arXiv:2304.05210v1 [cs.FL] 11 Apr 2023

具有案例间依赖的流程的精确与近似日志对齐

Dominique Sommers、Natalia Sidorova 和 Boudewijn van Dongen

埃因霍温理工大学数学与计算机科学系，荷兰埃因霍温
{d.sommers,n.sidorova,b.f.v.dongen}@tue.nl

摘要。 流程中不同案例的执行通常受到案例间依赖的限制，例如排队或共享资源。人们已提出多种高级佩特里网形式体系，用于建模和分析共同演化的案例。本文聚焦于一种专为通过对齐进行一致性检查而设计的形式体系；这种形式体系带来了相关挑战，包括模型应如何约束相互作用的流程实例和资源实例及其角色。我们提出了对资源受限流程进行建模和分析的要求，并比较了若干允许纳入案例间约束的佩特里网扩展。我们认为，资源受限 ν网是一种适用于一致性检查的形式体系；传统的一致性检查会单独对齐各案例，因而无法揭示案例间依赖上的偏差。我们基于偏序集理论，为全局对齐的事件日志提供了严格的数学基础，并提出一种基于单独对齐案例组合的近似技术，以在局部解决案例间违规问题。

关键词： Petri网 · 一致性检查 · 案例间依赖 · 共享资源。

1 引言

事件日志记录了每项活动在何时执行，此外通常还会说明哪些资源参与了哪些活动，并标明确切的人员或机器。此类事件日志的可用性使得能够对资源约束流程进行一致性检查，不仅分析单个实例的控制流视角，还检查事件日志中记录的实际流程行为是否以及在何处偏离了流程模型所规定的资源约束。

流程模型，尤其是具有精确语义的佩特里网，常用于描述流程的执行并对其进行推理。许多方法中的流程模型会将流程实例（一个案例）与其他案例 [1] 隔离开来加以考量。然而在实践中，流程实例通常会与其他案例和/或资源发生交互，而这些案例和资源的可用性会对流程执行施加额外约束。为了揭示由案例间依赖导致的工作流偏差，至关重要的是使用能够同时考量多个案例的模型。

目前有多种方法用于对具有案例间依赖的流程进行建模和分析。在 [7] 和 [12], 中，Petri网通过引入资源进行了扩展，以建模持久资源的可用性，多个案例通过占用和释放这些共享 资源展开竞争。为区分各个案例， ν-Petri网 [22] 将名称创建与管理作为对经典Petri网的最小扩展纳入其中，其优势在于，与更高级的Petri网扩展不同，可覆盖性和终止性仍然是可判定的。 ν-Petri 网的功能被其他扩展继承，例如目录Petri 网 [11], 同步过程单元模型 [10],资源和实例感知工作流网（RIAW 网） [18],DB 网 [19]以及资源约束的 ν-Petri 网 [24], , 它们都能够同时处理多个案例。对于后者，假设这些案例遵循相同的流程，并通过（抽象的）共享资源以一对多关系进行交互，即一个资源实例在同一时间只能被一个案例占用。更复杂的扩展允许从不同视角处理案例，并通过例如数据库、共享资源和过程单元通道中的概念进行多对多交互。然而，这可能会在一致性检查期间引发不可判定性问题，这正是我们在本文中讨论的内容。

许多一致性检查技术使用对齐 来揭示日志中记录的行为与模型在哪些地方一致、模型规定的哪些活动在日志中缺失，以及根据模型哪些日志活动不应被执行 [8,3]。通常关注的是流程的控制流。在更先进的技术中 [6,15,16,17], , 通过仅在控制流之后考虑数据和/或资源信息 [15],, 通过以可定制的方式平衡不同视角 [16] , 或通过同时考虑所有视角 [17], 将数据和/或资源信息额外纳入对齐。这三类技术均以逐案例的方式运行，在存在共享资源的情况下可能导致误导性的结果，例如多个案例同时请求同一资源时。

在我们之前的工作中，我们通过将完整事件日志与资源约束的 ν-佩特里网 [24]进行对齐，研究了所有流程实例的执行。在本文中，我们进一步开展了以下工作：(1) 我们通过提出此类模型的要求，比较现有佩特里网扩展在支持对具有案例间依赖的流程进行建模和分析方面的能力，并论证 ν-网是一种合适的形式化方法。(2)我们运用偏序集理论，为完整事件日志的对齐以及案例间依赖偏差的揭示提供数学基础；(3) 我们提出了一种用于实际计算最优对齐的近似方法，该方法解决了计算完整事件日志对齐时计算效率受限的问题。该近似方法首先基于为孤立案例组成对齐，然后在日志中局部解决案例间冲突和偏差。

本文组织如下：第 2 节介绍偏序集理论、佩特里网和事件日志的基本概念。第 3 节比较不同的 Petri网扩展。第 4 节给出完整事件日志对齐的数学基础。第 5 节介绍用于计算对齐的近似方法。第 6 节讨论本研究的影响。

2 预备知识

本节介绍与佩特里网和事件日志相关的基本概念，并说明本文将使用的符号。

2.1 多重集与偏序集

我们首先介绍有关多重集和偏序集的定义与记号。

定义 1. （多重集）集合 X 上的多重集 m m： $X \to \mathbb{N}$。$X^{\oplus}$ 表示 X上的所有多重集构成的集合。我们将多重集 m 的支集 $supp(m)$ 定义为集合$\{x \in X \mid m(x) > 0\}$。多重集的元素列作 $[m(x) \cdot x \mid x \in X]$；当上下文能够明确所指的多重集时，将$|x|$记作 $m(x)$。

对于 X上的两个多重集 m_1、m_2，我们记 $m_1 \leq m_2$，若 $\forall_{x \in X} m_1(x) \leq m_2(x)$；若 $m_1 \leq m_2 \wedge m_1 \neq m_2$，则记 $m_1 < m_2$。我们定义 $m_1 + m_2 = [(m_1(x) + m_2(x)) \cdot x \mid x \in X]$，并将 $m_1 - m_2 = [\max(0, m_1(x) - m_2(x)) \cdot x \mid x \in X]$ 记作 $m_1 \geq m_2$。

此外，$m_1 \sqcup m_2 = [\max(m_1(x)\ m_2(x)) \cdot x \mid x \in X]$，$m_1 \sqcap m_2 = [\min(m_1(x)\ m_2(x)) \cdot x \mid x \in X]$。

在某些情况下，我们将集合 X 上的多重集视为长度为 $|X|$的向量，并假定 X 的元素具有任意但固定的顺序。

定义 2。 （偏序、偏序集、反链）一个偏序集 $\bar{}$ $\bar{}$（偏序集） $X = (X, \prec_X)$ 是由集合 X 和偏序 $\prec_X \subseteq X \times X$ 构成的有序对。我们 $\bar{}$ 滥用记号，写作 $x \in X$，若 $x \in X$。对于 x、$y \in X$，若$x \not\prec y \wedge y \not\prec x$，我们写作 $x \|_X y$；若 $x \prec y \vee x = y$，则写作 $x \preceq y$。

给定 $\prec_X$，我们将 $\prec_X^+$定义为包含$\prec_X$ 的最小传递闭关系。因此，$\prec_X^+$是一个满足 $\prec_X \subseteq \prec_X^+$ 的偏序。

我们将并集、交集、差集和子集等标准集合运算扩展到偏序集：对于任意两个偏序集 X 和 Y，$X \circ Y = (\bar{X} \circ \bar{Y}, (\prec_X \circ \prec_Y)^+)$，若 $\circ \in$ $\bar{}$ $\bar{}$ $\bar{}$ $\bar{}$ $\{\cup, \cap, \setminus\}$ 且 $Y \subseteq X$ 当且仅当 $Y \subseteq X$ 且 $\prec_Y = \prec_X \cap (Y \times Y)$。

偏序集 A 是反链反链，如果 A 中没有任何元素可比较，即 $\forall_{x,y \in A}\ x \| y$。对于偏序集 X，$\mathcal{A}(X)$ 表示所有反链反链 $A \subseteq X$的集合，而 $\mathcal{A}^+(X)$ 是所有极大反链的集合：$\mathcal{A}^+(X) = \{A \mid A \in \mathcal{A}(X)\ \forall_{B \in \mathcal{A}(X)}\ B \subseteq A \implies B = A\}$。

两个特殊的极大反链是 X 的最小和最大元素，定义为 $\min(X) = \{x \mid x \in X, \forall_{y \in X} y \not\prec x\} \in \mathcal{A}^+(X)$ 和 $\max(X) = \{x \mid x \in X, \forall_{y \in X} x \not\prec y\} \in \mathcal{A}^+(X)$。

我们定义 $\prec_X \subseteq \prec_Y = \{(a, b \in Y, \prec_Y) \mid a \bar{\|} b\}$ $\bar{}$ 为遵循偏序的 X 的全序排列集合。

定义 3。 （偏序集中的区间、前缀和后缀）给定一个偏序集 X 和两个反链A, $B \in \mathcal{A}(X)$，从 A 到 B 的闭区间 是如下定义的子偏序集：$[A, B] = (AB, \prec_X \cap (AB \times AB))$ 其中 $AB = \{x \mid x \in X, A \preceq x \preceq B\}$，以及半开区间和开区间：$(A, B] = [A, B] \setminus A$, $[A, B) = [A, B] \setminus B$ 和$(A, B) = [A, B) \setminus A$。

人工最小元和最大元分别记为 $\perp$ 和 $\top$，即$\forall_{x \in X} \perp \prec x \prec \top$。$(\perp$、$A]$、$(\perp$、$A)$、$[A, \top)$ $(A$、$\top)$ 表示 X 中反链 $A \in \mathcal{A}(X)$ 对应的前缀和后缀。

2.2 Petri网

Petri网可用作表示、验证和验证工作流过程的工具，从而深入了解流程的行为方式[21]。

定义 4.（标记*Petri*网、前集、后集）一个带标记的 Petri 网 *[20]* 是一个元组$N=(P,T,\mathcal{F},\ell)$，其库所集和变迁集分别为 P和 T，满足$P\cap T=\emptyset$，且弧多重集为 $\mathcal{F}:(P\times T)\cup(T\times P)\to\mathbb{N}$，用于定义网的流。$\ell:T\to\Sigma^{\tau}=\Sigma\cup\{\tau\}$ 是一个标记 函数，为每个变迁 t 分配一个来自字母表 Σ 的标记$\ell(t)$，或者为静默变迁分配 $\ell(u)=\tau$。

我们假设，交集、并集和子集仅针对两个带标记的 *Petri* 网 N_1、N_2 定义，其中 $\forall_{t\in T_1\cap T_2}\ell_1(t)=\ell_2(t)$。

给定元素 $x\in P\cup T$，其前集和后集 ${}^{\bullet}x$ *(*$x^{\bullet}$*)* 定义为多重集，分别为 ${}^{\bullet}x=[\mathcal{F}(y,x)\cdot y\mid y\in P\cup T]$ 和 $x^{\bullet}=[\mathcal{F}(x,y)\cdot y\mid y\in P\cup T]$ 分别。

定义 5。（变迁的标识、使能与发射，可达标识）一个（带标签的）佩特里网 $N=(P,T,\mathcal{F},\ell)$ 的标识 $m\in P^{\oplus}$ 指定每个库所包含的令牌数量，并定义 N 的状态。

给定 m 和 N，当且仅当 $m\geq{}^{\bullet}t$时，变迁 $t\in T$ 才可发射 。我们用 $m\xrightarrow{t}m'$ 表示 t 的发射，其中 m' 是发射 t 后的结果标识，其定义为 $m'=m-{}^{\bullet}t+t^{\bullet}$。对于变迁序列 $\sigma=\langle t_1,\ldots,t_n\rangle$，我们写作 $m\xrightarrow{\sigma}m'$，表示依次发射 t_1 至 t_n。我们称 m' 可从 m 到达，并记作$m\xrightarrow{*}m'$，如果存在某个 $\sigma\in T^*$ 使得 $m\xrightarrow{\sigma}m'$。

