



(21) 申请号 201811151807.X

(22) 申请日 2018.09.29

(65) 同一申请的已公布的文献号
申请公布号 CN 109635270 A

(43) 申请公布日 2019.04.16

(30) 优先权数据
15/726,394 2017.10.06 US

(73) 专利权人 声音猎手公司
地址 美国加利福尼亚州

(72) 发明人 冯鹿 邢博纳

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258
专利代理师 林强

(51) Int.Cl.
G06F 40/253 (2020.01)
G06F 40/30 (2020.01)
G06F 40/58 (2020.01)
G06F 40/56 (2020.01)

(56) 对比文件

CN 1387650 A,2002.12.25
CN 103198149 A,2013.07.10
CN 105912521 A,2016.08.31
CN 104157285 A,2014.11.19
US 8812316 B1,2014.08.19
US 2006136196 A1,2006.06.22
US 2011288852 A1,2011.11.24
US 2008270110 A1,2008.10.30
US 2006224384 A1,2006.10.05

苏建雷.语音识别后处理中错误处理研究.
《中国优秀硕士学位论文全文数据库 信息科技
辑》.2015,

李伟 等.基于前向后向语言模型的语音识
别词图生成算法.《计算机应用》.2010,第30卷
(第10期), (续)

审查员 王悦

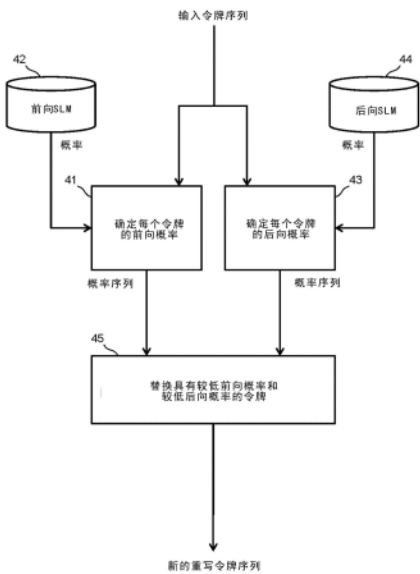
权利要求书2页 说明书17页 附图32页

(54) 发明名称

双向概率性的自然语言重写和选择

(57) 摘要

本公开涉及双向概率性的自然语言重写和选择。语音识别系统根据前向和后向统计语言模型 (SLM) 二者来在具有低概率的位置处执行令牌的插入、删除和替换编辑,以产生重写令牌序列。可以产生多个重写,并且分数取决于根据SLM的令牌的概率。可以根据自然语言语法来解析重写令牌序列以产生进一步加权的分数。可以使用基于树的搜索算法来迭代地重写令牌序列以找到最佳重写。输入令牌序列到重写令牌序列的映射可以被存储在缓存中,并且可以通过使用所缓存的重写(当存在时)来绕过搜索最佳重写。对产生相同的新的重写令牌序列的各种初始令牌序列的分析可用于改进自然语言语法。



CN 109635270 B

[接上页]

(56) 对比文件

Yik-Cheung Tam 等. ASR error detection
using recurrent neural network language

model and complementary ASR.《[https://
www.researchgate.net/publication/
262308966](https://www.researchgate.net/publication/262308966)》.2014,

1. 一种计算机实现的方法,用于在向用户提供查询结果时重写用户查询的输入令牌序列,所述方法包括:

通过计算机网络从在远程客户端设备上执行的应用接收来自系统用户的用户查询,所述查询包括所述输入令牌序列;

根据前向统计语言模型,确定所述输入令牌序列中的多个令牌的前向概率;

根据后向统计语言模型,确定所述输入令牌序列中的多个令牌的后向概率;

在具有低后向概率的第一令牌之后并且具有低前向概率的相邻第二令牌之前的位置处插入新令牌,以创建新的重写令牌序列;

处理所述新的重写令牌序列以产生结果;以及

向所述系统用户提供指示所述处理的结果的响应。

2. 一种重写输入令牌序列的计算机实现的方法,所述方法包括:

根据前向统计语言模型,确定所述输入令牌序列中的多个令牌的前向概率;

根据后向统计语言模型,确定所述输入令牌序列中的多个令牌的后向概率;

基于所述前向概率中的最低者和所述后向概率中的最低者,来计算所述多个令牌中的每一者的概率分数;以及

用新令牌替换具有最低分数的令牌,以创建新的重写令牌序列。

3. 一种重写输入令牌序列的计算机实现的方法,所述方法包括:

根据前向统计语言模型SLM,确定所述输入令牌序列中的多个令牌的前向概率;

根据后向统计语言模型,确定所述输入令牌序列中的多个令牌的后向概率;以及

用新令牌替换具有低于第一阈值的前向概率和低于第二阈值的后向概率的可疑令牌,以创建新的重写令牌序列。

4. 如权利要求3所述的方法,还包括:

在确定概率之前,替换所述输入令牌序列中的至少一个令牌的标记。

5. 如权利要求3所述的方法,还包括:

选择在根据所述前向SLM的最高概率令牌的列表和根据所述后向SLM的最高概率令牌的列表二者中的令牌,作为所述新令牌。

6. 如权利要求5所述的方法,还包括:

根据句法规则执行所述输入令牌序列的句法分析;以及

根据所述句法规则将所述选择限制于仅在相邻令牌的上下文中在句法上合法的令牌。

7. 如权利要求5所述的方法,还包括:

针对所述新的重写令牌序列计算重写分数,所述重写分数至少取决于所述新令牌在所述前向SLM中的概率以及所述新令牌在所述后向SLM中的概率。

8. 如权利要求7所述的方法,还包括:

基于所述新令牌在根据与各种主题相关的表述所构建的不同语料库SLM中的概率来缩放所述重写分数。

9. 如权利要求5所述的方法,还包括:

用替代的新令牌替换所述可疑令牌,以创建替代的重写令牌序列;

针对所述替代的重写令牌序列,将替代分数计算为所述替代的新令牌在所述前向SLM中的概率和所述替代的新令牌在所述后向SLM中概率的组合;以及

选择所述新的重写令牌序列和所述替代的重写令牌序列中具有较高分数的一者。

10. 如权利要求5所述的方法,还包括:

维护来自最近的连续语音的推测令牌的令牌缓冲器,

其中,所述输入令牌序列是所述令牌缓冲器中的令牌的序列。

11. 如权利要求3所述的方法,还包括:

存储最近的令牌序列中出现的令牌的历史缓存;

为了替换所述可疑令牌,从前向最可能的令牌的列表和后向最可能的令牌的列表中每一个列表选择所述新令牌;以及

将所述历史缓存中存在的至少一个令牌的概率分数提高。

12. 如权利要求3所述的方法,还包括:

使用自然语言解析器根据语法来解析所述新的重写令牌序列以产生解析分数。

13. 如权利要求12所述的方法,还包括:

用与所述可疑令牌不同的替代的可疑令牌进行替换以创建替代的重写令牌序列;

使用所述自然语言解析器根据所述语法来解析所述替代的重写令牌序列以产生替代的解析分数;并且

选择所述新的重写令牌序列和所述替代的重写令牌序列中具有较高解析分数的一者。

14. 如权利要求12所述的方法,还包括:

用替代的新令牌替换所述可疑令牌以创建替代的重写令牌序列;

使用所述自然语言解析器根据所述语法来解析所述替代的重写令牌序列以产生替代的解析分数;以及

选择所述新的重写令牌序列和所述替代的重写令牌序列中具有较高解析分数的一者。

15. 如权利要求12所述的方法,还包括:

使用基于树的算法迭代地执行重写,并计算每个重写的分数以产生一组重写,并从该组重写中选择具有最佳分数的一个重写。

16. 如权利要求3所述的方法,还包括:

将所述输入令牌序列存储在缓存中;

将所述新的重写令牌序列与所述输入令牌序列相关联地存储在所述缓存中;以及

针对所述输入令牌序列来对所述缓存进行搜索。

17. 如权利要求16所述的方法,还包括:

针对所述新的重写令牌序列,分析所述缓存以标识被重写到所述新的重写令牌序列的最频繁的输入令牌序列。

18. 如权利要求17所述的方法,还包括:

创建语法规则以覆盖被重写到所述新的重写令牌序列的最频繁的输入令牌序列。

19. 如权利要求17所述的方法,还包括:

调整语法规则以覆盖被重写到所述新的重写令牌序列的最频繁的输入令牌序列。

双向概率性的自然语言重写和选择

技术领域

[0001] 本发明属于语音识别和自然语言理解的领域。

背景技术

[0002] 自动语音识别(ASR)容易出错。

[0003] 给定输入音频的声学分析,ASR对音频信号执行频谱分析并提取特征,从中该过程推测出多个音素序列,其中,每个音素序列具有表示其是正确的的可能性的分数。ASR根据字典继续将音素序列推测令牌化(tokenize)为令牌序列推测,同时维持每个推测的分数。令牌可以是字母词,例如,英文词、诸如汉字之类的语素字符、或其他类型的书写系统的可辨别的元素单位。令牌化是不精确的,因为例如英文讲话者对表达“I scream(我呐喊)”和“ice cream(冰淇淋)”的发音几乎相同。为了处理这种模糊性,ASR系统使用对口语中常见的相邻词令牌的已知频率的统计来推测多个令牌序列推测中的哪一个是正确的。例如,词“ice(冰)”经常在词“eat(吃)”之后,如“eat ice cream(吃冰淇淋)”,但词“I(我)”很少跟在词“eat(吃)”之后。因此,如果词序列推测“I scream(我尖叫)”和“ice cream(冰淇淋)”在词“eat(吃)”之后,则具有“ice cream(冰淇淋)”的词序列推测的分数提升,而具有“I scream(我尖叫)”的词序列推测的分数降低。

[0004] 例如,普通话讲话者对表达“交代”和“胶带”的发音是相同的。因此,语音识别使用对口语中常见的相邻令牌的已知频率的统计来推测多个令牌序列推测中的哪一个是正确的。例如,词“胶”经常在词“卷”之后,如“卷胶带”,但词“交”很少在“卷”之后。因此,如果词序列推测“胶带”和“交代”在词“卷”之后,则具有“胶带”的词序列推测的分数提升,而具有“交代”的词序列推测的分数降低。

[0005] 传统的语音识别和自然语言理解系统是相对不准确且缓慢的。它们产生语法不正确的转录。它们使实时应用的延迟增加。此外,它们的语法规则很难创建和改进。此外,语法通常不会捕获用户表达他们自己的所有非正式和近似方式,因此覆盖范围不足。

发明内容

[0006] 本发明涉及用于编程系统来使用前向和后向SLM的组合重写由语音识别产生的令牌序列的方法、系统和软件。这可以涉及根据序列内的令牌的概率的组合来计算令牌序列的分数。

[0007] 本发明提供了对语音识别的准确性和速度、转录的语法正确性、实时转录的延迟、以及自然语言语法的增强的技术改进。

[0008] 这可以涉及通过删除、插入或替换特定令牌来编辑令牌序列。前向和后向概率均较低的令牌是针对删除、或者针对由令牌序列中的新令牌替换的候选者。这样的一对相邻的令牌,其中较早的令牌在后向方向上具有较低分数并且较后的令牌在前向方向上具有较低分数,是针对删除任一令牌或用于在其间插入新令牌的候选者。

[0009] 选择用于插入或替换的新令牌可以涉及检查前向SLM或后向SLM中的任一者以创

建最高概率令牌的列表,或者检查两个SLM以在两个SLM中找到具有较高概率的令牌。在基于两个SLM中的概率来选择新令牌的实施例中,最佳选择通常不是任一列表上最高的那一个,而是在两个列表上都相当高的一个。

[0010] 删除、插入或替换中的任何操作都创建具有新的组合分数的新令牌序列。一些实施例产生多个可能的重写令牌序列、计算每个重写令牌序列的分数、并选择输出具有最佳分数的重写令牌序列。一些实施例执行令牌序列的多次连续重写。

[0011] 一些实施例将令牌序列作为输入提交给自然语言解析器,该自然语言解析器根据一组语法规则来确定令牌序列的解析分数。在这样的实施例中,最佳令牌序列的最终选择(原始或任何特定重写)由语法解析分数或者语法解析分数与令牌序列概率分数的组合来确定。一些这样的语法分析器仅根据语法来产生指示令牌序列是有效还是无效的输出生。用于随机语法的解析器分配概率分数以进行解析。语义解析器使用语义语法,其中解析的分数可能受其语义的影响。

[0012] 一些实施例执行令牌序列的部分句法分析,例如,利用语音标记的部分和已知在口语中合法的语法模式的应用。

[0013] 一些实施例使用对于通常在一种或多种完整人类语言中使用的各种表述语料库通用的SLM、句法检查和语法规则中的一个或多个。一些实施例使用域特定SLM、句法检查、和语法,它们在根据特定的知识领域重写令牌序列可以是更准确的。

[0014] 一些实施例使用实体标记来用指示其通用类别的特殊标记令牌替换已知的特定令牌。一个示例是利用通常代表城市的标记来替换作为城市名称的每个标记。这可以改善满足模式的令牌序列的数量,从而使统计概率更准确。

[0015] 一些实施例在实时语音的延迟缓冲器内对来自实时ASR的令牌序列执行重写。一些实施例仅使用SLM概率来以较低延迟执行重写。一些实施例使用较大延迟缓冲器并尝试解析完整的表述短语。

[0016] 一些实施例存储最近使用的令牌的历史缓存,并且将历史缓存中存在的令牌的基于概率的分数提高。这通过利用主题特定的词在短期会话期间倾向于被多次说出的事实来提高准确性。

[0017] 一些实施例在缓存中存储原始输入令牌序列和所选择的一个或前几个重写二者。这可以通过缓存中进行的查找来完成未来的重写请求,而无需新搜索许多可能的重写。一些实施例存储重写的令牌序列和导致该重写的一个或多个输入令牌序列。这可以提供有用的信息,特别是在按发生频率进行排序时,用于语法开发者来创建和改进域特定语法。

[0018] 尽管不使用双向概率性重写的语音识别系统可能需要相对精细和复杂的语法,但是通过使用双向概率性重写,语音识别系统能够通过将近似表述映射到语法正确的一者来扩展相对简单的语法的覆盖范围。因此,经验较少的语法开发者能够依赖于采用用于双向概率性重写的编程规则作为支持创建满足其应用的需求的定制域语法的工具的语音识别系统。这具有更高性能且更精确的语音识别系统以及自然语言理解系统二者的效果。

附图说明

[0019] 图1示出了根据实施例的启用语音的自然语言理解和处理系统的流程图。

[0020] 图2示出了英文中的典型语音识别错误。

- [0021] 图3示出了普通话中的典型语音识别错误。
- [0022] 图4示出了根据实施例的重写的流程图。
- [0023] 图5A示出了英文表述的简单示例语料库。
- [0024] 图5B示出了根据图5A的语料库构建的前向SLM。
- [0025] 图5C示出了根据图5A的语料库构建的后向SLM。
- [0026] 图6A示出了普通话表述的简单示例语料库。
- [0027] 图6B示出了根据图6A的语料库构建的前向SLM。
- [0028] 图6C示出了根据图6A的语料库构建的后向SLM。
- [0029] 图7示出了英文中的简单示例自然语言域语法。
- [0030] 图8示出了普通话中的简单示例自然语言域语法。
- [0031] 图9示出了根据由图7的域语法解析的表述的不同语料库构建的前向SLM。
- [0032] 图10示出了根据由图7的域语法解析的表述的不同语料库构建的后向SLM。
- [0033] 图11示出了根据实施例的在令牌标签替换之后的重写的流程图。
- [0034] 图12A示出了英文中被正确识别的表述的示例。
- [0035] 图12B示出了根据实施例的图12A的表述的变体的示例,其中错误的令牌错误在所替换的令牌处导致较低的前向和后向概率。
- [0036] 图12C示出了根据实施例的图12A的表述的变体的示例,其中额外的令牌错误在所插入的令牌处导致较低的前向和后向概率。
- [0037] 图12D示出了根据实施例的图12A的表述的变体的示例,其中丢失的令牌错误导致令牌上的较低后向概率,然后是紧接着的令牌上的叫低前向概率。
- [0038] 图12E示出了根据实施例的图12A的表述的变体的示例,其中重复的令牌错误导致令牌上的较低后向概率,然后是紧接着的令牌上的较低前向概率。
- [0039] 图13A示出了普通话中被正确识别的表述的示例。
- [0040] 图13B示出了根据实施例的图13A的表述的变体的示例,其中错误的令牌错误在所替换的令牌处导致较低的前向和后向概率。
- [0041] 图13C示出了根据实施例的图13A的表述的变体的示例,其中额外的令牌错误在所插入的令牌处导致较低的前向和后向概率。
- [0042] 图13D示出了根据实施例的图13A的表述的变体的示例,其中丢失的令牌错误导致令牌上的较低后向概率,然后是紧接着的令牌上的较低前向概率。
- [0043] 图13E示出了根据实施例的图13A的表述的变体的示例,其中重复的令牌错误导致令牌上的较低后向概率,随后是紧接着的令牌上的较低前向概率。
- [0044] 图14示出了根据实施例的使用前向和后向SLM中的一组最可能的令牌进行重写的流程图。
- [0045] 图15A示出了根据实施例的英文表述中的替换错误。
- [0046] 图15B示出了根据实施例的使用2字母组 (gram) SLM的针对图15A的错误的一组最可能的预期令牌。
- [0047] 图15C示出了根据实施例的在后向方向 (具有接近每个方向的最高概率的公共令牌) 使用2字母组SLM的针对图15A的错误的一组最可能的预期令牌。
- [0048] 图16A示出了根据实施例的普通话表述中的替换错误。

