

# 面向复用的流程相似判定方法

胡俊<sup>1</sup> 孙瑞志<sup>1</sup> 向勇<sup>2</sup>

(中国农业大学计算机科学与技术系 北京 100083)<sup>1</sup> (清华大学计算机科学与技术系 北京 100084)<sup>2</sup>

**摘要** 业务流程的复用即利用已有的业务流程或者业务知识等建立新的流程模型,是 workflow 技术研究的一个热点。业务流程的复用能够降低流程定义的复杂度,提高流程定义的质量和效率。为实现流程复用,提出了相似流程的判定依据和一种面向复用的流程相似判定方法,详细阐述了面向复用的流程相似判定方法的原理和步骤。利用该方法能够判定流程是否相似并根据相似的流程自动归纳生成 workflow 模板,实现流程定义的复用。

**关键词** 本体,流程复用,流程相似,workflow 模板

**中图分类号** TP311 **文献标识码** A

## Similarity Determination Method of Workflow Process Oriented to Reuse

HU Jun<sup>1</sup> SUN Rui-zhi<sup>1</sup> XIANG Yong<sup>2</sup>

(Department of Computer Science and Technology, China Agricultural University, Beijing 100083, China)<sup>1</sup>

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)<sup>2</sup>

**Abstract** Business process reuse makes use of existing business processes or business knowledge to create a new process model. It's a hot research in workflow technology, business process reuse can reduce the complexity of process definition, improve the quality and efficiency of process definition. This paper presented the basis and method of determining similarity of process, and detailed the principles and steps of the Similarity Determination method of Workflow Process Oriented to Reuse. Using this method can determine whether the processer is similar, and produce a workflow template automatically according to the similar processes. It achieves the reuse of the process definition.

**Keywords** Ontology, Business reuse, Process similarity, Workflow template

## 1 引言

业务流程是指在某企业或机构中,能够实现业务目标和策略的相互连接的过程和活动集<sup>[1]</sup>。在企业构建业务流程时,因具体情况不同导致有成千上万流程,但往往相同领域里的流程间相似性<sup>[2]</sup>很强,在流程创建时若从零开发与已有流程相似的业务流程显然是不必要的。若能重复利用已积累的业务流程和知识,必能提高系统质量、开发生产率,降低整体成本。

工作流过程定义是业务流程的计算机化的形式表示。目前过程定义语言主要有 XPDL, BPML, BPEL 等。本文采用描述能力较强并获得工业界广泛认同的 XML 流程定义语言(XML Process Definition Language, XPDL)<sup>[3]</sup>来描述流程定义。workflow 模板是一种可通用的粗粒度的工作流过程定义,它是整个流程定义的基本框架,但是没有确定其中每个具体活动,可通过后期加工或修改加以确定,以满足具体业务的不同需要<sup>[4]</sup>。构建合适准确的工作流模板是流程复用的难点之一。

本文研究了工作流过程定义的复用问题,根据相似性原

理,提出了一种面向复用的流程相似性判定方法。包括 3 个方面:首先提出了流程相似的判定依据,然后获取与输入流程相似的流程,最后根据相似的流程归纳生成 workflow 模板。这种面向复用的流程相似判定方法简单可靠,效率较高,也易于在实际企业中应用。

## 2 相关研究

要实现业务流程的复用,关键是创建 workflow 模板。龚晓庆等<sup>[5]</sup>提出了一种基于领域业务本体检索 workflow 模板的方法。该方法对具体业务进行分析,构建领域业务本体,然后由领域专业人员开发 workflow 模板。用户可根据业务需求检索 workflow 模板并且进行修改或者实例化,这样 workflow 模板就转化为具体的工作流过程定义,以供使用。但这种方法需要由领域专业人员人工手动构建 workflow 模板,易受业务人员主观因素影响,而且效率不高。

