

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200710128348.9

[51] Int. Cl.

G06F 17/30 (2006.01)

G06F 17/27 (2006.01)

H04Q 7/32 (2006.01)

[43] 公开日 2008 年 3 月 5 日

[11] 公开号 CN 101136028A

[22] 申请日 2007.7.10

[21] 申请号 200710128348.9

[30] 优先权

[32] 2006.7.10 [33] CN [31] 200610106226.5

[71] 申请人 日电（中国）有限公司

地址 100007 北京市东城区东四十条甲 22 号  
南新仓国际大厦 B 栋 12 层 1222 室

[72] 发明人 丰强泽 齐红威 刘宇光 刘慧峰  
王亚波 刘晓炜 福岛俊一

[74] 专利代理机构 中科专利商标代理有限责任公司  
代理人 罗松梅

权利要求书 15 页 说明书 26 页 附图 24 页

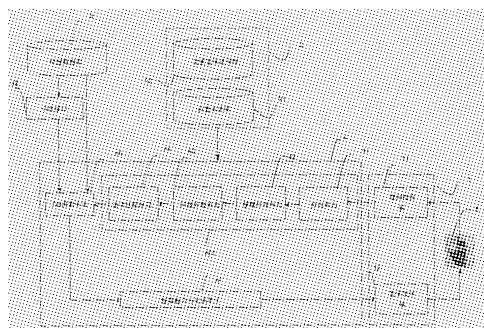
[54] 发明名称

基于自然语言的位置查询系统以及基于关键词的位置查询系统

合句查询和有语义错误的查询，从而提高了用户提问的自由度，使位置查询更加灵活、精确。

[57] 摘要

一种基于自然语言的位置查询系统和方法，所述系统包括：接收装置，用于从用户终端接收自然语言查询请求；访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；自然语言查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；和发送装置，用于将所述位置信息发送到用户终端。以及一种基于关键词的位置查询系统和方法与一种基于自然语言和关键词的位置查询系统和方法。本发明不但可以处理用户的模糊查询、间接查询，还可以处理复



1.一种基于自然语言的位置查询系统，包括：

接收装置，用于从用户终端接收自然语言查询请求；

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；

自然语言查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；和

发送装置，用于将所述位置信息发送到用户终端。

2. 如权利要求1所述的系统，其中所述自然语言查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理单元，用于通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

查询语言匹配单元，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

3. 如权利要求1所述的系统，其中所述自然语言查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配单元，将处理后的查询请求与位置查询语言库进行匹

配，并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

4. 如权利要求1所述的系统，其中所述自然语言查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理单元，用于通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配单元，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

5. 如权利要求2至4之一所述的系统，其中所述自然语言查询处理装置还包括：

复合句处理单元，用于将包括复合句的自然语言查询请求拆分为多个单一的自然语言查询请求并提供给分词单元。

6. 如权利要求2至4之一所述的系统，其中所述查询语言匹配单元在查询请求不满足匹配文法中的概念约束时，通过查找位置本体库中的映射本体，将相关描述转换为满足约束的概念。

7. 如权利要求5所述的系统，其中所述自然语言查询处理装置还包括：

错误诊断单元，用于在查询语言匹配单元匹配失败后按照预定的规则识别并分析查询请求中的语义错误，并将诊断出的错误信息经过答案融合

---

与生成单元发送给用户终端。

8.一种基于自然语言的位置查询方法，包括：

接收步骤，接收用户通过用户终端发送的自然语言查询请求；

自然语言查询处理步骤，通过查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送步骤，将搜索到的位置信息发送到用户终端。

9. 如权利要求8所述的方法，其中所述自然语言查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

10. 如权利要求8所述的方法，其中所述自然语言查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

11. 如权利要求 8 所述的方法，其中所述自然语言查询处理步骤包括：分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

12. 如权利要求 9 至 11 之一所述的方法，其中所述自然语言查询处理步骤还包括：

复合句处理步骤，用于将包括复合句的自然语言查询请求拆分为多个单一的自然语言查询请求并发送到分词步骤。

13. 如权利要求 9 至 11 之一所述的方法，其中所述查询语言匹配步骤还包括：

若查询请求不满足匹配文法中的概念约束，则通过查找位置本体库中的映射本体，将相关描述转换为满足约束的概念。

14. 如权利要求 12 所述的方法，其中所述自然语言查询处理步骤还包括：

错误诊断步骤，用于在查询语言匹配失败后按照预定的规则识别并分析查询请求中的语义错误，并将诊断出的错误信息发送给用户终端。

15. 如权利要求 8 所述的方法，其中还包括位置本体库生成步骤，其包括：

领域本体生成步骤，用于从电子地图和黄页信息的信息源中提取实体表与类别表、属性以及关系，并与预定的公理部分进行组合，从而生成针

对每个领域的领域本体；

映射本体生成步骤，用于根据同义词典、简称规则库、多语言词典和电子地图，抽取同义映射关系、语言映射关系和空间映射关系中的至少一个，从而生成映射本体。

16. 如权利要求 8 所述的方法，其中还包括位置查询语言库生成步骤，其包括：

领域查询语言生成步骤，用于收集各个领域的可能的问句，从各领域的用户问句集合中抽取文法、常量表，并将提取的文法以及文法对应的查询动作组合为领域查询语言；

公用查询语言生成步骤，用于计算各领域查询语言的文法间的相似度，并抽取公用查询语言。

17. 一种基于关键词的位置查询系统，包括：

接收装置，用于从用户终端接收关键词查询请求；

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；

关键词查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送装置，用于将所述位置信息发送到用户终端。

18. 如权利要求 17 所述的系统，其中所述关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理单元，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

文法部分匹配单元，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

19. 如权利要求 17 所述的系统，其中所述关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

文法部分匹配单元，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

20. 如权利要求 17 所述的系统，其中所述关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理单元，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

文法部分匹配单元，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询

---

结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

21. 如权利要求 18 至 20 之一所述的系统，其中所述关键词查询处理装置还包括：

错误诊断单元，用于在文法部分匹配单元匹配失败后按照预定的规则识别并分析查询请求中的语义错误，并将诊断出的错误信息经过答案融合与生成单元发送给用户终端。

22. 如权利要求 18 至 20 之一所述的系统，其中所述答案决策单元在查询请求不满足匹配文法中的概念约束时，通过查找位置本体库中的映射本体，将相关描述转换为满足约束的概念。

23. 一种基于关键词的位置查询方法，包括：

接收步骤，用于从用户终端接收用户的关键词查询请求；

关键词查询处理步骤，用于通过查找位置本体库和位置查询语言库，对用户终端发送的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送步骤，将所述搜索出的位置信息发送到用户终端。

24. 如权利要求 23 所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

文法部分匹配步骤，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策步骤，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发

送到用户终端。

25. 如权利要求 23 所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

文法部分匹配步骤，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策步骤，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

26. 如权利要求 23 所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

文法部分匹配步骤，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策步骤，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

27. 如权利要求 24 至 26 之一所述的方法，其中所述关键词查询处理步骤还包括：

错误诊断步骤，用于在文法部分匹配失败后按照预定的规则识别并分析查询请求中的语义错误，并将诊断出的错误信息经过答案融合与生成单元发送给用户终端。

28. 如权利要求 24 至 26 之一所述的方法，其中所述答案决策步骤还包括：

若查询请求不满足匹配文法中的概念约束，则通过查找位置本体库中的映射本体，将相关描述转换为满足约束的概念。

29. 如权利要求 23 所述的方法，其中还包括位置本体库生成步骤，其包括：

领域本体生成步骤，用于从电子地图、黄页信息的信息源中提取实体表与类别表、属性以及关系，并与预定的公理部分进行组合，从而生成针对每个领域的领域本体；

映射本体生成步骤，用于根据同义词典、简称规则库、多语言词典和电子地图，抽取同义映射关系、语言映射关系和空间映射关系中的至少一个，从而生成映射本体。

30. 如权利要求 23 所述的方法，其中还包括位置查询语言库生成步骤，其包括：

领域查询语言生成步骤，用于从各领域的用户问句集合中抽取文法、常量表，并将提取的文法以及文法对应的查询动作组合为领域查询语言；

公用查询语言生成步骤，用于计算各领域查询语言的文法间的相似度，并抽取公用查询语言。

31. 一种基于自然语言和关键词的位置查询系统，包括：

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；

自然语言和关键词查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的自然语言和/或关键词查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；

发送装置，用于将所述位置信息发送到用户终端。

32. 如权利要求31所述的系统，其中所述自然语言和关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对查询请求进行分词；

模糊处理单元，用于通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

判断单元，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

查询语言匹配单元，用于当查询请求是自然语言查询时，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

文法部分匹配单元，用于当查询请求是关键词查询时，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于当查询请求是关键词查询时，根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

33. 如权利要求31所述的系统，其中所述自然语言和关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

判断单元，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

查询语言匹配单元，用于当查询请求是自然语言查询时，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

文法部分匹配单元，用于当查询请求是关键词查询时，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于当查询请求是关键词查询时，根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

34. 如权利要求31所述的系统，其中所述自然语言和关键词查询处理装置包括：

分词单元，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理单元，用于通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理单元，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

判断单元，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

查询语言匹配单元，用于当查询请求是自然语言查询时，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

文法部分匹配单元，用于当查询请求是关键词查询时，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策单元，用于当查询请求是关键词查询时，根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索单元，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成单元，用于融合查询结果以便生成答案，并将答案发送到用户终端。

35. 如权利要求 32 至 34 之一所述的系统，其中所述查询语言匹配单

---

元和答案决策单元在查询请求不满足匹配文法中的概念约束时，通过查找位置本体库中的映射本体，将相关描述转换为满足约束的概念。

36. 一种基于自然语言和关键词的位置查询方法，包括：

判断步骤，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

在用户的查询请求是自然语言查询请求时，包括：

自然语言查询处理步骤，用于通过查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；

第一发送步骤，将所述位置信息发送到用户终端；和

在用户的查询请求是关键词查询请求时，包括：

关键词查询处理步骤，用于通过查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；和

第二发送步骤，将所述位置信息发送到用户终端。

37. 如权利要求 36 所述的方法，其中所述自然语言查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

38. 如权利要求 36 所述的方法，其中所述自然语言查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

39. 如权利要求 36 所述的方法，其中所述自然语言查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对自然语言查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

查询语言匹配步骤，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

40. 如权利要求36所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述

的模糊描述包括冗余描述和不完整描述中的至少之一；

文法部分匹配步骤，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策步骤，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

41. 如权利要求36所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

文法部分匹配步骤，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；

答案决策步骤，用于根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；

数据库搜索步骤，用于根据查询动作从位置数据库中搜索相应的查询结果；

答案融合与生成步骤，用于融合查询结果以便生成答案，并将答案发送到用户终端。

42. 如权利要求36所述的方法，其中所述关键词查询处理步骤包括：

分词步骤，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对用户的关键词查询请求进行分词；

模糊处理步骤，通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；

间接处理步骤，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；

---

文法部分匹配步骤,将处理后的查询请求与位置查询语言库进行部分匹配,并获得匹配文法集合;

答案决策步骤,用于根据预定的决策规则,从匹配文法集合中选择最优匹配文法并生成对应的查询动作;

数据库搜索步骤,用于根据查询动作从位置数据库中搜索相应的查询结果;

答案融合与生成步骤,用于融合查询结果以便生成答案,并将答案发送到用户终端。

43. 如权利要求 37 至 39 之一所述的方法,其中所述查询语言匹配步骤还包括:

若查询请求不满足匹配文法中的概念约束,则通过查找位置本体库中的映射本体,将相关描述转换为满足约束的概念。

44. 如权利要求 40 至 42 之一所述的方法,其中所述答案决策步骤还包括:

若查询请求不满足匹配文法中的概念约束,则通过查找位置本体库中的映射本体,将相关描述转换为满足约束的概念。

## 基于自然语言的位置查询系统以及基于关键词的位置查询系统

### 技术领域

本发明通常涉及位置搜索引擎领域，更具体地，本发明涉及一种基于自然语言的位置查询系统及其查询方法，一种基于关键词的位置查询系统及其查询方法以及一种基于自然语言和关键词的位置查询系统及其查询方法。

### 背景技术

随着移动通信网络的发展，特别是 3G 在中国开始的前夕，与人们日常生活密切相关的本地查询服务在不久的未来将具有巨大的市场。

此外，随着经济与技术的快速发展，人们的生活或工作越来越全球化。当到达一个陌生的环境时，人们会需要关于酒店，汽车租赁，或者医疗服务的信息。更特别地，人们甚至需要在新住处查找例如管道工人，饭店，会计师或花匠等等。

所以位置搜索非常重要。但是在当前的搜索引擎中，用户提交一个查询请求后，经常不得不从很多结果中选取自己想要的答案。特别是用户利用手机进行位置信息搜索的话，则因为手机的屏幕太小，手机键盘输入受限，从而需要花费大量的时间翻页查找合适的结果。当前的搜索引擎并不能够提供较为准确且高效的查询结果。首先，用户希望使用一种很自由的查询方式（如格式不受限的自然语言），其次，用户希望系统能返回准确简明的信息。

存在一些向用户提供本地信息服务的已有的搜索引擎。例如，Google 本地搜索允许用户在一个特定的区域内查找商店和商业信息。Yahoo 本地搜索可以用来查找饭店，娱乐场所和商业信息。它们针对移动用户和网络用户采用了相同的搜索机制，因此针对移动用户的本地搜索结果和利用电脑进行本地查询的结果相同。

Mobile Info Search 能自动从网页内容中收集位置相关的各种信息，如单位名称，地址等，然后把这些信息与地图进行关联，最后提供给用户用关键词的方式查询位置信息。

Answers Anywhere 是一个中间件平台，通过 SMS、MMS、语音等交互模式为无线设备的用户提供数据服务的直接访问。它提供了高级的自然语言和文法辨识技术。

此外，一些现有的专利专注于在信息服务系统中提供自然语言的处理方法。

在美国专利 US2002161587A1 中提出了一种用于在基于位置的服务系统中提供自然语言处理的方法。此系统能够接收语音请求并根据所述远程终端的地理位置生成应答。

在中国专利 CN1466367A 中提出了一种移动的人知交互系统和方法。用户可以通过手机来查询多领域的知识（比如天气信息，交通信息）。它有两个优点：1 此系统允许用户使用自然语言进行查询。2 此系统能够识别并纠正用户查询中的拼写错误。

虽然移动用户偏好自由并准确地查询本地信息，但是不同的用户的搜索偏好彼此不同。首先，一个用户可以间接地搜索本地信息（比如用“好吃”代替查询“饭店”）；其次，有些用户喜欢用自然语言搜索，另外的用户喜欢用关键字搜索；第三，用户的查询可以是复合句；第四，用户的查询可能较模糊。最后，当用户查询的时候，他可能添加或丢失一些字符，甚至是在查询中存在语义的错误。

由于以上的情况，现有技术中存在着缺点。

Google 和 yahoo 的本地搜索以及 Mobile Info Search 都是基于关键字查询的搜索引擎，并不能处理自然语言查询。

尽管 Answers Anywhere 提供了自然语言查询，但难以处理位置服务中灵活的间接查询和复合句查询，另外模糊处理和语义错误分析能力也较弱。

专利 US2002161587A1 只提供了英语自然语言查询，而且该专利只考虑了词性而并没有考虑较为丰富的语义信息。因此，在处理模糊查询和带有语义错误的查询时受到限制。

---

专利 CN1466367A 是一个汉语问答系统，该系统并不能处理复合句、关键词查询以及语义错误的分析。

## 发明内容

因此，考虑到上述问题提出本发明。本发明的目的是提供一种利用位置本体库和位置查询语言库进行位置查询的系统和方法。其中位置本体库提供了丰富的关于位置信息的语义知识。位置查询语言库提供了语言规则用来分析基于自然语言和基于关键字的查询。

根据本发明第一方面，提出了一种基于自然语言查询的位置查询系统，包括

接收装置，用于从用户终端接收自然语言查询请求；

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；

自然语言查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个并从位置数据库中搜索出对应的位置信息；和

发送装置，用于将所述位置信息发送到用户终端。

根据本发明第二方面，提出了一种基于自然语言查询的位置查询方法，包括：

接收步骤，接收用户通过用户终端发送的自然语言查询请求；

自然语言查询处理步骤，通过查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送步骤，将搜索到的位置信息发送到用户终端。