[0049] 图16B示出了根据实施例的在前向方向使用2字母组SLM的针对图16A的错误的一组最可能的预期令牌。

[0050] 图16C示出了根据实施例的在后向方向(具有接近每个方向的最高概率的公共令牌)使用2字母组SLM的针对图16A的错误的一组最可能的预期令牌。

[0051] 图17A示出了根据实施例的中文表述中的替换错误。

[0052] 图17B示出了根据实施例的在前向方向使用3字母组SLM的针对图17的错误的一组最可能的预期令牌。

[0053] 图17C示出了根据实施例的在后向方向(具有接近每个方向的最高概率的两个公共令牌)使用3字母组SLM的针对图17A的错误的一组最可能的预期令牌。

[0054] 图17D示出了根据实施例的针对错误令牌(每个公共令牌具有针对每个SLM的较高概率)的替换的重写分数的计算。

[0055] 图18A示出了根据实施例的英文表述中的替换错误。

[0056] 图18B示出了根据实施例的在前向方向使用3字母组SLM的针对图18A的错误的一组最可能的预期令牌。

[0057] 图18C示出了根据实施例的在后向方向(具有接近每个方向的最高概率但处于不同概率顺序的两个公共令牌)使用3字母组SLM的针对图18A的错误的一组最可能的预期令牌。

[0058] 图18D示出了根据实施例的针对错误令牌(每个公共令牌具有针对每个SLM的较高概率)的替换的重写分数的计算。

[0059] 图19示出了根据实施例的使用前向和后向SLM中的一组最可能的令牌进行重写以产生新的重写令牌序列和相应的分数的流程图。

[0060] 图20示出了根据实施例的使用前向和后向SLM中的一组最可能的令牌进行重写以产生新的重写令牌序列和由不同的语料库SLM引导的相应的分数的流程图。

[0061] 图21示出了根据实施例的使用前向和后向SLM中的一组最可能的令牌进行重写以产生多个新的重写令牌序列和相应的分数并选择具有最高分数的重写令牌序列的流程图。

[0062] 图22示出了根据实施例的使用句法分析的重写引导从前向和后向SLM中的集合选择最可能的令牌的流程图。

[0063] 图23示出了根据实施例的后跟使用域语法的自然语言解析的重写的流程图。

[0064] 图24示出了根据实施例的针对英文中的不同令牌插入的示例语法解析分数。

[0065] 图25示出了根据实施例的针对普通话中的不同令牌插入的示例语法解析分数。

[0066] 图26示出了根据实施例的后跟使用域语法的自然语言解析(选择最高分数的重写序列)的重写(产生多个新的重写令牌序列和相应的分数)的流程图。

[0067] 图27示出了根据实施例的后跟令牌历史缓存(以自动化重写)的重写的流程图。

[0068] 图28示出了根据实施例的用于将输入令牌序列转换为重写令牌序列的令牌历史缓存的内容。

[0069] 图29示出了根据实施例的用于将重写令牌序列转换为输入令牌序列的令牌历史缓存的内容。

[0070] 图30示出了根据实施例的与基于服务器的启用语音的自然语言理解和处理系统的用户交互。

- [0071] 图31A示出了根据实施例的盘非暂态计算机可读介质。
- [0072] 图31B示出了根据实施例的闪速存储器非暂态计算机可读介质。
- [0073] 图31C示出了根据实施例的片上系统处理器。
- [0074] 图31D示出了根据实施例的片上系统处理器介质。
- [0075] 图32示出了根据实施例的片上系统的框图。
- [0076] 图33示出了根据实施例的服务器系统。
- [0077] 图34示出了根据实施例的服务器系统的框图。

具体实施方式

[0078] 概述

[0079] 统计语言模型 (SLM) 捕获对给定表述语料库中的相邻词的统计。将SLM应用于令牌序列推测显著提高了ASR系统的准确性。

[0080] 前向SLM表示给定一个或一系列先前令牌的下一令牌的条件概率。后向SLM表示给定一个或一系列后续令牌的紧接的在先令牌的条件概率。任何给定的令牌对在前向和后向的中每个方向可能具有非常不同的概率。考虑著名的美国棒球运动员Joe DiMaggio。在后向SLM中, DiMaggio之前的词很可能是Joe, 因为很少有名人的姓DiMaggio。然而, 在前向SLM中, 词DiMaggio在词Joe之后的概率相当低, 因为有许多名人叫Joe。

[0081] 考虑中国三国时期的著名领导人诸葛亮。在前向SLM中, 在诸葛之后的词为亮的概率非常高, 因为很少有名人的姓诸葛。然而, 在后向SLM中, 在词亮之前的词为诸葛的概率相当低, 因为有许多名人叫亮。

[0082] 表述是令牌的序列。在各种实施例中, 令牌是例如英文中的字母词或诸如中文字符之类的语素 (logogram)。在一些实施例中, 令牌可以表示代表单个概念的多个词, 例如一起代表城市的两个词“New”和“York”或者一起代表城市的“纽”和“约”。在一些实施例中, 不同的令牌表示具有相同拼写的词的不同含义, 例如, 汽车上的“tire (轮胎)”和变为困倦的“tire (疲倦)”, 表示气愤的“生气”和表示活力和生气勃勃的“生气”。

[0083] 图1示出了用于启用语音的自然语言处理的系统的高级视图。语音识别模块11从发出口头表述的人接收语音音频。语音识别模块11产生列令牌序列。一些实施例在不完全清楚人所说的内容时产生多个替代令牌序列。一些这样的实施例为替代给出概率权重。

[0084] 一些实施例通过除语音音频之外的方式从人接收令牌序列, 例如, 在键盘上打字、敲击触摸屏、眼睛运动、用手臂、手或手指打手势、或神经电信令。

[0085] 重写模块12接收令牌序列作为输入令牌序列。它将如下所述的编程规则应用于令牌序列以执行重写。在多个令牌序列的情况下, 它顺序地或并行地处理每个输入令牌序列。重写模块使用前向SLM 13和后向SLM 14来产生一个或多个新令牌序列。一些实施例为每个新令牌序列产生相应的分数, 该分数指示所重写的令牌序列是人想要表述的内容的可能性。一些实施例将新令牌序列馈送回重写模块12并重新处理它们以试图产生甚至更好的重写令牌序列。例如, 如果令牌序列具有多个错误, 则这将非常有用。

[0086] 一些实施例使用自然语言解析模块15。解析模块接收新令牌序列, 或者在一些实施例中, 接收多个新令牌序列, 并根据域特定语法16解析 (一个或多个) 令牌序列以产生自然语言解释和与解释相对应的解析分数。一些实施例使用多个域特定语法16并根据每个语

法16解析新令牌序列。一些实施例将来自重写模块12的重写分数作为解析中的权重来应用。一些实施例通过重写分数来对解析分数进行加权,以产生该解释是人的意图的最佳估计。

[0087] 一些实施例将解析分数反馈回重写模块12中,该重写模块12使用解析分数来选择多个可能的重写中最可能正确的一个或多个重写。

[0088] 最终,处理模块17接收分数最高的解释,确定其意图并相应地对其进行处理,并产生最可能有用的结果。处理模块17产生结果并将其作为系统输出提供给用户。

[0089] 通过使用重写模块12,实施例更可能确定人的表述的正确令牌序列,并因此产生正确的解释和有用的结果。通过进一步使用重写分数反馈和解析分数反馈,实施例甚至更可能确定人的表述的正确令牌序列,并因此产生正确的解释和有用的结果。结果,这样的实施例通过将错误识别的或语法上不正确的输入令牌序列变换为语法正确的形式,来增加语法的语义覆盖(即由语法识别的语料库的部分),从而有效地允许语法支持有缺陷的变体并执行错误校正。

[0090] 双向概率性自然语言重写和选择扩展了语法可以解释的表述的范围,这实际上是语法规则压缩的一种形式。它允许较小语法支持否则需要大得多的语法的表述范围。相应地,它允许相同大小的语法来解释更广泛的人类表述。

[0091] 用于人类语言之间的自动机器翻译的系统也可以获得提升的翻译准确性。一些这样的实施例在ASR结果被输入到翻译模块之前对其应用重写。重写减少了可能需要翻译的表述的范围,从而允许更简单的翻译规则来实现翻译准确性的必要要求。

[0092] 图30示出了服务器实现的系统中的图1的元件,该服务器实现的系统经由网络通过被耦合到服务器的客户端从用户接收表述,这对于被实现为移动电话应用或家用设备的虚拟助理是典型的。一些实施例是独立的并且独立于网络来工作,这对于汽车系统、高安全性系统和极低功率系统是典型的。

[0093] 图2示出了英文词令牌序列的类型和示例,其中人想要进行与天气有关的查询,但是未能通过与天气相关的语法来正确地解析。

[0094] 图3示出了普通话字符令牌序列的类型和示例,其中人想要进行与天气有关的查询,但是未能通过与天气相关的语法来正确地解析。

[0095] 图2和图3中的每个示例序列由下列四种类型的错误中的一种或多种引起:错误令牌、额外令牌、丢失令牌和重复令牌。每种类型的错误可以由下列三种类型的编辑中的一种来校正:令牌插入、令牌删除和令牌替换。一些实施例不具有令牌替换类型的编辑,但利用令牌删除和令牌插入的组合实现相同的效果。令牌插入和令牌替换中的每一个都创建输入令牌序列中不存在的新令牌。如下所述,程序规则系统基于鉴于先前和后续令牌的令牌的条件概率,确定用于对令牌重写进行编辑的编辑类型和位置的选择。

[0096] 对于令牌替换,编辑的位置在将被替换的较低概率令牌处。对于令牌插入,编辑的位置在两个令牌之间,前者具有较低后向概率并且后者具有较低前向概率。对于令牌删除,编辑的位置可以在在两个方向上具有较低概率的令牌处,或者在前者具有较低后向概率并且后者具有较低前向概率的令牌的任一令牌处。

[0097] 一些实施例在单个重写中执行多个令牌替换。一些实施例在任何特定重写中执行单个令牌替换。一些这样的实施例迭代地执行多个重写。

[0098] 图4示出了用于令牌替换的令牌序列重写流程。前向令牌概率模块41接收输入令牌序列,并使用前向SLM 42来为序列中的每个令牌产生前向概率。后向概率模块43接收输入令牌序列,并使用后向SLM 44来为序列中的每个令牌产生后向概率。编辑模块45接收前向和后向概率序列,在前向和后向方向两者中找到具有较低概率的令牌位置,并将该令牌替换为另一令牌以产生新的重写令牌序列。如图4中的一些实施例有条件地执行令牌替换,仅在至少一个令牌具有足够低的概率时才执行。

[0099] 一些实施例如图4来操作,但不是在编辑模块45中执行替换,而是执行对具有较低概率的令牌的删除。

[0100] 一些实施例如图4来操作,但不是在编辑模块45中执行替换,而是标识一对相邻的令牌,其中前者在后向方向上具有较低概率并且后者在前向方向上具有较低概率。一些这样的实施例在该对中的令牌之间插入新令牌。一些这样的实施例替换该对的第一或第二令牌。

[0101] 阈值

[0102] 一些实施例通过将概率与阈值进行比较来确定较低概率。一些实施例简单地在具有最低概率的位置处进行重写,而不使用阈值。在一些实施例中,阈值在前向和后向方向上不同。在一些实施例中,阈值取决于相邻令牌的概率。

[0103] 良好阈值水平是语料库中所有令牌的概率分布中的一个点,其允许大多数令牌通过。较低阈值导致更积极的重写。积极的重写可以纠正更惊人的令牌错误,但有将输入令牌序列重写为错误地表示讲话者的意图的新令牌序列的风险。

[0104] 令牌序列的开始处的前向令牌概率比从开始处的序列中的若干令牌更广泛地分布。类似地,令牌序列的末端处的后向令牌概率比从末端处的序列中的若干令牌更广泛地分布。一些实施例针对序列中的第一或前几个令牌使用与序列中的其他令牌不同的阈值。

[0105] 在一些实施例中,阈值取决于用户简档的属性。具体地,每个用户都有根据他们已经做出的有效表述的语料库构建的个人SLM。这样的实施例使用阈值,该阈值是通常适合于公共SLM的阈值和根据用户特定SLM的阈值的组合。这对于针对仅给出简单表述的人给出最佳精确性提升,以及针对意图做出复杂表述的人缩减较低概率令牌的重写是有用的。

[0106] 一些实施例不使用阈值,而是选择以前向、后向、或前向方向和后向方向二者中的组合中的任一项中的最低概率来替换令牌。

[0107] 构建SLM

[0108] 选择最有用的重写取决于具有准确的SLM。如果根据与要重写的表述相同类型的表述的语料库来构建,则SLM是最准确的。例如, Twitter™推文中的表述的语料库具有与来自纽约时报™报纸上的文章的表述的语料库完全不同的SLM概率。同样地,用于虚拟助理的表述的语料库通常具有与特定于天气域的表述的语料库不同的SLM概率。

[0109] 为了提高准确性,用于构建SLM的语料库应仅包含正确的表述-即适当的自然语言语法可以解析的表述或者当由随机语法解析时分数较高的表述。此外,使用引用不同特定实体(例如地名)的表述的语料库来构建SLM创建许多较低概率序列,每个较低概率序列用于每个不同的特定实体。通过利用通用标记替换特定实体令牌,所得到的SLM对于具有被替代用于实体的标记的相同表述将具有更高的概率。为了利用在标记替代之后构建的SLM,在执行令牌序列重写之前执行相应的标记替代对于令牌重写系统是有用的。

[0110] 图5A和图6A示出了用于示例的目的的较小语料库。每个语料库具有五个表述，每个表述具有两个令牌。图5B和图6B示出了从语料库构建的前向SLM。每个行对应于最近的令牌的序列（仅支持1令牌序列）。列对应于序列后面的令牌。单元格（cell）指示概率。例如，在图5B的SLM中，在令牌“who（谁）”之后有1/2的机会下一令牌是“is（是）”，并且在图6B的SLM中，在令牌“是”之后有1/4的机会下一令牌是“哪里”。图5C和图6C示出了从语料库构建的后向SLM。每个行对应于后续令牌的序列（仅支持1令牌序列）。列对应于序列之前的令牌。例如，在图5C的SLM中，对于令牌“is（是）”有1/4的机会前面的令牌是“who（谁）”，并且在图6C的SLM中，对于令牌“哪里”有1/2的机会下一令牌为“是”。注意，被表示为矩阵的前向和后向SLM不是彼此的转置。

[0111] 图7示出了用于解释天气域中的英文表述的简单自然语言语法的示例。图8示出了用于解释天气域中的普通话表述的简单自然语言语法的示例。这些语法仅是简单的说明性示例。它们定义太少措词而无法支持人们表述关于天气的查询的多种方式。