针对目前 workflow 系统的两个重大问题:(1)对流程定义动态改变的支持较弱;(2)某些流程定义的综合信息难以获取。Aalst 和 Basten<sup>[4]</sup>提出了一种保留继承的转化规则,建立了 4 种流程定义间的继承关系和 10 种转换规则。若需要生成新

到稿日期:2009-12-16 返修日期:2010-03-15 本文受国家自然科学基金(60736020),国家高技术研究发展计划("863"计划)基金(2007AA010305)资助。

胡俊(1986-),女,硕士生,主要研究方向为 workflow、计算机支持的协同工作, E-mail: junhu1717@gmail.com; 孙瑞志(1964-),男,博士,教授,主要研究方向为 workflow、计算机支持的协同工作; 向勇(1967-),男,博士,副教授,主要研究方向为计算机网络。

的流程定义与已有流程定义符合某种继承关系,那么就能依照转换规则将已有流程定义转化为新的流程定义。这种方法能够避免传统的工作流系统动态改变时出现的问题。不过实际的业务流程间能满足这4种继承关系的相当有限,因此只能用于某些特殊领域业务,不适用于普遍的业务领域。

对于流程相似<sup>[2]</sup>,文献[2]提出,基于 Petri net 建模的流程可通过计算流程结构或流程行为的准确率与召回率来比较流程相似度。准确率是指以一个流程为原始流程,计算第二个流程的区块(结构或行为)出现在原始流程中的比例。准确率为1,说明第二个流程的所有区块都在原始流程中存在。召回率是指原始流程中的区块出现在第二个流程中的比例。这种方法采用 Petri net 建模,使用范围较为广泛,但需根据事件日志获取典型行为,限制了使用。

本文提出的面向复用的流程相似判定方法能够自动生成工作流模板,适用于普遍的业务领域。方法简单易行,并在实际使用中初步验证了本方法的可行性与有效性。

### 3 流程相似性判定依据

流程相似度<sup>[2]</sup>表明了一个流程在某些条件下是否能被另一个流程替换。流程相似度的取值范围可设为[0,1]。当两个流程完全相同时,认为相似度为1。

文献[6]对一批工作流技术方面的专家和开发者进行调查后认为流程是否相互包含、流程服务类型等6个因素影响流程之间的相似程度。通过对实际业务流程的研究发现,绝大部分相似的流程的活动和变迁数目以及活动之间的连接方式是相同的,活动的功能也很相似,只是一些活动的属性和变迁的条件有细微差别,并且在实际流程中流程的名称一般表明了流程的功能。与此相对应,流程内活动名称也代表该活动的功能。基于这些基础,流程相似度的确定可以从以下5个方面考虑:

(1)流程功能是否相似。流程功能表示该流程所能完成的任务,这从整体上决定了流程的运行目标,是流程定义最重要的属性。

(2)流程拓扑结构是否相同。业务流程的拓扑结构说明流程的各个业务活动是如何依照逻辑关系相互连接的,它决定了流程的业务数据与控制流的运转方向,是流程的重要属性。相似的流程拓扑结构也相同。

(3)流程的各个对应活动的功能是否相似。活动可以看作将流程功能分解开来的一个个小功能模块,这些功能按照一定顺序组合,最终促使总体功能实现。相似流程的对应活动功能也是相似的。

(4)流程的各个对应活动类型是否相同。活动类型是指活动执行的方式,可分为人工活动和自动活动等。活动类型决定了活动运行控制的详细特征,是活动的重要属性之一。相似活动运行控制的方式是一致的。

(5)流程的各个对应活动的参与者是否相似。参与者是指流程活动运行时所依赖的资源,它能够完成活动实例所对应的任务,包括人员、角色、组织机构等<sup>[7]</sup>。本文仅从角色考虑,在企业流程中,一般相似的角色承担的任务也相似。

可归纳总结得出两个流程相似的判定依据是:(1)流程拓扑结构相同;(2)流程信息相似。其中,流程信息相似是指1)流程的功能相似;2)流程里对应的活动相似,即活动的功能相

