Krauskopf, K., Bertram, J., … & van Aalst, J. (2012). Memetic processes as a conceptual framework for idea improvement in knowledge building. In J. van Aalst et al. (Eds.), Proc. 10th Int. Conf. Learning Sciences (pp. 157-161). Sydney, Australia: ISLS.

Memetic Processes as Conceptual Framework for Idea Improvement in Knowledge Building

Karsten Krauskopf, Johanna Bertram, Stefan Huber, Knowledge Media Research Center, Schleichstr. 6, 72076 Tübingen, Germany,
k.krauskopf@iwm-kmrc.de, j.bertram@iwm-kmrc.de, s.huber@iwm-kmrc.de
Ya Ping Hsiao, Open University of the Netherlands, Valkenburgerweg 177, 6419 AT, Heerlen , The Netherlands, amy.hsiao@ou.nlKatherine Panciera, University of Minnesota, 4-192A KHKH, 200 Union St SE, Minneapolis, MN 55455, USA, katpa@cs.umn.edu
Nicole Sträfling, University of Duisburg-Essen, Forsthausweg 2, 47057 Duisburg, Germany, nicole.straefling@uni-due.de
Astrid Wichmann, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany, astrid.wichmann@rub.de
Jan van Aalst, Faculty of Education, University of Hong Kong, 323 Runme Shaw Building, Pokfulam Road, Hong Kong, China, vanaalst@hku.hk

Abstract: Idea improvement is a key characteristic of knowledge building where ideas are conceptualized as improvable and epistemic objects to sustained inquiry. However, despite its importance in this theory, little research exists that focuses on the process of idea improvement. In this paper we examine the development of ideas within a community of learners by proposing a conceptual framework to understand ideas as memes and investigate the memetic processes affecting them. We apply this framework to students' discourse through a three week unit from a Knowledge Forum® data set by following the survival paths (fitness) of memes. We use a mixed methods approach to identify memes, define quantitative indicators to calculate the memes' fitness and elaborate on the improvement of the fittest memes. Benefits and open questions related to our suggested framework are discussed.

Introduction

Knowledge building focuses on the production and continual improvement of ideas of value to a community (Bereiter, 2002). The principle of idea improvement is a feature of knowledge building, where ideas are conceptualized as "real," improvable, and epistemic objects; once an idea has been contributed to a shared knowledge space, students can ask about the assumptions underlying the idea, and modify it in various ways (Bereiter, 2002; van Aalst, 2006). One influential aspect in this process lies with the learners and their original ideas. Another influential aspect is independent from their creators: the expressed ideas can also be described as something with an "out-in-the-world existence" and a public life (Zhang, Scardamalia, Lamon, Messina, & Reeve, 2007). Once an idea has been made public it changes based on the input of the whole community. In short, the development of an idea becomes an evolutionary process with the learners as the main driving force of idea improvement (Hong & Sullivan, 2009). Research on knowledge building has recently been focussing on assessing aspects of idea improvement by observing students' learning processes. Zhang and colleagues (2007), for instance, assess idea improvement by judging the scientific acceptability of an idea. The authors identify "inquiry threads", which are sequences of notes that address the same problem or topic, and measure scientific levels of ideas. But assessing idea improvement in terms of scientificness does not allow judgement of relevance of ideas for the community. We propose a framework that allows us to focus on the relevance of an idea in a community of learners. This framework conceptualizes ideas as memes and then analyzes the development of these ideas through the learners' discourse. The goal of this study is to explore idea development in analogy to the development of memes within the context of Knowledge Forum®, an online learning environment (Scardamalia & Bereiter, 2003). Therefore, we examine the development of ideas within a community of learners through the lens of memes and memetic processes. We analyze students' discourse by following the survival paths of memes through a three week unit in the Knowledge Forum® data. Based on this examination, memes are identified in the students' discourse and described by quantitative indicators for these memes' fitness. Second, we qualitatively describe the developmental paths for the fittest memes. Finally, benefits, limitations, and open questions related to our suggested framework are discussed.

