Towards A Semantic Quality Based Approach for Business Process Models Improvement

Sarah Ayad
CEDRIC-CNAM
292 Rue Saint Martin, F-75141 Paris Cedex03
Email: ayad_sa1@auditeur.cnam.fr

Samira Si-Said Cherfi
CEDRIC-CNAM
292 Rue Saint Martin, F-75141 Paris Cedex03
Email: samira.cherfi@cnam.fr

Isabelle Comyn-Wattiau2
CEDRIC-CNAM and ESSEC Business School
Email: wattiau@cnam.fr

Abstract

Business process (BP) modeling aims at a better understanding of processes, allowing deciders to improve them. We propose to support this modeling with an approach encompassing methods and tools for BP models quality measurement and improvement. In this paper we focus on semantic quality. The latter is evaluated by aligning BP model concepts with domain knowledge. The alignment is conducted thanks to meta-models. We also define validation rules for checking the completeness of BP models. A medical case study illustrates the main steps of our approach.

Keywords

Business process modeling, model quality, semantic quality, domain knowledge, domain ontology, quality evaluation, quality improvement.

INTRODUCTION

In recent years research related to modeling and improving business processes has been of growing interest. Thanks to their experience, companies are aware of the undeniable impact of a better tuning of business processes (BP) on the effectiveness, consistency and transparency of their business operations. This tuning requires a better understanding and an effective management of BP. However, to achieve the expected benefits it is necessary to rethink the approach of designing these processes. BP modeling is a prerequisite. It is now considered as an engineering activity aiming at providing the actors with a better understanding of the processes in which they are involved. The resulting models serve as concrete tools allowing stakeholders to understand i) how processes work, ii) why they may dysfunction, iii) which software could support them efficiently, and so on. As a consequence, the quality of these models is a hot topic.

Quality can be defined as the total of properties and characteristics of a product or service that are relevant for satisfying specific and obvious requirements (Tucker and al.,1998). Following Boehm and other quality precursors, ISO subdivides quality in a number of quality characteristics such that each one addresses a particular aspect of quality. The business process modeling approaches share many similarities with conceptual modeling activities, but are much more complex. Indeed, a business process model is a representation at a very high level of abstraction but it has also to integrate non-functional requirements such as flexibility and maintainability. Modeling these processes requires a high degree of pragmatic expertise: designers refer to a set of mainly empirical rules and heuristics. The latter are difficult to formalize and to share. Commercial tools for business process modeling activities mainly focus on the accuracy of models based on a set of syntactic criteria and provide little or no guide to guarantee the quality of produced models. Recent research work focusing on the quality of process models concentrates mainly on structural aspects of models taking into account two criteria of quality namely correctness and complexity. However, Krogstie et al. propose a semiotic framework for quality that distinguishes between syntactic, semantic, and pragmatic qualities (Krogstie and al., 1995). The syntactic quality addresses the structural relationships between modeling elements. The semantic quality is related to the relationship between the model and what it stands for (the domain). Finally, pragmatic quality considers the link between models and their interpreters. Our review of literature leads us to the conclusion that researchers mainly focused on syntactic quality aspects. Therefore, an effort on semantic and pragmatic quality definitions needs to be provided. Moreover evaluation is required. Our approach is a step forward to improve quality of BP modeling. Our contribution is threefold: 1) we define the semantic quality property for business process model evaluation; 2) we propose an approach based on domain knowledge to improve semantic quality of business process models, and 3) we enrich this approach by proposing guides to help BP modelers to improve their models and thus upgrade their semantic qualities.

STATE OF THE ART

Quality has been a subject of research in different disciplines. Accordingly, a variety of standards has been introduced to define, manage, measure, monitor, and improve quality. In particular, measurement allows the organization to improve its business processes. Considering quality early during the development process produces artifacts that are likely to be less error prone, easy to understand, maintain, and manage (Vanderfeesten, 2008). A business process is a set of related activities that take an input and transform it to create an ideal output (Johansson and al., 1993). The experts, both academic and professional, agree on the fact that the success of a company depends on its capability to understand its business processes (Aguilar and al., 2004). This explains the recent proliferation of proposals for BP quality management. We have identified three main directions in the literature: quality based on methodological guides, quality-based BP and quality-based BP model.

In the first direction, the approaches concentrate on providing advices and best practices to ensure the production of models with better quality. The underlying assumption is that the improvement of the development processes improves the quality of produced artifacts. In (Becker and al., 2000) the authors propose a set of guides for improving various characteristics of a process model such as clarity, comprehensibility or correctness of a model. Other authors focus on improving the comprehensibility of models by providing naming rules, documentation, and use of icons or symbols graphs (Mendling and al., 2010). Other approaches, such as (Van der alst and al., 2003), propose a set of best practices as design patterns, which can be reused in well-defined contexts.

