

Visualizing User Editing Behavior in Collaborative Ontology-Engineering Projects

Simon Walk, Tania Tudorache, and Mark A. Musen

Stanford Center for Biomedical Informatics Research
Stanford University
Stanford, CA 94305, USA
{lastname}@stanford.edu

Abstract. Over the last decade, ontologies have become the mainstay in the biomedical domain. Their size and complexity, as well as, the required expert domain knowledge to create these ontologies have increased significantly. In addition, many projects resort to collaborative approaches for building these ontologies, using the Internet as a cooperation platform. While online collaborative projects have become common, the processes that drive these collaborations are still not well understood. In this paper, we are investigating novel approaches and visualizations using Markov chains to improve our understanding of how users build large, real-world ontologies. Using our novel methods, we analyze, visualize and compare the editing behavior of users in two collaborative ontology-engineering projects from the biomedical domain. The contributions of our work are two-fold. First, we visually explore the editing behavior and dynamics of users in collaborative ontology-engineering projects; and second, we quantify the differences between the editing behaviors in these two projects. We also discuss the implications and potential applications of our findings, which, we believe, may be used to create adaptive user interfaces that better support the editing behaviors of the users.

Keywords: editing behavior, visualization, collaborative ontology engineering, Markov chain, transition matrix

1 Understanding Editing Behaviors in Ontology-Development Projects

In recent years, we have seen an increased adoption of ontologies, especially in the biomedical domain. These ontologies play a critical role in acquiring, representing and processing information about human health. With a steadily increasing relevance, ontologies had to cover new findings and application domains, which also triggered an increase in size and complexity. For example, the World Health Organization has embraced OWL as a representation language for the 11th revision of the International Classification of Diseases (ICD-11), which now consists of roughly 50,000 classes (diseases and causes of death), and it is edited in an online collaborative environment on the Web [18].

As no small group of domain experts, let alone single individuals, have the required expertise and resources to develop such highly-specialized and large-scale ontologies,

new requirements for the ontology-engineering process are emerging. In order to better support users in editing these large real-world ontologies collaboratively, and to support the project managers to get an overview of the entire process, we need to better understand the intricacies of the collaborative process for building ontologies.

Uncovering such new insights—which can potentially be used to adapt and improve existing ontology-engineering tools, or devise new and expand existing development and evaluation strategies—represents a very important first step towards overall improved and more easily maintainable structured knowledge representations.

In this paper, we focus on visually analyzing, exploring and comparing user editing behavior across two different collaborative ontology-engineering projects from the biomedical domain. To that end, we expand on our previous analyses [24–26], fit first-order Markov chain models on sequences of change-type actions, and visualize the resulting transition matrices. In contrast to most existing visualization techniques, the figures presented in this paper include information about the sequential nature of the change-logs. Further, we extend the utility of the presented approach by visually highlighting the differences and commonalities between the two investigated projects.

The remainder of the paper is structured as follows: We discuss the related work (Section 2), then we describe the datasets and methods to extract user-editing behaviors from the change history of collaborative ontology-engineering projects (Section 3). We present the results of our analyses in Section 4, and then discuss the implications and potential applications of the presented visualizations in Section 5.

2 Related Work

In general, research in the field of ontology engineering addresses all the tasks, actions, tools and processes required for developing and evaluating ontologies [5]. As a result, many researchers and practitioners have developed guidelines and methodologies [3, 6, 7, 14, 17] or compiled best practices [11] for engineering ontologies. Simperl and Luczak-Rösch [15] provide an exhaustive overview of different collaborative ontology-engineering methodologies and tools.

To learn more about the impact and implications of collaboration and consensus finding in collaborative ontology-engineering projects, researchers have analyzed several aspects of such projects. For example, Falconer et al. [4] investigated if contributors of collaborative ontology-engineering projects exhibit specific roles, and if these roles can be used to group and classify these users when contributing to the ontology. Strohmaier et al. [16] investigated the hidden social dynamics that take place in collaborative ontology-engineering projects from the biomedical domain and provided new metrics to quantify various aspects of the collaborative engineering processes. Pesquita et al. [12] showed that the location and specific structural features can be used to determine if and where the next change is going to take place in the Gene Ontology¹. To analyze user editing patterns, Wang et al. [27] used association-rule mining on the change-logs of collaborative ontology-engineering projects, and showcased the utility of the identified editing patterns in a prediction experiment.

