Towards Probabilistic Decision Making on Human Activities Modeled with Business Process Diagrams

(Extended Abstract)

Hector G. Ceballos
Tecnologico de Monterrey
Monterrey, N.L., Mexico
ceballos@itesm.mx

Victor Flores-Solorio
Tecnologico de Monterrey
Monterrey, N.L., Mexico
vmfsolorio@gmail.com

Juan P. Garcia-Vazquez
Universidad Autonoma de Baja California
Mexicali, B.C., Mexico
jpablo.garcia@uabc.edu.mx

ABSTRACT

Agent-based technologies, originally proposed with the aim of assisting human activities, have been recently adopted in industry for automating business processes. Business Process Model and Notation (BPMN) is a standard notation for modeling business processes that provides a rich graphical representation that can be used for common understanding of processes but also for automation purposes. We propose a normal form of Business Process Diagrams (BPDs) based on Activity Theory that can be transformed into a Causal Bayesian Network, which in turn can be used to tackle with uncertainty introduced by human participants. We illustrate our approach on an Elderly health care scenario obtained from an actual contextual study.

Categories and Subject Descriptors

D.2 [**Software Engineering**]: Requirements/Specifications; G.3 [**Mathematics of Computing**]: Probability and Statistics; I.2.11 [**Computing Methodologies**]: Artificial Intelligence—*Distributed Artificial Intelligence*

Keywords

BPMN, Engineering Agent-Based Systems, Bayesian Networks, Activity Theory.

1. INTRODUCTION

BPMN is a standard notation for modeling business processes that provides a rich graphical representation that can be used for common understanding of processes [5] and it has been also used for automating processes with support of agent technologies [4]. BPMN captures *data-based decisions* through the notion of gateways and conditional control flows, but it does not cope with uncertainty introduced by the participation of people. Approaches like [3] have proposed annotating edges with the probability of each alternative, but the reason of such variability is not related to other parts of the process (i.e. causal relationships between non-consecutive nodes).

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey.*

2. PROBABILISTIC BPMN WORKFLOWS

For this reason we propose the transformation of a normal form of BPMN workflows (BPDs) to a Bayesian Network that can be used for probabilistic decision making under uncertainty and discovering causal dependencies between events and actions. Figure 1 shows the BPD of the medical consultation activity of an elder person, taken from a contextual study based on Activity Theory [2]. In this activity, the *subject* is an older adult who has a medical appointment (the *object*). The *objective* of the activity is having a medical appraisal and its *outcome* includes getting a prescription, supply medicines and schedule a next appointment. The *community* involved in the activity includes a family member (optionally) and the doctor. It illustrates two alternative ways the elder may choose for getting to the hospital: going by himself, or being taken by a family member. Each lane represents a role that a User Agent (on behalf of a person) will follow to develop the activity. The purpose of the proposed transformation is to advise to the person on making a decision when reaching to a splitting gateway without conditional control flows. Otherwise, data-driven decisions can be automatically made by user agents.

2.1 A Probabilistic BPD normal form

We propose a BDP normal form that constitutes a subset of graphical elements of the BPMN 2.0 specification [5]. The Business Process Diagram **W** is constituted by a single pool (**P**), lanes (**L**), nodes (**N**) and control flows (**F**). Nodes (**N**) can be start events (N^S), intermediate events (N^I), end events (N^E), atomic actions (N^A) or gateways (N^G). It has a single start event ($s \in N^S$), i.e. the trigger, and multiple end nodes that represent activity outcomes.

All sequence flows are unconditional, denoted as $F(n_i, n_j) \in \mathbf{F}$ where $n_i, n_j \in N$. Each split or merge of control flows must be mediated by a splitting gateway ($N^G_S \subseteq N^G$) or a merging gateway ($N^G_M \subset N^G$). Gateways can be of type Parallel-AND (A), Optional-OR (O), or Exclusive-XOR (X). Splitting XOR gateways ($g \in N^G_S, \mathit{type}(g, X)$) must be followed by intermediate event nodes ($F(g, i) \in \mathbf{F}, i \in N^I$) or other XOR gateways, denoting alternative ways on which the activity can develop.

The graph G_N constituted by all $F(n_i, n_j) \in \mathbf{F}$ must not have any directed cycle or loop, i.e. it must be a Directed Acyclic Graph (DAG). Two consecutive action nodes must be mediated by at least one intermediate event node and as many gateways as needed. This means that action nodes