The second trend considers how quality can be improved during execution of business processes. In this category we find simulation techniques and process control-based approaches such as (Van der alst and al., 2003) where the authors present a set of simulation tools for business process and their evaluation. Others focus on the verification of certain characteristics when executing the process. In (Jansen-Vuller and al., 2006) for example, the authors present and discuss several techniques such as verification and process mining enabling the analysis of the process during its execution, etc.

The third category deals with the quality of BP models. The proposals rely mainly on the definition of quality metrics. Krogstie et al. distinguish three quality levels for BP models: syntactic, semantic, and pragmatic (Krogstie and al., 1995). *Syntactic quality* aims to ensure syntactic correctness of models, measuring how far the model conforms to the syntactic rules of the modeling language. In (Van der alst and al., 2007) the authors present a typology and an overall view on BP metrics. They mention five most important measures: coupling, cohesion, complexity, modularity, and size. The work presented in (Gruhun and al., 2006) investigates the importance of structuredness for process model correctness by introducing complexity metrics. These metrics help in evaluating whether the model has an appropriate size and is clearly structured. Furthermore the authors in (Resnik, 1995) propose a set of structural metrics for the evaluation of the complexity of BP models expressed in BPMN[1] such as the number of activities, precedences, dependencies, etc. *Semantic quality* includes 1) validity verifying that all statements made in the model are correct and 2) completeness checking that the model contains all relevant statements about the domain. In (Soffer and al., 2004) the authors propose a generic theory-based process modeling framework as well as criteria for validity evaluation of process models. They discuss and characterize causes for process invalidity and suggest ways to avoid these situations. Also semantic similarity between BP models is explored in (Ehrig and al., 2007) addressing the problem of process models interoperability. Rozinat et al. in (Rozinat and al., 2005) describe an approach for measuring the compliance between events logs and process models. This compliance measurement indirectly enables the evaluation of BP model completeness. These BP models represent how the system should be used.

Finally *pragmatic quality* refers to the interpretation of the models. It measures the understandability of models by the stakeholders. Mendling et al. in (Mendling and al.2007) analyzed the factors that impact the BP model understandability and concluded that the latter should include the purpose of the BP model and the field of knowledge that it exploits. On the other hand, authors in (Abd Ghani and al., 2008) defined complexity metrics for BP models and proposed an approach based on GQM[2] for measuring understandability and maintainability of BP models.

1. Business Process Modeling Notation
2. Basili, V. R., Caldiera G., and Rombach H. D., "The Goal Question Metric Approach", in Encyclopedia of Software Engineering, Wiley, 1994.

As a conclusion, business process model quality is a very active research topic. However, the evaluation and improvement of semantic quality remain difficult and open issues. We believe that it can be improved by the introduction of domain knowledge. Our contribution is a step forward in this direction.

MOTIVATION AND APPROACH

Our research aims to propose an approach and a software utility for business process models quality improvement. Therefore we employed design science as the principal research methodology. Much of the Information systems (IS) research can be classified into either behavioral science or design science research. The former seeks to develop and verify theories that explain or predict human or organizational behavior whereas the later seeks to extend the human and organizational capabilities by creating new and innovative artifacts (Hevner and al., 2004). Modeling activity in general and BP modeling in particular are creative activities conducted by modelers using a given notation or modeling language. The result is of course highly dependent on the modeler experience in the notation practice, on his/her interpretation of the reality, and on the decisions he/she makes regarding the choice of concepts and details to be modeled. However, these models are supposed to be faithful representations of the reality. Thus the definition of quality requirements for these models is, in fact, a mean to evaluate this modeling activity and ensure a better result. Many quality factors may be defined to characterize this quality. The semantic quality measures the degree of correspondence between the model and the domain. We consider semantic quality as the conformance of BP models to the domain knowledge. Domain knowledge is fundamental to all disciplines and is defined as the knowledge of the area to which a set of theoretical concepts is applied (Alexander and al., 1992). According to (Khatri and al., 2006) in Information Systems discipline, domain knowledge comprises two parts: 1) IS domain knowledge representing knowledge provided by methods, notations, and tools and 2) application domain knowledge referring to real-world problems. The authors presented in (Khatri and al., 2006) a study demonstrating that IS development is aided by both richer IS knowledge and richer application knowledge.

In our target problem, the IS domain knowledge represents the knowledge related to BP modeling notations, methods, and practices. However, as several notations exist for BP modeling we propose to build our reasoning on a BP modeling meta-model to make it independent of any specific notation. On the other hand, the application domain knowledge can be extracted from several sources such as user's requirements statements, domain expertise or existing models related to the same problem or to similar problems. We suppose that existing models are available, thanks to other modelers representing different views.

The inputs for our approach are the business process model under construction and the related domain knowledge. Our approach comprises three steps:

- The first step performs the mapping of elements from the BP model with elements from the domain knowledge. The purpose is to identify fragments from the BP model for which domain knowledge could provide improvements.
- The second step is the evaluation of quality using metrics. We have defined some semantic quality metrics related to completeness and validity characteristics.
- The final step is the quality improvement activity whose purpose is to correct the quality defects detected by the evaluation.

