BESCHWERDEKAMMERN DES EUROPÄISCHEN PATENTAMTS

BOARDS OF APPEAL OF THE EUROPEAN PATENT OFFICE

CHAMBRES DE RECOURS DE L'OFFICE EUROPEEN DES BREVETS

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

Datasheet for the decision of 13 May 2009

Case Number: T 1079/07 - 3.4.01

Application Number: 98914210.4

Publication Number: 0976092

IPC: G06K 9/20; G06F 17/24

Language of the proceedings: EN

Title of invention:
Method and arrangement for automatic data acquisition of forms

Patentee:
Readsoft AB

Opponent:
Top Image Systems Limited

Headword:
-

Relevant legal provisions:
EPC Art. 104(1)

Relevant legal provisions (EPC 1973):
EPC Art. 100(b)

Keyword:
"Sufficiency of disclosure (no; all requests)"
"Apportionment of costs (not equitable)"

Decisions cited:
T 0937/04; T 0544/94

Catchword:
-

Europäisches Patentamt | European Patent Office | Office européen des brevets

Beschwerdekammern | Boards of Appeal | Chambres de recours

Case Number: T 1079/07 - 3.4.01

D E C I S I O N
of the Technical Board of Appeal 3.4.01
of 13 May 2009

Appellant: (Proprietor of the patent)	Readsoft AB Garnisionsgaten 25A SE-254 66 Helsingborg (SE)
Representative:	Onn, Thorsten Zacco Sweden AB P.O. Box 23101 SE-104 35 Stockholm (SE)
Respondent: (Opponent)	TOP IMAGE SYSTEMS LIMITED 2 Habarzel Street, Ramat Hahayal Tel Aviv 69710 (IL)
Representative:	Hill, Justin John McDermott Will & Emery UK LLP 7 Bishopsgate GB-London EC2N 3AR (GB)
Decision under appeal:	**Decision of the Opposition Division of the European Patent Office posted 20 April 2007 revoking European patent No. 0976092 pursuant to Article 102(1) EPC 1973.**

Composition of the Board:

Chairman: B. Schachenmann
Members: H. Wolfrum
P. Fontenay

Summary of Facts and Submissions

I. The appellant (patent proprietor, ReadSoft AB) lodged an appeal against the decision of the opposition division, dispatched on 20 April 2007, revoking European patent No. 0 976 092 for the reason that the contested patent did not disclose the subject-matter of a main request and of two auxiliary requests on file in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art (Article 100(b) EPC 1973).

The notice of appeal was received on 18 June 2007 and the prescribed fee was paid on 19 June 2007. On 16 August 2007 a statement of grounds of appeal was filed. The requests on which the contested decision was based were maintained.

II. In a communication annexed to summons for oral proceedings, which were arranged upon corresponding requests from both parties, the board indicated that it was inclined to share the opposition division's judgement as to a lack of sufficiency of disclosure of the claimed subject-matter.

III. By facsimile dated 29 April 2009 the appellant informed the board that it would not attend the oral proceedings. No further substantive submissions were made.

On 6 May 2009 the board's registry forwarded the appellant's facsimile to the respondent (opponent, Top Image Systems Ltd.).

IV. Oral proceedings were held on 13 May 2009 in the absence of the appellant.

V. The appellant requested in writing, as its main request, that the decision under appeal be set aside and that the patent be maintained as granted. Alternatively, the appellant requested, as a first auxiliary request, that the patent be maintained in amended form on the basis of a set of claims 1 to 32 filed with letter of 18 August 2006, or, as a second auxiliary request, on the basis of claim 1 filed with letter of 7 February 2007 with the remaining claims to be amended *mutatis mutandis*.

VI. The respondent requested that the appeal be dismissed. Moreover, apportionment of costs for preparing and attending the oral proceedings was requested.

VII. Independent claims 1 and 17 of the appellant's **main request** read as follows:

"*1. Method for the automatic data acquisition (200), by means of a means for the same, of forms (10) whose design and information content (19) is not known in advance, by input into the said means, together with storage of patterns of the same, whereby the method is adaptive (224) and includes self-learning and registration of the design of forms not defining templates of forms in advance,* ***characterised in that*** *it includes the following steps to accomplish the adaptive registration (238):*

generation of a form map (18) based on the in advance unknown form's (10) design for identifying (210) information contained on the form;

searching and comparing (210, 220) the form map (18) with stored, registered maps in a means for storing form maps;
storage (238) of generated form maps (18) in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;
indication of agreement according to the limits for agreement when agreement is found;
and continued data acquisition (232) for identifying the information content of the form."

