Validation, Diagnosis and Decision-Making Support of Data in Business Processes

María Teresa Gómez-López(✉)

Departamento de Lenguajes y Sistemas Informáticos,
Universidad de Sevilla, Seville, Spain
maytegomez@us.es
http://www.idea.us.es

Abstract. Business processes involve data that can be modified and updated by various activities at any time. The data involved in a business process can be associated with flow elements or stored data. This data must satisfy the business compliance rules associated with the process, where business compliance rules are policies or statements that govern the behaviour of a company. To validate the correctness of a business process, it is necessary to validate the managed data before and during the process instantiation, since none of the activities of a process can work correctly using incorrect data. The analysis of the correctness of the business process is typically related to the activity executed according to the value of a data variable in each case, that verifies whether the model and the log conform to each other. The incorporation of the study of the correctness of the semantic of data value (called Business Data Constraints) is also consider essential. The execution of the correct activity according to the data is fundamental: it is no less important, however, to validate the correctness of the input data of a business process, and whether it affects the workflow and the compliance of the policies. In this paper, every special characteristic for the analysis of the correctness of data in a business process is studied, as is how the classic techniques can be improved for the validation, diagnosis and decision-making support concerning data in Business Processes.

Keywords: Business data constraints · Data validation · Data diagnosis · Decision-making data support

1 Introduction

Many companies have adopted Process-aware Information Systems (PAISs) to support their business processes in some form [1]. PAIS is a software system that manages and executes operational processes involving people, applications, and/or information sources on the basis of process models. It provides a way to manage data stored in a repository layer. The business process description includes a workflow model, a set of business rules or policies, and the data exchanged during the execution. The correctness of the business process implies

-

the correctness of these three aspects. The analysis of the correctness of business processes is typically related to the activity executed according to the value of a data variable in each case, by verifying whether the model and the log conform to each other [2]. The use of relational databases as a repository where the changed data is stored instead of log events is considered in [3], and papers such as [4] highlight the importance to validate the correctness of data in PAISs, although it remains a challenge as to how to find a fault in data instead of in a decision related to data. Therefore conformance checking analysis on log events is insufficient [5] to claim correctness in a business process.

We also consider it essential to incorporate the study of the correctness of the semantic of data values, especially in PAISs. The execution of the correct activity according to the data is fundamental: it is no less important, however, to verify the correctness of the input data of a process, and whether it affects the workflow and the compliance of the policies. It is crucial to validate the managed data before and during process instantiation, since none of the activities of a process can work correctly using incorrect data.

Although log events can include data information, the natural place to store data is in a database. Data must satisfy business process policies in terms of their possible values and relations during the process execution, and for every instance. The study of the evolution and change of data produced by a system has been addressed in various domains by using model-based diagnosis [6], although it cannot be applied directly in a business process since a BP has special characteristics that need to be analysed in order to adapt the classic techniques. These techniques must be improved to provide for the validation, diagnosis and decision-making support of data. The analysis of correctness of data in business processes holds a special complexity, since the data used is not always produced by the business process itself, the fault can be intermittent, and the data can be aggregated and used in different instances.

The study of data correctness according to the process implies several decisions, such as:

- **When should the problem be studied?: At design time, runtime or post-mortem time?** The data model correctness is essential to ascertain whether the data model is correct [7]. Furthermore, since, in business processes data is continuously created, updated or eliminated, a validation analysis is necessary to preserve high data quality at runtime. It is especially important in systems where there is a high degree of exchanged data, since it is necessary to maintain aspects such completeness and correctness according to the business goals.
- **What analysis is needed?: Validation, diagnosis or decision-making?** Depending on the aim of the data analysis, different parts of the model and data are necessary. The validation of an instant implies to analyse the current data, while the diagnosis include to study the data evaluation in the past, and the decision-making support provokes a reasoning of the possible paths that an instance can execute in the future.

– **What is the objective of the analysis?: Data, activities, compliance rules or a combination of a number of objectives?** Depending on the type of business process and the latency of change that a BP supports, the possible causes responsible for a malfunction can differ. For example, for very a mature process, the cause of a malfunction is typically the introduced data, although if the policies are very changeable, they can be a source of the problem, or if the process has a high level of human interaction, it is more likely that the input data is the responsible for a malfunction.

