



PATENTS ACT 1977

APPLICANT Lenovo (Singapore) PTE. LTD

ISSUE Whether application GB1501383.2 complies with
 Section 1(1) of the Patents Act 1977

HEARING OFFICER Stephen Brown

DECISION

Introduction

- 1 Patent Application GB1501383.2 was filed on 28th January 2015 claiming a priority date of 28th January 2014 from an earlier US application. It was published as GB 2524864 A on 7th October 2015. Despite several rounds of amendment the applicant has been unable to convince the Examiner that the application is novel and inventive. The applicant thus requested a hearing to resolve the matter.
- 2 This took place on 12th September by telephone with Benjamin Echterhoff of Schweiger and partners. I was assisted by Mr Nigel Hanley and the Examiner Dr Mark Lewney also attended. The hearing was observed by Mr Christopher Robertson and Mr Andrew Gillan.

Section 20

- 3 The original compliance date for this application was 28th July 2018. This has since been extended until 29th September 2018. If I find in favour of the applicant it will be necessary to extend this further.

The Application

- 4 The application lies in the field of automatic speech recognition (ASR), in particular in adjusting an ASR engine using contextual information. This contextual information may be related to the current operation of a device. So, by way of example, if the user is currently listening to music using a music player the ASR could be adjusted to 'expect' commands such as "play" or "stop". In effect, the ASR is biased towards words and/or commands that the user is more likely to use in the identified context.

- 5 Throughout the examination process the examiner has maintained that the claims lack novelty and/or the required inventive step. These were the issues addressed at the hearing and I will therefore consider both issues in this decision.
- 6 The current claims on file are those filed on 7th June 2018 and comprise three independent claims directed to a method, an information handling device and a program product:

Claim 1

A method, comprising

- a) Obtaining, using a processor, contextual information relating to an information handling device;*
- b) Adjusting, before receiving user speech input, using a processor, an automated speech recognition engine using the contextual information;*
- c) Thereafter receiving, at an audio receiver of the information handling device, user speech input;*
- d) Providing, using a processor, a communication to the user including estimated recognized speech based on the user speech input received and the contextual information adjustment to the automated speech recognition engine;*
- e) Receiving user input associated with the communication; and*
- f) Committing a predetermined action matching the recognized speech, wherein said committing a predetermined action matching the recognized speech proceeds responsive thereto.*

Claim 9

An information handling device, comprising:

An audio receiver

A processor operatively coupled to the audio receiver; and

A memory device storing instructions executable by the processor to:

- a) Obtain contextual information relating to an information handling device;*
- b) Adjust, before receiving user speech input, using a processor, an automated speech recognition engine using the contextual information;*
- c) Thereafter receive, at an audio receiver of the information handling device, user speech input;*

- d) *Provide a communication to the user including estimated recognized speech based on the user speech input received and the contextual information adjustment to the automated speech recognition engine;*
- e) *Receive user input associated with the communication; and*
- f) *Commit a predetermined action matching the recognized speech, wherein said committing a predetermined action matching the recognized speech proceeds responsive thereto.*

Claim 16

A program product, comprising:

A storage medium comprising device readable program code, the code being executable by a processor and comprising:

- a) *Code that obtains, using a processor, contextual information relating to an information handling device;*
- b) *Code that adjusts, before receiving user speech input, using a processor, an automated speech recognition engine using the contextual information;*
- c) *Code that thereafter receives, at an audio receiver of the information handling device, user speech input;*
- d) *Code that provides, using a processor, a communication to the user including estimated recognized speech based on the user speech input received and the contextual information adjustment to the automated speech recognition engine;*
- e) *Code that receives user input associated with the communication; and*
- f) *Code that commits a predetermined action matching the recognized speech, wherein said committing a predetermined action matching the recognized speech proceeds responsive thereto.*

Note: I have added the references a) – f) to assist with my later discussion of the claims.

