

Modeling Stories in the Knowledge Management Context to Improve Learning Within Organizations

Stefania Bandini and Federica Petraglia and Fabio Sartori

Abstract Knowledge Management has been always considered as a problem of acquiring, representing and using information and knowledge about problem solving methods. Anyway, the complexity reached by organizations over the last years has deeply changed the role of Knowledge Management. Today, it is not possible to take care of knowledge involved in decision making processes without taking care of social context where it is produced. This point has direct implications on learning processes and education of newcomers: a decision making process to solve a problem is composed by not only a sequence of actions (i.e. the know-how aspect of knowledge), but also a number of social interconnections between people involved in their implementation (i.e. the social nature of knowledge). Thus, Knowledge Management should provide organizations with new tools to consider both these aspects in the development of systems to support newcomers in their learning process about their new jobs. This paper investigates how this is possible through the integration of storytelling and case-based reasoning methodologies.

1 Introduction

Storytelling is a short narration through which an individual describes an experience on a specific theme. In this way, the human being is motivated to focus the attention on his/her own knowledge about the specific theme that is the subject of narration [5]. Within organizations, storytelling can be considered an effective way to treasure the knowledge that is produced from the daily working activities. For

Stefania Bandini
CSAI, Viale Sarca 336, 20126 Milan (ITALY), e-mail: bandini@csai.disco.unimib.it

Federica Petraglia
DISCO, Viale Sarca 336, 20126 Milan (ITALY), e-mail: federica.petraglia@csai.disco.unimib.it

Fabio Sartori
DISCO, Viale Sarca 336, 20126 Milan (ITALY), e-mail: sartori@disco.unimib.it

example, Roth and Kleiner [9] have analyzed how the adoption of storytelling allows an organization to be more conscious about its overall knowledge, to share knowledge among all the people involved in its generation, to treasure and disseminate new knowledge originated by the sharing of different stories. The adoption of storytelling can promote the development of new professional contexts where different professionals collaborate to solve common problems, share experiences, explicit and implicit assumptions and understandings in order to improve the global capability of the organization to transform, create and distribute knowledge. In this sense, Knowledge Management can profitably exploit the storytelling as a way to make explicit the individual experiences, skills and competencies, promote the negotiation processes through dialogues among people involved, support the reification of new knowledge in order to make it available for the future and help newcomers in the learning process about his/her job through the analysis of the problem-solving strategies and social context represented by the stories. In this paper, we present a conceptual and computational framework for supporting continuous training within wide organizations, in the learning by doing [12] context. This approach is based on the integration of storytelling and case-based reasoning [10] methodologies: the former allows to manage a decision making process like a story that describes problem characteristics and what kind of communications among people and problem solution strategies can be applied to solve it; the latter is a very useful and efficient mean to compare stories (i.e. cases) finding solutions to new problems by reusing past experiences. Next section is devoted to make clear how learning by doing, storytelling and case based reasoning can be put together; first, a brief introduction to learning by doing and historical/methodological motivations to adopt it as a good paradigm for supporting continuous learning in organization is given. Then, its relationship with storytelling and case based reasoning is explored in details, to show how storytelling is the theoretical bridge between the need to support learning by doing through computer-based tools and one of the most suitable computer science paradigm for this scope. In section 3, an application of the framework to the SM-MART (System for Mobile Maintenance Accessible in Real Time) project will be briefly introduced, to show its effectiveness in representing problem solving strategies of experts in the form of stories that can be archived as cases into a case base and used as pieces of experience to build newcomers training systems, according to the learning by doing approach. In particular, the domain of the SMMART project is the troubleshooting of trucks (thanks to the collaboration with Volvo Trucks), thus the stories involved concern the experience owned by expert mechanics and the system is devoted to support newcomers of a truck manufacturers after-sales department. Finally, conclusions and future work will be briefly pointed out.