根据本发明第三方面，提出了一种基于关键词查询的位置查询系统，包括：

接收装置，用于从用户终端接收关键词查询请求；

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询

---

语言库包括用于位置服务查询的语言模型；

关键词查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送装置，用于将所述位置信息发送到用户终端。

根据本发明第四方面，提出了一种基于关键词查询的位置查询方法，包括：

接收步骤，用于从用户终端接收用户的关键词查询请求；

关键词查询处理步骤，用于通过查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

发送步骤，将所述搜索出的位置信息发送到用户终端。

根据本发明第五方面，提供了一种基于自然语言和关键词的位置查询系统，包括：

判断单元，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

访问装置，用于访问系统内部或外部的位置本体库和位置查询语言库，其中所述位置本体库包括位置服务中的领域知识描述，所述位置查询语言库包括用于位置服务查询的语言模型；

自然语言查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；

关键词查询处理装置，用于通过访问装置查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；

发送装置，用于将所述位置信息发送到用户终端。

根据本发明第六方面，提供了一种基于自然语言和关键词的位置查询方法，包括：

判断步骤，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；

在用户的查询请求是自然语言查询请求时，包括：

自然语言查询处理步骤，用于通过查找位置本体库和位置查询语言库，对接收的自然语言查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；

第一发送步骤，将所述位置信息发送到用户终端；和

在用户的查询请求是关键词查询请求时，包括：

关键词查询处理步骤，用于通过查找位置本体库和位置查询语言库，对接收的关键词查询请求进行模糊处理和间接处理中的至少一个，并从位置数据库中搜索出对应的位置信息；和

第二发送步骤，将所述位置信息发送到用户终端。

此外，根据本发明第一方面的基于自然语言查询的系统还可以进行复合句分析和对包括有语义错误的查询的错误诊断处理。根据本发明第二方面基于关键词查询的系统还可以进行错误诊断处理。

因此，根据本发明的基于自然语言的位置查询系统不仅可以执行复合句查询、间接查询、模糊查询的分析处理，而且还能够实现对有语义错误的查询的分析处理。根据本发明的基于关键词的位置查询系统不仅可以执行对间接查询、模糊查询的分析处理，而且还能够实现对有语义错误的查询的分析处理。根据本发明的基于自然语言和关键词的位置查询系统允许用户使用自然语言或关键词的方式来访问同一接口，从而可以针对不同用户的偏好以及需求，更为灵活，准确和简明地提供用户所需的位置信息，同时提高了用户查询的自由度，并改进了已有的搜索系统中存在的缺点。

## 附图说明

通过以下结合附图，对优选实施例的描述，本发明的这些和/或其他目的和特征将变得更加清楚而且更为容易理解，其中：

图 1 是根据本发明的基于自然语言查询的位置查询系统的方框图；

图 2 是根据本发明的基于自然语言查询的位置搜索方法的流程图；

图 3 是根据本发明的基于关键词查询的位置查询系统的结构图；

图 4 是根据本发明的基于关键词查询的位置搜索方法的流程图；

图 5a 示出了根据本发明的位置本体库的示意性结构；

图 5b 示出了根据本发明的位置本体库的类别表和实体表的一个例子；

图 5c 示出了根据本发明的位置本体库的概念部分的一个例子；

图 5d 示出了根据本发明的位置本体库中的属性部分和关系部分的一个示例；

图 5e 示出了根据本发明的位置本体库中的公理部分的一个示例；

图 5f 示出了根据本发明的位置本体库中的位置概念空间的示意结构图；

图 6 示出了根据本发明的位置查询语言库的示意性结构；

图 7 是根据本发明的答案融合和生成单元融合并生成答案的示意图；

图 8 是根据本发明的答案模板的结构示意图；

图 9a 是根据本发明的自然语言查询处理装置对自然语言查询进行处理的一个示例；

图 9b 是根据本发明的自然语言查询处理装置对包括复合句的自然语言查询进行处理的另一个示例；

图 10 是根据本发明的关键词查询处理装置对关键词查询进行处理的一个示例；

图 11a 是基于自然语言的位置查询系统执行的查询的一个示例；

图 11b 是基于关键词的位置查询系统执行的查询的一个示例；

图 12 是根据本发明另一个实施例的基于自然语言查询的位置查询系统的方框图；

图 13 是根据本发明另一个实施例的基于关键词查询的位置查询系统的结构图；

图 14a 是根据本发明的基于自然语言和关键词的位置搜索系统的方框图；

图 14b 示出了基于自然语言和关键词的位置搜索系统中的处理装置的另一实施例；

图 15 是根据本发明的位置本体库生成方法的流程图；

图 16 是根据本发明的位置查询语言库生成方法的流程图。

## 具体实施方式

下面，将参考附图描述本发明的优选实施例。在附图中，相同的元件将由相同的参考符号或数字表示。此外，在本发明的下列描述中，将省略对已知功能和配置的具体描述，以避免使本发明的主题不清楚。

图 1 示出了根据本发明的基于自然语言查询的位置查询系统的方框图。参考图 1，位置查询系统包括用户接口 1，存储单元 2，位置数据库 3，GIS 接口 32 和自然语言查询处理装置 4。

位置数据库 3 包含位置服务中所有地理实体的详细数据，其中存放位置服务中的空间信息及普通信息。空间信息包括地图中所有实体的位置标签。一个点由经度和纬度描述。一条路或者一个地区由一序列的点来描述，其中每个点由经度和纬度描述。普通信息包括所有实体的静态信息(比如，地址，电话号码，产品/服务信息等等)。位置数据库 3 能够从电子地图，黄页以及一些重要的知识源中产生。

GIS 接口 32 用来计算位置数据库中的空间信息。GIS 接口 32 由 GIS 函数实现。目前，流行的 GIS (地理信息系统)平台包括 Supermap, Mapinfo, ArcInfo 等。基于空间信息的 GIS 函数用来计算和获取空间信息 (比如，空间位置，邻近信息)。一个 GIS 函数定义成 FuncName( $p_1, p_2, \dots, p_m$ )，其中 FuncName 是 GIS 函数的名字， $p_1, p_2, \dots, p_m$  为参数。一些 GIS 函数的例子如下：

1. GISLocation(X) 用作 GIS 位置函数，用来查找空间位置 X。比如，GISLocation(海龙大厦)的值为“中关村大街西南 137 米处，海淀中路东北 580 米处”。

2. GIS Near(X, Y, Z) 用作 GIS 邻近函数，用来查找目标位置附近信息。比如，“GISNear(创新大厦, 银行, 500)是查找创新大厦附近 500 米内有哪些银行。

3. GISPath(X, Y, Z, P) 用作 GIS 路径规划函数，基于实时或历史交通信息和交通规则，用来查找从 X 点到 Y 点最近、最快的路线。X 点是起始点，Y 点是终止点，Z 点是一序列的途经点，P 是交通方式“开车”或者“公交车”。

用户接口 1 包括查询接收器 11 和答案发送器 12，用户通过查询接收

器 11 向位置查询系统发送来自移动终端 5 的自然语言查询请求，并通过答案发送器 12 接收查询结果，移动终端 5 可以通过 SMS、MMS、WAP，语音等查询方式来查询位置信息。用户也可以通过 WEB 方式查询位置信息。此外，本发明并不限于移动终端 5，也可以是进行查询的其它的用户终端。存储单元 2 存储有位置本体库 21 和位置查询语言库 22。位置本体库 21 包括用于处理位置查询的领域知识，如图 5a 所示。位置查询语言库 22 包括用于处理位置查询的语言模型，图 6 示出了位置查询语言库的示例结构。

随后将对本发明的位置本体库和位置查询语言库进行详细描述。自然语言查询处理装置 4 包括自然语言查询分析装置 401，DB 搜索单元 46，答案融合与生成单元 47 和访问单元（未示出）。该访问单元设置在自然语言查询分析装置 401 和存储单元 2 之间，用于向自然语言查询分析装置 401 提供对位置本体库 21 和位置查询语言库 22 的访问。自然语言查询分析装置 401 用于利用访问单元访问位置本体库 21 和位置查询语言库 22 来处理用户的自然语言查询请求并返回查询动作，其包括分词单元 41，模糊处理单元 42，间接处理单元 43 和语言匹配单元 44。自然语言查询处理装置 4 对接收到用户的自然语言查询进行处理并向用户接口 1 发送查询结果。

在通过用户接口 1 接收到用户的自然语言查询请求之后，自然语言查询分析装置 401 中的分词单元 41 通过访问单元查找位置本体库的类别表（用于描述位置本体库中所有的地理类别）和实体表（用于描述位置本体库中所有的地理实体）以及位置查询语言库的常量表对用户的自然语言查询请求进行分词，从而获得查询请求的文法特征。模糊处理单元 42 基于句子分词获得的文法特征，对查询请求中出现的包括冗余描述和不完整描述的模糊描述进行模糊处理。

模糊处理采用的方法包括（1）冗余词的识别和处理（2）补充有用的字符和词。（3）基于用户的查询历史，利用上下文相关技术，等等。间接处理单元 43 在查询请求中包含间接描述时，通过访问单元从位置本体库的类别表中的同义词中查找与间接描述对应的类别名。语言匹配单元 44 用于将用户的查询请求和查询语言库中的查询语言进行匹配，从而得

到查询动作。

之后，DB 搜索单元 46 根据生成的查询动作，或直接搜索位置数据库 3，或利用 GIS 函数 32 间接地搜索位置数据库 3，从而找到相应的位置信息。答案融合与生成单元 47 将搜索到的位置信息进行融合并利用例如图 8 所示的答案模板生成文本查询答案并通过用户接口 1 将其发送到移动终端 5。图 7 示出了根据本发明的答案融合和生成单元 47 融合并生成答案的示例。

虽然图 1 示出位置本体库 21 和位置查询语言库 22 位于位置查询系统的内部，但是对于本领域的普通技术人员来说显而易见的是位置本体库 21 和位置查询语言库 22 可以位于该位置查询系统的外部。从而该位置查询系统通过访问单元访问外部的位置本体库 21 和位置查询语言库 22 来对自然语言查询进行分析处理。在图 1 所示的示例中，自然语言查询分析装置 401 可以同时执行模糊处理和间接处理，但是很明显的，自然语言查询分析装置可以只包括间接处理单元和模糊处理单元之一，从而仅执行模糊处理和间接处理中的一个。

由于基于本发明的新的位置本体库 21 和位置查询语言库 22，可以实现对模糊的自然语言查询请求或间接的自然语言查询请求进行处理，所以下面将分别参考图 5a 至图 5f 和图 6 对本发明的位置本体库和位置查询语言库进行描述，之后再结合位置本体库和位置查询语言库对本发明的位置查询系统执行的位置搜索过程进行详细描述。

图 5a 示出了根据本发明的位置本体库的示意性结构。如图 5a 所示，位置本体库由多个领域本体和一个映射本体组成。每个领域对应于一个领域本体，用来记录该领域的语义知识，例如地图领域对应于地图领域本体，黄页领域对应于黄页领域本体。映射本体用来记录不同领域本体中的概念之间的各种关联。

领域本体包含四部分：概念部分、属性部分、关系部分和公理部分。

### (1) 概念部分

概念部分用于描述该领域中所有的地理实体及地理类别，其存储在类别表和实体表中。图 5b 示出了类别表和实体表的一个例子。类别提供了位置信息的一个层次分类结构。如图 5b 所示，类别表中的每个类别表示

为(名称，类型，父类，同义词)。每个类别的类型按照从粗到细划分为：基础类、扩展类和连锁店。所述三种类型中的基础类是位置信息中最大的类别（如饭店、学校、大厦等）；扩展类是基础类的延伸或细化（如快餐店属于饭店）；连锁店是最细的类别（如肯德基属于快餐店）。类别表中的父类用来描述类别间的隶属关系，例如肯德基的父类是快餐店。另外由于大量的词具有相同的含义，所以每个类别还总结有一些同义词，例如，对于饭店这一类别来说，其同义词可以是餐厅，餐馆，酒楼，吃饭，好吃等等。此外，每个类别的同义词还可以例如是英文同义词。

如图 5b 的实体表所示，实体是一个具体的地点，每个实体表示为(名称，父类)，其中所述“父类”指示了实体和类别之间的层次关系。每个实体都隶属于一个类别，例如肯德基中关村店(实体)属于肯德基(类别)。所有的实体被定义在实体表中。

### (2) 属性部分

属性定义了概念的性质，用于描述地理实体的属性。例如，地址、电话等。每个属性至少有一个项‘类型’来表明属性的取值类型，例如电话的类型为数字。

### (3) 关系部分

关系描述了概念间的各种关联，即，定义了概念之间的语义关系和空间关系。如  $\text{isa}(x, y)$  用来描述类别和实体之间以及类别和类别之间的隶属关系，即  $x$  隶属于  $y$ ，例如，“ $\text{isa}(\text{肯德基}, \text{快餐店})$ ” 表示“肯德基”隶属于“快餐店”；再如  $\text{geo-part-of}(x, y)$  用来描述  $x$  在地理位置上是  $y$  的一部分，举例 NEC 中国研究院( $x$ )是创新大厦( $y$ )的地理部分（因为 NEC 中国研究院位于创新大厦内）。由于每个属性或者关系定义了概念的某个方面，从而多个属性和关系可以描述一个概念的整体。

### (4) 公理部分。

公理部分是基于概念及关系的规则，可进行进一步的推理。如  $\text{geo-part-of}(x, y) \& \text{south-of}(y, z) \rightarrow \text{south-of}(x, z)$ ，例如 NEC 中国研究院是创新大厦的地理部分，并且创新大厦在清华大学的南边，那么能推理出 NEC 中国研究院也在清华大学的南边。公理部分中的规则一般是有限的，可以根据需要进行扩充。公理通常由人工组织并确定。

映射本体只包含关系部分，具体包括同义映射关系、语言映射关系和空间映射关系。这些关系描述了不同领域本体中的概念之间的关系。

同义映射关系在同义词或缩略语之间进行映射，例如“synonymous(银行，储蓄所)”表示二者是同义词，其中银行和储蓄所分别是地图本体和黄页本体中的类别。语言映射关系在同一概念的不同语言翻译之间进行映射，例如“Chinese-English(路，Road)”表示后者是前者的英文翻译，其中路和Road分别是地图本体和自驾本体中的类别。空间映射关系在空间相关的概念之间进行映射，例如“near(银谷大厦，保福寺站)”表示二者距离很近，其中银谷大厦和保福寺站分别是地图本体和公交本体中的实体。

图 5c 示出了根据本发明的位置本体库的领域本体的概念部分的一个例子。例如，对于类别“路”，其类型是基础类，隶属于类别“路”的实体为“二环路”。以及，对于类别“大学”，其类型是扩展类，隶属于类别“大学”的实体为“清华大学”，而对于类别“家乐福”，其类型是“连锁店”，而隶属于类别“家乐福”的实体为“家乐福中关村店”。

图 5d 示出了根据本发明的位置本体库中的领域本体的属性和关系的一个示例。例如，对于“起点”，其类型为路的属性，属性值举例“学知桥”。对于“电话”，其类型为“属性”，属性值举例为 010-62705962，等等。以及对于  $isa(x, y)$ ，其类型为关系，关系值举例“ $isa(\text{中国银行}, \text{银行})$ ”。

图 5e 示出了根据本发明的位置本体库中的领域本体的公理部分的示例。根据图 5e 中所示的公理部分，就可以进行语义关系和空间关系等的推理。

图 5f 示出了位置本体库的部分位置概念空间。该位置概念空间是根据概念部分中的类别表和实体表自动生成的。如图 5f 所示，位置概念空间由一组概念（包括类别和实体）和概念间的关系组成，所述箭头指示了起点（概念）和终点（概念）之间的关系。

位置查询语言库由多个领域查询语言和一个公用查询语言组成。每个领域对应于一个领域查询语言，用来记录该领域的语言模型，如地图领域对应于地图领域查询语言，自驾领域对应于自驾领域查询语言。公用查询语言总结了各领域查询语言的共同查询特征，它可以被领域查询语言进行参数继承。因此位置查询语言库具有层次结构。

图 6 示出了位置查询语言库的一个示例。领域查询语言和公用查询语言的表示方法是一样的。查询语言包括两部分：文法部分和动作部分。

(1) 文法部分描述了位置服务中用户所有可能的查询方式，其提供了一个语法定义系统，文法部分中的文法描述类似于上下文无关文法，它记录了用于位置查询的所有文法。文法部分包括一个常量表，该常量表包含文法部分中的常数定义，包括各类名词、动词、疑问词等。