Ideas and Memetic Processes

Memes, as defined by Dawkins (1976), are units of cultural transmission and imitation that spread within a culture, as an equivalent of genes. Behavior, ideas, knowledge or fashion can be considered memes that propagate from person to person via imitation and variation (Dawkins, 1976). In some ways, the culture itself can be seen as a meme pool. Bereiter (2002) describes the cultural meme-pool as frequently changing over time, while new memes emerge, others mutate or their frequency changes. These processes can be seen as a cultural evolution. This concept of the cultural meme-pool contributes to Scardamalia and Bereiters' (2006) description of knowledge as an advancing concept with ideas emerging, others dying out, problems being solved and new

problems coming up. Because knowledge can thus be compared to the cultural meme pool which underlies evolutionary processes, idea improvement can be analyzed by comparing it to memes. The quality of an idea can be defined by evolutionary indicators such as its likelihood to survive. The overall survival rate of a meme can be understood as its *fitness* (Heylighen, 1999). This notion includes, however, that idea survival determined for an expressed meme by its repetition and variation as in evolution, does not necessarily progress but adapts. In analogy, the fitness of a meme cannot be used as a sole measure of quality but is a complementary step for focused qualitative analyses. With regard to the memetic processes co-determining a meme's fitness *transmission*, *variation*, and the resulting selection pressure are of relevance to our framework. Following the overview by Nye (2011) these can be defined as follows: "The core information of a meme is its semantic information. When semantic information changes, the meme has mutated or a new meme has been created. A meme reproduces when semantic information is replicated from one agent to another. [...] Conversely, identical physical transmissions change semantic meaning based on context and interpretation." (Nye, 2011, p. 14). Thus, a prerequisite for considering semantic information to be a meme is its ability to reproduce recursively within the respective environment. First, this implies that a meme must be expressed in behavior, or in the case investigated here, in written language. The transmission of this expressed meme can be considered most important to its reproduction. As Nye (2011, p. 18) puts it: "This definition is ontologically complete: semantic information is a meme within a society and environment if and only if it can recursively reproduce in that society and environment." Variation can be affected during expression and/or transmission by external factors, such as time pressure, or internal factors, such as limits to cognitive and motivational resources. Even though following a code, language, or procedure can reduce misunderstandings thanks to a given syntax, reducing the complexity in to a specific form of notation fosters ambiguous statements. This is why semantic variation appears likely to occur for symbolically expressed memes. Overall, this "noise" included in the transmission of memes and resulting variation puts the memes under a selection pressure and only certain memes survive. Therefore, analyzing the development of ideas, instead of the correctness of ideas, focuses more on the processes involved and less on the content. In sum, we assume that identifying memetic processes in knowledge building is a complementary approach that can help to tap more directly into the central process of idea improvement, i.e. how ideas spread and survive over time in a learners' discourse. We expect to (1) quantitatively describe indicators for idea improvement that can be inferred from the notion of memes and memetic processes. Furthermore, we expect, (2) based on this, to be able to qualitatively describe how ideas develop over time and to more directly describe their improvement as a process.

Method

Data

We analyzed a subset of the data from a study conducted by Niu and van Aalst (2009). Two classes of a tenth grade social studies course participated in that study. Each class was divided into groups of eight persons. In a short inquiry unit (three weeks) the students investigated general environmental problems such as pine beetle infestation. For this investigation they used the asynchronous online discourse environment Knowledge Forum® (Version 4.5, Scardamalia, 2003, see www.knowledgeforum.com). In Knowledge Forum®, students can contribute their ideas to the database in the form of written notes which make up discussion threads. Other students who have access to the database can revise these notes, reply to them, or contribute their own reflection. For this study we used a data set of eight students (one female, seven male). The participants worked on the topic of how to free a forest from a pine beetle infestation. They worked irregularly on the problem over a period of eight consecutive days and five additional posts were added a month later. A total of 128 posts were analyzed, consisting of all the posts from this group.

Data Analysis

The analytic procedure consisted of three steps. First, to identify the memes emerging in the data set a coding scheme was inductively developed and all notes were coded accordingly. Individual notes served as the unit of analysis. Secondly, quantitative information was extracted from the data in order to calculate a fitness score for each meme monitoring when memes were expressed and reproduced. Finally, after calculating and plotting the fitness scores over time, we "zoomed" in qualitatively on the fittest memes and followed their path in order to describe the development of the complexity and quality of the meme, and therefore, understand its variation.

