

TU Eindhoven al jaren actief op gebied van workflow

PROCESS MINING IS TOVEREN

Sinds begin jaren '90 richt de TU Eindhoven het zoeklicht op procesmodellering en workflowautomatisering. Dat resulteerde onder meer in de bouw van het eigen open-source workflowsysteem YAWL. Onlangs werd een geslaagde proef gedaan met het batchgewijs verifiëren van honderden SAP referentiemodellen. Daarnaast werk men aan nieuwe technieken voor het automatisch in kaart brengen van processen en een methodiek om te voorspellen of een model fouten zal bevatten, zonder het model zelf in te zien. Een gesprek met professor Wil van der Aalst.

Door Hans Lamboo

“**B**egin jaren '90 beschikten we al over modellen om processen, zij het vooral logistieke, te beschrijven en analyseren. Het viel ons op dat de nieuwere informatiesystemen steeds meer modelgedreven werden. De workflowmodellen die in informatiesystemen zitten, lijken veel op simulatiemodellen. Nu is het bij een logistiek proces zo, dat eventuele fouten zich vroeg of laat vanzelf openbaren. Maar ontwerpfouten in een administratief proces komen laat of soms helemaal nooit aan het licht – terwijl ze veel onnodig extra werk genereren.” Aan het woord is prof. Wil van der Aalst, hoogleraar aan de sectie Informatiesystemen van de TU Eindhoven.

Workflow-patronen

Het werkgebied van Van der Aalst c.s. wordt nog verder verduidelijkt aan de hand van een schema, afbeelding 1. “In het ontwerpgedeelte kijken we naar allerlei talen, naar de eigenschappen van de modellen en waar eventuele verbeteringen kunnen worden aangebracht. We timmeren al enige tijd aan de weg met ons werk op het gebied van workflowpatronen. We hebben zelfs een website, www.workflow-patterns.com, die dagelijks door honderden mensen wordt geraadpleegd. We vergelijken daar een groot aantal workflowsystemen op basis van die patronen. We hebben typische constructies geïdentificeerd en kijken naar wat wij noemen ‘de control flow’, wat voor mechanisme bepaalt de volgorde, we bekijken de data. We kijken ook – en dat is met name interessant – naar de mechanismen op het gebied van werkverdeling. Hier zijn grote verschillen te vinden, sommige systemen zijn sterk push-georiënteerd, andere meer pull-georiënteerd. Zo hebben we een collectie van ruim veertig workflowpatronen in elk van die domeinen op basis waarvan de workflowmanagementsystemen met elkaar kunnen worden vergeleken. Dat zit allemaal aan de ontwerpkaart.”

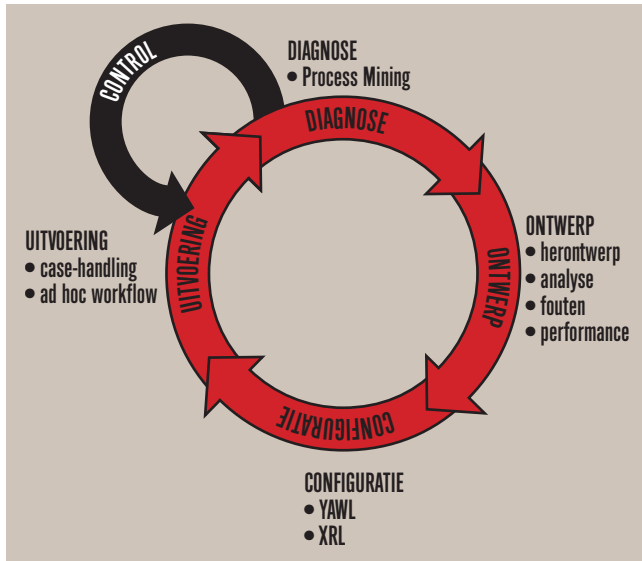
Process Mining

“In het gedeelte Uitvoering kijken we vooral naar flexibiliteit, naar case-handling en ad-hoc workflow. Met dat laatste bedoel ik situaties waarin elke individuele case in principe zijn eigen procesmodel krijgt. Als je dus met de uitvoering bezig bent van zo'n specifieke case, kun je die aanpassen. Je krijgt dan iets dat inzit tussen workflow en projectmatig werken.

In het gedeelte Diagnose werken we met vijftien promovendi aan iets dat we Process Mining hebben genoemd. En dat is toveren. Wij hebben in het Catharina-ziekenhuis in Eindhoven een log gepakt met gegevens van patiënten en welke behandelingen er zijn uitgevoerd. Op basis daarvan hebben we automatisch procesmodellen gegenereerd. Zonder enige voorkennis, gewoon op basis van wat er daadwerkelijk is gebeurd. We zijn begonnen met het toepassen van process mining op de logs van workflowsystemen. Nu zijn we in staat de logs van meer dan 20 verschillende systemen te analyseren, het gaat hierbij niet alleen om typische workflow systemen, maar ook case handling systemen, ERP systemen, enzovoorts. We hebben ontdekt dat we mining op vele gebieden kunnen toepassen uiteenlopend van patiëntstromen en administratieve processen tot complexe machines en e-business processen. We sporen op wat de knelpunten zijn. We kunnen zelfs een sociaal netwerk genereren, waarin je kunt zien wie nu met wie samenwerkt en dat dan weer gekoppeld aan allerlei aspecten van de uitvoering. Ik vind dit op dit moment iets dat heel spannend is; het staat qua toepassingsmogelijkheden nog in de kinderschoenen.”

