策略确定方法、装置、设备和存储介质

相关公开的交叉引用

本公开要求于 2025 年 03 月 6 日提交中国专利局的申请号为 2025102622587、名称为“策略确定方法、装置、设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及计算机技术领域，具体而言，涉及一种策略确定方法、装置、设备和存储介质。

背景技术

相关技术中，为实现大模型能够针对输入数据输出有效的策略，都需要对大模型进行训练，即，先构造大量的训练数据，然后再使用监督微调（Supervised Fine-Tuning，SFT）技术或者人类反馈强化学习（Reinforcement Learning from Human Feedback，RLHF）技术，来对模型进行训练。

但由于大模型的构建需要依赖预先构建的大量训练数据，因此大模型的构建耗时特别长，且模型训练也需要比较多的硬件设备资源。而且在将大模型应用到确定策略的场景下，针对每个策略，都需要人工构造比较多的数据量才能够让模型学会该策略；此外，大模型具备相同的输入条件的情况下，也必须要输出相同的策略，但如果训练数据中存在相同的输入条件但执行策略不同的情况下，会导致训练出来的大模型的处理过程出现混乱，进而影响大模型的策略输出效果。

公开内容

有鉴于此，为至少解决相关技术中因大模型需依赖大量训练数据才能正确输出策略，而导致训练耗时长和硬件设备资源占用多的技术问题，本公开的目的在于提供一种策略确定方法、装置、设备和存储介质。

为了实现上述目的，本公开实施例采用的技术方案如下：

本公开实施例的第一方面，提供了一种策略确定方法，包括：

通过大模型从当前输入的用户问题中抽取出目标条件信息；所述目标条件信息包括目标条件槽名和目标条件槽值；

从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径；所述决策路径池中的每条决策路径都包含条件信息和执行策略；

通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息。

在可选的实施方式中，所述从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径的步骤，包括：

从所述决策路径池中，获取条件槽名与所述目标条件槽名相同的候选决策路径；

将所述候选决策路径中与所述目标条件槽名对应的条件槽值和所述目标条件槽值进行匹配，以得到目标决策路径。

在可选的实施方式中，所述将所述候选决策路径中与所述目标条件槽名对应的条件槽值和所述目标条件槽值进行匹配，以得到目标决策路径的步骤，包括：

将所述候选决策路径中与所述目标条件槽名对应的条件槽值转换为第一槽值向量，并将所述目标条件槽值转换成第二槽值向量；

根据与所述第二槽位向量的相似度大于第一设定相似度阈值的第一槽位向量所在的候选决策路径，确定目标决策路径。

在可选的实施方式中，所述决策路径池的构建过程包括：

通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径；

通过所述大模型基于所述设定策略目标、所述设定用户问题和当前待测决策路径中的

条件信息，生成新待测决策路径，并将当前待测决策路径更新为所述新待测决策路径，重复执行本步骤，直至所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过所述大模型对所有待测决策路径进行评分，并对分数高于设定分数阈值的待测决策路径进行存储，以得到决策路径池。

在可选的实施方式中，所述方法还包括：

在收到策略更新请求的情况下，通过所述大模型基于所述策略更新请求中携带的新策略目标和参考问题，获得与所述新策略目标和所述参考问题相应的种子决策路径；

通过所述大模型基于所述新策略目标、所述参考问题和种子决策路径中的条件信息，生成新决策路径，并将所述种子决策路径更新为所述新决策路径，重复执行本步骤，直至所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过所述大模型对所有种子决策路径进行评分，并将分数高于设定分数阈值的种子决策路径存储于所述决策路径池。

在可选的实施方式中，所述通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息的步骤，包括：

根据所述目标决策路径中的执行策略，生成相应的引导文本；

通过所述大模型基于所述用户问题和所述引导文本，输出用于答复所述用户问题的策略信息。

在可选的实施方式中，所述方法还包括：

通过预先构建的训练样本集对所述大模型进行训练，直至所述大模型收敛；其中，所述训练样本集包括多组训练样本，每组训练样本均包括用户问题样本及其对应的标签；所述标签用于表征所述用户问题样本对应的真实条件信息。

本公开实施例的第二方面，提供了一种策略确定装置，包括：

提取模块，被配置为：通过大模型从当前输入的用户问题中抽取出目标条件信息；所述目标条件信息包括目标条件槽名和目标条件槽值；

筛选模块，被配置为：从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径；所述决策路径池中的每条决策路径都包含条件信息和执行策略；

输出模块，被配置为：通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息。

本公开实施例的第三方面，提供了一种电子设备，包括处理器和存储器，所述存储器存储有能够被所述处理器执行的机器可执行指令，所述处理器可执行所述机器可执行指令以实现上述第一方面任一项中提供的策略确定方法。