The rest of the paper is organised as follows: Sect. 2 depicts the components of a business process model where the stored data and compliance rules are analysed. Section 3 explains the various characteristics that need to be adapted from classic model-based diagnosis to the business process area. Section 4 explains the differences between the various types of analysis (validation, diagnosis and decision-making support). Section 5 details how the above-mentioned objectives are performed. Finally, conclusions are drawn and open issues are explained in Sect. 6.

2 Modelling the Business Processes for Correctness Analysis of Data

In order to compare the expected behaviour of the system with the real behaviour, it is necessary to design the process and to extract the observations produced. This implies to consideration of a pair $\{SD, OBS\}$, where SD is the system description and OBS is a set of values of the observable variables. The process model is formed of a workflow, a set of business rules that describe the semantic of the data values, and a relational database schema. In the following subsections, these parts of the model are given in detailed, together with the observational model extraction.

2.1 Process Model Workflow

The elements that can conform a process model to support data validation or a diagnosis process are:

– $\mathbb{SE}$, one start event to initialize the process.
– $\mathbb{EE}$, a set of end events, with at least one element.
– $\mathbb{A}$, a set of activities that defines the model of the process.
– $\mathbb{CF}$, a set of control flow patterns (AND, OR, XOR) that describes the possible branches to execute.
– $\mathbb{C}ond$, a set of conditions associated with the control flows OR and XOR, that describes the paths that the process can take depending on the values of the variables in the dataflow. These conditions are evaluated at runtime, when the values of the variables are known.

2.2 What Are Business Data Constraints?

The compliance rules which represent the semantic relation between the data values are called Business Data Constraints. BDCs are used to describe the possible correct data values that are introduced, read and modified during the business process instances [8]. BDCs are a fundamental part of the model to be validated or diagnosed, since they describe the correct value relations between the involved data. BDCs are presented as Numerical Constraints that involve stored information and dataflow, and can be expressed as a Boolean combination of arithmetic constraints for Natural, Integer and Float domains, following the grammar introduced in [9], where it is important to highlight that the variables participating in the constraints can come from the database or from the dataflow.

Regarding to how the rules are integrated in a framework to complement the business process model, in [10] there is an in-depth analysis concerning the integration of rules and process modelling and the shortcomings of existing solutions. The BDCs can be defined as either an invariant of the process, and therefore associated to the whole process, or as a contract (pre- or post-condition) for an activity or a set of activities. To support the storage, extraction and query of these constraints, the Constraint Database LORCDB is used [11,12]. The example shown in Fig. 1 presents a process example where BDCs are associated as pre-conditions, post-conditions, and invariants of the model.

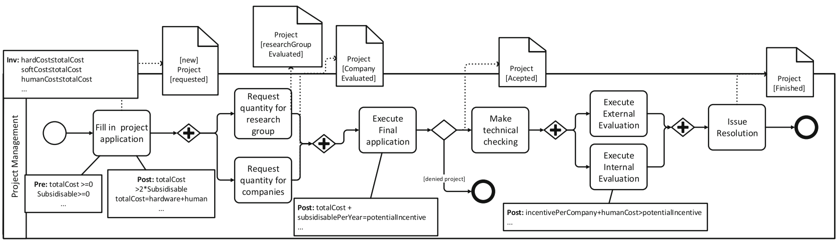

Fig. 1. Example of a process with Business Data Constraints and Data State Objects

2.3 Relational Database Model

A Relational Database is a collection of predicates over a finite set of variables described by means of a set of relations. A relation R is a data structure which consists of a heading and an unordered set of tuples which share the same type, where A_1, A_2, ..., A_n are attributes of the domains D_1, D_2, ..., D_n.

The set $\{A_1{:}D_1,\ A_2{:}D_2,\ \ldots,\ A_n{:}D_n\}$ is a relational-schema. A relation R defined over a relational-schema S is a set of assignments for each attribute for each domain. Therefore, the relation R is a set of n-tuples:

$\{A_1{:}d_1,\ A_2{:}d_2,\ \ldots,\ A_n{:}d_n\}$, where $d_1 \in D_1, d_2 \in D_2, \ldots, d_n \in D_n$.

