

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 7 月 3 日 (03.07.2025)



(10) 国际公布号
WO 2025/140047 A1

- (51) 国际专利分类号:
G06F 7/483 (2006.01)
- (21) 国际申请号: PCT/CN2024/141099
- (22) 国际申请日: 2024 年 12 月 20 日 (20.12.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202311811917.5 2023 年 12 月 25 日 (25.12.2023) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。
- (72) 发明人: 罗元勇 (LUO, Yuanyong); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。

张忠星 (ZHANG, Zhongxing); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。胡桂鹏 (HU, Guipeng); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。胡天驰 (HU, Tianchi); 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司 (TDIP & PARTNERS); 中国北京市西城区裕民路 18 号北环中心 A 座 2002 100029 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA,

(54) Title: FLOATING POINT NUMBER PROCESSING METHOD, APPARATUS, COMPUTING DEVICE AND STORAGE MEDIUM

(54) 发明名称: 一种浮点数处理方法、装置、计算设备和存储介质

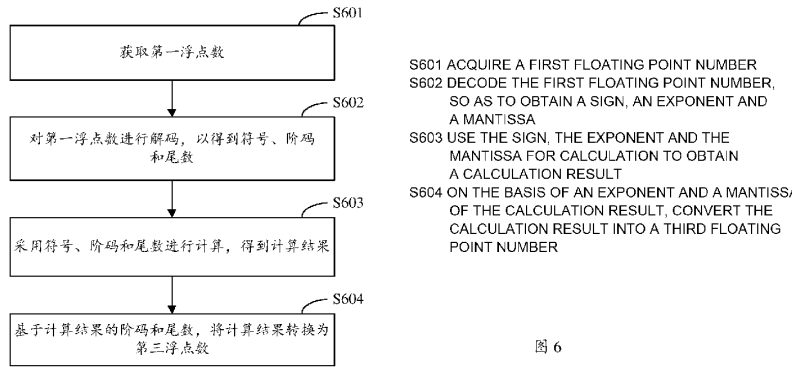


图 6

(57) Abstract: The present application relates to the technical field of computers. Disclosed are a floating point number processing method, an apparatus, a computing device and a storage medium. The computing device acquires a first floating point number, and decodes the first floating point number, so as to obtain a sign, an exponent and a mantissa. The first floating point number may comprise a first sign field, a bit-width indication field, a first exponent field and a first mantissa field. The bit-width indication field is used for indicating the bit width D occupied by the first exponent field in the total bit width N of the floating point number, and when the total bit width is fixed, the bit width of the first exponent field and the bit width of the first mantissa field may dynamically change along with a numerical value indicated by the bit-width indication field. The bit-width indication field may also indicate whether the first mantissa field represents an offset of the mantissa or of the exponent; and multiplexing the first mantissa field can ensure small impacts on the precision of the floating point number, and can increase the coding space of the exponent, thereby expanding the dynamic range of the numerical value represented by the floating point number, and satisfying the requirements of AI training and reasoning.

(57) 摘要: 本申请公开了一种浮点数处理方法、装置、计算设备和存储介质, 涉及计算机技术领域。

[见续页]

WO 2025/140047 A1

PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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)。

本国际公布：
一 包括国际检索报告(条约第21条(3))。

其中，计算设备获取第一浮点数，对第一浮点数进行解码，以得到符号、阶码和尾数。第一浮点数可以包括第一符号域、位宽指示域、第一阶码域和第一尾数域。位宽指示域用于指示第一阶码域在浮点数的总位宽N中所占用的位宽D，在总位宽固定时，第一阶码域的位宽和第一尾数域的位宽可以随着位宽指示域指示的数值动态变化。位宽指示域还可以指示第一尾数域表征的是尾数还是阶码的偏移量，通过复用第一尾数域，既可以保证对浮点数的精度影响较小，又可以增加阶码的编码空间，从而扩大浮点数所表示的数值的动态范围，可以满足AI训练和推理的需要。