

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 17/30 (2006.01)

G06F 17/27 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200510100419.5

[43] 公开日 2007年4月25日

[11] 公开号 CN 1952928A

[22] 申请日 2005.10.20

[21] 申请号 200510100419.5

[71] 申请人 梁威

地址 417100 湖南省涟源市六亩塘镇塘湾村
164号

[72] 发明人 梁威

[74] 专利代理机构 深圳市金阳行专利商标事务所
代理人 金辉

权利要求书2页 说明书21页 附图3页

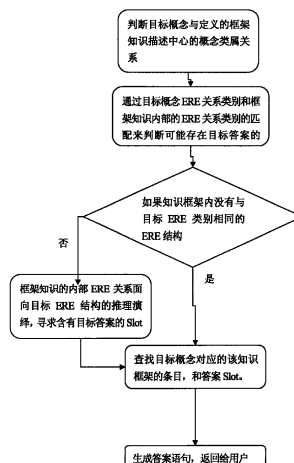
[54] 发明名称

建立自然语言知识库及其自动问答检索的计算机系统

[57] 摘要

本发明涉及一种基于 HNC 自然语言理解的知识库建立和自动问答检索系统，是一种基于 HNC 自然语言处理技术，在 HNC 对自然语言处理的基础上，从语句中抽取 ERE (Entity - Relation - Entity) 知识，建立 ERE 知识库，并在 ERE 知识的基础上定义和建立以知识框架的形式描述知识的框架知识库。基于这两种知识库，系统接受用户以自然语言方式的查询请求，通过对于疑问句的分析，目标 ERE 结构的抽取，待选篇章 ERE 的融合和推理演绎，面向目标 ERE 的相似度计算等处理后，返回给用户按照准确度降序排序的答案列表。本发明由于系统接受用户以自然语言提问的查询请求，使得用户可以方便而且精确地定义自己的查询意图；本发明由于在语言的复杂形式之上抽取 ERE 的语义表达式，使

得对于语义的处理能够脱离语言形式复杂性的麻烦。



得对于语义的处理能够脱离语言形式复杂性的麻烦。

1. 一种建立自然语言知识库的计算机系统，通过对获取的各种篇章文本进行 HNC 自然语言处理的语句标注，其特征在于，还根据从所述 HNC 自然语言处理的语句标注中提取 ERE 知识表达式，建立 ERE 知识库；所述 ERE 知识表达式是包括 E1、E2 和 R 的三元组形式，其中 R 相当于一个逻辑谓词，表示 E1 与 E2 之间的语义关系；E1 和 E2 可以表示任何语义元素，如语句、语义块、语义块的构成部分、词语的组合、词语、或者是另外一个 ERE 知识表达式；E1, E2 可以是单个的语义元素，也可以是多个语义元素的组合。

2. 根据权利要求 1 所述的建立自然语言知识库的计算机系统，其特征在于，所述根据 HNC 自然语言处理的语句标注中提取 ERE 知识表达式，包括抽取句类 ERE；抽取辅块与主块关系的 ERE；和抽取语义块内部各个词语概念之间关系的 ERE，通过记录所述 ERE 知识表达式的 E1,R,E2 与其所对应的语句部分的引用，ERE 来源的篇章，引用语句标识 ID 来建立 ERE 知识库。

3. 根据权利要求 2 所述的建立自然语言知识库的计算机系统，其特征在于，所述根据 HNC 自然语言处理的语句标注提取 ERE 知识表达式，还包括应用 ERE 推理演绎推理出新的隐含 ERE 知识，和同一个 ERE 知识的不同的表达形式；上述应用 ERE 推理演绎推理出新的隐含 ERE 知识表达式包括：句类 ERE 之间的推理演绎，按照推导特性而进行动词分组的句类 ERE 的推理演绎，和句类 ERE 向偏正类 ERE 的推理演绎；并将推理演绎所得到的新的 ERE 知识，按照 ERE 的结构定义填充到 ERE 知识库中。

4. 根据权利要求 1、2 或者 3 所述的建立自然语言知识库的计算机系统，其特征在于，还根据从所述 HNC 自然语言处理的语句标注中建立的 ERE 知识库的基础上建立知识框架结构，按照框架知识库定义的要求建立框架知识库，求解得到框架知识库所要求的特定的 ERE 知识；所述知识框架结构以一类事物为描述中心，并定义了该类事物的属性（Slot）结构，每个 Slot 与描述中心之间，不同 Slot 之间的 ERE 语义关系，所述框架知识库还定义了每个知识框架的 Slot 对应的目标 ERE 的特征和从目标 ERE 条目向 Slot 映射的方式。

5. 根据权利要求 4 所述的建立自然语言知识库的计算机系统，其特征在于，所述描述中心知识框架的属性（Slot）结构包括：单值结构、组合结构、列举结构、嵌套框

架结构和复合结构;单值结构就是一个 Slot 简单地对应着一个属性的值;组合结构指 Slot 对应多个值的组合;列举结构指 Slot 对应着一个相同的值结构的队列;嵌套框架结构指一个 Slot 的值由另外一个知识框架来表示;复合结构是指 Slot 的值由单值结构、组合结构、列举结构和嵌套框架结构四种结构混和组合而成。

6. 一种建立自然语言知识库自动问答检索的计算机系统,其特征在于,包括如下处理步骤:第一步,对疑问句进行分析获得包含目标答案要求的目标 ERE 结构;第二步:分析目标概念在待选篇章中的出现概率和待选篇章中含有目标概念的句群分布情况,对待选篇章含有目标概念的程度进行初步评价;第三步,发现篇章内部分散的可以通过指代、相同概念、语义块共享等方式融合的语义元素,并将相关的 ERE 做融合处理;第四步,通过综合计算篇章与目标 ERE 结构的相似程度,篇章内各个句群对于答案的回答程度得到篇章的综合得分;第五步,按照篇章的综合得分排序返回答案篇章的列表。

7. 根据权利要求 6 所述的建立自然语言知识库自动问答检索的计算机系统,其特征在于,在所述第三步结束后,还包括在融合的基础上做面向目标 ERE 结构的推理演绎分析,从待选篇章融合的关于目标概念的 ERE 结构不断向前推演,获得更多待选 ERE,直到推理的过程不能够产生新的 ERE 知识时停止,再执行第四步。

8. 根据权利要求 6 所述的建立自然语言知识库自动问答检索的计算机系统,其特征在于,所述第一步对疑问句分析包括:对疑问句的句类分析、对疑问句的疑问中心的分析、对疑问句的 ERE 结构的抽取,和对于疑问句的 ERE 结构应用推理转换规则,得到一个范围扩大的目标 ERE 结构。

9. 根据权利要求 6 所述的建立自然语言知识库自动问答检索的计算机系统,其特征在于,在完成第一步以后执行下述步骤:第二步,通过对目标 ERE 结构的概念分析来判断可能含有目标答案的各种框架知识类型;第三步,通过目标 ERE 结构的 ERE 关系和概念与框架知识库的知识条目的概念匹配得到可能含有目标答案的知识条目,并从指定知识条目的指定 Slot 得到答案值;第四步,生成答案语句,返回用户。

10. 根据权利要求 9 所述的建立自然语言知识库自动问答检索的计算机系统,其特征在于,完成所述第三步的方法是将可能含有目标答案的框架知识库 Slot 的 ERE 关系做面向目标 ERE 结构的演绎推理,再通过匹配识别得到知识库中含有目标答案的 Slot。

建立自然语言知识库及其自动问答检索的计算机系统

技术领域

本发明涉及一种计算机系统，尤其涉及建立自然语言知识库及其自动问答检索的计算机系统，基于 HNC 自然语言处理技术，通过对自然语言处理建立知识库，并接受以自然语言方式提问的查询请求，自动给出答案。

