Business Process Management #processscience #bpmatqut

Discovering Stochastic Process Models By Reduction and Abstraction

Adam Burke, Sander Leemans and Moe Thandar Wynn Petri Nets 2021



the university for the real world



Desire Paths

Complement top-down managerial understanding with bottom-up data-driven understanding

Find out what people and systems are doing in an organization





Client: Bob; Order 721; New Order; Buy 300 kg sugar Client: Jill; Order 431; New Order; Buy 100 iceblocks Client: Bob; Order 721; Acknowledge Order

Client: Bob; Order 722; New Order; Buy Toothpaste

Event Logs

- Timestamp(s) + Case ID + Activity + Other resource information
- The timestamp-ordered activity messages for a Case ID form a *trace*
- A *trace* is then a *sequence* of activities, eg
 - <a,c,e,a>
 - <Assess Claim, Initiate Payment, Advise Claimant, ...>

• A *log* is a multiset (bag) of traces, eg [[<a,c,e,a>²¹, <c,b,f>⁴, <b,b,b,e>⁸]]



Process Mining Discovery

- Dozens of process mining discovery algorithms including very successful ones
- Explicitly modeling probability much rarer



Stochastic Process Mining Discovery

- When stochastic models used, as in performance or simulation, output of post-discovery processing
- Instead, leverage power of Stochastic PNs



Toothpaste miner



Example - Starting Log

Log [<a,b>, <a,b>, <b,a>, <c,c,c>]

Toothpaste Example - Starting Trace Model

Log [<*a*,*b*>,<*a*,*b*>,<*b*,*a*>,<*c*,*c*,*c*>]



Toothpaste Example - Concurrent reduction



Toothpaste Example - Loop Roll CO.2 / FPL.2



Toothpaste Example - Loop Roll CO.2 / FPL.2





Toothpaste miner



Toothpaste miner

Probabilistic Process Trees



Concurrent Reduction - Probabilistic Process Trees



PPT ⊆ Generalized Stochastic Labelled Petri Nets



Transformation Rule (Choice Folding)





Toothpaste Miner Variations

- Batch miner
- Incremental miner
- K-retries

Prototype

- Batch miner prototype implemented in Haskell
- Rules nice fit with Haskell pattern matching

https://github.com/adamburkegh/ toothpaste

~	coothpaste/ loothpaste.hs at mi X +
\rightarrow	C û 🗊 ≜ https://github.com/adamburkegh/toothpa … 🗵 ☆ 🛓 🛝 🗊 💿 💿
254	loopChoiceSim (Node2 Choice y (Node1 PLoop x r w1) w2)
255	x =~= y = loopMean y (Nodel PLoop x r w1)
256	loopChoiceSim x = x
257	
258	choiceFold :: (Eq a, Ord a) => PRule a
259	choiceFold (Node2 Choice
260	(Node2 Seq x1 y1 w1)
261	(Node2 Seq x2 y2 w2) w3)
262	x1 =~= x2 = Node2 Seq (merge x1 x2) (choiceP y1 y2) w3
263	y1 =~= y2 = Node2 Seq (choiceP x1 x2) (merge y1 y2) w3
264	z1 =~= z2 = Node2 Seq (choiceP h1 h2) (merge z1 z2) w3
265	<pre>where (Node2 Seq h1 z1 w4) = seqSuffix (Node2 Seq x1 y1 w1)</pre>
266	(Node2 Seq h2 z2 w5) = seqSuffix (Node2 Seq x2 y2 w2)
267	choiceFold $x = x$
268	
269	
270	loopChoiceFold :: (Eq a, Ord a) => PRule a
271	loopChoiceFold (Node2 Choice (Node2 Seq x1 y1 w1)
272	(Node2 Seq (Node1 PLoop x2 r2 w2)
273	y2 w3) w4)
274	x1 =~= x2 = Node2 Seq (loopMean x1 lx)
275	(choiceP y1 y2) w4
276	where 1x = Node1 PLoop x2 r2 w2
277	loopChoiceFold (Node2 Choice (Node2 Seq (Node1 PLoop x1 r1 w1)
278	y1 w2)
279	(Node2 Seq x2 y2 w3) w4)
280	x1 =~= x2 = Node2 Seq (loopMean x2 lx)
281	(choiceP y1 y2) w4

Toothpaste on teleclaims [1]



Teleclaims discovered model (detail)



Experimental Evaluation

- Two real-life logs
 - incident management (BPIC2013)
 - Sepsis hospital process
- Established dataset teleclaims (call centre)
- Two other stochastic discovery techniques GDT_SPN and estimator
- Conformance measures
 - Earth movers' distance (*tEMSC*)
 - Entity count
- *k*-fold cross-validation

Evaluation Results

- Prototype
 - Trades off simplicity for quality
 - Good at preserving quality as represented by truncated Earth Movers' Distance (*tEMSC*)

Overall

- Direct Stochastic Process Discovery based on rules
- Polynomial
- Works directly on Generalized Stochastic Petri Nets
- Prototype shows viability and promising preservation of quality (*tEMSC*)

Business Process Management #processscience #bpmatqut

Discovering Stochastic Process Models By Reduction and Abstraction





the university for the real world

References

[1] van der Aalst - Process Mining: Data Science In Action