



中山大學  
SUN YAT-SEN UNIVERSITY

# 状态机 workflow 模型及分析技术研究

Research on State Machine Workflow Model and Analysis Technology

姓名：郑强

学号：14212980

专业：软件工程

学院：数据科学与计算机学院



# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望





# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望





# 研究背景与问题

- 课题来源

- “面向网络社会的工作流关键技术研究”（国家自然科学基金）

- 研究背景

- 互联网高速发展，新颖的业务模式层出不穷，例如众包等群体协作模式。
  - 企业业务扩展，企业中的业务流程越来越复杂，例如跨部门跨职能流程协作。

- 问题

- 新的业务模式，复杂的流程具有：**步骤不固定，业务实体数量不固定，无法预估流程结束时间等特点**。现有的工作流建模方式不能建模这类型的模式



无法确定步骤数量和  
与之关联的业务实体数量



# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望





# 研究现状分析—工作流建模方法

类型	代表	优点	缺点
以活动为中心的建模方法 关注点：步骤	流程图，活动图，Petri Net，WF-Net，状态图等	直观 便于理解	不能表达数据流、 不支持模块化、 不能控制流程粒度
以数据为中心的建模方法 关注点：数据	Artifact，Case Handling，ADoc等	数据流动清晰	不支持模块化， 不能控制流程粒度
以对象为中心的建模方法 关注点：协作	Procllet，FlexFlow等	支持模块化 支持流程粒度控制	不能表达数据流 流程协作的定义不全面

以信息为中心的建模方法 关注点：步骤+数据+信息	文献[31]	直观 便于理解	单模块表达能力不足，缺乏层次性 之间协作描述模糊
-----------------------------	--------	------------	-----------------------------

由于上述建模方法存在一定的不足，  
导致它们不能建模具有步骤不固定，流程  
结束时间不可预估等特性的流程





中山大學  
SUN YAT-SEN UNIVERSITY

# 研究目标

## 首要目标：

设计一种 workflow 建模方法能够建模具有步骤不固定，流程结束时间不可预估等特性的新型业务模式和复杂的协作流程。

## 引申目标：

- 分析模型表达能力，证明改模型的通用性
- 检验模型正确性，证明改模型的可靠性
- 模型实现，证明该模型的实际可行性





中山大學  
SUN YAT-SEN UNIVERSITY

# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机 workflow 模型
4. 状态机 workflow 表达能力分析
5. 状态机 workflow 合理性分析
6. 实验结果及分析
7. 总结与展望







# 状态机 workflow 模型—概念定义

## ➤ 定义：

- **业务实体**：业务实体是与关键业务相关的动态的概念性对象。它包含与业务相关的数据模型和生命周期模型，其生命周期模型可以通过状态机表达。
- **状态机 workflow**：以业务实体为粒度，通过状态机的执行语义从事件驱动的角度来描述其过程规则及协作的工作流模型。