Fouten in het SAP Referentiemodel

“Kijk, wij hebben te maken met talloze studenten die processen ontwerpen die daarna moeten worden nagekeken en geanalyseerd. We hebben daarvoor analysetools ontwikkeld. Zo kwamen we op het idee om het omvangrijke SAP Referentiemodel eens op fouten te onderzoeken. Onze doelstelling was tweeledig. Ten eerste: zijn onze analyse-toepassingen voldoende volwassen om op zo'n grote schaal



te worden toegepast? Wij waren daarvan al overtuigd, maar wilden het graag kunnen aantonen. En ten tweede: waarom maken mensen steeds ontwerpfouten? Het analyseren van processen is tijdrovend, maar je weet dat mensen altijd ergens fouten maken. Met Visio of PowerPoint ben je gauw klaar: je zegt dat het af is, je gooit het over de muur naar de programmeur en je komt ermee weg.”

Het SAP Referentiemodel wordt sinds midden jaren '90 meegeleverd. Het is een verzameling van ongeveer tienduizend modellen waarvan een groot aantal zich richt op de datakant, maar een aantal richt zich ook op de processen. Er zitten 604 niet-triviale processen in. Die 604 zijn voorgesteld in termen van EPC's, Event-driven Process Chains. “We hebben die EPC-modellen automatisch vertaald naar ons eigen workflowsysteem YAWL (acroniem voor Yet Another Workflow Language), dat is gebaseerd op workflowpatronen en qua uitdrukingskracht veel sterker dan alles wat er op de markt is,” zegt Van der Aalst. “Omdat de taal die YAWL gebruikt vele malen krachtiger is dan die van SAP is er sprake van een triviale vertaling die gevaarloos kan worden verricht. Het probleem bij EPC-modellen is dat er constructies in zitten die je zou kunnen omschrijven als ‘vaag’. Je moet gissen naar de intentie van de modelleerder. Dat wilden we uitsluiten: het was voor ons van groot belang dat als we een vermeende fout gevonden hadden, we er heel zeker van moesten zijn dat het ook echt een fout wás. We kunnen onze analysetools heel fijn afstemmen, zodat we heel veel fouten vinden, maar we kunnen ze ook heel robuust instellen zodat als we fouten vinden niemand kan ontkennen dat het ook echt fouten zijn. We hebben gekozen voor dat laatste, de conservatieve aanpak. We vertaalden in één klap het hele model naar YAWL en verifieerden vervolgens in één klap alle modellen in batch. Het resultaat: er zitten fouten in de referentiemodellen van SAP. 34 van de 604 modellen blijken evidente fouten te

bevatten, dat is 5,6 procent. We beschouwen dat als een absolute ondergrens, want het is onze inschatting dat we met een minder conservatieve aanpak op twintig procent zullen uitkomen. Dat is schrikbarend hoog, ja.”

Voorspellen van fouten

Het onderzoek heeft ook nog iets nieuws opgeleverd. “We vinden dus 34 foutieve processen in het SAP Referentiemodel. Waarom zijn die fouten überhaupt gemaakt?” vraagt Van der Aalst. “Dan komen andere aspecten zoals omvang en constructies aan de orde. Daar laat je een regressieanalyse op los. Daarmee konden we aantonen dat er een sterke correlatie bestaat tussen de omvang van het model en de fouten. Hoe groter het model, hoe groter de kans op fouten. Dat lijkt logisch. Maar we gaan verder: we hebben vijf kenmerken vastgesteld aan de hand waarvan we met 95 procent zekerheid kunnen voorspellen of er fouten in een model zullen zitten, zonder dat we naar het model zelf te kijken.”

“Ik heb ruim voor er iets over het rapport naar buiten kwam, een presentatie over onze bevindingen gegeven in het hoofdkwartier van SAP in Walldorf. Maar niemand vond het interessant of wilde er iets mee doen. Ook in de ARIS-toolset is het SAP Referentiemodel beschikbaar. Ik heb het rapport dus naar IDS Scheer gestuurd, maar ook daar wilde niemand er iets mee. Men wil de gevonden problemen dus als het ware negeren. Inmiddels is er in een toonaangevend Duits vaktijdschrift een artikel over ons rapport verschenen. Binnen SAP zijn diverse mensen erg kwaad geworden over het naar buiten komen van onze bevindingen. Men onderkent de problemen, maar geeft aan dat dit geen probleem is omdat men gestopt is met het uitleveren van het referentiemodel. Een tamelijk flauwe reactie, lijkt me”, besluit Van der Aalst. “Referentiemodellen zijn essentieel voor het verbeteren van de transparantie en configureerbaarheid van een complex systeem als SAP. Alle energie zou gericht moeten worden op het verbeteren van de bestaande modellen in plaats van het bedenken van allerlei nieuwe features die toch niet gebruikt worden.”

Hans Lamboo

is hoofdredacteur van Business Process Magazine (bpm@array.nl)

Prof. dr.ir. W.M.P. van der Aalst is verbonden als hoogleraar aan de sectie Informatiesystemen van de faculteit Wiskunde en Informatica van de TU Eindhoven. Daarnaast heeft een deeltijdaanstelling bij de faculteit Technologie Management, waar hij tot voor kort voorzitter van de capaciteitsgroep Informatie-systemen was, en heeft hij een deeltijdaanstelling aan de Queensland University of Technology in Australië. Ook is hij directeur van de opleiding Business Information Systems van de TU Eindhoven.