An 因果单元选择的算法与复杂度结果

Haiying Huang HHAIYING@CS.UCLA.EDU
加州大学洛杉矶分校计算机科学系
Adnan Darwiche DARWICHE@CS.UCLA.EDU
加州大学洛杉矶分校计算机科学系

编辑： Mihaela van der Schaar、Dominik Janzing 和 Cheng Zhang

arXiv:2302.14412v1 [cs.AI] 28 Feb 2023

摘要

单元选择问题旨在识别那些被称为单元的对象，它们在受到刺激后最有可能表现出期望的行为模式（例如，即将流失但在受到鼓励后可能改变主意的客户）。基于反事实目标函数的单元选择问题是近来才提出的，现有研究主要关注如何根据观测数据和干预数据，对一类特定的目标函数——收益函数——给出界限；这些研究假设没有可用于评估这些函数的完整模型。我们通过提出首个精确算法来补充这一研究方向，该算法可在给定广泛类别的因果目标函数和完整指定的结构因果模型（SCM）的情况下，找出最优单元。我们证明，在这类目标函数下，单元选择是 NP^{PP}完全的；但当单元变量对应于 SCM 中的所有外生变量时，该问题是 NP完全的。我们还为所提出的算法给出了基于树宽的复杂度界限，并将其与一个用于最大后验概率（MAP）推断的著名算法联系起来。

关键词： 单元选择、结构因果模型、反事实推理

1. 引言

过去几十年间，基于两个并行层级——信息层级和推理层级（通常称为因果层级）——逐渐形成了一套因果关系理论（Pearl 和 Mackenzie，2018年；Bareinboim 等，2021年）。在推理方面，该理论将推理划分为三个复杂程度逐渐提高、也越来越接近人类推理的层次：关联、干预和反事实，以下典型概率体现了这三个层次。关联 $\Pr(y|x)$：在已观察到 x 的情况下，y 的概率（例如，已知患者发烧时，其患流感的概率）。干预 $\Pr(y_x)$：通过干预确定 x 后，y 的概率；这不同于 $\Pr(y|x)$（例如，看到气压计指针下降可以让我们了解天气，但移动气压计指针并不会带来降雨）。反事实 $\Pr(y_x|y', x')$：在 x 和 y 均不成立的情况下，如果我们确定 x，y 成立的概率（例如，一名未接种疫苗且已经死亡的患者，如果接种了疫苗，他康复的概率）。在信息方面，这些推理形式需要不同层次的知识，并分别编码为关联模型、因果模型和函数式（机制）模型；每一类模型所包含的信息都多于前一类。在概率图模型的框架下（Koller 和 Friedman，2009年），这类知识由贝叶斯网络（Pearl，1988年；Darwiche，2009年）、因果贝叶斯网络（Pearl，2000年；Peters 等，2017年；Spirtes 等，2000年）以及函数型贝叶斯网络（Balke 和 Pearl，1995年）编码，后者也称为结构因果模型（SCM）。

这一理论的一项用途，最近通过单元选择问题得到了具体体现，该问题由Li 和 Pearl（2019年提出，他们借助选择目标客户的问题来阐释这一理论

© 2023 H. Huang 和 A. Darwiche。

，即通过提供激励来促使客户续订订阅。令 c 表示客户的特征， x 表示激励， y 表示续订。可以用反事实来描述不同类型的客户。响应者（y_x,$y'_{x'}$）会在受到激励时续订订阅，否则不会续订。总是接受者（$y_x, y_{x'}$）无论是否受到激励，都会续订。总是拒绝者（y'_x,$y'_{x'}$）无论是否受到激励，都不会续订。反向者（$y_{x'}, y'_x$）会在受到激励时不续订，否则会续订。随后，可以通过优化一个包含反事实概率的表达式来识别感兴趣的客户；Li 和 Pearl, 2019 将该表达式称为收益函数。在此例中，收益函数的形式为 βPr（响应者$|c$） $+ \gamma$Pr（总是接受者$|c$） $+ \theta$Pr（总是拒绝者$|c$） $+ \delta$Pr（反向者$|c$），其中 β、γ、θ、δ 是相应的收益。换言之，可以用这个表达式为具有特征 c 的客户打分，从而选出最有希望获得激励的客户。将上述收益函数与经典损失函数（例如用于训练神经网络的损失函数）进行比较，可以看出反事实推理发挥着根本作用，因为它使我们能够根据对象（例如人或情境）对刺激的反应来区分它们。这一区别使反事实推理（因果层级的第三层）有别于更常见但不够精细的关联推理（第一层）。它也使反事实推理有别于干预推理（第二层）；干预推理同样不足以做出此类区分。

关于单元选择的现有研究聚焦于一种非常实际的情形：我们只能获得结构因果模型（SCM）的结构，以及一些观测数据和实验数据（Li 和 Pearl，2019，2022a、b、c；Li 等人，2022b）。这类数据通常不足以确定一个完全指定的结构因果模型，因此无法获得收益函数的精确值。因此，近期研究着重于界定因果概率，并尽可能收紧这些界限（Dawid 等人，2017；Mueller 等人，2021），但较少关注如何基于这些界限优化收益；一个值得注意的例外见（Li 等人，2022b、a）。本文从不同的计算角度研究单元选择问题，以补充这条研究路线。我们尤其关注将单元选择应用于结构化单元（例如决策、政策、人、情境、地区、活动）；这些单元对应于多个变量（称为单元变量）的赋值。我们假设结构因果模型已完全指定，因此能够获得任意因果目标函数的精确值，如第 2 节所述。所谓因果目标函数，是指涉及因果层级（观测、干预和反事实）任一层面上的量的表达式。这样，我们便能寻找满足广泛条件的单元。例如：哪些活动组合最能有效应对特定的人道主义需求（人类苦难、疾病、饥饿、匮乏）？应重点关注哪些地区，才能减少难民人口流动？怎样的激励政策能让客户保持参与的时间最长？随后，我们在第 3 节考察一类特定但广泛的因果目标函数，并正式定义寻找能够优化这些函数的单元这一计算问题。第 4 节专门研究此设定下单元选择的复杂度，并证明其复杂度与经典的最大后验概率（MAP）问题相同。接着，我们在第 5 节通过将单元优化问题归约为一个称为 Reverse-MAP 的新问题，给出一种精确算法。我们还用树宽的概念刻画所提算法的复杂度，并分析其复杂度如何随所用目标函数的不同而变化。最后，我们在第 6 节给出总结性评论。部分结果的证明收录于正文，其余证明见附录。

2. 结构因果模型上的反事实查询

本节回顾结构因果模型（SCM），因为单元选择问题是在这类模型上定义的；全面介绍请参见（Galles 和 Pearl，1998年；Halpern，2000年）。我们用大写字母（例如 X）表示变量，用小写字母（例如 x）表示变量的状态。我们用粗体大写字母（例如$\mathbf{X}$）表示变量集合，用粗体小写字母（例如$\mathbf{x}$）表示这些变量的赋值。二元变量 X 的状态记为 x 和 x'。我们还用 $x \in \mathbf{x}$ 表示变量 X 在变量 x 的赋值$\mathbf{x}$ 中取状态；其中，变量集合为$\mathbf{X}$。

结构因果模型包含三个组件。首先是一个有向无环图，其节点表示变量。根节点称为外生 变量，内部节点称为内生变量。 其次，模型中的每个外生变量都有一个概率分布 $\theta(U)$。第三，对于每个父节点为$\mathbf{P}$的内生变量 V，结构因果模型都包含一个方程，称为结构方程，它针对其父节点$\mathbf{P}$的每种赋值$\mathbf{p}$ ，指定 V 的一个状态。设$\mathbf{U}/\mathbf{V}$ 为结构因果模型中的外生变量和内生变量。结构因果模型所指定的概率分布$Pr(\mathbf{U}, \mathbf{V})$如下：$Pr(\mathbf{u}, \mathbf{v}) = \prod_{u \in \mathbf{u}} \theta(u)$，如果$\mathbf{U} = \mathbf{u}$与结构方程蕴含$\mathbf{V} = \mathbf{v}$；否则，$Pr(\mathbf{u}, \mathbf{v}) = 0$。

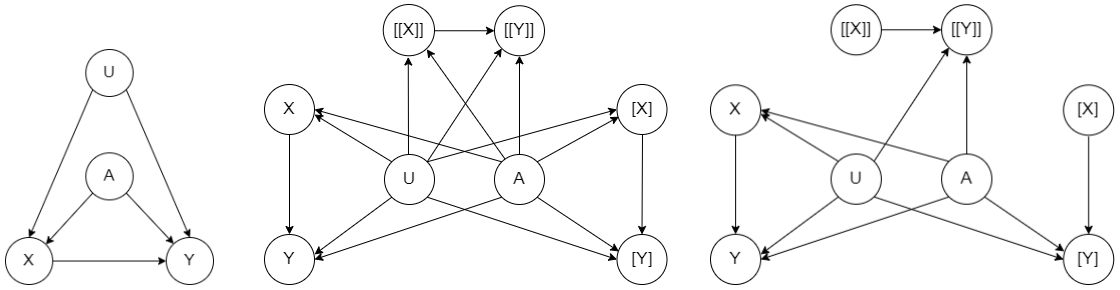
结构因果模型是一种特殊类型的贝叶斯网络（Pearl，1989；Darwiche，2009年），要求网络中的每个节点都配有条件概率表（条件概率表）。具体而言，对于父节点为$\mathbf{P}$的节点 V，条件概率表指定条件分布$Pr(V|\mathbf{P})$。结构方程可以编码为满足以下条件的条件概率表：$Pr(v|\mathbf{p}) \in \{0, 1\}$ 对所有 v 和$\mathbf{p}$。这样的条件概率表称为函数式，这也是结构因果模型有时被称为函数式贝叶斯网络的原因。

贝叶斯网络只能用于计算观测概率，例如$Pr(y|x)$，即在我们观测到 $X = x$的情况下，$Y = y$ 的概率。结构因果模型还可用于计算干预概率，例如$Pr(y_x)$，即设定$X = x$之后 $Y = y$ 的概率。结构因果模型还可用于计算反事实概率，例如$Pr(y_x, y'_{x'}|e)$，即在观测到 $E = e$的情形下（设定 $X = x$ 之后$Y = y$ 且设定 $X = x'$之后 $Y = y'$ ）的概率。[1] 我们尤其关注这种形式的反事实概率，因为它们将作为目标函数的组成部分。接下来，我们将展示如何通过在辅助模型上计算观测概率，来计算结构因果模型上的这类反事实概率。这对于本文后续使用的构造至关重要。

考虑图$Pr(y_x, y'_{x'}|x, y)$ 中结构因果模型上的反事实概率1(*a*)。此查询包含三个相互冲突的组件：y_x，$y'_{x'}$ 和(x, y)。前两个涉及冲突的行动（x 和x'）。此外，前两个组件中的行动和结果与第三个组件中的观测（x，y）相冲突。因此，计算反事实概率通常需要一个辅助模型，该模型包含多个共享相同因果机制（外生变量）的世界（真实世界和假想世界）。对于我们关注的反事实查询，正如接下来将要讨论的那样，包含三个世界的辅助模型就足够了。

给定一个结构因果模型 G，其三元组模型是另一个结构因果模型：通过创建 G 的三个副本 G^1、G^2 和 G^3，再将它们连接起来，使其共享外生变量；参见图 1(b)。如果 X 是 G^1 中的一个变量，我们用 $[X]$ 表示它在 G^2 中的副本，用 $[[X]]$ 表示它在 G^3 中的副本。三元组模型是平行世界模型（Avin 等，2005）的一种特例，平行世界模型还包括孪生模型（Balke 和 Pearl，1994）。[2] 现在，我们可以通过对三元组模型进行如下操作，计算结构因果模型 G 上的反事实概率 Pr $(y_x, y'_{x'}|x, y)$。首先，我们对副本进行残缺处理

1. 因果贝叶斯网络介于贝叶斯网络和函数型贝叶斯网络之间，可用于计算观测概率和干预概率，但无法计算反事实概率（Pearl 等人，2000年）。2. 孪生模型足以评估诸如 $Pr(y'_{x'}|x, y)$ 和 $Pr(y'_x, y_{x'})$ 这样的反事实概率，但无法评估诸如 $Pr(y_x, y'_{x'}|e)$ 这样的概率，而这正是我们感兴趣的；另见（Tian 和 Pearl，2000年；Pearl 等人，2000年）。



(*a*) 结构因果模型　(*b*) 三元组模型　(*c*) 残缺三元组模型

图 1：将模型 (a) 上的反事实概率 $Pr(y_x,\ y'_{x'}|x,\ y)$ 转化为模型 (c) 上的观测概率 $Pr([y],\ [[y']]\ |\ [x],\ [[x']],\ x,\ y)$。

