Towards Goal-Oriented Conformance Checking

Hiroki Horita, Hideaki Hirayama, Yasuyuki Tahara and Akihiko Ohsuga
Graduate School of Information Systems
The University of Electro-Communications
Tokyo, Japan
Email: h-horita@ohsuga.is.uec.ac.jp, hirayama968@ybb.ne.jp, {tahara,ohsuga}@is.uec.ac.jp

***Abstract*—Constructing a business process is important area between requirements engineering and business process management. Goal-oriented requirements analysis method is widely researched in requirements engineering and useful for reflecting organizational requirements to business process models, but actual business processes deviate from defined process models. Therefore, it is not sufficient for business process analysis only using model's information. It is important to analyze actual conducted business process logged data. Analyzing business process logged data is called process mining and detecting differences between models and logs is called conformance checking. A lot of conformance checking approaches mainly focus on process aspects of business process, but this is not sufficient for analysis whether actual business processes can satisfy organizational goals. In this paper, we propose a goal-oriented conformance checking approach which can detect deviations between logs and models, and can analyze the effects of the deviation. It is useful for evaluation of the detected deviation. We represent the effectiveness of our approach conducting a case study using the publicly available log.**

I. Introduction

In recent years, problems of business process complexity and rapidly changing business environments are needed to deal with. In that situation, constructing business process is used for discussion, verification, documentation and etc [1]. Using goal models are effective against constructing appropriate business process models. Goal models are researched in requirements engineering area, and used for requirements analysis [2]. Goal models have systematic and logical construction for representing requirements for information systems. These characteristics are useful for reflecting organizational requirements to business process models.

Using goal models for constructing business process models is effective, but it is not sufficient for organization, because acutual business often deviate from defined business process models [1]. In that situation, model-based analysis is not adequately effective. Therefore, in recent years, analyzing business process logged data is widely researched and it is called process mining. In process mining, conformance checking between models and logs are important topics. Conformance checking can detect deviations between normative models and actual logged data (it is called event logs). When the models and reality (logged data) have little in common, model-based analysis does not make much sense [1]. Therefore, analyzing logs and improving models are important.

Quite a lot of conformance checking researches are conducted. At first, these approaches only focus on control-flow perspective afterward, data (event related information) and resource (agent conducting the event) perspectives are focused. These approaches can check various perspectives relating business process. These perspectives are important, but it is not sufficient yet. Using these approaches against logs and models represents deviations between them, but it is not concrete what should we do for interpreting deviation and improving business process.

In this work, we propose a goal-oriented conformance checking approach. Goal-oriented aspects are effective against deviation interpretation. Goal models have systematic and logical construction. Therefore, it is possible to represent what is important in a business process and encourage efficient decision making. Our approach is constructed by two phases. The first phase checks deviation between a goal model and logs focusing control-flow, data and resources. Goal models are described using linear temporal logic, so verification is formally conducted. In addition, in case of deviation detected, goals are combined with a goal using goal model construction for evaluating the deviation. Therefore, a cross tabulation table relating these goals is constructed. In the second phase, we calculate significant difference between two goals in a cross tabulation table. Next, if these goals have significant differences, a relation between two goals are positive or negative is calculated. In this way, deviations are detected and evaluated the effects of the deviation. It is useful for interpretation of deviations between processes and logs.

II. Goal Model and Business Process Management

Goal models are used for requirements elicitation, evaluation, negotiation, eraboration, structuring, documentation, analysis and evolution [2] for system development. Goals should be achieved and refined into subgoals through AND/OR decompositions.