¹ <http://www.geneontology.org>

In 2014, Van Laere et al. [21] used k-means and the GOSPL methodology to classify users by analyzing and clustering the different interactions that users engage in, while collaboratively working on engineering an ontology.

To support the collaborative development of ontologies, the Semantic Web community has already developed a number of tools and visualizations. For example, Protégé, and its versions for collaborative ontology development, such as WebProtégé [20], iCAT [18] and Collaborative Protégé [19], are prominent stand-alone tools, which are used by a large community worldwide to develop ontologies for a variety of different projects.

Many of the existing ontology-engineering tools already provide built-in functionality or plug-ins to visualize and browse ontologies [1, 8, 10]. However, most of these visualizations lack the ability to visualize the dynamic nature of collaborative ontology-engineering projects. Falconer et al. [4] developed the Change Analysis plugin for Protégé that presented user-activity graphs and author dependency graphs from the change history of a project. In 2013, Pöschko et al. [13], and Walk et al. [22] developed new visualizations, which take the historical evolution of an ontology into account. The tool, called *PragmatiX*, allows users to browse aspects of the history of collaboratively-engineered ontologies. PragmatiX also provides quantitative insights, which allow for an easier monitoring of the progress of collaborative ontology-engineering projects. In 2015, Burch and Lohmann [2, 9] presented visualizations of dynamic and time-varying graphs, using the VOWL notation.

In our previous work, we have used Markov chains to visualize different aspects of the editing behavior of users in collaborative ontology-engineering projects [24–26]. In this paper, we expand on our previous work and present a new analysis and a novel visualization that extends the arsenal of available analysis tools for investigating social interactions in collaborative ontology-engineering projects.

3 Materials & Methods

3.1 Datasets

We used the change logs of two real-world ontology-engineering projects to conduct the analyses presented in this paper. Both projects were developed with iCAT [18], a customized version of WebProtégé [20], which records a log of all changes performed

Table 1. Characteristics of the ICD-11 and ICTM datasets used in our analyses.

	ICD-11	ICTM
Classes #	48,771	1,506
Changes #	439,229	67,522
Users #	109	27
First change date	2009/11/18	2011/02/02
Last change date	2013/08/29	2013/07/17
Editing period (ca.)	4 years	2.5 years

by each user. Each entry in the record stores additional metadata about the change, such as the user who performed the change, a brief description of the change, the timestamp, and all the classes and properties involved in the change. Additionally, we have removed all automatically executed changes.

Using these additional metadata, we can create sequences of change actions for each user, which can then be used to fit a (first-order) Markov chain model. A brief characterization of the datasets can be found below and in Table 1. Note that we have removed users with less than 2 changes from our analysis as no transitions between changes could be observed.

The International Classification of Diseases (ICD),² developed by the World Health Organization (WHO), is the international standard for diagnostic classification used to encode information relevant to epidemiology, health management, and clinical use in nearly all of the United Nations' 193 member states. The 11th revision of the classification, **ICD-11**,³ is currently in progress, with a planned finalization in 2018. In contrast to previous revisions, ICD-11 is developed as a rich OWL ontology [18], with over 48,000 classes and authored by over 100 domain experts.

The International Classification of Traditional Medicine (ICTM)⁴ is a project that aimed to produce an international standard terminology and classification for diagnoses and interventions in Traditional Medicine and was led by WHO. Analogously to ICD-11, ICTM was developed collaboratively as an OWL ontology with the goal of unifying the traditional medicine practices from China, Japan and Korea. The content of the ontology is authored in English, Chinese, Japanese and Korean. More than 20 domain experts from the three countries developed ICTM, using a customized version of WebProtégé, until the development of ICTM ended in 2012. Currently, there is ongoing work to make ICTM a chapter in ICD-11.

Table 2. Listing of all change-type actions in the change-logs.