G^2 和 G^3，方法是移除指向变量 $[X]$ 和 $[[X]]$ 的边，并将 $[X] = x$ 和 $[[X]] = x'$ 设定为给定值（因为我们要对这些变量进行干预）。所得结果是残缺三元组模型，如图 1(c) 所示。接着，我们可以通过计算残缺三元组模型上的观测概率 $\Pr([y], [[y']]\ |\ [x], [[x']], x, y)$，来评估结构因果模型$G$ 上的 $\Pr(y_x, y'_{x'}|x, y)$。直观来说，三元组模型可以视为刻画了三个世界：G^1、G^2 和 G^3。世界G^1 刻画观测 x, y；世界 G^2 刻画干预 $X = x$，世界 G^3 刻画干预 $X = x'$。上述处理可直接推广到形如 $\Pr(\mathbf{y_x}, \mathbf{w_v}|\mathbf{e})$ 的反事实查询，其中 E、X、Y、V、W 是变量集合。本文接下来将使用的正是这一类反事实查询，从下一节开始。

3. 因果目标函数与单元选择

因果目标函数可以是涉及观测概率、干预概率或反事实概率的任意表达式，而单元选择的目标是找到能够优化该函数的对象（单元）。然而，受（Li 和 Pearl，2019年）的启发，我们将研究范围限定于一类特定的因果目标函数，即形如 $\Pr(\mathbf{y}^i_{\mathbf{x}^i}, \mathbf{w}^i_{\mathbf{v}^i}|\mathbf{e}^i, \mathbf{u})$ 的反事实概率的线性组合，其中 $i = 1, \ldots, n$。我们将 U 称为单元变量，因为我们的目标是找到这些变量的赋值 u（即单元），使目标函数达到最优。[3] 变量 $\mathbf{X}^i\mathbf{V}^i$ 表示处理，变量 $\mathbf{Y}^i\mathbf{W}^i$ 表示结果，$(\mathbf{X}^i \cup \mathbf{V}^i) \cap (\mathbf{Y}^i \cup \mathbf{W}^i) = \emptyset$，变量 $\mathbf{E}^i$ 表示证据。目标函数的所有组件共享单元变量，但每个组件都可以有自己的处理、结果和证据变量。

我们进一步假设，单元变量**U** 在结构因果模型中是外生的（即根变量），而处理变量、结果变量和证据变量是内生的。不过，并非所有外生变量都必须是单元变量。这与（Li 和 Pearl，2022a）中的假设一致：单元变量（也称为特征）不能是处理变量或结果变量的后代。

3. 一位匿名审稿人指出，“单元”一词通常用于指代分析单位，也就是由随机变量刻画的实体。例如，在许多医学研究中，单元是患者；在许多管理学研究中，单元是公司；而在许多关于犯罪率的研究中，单元是城市或自治市。在此语境下，“单元选择”可能会被理解为选择分析单位，这与本文中该术语的用法不同。

这引出了如下形式的目标函数：[4]

$$L(\mathbf{u}) = \sum_{i=1}^{n} w_i \cdot Pr(\mathbf{y}^i_{\mathbf{x}^i}, \mathbf{w}^i_{\mathbf{v}^i} | \mathbf{e}^i, \mathbf{u}) \quad \text{where } w_i \geq 0 \text{ and } \sum_{i=1}^{n} w_i = 1 \tag{1}$$

现在，我们可以正式定义结构因果模型上的单位选择推断问题。

定义 1（单位选择）给定一个结构因果模型 G、其变量的一个子集$\mathbf{U}$ 以及一个目标函 L（$\mathbf{u}$）数，例如公式（*1*），单位选择推断问题就是计算$\text{argmax}_{\mathbf{u}}$ L（$\mathbf{u}$）。

（Li 和 Pearl，2019年）中讨论的收益函数形式如下：

$$L(u) = \beta Pr(y_x, y'_{x'}|u) + \gamma Pr(y_x, y_{x'}|u) + \theta Pr(y'_x, y'_{x'}|u) + \delta Pr(y'_x, y_{x'}|u) \tag{2}$$

通过设置 $n = 4$、$\mathbf{E}^i = \emptyset$、$\mathbf{X}^i = \mathbf{V}^i = \{X\}$和 $\mathbf{Y}^i = \mathbf{W}^i = \{Y\}$为 $i = 1$、. . .、4，其中 X、Y是二元变量，这类目标函数可作为公式（1）的一个特例。也就是说，目标函数的每个组件 i 都使用相同的单个处理变量 X 和相同的单个结果变量 Y。Li 和 Pearl（2022b）提出了一个更一般的形式，其中处理 X的取值为 x_1、. . .、 x_m，结果 Y的取值为 y_1、. . .、y_k，因此目标函数最多可以有 k^m 个组件，每个组件对应一种不同的响应类型，例如 $\Pr(y_{2x_1}, y_{1x_2}, y_{1x_3}, y_{3x_4}, y_{2x_5}|u)$，其中 $k = 3$且 $m = 5$。这类目标函数比公式（1）更一般，因为它允许表达更多响应类型，但假定只有一个处理变量和一个结果变量。我们在公式（1）中考虑的目标函数类别允许使用复合处理和结果，也允许我们从特定群体中选择单位。例如，如果 A和 B是两种药物（二元处理），而 T和 P分别表示高体温和高血压（二元结果），且 E 是取值为e_1、. . .、 e_4的年龄组，那么目标函数可以包含如下项：$\Pr((t, p')_{a,b}, (t', p')_{a',b} \mid e_3, u)$。该概率表示：如果给第三年龄组的一名成员同时服用两种药物，他会出现高体温和正常血压；如果只服用第二种药物，他则会出现正常体温和正常血压。此外，由于目标函数的各个组件可以具有不同的处理变量和结果变量，因此可以根据单位对不同刺激的响应来选择单位（例如，一种鼓励措施对会员续期的影响，以及另一种鼓励措施同时对购买增加的影响）。[5]

4. 单元选择的复杂度

接下来，我们证明，对于方程（PP1）中给出的因果目标函数类，单元选择是 NP完全的。我们还证明，当单元变量对应于结构因果模型中的所有外生变量时，这个问题也是 NP完全的。[6] 我们首先给出从单元选择到一种广为人知的最大后验推断问题变体的有效归约

4. 我们对权重 w_i 施加的条件是为了方便，并非限制性条件。5. 超越公式（1）的形式，还可以使用包含更一般要素的因果目标函数，例如：在患者并非反向者的条件下，该患者是响应者的概率， $Pr(y_x, y'_{x'}|\neg(y'_x, y_{x'}))$；或者，在患者既未节食（$y'_d$）也未锻炼（$y'_e$）的条件下，如果患者节食或锻炼，本来不会中风的概率（d', e'），即 $Pr(y'_d \vee y'_e|y, d', e')$。文献中尚未探讨过这类一般量，但一些讨论论证了其重要性，并研究了若干特殊情形；例如（Pearl，2017年）。6. 关于与贝叶斯网络推理相关的复杂性类，Shimony（1994）讨论了最可能解释决策问题为NP完全问题；Park（2002年；Park 和 Darwiche，2004a年）则讨论了最大后验概率决策问题为NP^{PP}完全问题。Roth（1996年）证明，在贝叶斯网络中计算节点边缘概率属于#P完全问题。关于这些复杂度结果的教科书式讨论，参见Darwiche（2009年，第11章）。

，我们将其称为 Reverse-MAP。随后，我们研究 Reverse-MAP 和单元选择的复杂度。

请回顾，我们的目标是找到能够使$\mathbf{u}$ 的目标函数值 $L(\mathbf{u})$ 最大化的单元。解决这一优化问题的第一步，是能够计算目标函数 $L(\mathbf{u})$。接下来，我们将展示一种构造，通过在扩展且经过删减的模型上评估涉及单元变量 $\mathbf{U}$ 的单个观测概率，来计算 $L(\mathbf{u})$ 。该构造有两个用途。首先，它将使我们能够刻画使用公式（1）形式的目标函数时单元选择的复杂度。其次，我们将在后文利用该构造，开发一种使用这些目标函数求解单元选择问题的具体算法。

考虑公式（$Pr(\mathbf{y}^i\mathbf{x}^i, \mathbf{w}^i\mathbf{v}^i|\mathbf{e}^i, \mathbf{u})$ 中的每一项1）。我们在第 2 节中介绍过，如何将这一量约化为三元组模型上的经典条件概率 G^i。下一步是将这些条件概率的线性组合编码为某个模型上的条件概率 G'。这可以通过以下构造实现。

定义 2（目标模型）考虑一个结构因果模型 G 其参数为 θ，目标函数为L 公式（*1*）中的目标模型 G' 对于 $\langle G, L\rangle$ 具有参数 θ'，其构造如下：

1. 构造一个三元组模型 G^i，对应于 G 中的每一项$Pr(\mathbf{y}^i\mathbf{x}^i, \mathbf{w}^i\mathbf{v}^i|\mathbf{e}^i, \mathbf{u})$，其中 L （参见第2节）。将G^1、.....、G^n 连接起来，使它们共享单元变量$\mathbf{U}$。由此得到模型 G'。

2. 添加节点 H 到 G'，使其成为所有结果节点的父节点$\mathbf{Z} = \{[\mathbf{Y}^i]、[[\mathbf{W}^i]]\}_{i=1}^n$。节点 H 的状态为 h_1、.....、h_n，先验概率为 $\theta'(h_i) = w_i$。现在，每个节点 $Z \in \mathbf{Z}$ 的父节点为 $\mathbf{P}_Z \cup \{H\}$，其中 $\mathbf{P}_Z$ 是添加节点 H 之前模型 G' 中节点 Z 的父节点。设 z^i 为目标函数 L在相应赋值$\mathbf{y}^i\mathbf{w}^i$ 中节点 Z 的状态。Z 的新条件概率表为：

$\mathbf{P}_Z$	H	Z	$\theta'(Z\|\mathbf{P}_Z,\ H)$
$\mathbf{p}$	h_i	z^i	$\theta(z^i\|\mathbf{p})$
$\mathbf{p}$	h_i	$\bar{z}^i$	$\theta(\bar{z}^i\|\mathbf{p})$
$\mathbf{p}$	$\bar{h}_i$	z^i	1.0
$\mathbf{p}$	$\bar{h}_i$	$\bar{z}^i$	0.0

这里，$\bar{z}^i$、$\bar{h}_i$ 表示变量 Z、H 中与状态 z^i、h_i不同的任意状态。

我们称目标模型 G' 具有 n 个组件，并将 H 称为混合变量，因为它编码了目标函数项的混合。模型 G' 中变量 $[\mathbf{Y}^i]$、$[[\mathbf{W}^i]]$ 的条件概率表，在 $H = h_i$时会还原为结构因果模型 G 中的原始条件概率表；当 $H \neq h_i$. 时，则意味着 $[\mathbf{Y}^i] = \mathbf{y}^i$、$[[\mathbf{W}^i]] = \mathbf{w}^i$。结构因果模型 G 中的目标 $L(\mathbf{u})$ 在目标模型 G' 中是一个经典概率（证明见附录A）。

定理*3* 设有一个包含单元变量 U 的结构因果模型 G。令 L 为公式（*1*）中的目标函数，令 G' 为 $\langle G, L\rangle$ 的一个目标模型。令 $X = \{[X^i]\}_{i=1}^n$、$Y = \{[Y^i]\}_{i=1}^n$、$W = \{[[W^i]]\}_{i=1}^n$、$V = \{[[V^i]]\}_{i=1}^n$ 以及 $E = \{E^i\}_{i=1}^n$。则有 $L(u) = Pr'(y, w|x, v, e, u)$，其中 y、w、x、v、e 是目标函数 L中变量 Y、W、X、V、E 的赋值。

考虑图 1(a) 中的结构因果模型和因果目标函数 $L(u) = w_1 \cdot \Pr(y_x, y'_{x'}|u) + w_2 \cdot \Pr(y_x, y_{x'}|u)$。图 2 展示了根据定义 2 构建的相应目标模型 G'。现在我们有 $L(u) = \Pr'([y^1], [[y'^1]], [y^2], [[y^2]] \mid [x^1], [[x'^1]], [x^2], [[x'^2]], u)$。

定理 3 表明，我们可以通过计算以下赋值来优化结构因果模型 $L(\mathbf{u})$ 上的目标函数 G 在目标模型 $\text{argmax}_{\mathbf{u}}\, Pr(\mathbf{y}, \mathbf{w}|\mathbf{x}, \mathbf{v}, \mathbf{e}, \mathbf{u})$ 上。它与模型 G'上的经典MAP问题相似，不同之处在于，经过优化的变量$\mathbf{U}$ 出现在条件运算符之后，而不是之前。这引出了我们对反向最大后验概率问题的定义。