在文法中定义了一些特殊的符号。

- “|”指的是“或”逻辑运算。
- “ $<X>$ ”指 X 是文法名称，并能够在文法中找到其定义。
- “ $<!X>$ ”指 X 是一常量，它可以被相应的词代替。在常量表中能找到 X 的定义，X 由类型，词性和包含的词集组成。
- “ $<?X>$ ”指 X 是一概念，它能被在位置本体库中任何的类别或实体代替。
- $<?X(cons_1|...|cons_m)>$ 指 X 是一概念， $cons_i(i=1\dots m)$ 是一约束条件，并且 X 只能被位置本体库中满足任意一个约束条件  $cons_i$  的概念代替。例如， $<?C1(geo-entity)>$ 只能被位置本体库中的实体代替。
- “[ ]”指的是在“[” 和 “]”中间的内容在当前文法中可有可无。
- “{ $<?X>$ }”是查询变量的集合。它可以匹配多个有并列关系的概念。
- “ $<\#X>$ ”指 X 是一个参数，可以由其他文法通过该参数的赋值来继承当前文法。
- “ $<X1>=<X2(Y=Z)>$ ”指 X1 的文法通过将 X2 中的参数 Y 赋值为 Z，来继承 X2 的文法。

(2) 动作部分描述了各查询文法对应的查询动作，其定义了一组查询处理规则，每个规则有一个条件和相应的查询动作，一般用于指明“当用户查询符合某文法时，生成什么样的查询动作”。动作部分中的查询动作是系统对用户查询的理解结果。

通常由人工定义每种文法的查询动作。例如，“isSyntax(x)”是最通用的条件，指的是用户查询是否与文法 x 匹配。

位置查询语言有四个重要的特征：

1. 使用较少的文法可以覆盖较大量查询。
2. 为自然语言查询分析和关键字查询分析提供语言模型。
3. 可以表达复合句。“{<?X>}”使得文法能够匹配一个查询句子中的复合概念。
4. 可以通过比较用户查询的内容和位置查询语言库中的查询文法，分析用户查询中的语义错误。

下面将结合位置本体库21和位置查询语言库22对本发明的位置查询系统执行的位置搜索过程进行描述。图2示出了基于自然语言查询的位置搜索方法的流程图。

如图2所示，在S201，在通过用户接口1接收到用户的自然语言查询请求之后，自然语言查询分析装置401中的分词单元41用于对用户的自然语言查询请求进行分词，具体地，通过访问单元查找位置本体库的类别表和实体表从而识别出自然语言查询请求中的概念并判断其类型，以及通过访问单元查找位置查询语言库的常量表识别出自然语言查询请求中的非概念并判断词性和其类型。

模糊处理单元42在S202基于句子分词获得的文法特征，对查询请求中包括冗余描述和不完整描述的模糊描述进行模糊处理，模糊处理采用的方法包括（1）冗余词的识别和处理，即，基于语法特征删除多余的词（包括请求词，辅助词，和无意义的副词）。（2）补充有用的字符和词。对于不完整的实体，基于部分匹配技术提供一种方法。本发明提出了一种部分匹配的方法来找到全名。首先，从用户查询的分词结果中找到未识别词。之后，对未识别词基于一个常用词典进行更细粒度的分词，然后，通过访问单元从位置本体库中找出包含上述词的所有实体。最后，考虑到移动终端例如手机的屏幕较小，如果存在多个选项实体时可以仅仅选择具有最短长度的实体。比如在用户查询“创新”时，由于“创新”是一个不完整的未登录词，所以用“创新大厦”来代替。（3）基于用户的查询历史，使用上下文相关技术。首先，检查当前的查询是否完整。之后，若发现查询不完整，则从用户的历史查询中找到最近的记录，并添加丢失的词组。

在S203，间接处理单元43在查询请求中包含间接描述时，通过访问

单元从位置本体库的类别表中的同义词中查找与间接描述对应的类别名。

在 S204，语言匹配单元 44 用于将用户的查询请求和查询语言库中的文法进行匹配，从而得到查询动作。查询语言匹配包括：从位置查询语言中获得与用户查询完全匹配的文法（不需要严格按照词的顺序）。查询语言的匹配过程可以是自顶向下：先匹配公用查询语言的文法，若有匹配的公用查询语言，则根据其中的参数值，再匹配可继承该语言的领域查询语言的文法；若没有可匹配的公用查询语言，则直接匹配各领域查询语言的文法。查询语言的匹配过程也可以是自底向上：先匹配各领域查询语言再匹配公用查询语言。应该注意到一组并列的概念可以与文法中的“{<?X>}”匹配。根据匹配的文法，为用户的查询请求创建动作。此外，在文法的概念约束判断中，有时需基于位置本体库来推理。例如某文法描述饭店的招牌菜，若用户查询“肯德基中关村店有啥招牌菜”，则需经过推理才能得知肯德基中关村店是饭店，满足该文法的概念约束。推理中需利用：

- ◆ 关系：isa(肯德基中关村店, 肯德基)
- ◆ 关系：isa(肯德基, 快餐店)
- ◆ 关系：isa(快餐店, 饭店)
- ◆ 公理：isa(x,y) & isa(y,z) → isa(x,z)

在匹配过程中，如果找到了匹配文法，但不满足匹配文法中的概念约束，则可以查找位置本体库中的映射本体，对查询中的相关描述进行转换，使其满足相应的概念约束。例如用户查询“从银谷大厦到中关村怎么坐公交”，其匹配文法为“<?C1(公交车站)>; <!到达介词>; <?C2(公交车站)>; <!怎么疑问词>; [<!乘坐动词>]; <!公交名词>”，但“银谷大厦”属于类别“大厦”，不属于“公交车站”，则需利用映射本体中的空间映射关系“near(银谷大厦, 保福寺站)”，将“银谷大厦”转换为“保福寺站”。

在 S205，DB 搜索单元 46 根据生成的查询动作，直接搜索位置数据库 3 或利用 GIS 函数间接地搜索位置数据库 3，从而找到相应的位置信息。如果用户查询的是一般的静态信息（比如，地址，电话号码，和企业的产品/服务信息，等等），可以直接地搜索位置数据库 3。如果用户查询空间信息（比如，位置，邻近或路线信息，等等），则根据查询动作，使用 GIS 函数搜索位置数据库 3。对于每个查询动作，规定了对应的查询方法。例

如：

- 1) QueryLocation(X):如果 X 的地址值不是空值，获得所述值，此外还调用 GIS 函数“GISLocation(X)”给用户提供更多的位置信息。QueryLocation(X) 的值包括 A1=GetValue(X, address) 和 A2=GISLocation(X)，其中 GetValue(X, A) 表示获得 X 的属性 A 的值。比如，QueryLocation(海龙电子大厦) 包括 A1=“北京市海淀区中关村大街 1 号”和 A2=“中关村大街西南 137 米处, 海淀中路东北 580 米处”；
- 2) QueryNear(X, Y): 在 X 附近可能存在多个可选实体，因此仅提供 GISNear(X, Y, 500) 函数的值；
- 3) QueryNearest(X, Y): 为了提供给用户更多的位置信息，除提供离 X 最近的实体的名称外，还提供其地址信息。QueryNearest(X, Y) 的值包括 A1=GISNearest(X) 和 A2=GetValue(A1, address)。比如，QueryNearest(创新大厦, 银行) 的值包括了 A1=“中国银行”以及 A2=“海淀区中关村东路，清华科技园 1 号”；
- 4) QueryPath(X, Y): 用户想到达 Y，此时他也需要 Y 的空间信息。因此 QueryPath(X, Y) 的值包括了 A1=GISPath(X, Y, NULL, driving) 和 A2=GISLocation(Y)。

执行数据库查询之后，需要融合搜索的结果并产生最终的位置查询答案。在 S206，答案融合与生成单元 47 将搜索到的位置查询答案进行融合，融合包括：融合多个搜索动作。一个查询动作可能包括多个搜索动作，从而需要针对每个查询动作融合搜索动作。例如 QueryNearest(X, Y) 包括两个搜索动作“GISNearest(X)”和“GetValue(A1, address)”。在答案融合与生成单元 47 融合搜索动作之后，在 S207 利用针对每个查询动作定制的多语言答案模板生成最终的位置查询答案并经用户接口 1 发送到移动终端 5 进行显示。图 7 示出了融合位置查询答案的示例。图 8 示出了答案模板的一个示例。

图 9a 是根据本发明的自然语言查询处理装置对自然语言查询进行处理的一个示例。下面，以用户输入的自然语言查询请求“请告诉我清华附近有什么好吃的”为例，对本发明的位置查询系统进行说明。在自然语言

查询分析装置 401 通过用户接口 1 接收到所述查询请求时，由分词单元 41 通过访问单元访问位置本体库 21 和位置查询语言库 22 对该查询请求进行分词：请告诉我(请求词) 清华(未识别词) 附近(附近副词) 有(拥有动词) 什么(什么疑问词) 好吃(类别) 的(助词)。之后，模糊处理单元 42 根据分词后的查询请求进行删词和补词处理，将请求词“请告诉我”和助词“的”删除，并通过访问单元查找位置本体库中的实体表将“清华”一词补充为“清华大学”，从而所述的查询请求变为“清华大学(实体) 附近(附近副词) 有(拥有动词) 什么(什么疑问词) 好吃(类别)”。间接处理单元 43 对上述结果进行间接分析，通过访问单元来搜索位置本体库 21 中的类别表，从类别表中的同义词一栏中找到“好吃”一词的同义词是“饭店”，从而向语言匹配单元 44 输出处理后的查询请求：“清华大学(实体) 附近(附近副词) 有(拥有动词) 什么(什么疑问词) 饭店(类别)”。语言匹配单元 44 将用户的查询语言与位置查询语言库 22 进行匹配，找到匹配文法“<NearNeighborQuery> = <CommonQuery2(<#副词>=[<!附近副词>])>”，其中“<CommonQuery2>=<?C1(geo-entity)>{<#副词>}{<?C2(geo-category|geo-entity)>}[<!位于动词|!拥有动词>][<!地点疑问词|!什么疑问词>]”，并生成查询动作 QueryNear(清华大学，饭店)。之后，DB 搜索单元 46 接收所述的查询动作并基于所述的查询动作直接或间接地搜索位置数据库中与查询动作相关的信息，并生成搜索结果例如“六道口郭林家常菜饭店，五道口必胜客餐厅，五道口肯德基快餐店”。答案融合与生成单元 47 将搜索结果进行融合，并生成答案：“五道口必胜客餐厅和肯德基快餐店，六道口郭林家常菜”。所生成的答案经过用户接口 1 发送到用户的移动终端上以便显示。图 11a 示出了基于自然语言的位置查询系统执行的查询的一个示例。

图 3 示出了根据本发明的基于关键词查询的位置查询系统的示意图。参考图 3，位置查询系统包括用户接口 1，存储单元 2，位置信息库 3，GIS 接口 32 和关键词查询处理装置 6。位置数据库 3 用于存放位置服务中的空间信息及普通信息，GIS 接口 32 用来计算位置数据库中的空间信息。

用户接口 1 包括查询接收器 11 和答案发送器 12。存储单元 2 存储有位置本体库 21 和位置查询语言库 22。位置本体库 21 存储了包括用于处

理位置查询的领域知识。位置查询语言库 22 包括用于定义位置服务的查询语言模型。关键词查询处理装置 6 包括关键词查询分析装置 402, DB 搜索部分 46, 答案融合与生成单元 47 和访问单元（未示出）。该访问单元设置在关键词查询分析装置 402 和存储单元 2 之间，用于向关键词查询分析装置 402 提供对位置本体库 21 和位置查询语言库 22 的访问。自然语言查询处理装置 402 通过访问单元查找位置本体库 21 和位置查询语言库 22 处理用户的关键词查询，并返回答案，包括分词单元 41, 模糊处理单元 42, 间接处理单元 43, 文法部分匹配单元 44'，答案决策单元 45。

关键词查询分析装置 402 中的分词单元 41 用于将用户的关键词查询请求进行分词，具体地，通过访问单元检索位置本体库的类别表和实体表从而识别出关键词查询请求中的概念并判断其类型，以及通过访问单元检索位置查询语言库中的常量表识别出关键词查询请求中的非概念并判断词性和其类型。

模糊处理单元 42 对接收的关键词中包括冗余描述和不完整描述的模糊描述进行模糊处理采用的方法包括（1）冗余词的识别和处理，即，基于语法特征删除多余的词（包括请求词，辅助词，和无意义的副词）。（2）补充有用的字符和词。当用户输入了关键词时，可能会丢失若干个汉字。本发明提出了一种部分匹配的方法来找到全名。首先，如果分词后得到了未识别词，则在此将关键词基于一个常用词典进行更细粒度的分词，之后，通过访问单元从位置本体库中找出包含上述词的所有实体。例如当用户查询“海龙大厦 存钱啊”，分词单元 41 得到的结果是“海龙大厦（未识别词） 存钱（类别） 啊（语气词）”，因为“海龙大厦”是未识别词，所以对其再度分词得到“海龙 大厦”，然后通过访问单元检索位置本体库，发现“海龙电子大厦”包含词“海龙”和“大厦”，至此部分匹配成功。最后，考虑到移动终端例如手机的屏幕较小，如果存在多个选项实体时可以仅仅选择具有最短长度的实体。

间接处理单元 43 在查询请求中包含间接描述时，通过访问单元从位置本体库的类别表中的同义词中查找与查询请求对应的类别名。

文法部分匹配单元 44' 用于通过访问单元访问位置查询语言库，得到与用户查询部分匹配（不是完全匹配）的文法集合。其包括：从位置查

询语言库中获得包含用户查询中所有关键词的文法。需要注意一组并列的概念可以匹配文法中的“{<?X>}”。

答案决策单元 45 用于根据预定的决策规则，选择最优匹配，生成查询动作或与用户进行交互。当用户利用关键词进行查询时，多个文法可能被部分匹配，以及多个文法可能存在相同动作，因此需要删除那些冗余的文法并确定最终的答案。如果文法和查询完全匹配，选择该文法并且产生相应的动作。如果文法和查询不能完全匹配，但存在若干个包含查询中的所有关键词的文法，则需要通过匹配的程度确定最优解（最优答案）。如果一个文法与其他文法相比较，具有最高的匹配程度，则选择该文法并且创建相应的动作，否则产生所有可能的查询，之后与用户交互来让用户自己选择。

同自然语言查询系统的语言匹配单元 44 类似，答案决策单元 45 在匹配过程中，如果找到了匹配文法，但不满足匹配文法中的概念约束，则可以查找位置本体库中的映射本体，对查询中的相关描述进行转换，使其满足相应的概念约束。

虽然图 3 示出位置本体库 21 和位置查询语言库 22 位于位置查询系统的内部，但是对于本领域的普通技术人员来说显而易见的是位置本体库 21 和位置查询语言库 22 可以位于该位置查询系统的外部。从而位置查询系统可以通过访问单元访问外部的位置本体库 21 和位置查询语言库 22 来对关键词查询进行分析处理。此外，图 3 中的实例示出关键词查询分析装置 402 可以同时执行模糊处理和间接处理，但是该关键词查询分析装置也可以仅包括间接处理单元和模糊处理单元之一，从而执行模糊处理和间接处理中的一个。

图 4 示出了根据本发明的基于关键词查询的位置搜索方法的流程图。在 S401，基于关键词查询的位置查询系统接收用户的移动终端 5 发送的关键词查询请求。

在 S402，分词单元 41 将接收到的查询请求分词。分词单元 41 通过访问单元检索位置本体库 21 的类别表和实体表识别出概念并判断其类型；以及通过检索位置查询语言库 22 中的常量表识别出非概念并判断其词性和类型。由于利用了位置本体库 21 和位置查询语言库 22，所以可以

分析出更多的文法特征，从而更准确地进行搜索。

在查询请求输入的概念模糊时，在 S403，模糊处理单元 43 利用分词后的句子获得的文法特征，通过访问单元利用位置本体库 21 对用户的查询请求中包括冗余描述和不完整描述的模糊描述进行模糊处理，包括冗余词的识别及处理，不完整词的判断及补全和上下文相关技术等。

在 S404，间接处理单元 43 通过访问单元从位置本体库 21 的类别表中查找与关键词查询中的间接描述对应的类别名。

在 S405，文法部分匹配单元 44' 通过访问单元将查询请求与位置查询语言进行部分匹配，得到匹配文法集合。

在 S406，答案决策单元 45 根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成查询动作；或者生成所有可能的查询以便通过与用户交互来让用户进行选择，并根据用户的选择生成对应查询动作。