本公开实施例的第四方面，提供了一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现上述第一方面任一项中提供的策略确定方法。

本公开实施例提供的策略确定方法、装置、设备和存储介质，通过利用大模型从用户问题中抽取出目标条件信息，然后再从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径，实现利用目标决策路径来干预大模型的推理过程，使得大模型能够针对用户问题输出更加准确的结果。其中，大模型实现从用户问题中抽取出目标条件信息、以及针对用户问题做相应答复，都可以采用大模型原有功能。因此，本公开实施例中，无需对大模型进行训练，也可以结合决策路径来实现大模型输出结果的准确性，具备较低的模型应用成本和较高的模型应用效率。此外，如果针对用户问题进行答复的策略存在更新需求，可以通过更新决策路径池中的决策路径，来使得大模型也具备针对新策略输出针对性的策略信息的能力，更加符合一些应用场景下的策略经常变动的需求，例如，营销场景下的营销策略。

为使本公开的上述目的、特征和优点能更明显易懂，下文特举较佳实施例，并配合所附附图，作详细说明如下。

附图说明

为了更清楚地说明本公开实施例的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本公开的某些实施例，因此不应被看作是对范围的限定，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他相关的附图。

图 1 示出了本公开实施例提供的一种电子设备的结构框图；

图 2 示出了本公开实施例提供的一种策略确定方法的流程图；

图 3 示出了本公开实施例提供的一种策略确定装置的功能模块图。

具体实施方式

下面将结合本公开实施例中附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本公开实施例的组件可以以各种不同的配置来布置和设计。

因此，以下对在附图中提供的本公开的实施例的详细描述并非旨在限制要求保护的本公开的范围，而是仅仅表示本公开的选定实施例。本公开的实施例，本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

需要说明的是，术语“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括要素的过程、方法、物品或者设备中还存在另外的相同要素。

为至少解决相关技术中因大模型需依赖大量训练数据才能正确输出策略，而导致训练耗时长和硬件设备资源占用多的技术问题，本公开提供了一种策略确定方法，通过利用大模型从用户问题中抽取出目标条件信息，然后再从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径，实现利用目标决策路径来干预大模型的推理过程，使得大模型能够针对用户问题输出更加准确的结果。其中，大模型实现从用户问题中抽取出目标条件信息、以及针对用户问题做相应答复，都可以采用大模型原有功能。因此，本公开实施例中，无需对大模型进行训练，也可以结合决策路径来实现大模型输出结果的准确性，具备较低的模型应用成本和较高的模型应用效率。此外，如果针对用户问题进行答复的策略存在更新需求，可以通过更新决策路径池中的决策路径，来使得大模型也具备针对新策略输出针对性的策略信息的能力，更加符合一些应用场景下的策略经常变动的需求，例如，营销场景下的营销策略。

本公开提供的策略确定方法可以应用于电子设备中，请参照图 1，是电子设备的结构框图。电子设备 100 包括存储器 110、处理器 120 及通信模块 130。存储器 110、处理器 120 以及通信模块 130 各元件相互之间直接或间接地电性连接，以实现数据的传输或交互。例如，这些元件相互之间可通过一条或多条通讯总线或信号线实现电性连接。

其中，存储器配置成存储程序或者数据。存储器可以是，但不限于，随机存取存储器（Random Access Memory，RAM），只读存储器（Read Only Memory，ROM），可编程只读存储器（Programmable Read-Only Memory，PROM），可擦除只读存储器（Erasable Programmable Read-Only Memory，EPROM），电可擦除只读存储器（Electric Erasable Programmable Read-Only Memory，EEPROM）等。

处理器配置成读/写存储器中存储的数据或程序，并执行相应地功能。

通信模块配置成通过网络建立电子设备与其它通信终端之间的通信连接，并配置成通

过网络收发数据。

应当理解的是，图 1 所示的结构仅为电子设备的结构示意图，电子设备还可包括比图 1 中所示更多或者更少的组件，或者具有与图 1 所示不同的配置。图 1 中所示的各组件可以采用硬件、软件或其组合实现。

以下结合图 2 对本公开实施例提供的策略确定方法进行说明，图 2 是本公开实施例提供的一种策略确定方法的流程图，策略确定方法包括：

在步骤 S100 中，通过大模型从当前输入的用户问题中抽取出目标条件信息；目标条件信息包括目标条件槽名和目标条件槽值；

在步骤 S200 中，从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径；决策路径池中的每条决策路径都包含条件信息和执行策略；

在步骤 S300 中，通过大模型基于用户问题和目标决策路径中的执行策略，输出用于答复用户问题的策略信息。

在一些应用场景下，如果用户提出问题后，需要触发一些策略性引导输出，以引导用户进行相关操作而完成既定的策略，例如，以营销场景为例，假设营销策略（也可以称为营销目标）为:“让用户填写手机号”,而用户问题为“宝马 X5 在深圳多少钱能够落地买到”,基于此，上述策略性的引导输出可以为以下其中一种：