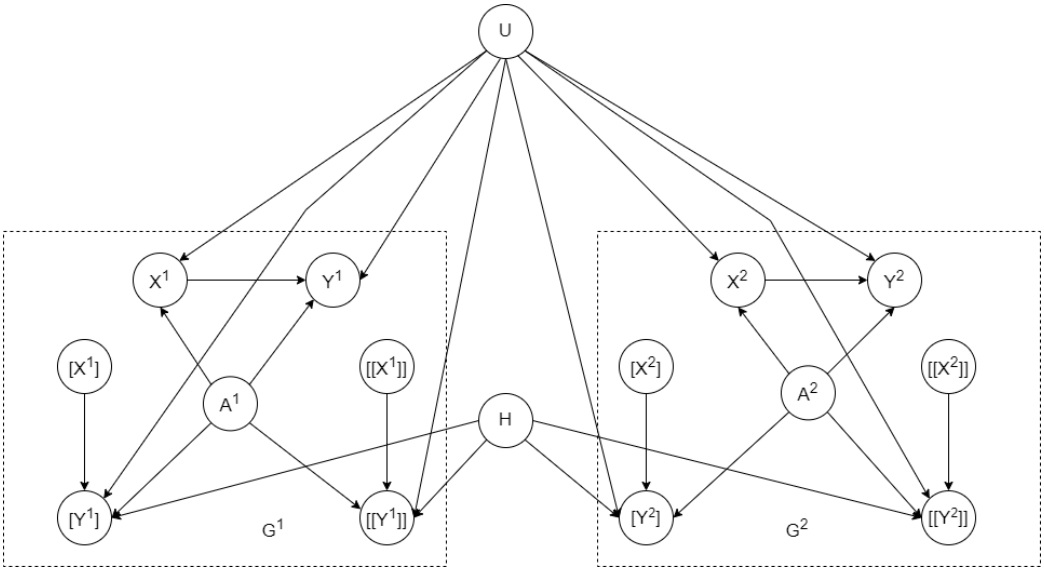


图 2：图 1（*a*） 中结构因果模型的双组件目标模型。

定义 4（Reverse-MAP）考虑结构因果模型 G，其概率分布为 Pr，并假设$\mathbf{U}$、$\mathbf{E}_1$,和$\mathbf{E}_2$是 G中互不相交的变量集。变量$\mathbf{U}$ 的反向最大后验概率赋值在给定赋值 $\mathbf{e}_1$, 和$\mathbf{e}_2$ 时定义如下：$\mathrm{RMAP}(\mathbf{U}, \mathbf{e}_1, \mathbf{e}_2) \triangleq \mathrm{argmax}_{\mathbf{u}} Pr(\mathbf{e}_1 \mid \mathbf{u}, \mathbf{e}_2)$。

为了了解反向最大后验概率与最大后验概率之间的联系，注意到 $\mathrm{argmax}_{\mathbf{u}} \Pr(\mathbf{e}_1|\mathbf{u}, \mathbf{e}_2) = \mathrm{argmax}_{\mathbf{u}} \Pr(\mathbf{u}, \mathbf{e}_1, \mathbf{e}_2)/\Pr(\mathbf{u}, \mathbf{e}_2)$，其中 $\mathrm{argmax}_{\mathbf{u}} \Pr(\mathbf{u}, \mathbf{e}_1, \mathbf{e}_2) = \mathrm{argmax}_{\mathbf{u}} \Pr(\mathbf{u}|\mathbf{e}_1, \mathbf{e}_2)$ 是众所周知的最大后验概率问题（Pearl，1989）。一般而言，最大后验概率赋值 $\mathrm{argmax}_{\mathbf{u}} \Pr(\mathbf{u}, \mathbf{e}_1, \mathbf{e}_2)$ 并非反向最大后验概率赋值，因为 $\Pr(\mathbf{u}, \mathbf{e}_2)$ 也取决于 U；附录 B 中给出了一个具体示例来说明这一点。我们现在得到以下结果，该结果已在附录 C 中证明。

推论5 反向最大后验概率问题与目标函数形式如公式（*1*）的单元选择问题之间存在多项式时间归约。

接下来，我们将刻画不同条件下反向最大后验概率问题的复杂度。考虑该问题的判定版本 D-Reverse-MAP，定义如下。

定义 6（D-Reverse-MAP）给定一个参数为有理数且诱导出概率分布 Pr 的结构因果模型、一些目标变量 $\mathbf{U}$、一些证据 $\mathbf{e}_1, \mathbf{e}_2$ 以及一个有理数阈值 p，*D-Reverse-MAP* 问题是要判断是否存在一个赋值 $\mathbf{u}$，使得 $Pr(\mathbf{e}_1|\mathbf{u}, \mathbf{e}_2) > p$。

下一个定理表明，D-Reverse-MAP 是 $\mathrm{NP}^{\mathrm{PP}}$完全问题，与经典 MAP 一样（Park 和 Darwiche，2004b）。其证明可见附录 D。

定理7 *D-Reverse-MAP* 是 NP^{PP}完全问题。

现在，我们可以利用定理 7 和推论 5来刻画单元选择的复杂度。

推论 8 单元选择是 NP^{PP}-完全的，前提是目标函数采用方程（*1*）。

在结构因果模型中，外生（根）变量表示模型中的所有不确定性，而内生（内部）变量则由外生变量唯一确定。当单元变量对应于结构因果模型的所有外生变量时，结构因果模型的这一特性会显著降低单元选择的复杂度。以下结果说明了这一点，其证明见附录 E。

定理 9 如果 *D-Reverse-MAP* 的目标变量全都是结构因果模型的根变量，那么它是 *NP*完全问题。

推论10 当单元变量全都是结构因果模型的外生（根）变量时，假设目标函数如方程（*1*）所示，单元选择是 *NP*完全问题。

5. 使用变量消元进行单元选择

第 4 节给出了将结构因果模型上的单元选择归约为目标模型上的 Reverse-MAP 的方法。在第5.1节中，我们给出了一个用于 Reverse-MAP 的变量消除（VE）算法，该算法可应用于目标模型以解决单元选择问题。在第5.2节中，我们分析了该方法的复杂度，并将其与底层结构因果模型上 Reverse-MAP 的复杂度进行比较。

5.1. 使用变量消元的 Reverse-MAP

用于 Reverse-MAP 的变量消元算法将采用经典 MAP 变量消元算法所使用的相同机制和技术（Dechter, 1999年）。因此，我们将首先依据（Darwiche，2009年，第 10 章）中的处理方法回顾 MAP 的变量消元算法，然后讨论 Reverse-MAP 算法。

变量消元算法基于因子 f(X) 的概念，该因子将变量 X 的每种赋值 x 映射为一个非负数 f(x)。变量消元算法会使用多种因子运算，包括将两个因子相乘（$f \cdot g$）、对因子中的变量求和消元（$\sum_X f$）、对因子中的变量取最大值消元（$\max_X f$），以及将两个因子相除（f/g）。设 G 是一个结构因果模型，并假定其变量 Z 被划分为三个互不相交的集合 U、 V、 E，其中 U 为目标变量，E 为证据变量。以下讨论中令 S $=$ Z\U。我们将把结构因果模型 G 中每个变量 Z 的条件概率表视为定义在该变量及其父节点 P 上的因子，记为 f_Z(ZP)。于是，结构因果模型的概率分布为 Pr (Z) $=\prod_{Z\in \mathbf{Z}} f_Z$。我们通过为每个$e \in$ e 创建证据因子 $\lambda_e(E)$ 来表示证据 e，其中 $\lambda_e(e')=1$ 当 $e'=e$，否则 $\lambda_e(e')=0$。MAP 概率为 [7]

$$\text{MAP}_p(\mathbf{U},\mathbf{e}) = \max_{\mathbf{u}} Pr(\mathbf{u},\mathbf{e}) = \max_{\mathbf{u}} \sum_{\mathbf{v}} Pr(\mathbf{u},\mathbf{v},\mathbf{e}) = \max_{\mathbf{U}} \sum_{\mathbf{S}} \prod_{Z\in\mathbf{Z}} f_Z \prod_{e\in\mathbf{e}} \lambda_e(E) \quad (3)$$

这与 *MAP*赋值形成对比，后者为 $\text{argmax}_{\mathbf{u}} Pr(\mathbf{u},\mathbf{e})$。通过一些简单的簿记处理，用于计算 MAP 概率的变量消元算法也可以返回一个 MAP赋值；例如，见（Darwiche，2009年，第10章）。因此，接下来我们将重点讨论如何计算 MAP 概率。

考虑图3 中的结构因果模型，并假设 $\mathbf{U}=\{A,B\}$，且证据 $\mathbf{e}$ 为 $\{E=e\}$。在这种情况下，最大后验概率$_p(AB,\ e)$ 将等于

$$\max_{AB} \sum_{CDE} f_A(A) f_B(AB) f_C(AC) f_D(BCD) f_E(CE) \lambda_e(E) \quad (4)$$

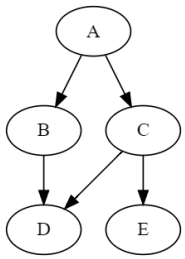


图3：结构因果模型

7. 方程 3 的左侧是一个标量（概率），而右侧是定义在空变量集上的一个因子，称为标量因子。这种因子仅将一个赋值，即空赋值，映射为一个标量。

算法1 MAP VE(G，$\mathbf{U}$，$\mathbf{e}$)

输入： 结构因果模型 G，目标变量$\mathbf{U}$，证据 $\mathbf{e}$
输出： 包含最大后验概率p($\mathbf{U}$，$\mathbf{e}$)的标量因子
1: **过程**MAIN
2: $\pi_{\mathbf{S}} \leftarrow$非目标变量的一个消元顺序
3: $\pi_{\mathbf{U}} \leftarrow$ 目标变量的消元顺序$\mathbf{U}$
4: $\mathcal{F} \leftarrow \{f:$ f是结构因果模型 $G\} \cup \{\lambda_e$ 的条件概率表：λ_e是 $e \in \mathbf{e}\}$的证据因子
5: $\mathcal{G} \leftarrow$ 消元($\sum$，$\mathcal{F}$，$\pi_{\mathbf{S}}$)
6: $p \leftarrow$ 消元(最大值 $\mathcal{G}$，$\pi_{\mathbf{U}}$)
7: **返回** p
8: </style}结束过程

算法2使用求和或最大值消去变量

输入： 运算 $\bigcirc \in \{\sum$,最大值$\}$、因子集 $\mathcal{F}$、全变量顺序 π
输出： 因子集
1: **过程** 消元($\bigcirc, \mathcal{F}, \pi$)
2: **对于** $i = 1$顺序的长度 π **执行**
3: $V \leftarrow i^{th}$ 顺序中的变量 π
4: $\mathcal{G} \leftarrow$ $\mathcal{F}$中包含变量 V的因子
5: $f_i \leftarrow \prod_{f \in \mathcal{G}} f$
6: $f_i \leftarrow \bigcirc_V f_i$
7: 将 $\mathcal{G}$ $\mathcal{F}$中的因子替换为因子 f_i
8: **结束循环**
9: **返回** $\mathcal{F}$
10: **结束过程**

对该表达式进行朴素求值时，会将所有因子相乘，得到一个关于所有变量的因子 f($ABCDE$)，然后计算 $\max_{AB} \sum_{CDE} f(ABCDE)$，其复杂度为 $O(n \exp(n))$，其中n 是模型变量的数量。变量消元算法尝试更高效地计算此表达式，其伪代码见算法1（MAP VE）。第 4 行的因子乘积 $\mathcal{F}$ 表示联合分布 Pr(Z, e)，因此我们首先在第 5 行对 $\mathcal{F}$中的变量 S 求和，计算得到因子集 $\mathcal{G}$，其乘积表示边缘分布 Pr(U, e)。然后在第 6 行对 $\mathcal{G}$ 中的变量 U 求最大值，得到标量因子 p，其中包含 MAP 概率（见脚注 7）。算法1使用算法2，按照全变量顺序 $\pi = \langle \pi_{\mathbf{S}}, \pi_{\mathbf{U}} \rangle$逐个消去变量，该顺序称为消元顺序。由于求和与最大化不可交换，MAP VE 要求变量 U 在顺序 π 中排在最后。满足此约束的顺序称为 U 约束消元顺序。MAP VE 的复杂度取决于所用的消元顺序 π。在算法2的每个消元步骤中，我们将所有涉及变量 $\pi(i)$ 的因子相乘，得到第 6 行的因子 f_i。f_i 中的变量称为簇 C_i，因此消去变量 $\pi(1), \ldots, \pi(n)$ 会产生团 $C1, \ldots, C_n$。消元顺序 π 的宽度 w是最大簇的大小减一，MAP VE 的复杂度为 $O(n \exp(w))$。

下表展示了在使用消元顺序 $\pi = E$、D、C、B、A 计算方程（4）中的 MAP 概率时，MAP VE 的执行轨迹。该轨迹表明，MAP VE 求值了

以下 owi因子化表达式，且顺序 π 的宽度为 2（最大簇的大小为 3）：

$$\text{MAP}_p(AB, e) = \max_A f_A(A) \left[\max_B f_B(AB) \left[\sum_C f_C(AC) \left[\sum_D f_D(BCD) \left[\sum_E f_E(CE)\lambda_E \right] \right] \right] \right]$$

i	被消去变量	因子 $\mathcal{G}$ （第4行，算法2）	新因子 f_i （第6行，算法2）	$\mathbf{C}_i$
1	E	$f_E(CE)$ λ_E	$f_1 = \sum_E f_E(CE)$ λ_E	CE
2	D	$f_D(BCD)$	$f_2 = \sum_D f_D(BCD)$	BCD
3	C	$f_C(AC)$ $f_1(C)$ $f_2(BC)$	$f_3 = \sum_C f_C(AC)$ $f_1(C)$ $f_2(BC)$	ABC
4	B	$f_B(AB)$ $f_3(AB)$	$f_4 =$ 最大值$_B$ $f_B(AB)$ $f_3(AB)$	AB
5	A	$f_A(A)$ $f_4(A)$	$p =$ 最大值$_A$ $f_A(A)$ $f_4(A)$	A