$\mathcal{M}(N)=P^{\oplus}$ 并且它表示网 N中的所有标识，以及 $\mathcal{R}(N,m)$可达标识的集合 于网 N 中从标识 m。

定义 6.（位置不变量）设 $N=(P,T,\mathcal{F},\ell)$ 为一个佩特里网。若行向量 $I:\mathbf{P}\to\mathbb{Q}$ 满足 $I\cdot\mathbf{F}=\mathbf{0}$，其中 $\mathbf{P}$ 和 $\mathbf{F}$ 分别是 P和 $\mathcal{F}$ 的向量表示，则称其为位置不变量*[14]*。我们将所有位置不变量的集合记作 $\mathcal{I}_N$，它是 $\mathbb{Q}^P$ 的一个线性子空间。

网中带有初始标识 m_i 的位置不变量 I 的主要性质是 $\forall_{m_1,\ m_2\in\mathcal{R}\ (N,\ m_i)}$ $I\cdot m_1=I\cdot m_2$。

定义 7。（网系统、执行偏序集和序列、语言）一个网系统 是一个元组 $SN=(N,m_i,m_f)$，其中 N是一个（带标签的）佩特里网，而 m_i和 m_f分别是初始标识和最终标识。网系统 $SN=(N,m_i,m_f)$ 中的一个执行序列 是从 m_i到 m_f 的发射序列。此外，一个执行偏序集 是一个变迁发射的偏序集，其中每个全序排列都是一个发射序列。网系统 SN 的语言 是 SN中所有执行序列的集合。

2.3 事件日志

事件日志将活动执行记录为事件，其中至少包括发生的活动、发生时间以及相应案例的案例标识符。通常，资源也会作为事件属性记录，例如执行操作的执行者。通常可以预先知道特定资源角色 R 会参与哪些活动

以及每种角色 $r \in R$在流程中涉及哪些资源实例 Id_r 。我们假设每个资源只有一个角色（职能），能够执行预先定义数量的任务，因此将所有角色的资源实例集合定义为各角色资源实例集合的不相交并集：$\mathrm{Id}_R = \uplus_{r \in R} \mathrm{Id}_r$。具有角色 $r \in R$ 的资源实例$\rho \in \mathrm{Id}_R$ 配备了容量，因此 Id_r 和 Id_R 都是多重集。

定义*8*。（事件、事件日志、轨迹）事件 e 是一个元组（a、t、c、Id'_R），其中包含活动 a = 活动（e） $\in \Sigma$、时间戳 t = $time$（e） $\in \mathbb{R}$、案例标识符c = 案例（e） $\in Id_c$ 以及资源实例的多重集Id'_R = Res（e） $\leq Id_R$。此类事件表示活动 a 发生在时间戳 t，针对案例 c，并由来自 Id'_R 且可能属于不同资源角色的资源实例执行。

事件日志 L 是一组带有偏序的事件 $\prec_L$，其遵循事件的时间顺序，即$\forall_{e_1, e_2 \in L} time(e_1) < time(e_2) \implies e_2 \not\prec_L e_1$。事件日志可以划分为轨迹，这些轨迹定义为投影，例如投影到案例标识符或资源名称上。对于每个 $c \in Id_c$， L_c 表示投影到案例标识符 c上的轨迹，该轨迹由 L_c =$(\{e \mid e \in L$,案例(e) = $c\}, \prec_{L_c})$定义，其中$\prec_{L_c} = \{(e, e') \mid (e, e') \in \prec_L$，案例$(e)$ = 案例(e') = $c\}$。

或者，我们写作 $\langle e_1, e_2, \cdots \rangle$ 表示完全有序的事件日志，而$a^{\mathrm{Id}'_R}$ 以及 $a^{\mathrm{Id}'_R}$ 表示事件，其中案例由活动颜色（以及条形位置）标识，且不考虑发生时间。

对于用于建模流程的（带标签）佩特里网，其变迁名称或标签对应于记录的事件日志中的活动名称。

3 案例间依赖下案例处理系统的建模、分析与仿真

经典Petri网通过变迁发射以及相应的标识变化来建模流程执行，而不区分建模系统中同时运行的不同案例。为创建案例视角，工作流网 [2] 从单个案例的视角对流程进行建模。案例之间会相互作用（例如通过排队或共享资源）的系统，需要采用不同方式建模。从建模角度来看，这归结为多个案例在佩特里网中竞争使用代表资源的共享令牌，因此需要扩展经典Petri网形式体系。在第3.1节中，我们通过示例说明这些要求；随后在第3.2节中，讨论现有Petri网扩展是否满足这些要求。最后，在第3.3节中，我们提出一种基于 ν-Petri网 [22] 的最小扩展，满足资源约束流程仿真与分析的各项要求。

3.1 Requirements imposedby inter-case dependencies

在对具有案例间依赖（即共享资源）的系统进行建模时，同时运行的案例可能会通过资源相互干扰处理过程，从而产生案例间依赖。为了对此类行为进行建模、仿真和分析，案例和资源，

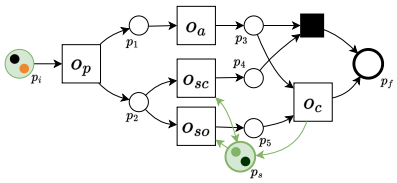

图 1：用于阐述这些要求的示例佩特里网 N_1 ，其中令牌颜色表示不同的实例。

在佩特里网中表示为令牌，应在流程模型中一并处理并同时运行。这就要求实现案例隔离（R1）和资源隔离（R2），并支持持久资源（R3）及案例-资源关联（R4），而普通Petri网无法做到这一点。在分析中，例如计算对齐（见第4节）时，不可逆函数可能导致状态空间爆炸（R5）。对于每项要求，我们都会展示在未满足该要求时，涉及多个同时运行案例的仿真和／或分析会如何失败：

R1 处理多个案例时，需要可区分的案例。参与一次变迁发射的令牌不应属于不同的案例，除非使用案例批处理。在经典Petri网中，可以混合来自不同案例的令牌，这可能导致模型行为在所建模的系统中无法实现：假设我们有一个由佩特里网 N_1建模的简单操作流程，如图1 所示，其中患者经历一次手术，该手术涉及准备 (o_p)、协助 (o_a)、闭合手术 (o_{sc}) 和开放手术(o_{so})等活动，随后进行收尾 (o_c)。我们假设案例令牌是不可区分的。(N_1, $[p_i, 2p_s]$, $[p_f, 2p_s]$) 的语言为 $\{\langle o_p, o_a, o_{sc}\rangle, \langle o_p, o_{sc}, o_a\rangle, \langle o_p, o_a, o_{so}, o_c\rangle, \langle o_p, o_{so}, o_a, o_c\rangle\}$ ，而具有充足资源的同一网处理两个案例时，其语言必须由属于单个案例的两条轨迹的所有可能交错组成。然而， $\{\langle o_p, o_a, o_{sc}, o_p, o_a, o_{so}, o_c\rangle\}$ 被包含在 (N_1, $[2p_i, 2p_s]$, $[2p_f, 2p_s]$) 的语言中，这是不可能通过两条单个案例的轨迹交错得到的，因为在 o_{sc} 发射后，o_c 永远不会被使能。在本例以及后文中，当引用示例中的第二个案例时，我们使用带下划线的符号。从现在起，我们假设案例令牌是可区分的，并且有 $m_i(p_i) = (c, c)$ ；

R2 可区分的资源是资源实例可被唯一识别时的必要条件。如果 p_s 中的令牌不可区分，那么 $\langle \ldots, o_{so}^{\{x\}}, o_{sc}^{\{x\}}, o_c^{\{x\}}\rangle$ 属于 (N_1, $[2p_i, 2p_s]$, $[2p_f, 2p_s]$) 的语言。然而，资源实例 x 只有在第一个案例释放该实例后（通过发射变迁 o_c），才能由第二个案例占用，因此不应将其纳入该语言。从现在起，我们假设资源令牌可区分，并且有 $m_i(p_s) = (x, y)$；

R3 当系统中同时存在数量可变的案例时，资源必须是持久的 。在 N_1 中， p_s 中的资源实例被建模为持久资源，因为这些实例在被占用后总会被释放。然而，如果移除弧 (o_c, p_s)，当观测行为涉及两个以上案例时，就会出现问题，因为变迁 o_{so} 触发两次后便不会再被使能，从而导致死锁；

R4 捕获案例-资源关联在处理多个可区分的案例和资源时是必需的，以便跟踪哪个资源处理哪个案例。否则，$(N_1, [2p_i, 2p_s], [2p_f, 2p_s])$ 会包含例如 $\langle \ldots, o_{so}^{\{x\}}, o_{so}{}^{\{y\}}, o_c{}^{\{x\}}, o_c^{\{y\}} \rangle$，这是不可取的，因为资源 x 和 y 在变迁 o_{so} 发射两次后交换了案例。在这种情况下，案例-资源关联应确保变迁 o_c 只能使用与发射变迁时所占用的同一资源进行发射 o_{so};

R5 为了使问题保持可判定，在对齐观测到的行为和建模行为时，对令牌值的操作（例如守卫、弧注释）应当可逆且可计算。举例来说，患者通过姓名和出生日期 v 进入流程，这些信息会在第一个变迁中通过对（o_p）执行操作 $f(v)$ 转换为标识符 c。当患者缺少活动 o_p 时，如果 f 不可逆，那么在发射 o_p 时应插入哪个值 v 就无法判定。

3.2 现有 Petri 网扩展

已有多种针对多案例和/或多资源流程的佩特里网扩展被提出，这些扩展能够处理（部分）案例间依赖。我们逐一介绍每种扩展，说明它们如何满足（以及违反）第3.1 节中列出的要求。我们提出了一种扩展，该扩展结合了所述扩展的概念，并满足所有要求。

资源约束工作流网(RCWF 网) [12] 是扩展了资源约束的 Petri 网，其中资源是持久单元：它们先被占用，之后再释放（R3）。它们为其正确性定义了结构标准。

定义 9. （资源约束工作流网 *[12]*）设 R 为一组资源角色。网系统 $N = (P_p \uplus P_r, T, \mathcal{F}_p \uplus \mathcal{F}_r, m_i, m_f)$ 是一个资源约束工作流网（RCWF 网），其生产库所集合为 P_p，资源库所集合为 $P_r = \{p_r \mid r \in R\}$，当且仅当

- $\mathcal{F}_p : (P_p \times T) \cup (T \times P_p) \to \mathbb{N}$ 且 $\mathcal{F}_r : (P_r \times T) \cup (T \times P_r) \to \mathbb{N}$;
- $N_p = (P_p, T, \mathcal{F}_p, [m_i(p) \cdot p \mid p \in P_p], [m_f(p) \cdot p \mid p \in P_p])$ 是一个网系统，称为 N 的生产网。

Petri网的语义通过在生产库所（R1）上使用带颜色的令牌进行了扩展；由于资源在所有案例之间共享，资源库所上的令牌不带颜色（¬R2，¬R4）。当且仅当其输入库所上存在足够数量的令牌，并且生产库所上的令牌颜色相同时，变迁才被启用。

ν-佩特里网 [22] 是佩特里网的一种扩展，支持纯名称创建和名称管理，其表达能力严格超越普通Petri网，且本质上对应于 [13]的最小面向对象Petri网。在 ν-Petri网中，可以创建、传递和匹配名称，从而处理身份认证问题 [23], 、关联或实例隔离 [9]。名称管理通过将普通令牌替换为可区分令牌来形式化，从而为Petri网添加颜色。