Relating goal models and business process models are important. Various researches are conducted in this area (e.g. transformation:[3],[4], validation: [5], integration: [6]). These approaches are mainly used in business process construction phase. It is corresponding to diagnosis/requirements and (re)design phases in the business process life cycle [1]. These phases are important, but business processes are life cycle and improved at various times. Therefore, it is needed to confirm desirable business process that can achieve organizational goals. In many cases, actual processes deviate from the normative business process model. In these cases, it is difficult to know the effect of deviations from business process models using only model information because models only have normative information. Therefore, it is needed to use event logs of business process and to confirm event logs can

achieve business goal or not. These are enactment/monitoring and adjustment phase in business process life cycle. It is needed to research using goal models, business process models and event logs for continual improvement of business process.

III. Approach

In this section, we present the details of the proposed approach. Figure 1 sketches the proposed Goal-oriented conformance checking method. It relies on two phases: Trace & Goal Processing phase to filter traces by a goal is achieved or not and Combine Goals to construct a cross tabulation table and a Statistical Analysis phase to measure significant differences between two goals in a cross tabulation table constructed in prior phase and evaluate the deviation have positive effects or negative effects.

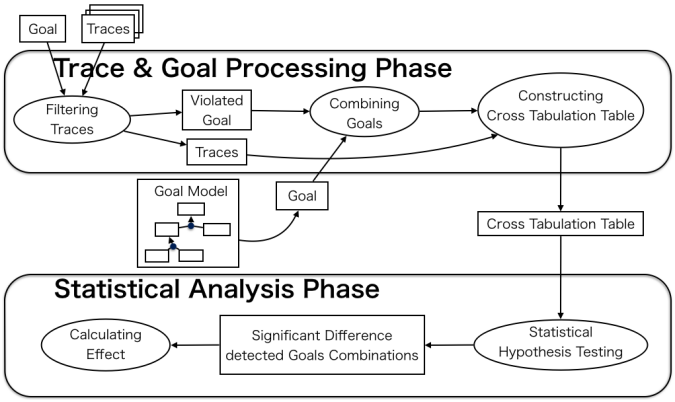

Fig.1. Overview of Our Proposed Method

A. Trace & Goal Processing phase

Trace & Goal Processing phase contains three processes and represented as an ellipse in Trace & Goal Processing phase of Figure 1. This phase conducts checking whether event logs satisfy goals of a goal model represented as logical formula and a constructing cross tabulation table for the next phase.

First step, Filtering Traces uses a goal of a goal model and traces as inputs and divides all traces to a goal satisfied traces or not satisfied traces. Goals are described using linear temporal logic, so it is possible to verify the trace satisfy the goal or not using LTL checker [7] on ProM. If any of these traces can not satisfy the goal, the goal is considered as violated goal. Second step, Combining Goals conduct combining the violated goal and a more upper or a high priority goal using information about a goal model configuration. more upper or high priority goals is more important than low level goals. Combining Goals is conducted for evaluating the violated goal influences other goals or not. Third step, Constructing Cross Tabulation Table use goals combined in the prior step and use these goals to construct a cross tabulation table. Table I is a cross tabulation table we want to construct. The table has two variables which represent a each combined goal is achieved or not. B in Table I represents a case when the violated goal is achieved, conversely, !B in Table I represents a case when the violated goal is not achieved. A in Table I represents a case when the combined goal is achieved, conversely, !A in Table I represents a case when the combined goal is not achieved. Therefore the cross tabulation table has 2×2 cells which represent traces numbers of $(A \wedge B)$, traces numbers of $(A \wedge !B)$, traces numbers of $(!A \wedge B)$, traces numbers of $(!A \wedge !B)$. The cross tabulation table represents the correlation between goal A and B. These numbers in cells are used in next statistical analysis phase. In these ways, the deviation from process defined by goal models are detected and preparation for the next phase are conducted.

TABLE I. Cross-tabulation table used for statistical hypothesis testing

	B	!B
A	trace numbers of $(A \wedge B)$	trace numbers of $(A \wedge !B)$
!A	trace numbers of $(!A \wedge B)$	trace numbers of $(!A \wedge !B)$

B. Statistical Analysis phase

Statistical Analysis phase contains two processes and represented as an ellipse in the Statistical Analysis phase of Figure 1. This phase analyzes whether goals achievement relation have significant differences and when the goal is not achieved, whether it have positive effects or negative effects against achieving the combined goal are evaluated.