选择良好的消元顺序对于变量消元算法的复杂度至关重要。结构因果模型的树宽定义为 G 任意消元顺序所达到的最小宽度。由于最大后验概率要求$\mathbf{U}$约束的顺序，$\mathbf{U}$约束树宽定义为 G 任意$\mathbf{U}$约束消元顺序所达到的最小宽度（Park 和 Darwiche，2004b）。

现在我们可以介绍用于 Reverse-MAP 的变量消元算法了。同样地，我们假设模型变量 $\mathbf{Z}$ 被划分为互不相交的集合 $\mathbf{U}$、$\mathbf{V}$、$\mathbf{E}$，其中 $\mathbf{U}$ 是目标变量，而$\mathbf{S} = \mathbf{Z} \setminus \mathbf{U}$。但我们进一步将证据变量 $\mathbf{E}$ 划分为 $\mathbf{E}_1$ 和 $\mathbf{E}_2$。同样地，我们关注于计算 Reverse-MAP 概率 RMAP$_p$($\mathbf{U}$、$\mathbf{e}_1$,$\mathbf{e}_2$)，而不是以下赋值:

$$\max_{\mathbf{u}} Pr(\mathbf{e}_1|\mathbf{u}, \mathbf{e}_2) = \frac{Pr(\mathbf{u}, \mathbf{e}_1, \mathbf{e}_2)}{Pr(\mathbf{u}, \mathbf{e}_2)} = \frac{\sum_{\mathbf{v}} Pr(\mathbf{u}, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2)}{\sum_{\mathbf{v}} Pr(\mathbf{u}, \mathbf{v}, \mathbf{e}_2)} = \max_{\mathbf{U}} \frac{\sum_{\mathbf{S}} \prod_{Z \in \mathbf{Z}} f_Z \prod_{e \in \mathbf{e}_1 \cup \mathbf{e}_2} \lambda_e}{\sum_{\mathbf{S}} \prod_{Z \in \mathbf{Z}} f_Z \prod_{e \in \mathbf{e}_2} \lambda_e}$$

我们的算法称为 RMAP VE，按照算法 3 所示执行两轮消除。第一轮（第 4 行）在证据 $\mathbf{e}_1$,$\mathbf{e}_2$ 下对变量 S 求和消除，第二轮（第 5 行）在证据 $\mathbf{e}_2$下对变量 S 求和消除。由此得到两个因子集 $\mathcal{G}_1$ and$\mathcal{G}_2$，分别对应边缘分布 Pr(U, $\mathbf{e}_1, \mathbf{e}_2$) 和 Pr(U, $\mathbf{e}_2$)。现在，我们需要将 Pr (U, $\mathbf{e}_1, \mathbf{e}_2$) 除以 Pr(U, $\mathbf{e}_2$)，以计算 Pr($\mathbf{e}_1$|U, $\mathbf{e}_2$)。接下来，我们将说明，按照第 8 行所示对 $\mathcal{G}_1$ 和 $\mathcal{G}_2$ 进行“相除”，即可高效地完成这一计算。关键思想是，如果我们按照相同的消元顺序执行两轮消除，那么 $\mathcal{G}_1$ 和 $\mathcal{G}_2$ 中的因子之间就存在一一对应关系。设 (g_1^i, g_2^i) 是 $i = 1, \ldots, k$ 的对应因子对，其中 $k = |\mathcal{G}_1| = |\mathcal{G}_2|$。我们需要的是 $\left(\prod_{i=1}^n g_1^i\right) / \left(\prod_{i=1}^n g_2^i\right)$，因为它表示 Pr($\mathbf{e}_1$|U, $\mathbf{e}_2$)。但由于上述对应关系，这等于 $\prod_{i=1}^n g_1^i / g_2^i$。因此，我们可以将每一对对应因子相除，得到因子集 $\mathcal{G}$，如第 8 行所示。最后，我们从 $\mathcal{G}$ 中对目标变量 U 求最大值消除，从而得到反向最大后验概率（第 9 行）。

如果两者使用相同的消元顺序，RMAP VE 与 MAP VE 具有相同的复杂度。假设有 k 个因子，其中最大的因子大小为 $\mathcal{G}_1/\mathcal{G}_2/\mathcal{G}$ c。第8行中的除法成本为 $O(k\exp(c))$，而第9 行中的最大化成本至少为 $O(k\exp(c))$，因此除法成本受最大化成本支配。因此，RMAP VE 的复杂度仍然为$O(n\exp(w))$，其中 n 是变量的数量，而 w 是所使用的$\mathbf{U}$约束顺序的宽度 π。

算法 3 RMAP VE(G, $\mathbf{U}$, $\mathbf{e}_1$,$\mathbf{e}_2$)

输入： 结构因果模型 G，目标变量$\mathbf{U}$，证据 $\mathbf{e}_1$ 和 $\mathbf{e}_2$
输出： 包含 RMAP$_P$($\mathbf{U}$，$\mathbf{e}_1$,$\mathbf{e}_2$) 的标量因子
过程MAIN
  $\pi_{\mathbf{S}} \leftarrow$非目标变量的一个消元顺序
  $\pi_{\mathbf{U}} \leftarrow$ 目标变量$\mathbf{U}$的消元顺序
  $\mathcal{F}_1 \leftarrow \{f:$ f是结构因果模型 $G\} \cup \{\lambda_e$ 的条件概率表：λ_e是 $e \in \mathbf{e}_1$,$\mathbf{e}_2\}$的证据因子
  $\mathcal{F}_2 \leftarrow \{f:$ f是结构因果模型 $G\} \cup \{\lambda_e$ 的条件概率表：λ_e是 $e \in \mathbf{e}_2\}$的证据因子
  $\mathcal{G}_1 \leftarrow$ 消元($\sum, \mathcal{F}_1, \pi_{\mathbf{S}}$)
  $\mathcal{G}_2 \leftarrow$ 消元($\sum, \mathcal{F}_2, \pi_{\mathbf{S}}$)
  $\mathcal{G} \leftarrow \{g_1/g_2:$ g_1, g_2是 $\mathcal{G}_1, \mathcal{G}_2\}$中对应的因子
  $p \leftarrow$ 消元(最大值 $\mathcal{G}$，$\pi_{\mathbf{U}}$)
  返回 p
结束过程

5.2. 使用变量消除界定单元选择的复杂度

如定理3所示，我们可以通过将 RMAP VE 应用于结构因果模型（SCM）的目标模型来解决单元选择问题。然而，与给定的 SCM 相比，RMAP VE（以及 MAP VE）应用于目标模型时预计会更加昂贵，因为前者比后者更大、更稠密。但究竟会昂贵多少？特别是，当在底层 SCM 上具有可处理性时，RMAP VE 在目标模型上是否始终也具有可处理性？接下来，我们将使用通常用于分析消元算法的树宽来研究这一问题。还要注意，当应用于相同的 SCM 并使用相同的目标变量时，MAP VE 和 RMAP VE 具有相同的复杂度。

我们的出发点是研究目标模型的树宽与其底层结构因果模型的树宽之间的关系。我们将基于（Han 等人，2022）中报告的研究反事实推理复杂性的技术和结果展开研究。具体而言，给定结构因果模型 π 的一个消元顺序 G，接下来我们将展示如何构造目标模型的一个消元顺序 π'，G' 同时根据顺序 π' 的宽度，给出顺序 π的宽度界。回顾一下，我们使用 $[X]$ 和 $[[X]]$ 表示三元组模型中变量 X 的副本，其中 $X = [X] = [[X]]$ 在X为外生变量时成立。此外，如果 U是单元变量，那么在目标模型中 $U = U^1 = \cdots = U^n$。

定义 11 设 G 是一个结构因果模型，G' 是由 n 个组件组成的相应目标模型。若 π 是 G 的消元顺序，则 G' 的相应消元顺序 π' 可通过将 π 中的每个非单元变量 X 替换为 X^1、.. .、X^n、$[X^1]$、.. .、$[X^n]$、$[[X^1]]$、.. .、$[[X^n]]$，然后将混合变量 H 追加到该顺序末尾来获得。

考虑图 $\pi = A$、X、Y、U 中结构因果模型的消元顺序 1(a)。图 π' 中的目标模型的相应消元顺序为 2 如下：

$$\pi' = A^1, A^2, X^1, X^2, [X^1], [X^2], [[X^1]], [[X^2]], Y_1, Y_2, [Y_1], [Y_2], [[Y_1]], [[Y_2]], U, H$$

以下界（定理14）由引理12 和定理13 得出，定理 13 讨论的是n世界模型。给定一个结构因果模型 G 及其根节点的子集$\mathbf{U}$，一个 n世界模型可通过创建 n 个 G 副本，并使它们共享节点$\mathbf{U}$ (Han 等人，2022) 得到。当$\mathbf{U}$ 包含结构因果模型 G的所有根节点时，这一概念对应于平行世界模型（Avin 等人，2005）。具有n 个组件的目标模型可视为一个 $3n$世界模型，但额外包含一个混合节点 H 以及

一些从 H出发的边。引理12 和定理14 的证明见附录G 和附录H。

引理 12 考虑一个结构因果模型 G，并假设结构因果模型 G' 是通过将一个根节点 H添加为 G中某些节点的父节点，从 G 得到的。如果 π 是 G 的一个消元顺序，且宽度为 w，那么 $\pi' = \langle \pi, H \rangle$就是 G' 的一个消元顺序，且宽度为 $w' \leq w + 1$。

定理 13（Han 等人 （2022）） 考虑一个结构因果模型 G、其根节点的子集$\mathbf{U}$ 以及相应的n世界模型 G'。如果 G 具有一个消元顺序 π，其宽度为 w，那么就存在一个相应的消元顺序 π'，属于 G'，且其宽度为 $w' \leq n\,(w+1)\, - 1$。

定理 14 考虑一个结构因果模型 G 及其对应的目标模型 G'，后者包含 n 个组件。令 π 为 G 的一个消元顺序，并令 π' 为相应的 G'的消元顺序。如果 π的宽度为 w，且 π' 的宽度为 w'，那么 $w' \leq 3n\,(w+1)$。

推论 15 如果 w 是结构因果模型 G 的树宽，且 w' 是对应目标模型 G' （含有 n 个组件）的树宽，那么 $w' \leq 3n(w+1)$。

如前所述，RMAP VE 和 MAP VE 需要满足 $\underline{\mathbf{U}}$ 约束的消元顺序，其中单元变量$\mathbf{U}$ 在顺序中最后出现。因此，目标模型的满足 $\mathbf{U}$ 约束的消元顺序必须将混合变量 H 排在$\mathbf{U}$之前。由此引出下一个定义。

定义 16 设 G 是一个结构因果模型，其中包含单元变量$\mathbf{U}$，并设 G' 是一个对应的目标模型，含有 n 个组件。若 π 是 G 的一个满足 $\mathbf{U}$ 约束的消元顺序，则相应的、针对 G' 的满足 $\mathbf{U}$ 约束的消元顺序 π'，可通过将 π 中的每个非单元变量 X 替换为X^1、.....、X^n、$[X^1]$、.....、$[X^n]$、$[[X^1]]$、.....、$[[X^n]]$，然后将混合变量 H 插入到$\mathbf{U}$之前而得到。

考虑$\mathbf{U}$约束顺序 $\pi = A$、X、Y、U对于图 1(*a*)中的结构因果模型。相应的$\mathbf{U}$约束消元顺序（针对图 2 中的

$$\pi' = A^1, A^2, X^1, X^2, [X^1], [X^2], [[X^1]], [[X^2]], Y_1, Y_2, [Y_1], [Y_2], [[Y_1]], [[Y_2]], H, U$$

我们现在得到关于目标模型的$\mathbf{U}$约束树宽的如下界限。与树宽的界限相比，这一结果有些出乎意料。具体而言，尽管树宽的界限随目标模型中组件数量的增加而线性增长，但$\mathbf{U}$约束树宽的界限与组件数量无关。此外，$\mathbf{U}$约束树宽的界限可能取决于单元变量的数量，而树宽则不具备这一情况。

定理*17* 设 G 是一个具有单元变量 U 的结构因果模型，并设 G' 是相应的目标模型。若 π 是 G 的 U 约束消元顺序，其宽度为 w，且 π' 是 G' 的相应 U 约束消元顺序，其宽度为 w'，则 $w' \leq$最大值*(*$3w+3$、$|U|$*)*。如果公式（*1*）中的目标函数只有一个结果变量（对于所有 i，$\boldsymbol{Y}^i = \boldsymbol{W}^i = \{Y\}$），则 $w' \leq 3w + 3$。

推论18 设 G 是一个具有单元变量 $\mathbf{U}$ 的结构因果模型，并设 G' 是相应的目标模型。如果 w 和 w' 分别是 $\mathbf{U}$-约束树宽 G 和 G'，那么 $w' \leq$最大值（$3w+3$，$|\mathbf{U}|$）。此外，如果公式（*1*）中的目标函数具有单一结果变量，那么 $w' \leq 3w+3$。

