Model Consolidation: A Process Modelling Method Combining Process Mining and Business Process Modelling

Ornela Çela[(✉)], Agnès Front, and Dominique Rieu

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France
{ornela.cela,agnes.front,
dominique.rieu}@univ-grenoble-alpes.fr

Abstract. Analysing the current state of a process is a crucial step when improving the process. Four elements are defined during the process analysis: the objective, the process model, the indicators and the blocking points. It is important to well-define all of them when characterizing a process state, however the process model has a more fundamental influence. This model provides a map upon which is done the analysis of the process state, prediction of its evolution and simulation of the impact of the change to be undertaken for improving the process.

In this work we are proposing a strategy for the model consolidation, that combines the process discovery and business process modeling approach. This strategy aims to merge the models derived by these existing approaches in one unique model that is complete, comprehensive, aligned to the reality and useful for in-depth analysis.

This article describes the model consolidation strategy by detailing the steps to be taken and illustrating its usage into a real-life process provided from our industrial collaborator Net Invaders [19].

Keywords: Process analysis · Business process modelling and process mining

1 Introduction

Nowadays companies are operating in a rapid changing environment that demands them to continually evolve their strategic, operational and organizational aspects. This need of continual evolution is impacting also the business processes, but their evolution is time and financially consuming and internal resistance from the actors of the process can be encountered. CEFOP method was introduced in [1, 2], to guide this continual evolution by:

1. offering a full coverage of the required functionalities for analyzing, diagnosing and continually evolving the process.
2. involving and motivating the actors of the process in participating in the process evolution, by facilitating their interaction.
3. offering the possibility to transform business and operational objectives to objectives related to the process evolving.
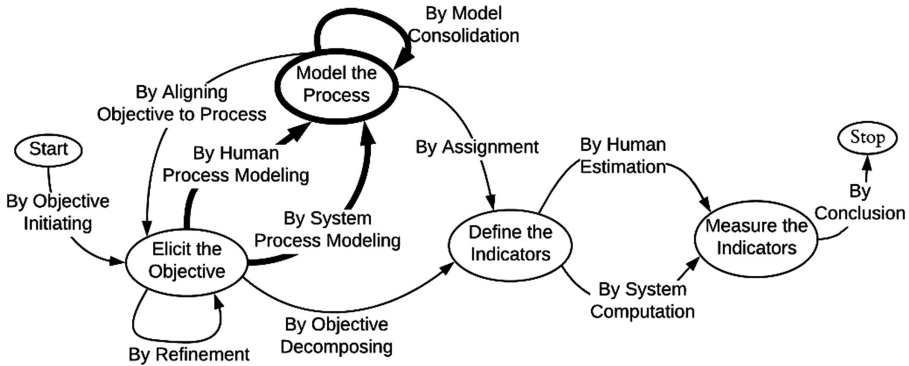
Fig. 1. The intentional map for the process analysis strategy of the CEFOP method [1].

In [1], the process analysis strategy was detailed by identifying four main intentions to be achieved as shown in Fig. 1.

The process model is one of the intentions to be attained during this analysis. This model is used for: the communication between the participants, detection of problems and solutions, prediction and simulation of process evolution. Taking into consideration these usages, the process model must be:

- **Comprehensive:** It should be easy to read and understand by the actors of the process and the process owner, respecting so the pragmatic quality criteria in [3].
- **Complete:** All the activities done by the actors during the process must be represented in the model.
- **Aligned to the reality:** It should describe how the process is really executed, illustrating what is happening, as defined in [4].
- **Useful:** The actors and process owner can perform in-depth analysis [4] and set-up indicators based upon their needs.

In this article, we are going over the existing approaches targeting the process modelling intention and introducing the model consolidation strategy. This strategy aims to merge the models derived by business process modelling and process mining techniques into one unique consolidated model that contains two levels of abstractions:

- The low-level abstraction illustrating the actor' activity at a fine-grained degree.
- The high-level abstraction describing the global activities of the process performed by the actors.

The steps proposed for consolidating the model are described and illustrated by using a real-life process provided by our collaborator, Net Invaders. The client management process covers the activities from the client inscription up to the purchase and use of the solution proposed by the company. The triggering activity is the client inscription. This request is validated by the client manager by a first telephonic contact and after that the client's testing environment is created. Once the environment is ready, the client may start testing the solution. At the end of the testing period the client can decide to purchase or not the solution. If the solution is purchased, a contract is

signed between the parties and the client can start using the solution and get charged for this service. Else the client environment is deleted, by rolling back all the configurations and erasing all the actions performed by the client during the testing period.

