1H3-5

Predicate-based knowledge elicitation for action in knowledge networking

Danilo Eidy Miura Yukio Ohsawa Kazuo Furuta

School of Engineering, The University of Tokyo

This paper reports the proceedings of the International Innovators Marketplace on Data Jackets (IMDJ) Workshop and presents the analysis results of the knowledge elicitation through verbalization of actions as predicates on data usage. This experiment revealed an effective reuse of data by reuse of the predicate. The participants used the same data to provide different solutions, using similar actions. As conclusion, we will explain how to make use of predicates to support formal specifications in service design.

1. INTRODUCTION

Innovators Marketplace (hereafter IMDJ) Workshop is a method for knowledge elicitation [1] in a role play approach [2] for creating innovative solutions based on the use of data. In the workshop, the elicitation of knowledge was observed as predicate verbalization in their solutions` proposals.

The predicate, in linguistics, is the part of the sentence that informs something about a subject. Those predicates can be formulated into logics and then, can be computed in a formal meaning.

In formal methods research, the importance of understanding how to compose methods [3], suggest the need of close look on the ideation process in solutions proposals. In this approach, the use of verbs as a logical predicate was observed as knowledge representation and element for knowledge reuse.

The object of this experiment is the actions that were explicit in the solution creation process.

2. INNOVATORS MARKETPLACE ON DATA JACKETS (IMDJ)

The Innovators` Marketplace is the method for aiding innovative thought and communication where participants interact and elicit their knowledge to create solutions for research and market requirements [4].

The Data Jacket (here after DJ) is a digest description of a dataset that allows a potential consumer to sense the value of them without revealing the dataset contents. We use the DJs to communicate the meaning of each dataset and produce the supportive board where participants analyze potential connections among those datasets [5].

The board mentioned above is a tool for visualizing datasets` relationships, called KeyGraph (KG) [6]. The visualization of a network graph where the nodes represent DJs and relevant keywords, and edges represent relationships.

The Innovators` Marketplace on Data Jackets (IMDJ) is the combination of the method and tools above in some gamified [7] procedures and rules. In the IMDJ, participants are called to share their knowledge, combine their datasets and create solutions for elicited requirements. The requirements work as constraints to stimulate participants to think creative solutions with their knowledge [8]. All participants take roles as data supplier, innovator and consumer. As result, we externalize knowledge to create ideas of use of data and stimulate the sense to perceive value of data in a given context.

The results of IMDJ become relevant elements of the Action Planning (AP) [9], where the ideas are serialized into concrete actions, including resources and time planning.

3. FOCUS GROUP

The experiments in focus group had the duration of about 2 hours, with total of 20 subjects. This group was composed by researchers in the field of resilience engineering of different nationalities (Japan, France, Italy, England, and Sweden).

The experiment procedures consist in the following steps:

1) **Visualization** of KeyGraph, which was produced with the content of DJs provided by participants.

2) **Elicit requirements**, asking what participants want to do or desire. Participants were stimulated to think their requirements related to the content of the KeyGraph.

3) **Elicit knowledge** to satisfy the requirements. To provide ideas to satisfy the requirements, the participants should use the datasets included in the KeyGraph or add new DJ to compose the solutions.

4) **Evaluate solutions**, negotiating the price of the solution using toy money. For solutions that satisfy the given requirement, participants paid toy money, for not satisfactory solutions, the toy money was not paid.

Participants were instructed to think on the theme "Foster the Urban Resilience" and follow the procedures above to provide on solutions with the DJs of their research datasets. All participants were called to take the three roles on IMDJ (data supplier, innovator and consumer).

Figure 1. KeyGraph board in the end of the workshop

Figure 1 shows the agglomeration of pieces of paper in some areas of KeyGraph, resulted by the elicited knowledge from participants. The data consists in techniques description (partially pink), DJs (green), new DJ (pink), requirements (yellow), and solutions (blue).

Participants, as consumers, negotiated and confirmed satisfaction of their requirements by buying solutions with toy money.

4. KNOWLEDGE ELICITATION

The knowledge of participants was elicited by the verbalization of actions that can be done with the given data. Thinking on the variables on a dataset, someone propose actions that are expected to produce some useful results, according to his or her own knowledge.