The approach is iterative for a continuous and incremental quality evaluation and improvement.

REPRESENTING DOMAIN KNOWLEDGE

As proposed by (Khatri and al., 2006), we represent the domain knowledge with its two facets. This allows us to build a basis for semantic quality evaluation and improvement.

Capturing IS Domain Knowledge through Meta modeling

Our assumption about IS domain knowledge is that each modeling language has a precise semantics that is not always clearly formulated. Several approaches in the literature propose means to represent this semantics using a multiplicity of techniques: Petri nets (Eshuis, 2002), graph reduction techniques (Sadiq and al., 2000), syntax highlighting (H.A. Reijers and al., 2011) etc. We aim at exploiting such knowledge along with ensuring independence of a given BP modeling notation. In order to reach this independence, we use a meta-model for BP modeling and a set of rules capturing semantic constraints on concepts defined at the meta-model level. An extract from the process meta-model is represented at Figure1. This meta-model is adapted from several literature proposals (Loja and al., 2010,Eriksson 2000). It expresses the fact that a business process model is a diagram composed of flow objects, artifacts, and connectors. A flow object can be a gateway, an event or an activity. An activity can be atomic (task) or non atomic (process). A task is defined as a minimal traceable work that can not be split up. A connector can be an association, a sequence, or a message flow. Activities require resources that can be information or things. Things can be physical or abstract. People are physical resources, responsible for activities or owning the process.

We have enriched the meta-model with OCL (OMG, 2010) rules expressing semantic validation. For example, the fact that, in the context of process modeling, at least a human resource is required to manage the process, is a validation rule. The OCL rule is the following:

```
context Activity inv Direct_allocation: self.HR_allocation-> notEmpty()
```

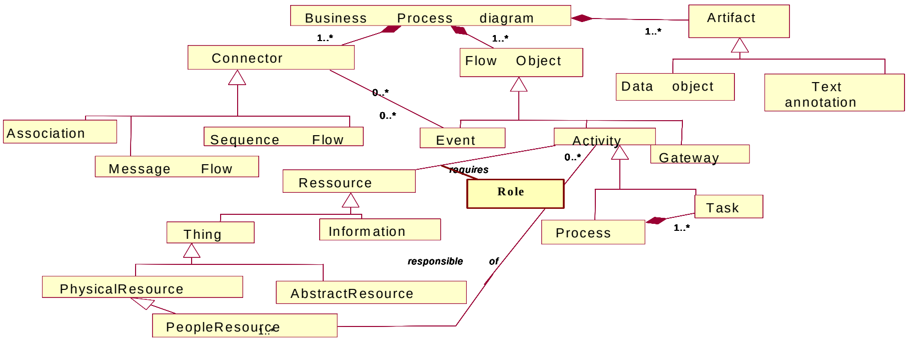

Figure 1. An extract from the BP meta-model

Capturing Application Domain Knowledge

Application domain knowledge covers 1) recurring and shared knowledge enabling the reuse of expertise and 2) application-specific knowledge required for development from scratch. We propose to capture the first category through domain ontology. The latter can be the result of expertise capitalization or standardization effort within a given domain such as the e-tourism ontology[3] for data exchange in tourism business domain. Such ontology can be built using requirements specification documents or requirements models. The definition of these ontologies is not in our scope. We suppose that they are available and we focus on their usage. Since our approach must fit ontologies and in order to formalize rules reasoning on them, our approach relies on a meta-model that serves as a structure for domain ontologies description. This meta-model is roughly depicted at Figure 2.

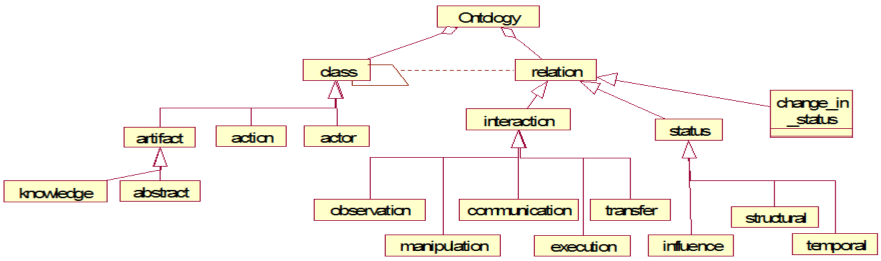

Figure 2. Ontology Meta-Model (extract)

An ontology is composed of classes and relations. For the class definition, we use the classification of entities (proposed in (Purao and al.,2005)) into actor, action and artifact. The relations are based on a classification adapted from (Purao and al.,2005). We consider three categories of relations namely: *status relationships* describing durable states induced by events such as structural relations(IS-A, part of, etc.), *interaction relationships* capturing, for example, communications among objects and *change in status relationships* which describe the transition life cycles. Examples from this category are intentions to do an action or attempts to perform an action, etc.

IDENTIFYING MODEL-ONTOLOGY SIMILARITIES

We formalize mapping rules by establishing links between the ontology meta-model and the process meta-model. This mapping is composed of two steps, namely *type-based mapping* aiming to avoid type mismatch errors, and *semantic mapping* based on the semantics of the concepts.