第一种：“宝马 X5 的指导价是 72 万，落地价还需考虑当地的优惠、购置税等内容，填下手机号后我们联系你”；

第二种：“所有车型所在城市的指导价、落地价和优惠等信息我们这边都有，填下手机号后我们联系你”。

可见无论是哪一种引导输出，其都能与营销策略匹配，即引导输出（相当于策略信息）都提及到了与营销策略相关的信息，都提出了指引用户填写手机号的内容。

因此，在问答场景下，如果要结合策略目标来输出用于引导用户行为的策略信息，则可以采用本公开实施例提供的策略确定方法，以实现相应策略信息的输出，进而实现策略目标，或者使得策略目标更加容易完成。

在执行本公开实施例提供的策略确定方法的过程中，如果检测到有用户问题输入，则执行步骤 S100，以利用大模型提取出用户问题中的目标条件信息。

例如，继续沿用上述营销场景下的示例，提取出的目标信息可以包括：

目标条件信息（也可以称为 Condition）：｛“车名称”:“宝马 X5”,“预算”:“”, “城市”:“深圳”, “意图 1 级”:“询问价格”, “意图 2 级”:“落地价”｝。

上述中，“车名称”:“宝马 X5”为一组子条件信息，“预算”:“”为一组子条件信息，“城市”:“深圳” 为一组子条件信息，“意图 1 级”:“询问价格” 为一组子条件信息，“意图 2 级”:“落地价”也为一组子条件信息。每组子条件信息中，“:”之前的信息为目标条件槽名、而“:”之后的信息为目标条件槽值。

上述目标条件信息的提取可以采用大模型已有的功能实现，无需对大模型进行训练。

在一些实施例中，为提高大模型提取目标条件信息的准确性，同时在一定程度上降低训练耗时和对硬件设备资源的要求，本公开实施例提供的策略确定方法还提供了一种用于训练大模型对目标条件信息的提取能力的训练方案，即，本公开实施例提供的策略确定方法还可以包括：

在步骤 S010 中，通过预先构建的训练样本集对大模型进行训练，直至大模型收敛。

其中，训练样本集包括多组训练样本，每组训练样本均包括用户问题样本及其对应的标签；标签用于表征用户问题样本对应的真实条件信息。

本公开实施例提供的上述训练方案能够达到降低训练耗时和对硬件设备资源的要求的原因在于：采用了决策路径中的 condition 辅助大模型训练，而 condition 中可以包括一个或一个以上的子条件信息，因此可以用数量较少且不同的子条件信息进行组合，即可得到更多数量的标签。接着可以再基于标签来构建相应的用户问题样本。并且，由于本公开

实施例在大模型推理过程中，还利用决策路径进行干预而提升大模型输出的准确性。因此无需如相关技术所记载，需要采用大量的训练数据，且针对每个用户问题都需构建一一对应的标签，才能实现大模型具备准确的输出。

相应的，步骤 S100 中所采用的大模型可以为经过步骤 S010 进行训练得到的大模型，可以进一步提升抽取得到的目标条件信息的准确性，利用更加准确的目标条件信息，也可以匹配到更加准确的决策路径，进而可以使得大模型输出的策略信息更加符合策略目标，具备更强的鲁棒性和泛化能力。

通过上述任一实施例实现目标条件信息的抽取之后，可以执行步骤 S200，以从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径。其中，可以通过相似度匹配算法来实现目标决策路径的获取。

在一些实施例中，为提高匹配效率和匹配精度，进而提高目标决策路径的获取效率和精度，本公开实施例提供的策略确定方法，还提供了一种匹配方案，即，上述步骤 S200 中，从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径的步骤，包括：

在步骤 S210 中，从决策路径池中，获取条件槽名与目标条件槽名相同的候选决策路径；

在步骤 S220 中，将候选决策路径中与目标条件槽名对应的条件槽值和目标条件槽值进行匹配，以得到目标决策路径。

可以理解的，在根据目标条件信息匹配出目标决策路径的过程中，先通过步骤 S210 来对目标条件槽名进行匹配，在目标条件槽名匹配成功后，再执行步骤 S220，以进一步做目标条件槽值的匹配。可见，先通过匹配条件槽名再匹配条件槽值，可以先排除掉不具备目标条件槽名的决策路径，以得到数量较少的候选决策路径，然后再进一步匹配条件槽值，实现从数量较少的候选决策路径中做匹配处理，相比于同时进行条件槽值和条件槽名的匹配方式，能够减少匹配处理中所需比较的数据量，而且也有利于保证匹配的准确性。