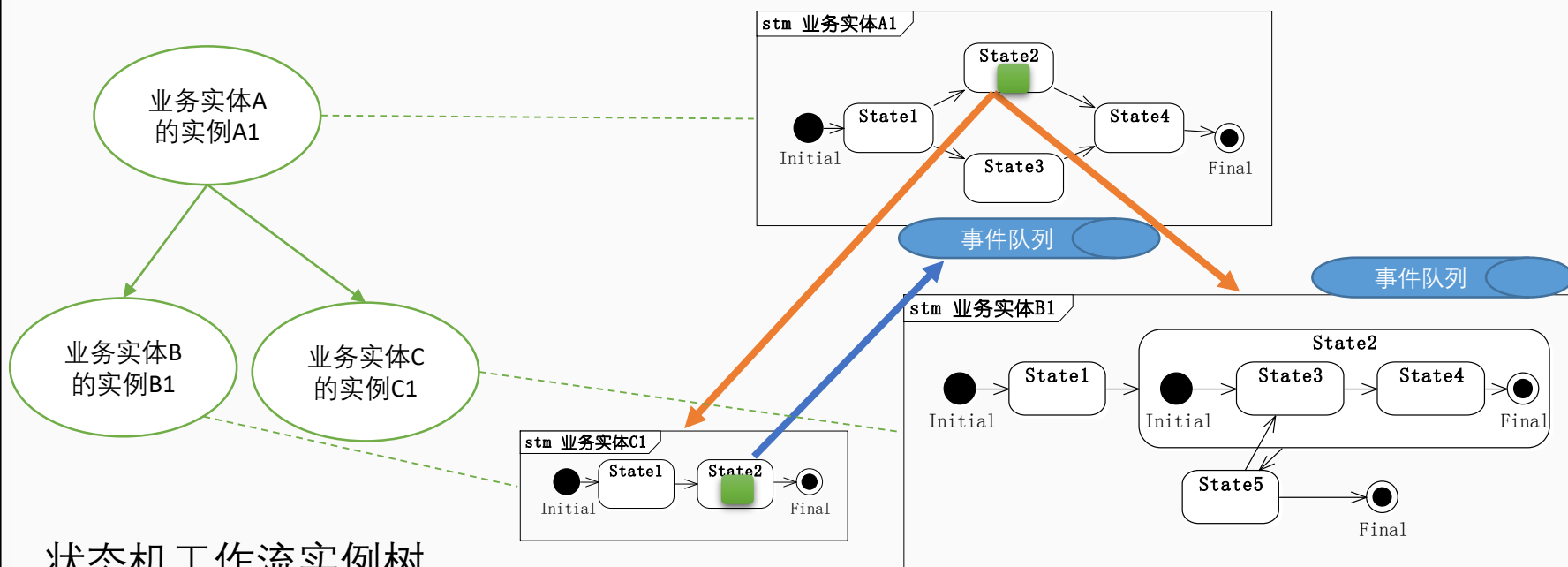




# 状态机 workflow 模型—模型概述

业务实体A, 业务实体B, 业务实体C

粒度



状态机 workflow 实例树

业务目标：业务实体的状态机之间通过协作共同完成

核心：状态机之间的协作

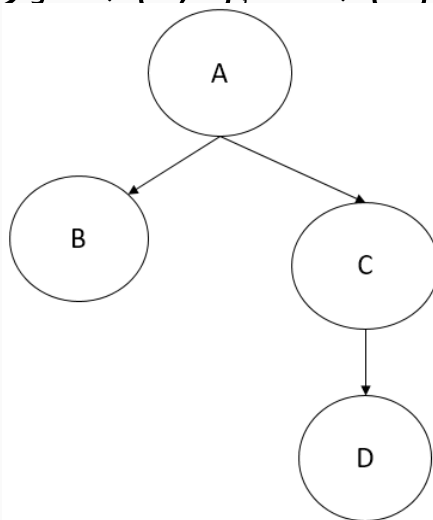
运行时创建  
对应关系  
发送事件





# 状态机 workflow 模型—业务实体之间的关系

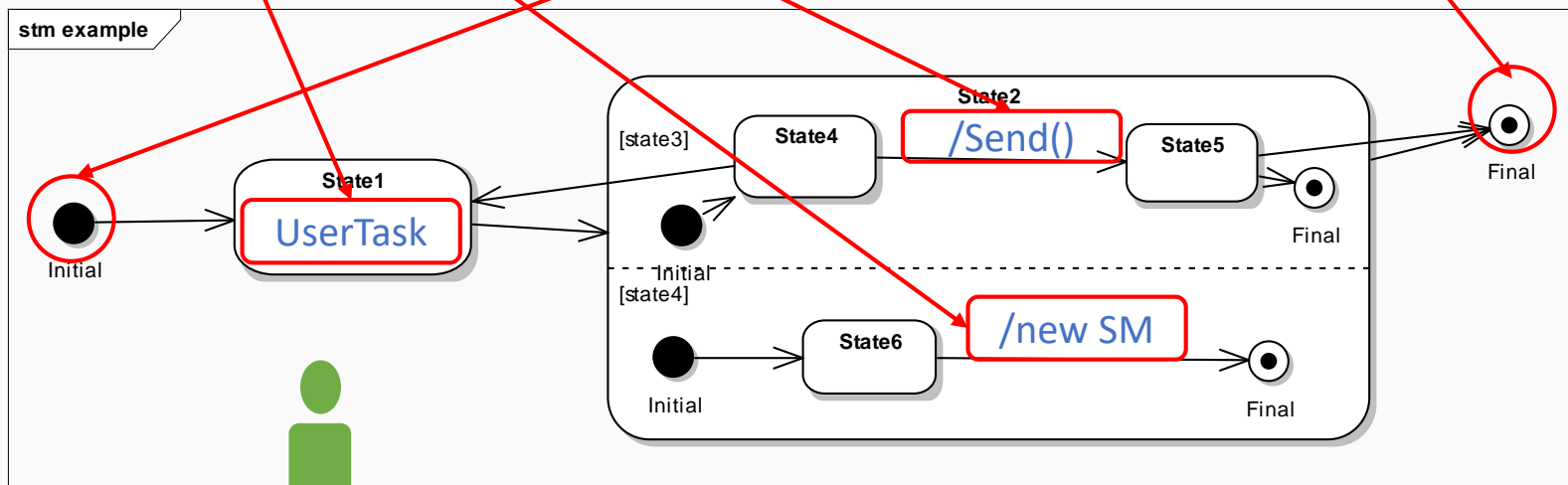
- 监督关系：业务实体A的状态机在其状态变化过程中启动了业务实体B的状态机，并且业务实体B的状态机结束了通知业务实体A的状态机，那么A,B两个业务实体之间是监督关系： $A \rightarrow \{B\}$ 。
- 如下图所示的监督关系为 $A \rightarrow \{B, C\}$ ， $C \rightarrow \{D\}$ 。





# 状态机 workflow 模型—单状态机

- 扩展ECA(Event-Condition-Action) 中的Action以及进入动作和退出动作
  - 任务分派
  - 创建新的状态机 workflow 实例
  - 在状态机 workflow 实例树上传递事件
- 限定一个状态机只有一个全局初始状态和一个全局结束状态。





中山大學  
SUN YAT-SEN UNIVERSITY

# 状态机工作流程模型—协作方式

- 事件目的地的关注点：

- ① 目标状态机实例类型
- ② 目标状态机所处的状态
- ③ 事件内容：{发送者状态机工作流实例ID，事件名，参数列表}
- ④ 事件的传播类型：
  - a) 广播：在整个状态机工作流实例树内传递
  - b) 多播：多播包括**toOffSpring**,**toChild**,**toSibling**,**toAncestor**，或者任意指定的一组树中的实例对象等
  - c) 单播：单播包括**toParent**，或者任意指定的一个树中的实例对象。

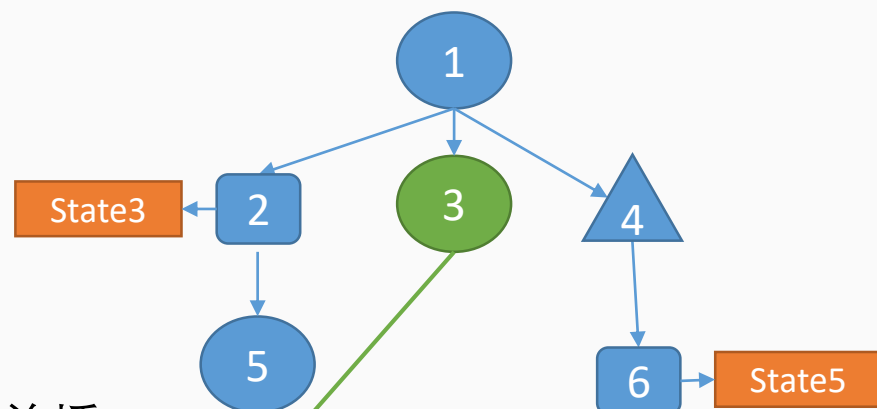


# 状态机工作流程模型—协作方式

- 事件目的地的关注点：

- ① 目标状态机实例类型
- ② 目标状态机所处的状态
- ③ 事件内容：{事件名，参数列表}
- ④ 事件的传播类型：广播，多播，单播。

某一个格局下的状态机工作流程实例树



使用下述发送规则发送事件

{传播类型:广播，目标状态机实例类型:方框，目标状态机当前状态:State5，事件:a\_event}

# 状态机 workflow 模型—形式化定义

- 一个状态机 workflow 模型是一个三元组  $(M, R, P)$ 
  - $M = \{sm_1, sm_2, \dots, sm_n\}$  是一组业务实体的状态机的集合
  - $R : M \rightarrow 2^M$  表示业务实体之间的监督关系,  $\forall e, R(e)$  表示  $e$  监督的业务实体集合, 如果  $e$  没有监督的业务实体则  $R(e) = \emptyset$
  - $P$  表示由  $M$  及  $R$  运行时生成的状态机 workflow 实例集合。
- 一个状态机可以定义为一个9元组,  $SM = \langle S, \rho, \Psi, V, C, E, A, T, R \rangle$ 
  - $S = \{s_1, s_2, \dots, s_n\}$  是所有状态的有序集合,  $n > 0$ , 其中,  $s_i < s_j (i < j)$  表示在流程定义文件中  $s_i$  位于  $s_j$  之前。由此可知,  $s_1$  为全局初始状态,  $s_n$  为全局结束状态
  - $T = \{t_1, t_2, \dots, t_n\}$  是所有转移的有序集合,  $n > 0$ ,  $t = e[c]/a$ , 其中  $e \in E, c \in C, a \in A$ 。其中  $t_i < t_j (i < j)$  表示在流程定义文件中  $t_i$  位于  $t_j$  之前。

状态机语义为 SCXML 语义, 由 W3C 组织于 2015 年正式发布  
<https://www.w3.org/TR/scxml/>





# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
- 4. 状态机工作流表达能力分析**
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望







中山大學

SUN YAT-SEN UNIVERSITY

# 表达能力分析—工作流模式

工作流模式是对一定上下文中连续发生的活动的抽象。

著名学者Alast根据过程原模型总结出了7类共43种工作流模式

1. 基础控制流模式
2. 高级分支与同步模式
3. 多实例模式
4. 基于状态的模式
5. 迭代模式
6. 终止模式
7. 触发器模式



# 表达能力分析—模型对比

建模方法/工作流产品	支持的工作流模式数量
WebSphere MQ	11
EPC	12
Staffware	14
iPlanet	15
FileNet	18
SAP Workflow	20
WebSphere BPEL	20
COSA	21
BPEL	21
Oracle BPEL	22
FLOWer	24
UML Ads	30
BPMN	33
XPDL	33
YAWL	31
State Machine Workflow(Ours)	36+5

数据来源：

文献[33]

[www.workflowpatterns.com](http://www.workflowpatterns.com)

Modern Business Process Automation[图书]