2 Learning by Doing, Storytelling and Case Based Reasoning

Contemporary socio-cultural context supports the idea of knowledge acquisition and management, not only as development of organisation, policy, methods of knowl-

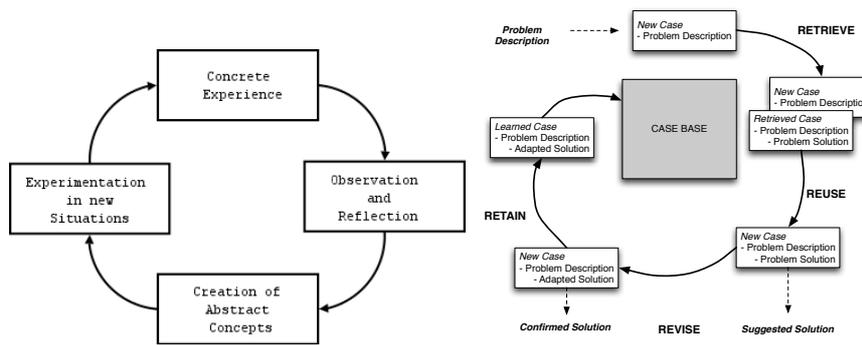


Fig. 1 On the left, the four steps in learning by doing methodology; on the right, the 4R's cycle of CBR applications

edge diffusion, but also as a community's benefit. Starting from these considerations, we reflect about the concept of continuous learning within organizations and how to support it. In particular, we focus the attention on learning by doing paradigm. Learning by Doing is based on well known psycho-pedagogical theories, like cognitivism and behaviourism, which are devoted to point out the role of practice in humans' intellectual growth and knowledge improvement. In particular, this kind of learning methodology refuses the typical idea that concepts are more fundamental than experience and, consequently, that only a solid set of theoretical notions allows to accomplish a given task in a complete and correct way. Learning by doing methodology states that the learning process is the result of a continuous interaction between theory and practice, between experimental periods and theoretical elaboration moments. Learning by doing can be articulated into four distinct steps (see the left part of Figure 1), where practical phases (i.e. Concrete Experience and Experimentation) are alternated with theoretical ones (i.e. Observation and Reflection and Creation of Abstract Concepts): starting from some kind of experience, this experience originates a mind activity that aims to understand the phenomenon; this step ends when a relation between the experience and its results (typically a cause-effect relation) is discovered that can be generalized to a category of experiences similar to the observed phenomenon. The result is a learned lesson that is applicable to new situations which will eventually occur in the future.

In our framework, a concrete experience can be represented by a story, which represents a decision making process about a problem to be solved. This story should give to a newcomer an idea of how a critical situation could be tackled, according to the knowledge owned by experts. Moreover, it could give indications about who could help him/her in case of need.

Stories can be archived as cases according to the case-based reasoning (CBR) paradigm. Case Based Reasoning is an Artificial Intelligence method to design knowledge management systems, which is based on the principle that *similar problems have similar solutions*. For this reason, a case based system doesn't require a

complete and consistent knowledge model to work, since its effectiveness in finding a good problem solving strategy depends typically on how a problem is described. Thus, CBR is particularly suitable to adopt when domains to tackle are characterized by episodic knowledge and it has been widely used in the past to build decision support systems in domain like finance [4], weather forecasting [8], traffic control [7], chemical product design and manufacturing [3], and so on. A case, is a complete representation of a complex problem and it is generally made of three components: description, solution and outcome [10]).

The main aim of CBR is finding solutions to new problems through the comparison of it with similar problems solved in the past, as shown in the right part of Figure 1, that is the well known 4Rs cycle by Aamodt and Plaza [1]: the comparison is made according to a *retrieval algorithm* working on problem features specified in the description component. When an old problem similar to the current one is retrieved, its solution is reused as a solving method for the new problem. The solution can be then *revised* in order to fit completely the new problem description and finally *retained* in the case base to become a sort of new lesson learned. In the retained case, the outcome component gives an evaluation about the effectiveness of the proposed solution in solving the problem. In this way, new cases (i.e. stories) can be continuously created and stored to be used in the future, building up a memory of all experiences that can be used as newcomer training tool.

Starting from concrete experiences newcomers can learn decision making processes adopted within the organization they are introducing quicker than studying manuals or attending courses. Moreover, the comparison between their own problem solving strategy and the organization one, represented by the collection of stories, stimulates the generalization of problems and consequently the reflection about general problem solving methods, possibly reducing the time period to make the newcomers able to find effective solutions.

