



(12) 发明专利申请

(10) 申请公布号 CN 104769584 A

(43) 申请公布日 2015. 07. 08

(21) 申请号 201380046721. X

A · N · 莫海迪恩佩瑟迪恩

(22) 申请日 2013. 07. 30

K · C · S · 雷迪

(30) 优先权数据

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

61/677, 895 2012. 07. 31 US

13/801, 812 2013. 03. 13 US

代理人 陈新

(85) PCT国际申请进入国家阶段日

(51) Int. Cl.

2015. 03. 09

G06F 17/30(2006. 01)

(86) PCT国际申请的申请数据

PCT/US2013/052637 2013. 07. 30

(87) PCT国际申请的公布数据

W02014/022345 EN 2014. 02. 06

(71) 申请人 韦韦欧股份有限公司

地址 美国马萨诸塞

(72) 发明人 M · 阿拉瓦穆丹 R · 巴尔维

S · 万卡塔拉曼 V · 阿加瓦尔

A · 萨瓦尔卡 G · 拉马默尔赛

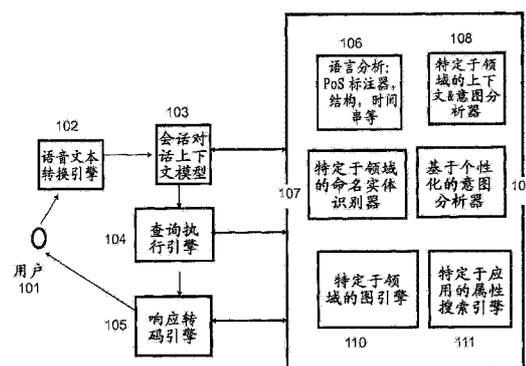
权利要求书2页 说明书16页 附图3页

(54) 发明名称

在对话交互中消除用户意图歧义

(57) 摘要

一种在用于信息检索的对话交互中消除用户意图歧义的方法。该方法包括提供对内容项集合的访问,其中内容项具有描述该内容项的元数据,并且提供对结构化知识的访问,其中结构化知识示出内容项之间的语义关系和链接。该方法还包括提供用户偏好签名、接收来自用户的第一输入、并且确定第一输入的歧义指数,其中用户想要第一输入识别至少一个期望的内容项。如果歧义指数高,则该方法基于第一输入以及结构化知识、用户偏好签名、用户的位置和第一输入的时间中的至少一个来确定查询输入,并且基于将该查询输入与和内容项相关联的元数据进行比较来选择内容项。



对话交互体系结构

1. 一种在用于信息检索的对话交互中消除用户意图歧义的方法,该方法包括:
提供对内容项集合的访问,每个内容项都与描述相应内容项的元数据相关联;
提供对结构化知识的访问,结构化知识表示内容项之间的语义关系和链接;
提供用户偏好签名,用户偏好签名描述用户对 (i) 特定内容项和 (ii) 与内容项相关联的元数据中的至少一个的偏好;
接收来自用户的第一输入,用户想要第一输入识别至少一个期望的内容项;
确定第一输入的歧义指数;
在歧义指数超过第一阈值的条件下,基于结构化知识、用户偏好签名、用户的位置和第一输入的时间中的至少一个以及第一输入来确定查询输入,并且基于将该查询输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择内容项子集;以及
在歧义指数没有超过第一阈值的条件下,基于将第一输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择内容项子集。
2. 如权利要求 1 所述的方法,还包括将该内容项子集呈现给用户。
3. 如权利要求 1 所述的方法,其中歧义指数是基于第一输入的多种可能的解释而确定的。
4. 如权利要求 1 所述的方法,还包括在歧义指数超过第一阈值的条件下,确定第一输入的哪个部分是有歧义的,其中查询输入的确定还基于输入的歧义部分。
5. 如权利要求 1 所述的方法,还包括确定输入的意图、实体和过滤器,意图是用户寻找的东西,实体是描述该意图的名词或代词,并且过滤器是该实体的限定词。
6. 如权利要求 1 所述的方法,还包括在其中歧义指数超过第二阈值的条件下,请求和接收来自用户的第二输入,其中查询输入的确定还基于第二输入。
7. 如权利要求 1 所述的方法,还包括请求和接收来自用户的第二输入,其中查询输入的确定还基于第二输入。
8. 如权利要求 7 所述的方法,还包括将该内容项子集呈现给用户。
9. 如权利要求 7 所述的方法,其中第二阈值高于第一阈值。
10. 如权利要求 7 所述的方法,还包括在歧义指数超过第一阈值的条件下,确定第一输入的哪个部分是有歧义的,其中查询输入的确定还基于输入的歧义部分。
11. 如权利要求 7 所述的方法,还包括确定输入的意图、实体和过滤器,意图是用户寻找的东西,实体是描述该意图的名词或代词,并且过滤器是该实体的限定词。
12. 一种在用于信息检索的对话交互中消除用户意图歧义的系统,该系统包括:
在非暂态计算机可读介质上编码的计算机可读指令,该计算机可读指令使计算机系统:
提供对内容项集合的访问,每个内容项都与描述相应内容项的元数据相关联;
提供对结构化知识的访问,结构化知识表示内容项之间的语义关系和链接;
提供用户偏好签名,用户偏好签名描述用户对 (i) 特定内容项和 (ii) 与内容项相关联的元数据中的至少一个的偏好;
接收来自用户的第一输入,用户想要第一输入来识别至少一个期望的内容项;
确定第一输入的歧义指数;
在歧义指数超过第一阈值的条件下,基于结构化知识、用户偏好签名、用户的位置和第

一输入的时间中的至少一个以及第一输入来确定查询输入,并且基于将该查询输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择该内容项子集;以及

在歧义指数没有超过第一阈值的条件下,基于将第一输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择内容项子集。

13. 如权利要求 12 所述的系统,其中计算机可读指令还使计算机系统将该内容项子集呈现给用户。

14. 如权利要求 12 所述的系统,其中歧义指数是基于第一输入的多种可能的解释而确定的。

15. 如权利要求 12 所述的系统,其中计算机可读指令还使计算机系统歧义指数超过第一阈值的条件下,确定第一输入的哪个部分是有歧义的,其中查询输入的确定还基于输入的歧义部分。

16. 如权利要求 12 所述的系统,其中计算机可读指令还使计算机系统确定输入的意图、实体和过滤器,意图是用户寻找的东西,实体是描述该意图的名词或代词,并且过滤器是该实体的限定词。

17. 如权利要求 12 所述的系统,其中计算机可读指令还使计算机系统请求和接收来自用户的第二输入,其中查询输入的确定还基于第二输入。

18. 如权利要求 12 所述的系统,其中计算机可读指令还使计算机系统在歧义指数超过第二阈值的条件下,请求和接收来自用户的第二输入,其中查询输入的确定还基于第二输入。

19. 如权利要求 18 所述的系统,其中计算机可读指令还使计算机系统将该内容项子集呈现给用户。

20. 如权利要求 18 所述的系统,其中第二阈值高于第一阈值。

21. 如权利要求 18 所述的系统,其中计算机可读指令还使计算机系统在歧义指数超过第一阈值的条件下,确定第一输入的哪个部分是有歧义的,其中查询输入的确定还基于输入的歧义部分。

22. 如权利要求 18 所述的系统,其中计算机可读指令还使计算机系统确定输入的意图、实体和过滤器,意图是用户寻找的东西,实体是描述该意图的名词或代词,并且过滤器是该实体的限定词。

在对话交互中消除用户意图歧义

[0001] 相关申请交叉引用

[0002] 本申请要求以下专利申请的权益,其内容通过引用被结合于此:

[0003] 于 2013 年 3 月 13 日提交的标题为“Disambiguating User Intent in Conversational Interaction System for Large Corpus Information Retrieval”的美国专利申请 No. 13/801,812 ;及

[0004] 于 2012 年 7 月 31 日提交的标题为“Disambiguating User Intent in Conversational Interaction System for Large Corpus Information Retrieval”的美国临时专利申请 No. 61/677,895。

技术领域

[0005] 本发明涉及在用于信息检索的对话交互系统中用于消除用户意图歧义的方法,并且更具体地涉及利用结构化信息和用户偏好的技术。

背景技术

[0006] 本发明涉及在用于大型语料信息检索的对话交互系统中用于“最优地”和“智能地”消除用户意图/输入的歧义的方法,其中意图/输入具有一种或多种的以下歧义:(1) 词汇歧义(多个合格响应在词汇上与用户输入匹配)或(2) 语义歧义,其中歧义在于时间(基于时间属性的多个合格响应),或者歧义在于位置(基于位置属性的多个合格响应),或者歧义在于任何内容属性或内容属性的组合(基于用户指定的一个或多个内容属性的多个合格响应)或者仅仅是在用户请求固有的非特殊性中产生的又导致多个合格响应的歧义(例如,宽泛的意图请求)。在本公开内容中描述的“最优”消除歧义方法的实现使得系统能够问最少量的澄清问题(在理想的情况下,根本不问任何问题)来理解用户意图。本公开内容中描述的“智能”消除歧义方法的实现使得系统能够自然地进行消除歧义的交流,类似于人类在对话中澄清歧义的方式。该系统利用特定于领域的结构化知识、时间、用户的位置(如果可以得到的话)、以及用户偏好签名(如果可以得到的话)来执行最优的和智能的歧义消除。本公开内容中描述的方法是语言无关的并且在特定于语言的模块的帮助下可以被应用到多种语言。此外,本文公开的方法尤其适于具有高度语义歧义和冲突的大型语料信息库,其中高度语义歧义和冲突是由于库中的给定实体或概念可以以多种方式被引用并且同一词语会在不同含义的上下文中出现的事实引起的。