之后，在 S407，DB 搜索单元 46 直接搜索位置数据库 3 或利用 GIS 接口搜索位置数据库 3，以查找与用户的查询请求对应的位置信息。

在 S408，答案融合与生成单元 47 将搜索到的位置信息进行融合并生成答案。

在 S409，答案融合和生成单元 47 通过用户接口 1 将答案发送到移动终端 5 进行显示。

图 10 是根据本发明的关键词查询处理装置对关键词查询进行处理的一个示例。例如，当用户的查询请求为“海龙大厦 存钱啊”时，由分词单元 41 通过访问单元利用位置本体库 21 和位置查询语言 22 将该查询请求进行分词，得到：海龙大厦（未识别词）存钱（类别）啊（语气词）。之后，模糊处理单元 42 通过访问单元查找位置本体库 21 和查询语言库 22 删除“啊”并将“海龙大厦”一词补充为“海龙电子大厦”。间接处理单元 43 通过访问单元查找位置本体库 21 中的类别表将查询请求中的间接描述“存钱”转换为直接描述“银行”，得到：海龙电子大厦（实体）银行（类别）。之后文法部分匹配单元 44' 从位置查询语言中找出匹配文法，包括：

```
<CommonQuery2>=<?C1(geo-entity)><#副词>{<?C2(geo-category  
|geo-entity)>}[<!位于动词|!拥有动词>][<!地点疑问词|!什么疑问词  
>]
```

<NearNeighborQuery>=<CommonQuery2(<#副词>=[<!附近副词>])

<NearestNeighborQuery>=<CommonQuery2(<#副词>=<!最近副词>)

答案决策单元 45 选择<NearNeighborQuery>的文法，因为可以完全匹配，并生成相应的查询动作 QueryNear(海龙电子大厦，银行)。DB 搜索单元 46 搜索位置数据库 3 中的位置信息，经答案融合和生成单元 47 对搜索到的位置信息进行融合，从而生成最终的答案文本，并通过用户接口 1 发送到移动终端 5 以便显示。图 11b 示出了基于关键词的位置查询系统执行的查询的一个示例。

图 12 是根据本发明另一个实施例的基于自然语言查询的位置查询系统的方框图，其与图 1 的区别在于该位置查询系统还包括复合句处理单元 48 以及错误诊断单元 49。由于图 12 中其它的单元与图 1 所示的相同的附图标记所表示的单元功能相同，所以省略对其的描述。通常，用户输入的自然语言查询请求可能是复合句，复合句可以是包括多个组合句子的复合句，例如，“告诉我创新在哪里，存钱哪儿最近？”。复合句也可以是包括多个并列的概念的复合句，例如，“创新大厦和海龙大厦在那里？”。复合句处理单元 48 利用位置查询语言库 22 对用户输入的自然语言查询请求中包括的复合句进行分析处理。首先复合句处理单元 48 根据标点符号及位置查询语言将复合句拆分为多个单句，再由分词单元 41 至语言匹配单元 44 进行下一步处理。图 9b 示出了根据本发明的自然语言查询处理装置对包括复合句的自然语言查询进行处理的一个示例。

错误诊断单元 49 通过访问单元利用位置本体库和位置查询语言依据预定的规则识别并分析用户查询中的语义错误。语义错误包括 1) 类别错误和 2) 不完全错误。对于 1) 类别错误，错误诊断单元 49 检查用户的查询请求中的每个变量是否满足其约束。对于用户的查询请求，首先查找最相似的文法，之后将查询请求和文法进行匹配并得到变量和其约束。如果一个变量不满足其约束，则错误诊断单元 49 就判定该查询请求是类别错误。可以向用户提供错误信息和帮助信息来与用户交互。比如，用户查询请求是“银行在哪里”，并且最相似的文法是“{<?C(geo-entity)>}<!位于动词><!地点疑问词>”，但是“银行”是一个类别，并不满足它的约束“geo-entity”(地理实体)，因此，此查询有语义错误，因为银行不是一

个具体的地理实体。对于 2) 不完整错误，错误诊断单元 49 基于位置查询语言检查用户查询请求是否完整。对于用户的查询，首先查找最相似的文法，如果用户查询是该文法的子集，此查询就不完整。如果在上下文或者用户的查询历史中或者其他地方都找不到缺少的信息，此查询请求就存在不完整错误。可以向用户提供错误信息和帮助信息来与用户交互。例如，用户的查询请求是“怎么去创新大厦”，并且最相似的文法是“<?C1(geo-entity)>[<!怎么疑问词>]<!到达介词><?C2(geo-entity)>”，但是此处缺少“?C1”。如果无法得到用户当前位置，而且起始点也不在上下文中，则错误诊断单元 49 确定此查询有语义错误，其错误信息为缺少起始点，并将错误信息发送到答案融合与生成单元 47，答案融合与生成单元 47 将诊断出的错误发送到移动终端 5。由于该系统可以对复合句查询请求进行处理，所以答案融合与生成单元 47 针对与用户的复合句查询请求对应的多个查询动作，在为每个查询动作融合搜索动作之后，融合所有查询动作的查询结果。例如查询请求“创新大厦和海龙大厦最近的饭店？”包含两个查询动作“QueryNearest(创新大厦，饭店)”和“QueryNearest(海龙大厦，饭店)”。答案融合与生成单元 47 需要对上述两个查询动作的查询结果进行融合。

图 13 示出了根据本发明另一个实施例的基于关键词查询的位置查询系统的示意图。其与图 2 的区别在于图 13 所示的位置查询系统还包括错误诊断单元 49，用于根据预定的规则通过访问单元利用位置本体库和位置查询语言识别并分析用户的关键词查询中的语义错误。错误诊断单元 49 将错误结果发送到答案融合与生成单元 47，答案融合与生成单元 47 将诊断出的错误发送到移动终端 5，以便与用户进行交互。

图 14a 和 14b 示出了根据本发明的基于自然语言和关键词的位置查询系统的两个实施例。如图 14a 所示，所述的系统包括用户接口 1，存储单元 2，位置数据库 3，GIS 接口 32，以及处理装置包括判断装置 7，自然语言查询处理装置 4 和关键词查询处理装置 6。所述存储单元 2 包括位置本体库 21 和位置查询语言库 22。此外，由于已经结合图 1，图 12 和图 3，图 13 分别对自然语言查询处理装置 4 和关键词查询处理装置 6 进行了描述，这里将省略对其的描述。判断装置 7 用于根据用户问句特点，判断

从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求。如果是自然语言查询请求，则由自然语言查询处理装置 4 对该自然语言查询请求进行处理并生成位置查询的答案，如果是关键词查询请求，则由关键词查询处理装置 6 对该关键词查询请求进行处理并搜索到位置查询的答案。之后，将答案通过用户接口 1 发送到用户终端 5。

在判断装置 7 判断用户查询是基于自然语言还是关键词时需要根据用户问句特点对用户的查询请求进行分类。一般来说，自然语言查询和关键词查询的问句特点有：

(1) 关键词查询可以含有一些逻辑运算符，例如“ ”(空格), “和”, “或”, “+” 和 “;” 等。关键词查询可以由多个以运算符间隔的字符串组成，每个字符串包括一个或多个连续的词。

(2) 自然语言查询是一个连续的字符串，通常包括多个词。中间也可能有逻辑运算符，但是做别的意思用，例如“和”可作连词用。此外，自然语言查询通常包括疑问词（例如，哪里，什么，何时等）。

判断装置 7 所采用的第一种判断方法是基于逻辑运算符的，包括

- 1) 检测用户查询中是否存在任何一个逻辑运算符；
- 2) 如果运算符附近的词可以构成一个完整的词，则删除该运算符；

3) 此时，如果在用户查询中仍然不存在逻辑运算符，则判断该用户查询是自然语言查询，否则判断出该用户查询是关键词查询。

例如，用户查询为“创新 大厦在哪里”。首先，在“创新”和“大厦”之间存在空格，但是空格附近的词“创新”和“大厦”可以构成一个完整的词“创新大厦”。其次，该用户查询是一个疑问句。因此，判断装置判断出该用户查询是自然语言查询。

判断装置 7 所采用的第二种判断方法是检查用户查询的完整性。自然语言查询通常是给出明确需求的疑问句，而关键词查询通常并不完整。

判断装置 7 还可以采用的判断方法是同时进行自然语言查询和关键词查询的处理后选择最优结果。此外，判断装置 7 还可以采用其它的已知的判断方法判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求。

以及，由基于自然语言和关键词的位置查询系统执行的方法包括确定步骤，如图 2 所示的基于自然语言的位置查询步骤和如图 4 所示的基于关键词的位置查询步骤。

此外，虽然以中文查询为例，示出了位置搜索系统，对于本领域显而易见的是本发明也可以用于其它语言的查询，例如，英语和日语等等。

图14b示出基于自然语言和关键词的位置查询系统中的处理装置的另一个实施例。由于该系统具有与图14a所示的相同的组件，包括用户接口1，存储单元2，位置数据库3， GIS接口32和答案发送器12，图14b仅示出了处理装置142。处理装置142包括：分词单元41，通过查找位置本体库中的类别表、实体表以及位置查询语言库中的常量表，对查询请求进行分词；模糊处理单元42，用于通过查找位置本体库、位置查询语言库和用户查询历史，对分词后的查询请求中的模糊描述进行删词或补词的处理，其中所述的模糊描述包括冗余描述和不完整描述中的至少之一；间接处理单元43，通过查找位置本体库中的类别表，将查询请求中的间接描述转化为位置本体库中的对应类别名称；判断单元7，判断从用户终端接收的查询请求是自然语言查询请求还是关键词查询请求；查询语言匹配单元44，用于当查询请求是自然语言查询时，将处理后的查询请求与位置查询语言库进行匹配，并生成对应的查询动作；文法部分匹配单元44'，用于当查询请求是关键词查询时，将处理后的查询请求与位置查询语言库进行部分匹配，并获得匹配文法集合；答案决策单元45，用于当查询请求是关键词查询时，根据预定的决策规则，从匹配文法集合中选择最优匹配文法并生成对应的查询动作；数据库搜索单元46，用于根据查询动作从位置数据库中搜索相应的查询结果；以及答案融合与生成单元47，用于融合查询结果以便生成答案，并将答案发送到用户终端。

虽然图 14b 示出了处理装置 142 包括模糊处理单元 42 和间接处理单元 43，其可以仅包括模糊处理单元 42 或者间接处理单元 43。此外，处理装置还可以包括复合句处理单元 48 和/或错误诊断单元 49。

图 15 是根据本发明的位置本体库生成方法的流程图。位置本体库生成方法包括领域本体生成步骤 233，映射本体生成步骤 234 和组合步骤 235。

领域本体生成步骤 233 用于对每个领域生成领域本体，具体包括地理实体提取、地理类别提取、属性提取和关系提取。

首先，领域本体生成步骤 233 从各领域的信息源中提取地理实体，生成实体表。具体提取方法因领域而异，例如，生成地图领域本体时利用现有的 GIS 函数抽取电子地图上标示的所有地物名，生成黄页领域本体时利用现有的未识别词识别算法从 web 上的黄页信息中提取机构名、地名等词。

其次，领域本体生成步骤 233 提取地理类别。现有的电子地图中提供了一些粗分类，本发明在此基础上进行了扩充。

首先将地图中的地物分类作为基础类；其次利用现有的统计和聚类算法，对实体表中的所有实体名进行类型特征提取（通常是词频很高的后缀，如某某快餐店），聚类得到所有扩展类（如快餐店）；然后再对所有实体名进行连锁店特征提取（通常是词频较高的前缀，如肯德基某某店），聚类得到所有连锁店（如肯德基）；最后根据上面的聚类结果和同义词典，得到各类别的同义词集，并生成类别表。

然后领域本体生成步骤 233 提取属性。具体提取方法因领域而异，例如，生成地图领域本体时提取地图数据库中的地物描述字段（如经度、纬度），生成黄页领域本体时查找 Web 上的结构化描述并利用现有的信息提取算法找出可能的属性集。最后对每个属性，人工标明属性的类型。

领域本体生成步骤 233 再次提取关系。主要提取类别间的隶属关系、实体和类别间的隶属关系、实体间的空间关系。类别间的隶属关系主要依据现有的地物分类法标准，并人工修正和总结。实体与类别间的隶属关系主要依据地理类别提取中的实体聚类结果来获得。而实体间的空间关系则可利用 GIS 函数来计算得到。

最后领域本体生成步骤 233 将抽取到的地理实体、地理类别、属性、关系以及人工组织的预定公理部分组合，形成最终的领域本体。

映射本体生成步骤 234 用于根据各领域本体生成映射本体，具体包括同义映射关系提取、语言映射关系提取和空间映射关系提取步骤之一。

首先，映射本体生成步骤 234 根据一个同义词典和一个简称规则库生成同义映射关系。同义词典包含概念间的同义关系，简称规则库包括各短

语的缩略语，例如“abbreviate(北京大学，北大)”表示“北京大学的简称是北大”。根据简称规则可以判断两个概念是否同义，例如根据abbreviate(北京大学，北大)”和“abbreviate(附属小学，附小)”，可以得到“北京大学附属小学”和“北大附小”是同义词。

其次，映射本体生成步骤 234 根据一个多语言词典生成语言映射关系。

最后，映射本体生成步骤 234 根据电子地图的 GIS 函数生成空间映射关系。

组合步骤 235 将领域本体和映射本体组合在一起，生成位置本体库 21。

图 16 是根据本发明的位置查询语言库生成方法的流程图。参考图 16，位置查询语言库生成方法包括领域查询语言生成步骤 241，公用查询语言生成步骤 242 和组合步骤 243。

领域查询语言生成步骤 241 用于对每个领域生成领域查询语言，具体包括问句收集、语料建立，问句聚类和文法提取。问句收集用于收集该领域的用户真实问句集合。语料建立用于利用现有的分词算法对各问句进行分词和标注（标注分为概念、疑问词、名词、动词等），从而生成问句语料库。问句聚类用于利用现有的统计和聚类算法，计算问句间的相似度并按相似度聚类。文法提取用于根据聚类结果来总结文法，具体如下：

- 1) 对每个聚类分类，定义一个文法名
- 2) 根据下列方法，根据当前聚类分类的问句的相似度，提取查询文法

- ◆ 若有多个文法，则中间用“|”隔开
- ◆ 一个文法可包含一个或多个部分，每一个部分 X 由“<X>”来表示
- ◆ 一组同义词可以归纳为一个常量，表示为<!常量名>。将所有的常量存放在常量表中
- ◆ 一组并列的概念可以归纳为一个变量，表示为<?变量名>。若该变量有类型约束，则表示为<?变量名(约束类型集)>
- ◆ 若文法中的某个部分可有可无，则加“[]”来表示

- ◆ 若文法中的某个部分可以是一个并列的概念集，则加“{ }”来表示

3) 对每种文法，人工定义它的动作。例如 isSyntax(<LocationQuery>) → QueryLocation(?C)描述如果用户查询符合<LocationQuery>的文法，则产生查询动作 QueryLocation(?C)。

公用查询语言生成步骤 242 计算各领域查询语言的文法的相似度，抽取出共同的文法，写入公用查询语言中。

组合步骤 243 用于将领域查询语言和公用查询语言组合在一起，生成位置查询语言库 22。

尽管已经示出和描述了本发明，本领域的技术人员可以理解的是，在不偏离本发明的原理和精神的前提下，可以在本实施例中进行改变，本发明的范围由所附权利要求及其等价物限定。