CBR is one of the most suitable Artificial Intelligence methods to deal with learning by doing [11], due to the perfect match between their cycles of life. In particular: the description of a new case can be a way to represent experimentation in new situations, since the aim of CBR is to solve a new problem exploiting old solutions to similar problems. Thus, a new case is the attempt to apply past experiences to a new concrete situation in order to validate a problem solving strategy, as the experimentation in new situations is a way in the learning by doing context to test the generation of abstract concepts starting from already validated concrete experiences; a retrieved case in the case base represents a concrete experience in the learning by doing framework; retrieval, reuse and revise are the CBR phases during which a solution to a new problem is found and reused by comparison with similar past problems and then adapted to fit completely the critical situation defined by problem description. Thus, they can be exploited to model the theoretical steps of learning by doing methodology (i.e. Observation/Reflection and Creation of abstract concepts), through which a newcomer finds a general way to tackle a problem starting from a set of existing examples; finally, the retained case in the CBR paradigm is the completion of the initial problem to be solved with the optimal solution ob-

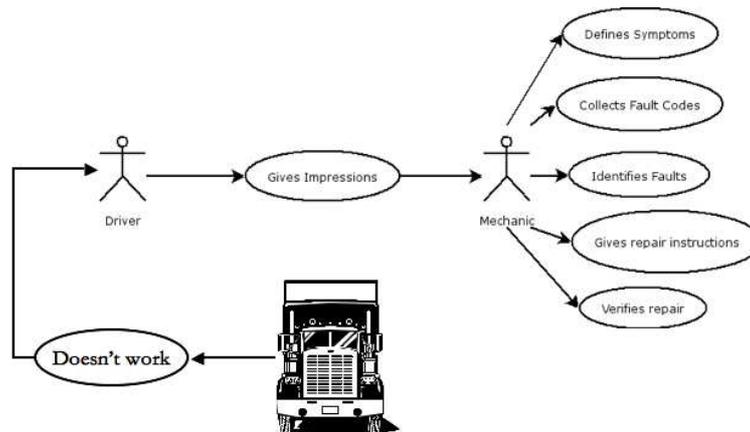


Fig. 2 A typical story about a truck troubleshooting session

tained at the end of the CBR cycle, thus it represents a new instance of the initial experimentation in new situations.

Moreover, since the concept of story can be used to describe both a case in the CBR paradigm and a concrete experience in the learning by doing methodology, in our opinion, storytelling is the optimal connection between a case-based support to the development of training systems for newcomers and the learning by doing context.

3 The SMMART Project

SMMART (System for Mobile Maintenance Accessible in Real Time) is a research project funded by the European Community¹ that aims to develop a decision support system for supporting experts of Volvo Truck², a world leader in the manufacturing of trucks in troubleshooting vehicle problems. To this aim, a case-based reasoning module of the final system is going to be designed and implemented in order to detect the most probable faulty engine component on the basis of a given set of information, which can be archived as a story.

The narration (see Figure 2) about the problem starts when a driver recognizes that a problem arose on his/her truck. For example, a light of the control panel turns on or some unpredictable event happens (e.g. smoke from the engine, oil loss, and noises during a break and so on). Thus, the driver contacts the truck after sale assistance to obtain problem solution. The mechanic who receives the truck is responsible for making a detailed analysis of it by taking care of driver impressions,

¹ Project number NMP2-CT-2005-016726

² <http://www.volvo.com/trucks/global/en-gb/>

testing it and collecting information coming from on-board computers. Then, he/she has to find the fault, repair it and verify that the problem has been solved before the truck leaves the workshop. In the following, a detailed description of how such stories have been represented and used in the context of SMMART is given in terms of case structure and similarity functions developed.