[0007] 对话系统的关键性能度量不是当用户意图/输入没有歧义时它的响应多好地匹配用户意图,而是当用户意图/输入具有歧义时它如何响应。好的对话系统在其可能的响应策略的技能中不具有喷出大量响应的奢侈条件——像搜索引擎将喷出结果那样,即使在用户输入/意图中存在歧义。对于搜索引擎而言,以某种相关性顺序为歧义的输入/意图显示所有的结果将会被赞美为好搜索引擎的标志。在对话系统中为歧义的用户输入/意图采用相同的方法将相当于在星巴克令人疑惑地遇到过分热情的销售人员,只因为用户健忘地不能描述用户脑中特定种类的咖啡(例如,Caffe Latte),销售人员就一口气地说出十种

咖啡选项。这里,尽管销售人员清楚地理解意图是咖啡(而不是茶),但是销售人员没有留意存在许多与意图匹配的选项的事实——更聪明的销售人员可能会说“噢,有多种咖啡可选,你想让我给你快速过一遍这些选项吗?”

[0008] 本公开内容在宽泛意义上使用术语“歧义”来捕获当存在匹配用户输入的多个合格响应(有一个以下提到的例外)时的情形。如在本公开内容中所使用的,术语“歧义”的含义的一个例子可以从以下例子中得到理解:尽管好的对话系统将致力于理解用户意图并产生最简洁的针对性的响应,其中取决于问题,该响应可以理想地只是一个简洁的响应(例如,问题“今晚 sox 比赛吗?”可能产生只显示 red sox 比赛的时间和地点的响应,其中 red sox 是根据用户的签名推断出的),但是这并不一定暗示所有的用户问题都产生单个响应。这也不暗示对问题提供多个选项将是次优的。例如,如果用户陈述“给我显示附近的星巴克”,则最好的响应将是显示标有所有靠近该用户的星巴克结果的地图,使得用户可以从可视化地图中毫不费劲地选出任何一个。即使对于诸如“给我显示附近的餐馆”之类的更宽泛的意图请求,在地图上显示多个响应也是系统可以提供的最好响应。

[0009] 在这两个例子中意图都是清楚的——但是响应在某种意义上是“歧义”的,因为它多于一个——系统不知道用户会喜欢哪个特定的餐馆。但如果存在用户偏好签名,则它可以产生从其它响应中突出的具有最优的星巴克/餐馆的响应。在以上提到的这些例子中,多个响应并不真正是歧义响应,而是都与用户意图匹配的一组“选项”(当然用户可能由于主观原因仍不选择星巴克或餐馆)。这里使用“选项”一词是为了与“响应”区分,以显示用户想要多个选项——而不是只有一个选项(即使系统有用户偏好签名,它仍将提供多个“选项”)。另一个例子是——“给我显示 Meryl Streep 的电影(movies)”。在这个例子中,用户想要 Meryl Streep 的多个电影以便从中“选择”。

[0010] 在本公开内容中描述的方法将着重于其中歧义(或多个合格响应)源于不能提供可通过良好的置信度被认为是与用户意图相匹配的一个清楚的“选项”或一组“选项”的情况。而且,当用户想要特定的一个或多个选项时,虽有挑选那个特定的选项或选项集合的词汇和/或语义歧义,但是负担落在系统上。这种歧义不是由于系统的缺陷或“缺少智能”,而正是由于用户提交的问题中的固有歧义(词汇或语义)。

[0011] 本公开内容中描述的方法将着重于为其中由于用户意图/输入中的固有歧义而不可能提供一组选项的这些歧义情况消除歧义的方法。星巴克/餐馆和“Meryl Streep”的响应是无需歧义解决方案的最好情况的情形。这些系统响应与上面提到的对问题“今晚有 sox 比赛吗”的简洁响应简直一样好——多个响应是“选项”而不是响应中有歧义。

[0012] 在本公开内容中还使用“歧义”一词来处理例外情况——当根本不存在与用户意图/输入相匹配的响应时。在这种边界条件下,歧义可能是由于多种多样的原因,从用户没有正确地表达意图到仅仅是在信息领域空间中不存在匹配。例如,如果用户问“今晚有 sox 比赛吗”,但是今晚没有任何 sox 比赛,那么这就是其中不存在与用户想看比赛的意图的匹配的例子。

[0013] 从严格的请求/响应来说,这里不存在歧义。但是在人类交互中,当用户表达不可被满足的意图时,会产生合理的问题“我能给用户可能接近满足原来意图的东西吗?”。通常,提供接近的替代选项的响应常会被赞赏。在“今晚有 sox 比赛吗”的例子中,响应“今晚没有,但是有一场昨天晚上你错过的比赛”(这种响应可以利用用户的偏好签名

和过往历史记录来创建)。本发明的实施例将这种无响应的情况视为“空 (null) 响应歧义”的情况,并且产生尽量接近满足用户意图的响应。另一个例子是,“X 和 Y 一起演过电视剧吗?”。假设 X 和 Y 从来没有一起演过电视剧,那么本发明的实现将利用特定于领域的结构化知识来产生“没有,他们没有一起演过电视剧,但是回到 1989 年他们确实在电影 Z 中一起主演过”。这里,特定于领域的结构化知识被用于产生对“空响应歧义”情况的响应。

[0014] 本公开内容中描述的大多数歧义的例子主要基于数字娱乐空间。但是,本公开内容中描述的方法可以应用到任何信息领域(娱乐、诸如电子邮件、通讯录等的个人语料库),并且也可以跨不同的信息领域。

[0015] 用户意图/输入中的歧义可以是不同类型的。一种可能性是用户输入中的词汇歧义,但是用户有明确的意图。例如,假如用户说“我想看那个 Beethoven 电影”。三部电影具有“Beethoven 电影”的资格——1936 年关于作曲家 Beethoven 的电影、1992 年关于名为 Beethoven 的狗的电影、或者 1994 年关于 Beethoven 的著名电影“Immortal Beloved”。用户的意图显然只是这些电影中的一个(基于请求中“那个”的使用),但是用户的输入词汇上与三个合格的响应相匹配。在这种情况下,好的对话系统绝不会提供这三个合格的响应作为三个平等的有效选项让用户从中选取一个。这样的系统将是其性能已退化到提供结果的搜索引擎性能的对话系统——很显然,除了可能的某个相关性度量外,该系统对词语 Beethoven 没有内部的理解。

[0016] 致力于更接近人类之间对话的对话系统将问用户“你是指关于作曲家的电影还是关于狗的电影?”——很像人在对话中将回应的。该消除歧义的问题本身是对话系统理解词语 Beethoven 的指示,更像人类所做的。例如,同样消除歧义的问题可能被构造成“你是指作曲家 Beethoven 还是狗 Beethoven?”。虽然这仍然是好的消除歧义的响应,但是之前的响应更接近正常的说话,其中正是有歧义的词语,即 Beethoven,被从给用户的消除歧义的响应中去除。总而言之,好的对话系统在其对词汇歧义的响应中将尤为敏感并且将产生更像人类的消除歧义响应,因为这种响应是决定系统水准的关键度量,其中系统水准的范围可能从“搜索引擎智能”到“自然对话智能”。

[0017] 歧义的另一形式在于时间或线性连续性的语义歧义。如果用户说“我想看 Borghia”(电视系列剧),则在用户想看的季中存在歧义,尽管当前季在大多数情况下会被认为是合理的响应。但是,如果用户已经从第一季观看该系列剧,则理想地,跟在最后观看的一季之后的季将是理想的。当用户在观看一系列连续内容(像 David Attenborough 的自然系列“Life on earth”)的过程中时,也可能出现这种形式的歧义。在这种情况下,歧义的解决方案理想地也通过以用户最后观看的剧集开始加以解决。在任何这两种情况下(季或线性系列),如果用户没有以时间或线性顺序观看,那么消除歧义的问题是不可避免的。但是,如果用户说“我想看下一个 Borghia”,那么用户意图可以被解释为指用户最后观看的一集后面跟着的剧集。

