be different, and the time point of transmitting the training results may also be different. Successive training rounds can be undertaken before all the client devices have transmitted their results. So, for example client device 1201 may respond with the results of the second round of testing before client device 1202 has responded to the first round of testing. The invention seeks to ensure that the training procedures of the host model are not excessively affected by the client device 1201 that performs the most training rounds.
- 6 To achieve this the invention uses a threshold value regarding a difference in number of training rounds between client devices to update the host model. In particular, when the difference in training rounds between client devices is not greater than the threshold, then the host updates the host model without results received from one of the clients which has a training round number corresponding to the minimum of the round numbers (i.e. it does not need to wait for the training result from the client with the lowest training round number in order to update the host model when the difference in training rounds is not larger than a threshold number).
- 7 The invention also provides that the host device stores a worklist 112 with corresponding identifiers for the clients, and when a time stamp (i.e. the time taken to receive a training result) for a client is larger than a preset time value, the host removes the identifier for the client device and no longer receives training results from that client.

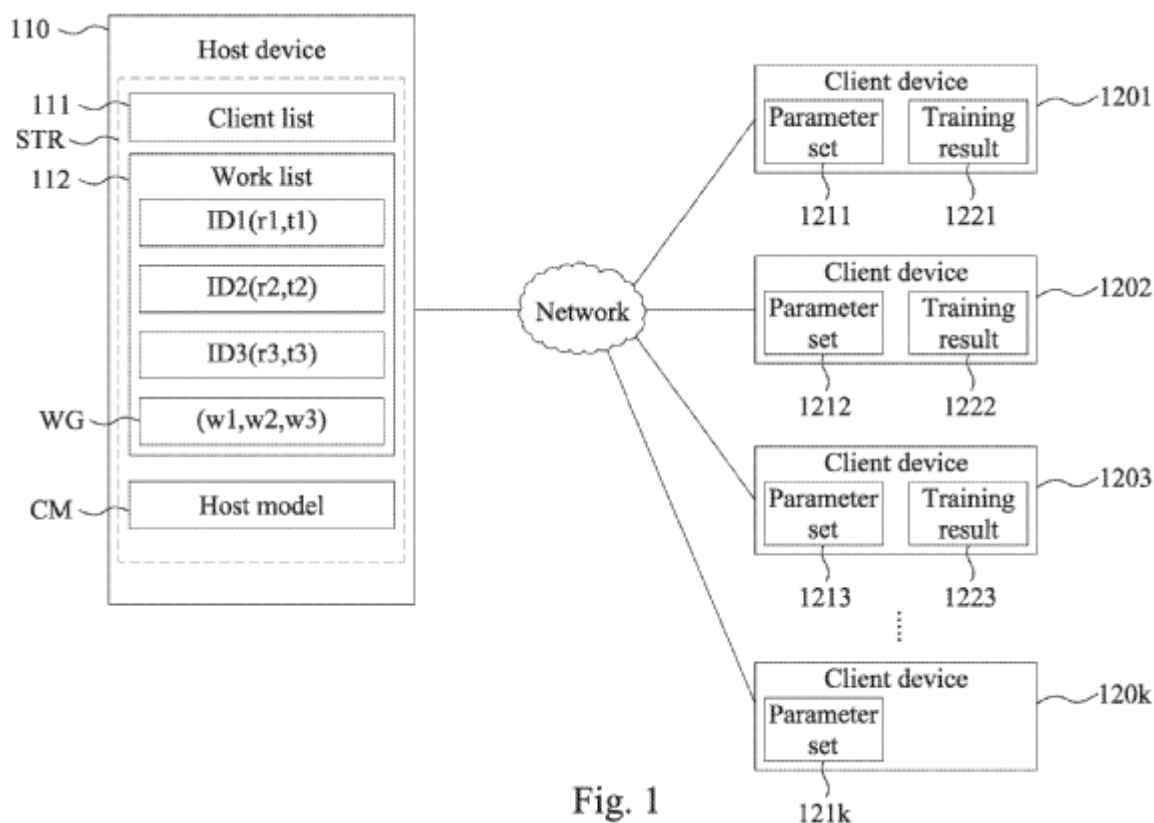


Fig. 1

- 8 The claims under consideration were filed on the 30th of September 2021. There are two independent claims, which are reproduced below:

Claim 1:

A machine learning system comprising:

a host device configured to store a host model, and configured to update the host model one or multiple times; and

a plurality of client devices configured to be communicatively coupled to the host device and receive the host model, the plurality of client devices comprising:

a first client device storing a first parameter set, and configured to train the received host model according to the first parameter set to generate a first training result; and

a second client device storing a second parameter set, and configured to train the received host model according to the second parameter set to generate a second training result;

wherein under a condition that the host device receives the first training result of the first client device corresponding to an m -th training round but not receive the second training result of the second client device corresponding to a n -th training round, when the difference between m and n is not larger than a threshold value, the host device updates the host model with the first training result corresponding to the m -th training round but not with the second training result corresponding to the n -th training round, wherein m is larger than n ,

wherein the host device comprises a host storage unit, and the host storage unit is configured to store at least one preset time value and a plurality of identifiers (IDs) respectively corresponding to the plurality of client devices, wherein

when the host device not receiving the first training result exceeding a corresponding one of the at least one preset time value, the host device removes one of the plurality of identifiers that corresponds to the first client device to generate updated identifiers, and no longer receives the first training result.

Claim 6:

A machine learning system comprising:

a host device configured to store and update a host model one or multiple times; and

a plurality of client devices configured to be communicatively coupled to the host device and receive the host model, wherein each of the plurality of client devices is configured to train the received host model according to a corresponding one of a plurality of parameter sets to generate a training result; wherein

the host device stores a plurality of round numbers respectively corresponding to the plurality of client devices, and when the host device receives the training result of one of the plurality of client devices, the host device updates a corresponding one of the plurality of round numbers, wherein

in an i-th update that the host device conducts to the host model, when a difference between a maximum and a minimum of the plurality of round numbers is not larger than a threshold value, and the host device does not receive the training result transmitted from one of the plurality of client devices corresponding to the minimum of the plurality of round numbers, the host device updates the host model according to the training results that are received in the i-th update, and i is a positive integer,

wherein a first client device of the plurality of client devices is configured to generate a first training result in each training round, and the host device comprises a host storage unit configured to store at least one preset time value and a plurality of identifiers respectively corresponding to the plurality of client devices, wherein

when the host device not receiving the first training result exceeding a corresponding one of the at least one preset time value, the host device removes one of the plurality of identifiers that corresponds to the first client device to generate updated identifiers, and no longer receives the first training result.

The Law

- 9 The examiner has raised an objection under section 1(2) of the Patents Act 1977 that the invention is not patentable because it relates a category of excluded matter. The relevant provisions of this section of the Act are shown with added emphasis below:

1(2) *It is hereby declared that the following (amongst other things) are not inventions for the purpose of the Act, that is to say, anything which consists of...*

(c) **...a scheme, rule or method for...doing business, or a program for a computer;**

but the foregoing provisions shall prevent anything from being treated as an invention for the purposes of the Act only to the extent that a patent or application for a patent relates to that thing as such.

- 10 As explained in the notice published by the IPO on the 8th December 2008¹, the starting point for determining whether an invention falls within the exclusions of section 1(2) is the judgment of the Court of Appeal in *Aerotel/Macrossan*².
- 11 The interpretation of section 1(2) has been considered by the Court of Appeal in *Symbian*³. *Symbian* arose under the computer program exclusion, but as with its previous decision in *Aerotel* the Court gave general guidance on section 1(2).

¹ <http://www.ipo.gov.uk/pro-types/pro-patent/p-law/p-pn/p-pn-computer.htm>

² *Aerotel Ltd v Telco Holdings Ltd and Macrossan's Application* [2006] EWCA Civ 1371; [2007] RPC 7

³ *Symbian Ltd v Comptroller-General of Patents*, [2009] RPC 1

Although the Court approached the question of excluded matter primarily on the basis of whether there was a technical contribution, it nevertheless (at paragraph 59) considered its conclusion in the light of the *Aerotel* approach. The Court was quite clear (see paragraphs 8-15) that the structured four-step approach to the question in *Aerotel* was never intended to be a new departure in domestic law; that it remained bound by its previous decisions, particularly *Merrill Lynch*⁴ which rested on whether the contribution was technical; and that any differences in the two approaches should affect neither the applicable principles nor the outcome in any particular case.