背景技术

目前人们用于查询信息的工具如查询软件，搜索引擎如 google 等，采用的主要是关键词匹配，网页链接分析，统计分析等技术在浩瀚的信息海洋中查找用户想要的内容。但是人们很难通过简单的关键词组合来准确定义自己的搜索意图，而且关键词匹配的搜索过程没有处理词意组合，语句内部的语义关系等语义因素。所以人们经常不得不花大量时间对数目巨大的结果网页列表再作人工辨识来寻找理想的答案。

所以一种使用自然语言定义搜索意图，由计算机直接给出自然语言的理想答案的查询软件将为人们节省大量的时间。

现有的自然语言处理技术将语句中的各个部分的语言学角色标注出来，这种语句和句群的标注对于知识的表示仍然保留着对于自然语言在形式上的依赖和自然语言形式的繁琐复杂，所以不利于建立一个高效，统一的知识处理的模型。

发明内容

本发明的目的在于建立一个高效，统一的知识处理的模型，生成一种建立自然语言知识库的计算机系统。

本发明公开的一种建立自然语言知识库的计算机系统，通过对获取的各种篇章文本进行 HNC 自然语言处理的语句标注，还根据从所述 HNC 自然语言处理的语句标注中提取 ERE 知识表达式，建立 ERE 知识库；所述 ERE 知识表达式是包括 E1、E2 和 R 的三元组形式，其中 R 相当于一个逻辑谓词，表示 E1 与 E2 之间的语义关系；E1 和 E2 可以表示任何语义元素，如语句、语义块、语义块的构成部分、词语的组合、词语、或者是

另外一个 ERE 知识表达式；E1, E2 可以是单个的语义元素，也可以是多个语义元素的组合。

本发明公开的一种建立自然语言知识库的计算机系统，还根据从所述 HNC 自然语言处理的语句标注中建立的 ERE 知识库的基础上建立知识框架结构，按照框架知识库定义的要求建立框架知识库，求解得到框架知识库所要求的特定的 ERE 知识；所述知识框架结构以一类事物为描述中心，并定义了该类事物的属性（Slot）结构，每个 Slot 与描述中心之间，不同 Slot 之间的 ERE 语义关系，以及从 ERE 知识中抽取 Slot 的方式，所述框架知识库还定义了每个知识框架的 Slot 对应的目标 ERE 的特征和从目标 ERE 条目向 Slot 映射的方式。

本发明公开的一种建立自然语言知识库自动问答检索的计算机系统，包括如下处理步骤：第一步，对疑问句进行分析获得包含目标答案要求的目标 ERE 结构；第二步：分析目标概念在待选篇章中的出现概率和待选篇章中含有目标概念的句群分布情况，对待选篇章含有目标概念的程度进行初步评价；第三步，发现篇章内部分散的可以通过指代、相同概念、语义块共享等方式融合的语义元素，并将相关的 ERE 做融合处理；第四步，通过综合计算篇章与目标 ERE 结构的相似程度，篇章内各个句群对于答案的回答程度得到篇章的综合得分；第五步，按照篇章的综合得分排序返回答案篇章的列表。

本发明公开的建立自然语言知识库自动问答检索的计算机系统，在完成上述第一步以后执行下述步骤：第二步，通过对目标 ERE 结构的概念分析来判断可能含有目标答案的各种框架知识类型；第三步，通过目标 ERE 结构的 ERE 关系和概念与框架知识库的知识条目的概念匹配得到可能含有目标答案的知识条目，并从指定知识条目的指定 Slot 得到答案值；第四步，生成答案语句，返回用户。。

本发明专利与以往相类似技术相比，由于系统接受用户以自然语言提问的查询请求，使得用户可以方便而且精确地定义自己的查询意图；由于系统对查询的问句做问句分析，能够识别问句的语义关系和目标答案的要求；系统在语言的复杂形式之上抽取 ERE 的语义表达式，使得对于语义的处理能够脱离语言形式复杂性的麻烦；系统通过 ERE 知识的演绎推理方式来发现语义之间的蕴涵关系，使得系统对于语言中隐知识的发现具有良好的可扩展性能；系统通过篇章内部的 ERE 的融合来整体地把握一个篇章对于目标语义的综合表达，增强了对于语义处理的深度和广度；系统通过 ERE 组合之间的相似度计算来匹配寻求目标答案，使得答案符合语义和知识的要求。

本发明建立自然语言知识库及其自动问答检索的计算机系统通过 HNC 自然语言处理, ERE 抽取, 框架知识库知识建立三个层次对自然语言进行处理, 使得信息逐步被结构化为知识, 将计算机难以处理的自然语言信息逐步转换为能够灵活被计算机处理的知识库。由于本发明采用的 ERE 三元表达式的方式统一简洁的表达多样化的语言形式所蕴涵内容, 不依赖于语言的外在表达形式, 有利于设计高效的针对 ERE 结构的知识处理系统。本发明基于 ERE 统一的知识表达模式, 定义一系列各种语义情况下的蕴涵推导的产生式规则, 从已有的语义 ERE 组合中推导得到新的语义 ERE, 从而加深计算机对于自然语言的语义的理解。本发明还提出了一种篇章内语义融合的技术, 通过发现篇章内的相同概念, 指代引用, 语义块共享等语言现象所蕴涵的语义元素之间的等语义性质, 将各个在篇章内部位置分散的语义 ERE 按照等语义元素进行融合, 从而给出准确的答案。本发明通过对用户问句进行疑问句分析, 目标 ERE 提取, 对待选篇章的初步评价, 在待选篇章中面向目标 ERE 结构的推导演绎和识别匹配, 在框架知识库中对于目标 ERE 的求解分析, 最终自动返回给用户符合语义、符合知识的答案。本发明通过统计目标概念在待选篇章的出现概率等方法, 来初步评价各个待选篇章含有目标答案可能性, 减小了计算量。

附图说明

本发明包括如下附图:

图 1 是篇章内相同概念融合处理方法示意图;

图 2 偏正类 ERE 与偏正类 ERE 语句比较流程图;

图 3 句类 ERE 与句类 ERE 之间的比较流程图;

图 4 基于框架知识库的自动问答流程图。

具体实施方式

下面结合附图对本发明做进一步详细说明。

本发明公开的技术方案是基于 HNC 自然语言理解的知识库建立和自动问答检索计算机系统, 自然语言是一种非结构化的信息形式, 通过对其使用 HNC 技术进行自然语言处理, 得到语句的句群, 句类, 语义块, 语义块结构等标注信息。本发明在此自然语言处理的基础上使用 ERE 表达式来表达语义知识。在 HNC 句类分析结果的基础上抽取

ERE 知识，进而构建 ERE 知识库，定义和填充框架知识库。

先对 ERE 知识表达式和框架知识库说明。ERE (Entity-Relation-Entity) 是一种三元组形式的知识表达方式，它的 R 相当于一个逻辑谓词，表示 E1 与 E2 之间的语义关系，由系统定义其含义和运算性质。E1 和 E2 可以表示任何语义元素，如语句，语义块，语义块构成部分，词语的组合，词语，或者是另外一个 ERE，所以 ERE 可以通过嵌套组合的方式来表达语义。

E1, E2 可以是单个的语义元素，也可以是多个语义元素按照与、或关系的组合。ERE 之间具有向上继承的层次结构，子类 ERE 如果不声明某个属性的值，则将隐式地继承父类 ERE 对应的属性值，如果显式地声明了自己的属性，则将以自己的属性值覆盖掉父类 ERE 的对应属性。ERE 的结构定义举例如下表：