[0018] 歧义的另一形式在于导致多个合格响应的位置的歧义。例如,请求:“给我显示在夏威夷拍摄的那个 Spielberg 电影”将产生多个电影 Jurassic Park 及其续集——Lost world 和 Jurassic Park III——所有这些都位置在夏威夷拍摄。这里,用户通过问“那个 Spielberg 电影”只想要一个电影。更接近人类响应的响应会是“Jurassic park 是在夏威夷拍摄的。它的续集“Lost World”和“Jurassic Park III”也是在那里拍摄的”。

[0019] 在另一个例子中,如果用户问“今晚有 tiger 的比赛吗”,那么用户可能指 Detroit Tigers 棒球队或 Louisiana Tigers 足球队 (Louisiana Tigers 足球队比使用同一名字的该棒球队更流行)。但是,如果用户的位置已知是在 Louisiana,那么很可能用户指 Louisiana 足球队。但是,如果用户的位置已知是在 Detroit,那么该问题可以映射到 Detroit 棒球队。在其中用户正在旅行并且位置未知的情况下,则在问题中存在需要加以解决的歧义,尤其当不存在关于用户对任何这两支球队的偏好的先前信息时。此外,如果该问题是在赛季期间提出的,那么除了位置之外,这也可能是消除歧义的因素。一般而言,在用户指定的任何属性中都可能存在歧义,而不仅仅是在位置和时间中——以上例子示出了在诸如位置和时间的属性中的歧义。

[0020] 在从意图的最宽泛程度来理解用户意图时也可能存在歧义。例如,如果用户说“我今晚想看电影”,即使用户偏好签名是已知的,用户也可能对动作片或悬疑电影感兴趣。因此,在这两种题材类型之间仍然存在需要加以解决的歧义。在一些现有对话系统中使用的消除歧义机制是使用户经历多层决策树,向用户提出问题以缩小选项。人类从来没有在自然对话中做过这种“算法树遍历方法”,这使得该策略对于致力于接近自然对话的对话系统来说是不可接受的。这种多层决策树遍历对一些诸如航班预订过程的领域来说在某种程度上是可以接受的,但是当应用在诸如娱乐空间的某些领域时,它将看起来滑稽愚蠢。

[0021] 歧义也可以从用户意图输入中的错误中产生,其中输入可以是语音或文本输入。为了本公开内容中所描述的方法的目的,这些错误被认为是词汇错误(尽管在一些情况下词汇错误可能实际上导致语义差异)。在本公开内容中描述的歧义的解决方案利用特定于领域的结构化知识、用户偏好签名(如果可以得到的话)、用户的位置(如果可以得到的话)和时间。但是,显然,如在以上例子中看到的,不是所有的歧义都是可解决的。

[0022] 总而言之,用户输入/意图中的歧义会导致合格的响应(除了“空响应”情况之外),这些合格的响应可能松散地彼此关联,就像词汇歧义的例子那样(例如 Beethoven 电影可以与关于音乐家或关于名为 Beethoven 的狗的电影相匹配)。在另一种极端情况下,用户输入/意图中的歧义会导致可紧密地彼此关联的合格响应,在某种程度上,该多个响应更像是“选项”——所有响应紧密关联并且具有高度的匹配用户意图的概率(例如对“给我显示附近的星巴克”的响应)。此外,当用户意图宽泛时,合格的响应潜在地非常大,因此有必要向用户提供消除歧义的响应。本发明中描述的对话系统的实施例基于歧义的本质(词汇或语义歧义)以及合格响应彼此的关联程度,通过利用特定于领域的结构化知识、时间、用户的位置(如果可以得到的话)和用户偏好签名(如果可以得到的话)在对话中响应用户。确保消除用户意图歧义的对话交流致力于接近人类对话流畅性的理想目标,在人类对话中歧义消除被无缝地织入交流的真实结构中,并且不由于因其机器产生来源的人造物造成的突兀而中断这种无缝流动。本公开内容中描述的对话系统的实施例还解决了“空响应歧义”的情况,使得用户不会带着未实现的意图被留在死胡同。

发明内容

[0023] 提供了在用于信息检索的对话交互中消除用户意图歧义的方法。该方法包括提供对内容项集合的访问。每个内容项都与描述相应内容项的元数据相关联。该方法还包括提供对示出内容项之间的语义关系和链接的结构化知识的访问,并且提供用户偏好签名,用

户偏好签名描述用户对 (i) 特定内容项和 (ii) 与内容项相关联的元数据中至少一个的偏好。该方法还包括接收来自用户的第一输入。用户想要第一输入来识别至少一个期望的内容项。该方法还包括确定第一输入的歧义指数。该方法包括：在其中歧义指数超过第一阈值的条件下，基于第一输入以及结构化知识、用户偏好签名、用户的位置和第一输入的时间中的至少一个来确定查询输入，并且基于将该查询输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择该内容项子集。该方法还包括，在其中歧义指数不超过第一阈值的条件下，基于将第一输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择该内容项子集。

[0024] 在另一个实施例中，该方法还包括将该内容项子集呈现给用户。

[0025] 在还有的另一个实施例中，歧义指数是基于第一输入的多种可能的解释而确定的。

[0026] 在不同的实施例中，该方法还包括：在其中歧义指数超过第一阈值的条件下，确定第一输入的哪个部分是有歧义的。查询输入的确定可以进一步基于输入的歧义部分。

[0027] 在还有的实施例中，该方法包括确定输入的意图、实体和过滤器。意图可以是用户寻找的东西，实体可以是描述该意图的名词或代词，并且过滤器可以是该实体的限定词。

[0028] 在还有的另一个实施例中，该方法包括：在其中歧义指数超过第二阈值的条件下，请求和接收来自用户的第二输入。查询输入的确定可以进一步基于第二输入。

[0029] 在另一个实施例中，该方法包括请求和接收来自用户的第二输入。查询输入的确定可以进一步基于第二输入。

[0030] 在还有的另一个实施例中，第二阈值高于第一阈值。

[0031] 在实施例中，提供了在用于信息检索的对话交互中消除用户意图歧义的系统。该系统包括在非暂态计算机可读介质上编码的计算机可读指令。计算机可读指令使计算机系统提供对内容项集合的访问，其中每个内容项都与描述相应内容项的元数据相关联，并且计算机可读指令使计算机系统提供对结构化知识的访问，其中结构化知识示出内容项之间的语义关系和链接。计算机可读指令还使计算机系统提供描述用户对 (i) 特定内容项和 (ii) 与内容项相关联的元数据中的至少一个的偏好的用户偏好签名、接收来自用户的第一输入、并且确定第一输入的歧义指数，其中用户想要第一输入识别至少一个期望的内容项。计算机可读指令还使计算机系统在其中歧义指数超过第一阈值的条件下，基于第一输入以及结构化知识、用户偏好签名、用户的位置和第一输入的时间中的至少一个来确定查询输入，并且基于将该查询输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择该内容项子集。计算机可读指令还使计算机系统在其中歧义指数不超过第一阈值的条件下基于将第一输入与和内容项子集相关联的元数据进行比较来从内容项集合中选择该内容项子集。

附图说明

[0032] 为了对本发明的各种实施例有更完整的理解，现在结合附图参考以下描述，附图中：

[0033] 图 1 说明作为本发明实施例的体系结构。

[0034] 图 2 说明特定于领域的结构化知识库的创建。

[0035] 图 3 说明创建特定于领域的结构化知识库的阶段。

[0036] 图 4 说明特定于领域的知识库实体以及实体之间关系的一部分的示意图。

具体实施方式

[0037] 本发明的优选实施例包括在对话交流中用于消除用户意图的歧义并且满足那个意图的方法和系统。本发明的优选实施例及其优点可以通过参考图 1-4 来理解,其中相同的参考标号指代相同的元素。

[0038] 信息库的创建

[0039] 用于消除用户意图 / 输入歧义的特特定于领域的信息库在不断地进化,命名实体的可扩展数据库通过把从不同来源收集到的许多结构化和非结构化信息放在一起进行整合。如图 2 中所示出的,由于结构化知识是从不同来源整合的,因此通过对命名实体可用的元内容上执行统计文本处理、链接分析和其它信号分析(对例如位置信息等),就在信息库自身的成员之间创建隐含的和明确的语义关系和链接。这些关系始终在进化(如在图 3 中所示出的),并随着时间的推移通过聚合使用分析、协同过滤以及其它技术而被增强。

[0040] 信息库中的每个命名实体都以类似于文本信息检索工作将文档表示为加权的文本短语向量的方式被表示为加权的文本短语(词语 term)的向量。因为简单的基于“tf-idf”(词频 / 逆文档频率)的方法单独在许多重要的情况下就本发明的实现目的来说是不够的。命名实体的向量表示中的权重计算被设计成利用以显示文本短语的方式呈现的更多的信息信号、在各种类型文本描述中文本短语的位置、以及还有与文本短语相关联的超链接的结构化和位置属性。权重计算因此基于文本、超链接以及从信息库中的元内容中挖掘的其它属性和关系的较丰富的统计和结构化分析。