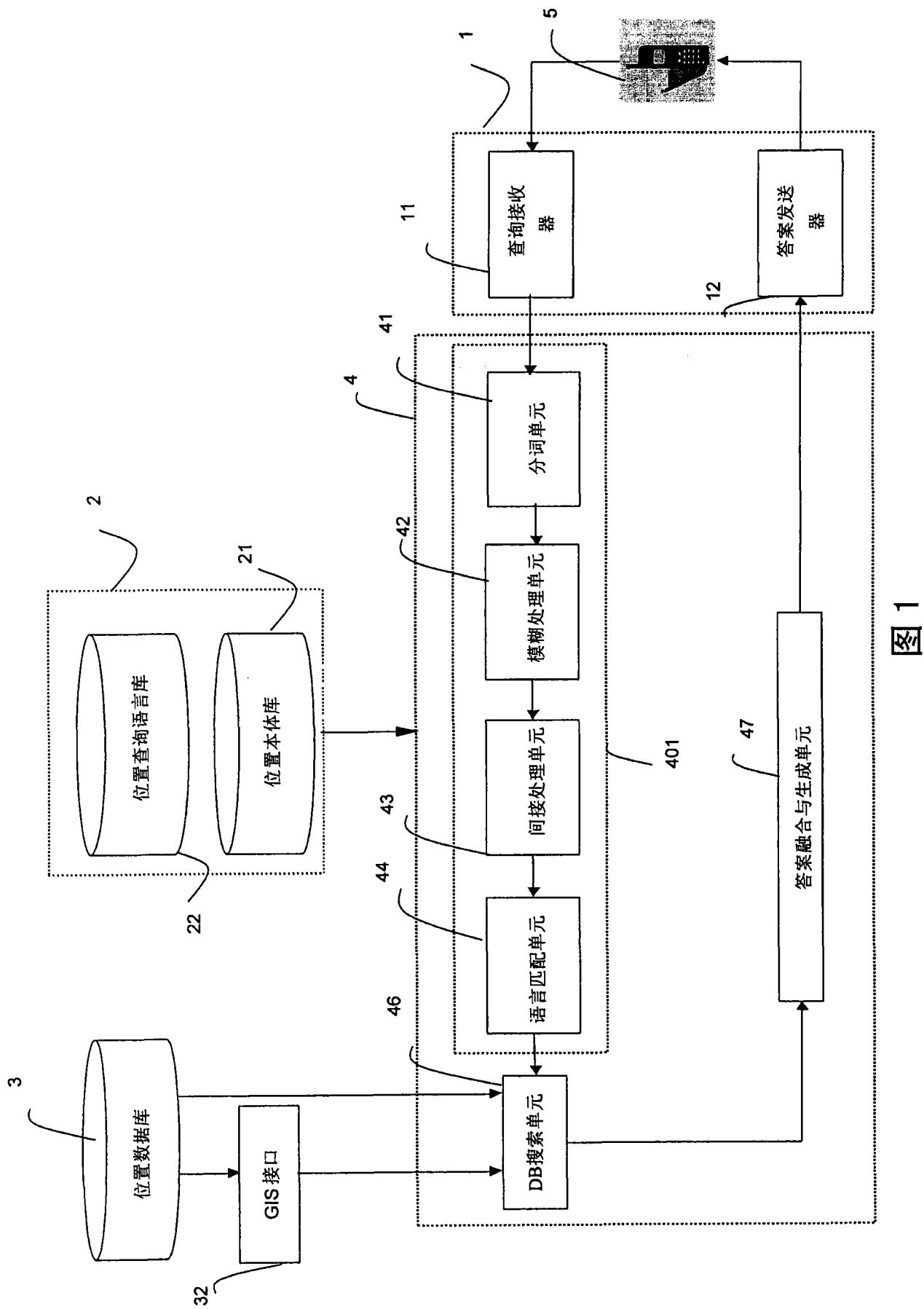


图 1

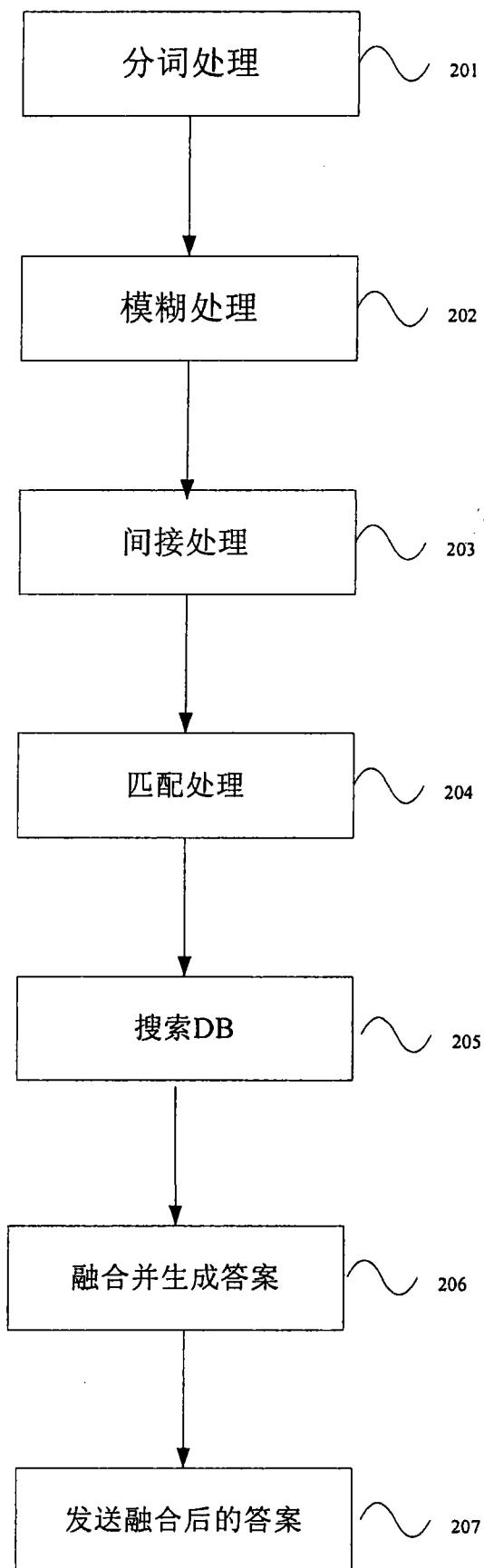


图 2

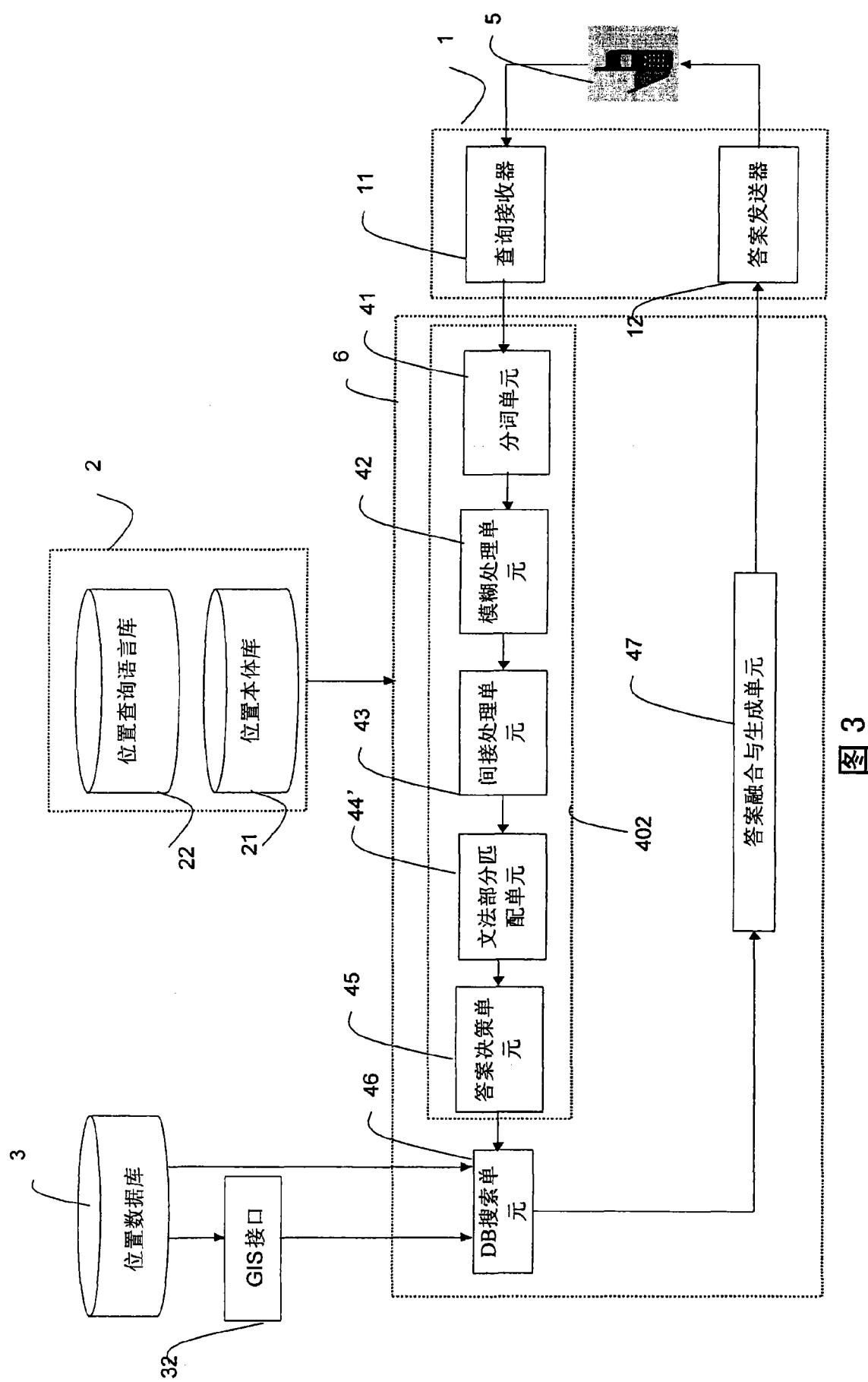


图 3

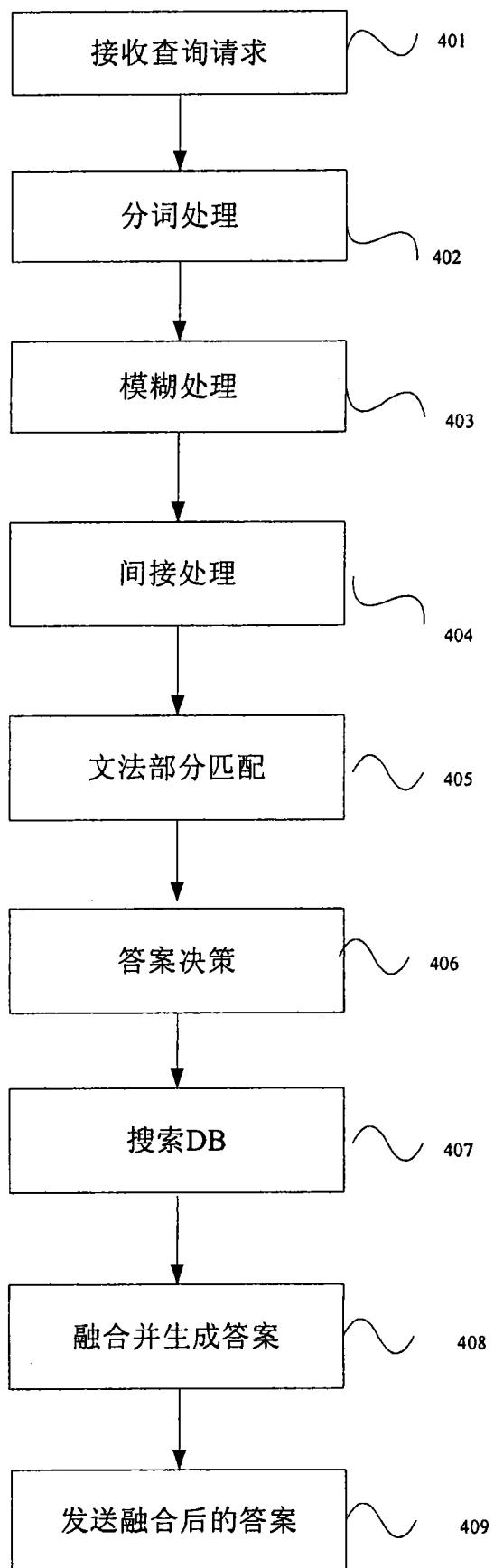


图 4

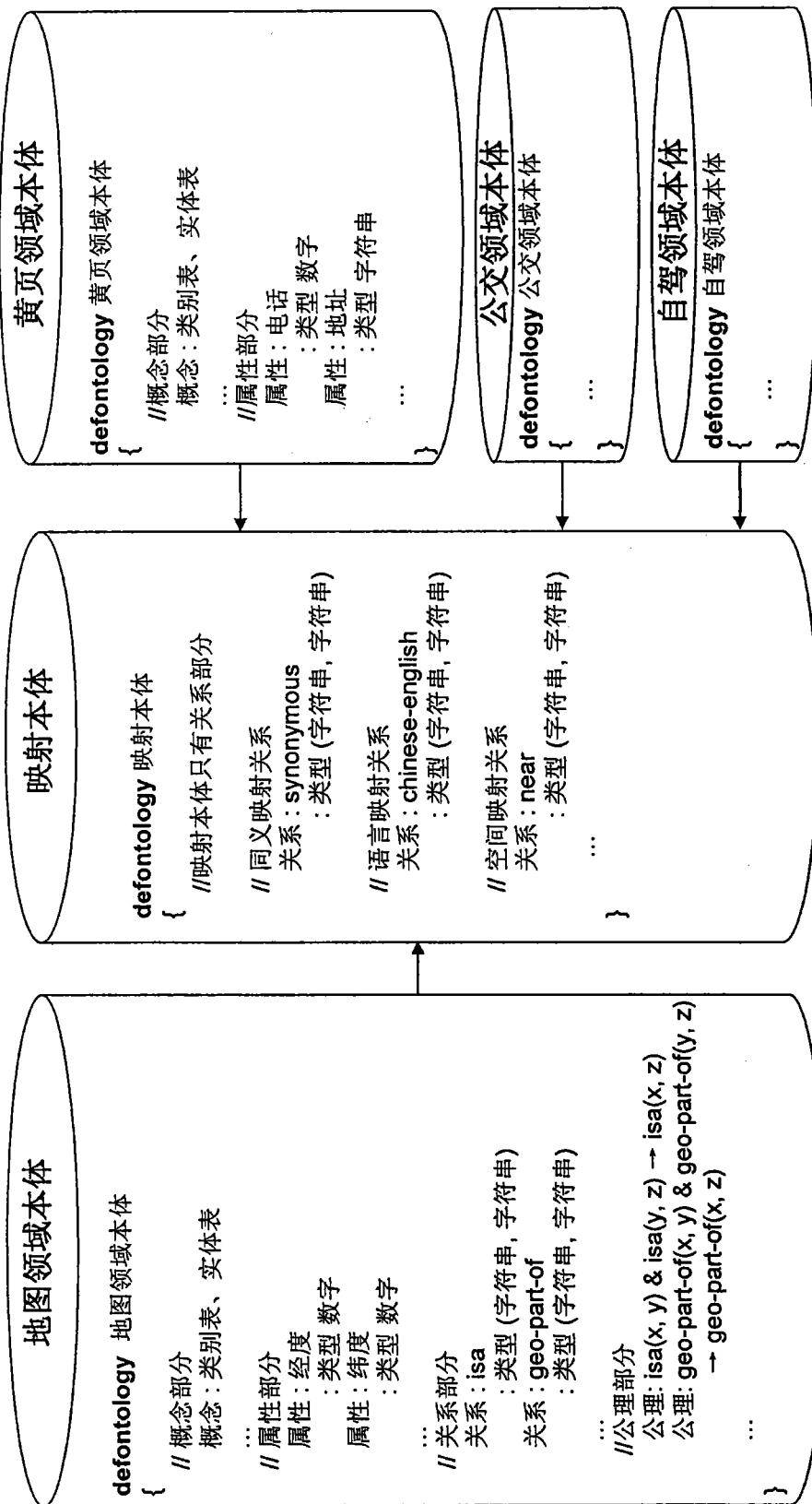


图 5a

| 名称  | 类型  | 父类  | 同义词(汉语)        | 同义词(英语)            |
|-----|-----|-----|----------------|--------------------|
| 大厦  | 基础  | 位置  | 大楼             | plaza building     |
| 学校  | 基础  | 位置  | 上学             | school             |
| 饭店  | 基础  | 位置  | 餐厅 餐馆 酒楼 吃饭 好吃 | restaurant eat     |
| 银行  | 扩展  | 金融  | 储蓄所 存钱 取钱      | bank save money    |
| 大学  | 扩展  | 学校  |                | university college |
| 快餐厅 | 扩展  | 饭店  | 快餐 速食          | snack fast food    |
| 肯德基 | 连锁店 | 快餐厅 |                | KFC                |
| ... | ... | ... | ...            | ...                |

类别表

| 名称      | 父类  |
|---------|-----|
| 清华大学    | 大学  |
| 创新大厦    | 大厦  |
| 肯德基中关村店 | 肯德基 |
| ...     | ... |

实体表

图 5b

| 类别(汉语)           | 类型  | 实体示例(汉语)                                |
|------------------|-----|---|
| Road(路)          | 基础  | Second Ring(二环路)                        |
| Region(区域)       | 基础  | Zhongguancun(中关村)                       |
| Bridge(桥)        | 基础  | Xi zhemen Bridge(西直门桥)                  |
| Plaza(大厦)        | 基础  | Innovation Plaza(创新大厦)                  |
| School(学校)       | 基础  | Zhongguancun High School(中关村中学)         |
| Restaurant(饭店)   | 基础  | Beijing Hotel(北京饭店)                     |
| Speedway(高速)     | 扩展  | Jing-Shi Speedway(京石高速)                 |
| University(大学)   | 扩展  | Tsinghua University(清华大学)               |
| Bank(银行)         | 扩展  | Bank of China(中国银行)                     |
| KFC(肯德基)         | 连锁店 | KFC Zhongguancun Store(肯德基中关村店)         |
| Macdonald's(麦当劳) | 连锁店 | Macdonald's Zhongguancun Store(麦当劳中关村店) |
| Carrefour(家乐福)   | 连锁店 | Carrefour Zhongguancun Store(家乐福中关村店)   |
| ...              |     | ...                                     |

