Knowing What to Explain and When

Jörg Cassens

Norwegian University of Science and Technology (NTNU),
7491 Trondheim, Norway,
jorg.cassens@idi.ntnu.no,
http://www.idi.ntnu.no/

Abstract. We have argued elsewhere that user goals should be taken into account when deciding what kind of explanation of its results a CBR system should give. In this paper, we propose the use of an Activity Theory based methodology for identifying different user goals and expectations towards explanations given by a system supporting a work process.

1 Introduction

Customized IT Systems are usually designed for specific purposes and tasks which the system has to support and in settings where comparable work was done also before the system was introduced. It is used by people with specific needs and qualifications, and it should preferably adapt to changes in these needs over time [1, 2]. Althoff et al. [3] have introduced an organizational view of the CBR cycle for the purpose of business process modeling. For the purpose of this paper, we are looking at CBR systems embedded in such a work context, but on a more general level.

2 Problems with Explanations in CBR

The term explanation can be interpreted in two different ways in AI [4, p. 59]. One interpretation deals with explanations as part of the reasoning process itself, for example used in the search for a diagnostic result in order to support certain hypotheses. The other interpretation deals with usage aspects: making the reasoning process, its results, or the usage of the result transparent to the user. Both interpretations can be found in CBR research. The ability to explain its results is often considered as one of the main advantages of CBR systems [5–7]. A knowledge-intensive CBR system may use explanations to guide the CBR process itself [8, 9].

One problem is that the question of what makes up a good explanation depends on the goals of the user [10]. This also means that we cannot be sure that we will match the user's needs by presenting the case alone as it is [11]. So it might not be as straightforward as it sounds to provide the user with an adequate explanation. Where explanations are used to support the CBR process, this problem reappears when the explanation is used to asses the users needs or wishes, e.g. in an adaptive CBR system.

Another important point is that it might not suffice to purely present the best matched case(s) to the user to give an explanation even when his goals are matched. McSherry [12] points out that the presented case(s) might contain both features supporting the given results and features opposing it. Smyth and McClave [13] strengthen the importance of giving a set of results with sufficient diversity for certain types of problems. McGinty and Smyth [14, 15] propose an adaptive way of presenting a set of cases adapted to the user's changing needs for diversity. All these works deal with the shortcomings of presenting a single (or to narrow set of) case(s).

Using only cases as explanations means further on to rely on the implicit assumption that by presenting the case to the user he will be able to do a similarity comparison himself. This may often be true, but is by no means guaranteed, especially when the case structure is complex or the similarity measure more convoluted. The problem increases when we start incorporating other AI technologies into the CBR process (as suggested by Watson, [16]), e.g. when using a neural network in similarity assessment.

3 Activity Theory

In this paper, we propose the use of Activity Theory (AT) to support the design of CBR systems which take these problems into account. We can use AT to analyze the use of intelligent systems as instruments for achieving a predefined (by the human) goal in the work process and especially to understand the transformation of the artifact itself or the socio-technical system during this process. This could help us understand which types of explanations are expected. On the other hand, our knowledge about the work process can help us understand problems showing up in the use of the CBR system, supporting our (implicit or explicit) user model.

3.1 Basic Properties of AT

In this section, we give a short summary of aspects of AT that are important for this work. See [17] for a short introduction to AT and [18, 19] for deeper coverage. The theoretical foundations of AT in general can be found in the works of Vygotsky and Leont'ev [20–22].

Activity Theory is a descriptive tool to help understand the unity of consciousness and activity. Its focus lies on individual and collective work practice. One of its strengths is the ability to identify the role of material artifacts in the work process. An activity (Fig. 1) is composed of a subject, an object, and a mediating artifact or tool. A subject is a person or a group engaged in an activity. An object is held by the subject and motivates activity, giving it a specific direction.

Some basic properties of the AT are:

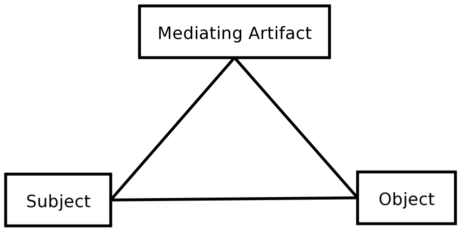

Fig. 1. Activity Theory: The basic triangle of Mediation.