全面的表达能力分析：

<https://sysuworkflower.github.io/BOOWorkflow/>

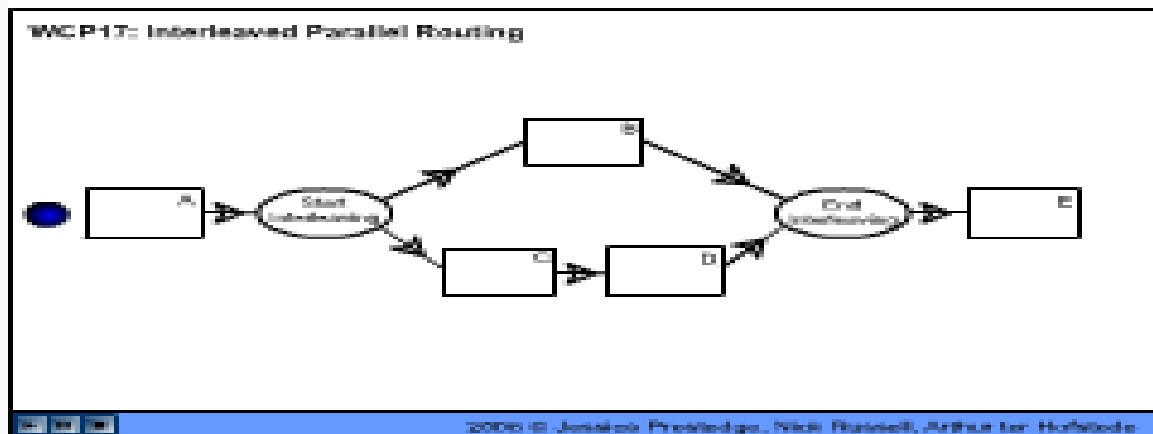
表达能力最强



# 表达能力分析—基于状态的模式

- **Interleaved Parallel Routing**模式的传统工作流示意

含义：所有的并发分支中任意时刻只有一个任务执行，任务执行的顺序可以是任意的。



文献[53]指出  
状态机不能表达此模式

本文通过增加扩展  
表达了此模式





# 表达能力分析—基于状态的模式

## • Interleaved Parallel Routing模式的状态机工作流程表达方式

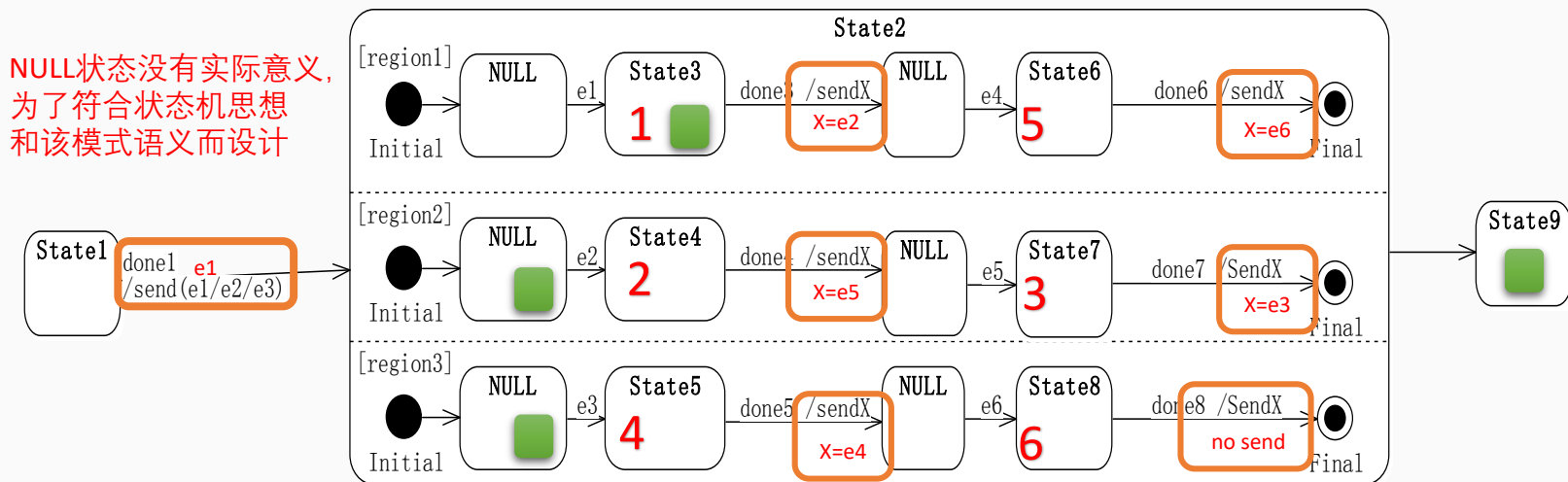
事件集合{ e1 e2 e3 e4 e5 e6 }

事件出集合规则是：随  
机选择一个事件出集合。

事件加入集合规则是：  
当前格局下可以发送的  
事件加入集合。

stm Interleaved Parallel Routing

NULL状态没有实际意义，  
为了符合状态机思想  
和该模式语义而设计





中山大學

SUN YAT-SEN UNIVERSITY

# 表达能力分析—不支持的模式

## 不支持的模式

1. Multi-Merge: Multi-Merge
  2. Structured Discriminator
  3. Blocking Discriminator
  4. Structured Partial Join
  5. Blocking Partial Join
6. General Synchronizing Merge
  7. Implicit Termination

## 原因

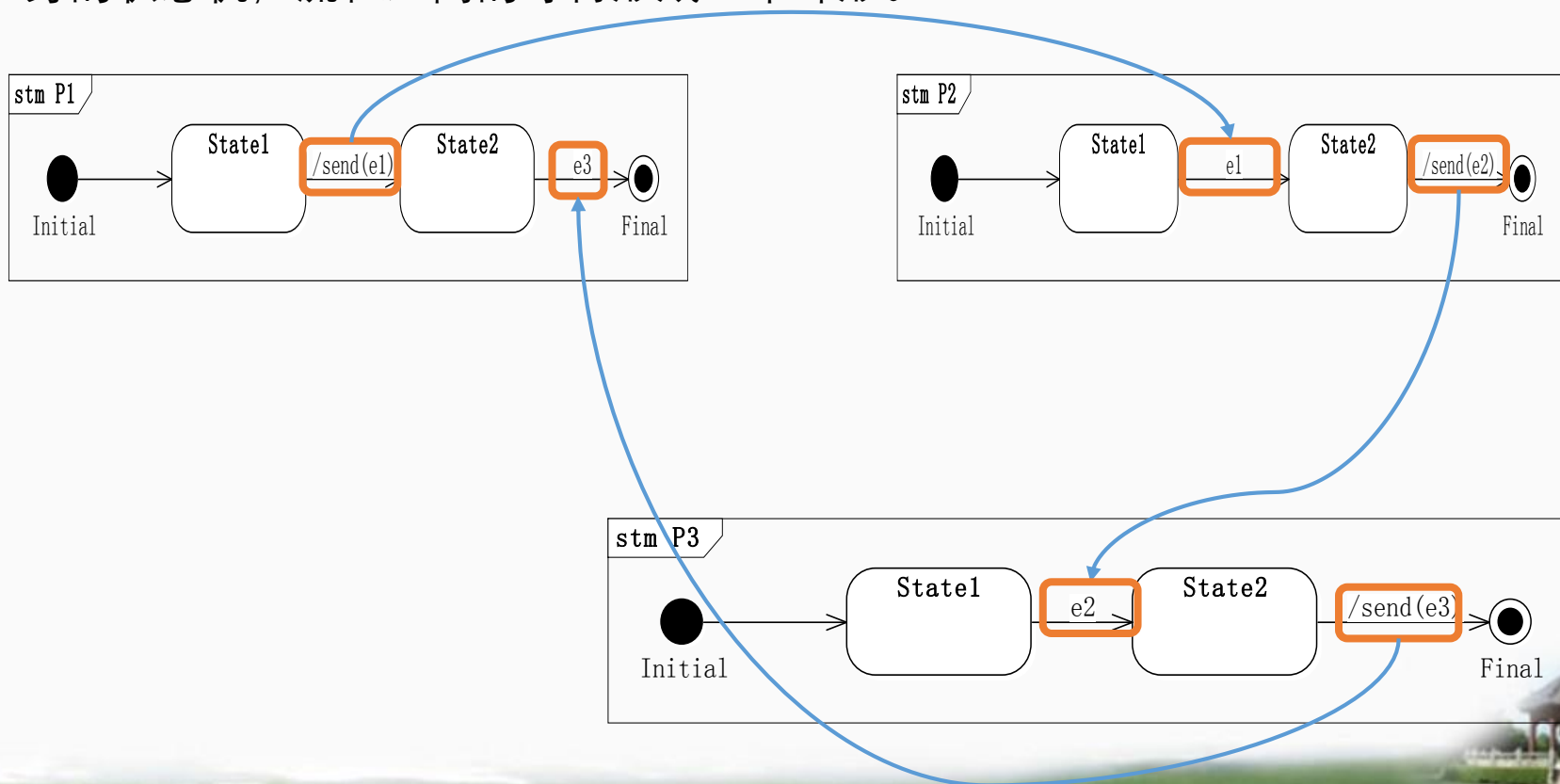
1. 状态机的语义限制
2. 状态机 workflow 模型合理性要求。



# 表达能力分析—特有的模式

## • 循环同步

含义：多个流程之间互相等待另一个流程传递的事件，从而激活自身的状态机，流程之间的等待形成一个环形。





# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望



# 合理性分析

workflows are correct or not are generally judged by analyzing their rationality.