[0112] 注意，语法使用标记<date（日期）>和<place（地方）>来表示日期术语和地名。能够执行实体标记的适当的自然语言处理系统检测这种实体，并针对表述中的任何特定日期术语或地名应用该语法规则。

[0113] 图9示出了从具有大量表述的语料库构建的前向SLM，每个表述匹配图7的语法中的措辞。每个行对应于最近的令牌的序列（仅支持1令牌序列）。其他实施例使用一个以上的令牌的序列来更准确地预测下一令牌。列对应于序列后面的令牌。单元格指示概率。符号<s>和</s>指示令牌序列的开头和结尾。例如，序列的开始处的第一个词为“what（什么）”的概率是0.69，并且序列的结尾处的最后一个词为日期实体的概率是0.67。词“what（什么）”在词“the”之后的概率是0.01，表明人们很少说“the what”。

[0114] 图10示出了从具有大量表述的语料库构建的后向SLM，每个表述匹配图7的语法中的措辞。每个行对应于后续令牌的序列（仅支持1令牌序列）。列对应于序列之前的标记。例如，它示出以词“weather（天气）”结尾的令牌序列的概率是0.128。词“will（将）”在词“what（什么）”之后的概率是0.51，表明人们说“what will（什么将）”是相对常见的。

[0115] 图5、图6、图9和图10的示例示出了2字母组（也称为双字母组）SLM。也就是说，它们示出每个令牌与特定的一个相邻令牌相邻的概率。一些SLM使用3字母组。也就是说，它们示出每个令牌与2个相邻令牌的特定序列相邻的概率。词“be（是）”在词“will（将）”之后的概率不同于词“be（是）”在2个词“what will（什么将）”之后的概率。一些SLM使用4字母组或更高阶的n字母组。

[0116] 一些实施例从与各种主题相关的表述的不同语料库构建SLM。这对于支持各种表述是有用的，例如特定人类语言中的所有可能的表述。

[0117] 一些实施例从与仅一个域相关的表述的语料库构建SLM。这样的语料库具有非常小的表述的多样性，这增强了用于域中的表述的n字母组之间的概率差异。结果，重写更准确地选择对于该域最不可能正确的令牌。从各种表述构建的语料库在选择用于编辑的最佳令牌时不太准确，因此，将需要更多的重写选择以针对特定域实现相同的精确性。

[0118] 一些实施例过滤域特定语法可以解释为有效或者域特定随机语法以高于阈值的分数解析的表述中的大量表述。这种实施例仅使用其SLM训练语料库的通过表述。一些这样的实施例进一步将语料库过滤为仅在一个或多个其他语法中不具有较高解释分数的表述。

一些实施例在执行解释时考虑对话状态历史、诸如位置和用户简档之类的其他上下文信息、或两者。

[0119] 通过使用一组语法规则可以解析的一组表述,可以在没有表述的语料库的情况下构建SLM。使用随机语法作为表述生成器可以产生这样的SLM,其中语法规则的权重确定令牌序列的概率。仅当语法设计者能够预期人们可以做出与语法域相关的表述的所有方式并估计语法权重时,仅从语法构建SLM才能成功。实际上,当从表述的语料库构建时,SLM概率更准确。从语法构建SLM通常仅用于引导系统,该系统将收集大量表述,它们稍后可以用来构建更准确的SLM。还可以组合单独仅从语法产生的SLM概率以及通过从适当的语料库进行构建而确定的概率。

[0120] 标记

[0121] 图11示出了用于令牌替换的令牌序列重写流程。该流程与图4中的流程相同,但标记模块116接收输入令牌序列,并用标记代替令牌以创建用于概率模块41和43的令牌序列。

[0122] “标记”是对于一类令牌通用的令牌。例如,<city(城市)>标记可以代替任何令牌“London(伦敦)”、“Beijing(北京)”和令牌序列“New(纽)”“York(约)”。<time(时间)>标记可以代替任何标记序列“1:23”、“1:23pm”、“01:23pm”、“13:23”和“下午一点二十三”。

[0123] 一些实施例使用域特定标记。例如,天气域可以具有标记<weather condition(天气条件)>。<weather condition(天气条件)>标记可以代替令牌“sunny(晴朗)”、“cloudy(多云)”、“foggy(有雾)”、“rainy(下雨)”、“windy(刮风)”、“hot(热)”和“cold(冷)”。实施例首先接受被奇怪措辞的令牌序列“it sunny or rainy will be(它晴朗或下雨将是)”;用标记实例<weather condition 1(天气条件1)>代替令牌“sunny(晴朗)”,并用将标记实例<weather condition 2(天气条件2)>代替令牌“rainy(下雨)”,同时存储标记<weather condition(天气条件)>的每个特定实例与其(一个或多个)原始令牌的关联;重写所标记的令牌序列以变为“will it be<weather condition 1>or<weather condition 2>(它将是<天气条件1>或<天气条件2>)”;并最终用其原始特定令牌来替换<天气条件>标记实例以产生经重写的令牌序列“will it be sunny or rainy(它将是晴朗或下雨)”。天气域语法不能解析输入令牌序列“it sunny or rainy will be(它晴朗或下雨将是)”,但可以解析所重写的令牌序列“will it be sunny or rainy(它将是晴朗或下雨)”。

[0124] 具有域特定语法的一些实施例将不在语法中的所有令牌标记为<unknown(未知)>。它们使用未知令牌的标记来执行重写,然后在重写之后,用其原始令牌替换未知令牌标记。用于SMS文本消息域的这种实施例在其语法中具有令牌“send(发送)”、“text(文本)”和“message(消息)”。当接收到令牌序列“send message cow horse and chicken as a text message(发送消息牛马和鸡作为文本消息)”时,该实施例将令牌序列标记为“send message<unknown 1><unknown 2>and<unknown 3>as a test message(发送消息<未知1><未知2>和<未知3>作为文本消息)”。实施例执行重写以“send<unknown 1><unknown 2>and<unknown 3>as a test message(发送<未知1><未知2>和<未知3>作为文本消息)”,并用其原始令牌替换标记以产生经重写的令牌序列“send cow horse and chicken as a text message(将牛马和鸡作为文本消息发送)”。

[0125] 由于英文表述的通用用法,这样的实施例经常接收具有“send<unknown>(发送<未知>)”的令牌序列并很少接收具有“message<unknown>(消息<未知>)”的令牌序列,它在令

牌“message(消息)”第一次出现时执行删除编辑。

[0126] 一些实施例使用从不同语料库构建的SLM。一些这样的实施例利用<low frequency(较低频率)>标记来标记以较低频率在语料库中出现的所有词。标记所有较低频率词增加了较高频率词的概率准确性。

[0127] 一些实施例执行语料库中的语音检测和令牌加标签的部分的过程。对于可以用作语音的不同部分的令牌，例如令牌“fast(快)”可以用作形容词(快速移动)、动词(不吃)和名词(不吃的时间段)，这些实施例在SLM中包括不同的条目，用于在用作语音的不同部分时的令牌。在计算用于重写的令牌序列上的令牌概率之前，这些实施例对令牌序列执行部分句法分析以标识序列内的每个令牌的语音的可能部分。这样做使得SLM中加标签的令牌概率更准确。

[0128] 一些实施例对语料库中的表述和输入令牌序列两者执行同义词替换。例如，这种实施例将令牌“fast(快)”替换为令牌“rapid(迅速)”、“quick(敏捷)”和“speedy(快速)”中的任一个。因此，对于给定的不同表述的语料库，SLM统计更准确。

[0129] 一些实施例对语料库中的表述和输入令牌序列两者执行同音异义词替换。例如，这种实施例用单个令牌来替换英文令牌“to”、“too”和“two”。

[0130] 一些域特定实施例具有它不能重写的域特定保留令牌的集合。例如，在用于SMS消息的域中，这样的实施例包括令牌“text(文本)”、“message(消息)”和“send(发送)”，这些令牌不能作为编辑操作来替换、删除、或插入。一些域语法格式允许定义保留关键字。

[0131] 从概率检测错误类型

[0132] 在前向和后向方向中的每个方向具有低概率的令牌的位置的共址位或邻近给出了关于可能的错误的类型(错误令牌、额外令牌、丢失令牌和重复令牌)，并因此用于重写令牌序列的编辑的最佳类型(令牌插入、令牌删除和令牌替换)的线索。

[0133] 图12A示出了用于其中后向和前向概率都是高(H)的表述式的英文词令牌的序列。图13A示出了用于其中后向和前向概率都是高(大)的表述式的普通话汉字令牌的序列。

[0134] 图12B是示出具有错误令牌错误的表述的语文词，并且图13B是示出具有错误令牌错误的表述的普通话汉字(例如，图12B中的“pin”而不是“in”以及图13B中的“跑的”而不是“的”)。错误令牌之后的所有令牌在后向方向上具有高概率，并且错误令牌之前的所有令牌在前向方向上具有高概率。错误令牌在后向和前向方向中的每个方向上具有低概率。

[0135] 图12C是示出具有额外令牌错误的表述的语文词，并且图13C是示出具有额外令牌错误的表述的普通话汉字。额外令牌之后的所有令牌在后向方向上具有高概率，并且额外令牌之前的所有令牌在前向方向上具有高概率。额外令牌在后向和前向方向中的每个方向上具有低概率。

[0136] 图12D是示出具有丢失令牌错误的表述的语文词，并且图13D是示出具有丢失令牌错误的表述的普通话汉字(例如，图12D中的“weather(天气)”和“Hawaii(夏威夷)”之间缺少令牌“in”并且在图13D中的“的”和“怎么样”之间缺少令牌“天气”)。紧接在丢失令牌之前的令牌在后向方向上具有低概率，但丢失令牌的位置之后的所有令牌在后向方向上具有高概率。丢失令牌之前的所有令牌在前向方向上具有高概率，但紧接在丢失令牌的位置之后的令牌在前向方向上具有低概率。

[0137] 图12E是示出具有重复令牌错误的表述的语文词，并且图13E是示出具有重复令牌

错误的表述的普通话汉字。第一重复令牌在后向方向上具有低概率,但第二和所有后续令牌在后向方向上具有高概率。第一重复令牌之前的所有令牌并且包括第一重复令牌在前向方向上具有高概率,但第二重复令牌在前向方向上具有低概率。

[0138] 因此,在后向和前向方向中的每个方向上具有低概率的令牌是可疑的。它指示其中删除或替换编辑可能适当的位置。其中前者在后向方向上具有低概率并且后者在前向方向上具有低概率的一对相邻的不同令牌是一对可疑令牌。它指示插入编辑可能适当的位置。其中前者在后向方向上具有低概率并且后者在前向方向上具有低概率的一对相邻的相同令牌是一对可疑令牌。它指示其中删除相同令牌之一是可能适当的位置。

[0139] 语音识别自然地是概率性的。令牌序列源自口头音素的一个或多个推测序列,其自身源自数字处理的声学信号样本的帧。一些实施例产生表示对音素序列推测中的每个音素的识别的准确性的置信度的分数。在一些实施例中,处理样本的帧产生可以对音素推测的分数做出贡献的值。一些实施例在将音素序列令牌化为令牌序列时,计算每个令牌的分数,该分数至少部分地基于令牌内的音素的声学分数。

[0140] 一些实施例至少部分地从后向和前向SLM概率计算后向和前向概率分数,通过令牌的声学分数来加权。结果,对于质量差的捕获语音音频(例如,在某些令牌期间具有吵闹背景声音事件的语音音频),这种实施例集中在这些令牌上重写编辑。

[0141] 评分和比较重写

[0142] 分数可以以各种方式表示,例如,整数、浮点数或字母数字。在一些实施例中,分数被表示为有效性的布尔值(有效/无效)。

[0143] 各种实施例通过标识后向概率低于阈值、前向概率低于阈值、或后向和前向概率的组合低于阈值的位置来创建重写。一些这样的实施例各自通过在概率分数低于阈值的位置处进行单个编辑来首先创建多个替代重写,然后通过在第一组重写令牌序列内的第二位置处进行编辑来进行进一步重写,以创建进一步的重写。结果是创建重写树或重写晶格(lattice),其解释了替换A后跟替换B(在不同位置)具有与B后跟A相同的结果。对于具有高阈值或质量差语音音频的实例的实施例,重写次数对于可用的系统资源可能变得过大。为了适应这样的限制,各种实施例为了找到最佳重写而基于对诸如最佳整体重写令牌序列分数之类的价值的测量,使用广度优先搜索算法、深度优先搜索算法或最佳优先算法来构建重写树或晶格。

[0144] 各种实施例通过在如上所述的相同或相邻令牌上的具有后向方向上的最低概率、在前向方向上的最低概率、或在后向和前向概率的最低组合的位置处执行编辑来创建重写。在各种实施例中,组合是诸如乘积、求和或某些等同物之类的功能。一些这样的实施例创建多个重写以产生重写推测的集合,并计算每个重写推测的重写分数以便指导选择最佳(最可能正确的)重写。一些实施例通过在最低和第二低位置中的每一者处进行编辑并通过对第一对重写的结果令牌序列进行编辑以进行进位,来创建多个重写。

[0145] 图14示出了用于令牌替换的令牌序列重写流程。该流程与图4中的流程相同,除了令牌替换模块145从前向SLM和后向SLM中的每一者读取最可能的令牌列表。对于先前的N个令牌,针对每个位置的前向最可能令牌的列表包括一组最可能的令牌及其概率。对于后续N个令牌,针对每个位置的后向最可能令牌的列表包括一组最可能的令牌及其概率。为了选择用于插入或替换的令牌,令牌替换模块145搜索在两个列表中存在的令牌(列表的交集)

并基于每个列表中的令牌的概率选择新令牌。

[0146] 如果未找到在两个列表中存在的令牌,则一些实施例将下一最可能词从SLM添加到列表。如果未找到在两个列表中存在的令牌,则各种实施例从前向列表中选择最高概率令牌、从后向列表中选择最高概率令牌、或者跳过重写令牌序列。

[0147] 图15A是示出其中中间令牌不正确的3令牌序列的英文词,并且图16A是示出其中中间令牌不正确的3令牌序列的普通话汉字。图15B和图16B示出了来自2字母组前向SLM的5个令牌的列表及其概率(每个令牌概率从一个先前令牌计算)。图15C和图16C示出了来自2字母组后向SLM的5个令牌的列表以及其概率。这些列表共享两个公共令牌。第一公共令牌是每个列表中第二高概率。第二公共令牌在前向列表中是第五个以及后向列表中是第三个。由于第一公共令牌在两个列表中都比较常见,因此它是用于令牌替换编辑的最佳选择。

[0148] 图17A是示出其中中间令牌不正确的5令牌序列的英文词,并且图18A是示出其中中间令牌不正确的5令牌序列的普通话汉字。图17B和图18B示出了来自3字母组前向SLM的5个令牌的列表及其概率(每个令牌概率从两个先前令牌计算)。图17C和图18C示出了来自3字母组后向SLM的5个令牌的列表及其概率。前向和后向SLM不必使用相同的n字母组长度。这些列表共享两个公共令牌。第一公共令牌是前向列表中的第三高概率和后向列表中的第二高概率。第二公共令牌在前向列表中的第五个以及后向列表中的第一个。由于两个令牌对于两个列表是公共的,但公共令牌的概率的顺序在两个列表中是不同的,因此有必要计算哪个是用于令牌替换编辑的更好选择。

[0149] 一些实施例将每个列表中的公共令牌的概率相加,并选择概率总和最高的令牌。一些实施例将每个列表中的公共令牌的概率相乘,并选择概率乘积最高的令牌。一些实施例将每个列表中的公共令牌的概率的对数相加,并选择概率的对数的总和最高的令牌,其给出与乘法相同的相对结果顺序,但具有不同的绝对值。

[0150] 图17D和图18D示出了针对每个公共令牌的概率相加的计算、概率相乘的计算、以及概率的对数相加的计算。相加方法找到了用于第二令牌的更高重写分数,但对数的相乘和求和方法找到了用于第一令牌的更高重写分数。对于一些实施例,对数相乘或相加方法是优选的。它具有有利于在两个方向上更合理的令牌的效果,而概率相加的方法有利于在一个方向上合理但在另一方向上不合理的令牌。