以下结合示例说明下步骤 S210 和步骤 S220 的技术原理：

假设提取出的目标条件信息包括：

Condition：｛“车名称”：“宝马 X5”，“预算”：“”，“城市”：“深圳”， “意图 1 级”：“询问价格”， “意图 2 级”：“落地价”｝。

为便于描述，将所有目标条件槽名“车名称”、“预算”、“城市”、“意图 1 级”和“意图 2 级”分别称为槽名 A1、A2、A3、A4 和 A5。而所有条件槽值分别称为槽值 B1、B2、B3、B4 和 B5。其中，A1 和 B1 对应， A2 和 B2 对应，例如，条件槽名“车名称”与条件槽值“宝马 X5”对应，其他以此类推。

假设决策路径池中有以下 3 条决策路径：

第一条决策路径：

Condition：｛“车名称”：“宝马 X5”，“预算”：“”， “城市”：“广州”， “意图 1 级”：“询问价格”， “意图 2 级”：“落地价”｝；

Action：｛“回复内容”：“指导价”，“引导内容”：“填写手机号”｝。

第二条决策路径：

Condition：｛“车名称”：“宝马 X5”，“预算”：“”， “城市”：“深圳”， “意图 1 级”：“询问价格”， “意图 2 级”：“落地价”｝；

Action：｛“回复内容”：“指导价”，“引导内容”：“填写手机号”｝。

第三条决策路径：

Condition：｛“车名称”：“宝马 X5”，“预算”：“100 万”， “城市”：“深圳”， “意图 1 级”：“保险购置”， “意图 2 级”：“落地价” ｝；

Action：｛“回复内容”：“指导价”，“引导内容”：“填写手机号”｝。

基于此，在执行步骤 S210 的过程中，会将所有目标条件槽名 A1~A5 和每条决策路径中的所有条件槽名进行一一比对，如果比对结果表示某个决策路径中未存在 A1~A5 中至少一个条件槽名的情况下，则可忽略该决策路径。反之，如果比对结果表示某个或某些决策路

径中存在 A1~A5 所有条件槽名的情况下，则将这些决策路径作为候选决策路径。候选决策路径可能有 1 条或 1 条以上。以上述示例可知，执行步骤 S210 之后，可以得到的候选决策路径有第二条决策路径和第三条决策路径。

得到候选决策路径之后，可以继续执行步骤 S220，以进一步进行条件槽值匹配，即，对于筛选得到的第二条决策路径和第三条决策路径，将与槽名 A1 相同的条件槽名对应的条件槽值和槽值 B1 进行匹配、将与槽名 A2 相同的条件槽名对应的条件槽值和槽值 B2 进行匹配、将与槽名 A3 相同的条件槽名对应的条件槽值和槽值 B3 进行匹配、将与槽名 A4 相同的条件槽名对应的条件槽值和槽值 B4 进行匹配、将与槽名 A5 相同的条件槽名对应的条件槽值和槽值 B5 进行匹配，如果上述任一项均匹配成功，则认为对应的候选决策路径可以作为目标决策路径。反之，只要存在一项匹配不成功，则排除掉对应的候选决策路径。

基于上述示例可知，由于第三条决策路径中“意图 1 级”对应“保险购置”，明显与目标条件槽名“意图 1 级”对应的条件槽值 “询问价格”不匹配，因此将第三决策路径排除。而第二条决策路径完全匹配，因此可以将第二条决策路径作为目标决策路径。

经过步骤 S220 处理得到的候选决策路径有多条的情况下，可以随机选取其中一条候选决策路径作为目标决策路径。

在一些实施例中，为减少匹配过程中的计算难度，本公开实施例针对上述步骤 S220 还提供了一种实现手段，即，上述步骤 S220 中，将候选决策路径中与目标条件槽名对应的条件槽值和目标条件槽值进行匹配，以得到目标决策路径的步骤，包括：

在步骤 S221 中，将候选决策路径中与目标条件槽名对应的条件槽值转换为第一槽值向量，并将目标条件槽值转换成第二槽值向量；

在步骤 S222 中，根据与第二槽位向量的相似度大于第一设定相似度阈值的第一槽位向量所在的候选决策路径，确定目标决策路径。

可以理解的，在进行槽值匹配的过程中，可以将槽值转换成向量后进行向量相似度计算，进而实现匹配。

其中，第一设定相似度阈值可以根据经验或实验或实际需求设定，在此不进行限定。

而在步骤 S210 中进行槽名匹配的过程中，也可以将槽名转换成向量后进行向量相似度计算，进而实现匹配。其与槽值匹配的区别在于：采用的相似度阈值不一样，因为对槽名的要求是相同，因此可以理解为其采用的相似度阈值为 100%。