上述界限的具体紧致程度取决于目标函数的性质。推论18指出了其中一种性质，(Li 和 Pearl, 2019)中的收益函数满足该性质；参见公式 (2)。此外，这些界限中的因子3是使用三元组模型所导致的，而使用三元组模型可能并非必要。考虑公式（1）的目标函数中的组件 $\Pr(y^i x_i, w^i v_i | e^i, u)$。如果对于所有 i都有 $E^i=\emptyset$，那么在构建目标模型时，孪生模型就足够了（如果 $Y^i=X^i=\emptyset$或 $W^i=V^i=\emptyset$，情况类似）。(Li 和 Pearl，2019)中的公式 (2) 的目标函数对于所有 i 都有 $E^i=\emptyset$，因此可得到更紧的界限 $w' \leq 2w+2$。更一般地，如果目标函数的性质使界限中的 $|U|$ 依赖关系得以移除，那么，当底层结构因果模型上的 RMAP VE（MAP VE）是可处理的时，目标模型上的 RMAP VE 也是可处理的。否则，推论18中的界限无法保证这一点。回顾一下，正如前文所示，使用公式（1）进行 MAP、Reverse-MAP 和单元选择均为 NP^{PP}-完全问题。

我们在附录J 中给出了一项初步实验，并详细讨论了三种算法的复杂度：（1）MAP VE（算法1），通过对结构因果模型进行操作来求解最大后验概率；（2）RMAP VE（算法3），通过对目标模型进行操作来求解单元选择；（3）一种基线暴力方法，通过对孪生模型或三元组模型进行操作来求解单元选择（具体取决于目标函数）。实验的主要发现是，随着问题规模增大，[8] MAP VE 与 RMAP VE 的复杂度差距逐渐缩小，而 RMAP VE 与暴力方法的复杂度差距则逐渐扩大（暴力方法的表现明显更差，很快便会变得不切实际）。

附录J 还指出了一类结构因果模型结构（及单元变量），对于这类结构，单元变量的数量没有上限，但在目标模型上运行 RMAP VE 的复杂度有上界。

6. 结论

我们在一种与现有研究互补的计算设置下研究了单元选择问题。我们假设结构因果模型已被完全指定，因此可以计算因果目标函数的点值，从而能够考虑比通常研究的函数更广泛的函数类。我们证明了具有这类目标函数的单元选择问题是 NP^{PP}-完全的，与经典最大后验概率问题类似，并确定了使其成为 NP-完全问题的一个直观条件。进一步地，我们基于变量消除为单元选择问题提供了一种精确算法，并根据树宽刻画了其复杂度，同时将这一复杂度与最大后验推断的复杂度联系起来。在此过程中，我们定义了一个新的推断问题 Reverse-MAP，该问题同样是 NP^{PP}-完全的，但相比最大后验推断，它更能捕捉单元选择的本质。

致谢

感谢 Yizuo Chen、Yunqiu Han、Ang Li 和 Scott Mueller 为本文早期版本提供了宝贵意见。本研究部分得到了美国海军研究办公室资助项目 N000142212501 的支持。

参考文献

Chen Avin、Ilya Shpitser 和 Judea Pearl。路径特异性效应的可识别性。发表于*IJCAI*，第357–363 页。Professional Book Center，2005年。

8. 问题规模以结构因果模型中的节点数和单元变量数来衡量。

Alexander Balke 和 Judea Pearl。反事实查询的概率评估。发表于*AAAI*，第 230–237 页。AAAI Press / 麻省理工学院出版社，1994年。Alexander Balke 和 Judea Pearl。结构模型中的反事实与策略分析。发表于*UAI*，第 11–18 页。Morgan Kaufmann，1995年。E. Bareinboim、Juan David Correa、D. Ibeling 和 Thomas F. Icard。论 Pearl 层级与因果推断的基础。2021年。技术报告 R-60，哥伦比亚大学。Adnan Darwiche。 贝叶斯网络建模与推理。剑桥大学出版社，2009年。Philip Dawid、Monica Musio 和 Rossella Murtas。因果概率。《法律、概率与风险》，16(4)：163–179，2017年。Rina Dechter。桶消元：一种统一的推理框架。《人工智能》，113(1-2)：41–85，1999年。David Galles 和 Judea Pearl。因果反事实的公理化刻画。《科学基础》，3(1)：151–182，1998年。Joseph Y Halpern。因果推理的公理化。《人工智能研究杂志》，12：317–337，2000年。Yunqiu Han、Yizuo Chen 和 Adnan Darwiche。反事实推理的复杂性。发表于动力系统的因果视角，*NeurIPS 2022*年研讨会，2022年。网址：https://arxiv.org/abs/2211.13447。Uffe Kjærulff。图的三角化——实现较小总状态空间的算法。技术报告 R 90-09，数学与计算机科学系，Strandvejen，DK 9000 奥尔堡，丹麦，1990年。Daphne Koller 和 Nir Friedman。概率图模型——原理与技术。麻省理工学院出版社，2009年。A. Li、S. Jiang、Y. Sun 和 J. Pearl。从有限总体数据中学习因果概率。技术报告 R-519， <http://ftp.cs.ucla.edu/pub/statser/r519.pdf>，计算机科学系，加利福尼亚大学洛杉矶分校，加利福尼亚州洛杉矶，2022a。Ang Li 和 Judea Pearl。基于反事实逻辑的单元选择。发表于*IJCAI*，第 1793–1799 页。ijcai.org，2019年。Ang Li 和 Judea Pearl。基于因果图的个体选择。发表于*AAAI*，第 5765–5772 页。AAAI Press，2022a。Ang Li 和 Judea Pearl。非二元处理与效应的单元选择。*CoRR*，abs/2208.09569，2022b。Ang Li 和 Judea Pearl。单元选择：案例研究及与 A/B 测试启发式方法的比较。*CoRR*，abs/2210.05030，2022c。

Ang Li、Song Jiang、Yizhou Sun 和 Judea Pearl。单元选择：从有限总体数据中学习收益函数。*CoRR*，abs/2210.08203，2022b。Michael L Littman、Judy Goldsmith 和 Martin Mundhenk。概率规划的计算复杂度。《人工智能研究杂志》，9:1–36，1998。S. Mueller、A. Li 和 J. Pearl。结果的原因：从总体数据中学习个体响应。技术报告 R-505，<http://ftp.cs.ucla.edu/pub/stat ser/r505.pdf>，加利福尼亚大学洛杉矶分校计算机科学系，加利福尼亚州洛杉矶，2021。即将发表于 IJCAI-2022 论文集。James D Park。MAP 复杂度结果和近似方法。载于人工智能不确定性第十八届会议论文集，第 388–396 页，2002。James D. Park 和 Adnan Darwiche。MAP 解释的复杂度结果和近似策略。人工智能研究杂志（*JAIR*），21:101–133，2004a年。James D. Park 和 Adnan Darwiche。MAP 解释的复杂度结果和近似策略。人工智能研究杂志，21:101–133，2004b。Judea Pearl。智能系统中的概率推理：合理推断网络。MK，1988年。Judea Pearl。智能系统中的概率推理——合理推断网络。《表征与推理》系列，Morgan Kaufmann，1989。Judea Pearl。因果关系。剑桥大学出版社，2000年。Judea Pearl。物理与形而上学反事实：评估析取行动。因果推断杂志，5(2)，2017年。doi：https://doi.org/10.1515/jci-2017-0018。Judea Pearl 和 Dana Mackenzie。《为什么之书：因果关系的新科学》。Basic Books，2018年。Judea Pearl 等人。模型、推理与推断。剑桥，英国：剑桥大学出版社，19(2)，2000年。Jonas Peters、Dominik J anzing 和 Bernhard Sch¨olkopf。因果推断要素：基础与学习算法。麻省理工学院出版社，2017年。Dan Roth。论近似推理的困难性。《人工智能》，82(1-2):273–302，1996年。Solomon Eyal Shimony。信念网络的 MAP 求解是 NP 困难的。《人工智能》，68(2):399–410，1994。Peter Spirtes、Clark Glymour 和 Richard Scheines。《因果、预测与搜索》第二版。自适应计算与机器学习。麻省理工学院出版社，2000年。Jin Tian 和 Judea Pearl。因果概率：界限与识别。《数学人工智能年刊》，28(1-4):287–313，2000年。

附录 A. 定理3的证明

该定理的证明需要一个引理，而该引理又需要以下定义。我们称变量集**Z** 分解一个有向无环图，若移除**Z** 的出边后，该有向无环图被分成至少两个非连通分量。

引理*19* 考虑一个服从概率分布 *Pr* 的结构因果模型 G，以及三个互不相交的变量集 *X*、*Y*、*Z*。假设 *Z* 可分解为 G 个非连通分量 G_1 和 G_2。如果 X_1,Y_1 是与 G_1 相关的 *X*、*Y* 的子集，而 X_2, Y_2 是与 G_2 相关的 *X*、*Y* 的子集，那么 *Pr(y|x, z)* $=$*Pr* $(y_1|x_1, z)$*Pr* $(y_2|x_2, z)$。

证明 由于 **Z** 可分解为 G，因此有 dsep$_G$(**X**$_1$,**Z**，**X**$_2$) 和 dsep$_G$(**X**$_1$**Y**$_1$, **Z**，**X**$_2$**Y**$_2$)。我们有：

$$\begin{aligned} Pr(\mathbf{y}|\mathbf{x},\mathbf{z}) &= \frac{Pr(\mathbf{y},\mathbf{x}|\mathbf{z})}{Pr(\mathbf{x}|\mathbf{z})} = \frac{Pr(\mathbf{y}_1,\mathbf{y}_2,\mathbf{x}_1,\mathbf{x}_2|\mathbf{z})}{Pr(\mathbf{x}_1,\mathbf{x}_2|\mathbf{z})} = \frac{Pr(\mathbf{y}_1,\mathbf{x}_1|\mathbf{z})Pr(\mathbf{y}_1,\mathbf{x}_1|\mathbf{z})}{Pr(\mathbf{x}_1|\mathbf{z})Pr(\mathbf{x}_2|\mathbf{z})} \\ &= \left[\frac{Pr(\mathbf{y}_1,\mathbf{x}_1|\mathbf{z})}{Pr(\mathbf{x}_1|\mathbf{z})}\right]\left[\frac{Pr(\mathbf{y}_2,\mathbf{x}_2|\mathbf{z})}{Pr(\mathbf{x}_2|\mathbf{z})}\right] = Pr(\mathbf{y}_1|\mathbf{x}_1,\mathbf{z})Pr(\mathbf{y}_2|\mathbf{x}_2,\mathbf{z}) \end{aligned} \tag{5}$$

方程（5）由 dsep$_G$（**X**$_1$,**Z**，**X**$_2$）和dsep$_G$（**X**$_1$**Y**$_1$,**Z**，**X**$_2$**Y**$_2$）得出。这就完成了证明。虽然这里我们只考虑两个子网络的情形，但很容易看出，这个引理可以推广到由**Z**分解出的任意数量的子网络。■

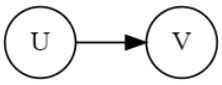
现在我们可以证明定理 3。根据 G' 的构造，$\mathbf{U} \cup \{H\}$ 分解 G' 为其 n 组件 G^1，G^2，$\ldots G^n$。我们有：

$$\begin{aligned} Pr'(\mathbf{y},\mathbf{w} \mid \mathbf{x},\mathbf{v},\mathbf{e},\mathbf{u}) &= \sum_{i=1}^{n} Pr'(\mathbf{y},\mathbf{w},h_i \mid \mathbf{x},\mathbf{v},\mathbf{e},\mathbf{u}) \\ &= \sum_{i=1}^{n} Pr'(\mathbf{y},\mathbf{w} \mid \mathbf{x},\mathbf{v},\mathbf{e},\mathbf{u},h_i)Pr'(h_i \mid \mathbf{x},\mathbf{v},\mathbf{e},\mathbf{u}) \\ &= \sum_{i=1}^{n} Pr'(\mathbf{y},\mathbf{w} \mid \mathbf{x},\mathbf{v},\mathbf{e},\mathbf{u},h_i)Pr'(h_i) && (6) \\ &= \sum_{i=1}^{n} \left[\prod_{j=1}^{n} Pr'(\mathbf{y}^j,\mathbf{w}^j \mid \mathbf{x}^j,\mathbf{v}^j,\mathbf{e}^j,\mathbf{u},h_i)\right] Pr'(h_i) && (7) \\ &= \sum_{i=1}^{n} \left[\prod_{j\neq i} pr'(\mathbf{y}^j,\mathbf{w}^j \mid \mathbf{x}^j,\mathbf{v}^j,\mathbf{e}^j,\mathbf{u},h_i)\right] Pr'(\mathbf{y}^i,\mathbf{w}^i \mid \mathbf{x}^i,\mathbf{v}^i,\mathbf{e}^i,\mathbf{u},h_i)Pr'(h_i) && (8) \\ &= \sum_{i=1}^{n} \left[\prod_{j\neq i} 1.0\right] Pr'(\mathbf{y}^i,\mathbf{w}^i \mid \mathbf{x}^i,\mathbf{v}^i,\mathbf{e}^i,\mathbf{u})\, w_i && (9) \\ &= \sum_{i=1}^{n} w_i\, Pr(\mathbf{y}^i_{\mathbf{x}^i},\mathbf{w}^i_{\mathbf{v}^i} \mid \mathbf{e}^i,\mathbf{u}) \\ &= L(\mathbf{u}) && (10) \end{aligned}$$