似、活动的类型相同、活动的参与者角色相同。

### 4 流程相似判定方法

根据流程相似判定依据可提出一种流程相似判定方法。方法首先对流程进行结构相似性判定,然后对结构相似的流程进行流程信息相似性判定,最后由相似的流程自动归纳生成工作流模板和对应的初始化表单。流程信息指流程除结构之外的信息。在流程定义中,这些信息是通过语义来描述的。可引入流程知识本体来表示流程语义信息,进而判定流程信息相似性。

#### 4.1 构建流程信息本体

构建流程信息本体首先需要对业务知识进行分析,抽象出相关的概念类以及概念之间的关系。限定从以下几个方面刻画流程信息:

- (1)流程功能,描述流程定义能达到的目标。
- (2)活动功能,描述流程定义中活动的任务。
- (3)活动类型,描述流程定义中活动的执行方式。
- (4)参与者,描述流程定义中完成任务的人员、角色、服务等。

由此可抽象出流程、活动、流程功能、活动功能、活动类型、参与者6个概念类型。根据流程定义,这6个概念间的关系是:流程功能是流程的一个属性,活动是流程的一部分,而活动功能、活动类型是相应活动的属性,参与者则是活动的一部分。可以建立相应的本体模型,包括:

##### (1)建立本体概念类

识别流程定义的信息,可采用自顶向下的方法。首先将流程加入概念集合,然后获取流程定义的流程功能和活动并加入相应的概念集合中,最后获取活动类型和活动功能、参与者并加入概念集。

##### (2)定义本体概念类的属性

针对概念集合中每个概念的属性,确定其属性类型、取值范围以及约束<sup>[8]</sup>。这与构建本体概念类相辅相成,两者反复抽取知识,同时进行。

##### (3)定义本体概念类间的关系

根据抽象出来的概念类之间的关系,采用自顶向下的方法从最顶层的类再到底层的概念类。也可以采用自下而上的方法,从最底层的局部类开始,最后定义最顶层的类。实际应用中可将这两种方法综合运用。

##### (4)添加本体概念类的实例

实例的来源一般是领域中已经成功应用的工作流系统。可采用 OWL 描述领域业务本体,使用 Jena 中的方法来获取相应的概念。流程功能、活动功能和活动类型以及活动参与者都添加了相应的个体实例,分别为送件流程功能、审核活动功能、路由活动类型、经理参与者。构建出来的本体模型如图1所示。

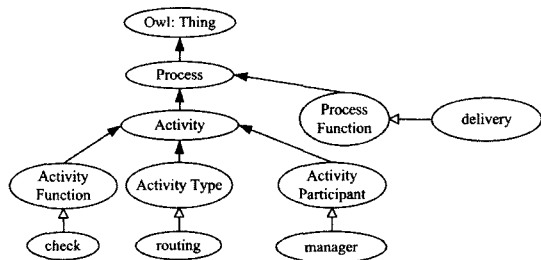


图1 流程信息本体模型

## 4.2 流程结构相似性判定

流程结构相似性判定即判定流程拓扑结构是否相同。首先需要解析输入的流程定义,以提取出有向图。具体的解析方法如下:流程定义里的活动解析为图的顶点,变迁解析为图的有向边,流程定义即被解析为有向图。对流程结构的相似性判定即转化为对有向图的同构判定。

图同构问题可分为4类:精确图同构、精确子图同构、不精确图同构、不精确子图同构<sup>[9]</sup>。本文的有向图同构是指精确的图同构。目前图的同构判定方法有很多,根据调查发现,在实际企业中,一般业务流程活动和变迁数目并不庞大,即业务流程构成的流程有向图顶点和边的数目较少。根据这个特点,可采用效率较高而且比较简单的出入度序列法<sup>[10]</sup>来判定图是否同构。具体算法如下:

(1) 获取流程有向图,构建流程有向图的邻接矩阵。

(2) 遍历流程的邻接矩阵,将这些流程进行初步分类,得到  $N$  阶同型矩阵列表(同型矩阵即这些流程的活动数和变迁数目都相同)。

(3) 在同类中获取所对应的流程图的  $C$  阶子图,  $C$  为子图的顶点数目( $C$  的初始值为 1)。

(4) 构建流程图的子图的邻接矩阵,并求其可达矩阵。

(5) 如果该子图是连通子图,求取该连通子图的出入度。子图的出入度等于原流程图的邻接矩阵中与子图顶点对应节点的出入度减去该子图的邻接矩阵的出入度。

(6) 获取该流程图的所有  $C$  阶连通子图,求取这些连通子图的出入度,构建该流程图的  $C$  阶连通子图出入度序列表。

(7)  $C=C+1$ , 转至(3), 直至  $C$  等于原流程图中活动数目  $N$ 。

(8) 遍历  $N$  阶同型矩阵列表。获取每个矩阵的  $C$  ( $C$  从 1 到  $N$ ) 阶连通子图的出入度序列列表,将所有出入度序列列表都相同的流程图归作一类,同类的流程有向图即为同构图。

(9) 返回最终分类后的流程图和对应的节点映射序列,至此流程有向图同构判定完毕。

对于同构的流程定义继续进行流程信息相似性判定。

## 4.3 流程信息相似性判定

流程结构相似判定之后,根据返回的顶点以及边的映射序列,将流程定义的活动和变迁进行一一映射,即可以进行流程信息相似性判定。由于流程语义信息是通过本体来描述的,因此流程信息相似性判定转化为基于本体的语义相似度判定,包括流程功能、类型、参与者信息相似度计算和流程相似性计算。

(1) 语义相似度

语义相似度可以用语义距离来衡量。两个概念之间语义距离越大,那么它们语义相似度就越低,反之则越高<sup>[11]</sup>。由此可建立语义距离到语义相似度的转换公式,设概念  $A$  和  $B$  相似度为  $Xsim(A, B)$ ,  $A$  和  $B$  的语义距离为  $Dist(A, B)$ ,  $A$  和  $B$  在层次网络深度分别为  $Dep(A)$  和  $Dep(B)$ , 语义相似度为<sup>[12]</sup>

$$Xsim(A, B) = 1 - \sqrt{\frac{\alpha-1}{\alpha} \times \beta \times Dist(A, B)} \quad (1)$$

式中,  $\beta = \frac{Dep(B)}{Dep(A) + Dep(B)}$ ;  $\alpha$  为可调节的参数,  $\alpha \geq 2$ 。  $Dist(A, B)$  为连接  $A$  和  $B$  的最短路径上所有  $n$  条边的权值之和,

即

$$Dist(A, B) = \sum_{i=1}^n weight_i \quad (2)$$

式中,  $weight_i$  为第  $i$  条边的权值,其计算公式为

$$weight(A) = \frac{1}{wid(A)} \times \frac{1}{\vartheta} \times weight(parent(A)) \quad (3)$$

$wid(A)$  为概念  $A$  的孩子节点的数目。

(2) 流程功能相似度和活动功能相似度

设  $Pdsim(A, B)$  为流程  $A, B$  的功能相似度,  $Dsim(A_i, B_i)$  为  $A_i, B_i$  活动功能相似度。根据定义,流程功能相似度和单个活动功能相似度判定可以用语义相似度来判定。当  $Pdsim(A, B) > T_1$  时,  $A, B$  的流程功能相似,返回 1, 否则返回 0。  $T_1$  为阈值, 可据实际经验调节。  $Dsim(A_i, B_i) > T_1$  时,  $A_i, B_i$  活动功能相似,返回 1, 否则返回 0。

(3) 活动类型相似度

活动类型信息相似度的计算只需做简单判断。设  $Tsim(A_i, B_i)$  是  $A_i, B_i$  活动类型的相似度, 若活动类型一致, 则  $Tsim(A_i, B_i)$  返回 1, 否则返回 0。