图 5c

| 属性(汉语)               | 类型   | 示例值(汉语)  |
|----------------------|------|--|
| StartPoint (起点)      | 路的属性 | Xuezhi Bridge (学知桥)  |
| EndPoint (终点)        | 路的属性 | East of Chengfu Road (成府路东口)   |
| Address (地址)         | 属性   | 11F, Building A, Innovation Plaza, Tsinghua Science Park.No1, Zhongguancun East Road, Haidian District, Beijing (北京市海淀区中关村东路清华科技园1号创新大厦A座11层 ) |
| PhoneNumber (电话)     | 属性   | 010-62705962   |
| ZipCode(邮政编码)        | 属性   | 100080   |
| BusinessService (业务) | 属性   | Chinese food (中餐); Computer (计算机);   |
| Linkman (联系人)        | 属性   | Miss Wang (王小姐)  |
| geo-part-of(x, y)    | 关系   | geo-part-of(创新大厦, 清华科技园)   |
| isa(x, y)            | 关系   | isa(中国银行, 银行 )   |
| north-of(x, y)       | 关系   | north-of( 西直门桥, 复兴门桥 )   |
| between(x, y, z)     | 关系   | between( 西二环中路, 西二环北路, 西二环南路 )   |
| ...                  |      | ...  |

图 5d

| 公理<br>示例  |  |
|---|--|
| $\text{isa}(x, y) \& \text{isa}(y, z) \rightarrow \text{isa}(x, z)$                         | $\text{isa}(\text{KFC}, \text{快餐店}) \& \text{isa}(\text{快餐店}, \text{饭店}) \rightarrow \text{isa}(\text{KFC}, \text{饭店})$  |
| $\text{geo-part-of}(x, y) \& \text{geo-part-of}(y, z) \rightarrow \text{geo-part-of}(x, z)$ | $\text{geo-part-of}(\text{NEC 中国研究院}, \text{创新大厦}) \& \text{geo-part-of}(\text{NEC 中国研究院}, \text{清华科技园}) \rightarrow \text{geo-part-of}(\text{NEC 中国研究院}, \text{清华科技园})$ |
| $\text{north-of}(x, y) \rightarrow \text{south-of}(y, x)$                                   | $\text{north-of}(\text{西直门桥}, \text{复兴门桥}) \rightarrow \text{south-of}(\text{复兴门桥}, \text{西直门桥})$  |
| $\text{geo-part-of}(x, y) \& \text{south-of}(y, z) \rightarrow \text{south-of}(x, z)$       | $\text{geo-part-of}(\text{NEC 中国研究院}, \text{创新大厦}) \& \text{south-of}(\text{创新大厦}, \text{清华大学}) \rightarrow \text{south-of}(\text{NEC 中国研究院}, \text{清华大学})$              |
| ...   | ...  |

图 5e

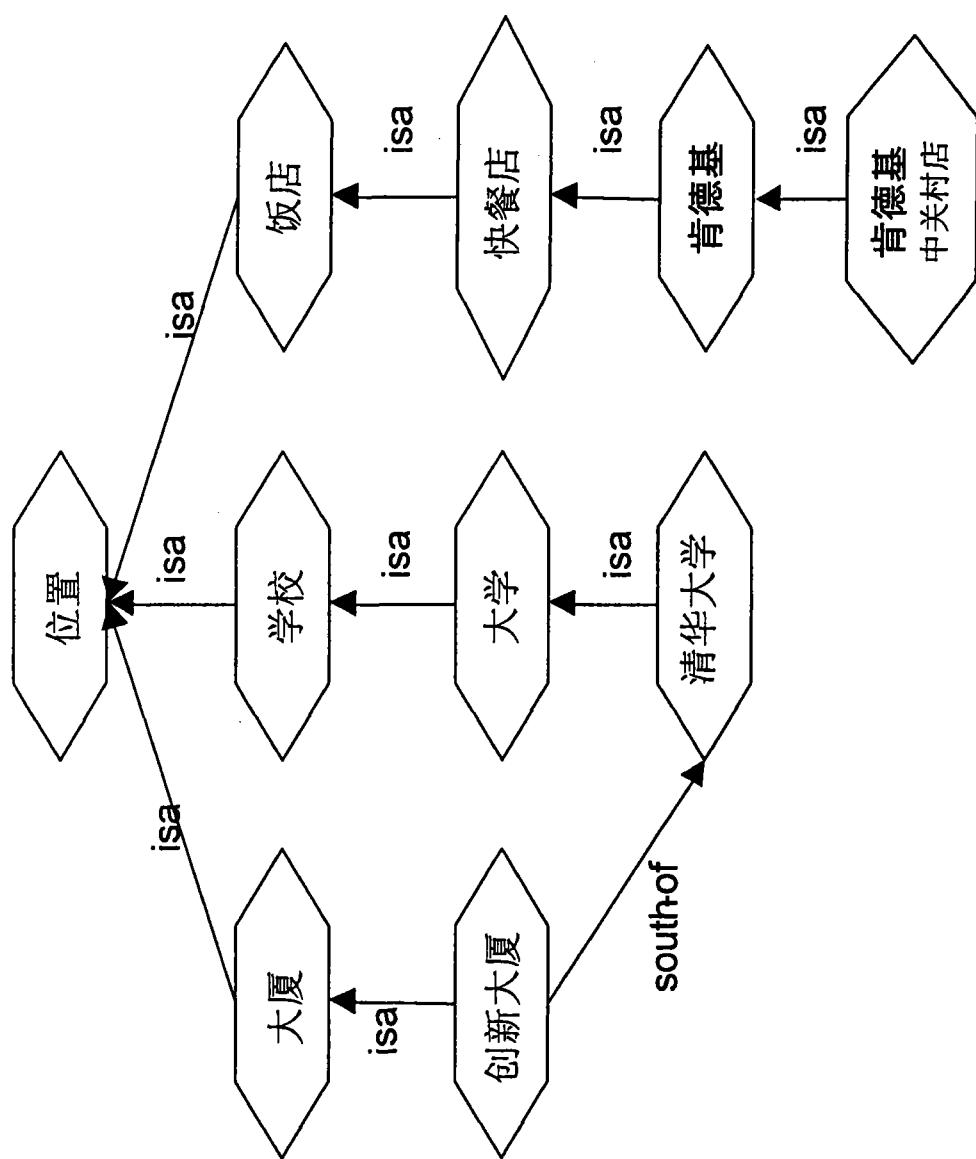


图 5f

```

Defquery 公用查询语言
{
    //文法部分
    文法: <CommonQuery1>=?C1(geo-entity)><#名词>
    文法: <CommonQuery2>=?C1(geo-entity)><#副词>{<?C2(geo-category|geo-entity)>}<!位于动词|拥有动词>|<!地点疑问词|!什么疑问词>
}

```

```

Defquery 地图领域查询语言
{
    //文法部分
    文法: <LocationQuery>=?C1(geo-entity)><!位于动词><!地点疑问词> | <CommonQuery1(<!名词>=<!位置名词>)>
    文法: <NearNeighborQuery>=<CommonQuery2(<!副词>=?[附近副词])>
    文法: <NearestNeighborQuery>=<CommonQuery2(<!副词>=?[最近副词])>

//动作部分
动作: isSyntax(<LocationQuery>)→QueryLocation(?C)
动作: isSyntax(<NearNeighborQuery>)→QueryNear(?C1, ?C2)
动作: isSyntax(<NearestNeighborQuery>)→QueryNearest(?C1, ?C2)
}

```

```

Defquery 自驾领域查询语言
{
    //文法部分
    文法: <PathPlanningQuery>=?C1(geo-entity)>[<!怎么疑问词>]<!到达介词><!C2(geo-entity)>

//动作部分
动作: isSyntax(<PathPlanningQuery>)→QueryPath(?C1, ?C2)
}

```

| 词表  |       |     |                         |
|-----|-------|-----|-------------------------|
| ID  | 类型    | 词性  |                         |
| 1   | 位于动词  | 动词  | 在 位于 地处                 |
| 2   | 位置名词  | 名词  | 位置 方位 地址                |
| 3   | 附近副词  | 副词  | 附近 邻近 周围 旁边 周边          |
| 4   | 最近副词  | 副词  | 最近 距离最短                 |
| 5   | 什么疑问词 | 疑问词 | 什么 哪些 哪个 哪个 啥           |
| 6   | 地点疑问词 | 疑问词 | 什么地方 啥地方 何处 何地 哪里 哪块 哪儿 |
| 7   | 拥有动词  | 动词  | 有 拥有 包括                 |
| ... | ...   |     | ...                     |

图 6

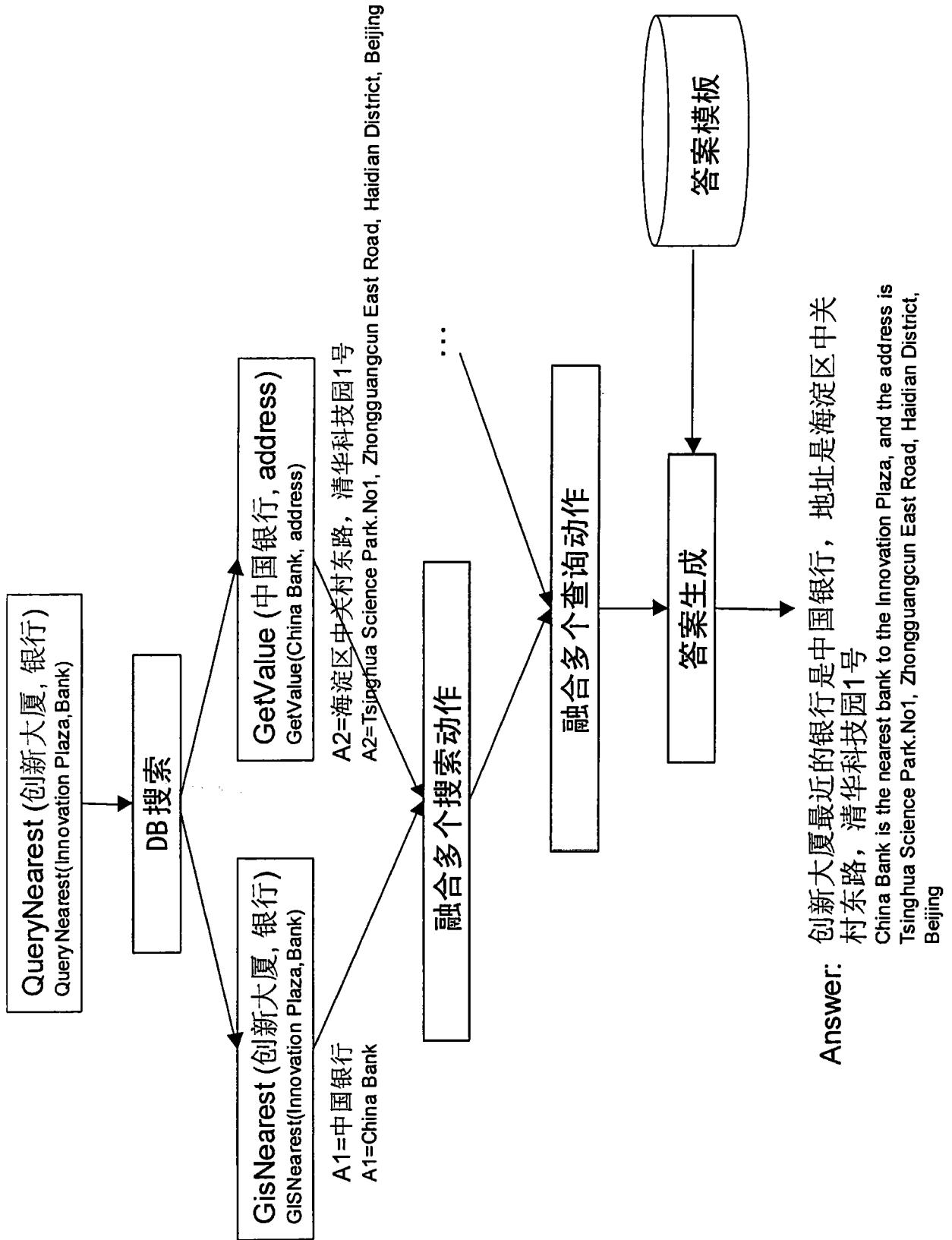


图 7

```
DefAnswerTemplate QueryNearest(X, Y)
{
    Entity=<GetNearest(X, Y)>
    Address=<GetValue(Entity,address)>
    EnglishAnswer=<Entity> is the nearest <Y> to the <X>, and the address is <Address>
    ChineseAnswer=<X>最近的<Y>是<Entity>, 地址是<Address>
    JapaneseAnswer=...
}
```

