



(51) International Patent Classification:

G06F 17/27 (2006.01) G06F 3/0488 (2013.01)
G06F 3/023 (2006.01)

(21) International Application Number:

PCT/US2014/011893

(22) International Filing Date:

16 January 2014 (16.01.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/759,274 31 January 2013 (31.01.2013) US
13/774,656 22 February 2013 (22.02.2013) US

(71) Applicant: **GOOGLE INC.** [US/US]; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US).

(72) Inventors: **OUYANG, Yu**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **ZHAI, Shumin**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **CHELBA, Ciprian Ioan**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US).

(74) Agent: **VREDEVELD, Albert W.**; 1625 Radio Drive, Suite 300, Woodbury, Minnesota 55125 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

[Continued on next page]

(54) Title: CHARACTER AND WORD LEVEL LANGUAGE MODELS FOR OUT-OF-VOCABULARY TEXT INPUT

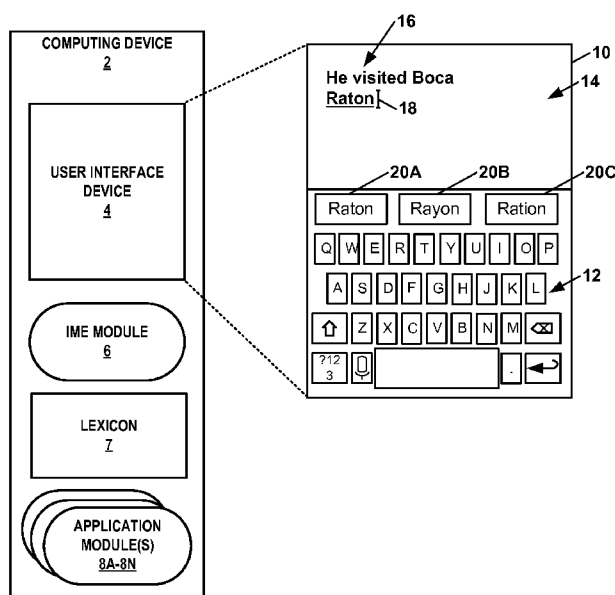


FIG. 1

(57) Abstract: A computing device determines, based at least in part on indications of user input, scores for a first set of candidate strings and a second set of candidate strings. Each candidate string from the first set of candidate strings is in a lexicon. Candidate strings from the second set of candidate strings are not necessarily in the lexicon. The computing device determines the scores for the first set of candidate strings based on probabilities of the candidate strings being entered. For each candidate string from the second set of candidate strings, the computing device determines the scores for the candidate string based on probabilities of characters of the candidate string being entered. The computing device selects a candidate string based on the scores for the first and second sets of candidate strings and outputs, for display at the display device, the selected candidate string.



- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))* — *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

Published:

- *with international search report (Art. 21(3))*

(88) Date of publication of the international search report:

23 October 2014

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/011893

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/27 G06F3/023 G06F3/0488
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2010/315266 A1 (GUNAWARDANA ASELA J [US] ET AL) 16 December 2010 (2010-12-16) abstract paragraphs [0051] - [0070] -----	1-20
E	WO 2014/066106 A2 (GOOGLE INC [US]) 1 May 2014 (2014-05-01) abstract claims 1,2,6,7,12,15 paragraphs [0033], [0038] - [0057]; figures 3-6 -----	1-5, 7-10, 18-20
A	US 6 654 733 B1 (GOODMAN JOSHUA [US] ET AL) 25 November 2003 (2003-11-25) abstract column 9, line 58 - column 10, line 52 ----- -/-	11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 August 2014

Date of mailing of the international search report

26/08/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Woods, Justin

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/011893

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009/216690 A1 (BADGER ERIC [US] ET AL) 27 August 2009 (2009-08-27) abstract paragraphs [0034] - [0037] -----	12
A	Tommi A Pirinen ET AL: "Creating and Weighting Hunspell Dictionaries as Finite-State Automata", 1 January 2010 (2010-01-01), XP55135123, Retrieved from the Internet: URL:https://tuhat.halvi.helsinki.fi/portal /services/downloadRegister/20974061/Pirine n_il_2010.pdf [retrieved on 2014-08-18] abstract Sections 1, 4 -----	16,17
A	Joshua Goodman ET AL: "Language Modeling for Soft Keyboards", Microsoft Technical Report MSR-TR-2001-118, 28 November 2001 (2001-11-28), XP055134921, Microsoft Research, Redmond, WA 98052, USA Retrieved from the Internet: URL:http://research.microsoft.com/pubs/699 03/tr-2001-118.pdf [retrieved on 2014-08-15] abstract Sections entitled "Introduction", "Mathematical models", "Letter sequence language model" and "Pen down position model" on pages 1-3 -----	1-20
A	YOUSSEF BASSIL: "Parallel Spell-Checking Algorithm Based on Yahoo! N-Grams Dataset", INTERNATIONAL JOURNAL OF RESEARCH AND REVIEWS IN COMPUTER SCIENCE (IJRRCS), vol. 3, no. 3, 15 February 2012 (2012-02-15), XP055060899, ISSN: 2079-2557 abstract Section 2 -----	1-20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/011893

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 1-18(partially)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 1-18(partially)

A repetition of the claims in the summary notwithstanding, the description, supported by the drawings shows a string being input by the user, see for example the input string "Raton" in figure 1. Claim 1, however, sets out "the first set of candidate strings being entered" and then "the second set of candidate strings being entered". As the description does not disclose such first and second sets of strings, claim 1 is not clear and concise and fully supported by the description as required by Article 6 PCT.

Moreover, while first and second sets of candidate strings are indeed computed (see claim 1 and, figure 1 and description as a whole), they are not entered as would be implied by the sections of the claim cited above, again rendering claim 1 unclear in the sense of Article 6 PCT.

As the first and second candidate strings are not actually entered, it is not at all meaningful to search for such a constellation. Accordingly, claim 1 has been interpreted in accordance with the application as a whole, whereby the words "first set of candidate" on line 5 of claim 1 and "second set of candidate" on line 9 of claim 1 are not considered. Such an interpretation is not only consistent with the remainder of the description and the drawings, but also claims 19 and 20, in which a single input is implied.

In as far as claims 2-18 are dependent on claim 1, they are likewise unclear (Article 6 PCT). They have been searched in the light of the interpretation of claim 1 set out above.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guidelines C-IV, 7.2), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2014/011893

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2010315266	A1	16-12-2010	NONE
WO 2014066106	A2	01-05-2014	US 2014122057 A1 01-05-2014 WO 2014066106 A2 01-05-2014
US 6654733	B1	25-11-2003	NONE
US 2009216690	A1	27-08-2009	US 2009216690 A1 27-08-2009 US 2011270786 A1 03-11-2011