[0041] 在本发明的优选实施例中,信息库的创建受命名实体整合引擎的驱动,命名实体整合引擎本质上基于每个内容项的文本元内容计算它的简单的加权文本短语向量表示,接着高效地利用对应于所有命名实体的文本短语向量计算该项的文本短语向量的“点积”,接着收集对应于跨越阈值的点积的所有命名实体的列表,应用进一步的过滤以及重新整理(re-ordering)标准(其可以包括该项的非文本元内容和实体),并且接着最终输出与该项关联的实体的最终列表。该过程类似于网络搜索引擎将搜索查询当作向量并执行一种点积计算以根据其索引排列有意义的文档的方式。

[0042] 用于创建信息库的技术使得本发明的实施例能够产生任何不可映射到某个单个命名实体的有利可图领域的丰富的加权文本短语向量表示,并且还可以发现现有实体之间的新的关系。概括地说,利用上述方法建造的信息库用作消除用户意图 / 输入的词汇和语义级别歧义的基础并且帮助图 1 体系结构中描述的诸多模块。依靠这个库来建造其自己的表示(本公开内容中描述的消除歧义机制的一部分)的重要模块是以下描述的图引擎 110。

[0043] 适用于本发明实施例的信息库

[0044] 一些信息库包括实体以及实体之间的关系。每个实体 / 关系分别具有来自类型集合中的类型。此外,与每个实体 / 关系相关联的是一组属性,在一些实施例中,这些属性能够被捕获,作为定义的名称 - 值字段的有限集合。由于实体 / 关系映射提供了描述各个内容项的信息,因此实体 / 关系映射也用作与内容项相关联的一组元数据。换句话说,特定实体将具有与其它实体的关系,并且这些“其它实体”用作对“特定实体”的元数据。此外,映

射中的每个实体可以具有分配给它或分配给在映射中将该实体连接到其它实体的关系的属性。共同地,这构成了与实体 / 内容项相关联的元数据。一般而言,这种信息库被称为结构化的信息库,并且由结构化的信息库提供的信息被称为结构化知识。在一些实施例中,本发明使用结构化的信息库来访问用于信息检索的结构化知识。

[0045] 一些信息库与领域相关联,领域是相似类型的信息和 / 或特定类型的内容项的分组。这些特定于领域的结构化的信息库包含特定于领域的结构化知识。本发明使用的结构化的信息库可以是特定于领域的信息库。下面接着是与领域相关联的信息库的例子。

[0046] 媒体娱乐领域包括实体,诸如电影、电视节目、剧集、剧组、角色 / 人物、演员 / 名人、运动员、比赛、团队、联赛和锦标赛、体育人士、音乐艺术家和表演者、作曲家、专辑、歌曲、新闻名人和 / 或内容分销商。这些实体具有在信息库中捕获的关系。例如,电影实体经“在…中扮演”关系关联到一个或多个演员 / 名人实体。类似地,电影实体可以经“原声曲目”关系关联到音乐专辑实体,其又可以经“专辑中的曲目”关系关联到歌曲实体。同时,名称、描述、日程信息、评论、等级、成本、到视频或音频的 URL、应用或内容存储句柄、分值等可以被视为属性字段。

[0047] 个人电子邮件 (email) 领域包括实体,诸如电子邮件、电子邮件线程 (email-thread)、联系人、发件人、收件人、公司名称、企业中的部门 / 业务单位、电子邮件文件夹、办公室位置和 / 或对应于办公室位置的城市和国家。关系的说明性例子包括关联到电子邮件发件人实体 (以及到实体、cc 实体、bcc 实体、收件者实体和电子邮件线程实体) 的电子邮件实体。同时,可以存在联系人与他或她的公司、部门、办公室位置之间的关系。在这个库中,与实体相关联的属性字段的实例包括联系人的姓名、职称、电子邮件句柄、其它联系信息、电子邮件发送 / 接收时间戳、主题、正文、附件、优先级、办公室的位置信息和 / 或部门的名称和描述。

[0048] 旅行相关的 / 宾馆和景点领域包括实体,诸如城市、宾馆、宾馆品牌、各个兴趣点、兴趣点类别、面向消费者的零售连锁店、汽车租赁点和 / 或汽车租赁公司。这些实体之间的关系包括位置、连锁店的会员关系、和 / 或类别。此外,名称、描述、关键词、成本、服务类型、等级、评论等都相当于属性字段。

[0049] 电子商务领域包括实体,诸如产品条目、产品类别和子类别、品牌、商店等。这些实体之间的关系可以包括产品条目之间的兼容性信息、“由…(商店)出售”的产品等。属性字段包括描述、关键词、评论、等级、成本和 / 或可获得性信息。

[0050] 地址簿领域包括实体和诸如联系人姓名、电子邮件地址、电话号码、物理地址以及雇主的信息。

[0051] 本文列出的实体、关系和属性只是说明性的,并不是要成为详尽的列表。

[0052] 本发明的实施例还可以使用不是如上所描述的结构化信息库的库。例如,对应于基于网络的文档的信息库 (例如,因特网 / 万维网) 可以被认为是链接的文档 (实体) 的关系网。但是,一般而言,没有直接适用的类型结构可以在上述结构化信息库的意义上以非平凡 (nontrivial) 的方式有意义地描述与因特网的元素相关联的所有各种实体和关系以及属性。但是,诸如域名、互联网媒体类型、文件名、文件扩展名等的元素可以用作带有这种信息的实体或属性。

[0053] 例如,考虑由一组非结构化文本文档构成的语料库。在这种情况下,没有直接适用

的类型结构可以枚举一组实体和关系,它们有意义地描述文档内容。但是,作为预处理步骤的语义信息提取处理技术的应用可以产生可部分地从这种语料库中发现结构的实体和关系。

[0054] 根据本发明的某些实施例的访问信息库的说明性例子

[0055] 以下描述说明在如上所描述的结构化和非结构化信息库的上下文中进行的信息检索任务的例子。

[0056] 在一些情况下,用户对一些类型(本文一般地称为意图类型)中的一个或多个实体感兴趣,用户希望通过只指定实体必须满足的属性字段约束来发现这些实体。注意,有时当用户想要某种类型实体的一些属性时,意图可以是(类型,属性)对。例如,如果用户想要电影的等级,则意图可以被视作(类型,属性)=(电影,等级)。这种查询约束在本文一般地称为只有属性的约束(attribute-only constraints)。

[0057] 每当用户指出实体名称或指定足够的信息来直接匹配期望意图类型实体的属性时,它就是只有属性的约束。例如,当用户通过名称和一些附加属性(例如,在60年代制作的“Cape Fear”)识别电影时,或者当他指定与他想要找到的电子邮件匹配的主题时,或者当他基于价格区间询问宾馆时,或者当他指定他想要32GB,黑色的iPod touch时。

[0058] 但是,在一些情况下,用户不仅通过指定关于意图类型实体的属性字段约束,而且还通过在其它实体上指定属性字段约束或指出其它实体的名称而对意图类型中的一个或多个实体感兴趣,其中该意图类型实体以一些明确定义的方式经关系连接到所述其它实体。这种查询-约束在本文被一般地称为面向连接的约束。

[0059] 面向连接的约束的例子是当用户想根据指定电影的两个或更多个演员找出电影(意图类型)或根据演员和该电影赢得的奖项找出电影时。另一个在电子邮件上下文中的例子是如果用户想要最近七天里从特定公司的某些发件人接收到的电子邮件(意图类型)的情况。类似地,还有的例子是如果用户想要预定火车站以及星巴克连锁店附近的宾馆房间(意图类型)的情况。还有的另一个例子是如果用户想要Samsung制造的还与Nintendo Wii兼容的电视机(意图类型)的情况。所有这些都是面向连接的约束查询的实例。

[0060] 在以上面向连接的约束的例子中,用户明确地描述或指定连接到意图实体的其它实体。这种约束在本文被一般地称为面向明确连接的约束并且这些实体被称为明确的实体。

[0061] 同时,其它查询包含包括未指定或隐含实体作为约束规范的一部分的面向连接的约束。在这种情况下,用户试图识别不通过未知项与用户确实知道的项之间的关系而知道的一些信息、实体、属性等。这种约束在本文被一般地称为面向隐含连接的约束并且未指定的实体在本文被一般地称为约束的隐含的实体。

[0062] 例如,用户可能希望通过指出电影中两个人物的名字来识别她正在寻找的电影。但是,该用户不记得其中一个人物的名字,但她确实记得扮演该人物的特定演员。因此,在她的查询中,她通过名字指出一个人物并且通过指出未知人物是由特定演员扮演的来识别该未知人物。但是,考虑以下对特定信息检索目标的用户约束:用户想知道在关于指定角色(例如,人物“Tony Montana”)的未指定的电影中由指定的演员(例如,“Michelle Pfeiffer”)扮演的角色(意图)。在这个例子中,用户的约束包括对应于电影“Scarface”的未指定的或隐含的实体。类似地,假如用户想知道由指定的演员“Scarlett Johansson”

