ProcessGold: Enterprise Process Mining

1st Roeland Scheepens
Head of Research
ProcessGold
Eindhoven, The Netherlands
roeland.scheepens@processgold.com

2nd Erik-Jan van der Linden
Non-Executive Director
ProcessGold
Eindhoven, The Netherlands
erik-jan.vanderlinden@processgold.com

***Abstract*—The *ProcessGold Enterprise Platform* offers innovative features that are required for mature, scalable, and secure use of Process Mining throughout large enterprises. Processes are displayed using process graphs, which require stable layouts to ensure continued use by business users. The platform uses a novel graph layout algorithm, called *Tracy*, that ensures this stability over user interaction.**

Relevance of enterprise features in general is evident from the dispersion of process mining in enterprises, from use cases, and from broad adoption of the Platform.

***Index Terms*—Process mining, Enterprise process mining, stable graph layout**

I. Introduction

Since its inception around the turn of the century [AGL98], the first decade of Process Mining [Aal16] saw the development from academic use and academic tools, such as ProM [vDdMV*05], to stand-alone tooling for use by individual analysts in businesses. The second decade, since around 2010, has been characterized by increasing use of and interest in process mining of non-technical business users in large enterprises.

II. Common features

The above development has led to a number of properties that are by now common in Process Mining platforms.

Firstly, Process Mining tools have become part of the system landscape of enterprises, which requires server- and web-based architectures, and full compliance with standard technical enterprise requirements such as data preparation on board, collaborative development, controlled deployment, and security.

Secondly, dispersion to non-technical business users requires easy point-and-click applications for everyday use, and hence Process Mining has naturally integrated with Data Analytics. This way, users have full access to required functionality from the field of Data Analytics. This includes any kind of business graphic besides process graphs, and full navigation options for selection, filtering, drill-down, etc. [Kei02].

Thirdly, for non-technical users, Process Mining is not their primary task, and they require an optimal balance between time they spend working with the application on the one hand and business value on the other hand. In practice, users expect response times they experience in other web-based applications. This performance requirement needs to be aligned with the need for flexibility: when users apply, for example, edge or event sliders, apply different process mining algorithms, filter, or select different case ID's for multidimensional process mining, they are making different selections from the full event data. Round trips to databases then hamper performance, and hence *in-memory* architectures have become a standard feature of enterprise process mining. Data selection, visualization, user navigation, as well as process mining algorithms all run in-memory.

All features mentioned are part of the ProcessGold Enterprise Platform. For a full overview of features see the ProcessGold website [Pro] and sources below.

III. Novel feature: Tracy process graph layout

Besides these, by now common, features that meet requirements for enterprise use, *graph layout* has become a topic which requires attention to ensure process mining is continuously usable by business users. The ProcessGold Platform includes a novel algorithm for process graph layout, *Tracy*[1]. A full technical paper about Tracy is available [MSW19].

A. Process layout

Since its inception, Process Mining has mainly relied on existing algorithms for graph layout [GKN15]. However, these only use the graph topology to compute a layout and do not make use of the particular properties of processes. For example, most processes have some sort of a main path [RBRB06], [AEHK10] that is often the most frequent behavior. To fit with intuitive semantics of process graphs, this path needs to be centralized in the layout and depicted as a straight path through the graph. While existing techniques fail to properly do this [GKN15], the Tracy approach implements this.

B. Stable process graphs

Interaction with process graphs, required by business users, by means of sliders, filtering, drill-down and alike leads to changes in the layout of process graphs. Users build mental maps of these graphs. When the graph changes, users need to fit their mental model with the altered graph. Conventional algorithms [GKN15] lead to smaller or larger changes, and require users to at least put in significant cognitive effort to adapt their mental map, but in the worst case even to rebuild their mental map altogether. Graph layout stability

[1]Patent pending.