定义 10. *(ν-佩特里网 [22])* 设 Var 为一个固定的变量集合。一个 ν-佩特里网 是一个元组 ν-$N = \langle P, T, \mathcal{F} \rangle$，包含库所集合 P、变迁集合 T（其中 $P \cap T = \emptyset$）以及流函数 $\mathcal{F}$: $(P \times T) \cup (T \times P) \to Var^{\oplus}$，满足 $\forall_{t \in T}, \Upsilon \cap {}^\bullet t = \emptyset \wedge t^\bullet \backslash \Upsilon \subseteq {}^\bullet t$，其中 ${}^\bullet t = \bigcup_{p \in P} supp(\mathcal{F}(p,t))$ 且 $t^\bullet = \bigcup_{p \in P} supp(\mathcal{F}(t,p))$。$\Upsilon \subset Var$ 表示一个由 $p \in P$ 特殊变量构成的集合，这些变量记作 $\nu, \nu_1, \ldots$，用于实例化新名称。

ν-N 的一个标识是函数 $m: P \to Id^{\oplus}$。$Id(m)$ 表示m中的名称集合，即 $Id(m) = \bigcup_{p \in P} supp(m(p))$。

变迁 t 的模式 μ 是一个单射 $\mu: Var(t) \to Id$，它将每个变量实例化为一个标识符。

对于采用模式 μ 发射变迁 t，我们记作 $m \xrightarrow{t_\mu} m'$。当满足以下条件时，t 在模式 μ 下可发射：对于所有 $p \in P$，都有 $\mu(\mathcal{F}(p,t)) \subseteq m(P)$；并且对于所有 $\nu \in \Upsilon \cap Var(t)$，都有 $\mu(\nu) \notin Id(m) = supp(\cup_{p \in P} \mathcal{F}(p,t))$。采用模式 μ 发射 t 后到达的状态是标识 m'，其定义如下：

$$m'(p) = m(p) - \mu(\mathcal{F}(p,t)) + \mu(\mathcal{F}(t,p)) \; for\ all\ p \in P \tag{1}$$

我们将 T_μ 记为所有可能的变迁发射的集合。

ν-Petri网通过要求参与转换发射的令牌具有匹配的颜色（R1、R2），支持案例和资源的实例隔离。由于令牌具有唯一标识符，因此无法捕获案例与资源之间的关联（¬R4）。

资源与实例感知工作流网 （RIAW 网） [18], 是一类佩特里网，它结合了上述概念：在 ν-Petri网的基础上，为处理资源约束定义了类似的结构标准。不过，资源库所被假定只能容纳黑色令牌，因此无法实现资源隔离，也无法妥善捕获案例-资源关联。

同步过程小网 [10] 是一类用于描述具有多对多交互的流程行为的佩特里网：包括转换的无界动态同步、限制同步规模的基数约束，以及基于历史的令牌标识关联（R1、R2）。这种关联通过基于消息的交互来捕获，即将消息的属性指定为关联属性（R4）。关联约束包括 C_{init}、$C^{\subseteq}_{match}$ 和 $C^{=}_{match}$，分别用于初始化属性、部分匹配属性和完全匹配属性。ν-Petri网是处理多个对象的过程小网的基础，它通过将对象各自的子网分离开来实现这一点。尽管过程小网形式体系足以满足上述所有要求，但它还能扩展到多对多关系，从而取消资源只能由单个案例占用的限制。

对象中心*Petri*网 [4], 与同步过程小网类似，用于描述具有多个视角的流程，以及不同对象类型之间的一对多和多对多关系。这些网是有色Petri网的一种受限变体，其中库所具有类型，令牌是可识别的并指代对象（R1、R2），而变迁可以消耗和产生数量可变的令牌。通过增加组合类型的库所，可以实现关联（R4）。同样，由于存在多对多关系，我们对资源施加的一对多限制被解除。

数据库 *Petri*网 （DB 网） [19] 是 ν-佩特里网的扩展，具有多颜色令牌，允许使用多种类型的对象及其关联（R1、R2、R4）。此外，它们支持由关系型构成的底层读写持久化存储。

具有完备约束的数据库。网中的特殊“视图”库所用于检查底层数据的内容，而变迁则配备了数据库更新操作。这些操作在一般意义上不可逆，会导致不可判定性（¬R5）。

目录 *Petri* 网(CLog 网) [11] 与 DB 网相似，但不包含“写入”操作(R1,R2,R4)。DB 网中来自视图库所的查询已移至变迁守卫，仅依赖持久存储的“只读”模式；然而，由于这些守卫通常不可逆，因此同样面临不可判定性问题(¬R5)。

3.3 具有固定颜色类型的资源受限 ν-佩特里网

我们结合上述扩展中提出的概念性思路，对 RIAW 网进行扩展。RIAW 网继承了 RCWF 网的建模限制和 ν-佩特里网的名称管理，并采用了 DB 网和 CLog 网中的概念。

RCWF 网中的资源库所通过令牌对资源实例的可用性进行建模，但这不足以捕捉案例之间的关联，而资源实例正是依据这种关联被占用和释放的。我们提出一种最小扩展，称为资源受限 ν-佩特里网 (RC ν-网)，其中还为每种资源角色设置忙碌库所 $P_r = \{\bar{p}_r \mid r \in R\}$。令牌从 p_r 移动到 $\bar{p}_r$ 表示资源被占用，从 $\bar{p}_r$ 移动到p_r 则表示资源变为可用状态。此外，还可以对资源是否空闲/被占用进行测试建模。我们对网施加一个结构条件，以保证资源具有持久性，也就是说，资源既不能被创建，也不能被销毁。这还意味着，在具有初始标识和最终标识 m_i和 m_f的相应网系统中，对于任意资源角色 $r \in R$， $m_i(p_r) = m_f(p_r)$ 和 $m_i(\bar{p}_r) = m_f(\bar{p}_r)$。

此外，与DB 网和CLog 网类似，我们将令牌从携带单个数据值扩展为携带多个数据值。DB 网和CLog 网允许使用数量可变的预定义颜色类型，而我们将其限制为两种严格类型化的颜色类型，以区分案例和资源。

定义 *11*。（资源约束的 ν-佩特里网）设 C^{ε} 为扩展了普通令牌的案例 *ID* 集合，即 $\varepsilon \in Id_c$，且 Id_R^{ε} 为扩展了普通令牌的资源 *ID* 集合。资源约束的 ν-佩特里网 $N = (P, T, \mathcal{F}, m_i, m_f)$ 是一个具有 $\mathcal{F}: (P \times T) \cup (T \times P) \to (Var_c^{\varepsilon} \times Var_r^{\varepsilon})^{\oplus}$ 的 *Petri* 网系统，其中 Var_c表示案例变量，Var_r 表示资源变量，允许使用两种有色令牌。$\_P = (P_p \uplus P_r \uplus P_r)$，其中生产库所为 P_p，资源可用性和忙碌库所为 $P_r = \{p_r \mid r \in R\}$ 和 $P_r = \{\bar{p}_r \mid r \in R\}$。对 N 中的每个 $r \in R$施加以下建模限制：

1. ${}^\bullet p_r + {}^\bullet \bar{p}_r = p_r^\bullet + \bar{p}_r^\bullet$, *i.e.*, $\forall_{t \in T}\ \mathcal{F}(p_r, t) + \mathcal{F}(\bar{p}_r, t) = \mathcal{F}(t, p_r) + \mathcal{F}(t, \bar{p}_r)$;
2. $m_i(p_r) = m_f(p_r)$ *and* $m_i(\bar{p}_r) = m_f(\bar{p}_r) = 0$;

N 的标识是一个函数 $m: P \to (C^{\varepsilon} \times R^{\varepsilon})^{\oplus}$，其具有案例*ID* C 和资源R，它是从库所到带颜色令牌的多重集的映射。

变迁 t的一种模式是一个单射 $\mu: (Var_c^{\varepsilon} \times Var_r^{\varepsilon})(t) \to (C^{\varepsilon} \times R^{\varepsilon})$，它将每个变量实例化为一个标识符。

命题 1。 定义 *11* 中定义的资源约束型 ν-佩特里网满足要求 *R1-R5*，即它们能够区分案例和资源实例，且这些对象具有持久性；同时，它们能够捕获案例-资源关联，并将操作限制为可逆操作。

证明。 双色严格类型化令牌在系统中区分案例 (R1) 和资源实例 (R2)。对 RC ν-网施加的建模限制确保对于每个资源角色 $r \in R$，令牌只能在 p_r 和 $\bar{p}_r$ 之间移动，即，我们在 p_r and $\bar{p}$ 上有位置不变量 (1，1)，这意味着对于任何可达标记 m，$m(p_r) + m(\bar{p}_r) = m_i(p_r)$，并且当网到达其最终标识时，所有资源令牌都会返回到 p_r，从而确保资源具有持久性 (R3)。位于 $\bar{p}$ 中的令牌上的两种颜色捕获案例与资源实例之间的关联，表示资源实例在整个交互过程中由哪个案例认领。由于变迁发射的模式是双射函数，N 上的每个操作都是可逆的 (R5)。□

请注意，RC ν-网形式化是 DB 网、CLog-net 和同步过程小网的受限版本，因为这三者都能捕获可由 RC ν-网建模的行为。DB 网和 CLog 网还支持数据库操作，而我们认为这些操作与本文目的无关。同步过程小网支持多对多交互，而我们假设一个资源实例不能同时被多个案例共享。

4 完整事件日志对齐

一致性检查领域的多种最先进技术使用对齐，将流程的记录的执行与该流程的模型联系起来 [5]。对齐展示了如何在流程模型中重放日志或轨迹，从而揭示偏差，解释流程模型为何不符合现实，或现实为何不同于本应发生的情况。

传统上，这一过程是针对单个轨迹计算的，然而，正如我们在先前的工作 [24], 中所展示的，对于第3.3节所述的具有案例间依赖的流程，这无法揭示多案例和多资源层面的偏差。在本节中，我们将在第 4.1 节介绍对齐基础，并在第 4.2 节展示如何扩展这一方法，以计算完整事件日志的对齐。

4.1 对齐基础

对齐的核心是三种移动：日志移动、模型移动和同步移动（参见定义 12），分别表示日志中的某项活动无法在流程模型中模拟、模型要求执行某项日志中未观察到的活动，以及某项活动的观测行为与建模行为一致。

定义 12。 （日志、模型和同步移动）设 L 为一个事件日志，$N = (P, T, \mathcal{F}, \ell, m_i, m_f)$ 为一个带标记的 ν-佩特里网，其中 N中的所有可能发射构成集合 T_μ。我们定义日志移动集合 $\Gamma_l = \{(e, \gg) \mid e \in L\}$、模型移动集合 $\Gamma_m =$

$\{(\gg, t_\mu) \mid t_\mu \in T_\mu\}$ 以及同步移动集合 $\Gamma_s = \{(e, t_\mu) \mid e \in L, t_\mu \in T_\mu$活动$(e) = \ell(t)\}$。作为缩写，我们记作 $\Gamma_{ls} = \Gamma_l \cup \Gamma_s$, $\Gamma_{lm} = \Gamma_l \cup \Gamma_m$,$\Gamma_{ms} = \Gamma_m \cup \Gamma_s$ 和 $\Gamma_{lms} = \Gamma_l \cup \Gamma_m \cup \Gamma_s$。

日志移动和模型移动可以通过网（N、m_i、m_f）与事件日志 L（可能是单条轨迹）上的对齐（参见定义 13），揭示实际行为与模型之间的偏差。该对齐是定义 12 中的移动偏序集，包含事件日志以及从 m_i 到 m_f的执行序列，位于 N 中：

定义 13。（对齐）一个对齐 $\gamma = \text{align}(N, L)$ 事件日志 $L = (L, \prec_L)$ 和带标记的 *Petri* 网 $N = (P, T, \mathcal{F}, \ell, m_i, m_f)$ 是一个偏序集 $\gamma = (\bar{\gamma}, \prec_\gamma)$，其中$\bar{\gamma} \subseteq (\Gamma_l \cup \Gamma_s \cup \Gamma_m^{\oplus})$ 具有以下性质：