(4) 活动参与者相似度

设  $Csim(A_i, B_i)$  是  $A_i, B_i$  活动参与者的相似度, 若活动参与者来自全局变量且该全局变量所对应的参与者表达式相同, 那么  $Csim(A_i, B_i)$  返回 1, 否则返回 0。若活动参与者是角色且角色一致, 则  $Csim(A_i, B_i)$  返回 1, 否则返回 0。

(5) 单个活动相似度

单个活动的相似度  $Sim(A, B)$  计算公式如下:

$$Sim(A, B) = \omega_1 \times Dsim(A, B) + \omega_2 \times Tsim(A, B) + \omega_3 \times Csim(A, B) \quad (4)$$

式中,  $\omega_1, \omega_2, \omega_3$  为可调节的参数。当  $Sim(A, B) > T_2$  时返回 1, 否则返回 0。  $T_2$  为阈值, 可据经验调节。

(6) 流程定义的所有活动相似度

$$Tasim(A, B) = \frac{\sum_{i=1}^n Sim(A_i, B_i) - S_{oken}}{\sum_{i=1}^n A_i - S_{oken}} \quad (5)$$

式中,  $S_{oken}$  为这两个流程中信牌箱活动的数目,  $\sum_{i=1}^n A_i$  为流程中活动总数,  $\sum_{i=1}^n Sim(A_i, B_i)$  为两个流程中相似活动的数目。信牌箱活动是活动间传递消息的驿站, 与活动相似无关, 应消去其影响。当  $Tasim(A, B) > T_3$  时, 返回 1, 否则返回 0。  $T_3$  为阈值, 可据经验调节。

(7) 流程信息的相似度

$$Psim(A, B) = \omega_4 \times Pdsim(A, B) + \omega_5 \times Tasim(A, B) \quad (6)$$

式中,  $\omega_4$  和  $\omega_5$  为可调节的参数。当  $Psim(A, B) > T_4$ , 说明两流程相似, 返回 1, 否则返回 0。  $T_4$  为阈值, 可调节。

据以上定义, 构建流程信息相似性判定算法如下。

(1) 对单个活动进行活动功能相似度量。获取其功能概念, 根据式(1)计算其功能相似度。递归获取需比较活动的功能概念及概念的父类(包括直接父类和间接父类), 比较是否相同。一旦相同, 则停止比较。将该相同的概念存入对应的活动功能列表。

(2) 获取单个活动公共的输入输出参数。提取活动的公共输入输出参数, 存入相应的活动输入输出参数列表。输入输出参数为流程全局变量的一部分。

(3)对单个活动进行活动类型相似度量。若比较活动的类型一致,则将该活动类型存入活动类型表中,否则将默认活动类型(手工活动)存入对应的活动类型列表。

(4)对单个活动进行活动参与者相似度量。若比较的活动参与者角色相同,则将该角色存入活动参与者列表中,否则列表中该活动的参与者为空。

(5)依照式(4)计算单个活动的总相似度。

(6)对流程中所有活动都进行相似性度量,并将相应信息存入对应列表中,最后依照式(5)计算总活动相似度。

(7)根据式(1)计算两个流程功能的相似度,并获取相同的流程功能概念,存入对应的流程功能列表。

(8)获取两流程的全局变量信息,并提取出公共的全局变量,存入相应全局变量列表。

(9)根据式(6)计算两个流程的相似度。

(10)返回结果。若流程相似,则将相似流程和列表信息也返回。至此流程相似性判定结束。

两流程经过判定之后若相似,则根据相似流程和相应的列表信息构建 workflow 模板。

#### 4.4 构造 workflow 模板和表单

workflow 模板构造方法是根据相似流程和对应信息,构建出新的 workflow 模板,并根据公共参数生成相应的初始化表单以供交互,用户使用时可以根据具体需求加以修改。具体分为以下两个部分。