In Sect. 2, we are overviewing several existing methods and techniques in disciplines of the process discovery and business process model and illustrating the models of the process obtained for the case study. Section 3 details the steps to be followed for consolidating the model and mapping the two levels of abstractions. These steps are illustrated with the results perceived during the model consolidation in the client management process. We conclude this article and discus over perspectives on Sect. 4.

2   Related Works

Business process modelling and process discovery are the two principal disciplines identified in [3] that aim modelling the business process. These disciplines follow two distinct strategies for attaining the intention of process modelling:

The business process modelling discipline follows a human process modelling strategy. A team of persons composed of business analysts, actors of the process or both is charged to model the process.

The automate process discovery, on the other hand, follows a strategy of process modelling by system. The model of the process is automatically discovered by applying process mining techniques over the traces left during the process execution.

Several methods and techniques are proposed by both disciplines for modelling a process based on the context. In this section we are going over some of the most relevant methods proposed by each of these disciplines and illustrating the models obtained by them for the client management process.

2.1   Business Process Modelling

The business process modelling discipline proposes different methods and techniques for attaining the process model intention. These methods and techniques are grouped into two principal categories based upon the participation of the actors of the process into the modelling: participatory and non-participatory. We are focusing principally on the participative methods, the actor's participation into the process modelling reduces the intervention of an external expert and increases their understanding of the model. This last element reduces the internal resistance when it comes to diagnosing and evolving the process.

The intervention of the actors of the process into the modelling process generate a more comprehensive and complete model of the process as stated in [5]. In order to facilitate their participation and motivation, the concept of serious games was used in several methods. In [7], virtual modelling technique introduced the usage of avatars in the 3D environment to model the process. A similar technique was used also in ImProve presented in [8], for helping the actor of the process to better understand and improve their process. These methods provide a visual interface so that the actor can better interact with the model but a pre-training is required for taking into hand the modelling process. The usage of tangible artifacts lowers this entrance barrier as

presented in CoDesign [10]. CoDesign and ISEA [9] propose the usage of posted and other tangible tools to model the process. The actors are encouraged to exchange between them and better understand the model. ISEA provides a more structured guide to be followed when modelling the process and a mean of converting the model into a formal format. The actors of the process are guided in modelling the process themselves, by taking turn and describing their activities and the documents used during these activities.

We used ISEA method to model the process of the client management. The actors described their activities and the used tools instead of the documents. The model generated is shown in Fig. 2.
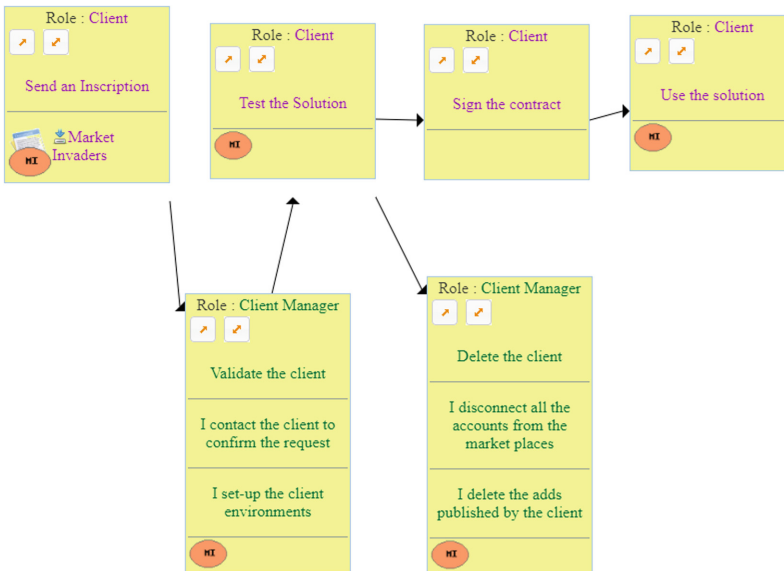

Fig. 2. The process of the client management modelled by its actors using the ISEA method.