Change Type	Description
<i>Add Condition</i>	A restriction is added to a class.
<i>Add Direct Type</i>	A direct type is added to an entity.
<i>Add Property Value</i>	A new value is added to a property.
<i>Create Class</i>	A new class is created.
<i>Create Reference</i>	A new reference is created.
<i>Delete Class</i>	A class is deleted.
<i>Delete Condition</i>	A restriction is deleted from a class.
<i>Delete Property Value</i>	A property value is deleted.
<i>Edit Property Value</i>	A property value is edited.
<i>Import Property</i>	A property value is imported from an external ontology.
<i>Move Class(es)</i>	One or more classes are moved in the class hierarchy.
<i>Remove Superclass</i>	A superclass of a class is removed.
<i>Replace Reference</i>	A reference is replaced.
<i>Retire Class</i>	A class is retired.
<i>BREAK</i>	30 minutes of inactivity between two actions.

² <http://who.int/classifications/icd/en/>

³ <http://who.int/classifications/icd/ICDRevision/>

⁴ http://who.int/mediacentre/news/notes/2010/trad_medicine_20101207/en/

3.2 Extracting Editing Behaviors

In order to analyze the editing behavior of users in two collaborative ontology-engineering projects, we have extracted and aggregated several different types of changes (see Table 2). Each change-type represents one action that a user performed in the user interface. For example, *Edit Property Value* describes a change, where an existing value of a property of an entity is edited by a user. In contrast, *Add Property Value* describes the action of adding a value to a property of an entity, which was previously empty. We extracted these change types from the change history recorded by iCAT. Note that even though ICD-11 and ICTM were both created using iCAT, not all types of changes were performed in both projects during our observation periods.

By extracting and analyzing the sequences of change-actions of all users in ICD-11 and ICTM, we can calculate the “average” editing behavior of all users involved in the corresponding project—in the form of a first-order Markov chain—and identify commonly followed workflows. Further, whenever more than 30 minutes have passed between two consecutive changes of one user, we have added a *BREAK* state. This allows us, not only to learn more about common workflows among users, but also which actions are conducted before users take a break, and which actions are conducted, once they resume work.

It is important to understand that the types of changes listed in Table 2 are manually aggregated change-types of all the different types of changes created by iCAT. The visualizations and analyses are not limited to these, and could be easily applied to a different set of change-logs created by a different ontology-development tool that provides similarly granular log-information. Aggregating similar change types into one action helps make the visualizations easier to read.

3.3 Fitting Markov Chains

A Markov chain consists of a finite state-space S , where each state $s_1, s_2, \dots, s_n \in S$ with $n = |S|$ and a transition matrix P , which lists all probabilities p_{ij} to traverse from the state s_i to the state s_j , and for each i , $\sum_j p_{ij} = 1$.

For example, if the state space of one specific Markov chain consists of all the types of changes that can be conducted on an ontology, the values listed in the transition matrix P reflect the probabilities of a user to conduct a specific type of change immediately after that user conducted the same or a different type of change.

Similarly, Markov chains can be used to model the transition probabilities between multiple transitions or higher orders. This means that the next state does not only depend on the current state, but on a sequence of k previous states, as well.

In our previous work, we have already demonstrated that higher-order Markov chain models can be extracted from the logs of changes of different collaborative ontology-engineering projects [24–26]. However, for the purpose of creating visualizations to manually inspect these transitions, higher-order Markov chains quickly become cluttered, as all permutations of possible states up to length k have to be included in the state space as well. Hence, in this paper, we focus on the visualization of first-order Markov chains.

To be able to calculate the transition matrix P , we first have to define W with each element w_{ij} representing the number of transitions between states s_i and s_j . Hence, $\sum_j w_{ij}$ describes the absolute number of occurrences of state s_i in the corresponding dataset. Finally, we calculate P by normalizing each row of W by its corresponding ℓ_1 -norm, and use it for our visualizations.

To be able to compare two projects, we calculate the transition matrix Q_{abs} by subtracting W_{ICTM} from W_{ICD-11} , and again normalizing each row by its ℓ_1 -norm, as described in Equation 1.

$$Q_{abs} = \frac{W_{ICD-11} - W_{ICTM}}{\|W_{ICD-11} - W_{ICTM}\|_1} \quad (1)$$

Note that the transition probabilities in Q_{abs} reflect the differences between the absolute numbers of occurrences of change-type transitions of the two datasets. After normalization, frequent transitions, which are mostly appearing in only one of the two dataset, will either be close to 1.0 for the minuend (W_{ICD-11}) or -1.0 for the subtrahend (W_{ICTM}).