The Law

7 Section 1(1) of the Act sets out what is required of a patentable invention as follows:

1(1) A patent may be granted only for an invention in respect of which the following conditions are satisfied, that is to say-

- (a) the invention is new;*
- (b) it involves an inventive step;*
- (c)*

and references in this Act to a patentable invention shall be construed accordingly

8 Section 2 of the Act sets out what 'new' means as follows:

2(1) An invention shall be taken to be new if it does not form part of the state of the art.

2(2) The state of the art in the case of an invention shall be taken to comprise all matter (whether a product, a process, information about either, or anything else) which has at any time before the priority date of that invention been made available to the public (whether in the United Kingdom or elsewhere) by written or oral description, by use or in any other way

2(3) The state of the Art in the case on an invention to which an application for a patent or a patent relates shall be taken also to comprise matter contained in an application for another patent which was published on or after the priority date of that invention, if the following conditions are satisfied, that is to say:-

- (a) That matter was contained in the application for that other patent both as filed and as published; and*
- (b) The priority date of that matter is earlier than that of the invention*

9 Section 3 of the Act states:

An invention shall be taken to involve an inventive step if it is not obvious to a person skilled in the art, having regard to any matter which forms part of the state of the art by virtue only of section 2(2) above (and disregarding section 2(3) above).

10 In addition to statute, the courts have long used the so called *Windsurfing* test to assess issues of inventive step. This test was reformulated by the Court of Appeal in *Pozzoli*¹. Paragraph 23 of this decision lays out the test as:

- (1) (a) Identify the notional "person skilled in the art"*
- (b) Identify the relevant common general knowledge of that person;*

¹ *Pozzoli Spa v BDMO SA & Anor* [2007] EWCA Civ 588

- (2) *Identify the inventive concept of the claim in question or if that cannot readily be done, construe it;*
- (3) *Identify what, if any, differences exist between the matter cited as forming part of the "state of the art" and the inventive concept of the claim or the claim as construed;*
- (4) *Viewed without any knowledge of the alleged invention as claimed, do those differences constitute steps which would have been obvious to the person skilled in the art or do they require any degree of invention?*

Analysis

- 11 I can see no issues of plurality with the claims so it follows that my analysis will apply to all the independent claims mutatis mutandis.
- 12 The Examiner has argued that the application does not make the requisite inventive step over the prior art. He has further argued that he considers the features I have labelled (a-c) and features I have labelled (d-f) to form two distinct inventions – i.e. that they are a collocation with no synergy between them.
- 13 Naturally, Mr Echterhoff took a different view. He argued that the order of the steps provides the required inventive step. Furthermore, he argued that there is synergy between the two sets of features. Specifically, he stated that the ASR engine being first adjusted and then used provided the required link between the two sets of features.
- 14 While I stated in the introduction that the issue of novelty was also under consideration, in this case I will consider inventive step first, beginning with the issue of collocation.

Collocation

- 15 The law regarding collocation is set out in the decision in *SABAF*². Briefly, this decision held that if a claim comprises two, or more, integers one must consider if they interact upon each other. If there is synergy between them, then they constitute a single invention and one considers the combined integers under section 3 of the Act. However, if each integer performs its own function independently of any of the others then each is considered a distinct invention for the purposes of section 3 and assessed separately.
- 16 *SABAF*² explained that two features are considered to interact synergistically if their functions are interrelated and lead to an additional effect that goes beyond the sum of the effects of each feature taken in isolation. It is not enough that the features solve the same technical problem or that their effects are of the same kind and add up to an increased but otherwise unchanged effect. Specifically, in the second half of paragraph 24, Lord Hoffmann held:

² *SABAF SpA v MFI Furniture Centres Ltd* [2004] UKHL 45

but before you can apply section 3 and ask whether the invention involves an inventive step, you first have to decide what the invention is. In particular, you have to decide whether you are dealing with one invention or two or more inventions. Two inventions do not become one invention because they are included in the same hardware. A compact motor car may contain many inventions, each operating independently of each other but all designed to contribute to the overall goal of having a compact car. That does not make the car a single invention.