##### (1)自动构造流程定义

根据相似流程的结构和流程信息,按照流程定义的描述方法构造 workflow 模板。具体构建的方法如下:

1)根据相似流程解析而成的有向图,生成 workflow 模板对应的活动和变迁。

2)获取流程功能列表中的流程功能概念,该概念即 workflow 模板的流程功能概念,获取相似流程的全局变量列表,workflow 模板的全局变量与之相同。

3)对于 workflow 模板中的单个活动,获取活动功能列表、活动类型列表、活动参与者列表和输入输出参数列表中对应的信息,设置该活动的活动功能、活动类型、活动参与者和输入输出参数。

至此可复用的 workflow 模板生成,可对其进行加工修改,以适应具体业务的要求。

##### (2)动态构建表单

表单既是流程运行信息记录的载体,又是原始统计数据输出的基础,也是系统和用户交互的主要方式<sup>[13]</sup>,一般一个活动对应一个表单。可采用动态构建的方法生成 workflow 模板中活动对应的初始化 HTML 表单,具体步骤如下:

1)对于 workflow 模板的每个活动,获取其输入输出参数。若输入输出参数为空,那么该活动不需要构建表单。

2)遍历输入参数列表,将每个参数根据参数类型生成对应的表单项。布尔类型的参数的表单项为单选框,其他类型的参数表单项表现为文本框。

3)遍历输出参数列表。每一个输出参数生成对应的表单项,设置为不可更改。

4)表单生成后,将表单上传至服务器并保存。

5)将表单与对应的活动绑定。至此与 workflow 模板相对应的表单构建完毕。

这样自动生成的表单虽然比较简单,可以满足流程基本运行需求,可以根据具体业务需求对其加以修改充实。

## 5 试验结果

采用某公司的部分业务流程共 33 条数据进行实验,具体方案如下:首先由业务专家人工对这些流程进行相似性评估,将流程按照相似性归类,并且手工构建对应的工作流模板。然后应用流程相似判定方法对这些流程进行相似判定并自动生成对应的工作流模板,对结果进行比较。实验时选取了多组参数和阈值。当试验相关参数与阈值的取值为  $\alpha=2, w_1=0.5, w_2=0.25, w_3=0.25, w_4=0.4, w_5=0.6$ , 阈值  $T_1=0.8, T_2=0.6, T_3=0.8, T_4=0.75$  时,部分结果如表 1 所列。

表 1 流程相似性部分判定结果比较

流程 1	流程 2	流程相似判定方法	专家人工判定
长治 IP 接入	阳泉 IP 接入	相似	相似
朔州撤销合同	晋中撤销合同	相似	相似
运城终止合同	太原终止网络合同	不相似	不相似
晋城终止合同	大同终止合同	相似	相似
朔州收验货	忻州工程物资收验货	不相似	相似
吕梁收验货	阳泉采购任务	不相似	不相似

专家人工判定这 33 条流程按照相似性分为 5 类,分别为分公司收验货(10 条)、分公司 IP 接入(5 条)、分公司采购任务书(5 条)、分公司终止合同(7 条)、分公司撤销合同(6 条)。以专家人工判定结果为标准,原始流程经过流程相似判定算法之后也分成了 5 类,生成 5 个工作流模板,33 条流程中有 1 条流程误判。设  $s_e$  为流程相似判定算法误判的流程数,  $s_r$  为原始流程综述,则流程相似判定算法结果与人类主观判断结果的相对误差  $E$ :

$$E = \frac{s_e}{s_r} \quad (7)$$

根据式(7)可知此时流程相似判定算法结果与人类主观判断结果的相对误差率为 3%。表 2 显示的是当选取不同的参数与阈值对时计算出来的相对误差率,可以看出参数与阈值按照表 1 所取值时相对误差率最小。