3. E-tourism ontology : http://sib.deri.ie/fileadmin/documents/e-tourism.owl

Type-based mapping

This mapping involves the types of concepts in order to establish correspondences between the concepts at the meta-level. These correspondences enable reconciliation based on the types of concepts independently of their meaning. These rules are very important to avoid typing errors. Similarly, we have established mappings between meta-model relations of BPM and those of the ontology meta-model (Table 1).

Table 1. Concept alignment

BP model meta-model concept	Domain Ontology meta-model concept
People resource	Actor
Abstract resource	Abstract
Information resource	Knowledge
Sequence flow	Temporal
Message flow	Transfer
Role	Execution, manipulation, observation or influence

Semantic-based mapping

The semantic based mapping is richer, being based on the semantics of concepts. Let O be a domain ontology and $o \in O$ a concept from this ontology. Each concept has a set of synonyms (set of words), hyponyms, hypernyms and keywords related to it. There are four classes of matching rules. The rules are all defined as functions having as input a BP model concept and returning one or several concepts from the domain ontology.
The similarity computation uses the names of concepts, the synonyms, and keywords associated to ontology concepts. It is based on WordNet and distance computation algorithms from literature such as Resnik information content (Resnik,1995), Wu & Palmer path length (Wu and al. 1994), Purandare & Pedersen context vectors (Purandare, 2004). We call these algorithms through equivalence functions (apply when the names are composed of one word) and partial equivalence functions (apply when names are composed of several words). Our approach uses five types of semantic similarity functions:

- Name based similarity: returns a set of ontology concepts having the same names or common names with the BP model concept.
- Synonym based similarity: returns a set of name ontology concepts having at least one synonym syntactically equivalent to the BP model concept. Additionally it returns ontology concepts that have semantic relatedness with the BPM element name based on the algorithms mentioned above.
- Hypernym similarity: based on the results obtained by name based and synonyms based similarity, this function returns for each ontology concept of the result the set of its hypernyms. It also returns the range of the "structural" relation for each ontology concepts.
- Hyponyms Similarity: based on the results obtained by name based and synonym based similarity, this function returns, for each ontology concept, the set of its hyponyms and the kind of "structural" relation for each ontology concept of the name based and synonym based results.
- Related concepts: returns a set of ontology concepts linked to the by a "require" or "assigned to" relation to a concept synonym of a BP model element.

EVALUATING SEMANTIC QUALITY

In order to evaluate semantic quality we have identified a set of what we call quality deficiencies. The latter result from modeling choices which produce models that do not meet the requirements or models with low expressiveness. Such models lead to inadequate systems due to incompleteness or to misunderstanding of the models during their implementation. According to (Wand and al., 1994), there are three generic categories of deficiencies related to representation of real word namely: ambiguous representations, incomplete representations, and meaningless states. These deficiencies have a direct impact on quality.

Ambiguous representations

Ambiguity results from using different names and constructs to express the same reality. This makes models unclear and creates confusion when trying to understand them. Additionally, it may also be due to a non adequate concept abstraction level. Indeed, in some cases, using general concepts instead of specific and precise ones can decrease the efficiency of the processes. On the contrary, using very specialized terms may decrease the understandability of the models. The relevant choice of an abstraction level depends on several factors among which we can mention the nature of audience (developers or users), the objective of the model (explanation or implementation), etc. We capture this aspect through three quality attributes:

Semantic clarity (c) = $1 / |synonyms(c)|$ if |synonyms(c)| ≠ 0 and *unknown* otherwise

Where c is a modeling element from the BP model and |synonyms(c)| is the number of concepts of the ontology belonging to the corresponding type (according to Table 1) considered as synonyms of c.

The more the concept has synonyms, the more ambiguous it is. When semantic clarity equals 1 for a given concept, this means that there is only one corresponding concept from the domain ontology and thus c is not ambiguous. The value of semantic clarity decreases when the number of corresponding elements in the ontology increases. When semantic clarity metric returns unknown, this means that there are no corresponding elements found for c. In this case, c can be a meaningless state. A validation from the modeler has to be performed since the domain ontology is not the exclusive source of knowledge.

Generality (c) = $1 / |hyponyms(c)|$ if |hyponyms(c)| ≠ 0 and unknown otherwise

Where c is a modeling element from the BP model and |hyponyms(c)| is the number of hyponyms of c. The more hyponyms the concept has, the more general it is. The generality value decreases when the number of hyponyms increases. When generality metric returns unknown, this means that there are no hyponyms found for c.

Specificity (c) = $1 / |hypernyms(c)|$ if |hypernyms(c)| ≠ 0 and *unknown* otherwise

Where c is a modeling element from the BP model and |hypernyms(c)| is the number of hypernyms of c.

The more hypernyms the concept has, the more specific it is. The value of specificity decreases when the number of hypernyms increases. When specificity metric returns unknown, this means that there are no hypernyms found for c.