1. $\overline{\gamma\restriction_L} = \bar{L}$ *and* $\prec_L \subseteq \prec_{\gamma\restriction_L}$
2. $m_i \xrightarrow{\gamma\restriction_T} m_f$, *i.e.*, $\forall_{\sigma \in (\gamma\restriction_T)^<}, m_i \xrightarrow{\sigma} m_f$

在日志事件 $\gamma\restriction_L$ 和变迁发射 $\gamma\restriction_{T_\mu}$上的对齐投影：

$$\gamma\restriction_L = \left(\{e \mid (e, t_\mu) \in \gamma \cap \Gamma_{ls}\}, \{(e, e') \mid ((e, t_\mu), (e', t'_\mu)) \in \prec_\gamma \cap (\Gamma_{ls} \times \Gamma_{ls})\}\right) \quad (2)$$

$$\gamma\restriction_T = \left(\{t_\mu \mid (e, t_\mu) \in \gamma \cap \Gamma_{ms}\}, \{(t_\mu, t'_\mu) \mid ((e, t_\mu), (e', t'_\mu)) \in \prec_\gamma \cap (\Gamma_{ms} \times \Gamma_{ms})\}\right) \quad (3)$$

请注意，与我们先前在 [24], 的工作相比，对齐的定义略有不同：我们将对齐从分布式运行简化为移动的偏序集。可以从对齐中提取过程的状态（标识）历史，反映其据称在现实中发生的情况。对于一般情形，我们引入对过程模型中与对齐的非日志移动相对应的变迁进行伪触发，以获得一个伪标识；该伪标识可能不可达，也可能包含负数个令牌：

定义 14.（伪标识）佩特里网 $N = (P, T, \mathcal{F})$的一个伪标识 m 是一个多重集 $P \to \mathbb{Z}$, 即库所中所含令牌的数量可以为负数。$\mathcal{M}(N)$ 表示 N中所有伪标识的集合。

定义 15.（偏序集的伪触发）设 $N = (P, T, \mathcal{F}, m_i, m_f)$ 是一个 ***RC*** ν-网，且$\sim\gamma$是 N 上的一个对齐。我们定义函数 $\widetilde{m}$: $\mathcal{P}(\gamma) \to \mathcal{M}(N)$，其幂集为 $\mathcal{P}$，以获得 γ每个子偏序集对应的模型伪标识。对于每个子偏序集 $\gamma' \subseteq \gamma$，对每个 $p \in P$，有：

$$\widetilde{m}(\gamma')(p) = m_i(p) + \sum_{(e, t_\mu) \in \gamma' : t_\mu \neq \epsilon} (\mu(\mathcal{F}(t, p)) - \mu(\mathcal{F}(p, t))) \quad (4)$$

也就是说，伪标识是通过以相应模式发射 γ' 的所有变迁而获得的。请注意，它不一定是可达的。

对齐中的反链表示一个可能的时间点，因此也表示一个过程状态。通过对该反链相应的（开放）前缀进行伪触发，我们可以得到相应的前置（或后置）反链标识：

定义 16。（前置和后置反链标识）设 γ 为一个对齐，$G \in \mathcal{A}(\gamma)$为 γ中的一个反链。前置（后置）反链标识定义为对$(\perp, G)$（$(\perp, G]$)进行伪触发后到达的标识，即 $\widetilde{m}((\perp, G))$（$\widetilde{m}((\perp, G])$）。

4.2 扩展对齐以纳入案例间依赖关系

关于构造对齐的基础性工作见 [5]，其依赖于过程建模佩特里网 $N=(P,T,\mathcal{F},\ell,m_i,m_f)$ 与轨迹佩特里网 $N_\sigma=(P^{(\sigma)},T^{(\sigma)},\mathcal{F}^{(\sigma)},\ell^{(\sigma)},m_i^{(\sigma)},m_f^{(\sigma)})$ 的同步积（后者是事件日志中轨迹的佩特里网表示）。同步积由 N 和 N_σ 的并集构成，并且对于每一对变迁 $(t_m,t_l)\in T\times T^{(\sigma)}$，若 ${}^\bullet t_s={}^\bullet t_m+{}^\bullet t_l$ 和 $t_s^\bullet=t_m^\bullet+t_l^\bullet$ 满足以下条件，则存在一个变迁 t_s：t_m 和 t_l 具有相同的标签以及输入弧上的变量，即 $\ell(t_m)=\ell^{(\sigma)}(t_l)$ 且 $Var(t_m)=Var(t_l)$。随后，使用 A^* 算法，在同步积网中从 $m_i+m_i^{(\sigma)}$ 到 $m_f+m_f^{(\sigma)}$ 进行深度优先搜索来计算对齐；根据定义 12 [5]，变迁在 $T^{(\sigma)}$、T 和 $T(s)$ 中的发射分别对应日志移动、模型移动和同步移动。

对于 $c:\ \Gamma_{lms}\to\mathbb{R}^+$ 成本函数，通常针对每个（$e,\ t_\mu$）定义，具体如下：

$$c((e,t_\mu))=\begin{cases}0 & (e,t_\mu)\in\Gamma_s\\ 1 & (e,t_\mu)\in\Gamma_{lm}\wedge\ell(t)\neq\tau\\ \epsilon & \ell(t)=\tau\end{cases}\tag{5}$$

最优对齐是指满足以下条件的对齐：对于任意对齐 γ'，γ 使得 $\sum_{g\in\gamma}c(g)\leq\sum_{g\in\gamma'}c(g)$ 成立；它优先选择同步移动，而非模型移动和日志移动。从一致性检查和揭示现实偏差的角度来看，最优对齐为观测行为与建模行为之间的关系提供了“最佳”解释。

在第 3.3 节中，我们展示了 RC ν-网是一种佩特里网形式体系，能够对案例间依赖关系建模，且适用于一致性检查。为了揭示案例间偏差，我们将同步乘积网调整为 ν-网，从而扩展对齐问题：一个 RC ν-网和一个日志 ν-网：

定义 17.（日志 ν-佩特里网）给定事件日志 L，日志 ν-佩特里网 $N^{(L)}=(P_{(L)},T^{(L)},\mathcal{F}^{(L)},\ell^{(L)},m_i^{(L)},m_f^{(L)})$ 是按如下方式构造的带标签 ν-网。对于每个 $e\in L$，我们创建一个变迁 $t_e\in T^{(L)}$，其具有 $\ell^{(L)}(t)=$ 活动(e)；对于每个资源实例 $\rho_r\in supp(\mathrm{Res}(e))$，我们创建一个库所 $p\in P^{(L)}$，其具有 ${}^\bullet p=\emptyset$，

$p^\bullet=[|\rho_r|\cdot t]$，$\mathcal{F}_{(L)}(p,t)=[|\rho_r|\cdot(\varepsilon,r)]$ 以及 $m_{(L)\,i}(p)((\varepsilon,\rho))=|\rho|$。此外，对于每一对 $(e_1,e_2)\in\prec_L$，我们创建一个库所 $p\in P_{(L)}$，其具有 ${}^\bullet p=[t_{e_1}]$，$p^\bullet=[t_{e_2}]$ 以及

$$\mathcal{F}^{(L)}(t_{e_1},p)=\mathcal{F}^{(L)}(p,t_{e_2})=\begin{cases}[(c,\varepsilon)] & \mathrm{case}(e_1)=\mathrm{case}(e_2)\\ [(\varepsilon,\varepsilon)] & \textit{otherwise}\end{cases}\tag{6}$$

对于每个 $e^-\in\min(L)$，我们创建一个库所 $p^-\in P^{(L)}$，其中包含 ${}^\bullet p^-=\emptyset$、$p^{-\bullet}=[t_{e^-}]$ 和 $m_i^{(L)}\,(p^-)((case(e^-)\ \varepsilon))=1$。同样，对于每个 $e^+\in\max(L)$，我们创建一个库所 $p^+\in P^{(L)}$，其中包含 ${}^\bullet p^+=[t_{e^+}]$、${p^+}^\bullet=\emptyset$ 和 $m_f^{(L)}(p^+)((case(e^+)\ \varepsilon))=1$。

计算完整事件日志对齐，同样是要在同步积网中寻找一条从初始标识到最终标识的路径，即从 $m_i+m_i^{(L)}$ 到 $m_f+m_f^{(L)}$。对此，我们可以使用此前所述的任何现有方法

。最优对齐同样是成本最低的对齐。从复杂度来看，空事件日志和全零成本函数下的对齐问题，可以归约为有界佩特里网从 m_i 到 m_f的可达性问题，其最坏情况下的复杂度为指数级[20]。向日志中添加事件 ν-佩特里网，并为移动设置非零成本，会使问题的复杂度严格增加。

请注意，尽管 ν-Petri网通常由于生成新的令牌而本质上是无界的，但在对齐的上下文中我们可以保留有界性，因为界限由事件日志决定，并且我们可以通过预处理事件日志来获取这一信息。

对于图 2 中所建模的运行示例，我们在图 1 的小型操作流程基础上进行了扩展：在操作过程中加入一名助理资源、一个由全科医生（GP）参与的接诊子流程（i_s，i_p），以及一个带有 FIFO 等待室的处方子流程（p_w，w_e，w_l，p_r）。在适当情况下，处方只能由参与接诊的全科医生开具。接诊和操作子流程都可以分别通过 N 中的静默变迁 τ_1 和 τ_3 跳过。图 3 展示了该流程记录的事件日志 L，其中涉及两名患者。通过上述方法计算得到的、关于 N 和 L 的最优完整事件日志对齐结果如图 4 所示。

5 通过组合与局部重新对齐进行近似

由于多个案例并行执行，如第 4 节所述，在完整事件日志 L 上计算对齐是一项计算开销很大的任务。同时可以看出，多案例与多资源对齐只有在案例间依赖出现违规时，才会偏离经典的单独对齐，例如，某项资源已达到最大容量时仍被占用。

我们可以通过组合各个案例的单独对齐，来近似完整事件日志 L 与佩特里网 N 的对齐。该方法概览如图 7 所示，可分为两部分，分别在第 5.1 节和第 5.2 节中介绍。

1. L 被分解为各个单独案例（L_c，$L_{\bar{c}}$），并与 N（γ_c，$\gamma_{\bar{c}}$）进行对齐，再利用事件日志的偏序 $\prec_L$（$\tilde{\gamma}$）进行组合。所得结果不一定是对齐，因为案例间偏差可能尚未解决；2. 我们通过对（$\tilde{\gamma}'$）进行排列，并对事件日志的局部部分（$[A_1, B_1]$，$[A_2, B_2]$，$[A_3, B_3]$）重新对齐以解决违规，将组合对齐转换为有效对齐。通过替换重新对齐结果（γ_{AB_1}，γ_{AB_2}，γ_{AB_3}），得到近似对齐（γ^*）。

用于计算完整事件日志对齐的 [24] 原始方法和近似方法的实现可在 gitlab.com/dominiquesommers/mira 获取，其中包括本文使用的示例以及一些额外示例。

5.1 组合各个对齐

对于每个案例 $c \in \mathrm{Id}_c$，我们都有投影到案例标识符c上的轨迹 L_c（参见定义 8）。如第 4 节所述，最优完整事件日志对齐 γ_L 由针对每个 $c \in \mathrm{Id}_c$ 的单独对齐 γ_c 组成，分别基于 N 和 L_c，并按照.....组合在一起