ERE 名称:	ModifyERE(偏正关系)
父类 ERE	RootERE
E1 说明	偏正语，表示对于某种属性的偏正，如颜色，形状，性质，状态等的偏正。
R 说明	偏正关系，ModifyERE 有子类表示对于 E2 的某个特定属性的偏正。
E2 说明	被偏正的 entity，可以是另外一个 ERE，或者 ERE 的组合。
整个 ERE 的意义描述:	ModifyERE 表示了 E1 所表示的属性特性对于 E2 的偏正。
ERE 抽取来源的结构特征描述	ModifyERE 抽取来源于偏正关系的语言结构，以“的” ¹⁴¹ 概念为特征的偏正关系。 偏正关系的偏正词语为 u, x, z 等类型的概念，被偏正语为 p, w, g 等类型概念。
ERE 抽取映射的规则	语义结构的偏正关系中偏正语对应 E1，被偏正语（核心语）对应 E2。
相关的 ERE 推导规则	

从 HNC 句类分析结果中抽取 ERE 知识的处理过程包括：首先获取网上或其他内容源的篇章文本，再进行 HNC 自然语言处理的语句标注；从 HNC 句类处理结果中抽取 ERE 知识；应用 ERE 推理演绎规则，推理出新的隐含 ERE 知识，ERE 推理规则在一个基于规则(Rule Based)推理机系统上运行；系统应用 ERE 推理演绎推理出新的隐含 ERE 知识，或者和不同的表达形式。

从 HNC 句类处理结果中抽取 ERE 知识包括：

句类 ERE 的抽取方法:

句类按照组合结构分为基本句类, 混合句类, 复合句类。

对于基本句类, 按照系统定义的句类与 ERE 关系抽取。

对于混合句类, 分别取两个句类的各自对应的语义块构造两个句类 ERE。

如: 混合句类 $XP01*211J = A + XP01 + PBC$

将抽取得到两个句类 ERE: $\langle A \ X \ NULL \rangle$, $\langle NULL \ P01 \ PBC \rangle$

对于复合句类, 分别取两个句类的对应部分语义块构造各自的句类 ERE, 如果复合句类两个句类间共享语义块, 那该语义块也将被两个句类 ERE 共享。

如: 复合句类: $(T2b+Y0)*1J$

$$T2bJ = TA + T2b + TB2$$

$$Y0J = YB + Y0 + YC$$

将抽取得到两个句类 ERE: $\langle TA \ T2b \ TB2 \rangle$, $\langle YB \ Y02 \ YC \rangle$ 其中第二个 ERE 共享第一个 ERE 的第一个语义块。

主辅语义块关系 ERE 抽取:

按照辅语义块的类别, 系统定义了手段, 工具, 途径, 条件, 参照, 因果几种辅块及其子类偏正主块的偏正类 ERE。按照句类分析结果对应抽取。

词语概念搭配 ERE 抽取方法:

词语概念搭配关系有偏正类 ERE, 逻辑组合类 ERE。

对于偏正类 ERE, 系统定义了偏正类 ERE 的若干子类, 按照句类分析结果对应地抽取偏正类 ERE。

逻辑组合类 ERE 有与、或关系和其他逻辑类概念(1 类概念)所对应的逻辑组合 ERE 知识表达式。

HNC 类 ERE:

系统定义了作用者, 对象, 内容, 指代, 句蜕, 块扩及其子类几种 HNC 类 ERE, 按照句类分析结果抽取 ERE。

ERE 的推理有如下几类:

句类 ERE 之间的推理规则:

句类 ERE 之间的推理规则定义了从一个句类 ERE 知识推导出隐含的另外一个句类 ERE 知识的方法。这类规则适用与该句类的所有语句的语义推导。

按照推导特性进行动词(LV 类概念)分组的句类 ERE 的推理规则:

某些动词具有相同的语义蕴涵特性, 系统将这类动词按照语义蕴涵推导特性进行分组, 分组后该组的推理规则的定义与句类 ERE 的推理规则定义相同。

句类 ERE 向偏正类 ERE 的推理演绎:

句类 ERE 向偏正类 ERE 的转换的主要目的是为了实现在句类 ERE 与偏正类 ERE 以及词语概念的直接比较, 以计算语义相似度。

就如语句都可以表现为一个要素句蜕的形式, 句类 ERE 都可以转换为相近语义的偏正类 ERE (或 ERE 组合)。

如语句: 超音速协和飞机提高了人们的旅行速度。

句类 ERE: <A X B>

可以转换为偏正类 ERE:

<<A 作用者对象 B> 偏正 X>: 超音速协和飞机将人们的旅行速度的提高。

作用者对象 ERE 是 HNC 类 ERE 的一种, 表示 E1, E2 是同一个 V 概念的作用者与对象的关系。

ERE 知识库建立包括系统将抽取的 ERE 条目标记在数据库之中, 可以方便地按照 ERE 类别, 来源的篇章, E1, R, E2 的概念建立各种索引。ERE 条目将记录 E1,R,E2 与其所对应的语句部分的引用, 和 ERE 条目来源的篇章, 语句标识 ID 的引用。

首先需要进行疑问句分析, 将得到一个包含疑问中心要求的目标 ERE 结构。包括如下步骤: 对疑问句的 HNC 句类分析; 对疑问句的疑问中心的分析和对疑问句的 ERE 结构的抽取。

疑问句的 HNC 句类分析使用 HNC 句类分析方法, 疑问句的 ERE 结构抽取的方法与从 HNC 句类分析结构中抽取 ERE 知识相同。

疑问句的疑问中心的分析: 不同的疑问词所引导的疑问句结构对目标答案具有不同的要求, 为了便于针对分析, 系统定义了两个概念: 疑问中心, 疑问中心词。疑问中心词是指疑问句中疑问词所引导和偏正的词语。如问句: 他去了什么地方? 其疑问中心词为地方; 疑问中心是指疑问词和疑问中心词所组成的结构。如问句: 他去了什么地方? 疑问中心为什么地方。系统通过对疑问中心和疑问中心词的分析来得到目标答案的概念和所要求的语义结构, 并以此与待选句的对应结构匹配比较, 作为计算待选句答案准确

度的一个重要因素。

有关疑问句分析如下表:

疑问词	频率	典型结构	疑问中心, 目标答案描述
什么	913	[陈述表达 J][是 j111]什么 [类别, 国家, 时候, 内容等]?	疑问词“什么”和疑问中心词[类别, 国家, 时候, 内容等]充当句子的 JK。表达对其所替代的语义块的疑问。 目标答案: 与疑问中心词符合 <i>概念相似度</i> 要求。
谁	214	[偏正语][的 h\$141, h\$ug][人 p 类概念][是 j111]谁?	“谁”在句中充当一个 JK, 目标答案:p, pe
多少	166	[偏正语][数量属性概念: 长, 高, 速度][有, 是 j111]多少[数量单位概念 zz]?	疑问词“多少”代替数量的偏正, 表示对数量的疑问。 目标答案: 数 j3
多, [多么]	112	[偏正语][Jkn][有, 是 j111]多[属性概念: 长, 高, 大, 久, 快等 u]?	疑问词“多”替代数量描述对疑问中心词进行偏正, 表示对数量, 程度的疑问。 目标答案: 数 j3, 或表示量的概念。j41, jzu41 且目标答案的数量单位概念需要和疑问中心的要求相对应。
哪	58	哪[量词 zz][p, pe, w, pw, jw 类概念, 或静态 g, 效应 r 概念, 或类别概念][陈述表达 J]?	疑问词“哪”和所偏正的疑问中心词组合常充当问句的一个 JK。 目标答案: “哪”所引导的疑问中心词通常表示一个概念范围, 一个类别的概念, 目标答案通常是具体概念, 专有名词等。
哪里	105	哪[些][具有类别意义的概念]?	“哪些”是个特殊的疑问词, 其要求的答案不是一个, 而是符合要求的多个答案。
哪里	105	[陈述表达 J][在 v50001]哪里?	疑问中心充当地点辅块 FK, 目标答案: 概念类型为 wj2 类概念。
为什么	15	[为什么][陈述表达 J]?	疑问词“为什么”在句中替代对于 E 块的原因, 目的等的偏正, 表示对于原因 Pr, 目的 Rt 的疑问。 目标答案: 与问句具有相对应的语义关系结构的语义部分。
怎样, 怎么, 如何	35	[怎样, 怎么, 如何][陈述表达 J]?	表示对于手段 Ms, 途径 Wy, 工具 In, 条件 Cn 等偏正 E 块的疑问。 目标答案: 与问句具有相对应的语义关系结构的语义部分。