[0151] 检测若干同音异义词中的哪一个对于转录一些给定语音是正确的挑战性的问题。一些实施例计算重写分数,该重写分数对作为输入令牌序列中的可疑令牌的同音异义词的令牌的概率应用有利权重。结果,这些实施例在选择最可能在其周围令牌的上下文中的同音异义词时更加准确。

[0152] 如果令牌最近在对话中已经说过,则它们具有高得多的概率。一些实施例缓存最近说过的令牌。一些这样的实施例仅对具有高声学分数的令牌这样做,以便最小化促使错误转录的正反馈回路的可能性。一些实施例缓存由对话中的多个伙伴最近说出的令牌,因为由一个伙伴说出的令牌更有可能被另一伙伴说出。

[0153] 图19示出了用于令牌替换的令牌序列重写流程。该流程与图14中的流程相同,但令牌替换模块195产生重写令牌序列和相应重写分数两者。在一些实施例中,重写分数是前向和后向方向中的每个方向中的每个令牌的概率的乘积。在一些实施例中,重写分数是前向和后向方向中的每个方向中的每个令牌的概率的乘积,每个令牌由声学分数加权。不使

用实体标记的一些实施例使得可以替换实体值,通过特定实体相对于环境参数值(例如,特定人的位置或存在)的接近度来加权重写分数。

[0154] 图20示出了用于令牌替换的令牌序列重写流程。该流程与图19中的流程相同,但令牌替换模块205从不同语料库SLM 207中读取令牌条件概率。前向SLM 42和后向SLM 44是根据特定于一个谈话域的表述构建的。不同语料库SLM 207是根据跨许多域的表述构建的,例如,根据仅受限于它们使用单一人类语言的事实的不同表述源。一些实施例基于不同语料库SLM 207中的每个令牌的相应条件概率来缩放输入令牌序列中的每个令牌的域特定条件概率。一些实施例基于不同语料库SLM 207中的每个令牌的相应条件概率来缩放最可能的令牌的列表中的每个替换令牌的域特定条件概率。一些实施例使用单个方向不同语料库SLM,并且一些实施例针对每个方向使用不同的语料库SLM。

[0155] 图21示出了用于令牌替换的令牌序列重写流程。该流程与图19中的流程相同,但令牌替换模块215产生一组多个重写令牌序列和与每个重写令牌序列相关联的重写分数。选择模块218使用重写分数来选择多个替代重写中的最佳重写令牌序列。一些替代方案可能在同一位置具有不同的编辑。一些替代方案可能在不同位置具有编辑。

[0156] 进行解析以计算解析分数

[0157] 一些实施例使用超过SLM概率和加权分数的更高级处理来确定最佳重写。图22示出了使用句法分析的实施例。所示实施例与图14相同,除了令牌替换模块225产生多个重写并选择分数最高的一个重写以产生新的重写令牌序。此外,令牌替换模块225使用句法分析输入来加权重写分数。

[0158] 句法分析模块229接收输入令牌序列,并确定每个令牌可以是语音的那些部分以及令牌用作语音的每个部分的频率。令牌替换模块225应用关于语音的不同部分的令牌的布置的合法性的句法规则。令牌替换模块225将其对新令牌的选择限制为仅在相邻令牌的上下文中在句法上合法的令牌。对于每个重写,令牌替换模块225根据语音的特定部分的令牌的布置的合法性来对重写令牌序列进行评分。

[0159] 一些实施例使用自然语言解析器来根据语法解析新的重写令牌序列以产生解析分数。图23示出了用于令牌替换的令牌序列重写流程。该流程与图4中的流程相同。然而,自然语言解析器236接收从令牌替换模块45输出的新的重写令牌序列。自然语言解析器根据由自然语言语法237定义的语法规则来解析新的重写令牌序列。

[0160] 一些实施例使用产生布尔值分数的简单语法。也就是说,语法是否能够解析令牌序列。一些实施例使用指定用于解析整数或浮点分数的随机语法。

[0161] 在一些实施例中,语法规则定义不同措辞的似然。语法能够解析两个不同的表述以导出同一含义(产生相同的结果解释数据结构),但具有对哪个是更常用的表述、或表述中的短语或词的指示。一些实施例通过考虑表述的措辞及其在讲话者中的使用似然来计算所解析的表述的分数。一些这样的实施例通过将语法规则内的短语或词的权重相乘来计算这样的解析分数。一些实施例将句法解析分数与语义分数相组合以确定解释的分数。

[0162] 图24示出了根据天气域随机语法具有同一含义的两个英文表述。一个使用词“for”并且另一个使用词“in”。由于更多的讲话者更喜欢以“in”来措辞,因此它分数更高。图25示出了根据天气域随机语法具有同一含义的两个普通话中文表述。一个使用词“里”并且另一个使用词“儿”。由于更多的讲话者更喜欢以“儿”来措辞,因此它分数更高。

[0163] 具有语法的一些实施例维持最近对话状态历史信息的缓存。例如,会话状态可以包括自然语言语法规则内的实体时隙的语义值。“New York(纽约)”是响应于表述“what is the weather in X(X的天气怎么样)”(其中,X是地方)的天气相关语法规则的“place(地方)”时隙内的“place(地方)”类型的实体的语义值。因此,在表述“what is the weather in New York(纽约的天气怎么样)”之后,存储会话状态的实施例将具有作为最近引用的地方的“New York(纽约)”的记录。

[0164] 会话状态还可以包括先前被询问的问题的类型。它还可以包括实体的相关参数的标识符。它还可以包括关于限定符和信息过滤信息的信息。一些实施例根据在会话状态的上下文中的令牌的语义显著性(其可以表示为解析分数)来加权他们的解析分数。

[0165] 一些实施例尝试根据多个域特定语法来解析重写令牌序列。如果至少一个域语法以可接受的分数解析重写令牌序列,则重写成功。

[0166] 图26示出了用于令牌替换的令牌序列重写流程。该流程与图19中的流程相同,但自然语言解析器266接收该组重写令牌序列和相关联的分数,尝试根据语法237来解析每个令牌序列以产生解析分数,并选择具有解析分数和重写分数的最佳组合的重写令牌序列作为其输出。一些这样的实施例根据多个语法进行解析。因此,一些这样的实施例产生实际上是声学分数、根据SLM的重写分数和语法解析分数的组合的输出令牌序列。

[0167] 迭代地创建重写的集合

[0168] 一些实施例在找到成功重写时,完成它们的处理并提供成功重写令牌序列作为输出。一些实施例产生输入令牌序列的可能的重写的列表。这可以通过在同一可疑位置进行不同的编辑、在不同的可疑位置进行编辑、或两者。

[0169] 一些实施例为了校正导致多个额外词、丢失词或重复词的多个错误词或错误,通过重写流程的附加迭代来馈送重写令牌序列。迭代可以通过创建简单的重写列表、或者基于重写的重写分数、语法解析分数或两者的组合选择重写来进行。选择可以通过尝试深度优先、广度优先或最佳分数优先搜索树算法来工作。

[0170] 图1示出了迭代地重写新令牌序列的过程。一些实施例在重写阶段12之后,根据新令牌序列的重写分数来重新处理新令牌序列。一些实施例在解析阶段14之后,根据新重写令牌序列的解析分数来重新处理新重写令牌序列。

[0171] 连续转录

[0172] 令牌序列的重写在连续转录应用中可能是有用的,例如,直播电视广播或听写的自动加字幕。

[0173] 双向概率重写在存在从其计算后向SLM概率的端点时是有用的。这利用令牌缓冲器是可能的,该令牌缓冲器使得具有足够高令牌化推测分数的序列中的最近的N个令牌是可用的,其中,N是在后向SLM概率计算中使用的令牌的数目。这需要至少N个令牌时间的转录延迟。

[0174] 英文讲话者倾向于每分钟说出150个词令牌。普通话中文讲话者倾向于每分钟说出大约300个字符。由于许多普通话令牌包含两个或三个字符,因此讲话的令牌速率是每分钟大约150个令牌。也就是说,每秒2.5令牌。因此,使用2字母组后向SLM的双向重写将需要至少1秒的缓冲延迟。

[0175] 对于连续语音应用,为了将语法解析应用于双向概率重写,需要完整表述或至少

完整短语的所有令牌,以便确定可以解析序列。这种应用需要在重写和选择上游的短语边界检测算法。各种这样的算法在本领域中是已知的。一些短语边界检测算法是基于语法的。通过检测讲话嗓音的变化可以获益。一些益处来自检测相应的视频场景变化。足够长度的暂停也可以强加保证。

[0176] 由于人们倾向于以每分钟150个令牌的令牌速率讲出具有多达25个令牌的表述,因此需要缓冲器和长达10秒的潜在延迟。尽管这对于电影或预先录制的电视节目是无法接受的字幕延迟,但对于电视新闻馈送和一些体育广播是合理的。通过一次处理完整的表述,字幕显示将逐个表述添加文本,这在电影和预先录制的电视节目中是常见的,而不是在直播电视广播中常见的传统逐个词显示。

[0177] 一些实施例可以将视频流从实时广播延迟,以使其与通过表述分离和基于语法的评分和选择的重写所需的延迟同步。

[0178] 一些实施例使用诸如音素推测、令牌化推测和短语边界推测之类的推测来执行语音转录的过程。这种实施例构建推测树,其具有树的每个分支的分数。一些这样的实施例对各种令牌化表述推测中的每一个执行双向概率自然语言重写、计算得到的重写解析分数、并根据每个分支的最佳重写解析分数来选择最佳树分支。

[0179] 历史缓存和反向缓存

[0180] 一些实施例通过使用一种或多种缓存技术来提高性能。一些实施例至少针对有限持续时间来缓存具有强声学推测分数的令牌。当从前向和后向SLM最高概率令牌的列表中选择令牌时,如果令牌存在于缓存中,则这些实施例将令牌的概率分数提高。这利用了会话中的词的时间局部性,具有提高识别准确性的效果。各种实施例在实体标记之前缓存词以改进对实体名称的识别,或在实体标记之后缓存词以便于与独立于实体的重写相集成。

[0181] 一些实施例缓存输入令牌序列到新的重写令牌序列的映射。图27示出了与图4中的流程相同的令牌序列重写流程。然而,序列检查模块278接收输入令牌序列。序列检查模块继续从缓存279进行读取,该缓存279包括输入令牌序列到新的重写令牌序列的映射。如果序列检查模块278在缓存中找到输入令牌序列,则序列检查模块278将相关联的新的重写令牌序列提供给令牌替换模块275。在从序列检查模块278接收到新的重写令牌序列时,令牌替换模块275直接输出新的重写令牌序列。否则,令牌替换模块275执行适当的重写分析,包括使用高概率SLM令牌的列表的重写。结果,流程不需要执行处理密集型树搜索以获得过去接收的令牌序列的最佳重写。

[0182] 图28示出了令牌重写缓存的七个条目。每个输入令牌序列(具有标记实体)仅与一个新的重写令牌序列相关联。

[0183] 一些实施例使用反向缓存,其可以将新的重写令牌序列与多个输入令牌序列相关联。图29示出了反向令牌重写缓存的五个条目。最后一个缓存条目是通常从五个不同的输入令牌序列中的每一个所创建的重写。在括号中示出了在缓存中创建最后一个新令牌序列的每个输入令牌序列的频率。一些实施例分析缓存内容以找到被重写到特定的新的重写令牌序列的最频繁的输入令牌序列。这对于语法开发者发现常用替代措辞以便改善它们的语法并将它们添加到语法规则、或创建不同的语法规则来处理针对其输入令牌序列不匹配表述的意图的重写是有用的。例如,如果“will the weather be good(天气会好吗)”被重写为“what will the weather be(天气怎么样)”,则系统工程师可以创建支持前一措辞新的

语法规则,并基于预测天气状况以是或否来响应。

[0184] 一些实施例包括自动语法诱导模块,其使用反向重写缓存来通过推断新语法规则或扩展现有语法规则来改善覆盖范围。

[0185] 通过调整特定令牌的权重或改变规则来改进随机语法,提升了系统操作员生成准确SLM的能力。此外,改进随机语法改进了语料库的过滤,从而构建准确的域特定SLM。

[0186] 其他考虑因素

[0187] 本文示出和描述的实施例是许多可选特征的示例。利用所示出和所描述特征的不同组合,许多实施例是可能的。所示出和描述的示例使用英文和普通话中文语言。类似地,各种实施例可用于其他语言或语言组合。所示出和描述的示例使用天气域。各种实施例类似地可用于其他域或域的组合作。所示出和描述的示例示出了用于令牌替换的模块。类似布置的模块可以执行令牌删除和令牌插入。

[0188] 图30示出了使用客户端-服务器类型的架构的虚拟助理的实施例。用户301向智能电话302说出表述。智能电话302在运行虚拟助理应用时表现为用于基于服务器的虚拟助理的远程客户端设备。智能电话302发送来自用户301的表述作为对虚拟助理的查询。客户端设备302通过计算机网络303传送查询音频以将表述音频发送到服务器304。服务器304使用计算机处理器来执行存储在非暂态计算机可读介质上的代码。通过处理指令,计算机处理器对表述音频执行自动语音识别305以产生令牌序列推测。计算机处理器使用令牌序列推测作为输入来执行重写306以产生新的重写令牌序列推测。计算机处理器尝试使用自然语言处理(NLP) 307来解析新的重写令牌序列推测。重写使用NLP解析307的结果来决定是否执行另一重写或输出来自服务器304的新的重写令牌序列。在选择具有可接受的高分数的解析之后,计算机处理器处理308所选择的解释以产生期望的结果。例如,处理可以在服务器304内本地执行,或者通过使用第三方系统的服务来执行。例如,对于诸如“what is the weather? (天气怎么样?)”之类的查询,处理308将包括电话的当前位置的应用编程接口(API) 请求发送到第三方天气服务。这样的天气信息服务以天气状况的文本描述来响应API请求。许多其他信息来源和服务类型对于虚拟助理是公知的。服务器304使用处理结果(例如对API请求的文本响应)来为用户301生成口头响应作为文本。计算机使用文本到语音(TTS) 309处理口头响应文本以产生针对用户的音频响应。服务器304通过网络303将音频发送到客户端302,该客户端302通过扬声器将其输出给用户。许多其他变体在其他实施例中是可能的。例如,在其他实施例中,服务器304代替音频响应或者在音频响应之外还向客户端302提供文本响应。

[0189] 图31A示出了旋转盘311,其是根据一些实施例的非暂态计算机可读介质。

[0190] 图31B示出了闪存芯片312,其是根据一些实施例的非暂态计算机可读介质。

[0191] 图31C示出了封装片上系统器件313的底侧,其具有用于焊接到印刷电路板的球栅阵列。片上系统313是根据一些实施例的计算机处理器。

[0192] 图31D示出了片上系统器件313的顶侧。

[0193] 图32示出了根据一些实施例的片上系统芯片313内的功能组件的框图。计算机处理器核321和图形处理核322通过经由互连323与临时数据存储装置的随机存取存储器(RAM) (通过RAM接口324)、闪存RAM(通过闪存接口325)以及用于与服务器通信的网络接口328来执行处理。显示接口326向用户提供对显示信息的访问,并且I/O接口327提供接收用

户输入并向用户提供输出。

[0194] 图33示出了根据一些实施例的基于机架的多处理器服务器303。

[0195] 图34示出了根据一些实施例的服务器304内的功能组件的框图。计算机处理器341和图形处理器342通过经由互连343与临时数据存储装置的RAM 344进行通信以及用于与其他服务器和设备通信的网络接口345来执行处理。