In contrast, $Q_{rel} = P_{ICD} - P_{ICTM}$ represents the difference between the relative transition probabilities of the change-type transitions. Analogously, after normalization, transition probabilities will range between 1.0 and -1.0 . Hence, we can either decide to compare the absolute number of transitions between the different change-type actions (Q_{abs}), or the relative transition probabilities (Q_{rel}).

A detailed description of all the steps necessary to apply Markov chains on collaborative ontology-engineering projects, as well as additional results that complement the analysis presented in this paper can be found in Walk et al. [25]. Please note that the analyses presented in this paper extend the previously presented analyses [23, 26], and include a novel approach to visualize the differences in the editing behaviors of users in ICD-11 and ICTM.

4 Results

Figures 1(a) and 1(b) depict the results of our Markov chain analyses for ICD-11 and ICTM, respectively. According to the histograms, depicted on top of the transition matrices, the top changes that are conducted in ICD-11 and ICTM are related to adding or editing values of property values (*Edit Property Value* and *Add Property Value*). Additionally, users in ICD-11 have also focused on creating and moving classes (*Create Class* and *Move Class(es)*) in the ontology, while users in ICTM focused on conducting *Import Property* and *Create Reference* changes.

For both projects, we can observe a strong tendency for all users to concentrate on a particular type of change, evident in the higher transition probabilities in the diagonal (between the same change actions). Further, we are also able to identify specific “workflow” patterns in the visualizations of both projects. For example, users have a tendency to create a class, then edit an existing property value, followed by adding a new property value to an empty property.