The obtained model is quite comprehensive: It is easy to read and understand the principal activities and the paths between them. The actors were able to describe all the activities that they performed assuring so the model completeness.

This model however is descriptive, the actors outlined the process and it is not sure that all the paths taken during the process execution are illustrated. When modelling the process, the actors indeed describe the general paths, neglecting possible deviations in the model. Another disadvantage of this model is that it doesn't offer a lower level of abstraction where it is possible to analyse the actions performed by the actors within their activities. Also no mean of quantifying the process performance is provided since there is no correlation between the models and the data upon which the indicators could be calculated.

2.2   Process Mining

Several techniques and methods are proposed by the process mining discipline. One of the principal method of this discipline is PM2 [11]. PM2 is composed of seven steps, guiding from data collection up to the identification of the problems in the current state of the process. The model discovered from process mining is based upon the event logs that are generated during the process execution, thus it is aligned to the reality. The usage of event logs to model a process allow to perform a process analysis at the action level and to set-up and compute indicators based on the traces of the process execution. However, the dependence from the event logs, limits process mining techniques in the discovery of activities that are not reflected in the event logs, such as manual activities performed in the process.

Moreover, real-life processes generate complex event logs, containing large number of events on a very fine-grained state and multiple paths. The models discovered in such case are complex and referred to as spaghetti models [15]. To discover a model for our case study, an event logs file containing 70 event names [14] and around two hundred cases was extracted from the tool used during this process. We use the DISCO [13] process mining tool to discover the model shown in Fig. 3.
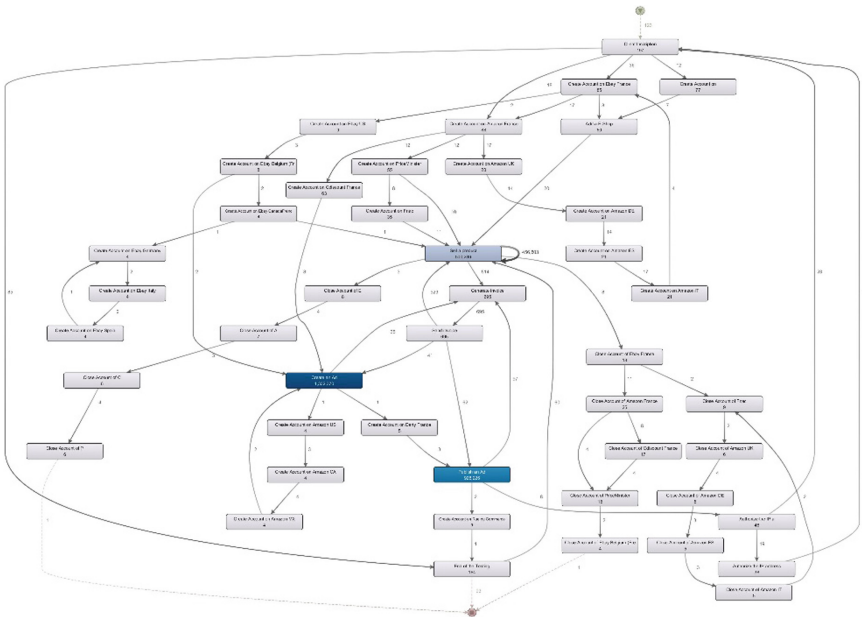

Fig. 3. The model of the client management process discovered by using DISCO.

The discovered model is not comprehensive and no analyse of the process can be performed upon it. Several techniques are proposed for structuring these spaghetti models. L* method introduced in [12] suggests to decompose such processes into sub-processes that are easier to understand and analyse. The disadvantages of such

solution are the impossibilities to have a global view over the process and to choose the right way of decomposing the process that fits the focus of the analyse.

Most recent techniques in process mining permit to deal with spaghetti models by creating a higher abstract level by using supervised [16] and unsupervised [17] learning techniques. These learning techniques are applied over the event logs file in order to create a higher level of abstraction within the events. It is upon the new created level of abstraction of the event logs that a more structured model is discovered. In [16], the abstraction of the events is performed by using a training dataset. This dataset is a set of events already labelled by a domain specialist as belonging to a given activity on the higher level of abstraction. It is up to the supervised learning technique to assign the rest of the events on these activities based on their similarities to the trained dataset. The disadvantage of this technique is that a training dataset has to be prepared by a domain expert and the activities belonging to the high levels must be already defined.