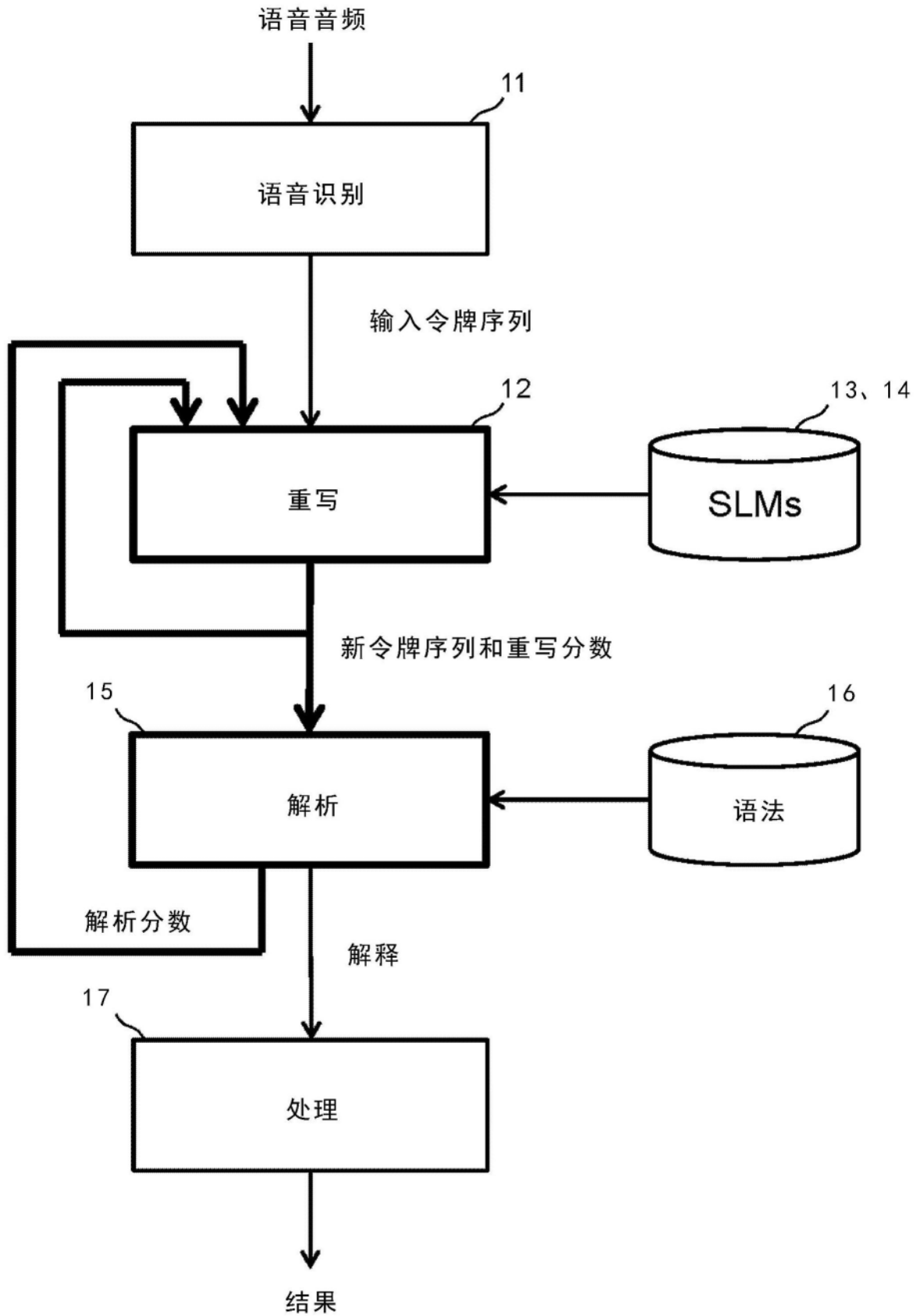


图1

不正确词语

will there be andy rain today

无关词语

will there be like rain today

will there the be rain today

中断

is it going to ____

____ going to rain today

口吃/重复

what what is the weather

what how is the weather

扩展/冗长

do we have rain for today

背景语音

it is working is it going to rain
today

不支持措辞

today will it rain

图2

不正确字符

今天活下雨吗

无关字符

今天会去下雨吗

今天的的会下雨吗

中断

今天会下——

——雨吗

口吃/重复

今天天气怎么怎么样

今天天气什么怎么样

扩展/冗长

我们今天是不是要会有雨

背景语音

他不会今天会下雨吗

不支持措辞

下雨今天会吗

图3

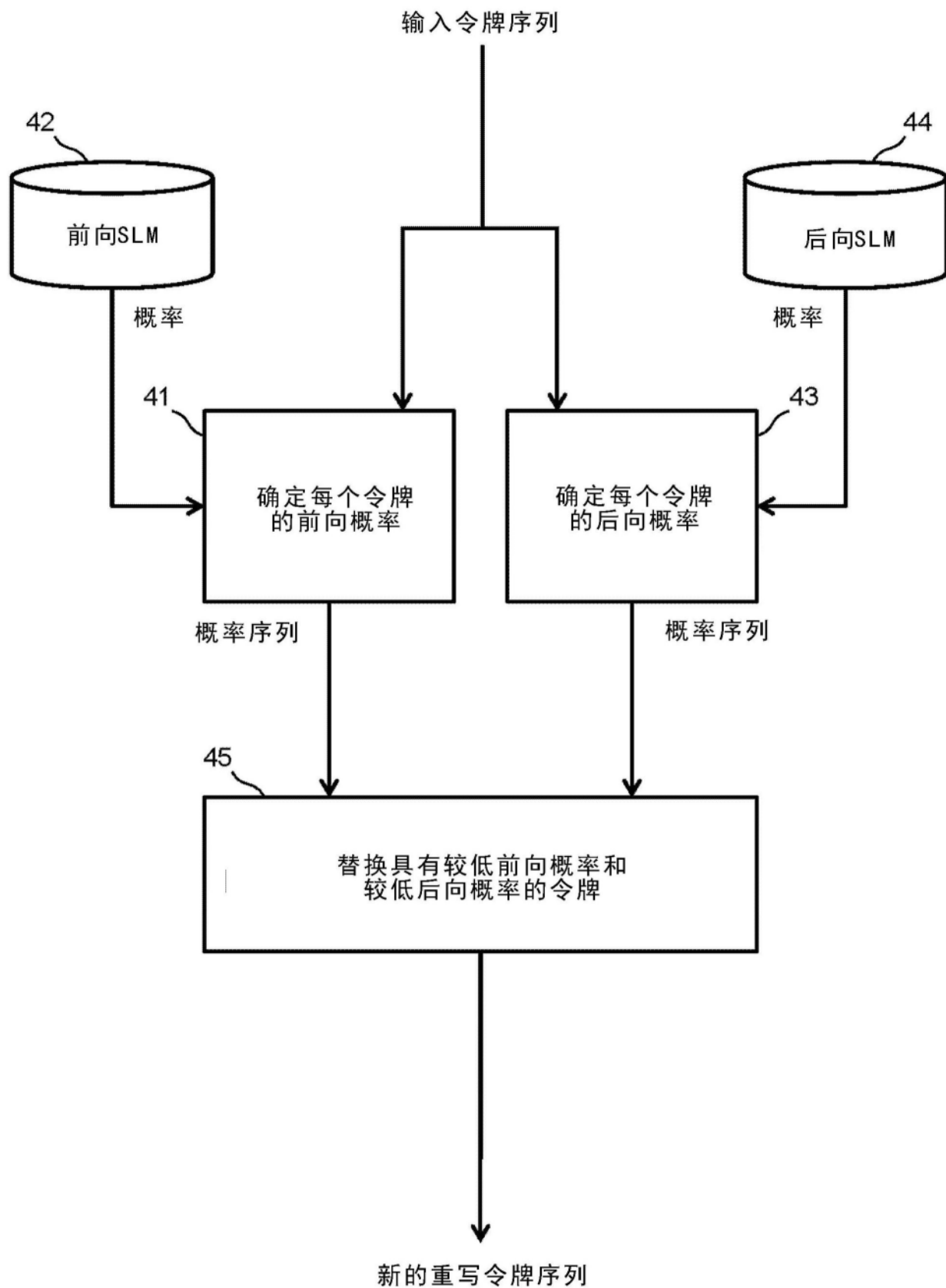


图4

语料库
what is
who is
when is
what is
who did

图5A

前向SLM概率

	is	did
who	$\frac{1}{2}$	$\frac{1}{2}$
what	1	0
when	1	0

图5B

后向SLM概率

	who	what	when
is	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
did	1	0	0

图5C

语料库
是什么
是谁
在哪里
是什么
去哪里

图6A

前向SLM概率

	哪里	什么	谁
是	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
去	1	0	0

图6B

前向SLM概率

	是	去
哪里	$\frac{1}{2}$	$\frac{1}{2}$
什么	1	0
谁	1	0

图6C

语法接受

what is the weather
what will the weather be on <date>
what is the weather going to be <date>
what is the weather in <place>
what is the weather at <place>

图7

语法接受

天气怎么样
<日期>的天气如何
<日期>会是什么天气
<地点>的天气如何
在<地点>是什么天气

图8

	what	is	the	weather	will	be	on	<date>	going	to	in	<place>	at	</s>
<s>	0.69	0.02	0.02	0.041	0.02	0.02	0.02	0.019	0.02	0.02	0.02	0.018	0.02	0.03
what	0.02	0.48	0.02	0.022	0.19	0.02	0.03	0.023	0.02	0.02	0.02	0.022	0.02	0.04
is	0.02	0.02	0.73	0.017	0.02	0.02	0.02	0.018	0.02	0.02	0.02	0.017	0.02	0.03
the	0.01	0.01	0.01	0.825	0.01	0.01	0.01	0.011	0.01	0.01	0.01	0.011	0.01	0.02
weather	0.03	0.03	0.03	0.025	0.03	0.17	0.07	0.026	0.17	0.03	0.12	0.025	0.12	0.08
will	0.03	0.03	0.51	0.03	0.03	0.03	0.04	0.032	0.03	0.03	0.03	0.03	0.03	0.06
be	0.05	0.05	0.05	0.045	0.05	0.05	0.2	0.048	0.05	0.05	0.09	0.045	0.05	0.13
on	0.02	0.02	0.02	0.022	0.02	0.02	0.03	0.652	0.02	0.02	0.02	0.022	0.02	0.04
<date>	0.02	0.02	0.02	0.022	0.02	0.02	0.03	0.023	0.02	0.02	0.02	0.022	0.02	0.67
going	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.032	0.03	0.51	0.03	0.03	0.03	0.06
to	0.03	0.03	0.03	0.03	0.03	0.51	0.04	0.032	0.03	0.03	0.03	0.03	0.03	0.06
in	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.032	0.03	0.03	0.03	0.51	0.03	0.06
<place>	0.02	0.02	0.02	0.022	0.02	0.02	0.23	0.023	0.02	0.02	0.02	0.022	0.02	0.47
at	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.054	0.05	0.05	0.05	0.184	0.05	0.1

图9

	what	is	the	weather	will	be	on	<date>	going	to	in	<place>	at	<s>
</s>	0.03	0.03	0.03	0.128	0.03	0.04	0.03	0.285	0.03	0.03	0.03	0.231	0.03	0.03
what	0.01	0.01	0.01	0.036	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.012	0.01	0.83
is	0.73	0.02	0.02	0.056	0.02	0.01	0.02	0.016	0.02	0.02	0.02	0.018	0.02	0.02
the	0.02	0.48	0.02	0.072	0.19	0.02	0.02	0.02	0.02	0.02	0.02	0.024	0.02	0.02
weather	0.02	0.02	0.69	0.059	0.02	0.01	0.02	0.016	0.02	0.02	0.02	0.019	0.02	0.04
will	0.51	0.03	0.03	0.1	0.03	0.02	0.03	0.028	0.03	0.03	0.03	0.033	0.03	0.03
be	0.03	0.03	0.03	0.34	0.03	0.02	0.03	0.028	0.03	0.27	0.03	0.033	0.03	0.03
on	0.03	0.03	0.03	0.18	0.03	0.22	0.03	0.026	0.03	0.03	0.03	0.259	0.03	0.03
<date>	0.02	0.02	0.02	0.072	0.02	0.02	0.65	0.02	0.02	0.02	0.02	0.024	0.02	0.02
going	0.03	0.03	0.03	0.58	0.03	0.02	0.03	0.028	0.03	0.03	0.03	0.033	0.03	0.03
to	0.03	0.03	0.03	0.1	0.03	0.02	0.03	0.028	0.51	0.03	0.03	0.033	0.03	0.03
in	0.04	0.03	0.03	0.436	0.03	0.1	0.03	0.032	0.03	0.03	0.03	0.038	0.03	0.04
<place>	0.04	0.04	0.04	0.126	0.04	0.03	0.04	0.035	0.04	0.04	0.34	0.041	0.09	0.04
at	0.02	0.01	0.01	0.585	0.01	0.01	0.01	0.014	0.01	0.01	0.01	0.017	0.01	0.02

图10

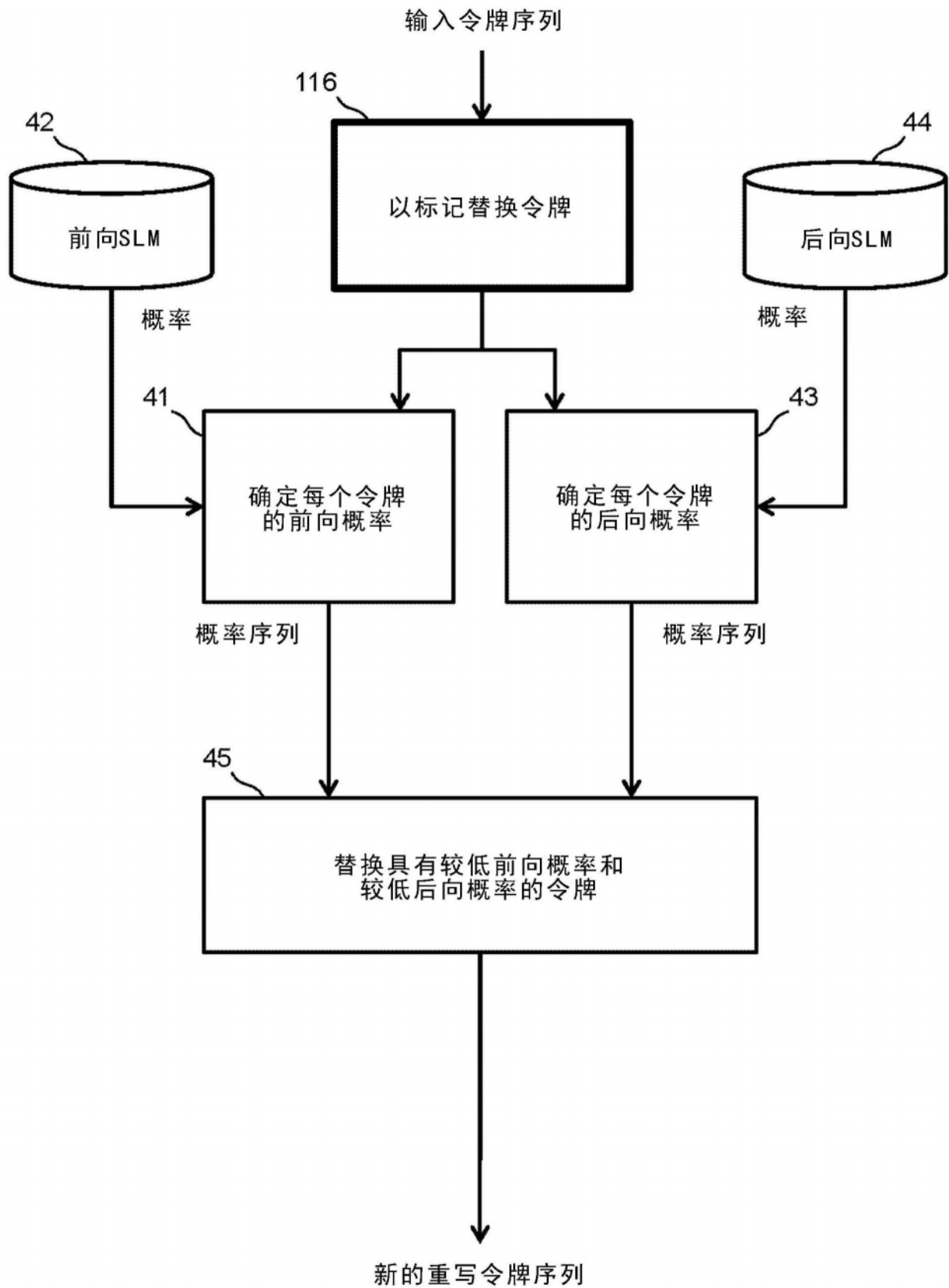


图11

what is the weather in Hawaii
 H---H---H-----H----H----H----BWD
 FWD--H---H---H-----H----H----H