传统工作流合理性的定义，工作流过程是合理的如果它满足如下条件：

1. 根据过程处理的案例总是可能完成的，这一条件保证了不存在死锁和活锁；
2. 工作流过程发出案例完成的信号时不可能仍有任务在执行；
3. 对于每一个任务，在工作流执行过程中它都有执行的机会，这一约束意味着在工作流过程中每一个任务都承担有意义的作用。







# 合理性分析—合理性定义

等价地，状态机工作流程是合理的当且仅当：↵

1. 任意一个状态机工作流实例，对于从初始状态可达的状态  $s$ ，都存在从  $s$  到结束状态的迁移序列。形式地：↵

$$\forall s \in S: (i \xrightarrow{*} s) \Rightarrow (s \xrightarrow{*} o) \quad \leftarrow$$

2. 状态机实例结束时不可能仍有任务执行。形式地：↵

$$\forall t \in T, \exists s, s' \in S: i \xrightarrow{*} s \xrightarrow{t_i} s' \quad \leftarrow$$

3. 考虑到状态机实例运行过程中的通信问题，所有的状态机实例最终都能达到结束状态。↵

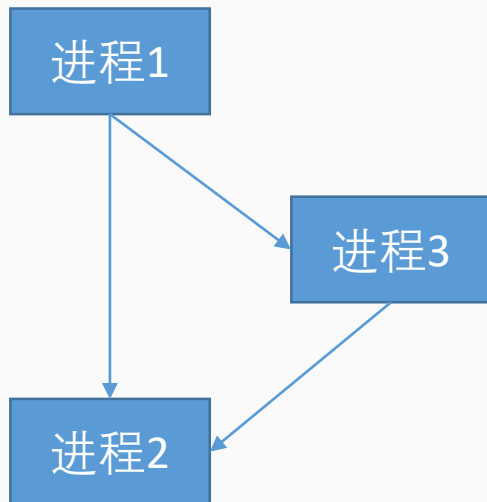
$$\forall p \in P, \exists s, s' \in S_p, e \in EP: i_p \xrightarrow{*} s \xrightarrow{e_i} s' \xrightarrow{*} o_p \quad \leftarrow$$

核心在于：流程中启动起来的所有状态机工作流实例，最终都能处于结束状态

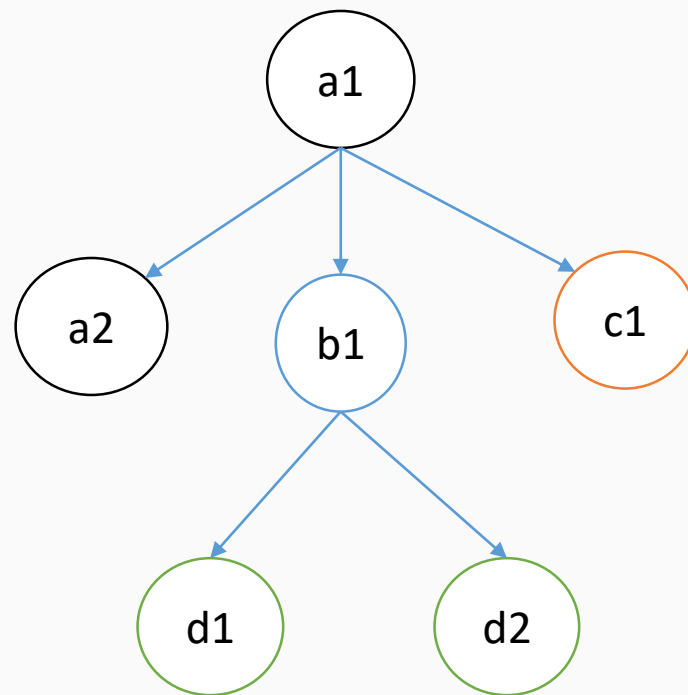


# 合理性分析—模型特点

## 进程间通信



## 状态机工作流通信



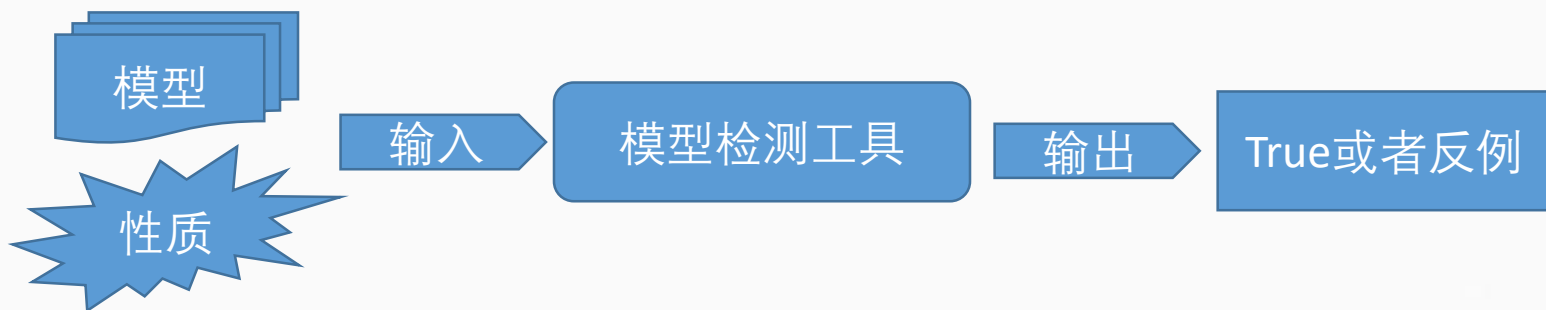


中山大學

SUN YAT-SEN UNIVERSITY

# 合理性分析—模型检验步骤

1. 将状态机 workflow 定义的XML代码转化成具有同等语义的Promela程序。
2. 使用线性时态逻辑(LTL)公式表示合理性定义
3. 使用Spin来验证该Promela程序是否满足上述LTL公式





中山大學

SUN YAT-SEN UNIVERSITY

# 合理性分析一难点

1. 如何表示与状态机同等语义的Promela结构

参考历史文献找到  
相关表示方法





中山大学  
SUN YAT-SEN UNIVERSITY

# 合理性分析— 同等语义 Promela程序

文献[55]：郭伟. UML 状态图模型检查技术及工具实现[D]. 硕士论文. 湖南大学, 2008.

```
所有变量定义;
active proctype ourtype(){
do
  ::(迁移1发生的条件)->
  atomic{
    退出迁移1的源状态，状态变化，设置迁移1发生，退出动作，变量赋值
  }
  ....
  ::(迁移n发生的条件)->
  atomic{
    退出迁移n的源状态，状态变化，设置迁移n发生，退出动作，变量赋值
  }
::else->{
  if
    ::(状态发生了变化)->
    临时变量归零，设置并发状态能否退出
  do
    ::(迁移1是否发生)->
    atomic{
      激活迁移1的目的状态，进入动作
    }
    ....
    ::(迁移n是否发生)->
    atomic{
      激活迁移n的目的状态，进入动作
    }
  od;
  临时变量归零，复位迁移x是否发生
::else->atomic{
  从当前状态机事件队列中取出一个事件
}
fi;
}
od;
DeadLock: printf("DeadLock");//没有事件可以派遣，程序不可继续运行
TheEnd: skip;当此到达结束状态的时候指向此语句
}
```