对于疑问句的 ERE 结构应用推理演绎规则, 得到一组新的目标 ERE 结构。

应用推理演绎规则推导出新的等语义的疑问句目标 ERE 结构的过程中, 所应用的推

理演绎规则必须是符合对称性要求的。也即推导出的新的 ERE 结构是与疑问句的 ERE 结构是具有相同语义，不同形式的 ERE，而不是蕴涵的新 ERE 知识。

然后对篇章内相同概念融合处理，如图 1 是融合处理的示意图。融合的过程就是使用目标概念 A 替换篇章内部这些具有相同的语义表示的 ERE 语义元素，处理过程包括：发现以指代，语义块共享，不同位置的相同词语概念等方式的相同概念。将与此相同的概念所相关的 ERE 都使用目标概念来替换这个相同概念。

在融合的基础上再做面向目标 ERE 结构的推理演绎分析。方法如下：从待选篇章融合的关于目标概念的 ERE 结构中不断向前推演，直到推理的过程不能够产生新的 ERE 知识时停止。推理方法与 ERE 抽取过程中的推理相同。

然后在在融合并推理演绎的 ERE 结果中寻找匹配与目标 ERE 结构相同、相似的 ERE 结构，并计算待选 ERE 与目标 ERE 之间的相似度。ERE（组合）之间的相似度计算方法如下：

自动问答过程中的 ERE 相似度比较，将目标 ERE（组合）作为比较目标，通过计算比较待选 ERE（组合）与目标 ERE 组合之间的差异来得到待选 ERE（组合）相对于目标 ERE（组合）之间的相似度。进行相似度计算的语义 ERE 有两种，偏正类 ERE 和句类 ERE。单个 ERE 相互比较的方式有：偏正类 ERE 与偏正类 ERE 的比较，句类 ERE 与句类 ERE 之间的比较，句类 ERE 与偏正类 ERE 之间的比较。以下分别论述之：

偏正类 ERE 与偏正类 ERE 的比较流程如图 2。计算方法如下：

$$\text{simLean}(t, b) = \text{recur}(\text{simComp}(t\text{Core}, b\text{Core}) * \beta_{\text{core}} + \text{simComp}(t\text{Lean}, b\text{Lean}) * \beta_{\text{lean}})$$

simLean: 偏正类 ERE 的相似度。

t: 目标 ERE。

b: 待选 ERE。

recur: 递归计算嵌套偏正 ERE 的函数。

β_{core} : 该偏正类 ERE 的核心部分的权重参数。

β_{lean} : 该偏正类 ERE 的偏正部分的权重参数。

simComp: b 相对于的概念相似度。

句类 ERE 与偏正类 ERE 之间的比较。在进行句类 ERE 与句类 ERE 之间的比较的时候，当其中一方的某个语义块由句蜕、块扩等句类 ERE 嵌套充当时，就需要进行句类 ERE 与偏正类 ERE 的比较。由于从句类 ERE 转换而来的偏正 ERE 的语义和结构都不同

于普通的偏正 ERE，所以系统将这些偏正 ERE 统称为广义偏正 ERE。系统从作用者、对象、内容和 E 块组合交叉的角度定义了句类 ERE 向广义偏正类 ERE 转换的规则。这些作用者、对象、内容和 E 块之间关系的 ERE 也归类于广义偏正类 ERE 的原因是这类 ERE 都可以使用“XX 的 XX”的偏正语言形式来表达。如：她打我。

<A X B>

对应规则可以转换为：

<<E EObject B> subjectIs A>

打 我 的 她。

转换以后的比较方法与偏正 ERE 之间的比较方法相同。

句类 ERE 与句类 ERE 之间的比较。

每个 HNC 句类表达式都可以转换为一个（或一组）的句类 ERE 表达式，句类 ERE 之间的比较流程图如图 3，其过程是：

E 块的上装按照其语义作用可以分为：

基本判断逻辑修饰。

语言逻辑修饰。

时态修饰。

空间或效应的说明。

属性修饰。

E 块的下装按照其语义作用分类有：

时态修饰。

效应和空间的说明。

属性修饰。

两个 E 块的上装之间，下装之间，以及上装和下装之间的可比性由二者的语义作用类别是否相同来判断。

E 块的 Ek 的组合构成情况有：

$E_k = E$

$E_k = EQ + EH$

$E_k = EQ + E$

$E_k = E + EH$

对于组合结构的 E 块, 比较的原则是, 将 EQ 与 EQ 比较, E 与 E 比较, EH 与 EH 比较。对于 EQ+EH 的组合方式与带有 E 的 Ek 组合方式的比较, 如果 EQ+EH 是动静搭配, 则将动部分与 E 比较。

句类 ERE 中的 R 部分(对应 HNC 句类分析的 E 块)的相似度计算方法如下:

$$\text{simE}(t, b) = \left(\sum \text{simUD}(t, b) * \beta_{ud} - \sum \text{difUD}(t, b) * \beta_{ud} \right) + \left(\sum \text{simEk}(t, b) * \beta_{ek} - \sum \text{difEk}(t, b) * \beta_{ek} \right) \text{simE}$$

两个 E 块之间的相似度比较方法如下:

t: 目标 E 块。

b: 待选 E 块。

simUD: E 块上装或下装之间的相似度比较计算函数。

β_{ud} : 上装或下装的权重参数。

difUD: E 块上装或下装之间的差异计算函数。

simEk: Ek 的相似度计算函数。

β_{ek} : Ek 各个组成部分的权重参数。

dif: k 的差异计算函数。

句类 ERE 的计算方法如下:

$$\text{simSERE}(t, b) = \text{simE}(t, b) + \sum \text{simC}(t, b) * \beta_c$$

simSERE: 两个句类 ERE 的相似度。

t: 目标句类 ERE。

b: 待选句类 ERE。

simC: 广义对象语义块之间的相似度, 由偏正类 ERE 相似度 $\text{simLean}(t, b)$ 计算得到。

β_c : 该语义块的权重参数。

下一步计算篇章内各个句群对于答案的回答程度, 计算篇章与目标 ERE 结构的相似程度, 篇章整体对于目标答案的回答程度。通过综合计算篇章与目标 ERE 结构的相似程度与篇章整体对于目标答案的回答程度能够得到篇章的综合得分。

$$\text{synAtc}(t, b) = \text{syn} \left(\sum \text{simERE}(t, b) * \beta_{ere}, \text{satfAns}(t, b), \text{quaAtc} \right)$$

synAtc: 一个篇章的对于疑问句目标 ERE 结构的综合得分。

t: 疑问句目标 ERE 结构。

b: 待选篇章。

syn: 综合计算篇章得分函数。

simERE: 各个待选 ERE 相对于对应的目标 ERE 的相似度, 如句类 ERE, 偏正类语义块 ERE 的计算结果。

β_{ere} : 该 ERE 的权重参数。

satfAns: 篇章内 ERE 组合对于目标答案的回答程度。

quaAtc: 篇章内目标概念分布质量的初步评价。

最后按照篇章的综合得分降序排列返回给用户答案篇章的列表, 并在每个篇章链接下显示该篇章对于答案回答程度较高的几个句群。

对选篇章含有目标答案程度的初步评价的步骤为:

在疑问句所有的词语概念中去除 1 语言逻辑类概念, 保留其他概念作为目标概念。

查找目标概念在待选篇章的出现概率和在语句中的分布情况。

评价每个含有目标概念的句群。

此处含有目标概念的句群的定义为: 前后紧邻的含有目标概念的语句集合。

$$\text{quaSC} = \text{tgtSC} / \text{tTgt} * \beta_s + \beta_f$$

quaSC: 句群含有目标概念分布情况的质量。

tgtSC: 含有目标概念的个数。

tTgt: 目标概念总个数。

β_s : 由该句群语句数目得到的参数。

β_f : 由该句群出现的目标概念的总次数得到的参数。

分析位置分散的含有目标概念的句群通过指代方式融合的可能性。

$$\text{proSC} = (\text{quaSC1} + \text{quaSC2}) * p(\text{distS})$$

proSC: 两个含有目标概念的句群通过指代方式互相引用具有可能性,

proSC 表示这种可能性对于篇章含有目标答案的质量的增益。

p: 通过统计得到的相互间隔的句群通过指代相互引用的可能性计算函数。

distS: 两个含有目标概念的句群之间所间隔的语句的数目。

综合评价篇章的目标概念分布。

$$\text{quaAtc} = [\Sigma(\text{quaSC}) + \Sigma(\text{proSC})] * \text{dTgt}(\text{tgtS}, \text{tS})$$

quaAtc: 篇章含有目标答案的质量。

dTgt: 篇章内部的目标概念密度计算函数。

tgtS: 含有目标概念的语句数。

tS: 篇章总共的语句数。

按照篇章质量排序。

框架知识结构定义一类事物的属性列表，将该事物称为描述中心，该事物的属性表示为 Slot。Slot 的类型方式有：单值结构、组合结构、列举结构、嵌套框架结构和复合结构。

单值结构就是一个 Slot 简单地对应着一个属性的值；组合结构指 Slot 对应多个值的组合；列举结构指 Slot 对应着一个相同的值结构的队列；嵌套框架结构指一个 Slot 的值由另外一个知识框架来表示；复合结构指 Slot 的值由以上四种结构混和组合而成。

知识框架定义了每个 Slot 与描述中心之间，不同 Slot 之间的 ERE 语义关系。知识框架定义举例如下表：

知识框架名称	人物				
Slot 定义列表	Slot 名称	Slot 的结构定义:	Slot 抽取识别规则	ERE 知识源向 SLOT 映射的规则	与描述中心, 其他 Slot 的关系的 ERE 表示。
	姓名	组合结构 姓: __名__	符合姓名识别特征		
	民族	单值结构 民族__	Pj52 类概念		
	教育经历	复合结构 1.起始时间__结束时间__学校__学位__ 2.起始时间__结束时间__学校__学	表示描述中心(人物)教育经历的语句。		

		位__ (学位是另一个框架 知识结构)			
--	--	---------------------------	--	--	--

基于框架知识库的自动问答处理过程为：

疑问句分析。

疑问句分析包括如下步骤：

对疑问句的 HNC 句类分析。

对疑问句的疑问中心的分析。

对疑问句的 ERE 结构的抽取。

疑问句的分析结构将得到一个包含疑问中心要求的目标 ERE 结构。

对于疑问句的 ERE 结构应用推理演绎规则，得到一组新的等语义的目标 ERE 结构。

通过对目标 ERE 结构的概念分析来判断可能含有目标答案的框架知识类型。

要求目标概念必要与知识框架的描述中心具有类属关系。

通过目标 ERE 与知识框架内部 Slot 之间的 ERE 关系类别匹配来判断可能含有目标答案的框架知识的 SLOT。

通过将知识框架的 SLOT 的 ERE 关系做面向目标 ERE 结构的的演绎推理来寻求目标 ERE 结构。

通过目标 ERE 结构的概念与框架知识库的知识条目的概念匹配得到可能含有目标答案的知识条目。

从指定知识条目的指定 SLOT 得到答案值。

生成答案语句，返回用户。

本发明采用了一系列技术来解决基于自然语言理解技术的知识库建立和自动问答中遇到的问题：

将自然语言表达为形式化知识。本发明采用了一种 ERE 三元表达式的方式来表达自然语言的语义知识。单个的 ERE 表达式和 ERE 组合具有强大的语义表达能力，且能够方便地从 HNC 的句类分析结果中抽取映射而得到语句的 ERE 语义表达式。ERE 表达式不依赖于语言的外在表达形式，而使用统一简洁的方式表达多样化的语言形式所蕴涵内在语义。ERE 以及 ERE 组合是一种统一的语义表达结构，有利于设计高效的针对 ERE

结构的知识处理系统。

语义知识之间的相互蕴涵推导。人脑理解语言的时候，能够很习惯地认识到一个语义当然地蕴涵着另外一个语义，比如语句：他那天下午坐火车到达了上海。其蕴涵的语义有：他那天下午的空间位置是在上海。基于 ERE 统一的知识表达模式，系统可以定义一系列各种语义情况下的蕴涵推导的产生式规则，从已有的语义 ERE 组合中推导得到新的语义 ERE。从而加深计算机对于自然语言的语义的理解。

篇章内 ERE 知识的融合。一个篇章对于某事物的描述的内容往往分布于篇章的各个部分，比如在开头段落说明该事物的时间，间隔数段落后交待了它的其他特性。人们可以综合地理解篇章的各个部分对于该事物所表达的语义。本发明提出了一种篇章内语义融合的技术，通过发现篇章内的相同概念，指代引用，语义块共享等语言现象所蕴涵的语义元素之间的等语义性质，将各个在篇章内部位置分散的语义 ERE 按照等语义元素进行融合。

自动回答用户以自然语言提出的问题。信息查询用户希望使用更自然，更准确的方式定义自己的查询要求，希望能够得到在语义，知识层面上准确地符合查询意图的答案，而不仅是匹配到关键词的答案列表。本发明通过对用户问句进行疑问句分析，目标 ERE 提取，对待选篇章的初步评价，在待选篇章中面向目标 ERE 结构的推导演绎和识别匹配，在框架知识库中对于目标 ERE 的求解分析，最终返回给用户符合语义，符合知识的答案。

待选篇章含有目标答案程度的初步评价。由于对一个篇章做面向目标 ERE 结构的推导演绎和识别匹配的过程计算量比较大的，所以系统通过对待选篇章的一次初步评价处理来得到一个按照含有目标答案可能性排序的篇章序列，以针对最有可能含有目标答案的篇章做深入的语义分析。本发明通过统计目标概念在待选篇章的出现概率，分析含有目标概念的句群情况，分散的目标概念通过指代、相同概念、语义块共享等方式融合的可能性，来初步评价各个待选篇章含有目标答案可能性。

ERE 结构之间的识别匹配。系统使用 ERE 表达式来表达知识，所以自动问答过程中对于目标知识寻找的过程就是 ERE 结构之间的识别匹配的过程，本发明对各种不同的 ERE 采用符合语义要求的不同的匹配计算方式，对于不同的语言结构部分采用不同的计算方式，对于各种嵌套组合的 ERE 结构提出了计算处理的方法，从而计算待选 ERE 结构与目标 ERE 结构之间的语义相似度。

句类 ERE 向偏正类 ERE 的转换和比较。在 ERE 的匹配过程中，按照语义结构的类

别可以分为句类 ERE 和偏正类 ERE。所以当需要将句类 ERE 与偏正类 ERE 进行语义相似度比较的时候，需要将二者转换为相同的结构形式以便于比较，本发明提出了一种将句类 ERE 转换为偏正类 ERE 的方法，通过定义各种句类 ERE 向偏正类 ERE 转换的可否性，和转换的方式，将句类 ERE 转换为一个或者一组 ERE 结构。所以偏正类 ERE 可以通过部分地或整体地与句类 ERE 进行匹配计算得到偏正类 ERE 与句类 ERE 之间的语义相似度。