"17. Arrangement for the automatic data acquisition, by means of a means for the same, of forms (10) whose design and information content (19) is not known in advance, by input into the said means together with storage of patterns of the same it learns adaptively and registers the design of forms (10) not defining templates of forms in advance, and includes a computer, ***characterised by*** *the following means for carrying out the adaptive registration (238):*
means for generating a form map (18) based on the in advance unknown form's (10) design for identifying information contained on the form;
means for searching and comparing the form map with stored, recognised maps in a means for storing form maps;
means for storage of generated form maps (18) in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;
means for indicating agreement according to the limits for agreement when agreement is found; and
means for identification and continued data acquisition of the information content of the form."

Claims 2 to 16 and 18 to 32 are dependent claims.

Claims 1 and 17 of the **first auxiliary request** further define the step of generation of a form map and the corresponding means, respectively, by the respective feature "*wherein the step of generation comprises filtering the form map (18) from other objects*" and "*wherein generation comprises filtering the form map (18) from other objects*".

Claim 1 of the second auxiliary request reads:

"*1. Method for the automatic data acquisition (200), by means of a means for the same, of forms (10) whose design and information content (19) is not known in advance, by input into the said means, together with storage of patterns of the same, whereby the method is adaptive (224) and includes self-learning and registration of the design of forms not defining templates of forms in advance, c h a r a c t e r i s e d in that it includes the following steps to accomplish the adaptive registration (238):*
generation of a form map (18) based on the in advance unknown form's (10) design for identifying (210) information contained on the form;
wherein said form map (18) is a line map comprising horizontal and/or vertical line keys, representing line elements (14,15) from the form (10);
wherein the horizontal line keys are generated by dividing the form into a predetermined number of horizontal segments along the y-axis in a cartographic system of coordinates, whereby each segment is equivalent to a position in the horizontal key;
wherein the vertical line keys are generated by dividing the form into a predetermined number of

vertical segments along the y-axis in a cartographic system of coordinates, whereby each segment is equivalent to a position in the vertical key;

wherein at least one line element (14, 15) that is included in a vertical or horizontal segment is marked in the equivalent key position, and segments that lack line elements remain unmarked in the equivalent key position;

wherein the step of generation comprises filtering the form map (18) from other objects;

searching and comparing (210, 220) the form map (18), with stored, registered maps in a means for storing form maps;

if the identification (216) is unsuccessful, and there are no more line key candidates, self-learning (224) is accomplished with a form definition, said form definition consisting of a template or set of rules describing the common elements of a specific collection of forms;

confirmation of the identity of the form by the data acquisition of a ReCoGnition (RCG) value (214), which is a value at a given position that is unique for a certain form;

storage (238) of generated form maps (18) in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

indication of agreement according to the limits for agreement when agreement is found;

and continued data acquisition (232) for identifying the information content of the form."

VIII. In its decision the opposition division found the patent documents to lack viable information in particular as to how a form map could be automatically generated of and

information retrieved from a hitherto unknown form whose design and information content was not known in advance. In this context, specifically the process of "self-learning", which was mentioned in the description and involved undefined rules or templates, could not be implemented by a skilled person.

IX. The appellant essentially relied on the following submissions:

The method and system according to the invention needed, as was stated in paragraph [0012] of the patent specification "no prior knowledge of what the form looks like or where on the form the information is to be found". On the other hand, as was stated in paragraph [0055], for self-learning the system made use of a form definition which "consists of a template or a set of rules that describe the common elements of a specific collection of forms, for example, Swedish invoices". This meant that the system had some knowledge beforehand regarding how documents of a particular form category generally looked like and what kind of information to expect on a particular form category.
The skilled person was an experienced scientific programmer within image processing and pattern recognition and thus was fully capable to understand the term "self-learning" in the context of the present invention. In fact, self-learning was a known technique at the time of filing the present application, as was evidenced by each one of documents:

A1 : EP-A-0 809 219;
A2 : Claudia Wenzel, “Supporting Information Extraction from Printed Documents by Lexico-Semantic Pattern

Matching", Fourth International Conference on Document Analysis and Recognition 1997 (ICDAR'97), 18-20 Aug. 1997, vol. 2, pages 732 - 735; and
A3 : F. Cesarini, M. Gori, S. Marinai, G. Soda, "A System for Data Extraction from Forms of Known Class", Third International Conference on Document Analysis and Recognition 1995 (ICDAR'95), 14-16 Aug. 1995, vol. 2, pages 1136 - 1140.

By ignoring the knowledge of the skilled person, to whom it was clear how to set up the necessary rules and templates for self-learning in the context of evaluating the form map of an unknown form, the decision to revoke the patent was not taken on the appropriate grounds.