12 Subject to the clarification provided by *Symbian*, it is therefore appropriate to proceed on the basis of the four-step approach explained at paragraphs 40–48 of *Aerotel* namely:

- (1) *Properly construe the claim.*
- (2) *Identify the actual contribution (although at the application stage this might have to be the alleged contribution).*
- (3) *Ask whether it falls solely within the excluded matter.*
- (4) *If the third step has not covered it, check whether the actual or alleged contribution is actually technical.*

Applying the *Aerotel* test

Step 1 – Properly construe the claim

13 Neither the examiner nor the applicant have suggested that there is any difficulty in construing the claims. I agree. I would note however, that claims 1&6 relate only to circumstances when the difference in the number of rounds of training results received is *not larger* than a threshold. There is no discussion in the claims regarding the circumstances of when the number of rounds of training results received is *larger* than the threshold.

Step 2 – Identify the actual or alleged contribution

14 Jacob LJ addressed this step in *Aerotel/Macrossan* where he noted:

“43. The second step — identify the contribution — is said to be more problematical. How do you assess the contribution? Mr Birss submits the test is workable — it is an exercise in judgment probably involving the problem said to be solved, how the invention works, what its advantages are. What has the inventor really added to human knowledge perhaps best sums up the exercise.”

15 Jacob LJ also adds in paragraph 44:

“ Mr Birss added the words "or alleged contribution" in his formulation of the second step. That will do at the application stage – where the Office must generally performe accept what the inventor says is his contribution”

16 Both the applicant and the examiner agree that the problem addressed by the invention lies in combating the effect of excessive elapsed time or excessive training round differences. In particular, the problem addressed by the claims is that due to

⁴ *Merrill Lynch's Appn.* [1989] RPC 561

client devices possibly having difference device specifications and capabilities, the training speed of the client devices will be different. This could result in problems such as network delays and unexpected client disconnection during the training process which may lead to an excessive elapsed time or training round difference between each of the client devices.

- 17 The applicant notes that an excessive round difference would cause an inaccurate update result due to the host model being affected too much by a certain client device. This would cause negative effects for the federated learning system such as reduced training speed and stability for the whole training process.
- 18 Furthermore, the applicant notes that when a training result is not received and exceeds a preset time, this serves to identify client devices which are potentially abnormal or interrupted. Similarly, such an interruption could be caused by differing device specifications, or additionally by network delays or unexpected disconnection. The identified client devices are therefore removed from the work list to prevent the learning process from being negatively affected.
- 19 It seems to me there are three potential effects to the contribution here. Firstly, it would appear that the model gets updated at the host in less time for each training round – as it is not ‘waiting’ for the results from the client with the lowest or minimum number of training rounds. Therefore, looking at the learning process as a whole, each client device receives the updated host model in less time per round (than without the threshold) and thus there is a reduction in the overall time waste caused by each client device waiting for each other – or at least waiting for the device with the lowest round number. Therefore, the overall training process is arguably more efficient in terms of time or speed to update the host model.
- 20 Secondly, it is alleged that the contribution “avoids excessive training round difference between each of the client devices, and further prevents an inaccurate update result caused by the host model being affected too much by a certain client device”. I struggle to see how this effect is produced by the claimed invention which, as discussed above in step one, is about circumstances when the host device does not receive training results from the device with the lowest or minimum number of rounds and the difference in the number of rounds of training results received is *not larger* than a threshold. Nor is it implicit that if the number of rounds is larger than the threshold, that the system waits for the client device with the minimum number of rounds so as to “avoid excessive training round differences” (although such an arrangement/effect would appear to be outlined in paragraphs 20,47 and step 470 in fig 3). For example, the system might just disregard all training results if all of them are not received and the difference in the number of rounds is larger than a threshold. Therefore, I do not see how this forms part of the contribution of the claimed invention.
- 21 Thirdly, by removing the corresponding client device from the work list when the time stamp is greater than the preset value, there is the alleged effect that “the whole learning process may be prevented from being affected when the client device is abnormal or interrupted” and that “the federated machine learning process may be prevented from being interrupted due to disconnections of the client devices, and thus enhances efficiency and stability of the learning process”. The overall learning

process is thus arguably more stable in that it is not subjected to interruptions or excessive wait times for results from a particular client.