2.4 Observational Model

The most important data changes in a business process are made persistent in a relational database. For this reason, the observational model contains the tuples of the database that refer to the variables included in the BDC. Depending on how these variables are introduced into the process models and hence into the database (such as input variables, triggered in the database, or derived by an activity), they play different roles in the diagnosis processes. These roles include:

- **Input Variable.** Variable whose value is introduced by the user in one of the activities of the business process model. It is a possible origin of problems if the data is incorrect.
- **Derived Variable.** Variable whose value depends on the value of other variables (input or derived variables). This means that its value is not introduced as an input, but it is calculated/created from other values. Therefore a derived variable cannot be responsible for a malfunction itself, although it can be involved in an incorrect BDC or it can be produced from incorrect input variables.
- **Key variables.** Set of variables that differentiate one tuple from another. They correspond to the primary key attributes of the relational model. These variables form a very important part in the diagnosis process since they enable one tuple to be isolated from another, but can seldom be determined as the origin of a problem.
- **Query variables.** Set of input variables and subset of key variables used to describe the set of tuples introduced and/or modified in an activity. These variables depend on the activity, and on the moment when the validation/diagnosis analysis is executed. These variables can be transferred through the activities as dataflow while they do not change in a process instance.

  A tuple is a set of attributes $\{A_1, A_2, \ldots, A_n\}$, Input Variables (IV), Derived Variables (DV), Key Variables (KV) and Query Variables (QV), and the properties of these sets of variables are:
  - $\{A_1, A_2, \ldots, A_n\} = \text{IV} \cup \text{DV} \cup \text{KV} \cup \text{QV}$
  - $\text{QV} \subseteq \text{KV}$
  - $\text{QV} \subseteq \text{IV}$
  - $\text{IV} \cap \text{DV} = \emptyset$

3 Special Characteristics of Data Management Problems

Although business process analysis fortunately does not consider component degradation or deterioration as physical systems, it does contain other complexities and special characteristics, the most important of which are listed in the following subsections.

3.1 High Level of Human Interaction and Intermittence of Faults

None of the activities of a process can work correctly using incorrect data. Although activities may work correctly, the people that interact with them can work incorrectly and introduce incorrect values. This is especially important when there is a high level of human interaction, since humans can introduce intermittent faults in the system. The intermittence of faults makes the detection and the diagnosis of faults difficult, since an inconsistence detected during the execution of an activity does not imply a malfunction in the activity, or that this fault may appear in the future.

Some typical examples where intermittent faults in data can be found are:

- Financial applications, where several items of data are introduced by hand from application forms.
- Medical applications that use data introduced by different types of medical staff and places, for example data concerning blood test results.
- Electronic equipments, which produce a major quantity of data, that can cause intermittent faults, such as loose or corroded wire wraps, cracked solder joints and broken wires.

3.2 High Use of Stored Data

As mentioned before, data involved in the diagnosis is not strictly limited to that flowing in the process, some data is also stored in databases. This implies that the quantity of data involved in an instance can be very large, and problems of poor data quality must be detected and determined.

Some of the attributes of a relation can be described as *Primary Key Attributes* which means that "two tuples of a relation cannot have the same values for their primary key attributes". The relation between two tables is described by referential integrity. Two tables can be related by means of their *Primary* and *Foreign Key Attributes*, described in the literature as the relational model. Referential integrity is a database concept which ensures that relationships between tables remain consistent. When one table has a foreign key to another table, the concept of referential integrity states that a record may not be added to the table that contains the foreign key unless there is a corresponding record added to the linked table. The division of the related data into several tables helps in data integrity but makes the validation of the relation of the data values more difficult, since they are disseminated in various tables. The data involved in a BDC can therefore be found in various tables, and it is necessary to *denormalize* the information using the primary and foreign keys.

Moreover, the relational model is very detailed and is therefore difficult to understand and query by non-expert users, because it is isolated from how the data objects are modified during the process execution. BPMN 2.0 [13] and further improvements of its capacity to describe the data stored states give the capacity to enrich the model [14,15]. Thanks to these proposals, instead of a relational model, the Conceptual Model and Object-Relational Mapping (ORM) are used to facilitate the description and management of the stored business data objects, as shown in Fig. 1.

3.3 Combination of Business Rules and Data

Data is not the only aspect that can be modified during a business process execution. The compliance rules that describe the policies and the goals of the process also tend to be modified and updated to represent the necessities of the companies. The changeability of the policies makes the process adaptable to new conditions at runtime. For this scenario, the question becomes What type of element of the model can be held responsible for a malfunction? Only activities, only data, only business rules, or a combination of data and business rules?