图 8

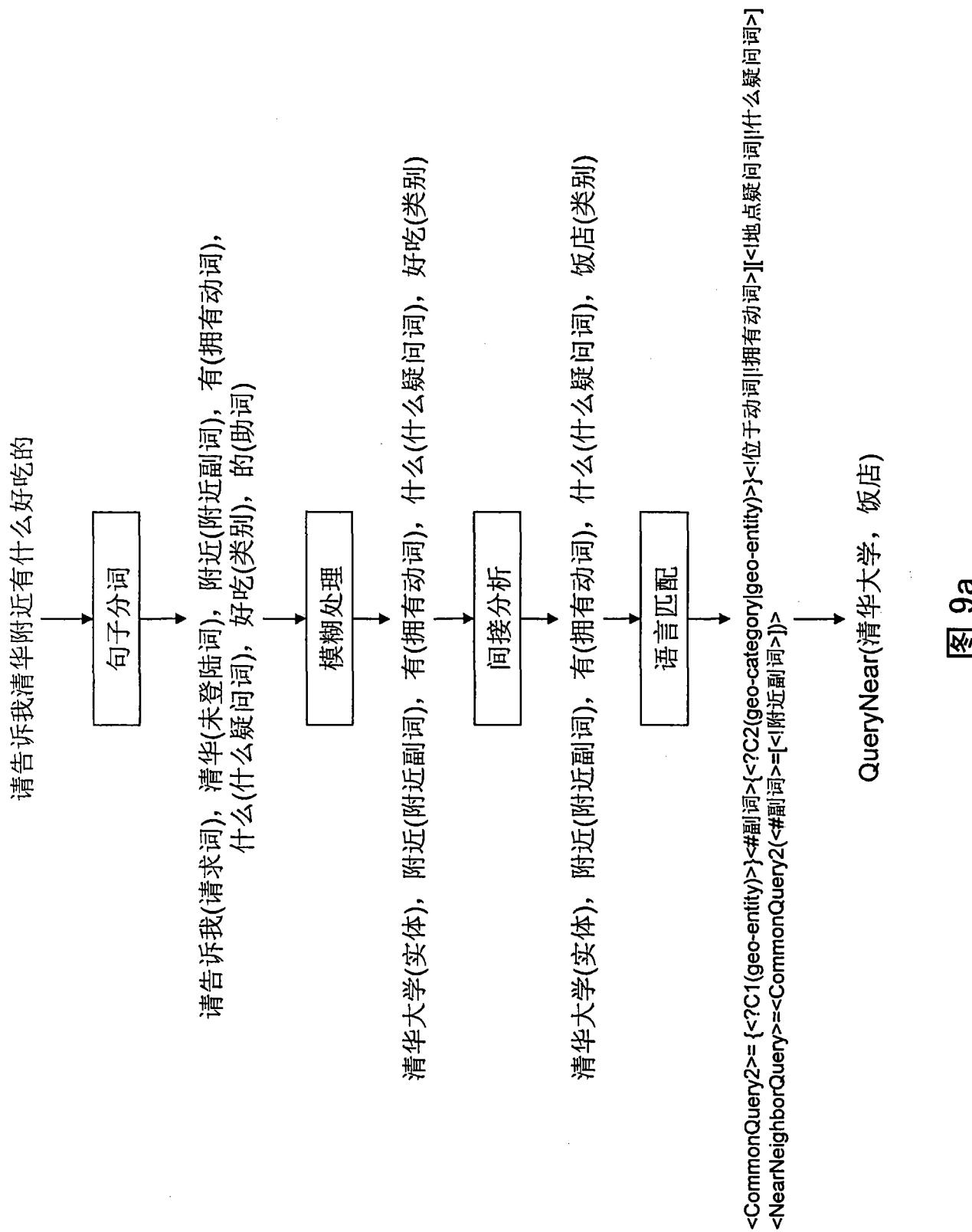


图 9a

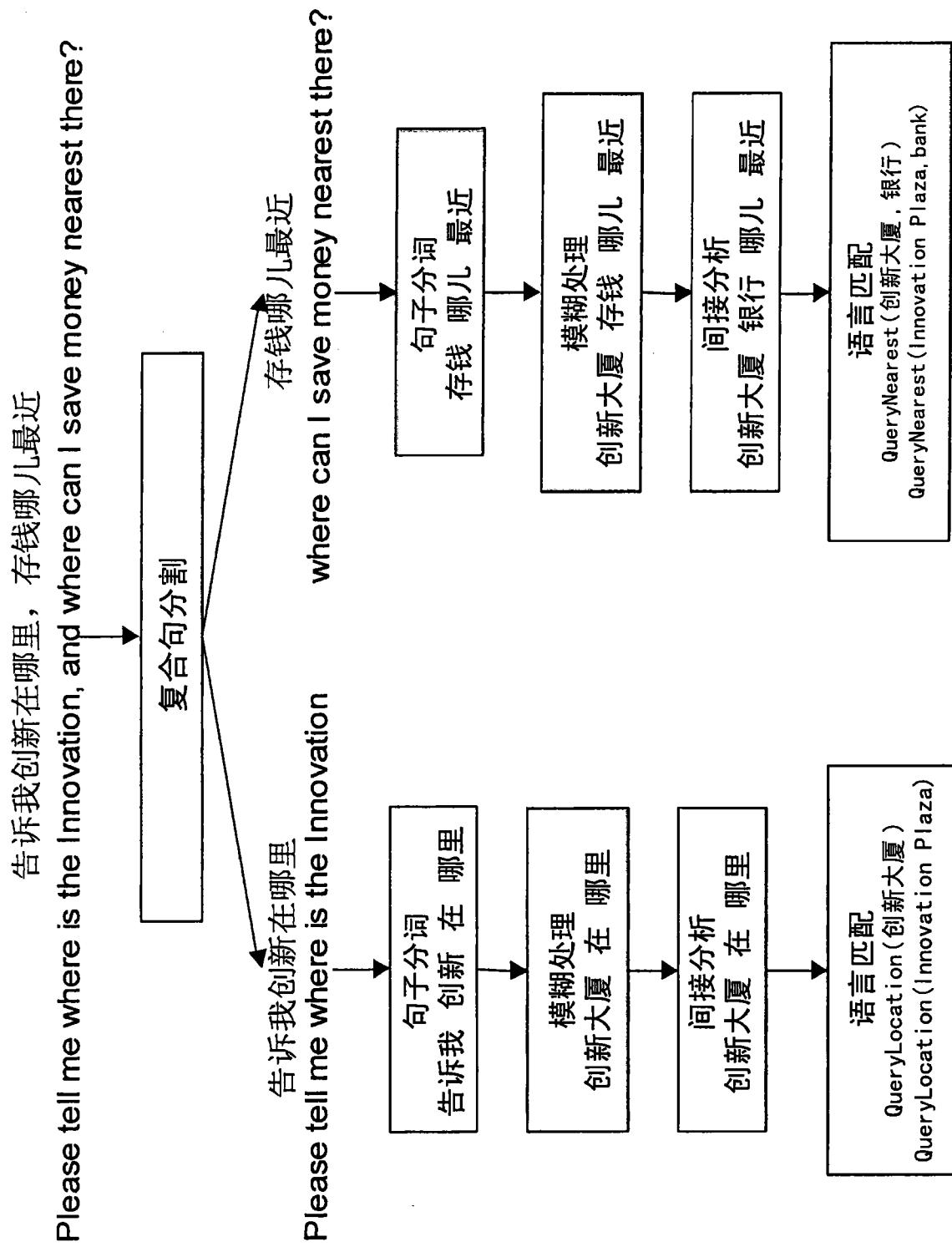


图 9b

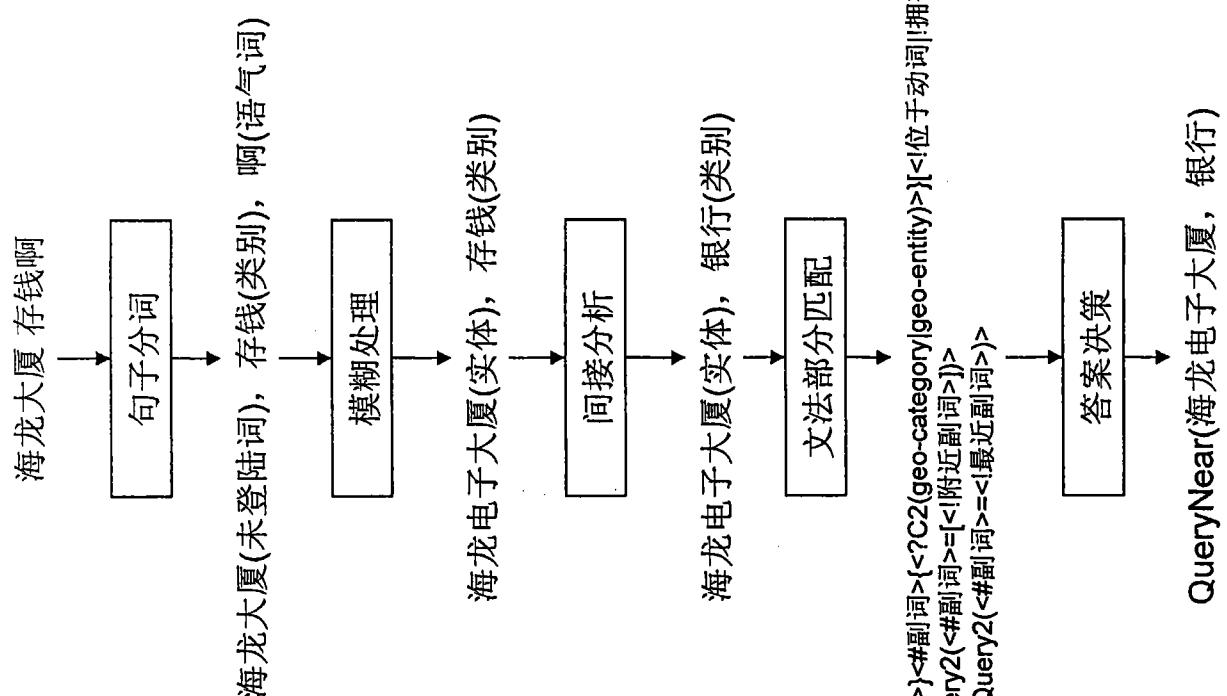
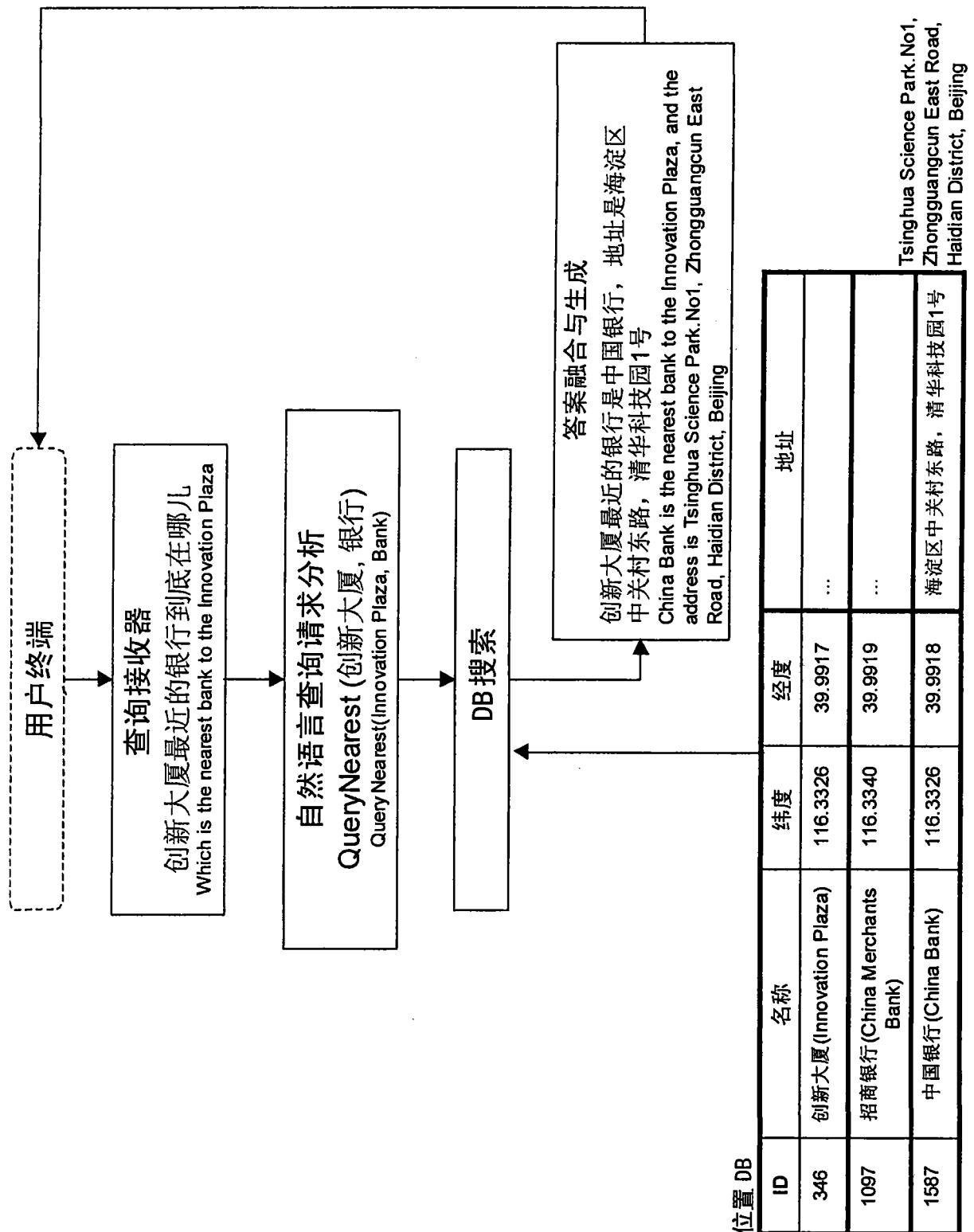