图12A

↔
 what is the weather pin Hawaii
 L----H---BWD
 FWD--H---H---H-----H-----L

图12B

↔
 what is the weather under in Hawaii
 L---H----H---BWD
 FWD--H---H---H-----H-----L

图12C

← →
 what is the weather Hawaii
 L-----H---BWD
 FWD--H---H---H-----H-----L

图12D

← →
 what is the weather in in Hawaii
 L--H-----BWD
 FWD--H---H---H-----H---H--L

图12E

夏威夷 的 天气 怎么样
 大--大--大---大---BWD
 FWD---大--大--大---大

图13A

↔
 夏威夷 跑的 天气 怎么样
 小--大---大---BWD
 FWD---大---小

图13B

↔
 夏威夷 下 的 天气 怎么样
 小-大--大---大---BWD
 FWD---大---小

图13C

← →
 夏威夷 的 怎么样
 小--大-大---大---BWD
 FWD--大--小

图13D

← →
 夏威夷 的 的 天气 怎么样
 小-大--大---大---BWD
 FWD--大---大-小

图13E

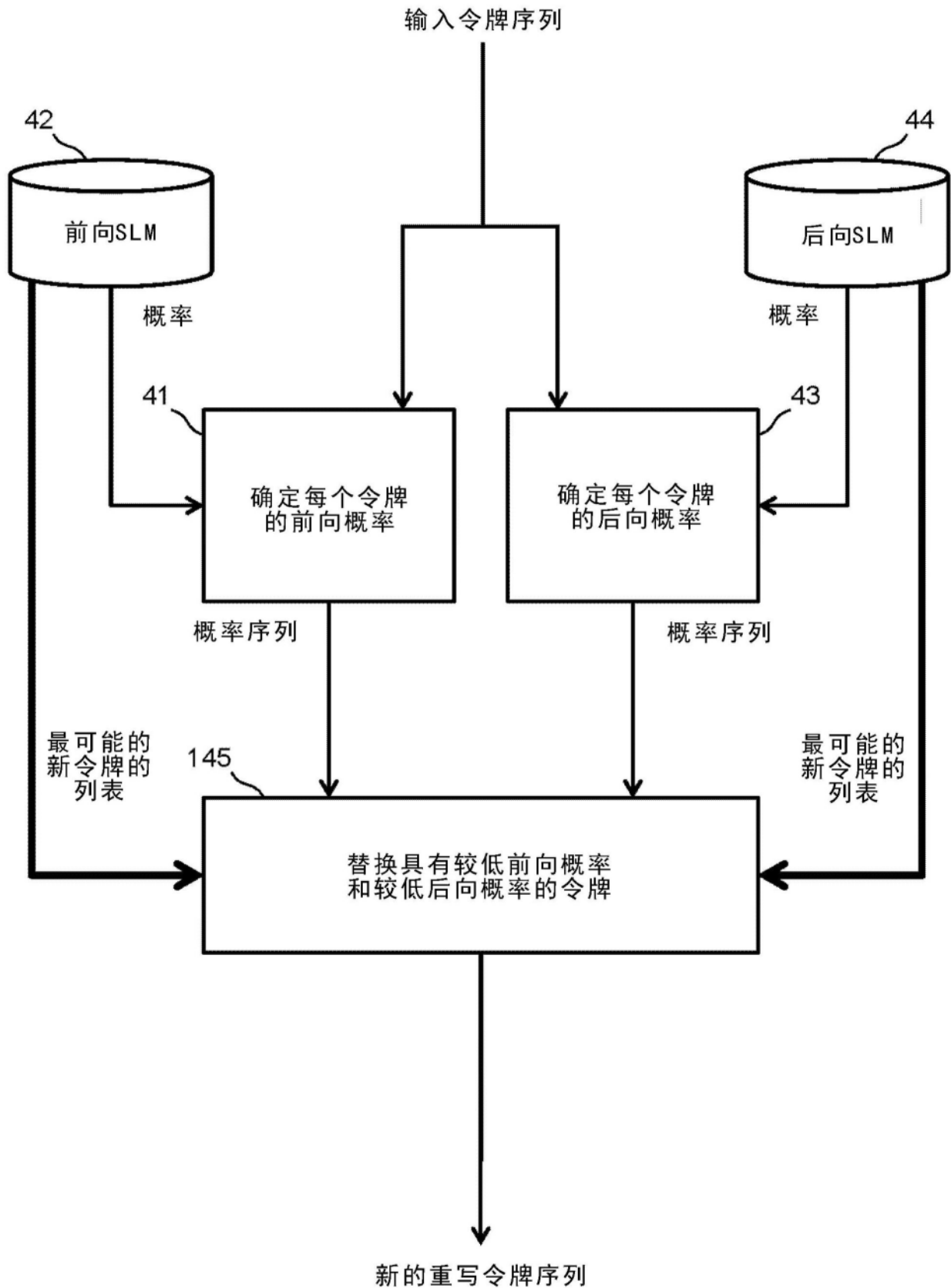


图14

weather pin Hawaii

图15A

weather _____
forecast .073
in .**.069**
Channel .033
data .013
and .**.005**

图15B

_____ Hawaii
to .072
in .**.059**
and .**.036**
Honolulu .012
love .003

图15C

夏威夷 过 天气

图16A

夏威夷 _____
岛 .073
的 .**.069**
旅游 .033
海滩 .013
和 .**.005**

图16B

——天气
 查询 .072
的 .059
和 .036
 多云 .012
 恶劣 .003

图16C

the weather pin Hawaii today

图17A

the weather ——
 today .083
 forecast .076
for .052
 now .038
in .029

图17B

—— Hawaii today
in .102
for .071
 to .054
 over .032
 and .011

图17C

重写分数

```

in   .029 + .102 = 0.131
      .029 * .102 = 0.0030
      log(.029) + log(.102) = -2.53
for  .052 + .071 = 0.123
      .052 * .071 = .0037
      log(.042) + log(.071) = -2.43

```

图17D

今天 赞 夏威夷的天气

图18A

今天—
 是 .083
 运势 .076
 在 .052
 日期 .038
 去 .029

图18B

 夏威夷
 —
 去 .102
 在 .071
 来 .054
 感受 .032
 漫步 .011

图18C

重写分数

$$\begin{array}{lcl} \text{去} & .029 + .102 & = 0.131 \\ & .029 * .102 & = 0.0030 \\ & \log(.029) + \log(.102) & = -2.53 \\ \text{在} & .052 + .071 & = 0.123 \\ & .052 * .071 & = .0037 \\ & \log(.042) + \log(.071) & = -2.43 \end{array}$$