– **Hierarchical structure of activity:** Activities (the topmost category) are composed of goal-directed actions. These actions are performed consciously. Activities, in turn, consist of non-conscious operations.
– **Object-orientedness:** Objective and socially or culturally defined properties. Our way of doing work is grounded in a praxis which is shared by our co-workers and determined by tradition. Praxis forms the look of artifacts, and by these the artifacts are passing on a specific praxis.
– **Mediation:** Human activity is mediated by tools, language, etc. The artifacts as such are not the object of our activities, but appear already as socio-cultural entities.

Taking a closer look on the hierarchical structure of activity, we can find the following levels:

– **Activity:** This is the topmost level. An individual activity is for example to check into a hotel, or to travel to another city to participate at a conference. Individual activities can be part of collective activities, e.g. when someone organizes a workshop with some co-workers.
– **Actions:** Activities consist of a collections of actions. An action is performed consciously, the hotel check-in, for example, consists of actions like presenting the reservation, confirmation of roomtypes, and handover of keys.
– **Operations:** Actions consist themselves of collections of non-conscious operations. To stay with our hotel example, writing your name on a sheet of paper or taking the keys are operations. That operations happen non-consciously does not mean that they are not accessable.

It is important to note that this hierarchical composition is not fixed over time. If an action fails, the operations comprising the action can get conceptualized, they become conscious operations and might become actions in the next attempt to reach the overall goal. This is referred to as a breakdown situation. In the same manner, actions can become automated when done many times and thus become operations. In this way, we gain the ability to model a change over time.

3.2 Action Cycle

Fjeld et al. [23] describe the notion of an action cycle for goal-directed pragmatic action. Their model is based on Hacker's work on Activity Theory and occupational psychology [24]. An action cycle (Fig. 2) consists of:

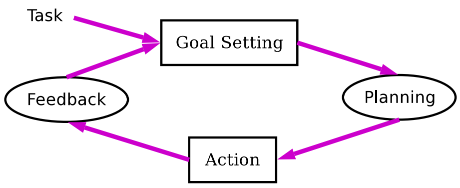

Fig. 2. The Action Cycle (according to [23]).

1. **Goal setting:** The initial goal for performing a task is set.
2. **Planning:** Plan how to achieve the goal set, including the selection of tools and choice of actions required.
3. **Action:** Consciously performed mental or physical actions.
4. **Feedback:** Controlling whether the anticipated goal was achieved, and if not, identifying the reasons for failure.

The starting point for this cycle is the activity, here identified as a task to perform.

We have a twofold interest in the notion of an action cycle when designing CBR systems. First, it is useful for identifying and modeling parts of the workplace activity where the CBR system is performing a task "on its own" with basically no human interaction. Here we try to translate the human activity as a whole into the artifact. Second, the CBR system can act as a supportive agent for the human performing the task. This is of special importance in situations where both the CBR system acquires knowledge and the human user gains an insight into the working of the artifact. For example, when the CBR system is introduced into a new workplace situation and has to acquire case knowledge at the same time as the user is gaining confidence in using it.

This second aspect is important for the issue at hand: identifying different explanatory needs. Given enough data, an analysis of the workplace situation based on the notion of an action cycle will explicitely model the goal settings (thereby identifying different user goals) as well as typical problems in the execution of the cycle (thereby identifying different needs for explanation in breakdown situations).

4 Using AT in Knowledge-Intensive CBR

We are considering knowledge-intensive CBR systems which incorporate both case knowledge, task knowledge, and domain knowledge in one single system.

Later in this section we will show how AT helps us in defining what to include in the design of the system's various knowledge containers [25, 26].

4.1 Choosing a Helpful Explanation

As pointed out before, the presentation of a single case may not suffice to give the user an explanation of the solution found. What constitutes a good explanation is to a large degree dependent on the goals of the user [11]. We therefore have to identify the possible goals of the user to the largest extent possible. This might be based on stereotypes of users, [27], for example in recommender systems, or on a survey of the concrete work situation the system is going to be embedded into [28], for example for experience management.

When using Activity Theory to model the work process, [29, 23], we can identify different types of activities the users are involved in. This helps us to understand the goals the user has when accessing the system, thereby also identifying the type of explanations necessary. We can use this analysis to guide us when defining the knowledge model of the system. This makes sure that a useful explanation can be given when requested by the user. For example, if in an AT based analysis of a learning situation we see that students tend to relate the knowledge to be acquired to a different domain, then the system should be capable of using these analogies as a basis for the explanation given.