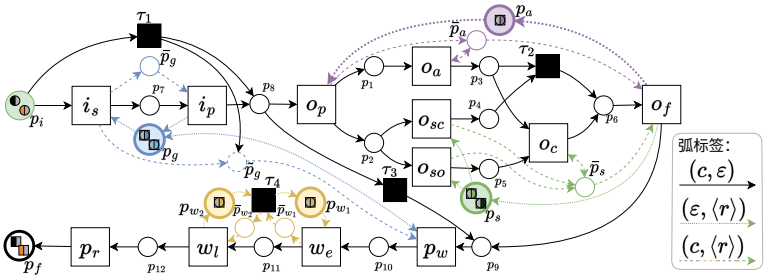

图 2：流程模型 RC ν-网 N，带有初始标识和最终标识，分别以圆形和方形令牌标注。

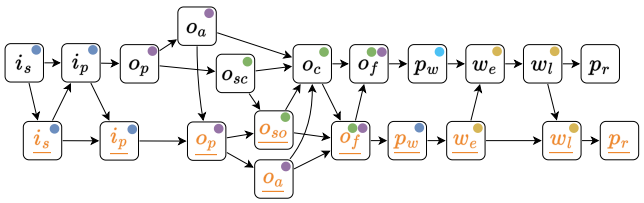

图 3：事件日志 L。

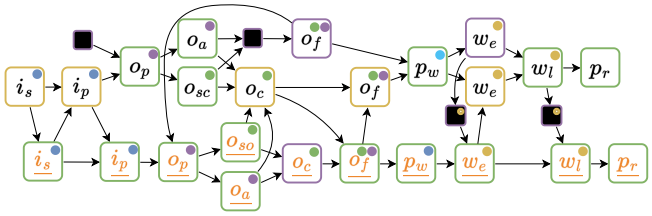

图4：完整事件日志对齐 γ，颜色表示移动类型；绿色、紫色和黄色分别表示同步移动、模型移动和日志移动。

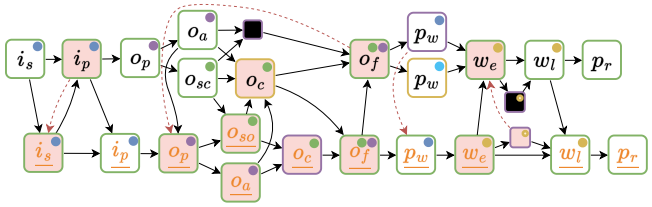

图 5：组合对齐 $\tilde{\gamma}$及带注释的排列和重新对齐区间。

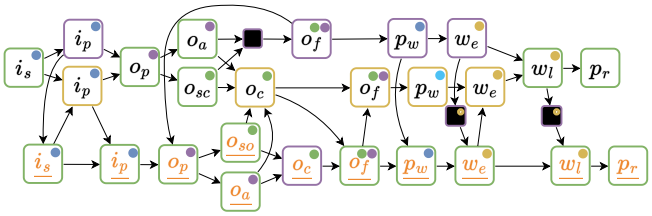

图 6：近似对齐 γ^*。

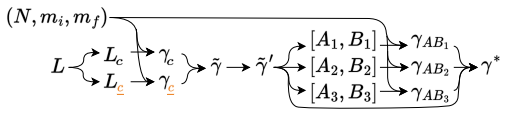

图 7：近似方法概览。

事件日志的偏序 $\prec_L$，其中每个 γ_c 并不一定是关于 L_c的最优解。

计算每个 $c \in \mathrm{Id}_c$ 的最优对齐 $\gamma_c = \mathbf{align}(N,\ L_c)$，再对 γ_L进行近似，计算成本更低。我们根据定义 18，将最优单独对齐与事件日志的偏序组合，得到组合对齐 $\widetilde{\gamma}$。图 5 展示了运行示例的组合对齐，并附有额外的注释（红色），我们稍后会讨论这些注释。

定义 18。（组合对齐）给定一个佩特里网 N 和一个包含轨迹 L_c 的事件日志 L，其中 $c \in \mathrm{Id}_c$，令 $\gamma_c = \mathrm{align}(N, L_c)$ 为对应的最优单独对齐。组合对齐 $\widetilde{\gamma} = \uplus_{c \in \mathrm{Id}_c} \gamma_c$ 是单独对齐的并集与同步移动上的扩展偏序；后者定义为各单独对齐中的偏序并集，以及事件日志偏序 $\prec_L$ 所施加的移动偏序的传递闭包：

$$\prec_{\widetilde{\gamma}} = \left(\bigcup_{c \in \mathrm{Id}_c} \prec_{\gamma_c} \cup \prec_{\gamma_L} \right)^+ \tag{7}$$

与 $\prec_{\gamma_L} = \{((e,\ t_\mu),\ (e',\ t'_\mu)) \mid e \prec_L e',\ (e,\ t_\mu),\ (e',\ t'_\mu) \in (\gamma \cap \Gamma_{ls})\}$。

回顾一下，对于对齐 $\widetilde{\gamma}$ 的每个序列 $\sigma \in \widetilde{\gamma}|_T$，都有 $m_i \xrightarrow{\sigma} m_f$，也就是说，$\sigma$ 是 N 中的一个发射序列。即使不存在案例间偏差，组合对齐也无法保证这一性质。若存在案例间偏差，我们称组合对齐为违反约束的，因为不存在这样的序列。

定义 19。（违反约束的组合对齐）设 $\rho_r \in \mathit{supp}(\mathrm{Id}_R)$ 为一个资源实例，$\widetilde{\gamma} = \uplus_{c \in \mathrm{Id}_c} \gamma_c$ 为一个组合对齐。我们定义

$$\mathcal{S}(\widetilde{\gamma}) = \{(\widetilde{\widetilde{\gamma}}', \prec_{\widetilde{\gamma}'}) \mid \widetilde{\widetilde{\gamma}} = \widetilde{\widetilde{\gamma}}', \prec_{\widetilde{\gamma}} \subseteq \prec_{\widetilde{\gamma}'}, \prec_{\widetilde{\gamma}'} = (\prec_{\widetilde{\gamma}'})^+, \forall_{g \in \widetilde{\gamma}} g \not\prec_{\widetilde{\gamma}'} g\} \tag{8}$$

作为满足偏序 $\prec_{\widetilde{\gamma}}$的传递闭包无环反链排列 $\widetilde{\gamma}$ 的集合。

当且仅当满足以下条件时，$\widetilde{\gamma}$ 与任意资源实例均发生冲突：

$$\forall_{\widetilde{\gamma}' \in \mathcal{S}(\widetilde{\gamma})} \exists_{G \in \mathcal{A}^+(\widetilde{\gamma}')}\ \mathrm{viol}(G) \tag{9}$$

以及违反判据 违反：$\mathcal{A}^+(\widetilde{\gamma}) \to \mathbb{B}$ 针对每个极大反链 $G \in \mathcal{A}^+(\widetilde{\gamma})$ 定义如下：

$$\mathrm{viol}(G) = \exists_{\rho_r \in supp(\mathrm{Id}_R)} \left[\widetilde{m}((\bot, G))(p_r)((\varepsilon, \rho_r)) < \sum_{(e, t_\mu) \in G} \mathcal{F}(p_r, t)(\mu^{-1}((\varepsilon, \rho_r))) \right] \tag{10}$$

也就是说，不存在一种发射对齐中所有变迁的方式，使得在任何时刻都有足够的容量可用。

在图 5 中，满足违反判据的反链是带有一条入射红色弧的单个移动。在定理 1 中，我们证明，对于违反约束的组合对齐 $\widetilde{\gamma}$ 中的每个变迁序列 $\sigma \in \widetilde{\gamma}\restriction_T$，都有 $m_i \not\xrightarrow{\sigma} m_f$，即 $\widetilde{\gamma}$ 不可发射。

定理 1.（违反约束的组合对齐不可发射）设 $\widetilde{\gamma} = \uplus_{c\in \mathrm{Id}_c}\gamma_c$ 是 *RC* ν-网 $N=(P,T,\mathcal{F},m_i,m_f)$ 上的一个组合对齐，事件日志为 L，且 $\widetilde{\gamma}$ 违反约束。那么，不存在 $\widetilde{\gamma}$ 中满足 $m_i \xrightarrow{\sigma} m_f$ 的发射序列 σ。

证明。 $\widetilde{\gamma}$ 违反约束，因此，对于每个 $\widetilde{\gamma}' \in \mathcal{S}(\widetilde{\gamma})$，都存在一个极大反链 $G \in \mathcal{A}^+(\widetilde{\gamma}')$ 和一个资源实例 $\rho_r \in supp(\mathrm{Id}_R)$，使得

$$\widetilde{m}((\bot,G))(p_r)((\varepsilon,\rho_r)) < \sum_{(e,t_\mu)\in G} \mathcal{F}(p_r,t)(\mu^{-1}((\varepsilon,\rho_r))) \tag{11}$$

$$\widetilde{m}((\bot,G))(p_r)((\varepsilon,\rho_r)) - \sum_{(e,t_\mu)\in G} \mathcal{F}(p_r,t)(\mu^{-1}((\varepsilon,\rho_r))) < 0 \tag{12}$$

因此，发射 G 中的变迁会在库所 p_r中为（ε，ρ_r）产生负标记，这是无效的。□

对于反链 $G \subseteq \mathcal{A}(\widetilde{\gamma})$，我们在引理 1 中证明，$\widetilde{m}((\bot,G))$（并且 $\widetilde{m}(\bot,G])$ 可达，当且仅当前缀 $(\bot,G)$ $((\bot,G])$ 不违规。

引理 1.（组合对齐中的前置（和后置）反链标识可达，当且仅当对应的前缀不违反约束。）设 $\widetilde{\gamma} = \uplus_{c\in \mathrm{Id}_c}\gamma_c$ 是 *RC* ν-网 $N=(P,T,\mathcal{F},m_i,m_f)$ 和事件日志 L 上的组合对齐，且 $G \in \mathcal{A}(\widetilde{\gamma})$ 是一个反链。那么，前置（和后置）反链标识 $\widetilde{m}((\bot,G))$（$\widetilde{m}((\bot,G])$）可达，当且仅当 $(\bot,G)$（$(\bot,G]$）不违反约束。

证明。 我们通过证明双条件的两边来证明该引理：

（$\Longrightarrow$）$m_G = \widetilde{m}((\bot,G))$ 可达，因此存在一个序列 $\sigma \in (\bot,G)^*$，满足 $\prec_{(\bot,G)} \subseteq \prec_\sigma$，使得 $m_i \xrightarrow{\sigma} m_G$。令 $\widetilde{\gamma}' \in \mathcal{S}(\widetilde{\gamma})$ 为一个反链排列，且满足 $\prec_\sigma \subseteq \prec_{\widetilde{\gamma}'}$。那么根据可达标记的定义，对于每个极大反链 $G \in \mathcal{A}^+(\widetilde{\gamma}')$ 和每个资源实例 $\rho_r \in supp(\mathrm{Id}_R)$，我们有 $\widetilde{m}((\bot,G))(p_r)((\varepsilon,\rho_r)) \geq \sum_{(e,t_\mu)\in G} \mathcal{F}(p_r,t)(\mu^{-1}((\varepsilon,\rho_r)))$。因此，$(\bot,G)$ 不违反条件。

（$\Longleftarrow$）$(\bot, G]$ 并未违反，因此存在一个 $\widetilde{\gamma}' \in \mathcal{S}((\bot, G])$，使得对于所有 $G' \in \mathcal{A}^+(\widetilde{\gamma}')$和所有 $\rho \in supp(\mathrm{Id}_R)$，我们都有：

$$\widetilde{m}((\bot,G))(p_r)((\varepsilon,\rho_r)) \geq \sum_{(e,t_\mu)\in G} \mathcal{F}(p_r,t)(\mu^{-1}((\varepsilon,\rho_r))) \tag{13}$$

$m_i \xrightarrow{\sigma} \widetilde{m}((\bot, G])$，并且 σ 遵循偏序 $\prec_{\widetilde{\gamma}'}$。 □

5.2 解决组合对齐中的违规

令 $\widetilde{\gamma}' \in \mathcal{S}(\gamma)$ 为 $\widetilde{\gamma}$ 的反链排列。于是，根据定义 19，我们得到一个违规极大反链集合（当 $\widetilde{\gamma}$ 不违规时，该集合为空），其中对应的变迁未启用。我们无需对完整事件日志进行对齐，而是展示如何在此类反链附近局部解决违规问题。对于每个违规反链 G，存在一个区间 $[A, B] \subseteq \widetilde{\gamma}'$，满足 $A \preceq G \preceq B$，使得 $[A, B]$ 可对齐，具体定义见定义 20。