基于框架知识库中寻求目标答案的问题。框架知识库是一种更加结构化的知识表达模型，相比 ERE 库松散的知识表达结构其从各个角度描述某类事物的知识。系统通过定义框架知识库的描述中心和知识框架中各个槽(Slot)之间的 ERE 关系来表述知识框架的语义结构。在基于框架知识库的自动问答的过程中，系统通过匹配知识目标 ERE 结构与框架知识库内部的 ERE 关系来寻求目标答案。

本系统的自然语言处理方法如下：

1.1.1 概念层次网络语言层面知识库的表述步骤如下：

1.1.1.1 将语句按语义划分成 7 个基本句类，作用句，过程句，转移句，效应句，关系句，状态句，判断句；根据语义块与句类的依赖性强弱，将语义块分为主语义块和辅语义块，其中辅语义块包括：条件，手段，工具，途径，参照，因，果；从其共性特征将主语义块分成：特征语义块，作用者，对象，内容；建立语义块的一般物理表示式： $SK = \text{个性} + \text{共性} = \text{句类信息} + \text{语义块类型信息}$ ；当句子的特征语义块包含两个基本句类的特征时，构成混合句；当句子中用两个或多个特征语义块表述两个或多个基本句类的特征时，构成复合句类；以符号的形式将上述信息表达出来，形成知识库。

1.1.1.2 对于知识库中的词汇，如果它的概念类别含 V，根据它本身的语义知识确定它对应的节点作用句 Φ_0 ，过程句 Φ_1 ，转移句 Φ_2 ，效应句 Φ_3 ，关系句 Φ_4 和状态句 Φ_5 以及一般判断句子类 Φ_8 与判断句的其他子类 j_{11} 中为代表的主要内容，根据对应情况确定词汇属于 7 个基本句类的哪一种；在对应 57 个子类中的句类代码；如果该词语的语义只是主要内容包括两个对应的前述节点，则按混合句类处理，混和句类

的代码约定为，以构成混合句类两个基本句类中第一个语义块开始，K 表示非 E 语义块的总个数，n 表示从第二个基本句类 E2 中取出的语义块的起始序号，当 $n=m+1$ 时，n 可以不写，对引起复合句类的词汇，以形成复合句类的两个基本句类的代码 E1 与 E2 中间加*号，填写句类信息：在分析时，可以根据 E1 和 E2 的指示，从概念层面句类表示式知识库中取出两个句类的格式知识。

1.1.1.3 当句类代码有效时，根据 (1.2.1.2) 中的句类代码，具体确定句类属于两从句，三从句及四从句：具体确定办法如下：语句的一般数学表达式可写成：

第一号广义对象语义块 JK 后接句子的特征语义块，再后接第二个广义对象语义块，再后接第三个广义对象语义块，其余广义对象语义块顺序列出：

表示式中并未限定广义对象语义块 JK 的个数，但对于基本句类，实际的自然语言只需要考虑 JK 个数为 1, 2, 3 的情况，它们分别相应于两主从句，三主从句，和四主从句：

对于四主从句，JK2 一定以对象 B 为主体，JK3 一定以内容 C 为主体，对于三主从句，B 或 C 都可以充当 JK 的主体：对于两主从句，可以没有 E，但这时 JK2 必须以 C 为主体，根据该词语组句时主语义块经常采用的叙述顺序确定其具体的格式代码；以代码的形式给出该词组成句子时经常采用的格式：当有多个格式的时候，用[1][2].....的形式标号，以便在下面各项中对应表示不同格式的不同情况：如组成句子时经常采用标准格式和规范格式，该项可以不填。

1.1.1.4 当句类代码有效时，该词语按照(b)中句类代码组句时，如果该词语与广义对象语义块之间有预期关系，即该词语要求特定的概念来充当它的某个广义对象语义块，则将这个特定的，优先与该词语搭配的概念按 $F = \Sigma(\text{字母串})(\text{数字串})$ 所叙的方法给出：这种预期，包括对广义对象语义块结构中的某个成分的预期：这时，首先对语义块的构成信息进行描述，然后给出对应成分的优先概念，语义块的构成知识及各构成部分的优先概念单位中以@S 代表，JK 语义块的构成知识，

用=和+填写在这一项中：各部分的优先概念知识，用：表示，也填写在这一项中：如果 v 概念构成的句子经常要求一个句子成为其中的语义块，如果词汇有这个知识，就在知识库中此项以 JK=J 和 JK: =J 分别表示某一语义块 JK 必须扩展成为句子或可能扩展成为句子：约定见 (vi)。

一个语义块或语义块的构成部分可以从内涵上分成对象 B 和内容 C 两个部分，也可以从形式上分成前 Q，后 H 的两部分：对于这种构成属于约定，无需在显式第写出表达式，只在这个语义块或构成部分后加上上述四各字母 B,C,Q,H 给出其优先概念，就表示这种构成存在，同时说明这部分的优先概念。

1.1.1.5 如果(iv)中描述的广义对象语义块组成结构中的一部分与其他部分不是紧接在一起，而是分别出现在语句的两个分开的位置上，对这种情况在语义块构成中进行表示，分别以[]和()表示语义块可能分离和语义块一定分离出去的部分

1.1.1.6 当句类代码有效时，同时按照给出的句类代码组句时，要求一个句子充当它的某一广义对象语义块，对这一情况进行指示，即给出该词汇引发的某些语义块扩展为语句的知识，

1.1.2 确定句类分析的具体步骤如下：

1.1.2.1 对输入的句子，进行词库匹配，切分出句子中遇到的词，从知识库中取得这些词汇的语义知识：

1.1.2.2 根据概念类别信息的指示，以语义块区分指示符 10 类概念和动词 v 概念为依据，形成语义块雏形，形成 E 假设：

1.1.2.3 如果未能形成 E 假设，转向(1.2.2.9)；否则继续；

1.1.2.4 对全部 E 假设进行筛选合排队，主要利用信息为：句类代码，格式代码。

1.1.2.5 按照入选 E 假设的排序，依次进行句类检验；主要利用信息为：语义块核心的概念优选性知识：如果全部检验失败，转向(1.2.2.10)；否则继续；

1.1.2.6 进行语义块构成检验；主要利用的信息为：语义块构成知识和构成语

义块各部分优先概念的知识，如果全部检验失败，转向(1.2.2.10)；否则继续；

1.1.2.7 在必要时进行语义块分离检验，主要利用的信息为：词汇引发的句类转换知识。

1.1.2.8 进行无 E 语义块句类检验：如果失败继续，否则转向(1.2.2.10)：

1.1.2.9 重新做 E 假设，成功转向(1.2.2.4)。

1.1.2.10 记录最终分析结果。

本系统的自然语言处理方法如下：

1.1.3 概念层次网络语言层面知识库的表述步骤如下：

1.1.3.1 将语句按语义划分成 7 个基本句类，作用句，过程句，转移句，效应句，关系句，状态句，判断句；根据语义块与句类的依赖性强弱，将语义块分为主语义块和辅语义块，其中辅语义块包括：条件，手段，工具，途径，参照，因，果；从其共性特征将主语义块分成：特征语义块，作用者，对象，内容；建立语义块的一般物理表示式： $SK = \text{个性} + \text{共性} = \text{句类信息} + \text{语义块类型信息}$ ；当句子的特征语义块包含两个基本句类的特征时，构成混合句；当句子中用两个或多个特征语义块表述两个或多个基本句类的特征时，构成复合句类；以符号的形式将上述信息表达出来，形成知识库。