In paragraph 26 he added:

The EPO guidelines say that "the invention claimed must normally be considered as a whole". But equally, one must not try to consider as a whole what are in fact two separate inventions. What the Guidelines do is to state the principle upon which you decide whether you are dealing with a single invention or not. If the two integers interact upon each other, if there is synergy between them, they constitute a single invention having a combined effect and one applies section 3 to the idea of combining them. If each integer "performs its own proper function independently of any of the others", then each is for the purposes of section 3 a separate invention and it has to be applied to each one separately. That, in my opinion, is what Laddie J meant by the law of collocation.

- 17 I will begin my analysis by considering whether steps (a, b & c) and steps (d, e & f) are a collocation. Steps (a-c) concern identifying a context and adapting an ASR *before* receiving user speech. At the hearing, Mr Echterhoff argued that this is a key feature of the invention. Prima facie, steps (d-f) relate to no more than relaying the ASR's interpretation of received speech to the user and then actioning a command once the user confirms that the interpretation is correct. The question is whether having adjusted the ASR in the first set of steps the use of it in the second set of steps provides the required synergy.
- 18 At the hearing, Mr Echterhoff argued that given that the adjusted ASR is used in the second set of steps this is enough. Ultimately, he asserted, it will result in a better method of speech recognition. His argument certainly appears to carry some merit since the ASR of the second set of steps is functionally different from a 'standard' ASR because of the action of the first set of steps. However, looking more closely at step (d), as I have defined it, I believe that there are actually two actions implicit in this step of the claims. Using my own words to make explicit what the applicants have left implicit, I believe step (d) may be expressed as:

(d1): determining the estimated recognized speech based on the user speech input received and the contextual information adjustment to the automated speech recognition engine;

and:

(d2) providing a communication to the user including the estimated recognized speech produced by the automated speech recognition engine;

- 19 If I now group the two sets of steps as (a-d1) and (d2-f) I believe that it is very clear that the 'new' second set of steps relates to simply confirming that the estimated speech is correct before actioning a command. I also believe that these steps will be performed identically irrespective of how the ASR arrived at its estimate of the speech. That is, I can see no synergy between the first set of steps and the second. I will thus apply Section 3 separately to the invention defined in steps (a-d1) and the invention defined in steps (d2-f).

Inventive Step - Applying the Pozzoli Test

- 20 At the hearing Mr Echterhoff agreed with the examiner's assessment that the skilled person was someone who worked with the speech engines. However, he argued that said skilled person would not have any knowledge of user interfaces. Indeed, he argued that including such knowledge as relevant would constitute the use of hindsight.
- 21 I am afraid that I do not agree with this assessment of the skilled person's common general knowledge. Speech recognition is a maturing technology and there are many commercially available products that make use of it in a variety of applications. ASR engines do not operate in isolation, they must perforce interface with their users in some manner. I thus conclude that the skilled person would have some knowledge, either individually or as part of a team, of how a given ASR might interface with its users.
- 22 I will now turn to the construction of the claim which is necessary in order to identify the inventive concept. As reasoned above, I consider the claim to consist of two separate integers. I have identified a first integer, consisting of steps I have labelled (a-d1), and a second integer consisting of steps (d2-f). Having concluded that these two integers are a collocation I will now consider the first integer separately.
- 23 Step (a) of the claim refers to "*obtaining contextual information relating to an information handling device*". I see little issue in what is meant by contextual information in this case. The application makes it clear that it relates to the current use of the device and provides examples. It could simply be that the user is listening to music on a music player (see paragraph 18 of the description) or related to what the device has done recently such as a text message (see paragraph 16).
- 24 Step (b) of the claim refers to "*Adjusting, before receiving user input, an automated speech recognition engine using the contextual information*". The key to understanding this phrase lies, I believe, in what is meant by adjusting the speech engine. The Examiner has argued that it involves selecting an appropriate dictionary whereas Mr Echterhoff argued for a more subtle view where the adjustment occurs within the engine rather than just replacing dictionaries.
- 25 I note that paragraph 31 of the specification talks about the adjustment of the ASR. It states that adjusting may take the form of "selecting a knowledge domain or template e.g. maps for a navigation context, telephony for received communication" or "selecting a lexicon". This latter option appears synonymous with the Examiner's interpretation. However, paragraph 31 also makes it clear that the adjusting step

“may comprise weighting one or more words”. This option clearly accords more with Mr Echterhoff’s view. However, I note that there is no detail as to how this is done other than that it uses a program. On this basis, and in the absence of any further detail, I will take a broad approach to construing the phrase “adjusting the ASR”. That is, I shall construe it to *include* selecting an appropriate dictionary, or lexicon, appropriate to an identified context.