Suppose that there are three different possible diagnoses, such as: 15 input data; 3 BDCs; or 2 BDCs and 5 input data at the same time. It is not possible to ascertain which is the most probable fault, since this can depend on:

- **The number of introduced items of data.** For example, if 16 items of data are introduced from the last validation or diagnosis, it is high improbable that 15 values are incorrect, but it is less improbable when $10,000$ items of data have been introduced.
- **Are the rules tested sufficient?** When a business rule has just been included in a model, it is likely that is has been insufficiently tested, and therefore it is more probable that other rules are incorrect.
- **Who has introduced the data?** Not all people in a company are equally reliable, and therefore this is an important factor when deciding between two possible diagnoses.
- **Who has introduced the rules?** A similar thing occurs with the rules, since different users can be more trustworthy when introducing rules.
- **The complexity of typing the data or rules.** Regardless of the person who introduces the process information, not every item of data or rule has the same complexity to be introduced, for example it is easier to introduce a fault into a number of 20 digits than into a number of 2 digits.

3.4 Data Shared Between Cases

The data of an instance is not independent from the data of other instances, since the same relational database is shared. This implies that the same data and rules can be involved in different validation or diagnosis processes for different instances. This implies that the final diagnosis needs to consider that the explanation of a malfunction in a case must also explain the good or wrong behaviour in the others.

For this reason, in a diagnosis process, it is necessary to incorporate every items of data involved in the independent-variable cluster. An independent-variable cluster is a set of BDCs whose variables are not involved directly or indirectly in another independent-variable cluster, thereby assuring that every variable involved is taken into account in the problem. A formal definition is:

Definition: Cluster of BDCs with Independent Variables. Let BC represent all the BDCs of a process, let B be a set of BDCs where $B \subseteq BC$, and let $V(B)$ be the set of variables involved in B. Therefore B is a Cluster of BDCs with Independent Variables iff $V(B) \cap V(BC{-}B) = \emptyset$, and B is minimal, which implies that $\nexists / B' \subset B \mid B'$ is a Cluster of BDCs with Independent Variables.

3.5 Data Uncertainty and Exoneration Principle

In the tables that form the database, it is usual to find unknown values during a process instantiation (*null* information in the database). This means that this data has not yet been introduced, because the instance has not finished or the data is not always necessary. This uncertainty needs to be included in the model to be validated or diagnosed, and this complicates the process since the exoneration principle, typically used in classic diagnosis, cannot be taken into account. The exoneration principle determines that a satisfiable BDC cannot contain incorrect variables and be correct itself at the same time. For business processes, this principle is not applicable since variables can be uninstantiated (*null* values), and the BDCs do not always determine strong relations between the variables (such as equal relations), since they can use soft data variable relations, such as $<$. For this reason, it is only possible to deduce incorrectness of BDCs for the tuples, but not correctness of BDCs applied to specific data.

4 Past, Present and Future of the Data in a Process Instance

The model described in Sect. 2 can be used for various purposes: the analysis of the validation of the model according to the data stored and data life-cycle (Subsect. 4.1); the diagnosis of the cause responsible for a found malfunction (Subsect. 4.2); or aid in the data introduction in the process to prevent future errors (Subsect. 4.3).

4.1 Data Validation

It is extremely complicated to determine how companies can adequately integrate their business objects stored in databases into a process model. Furthermore, it is possible to find non-conformances between the data model and the evolution of the states included in the workflow model [15]. The importance of data life-cycle validation has been the focus of several papers, where the data evolution during the business execution has been modelled as artifacts [7,16,17]. Artifact-centric proposals permit verification of the data model correctness according to their pre- and post-conditions, and the study of the possible data evolution. Artifact-centric solutions are especially focused on complex data models, such as those for the $n : m$ relation between objects [18].

In order to ascertain whether the data model and the stored data are consistent with the business process model, a systematic analysis is necessary on what activity modifies the state of the data, and which is the state of the stored data. It implies determining the relation between the data objects stored in the database and the activities that compose the business process model that modifies their states. However it is not only important to analyse the business data life-cycle, since the states of the objects are modified for activities of the model, and therefore the data life-cycle needs to be consistent with respect to the workflow. The possible states that an object can satisfy, and where it is read or written must be analysed. Certain analyses must assure that every object is in at least one state, or prevent that an object from satisfying two incompatible states.