表 2 相对误差率表

$\alpha$	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$T_1$	$T_2$	$T_3$	$T_4$	E
2	0.5	0.25	0.25	0.4	0.6	0.8	0.6	0.8	0.75	3%
3	0.6	0.2	0.2	0.3	0.7	0.75	0.7	0.85	0.9	6%
4	0.6	0.2	0.2	0.25	0.75	0.7	0.7	0.75	0.7	15%

图 4 是对图 2 和图 3 所示流程采用流程相似性判定方法后自动生产的工作流模板。经分析,该工作流模板与经过专家手工创建的工作流模板比较吻合,这说明方法是有效的。

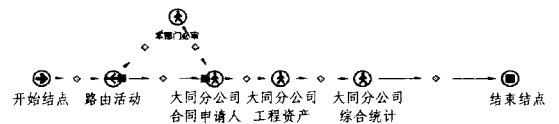


图 2 大同分公司收验货流程

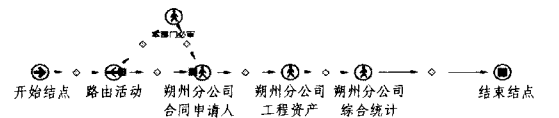


图 3 朔州分公司收验货流程

(下转第 171 页)

不会导致不可判定推理的 OCL 约束并研究其形式化的方法。

(2)本文中用于检验该方法的元数据中的冲突信息是人为引入的。由于真实系统的元数据具有更自然的冲突情况和演化过程,因此下一步工作是在更真实的环境中检验我们的方法。

(3)对于每种特定的冲突,都有多种可能的消解方法。如何有效地管理多种冲突以及它们的消解活动,是下一步要研究的问题。

### 参考文献

[1] Object Management Group. Common Warehouse Metamodel (CWM) Specification Version 1.1[M]. 2003

[2] Baader F, Calvanese D, McGuinness D, et al. The description logic handbook: Theory, implementation, and applications (2nd edition)[M]. Cambridge: Cambridge University Press, 2007

[3] University of Southern California. Loom Knowledge Representation System 4.0[M]. 2004

[4] Haarslev V, Moller R, Wessel M. RacerPro User's Guide Version 2.0[M]. 2009

[5] Horrocks I. FaCT and iFaCT[C]//Proceedings of the 1999 International Workshop on Description Logics(DL99). 1999:133-135

[6] Calvanese D, De Giacomo G, Lenzerini M. Identification constraints and functional dependencies in description logics[C]//Proceedings of the 17th International Joint Conference on Artificial Intelligence(IJCAI 2001). 2001

[7] MacGregor R. Inside the LOOM description classifier [J]. SIGART Bull, 1991, 2(3): 88-92

[8] Margaritopoulos T, Margaritopoulos M, Mavridis I, et al. A Conceptual Framework for Metadata Quality Assessment[C]//Proceedings of the 2008 Dublin Core and Metadata Applications Conference: Metadata for Semantic and Social Applications(DC-MI). Berlin: Universitätsverlag Gottingen, 2008: 104-113

[9] Hauch R, Miller A, Cardwell R. Information intelligence; meta-data for information discovery, access, and integration[C]//Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data. New York: ACM Press, 2005: 793-798

[10] Finkelstein A. A Foolish Consistency : Technical Challenges in Consistency Management[C]//Proceedings of the 11th International Conference on Database and Expert Systems Applications (DEXA). Heidelberg: Springer, 2000: 1-5

[11] Paige R F, Brooke P J, Ostroff J S. Metamodel-based Model Conformance and Multi-view Consistency Checking[J]. ACM Transactions on Software Engineering and Methodology (TOSEM), 2007, 16(3): 1-48

[12] Calvanese D, De Giacomo G. Expressive description logics[A]//The Description Logic Handbook: Theory, Implementation and Applications [C]. Cambridge: Cambridge University Press, 2003: 178-218

[13] Simmonds J. Consistency maintenance of UML models with description logic[D]. Brussel ; Vrije Universiteit Brussel, 2003