得到目标决策路径之后，可以执行步骤 S300，以通过大模型基于用户问题和目标决策路径中的执行策略，得到用于答复用户问题的策略信息。

在一些实施例中，为降低目标决策路径干预大模型的实现难度，并便于大模型理解目标决策路径的执行策略，以更好地生成满足策略目标的策略信息，本公开实施例提供的策略确定方法还提供了步骤 S300 的一种实现方式，即，上述步骤 S300 中，通过大模型基于用户问题和目标决策路径中的执行策略，输出用于答复用户问题的策略信息的步骤，包括：

在步骤 S310 中，根据目标决策路径中的执行策略，生成相应的引导文本；

在步骤 S320 中，通过大模型基于用户问题和引导文本，输出用于答复用户问题的策略信息。

可以理解的，在执行步骤 S300 的过程中，可以先通过步骤 S310 来将目标决策路径中的执行策略，生成相应的引导文本。

例如，继续沿用上述第二条决策路径，可见其执行策略为：

Action：｛“回复内容”：“指导价”，“引导内容”：“填写手机号”｝。

基于此生成的引导文本可以为：“请告知用户指导价并引导用户填写手机号码”。

由此，得到引导文本之后，可以执行步骤 S320，来通过大模型基于用户问题和引导文本，来生成相应的策略信息，例如，该策略信息可以为：

“宝马 X5 的指导价是 72 万，落地价还需考虑当地的优惠、购置税等内容，填下手机号后我们联系你”。

在一些实施例中，决策路径池中的决策路径可以由人工构建。

在一些实施例中，为减少人工构建决策路径池的工作量和负担，本公开实施例提供的策略确定方法还提供了一种利用大模型实现决策路径池的构建方案，即，决策路径池的构建过程可以包括：

在步骤 S020 中，通过大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与设定策略目标和设定用户问题相应的当前待测决策路径；

在步骤 S030 中，通过大模型基于设定策略目标、设定用户问题和当前待测决策路径中的条件信息，生成新待测决策路径，并将当前待测决策路径更新为所述新待测决策路径，重复执行本步骤，直至大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在步骤 S040 中，在大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过大模型对所有待测决策路径进行评分，并对分数高于设定分数阈值的待测决策路径进行存储，以得到决策路径池。

对于步骤 S020，设定策略目标和设定用户问题可以有多个。在执行步骤 S020 的过程中，可以利用提示词工程来引导大模型基于设定策略目标和设定用户问题来生成符合决策路径样例的当前待测决策路径。其中的决策路径样例可以参考上文第一条决策路径~第三条决策路径中的其中一个。

在一些实施例中，可以直接使用提示词来引导模型生成一个新的决策路径，示例性地：

用户问题为：宝马 x5 在深圳多少钱能够落地买到？

营销目标为：让用户填写手机号。

那么可以生成如下指令输入给大模型（下述指令仅仅只是示例）：

Goals

你现在是在一个营销场景中，你的最终目标是针对用户达成营销目标。用户会提问问题，你要决定如何对该问题进行回答才能够达成营销目标。

Workflow

1. 对于用户问题，你要识别出其中的各个槽位，输出格式按照 OutputFormat 中要求进行输出。

2. 你可以选择不回答用户问题、仅回答用户问题中的少部分、完全回答用户问题。你认为哪种更有可能达成营销目标就选择哪个。

3. 为了达成营销目标，你需要对用户进行引导。

OutputFormat

Condition：{“车名称”：“”,“预算”：“”,“城市”：“”,“意图 1 级”：“”,“意图 2 级”：“”}

Action：{“回复内容”：“”,“引导内容”：“”}

Input

用户问题为：宝马 x5 在深圳多少钱能够落地买到？

营销目标为：让用户填写手机号。

Output

其中，设定策略目标、设定用户问题和决策路径样例可以由人工输入。

在执行步骤 S020 的过程中，大模型可以基于设定用户问题、设定策略目标和决策路径样例，来从回复用户问题深浅程度的方面循环调用大模型，以使大模型可以生成完全不回复用户问题的至少一条待测决策路径、回复用户问题中部分问题的至少一条待测决策路径、以及回复用户问题中所有问题的至少一条待测决策路径，实现大模型进行深度搜索，并得到多条待测决策路径。