图18D

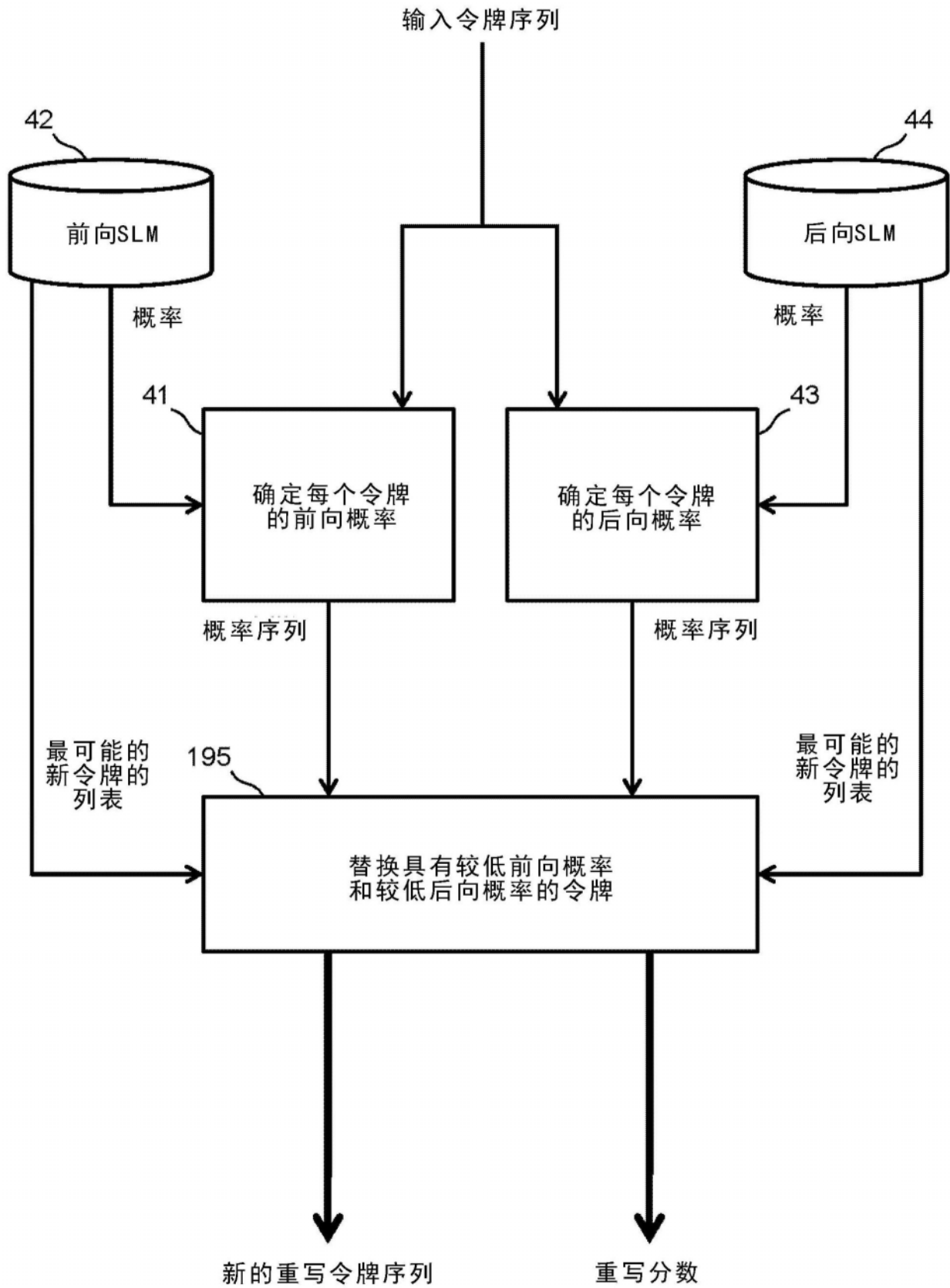


图19

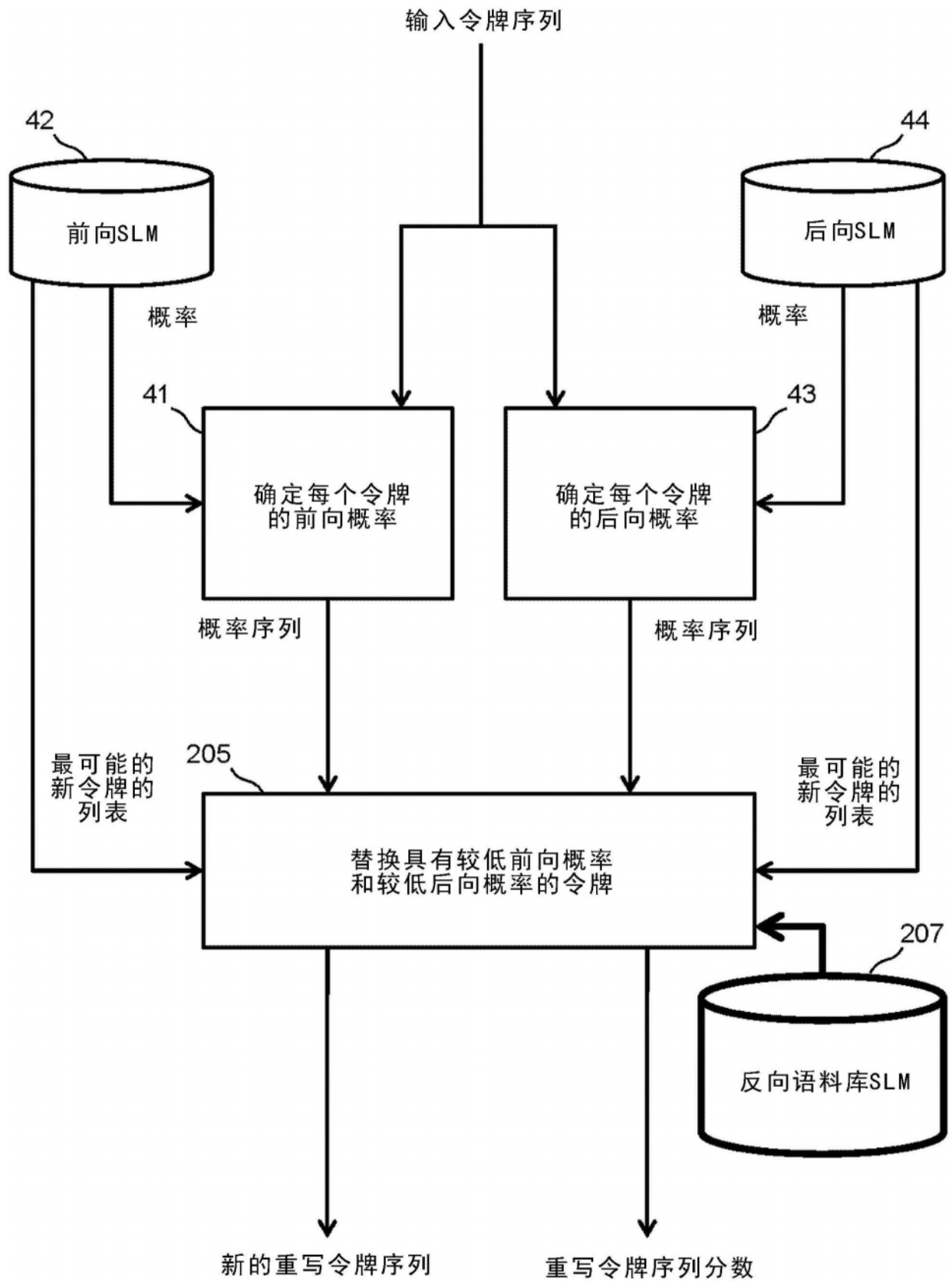


图20

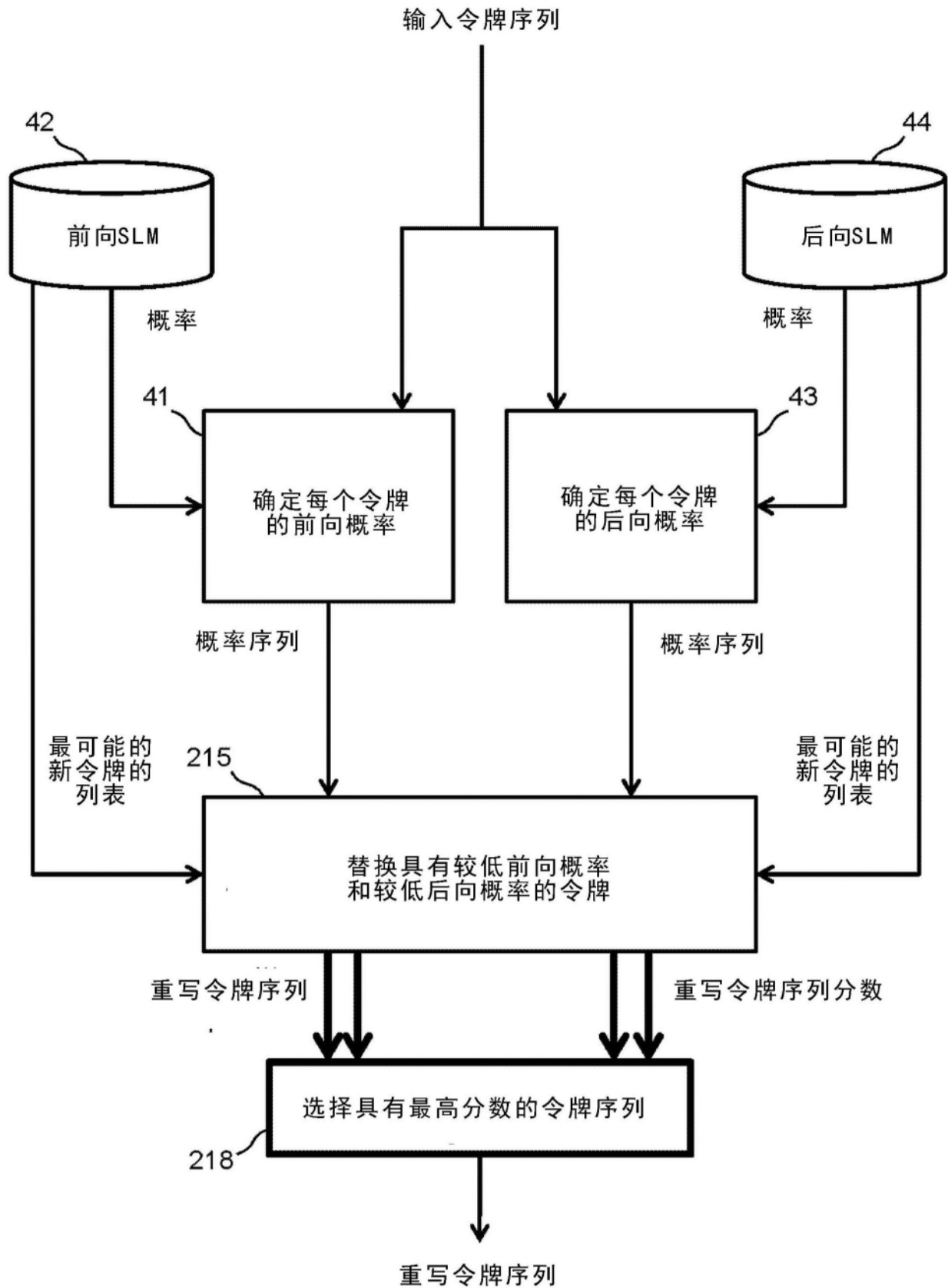


图21

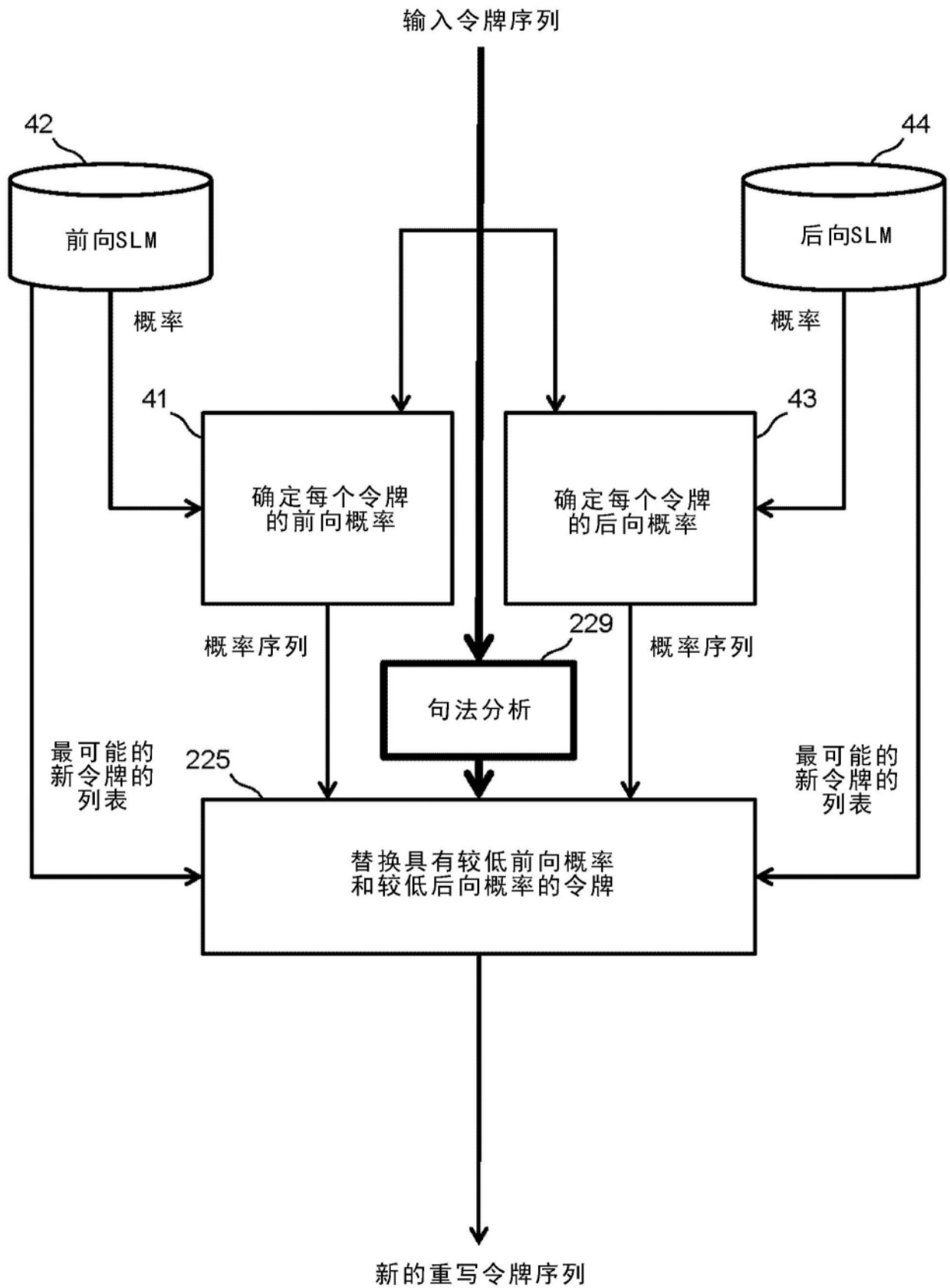


图22

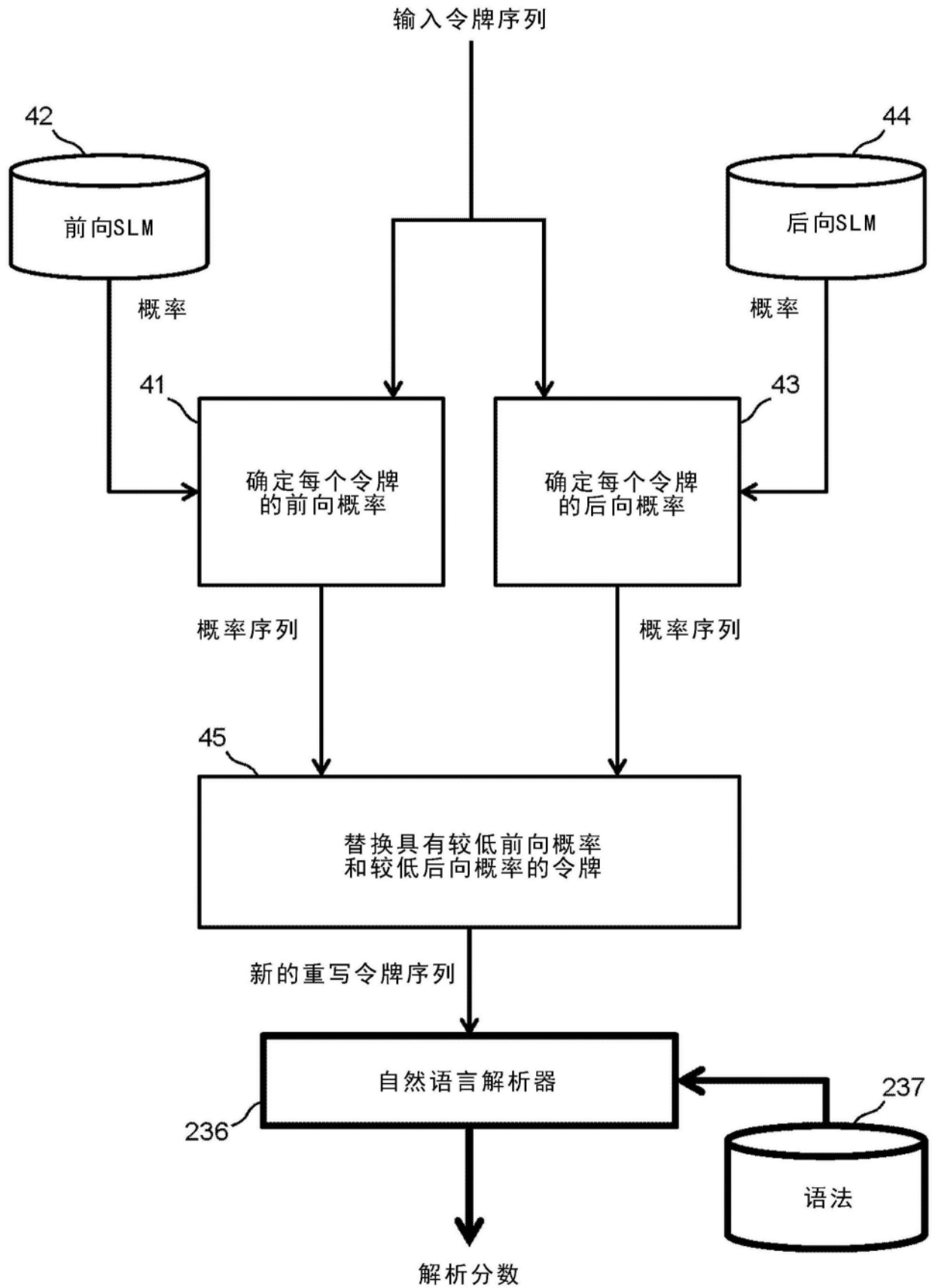


图23

what's the weather for Hawaii today
解析分数 .29

what's the weather in Hawaii today
解析分数 .35

图24

夏威夷在哪儿?
解析分数 .29

夏威夷在哪里?
解析分数 .35

图25

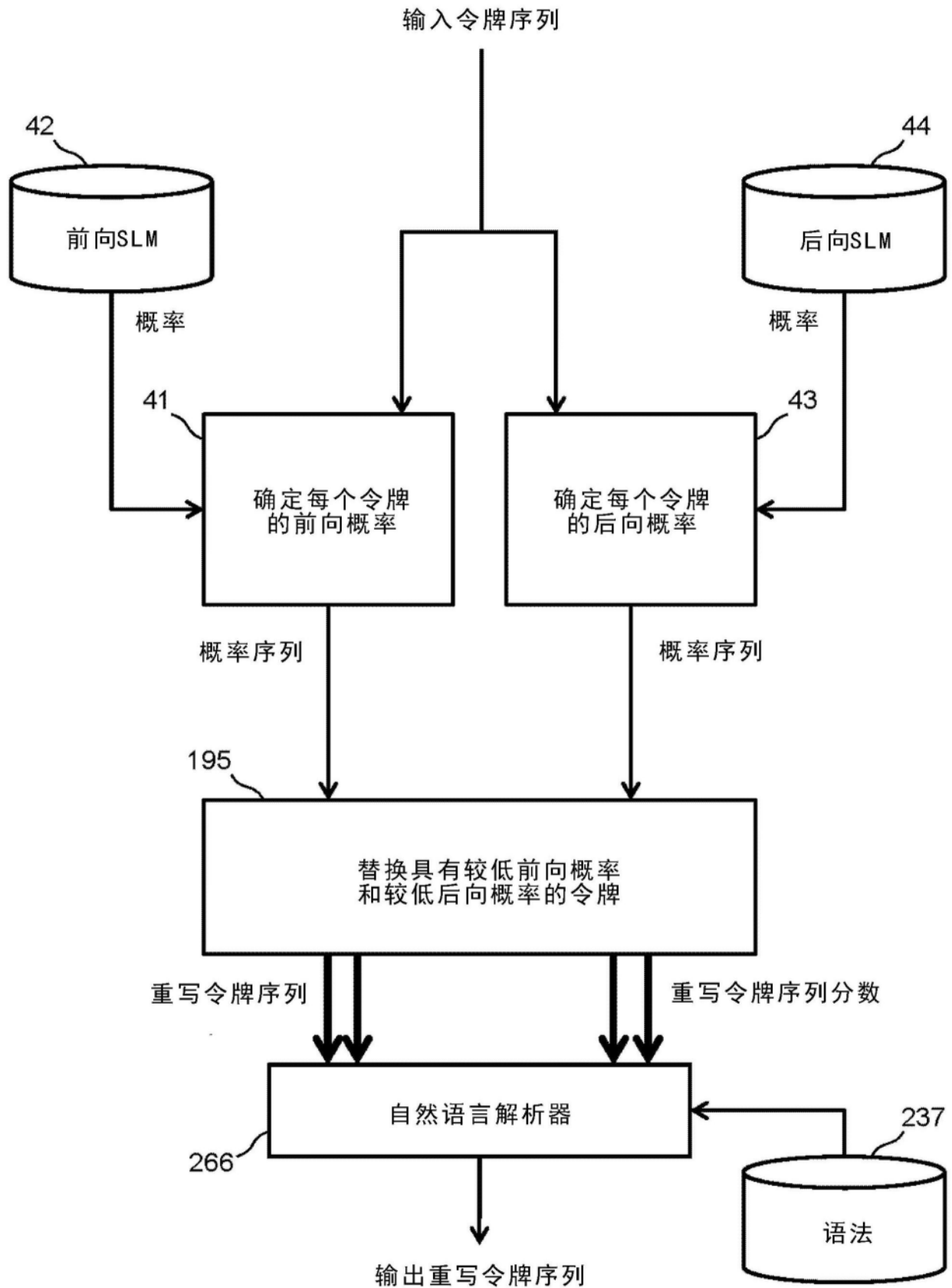


图26

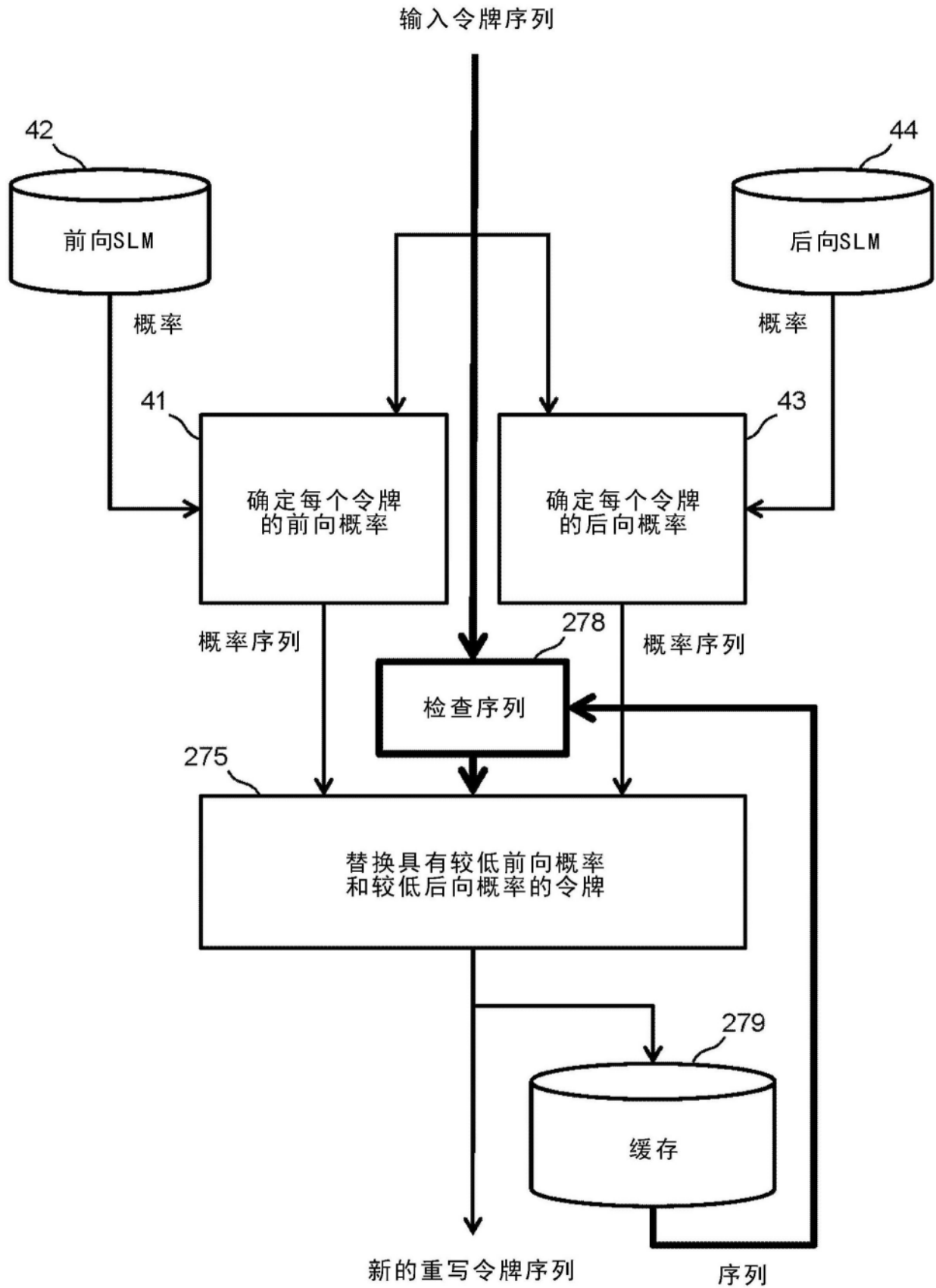


图27

输入：what the weather pin <place>
<date>

新：what's the weather in <place>
<date>

输入：will there be any rain <date>
新：will there be any rain <date>

输入：is it going to
新：is it going to rain <date>

输入：what the weather pin <place>
<date>
新：what's the weather in <place>
<date>

输入：do we have rain for <date>
新：will there be rain on <date>

输入：what the weather in <place>
<date>
新：what's the weather in <place>
<date>

输入：<date> will it rain
新：will it rain <date>

图28

新：will there be any rain <date>
输入：will there be any rain <date>

新：is it going to rain <date>
输入：is it going to

新：will there be rain on <date>
输入：do we have rain for <date>

新：will it rain <date>
输入：<date> will it rain

新：what's the weather in <place>
<date>
输入：what's the weather for <place>
<date> (3)
输入：what's the weather on <place>
<date> (2)
输入：what the weather pin <place>
<date>
输入：what the weather in <place>
<date>
输入：what the weather for <place>
<date>