Incomplete representations

Completeness is related to an incomplete representation of the real world. It occurs when the process model misses some requirements. This incompleteness can result from the complexity of concepts for which only a subset of the description is captured within the process model. It may also be due to requirements misunderstanding. This is generally captured by completeness quality attribute. There are several kinds of incompleteness. Incompleteness detection for an activity requires looking for missing or incomplete responsibility, documentation, input and/or output definitions. As an illustration, we propose two metrics for incompleteness evaluation:

- Responsibility completeness: detects activities (tasks and processes) that don't have an associated resource of type Personal Resource responsible for their execution.

Responsibility completeness(c) = |Human Resources (c)|
Where c is an activity element from the BP

|Human Resources (c)| computes the number of Human resources, responsible of *c* execution. If it equals zero, the analyst have to assign at least one human resource to *c*.

- Input completeness: helps detecting activities (tasks and processes) that could have incomplete input definition.

Input completeness (c) = |required (*oc*) ∩ input (c)| /|required (*oc*)|
Where c is an activity element from the BP and *oc* ∈ synonyms(c).

|input (c)| computes the number of resources required by activity *c*. *oc* designates an ontology concept validated as a synonym of *c*. |required (oc)| computes the number of concepts from the ontology related to oc by a "requires" relationship.

Meaningless states

Finally, meaningless states correspond to states and constructs from the models for which no correspondence is found in the real world. This decreases the relevance of models and has an impact on its intelligibility.

QUALITY IMPROVEMENT

The quality improvement activity consists in suggesting to the analyst or the quality expert a set of guidelines to improve the quality of their models. Below we list some quality improvement guidelines related to the quality attributes described above.

- Correcting ambiguity defects: consists in replacing the similar concepts in the model by a unique concept name. The ontology helps by providing the analyst with the list of synonyms and the analyst has to choose among them the most suitable term. Likewise, the user can rely on the knowledge provided by the ontology (Hypernyms/hyponyms) to choose the adequate abstraction level of the concept that will make the model more comprehensible.
- Correcting incompleteness defects assumes that the analyst can rely on the knowledge provided by the ontology to complete the missing parts of the model.
- In case of meaningless states defects, the analyst confirms whether the modeling element has no corresponding requirements. If so, he/she can discard the element from the model.

ILLUSTRATING THE APPROACH: AN EXAMPLE

To illustrate our approach we consider an initial BP model modelled in BPMN presented at Figure 3. This model is related to a patient care business process. When the patient arrives, the secretariat registers him/her by searching for his/her medical file or creating a new one. Then the nurse classifies the problem based on the symptoms reported by the patient and possibly on the laboratory exams results. The latter are notified to the medical team to be analyzed. At the same time the nurse keeps observations about the patient health status. Based on the transmitted results, the doctor gives his/her orders (medication, hospitalization, observation, or surgery). At the end of consultation, the patient pays. He/she buys prescribed drugs from pharmacy and the process ends.

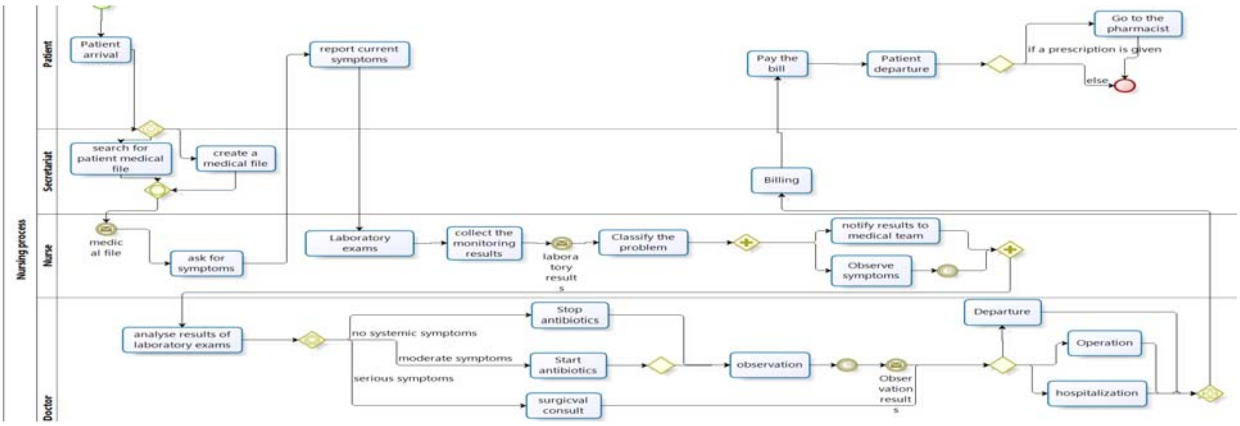

Figure 3. BP model from patient homecare case study

In addition our approach requires a domain ontology. An extract of the domain ontology "Clinical Ontology" is represented in Figure 4 It was built from collected clinical practices and documents. The ontology is an instantiation of the meta-model described at Figure 2.