公式（6）成立，因为辅助根节点 H 与 X $\cup$V $\cup$ E $\cup$U d-分离。公式（7）根据引理19成立，因为 U$\cup\{H\}$ 分解 G 后，所有三元组模型彼此断开。公式（8）根据新条件概率表的构造成立，其中 Y^i 和 W^i的构造如下：若 $H = h_i$，则保留 Y^i 和 W^i 原有的条件概率表，并将所有 $j \neq i$对应的 Y^j 和 W^j 的值固定为 y^j 和 w^j。公式（9）由三元组网络的性质得出。

附录 B. 最大后验概率和反向最大后验概率示例

考虑图4中的简单模型。对于最大后验概率，我们有$\text{argmax}_u Pr(u, v_1) = u_2$；而对于反向最大后验概率，我们有$\text{argmax}_u Pr(v_1|u) = u_1$。



(a)

U	$\theta(A)$
u_1	0.2
u_2	0.8

(b)

U	V	$\theta(V\|U)$
u_1	v_1	0.6
u_1	v_2	0.4
u_2	v_1	0.3
u_2	v_2	0.7

(c)

图4：用于说明最大后验概率与反向最大后验概率之间差异的示例。

附录 C. 推论5

我们可以通过选择形式为公式（1）的目标函数，并进行以下设置：$n = 1, w_1 = 1$, X $^1 = \{\}$, $Y^1 = E_1$ and $E^1 = E_2$，将 Reverse-MAP argmax_u Pr (e_1|u, e_2) 归约为单元选择。这显然是一个多项式时间归约。我们已在定理3中证明了从单元选择到 Reverse-MAP 的归约。令 $|G|$ 表示结构因果模型 G 的大小，n 表示因果目标函数中组件的数量。通过检查定义 2，我们可以立即看出，构建目标模型 G' 所需的时间为 $O(n \cdot |G|)$。此外，目标模型 G' 的大小也为 $O(n \cdot |G|)$。

附录 D. 定理7 的证明

属于 NP^{PP} 显而易见。给定 U 的一个赋值 u，通过查询 PP 预言机来判断 u 是否为解很容易：查询 Pr (e_1|u, e_2) $> p$。这是一个称为 D-MAR 的问题（Darwiche，2009年）。为证明其困难性，我们展示 E-MAJSAT（Littman 等人，1998年）可在多项式时间内归约到 D-Reverse-MAP；此归约基于对 (Park，2002年；Park 和 Darwiche，2004b) 中提出的经典MAP归约所作的轻微修改。
E-MAJSAT 问题定义如下：给定一个关于布尔变量 Z $=$U$\cup$V 的布尔公式 α，是否存在 U 的一个赋值 u，使得 V 的多数赋值 v 满足 uv $\models \alpha$（即公式 α 在 uv 处成立）？我们将证明，可以通过对一个模拟公式 α、且能高效构造的结构因果模型 G_α 回答 D-Reverse-MAP 查询，从而回答 E-MAJSAT。该结构因果模型 G_α 按归纳方式构造，如下所示：

9. 结构因果模型的大小是存储其结构和参数所需的空间。例如，如果结构因果模型由函数式贝叶斯网络表示，其大小通常为网络中条件概率表条目的总数。

如 (Park 和 Darwiche，2004b) 所示，[10] 它始终只有一个叶节点，记作 S_α.。该构造基于以下三条规则：(1) 若 $\alpha = X$，则 G_α 只有一个二元节点 X，其取值为$\{0, 1\}$，且具有均匀先验，因此 $S_\alpha = X$；(2) 若 $\alpha = \neg\beta$，则通过添加一个二元节点 S_α 作为 $S_\beta \in G_\beta$ 的子节点，并令其结构方程为 $S_\alpha = 1 - S_\beta$，从而由 G_β 构造出 G_α；(3) 若 $\alpha = \beta \wedge \gamma$（$\alpha = \beta \vee \gamma$），则通过添加一个二元节点 S_α 作为$S_\beta \in G_\beta$ 和 $S_\gamma \in G_\gamma$ 的子节点，并令其结构方程为 $S_\alpha = S_\beta \cdot S_\gamma$（$S_\alpha = S_\beta + S_\gamma$），从而由 G_β 和 G_γ 构造出 G_α。现在可以进行证明的最后一步。给定一个关于变量 Z $= \mathbf{U} \cup \mathbf{V}$ 的布尔公式 α，以及其对应的、具有概率分布 Pr 的结构因果模型 G_α，接下来我们证明：存在一个赋值 u，使得 Pr ($S_\alpha = 1$|u) $> 1/2$（D-Reverse-MAP 查询），当且仅当存在一个赋值 u，使得 V 的多数赋值 v 满足 uv $\models \alpha$（E-MAJSAT 查询）。令 |Z| $= n$ 和 s_α 表示 $S_\alpha = 1$。根据 G_α 的构造（Park 和 Darwiche，2004b），对于所有赋值 z，都有 Pr (z) $= 1/2^n$；若 z $\models \alpha$，则 Pr (s_α|z) $= 1$，否则 Pr(s_α|z) $= 0$。因此，若 z $\models \alpha$，则 Pr (z s_α) $=$Pr(s_α|z)Pr (z) $= 1/2^n$；否则 Pr(z s_α) $= 0$。最后，我们得到：

$$Pr(s_\alpha|\mathbf{u}) = \frac{Pr(\mathbf{u}, s_\alpha)}{Pr(\mathbf{u})} = \frac{\sum_{\mathbf{v}} Pr(\mathbf{u}, \mathbf{v}, s_\alpha)}{\sum_{\mathbf{v}} Pr(\mathbf{u}, \mathbf{v})} = \frac{\sum_{\mathbf{v}:\mathbf{uv}\models\alpha}(1/2^n)}{\sum_{\mathbf{v}}(1/2^n)} = \frac{\mathbf{card}(\{\mathbf{v} \in \mathcal{V} : \mathbf{uv} \models \alpha\})}{\mathbf{card}(\mathcal{V})}$$

No由此我们已证明其成员资格和困难性，D-Reverse-MAP 是 NP^{PP}-完全的 e.

附录 E. 定理9 的证明

回顾 D-Reverse-MAP 的定义 6：是否存在一个赋值 u，使得 Pr (e_1|u, e_2) $> p$? 其属于 NP 显而易见。由于 U 是外生变量的集合，证据变量 E 由 U 函数决定。因此，可以先计算由 u（通过结构方程）所蕴含的 E 的赋值 e，再检查 e 是否与 e_1, e_2一致，从而轻松判断赋值 u 是否为解。如果答案为是，则 Pr (e_1|u, e_2) $= 1$；否则，Pr (u, e_2) $= 0$ 或 Pr (e_1|u, e_2) $= 0$。为证明困难性，我们证明，在定理所述条件下，SAT 可归约为 D-Reverse-MAP。给定一个关于变量 U 的布尔公式α，我们按照定理7的证明构造一个结构因果模型 G_α。结构因果模型 G_α 有一个叶节点 S_α，其根节点为 U。根据 G_α 的构造，我们有：若 u 满足 α，则 Pr ($S_\alpha = 1$|u) $= 1$；否则 Pr ($S_\alpha = 1$|u) $= 0$。通过选择 $p = 0$、$e_1 = \{S_\alpha = 1\}$ 和 $e_2 = \emptyset$，D-Reverse-MAP 查询（是否存在 u，使得 Pr ($S_\alpha = 1$|u) > 0）的答案为是，当且仅当存在满足公式 α 的赋值 u（SAT 查询）。至此，证明完毕。

附录 F. 消元概念回顾

此处我们回顾一些标准概念，包括消元过程、簇和道德图，这些概念将在接下来的一些证明中用到；关于详细介绍，请参见（Darwiche，2009年，第 9 章）。

结构因果模型 G 的道德图 通过在每对共同父节点之间添加一条无向边，然后将所有边转为无向边，从 G 得到。要从图 G中消去 变量 X，需连接 G中 X 的每对邻居，然后从 G中移除节点 X。从结构因果模型 G 中消去变量，即从其道德图 G'中消去变量。按照变量顺序 π 从道德图 G' 中消去变量，会生成一个图序列 $G' =$

10.（Park 和 Darwiche，2004b）旨在构造一个贝叶斯网络，但由于该网络中的所有内部节点都具有函数型条件概率表，他们构造的实际上是一个结构因果模型。

G_1、……、G_n，其中图 G_{i+1}是通过从 G_i中消去变量 $\pi(i)$ 得到的。我们用 $G_i(X)$ 表示图 G_i 中的 X 及其邻居。我们还用$\mathbf{C}(X)$表示变量 X的簇，即消去 X之前的 X 及其邻居。如果 $X = \pi(i)$，那么$\mathbf{C}(X) = G_i(X)$。在这种情况下，我们还用$\mathbf{C}_i$ 表示 X的簇$\mathbf{C}(X)$。

附录 G。引理的证明 12

该证明使用附录 F中回顾的消元概念和记号。

设 G'_m 是 G' 的道德图，且 $\mathbf{C}'_i$ 是从 X_i 中消去变量 G'所产生的簇。设 n 是 G中的变量数量。要证明引理 12，只需证明以下陈述：$\mathbf{C}'_i \subseteq \mathbf{C}_i \cup \{H\}$ 对于 $i = 1, \ldots\ldots, n$，下面通过归纳法证明该陈述。

令 neigh(X) 表示 X 在 G_m 中的邻居，令 neigh′(X) 表示 X在 G'_m中的邻居。令 Z 表示 H 在 G'中的子节点。对于每个 $Z \in \mathbf{Z}$，令 $\mathbf{P}_Z$ 表示 Z 在G中的父节点。首先，我们证明当 $i = 1$ 时该陈述成立。从 G'创建 G'_m 时，引入节点 H 会导致两类在 G_m 中不存在的边被添加到 G'_m中：1) 对于$Z \in \mathbf{Z}$，添加 (Z, H)；2) 如果 Y是某个节点 $Z \in \mathbf{Z}$ 的父节点，则添加 (Y, H)，也就是说，Y和 H 是某个节点 Z的共同父节点。这意味着在消除开始之前，对于任意节点 X，如果 $X \in \mathbf{Z} \cup_{Z \in \mathbf{Z}} P_Z$，则有 neigh′$(X)$ = neigh$(X) \cup \{H\}$；否则 neigh′(X) = neigh(X)。因此，$C'_1 \subseteq \{C_1\} \cup \{H\}$。现在考虑消除 X_{i+1}，假设该陈述对于 1、2、……、i均成立。我们注意到，如果 $\mathbf{C}'_j \subseteq \mathbf{C}_j \cup \{H\}$，则消除 X_j 只会导致一类额外的边被添加到 G'_m中，即对于 $Y \in$ neigh(X_j)，添加 (Y, H)。这是因为消除 X_j 会在 G'_m中使 neigh$(X_j) \cup \{H\}$ 形成一个团，但在从 G_m中消除 X_j 后，neigh(X_j)已经形成一个团。这意味着消除 X_j $(j \leq i)$绝不会导致任意两个都不是 H 的节点之间添加额外的边（换言之，所有添加的额外边都与 H相连）。因此，在消除 X_{i+1}之前，有 neigh′(X_{i+1}) $\subseteq$neigh$(X_{i+1}) \cup \{H\}$，这意味着 $\mathbf{C}'_{i+1} \cap \mathbf{C}_{i+1} \cup \{H\}$，从而完成证明。

附录 H. 定理 14

一个 N世界模型通过创建 N 个有向无环图（有向无环图）副本并将它们连接起来，使其部分根节点共享而得到（子集 （Han 等人，2022）。因此，定义 2 中的目标模型对应于一个 N世界模型，区别在于增加了混合节点 H 及其出边。在添加混合节点 H之前，具有 n 个组件的目标模型对应于一个 $3n$世界模型，因此根据定理 $\leq 3n(w+1) - 1$，其树宽为 13。添加节点 H 后，根据引理 $\leq 3n(w+1)$，其树宽为 12。

附录 I. 定理 17 的证明

本证明使用了附录 F中回顾的消元概念和记号。

对于 SCM G 中的节点 X，我们用 $[X]^k$ 表示它在 G 的 n 世界模型中的 k^{th} 副本。如果节点 X 在所有 n 个世界中共享，那么对于所有 k，$[X]^k = X$。对于变量集合 $\mathbf{X}$，我们用 $[\mathbf{X}]^k$ 表示 $\{[X]^k : X \in \mathbf{X}\}$。

设 G 为一个结构因果模型，$\mathbf{U}$ 为其根节点（单元变量）的一个子集，并设 G' 为相应的目标模型，其中包含 n 个组件。我们的证明基于构造一个增广目标模型 G''，方法是向 G' 添加边，然后证明定理17 的界限对 G''成立。我们的证明基于引理20 和22，我们稍后将正式陈述并证明这两个引理：