图 10



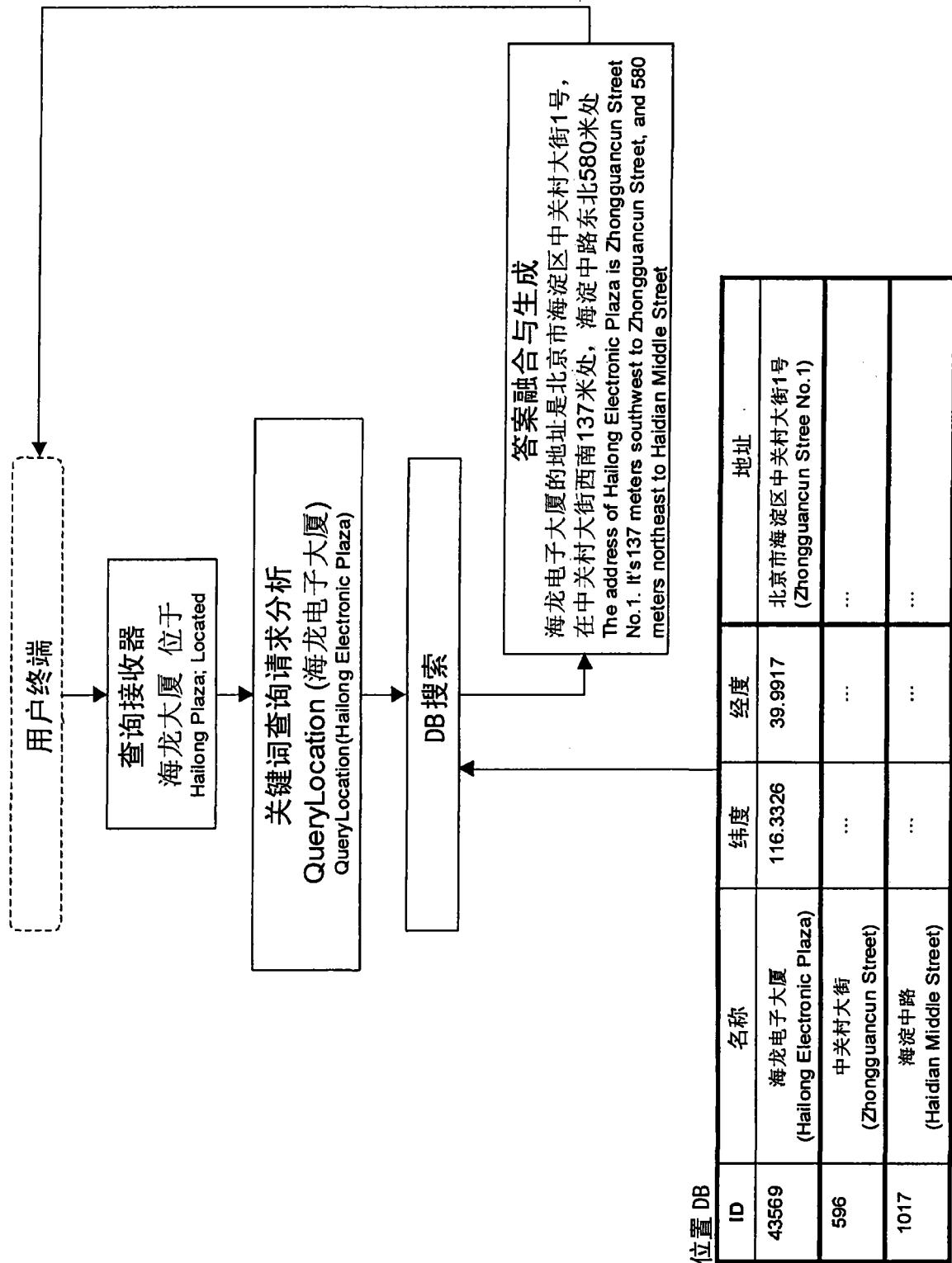


图 11b

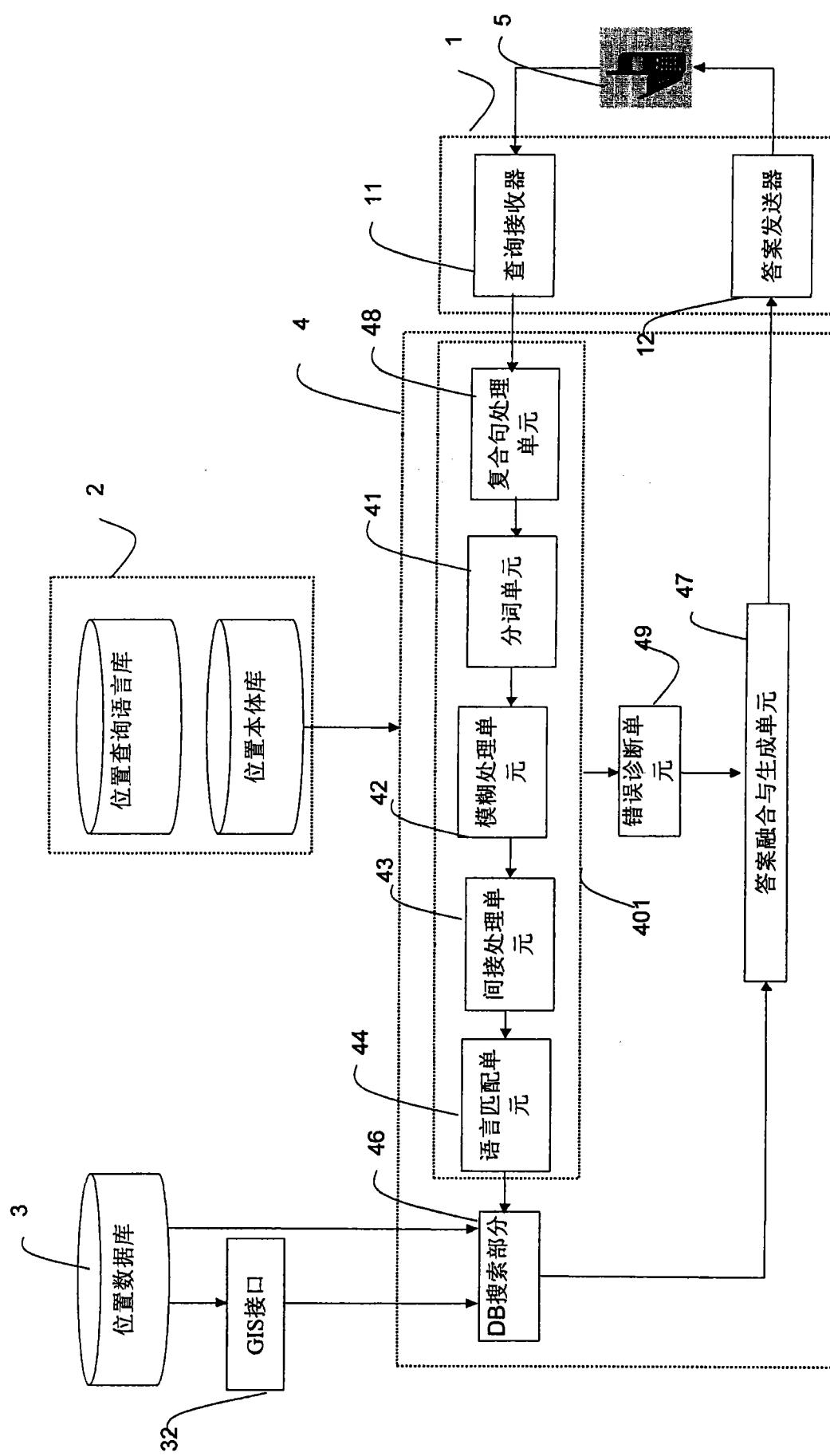


图 12

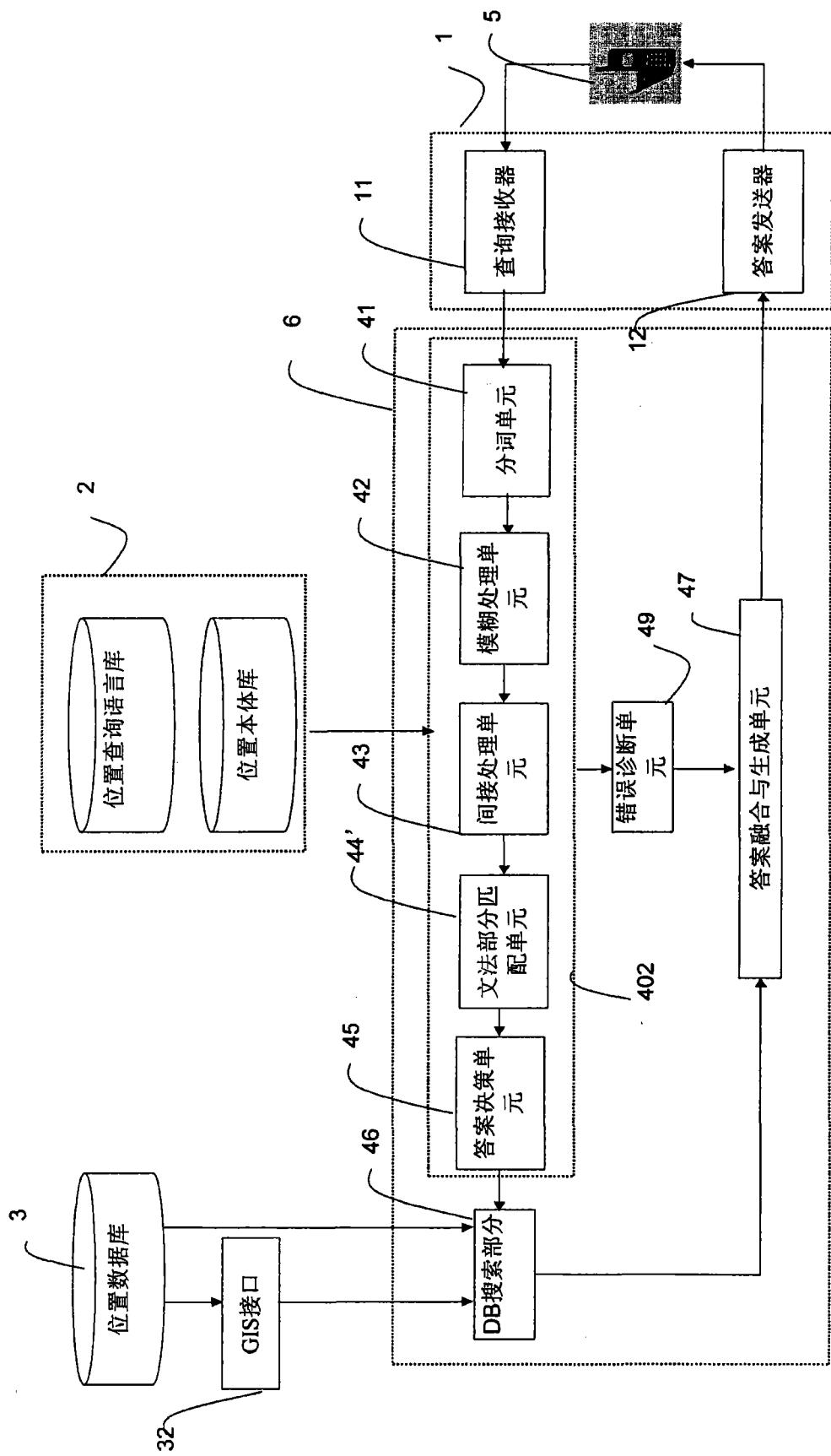


图 13

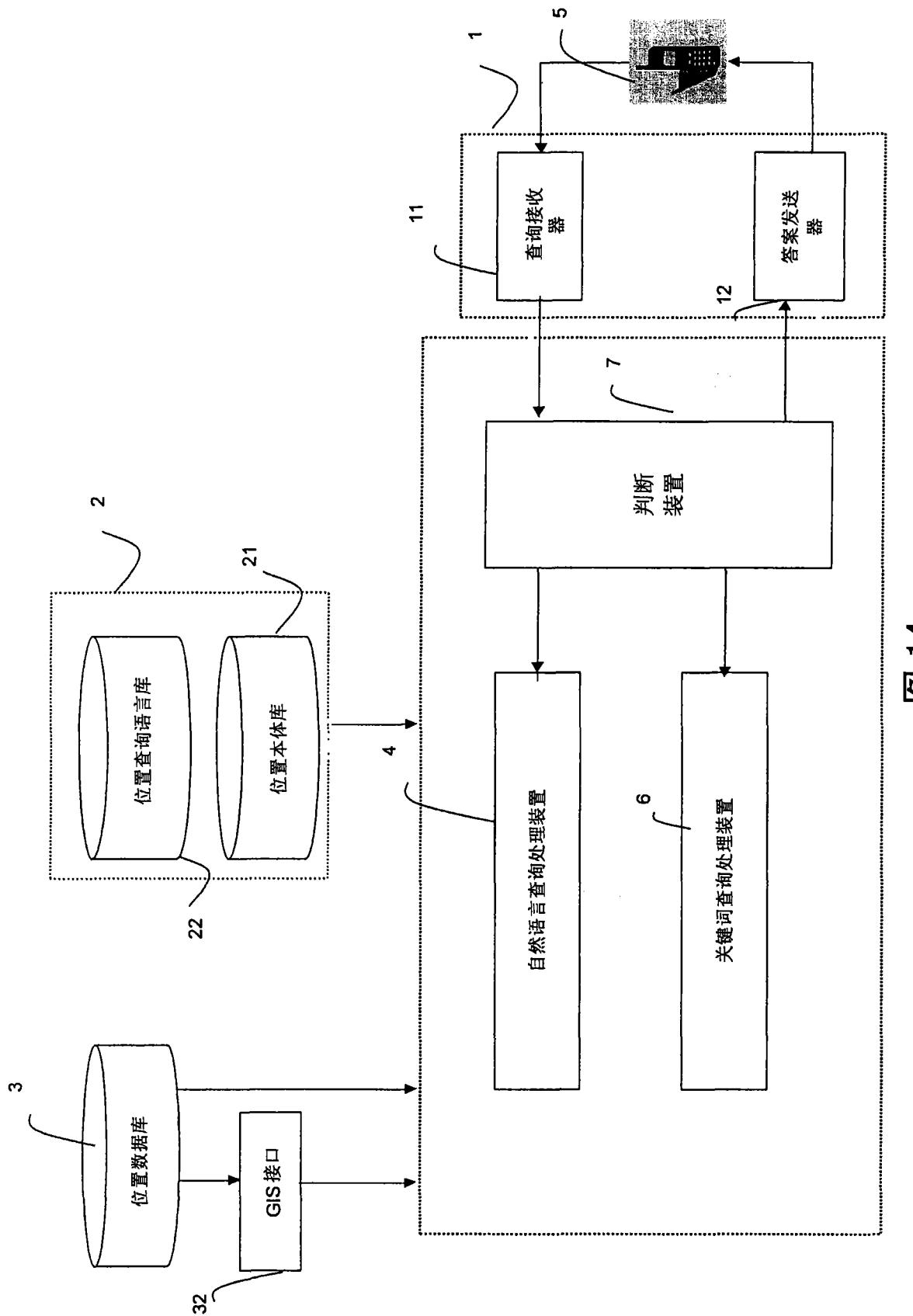


图 14a

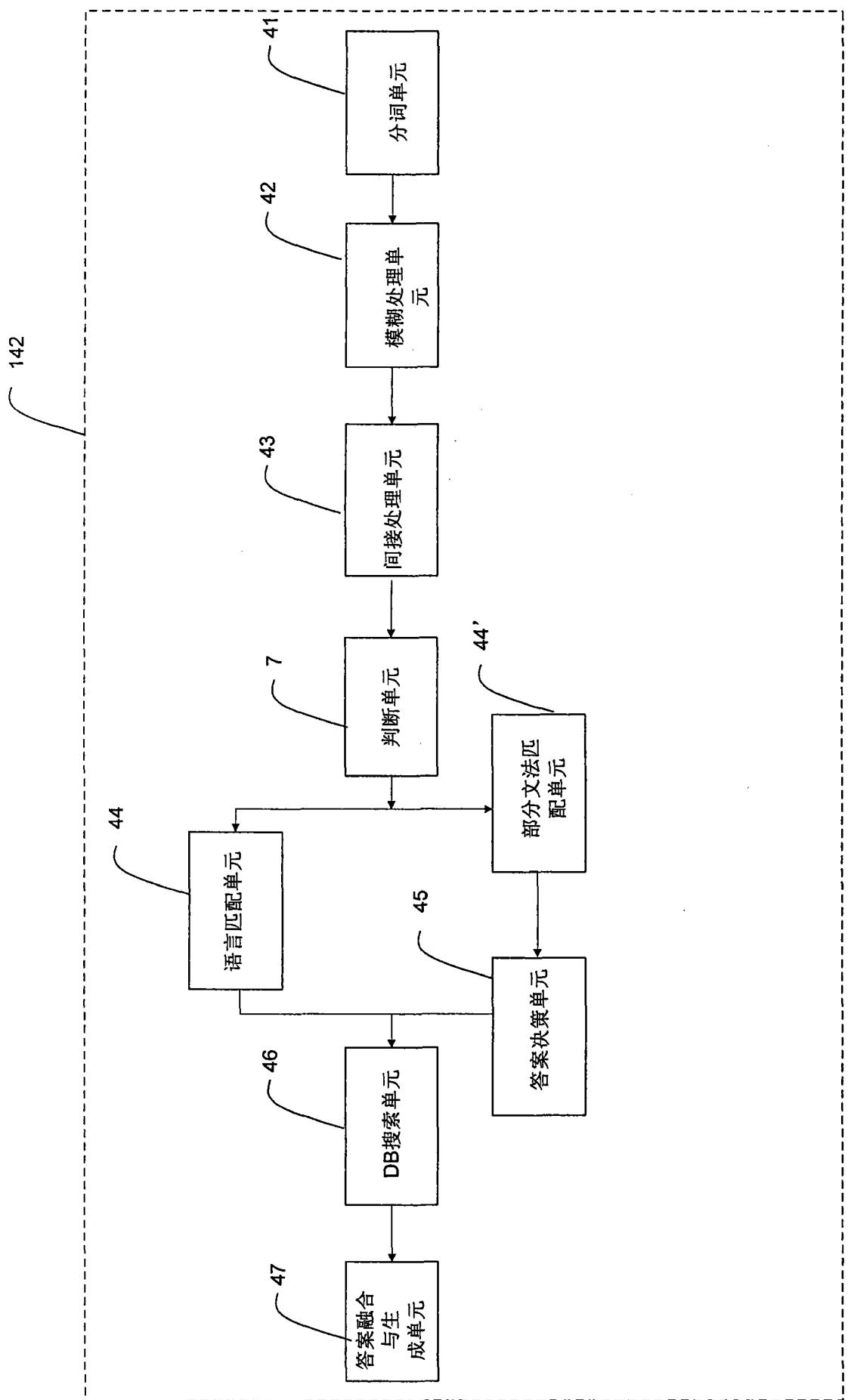


图 14b

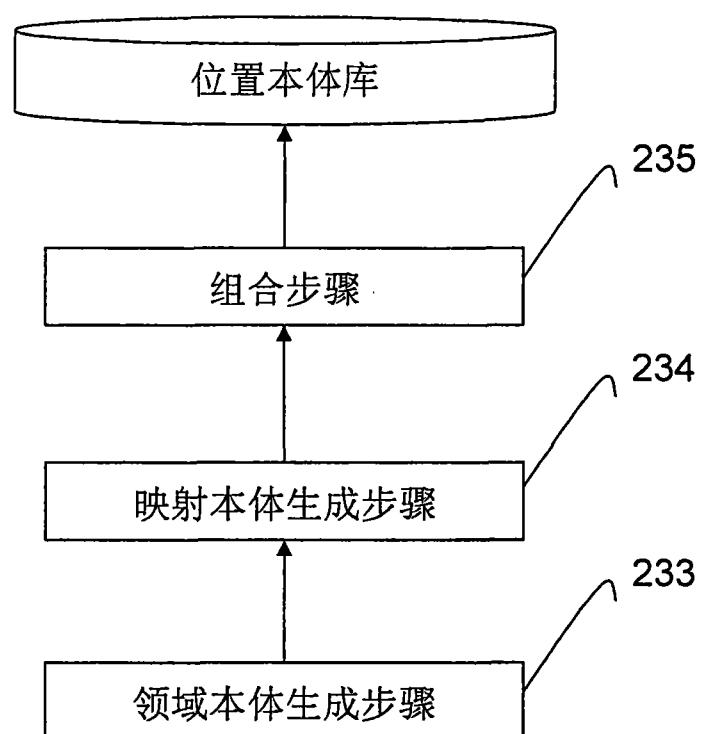


图 15

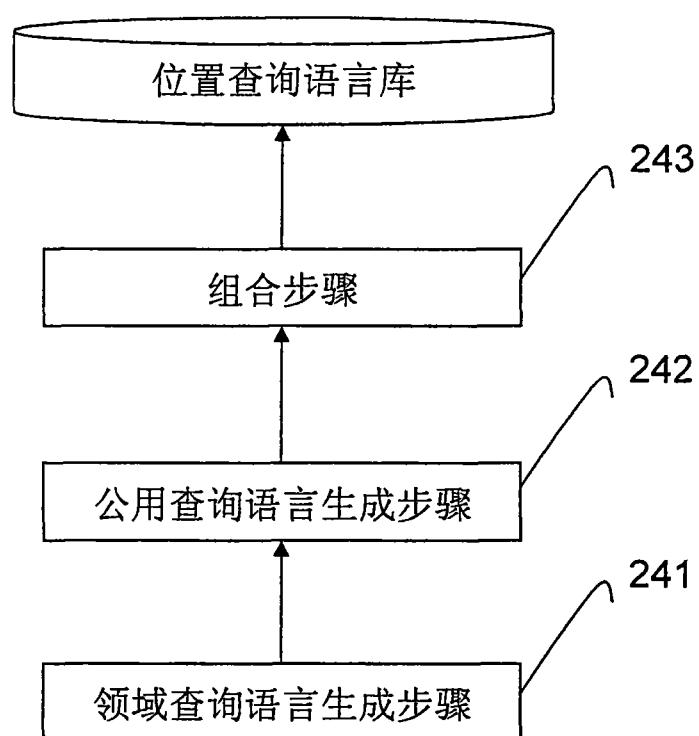


图 16