定义 20. *(可对齐区间)* 设 $\gamma = \uplus_{c \in \mathrm{Id}_c} \gamma_c$ 为 *RCν*-网 $N = (P, T, \mathcal{F}, m_i, m_f)$ 上的一个组合对齐，且事件日志为 L，并设 $A, B \in \mathcal{A}(\gamma)$ 为两个反链。当且仅当区间 $[A, B]$ 可对齐时，我们称其为可对齐区间；也就是说，若 m_A 可达，则 $m_B = \widetilde{m}((\bot, B])$ 可从 $m_A = \widetilde{m}((\bot, A))$到达，即 $m_A \xrightarrow{*} m_B$。

注意，$[\min(\gamma'), \max(\gamma')]$ 始终是一个可对齐区间。我们用贯穿全文的示例说明，也可以在 G 附近局部地选取它，例如，$[\{i_s\}, \{i_p\}]$ 采用$G = \{i_s\}$ (参见图 5)。请注意，可以通过将 $[A, B]$ 替换为从 $m_A = \widetilde{m}((\bot, A))$ 到 $m_B = \widetilde{m}((\bot, B])$的子对齐，来解决该违规。

为了证明不依赖于所选重新对齐机制的命题，我们现在假设存在一个函数 $f_{\widetilde{\gamma}}: \mathcal{A}^+(\widetilde{\gamma}) \to \mathcal{P}(\widetilde{\gamma})$，它能为任意 $G \in \mathcal{A}^+(\widetilde{\gamma})$ 生成一个可对齐区间 $[A, B]$。

$$W(\widetilde{\gamma}'_V) = \{[\min(\gamma_v), \max(\gamma_v)] \mid \gamma_v \subseteq \widetilde{\gamma}'_V, \forall_{g \in \gamma_v, g' \in \widetilde{\gamma}'_V \setminus \gamma_v} g \|_{\widetilde{\gamma}'_V} g', \quad \forall_{g \in \gamma_v} \exists_{g' \in \gamma_v} g \not\| _{\gamma_v} g'\} \tag{14}$$

其中 $\widetilde{\gamma}'_V = \bigcup_{G \in \mathcal{A}^+(\widetilde{\gamma}')} f_{\widetilde{\gamma}'}(G)$ 表示覆盖 $\widetilde{\gamma}'$ 中每个违例反链的可对齐区间集合。在图 5 的运行示例中，该集合以红色标注，三个区间$[\{i_s\}, \{i_p\}]$、$[\{o_p\}, \{o_f\}]$和 $[\{w_e\}, \{\tau\}]$ 分别覆盖违例反链$\{i_s\}$、$\{o_p\}$、$\{o_{so}\}$和 $\{w_e\}$。

我们通过将每个区间 $[A, B] \in W(\widetilde{\gamma}'_V)$ 替换为 N 上从 $m_A = \widetilde{m}((\bot, A))$ 到 $m_B = \widetilde{m}((\bot, B])$ 的对齐 γ_{AB}，来解决 $\widetilde{\gamma}'$ 中的违规，并且 $[A, B]\restriction_L$。

由于目前我们假设每个区间 $f(G)$ 都可以对齐，因此存在一个子对齐γ_{AB}。近似对齐 $\gamma^* = (\bar{\gamma}^*, \prec_{\gamma^*})$ 定义如下：

$$\bar{\gamma}^* = \bigcup_{[A,B] \in W(\widetilde{\gamma}'_V)} \bar{\gamma}_{AB} \cup (\bar{\widetilde{\gamma}} \setminus \bar{\widetilde{\gamma}}'_V) \tag{15}$$

$$\prec_{\gamma^*} = \left(\bigcup_{[A,B] \in W(\widetilde{\gamma}'_V)} \prec_{\gamma_{AB}} \cup \{(g_1, g_2) \mid g_1, g_2 \in \gamma \setminus \widetilde{\gamma}'_V, g_1 \prec_{\widetilde{\gamma}'} g_2\} \right)^+ \tag{16}$$

γ^* 运行示例见图 6，其中对图 5 中以红色标注的区间进行了替换式重新对齐。注意，由于局部重新对齐，γ^* 是图 4 中最优对齐 γ 的近似，即 $c(\gamma^*) \geq c(\gamma)$。定理 2 将证明，$\gamma^*$ 是一个有效的对齐。

定理 2. *(γ^* 是一个对齐。)* 令 $\widetilde{\gamma} = \biguplus_{c \in \mathrm{Id}_c} \gamma_c$ 为 *RCν*-网 $N = (P, T, \mathcal{F}, m_i, m_f)$ 上的组合对齐，L 为事件日志，令 $\widetilde{\gamma}' \in \mathcal{S}(\widetilde{\gamma})$ 为 $\widetilde{\gamma}$ 的一个反链排列，其中 $W(\widetilde{\gamma}'_V)$ 是可对齐区间的集合，覆盖 $\widetilde{\gamma}'$ 中的每个违例反链。

$\gamma^* = (\bar{\gamma}^*, \prec_{\gamma^*})$，根据公式 *15* 和*16*，是一个有效的对齐，即它具备定义 *13* 中的性质 *(1)*、*(2)* 和 *(3)*。

证明。我们通过对 $W(\widetilde{\gamma}'_V)$ 的大小进行归纳，证明 γ^* 是一个对齐。对于基础案例 $|W(\widetilde{\gamma}'_V)| = 0$，我们有 $\bar{\gamma}^* = \bar{\widetilde{\gamma}}$ 和 $\prec_{\gamma^*} = \prec_{\widetilde{\gamma}'}$。根据定义，$\widetilde{\gamma}|_L = L$和 $\prec_L \subseteq \prec_{\widetilde{\gamma}|_L}$。此外，由于 $|W(\widetilde{\gamma}'_V)| = 0$，我们知道对于所有 $G \in \mathcal{A}^+(\widetilde{\gamma}')$，都有 $\neg\mathrm{viol}(G)$，这意味着 $m_i \xrightarrow{\widetilde{\gamma}'} m_f$。

假设 γ^* 是 $|W(\widetilde{\gamma}'_V)| = w$ 的一个对齐。我们对 $W'(\widetilde{\gamma}'_V) = W(\widetilde{\gamma}'_V) \cup \{[A, B]\}$ 且 $|W'(\widetilde{\gamma}'_V)| = w+1$ 和 $[A, B] \in \min(W'(\widetilde{\gamma}'_V))$ 证明该命题。对于 A 之前的每个极大反链 $G \in \mathcal{A}^+((\bot, A))$，即 $G \prec A$，我们有 $\neg\mathrm{viol}(G)$，对此我们用反证法证明。假设 $\mathrm{viol}(G)$，那么根据我们关于存在 $f_{\widetilde{\gamma}'}$ 的假设，存在一个可对齐区间 $[A', B'] \subseteq \widetilde{\gamma}'$，满足 $A' \preceq G \preceq B'$；因此，根据 $G \prec A$，我们有 $[A', B'] \prec [A, B]$，这意味着 $[A, B] \notin \min(W'(\widetilde{\gamma}'))$，矛盾。根据引理 1，以及 $f_{\widetilde{\gamma}'}(G)$ 是一个可对齐区间这一假设，$m_i \xrightarrow{*} m_A \xrightarrow{*} m_B$ 和 $[A, B]$ 可以由 γ_{AB} 替代，而不会在 $(\bot, B]$ 中产生违规，从而完成证明。□

5.3 获取最小局部可对齐区间

我们提出了一种方法，用于查找组合对齐 $\widetilde{\gamma}$ 以及区间 $W(\widetilde{\gamma}_V)$ 的反链排列，使得所有违规都可以通过按照第 5.2 节所述重新对齐这些区间来解决。为提高计算效率，我们选择最小化需要重新对齐的区间中的移动数量。

我们将其表述为一个整数线性规划（ILP）问题。整数线性规划问题的目标是调整 $\widetilde{\gamma}$ 的偏序，使得能够围绕违反约束的反链识别可对齐区间，优先选择移动更少的区间。

设 $\widetilde{\gamma}$ 和 Id_R 中存在一个（可能任意的）固定顺序，使得每个元素都有唯一的索引，即对于每个 $1 \le i \le n_{\widetilde{\gamma}}$，$\widetilde{\gamma}(i)$ 和 $(e(i)\ t_\mu(i))$ 均表示 $\widetilde{\gamma}$ 中的第 i^{th} 个移动，其中 $n_{\widetilde{\gamma}} = |\widetilde{\gamma}|$。此外，对于每个 $1 \le j \le n_r$，$\mathrm{Id}_R(j)$ 表示第 j^{th} 个资源实例，其中 $n_r = |supp(\mathrm{Id}_R)|$。

令$\mathbf{R}$为一个 $n_{\widetilde{\gamma}} \times n_{\widetilde{\gamma}}$矩阵，其中$\mathbf{R}$对任意两个索引 $1 \le i,\ j \le n_{\widetilde{\gamma}}$均有定义，使得$\mathbf{R}_{ij}$是一个二进制值，表示（$\widetilde{\gamma}(i)\ \ \widetilde{\gamma}(j)$）$\in \prec_{\widetilde{\gamma}}$。对于每个 $c \in \mathrm{Id}_c$，我们引入与 $\widetilde{\gamma}|_{\gamma_c}$ 中的移动相对应的索引集合 I_c。此外，我们使用 $[1..n] = \{1, \ldots, n\}$ 作为从 1 到 n 的所有索引集合的缩写。

包含所有违规的最小可对齐区间集合，记为 $W(\widetilde{\gamma}'_V)$，其中 $\widetilde{\gamma}'_V$ 定义为

$$\widetilde{\gamma}'_V = \bigcup_{i,j \in [1..n_{\widetilde{\gamma}}] : \mathbf{X}_{ij} - \mathbf{R}_{ij} = 1} [\widetilde{\gamma}(j), \widetilde{\gamma}(i)] \tag{17}$$

其中 $\mathbf{X}$ 表示对齐移动之间新的偏序关系，该关系遵循资源容量，并为以下问题提供解决方案

$$\text{Minimize} \sum_{i,j\in[1..n_{\tilde{\gamma}}]} (1-\mathbf{R}_{ij})\mathbf{R}_{ji}\mathbf{X}_{ij} + \epsilon\cdot(1-\mathbf{R}_{ij})(1-\mathbf{R}_{ji})\mathbf{X}_{ij} \tag{18}$$

受约束于

$$\forall_{i,j\in[1..n_{\tilde{\gamma}}]} \quad \mathbf{X}_{ij} \in \{0,1\} \tag{19}$$

$$\forall_{c\in \mathrm{Id}_c}\forall_{i,j\in I_c} \quad \mathbf{X}_{ij} = \mathbf{R}_{ij} \tag{20}$$

$$\forall_{i,j\in[1..n_{\tilde{\gamma}}]} \quad \mathbf{R}_{ij} + (1-\mathbf{X}_{ij}) - \mathbf{X}_{ji} \leq 1 \tag{21}$$

$$\forall_{i,j,k\in[1..n_{\tilde{\gamma}}]} \quad \mathbf{X}_{ij} + \mathbf{X}_{jk} - \mathbf{X}_{ik} \leq 1 \tag{22}$$

$$\forall_{i\in[1..n_{\tilde{\gamma}}]} \quad (1-\mathbf{X}_{i\bullet})\mathbf{C}^{\downarrow} - \mathbf{X}_{\bullet i}^{T}\mathbf{C}^{\uparrow} \leq \mathbf{k} \tag{23}$$