4.2 Example Application

As mentioned in the introduction, an Intelligent System should adapt to its usage over time. Lets consider a CBR system for decision support. The user gets filled-out applications for credit cards and has to rate the creditability of the applicants. In the beginning, the user is likely to be interested in a detailed description of the results found. This is both due to the fact that the user has to learn to trust the system's capabilities, and that working with the system is quite new. In the language of Activity Theory, the user will perform mostly conscious actions and has not yet operationalized parts of the work process and/or the interaction with the system.

Over time, when the system offers correct or useful solutions, the user will both trust the system more, thereby eliminating his need to assess the reasoning process, and operationalize the work with the system. A lengthy presentation of the results by our CBR system will disrupt this process of operationalization. A mixed-initiative system will probably decrease its own activity: it no longer has to remind the user to do certain actions since they have become part of automated operations.

Let us now consider what happens when a breakdown situation occurs. For example, the user is now involved with free-form applications, e.g. by telephone. The problem viewed from the CBR system remains the same at first, but the work-flow of the user is disturbed so that he conceptualizes the operations he performs again. The CBR system must recognize this problem and change its behavior.

If we include relevant parts of the AT dealing with breakdown situations explicitly in our (general) user model, we can discover that such a situation has occurred (e.g. because the user is increasingly requesting explanations in situations where he has not done so before). The system can now adapt itself to give more detailed explanations and be more pro-active again. Likewise, the breakdown situation is a hint that the system's knowledge may no longer be adequate, and the search for solutions might have to be broadened until enough new (case) knowledge is acquired.

5 Ongoing and Future Work

The integration of an a posteriori method of analysis with design methodologies is always challenging. One advantage AT has is that it is process oriented, which fits nicely to a view on systems design where the deployed system itself is not static and where the system is able to incorporate new knowledge over time [30]. Activity Theory has its blind spots, and our goal is therefore to combine AT with other theories into a framework of different methods supporting the systems design process [31].

Focussing on AT, the relationship between the action cycle and the CBR process has to be examined further. Likewise, a methodological approach to integrate the findings of a work process analysis into the different knowledge containers of a CBR system has to be developed. Further on, a real world application of the outlined approach is necessary to asses its practicability. The method of choice is a qualitative study where the methodology has to be co-evolved with the ongoing project.

6 Conclusion

We have pointed out the importance of modeling the user's potential goals when defining which types of explanation an intelligent system can give. We have further introduced Activity Theory as a means of achieving this objective. Likewise we have suggested that the hierarchichal model of activity can be modeled in such systems to enable it to adapt to changes in usage over time. The action cycle as a model for goal directed pragmatic action can help identifying possible breakdown situations and resulting needs for specific types of explanation from the supporting intelligent system.

In our opinion, CBR system design methodologies will in the long run benefit from the integration of theories from occupational psychology and information systems design. They offer a supplement to cognitive science based approaches and integrate an understanding of organizational issues into the CBR process itself. What is and what is not a good explanation is dependent both on the individual user and her capabilities and on the organizational context. Therefore, we think it is necessary to achieve an understanding of the workplace situation the CBR system is going to be embedded into to deliver explanations of results which satisfy the user's needs.