[14] Cabot J, Clariso R, Riera D. Verification of UML/OCL Class Diagrams Using Constraint Programming[C]//Proceedings of the 2008 IEEE International Conference on Software Testing Verification and Validation. New York: IEEE Press, 2008: 73-80

[15] Satoh K, Kaneiwa K, Uno T. Contradiction Finding and Minimal Recovery for UML Class Diagrams[C]//Proceedings of the 21st IEEE/ACM International Conference on Automated Software Engineering. New York: IEEE Press, 2006: 277-280

[16] Baader F, Lutz C. Description Logic [A]// The Handbook of Modal Logic[C]. Elsevier, 2006: 757-820

[17] Motik B, Grau B C, Horrocks I, et al. Representing Structured Objects Using Description Graphs[C]//Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR 2008). California: AAAI Press, 2008: 296-306

(上接第 144 页)

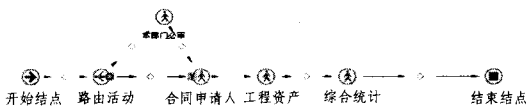


图 4 自动生成的分公司收验货工作流程模板

**结束语** 本文研究了工作流程定义的复用问题,提出了流程相似的判定依据,继而提出了流程相似性判定方法。本方法对流程进行结构和信息相似判定后自动生成 workflow 模板和表单。经测试,本方法运行的结果同专家手工创建的结果较为接近,证明流程相似性判定方法是有效的。在后续研究中,我们可以扩充流程定义的本体模型,细化流程相似的判定依据,以提高流程相似判定方法的准确率。

### 参考文献

[1] 岳晓丽,杨斌,郝克刚. 信牌驱动式 workflow 计算模型[J]. 计算机研究与发展, 2000, 37(12): 1513-1519

[2] van der Aalst W M P, Medeiros A K A, Weijters A J M M. Process Equivalence: Comparing Two Process Models Based on Observed Behavior[C]//International Conference on BPM'06. Lecture Notes in Computer Science, 2006(4102): 129-144

[3] Coalition W M. Workflow process definition interface-xml process definition language[R]. WFMC-TC-1025. Workflow Mana-

gement Coalition, 2002

[4] van der Aalst W M P, Basten T. Inheritance of workflows: an approach to tackling problems related to change[J]. Theoretical Computer Science, 2002(270): 125-203

[5] 龚晓庆,刘锋,葛玮,等. 基于复用的 workflow 过程定义工具——PDTBR[J]. 计算机应用, 2009, 20(1): 315-318

[6] Goderis A, Li P, Goble C. Workflow discovery: the problem, a case study from e-Science and a graph-based solution[C]//2006 IEEE International Conference on Web Services. 2006: 312-319

[7] 张少华,向勇,沈浴竹,等. 网格 workflow 生成中的一种流程执行机制[J]. 通信学报, 2008, 29(6): 43-50

[8] 韩轲,黄永忠,刘振林,等. OWL 本体构建方法的研究[J]. 计算机工程与设计, 2008, 29(6): 1397-1400

[9] Toran J. On the hardness of graph isomorphism[C]//Proc. 41st Annual Symposium on Foundations of Computer Science (FOCS). Redondo Beach, CA, Nov 2000: 180-186

[10] 李锋,商慧亮. 有向图的同构判定算法: 出入度序列法[J]. 应用科学学报, 2002, 20(3): 258-262

[11] 刘群,李素建. 基于知网的词汇语义相似度的计算[J]. 计算机语言学及中文信息处理, 2002(7): 59-76

[12] 徐德智,王怀民. 基于本体的概念间语义相似度计算方法研究[J]. 计算机工程与应用, 2007, 43(8): 154-156

[13] 陶望龙,邵新宇,张国军,等. Web 环境下基于表单的 workflow 管理系统研究[J]. 计算机应用研究, 2003, 12: 55-57