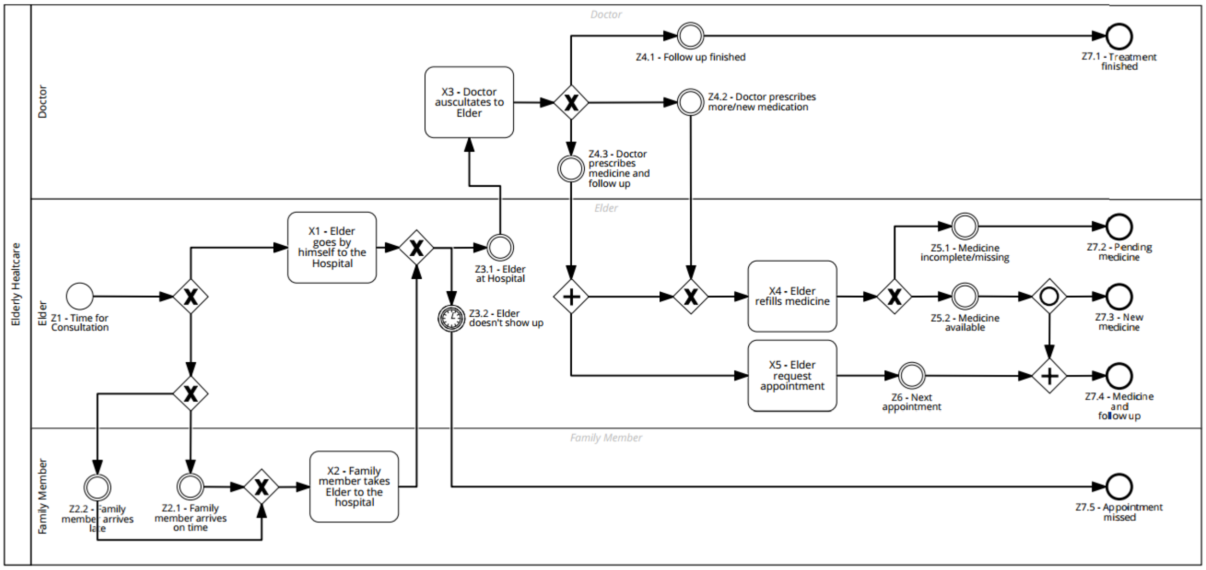

Figure 1: Business Process Diagram of the Medical Consultation Activity.

are not connected directly through sequence flows. Observable intermediate events will permit monitoring the activity development and introducing agent assistance [1].

2.2 Translation to a Bayesian Network

The translation consists on defining mappings of event and action nodes n to realizations of random variables ($V_i = v_i$), denoted $map(n, V_i = v_i)$. The start node is mapped to $Z_S = True$, every end node e_i is mapped to $Z_E = e_i$. Every intermediate event i preceded by another event or action is mapped to $Z_i = True$, whereas all those preceded by a splitting XOR gateway g are mapped to $Z_g = i$. The execution of an action a is mapped to $X_a = True$. The omission of an action a, or event n, is mapped to $X_a = False$, respectively $Z_n = False$.

In order to identify conditional dependencies between random variables, we use control flows incoming and outgoing to the corresponding event and action nodes. Gateways are ignored in this process, this is, a gateway g is replaced by a set of arcs outgoing from nodes preceding g and incoming in nodes following g. The resulting graph also codifies temporal precedence between random variables.

Gateways, on the other hand, codify how likely is that two or more events/actions occur during process execution. All valid scenarios generated observing gateway constraints constitute the joint probabilistic distribution of the process, which in turn can be used along with G_N for learning the Conditional Probabilistic Distribution of the corresponding Bayesian Network, i.e. $P(v_i|pa_i)$.

Definition 1. An Activity Causal Bayesian Network (ACBN) is represented by $D = \langle G_V, X, Z, Z_S, Z_E, P(v_i|pa_i)\rangle$, where G_V is a minimal DAG which arcs denote temporal precedence and conditional dependence between events (Z) and actions (X), $P(v_i|pa_i)$ encodes conditional probabilistic dependencies between random variables $V = Z \cup X$, and G_V has at least one directed path from the initial condition $Z_S \in Z$ to the outcome variable $Z_E \in (Z \setminus Z_S)$.

3. CONCLUSIONS

The resulting ACBN captures conditional dependencies established in the workflow, permitting to use Bayesian inference for deducing and predicting human actions based on observed events. Furthermore, learning causal dependencies from actual process instances [6] will permit to assess human decision making (e.g. going alone to the hospital if the family member is getting late).

Acknowledgments

This research was supported by CONACyT through the Basic Science grant (CB-2011-01-167460).

4. REFERENCES

[1] H. G. Ceballos, J. P. Garcia-Vazquez, and R. Brena. Using activity theory and causal diagrams for designing multiagent systems that assist human activities. In *Proceedings of the 12th Mexican International Conference on Artificial Intelligence, MICAI 2013*, pages 185–198, Mexico City, Nov. 2013.

[2] J. P. Garcia-Vazquez, M. D. Rodriguez, M. E. Tentori, D. Saldana, A. G. Andrade, and A. N. Espinoza. An agent-based architecture for developing activity-aware systems for assisting elderly. *Journal of Universal Computer Science*, 16(12):1500–1520, jun 2010.

[3] L. Herbert and R. Sharp. Precise Quantitative Analysis of Probabilistic Business Process Model and Notation Workflows. *Journal of Computing and Information Science in Engineering*, 13(1):011007(1–9), 2013.

[4] T. Kuster, M. Lutzenberger, A. Hessler, and B. Hirsh. Integrating process modelling into multi-agent systems engineering. In *Multiagent and Grid Systems*, pages 105–124, 2012.

[5] OMG. Business Process Model and Notation (BPMN), Version 2.0, January 2011.

[6] J. Pearl. *Causality. Models, Reasoning, and Inference.* Cambridge University Press, 2000.