可选地，初始时可设定让模型生成一个不回答用户问题的决策路径。然后在这个路径的基础之上使用如下指令进行迭代。然后这个指令生成的新的决策路径再次输入，这样多

次循环。示例性地：

Goals

你现在是在一个营销场景中，你的最终目标是针对用户达成营销目标。用户会提问问题，你要决定如何对该问题进行回答才能够达成营销目标。

你现在已经有了一个决策路径，经验证这个决策路径不能够达成营销目标。原因是回答用户问题后，用户得到的信息太少了。你需要修改这个决策路径，更多回答一些用户问题。

值得注意的是，需要尽量少的增加回答用户问题的信息。

Input

用户问题为：宝马 x5 在深圳多少钱能够落地买到？

营销目标为：让用户填写手机号。

无效的决策路径：

Condition：{“车名称”：“宝马 x5”,“预算”：“”,“城市”：“深圳”,“意图 1 级”：“询问价格”,“意图 2 级”：“落地价”}

Action：{“回复内容”：“”,“引导内容”：“填写手机号”}

Output

接着，可以针对通过步骤 S020 获得的每条待测决策路径，都执行步骤 S030，不断以大模型输出的新待测决策路径作为种子，来循环调用大模型进行广度搜索，在每次搜索后都生成不同的执行策略，直至连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值。其中，第二设定相似度阈值也可以根据经验或实验或需求设定，在此不进行限定。

可见，针对步骤 S020 获得的每条待测决策路径，都可以通过步骤 S030 来得到其对应的多条新待测决策路径。

在大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，可以执行步骤 S040，来对步骤 S030 得到的所有待测决策路径进行评分，将分数高于设定分数阈值的待测决策路径进行存储，即可得到相应的决策路径池。其中，评分的技术原理可参见相关技术，在此不展开说明。此外，设定分数阈值也可以根据经验或实验或需求设定，在此不进行限定。

因此，通过利用大模型构建决策路径池，不仅可以从深度和广度的方面来生成多条决策路径，使得决策路径更加丰富和全面，可以支持相同用户问题下的不同策略输出；而且由于在决策路径生成过程中，都会考虑到设定策略目标，因此还可以保证生成的决策路径是符合策略目标的要求的。此外，无需人工大量干预，可以大大减少人工工作量。

此外，由于在一些应用场景中，策略目标会经常性变动，例如营销策略，因此，为实现能够快速生成新策略并使该策略生效，进而更好地提高本公开实施例提供的策略确定方法的输出准确性、有效性和实用性，在一些实施例中，本公开实施例提供的策略确定方法还提供了一种决策路径池的更新方案，即，公开实施例提供的策略确定方法还可以包括：

在步骤 S400 中，在收到策略更新请求的情况下，通过大模型基于策略更新请求中携带的新策略目标和参考问题，获得与新策略目标和参考问题相应的种子决策路径；

在步骤 S500 中，通过大模型基于新策略目标、参考问题和种子决策路径中的条件信息，生成新决策路径，并将种子决策路径更新为新决策路径，重复执行本步骤，直至大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在步骤 S600 中，在大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过大模型对所有种子决策路径进行评分，并将分数高于设定分数阈值的种子决策路径存储于决策路径池。

步骤 S400 中的策略更新请求可以由相关人员输入，以触发本公开实施例提供的策略确定方法执行上述步骤 S400~步骤 S600，实现决策路径池的更新，进而使得更新的决策路径

可以在需要的时候，对大模型进行干预，进而实现更新后的策略的生效。

上述步骤 S400~步骤 S600 的技术原理可参见步骤 S020~步骤 S040 的相关记载，在此不进行赘述。

值得注意的是，本公开上述任一实施例中的技术特征或技术方案可以相互组合，只要不存在组合矛盾即可。

为了执行上述实施例及各个可能的方式中的相应步骤，下面给出一种策略确定装置的实现方式，可选地，该策略确定装置可以采用上述图 1 所示的电子设备的器件结构。可选地，请参阅图 3，图 3 为本公开实施例提供的一种策略确定装置的功能模块图。需要说明的是，本实施例所提供的策略确定装置，其基本原理及产生的技术效果和上述实施例相同，为简要描述，本实施例部分未提及之处，可参考上述的实施例中相应内容。该策略确定装置 300 包括：

提取模块 310，被配置为：通过大模型从当前输入的用户问题中抽取出目标条件信息；目标条件信息包括目标条件槽名和目标条件槽值；

筛选模块 320，被配置为：从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径；决策路径池中的每条决策路径都包含条件信息和执行策略；

输出模块 330，被配置为：通过大模型基于用户问题和目标决策路径中的执行策略，输出用于答复用户问题的策略信息。

在一些实施例中，筛选模块 320 从预先构建的决策路径池中确定出与目标条件信息匹配的目标决策路径的过程，被配置为：

从决策路径池中，获取条件槽名与目标条件槽名相同的候选决策路径；

将候选决策路径中与目标条件槽名对应的条件槽值和目标条件槽值进行匹配，以得到目标决策路径。

在一些实施例中，筛选模块 320 将候选决策路径中与目标条件槽名对应的条件槽值和目标条件槽值进行匹配，以得到目标决策路径的过程，被配置为：

将候选决策路径中与目标条件槽名对应的条件槽值转换为第一槽值向量，并将目标条件槽值转换成第二槽值向量；

根据与第二槽位向量的相似度大于第一设定相似度阈值的第一槽位向量所在的候选决策路径，确定目标决策路径。

在一些实施例中，该策略确定装置 300 还包括：

构建模块，被配置为：

通过大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与设定策略目标和设定用户问题相应的当前待测决策路径；