其中 $\mathbf{C}^{\downarrow}$ 和 $\mathbf{C}^{\uparrow}$ 均为矩阵，分别用于统计每个对齐移动所占用和释放的资源实例数量。它们对每个$i \in [1..n_{\tilde{\gamma}}]$ 和 $k \in [1..n_r]$ 均有定义，且满足 $(e, t_\mu) = \tilde{\gamma}(i)$ 和 $\rho_r = \mathrm{Id}_R(k)$:

$$\mathbf{C}^{\downarrow}_{ik} = \mathcal{F}(p_r,t)((\varepsilon,\mu^{-1}(\rho_r))) \text{ and } \mathbf{C}^{\uparrow}_{ik} = \mathcal{F}(t,p_r)((\varepsilon,\mu^{-1}(\rho_r))) \tag{24}$$

以及容量向量$\mathbf{k}$，长度为 n_r，定义为$\mathbf{k}_k = |\mathrm{Id}_R(k)|$，对于每个 $k \in [1..n_r]$。

$\mathbf{X}$ 提供了 $\tilde{\gamma}$ 中新的移动偏序的解决方案，使所有违规均得到解决，并且移除的偏序关系数量最少。对于该运行示例，解决方案中的附加弧$\mathbf{X}$ 在图 5中以红色显示。

正确性证明详见附录 A，其中我们证明了：(1) 每个约束都有效；(2) 始终存在可行解；(3) 最优解的成本为零，当且仅当组合对齐不存在违规；(4) 在 $W(\tilde{\gamma}'_V)$ 中得到的每个区间都可对齐。

6 结论

我们提出了对具有案例间依赖的流程进行建模和分析的要求，并论证了我们此前提出的名为资源受限 ν-Petri 网的佩特里网扩展满足这些要求。本文延续了 [24], 中的工作；我们在该工作中指出，在处理共同演化的案例时，传统的观测行为与模型行为对齐方法存在不足，因为这些方法只考虑孤立的案例。本文提出的技术能够同时对齐多个案例，从而揭示案例间依赖上的违规。我们开发并实现了一种近似技术，该技术基于单独对齐的组合以及对违规的局部消解，这是推动该技术在实践中应用的一项重要进展。

对于所暴露违规的解释可能存在歧义，例如：活动是已执行但未记录，还是由“错误的”资源实例执行，或者根本未执行？在 [24], 中，我们简要提到了对同步乘积模型进行放宽，将其作为提高偏差可解释性的一种手段。其中一种放宽方式

有助于检测流程执行中模型要求的步骤被跳过，且执行该步骤所需的资源在应执行时不可用的情况。为变迁添加“无资源”模型移动，可以捕捉此类偏差。当对齐中存在此类特殊移动时，可以减少歧义并提供更好的解释，例如，该活动根本没有执行，而非可能执行了但未记录。未来，我们计划扩展并形式化这些放宽条件，并基于真实案例研究评估通过对齐获得的见解。

致谢。 本工作在“通过人工智能认证生产过程质量（CERTIF-AI）”项目ფარგლებში完成，该项目由 NWO 资助（项目编号：17998）。

参考文献

1. Wil M.P. van der Aalst. Petri网在工作流管理中的应用。电路、系统与计算机杂志，8(01):21–66，1998。 2. Wil M.P. van der Aalst. 行动中的数据科学。载于流程挖掘，第3–23页。施普林格，2016。 3. Wil M.P. van der Aalst、Arya Adriansyah 和 Boudewijn F. van Dongen. 在流程模型上重放历史，以进行一致性检查和性能分析。《*Wiley*跨学科评论：数据挖掘与知识发现》，2(2):182–192，2012。 4. Wil M.P. van der Aalst 和 Alessandro Berti. 对象中心Petri网的发现。*Fundamenta Informaticae*，175(1-4):1–40，2020。 5. Arya Adriansyah. 对齐观测行为与建模行为。博士论文，数学与计算机科学，2014。 6. Mahdi Alizadeh、Xixi Lu、Dirk Fahland、Nicola Zannone 和 Wil M.P. van der Aalst. 连接数据视角与流程视角以进行一致性分析。《计算机与安全》，73:172–193，2018。 7. Kamel Barkaoui 和 Laure Petrucci. 具有共享资源的工作流网的结构分析。1998。 8. Josep Carmona、Boudewijn F. van Dongen、Andreas Solti 和 Matthias Weidlich. 一致性检查。施普林格，2018。 9. Gero Decker 和 Mathias Weske. 面向服务架构的实例隔离分析。载于*2008*年*IEEE*国际服务计算会议，第1卷，第249–256页。IEEE，2008。 10. Dirk Fahland. 描述具有多对多交互的流程行为。载于佩特里网与并发理论及应用国际会议，第3–24页。施普林格，2019。 11. Silvio Ghilardi、Alessandro Gianola、Marco Montali 和 Andrey Rivkin. 带参数化数据的佩特里网：建模与验证（扩展版）。*ar Xiv*预印本 *arXiv:2006.06630*，2020。 12. Kees van Hee、Natalia Sidorova 和 Marc Voorhoeve. 资源受限工作流网。*Fundamenta Informaticae*，71(2, 3):243–257，2006。 13. Olaf Kummer. 面向对象Petri网中的不可判定性。载于《佩特里网通讯》。Citeseer，2000。 14. K. Lautenbach. *Petri*网中的活性。波恩内部报告 ISF。GMD自出版，1975。

15. Massimiliano de Leoni 和 Wil M.P. van der Aalst. 对齐事件日志与流程模型以进行多视角一致性检查：一种基于整数线性规划的方法。载于业务流程管理，第113–129页。施普林格，2013。 16. Felix Mannhardt、Massimiliano de Leoni、Hajo A Reijers 和 Wil M.P. van der Aalst. 流程一致性的平衡式多视角检查。计算，98 437, 2016 4 407( ): – 。 17. Azadeh S Mozafari Mehr、Renata M de Carvalho 和 Boudewijn F. van Dongen. 检测业务流程中的隐私、数据和控制流偏差。载于高级信息系统工程国际会议，第82–91页。施普林格，2021。 18. Marco Montali 和 Andrey Rivkin. 使用以数据为中心的动态系统对带名称的Petri网进行模型检查。计算形式化方面，28(4):615–641，2016。 19. Marco Montali 和 Andrey Rivkin. DB网：论彩色Petri网与关系数据库的结合。载于《佩特里网及其他并发模型汇刊》第十二卷，第91–118页。施普林格，2017。 20. Tadao Murata. Petri网：性质、分析与应用。《*IEEE*会刊》，77(4):541–580，1989。 21. James Lyle Peterson. *Petri*网理论与系统建模。Prentice Hall PTR，1981。 22. Fernando Rosa-Velardo 和 David de Frutos-Escrig. 带名称Petri网的判定问题。*arXiv*预印本 *arXiv:1011.3964*，2010。 23. Fernando Rosa-Velardo、David de Frutos-Escrig 和 Olga Marroqu´ín-Alonso. 关于移动同步Petri网的表达能力。《理论计算机科学电子笔记》，180(1):77–94，2007。 24. Dominique Sommers、Natalia Sidorova 和 Boudewijn F. van Dongen. 将事件日志与资源受限的 ν-Petri网对齐。载于佩特里网与并发理论及应用国际会议，第325–345页。施普林格，2022。

A 整数线性规划问题的正确性

我们首先展示每个约束的有效性：

– 约束 20 确保保留单独对齐的原始偏序。请注意，这也确保了对于每个 $g \in \widetilde{\gamma}$ 都有 $g \not\prec g$ ； – 约束 21 强制规定，当移除关系 $g \prec g'$ 时，必须添加其相反关系 $g' \prec g$：

$$\mathbf{R}_{ij} = 1 \wedge \mathbf{X}_{ij} = 0 \implies \mathbf{X}_{ij} = 1 \tag{25}$$

– 约束 22 强制确保传递闭包得到覆盖：

$$\mathbf{X}_{ij} = \mathbf{X}_{jk} = 1 \implies \mathbf{X}_{ik} = 1 \tag{26}$$

结合约束 22 以及约束 20 中对每个 $g \in \widetilde{\gamma}$ 的 $g \not\prec g$，整数线性规划问题的解中不可能存在环；
– 约束 23 确保解不会违反任何资源实例的容量限制，这一点我们将在引理 2 中证明。

引理 2. （约束 *23* 确保不存在违规）对于上述整数线性规划问题的一个复合对齐 $\widetilde{\gamma} = \Cup_{c \in \mathrm{Id}_c} \gamma_c$ *and* $\mathbf{X}$ 解，其遵循偏序 $\mathbf{X}$ 的排列 $\widetilde{\gamma}' = (\widetilde{\bar{\gamma}}', \prec_{\widetilde{\gamma}'})$ $\widetilde{\gamma}$，即 $\widetilde{\bar{\gamma}}' = \widetilde{\bar{\gamma}}$ *and* $\prec_{\widetilde{\gamma}'} = \{(\widetilde{\gamma}'(i)\ \widetilde{\gamma}'(j)) \mid i, j \in [1..n_{\widetilde{\gamma}}], \mathbf{X}_{ij} = 1\}$，不会产生违规（参见定义 *19*）。

证明。 首先，我们将常量 23 改写为

$$\forall_{i \in [1..n_{\widetilde{\gamma}}]} \forall_{k \in [1..n_r]} (1 - \mathbf{X}_{i\bullet}) \mathbf{C}^{\downarrow}_{\bullet k} - \mathbf{X}^T_{i\bullet} \mathbf{C}^{\uparrow}_{\bullet k} \leq \mathbf{k}_k = |\mathrm{Id}_R(k)| \tag{27}$$

对于每个移动索引 $i \in [1..n_{\widetilde{\gamma}}]$，我们可以将 $G_i = \{\widetilde{\gamma}(j) \mid j \in [1..n_{\widetilde{\gamma}}], \mathbf{X}_{ij} = \mathbf{X}_{ji} = 0\}$ 定义为 $\mathbf{X}$ 中包含 $\widetilde{\gamma}(i)$ 的极大反链。$(1 - \mathbf{X}_{ij})$ 和 $\mathbf{X}^T_{ij}$ 与 G_i 的关系如下：

$$(1 - \mathbf{X}_{ij}) = 1 \iff \widetilde{\gamma}(j) \in (\bot, G_i] \text{ and } \mathbf{X}^T_{ij} = 1 \iff \widetilde{\gamma}(j) \in (\bot, G_i) \tag{28}$$

现在，我们可以将式 27 改写为符合式 9 中性质的形式，简写为$C^{\downarrow}_{\rho_r}((e, t_\mu)) = \mathcal{F}(p_r, t)((\varepsilon, \mu^{-1}(\rho_r)))\}$和 $C^{\downarrow}_{\rho_r}((e, t_\mu)) = \mathcal{F}(t, p_r)((\varepsilon, \mu^{-1}(\rho_r)))$。注意，对于每个 $j \in [1..n_{\widetilde{\gamma}}]$ 和$k \in [1..n_r]$，$\mathbf{C}^{\downarrow}_{jk} = C^{\downarrow}\ \mathrm{Id}_R(k)(\widetilde{\gamma}(j))$以及$\mathbf{C}^{\uparrow}_{jk} = C^{\uparrow}\ \mathrm{Id}_R(k)(\widetilde{\gamma}(j))$均成立。对于每个满足 $\rho = \mathrm{Id}_R(k)$的 $i \in [1..n_{\widetilde{\gamma}}]$ 和 $k \in [1..n_r]$，如下所示。