In [17], on the other hand, the abstraction of the events is done by an unsupervised technique. The events are grouped based upon patterns of executions. Once the abstraction is performed, it is up to the domain specialist to properly name the high-level events. In both cases on [16, 17], the intervention of a domain expert is required for completing the higher level of abstraction and the actors of the process are left aside.

2.3   Overview of the State of Art and Positioning

Two main strategies are followed in the existing state of the art for attaining the intention of the process model: the human and system process modelling. Figures 2 and 3 show the models of the client management process derived by these strategies.

The ISEA method was used in the case of the human process modelling strategy since it puts the actors of the process in the centre of the modelling process and offers a full guide for converting this model into a formal format. The actors are gather around the table and charge to model the process by exchanging between them. The derived model is comprehensive and complete, but it lacks alignment to the reality, uses for performing in depth analysis and means of measuring the performance of the process.

We used the fuzzy miner to discover a model of the process when undertaking the strategy of system process modelling. The model derived by this approach is aligned to the reality but not complete. The activity of *"Sign the contract"* is not represent in this model since it is a manual activity and it produces no trace in the system. Another disadvantage is that this model is uncomprehensive and cannot be used for analysis. Even though the possibility of creating a higher level of abstraction in the model is proposed, the mapping of the transition between the two levels of abstraction requires the intervention of an external expert and leaves aside the actors of the process.

In order to have a model of the process that complies to the four criteria and fill the gaps in the state of the art summarized in Table 1, we are introducing in the following section the model consolidation strategy.

This strategy aims to merge the models derived by the two existing approaches into one unique model. Since these models are used as starting points for the model consolidation, the model consolidation strategy inherits the context's requirement of:

Table 1. The overview of the state of the art illustrating the existing gaps in the related works.

	Criteria				
	Complete	Comprehensive	Aligned to reality	Useful	
				Analysis	Indicators
Model of the process ISEA	Yes	Yes		High-level	No
Model of the process Heuristic Miner	No	No	Yes	Low-level	Yes

The existence of event logs: This element is a restriction for undertaking the system modelling strategy, since the model is discovered based on the traces left during the process execution.

The actor's participation: This restriction is an essential criterion for assuring the continual evolution of the process in the CEFOP method, since it limits the intervention of an external expert and the internal resistance in the process evolution.

3   Process Model Consolidation

We are enriching the current state of the art, with the model consolidation strategy as show in Fig. 1. The aim of this strategy is to merge the models derived by the human and system process modelling strategies into one unique model, by structuring them as two different levels of abstractions for the same process model, where:

– The low-level abstraction is derived by system modelling strategy. The model illustrates the activities of the actors at a fine-grained degree, by illustrating the performed actions.
– The high-level abstraction is derived by the human modelling strategy, showing the global activities performed by the actors.

The model consolidation strategy guides the actor of the process into creating a map for travelling from the low-level to the high-level of model abstraction, allowing so the mutual enrichment between the models. This strategy permits to attain a model of the process that is comprehensive, complete, aligned to the reality and useful for analysis and process performance measuring. The model consolidation strategy is carried out in four steps. The tasks performed in each of these steps are detailed in the following subsection and illustrated on the client management process. We are using the terms:

• The process model to refer to the BPMN format of the described model of the process (Fig. 2), automatically retrieve by ISEA. This model is printed on a big format such that action cards can be placed within the activities. The process model of the client management is shown in Fig. 4.
• Activity to refer to an activity described by the actor when modelling the process by human. In the process model, the activities are represented as coloured rectangle,

where the colour defines the role performing it. Each role is represented by a unique colour into the model.