References

1. McSherry, D.: Mixed-Initative Dialogue in Case-Based Reasoning. In: Workshop Proceedings ECCBR-2002, Aberdeen (2002)
2. Totterdell, P., Rautenbach, P.: Adaptation as a Design Problem. In: Adaptive User Interfaces. Academic Pres (1990) 59–84
3. Althoff, K.D., Wilke, W.: Potential Uses of Case-Based Reasoning in Experience Based Construction of Software Systems and Business Process Support. In Bergmann, R., Wilke, W., eds.: Proc. of the Fifth German Workshop on Case-Based Reasoning, Centre for Learning Systems and Applications. Number LSA-97-01E (Technical Report), University of Kaiserslautern (1997) 31–38
4. Aamodt, A.: A Knowledge-Intensive, Integrated Approach to Problem Solving and Sustained Learning. PhD thesis, Norwegian Institute of Technology, Department of Computer Science, Trondheim (1991)
5. Leake, D.B.: CBR in Context: The Present and Future. AAAI Press/MIT Press, Menlo Park (1996) 3–30
6. McSherry, D.: Interactive Case-Based Reasoning in Sequential Diagnosis. Applied Intelligence **14** (2001) 65–76
7. Cunningham, P., Doyle, D., Loughrey, J.: An Evaluation of the Usefulness of Case-Based Reasoning Explanation. [34] 122–130
8. Aamodt, A.: Explanation-Driven Retrieval, Reuse, and Learning of Cases. In Richter, M., Wess, S., Althoff, K.D., Maurer, F., eds.: EWCBR-93: First European Workshop on Case-Based Reasoning. Number SR-93-12 (Technical Report), Kaiserslautern, University of Kaiserslautern (1993) 279–284
9. Aamodt, A.: Explanation-driven Case-Based Reasoning. In: Topics in Case-Based Reasoning: Proceedings EWCBR 1994. LNAI, Springer (1994) 274–288
10. Leake, D.B.: Goal-Based Explanation Evaluation. In: Goal-Driven Learning. MIT Press, Cambridge (1995) 251–285
11. Frode Sørmo and Jörg Cassens: Explanation Goals in Case-Based Reasoning. This volume (2004)
12. McSherry, D.: Explanation in Case-Based Reasoning: an Evidential Approach. In Lees, B., ed.: Proceedings of the 8th UK Workshop on Case-Based Reasoning, Cambridge (2003) 47–55
13. Smyth, B., McClave, P.: Similarity vs. Diversity. [33] 347–361
14. McGinty, L., Smyth, B.: On the Role of Diversity in Conversational Recommender Systems. [34] 276–290
15. Smyth, B., McGinty, L.: The Power of Suggestion. [35] 127–132
16. Watson, I.: Case-Based Reasoning is a Methodology, not a Technology. Knowledge-Based Systems (1999) 303–308
17. Nardi, B.A.: A Brief Introduction to Activity Theory. KI – Künstliche Intelligenz (2003) 35–36
18. Bødker, S.: Activity theory as a challenge to systems design. In Nissen, H.E., Klein, H., Hirschheim, R., eds.: Information Systems Research: Contemporary Approaches and Emergent Traditions. North Holland (1991) 551–564
19. Nardi, B.A., ed.: Context and Consciousness. MIT Press, Cambridge, MA (1996)
20. Vygotski, L.S.: Mind in Society. Harvard University Press, Cambridge, MA (1978)
21. Vygotski, L.S.: Ausgewählte Schriften Bd. 1: Arbeiten zu theoretischen und methodologischen Problemen der Psychologie. Pahl-Rugenstein, Köln (1985)
22. Leont'ev, A.N.: Activity, Consciousness, and Personality. Prentice-Hall (1978)

23. Fjeld, M., Lauche, K., Bichsel, M., Voorhoorst, F., Krueger, H., Rauterberg, M.: Physcial and Virtual Tools: Activity Theory Applied to the Design of Groupware. CSCW **11** (2002) 153–180
24. Hacker, W.: Allgemeine Arbeitspsychologie. Huber (1998)
25. Leake, D.B., Kinley, A., Wilson, D.C.: Learning to integrate multiple knowledge sources for case-based reasoning. [32] 246–251
26. Richter, M.: The Knowledge contained in Similarity Measures. Invited talk at ICCBR-95 (1995)
27. Rich, E.: User Modeling via Stereotypes. Cognitive Science (1979) 329–354
28. Tautz, C., Althoff, K.D., Nick, M.: Learning from Project Experience – A Case Study on Experience Factory. In: Proc. GI-Fachgruppentreffen Maschinelles Lernen (FGML'00), Sankt Augustin, GMD (2000)
29. Korpela, M., Mursu, A., Soriyan, H.A.: Information Systems Development as an Activity. CSCW **11** (2002) 111–128
30. Aamodt, A.: Knowledge Acquisition and Learning by Experience – The Role of Case-Specific Knowledge. In Tecuci, G., Kodratoff, Y., eds.: Machine Learning and Knowledge Acquisition – Integrated Approaches. Academic Press (1995) 197–245
31. Cassens, J.: A Work Context Perspective on Mixed-Initiative Intelligent Systems. In: Proceedings of the IJCAI 2003 Workshop on Mixed-Initiative Intelligent Systems, Acapulco (2003) 30–35
32. Pollack, M., ed.: Proceedings of the 15th International Conference on Artificial Intelligence (IJCAI 1997), Nagoya, Morgan Kaufmann (1997)
33. Aha, D.W., Watson, I., eds.: Case-Based Reasoning Research and Development: Proceedings ICCBR 2001. Number 2080 in LNAI, Vancouver, Springer (2001)
34. Ashley, K.D., Bridge, D.G., eds.: Case-Based Reasoning Research and Development: Proceedings ICCBR 2003. Number 2689 in LNAI, Trondheim, Springer (2003)
35. Gottlob, G., Walsh, T., eds.: Proceedings of the 18th International Conference on Artificial Intelligence (IJCAI 2003), Acapulco, Morgan Kaufmann (2003)