毕设格式：

相关工作和预备知识可以放到第二章

第一章：介绍

1.研究背景和意义

2.研究内容和目标

3.研究成果贡献

4.章节组织

第二章：相关工作

*本章节将介绍过程挖掘和日志修复技术的现有相关工作。随着科学技术的发展，从静态到变态，除了日志数量出现了爆炸性的增长，日志质量也参差不齐，这给相关数据发现、挖掘和管理等带来了新的难题，由此衍生的一些技术方向：一致性分析等等，可以很好的解决业务流程模型和日志事件之间的关联映射，从而对日志进行修复。*

过程挖掘

*随着大数据时代到来，人类社会拥有数据规模增长很快，每时每刻都有大量数据被产生和存储下来，以数据科学为核心理念的数据挖掘和机器学习技术可以有效的从数据中挖掘有价值的信息，从而从而帮助业务运营、改进产品以及帮助企业做更好的决策，但是仅仅关注数据挖掘和分析是不够的，还需要将数据与过程分析关联起来。过程挖掘旨在从现代信息管理系统中获取流程模型信息，发现、监控和改进现有系统行为模式，在传统的基于模型的过程分析和现代以数据为核心的挖掘技术之间搭建了一座桥梁，正在发挥着越来越重要的作用。过程挖掘技术划分为三个方面：过程发现、合规性检查和流程增强。*

*合规性检测*

*过程发现使用现存信息系统中的日志数据，抽象出事件之行的过程模型，常见方法如阿尔法算法等，介意大致介绍一下。*

*合规性检测关注过程模型和日志事件之间的关联性，过程模型在企业中发挥重要作用，但是因为实际业务过程的复杂性和可变性，实际业务流程和原始过程模型之间存在一定误差，导致过程模型不能很好发挥作用。合规性检测可以判断日志事件和过程模型之间的一致性，从而对流程模型进行修复，除此之外，还可以对比多个流程模型与日志事件的一致性，从而选择最优流程模型。Rozinat【Conformance checking using cost-based fitness analysis】等人首先提出基于模型-日志序列对齐的合规性检测方法，通过将日志事件与模型迁移进行映射，判断两者之间的差异，此外还提出了四种度量维度用来衡量事件日志和过程模型的匹配度：Fittness、Precision、Generalienation和Structurenedss.*

*基于对齐技术，可以在不舍弃原始过程模型的基础上，对模型进行修复，\*\*等人提出的基于重放日志事件，寻找模型关联差异最大的部分对模型进行修复【*Model repair-aligning process models to reality*】。基于成本技术，寻找过程模型和序列日志的最小对齐，并提供过程模型的偏差诊断信息【Aligning Event Logs and Declarative Process Models for Conformance Checking】，以此对原始过程模型进行修复。除此之外，基于对齐的日志修复也是很主要的应用方面，*

*日志修复？？？？补充*

*日志修复：*

*随着大数据时代的到来，信息集成成为了企业信息化管理的核心，通过信息管理系统把企业的设计、采购、生产、制造、财务、营销、经营、管理等各个环节集成起来，共享信息和资源，同时利用现代的技术手段来寻找自己的潜在客户，有效地支撑企业的决策系统，达到降低库存、提高生产效能和质量、快速应变的目的，增强企业的市场竞争力。随着企业各个生产环节的信息数字化，产生了海量的日志数据。与此同时，因为环节的多变性和环境的复杂性，日志中会存在异常事件，即脱离业务模型约束的“噪音”，按照时间异常的种类，可划分为：事件缺失、时间冗余和事件错位。为了提升日志数据的质量，为企业数据挖掘和智能决策提供支持，就需要日志修复技术，将日志序列中的异常事件进行修复，使其符合流程模型的约束，*

*目前日志修复按照修复维度划分为单维度日志修复和多维度日志修复，按照修复方法划分为：对齐法和分支法。*

*单维度日志修复：*

*单维度日志修复，通过删除、添加和更改原有日志事件，使其发生次序符合业务流程约束，[Aligning Event Logs and Declarative Process Models for Conformance Checking]提出了一种基于对齐的日志修复技术，通过自定义代价函数，使用A\*算法,寻找原始日志最小修复方案，[] 会构造两条路径，一条为实际业务场景中的轨迹序列，另一条代表模型中的路径，通过向两条路径中插入修改操作，对两条路径进行对齐操作。在企业环境中，A\* 算法修复长序列是，效率偏下，因为它会搜索和遍历整个修复空间中的所有修复，因此[*Conformance checking in the large: Partitioning and topology*][xiao 33] 提出了分治策略来提升对齐算法效率，通过将原始模型拆解为多个子模型，并利用启发式方法对修复空间进行剪枝，加速算法之行过程。针对日志事件缺失问题，[song wang]提出了一种分支修复算法，通过对比两个活动的后继节点和前驱节点集合，判断两个活动节点间是否存在缺失节点，并进行节点填补，最终获得完整修复方案，同时通过分支索引和可达索引等剪枝技术，保证了算法性能，对比A\*算法，效率提上了至少3个数量级。*

*多维度日志修复：*

*分类*

日志修复

第三章：预备知识和问题定义一

预备知识

*1.预备知识*

*事件日志*

*过程型*

*Petri网*

*流程定义*

*触发序列-我需要借助触发序列来完成模型优先级的声明*

*2.问题声明*

第四章：方法主体

对齐修复算法

*模型活动优先级*

*事件优先级m(e)*

*候选修复事件：一个事件：后面的元素出现在前面 或者。前面的元素出现在后面*

*Alignment Action: 选一个候选世界进行修复*

*搜索空间*

*剪枝*

最小预估代价对齐修复方案（把top放到最后处理）

基于结构的对齐修复算法

回放或者 分解结构

论文格式：  
1介绍

a. 研究背景和意义

b. 内容和目标

c. 贡献

2.相关工作

3.预备知识

4.问题声明

5.算法主体

6.实验

修复方法：

分支：

对齐：

聚类：Trace alignment in process mining: Opportunities for process diagnostics,

对齐方法：

代价对齐：Conformance checking using cost-based fitness analysis

对齐：Replaying history on process models for conformance checking and performance analysis

对齐：Aligning event logs and declarative process models for conformance checking

流程分解技术：

Conformance checking in the large: Partitioning and topology

对齐是NP-hard问题：

Aligning event logs and process models for multi-perspective conformance checking: An approach based on integer linear programming

日志修复：

Efficient recovery of missing events

Improving documentation by repairing event logs

Heuristic recovery of missing events in process logs,

模型修复：

Model repair-aligning process models to reality

Process Mining-Discovery, Conformance and Enhancement of Business Processes. Berlin, Germany: Springer

英文：夏小旭相关

第一节：数据挖掘主要三种分类方向，后两种需要建立模型和序列之间的关联，有了对齐，很多其他的方向工作，得以开展

第二节：引出对齐，对齐是一个NPhard问题，所以A\* 是一个不错的选择，但是存在缺点，和改进的方法

第三节：对齐和日志修复和模型修复都相关，日志修复的技术以及相关文章：王老师