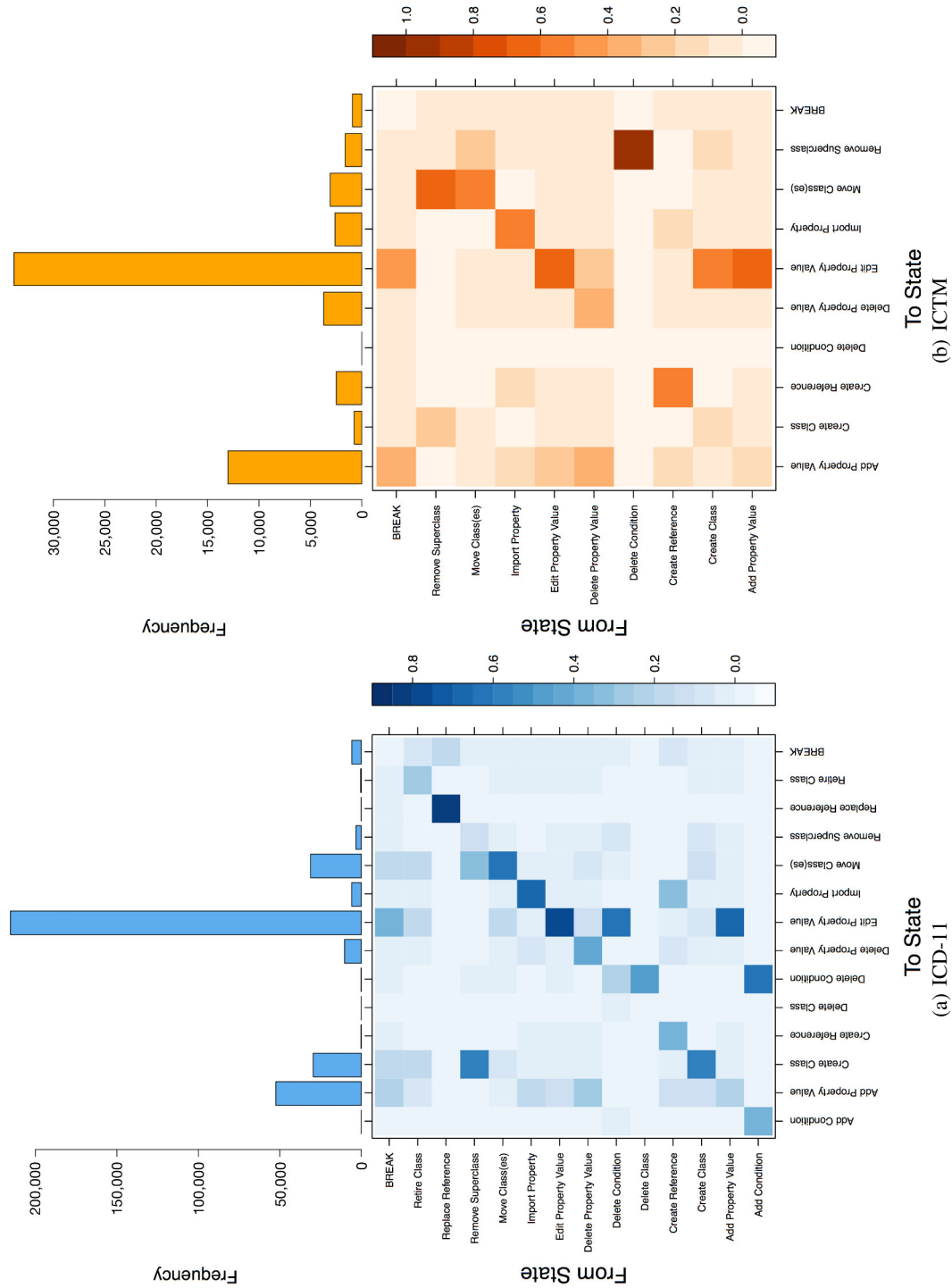


Fig. 1. Visualization of User Editing Behaviors: The top of the figures depict the histograms of the absolute occurrences of the corresponding change type actions. The transition maps are depicted on the **bottom** and visualize the probabilities (darker means higher probability) to transition *From State* (left) to *To State* (bottom) for the different change type actions in ICD-11 (a) and ICTM (b).

Additionally, we can observe that users who delete a property value are equally likely to delete another property value or add a new property value to a class. For both projects, we can further see that users, who conducted a *Remove Superclass* change, are very likely going to create a new class next or move classes in the hierarchy (*Create Class* and *Move Class(es)*). However, after moving a class, users in ICD-11 exhibit a higher tendency to edit a property value, while users in ICTM are likelier to remove another superclass.

For both projects, there does not appear to be a specific change action that users conduct before taking a break from work (see columns of *BREAK* in Figure 1), and the probabilities to conduct specific types of changes when returning from a break are according to the frequencies of the different types of changes depicted in the histograms (see rows of *BREAK* in Figure 1).

To further highlight the differences between the editing behaviors of the users of the two projects, we visualized Q_{abs} (see Figure 2(a)) and Q_{rel} (see Figure 2(b)). As ICD-11 exhibits roughly six times the amount of changes compared to ICTM (see Table 1), transition probabilities in Q_{abs} are mostly positive. The only exceptions are *Create Reference* and *Import Property*, which were performed more often by the users of ICTM. As described before, we can now see that users in ICTM have a higher tendency to first conduct a *Remove Superclass* change, immediately followed by a *Move Class(es)* change than users who work on ICD-11. In contrast, users in ICD-11 are likelier to conduct a *Create Class* change after removing a superclass (*Remove Superclass*).

When looking at Q_{rel} , this difference becomes even stronger (Figure 2(b)). We can also see that specific workflows, such as creating a class, editing a property value and then adding a property value are more dominant for ICTM, which is likely caused by the overall lower number of changes and thus, their increased relative importance. In contrast, users in ICD-11 exhibit a higher tendency than users of ICTM to consecutively perform the same change-type actions, as depicted by the diagonal in Figures 2(a) and 2(b).

5 Discussion

The analyses of the presented visualizations provide important insights into the editing behavior of users in collaborative ontology-engineering projects, which can not be inferred directly from static data. We have demonstrated that common workflows can be identified when exploring the editing behavior visualizations presented in Section 4. We believe that this newly obtained information could be used by ontology-engineering tool developers to identify opportunities for improving the ontology-development tool. For example, the user interface of the tool can be better aligned to the observed workflows, and tool developers could use the information about the obtained sequences of changes to adapt the interface with the goal of reducing the number of clicks that are required to perform certain workflows.