成立:

$$(1-\mathbf{X}_{i\bullet})\mathbf{C}^{\downarrow}_{\bullet k}-\mathbf{X}^{T}_{i\bullet}\mathbf{C}^{\uparrow}_{\bullet k}\leq|\rho| \tag{29}$$

$$\iff \sum_{j\in[1..n_\gamma]}\Big((1-\mathbf{X}_{ij})\mathbf{C}^{\downarrow}_{jk}-\mathbf{X}^{T}_{ij}\mathbf{C}^{\uparrow}_{jk}\Big)\leq|\rho| \tag{30}$$

$$\iff \sum_{g\in(\bot,G_i]}C^{\downarrow}_{\rho}(g)-\sum_{g\in(\bot,G)}C^{\uparrow}_{\rho}(g)\leq|\rho| \quad \text{(By Eq. 28)} \tag{31}$$

$$\iff \sum_{g\in(\bot,G_i)}\big(C^{\downarrow}_{\rho}(g)-C^{\uparrow}_{\rho}(g)\big)+\sum_{g\in G}C^{\downarrow}_{\rho}(g)\leq|\rho| \tag{32}$$

$$\iff |\rho|-\sum_{g\in(\bot,G_i)}\big(C^{\downarrow}_{\rho}(g)-C^{\uparrow}_{\rho}(g)\big)+\sum_{g\in G}C^{\downarrow}_{\rho}(g)\geq 0 \tag{33}$$

$$\iff \widetilde{m}((\bot,G))(p_r)((\varepsilon,\rho_r))+\sum_{g\in G}C^{\downarrow}_{\rho}(g)\geq 0 \quad \text{(By Def. 15)} \tag{34}$$

□

接下来，我们将介绍整数线性规划问题的两个重要性质。引理 3 证明，对于任意组合对齐 $\widetilde{\gamma}$，始终存在一个满足约束条件的解；引理 4 则证明，当且仅当 $\widetilde{\gamma}$ 未违反定义 19 中的规定时，该解的代价为 0。

引理 3. （整数线性规划问题始终存在解。）对于组合对齐 $\widetilde{\gamma}=\uplus_{c\in\mathrm{Id}_c}\gamma_c$ 以及如上所述构建的整数线性规划问题，存在一个解$\mathbf{X}$，使得所有约束条件均成立。

证明。我们通过构造证明，整数线性规划问题始终存在一个解$\mathbf{X}'$，且该解满足所有约束条件。设 Id_c 中存在一个（可能任意的）顺序，使得c_x 表示每个 $x\in[1..n_c]$对应的 x^{th} 案例标识符，其中 $n_c=|\mathrm{Id}_c|$。

$$\mathbf{X}'=\begin{bmatrix}\mathbf{R}_{c_1} & \mathbf{1} & \cdots & \mathbf{1}\\ \mathbf{0} & \mathbf{R}_{c_2} & \cdots & \mathbf{1}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_{c_{n_c}}\end{bmatrix} \tag{35}$$

对于每个 $c\in\mathrm{Id}_c$，取仅包含该案例元素的 $|\gamma_c|\times|\gamma_c|$ 子矩阵。

构造20$\mathbf{X}'$ 显然满足该约束，因为对于每个 $c\in\mathrm{Id}_c$，它都包含$\mathbf{R}_c$；构造21 设 $i\in I_{c_x}$ 和 $j\in I_{c_y}$ 为两个索引。如果 $x=y$，则显然成立，因为$\mathbf{X}'$包含$\mathbf{R}_{c_x}$。否则，根据$\mathbf{X}_{ij}$的构造，$\mathbf{R}_{ij}=1\wedge\mathbf{X}_{ij}=0\implies j<i\implies\mathbf{X}_{ji}=1$;

常量 22 分别令 $i,j,k\in I_{c_x},I_{c_y},I_{c_z}$。当 $x=y=z$时，这显然成立，因为$\mathbf{X}'$ 包含 $\mathbf{R}_{c_x}$，而后者是传递闭包。否则，根据$\mathbf{X}'$ 的构造，我们知道，对于 $x\neq y$，有$\mathbf{X}'_{ij}=1\iff x<y$ 和$\mathbf{X}'_{ij}=0\iff x>y$，我们用它们来证明这两种情况:

(1) $\mathbf{X}'_{ij} = 1 \wedge x \neq y \implies x < y$。$\mathbf{X}$
$'_{jk} = 1 \implies y \leq z \implies x < z \implies \mathbf{X}'_{ik} = 1$; (2) $\mathbf{X}$
$'_{jk} = 1 \wedge y \neq z \implies y < z$。$\mathbf{X}'_{ij} = 1 \implies x \leq y \implies x < z \implies \mathbf{X}$
$'_{ik} = 1$。因此$\mathbf{X}'$具有传递闭包；

构造 23 对于每个 $x \in [1..n_c]$ 以及每个 $i \in I_{c_x}$，令 $J_i = \{j \mid j \in [1..n_{\widetilde{\gamma}}], 1 - \mathbf{X}'_{ij} = 1\}$以及 $J'_i = \{j \mid j \in [1..n_{\widetilde{\gamma}}], \mathbf{X}'^T_{ij} = 1\}$，注意到 $J' \subseteq J$。根据$\mathbf{X}'$的构造，我们知道对于每个 $y \in [1..n_c]$ 以及每个 $j \in I_{c_y}$，$1 - \mathbf{X}'_{ij} = 1 \implies y \leq x$ 以及$\mathbf{X}'^T_{ij} = 1 \implies y \leq x$。因此

$$\forall_{y \in [1..n_c] \setminus \{x\}} I_{c_y} \subseteq J' \vee I_{c_y} \cap J' = \emptyset \tag{36}$$

成立。由于γ_{c_y}对于每个 $y \in [1..n_c]$都是一个对齐，我们对于每个资源索引 $k \in [1..n_r]$都有 $\sum_{j \in I_{c_y}} (\mathbf{C}^{\downarrow}_{jk} - \mathbf{C}^{\uparrow}_{jk}) = 0$，这些共同意味着

$$\begin{aligned} \sum_{j \in J} \mathbf{C}^{\downarrow}_{jk} - \sum_{j \in J'} \mathbf{C}^{\uparrow}_{jk} &= \sum_{y \in [1..n_c]} \left(\sum_{j \in J \cap I_{c_y}} \mathbf{C}^{\downarrow}_{jk} - \sum_{j \in J' \cap I_{c_y}} \mathbf{C}^{\uparrow}_{jk} \right) \\ &= \sum_{j \in J \cap I_{c_x}} \mathbf{C}^{\downarrow}_{jk} - \sum_{j \in J' \cap I_{c_x}} \mathbf{C}^{\uparrow}_{jk} \leq |\mathrm{Id}_R(k)| \end{aligned} \tag{37}$$

因为 γ_{c_x} 是不违反约束的对齐。

□

引理4。（成本=*0* 当且仅当组合对齐不违反约束。）对于组合对齐 $\widetilde{\gamma} = \Cup_{c \in \mathrm{Id}_c} \gamma_c$，上述所构造的整数线性规划问题的解的成本为零，当且仅当根据定义*19*，$\widetilde{\gamma}$ 不违反约束，即公式*9*对 $\widetilde{\gamma}$不成立。

证明。 我们证明双条件的两边:

($\implies$) 假设成本为零，根据目标函数，对于每个 $i,\ j \in [1..n_{\widetilde{\gamma}}]$，$\mathbf{R}_{ij} = 1 \implies \mathbf{X}_{ij} = 1$，因此 $\{(\widetilde{\gamma}(i)\ \widetilde{\gamma}(j)) \mid i,\ j \in [1..n_{\widetilde{\gamma}}],\ \mathbf{R}_{ij} = 1\} = \prec_{\mathbf{R}} \subseteq \prec_{\mathbf{X}} = \{(\widetilde{\gamma}(i)\ \widetilde{\gamma}(j)) \mid i,\ j \in [1..n_{\widetilde{\gamma}}],\ \mathbf{X}_{ij} = 1\}$。由于 $\mathbf{X}$ 遵守约束 23，$\mathbf{R}$也没有违反约束；

($\impliedby$) 假设 $\widetilde{\gamma}$ 不违反，则存在一个 $\widetilde{\gamma}' \in \mathcal{S}(\widetilde{\gamma})$，使得资源容量得到满足。任何$\mathbf{X}$ 使得 $\prec_{\widetilde{\gamma}} \subseteq \prec_{\widetilde{\gamma}'} \subseteq \prec_{\mathbf{X}}$ 的结果为零成本，因为对于每个 $i, j \in [1..n_{\widetilde{\gamma}}]$，$\mathbf{R}_{ij} = 1 \implies \mathbf{X}_{ij} = 1$。此外，成本不可能为负，因为对于每个 $i, j \in [1..n_{\widetilde{\gamma}}]$，$(1 - \mathbf{R}_{ij})\mathbf{X}_{ij} \geq 1$。

□

最后，为满足第 5节中关于违反反链周围的区间是可对齐的假设，我们在引理5 中表明，对于$W(\widetilde{\gamma}'_V)$中的每个区间，情况确实如此。

引理*5*。（每个区间 $[A, B] \in W(\widetilde{\gamma}'_V)$ 都是可对齐的）设 $\widetilde{\gamma} = \Cup_{c \in \mathrm{Id}_c} \gamma_c$ 为组合对齐，且 $W(\widetilde{\gamma}'_V)$ 与$\widetilde{\gamma}'_V = \cup_{i,j \in [1..n_{\widetilde{\gamma}}], X_{ij} - R_{ij} = 1} [\widetilde{\gamma}(j)\ \widetilde{\gamma}(i)]$ 为根据上述整数线性规划问题及其解 X 得到的区间集合。对于每个 $[A, B] \in W(\widetilde{\gamma}'_V)$，$[A, B]$ 都是可对齐的，即（根据定义*20*） $m_A \xrightarrow{*} m_B$ 与$m_A = \widetilde{m}((\perp, A))$以及 $m_B = \widetilde{m}((\perp, B])$。

证明。令 $[A,B] \in W(\widetilde{\gamma}'_V)$为 $W(\widetilde{\gamma}'_V)$ 中的任意区间。我们通过如下构造的子对齐 $\gamma_{AB} = (\bar{\gamma}_{AB},\ \prec_{\gamma_{AB}})$ 来证明 $m_A \xrightarrow{*} m_B$：

$$\bar{\gamma}_{AB} = \bigcup_{(e,t_\mu)\in[A,B]\cap\Gamma_s} \{(\gg, t_\mu), (e, \gg)\} \tag{38}$$

$$\prec_{\gamma_{AB}} = (\{((\gg, t_\mu(i)), (\gg, t_\mu(j))) \mid i,j \in [1..n_\gamma], \gamma(i), \gamma(j) \in [A,B] \setminus \Gamma_l, \mathbf{X}_{ij} = 1\} \tag{39}$$
$$\cup \{((e,t_\mu),(e',t'_\mu)) \mid (e,t_\mu),(e',t'_\mu)) \in \prec_{\gamma'_L}, (e,\gg),(e',\gg) \in \gamma_{AB}\})^+$$

也就是说，所有同步移动都拆分为相应的模型移动和日志移动，并且模型移动的偏序由 $\mathbf{X}$ 定义，日志移动的偏序则由事件日志的偏序 $\prec_L$定义。存在一个 $\sigma \in N(\gamma_{AB})^*$，使得 $m_A \xrightarrow{\sigma} m_B$，这是因为 γ_{AB}具有两个性质：

– 根据构造 23 和引理 2，$\forall_{G\in\mathcal{A}^+(\gamma_{AB})}\ \rho\in supp(\mathrm{Id}_R)$¬违反$(G)$；– 对于每个 $c \in \mathrm{Id}_c$，根据 γ_{AB} 的构造和 ILP 的约束，我们有$\gamma_{AB}{\restriction}_{T_c} = [A,B]{\restriction}_{T_c}$，其中 $T_c = \{t_\mu \mid t_\mu \in T_\mu, \exists_{v_c\in Var_c, v_r\in Var_r}\mu((v_c,v_r)) = (c,*)\}$。因此，$\gamma_{AB}{\restriction}_c = \gamma_{AB} \cap \Gamma_c$ 是一个对齐，其中 $\Gamma_c = \{(e,t_\mu) \mid (e,t_\mu) \in \gamma_{AB},$ 案例$(e) = c \vee t_\mu \in T_c\}$。

□