A medical team member could be a "doctor", a "nurse", or a "pharmacist", each having specific responsibilities and abilities. In fact a nurse is responsible for "collect the results" and "laboratory exams" actions. Hence, it is related to these actions by a "perform" relation. "Laboratory exams" that could be "blood tests", "cholesterol tests" or "X-ray reports" are "clinical results". "Medical symptoms" concept is related to "fever", "diabetes" and "fatigue" concepts by a structural is_a relation. Doctors could be specialists ("surgeon", "pediatrician", "cardiologist"). A pediatrician for example has the ability to take care of babies etc. The ontology describes also details about patient's information (insurance company name, identification number etc.).

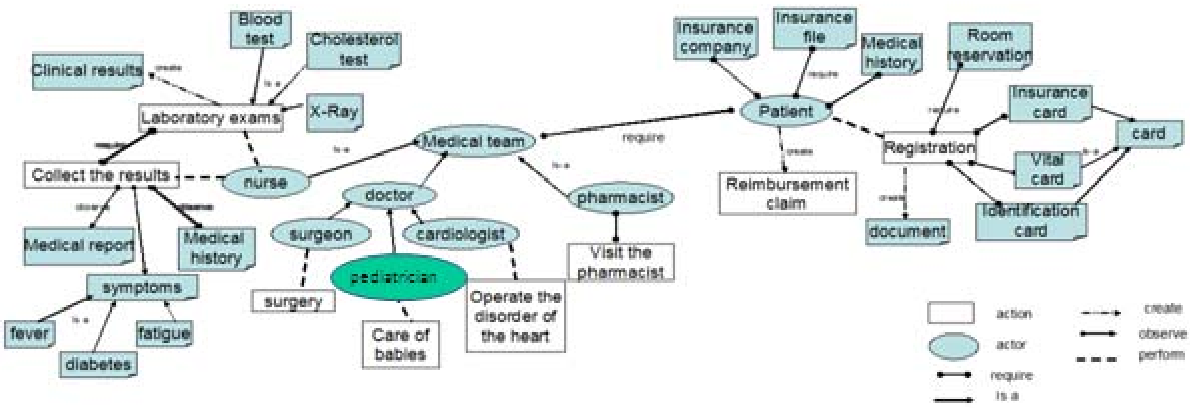

Figure 4. An extract from the clinical ontology

Similarities detection

The first step in our approach fires mapping rules between the BP model elements and the domain ontology concepts. Based on type and semantics mappings, our prototype computes a list of proposals. An extract of the results is listed in Table 2. Due to space limitations, only an extract is presented.

Table 2. Extract of mappings results

BP model concept	Domain Ontology concept	Synonyms/Hyponyms/ Hypernyms/keywords	Type of relation
Search for patient medical file	Investigate Patient's clinical history		
	Investigate Patient's family history		
Patient arrival	Room reservation, Registration, Enrollment		
Laboratory exams	Laboratory exams, Laboratory procedure, Medical report		
Medical file	Clinical results, Laboratory exams Registration	Clinical document, Medical report	is-a
	card	Insurance card, Identity card	is-a
Laboratory results	Patient exams	Notified results	hyponym
		Serious conditions, Mild conditions	is-a
		Symptoms	member of
		Blood test, Analyze hormone etc.	is-a
Patient arrival	Visit the pharmacy	Document	create
		Medical department	Transmit
		Insurance card, Identity card	Requires
Patient		Medical history, Insurance company	Requires
		Reimbursement claim	create
Go to the pharmacist		Medicine	assigned to
		Medical, medicinal drug	transmit

Quality defects detection

The second step is the evaluation of quality aiming at detecting semantic defects. The quality evaluation uses mapping results and quality metrics. Some quality values are listed at Table 3.

Table 3. Extract of quality values

Quality defect	BP model element	Quality value
Semantic clarity	Patient arrival	0.66
Semantic clarity	Collect the monitoring results	1
Generality	Symptom	0.099
Input incompleteness	Patient arrival	0

From semantic clarity point of view, "Patient arrival" action has three synonyms leading to 0.66 quality value. The action "Collect the monitoring results" has one synonym and has therefore 1 as semantic clarity value. Moreover, the concept "symptom" has more then eleven hyponyms. Generality quality metrics is equal to 0.09. This low value means that this concept is too general and should be replaced by one of its hyponyms suggested by the ontology. Finally, from completeness point of view, "patient arrival" activity should have as input "identification card" and "insurance card". Its input completeness value equals 0 accordingly.

Quality improvement

The last step of the approach tackles quality improvement according to the domain ontology knowledge. The correction actions are performed by the process modeler, supported by the approach. The resulting model is presented at Figure 5. In fact the patient's arrival requires an "identification card", and an "insurance file" based on the keyword function results. These resources enrich the model with added details. Also notice that the patient report contains a specific symptom "chest discomfort" based on one of the symptoms hyponyms. Then relying on this symptom, the doctor is replaced by a "cardiologist" responsible for the heart disorder. The activity "laboratory exams" is replaced by three more specialized exams as "blood test","cholesterol test" and "X-ray". That helps the analyst to choose the most adequate abstraction level. A temporal relation links "visit the pharmacy" and "create a claim" actions in the ontology, which allows the analyst to enrich his model by adding new activities such as "create a reimbursement" and "study the claim". This in turn leads to create a new actor "Insurance company".