和在指定的电影“Star Wars”中扮演指定角色“Obi-Wan Kenobi”的未指定的演员主演的电影（意图）。在这个例子中，隐含实体是演员“Ewan McGregor”并且意图实体是由“Scarlett Johansson”和“Ewan McGregor”主演的电影“The Island”。

[0063] 在电子邮件库的上下文中，例子包括用户想要得到来自指定公司“Intel”的未指定男士的最后的电子邮件（意图），其中该用户在上星期经电子邮件（属性指示器）被介绍给该未指定的男士。在这个例子中，隐含实体是可经员工/公司关系通过检查来自“Intel”的联系人发现的、上周第一次与该用户是共同电子邮件收件人的联系人。

[0064] 以上三个例子是面向连接的约束，但是它们包括未指定的或隐含的实体作为约束规范的一部分——我们把这种约束称为面向隐含连接的约束并且把未指定的实体称为约束的隐含实体。

[0065] 在面向连接的约束的上下文中，将信息库的实体和关系基本地映射到图的节点和边是有用的。特别地采用图模型代替实体关系模型的动机是观察到自然语言对话中的相关性、近似性和关联性可简单地通过诸如链接-距离以及在一些情况下的最短路径和最小重量树的概念来建模。在对话过程中，当用户对话涉及与实际寻找实体相关的其它实体时，将信息检索作为简单的图搜索问题的子程序有效地帮助减少对句子结构深入明确理解的依赖，这对实现会有巨大的好处。即使用户意图计算是歧义的或不确定的，只要实体已经在用户话语中被识别出，则基于图解释的问题处理就使得我们的系统能够以比其它可能的方式智能得多的方式响应。

[0066] 用于消除用户意图/输入歧义的对话交互界面

[0067] 我们现在描述用于消除用户意图/输入歧义的本发明实施例的对话交互界面。如果用户能够通过利用对信息检索系统说话并且可选地利用触摸或利用键盘或鼠标选择选项来提交查询或指令与信息检索系统交互，则认为它是对话交互界面。对用户查询的响应可以通过机器产生的口语文本到语音的转换来完成，并且可以通过在用户屏幕上显示的信息进行补充。对话交互界面一般而言几乎始终允许用户针对信息检索系统对之前查询的响应做出回应来提交他的下一个信息检索查询或指令，使得信息检索会话是一系列的操作，其中每个操作都是用户首先提交查询或指令接着系统向用户做出响应。

[0068] 本质上，本公开内容中描述的对话交互界面的实现是比用于消除用户输入/意图歧义的图形UI更有效和更令人印象深刻的范例。在许多情况下，尤其当涉及从大量可能的属性或存在的明确和隐含连接的节点中灵活选择时，图形UI方法不会很好地工作或者根本不工作。在这种情况下，对话交互界面是更自然的匹配，此外，随着改进的语音识别技术的来临，这将取悦用户。

[0069] 现在，我们描述用于对话交互的信息检索系统的体系结构、组件以及实现。

[0070] 对话系统体系结构

[0071] 图1表示本发明实施例的整体系统体系结构和基本信息流。用户101说出他/她的问题，该问题被送到语音文本转换引擎102。虽然输入可以是语音，但是本发明并不排除输入是直接的文本输入。用户输入的文本形式被送到会话对话内容模块103。该模块起跨对话维护状态的作用，它的一个用处是在对话过程中帮助理解用户意图，如以下所描述的。会话对话连同语言分析器（或语音标注器（speech tagger）的一部分）106以及以下描述的其它实体识别器模块将句子分解为它的组成部分，组成部分可以被广义地归类为（1）意

图——用户的实际意图,诸如寻找电影、播放歌曲、调到频道、响应电子邮件等 (2) 实体——描述意图的名词或代词短语,及 (3) 属性——对实体的限定词,诸如“最近的”电影、“更少的”暴力等。在以提供智能和有意义的对话为目标的环境中,意图有时是所有三个类别中最重要。任何好的搜索引擎仅仅通过从句子中提取实体——而无需理解语法或意图,就可以相当好地执行信息检索任务。

[0072] 例如,当假定用户询问“我的女儿能和我看 pulp fiction 吗?”,大多数搜索引擎会显示对 pulp fiction 的链接,如果通过那个链接就可以得到等级的话,那么这可能就够了。但是在对话界面中,期望明显更高——系统理想地必须理解与电影等级以及它适合的年龄组的期望响应对应的(电影,等级)意图。退化到搜索引擎那种响应的对话界面响应从用户的角度看相当于系统的失败。当意图未知或不能清楚地辨认时,意图的确定、以及更重要的看起来更接近人类响应的对用户问题的响应对于致力于比搜索引擎更接近人类交互的对话界面来说是关键。

[0073] 意图分析器 108 是分析意图并将意图分类到领域的特定于领域的模块,并且结合其它模块一起工作,这些其它模块包括特定于领域的实体识别器 107、基于用户的个人偏好将意图分类的基于个性化的意图分析器 109 以及特定于领域的图引擎 110。特定于属性的搜索引擎 111 帮助识别属性,并且这些属性的权重影响它们限定的实体。虽然图 1 是显示用于特定领域的模块的对话体系结构,但是本发明的实施例包括可以接受用户输入并进行其中用户意图可以跨领域的对话的对话界面。在本发明的实施例中,这是通过具有多个图 1 中示出的特定于领域的体系结构的实例并且跨领域对意图的权重进行评分以确定用户意图来实现。这种评分机制还用于隐含地确定对话主题切换(例如,在娱乐信息检索会话的过程中,用户可能正好说“我饿了”)。另一个新对话开始被隐含地确定的例子:

[0074] 用户:我们什么时候进行 Yankees 比赛

[0075] 响应:New York Yankees 在 7 月 6 日星期五 (7pm) 与 Boston Red Sox 比赛,可以在 NESN HD 观看

[0076] 用户:你能为我录制这场比赛吗

[0077] 响应:已安排录制 New York Yankees 对 Boston Red Sox 在 7 月 6 日星期五 (7pm) 的比赛。

[0078] 用户:Netflix 上有 Iron Man 电影吗

[0079] 响应:Iron Man 和 Iron Man 2 马上就可以在 Netflix 上得到。你想将它们添加到你的即时队列吗?

[0080] 在其中来自不同领域的响应同样可能的情况下,本公开内容中描述的系统的实施例向用户提出消除歧义的问题:

[0081] 用户:我在寻找好的音乐应用

[0082] 响应:你想要流音乐还是想要关于音乐的新闻和信息?

[0083] 在这个例子中,本公开内容中描述的说明性系统识别出题材“音乐”并且使用知识图跨音乐题材中的流行应用计算两个区分因素:流广播类型产品以及新闻和信息应用。

[0084] 用户:我想要更多关于不同艺术家和乐队的信息

[0085] 响应:这里是流行应用的列表:last.fm, allmusic 等。

[0086] 系统识别短语信息并且将应用结果列表过滤到那些更多基于新闻和信息的应用。

[0087] 语言分析模块 106 在体系结构中是可插拔模块,以使得系统能够支持多种语言。尽管对用户输入的语义的理解不限于语言模块 106 自身,但是诸如对话上下文模块 103 或图引擎 110 的体系结构的核心模块是语言无关的。如前面所提到的,除了句子分析和执行诸如将代词关联到其主语 / 宾语等任务之外,语言模块自身不可以做得更多。(“在近期的 Daniel Craig 电影的原声曲目 (Original Sound Track) (OST) 中的 Led Zeppelin 歌曲…谁演唱的它?”);独自地跨交流关联代词是无效的。是与会话对话上下文模块 103 的交互使得能够如下跨交流地进行代词的解析:

[0088] Q1:谁在新的 star wars 中扮演 obi-wan Kenobi?

[0089] A1:Ewan McGregor

[0090] Q2:他与 Scarlet Johansson 的电影怎么样?

[0091] 尽管乍看起来对话会话上下文可能像是对最后一个或多个活动实体的简单状态维护,但是下面的例子示出了对话会话上下文中潜在的复杂性:

[0092] Q1:谁在 Kramer vs. Kramer 中扮演主角?

[0093] A1:Meryl Streep 和 Dustin Hoffman

[0094] Q2:他的更多电影有哪些?

[0095] A2:这里是一些 Dustin Hoffman 的电影… [Dustin Hoffman 电影列表]

[0096] Q3:她的更多电影有哪些?

[0097] A3:[如果有的话,列出电影列表]

[0098] Q4:他的早期电影有哪些?

[0099] A4:[如果有的话,列出电影列表]

[0100] Q5:她的近期电影有哪些?

[0101] A5:[如果有的话,列出电影列表]

[0102] Q6:在最近过去的的时间里他们再次一起表演过吗?

