

(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



(10) 国际公布号
WO 2025/020611 A1

(43) 国际公布日
2025 年 1 月 30 日 (30.01.2025)

(51) 国际专利分类号:
G06F 16/332 (2019.01) **G06F 40/35** (2020.01)
(21) 国际申请号: PCT/CN2024/090401
(22) 国际申请日: 2024 年 4 月 28 日 (28.04.2024)
(25) 申请语言: 中文
(26) 公布语言: 中文
(30) 优先权:
202310913820.9 2023 年 7 月 24 日 (24.07.2023) CN

(71) 申请人: 京东科技控股股份有限公司
(JINGDONG TECHNOLOGY HOLDING CO., LTD.)
[CN/CN]; 中国北京市大兴区北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。

(72) 发明人: 王秋霖(WANG, Qiulin); 中国北京市大兴区北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。张立平(ZHANG, Liping); 中国北京市大兴区北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。曹俊豪(CAO, Junhao); 中国北京市大兴区北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。单新媛(SHAN, Xinyuan); 中国北京市大兴区北京经济技术开发区科创十一街 18 号 C 座 2 层 221 室, Beijing 100176 (CN)。

(74) 代理人: 北京律智知识产权代理有限公司 (BEIJING INTELLEGAL INTELLECTUAL PROPERTY AGENT LTD.); 中国北京市朝阳区慧忠路 5 号 B1605、B1606、B1607, Beijing 100101 (CN)。

(54) Title: SESSION RESPONSE METHOD AND APPARATUS, ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 会话应答方法和装置、电子设备、存储介质

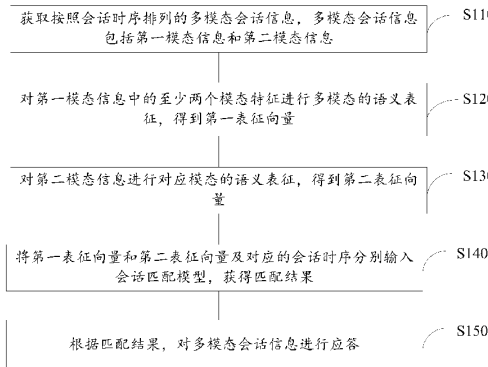


图 1

- S110 Acquire multi-modal session information arranged according to a session time sequence, wherein the multi-modal session information comprises first modal information and second modal information
- S120 Perform multi-modal semantic representation on at least two modal features in the first modal information to obtain a first representation vector
- S130 Perform semantic representation of a corresponding modal on the second modal information to obtain a second representation vector
- S140 Respectively input the first representation vector, the second representation vector and a corresponding session time sequence into a session matching model to obtain a matching result
- S150 Respond to the multi-modal session information according to the matching result

(57) Abstract: Provided are a session response method and apparatus, an electronic device and a storage medium. The method comprises: acquiring multi-modal session information arranged according to a session time sequence, wherein the multi-modal session information comprises first modal information and second modal information (S110); performing multi-modal semantic representation on at least two modal features in the first modal information to obtain a first representation vector (S120); performing semantic representation of a corresponding modal on the second modal information to obtain a second representation vector (S130); respectively inputting the

[见续页]

WO 2025/020611 A1

(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, 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:

— 包括国际检索报告(条约第21条(3))。

first representation vector, the second representation vector and a corresponding session time sequence into a session matching model to obtain a matching result (S140); and responding to the multi-modal session information according to the matching result (S150).

(57) 摘要: 提供一种会话应答方法和装置、电子设备、存储介质。方法包括: 获取按照会话时序排列的多模态会话信息, 多模态会话信息包括第一模态信息和第二模态信息(S110); 对第一模态信息中的至少两个模态特征进行多模态的语义表征, 得到第一表征向量(S120); 对第二模态信息进行对应模态的语义表征, 得到第二表征向量(S130); 将第一表征向量和第二表征向量及对应的会话时序分别输入会话匹配模型, 获得匹配结果(S140); 根据匹配结果, 对多模态会话信息进行应答(S150)。