First step, Statistical Hypothesis Testing uses cross tabulation table constructed in the prior phase as an input. We use Chi-squared test and Fisher's exact test. These statistical hypothesis testing methods are suitable for cross tabulation tables including categorical data and used for testing independence between two variables. If goal relations have significant differences in the significance level of 0.05 (p-value < 0.05), we consider these goals having relations. If some cells have lower values, Fisher's exact test is used. Second step, Calculating Effect are conducted for goals having significant differences in the significance level of 0.05 (p-value < 0.05). In this step, it is evaluated that violated goal influences positive effects or negative effects in a combined goal. Therefore, we use below equation (1). A is the number of traces when a combined goal is achieved. B is the number of traces when a violated goal is achieved. !A and !B are the number of traces when a each goal is not achieved. The equation represents the effect of not achieving goal B concerning goal A. If the value of *effect* is positive, not achieving goal B concerning goal A has positive relation. If value of *effect* is negative, not achieving goal B concerning goal A has negative relation. In this way, a relation between two goals are evaluated.

$$effect = \frac{A \wedge !B}{A} - \frac{!A \wedge !B}{!A} \quad (-1 \leq effect \leq 1) \qquad (1)$$

IV. Case Study

We have evaluated our approach on an event log which taken from a phone repair process and is publicly available and used in some researches for evaluations. The log contained 11855 events from 12 different events in 1104 cases, each case representing a phone terminal repair process (register, analyze defect, repair, test repair, archive and etc.). The log deta format is XES. Each trace describes a sequential list of events corresponding to a particular case. The log, its traces, and its events may have any number of attributes [1]. Attributes are standard (case id, time and etc.) or domain specific (phone

TABLE II. GOALS AND FORMALLY DEFINED GOALS OF GOAL MODEL

goal name	formal defined goal	goal type	priority
Achieve [repairing phone]	◇Repair (S) ∨ ◇Repair (C)	event	low
Achieve[archiving information]	◇(Archive Repair)	event	low
low repair numbers	numberRepairs $<$ 3	constraint	medium

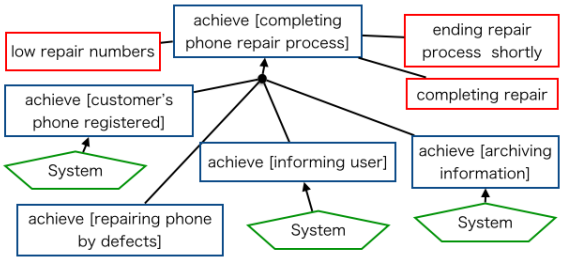

Fig.4. partial goal model of phone repair process

type, defect type and etc.). We constructed a goal model and use the model in Figure 4 for evaluations. Blue rectangles represent event and process related goals. Red rectangles represent constraint goals. Green pentagon represents agent that should achieve the goal devoted to the arrow. These goals are partially and formally described in Table II using linear temporal logic.

A. Goal: low repair numbers & other goals

In this section, we explain cases when goal: low repair numbers are detected as a deviation. First, in Trace & Goal Processing phase, goal: low repair numbers is combined with goal: completing repair using goal model construction. Next, a cross tabulation table is constructed. This is represented in the table III. This table represents 4 cases which show the number of both goals are achieved, only one goal is achieved and both goals are not achieved. Next, Statistical analysis phase uses this cross tabulation table. Fisher's exact test is used for calculating significant differences between these goals and p-value are calculated. The results are described in table IV. Calculated p-value is $2.2e^{-16}$, so this has significant differences in significance level of 0.05. Therefore, next, the effect is calculated using equation (1). The equation can represent that not achieving goal: low repair numbers is positive effects or negative effects against Goal: completing repair. The result is represented in table IV. Effect value is -0.903 (truncate a number to 3 decimal places). Therefore, not achieving Goal: low repair numbers detected as deviations are negative effects against Goal: phone repair completed.