通过大模型基于设定策略目标、设定用户问题和当前待测决策路径中的条件信息，生成新待测决策路径，并将当前待测决策路径更新为新待测决策路径，重复执行本步骤，直至大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过大模型对所有待测决策路径进行评分，并对分数高于设定分数阈值的待测决策路径进行存储，以得到决策路径池。

在一些实施例中，该策略确定装置 300 还包括：

更新模块，被配置为：

在收到策略更新请求的情况下，通过大模型基于策略更新请求中携带的新策略目标和参考问题，获得与新策略目标和参考问题相应的种子决策路径；

通过大模型基于新策略目标、参考问题和种子决策路径中的条件信息，生成新决策路径，并将种子决策路径更新为新决策路径，重复执行本步骤，直至大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值

的情况下，通过大模型对所有种子决策路径进行评分，并将分数高于设定分数阈值的种子决策路径存储于决策路径池。

在一些实施例中，输出模块 330 通过大模型基于用户问题和目标决策路径中的执行策略，输出用于答复用户问题的策略信息的过程，被配置为：

根据目标决策路径中的执行策略，生成相应的引导文本；

通过大模型基于用户问题和引导文本，输出用于答复用户问题的策略信息。

在一些实施例中，该策略确定装置 300 还包括：

训练模块，被配置为：

通过预先构建的训练样本集对大模型进行训练，直至大模型收敛；其中，训练样本集包括多组训练样本，每组训练样本均包括用户问题样本及其对应的标签；标签用于表征用户问题样本对应的真实条件信息。

可选地，上述模块可以软件或固件（Firmware）的形式存储于图 1 所示的存储器中或固化于电子设备的操作系统（Operating System，OS）中，并可由图 1 中的处理器执行。同时，执行上述模块所需的数据、程序的代码等可以存储在存储器中。

在本公开所提供的几个实施例中，应该理解到，所揭露的装置和方法，也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的，例如，附图中的流程图和框图显示了根据本公开的多个实施例的装置、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分，模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现方式中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

另外，在本公开各个实施例中的各功能模块可以集成在一起形成一个独立的部分，也可以是各个模块单独存在，也可以两个或两个以上模块集成形成一个独立的部分。

功能如果以软件功能模块的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。这样的理解，本公开的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本公开各个实施例方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

以上仅为本公开的优选实施例而已，并不用于限制本公开，对于本领域的技术人员来说，本公开可以有各种更改和变化。凡在本公开的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本公开的保护范围之内。

工业实用性

采用上述方案，无需对大模型进行训练，也可以结合决策路径来实现大模型输出结果的准确性，具备较低的模型应用成本和较高的模型应用效率。

权利要求书

1、一种策略确定方法，其特征在于，包括：

通过大模型从当前输入的用户问题中抽取出目标条件信息；所述目标条件信息包括目标条件槽名和目标条件槽值；

从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径；所述决策路径池中的每条决策路径都包含条件信息和执行策略；

通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息。

2、根据权利要求 1 所述的方法，其特征在于，所述从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径的步骤，包括：

根据所述决策路径池中各决策路径的条件槽名和条件槽名对应的条件槽值，以及所述目标条件信息，从所述决策路径池中确定出与所述目标条件信息匹配的目标决策路径。

3、根据权利要求 2 所述的方法，其特征在于，所述根据所述决策路径池中各决策路径的条件槽名和条件槽名对应的条件槽值，以及所述目标条件信息，从所述决策路径池中确定出与所述目标条件信息匹配的目标决策路径的步骤，包括：

从所述决策路径池中，获取条件槽名与所述目标条件槽名相同的候选决策路径；

将所述候选决策路径中与所述目标条件槽名对应的条件槽值和所述目标条件槽值进行匹配，以得到目标决策路径。

4、根据权利要求 3 所述的方法，其特征在于，所述将所述候选决策路径中与所述目标条件槽名对应的条件槽值和所述目标条件槽值进行匹配，以得到目标决策路径的步骤，包括：