3.1 The Case Structure: a story in the SMMART Context

The final scope of the CBR tool is to identify the most probable truck faulty component (e.g. engine, gearbox), namely *High Level Component* (HLC). The HLC is an indication where the real cause of the truck malfunction is: this is the *root cause* and it is detected by the mechanic manually or through the exploitation of traditional softwares used by Volvo repair shops. Anyway, the CBR systems archives all the information about the problem, in order to give a complete representation of the story involved, as shown in Figure 3: HLC and root cause represent the *solution* part of the case, while the problem analysis made by mechanic, that is represented as the *case description*, considers four main categories of information : *symptoms*, *fault codes*, *general context* and *vehicle data*.

Symptoms give qualitative descriptions of truck problems and their context. For example, the sentence “The truck cruise control fails to maintain set speed while driving uphill at -20C under heavy charge” specifies that a possible fault of the cruise control (i.e. the symptom) is detected when the road is not plane, the temperature is very low, and the truck is transporting a big load (i.e. the context). The same problem could be not detected under different conditions. Symptoms are grouped into a tree structure within the SMMART case description: currently, five levels are considered, but they could increase in the future.

Fault codes are quantitative information coming from on-board computers: when some event happens that possibly causes malfunctions, a fault code is generated and memorized to be used during troubleshooting sessions. A fault code is characterized by many fields, the most important are (1) the *Message Identifier* (MID), that specifies the on-board computer generating the error code (for example, the entries with MID 128 in Figure 3 identify the on-board computer monitoring the engine; for this reason, it can be deduced that the MID indirectly identifies a HLC), (2) the *Parameter Identifier* (PID), that specifies which component of the on-board computer has generated the Fault Code (this means that an on-board computer is characterized by the presence of many sensors, each of them devoted to monitor a specific part of the HLC under control), and (3) the *Failure Mode Identifier* (FMI) that identifies the category of the fault (electrical, mechanical, and so on). The main activity of the mechanic during the truck analysis is the correlation between symptoms and their fault codes: in this way, it is possible to identify the faulty component, to repair it and trying to verify if the problem has been solved by controlling if fault codes disappear when the truck is turned on.

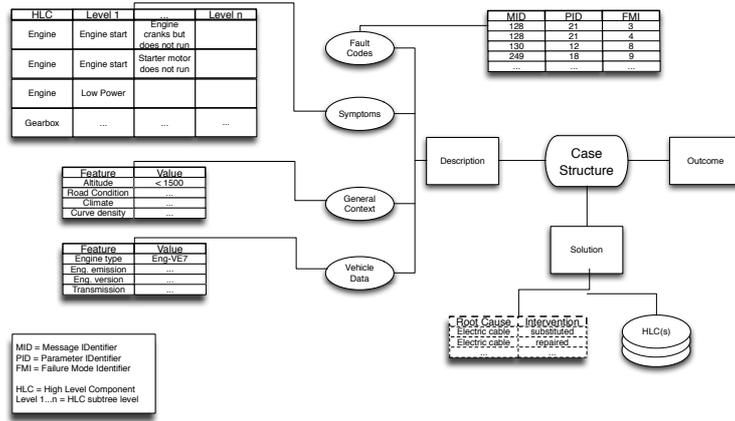


Fig. 3 The case structure of the SMMART project

Finally, general context and vehicle data contain information about driving conditions and truck characteristics respectively. These two kinds of information are not directly related to the fault generation, but they can be useful during the similarity calculus. For this reason, they have been included in the case description.

3.2 The Similarity Function: Retrieving Stories in the SMMART Context

When a new story is generated that represents the current problem (i.e. a problem without solution), it is represented as a case and properly described in terms of symptoms, fault codes and context. Then, it is compared with other cases already solved in the past in order to find similar story descriptions: the solution of most similar story is then reused as a starting point for deriving the solution to the current problem, suggesting in this way how to detect the most probable root cause. The comparison between stories is done according to a retrieval algorithm based on the K-Nearest Neighbor approach [6].

Given the current case C_c , for which no solution is given, the goal of the retrieval algorithm is to propose a possible solution (i.e. a HLC together with a possible root cause) by comparing its description C_c^d with the descriptions C_p^d of each case C_p solved in the past and included in the case base. The similarity among cases is calculated with a composition of sub functions, as described by the following formula

$$SIM(C_c, C_p) = \frac{k_1 * SIM_S + k_2 * SIM_{FC} + k_3 * SIM_{Vehicle} + k_4 * SIM_{GenContext}}{4}$$

where:

- $k_1...k_4$ are configurable weights and $k_1 + k_2 + k_3 + k_4 = 1$;
- SIM_S , SIM_{FC} , $SIM_{Vehicle}$ and $SIM_{GenContex}$ are in $[0.0 \dots 1.0]$.