1.1.3.2 对于知识库中的词汇，如果它的概念类别含 V，根据它本身的语义知识确定它对应的节点作用句 Φ_0 ，过程句 Φ_1 ，转移句 Φ_2 ，效应句 Φ_3 ，关系句 Φ_4 和状态句 Φ_5 以及一般判断句子类 Φ_8 与判断句的其他子类 j_{11} 中为代表的主要内容，根据对应情况确定词汇属于 7 个基本句类的哪一种；在对应 57 个子类中的句类代码；如果该词语的语义只是主要内容包括两个对应的前述节点，则按混合句类处理，混和句类的代码约定为，以构成混合句类两个基本句类中第一个语义块开始，K 表示非 E 语义块的总个数，n 表示从第二个基本句类 E2 中取出的语义块的起始序号，当 $n=m+1$ 时，n 可以不写，对引起复合句类的词汇，以形成复合句类的两个基本句类的代码 E1 与 E2 中间加*号，填写句类信息；在分析时，可以根据 E1 和 E2 的指示，从概念层面句类表示

式知识库中取出两个句类的格式知识。

1.1.3.3 当句类代码有效时，根据（1.2.1.2）中的句类代码，具体确定句类属于两从句，三从句及四从句：具体确定办法如下：语句的一般数学表达式可写成：

第一号广义对象语义块 JK 后接句子的特征语义块，再后接第二个广义对象语义块，再后接第三个广义对象语义块，其余广义对象语义块顺序列出：

表示式中并未限定广义对象语义块 JK 的个数，但对于基本句类，实际的自然语言只需要考虑 JK 个数为 1, 2, 3 的情况，它们分别相应于两主从句，三主从句，和四主从句：

对于四主从句，JK2 一定以对象 B 为主体，JK3 一定以内容 C 为主体，

对于三主从句，B 或 C 都可以充当 JK 的主体：对于两主从句，可以没有 E，但这时 JK2 必须以 C 为主体，根据该词语组句时主语义块经常采用的叙述顺序确定其具体的格式代码；以代码的形式给出该词语组成句子时经常采用的格式：当有多个格式的时候，用[1][2].....的形式标号，以便在下面各项中对应表示不同格式的不同情况：如组成句子时经常采用标准格式和规范格式，该项可以不填。

1.1.3.4 当句类代码有效时，该词语按照(b)中句类代码组句时，如果该词语与广义对象语义块之间有预期关系，即该词语要求特定的概念来充当它的某个广义对象语义块，则将这个特定的，优先与该词语搭配的概念按 $F = \Sigma(\text{字母串})(\text{数字串})$ 所叙的方法给出：这种预期，包括对广义对象语义块结构中的某个成分的预期：这时，首先对语义块的构成信息进行描述，然后给出对应成分的优先概念，语义块的构成知识及各构成部分的优先概念单位中以@S 代表，JK 语义块的构成知识，用 = 和 + 填写在这一项中：各部分的优先概念知识，用：表示，也填写在这一项中：如果 v 概念构成的句子经常要求一个句子成为其中的语义块，如果词汇有这个知识，就在知识库中此项以 JK = J 和 JK：= J 分别表示某一语义块 JK 必须扩展成为句子或可能扩展成为句子：约定见 (vi)。

一个语义块或语义块的构成部分可以从内涵上分成对象 B 和内容 C 两个部分，也可以从形式上分成前 Q，后 H 的两部分：对于这种构成属于约定，无需在显式第写出表达式，只在这个语义块或构成部分后加上上述四各字母 B,C,Q,H 给出其优先概念，就表示这种构成存在，同时说明这部分的优先概念。

1.1.3.5 如果(iv)中描述的广义对象语义块组成结构中的一部分与其他部分不是紧接在一起，而是分别出现在语句的两个分开的位置上，对这种情况在语义块构成中进行表示，分别以[]和()表示语义块可能分离和语义块一定分离出去的部分

1.1.3.6 当句类代码有效时，同时按照给出的句类代码组句时，要求一个句子充当它的某一广义对象语义块，对这一情况进行指示，即给出该词汇引发的某些语义块扩展为语句的知识，

