DISS. ETH NO 22266

Detecting, Understanding, and Fixing Control-Flow Errors in Business Process Models

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by
Cédric Favre

MSc in Computer Science, EPFL

born on 20.04.1984

citizen of Bavois (VD)

accepted on the recommendation of
Prof. Dr. Peter Müller
Prof. Dr. David Basin
Prof. Dr. Wil M.P. van der Aalst
Dr. Hagen Völzer

2014

Abstract

Business process management targets the continuous optimization of business processes within an organization using information systems. This management discipline uses business process models as central artifacts in numerous use cases including simulation, specification and code generation of an IT solution, or the direct execution of a business process model on a workflow engine.

Business process models frequently contain control-flow errors, such as deadlocks. These errors have severe consequences on the executability of the business process model: A control-flow error can lead to unexpected results, cause runtime exceptions, or completely block the process execution. Being able to detect these errors automatically during the modeling of the process allows companies to save time and money.

A business process is modeled by a graph of activities arranged to achieve a specific business goal. The control-flow of a business process model is defined by the structure of its underlying graph. Consequently, we characterize control-flow errors in terms of structural error patterns, i.e., in terms of three graph structures that are indicative of a control-flow error. We present a technique to detect these error patterns in polynomial time. Upon detecting a control-flow error, the technique delivers to the user a visualization of the error pattern in the business process model and a reduced error trace.

The business processes are modeled using industrial languages which are more complex than the control-flow models used for their analysis. We prove that the control-flow models used by many existing analysis techniques, including the structural approach mentioned previously, cannot represent in a satisfactory manner a particular type of synchronization logic, viz. the inclusive OR-join, used in some popular industrial languages.

We propose a second control-flow analysis technique to check acyclic business process models that may contain inclusive OR-joins. This technique computes the control-flow relationships between the elements of the business process models in quadratic time. We use these relationships to derive a control-flow analysis technique and as a means for the user to reason about the control-flow errors.

Finally, we show how the presented techniques can be combined to obtain a control-flow analysis tool allowing a user without a verification background to perform a control-flow analysis at modeling time. The tool is designed to support the user in detecting, locating, understanding, and fixing control-flow errors. We validate the applicability of this tool on a set of 1350 industrial business process models.

Résumé

Business process management (fr. la gestion des processus commerciaux) vise à continuellement optimiser les processus au sein d'une organisation au moyen de systèmes d'information. Cette discipline de management est basée sur des modèles de processus. Ces modèles ont de nombreuses applications comprenant la simulation, la spécification et la génération d'une solution informatique, ou encore d'être directement exécutés par un moteur de processus.

Les modèles de processus métiers contiennent fréquemment des erreurs de flux de contrôle, comme une deadlock par exemple. Ces erreurs ont de graves conséquences sur l'exécutabilité du modèle: Une erreur de flux de contrôle peut engendrer des résultats inattendus, provoquer des exceptions, ou bloquer l'exécution du processus. La détection précoce d'une erreur de flux de contrôle peut permettre des économies considérables.

Un processus est modélisé par un graphe composé d'activités organisées de manière à atteindre les objectifs de l'entreprise. Le flux de contrôle d'un modèle de processus est défini par la structure de son graphe sous-jacent. Par conséquent, nous caractérisons les erreurs de flux de contrôle en fonction de trois structures d'erreur, c'est dire, trois structures graphiques qui sont indicatives d'erreurs de flux de contrôle. Nous présentons une technique permettant de détecter ces structures d'erreur en temps polynomial. Lorsqu'une erreur de flux de contrôle est détectée, la technique offre deux types d'information: une visualisation de la structure d'erreur et un contre-exemple.

Les processus sont modélisés au moyen de langages industriels qui sont plus complexes que les modèles utilisés pour leur analyse. Nous prouvons que les modèles utilisés par de nombreuses techniques d'analyse existantes, incluant la technique mentionnée plus tôt, ne peuvent pas représenter un type de synchronisation particuler, appellé 'IOR-join', qui est utilisé par certains langages industriels communs.

Nous proposons une seconde technique d'analyse de flux de contrôle adaptée à l'analyse de processus acycliques qui peuvent contenir des 'IOR-joins'. Cette technique calcule les relations de flux de conrôle entre les éléments d'un processus. Ces relations sont utilisées pour obtenir une technique d'analyse et comme outil de raisonnement pour l'utilisateur.

Finalement, nous montrons comment les techniques présentées peuvent être combinées pour obtenir un outil d'analyse de flux de contrôle, qui permet à un utilisateur sans connaissances préalables en vérification d'effectuer une analyse de flux contrôle lors de la modélisation d'un processus. L'outil

est conçu pour assister l'utilisateur lors de la détection, la localisation, la compréhension et la correction d'une erreur de flux de contrôle. Nous validons cet outil sur un ensemble de 1350 modèles de processus.