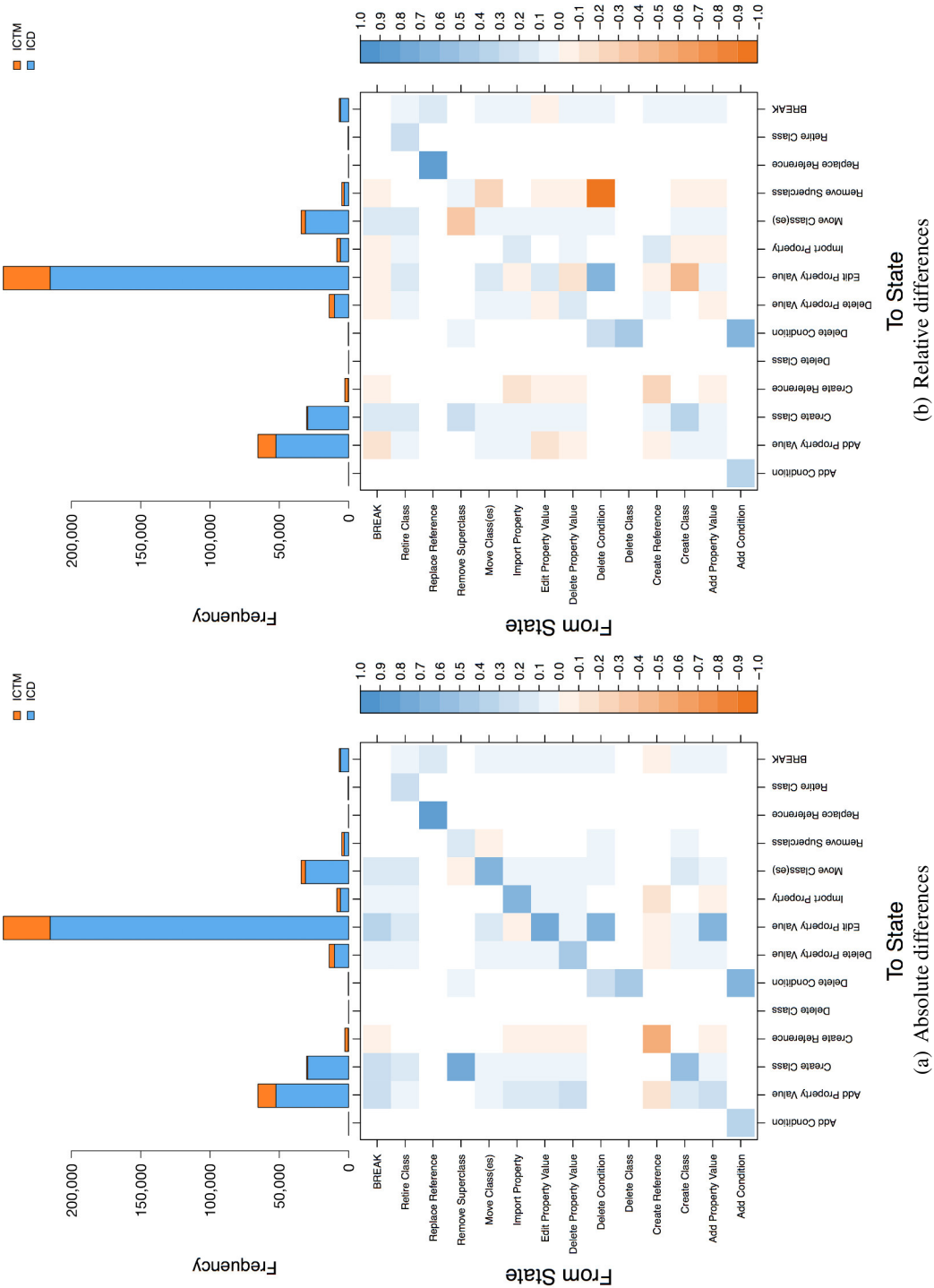


Fig. 2. Visualization of User Editing Behavior Differences: The top of the figures depict the stacked histograms of the absolute occurrences of the corresponding change type actions of both projects. The transition maps are depicted on the **bottom** and visualize the probabilities to transition *From State* (left) to *To State* (bottom) for the different change type actions in ICD-11 (blue; 1.0) and/or ICTM (orange; -1.0). The closer the transition probability to zero the lighter the color in the transition map, meaning that the transition is equally important in both datasets.