22 The applicant has alleged that the contribution resides in:

“an improved federated machine learning system in which time wastage caused by client devices waiting for one another is reduced, and in which the problem of the accuracy of the host model being affected by excessive differences in operational speed and throughput of the client device is mitigated.”

23 Whilst there is some doubt regarding how the claimed invention necessarily effects the “accuracy” of the host model, I am willing to perforce accept that by updating the model without the training result transmitted from a client device(s) corresponding to the minimum of the plurality of round numbers, when the number of rounds is not larger than a threshold, this may have some effect on the accuracy of the model. Similarly, by removing and not using results from a client device which has a time stamp greater than the preset value, the whole learning process may be prevented from being affected, such as to have at least some effect on the accuracy of the model.

24 I would add that the system itself is conventional – i.e. a host computer and clients – and the contribution lies in the functionality or the way in which the system is programmed.

25 Therefore, I shall proceed to step 3 based on the applicant’s assessment of the contribution, with the caveat that the scenario of avoiding “excessive training round difference between each of the client devices” has not been mitigated (see paragraph 17 above).

Steps 3 and 4 – Ask whether it falls solely within the excluded matter and check whether the actual or alleged contribution is actually technical

26 I will consider steps 3 and 4 together.

27 Although the description does not refer to computers, it seems clear that the invention is computer implemented and that it includes one or more computers suitably programmed. Lewison J. (as he then was) set out five signposts *AT&T/CVON*⁵ that he considered to be helpful when considering whether a computer program makes a technical contribution. In *HTC*⁶ the signposts were reformulated slightly in light of the decision in *Gemstar*⁷. The signposts are:

- i. Whether the claimed technical effect has a technical effect on a process which is carried on outside the computer.
- ii. Whether the claimed technical effect operates at the level of the architecture of the computer; that is to say whether the effect is produced irrespective of the data being processed or the applications being run.

⁵ *AT&T Knowledge Venture/CVON Innovations v Comptroller General of Patents* [2009] EWHC 343 (Pat); [2009] FSR 19

⁶ *HTC v Apple* [2013] EWCA Civ 451

⁷ *Gemstar-TV Guide International Inc v Virgin Media Ltd* [2009] EWHC 3068 (Pat); [2010] RPC 10

- iii. Whether the claimed technical effect results in the computer being made to operate in a new way.
- iv. Whether the program makes the computer a better computer in the sense of running more efficiently and effectively as a computer.
- v. Whether the perceived problem is overcome by the claimed invention as opposed to merely being circumvented.

- 28 It is important to stress that these signposts are just that. They are not barriers or hurdles that need to be individually or collectively overcome by the applicant. They are rather a non-exhaustive list of some of the factors that can indicate in some cases whether a particular contribution may be technical.
- 29 The applicant has focussed their submissions on signposts (i) and (v). With regard to signpost (i), the applicant submits that the functionality of the system is fundamentally different as a client device is excluded from the machine learning process. In particular, the federated machine learning system is made up of a set of devices (i.e. the host device and a plurality of client devices). That set of devices that make up the system is changed when a client device is added or removed. This necessarily requires that the system functions differently, since data involved in the machine learning process will be transmitted and processed in dependence on which devices form a part of the system.
- 30 The applicant adds that the host device is configured to dynamically control which client devices form a part of the machine learning system by removing the identifier of the client device from the storage unit. Clearly, the applicant submits, the removal of a device from a system is an effect in the outside world.
- 31 The examiner has referred to the decision in *Lantana*⁸ in his assessment of the first signpost. The applicant argues that the examiner has “overstepped the substantive point made in *Lantana*” and “considers that it is unsuitable to equate the claimed federated machine learning system to the mere internal operations of a single computer”. However, it was established in *Lantana* that a system operating as a network can be considered as “the computer” for the purposes of the first signpost. Paragraph 30 of *Lantana* notes that:

“30. I start by noting that this invention consists entirely of software running on a conventional computing arrangement. I use the term “computing arrangement” rather than computer because the applicant is at pains to point out that this system requires two computers connected by a “telecommunications network”. So it does but at the relevant date (2008) two computers connected across the internet was an entirely conventional computing arrangement. The fact that two computers and the internet are required is not what makes a software invention patentable.”