CEUR-WS.org/Vol-2374/paper5.pdf

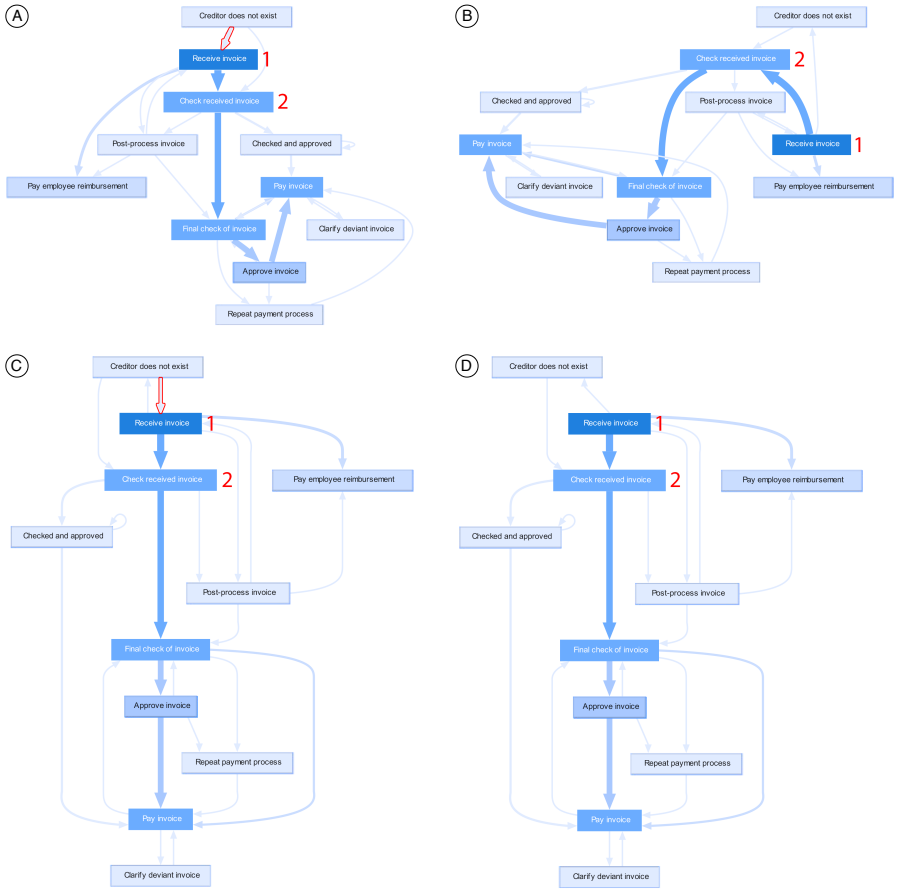

Fig. 1. Graphs A and B show the same process as graphs C and D, respectively. The layouts in A and B are computed by the industry standard [GKN15], where the actual process is poorly represented, while graphs C and D are computed by our novel graph layout algorithm where the process is easy to follow. Graphs B and D are obtained after removing the edge highlighted in red from A and C. As we can see, B differs significantly from A; especially note how nodes 1 and 2 swap vertically. Consequently, the mental map of the user is lost. On the other hand, C and D barely differ, preserving the mental map of the user.

helps preserve the mental map of the user [PHG06], [ZKS11], significantly reducing cognitive load.

Therefore, we need to preserve the mental map when a new graph layout is displayed. The three models of Misue [MELS95] to represent the mental map are our starting point to accomplish this. These models state that a layout adjustment should preserve the direction of node n to node m for each pair of nodes n and m, that nodes that are close together should remain close together, and that graphical objects in a region should stay in that region.

In *online* dynamic graph drawing [BBDW17], graph layouts need to be computed for a sequence of graphs without knowing the full sequence from the beginning. In contrast, in *offline* dynamic graph drawing [BBDW17], the whole sequence of graphs is known up front. Interactive Process Mining is offline in the sense that we know which graphs we can potentially encounter, since any filtered graph will be a subset of the data found in the event log, but it is also online because we do not know the exact sequence of graphs beforehand. The Tracy-approach presents a novel way to take this into account, and results in layouts that are stable for the user.

Stability vs quality: Besides stability, we also require layouts to be of high *quality*. Stability and quality are two conflicting requirements: graph layout stability helps preserve the mental map of the user, but also restricts the graph layout algorithm in optimizing layout quality. A way of dealing with this conflict is to allow somewhat larger changes to the layout and to make use of animation and transitioning as a secondary approach to mental map preservation. The combination of layout stability and transitioning provides the best approach to preserving the mental map.

C. Graph layout: Tracy

Tracy is a novel stable layout algorithm for process graphs that computes layouts that intuitively represent the semantics of the process. Our algorithm is based on the Sugiyama framework [STT81] but includes:

- A novel ranking algorithm;
- A novel order constraint computation algorithm;
- A novel crossing minimization algorithm.

In Figure 1, we present a comparison of Tracy and conventional layout.

Tracy is part of the production version of ProcessGold starting from version 16, and will be publicly announced at ICPM 2019.

IV. Maturity

The ProcessGold Platform is a mature enterprise Process Mining Platform. Customers of the platform include global consulting firms, and global enterprises with over 1 billion euro in annual turnover.