4.2 Diagnosis to Find the Origin of the Problem

Not only is it crucial to ascertain that something is wrong, it is also essential to diagnose where the problem is in order to solve it and for it to be prevented in the future. Another major issue in diagnosis is to find the minimal diagnosis: the minimal possible number of conflicts that explain a malfunction in an efficient way. Minimal explanation of the problem is desirable, since model-based diagnosis is based on the parsimony principle [19]. It states that, among competing hypotheses, that with the fewest assumptions should be selected. A malfunction is visible when a discrepancy between the expected behaviour and the observed model is found. The objective is to determine what minimal part can explain the problem.

In relation to the persistence layer and dataflow, relational databases have been used in the business process, for example in [20] which presents a solution where data is audited and stored on a relational database. However, no validation of the semantics is performed for this persistence layer and the business rules. In papers such as [21], the necessity to resolve a fault after detection is identified, unfortunately the data aspect is not included. In previous papers, the possibility of finding incorrectness in input data is introduced [9,22], although many challenges related to this problem are ready to be resolved as mentioned in the section future work.

4.3 From Data-Based Decisions to Decision-Making About Data

Several research studies have been published to improve the decision-making support during a business process execution. These decisions tend to be related to the execution of an activity analysing the values of the data involved in the process [23–25]. But when the goal of a process is oriented towards the decision-making support about data, instead of concerning the decision based on the data value of a process instance, we are faced with a completely different problem. The decision-making support about data is especially important when the information that flows between the activities can be introduced by the users. The user has sometimes to decide on which value to introduce while taking into account the potential actions in order to make the process instance correct [26], or for process outcome optimization [27]. In the business process scenario, this implies the analysis of all the possible branches that can be executed, and the

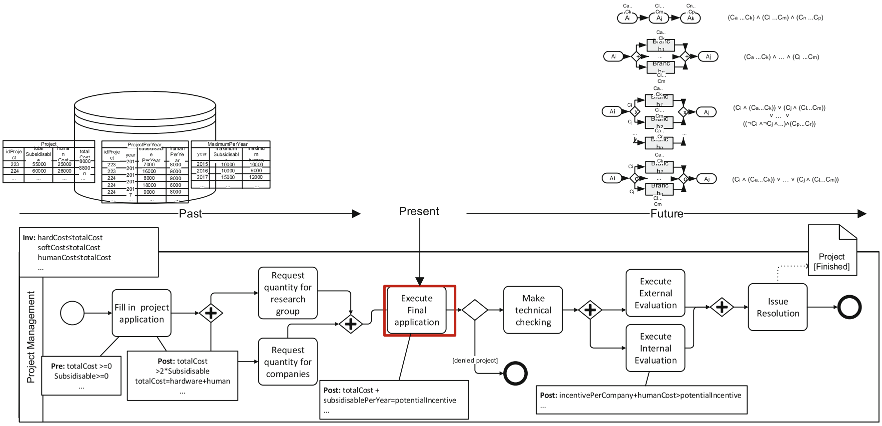

Fig. 2. Diagnosis, Validation and Decision-making in BP

decisions that can be taken in the future. If the decision made is incorrect, it will affect other decisions in the future, or it may even make it impossible to finish the instance correctly. The requirement for decision support always arises when decisions have to be made in complex, uncertain, and/or dynamic environments. From the point of view of input data values, the correctness of a business process is based on the correctness of the compliance rules that describe the policy of the company. For this reason, we propose a solution where the decision-making support for input data can be integrated into the business process instances, to inform the user about the possible values of the input variables, thereby rendering the instance of the business process consistent.

Figure 2 graphically depicts what parts of the model are used in accordance with the Validation analysis (Present), Diagnosis (Past including stored data), and Decision-making about data (analysing every possible paths that an instance can take in the future). In the three contexts, business data constraints are needed, and are incorporated into the analysed model.

5 Solving the Challenges

In order to transfer the capacity to manage data as a fundamental part of the business process enactment, it is necessary to propose a methodology that supports the solution of the problems introduced. The steps that are followed to extract the necessary part of the model (Fig. 3), and the observational model include: (1) Building a referential integrity graph that relates the data involved in the database to ascertain the independent cluster variables; (2) Extracting the BDCs involved in each case; (3) Creating the observational model joining the tuples in accordance with the stored data and the BDCs; and (4) Creating and solving using Constraint Programming Techniques [28] and a Max-CSP to find the minimal-conflict sets [9], or creating a CSP to validate the current process