- 32 In a similar manner to the arrangement in *Lantana*, the system of the present invention comprises a new program providing different functionality (in that a host no longer receives a training result from a client when they take too long to complete a training round) implemented using conventional devices communicating in conventional ways. As with *Lantana*, everything is going on within the computing arrangement. Hence notwithstanding that that system of computers here functions

⁸ *Lantana v Comptroller-General of Patents* [2013] EWHC 2673 (Pat)

differently to the prior art, as so did the arrangement in *Lantana*, signpost i) cannot assist the applicant. Or to put it in other words choosing to no longer receive training results from a client device – or “removing” this device from the machine learning system – has no effect on the world outside of this system.

- 33 The applicant also contends regarding signpost v) that the invention provides that problems arising from network delays and unexpected client disconnection during the training process can be avoided. The applicant argues that the problem does not solely relate to training a machine learning model. The host device can dynamically manage client devices that are suitable for training in order to improve the overall training speed and stability of the system.
- 34 The applicant further points out that the present invention is directed to identifying and excluding client devices that present issues arising from different training speeds, network delays, or unexpected disconnection. This mechanism seeks to improve the host model that results from the federated learning process. Whether it makes more efficient use of resources is not relevant. Indeed, in some examples, removing a client device could result in a less efficient machine learning process due to fewer training results in each training round. However, this is done in exchange for the advantage of a better end result.
- 35 The examiner has argued that improving a model is a mathematical concern. The applicant suggests that, while this may be so *per se*, the technical domain of the present invention relates to dynamically configuring the machine learning system (i.e. by selectively removing client devices, when it is deemed beneficial to do so) in order to produce an improved model. The applicant also adds that, clearly, configurational changes to a federated system in which participant devices are included or excluded in dependence on performance is a technical consideration, and not merely an abstract mathematical concern.
- 36 I accept that the invention here does not relate to a mathematical method as such.
- 37 I also accept that the method of machine learning proposed is potentially better than the prior art. However, it is important to understand why that is. At the heart of the invention is what to do if one or more of the federated client devices does not respond as quickly as expected. This may be because of problems regarding client speeds, wait times, network delays, or unexpected disconnection. What the invention does however is not to solve any of those problems, which may be technical in nature, but rather circumvents them by choosing to no longer receive a training result from a client and/or not waiting for training results. Whilst this may make the training more reliable, that is not because the invention has solved any technical problems. Rather it results from a conventional network running a better computer program and that on its own is not enough. I would add that I could see no technical contribution in the manner of deciding which clients to ignore; this being based on a conventional assessment of any delay in receiving the results.
- 38 Therefore, it is my conclusion that signpost (v) is not satisfied.
- 39 Whilst there have been no explicit arguments regarding the remaining signposts, I do not consider these to have been satisfied either. The program of the present invention may provide a more efficient or quicker process for updating a model (as

the system does not have to wait for other clients before updating a host model), and may provide a more accurate model in a more stable (or less interrupted) way. But these benefits are only felt when the computer system is concerned with executing the inventive program. It is not an improvement made to the computer or network irrespective of the application(s) being run. In other words, the contribution is no more than a better computer program.

- 40 Taking a step back and looking at the invention as a whole, I am satisfied that the contribution falls solely within the computer program exclusion. In particular, updating a host model based on a difference in the number of training rounds and the time taken to receive training results are policy based decisions regarding when to update a model and from which clients to receive training data, rather than being concerned with a technical process. Thus, the improved machine learning system is not technical in nature and is no more than a better computer program.
- 41 Finally, I would also note that even if the claimed invention were limited to mitigating the problem of “excessive training round difference between each of the client devices” (i.e. by waiting for a client with minimum rounds when the difference in round numbers is larger than a threshold), this would not provide the required technical contribution. No technical problem regarding the client computers or network has been solved and the result is a better program for updating a model.

Conclusion

- 42 Having carefully considered the arguments, I am of the view that the invention falls solely within the matter excluded under section 1(2) as a program for a computer as such. I can see nothing in the specification that could reasonably be expected to form the basis of a valid claim. I therefore refuse this application under section 18(3).

Appeal

- 43 Any appeal must be lodged within 28 days after the date of this decision.

Phil Thorpe

Deputy Director, acting for the Comptroller