中山大學

SUN YAT-SEN UNIVERSITY

# 合理性分析—难点

1. 如何表示与状态机同等语义的Promela结构

参考历史文献找到  
相关表示方法

2. 如何使用Promela描述状态机 workflow 中的树形结构

3. 如何使用树形结构对协作模式建模（很重要但只关乎编程问题）

4. 如何将状态机 workflow 定义转化到Promela结构

本文的贡献





中山大学

SUN YAT-SEN UNIVERSITY

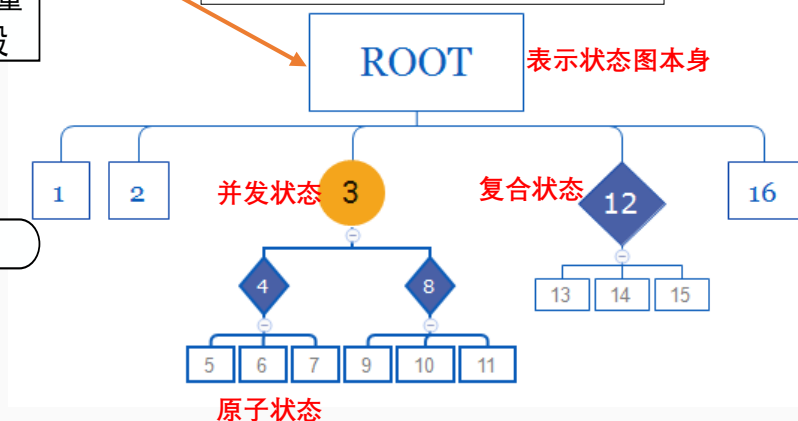
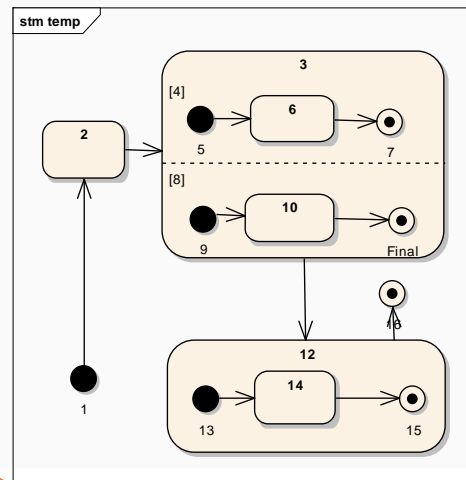
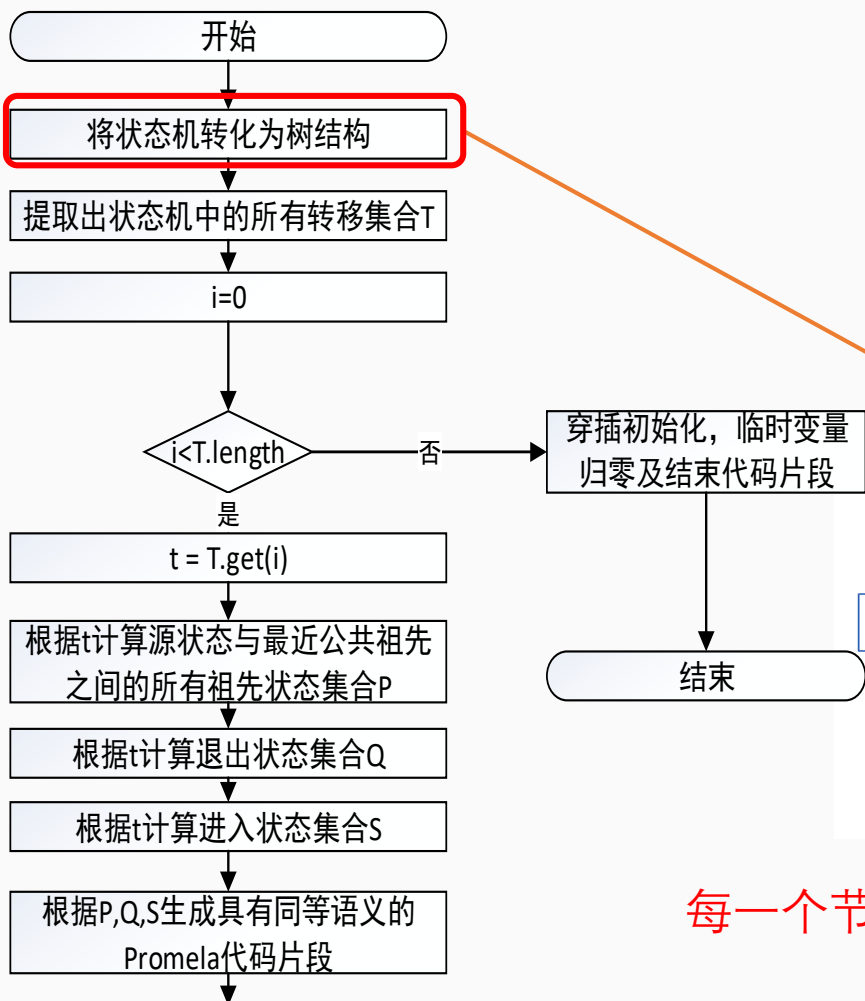
# 合理性分析—树形结构描述

1. 通过在Promela程序中嵌入c程序片段，利用结构体和指针描述树形结构。  
优点：程序无冗余；缺点：出错难排查。
2. 使用如下表所示的promela片段表达表达树形结构。  
优点：易于编程和排错；缺点：程序有一定的冗余和限制。

```
typedef TreeNode{  
    int instanceId;//Promela进程实例id 对应于状态机 workflow实例ID  
    mtype instanceType;//状态机 workflow实例类型  
    mtype currentState;//状态机 workflow当前所处状态  
    chan currentChan = [255] of { mtype };//当前状态机的外部事件  
    队列  
    int childrenNumber;//当前状态机的孩子数目  
    int parent;//父状态机所在的索引位置  
    int children[4];//当前状态机的孩子状态机所在的索引位置  
}  
TreeNode instanceTree[255];//状态机 workflow实例树
```



# 合理性分析—转化算法



每一个节点都包括其孩子节点，父亲节点，转移信息





中山大学

SUN YAT-SEN UNIVERSITY

# 合理性分析—转化算法

所有变量定义;

```
active proctype ourtype(){
```

```
do
```

```
::(迁移1发生的条件)->
```

```
atomic{
```

退出迁移1的源状态, 状态变化, 设置迁移1发生, 退出动作, 变量赋值

```
}
```

```
.....
```

```
::(迁移n发生的条件)->
```

```
atomic{
```

退出迁移n的源状态, 动作

```
}
```

```
::else->{
```

```
if
```

```
::(状态发生了变化)->
```

临时变量归零

```
do
```

```
::(迁移1是否发生)->
```

```
atomic{
```

激活迁移1的目的状态, 进入动作

```
}
```

```
.....
```

```
::(迁移n是否发生)->
```

```
atomic{
```

激活迁移n的目的状态

```
}
```

```
od;
```

临时变量归零

```
::else->atomic{
```

从当前状态机事件队列中取出一个事件

```
}
```

```
fi;
```

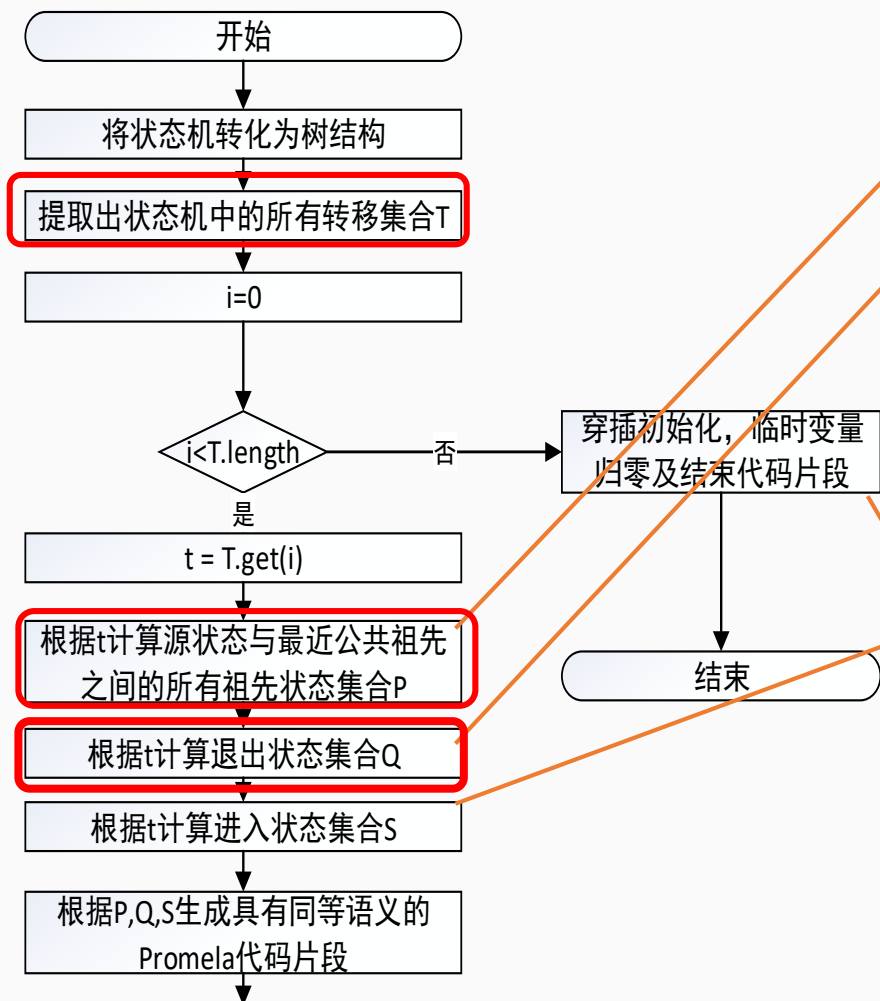
```
}
```

```
od;
```

DeadLock: printf("DeadLock");//没有事件可以派遣, 程序不可继续运行

TheEnd: skip;当此到达结束状态的时候指向此语句

```
}
```





中山大學

SUN YAT-SEN UNIVERSITY

# 合理性分析—合理性的公式表达

## ➤公式

$[\langle \rangle\_pr\_nr == 1]$ ：即系统中是否最终只存在该公式对应的NeverClaim这一个进程，其他 workflow 实例的进程最终都会到达终止状态。