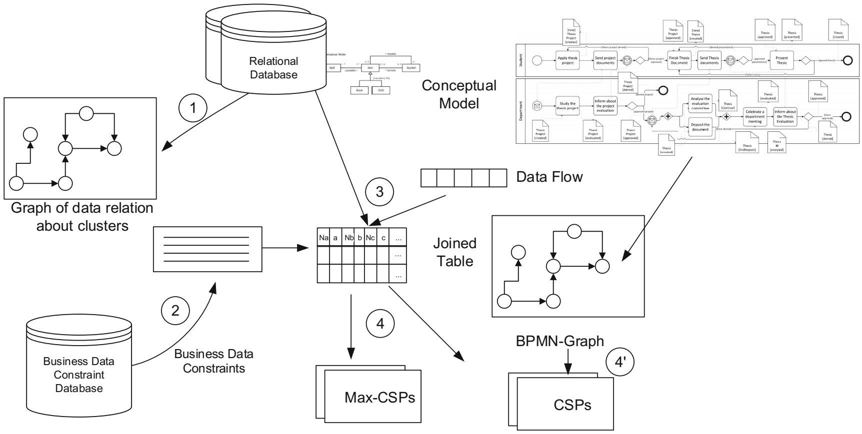

Fig. 3. Steps to evaluate the stored data

state. If decision-making support is performed, in Step 4′, a BPMN-Graph [29] needs to be included to ascertain the possible paths of the instance in the future.

6 Conclusions and Future Work

This paper emphasises the importance of changing the point of view in business processes: from data-based decisions to decisions about data. It implies the validation, diagnosis and decision-making in business processes where the aim of the analysis is data-centric, especially when there is a database with a great quantity of shared information. For this reason, the required composition of the model (workflow, database schema, BDCs and data states) it is detailed, and the observational model is extracted from the relational database. The special characteristics of data management for validation and diagnosis have been listed, and the steps to perform the different scenarios are presented. Several open issues have also been included in the paper, such as: the possibility of combining data and compliance rules at the same time as making a minimal diagnosis; the necessity to validate the data model of the stored information and the workflow; the uncertainty related to the unknown variables produced by the *null* values; or the challenge that supposes the incorporation of the trustworthiness of the users into the model.

The difficulties of stored data management have been applied to local relational databases, but obviously the complexity increase exponentially with big data, that implies volume, variety, velocity, and veracity of data sources.

Acknowledgement. This work has been partially funded by the Ministry of Science and Technology of Spain (TIN2015-63502-C3-2-R) and the European Regional Development Fund (ERDF/FEDER).

References

1. Weber, B., Sadiq, S.W., Reichert, M.: Beyond rigidity - dynamic process lifecycle support. Comput. Sci. R&D **23**(2), 47–65 (2009)
2. van der Aalst, W., Adriansyah, A., van Dongen, B.: Replaying history on process models for conformance checking and performance analysis. Wiley Int. Rev. Data Min. and Knowl. Disc. **2**(2), 182–192 (2012)
3. van der Aalst, W.M.P.: Extracting event data from databases to unleash process mining. In: Brocke, J.V., Schmiedel, T. (eds.) BPM - Driving Innovation in a Digital World. Management for Professionals, pp. 105–128. Springer, Heidelberg (2015)
4. Ly, L.T., Rinderle-Ma, S., Dadam, P.: Design and verification of instantiable compliance rule graphs in process-aware information systems. In: Pernici, B. (ed.) CAiSE 2010. LNCS, vol. 6051, pp. 9–23. Springer, Heidelberg (2010)
5. Rozinat, A., van der Aalst, W.M.P.: Conformance checking of processes based on monitoring real behavior. Inf. Syst. **33**(1), 64–95 (2008)
6. Ceballos, R., Gómez-López, M.T., Gasca, R.M., Del Carmelo, V.: A compiled model for faults diagnosis based on different techniques. AI Commun. **20**(1), 7–16 (2007)
7. Borrego, D., Eshuis, R., Gómez-López, M.T., Gasca, R.M.: Diagnosing correctness of semantic workflow models. Data Knowl. Eng. **87**, 167–184 (2013)
8. Gómez-López, M.T., Gasca, R.M.: Run-time monitoring and auditing for business processes data using contraints. In: International Workshop on Business Process Intelligence, BPI 2010, pp. 15–25. Springer, Heidelberg (2010)
9. Gómez-López, M.T., Gasca, R.M., Pérez-Álvarez, J.M.: Compliance validation and diagnosis of business data constraints in business processes at runtime. Inf. Syst. **48**, 26–43 (2015)
10. Muehlen, M.Z., Indulska, M.: Indulska.: Modeling languages for business processes and business rules: A representational analysis. Inf. Syst. **35**(4), 379–390 (2010)
11. Gómez-López, M.T., Gasca, R.M.: Using constraint programming in selection operators for constraint databases. Expert Syst. Appl. **41**(15), 6773–6785 (2014)
12. Gómez-López, M.T., Ceballos, R., Gasca, R.M., Del Valle, C.: Developing a labelled object-relational constraint database architecture for the projection operator. Data Knowl. Eng. **68**(1), 146–172 (2009)
13. OMG. Object Management Group, Business Process Model and Notation (BPMN) Version 2.0. OMG Standard (2011)
14. Meyer, A., Pufahl, L., Fahland, D., Weske, M.: Modeling and enacting complex data dependencies in business processes. In: Daniel, F., Wang, J., Weber, B. (eds.) BPM 2013. LNCS, vol. 8094, pp. 171–186. Springer, Heidelberg (2013)
15. Gómez-López, M.T., Borrego, D., Gasca, R.M.: Data state description for the migration to activity-centric business process model maintaining legacy databases. In: Abramowicz, W., Kokkinaki, A. (eds.) BIS 2014. LNBIP, vol. 176, pp. 86–97. Springer, Heidelberg (2014)
16. Hull, R.: Artifact-centric business process models: brief survey of research results and challenges. In: Tari, Z., Meersman, R. (eds.) OTM 2008, Part II. LNCS, vol. 5332, pp. 1152–1163. Springer, Heidelberg (2008)
17. Eshuis, R., Kumar, A.: Kumar.: An integer programming based approach for verification and diagnosis of workflows. Data Knowl. Eng. **69**(8), 816–835 (2010)
18. Borrego, D., Gasca, R.M., Gómez López, M.T.: Automating correctness verification of artifact-centric business process models. Inf. Softw. Technol. **62**, 187–197 (2015)