- Role to refer to an actor or a group of actors having the same role in the process. Each role participating in the process must be represented by at least one actor during the model consolidation.
- Action to refer to an event name [14], an activity found in the dataset used for modelling the process by system. The actions are represented by action card during the experiment as illustrated in Fig. 5. The action card can be coloured (with the role corresponding colour) or white (if the role is undefined in the dataset).
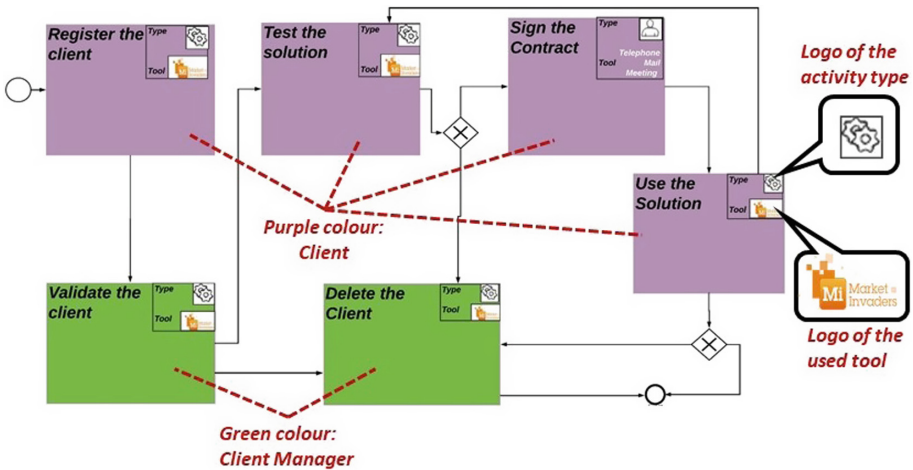

Fig. 4. The process model projected from the model describe by the actors in Fig. 2. The green and purple color represent respectively the role of the client manager and the client. (Color figure online)

3.1   Step 1: Correlate Activities and Actions

This is the first step to be performed when undertaking the model consolidation strategy. In this step, the roles have to associate their actions with the activities they have described. Table 2 describes the tasks to be performed in this step.

When undertaking this step in the process of the client management, the set of 70 action cards were distributed to the participants. Once the roles associated the action cards with the activities, two action cards were places aside one of the existing activity, creating so a new activity as illustrated in Fig. 5.

3.2   Step 2: Enriching the Set of Actions

Once the actions are correlated to the activities, the second step of the model consolidation can take place. The enriching of the set of actions is done in two tasks as detailed in Table 3. The time for the tool analyzing task differs depending on the time required for: verifying if the proposed tool stores traces of the process execution,
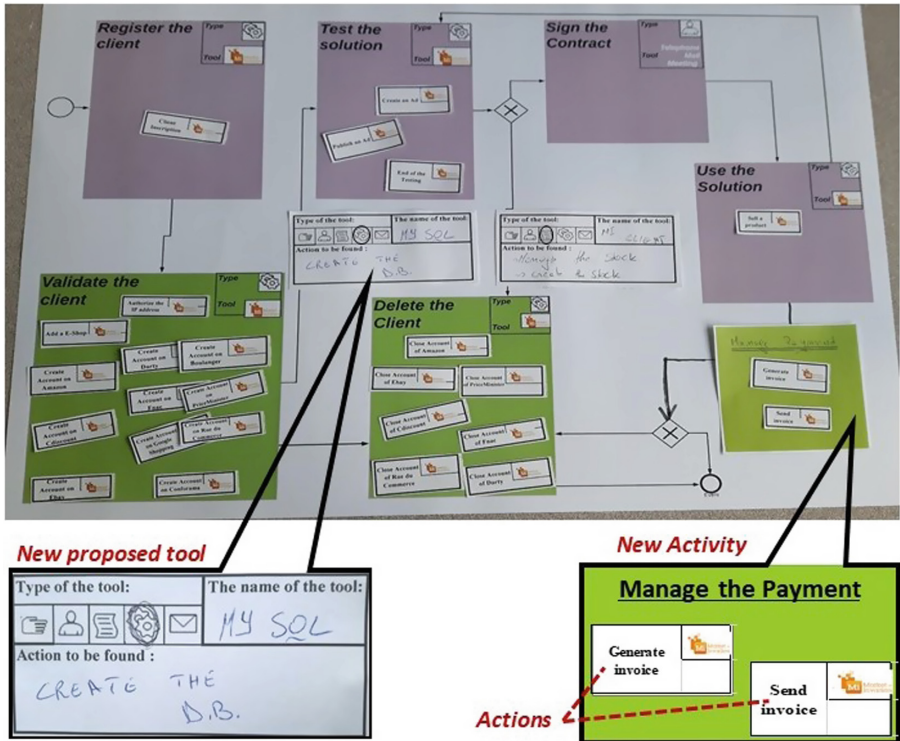
Fig. 5. The figure illustrates the result perceived during the two first steps of the model consolidation on the client management process.

extracting the traces and adapting them to the existing event logs. In case when a new set of traces is extracted by the proposed tool, the correlation of activities with actions must be retaken.