## ➤检验方法

将Promela程序和上述公式输入到Spin中进行验证，如果显示no error found则表示模型满足合理性，否则给出不满足合理性的一个执行路径。





中山大學  
SUN YAT-SEN UNIVERSITY

# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望





# 状态机 workflow 模型—引擎原型实现

## Apache Commons SCXML API

Core API

Advance API

SCXML  
Parser

SCXML  
Data  
Model

Content  
and  
Evaluators

SCXML  
Executor

Triggering  
Event

Customer  
Actions

Customer  
Semantics

状态机 workflow 引擎运行时环境

流程解  
析器

任务派发  
器

实例管理  
器

自定义  
语义

多状态机事件转发器

状态机 workflow 实例

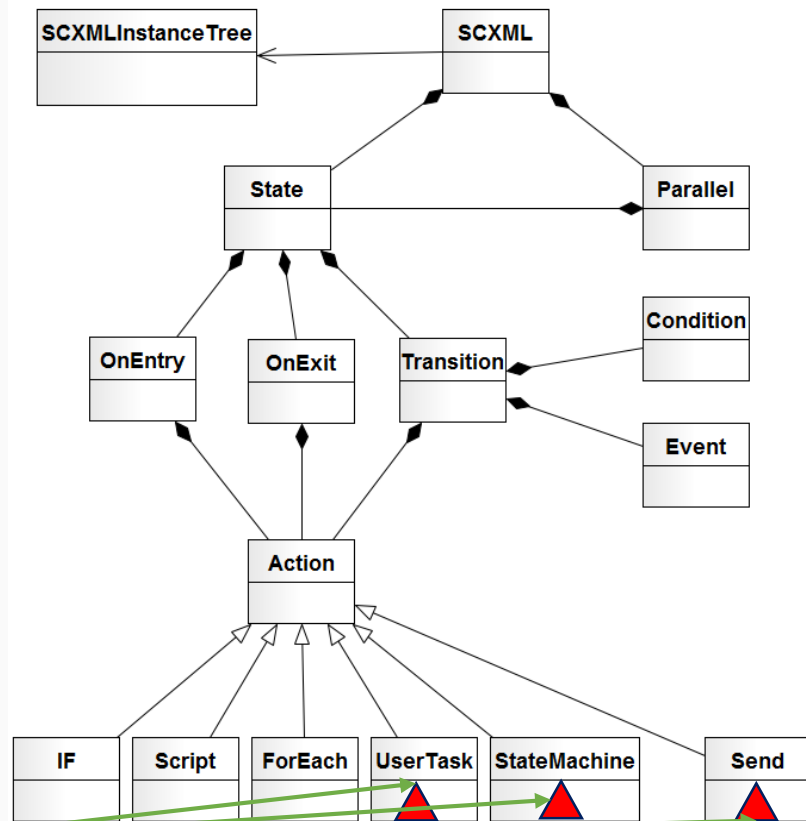
数据模  
型

流程上下  
文

实例执  
行器

状态机  
实例树

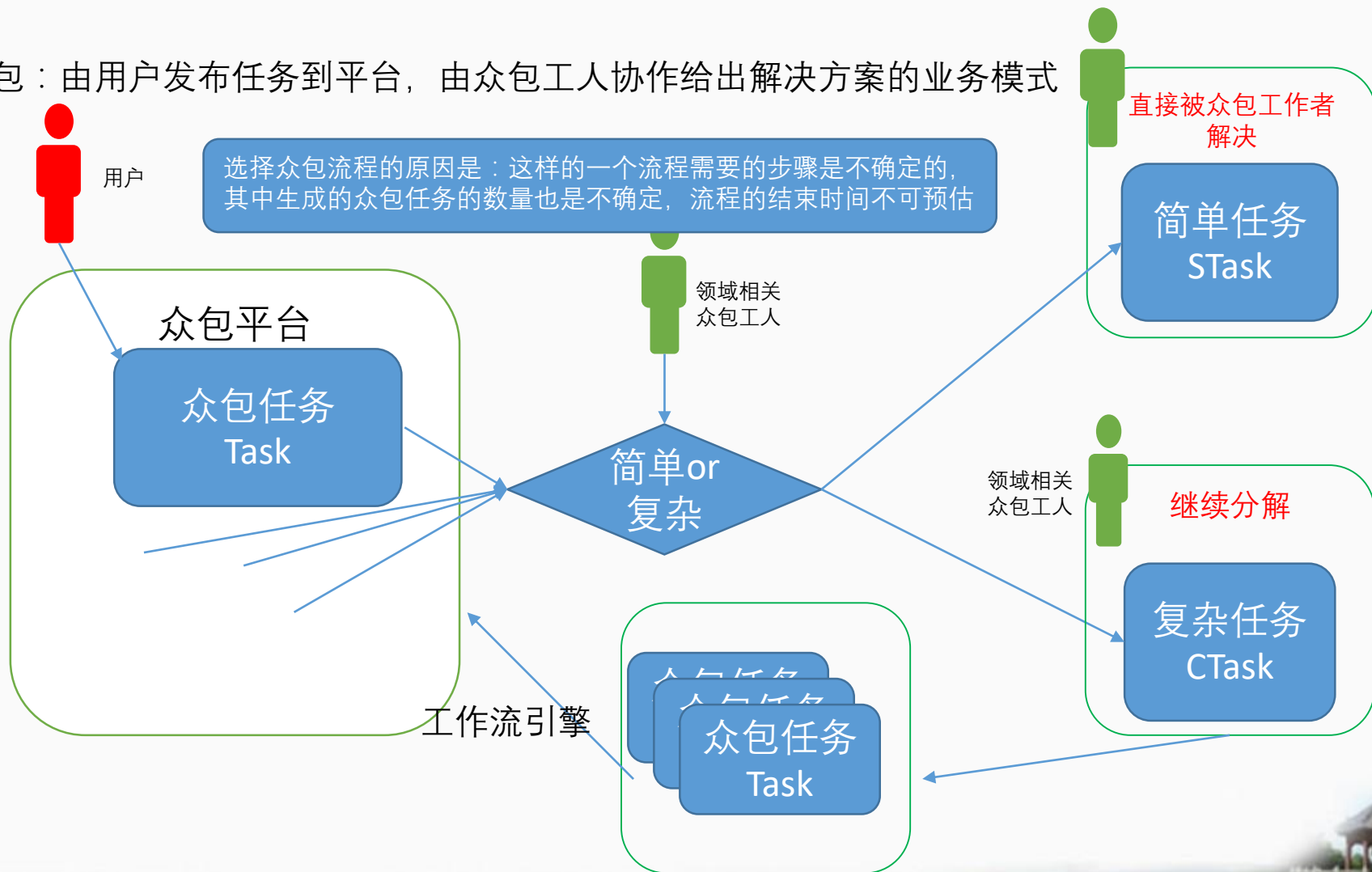
自定义动  
作





# 案例与实验结果分析—众包介绍

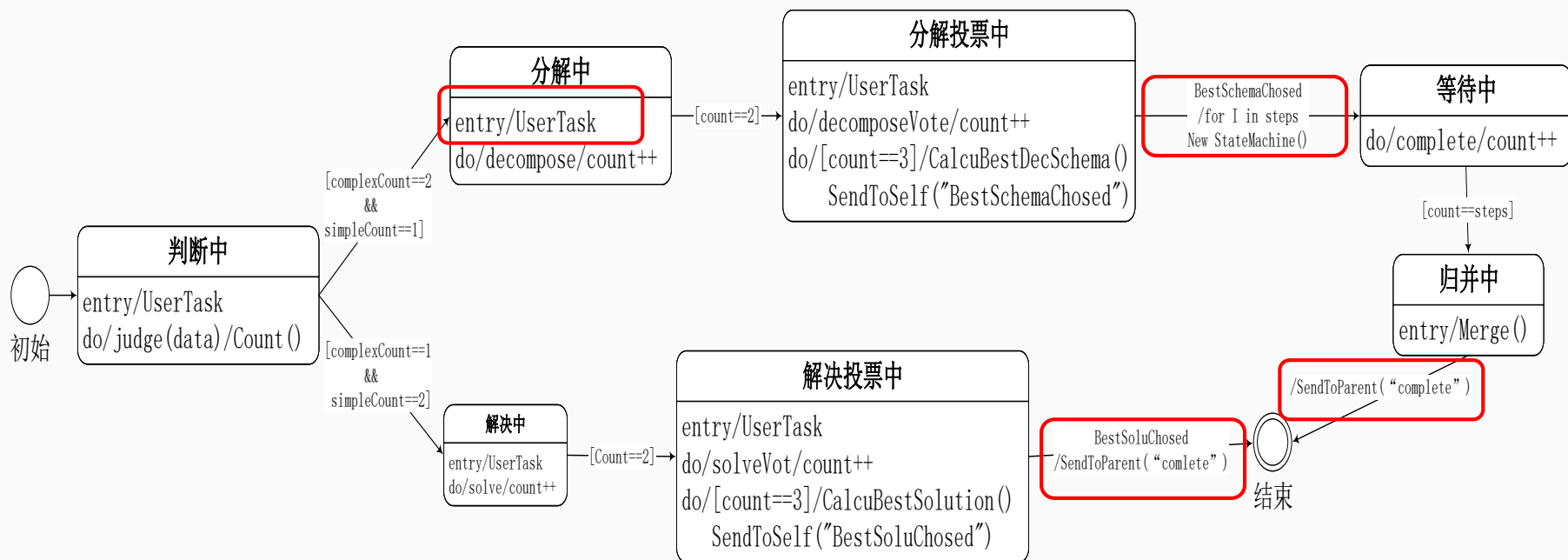
众包：由用户发布任务到平台，由众包工人协作给出解决方案的业务模式





中山大学  
SUN YAT-SEN UNIVERSITY

# 案例与实验结果分析—众包任务状态机





# 案例与实验结果分析—众包流程定义模型检验

```
spin -a crowdsourcing1.pml
ltl finalend: [] (<> ((nr_pr==1)))
gcc -DMEMLIM=4096 -O2 -DVECTORSZ=100000 -DNFAIR=3 -DXUSAFE -w -o pan pan.c
./pan -m10000 -a -N finalend
Pid: 9828
warning: only one claim defined, -N ignored
pan: out of memory
      4.17711e+09 bytes used
      1e+07 bytes more needed
      4.29497e+09 bytes limit
hint: to reduce memory, recompile with
      -DCOLLAPSE # good, fast compression, or
      -DMA=75232 # better/slower compression, or