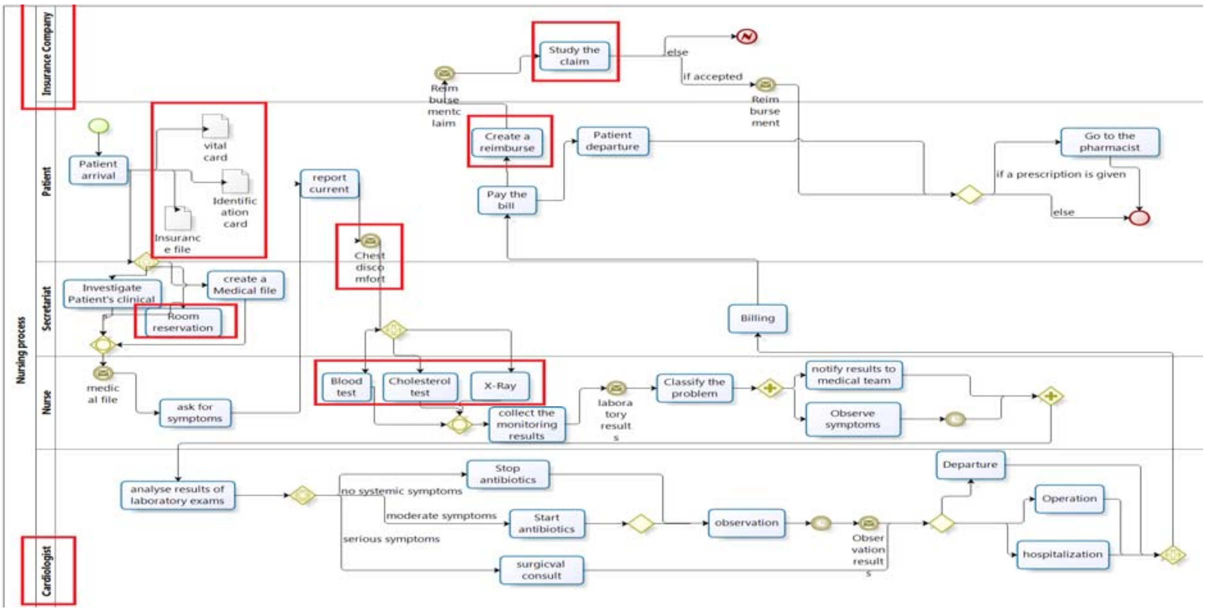

Figure 5. The improved BP model

CONCLUSION

The research presented in this paper is a step forward in semantics based quality evaluation and improvement using domain knowledge. One advantage of our approach is that it is not domain-specific or notation specific. The alignment process can be applied to each domain for which such an ontology is available. It may highly facilitate the task of BP modelers and lead to a significant improvement of BP models. The other advantage is that it encourages capitalization of expertise. Indeed, in many fields, there is an effort of definition of structured and shared knowledge in several areas: medical practices, HR processes, e-learning etc. The proposed approach is an actual usage of such knowledge. Finally, the quality measurement metrics provide the BP modelers with a quantified evaluation of the quality of their process models regarding a given domain ontology. The approach suffers from some limits. The main one is the fact the domain knowledge in hand is not necessarily complete. This is the reason why we propose to enrich the sources of domain knowledge with requirements and domain models. Further research will also include the enrichment of quality defects detection rules and the development of a metrics suite for semantic quality evaluation. We are also developing a prototype implementing the proposed approach for BP model quality evaluation and improvement. There is also need for more validation effort by conducting real-life case studies with practitioners.

REFERENCES

Abd Ghani A., Wei Tieng, K., Muketha G. and Wen, W. P. 2008. "Complexity Metrics for Measuring the Understandability and Maintainability of Business Process Models using Goal-Question-Metric (GQM)." Journal of Computer Science and Network security ,8, 219-225.

Aguilar-Savén R. S. 2004. "Business process modeling: Review and framework", International Journal of Production Economics, Volume 90, Issue 2, , Pages 129-149, ISSN 0925-5273.

Alexander, P. A. 1992." Domain knowledge: evolving themes and emerging concerns",Educational Psychologist, vol. 27, no. 1, pp. 33-51.

Becker J, Rosemann M. ,Uthmann C. V. 2000 "Guidelines of Business Process Modeling. Business Process

Management "

Blanchard E., Harzallah M.,Briand H. and Kuntz P. 2005. "A typology of ontology-based semantic measures", in proc of EMOI-INTEROP'05 workshop, at CAiSE'05

Cardoso J. Business Process quality Metrics. 2007 "Log-based Complexity of Workflow Patterns", In: OTM'07 Proceedings of the OTM Confederated international conference.