Two new tools were introduced by the actors during the experimentation on the case study as shown in Fig. 4. These sources were analyzed and no traces were able to be extracted since the tools didn't store traces of the process execution. However, a procedure for enabling the generation of logs in both tools is planned for enriching the model in the future.

3.3   Step 3: Paths Consolidation

The goal of this step is to identify deviations between the paths discovered by the system process modelling strategy and the paths described by the human process modelling strategy. The deviations between the paths are identified by the animator and validated by the actors of the process by enriching both models mutually. These tasks are described in Table 4.

In the client management process, three deviations were identified and only two of them were validated by the actors as shown in Fig. 6. The deviations were due to:

Table 2. The tasks performed for correlating the activities with the actions: the supports and the duration previewed for each of them.

Task and Instructions	C/I*	Supports and duration
Preparation: The animator brings the process model and describe the color code to the participants. The roles are asked to: *"Can you describe your activities in 2–3 sentences by following the order of the activities' execution within the process?"*	C	The process Model 10 min
The animator places the action card on the table and roles have to: - If the roles are defined: *"Take the action cards which have the color that was assigned to you"* - If the roles are not defined: *"Go over the action cards and pick the one's that represent actions that your perform"*		The action cards 5 min
The roles are asked to: "*Distribute the action cards over the process model"* **Attention:** *The action cards can be placed only:* *– inside the activity if you do the action within an activity,* *– over the input/output arrows if the action is not performed within an existing activity, but it is performed before/after it*		10 min
The animator asks the roles: "*Please give me the action cards that you have not placed in the process model?"* For each left aside action card, the animator filters the cases containing the action and discovers the model for this set of cases using Pro-M [18]. The simplified model is used to narrates the path of the action execution by describing the actions performed before and after. Once the path is described, the roles are asked: *"Do you want to place this action card into the model or discard it definitively*?"	C	The discovered process model 3 min/action card
The roles are equipped with markers and empty activities cards on their corresponding color and asked to: "*You can use the empty activity cards to create new activities in the model in order to group the action cards left outside the existing activities, for this:* **1.** *Go over the action cards placed on each side of the existing activities and see if they can be grouped at the same activity* **2.** *Once the grouping is done, place an empty activity card under the action cards and title this new activity* **3.** *Using the marker, outline the paths between the new and the exiting activities*	I	Empty activity card Black markers 10 min

*Collective/Individual

The first validated deviation, from *"Test the solution"* to *"Use the solution"* is caused by the absence of the actions reflecting the manual activity *"Sign the contract"* performed in between.

Table 3. Describes the tasks to be performed for enriching the set of actions.

Task and Instructions	C/I*	Supports and duration
The animator distributes tool cards to the roles and asks them: *Are there actions that you performed in this process that are not yet represented in the process model? If yes, use a tool card to describe the tool that you use when performing these actions?*	C	Tool cards 10 min
The animator goes over the tool cards proposed by the role and 1. Identify the new tools to be analyzed, by filtering the tools types (non-manual) that can be further inspected for traces 2. Extract the traces from the tool if possible	I	The required time is variable

*Collective/Individual

Table 4. The tasks performed for identifying deviations between the paths and validating them.

Task and Instructions	C/I*	Supports and duration
The animator is charged to detect deviations between the paths in the two models. The deviations are detected by following the steps: 1. Identify the paths between actions. All the paths discovered by using Pro-M, heuristic miner algorithm over the dataset are identified 2. Abstract the paths. The paths between actions are abstracted to the higher level, between activities A and B when: It exists at least one discovered path between events a and b, where a is the set of events belonging to activity A and b is the set of event belonging to activity B 3. Find the deviations by comparing the discovered paths with the described paths in the process model	I	Pro-M Dataset of actions 10 min
The deviations are overviewed one by one. The animator, using a red marker, draws the deviation path into the process model and asks the roles: *"Do you know this path or you need to decompose it and inspect it at the actions level?"* Once the path is clear, the actor must decide to validate it or not: *"Which of the following reason better describe this deviation and do you want to keep this path or not? (a) The path was forgotten when the model was described (b) The path is an exception; taken in exceptional cases (c) Actions are missing in the model. An action is performed between the two edges but it is not represented by an action card (d) Existing actions must be duplicated (or moved): The edge actions for this path are present in the model, but their position need to be reviewed* If the path is validated, the animator uses a black highlighter to overwrite the red path In cases where the reason of the deviations are (c) or (d), step 1 must be retaken	C	The discovered process model Black and red highlighter The deviations