19. Peng, Y., Reggia, J.A.: Abductive Inference Models for Diagnostic Problem-solving. Symbolic computation. Springer, Heidelberg (1990)
20. van der Wil, M.P., van der Aalst, W.M.P., van Hee, K.M., van der Werf, J.M.E.M., Kumar, A., Verdonk, M.: Conceptual model for online auditing. Decis. Support Syst. **50**(3), 636–647 (2011)
21. Awad, A., Smirnov, S., Weske, M.: Towards resolving compliance violations in business process models. In: Proceedings of the 2nd International Workshop on Governance, Risk and Compliance - Applications in Information Systems, Amsterdam, The Netherlands, June 2009
22. Gómez-López, M.T., Gasca, R.M.: Fault diagnosis in databases for business processes. In: 21th International Workshop on Principles of Diagnosis, pp. 1–8 (2010)
23. Conforti, R., de Leoni, M., La Rosa, M., van der Aalst, W.M.P., ter Hofstede, A.H.M.: A recommendation system for predicting risks across multiple business process instances. Decis. Support Syst. **69**, 1–19 (2015)
24. van der Aa, H., Leopold, H., Batoulis, K., Weske, M., Reijers, H.A.: Integrated process and decision modeling for data-driven processes. In: 3th International Workshop on Decision Mining & Modeling for Business Processes, DeMiMo, pp. 15–25. Springer, Heidelberg (2015)
25. Bazhenova, E., Weske, M.: A data-centric approach for business process improvement based on decision theory. In: Bider, I., Gaaloul, K., Krogstie, J., Nurcan, S., Proper, H.A., Schmidt, R., Soffer, P. (eds.) BPMDS 2014 and EMMSAD 2014. LNBIP, vol. 175, pp. 242–256. Springer, Heidelberg (2014)
26. Gómez-López, M.T., Gasca, R.M., Pérez-Álvarez, J.M.: Decision-making support for the correctness of input data at runtime in business processes. Int. J. Coop. Info. Syst. **23**(2), 29 (2014)
27. Parody, L., Gasca, R.M., López, M.T.G.: Hybrid business process modelling for the optimization of outcome data. Inf. Softw. Technol. **70**, 68–70 (2015)
28. Rossi, F., van Beek, P., Walsh, T.: Handbook of Constraint Programming. Elsevier, New York (2006)
29. Weber, I., Hoffmann, J., Mendling, J.: Semantic business process validation. In: 3rd international workshop on Semantic Business Process Management (2008)