[0103] A6:[如果有的话,列出电影列表]

[0104] Q7:他们到底曾经一起再次表演过吗?

[0105] 在以上例子中,实体 Meryl Streep 和 Dustin Hoffman 被间接地在六个问题中引用,有时一起并且有时分开。以上例子还说明了本发明实施例与围绕中心主题进行探索式交流的简单请求响应系统的一个重要区别。本发明的实施例不仅解决了交流中的歧义,它们还同时促进了与对对话中早前提到的实体和 / 或意图的隐含引用的探索式交流——那些在丰富的人类交互中自然而然所做的一些事情。

[0106] 以下例子说明用户引用甚至没有在之前交流中明确指出的实体 (面向隐含连接的约束)。

[0107] Q1:哪个节目里有那个提到基于非碳生命形式的可能性的有生理缺陷的科学家?

[0108] A1:那是 Stephen Hawking 的关于外星人的发现频道的节目。

[0109] Q2:他在 David Attenborough 制作的另一个节目中吗?

[0110] 另一个对话状态维护不限于仅实体及其属性的例子是当系统也维护意图的状态时,使得意图被跨对话交流携带,如在以下例子中很明显的:

[0111] Q1:“我的女儿能和我看 pulp fiction 吗”

[0112] A1:Quentin Tarantino 的 pulp fiction 由于图像暴力和裸露被评为 R 级。

[0113] Q2:他的其它电影呢怎样?

[0114] A2:它们都被评为 R 级 - Reservoir Dogs, Jackie Brown, Kill Bill, Death Proof。

[0115] 在这个例子中,除了维护实体“Quentin Tarantino”的状态使得系统能够在 Q2 中理解对他(以“他的”的形式)的代词引用之外,系统还跨交流跟踪用户意图——用户意图是“等级”。正是这个维护促进了如 A2 中的简洁和定向的响应,几乎与人类交互相配。

[0116] 在特定于领域的意图和实体分析器 108, 107 与基于个性化的意图分析器 109 密切协同工作下,以上说明的定向的响应是可能的。这些模块都受到特定于应用的属性搜索引擎 111 的帮助,特定于应用的属性搜索引擎 111 帮助确定相关的属性(例如,最近的、较少的暴力、更多的动作)并且给它们分配权重。因此,在上述所有模块协同工作(查询执行引擎 104 起协调作用)的处理之后,来自语音文本转换引擎 102 的用户输入交流将产生用户输入的一个或多个候选解释。例如,响应于问题“你有关于 Bombay bomb blast 的电影 Kay Menon 吗?”,系统可能有两种可替换的候选表示,其中一种候选表示使“Bombay”作为实体(有叫作 Bombay 的电影)并使“bomb blast”作为另一个属性,另一种候选表示使“Bombay bomb blast”作为单个实体。系统接着通过进行与用户的对话,根据作为演员的另一个识别的实体 Kay KayMenon 的存在,试图在这些候选表示之间进行解析。

[0117] 在某些情况下,歧义的解决可以在不进行对话的情况下通过已知的用户偏好来完成。例如,用户可能询问“今晚有 sox 比赛吗?”,虽然这个问题具有歧义部分——球队是 Boston Red Sox 还是 Chicago White Sox 的歧义——但是如果系统意识到用户的偏好是 Red Sox,那么如果那天晚上有一场 Red Sox 的比赛的话,响应可以定向到显示 Red Sox 的比赛时间表。在其中存在多个跨领域的匹配的情况下,结果产生较高整体信心分值的领域匹配将胜出。当适用时,也可以基于查询的性质做出结果的个性化。例如,如果用户声明“给我显示今晚的 Tom Cruise 的电影”,那么这个查询不应该应用个性化,而是只返回 Tom Cruise 的最近的电影。但是,如果用户声明“给我显示今晚的体育运动”,则系统应当应用个性化并基于用户的明确偏好或从用户活动信息的各种来源捕获到的隐含动作来显示已知的该用户感兴趣的体育运动和比赛。

[0118] 与其中用户的偏好(隐含或明确地推断的)以二进制的方式被应用(像开或关的“开关”)的现有系统不同,本发明的实施例以上下文相关的方式使用用户偏好签名(在本公开内容中也称作个人图,其捕获用户活动和兴趣,包括隐含地确定的和明确地确定的两者)来解决用户输入中的歧义,并且如果适用的话,还对结果选择应用个性化以提供具有高可能性匹配用户意图的最佳响应。如果可用的话,本发明的某些实施例使用用户偏好签名来解决用户输入中的歧义。但是,在刚提到的消除歧义步骤之后,为了裁剪结果而使用签名非常依赖于用户输入中指定的实体的定义的精确度水平。

[0119] 用户偏好签名可以由系统利用已知的用于发现和存储这种用户偏好信息的技术来提供。例如,在于 2010 年 8 月 10 日授权的标题为 Methods and Systems for Selecting and Presenting Content Based on Learned Periodicity of User Content Selections 的美国专利 No. 7, 774, 294、于 2010 年 11 月 16 日授权的标题为 Methods and Systems for Selecting and Presenting Content on a First System Based on User Preferences Learned on a Second System 的美国专利 No. 7, 835, 998、于 2008 年 12 月 2 日授权的标题

为 User Interface Methods and Systems for Selecting and Presenting Content Based on User Navigation and Selection Actions Associated with the Content 的美国专利 No. 7, 461, 061 以及于 2012 年 2 月 7 日授权的标题为 Methods and Systems for Ordering Content Items According to Learned User Preferences 的美国专利 No. 8, 112, 454 中阐述的方法和系统可以与本文所公开的技术一起使用, 其中每个专利都通过引用被结合于此。但是, 用户偏好签名和 / 或信息的使用不限于在所结合的申请中阐述的技术。

[0120] 关系或连接引擎 110 是起理解用户输入以提供定向的响应作用的模块之一。该关系引擎可以以多种方式实现, 图数据结构是其中一个例子, 使得我们可以通过名称图引擎来称呼关系引擎。图引擎在实体之间已知的加权连接的背景下评估用户输入。

[0121] 实体定义的精确度水平在图引擎 110 的节点中捕获。图引擎 110 中的每个实体节点都被分配有歧义指数, 这是为实体统计上确定的分值——并且可以是连续区间的值, 比如从“低”值到“高”值, 其中“低”意味着低歧义性并且“高”意味着高歧义性, 以及这些端点限定值之间的所有中间值。这个歧义指数用于确定何时可以利用个人图 (如果有的话)。一个例子是下面的对话:

[0122] 用户: 今晚有比赛吗? (或) 今晚我们进行比赛吗?

[0123] 响应: Boston Red Sox 与 Florida Marlins 今晚 (7pm) 进行比赛, 可以在 ESPN HD 观看。

[0124] 在这个例子中, 用户输入的“sox”具有高歧义指数。系统将动词短语“有比赛吗”映射到实体类型体育运动并且基于该用户的个人图为他 / 她关联实体 Boston Red Sox。个性化的决定受“sox”歧义指数的驱动。形容词“今晚”充当时间指示器用于改进查询。注意, 尽管用户的输入有歧义, 但是在基于个人偏好将用户的意图解析为“red sox”之后, 输入不再有歧义 (假定“red sox”的“低”歧义分值)。因此, 由于歧义指数现在低了 (在映射到 Boston red sox 之后), 所以结果没有被个性化。在可替换的变化中, 代词“我们”被与实体 Boston Red Sox 相关联。

[0125] 另一个例子如下:

[0126] 用户: 我们什么时候进行 Yankees 比赛

[0127] 响应: New York Yankees 在 7 月 6 日星期五 (7pm) 与 Boston Red Sox 比赛, 可以在 NESN HD 观看

[0128] 用户: 你能为我录制该比赛吗

[0129] 响应: 已安排录制 New York Yankees 对 Boston Red Sox 在 7 月 6 日 (7pm) 星期五的比赛。

[0130] 在这个例子中, 系统提取实体 New York Yankees 并且基于用户个人图使代词“我们”归到实体 Boston Red Sox。

[0131] 以上执行的个性化是基于用户过去活动的签名, 包括在社交网站、媒体消费、短信、推特 (tweeting) 活动等中的, 并且还包含电子邮件、日历预约项、任务 / 待办事项列表、文档等用户个人语料库的签名。如前面所描述的, 虽然用户的偏好签名在一些情况下用于解决用户输入中的歧义 (例如, 今晚有“sox”比赛吗, “我们”今晚比赛吗), 但是在解决用户输入歧义之后, 仍然具有高歧义指数的实体的出现确定是否应该利用用户的个人图来剪裁结果来匹配用户的意图。例如, 即使用户在他的个人图中有 Tom Cruise 和 Demi Moore,

以下的查询也将不会触发结果个性化——这是因为用户的意图是明确的和无歧义的。

[0132] 用户:Tom Cruise 与 Demi Moore 合作过吗?