SIM_S is the similarity among the two sets of symptoms of current case and past case, named S_c and S_p respectively: for each symptom A in the current case, the algorithm finds the closest symptom B (possibly the same as symptom A) in the past case, belonging to the same sub-tree, having the HLC name as its root.

The function $dist(A,B)$ gives the minimum number of arcs that separates A and B in the symptoms tree and it is used for calculating the similarity. Similarity between symptom A and symptom B ($A \in S_c$ and $B \in S_p$) is

$$sim(A,B) = (1 - dist(A,B)/dmax)$$

where dmax is the constant maximum distance possible between two nodes in the tree (in the current symptom tree dmax=5). Similarity between symptom A and symptom B is modified by the conditions under which the symptoms occurred; the algorithm evaluates the degree of similarity between the two sets of conditions and modifies the value of $sim(A,B)$ consequently.

The similarity among symptoms SIM_S is the sum of all the $sim(A,B)$ normalized with the number noc of couples of symptoms considered and eventually penalized if the two cases are different in number of symptoms. The final formula is:

$$SIM_S = (SIM_S/noc) * (1 - Penalty)$$

$$\text{where } Penalty = \frac{(\#S_c + \#S_p - 2 * noc)}{\#S_c + \#S_p}$$

SIM_{FC} is the similarity among the two sets of fault codes (FCs) calculated on each HCL group of FCs (FCs grouped by high level component): the relation between FCs and HLCs is given by mapping the MID of each FC to the HLC name. Doing so, different MIDs (that means FCs coming from different processing units) can be associated to the same HLC. If a FC has not any MID-HLC mapping entry, the FC will be related to a fictitious HLC, called HLC_0 : in this way, also Fault Codes which cannot be linked directly to a specific HLC can be compared, with benefits from the final similarity point of view.

When all the Fault Codes of both C_c and C_p have been grouped in the FC_c and FC_p sets respectively, the algorithm compares the information they contains: the similarity $sim(A, B)$ between two Fault Codes belonging to C_c and C_p depends on their PID and FMI values. The similarity values are fixed and they have been determined with the collaboration of Volvo Truck experts. The similarity among fault codes SIM_{FC} is the sum of all the $sim(A, B)$ normalized with the number noc of couples of fault codes considered and eventually penalized if the two cases are different in the number of fault codes; The final formula is:

$$SIM_{FC} = (SIM_{FC}/noc) * (1 - Penalty)$$

$$\text{where } Penalty = \frac{(\#FC_c + \#FC_p - 2 * noc)}{\#FC_c + \#FC_p}$$

$SIM_{Vehicle}$ is the similarity among the two vehicle characteristics: each possible feature involved in vehicle description is linked to a weight. These weights are used in the computation of the similarity between vehicle descriptions given in the current case and in the past case.

$SIM_{GenContext}$ is the similarity among the two general contexts. Since items describing general contexts are assigned qualitative values (i.e. strings), these values are preprocessed according to an opportune mapping function to be converted an integer values.

4 Conclusions

This paper has presented a framework to support learning by doing within organizations; this framework is based on the integration of storytelling and case based reasoning methodologies. Storytelling has been chosen due to its capability of taking care of different kinds of knowledge in the description of working experiences and presenting important pieces of expertise to newcomers in wide organizations; according to Atkinson [2]:

Storytelling is a fundamental form of human communication [...] We often think in story form, speak in story form, and bring meaning to our lives through story. Storytelling, in most common everyday form, is giving a narrative account of an event, an experience, or any other happening [...] It is this basic knowledge of an event that allows and inspires us to tell about it. What generally happens when we tell a story from our life is that we increase our working knowledge of ourselves because we discover deeper meaning in our lives through the process of reflecting and putting the events, experience, and feelings that we have lived into oral expression.