Ehrig M., Koschmider A., and Oberweis A. 2007. "Measuring similarity between semantic business process models". In Proceedings of the fourth Asia-Pacific conference on Conceptual - Volume 67 (APCCM '07), John F. Roddick and Annika Hinze (Eds.), Vol. 67. Australian Computer Society, Inc.

Eriksson H.E., Magnus Penker, 2000."Business Modeling With UML: Business Patterns at Work", John Wiley & Sons, Inc. New York, NY, USA ,ISBN:0471295515

Eshuis, H. 2002. "Semantics and Verification of UML Activity Diagrams for Workflow Modelling thesis". http://purl.utwente.nl/publications/64834 .

Gruhn V., Laue R.. 2006. "Complexity metrics for business process models" ; 9th international conference on business information systems , volume 85 of Lecture Notes in Informatics

H.A. Reijers, et al..2011. "Syntax highlighting in business process models", Decis. Support Syst.

Hevner, A.R., March, S.T., Park, J., Ram, S. 2004. "Design Science in Information Systems Research". MIS Quarterly 28(1), 75–105 (2004)

Jansen-Vullers M.and Netjes M. 2006. "Business Process Simulation: A Tool Survey". In Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools, Aarhus, Denmark.

Johansson H. J., McHugh P., Pendlebury A. J., Wheeler W. A. 1993. "Business Process Reengineering: BreakPoint Strategies for Market Dominance", John Wiley & Sons

Khatri V. , Vesseey I., Ramesh V., Clay P., 2006. "Understanding Conceptual Schemas: Exploring the Role of application and IS domain knowledge" In: Information Systems Research, vol 17, No.1, pp81-99.

Krogstie J., Lindland O. I., Sindre G. 1995. "Defining quality aspects for conceptual models". In Proceedings of the IFIP international working conference on Information system concepts, pp 216-231

Loja L., Neto V., Costa S., Oliveira J. 2010. "A business process metamodel for enterprise information systems automatic generation". In Brazilian Workshop on Model-Driven development. Brazil.

Mendling J., Recker J. and Reijers H.A. 2010. " On the Usage of Labels and Icons in Business Process Modeling" in IJISMD 1(2): 40-58

Mendling J., Rejers, Cardoso. 2007. "What Makes Process Models Understandable?" In: Lecture Notes in Computer Science, Volume 4714/2007, 48-63

Object Management Group: UML 2.2 OCL specification .2010. "Available as OMG document formal"

Purao S. and Storey Veda C. 2005."A multi-layered ontology for comparing relationship semantics in conceptual models of databases". Appl. Ontol. 1, 1 , 117-139.

Purandare A. and Pedersen T..2004. "Word Sense Discrimination by Clustering Contexts in Vector and Similarity Spaces", HLT-NAACL 2004 Workshop: 8th Conf. on Computational Natural Language Learning.

Resnik P.1995. "Using Information Content to Evaluate Semantic Similarity in a Taxonomy", International Joint Conference for Artificial.

Rolon E., Ruiz, Garcia , Piattini M. 2006. "Applying Software metrics to evaluate Business Process Models. " In: CLEIEl Electronic Journal, volume 9, number1, paper 5.

Rozinat A. and van der Aalst W.M.P.2005. "Conformance Testing: Measuring the Fit and Appropriateness of Event Logs and Process Models". In C. Bussler et al., (Workshop on Business Process Intelligence).

Sadiq W. and Orlowska M.E. 2000."Analyzing process models using graph reduction techniques." Information Systems, 25(2):117–134.

Soffer P., Wand Y. 2004. "Goal-driven analysis of process model validity". In: PROCEEDINGS OF CAISE'04

Tucker Taft Robert A.,Duff Randall L., Brukardt Erhard Pascal Leroy. 1998. "The International Standards Organisation ISO 8601"

Van der Aalst W.M.P, ter Hofstede A.H.M., Kiepuszewski B, and Barros and A.P. 2003."Workflow Patterns. Distributed and Parallel Databases".

Van der Aalst W. M. P. 2007." Challenges in Business Process Analysis" in ICEIS (Selected Papers).

Vanderfeesten I., Cardoso J., Mendling J., Reijers, Alast. 2007. "Quality Metrics for Business Process Models." In: Fischer, L. (ed.) BPM and Workflow Handbook, pp. 179-190.

Vanderfeesten I. , Reijers H.A., Mendling J. , Cardoso J. 2008. "On a Quest for Good Process Models: The Cross-Connectivity Metric". In Proceedings of the 20th international conference on Advanced Information Systems Engineering (CAiSE '08),

Wand Y. and Wang. R. Y. 1996. "Anchoring data quality dimensions in ontological foundations". Commun. ACM 39.

Wu Z. and Palmer M. 1994. "Verb semantics and lexical selection". In 32nd Annual Meeting of the Association for Computational Linguistics, pages 133–138.