[0133] 由于在这个查询中的实体具有低歧义指数,因此对它的响应将不应用用户的个人图信息。但是,对于以下查询:

[0134] 用户:今晚有 sox 比赛吗?

[0135] 由于“sox”具有与它相关联的歧义指数,因此将应用个性化。下面提供了帮助解决用户输入中的歧义并且个性化结果以匹配用户兴趣的用户个人图的更多例子:

[0136] 用户:“sox 何时与 San Fransisco Giants 比赛”

[0137] 情况 1:Red Sox 在用户的签名中

[0138] 响应:“Boston Red Sox 这个赛季不与 San Francisco Giants 比赛”

[0139] 情况 2:Red Sox 不在用户的签名中

[0140] 回答:“你是指 Boston Red Sox 还是 Chicago White Sox”

[0141] 注意,虽然明确指出了其中一个实体 San Francisco Giants(歧义指数是“低”),但是我们仍然需要使用个性化来消除其它实体“sox”(其基于“高”歧义指数)的歧义。总而言之,“高”的“歧义指数”意味着使用“个人图”来解决歧义,但是一旦歧义被解决,如果它变成“非常精确指定的实体”的情况,那么就不使用“个性化”来计算回答。但是,如果甚至在消除歧义步骤之后歧义指数仍然高,则应用个性化。

[0142] 歧义消除的例子

[0143] 图 4 说明命名实体之间的连接和链接的图的一部分。在用户提出请求“给我显示 Beethoven 电影”的例子中,歧义本质上是词汇的——该问题与关于作曲家 Beethoven 的电影和有名为 Beethoven 的狗的电影相匹配。图 4 示出了命名的实体“Beethoven”,其具有到这个节点的链接,链接在抽象意义上表示作曲家 Beethoven 的系统“心智模型(mental model)”。存在一个等效的子图,其具有称为“Beethoven”的实体,它的链接和属性清楚地将它识别为犬。当用户输入与这些不同实体相匹配时(这两者之间的图距离连同这些节点的属性的相关性是这些实体的接近程度的量度),该系统意识到歧义更可能是词汇上的并且利用这两个节点的关键的区分属性——人和狗——来提交消除这两个节点的歧义的适当响应。一旦这个关键的消除歧义的差异被系统推断出,则产生的响应可以变化。在一个实施例中,系统可以提交消除歧义的问题“你是指作曲家还是名为 Beethoven 的狗?”,接着针对用户的反馈进行响应。

[0144] 在另一个实施例中,通过说“如果你指音乐家,则这里是两个关于他的电影<并且列出/说出电影>。如果你指狗,则这里是关于这个狗的电影<并且列出/说出电影>”,系统可以将歧义消除和回答合并到单个响应中。这种响应与“搜索引擎级别的智能”明显不同,其中后者会以某种相关性的顺序列出这两种结果,并且甚至是混合的(作曲家 Beethoven 的结果,后面跟着关于狗的电影,后面跟着另一个 Beethoven 的电影),系统没有展示理解在用户的问题中存在人和狗之间的歧义。

[0145] 在另一个例子中,“谁是 Micheal Jones?”,用户想知道特定的 Micheal Jones,并且存在对用户问题的多个词汇匹配。其中一个匹配是美国医疗保健执行官和保守党策略分析师。另一个是参加了电视节目“American idol”的澳大利亚歌手-歌曲作者。由于用户的真正意图是“谁是 Micheal Jones?”,因此系统用消除歧义的问题“你是指 American

idol 参加者还是保守党策略分析师”响应看起来会很滑稽——因为用户的真正问题指示用户也不知道。在这个例子中，本实施例将起草消除歧义响应，其兼作对用户所问问题的真正答案：“有两个杰出人物有那个名字——一个是参加了 American Idol 的歌手 - 歌曲作者，另一个是保守党策略分析师”。

[0146] 对于另一个例子——“我想看谁想成为百万富翁”，存在三个合格的匹配——美国比赛节目、英国比赛节目和新加坡比赛节目。在一种情景中，其中用户偏好签名是未知的但是用户位置是已知的，正确的比赛节目可以被隐含地选取（例如对在美国的人来说选美国的比赛节目等）。但是，如果用户的偏好签名存在，则签名可以超越用户位置。例如，用户可能在出差访问英国，但是仍然想看来自他国家的版本，比如美国。如果位置和用户偏好都不可用，则系统将提交以下类型的消除歧义问题：“在三个国家有三个不同的节目具有这相同的名字。你想看哪一个？ < 并且显示 / 说出这三个节目 >”。

[0147] 在另一个例子中，“给我讲讲澳大利亚”，在本公开内容中描述的特定于领域的结构化知识帮助系统识别在国家澳大利亚与具有相同名称的两个电影（在 1989 年拍摄的英国电影“澳大利亚”和在 1992 年拍摄的具有相同名字的马拉雅拉姆语 (Malayalam) 的印度电影）之间存在冲突。在这个例子中，用户偏好签名以及历史，特别是从用户的个人语料库中收集的信息可以指示用户将在不远的将来去澳大利亚旅行。这个信息结合对澳大利亚是指国家的理解，将被系统用来隐含地消除用户问题的歧义，以表示国家澳大利亚，并直接显示关于该国家的响应。这里，利用个人语料库和特定于领域的知识的消除歧义甚至消除了澄清交流，使得系统响应更接近于人类交互。

[0148] 本文公开的技术和系统可以实现为与计算机系统或计算机化的电子设备使用的计算机程序产品。这种实现可以包括一系列的计算机指令或逻辑，这些指令或逻辑或者固定在有形介质上，诸如计算机可读介质（例如，盘、CD-ROM、ROM、闪存存储器或其它存储器或固定盘），或者可以经诸如通过介质连接到网络的通信适配器之类的调制解调器或其它接口设备发送到计算机系统或设备。

[0149] 介质可以是有形介质（例如，光学或模拟通信线路）或利用无线技术（例如，WiFi、蜂窝、微波、红外或其它传输技术）实现的介质。该一系列计算机指令体现本文所描述的关于该系统的至少一部分功能。本领域技术人员应当理解，这种计算机指令可以用与多种计算机体系结构或操作系统一起使用的多种编程语言编写。

[0150] 此外，这些指令可以存储在任何有形的存储器设备中，诸如半导体、磁、光或其它存储器设备，并且可以利用任何通信技术（诸如光、红外、微波或其它传输技术）进行传输。

[0151] 可以预期，这种计算机程序产品可以分发为具有伴随的打印或电子文档（例如，压缩打包软件）的可拆卸介质、用计算机系统（例如，在系统 ROM 或固定盘上）预加载、或通过网络（例如，因特网或万维网）从服务器或电子公告板分发。当然，本发明的一些实施例可以实现为软件（例如，计算机程序产品）和硬件两者的组合。本发明还有的其它实施例完全实现为硬件、或完全实现为软件（例如，计算机程序产品）。

[0152] 此外，本文公开的技术和系统可以与多种多样的移动设备一起使用。例如，本文所讨论的移动电话、智能电话、个人数字助理和 / 或能够接收信号的移动计算设备可以在本发明的实现中使用。