X. In the respondent's view, the appellant had admitted that the system of the patent under consideration was not able to "self-learn" a form when form design and information content were not known in advance. Moreover, rather than demonstrating that self-learning was common general knowledge at the priority date of the patent in suit, documents A1 to A3 merely established that self-learning of and information retrieval from a form required information about the form in advance. Besides, none of documents A1 to A3 formed part of the common general knowledge, with A1 and A2 having been published even after the priority date of the present patent.

In support of its request for an apportionment of costs which were incurred by preparing and attending the oral proceedings, the professional representative of the respondent submitted that she had received a copy of the appellant's announcement of 29 April 2009, not to attend the oral proceedings, on 11 May 2009, *ie* only two days

before the oral proceedings. At that time the flight ticket to Munich could no longer be cancelled and costs for preparing the oral proceedings had already been incurred. This unfortunate situation could have been avoided if the appellant had faxed its announcement on 29 April 2009 not only to the EPO but also to the respondent. In view of the board's critical appraisal of the merits of the appeal in the annex to the summons to oral proceedings, the representative would not have prepared and attended the oral proceedings if she had been informed in good time about the appellant's intention to stay absent. Hence for the reasons given in decision T 937/04, an apportionment of costs in favour of the respondent should be considered equitable.

Reasons for the Decision

1. In the light of the entry into force of the EPC 2000, reference is made to Article 7(1), 2nd sentence of the Revision Act of 29 November 2000 ("Act revising the Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973, last revised on 17 December 1991") and the transitional provisions for the amended and new provisions of the EPC (Decision of the Administrative Council of 28 June 2001), from which it may be derived which Articles of the EPC 1973 are still applicable and which Articles of the EPC 2000 shall apply.

2. The appeal complies with the requirements of Articles 106 to 108 and Rule 64 EPC 1973 and is, therefore, admissible.

3. Sufficiency of disclosure (Article 100(b) EPC 1973)

3.1 The system (method and arrangement) for the automatic data acquisition according to all independent claims of the appellant's requests on file aims at acquiring data of a (filled-in) form whose design and information content is not known in advance. A hitherto unknown form is scanned and, as a first measure, the design of the form (ie the pattern of lines and of object areas of the form without filled in text) is identified by generation of a form map. Once the form map is established the information content of the form is obtained by interpretation and validation.

As is explained in more detail in paragraph [0053] of the patent specification, identification of a form by means of the generation of a form map consists of generating a line map by identifying horizontal and vertical lines and, possibly, of generating an object area list. In the course of this process a so-called RCG-value (ReCoGnition) is acquired which uniquely identifies the design of a form. The retrieved form map is compared with known maps which have already been stored in a form database of the system. Subsequent to a successful identification, which is confirmed by a comparison of the RCG-value, the form's information content is interpreted and validated (Figure 2; paragraph [0054]).

As to how information retrieval would be accomplished in case the obtained form map is not yet present in the system's form database so that its identification by a comparison with the stored maps proves unsuccessful, the patent refers to a self-learning process to be performed

(see Figure 2; paragraph [0055]). The self-learning is explained to be accomplished with a "form definition" which consists of a template or a set of rules that describes the common elements of a specific collection of forms, for example, Swedish invoices. The RCG-value obtained in the course of this process is again checked for conformity with those of already existing forms in the system's form database. In case this comparison still failed, the patent specification states, without further explanation, that validation commences, after which the form is saved in the form map database. In this context, an operator can assist with the self-learning process (see Figure 2; paragraph [0056]).

3.2 It is immediately apparent from these pieces of information that the process of self-learning, which has to be applied when a form under examination is not yet known to the system and thus not yet stored in the system's form database, requires knowledge about a template or a set of rules each of which relate to "common elements of a specific collection of forms". This means, as is acknowledged by the appellant (see page 2, last paragraph of the statement setting out the grounds of appeal), that in the process of self-learning, as far as it is explained in the patent description, the system has in fact some *a priori* knowledge as to what documents of a particular form category are expected to look like, and what kind of information is to be expected on that particular form category.

However, as has been pointed out by the appellant and in the contested decision, such a procedure is incompatible with the claimed target that data acquisition is achieved in an automated manner from forms whose design

and information content is not known in advance. Therefore, the description of the self-learning process in the patent specification does not explain how the claimed ability to acquire data by an adaptive method including self-learning and registration of the design of forms not defining templates of forms would be put into practice.

In fact, the patent documents are silent as regards self-learning by templates or rules which do not imply some kind of *a priori* knowledge about potential forms. In particular given the fact that a form may include object areas which contain form text, it is not imaginable what kind of "templates" or "rules" for establishing a form map would allow to distinguish such form text from filled-in text so as to retrieve, in an automatic manner (*ie* without the assistance of an operator), novel form map information as well the form's data content, as is implied in paragraph [0056] of the patent specification.