The question to instruct the elicitation is "*What can you do with this dataset to solve a given requirement?*".Given the 14 requirements, the participants proposed 20 solutions using the datasets visualized in the KeyGraph. From all proposed solutions, 9 were successfully negotiated, indicating satisfaction of the addressed requirements. For example, the solution "An application to keep collecting the bio-information" was proposed to solve the requirement "Want to know the safety of my family [immediately, in case of disaster]". The solution "keep collecting bio-information" is the elicited knowledge from the participant.

5. SOLUTIONS TO PREDICATES

To make the analysis of the elicited knowledge in the written solutions, we assumed a linguistic approach to analyze the dependency of terms in the sentences. We identify the transitive verb in the sentence that link the subject to the object and assign it as predicate of the solution. To explain this representation we omitted the subject and defined the general model:

Predicate (object)

In this experiment, we observed the use of the actions related to 'Monitor'. For each solution which presented action related to monitor. It was assigned the predicate "Monitor", as the proposed action for the solution. In this paper, we defined "Monitor" as "keep a continuous record of something", where something is the object of the verb monitor. As the example mentioned above, the proposed solution was predicated as

Monitor (bio-information)
$\models$ Confirm (user, family_safety)

In this solution, the DJ named 'Mobile Phone Activity Log' contains 'time' as variable and it allows the participants to think to monitor other variables in the dataset. We applied the predicate 'Monitor' as part of the formal representation of the action in the solution.

According to the participants` knowledge, for dataset which the variable 'time' is included, it is possible to monitor the any data in the dataset.

For $\forall D$, where $t \in D$, $x \in D \rightarrow$ Monitor(x),
where, D: Dataset t: variable time x: any variable

According to this assumption, we applied the predicate "Monitor" to all successfully negotiated solutions that contained the variable "time" in the input dataset. From those 9 successfully negotiated solutions, 5 contained the variable "time" in the input dataset.

6. REUSE OF KNOWLEDGE

Once the knowledge of action in the solution is computed in the predicate formulation, the reuse of this knowledge depends on the requirements. To identify the satisfied requirement, we applied the predicate that shows the condition that is satisfied by the proposed solution. In the example provided above "Want to know the safety of my family [immediately, in case of disaster]", we observe two specific conditions to the requirement "safety of my family": 1) immediately and 2) in case of disaster:

Immediately (family_safety)
In_disaster (family_safety)

The proposed solution offers the value of association of a variable to the time. Therefore, we added the predicate "Real-time" to represent the dependency of the required object on time. In the example mentioned above, the required data about the family`s safety should be updated all the time to satisfy the user`s need. We added the predicate "Real-time" to the requirement to represent the object`s dependency on time, and requires an action to update the data to the time in which it is required. In this experiment, we defined "real-time' as an attribute that states a recent status of the object that it is referred.

Real-time (object)

From the 9 successfully satisfied requirements, 6 presented some dependency on the time delay to be satisfied. Such as the example above, the requirement would not be satisfied if the solution provides the information after a long time.

Computing the predicate "Real-time" into the requirement, we identify the need of the action "Monitor" the data to satisfy the given requirement. In this experiment, we defined "monitor" as the action of keep recording data according to the time.

We formulated the proposition on the solution proposal Monitor(data) satisfies requirements on Real-time(data) as:

Solution: Monitor(data) $\models$ *Requirement: Real-time(data)*

In the total of 9 successfully negotiated solutions, 5 of those followed the rule above, where solution had contained the predicate 'Monitor' and requirement had predicate 'Real-time'.

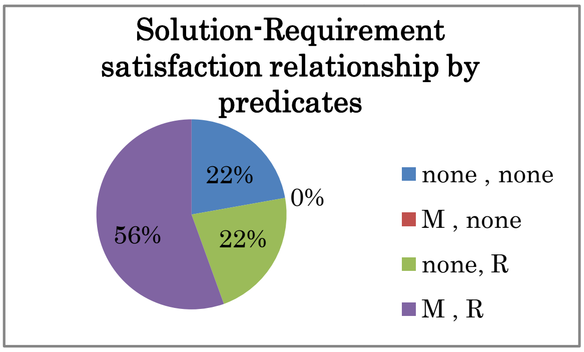

Figure 2. Satisfaction relationships according to predicates.