[0153] 本发明的各个方面和实施例可以与以下申请中阐述的技术结合使用，其中所有申

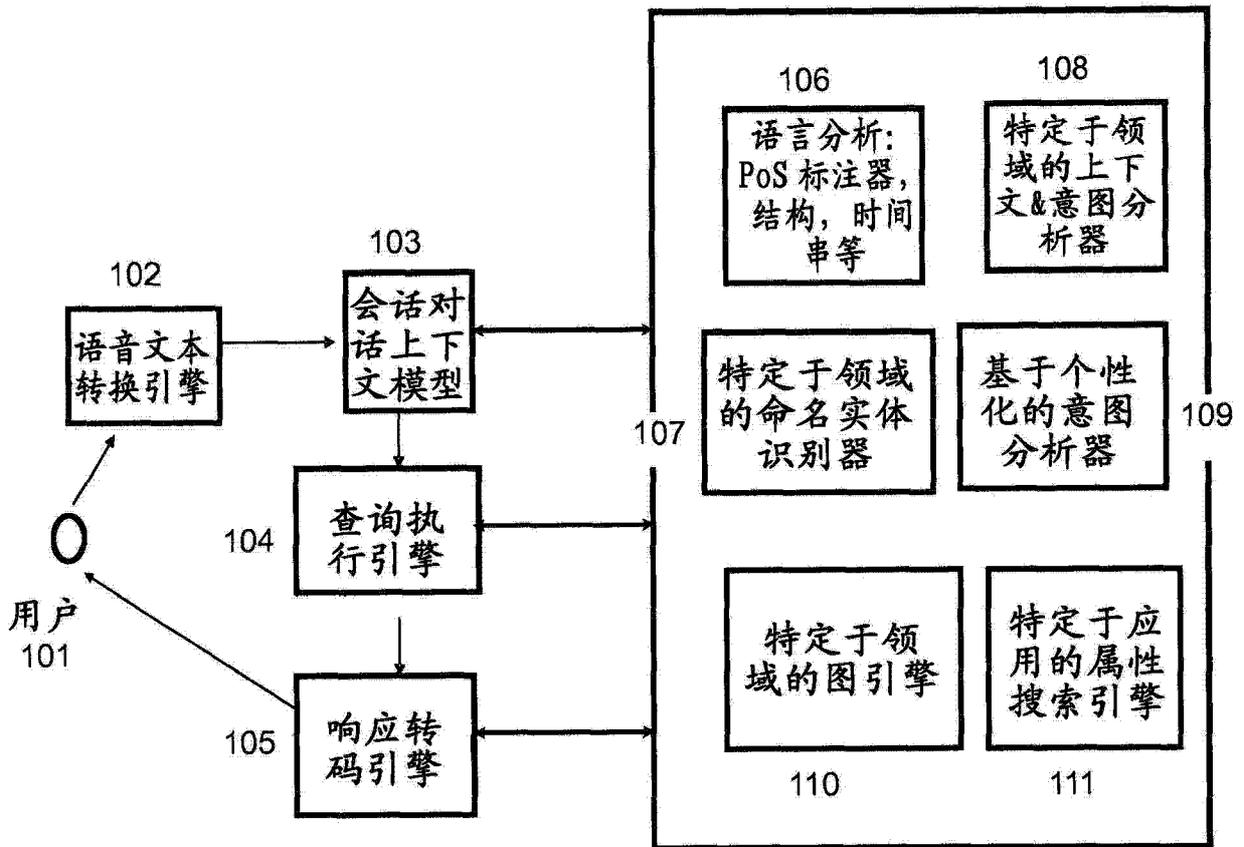
请都通过引用被结合于此：

[0154] 于2012年7月20日提交的标题为“A Conversational Interaction System for Large Corpus Information retrieval”的美国临时申请 No. 61/673,867；

[0155] 于2010年9月10日提交的标题为“Method of and System for Presenting Enriched Video Viewing Analytics”的美国专利申请 No. 12/879,141；及

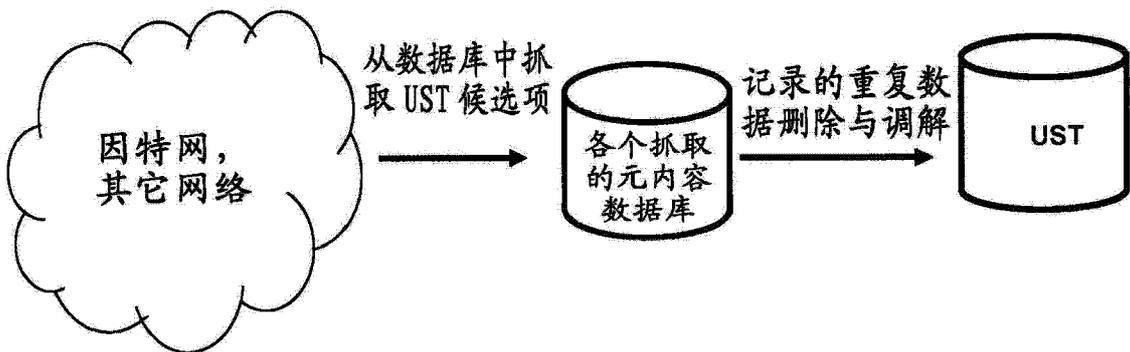
[0156] 标题为“Methods and Systems for Selecting and Presenting Content Based on Learned Periodicity of User Content Selections”的美国专利 No. 7,774,294。

[0157] 如将对阅读本公开内容的本领域普通技术人员显而易见的，本公开内容可以体现为不同于那些以上具体公开的形式。上述特定的实施例因此要被认为是说明性的而不是限制性的。本领域技术人员将仅仅利用常规实验就认识到或者能够确定本文所描述的具体实施例的多种等价物。本发明的范围是如在权利要求及其等价物中阐述的，而不是被限制到前面描述中所包含的例子。



对话交互体系结构

图 1



抓取累积以及创建去重复数据的库 UST

图 2

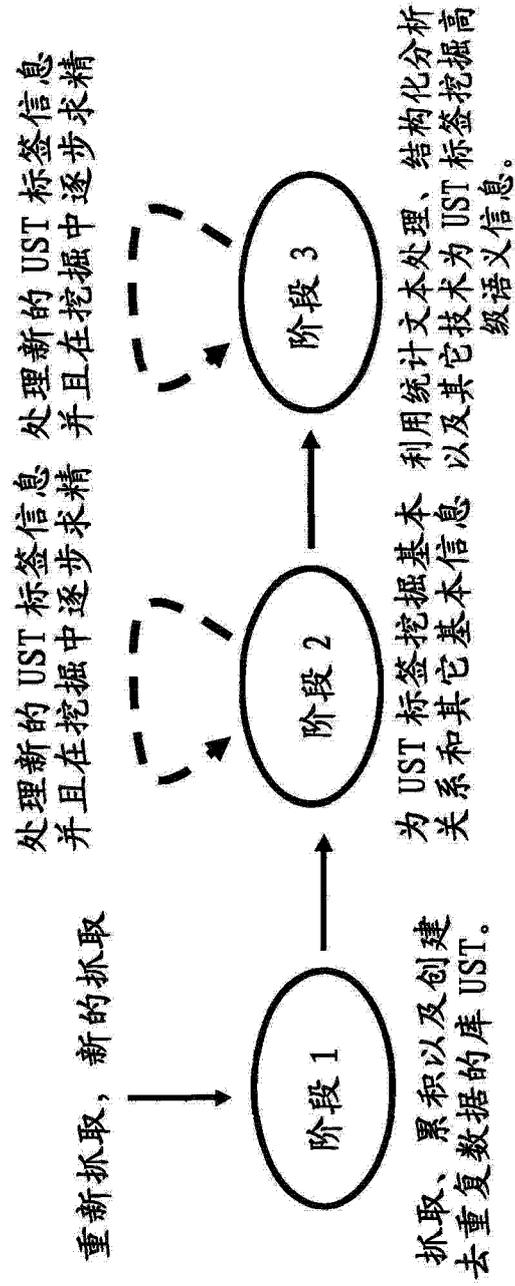
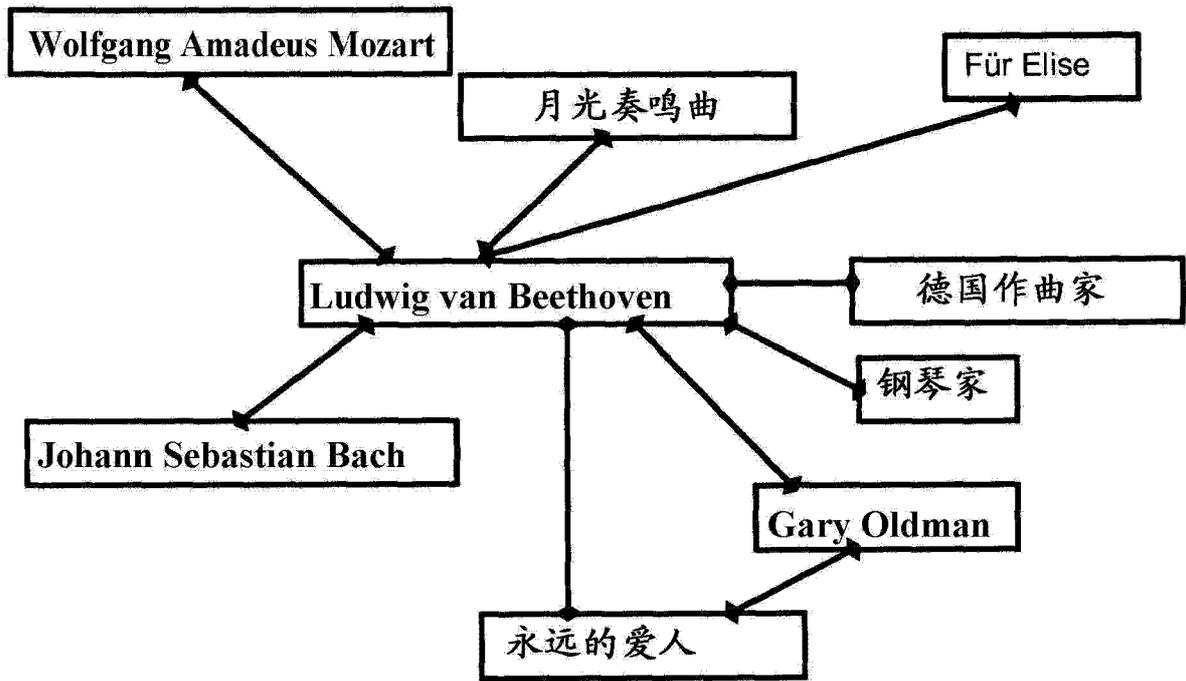


图 3



连接到 Beethoven 的一些 UST 记录

图 4