1.1.4 确定句类分析的具体步骤如下：

1.1.4.1 对输入的句子，进行词库匹配，切分出句子中遇到的词，从知识库中取得这些词汇的语义知识：

1.1.4.2 根据概念类别信息的指示，以语义块区分指示符 10 类概念和动词 v 概念为依据，形成语义块雏形，形成 E 假设：

1.1.4.3 如果未能形成 E 假设，转向(1.2.2.9)；否则继续；

1.1.4.4 对全部 E 假设进行筛选合排队，主要利用信息为：句类代码，格式代码。

1.1.4.5 按照入选 E 假设的排序，依次进行句类检验；主要利用信息为：语义块核心的概念优选性知识：如果全部检验失败，转向(1.2.2.10)；否则继续；

1.1.4.6 进行语义块构成检验；主要利用的信息为：语义块构成知识和构成语义块各部分优先概念的知识，如果全部检验失败，转向(1.2.2.10)；否则继续；

1.1.4.7 在必要时进行语义块分离检验，主要利用的信息为：词汇引发的句类转换知识。

1.1.4.8 进行无 E 语义块句类检验：如果失败继续，否则转向(1.2.2.10)：

1.1.4.9 重新做 E 假设，成功转向(1.2.2.4)。

1.1.4.10 记录最终分析结果。

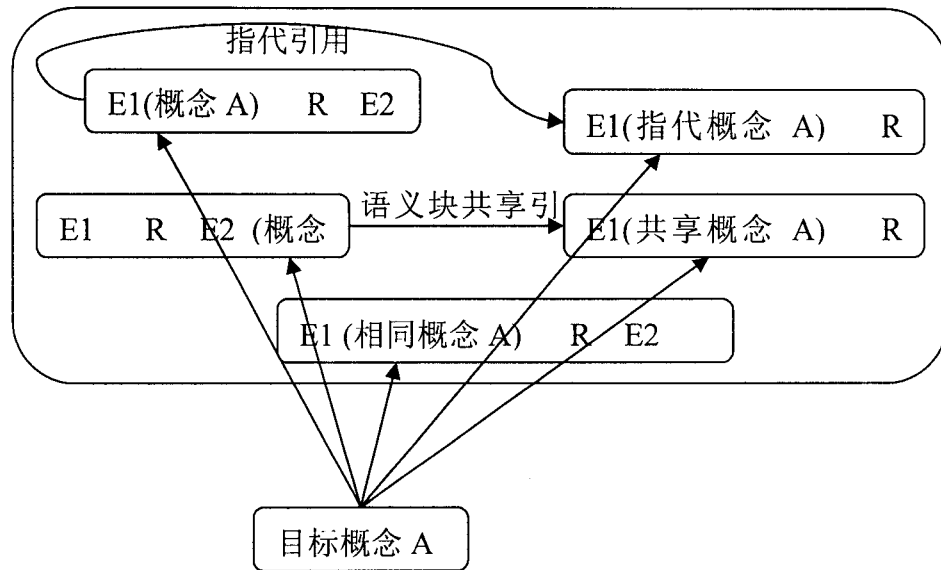


图 1

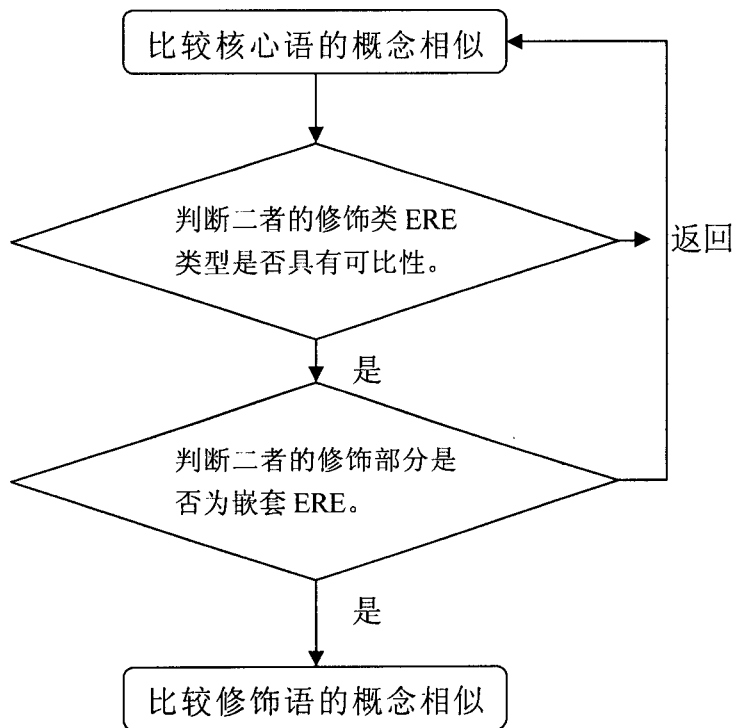


图 2

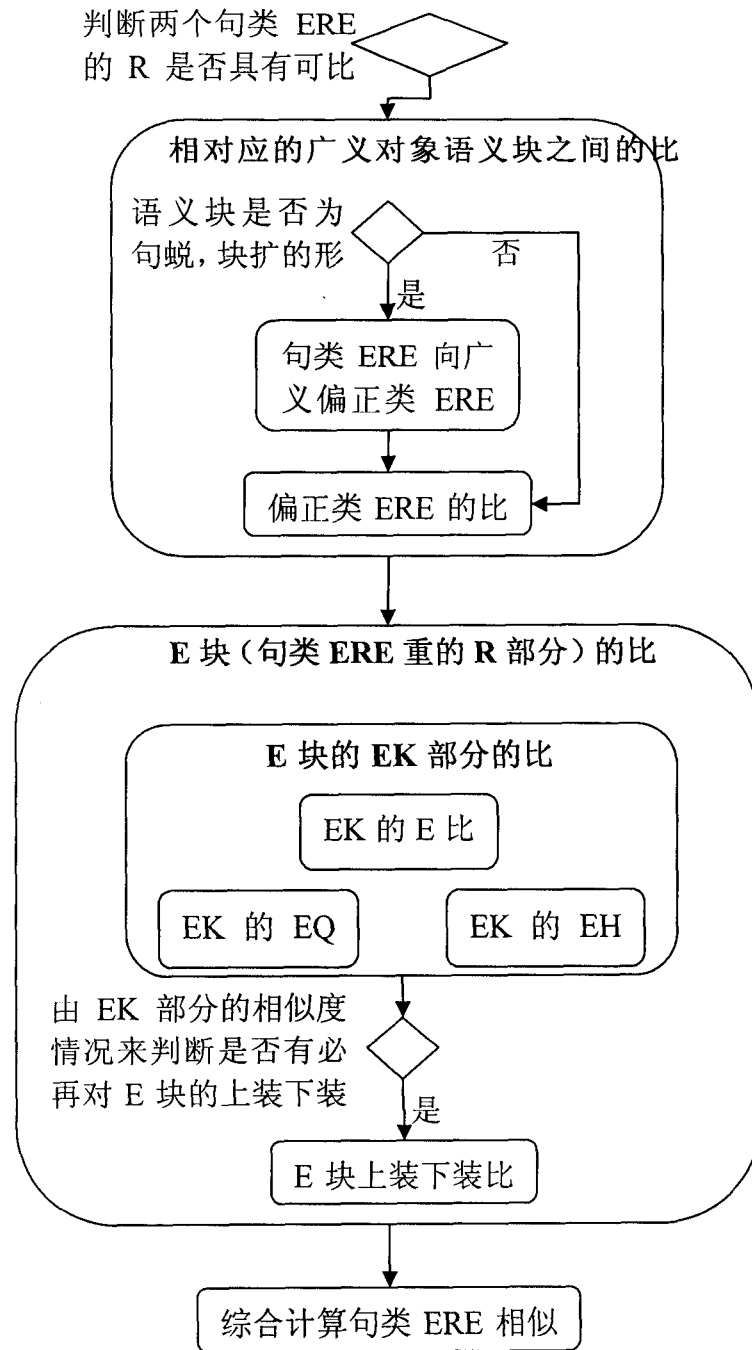


图 3

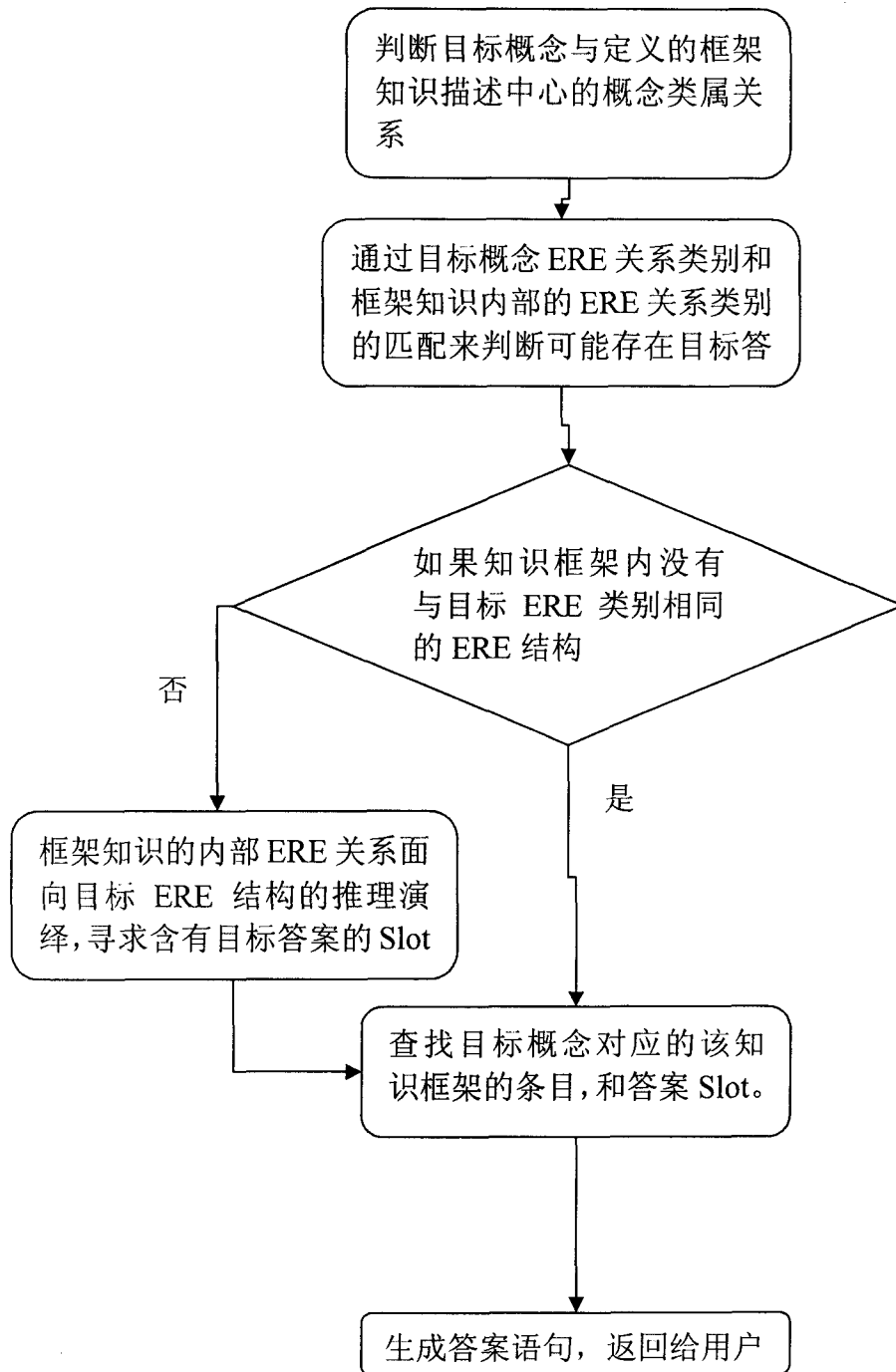


图 4