– 引理20 补充了定理13 ，证明任何针对结构因果模型的**U**约束消元顺序都可以转换为相应**U**约束的n世界模型的消元顺序，同时保持该顺序的宽度不变。

– 引理 22 讨论通过添加一个根节点 H 以及一些起源于 H的边来扩增结构因果模型 H。具体而言，给定结构因果模型的一个宽度为 w 的 U 约束消元顺序，该引理展示了如何构造其扩增模型的 U 约束消元顺序，其宽度为**U**-约束消元顺序，宽度为 $\leq$ 最大值（$w+1$，$|\mathbf{U}|$）。

我们首先说明如何从 G' 构造增广目标模型 G''。设 H 为 G' 和 $\mathbf{Z} = \{\mathbf{Y}^i、\mathbf{W}^i\}_{i=1}^n$ 的混合节点，其中后者是公式（1）的目标函数中的所有结果变量的集合。如果对于每个 $Z \in \mathbf{Z}$，这样一条边在 G' 中尚不存在，我们就在 G' 中添加边 $H \rightarrow Z$，由此得到 G''。G'' 的边是 G' 的边的超集，因此只需证明定理17的界限对 G'' 成立。接下来，我们用 G^t 表示 G 的一个三元组（3世界）模型。我们还用 G^b 表示通过为 G^t 添加混合节点 H 以及针对 $Z \in \mathbf{Z}$ 的边 $H \rightarrow Z$ 所得到的扩增模型。注意，增广目标模型 G'' 对应于 G^b 的 n 个副本，这些副本共享根节点 $\mathbf{U} \cup \{H\}$。因此，G'' 是 G^b 的一个 n 世界模型。

设 π 是 G 的一个 U 约束消元顺序，其宽度为 w。由于 G^t 是 G 的一个三元组（3 世界）模型，根据定理 13，存在 G^t 的一个消元顺序 π^t，其宽度为 w^t，使得 $w^t \leq 3w+2$（顺序 π^t 也将满足 U 约束）。回顾一下，G^b 是在 G^t 的基础上添加根节点 H 以及从 H 发出的一些边得到的。根据引理 22，存在 G^b 的一个消元顺序 π^b，其宽度为 w^b，使得 $w^b \leq \max(w^t+1, |\mathbf{U}|) = \max(3w+3, |\mathbf{U}|)$。此外，如果目标函数只有一个结果变量 Y，那么在 G^b 中，H 只有一个子节点 Y，因此根据引理 22，我们还可得 $w^b \leq 3w+3$。由于 G'' 是 G^b 的一个基于 G^b 的根节点 $\mathbf{U} \cup \{H\}$ 的 n-世界模型，根据引理 20，我们有 $w'' = w^b$。总之，我们有 $w' \leq w'' \leq \max(3w+3, |\mathbf{U}|)$。如果目标函数只有一个结果变量，则有 $w' \leq 3w+3$。定理17的证明至此结束。

接下来，我们将正式陈述并证明引理 20 和22 ，这两个引理在上述证明中用到了。

引理 20 考虑一个结构因果模型 G、其根节点的子集 **U** ，以及一个对应的 n-世界模型 G'，它与 G 共享**U**。若 w 是针对 G的一个受**U**约束的消元顺序 π 的宽度，而 w' 是针对 G'的对应**U**约束消元顺序 π' 的宽度，则 $w = w'$。

给定结构因果模型 G的消元顺序 π，我们可以按照 (Han 等人，2022) 中的定义 2，将其中每个变量$X \notin \mathbf{U}$ 替换为其副本 $[X]^1$、$[X]^2$、.....、$[X]^n$，从而将其转换为相应的 π' 消元顺序，适用于其 n-世界模型 G'（上面的引理中提到）。若π 满足 U 约束，则 π' 也满足 U 约束。此外，我们为 n-世界模型定义一个图序列 G'_1、G'_2、.....、G'_n，其中 G'_1 是 G'的道德图，而 G'_{i+1}是通过从 G'_i中消除变量 π的所有副本（i），即 $[\pi(i)]^1$、$[\pi(i)]^2$、.....、$[\pi(i)]^n$，得到的。

证明 假设我们按照顺序 π/π' 消去 G/G' 中的变量。我们声称，在每个消元步骤 i中，以下性质均成立：

A) 对于每个节点 $X \notin \mathbf{U}$，$G'_i([X]^k) = [G_i(X)]^k$

B) 对于每个节点 $U \in \mathbf{U}$，$G'_i(U) = \bigcup_{k=1}^n [G_i(U)]^k$

接下来我们证明性质 A)、B) 蕴含 $w' = w$，然后证明这些性质。设 $Y = \pi(i)$。若 $Y \notin \mathbf{U}$，则当从 G' 中消元其副本 $[Y]^k$ 时，有 $\mathbf{C}'([Y]^k) = [\mathbf{C}(Y)]^k$。若 $Y \in \mathbf{U}$，则当从 G' 中消元 Y 时，有 $\mathbf{C}'(Y) = \bigcup_{k=1}^{n} [G_i(Y)]^k = G_i(Y) = \mathbf{C}(Y)$，因为所有非共享节点都已在 Y 之前被消元，即 $[G_i(Y)]^k = G_i(Y)$。这意味着，从 G' 中消元一个变量所诱导的簇，其大小始终与从 G 中消元对应变量所诱导的簇相同，由此可得 $w' = w$。

接下来我们通过归纳法证明性质 A)、B)。根据 n-世界模型的定义，这些性质最初对于 G'_1成立。假设它们对于 G'_i 成立，并考虑 G'_{i+1}。令 $Y = \pi(i)$。那么， G'_{i+1}是从 G'_i中消除节点 $[Y]^1$、.....、$[Y]^n$ 的结果。我们分两种情形讨论。

情形： $Y \notin \mathbf{U}$。考虑 G'_i中的每个节点 Z。如果在 G'_i中， Z 不是 $[Y]^1$、.....、$[Y]^n$ 的邻居，那么 $G'_{i+1}(Z)$不会受到消除 $[Y]^1$、.....、$[Y]^n$ 的影响，并且根据归纳假设，这些性质成立。否则， 节点 Z 分为两种情形：A) 节点 $X \notin \mathbf{U}$的副本 $[X]^k$；B) 共享节点 $U \in \mathbf{U}$。

A) 根据归纳假设， G'_i 中 $[X]^k$ 的邻居必须属于第 k个世界，因此 $[X]^k$只能是第 k个副本 $[Y]^k$的邻居。这意味着 $G'_{i+1}([X]^k)$只能受到消除 $[Y]^k$的影响。根据变量消除的定义，我们有：

$$\begin{aligned}
G'_{i+1}([X]^k) &= G'_i([X]^k) \cup G'_i([Y]^k) \setminus \{[Y]^k\} \\
&= [G_i(X)]^k \cup [G_i(Y)]^k \setminus \{[Y]^k\} && \text{by the induction hypothesis} \\
&= [G_i(X) \cup G_i(Y) \setminus \{Y\}]^k \\
&= [G_{i+1}(X)]^k && \text{by definition of variable elimination}
\end{aligned}$$

这证明了性质 A)。

B) 根据归纳假设， U必须是所有副本 $[Y]^1$、.....、 $[Y]^n$的邻居。我们有：

$$\begin{aligned}
G'_{i+1}(U) &= G'_i(U) \bigcup_{k=1}^{n} G'_i([Y]^k) \setminus \{[Y]^k\}_{k=1}^{n} \\
&= \Big(\bigcup_{k=1}^{n} [G_i(U)]^k\Big)\Big(\bigcup_{k=1}^{n} [G_i(Y)]^k\Big) \setminus \{[Y]^k\}_{k=1}^{n} && \text{根据归纳假设} \\
&= \bigcup_{k=1}^{n} [G_i(U)]^k \cup [G_i(Y)]^k \setminus \{[Y]^k\} \\
&= \bigcup_{k=1}^{n} [G_i(U) \cup G_i(Y) \setminus \{Y\}]^k \\
&= \bigcup_{k=1}^{n} [G_{i+1}(U)]^k && \text{根据变量消除的定义}
\end{aligned}$$

这证明了性质 B)。

情形： $Y \in \mathbf{U}$。在此情形下， G'_i 仅包含$\mathbf{U}$中的节点。性质 A) 显然成立。而性质 B) 中的关系简化为 $G'_i(U) = \bigcup_{k=1}^{n} [G_i(U)]^k = G_i(U)$。根据归纳假设，我们知道 $G'_i = G_i$，因此 $G'_{i+1} = G_{i+1}$。性质 B) 成立。证明至此完成。 ■

The 引理 22 的证明需要以下关于从图中消去变量的结果 .

引理 21 考虑一个有向无环图 G、其节点的一个子集 $\mathbf{U}$ 以及一个节点 H，其中 G 中的 $H \notin \mathbf{U}$。令G_1为 G 的道德图，并令 G_2 为从G_1中消去除 $\{H\} \cup \mathbf{U}$ 之外所有节点后得到的结果。对于任意节点 $X \in \mathbf{U}$，当且仅当在 G_1 中存在一条连接 X 与 H 的路径，且该路径不包含 $\mathbf{U} \setminus \{X\}$ 中的节点时，X 与 H 在 G_2 中相邻。

证明 我们先证明充分性方向。假设在 G_1 中存在这样一条路径（$X, \ldots, Z_1, Y, Z_2, \ldots, H$）。从 G_1 中消去节点 Y 将得到一条路径（$X, \ldots, Z_1, Z_2, \ldots, H$）。由于该路径上不会出现 $\mathbf{U} \setminus \{X\}$ 中的节点，因此消去除 $\{H\} \cup \mathbf{U}$ 之外的所有节点后，将在 G_2 中得到边（X, H）。接下来我们用逆否命题证明必要性方向。假设在 G_1 中不存在一条连接 X 与 H 且不包含 $\mathbf{U} \setminus \{X\}$ 中节点的路径。分两种情形：1） X 与 H 之间不存在路径；2） X 与 H 之间的每条路径都至少包含一个节点 $U \in U \setminus \{X\}$，其形式为（$X, \ldots, U, \ldots, H$）。在第一种情形下，X 与 H 在 G_2 中将不连通。在第二种情形下，从这样的路径中消去除 $\{H\} \cup \mathbf{U}$ 之外的所有节点将得到 $X \to U \to H$，因此 X 不可能与 H 在 G_2 中直接相邻。证明完毕。 ■

引理 22 考虑一个结构因果模型 G 及其根节点的子集 $\mathbf{U}$。假设结构因果模型 G' 由G 通过添加一个根节点 H 作为 G 中某些节点 $\mathbf{Z}$ 的父节点而得到，其中 $\mathbf{Z} \cap \mathbf{U} = \emptyset$。设 π 是 G的一个受 $\mathbf{U}$约束的消元顺序，且 π' 是 G'的一个受 $\mathbf{U}$约束的消元顺序，它由 π 通过将 H 放在变量$\mathbf{U}$之前得到。若 π 的宽度为 w，且 π' 的宽度为 w'，则$w' \leq \max(w+1, |\mathbf{U}|)$。此外，若 H 在 G'中有一个单个子节点，则 $w' = w + 1$。

证明 令$\mathbf{X}$ 表示 G中除$\mathbf{U}$ 以外的变量，并令$\mathbf{U}' = \mathbf{U} \cup \{H\}$。假设我们先按照顺序 π'从 G' 中消除变量$\mathbf{X}$，然后消除 H，最后消除$\mathbf{U}$。这会得到图序列$G'_1, \ldots, G'_j, G'_{j+1/2}, G'_{j+1}, \ldots, G'_{j+k}$，其中 $j = |\mathbf{X}|$ 且 $k = |\mathbf{U}|$。这里，$G'_{j+1/2}$是通过从 G'_1中消除所有变量$\mathbf{X}$ 得到的，而 G'_{j+1} 是通过从 $G'_{j+1/2}$中消除 H 得到的。我们声称：

- 若 $i \leq j$，则对于 G'_i中的每个节点 $X \neq H$，都有 $G'_i(X) \subseteq G_i(X) \cup \{H\}$。
- 若 $i > j$，则对于 G'_i中的每个节点 $X \neq H$，都有 $G'_i(X) \subseteq \mathbf{U}'$。此外，若 H 在 G'中有一个单个子节点，则 $G'_i(X) = G_i(X)$。

我们首先证明上述断言蕴含该引理，然后通过证明该断言得出结论。假设我们要从 G'中消去变量 Y。若 $Y \notin \mathbf{U}'$，则 $i \leq j$，且上述断言蕴含$\mathbf{C}'(Y) \subseteq \mathbf{C}(Y) \cup \{H\}$。若 $Y \in \mathbf{U}'$，则 $i > j$，且上述断言蕴含 $\mathbf{C}'(Y) \subseteq \mathbf{U}'$；并且当 H 有一个单个子节点时，$\mathbf{C}'(Y) = \mathbf{C}(Y)$。这保证了引理中的陈述：$w' \leq \max(w+1, |\mathbf{U}|)$，并且 $w' = w + 1$ 若 H 在 G'中有一个单个子节点。