Further, the presented visualizations could act as source of information for automatically adapting interfaces, which guide users through the process of developing an ontology by considering their historical editing behavior. We have also observed that users tend to edit “vertically” in the ontology (e.g., if they edit the property value of a class, they will do so for all the subclasses). Such a workflow suggests that a tabular editing interface, similar to spreadsheets, presenting one class per row, would speed up the content entry by the users.

The visualization of the differences in editing behaviors between the two projects may be used by project administrators to identify and analyze differences in the workflows between projects, and especially between the ones that use different tools or guidelines for creating the ontology. In contrast to existing visualizations, our presented method specifically concentrates on the sequential nature of the change-log data, and provides novel insights into the dynamic nature of collaborative ontology-engineering projects.

For future work, we plan on further refining the presented analyses to compare the editing behavior of different groups of users. For example, biologists might exhibit different workflows and patterns, when developing an ontology than chemists or computer scientists do. To extend the utility of the Markov chain visualization, we plan on making the visualizations interactive. This extension would allow us to dynamically aggregate similar types of changes into abstract classes of changes, which could be expanded and collapsed by users while exploring the visualization. Additionally, this extension would make it possible to visualize higher-order Markov chains, and avoid visual clutter due to the increased number of states. Finally, we are also interested in comparing the editing behavior of users across different ontology-development tools to assert the extent to which the tool influences the editing behavior.

6 Conclusions

In this paper we have extended the analyses from our previous work, and demonstrated how to visualize and compare the editing behavior of users of two different collaborative ontology-engineering projects from the biomedical domain. In that process, we have uncovered and discussed several editing workflows, and we have presented a novel visualization, which highlights the differences in editing behaviors between two projects. Finally, we have discussed the implications of our findings for ontology tool developers and project administrators, and we have outlined potential applications of the visualizations for future work.

Acknowledgments This work was supported by Grant GM086587 from the U.S. National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health. The Protégé project is support by NIGMS grant GM103316.