In the figure 2, we can see the distribution of the negotiated solution-requirement relationship. In the subtitles, "M" is Monitor predicated solution and "R" is Real-time predicated requirement.

7. CONCLUSIONS

As showed above, the elicited knowledge about the action "monitoring", represented by the predicate "Monitor", satisfied most of the requirements which were predicated as "Real-time". This experiment reveals the use of predicates to identify actions and features that present a satisfaction relationship. Those features can be expressed in the written language and represented as predicate form to identify the knowledge and problem in a stored data.

The predicate can be used as a representation of knowledge by the feature action, and representation of requirements by the feature expected condition. The satisfaction relationship of knowledge and requirements can be identified by those features. And the use of predicate in the requirements can be used to identify the potential reuse of the knowledge.

8. DISCUSSIONS

In this workshop, we observed verbalization of actions, where the verbs can be used as predicates to share one`s reusable knowledge to create new solutions.

In this experiment, the action was extracted from the proposed solution of a participant and reused into another participant`s solution. We found a limitation of analysis of the explicit contents which were presented in different terms and share similar meaning that satisfies the definition of the predicates in this experiment.

The 6 explicit requirements that depend on time were not explicitly described with the term "real-time". But due to the dependency on time being implicit in the participants knowledge, we could infer this condition and define a term to represent the shared sense of real-time.

In the other hand, terms such "keep collecting", "simulate", or "judge happiness" in the elicited knowledge allow us to infer the implicit action of keep recording data, due to the dynamic nature of the data that they refer to. These contents suggest an **implicit shared knowledge that is codified in different terms**.

As future work, we will apply the formal method to externalize the participants` knowledge in the form of action over the data. The knowledge should be formulated and included in the knowledge base for reuse. The experiment should be extended to other knowledge and predicates.

REFERENCES

[1] Sagheb-Tehrani, M., "A Conceptual Model of Knowledge Elicitation", in "Proc CONISAR 2009, v2 (Washington DC): §1542 (refereed) c 2009 EDSIG, pp. 1 (2009).

[2] Cooke, N., J., "Varieties of knowledge elicitation techniques", in Int. J. Human-Computer Studies 41, pp. 801-849, (1994).

[3] Clarke, E. M., J. M. Wing, Et Al., "Formal Methods: State of the Art and Future Directions", in "ACM Computing Surveys", Vol. 28, No. 4, December (1996).

[4] Ohsawa, Y., and Nishihara, Y., "Theories for Innovative Thought and Communication", in "Innovators` Marketplace: Using Games to Activate and Train Innovators", Springer, pp. 59-76, (2012).

[5] Ohsawa, Y., Liu C., Hayashi, T., Kido, H., "Data jackets for externalizing use value of hidden dataset", in "18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014", Procedia Computer Science 35. Elsevier, (2014).

[6] Ohsawa, Y.; Benson, N., E., Yachida, M., "KeyGraph: automatic indexing by co-occurrence graph based on building construction metaphor", in "Research and Technology Advances in Digital Libraries", in "ADL 98. Proceedings. IEEE International Forum", pp.12-18, 22-24 Apr (1998).

[7] Wang, H., Ohsawa, Y., "Data-Driven Innovation Technologies for Smarter Business from Innovators` Market Game to iChance Creativity Support System" in "Advances in Chance Discovery: Extended Selection from International Workshops". Springer, pp. 107-126, (2013).

[8] Ward, T. B., Smith M. S., and Finke A. R. (1999), Creative Cognition. In: Robert J. Sternberg (Ed.) Handbook of Creativity, pp. 189-212. New York: Cambridge University Press, 1999.

[9] Hayashi, T., Ohsawa, Y., "Relationship between Externalized Knowledge and Evaluation in the Process of Creating Strategic Scenarios" in "Open Journal of Information Systems (OJIS)", 2 (1), pp. 29-14, (2015).