On the other hand, case based reasoning is one of the most suitable Artificial Intelligence paradigms to deal with episodic and heterogeneous knowledge and consequently, in our opinion, it is probably the best approach to manage unstructured narrations about expertise and problem solving strategies. The proposed framework provides newcomers with a complete representation of the competencies developed by experts over the years. Thus, they can increase their experience about the problem solving strategy used inside the organization as well as the understanding about who are the people to contact in case of need (i.e. the experts who solved similar problem in the past).

In order to test the effectiveness of our approach, its application in the context of the SMMART project has been briefly introduced. It is important to highlight that the SMMART projects aims at the development of a CBR module to identify the most probable faulting component of a truck by means of a specific retrieval algorithm: the solution of the CBR engine is not subject to adaptation, since it is not the real solution of the mechanic troubleshooting session. A mechanic exploits this solution as a starting point to make deeper analysis looking for the root cause. Anyway, once the mechanic detect the real cause(s) of the problem, the CBR module retains it in the case base together with all other information, in order to give

a complete representation of the story related to that troubleshooting session. From the learning by doing point of view, the case base composed of all the stories about past troubleshooting sessions is a very important source of knowledge for newcomers; they could be solicited to solve a problem by specifying what are the symptoms and the related fault codes. Then they could try to identify faulty components and then compare their solution with the one proposed by the system, with an immediate evaluation of their own capability to learn expert mechanics decision making processes and identification of points they have to work on, maybe asking directly to the people who solved past problems. In this way, experience and knowledge created by the organization over the years and captured by the CBR system could be used as a very important training method alternative to the more traditional ones.

Future works are devoted to verify the applicability of the proposed methodology in building supporting systems for learning by doing in other complex contexts.

References

1. Aamodt, A. and Plaza, E.: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*, Vol. 7, No. 1, pp. 39-59, (1994).
2. Atkinson, R.: *The Life Story Interview*, Sage University Papers Series on Qualitative Research Methods, vol. 44, SAGE Publications, Thousand Oaks, CA, (1998).
3. Bandini, S., Colombo, E., Sartori, F., Vizzari, G.: Case Based Reasoning and Production Process Design: the Case of P-Truck Curing, In: *ECCBR Proceedings*. Volume 3155 of Lecture Notes in Computer Science, pp. 504517. Springer, Heidelberg (2004).
4. Bonissone, P. P., Cheetham, W.: Financial Application of Fuzzy Case-Based Reasoning to Residential Property Valuation. In *Proceedings of the 6th IEEE International Conference on Fuzzy Systems*, Vol. 1, pp 37-44, (1997).
5. Bruner, J.: The Narrative Construction of Reality. *Critical Inquiry*, 18, pp. 1-21, (1991).
6. Finnie, G., Sun, Z.: Similarity and Metrics in Case-based Reasoning. *Intelligent Systems*, 17(3), pp 273-285, (2002).
7. Gomide, F., Nakamiti, G.: Fuzzy Sets in Distributed Traffic Control, 5th IEEE International Conference on Fuzzy Systems - FUZZ-IEEE 96, pp 1617-1623, New Orleans - LA - EUA, (1996).
8. Hansen, B.K., Riordan, D., Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory, Workshop on Soft Computing in Case-Based Reasoning, 4th International Conference on Case-Based Reasoning (ICCB01), Vancouver, (2001).
9. Kleiner, A. and Roth, G., How to Make Experience Your Companys Best Teacher, *Harvard Business Review*, Vol. 75, No. 5, p 172, (1997).
10. Kolodner, J., *Case-Based Reasoning*, Morgan Kaufmann, San Mateo (CA), (1993).
11. Petraglia, F. and Sartori, F., Exploiting Artificial Intelligence Methodologies to Support Learning by Doing Within Organisations., In G. Hawke and P. Hager (eds.): *Proceedings of RWL405-The 4th Int. Conference on Researching Work and Learning*, Sydney, December 2005, (2005).
12. Wenger, E., (1998) *Community of practice: Learning, meaning and identity*, Cambridge University Press, Cambridge, MA, (1998).