*Collective/Individual

The validated deviation, from *"Use the solution"* to *"Test the solution"* reflects a deviation that is currently happening and must be corrected in the future.

The non-validated path, shown in red in Fig. 6 represents a deviation previously detected and already rectified in the system, so there is no need to keep it for further analysis.
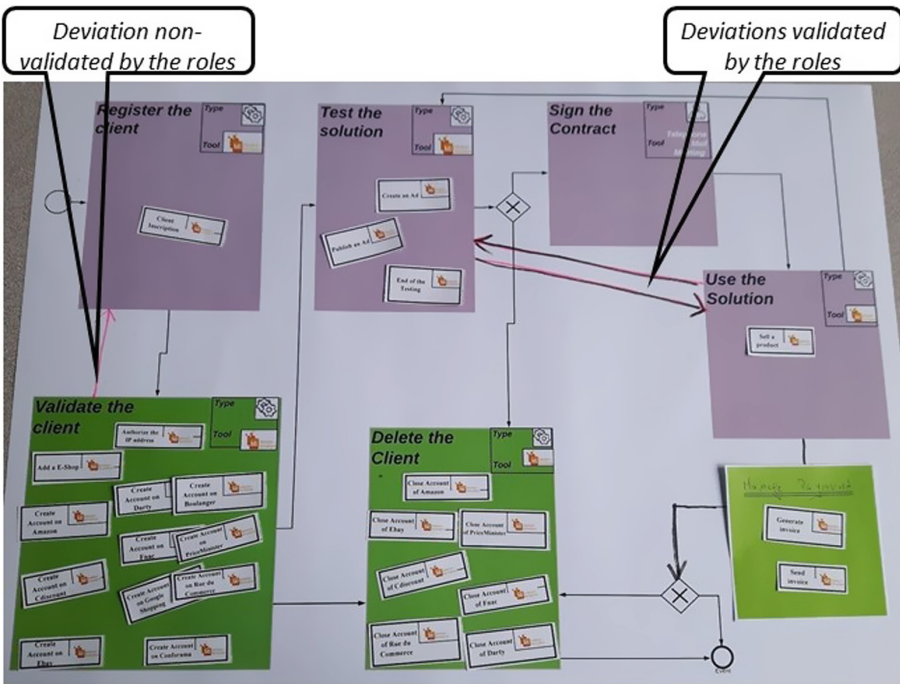

Fig. 6. The result of the path consolidation step in the client management process is illustrated, by showing the validated and the non-validated paths.

3.4   Step 4: Consolidate the Models

This is the last step in the process consolidation. The aim of this step is to mutually consolidate both models, for this, two tasks are performed in parallel:

1. Cleaning the dataset: The aim of this task is to remove from the dataset all the non-used actions and all the cases containing the non-validated paths.
2. Enriching the high-level model: All the validated deviations and new activities emerged during the process consolidation must be inserted into the high-level model that was perceived through the actor's description.

In our case study, the event logs were cleaned by removing all the cases containing the path from *"Validate the client"* to *"Register the client"* and the described model were enriched with the new activity *"Manage payment"* and with the path from *"Use the solution"* to *"Test the solution"*.

4   Conclusions and Perspectives

Two main strategies exist nowadays for modelling a process: human process modelling and system process modelling. Neither of them is able to provide a process model that is complete, comprehensive, aligned to the reality and useful for analysis. In order to attain such objective, we introduced in this article a model consolidation strategy that merge the models derived by these strategies into one unique model with two different levels of abstraction. This consolidation permits to have a model that is:

Comprehensive since the high-level model abstraction is created by the actors of the process and it is understandable and readable by them.

Complete: All the activities performed by the actors are reflected at the high-level of abstraction.

Aligned to the reality since the mapping between the levels of abstraction enriches and correlates both models mutually.