TABLE III. CROSS-TABULATION TABLE OF GOAL (COMPLETING REPAIR) & GOAL (LOW REPAIR NUMBERS)

	low repair numbers	!(low repair numbers)
completing repair	1014	50
!(completing repair)	2	38

V. RELATED WORK

In this section, we explain related work about conformance checking. Related works excepted for conformance checking are lined up in section 2. Rozinat et al proposed token replay conformance checking method [8]. It measures the fitness

TABLE IV. P-VALUE & EFFECT

A: combined goal & B: violated goal	testing method	p-value	effect
A: completing repair & B: low repair numbers	Fisher's exact test	$2.2e^{-16}$	-0.903
A: ending repair process shortly & B: low repair numbers	Pearson's Chi-squared test	$2.2e^{-16}$	-0.317

of the process model and event log. Adriansyah proposed alignment based conformance checking method [9]. It makes it possible to check conformance more precise. Leoni et al proposed multi perspective conformance checking method [10]. This technique deal with control flow, data and resource for conformance checking.

These researches mainly focus on deviation between process models and event logs. This is an important aspect in process improvements, but detecting deviation between process models and event logs are not conclusive destination. Utilizing the results of conformance checking as a means to improve business process models, business rules and to reconfigure business goals should be conducted. In this perspective, our proposed conformance checking method is suitable for this.

VI. CONCLUSION

Today's organizations need to comply with a rapidly changing business environments and need to set goals against the change. In this paper we proposed a goal-oriented conformance checking method. The method can detect deviations between goal models and logs using the verification method based on linear temporal logic and can evaluate detected deviations using statistical analysis. It is useful for constructing more appropriate business processes and organizational goals.

REFERENCES

[1] W.M.P. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer, Verlag, 2011.

[2] A. van Lamsweerde. Requirements Engineering :From System Goals to UML models to software specications. WILEY, 2009.

[3] H. Horita, K. Honda, Y. Sei, H. Nakagawa, Y. Tahara and A. Ohsuga " Transformation Approach from KAOS Goal Models to BPMN Models Using Refinement Patterns, " in SAC 2014, ACM, Gyeongju, pp. 1023-1024, 2014.

[4] J.L. de la vara, J. Sánchez, and Ó. Pastor. "On the Use of Goal Models andBusiness Process Models for Elicitation of System Requirements," in BPMDS 2013, Springer, Valencia, pp.168-181, 2013.

[5] G. Gröner, M. Asadi, B. Mohabbati, D. Gašević and F. Silva Parreirasand Marko Bošković. "Validation of User Intentions in Process Models," in CAiSE 2012, Springer, Berlin, pp. 366-381, 2012.

[6] M. Ruiz, D. Costal, S. España, X. Franch, Ó. Pastor, "Integrating the goal and business process perspectives in information system analysis," in CAiSE 2014, Springer, Heidelberg, pp. 332-346, 2014.

[7] W.M.P. van der Aalst, H. de Beer and B. van Dongen, "Process Mining and Verification of Properties: An Approach Based on Temporal Logic," in OTM 2005, Heidelberg, pp. 130-147, 2005.

[8] A. Rozinat and W.M.P. van der Aalst. Conformance Checking of Processes Based on Monitoring Real Behavior. Information Systems, 33 : pp. 64-95, March 2008.

[9] A. Adriansyah, "Aligning Observed and Modeled Behavior", PhD Thesis. Technische Universiteit Eindhoven, The Netherlands, 2014.

[10] M. de Leoni and W.M.P. van der Aalst, "Aligning event logs and process models for multi-perspective conformance checking: An approach based on integer linear programming," in BPM 2013, Splinger, pp. 113-129, Beijing, 2013.