      -DHC # hash-compaction, approximation
      -DBITSTATE # supertrace, approximation

(Spin Version 6.4.5 -- 1 January 2016)
Warning: Search not completed
        + Partial Order Reduction

Full statespace search for:
      never claim          + (finalend)
      assertion violations + (if within scope of claim)
      acceptance cycles    + (fairness disabled)
      invalid end states   - (disabled by never claim)

State-vector 75232 byte, depth reached 3194, errors: 0
      54956 states, stored (108704 visited)
      58150 states, matched
      166854 transitions (= visited+matched)
      1862 atomic steps
hash conflicts:      108 (resolved)

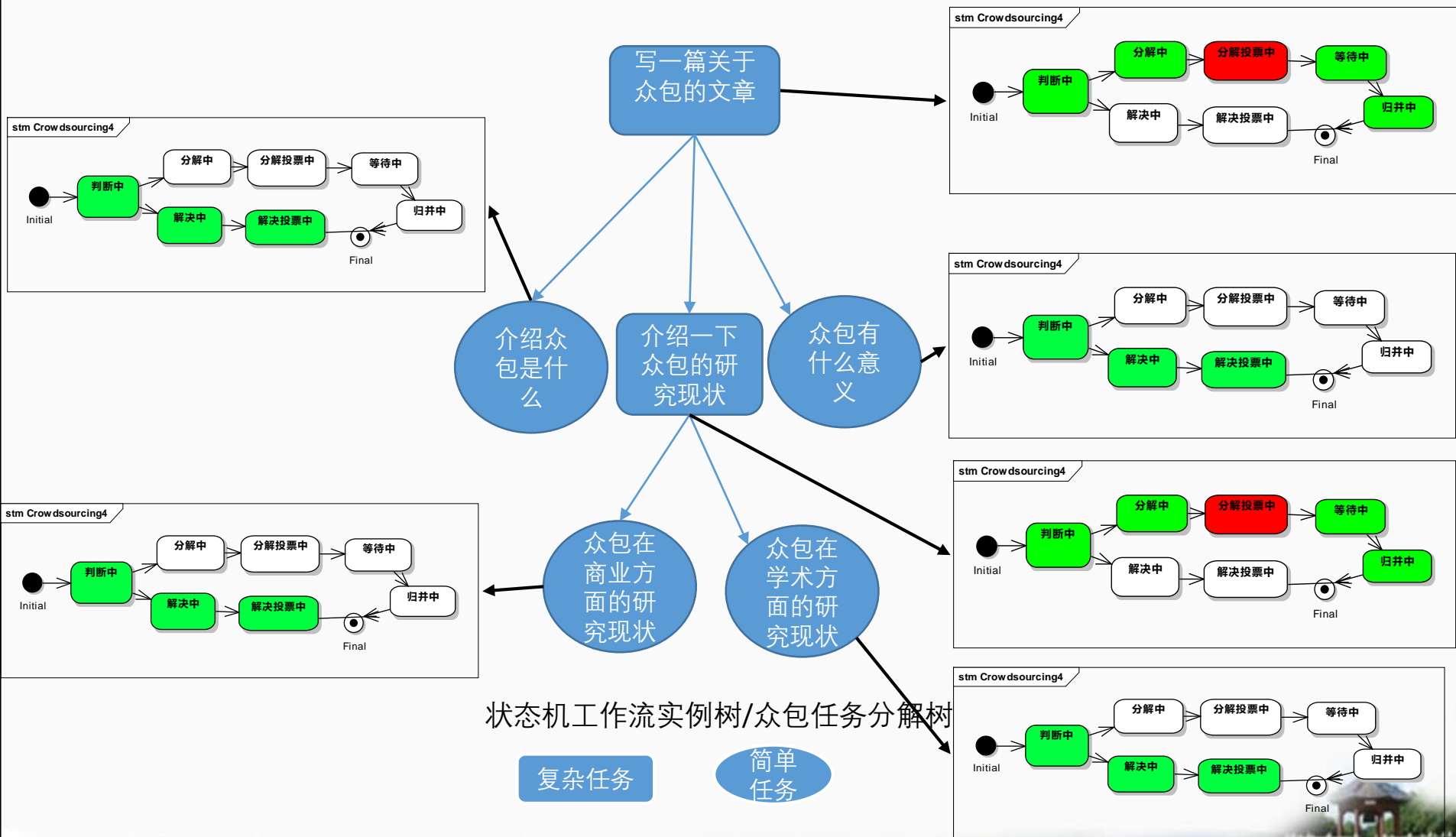
Stats on memory usage (in Megabytes):
3943.757      equivalent memory usage for states (stored*(State-vector + overhead))
3942.921      actual memory usage for states (compression: 99.98%)
              state-vector as stored = 75216 byte + 16 byte overhead
        64.000      memory used for hash table (-w24)
         0.343      memory used for DFS stack (-m10000)
        23.741      memory lost to fragmentation
3983.601      total actual memory usage

pan: elapsed time 45 seconds
No errors found -- did you verify all claims?
```