A further indication of maturity is the business value that customers get from implementing and using applications on the platform. We present two use cases which illustrate this.

A. Use cases

Telecom provider: A telecom provider is faced by increasing competition on price, requiring the need to reduce cost. A process mining effort identifies and resolves a number of process issues, in particular related to manual exception handling, communication with strategic suppliers. This leads to 20 percent personnel costs savings, and more broadly to 1+ M cost savings in general, which equals 300 percent return on investment of the effort within 3 months.

Insurance company: An insurance company is faced with customers that are dissatisfied by untimely customer contacts. Firstly, 71 percent of regular customer questions are not responded to within the agreed service level, two days. A process mining effort identifies incorrect routing of questions, leading to delays. Once resolved, this dramatically improves the timeliness. Secondly, in case of disease of a person having a pension insurance, the company witnesses highly negative comments from relatives on social media. These need to be prevented for their negative impact on the reputation of the company. In the process mining effort, the 15 percent of cases where communication is not timely handled are identified as well as the root causes, which are found in hand-over over cases and consulting secondary systems. By-catch is identification of 11 percent cost reduction.

B. Scale

Since 2016, customers and partners have acquired licenses for use of the ProcessGold Platform for 10.000+ potential users.

Acknowledgment

Research and development presented in this demo paper was carried out at ProcessGold. The authors acknowledge all colleagues at ProcessGold that contributed to this work, as well as academic partners at the Universities of Eindhoven (the Netherlands) and RWTH Aachen (Germany) for productive and long-lasting cooperation.

References

[Aal16] Aalst W. M. P. v. d.: *Process Mining: Data Science in Action*. Springer, Apr. 2016.

[AEHK10] Albrecht B., Effinger P., Held M., Kaufmann M.: An automatic layout algorithm for BPEL processes. In *Proceedings of the 5th International Symposium on Software Visualization* (New York, NY, USA, 2010), SOFTVIS '10, ACM, pp. 173–182.

[AGL98] Agrawal R., Gunopulos D., Leymann F.: Mining process models from workflow logs. In *International Conference on Extending Database Technology* (1998), Springer, pp. 467–483.

[BBDW17] Beck F., Burch M., Diehl S., Weiskopf D.: A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum 36*, 1 (Jan. 2017), 133–159.

[GKN15] Gansner E. R., Koutsofios E., North S.: *Drawing graphs with dot*. Tech. rep., Graphviz, Jan. 2015.

[Kei02] Keim D. A.: Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics 8*, 1 (2002), 1–8.

[MELS95] Misue K., Eades P., Lai W., Sugiyama K.: Layout adjustment and the mental map. *Journal of Visual Languages & Computing 6*, 2 (June 1995), 183–210.

[MSW19] Mennens R. J., Scheepens R., Westenberg M. A.: A stable graph layout algorithm for processes. *Eurographics Conference on Visualization (EuroVis)* (June 2019).

[PHG06] Purchase H. C., Hoggan E., Görg C.: How important is the "mental map"? – an empirical investigation of a dynamic graph layout algorithm. In *Graph Drawing* (Sept. 2006), Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, pp. 184–195.

[Pro] ProcessGold: ProcessGold - business intelligence and process mining platform. https://processgold.com/en/. Accessed: 2019-23-02.

[RBRB06] Rinderle S. B., Bobrik R., Reichert M. U., Bauer T.: Business process visualization – use cases, challenges, solutions. In *Proceedings of the Eighth International Conference on Enterprise Information Systems (ICEIS'06): Information System Analysis and Specification* (May 2006), INSTICC PRESS.

[STT81] Sugiyama K., Tagawa S., Toda M.: Methods for visual understanding of hierarchical system structures. *IEEE Transactions on Systems, Man, and Cybernetics 11*, 2 (1981), 109–125.

[vDdMV*05] van Dongen B. F., de Medeiros A. K. A., Verbeek H. M. W., Weijters A. J. M. M., van der Aalst W. M. P.: The prom framework: A new era in process mining tool support. In *Applications and Theory of Petri Nets 2005* (Berlin, Heidelberg, 2005), Ciardo G., Darondeau P., (Eds.), Springer Berlin Heidelberg, pp. 444–454.

[ZKS11] Zaman L., Kalra A., Stuerzlinger W.: The effect of animation, dual view, difference layers, and relative re-layout in hierarchical diagram differencing. In *Proceedings of Graphics Interface 2011 (GI'11)* (2011), Canadian Human-Computer Communications Society, pp. 183–190.