Usable for depth analysis: During the mapping of the two levels of model abstraction a correlation between event logs and activities is settled. This correlation facilitates setting up indicators and their computation.

During the undertaking of model consolidation strategy beside the model's correlation, the actors of the process are also:

Cleaning and enriching the event logs: The identification of non-used actions or paths cleans the event logs from noisy data that might be present in the event log. The consolidation incites them into identifying new sources of traces that can further enrich the model and the analysis.

Apprehend the models: The implication of the actors in the model consolidation, permits them to better understand the models, its deviations and the need of changing the process.

In this article, we illustrated an experimentation of the model consolidation performed over a process having a set of 70 event names. In order to facilitate the undertaking of the model consolidation strategy for other process having hundreds of event names, in the future works we want to provide an interface that facilitate the correlation between the actions and activities. The automatically detect the deviations between the models is also a step to be put in place, since currently this task is performed manually demanding so that the animator must have process mining knowledge.

The process model is one of the four intentions aimed to be achieved when analysing the state of a process as shown in the intentional map in Fig. 1. In our future works, we aim to define how to attain the three remaining intentions: define the objective, define the indicators and measure the indicators.

References

1. Çela, O., Front, A., Rieu, D.: CEFOP: a method for the continual evolution of organisational processes. In: 11th International Conference on Research Challenges in Information Science (RCIS) (2017)
2. Front, A., Rieu, D., Cela, O., Movahedian, F.: Les méthodes d'évolution continue au sein des organisations: le cadre As-Is/As-If. In: INFORSID (2017)
3. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.: Fundamentals of Business Process Management, 1st edn. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-33143-5
4. Van der Aalst, W.M.P.: What makes a good process model? Softw. Syst. Model **11**(4), 557–569 (2012)
5. Rittgen, P.: End-user involvement and team factors in business process modeling. In: 45th Hawaii International Conference on System Sciences, Hawai, USA (2012)
6. Buur, J., Ankenbrand, B., Mitchell, R.: Participatory business modelling. CoDesign: Int. J. CoCreation Des. Arts **9**(1), 55–71 (2013)
7. Brown, R., Recker, J., West, S.: Using virtual worlds for collaborative business process modeling. Bus. Process Manag. J. **17**(3), 546–564 (2010)
8. Ribeiro, C., Fernandes, J., Lourenço, A., Borbinha, J., Pereira, J.: Using serious games to teach business process modelling and simulation. In: International Conference of Modelling, Simulation and Visualisation Methods, Las Vegas, USA (2012)
9. Front, A., Rieu, D., Santorum, M., Movahedian, F.: A participative end-user method for multi-perspective business process elicitation and improvement. Softw. Syst. Model., 1–24 (2017)
10. Herzberg, N., Kunze, M.: The business process game. In: 7th Central-European Workshop on Services and their Composition (2015)
11. van Eck, M.L., Lu, X., Leemans, S.J.J., van der Aalst, W.M.P.: PM^2: a process mining project methodology. In: Zdravkovic, J., Kirikova, M., Johannesson, P. (eds.) CAiSE 2015. LNCS, vol. 9097, pp. 297–313. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19069-3_19
12. Van der Aalst, W.M.P.: Process mining: discovering and improving Spaghetti and Lasagna processes. In: IEEE Symposium on Computational Intelligence and Data Mining (2011)
13. DISCO tool Homepage. https://fluxicon.com/products/
14. van der Aalst, W.M.P.: What kind of data does process mining require? http://www.processmining.org/logs/start
15. Van der Aalst, W.M.P.: Discovery, Conformance and Enhancement of Business Processes, 1st edn. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-19345-3
16. Tax, N., Sidorova, N., Haakma, R., van der Aalst, W.M.P.: Event abstraction for process mining using supervised learning techniques. In: Bi, Y., Kapoor, S., Bhatia, R. (eds.) IntelliSys 2016. LNNS, vol. 15, pp. 251–269. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-56994-9_18
17. Mannhardt, F., Tax, N., Unsupervised event abstraction using pattern abstraction and local process models. In: Enabling Business Transformation by Business Process Modeling, Development, and Support Working Conference-BPMDS (2017)
18. Pro-M tool Homepage. http://www.promtools.org/doku.php
19. Net Invaders company homepage. https://en.net-invaders.com/