将所述候选决策路径中与所述目标条件槽名对应的条件槽值转换为第一槽值向量，并将所述目标条件槽值转换成第二槽值向量；

根据与所述第二槽位向量的相似度大于第一设定相似度阈值的第一槽位向量所在的候选决策路径，确定目标决策路径。

5、根据权利要求 3 或 4 所述的方法，其特征在于，所述从所述决策路径池中，获取条件槽名与所述目标条件槽名相同的候选决策路径的步骤，包括：

将所述决策路径池中各决策路径的条件槽名转换为第一槽名向量，并将所述目标条件槽名转换为第二槽名向量；

根据与所述第二槽名向量的相似度大于第二设定相似度阈值的第一槽名向量所在的决策路径，确定候选决策路径。

6、根据权利要求 2-5 任一项所述的方法，其特征在于，所述根据所述决策路径池中各

决策路径的条件槽名和条件槽名对应的条件槽值，以及所述目标条件信息，从所述决策路径池中确定出与所述目标条件信息匹配的目标决策路径的步骤，包括：

将所述决策路径池中各决策路径的条件槽名与所述目标条件槽名进行匹配，并将各决策路径的条件槽名对应的条件槽值与所述目标条件槽值进行匹配，得到目标决策路径。

7、根据权利要求 1-6 中任一项所述的方法，其特征在于，所述决策路径池的构建过程包括：

通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径；

通过所述大模型基于所述设定策略目标、所述设定用户问题和当前待测决策路径中的条件信息，生成新待测决策路径，并将当前待测决策路径更新为所述新待测决策路径，重复执行本步骤，直至所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过所述大模型对所有待测决策路径进行评分，并对分数高于设定分数阈值的待测决策路径进行存储，以得到决策路径池。

8、根据权利要求 7 所述的方法，其特征在于，所述通过所述大模型基于所述设定策略目标、所述设定用户问题和当前待测决策路径中的条件信息，生成新待测决策路径，包括：

针对设定的相同的用户问题，多次调用所述大模型生成当前新的决策路径，直至多次生成的当前新的决策路径发生重复，则将当前新的决策路径作为所述新待测决策路径。

9、根据权利要求 7 所述的方法，其特征在于，所述通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径的步骤，包括：

通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，基于提示词工程，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径。

10、根据权利要求 7 所述的方法，其特征在于，所述通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径的步骤，包括：

通过所述大模型基于输入的设定策略目标、设定用户问题和决策路径样例，根据用户问题的深浅程度，获得与所述设定策略目标和所述设定用户问题相应的当前待测决策路径的步骤。

11、根据权利要求 1-10 中任一项所述的方法，其特征在于，所述方法还包括：

在收到策略更新请求的情况下，通过所述大模型基于所述策略更新请求中携带的新策

略目标和参考问题，获得与所述新策略目标和所述参考问题相应的种子决策路径；

通过所述大模型基于所述新策略目标、所述参考问题和种子决策路径中的条件信息，生成新决策路径，并将所述种子决策路径更新为所述新决策路径，重复执行本步骤，直至所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值；

在所述大模型连续两次输出的新决策路径中的执行策略的相似度达到第二设定相似度阈值的情况下，通过所述大模型对所有种子决策路径进行评分，并将分数高于设定分数阈值的种子决策路径存储于所述决策路径池。

12、根据权利要求 1-11 任一项所述的方法，其特征在于，所述通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息的步骤，包括：

根据所述目标决策路径中的执行策略，生成相应的引导文本；

通过所述大模型基于所述用户问题和所述引导文本，输出用于答复所述用户问题的策略信息。

13、根据权利要求 12 所述的方法，其特征在于，所述根据所述目标决策路径中的执行策略，生成相应的引导文本的步骤，包括：

根据所述目标决策路径中的执行策略所指示的回复内容及引导内容，生成相应的引导文本。

14、根据权利要求 1-13 任一项所述的方法，其特征在于，所述方法还包括：

通过预先构建的训练样本集对所述大模型进行训练，直至所述大模型收敛；其中，所述训练样本集包括多组训练样本，每组训练样本均包括用户问题样本及其对应的标签；所述标签用于表征所述用户问题样本对应的真实条件信息。

15、一种策略确定装置，其特征在于，包括：

提取模块，被配置为：通过大模型从当前输入的用户问题中抽取出目标条件信息；所述目标条件信息包括目标条件槽名和目标条件槽值；

筛选模块，被配置为：从预先构建的决策路径池中确定出与所述目标条件信息匹配的目标决策路径；所述决策路径池中的每条决策路径都包含条件信息和执行策略；

输出模块，被配置为：通过所述大模型基于所述用户问题和所述目标决策路径中的执行策略，输出用于答复所述用户问题的策略信息。

16、一种电子设备，其特征在于，包括处理器和存储器，所述存储器存储有能够被所述处理器执行的机器可执行指令，所述处理器可执行所述机器可执行指令以实现权利要求 1-14 中任一项所述的方法。

17、一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程

序被处理器执行时实现如权利要求 1-14 中任一项所述的方法。