图29

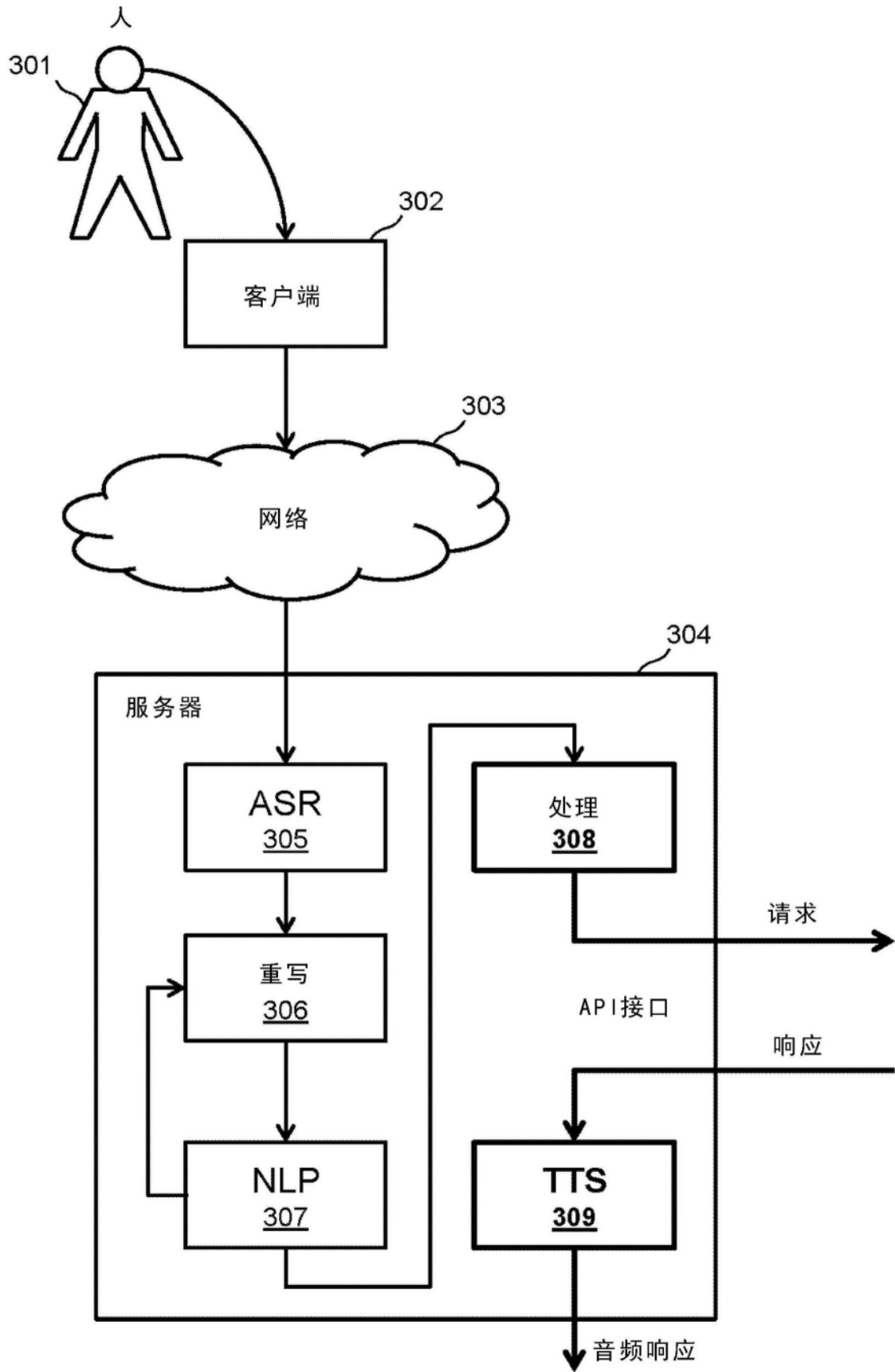


图30

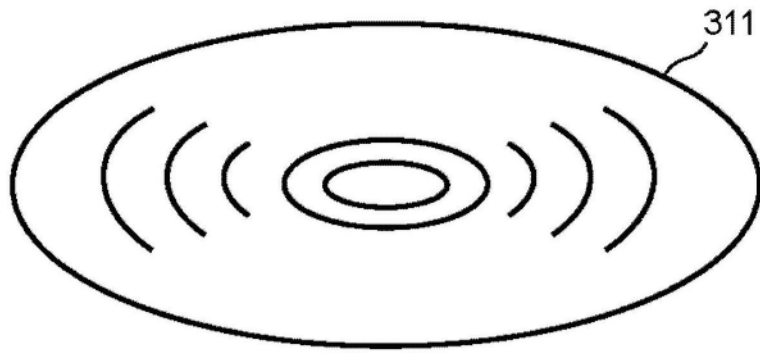


图31A

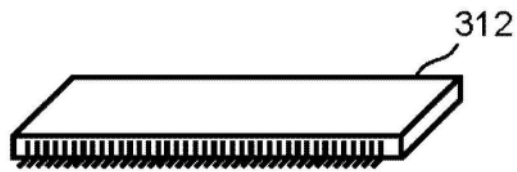


图31B

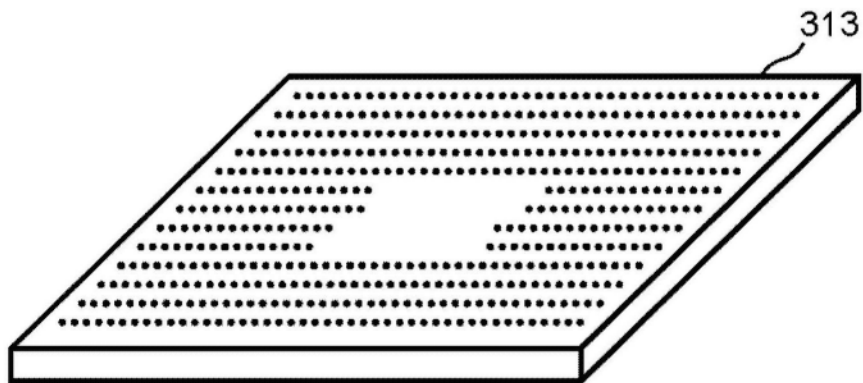


图31C

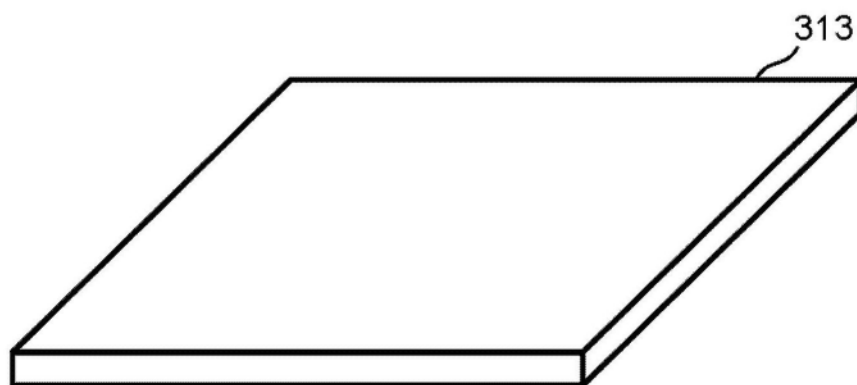


图31D

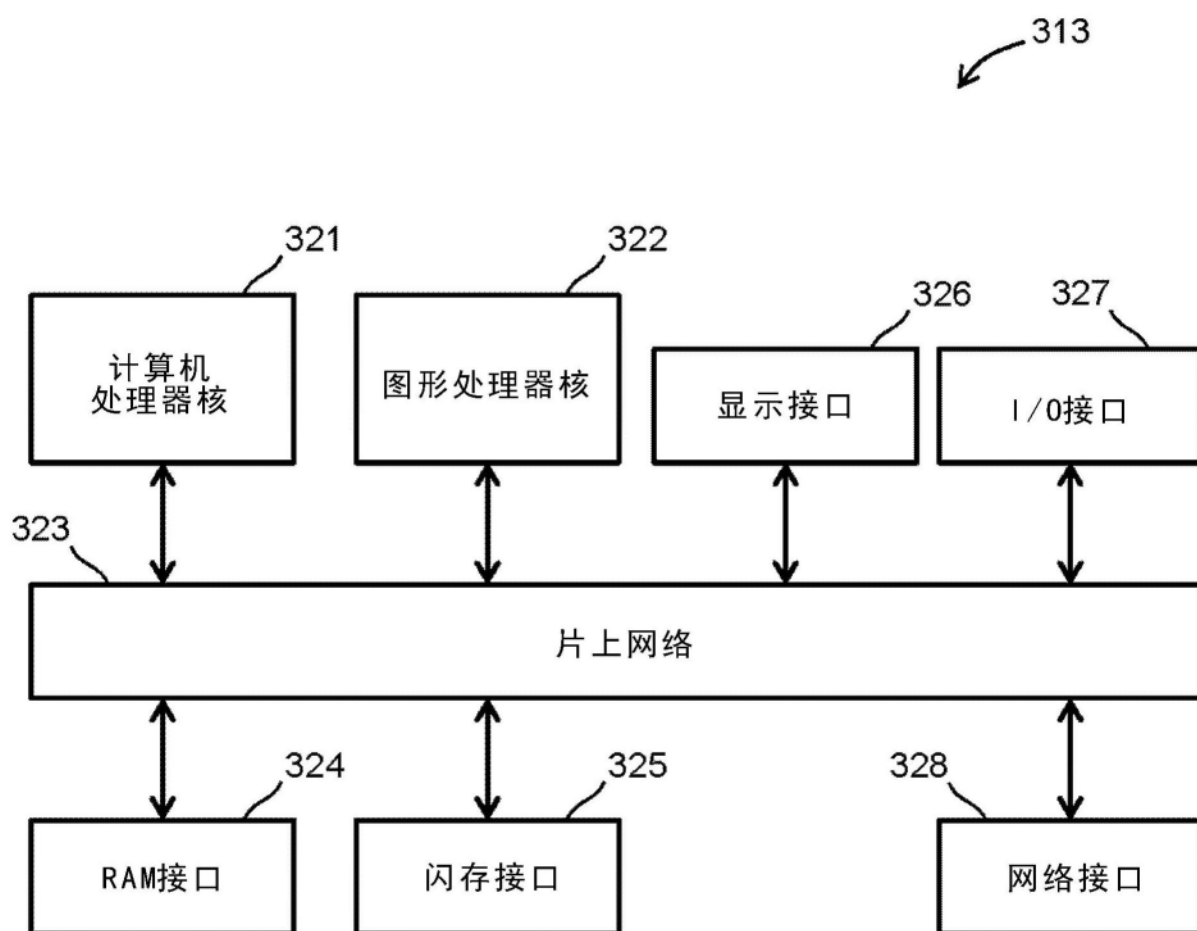


图32

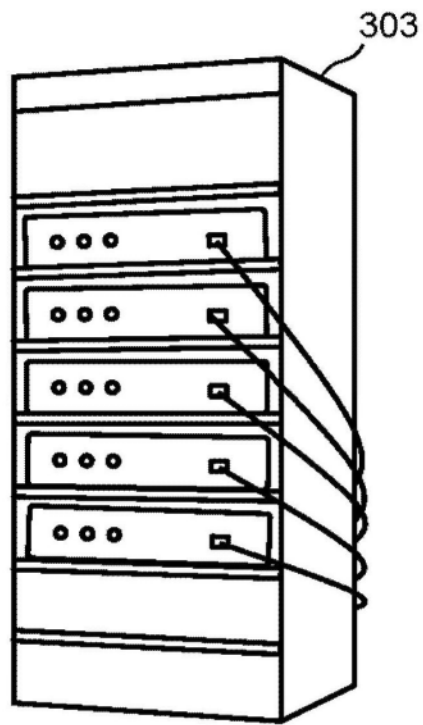


图33

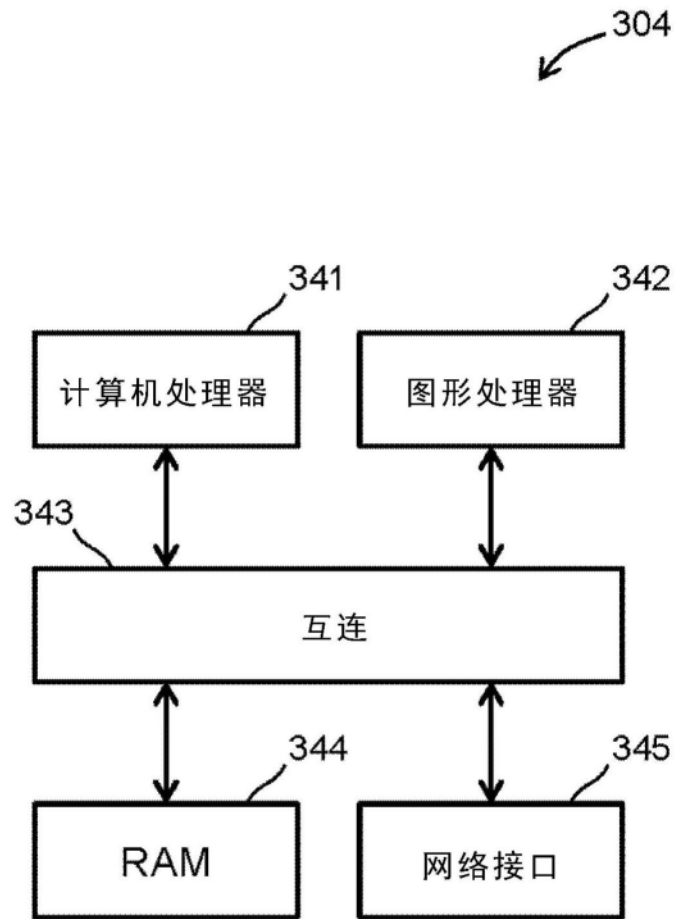


图34