接下来，我们通过归纳法证明这一论断。令 $\mathbf{Z}$ 表示 G' 中 H 的子节点。在根据 G' 构造道德图 G'_1 时，引入节点 H 会使两类在 G_1 中不存在的边被添加到 G'_1中：对于 $Z \in \mathbf{Z}$，添加 (Z, H)；如果 Y是某个节点 $Z \in \mathbf{Z}$的父节点，即 Y和 H 是某个节点 Z的共同父节点，则添加 (Y, H)。所有这些额外边都与 H 相 incident，这意味着对于 G'_1中的任意节点 X，$G'_1(X) \subseteq G_1(X) \cup \{H\}$。因此，我们的论断对 G'_1成立。接下来，假设我们的论断对 G'_i 成立（归纳假设），并考虑 G'_{i+1}。我们分两种情形。

情形： $i \leq j$。令 $Y = \pi(i)$。考察 G'_{i+1} 中的每个节点 X。如果节点 X 不是 G_i/G'_i 中 Y 的邻居，那么消除 Y 不会影响 X，即 $G'_{i+1}(X) = G'_i(X)$ 且 $G_{i+1}(X) = G_i(X)$。因此，根据归纳假设，论断成立。否则，我们可以如下界定 $G'_{i+1}(X)$：

$$
\begin{aligned}
G'_{i+1}(X) &= G'_i(X) \cup G'_i(Y) \setminus \{Y\} && \text{根据消除的定义} \\
&\subseteq \big(G_i(X) \cup \{H\}\big) \cup \big(G_i(Y) \cup \{H\}\big) \setminus \{Y\} && \text{根据归纳假设} \\
&\subseteq \big(G_i(X) \cup G_i(Y) \setminus \{Y\}\big) \cup \{H\} \\
&\subseteq G_{i+1}(X) \cup \{H\}
\end{aligned}
$$

情形：$i > j$。对于此情形，G'_i 仅包含 U′ 中的节点。显然，对每个属于 G'_i 的节点 X，$G'_i(X) \subseteq$ U′。回想一下，从 $G'_{j+1/2}$ 中消去 H 会得到 G'_{j+1}。根据归纳假设，$G'_{j+1/2}$ 中所有在 G_{j+1} 中不存在的额外边都必须与 H 相 incident。考虑 H 在 G' 中只有一个子节点的特殊情形。我们声称，在此情形下，$G'_{j+1/2}(H)$ 中任意两个节点在 $G'_{j+1/2}$ 中都相邻，也就是说，H 的邻居在 $G'_{j+1/2}$ 中已经构成一个团。因此，从 $G'_{j+1/2}$ 中消去 H 不会在 G'_{j+1} 中添加任何填充边。这保证了 $G'_{j+1} = G_{j+1}$，即对所有 $i >= j+1$，$G'_i(X) = G_i(X)$。最后，我们通过反证法证明这一断言。假设节点 U_1 和 U_2 是 $G'_{j+1/2}$ 中 H 的邻居，但在 $G'_{j+1/2}$ 中不相邻。根据引理 21，在 G'_1 中，必定存在一条连接 U_1 和 H 的路径 P_1，且不包含 U $\setminus \{U_1\}$ 中的节点；还必定存在一条连接 U_2 的路径 P_2，且不包含 U $\setminus \{U_2\}$ 中的节点。由于 H 是根节点，且在 G' 中只有一个子节点 Z，P_1 在 G'_1 中必定具有 $(U_1, \ldots, Z, H)$ 的形式，而 P_2 在 G'_1 中必定具有 $(U_2, \ldots, Z, H)$ 的形式。因此，G'_1 中必定存在一条路径 $(U_1, \ldots, Z, \ldots, U_2)$，且不包含 U′ $\setminus \{U_1, U_2\}$ 中的节点。根据引理 21，从 G'_1 中消去除 U' 以外的所有节点后，U_1 和 U_2 必须在 G'_{j+1} 中相邻。这就导致了矛盾。 ■

附录 J. 初步实验

接下来，我们通过一项初步实验比较三种算法的复杂度：(1) MAP VE（算法 1），用于计算最大后验概率（在结构因果模型上运行）；(2) RMAP VE（算法 3），用于求解单元选择（在目标模型上运行）；以及 (3) 一种基线暴力方法，用于求解单元选择（在双胞胎模型上运行）。我们考察这些算法在随机结构因果模型上的复杂度，这些模型采用（Han 等人，2022）中的方法生成。该方法生成一个随机有向无环图，然后通过添加额外的根节点，确保图中的每个内部节点至少有一个作为根节点的父节点（以模拟结构因果模型的结构）。这类有向无环图往往包含许多根节点，对于复杂度随约束树宽呈指数增长的算法尤其难以处理，例如 MAP VE 和 RMAP VE，正如我们稍后将展示的那样。

给定一个随机的结构因果模型结构，我们随机选取不同比例的根节点作为单元变量 **U**。我们假设目标函数为（Li 和 Pearl，2019年）中方程（2）所给出的函数。该函数包含一个单一结果变量，我们从结构因果模型叶节点中随机选取该变量。此外，如前所述，由于该函数不包含证据变量，构造目标模型时仅需使用一个孪生模型。我们不对最大后验概率 VE 使用的结构因果模型、RMAP VE 使用的目标模型，或暴力方法使用的孪生模型进行剪枝（见（Darwiche，2009年，第 6 章）），因此，干预变量的选择不会影响我们的复杂度分析（方程（2）的目标函数中没有证据变量）。时间

最大后验概率 VE 的时间复杂度为 $O(n \cdot \exp(w))$，其中 n 是结构因果模型节点的数量，w 是结构因果模型的 U 约束消元顺序的宽度。RMAP VE 的时间复杂度为$O(n_1 \cdot \exp(w_1))$，其中 n_1 是目标模型中的节点数，w_1 是目标模型的 U 约束消元顺序的宽度。暴力方法枚举每个赋值 u，并返回使目标 $L(\mathbf{u})$ 最大化的赋值。其时间复杂度为$O(n_2 \cdot \exp(w_2))$，其中 n_2 是用于计算 $L(\mathbf{u})$ 的孪生模型中的节点数，$w_2 = |\mathbf{U}|+$ 是该孪生模型无约束消元顺序的宽度。因此，我们通过报告节点数 n、n_1、n_2 及其对应的宽度 w、w_1、w_2，比较这三种算法的复杂度。这些数据见表 1，其中还报告了结构因果模型根节点的数量（R）以及作为单元变量使用的根节点所占百分比（ur）。

在介绍本实验的结果之前，我们先对所用结构因果模型的类别及其难点作一些说明。接下来，我们将刻画一类问题：其 $\mathbf{U}$-约束树宽不小于单元变量的数量，即 $|\mathbf{U}|$。鉴于我们所用随机结构因果模型的构造方式，它们与这类问题相似。

定义 23 考虑一个连通有向无环图 G 及其根节点的一个子集 $\mathbf{U}$。如果移除节点$\mathbf{U}$ 及其所有关联边后，该有向无环图仍然连通，则称 $\mathbf{U}$ 是 G 的外部节点。

马尔可夫结构因果模型（每个根节点都有一个子节点）满足上述条件。

引理 24 考虑一个连通有向无环图 G、其根节点的一个子集 $\mathbf{U}$ 以及其道德图 G'。令 $\mathbf{S}$为 $\mathbf{U}$ 的子集，使得对于 $\mathbf{S}$中的任意两个节点 U_1 和 U_2，在 G' 中都存在一条连接 U_1 和U_2 的路径，且该路径不包含 $\mathbf{U} \setminus \{U_1, U_2\}$中的任何节点。如果 π 是 G 的一个宽度为 w的$\mathbf{U}$-约束消元顺序，那么有 $w \geq |\mathbf{S}|$。

证明 根据引理21，任意两个节点 U_1 和 U_2 在 $\mathbf{S}$ 在从$\mathbf{U}$ 中消除所有其他节点后都会相邻 G'。因此，在消除所有$\mathbf{U}$ 之外的节点后，$\mathbf{S}$ 中的节点将形成一个团。这会在消元过程中形成一个大小为 $|\mathbf{S}|$ 的簇，因此 $|\mathbf{S}|$ 是任何$\mathbf{U}$-约束消元顺序宽度的下界。 ■

我们的核心洞见由以下推论阐述：对于定义 23所确定的结构因果模型（SCM）类别（以及单元变量），MAP VE 和 RMAP VE 的复杂度必然随单元变量的数量呈指数增长。基线方法的表现可能差得多，因为其复杂度随单元变量数量与孪生模型的无约束树宽之和呈指数增长。

推论 25 考虑一个连通的结构因果模型 G、其根节点的一个子集 $\mathbf{U}$，以及一个宽度为 w 的 $\mathbf{U}$ 约束消元顺序 π。如果 $\mathbf{U}$ 是外部于 G的，那么 $w \geq |\mathbf{U}|$。

证明 考虑 U 中任意两个节点 U_1 和 U_2。假设 X_1 是 U_1 的子节点，且 X_2 是 U_2 的子节点。由于 U 位于 G 的外部，道德图 G 中存在一条路径 $(X_1, \ldots, X_2)$，且不包含 U 中的任何节点。因此，存在一条路径 $(U_1, X_1, \ldots, X_2, U_2)$，且不包含 $\mathbf{U} \setminus \{U_1, U_2\}$ 中的任何节点。根据引理 24，这意味着 $w \geq |\mathbf{U}|$，因为 $\mathbf{S} = \mathbf{U}$。 ■

现在，我们可以着重分析表1中的模式。MAP VE 和 RMAP VE 的复杂度比较接近，后者的开销略高于前者。此外，随着结构因果模型变量的数量（n）和单元变量的数量（ur）增加，两者之间的差距会缩小。

ur			20%				40%				60%			
n	n_2	R	n_1	w	w_1	w_2	n_1	w	w_1	w_2	n_1	w	w_1	w_2
10	14	6	52	5.5	7.2	7.3	49	5.5	7.4	8.3	46	5.5	7.5	9.3
15	21	9	82	7.4	10.0	10.6	76	7.4	10.2	12.6	70	7.5	11.0	14.6
20	30	12	116	10.1	14.0	16.0	110	10.1	14.5	18.0	101	10.1	15.4	21.0
25	37	15	140	11.0	16.8	19.7	131	11.0	17.4	22.7	122	11.2	18.4	25.7
30	43	17	163	11.3	18.8	21.6	154	11.4	19.2	24.6	142	11.8	20.7	28.6
35	50	19	190	12.4	21.4	24.2	178	12.4	21.9	28.2	166	12.8	23.5	32.2
40	57	22	218	13.3	24.3	27.8	206	13.6	24.9	31.8	191	14.3	26.6	36.8
45	64	24	246	14.3	26.6	29.8	231	14.4	27.2	34.8	216	15.6	28.9	39.8

ur			80%				100%			
n	n_2	R	n_1	w	w_1	w_2	n_1	w	w_1	w_2
10	14	6	43	5.5	7.7	10.3	37	6.3	8.3	12.3
15	21	9	64	8.2	11.8	16.6	58	9.6	12.6	18.6
20	30	12	95	10.5	15.6	23.0	86	12.8	16.4	26.0
25	37	15	113	12.8	18.8	28.7	104	15.9	19.9	31.7
30	43	17	133	13.8	21.2	31.6	121	18.0	21.6	35.6
35	50	19	154	15.6	24.0	36.2	142	19.6	23.6	40.2
40	57	22	179	17.6	27.0	40.8	164	23.0	26.6	45.8
45	64	24	201	20.1	30.6	44.8	186	25.6	29.2	49.8

表 1：比较使用 MAP VE 求解 MAP（$n\exp(w)$）、使用 RMAP VE 求解单元选择（$n_1\exp(w_1)$）以及使用暴力方法求解单元选择（$n_2\exp(w_2)$）的复杂度。每个数据点是 25 次运行的平均值。所有消元顺序均使用最小填充启发式计算（Kjærulff，1990年）。

请注意，根据定理17，$w_1/w \leq 2$ 然而，表1 显示，这一比值可能显著小于2。最后，暴力方法的表现明显逊于 RMAP VE，而且随着结构因果模型变量的数量（n）和单元变量的数量（ur）增加，_ 两者之间的差距也会扩大。

最后，我们通过指出一类单元变量数量无界、**U** 约束树宽却有界的问题，来结束本讨论。这类问题如图 **5** 所示。其**U**-约束树宽为 3，使用**U**-约束消元顺序 $S_1, \ldots, S_n, \ldots, U_1, \ldots, U_n$. 即可证明。对于这类问题，即使单元变量数量无界，只要采用合适的目标函数（例如方程（Li 和 Pearl，2019年）中给出的收益函数，见公式（2）），使用 RMAP VE 进行单元选择仍然是可处理的。

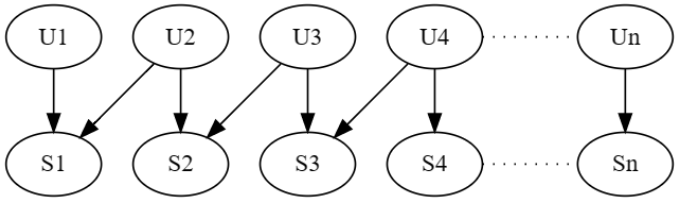


图 5：单元变量为 $\mathbf{U} = U_1, \ldots, U_n$。**U** 约束树宽为 3 .