- 26 I see no difficulty in construing steps (c) and (d1). “*receiving... user speech input*” and then “*determining the estimated recognized speech based on the user speech input received*” is the central thing that all ASR engines do. What I believe is key in this case though is that the estimated recognized speech is determined “*based on the user speech input received and the contextual information adjustment to the automated speech recognition engine*”. Again, I see no difficulty in construing this part of the invention.
- 27 What then is the inventive concept of the first integer of the claim? I consider that it is obtaining a context in which the device is operating and adjusting the ASR *before* speech is received from the user and processed using the *adjusted* ASR. At the hearing, Mr Echterhoff was keen to emphasise that the detail that the invention adjusted the ASR before user speech was received, and not merely in response to it, was a key feature. I agree with Mr Echterhoff on this point.
- 28 The next step is to consider the matter cited as forming part of the state of the art and to identify what differences exist between it and the inventive concept identified immediately above. The Examiner cited three documents:
- US 2009/0150156 KENNEWICK
- US 2007/0294084 CROSS
- US 2014/0025377 KOCH
- 29 CROSS and KOCH were only discussed briefly at the hearing but Mr Echterhoff convinced me that they assist only in so far as they show a context based grammar or dictionary. I agree that in both cases the order of the steps is unclear but the more likely order is adjustment post utterance. As such they omit a key feature of the current invention and thus do not, in my view, constitute useful jumping off points to consider the inventiveness of the current invention.
- 30 I believe that KENNEWICK, on the other hand, is much more relevant. Prior to the hearing, the Examiner has argued that Figure 3 of KENNEWICK shows a system where the ASR is adjusted before any user input. On the other hand, before and during the hearing, Mr Echterhoff argued that KENNEWICK does not show an explicit situation where ASR adjustment clearly occurs before user input.
- 31 Mr Echterhoff also argued that adjusting the ASR in the current application is not just the selection of a dictionary or, to use the term in KENNEWICK, ‘a grammar’. In his view it must also include the weighting of words in a grammar. On this latter point I have already construed the claims, in light of the description, to include adjusting an ASR by replacing a ‘grammar’, so I am afraid that I do not agree with this part of Mr Echterhoff’s argument.

32 At the hearing, Mr Echterhoff conceded that KENNEWICK shows a grammar being adjusted according to a context. In paragraph 0064 an example is given where the size of the dynamic grammar is constrained by the availability of system resources. Paragraph 0065 further refines this constraining to an example based on the user's location.

33 In the hearing I drew attention to paragraph 0066 of the citation. The end of this paragraph states:

As a result, by mapping the user's geographic proximities to one or more topological domains, dynamic grammars may be pruned, extended, swapped in or out of memory, or otherwise generated and/or loaded to provide optimal recognition based on location, time, travel, or other factors (e.g., information may be swapped in and out of a grammar as a user drives from one area to another, optimizing the information along the drive to ensure that the system resources utilize information presently relevant for a given location).

34 In my view, this section clearly discloses that a grammar used by an ASR may be adjusted (pruned, extended or swapped) in response to one or more of a number of factors *other than the input of user speech*. More specifically, the text in brackets quoted above shows that it was known to adjust an ASR in response to driving from one geographical area to another.

35 I think I may usefully illustrate my understanding of this disclosure by way of an example. I take the text to mean that if I were to drive from Newport towards London then the grammar in the ASR engine would change as I moved. As I approached Bristol, for example, it would be related to that area, with the same happening as I drove past Swindon then Reading. That is, the grammar is adjusted as I move from one area to another in anticipation of me making geographically relevant speech.