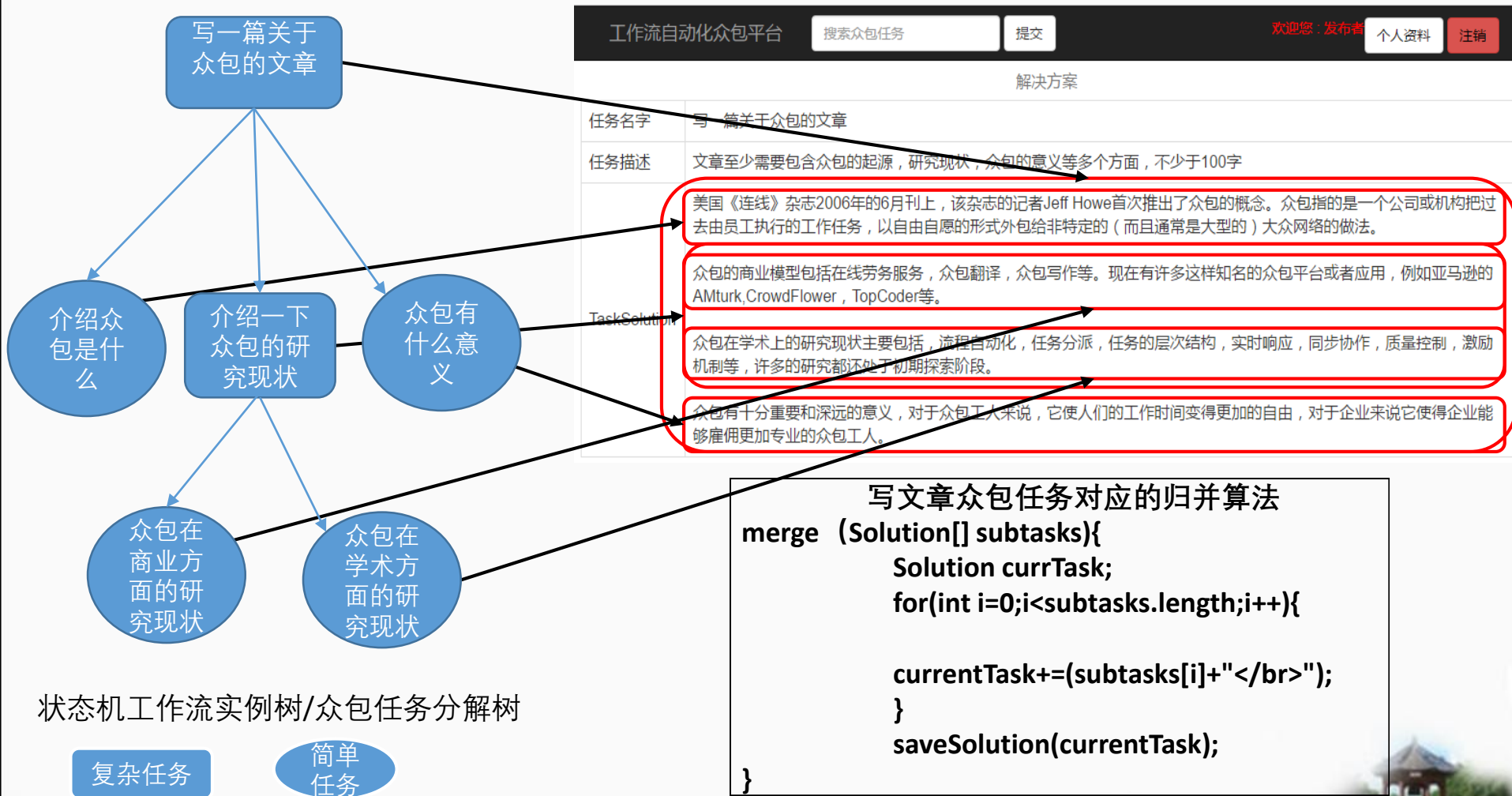
# 案例与实验结果分析—实验结果



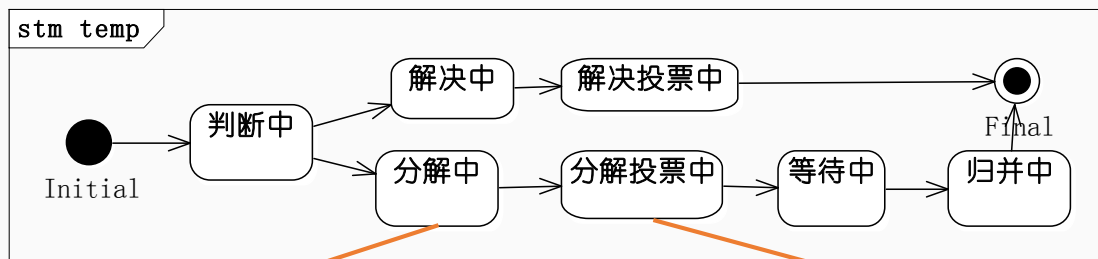




# 案例与实验结果分析—实验结果



# 案例与实验结果分析—众包图形化界面



workflows 众包平台 搜索众包任务 提交 欢迎您: 分解者1 个人资料 注销

我的所有任务 发布的任务 判断任务 分解任务 解决任务 投票任务 归并任务 其他任务

任务分派时间: 2017-3-23 16:09:36 执行者: 分解者1

任务名字	写一篇关于众包的文章
任务描述	文章至少需要包含众包的起源, 研究现状, 众包的意义等多个方面, 不少于100字
步骤数量	3
步骤1	众包是什么 请简要介绍众包的起源, 以及众包的概念
步骤2	众包的研究现状 请从多个角度去介绍一下众包的研究现状, 不少于2个角度, 不少于50字
步骤3	众包的意义 请简要描述众包的意义及其对人们生活的影响

完成

workflows 众包平台 搜索众包任务 提交 欢迎您: 投票者1 个人资料 注销

我的所有任务 发布的任务 判断任务 分解任务 解决任务 投票任务 归并任务 其他任务

任务分派时间: 2017-3-23 16:56:32 执行者: 投票者1

任务名字	写一篇关于众包的文章
任务描述	文章至少需要包含众包的起源, 研究现状, 众包的意义等多个方面, 不少于100字

上述任务的分解结果如下, 请选出您认为的最佳分解方案

执行者	分解者2	<input type="radio"/>
步骤1	众包的历史	请对众包的历史做一个简要的描述
步骤2	众包的意义	请对众包的意义做一个简要的描述
执行者	分解者1	<input type="radio"/>
步骤1	众包是什么	请简要介绍众包的起源, 以及众包的概念
步骤2	众包的研究现状	请从多个角度去介绍一个众包的研究现状, 不少于两个角度, 不少于50字
步骤3	众包的意义	请简要描述众包的意义及其对人们生活的影响

Complete



# 主要内容

1. 研究背景与问题
2. 研究现状与目标
3. 状态机工作流模型
4. 状态机工作流表达能力分析
5. 状态机工作流合理性分析
6. 实验结果及分析
7. 总结与展望



## ➤总结

1. 提出了一种新的状态机工作流建模方法。
2. 分析了状态机工作流的表达能力。
3. 给出状态机工作流模型的合理性定义及其检验方法。
4. 实现了一个支持此状态机工作流建模的状态机工作流引擎原型。

## ➤展望

1. 状态机工作流模型协作方法能够推而广之运用到任意类型的流程建模，需要实践检验。
2. 使用嵌入c语言的方式来表达状态机工作流树形结构。
3. 使用更灵活的众包流程，利用完善过后的工作流引擎以支持其运行。
4. 总结状态机工作流的特点，探索面向业务对象的工作流。





# 攻读硕士学位期间研究成果

- [1] Crowdsourcing Complex Task Automatically by Workflow Technology.  
International Workshop on Management of Information, Processes and Cooperation,  
Hangzhou, China, Sep, 23, 2016, pp: 17-30. (EI索引, 本人第一作者, 与论文第六章相关)
- [2] 一种基于SCXML标准的状态机工作流引擎[P]. 中国: CN106528157A,  
2017.03.22 (导师第一作者, 本人第二作者, 与论文第三章相关)



# 致谢

- 衷心感谢：
  - 我的导师
  - 在座各位专家、老师
  - 参与论文评审的老师

请评委老师提问和指正，谢谢！