3.3 The appellant did not respond to this specific objection which was raised in the boards communication.

Instead, in the statement setting out the grounds of appeal, the appellant made reference to documents A1 to A3 in an attempt to prove that self-learning involving rules or templates was a technique which belonged to the common general knowledge in the technical field at issue.

However, the appellant's submission is not convincing. Already from a formal point of view, documents A1 and A2 having publication dates in November 1997 and August 1997, respectively, cannot form part of existing

common general knowledge at the priority date of the patent in suit, which is 1 April 1997.

Moreover, as has been pointed out by the respondent in its letter of 20 December 2007 and uncontested by the appellant, documents A1 to A3 all discuss a system which is capable of self-learning in the context of information retrieval from forms only on the basis of some kind of *a priori* knowledge about the form. Reference is made in this respect to page 6, lines 28 to 33 of document A1, to page 732, right hand column, fourth paragraph of A2, and page 1137, left hand column, second paragraph in A3. Thus, none of documents A1 to A3 teaches how information could be retrieved from forms when nothing was known of the forms' design or information content in advance.

3.4 In conclusion, the board has found that, on the basis of the information given in the patent in suit, it is not conceivable how, in an automated manner and without any *a priori* knowledge, a form map of a hitherto unknown form can be generated and at the same time the information content of the form can be retrieved.

This deficiency does not only apply to the appellant's main request but also to the two auxiliary requests on file, the independent claims of which define in increasing detail the step of generating a form map but do not provide any instructions which would enable the skilled person to retrieve the information content from an unknown form.

Accordingly, the Board has come to the conclusion that none of the appellant's requests on file is allowable.

4. Apportionment of costs (Article 104 EPC)

4.1 Article 104(1) EPC provides that "each party to the opposition proceedings shall bear the costs it has incurred, unless the Opposition Division, for reasons of equity, orders, in accordance with the Implementing Regulations, a different apportionment of costs".

By virtue of Article 111(1) EPC this provision applies also to opposition appeal proceedings.

4.2 In support of its request for an apportionment of costs, the respondent has made reference to decision T 937/04 of the boards of appeal, in which it was held that an appellant, by informing solely the European Patent Office and not the other parties only one working day before the oral proceedings of its intention not to attend the oral proceedings, had failed to exercise all due care required. The then deciding board based this finding on the understanding that, if a party decided only shortly before the date scheduled for oral proceedings that it was not going to attend them, its equitable obligations extended to informing any other parties to the appeal proceedings of its decision not to attend oral proceedings (see point 5.1 of the reasons for the decision).

In the present board's view, the factual situation of the case at issue is decisively distinguished from that of case T 937/04 because of the fact that the appellant had informed the EPO of its intention not to attend the oral proceedings already two weeks before the scheduled date of the oral proceedings. Under these circumstances,

the appellant cannot be held responsible for an unfortunate course of actions as a consequence of which the information reached the respondent's representative only two working days before the scheduled date, nor for the fact that travel arrangements for the representative had been such that they could no longer be cancelled at that time. Already for this reason a different apportionment of costs in favour of the respondent is not considered to be equitable.

4.3 Besides, the board does not share the standards based on which the discretion in case T 937/04 was exercised. In the present board's view, a party to proceedings before the EPO has a right to be heard in oral proceedings but has no obligation to attend oral proceedings to which it had been summoned (see for instance T 544/94, point 5 of the Reasons). Thus, to inform the EPO and any other party to the proceedings about the intention not to attend oral proceedings in due time before the scheduled date is a matter of courtesy and respect rather than a procedural obligation to be met. In order to judge a late announcement of an intended absence to oral proceedings as constituting a "culpable action of an irresponsible or even malicious nature" (see T 937/04, point 5.1 of the Reasons) strong supporting evidence for such an allegation would be required.

Moreover, the mere fact that a party does not attend oral proceedings does not mean by itself that its case will be lost, making the presence of the other party unnecessary. The outcome of a case cannot be certain until the board has taken its final decision. There can be no doubt in this context that views expressed in a board's preliminary communication are nothing but

provisional assessments which are by no means binding. Thus, however critical such views would be, there is no guarantee that they reflect the final decision. Therefore, each party is obliged to decide on its own and independently from the behaviour of another party to the proceedings whether or not to attend oral proceedings.

Order

For these reasons it is decided that:

1. The appeal is dismissed.

2. The request of the respondent for apportionment of costs is rejected

The Registrar | The Chairman

R. Schumacher | B. Schachenmann