References

1. Alani, H.: Tgviztab: an ontology visualisation extension for protégé (2003)

2. Burch, M., Lohmann, S.: Visualizing the evolution of ontologies: a dynamic graph perspective. In: Proceedings of the International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data (VOILA 2015). CEUR-WS. vol. 1456, pp. 69–76 (2015)
3. Debruyne, C., Reul, Q., Meersman, R.: Gospl: Grounding ontologies with social processes and natural language. In: 2010 Seventh International Conference on Information Technology. pp. 1255–1256. IEEE (2010)
4. Falconer, S.M., Tudorache, T., Noy, N.F.: An analysis of collaborative patterns in large-scale ontology development projects. In: Musen, M.A., Corcho, O. (eds.) K-CAP. pp. 25–32. ACM (2011)
5. Gomez, A., Fernandez, M., Corcho, O.: Ontological engineering. 2 nd Edition, Springer-Verlag (2004)
6. Gómez-Pérez, A., Suárez-Figueroa, M.C.: Scenarios for building ontology networks within the neon methodology. In: Proceedings of the fifth international conference on Knowledge capture. pp. 183–184. ACM (2009)
7. Grüninger, M., Fox, M.S.: Methodology for the design and evaluation of ontologies (1995)
8. Horridge, M.: Owlviz-a visualisation plugin for the protégé owl plugin. the university of manchester, 2004 (2008)
9. Lohmann, S., Link, V., Marbach, E., Negru, S.: WebVOWL: Web-based visualization of ontologies. In: Proceedings of EKAW 2014 Satellite Events. LNAI, vol. 8982, pp. 154–158. Springer (2015)
10. Lohmann, S., Negru, S., Bold, D.: The protégéowl plugin: ontology visualization for everyone. In: European Semantic Web Conference. pp. 395–400. Springer (2014)
11. Noy, N.F., McGuinness, D.L., et al.: Ontology development 101: A guide to creating your first ontology (2001)
12. Pesquita, C., Couto, F.M.: Predicting the extension of biomedical ontologies. *PLoS Comput Biol* 8(9), e1002630 (09 2012), <http://dx.doi.org/10.1371/journal.pcbi.1002630>
13. Pöschko, J., Strohmaier, M., Tudorache, T., Noy, N.F., Musen, M.A.: Pragmatic analysis of crowd-based knowledge production systems with icat analytics: Visualizing changes to the icd-11 ontology. In: Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium: Wisdom of the Crowd. Stanford, CA, USA (2012)
14. Schreiber, G., Wielinga, B., de Hoog, R., Akkermans, H., Van de Velde, W.: Commonkads: A comprehensive methodology for kbs development. *IEEE expert* 9(6), 28–37 (1994)
15. Simperl, E., Luczak-Rösch, M.: Collaborative ontology engineering: a survey. *The Knowledge Engineering Review* 29(01), 101–131 (2014)
16. Strohmaier, M., Walk, S., Pöschko, J., Lamprecht, D., Tudorache, T., Nyulas, C., Musen, M.A., Noy, N.F.: How ontologies are made: Studying the hidden social dynamics behind collaborative ontology engineering projects. *Web Semantics: Science, Services and Agents on the World Wide Web* 20(0) (2013), <http://www.websemanticsjournal.org/index.php/ps/article/view/333>
17. Suárez-Figueroa, M.C., Gómez-Pérez, A., Motta, E., Gangemi, A.: Ontology engineering in a networked world. Springer Science & Business Media (2012)
18. Tudorache, T., Falconer, S.M., Nyulas, C.I., Noy, N.F., Musen, M.A.: Will Semantic Web technologies work for the development of ICD-11? In: Proceedings of the 9th International Semantic Web Conference (ISWC 2010). ISWC (In-Use), Springer, Shanghai, China (2010)
19. Tudorache, T., Noy, N.F., Tu, S., Musen, M.A.: Supporting collaborative ontology development in Protégé. Springer (2008)
20. Tudorache, T., Nyulas, C., Noy, N.F., Musen, M.A.: WebProtégé: A Distributed Ontology Editor and Knowledge Acquisition Tool for the Web. *Semantic Web Journal* 4(1/2013), 89–99 (2013)

21. Van Laere, S., Buyl, R., Nyssen, M.: A Method for Detecting Behavior-Based User Profiles in Collaborative Ontology Engineering. In: *On the Move to Meaningful Internet Systems: OTM 2014 Conferences*. pp. 657–673. Springer (2014)
22. Walk, S., Pöschko, J., Strohmaier, M., Andrews, K., Tudorache, T., Noy, N.F., Nyulas, C., Musen, M.A.: Pragmatix: An interactive tool for visualizing the creation process behind collaboratively engineered ontologies. *International journal on Semantic Web and information systems* 9(1), 45 (2013)
23. Walk, S., Singer, P., Noboa, L.E., Tudorache, T., Musen, M.A., Strohmaier, M.: Understanding how users edit ontologies: Comparing hypotheses about four real-world projects. In: *The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference*, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part I. pp. 551–568 (2015), http://dx.doi.org/10.1007/978-3-319-25007-6_32
24. Walk, S., Singer, P., Strohmaier, M.: Sequential action patterns in collaborative ontology-engineering projects: A case-study in the biomedical domain. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, CIKM 2014, Shanghai, China, November 3-7, 2014*. pp. 1349–1358 (2014), <http://doi.acm.org/10.1145/2661829.2662049>
25. Walk, S., Singer, P., Strohmaier, M., Helic, D., Noy, N.F., Musen, M.A.: How to apply markov chains for modeling sequential edit patterns in collaborative ontology-engineering projects. *International Journal of Human-Computer Studies* 84, 51 – 66 (2015), <http://www.sciencedirect.com/science/article/pii/S107158191500124X>
26. Walk, S., Singer, P., Strohmaier, M., Tudorache, T., Musen, M.A., Noy, N.F.: Discovering beaten paths in collaborative ontology-engineering projects using markov chains. *Journal of Biomedical Informatics* 51, 254–271 (2014), <http://dx.doi.org/10.1016/j.jbi.2014.06.004>
27. Wang, H., Tudorache, T., Dou, D., Noy, N.F., Musen, M.A.: Analysis of user editing patterns in ontology development projects. In: *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*. pp. 470–487. Springer (2013)