36 At the hearing Mr Echterhoff argued that KENNEWICK does not explicitly disclose ASR adjustment happening *before* receiving user input. Indeed, he drew my attention to paragraph 0069, the beginning of which states (my emphasis added) that:

Thus, the grammar information downloaded at operation 350 may be based on a variety of factors, including information in a current voice-based input, current geographical proximities of the user, topological domains or geographical chunks related to the current geographical proximities, or various kinds of affinities, among other things.

37 Mr Echterhoff argued that this clearly discloses that adjustment may happen in response to, i.e. *after*, user input. I agree that this is indeed one option shown in KENNEWICK however it is not the only option. The rest of the quoted passage in combination with paragraph 0066 discloses that it is equally possible to adjust an ASR in response to driving from one geographical area to another. I can see no suggestion that *this* adjustment must wait for the user to speak and thus I believe it can happen independently and therefore *before* the next time the user may speak. I thus conclude that KENNEWICK discloses adjusting an ASR before, and/or after, a user input.

- 38 Refocussing now on the third step of the Pozzoli test I must ask what differences are there between KENNEWICK and the inventive concept of the first integer of the current claims? As discussed above, I believe that KENNEWICK discloses, among other things, adjusting the grammar of an ASR engine based on a geographical context before a user speech input. Any subsequent speech input will then be recognised based upon the adjusted engine. Thus there appears to be no differences between the state of the art and the inventive concept of the first integer. Thus the first integer appears to lack novelty with respect to KENNEWICK.
- 39 If I were to concede Mr Echterhoff's point that adjusting the ASR engine includes the weighting of words in a grammar then this might constitute a difference. However, even if I do construe the claims as Mr Echterhoff would prefer, I believe that this difference would be obvious to the skilled person when considering KENNEWICK's disclosure of pruning, extending, or otherwise generating dynamic grammars. Thus, at best, the first integer is obvious.
- 40 I will now turn to the second integer identified above. I see no difficulty in construing steps (d2-f). To my mind "providing estimated recognized speech... to the user" and "receiving user input associated with the communication" is simply the well-known process of an ASR engine relaying to the user what it thinks it has heard and the user confirming, or not, that it is correct. Likewise, I believe that step (f) is straightforward to construe and also very well known. In my view it amounts to no more than performing a predetermined action matching the recognized speech once that speech has been confirmed by the user.
- 41 Mr Echterhoff did not disagree with my construction but did argue that to include it as part of the common general knowledge was to indulge in hindsight. Again, I am afraid I must disagree. It has been well known for many years for voice recognition systems to seek confirmation before taking action. My own telephone banking uses a confirmation technique to check what I have said. I am also aware of several comedians and at least one film that have found humour in scenarios where 'computers' mishear spoken instructions and will not accept denial of their interpretation by the user before acting. I thus firmly believe that such a 'confirm then act step' formed part of the common general knowledge of the skilled person at the priority date of the application. I thus conclude that the second integer lacks novelty.
- 42 I will now briefly return to the issue of collocation. If I am correct in my conclusion that the two integers should be assessed separately then the invention is no more than a collocation of two known things. Even if I am not correct and assess the claim as a whole then the first integer is disclosed in KENNEWICK. Thus the difference between KENNEWICK and the whole claim is merely what I have called the second integer. Since I have concluded that this was common general knowledge in the field of voice recognition it follows that the claim as a whole would have been obvious to the skilled person. That is, I believe that the skilled person would consider it obvious to add a 'confirm then act step' to the device disclosed in KENNEWICK. Thus if even if there is no collocation, I conclude that the independent claims would lack the required inventive step.

Other Claims

- 43 I have given due consideration to the other claims of the application and can see nothing in them that could be considered inventive over what is shown by the cited prior art in the light of the common general knowledge of the skilled person.

Decision

- 44 I have found that the invention set out in the independent claims is a collocation of two separate integers, both lacking novelty. Even if I were to not consider the invention to be a collocation, I have concluded that it would lack an inventive step.
- 45 Accordingly, I decide that the invention set out in the claims does not satisfy section 1(1) and I therefore refuse this application under Section 18(3).

Appeal

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

Stephen Brown

Deputy Director, acting for the Comptroller