Qualitative Analysis - Identifying Memes

We based our analysis on the notion that a meme can be expressed in written language, e.g. the notes posted by students. To identify the memes emerging in the students' discussion two independent raters (two of the authors) performed a qualitative content analysis following Mayring (2000) to inductively develop categories that would capture the central memes emerging in the data set. These initial sets of categories were discussed by the two raters and converged. Then, a third independent rater (first author) double checked the categories.

Disagreements were resolved by discussion resulting in the final 18 categories, example categories/memes are described in Table 1. Finally, all posted notes were coded by applying one or more of these categories. We differentiated between the repetition of the essential meme and variation, i.e. emergence of a different meme. *Repetition* (the same category applied) was coded when the analyzed note resembled the core idea of the previous note. Variation *within* a meme was also coded as repetition if the main message had not been changed. If two or more memes were contained in a note, multiple codes were applied. Coders followed the rule to apply the same code again if the content of the note resembled (copied) the same idea as the previous note. *Variation* (new category applied) was coded when different ideas emerged spontaneously or existing ones were integrated in a way that the original idea did not resemble the final idea. Coders followed the rule to apply a different code if individual ideas mutated (at the group level) or the new change to existing elements was introduced, also if the preceding idea was lost during the developmental process and the following idea did not resemble the preceding idea.

Table 1: Coding scheme - examples of memes.

Meme	Definition
Forest ecosystem	Impact of beetle infestation and counter measures on the forest ecosystem, question of balance between environmental costs and benefits
Pest control	Repetition of initial task or problem, the question of how to resolve the beetle infestation is discussed
Predators	Beetles should be killed by predators, from within (parasites) or outside (woodpeckers) and discussion about these predators

Quantitative Analysis

After we had identified memes in the qualitative coding procedure described above, we extracted the respective notes in which each meme occurred. Because the notes were posted over a period of 8 days at varying points in time, we defined a fixed time interval for which we aggregated the quantitative indicators. We performed the analysis with MS Excel. Due to the fact that the results do not differ meaningfully, we chose a 15 minutes interval, mainly to represent our analyses in an economic way (see Figure 1). The results did not differ meaningfully between smaller time intervals, but contained more specific information than larger time intervals. However, we are aware that the issue of conceptualizing time in asynchronous communication is a complex issue, in part because the time scale is a different one for every participant (Suthers, Dwyer, Medina, & Vatrapu, 2010). Thus, the time intervals created here are defined by our analysis not by the original time line of the data and therefore not readily interpretable.

Quantitative Indicators. As mentioned above, we adopted the formula proposed by Heylighen (1999) for the two external stages: *expression* and *transmission* (see Formula 1). For each of the two components a value can be computed and then combined into a fitness score for the respective meme a predefined time interval. The simplified formula for our combined fitness measure f has the following form

$$f(m, t) = E(m, t) * T(m, t) \quad (1)$$

The fitness f of a meme m for the expression stage in time interval t is denoted by $E(m, t)$ (expression) and the fitness in the transmission stage by $T(m, t)$ *(transmission)*. $E(m, t)$ describes expression, or how often a meme has been expressed by saving a new note or changes in a note that contain the respective meme m for time interval t. $T(m, t)$ describes transmission, or how often a posted note containing the respective meme m has been read by others during time interval t. The values of f are not interpretable in an absolute sense but only relative to other memes. Both terms can be larger than 1 and if one the terms reaches 0, the meme has been eliminated. This happens when a meme is not replicated further or when posted messages containing a meme are not read by other students anymore. In our case, however, notes were not deleted and stayed present in the Knowledge Forum® database to potentially be read. Therefore, in this study a meme can only be eliminated when $T(m, t)$ equals 0. To get a global fitness measure f_g, we calculated the mean fitness of a meme (Formula 2). In this formula the fitness of all time intervals is summed up and divided by the number of intervals n:

$$f_g(m) = \left(\sum_{i=1}^{n} E(m, t_i) * T(m, t_i) \right) \quad (2)$$

Results

Quantitative – Describing the Paths of Memes

Figure 1 depicts the results for the analysis applying Formula 1 to three sample memes (see Table 1). The last time interval available was excluded from analysis, because reading activity had ceased at the end of the course. What is most prominent in the data is that most of the memes have low fitness values over all the students' discourse and that there are three dominant memes, which also repeatedly show higher spikes than any other meme: *Forest Ecosystem, Pest Control*, and *Predators*. Towards the end of the discussion, however, after showing two peaks clearly higher than any other meme (time intervals 20 and 24), the forest ecosystem becomes visible as the fittest meme, which is also the only meme that is still "alive" at the very end. In contrast, in the beginning of the discussion other memes - concerned with describing pine beetles in general and as a threat - show higher fitness values (not included in Figure 1).

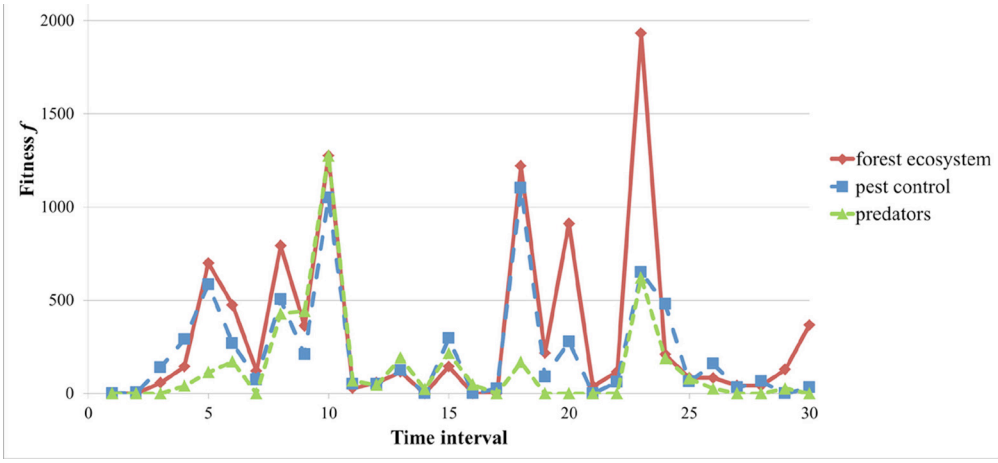

Figure 1. Fitness of sample memes within predefined 15 minute time intervals (see sect. *quantitative analysis*).

The average fitness values derived from Formula 2 mirror this data pattern. Here, also forest ecosystem, and pest control show high mean fitness. Additionally, the *Predator* meme had a high mean fitness, however as shown in Figure 1 it did not survive until the end of the discussion. Overall, we see from the quantitative data that the different solutions to the infestation problem are the fittest memes but rather we find a "struggle" between the elaboration on the initial task (pest control) and a counter argument; the impact of any counter measure on the forest ecosystem.

Qualitative Interpretation – How do the Fittest Memes Develop (Variation)

To qualitatively zoom in on the idea development in our data, we tried to describe the variation within a meme and the interactions between memes. The meme *Pest Control* was coded 25 times in the data. Different kinds of pest control mechanism were discussed, starting with a list of various pest control options, which were not discussed any further by the students. Some options from this list of measures were discussed and were judged to be no solution to the pine beetle problem, mainly again due to their impact on the forest as a whole (i.e. *forest ecosystem)*. The most prominently discussed solution was the idea of having woodpeckers prey on the beetles (i.e. *predators*). Furthermore, it was mentioned that these measures can kill the beetles, but can not prevent further epidemics. In other words, it was discussed that long-term solutions must be sought; for the current pest control there will be no solution which not has an impact on other animals or environment - the forest ecosystem. The meme *Forest Ecosystem* was coded 32 times in the data and it clearly interacts with pest control. The students' notes show that the students tried to optimize their ideas regarding the problem of the pine beetles. They not only tried to eliminate the beetles, but also considered the environmental and economic impact of the possible solutions. Every single idea regarding the elimination of the pine beetles was tested against the impact on the environment and most were also tested against the cost of the solution. Thus showing that the students wanted to improve their ideas.

Overall, the qualitative results show that the development of the solution of the pine beetle problem was not a straightforward one. The idea for the solution improved from simple versions like "kill beetles" to sophisticated ones like looking for long-term solutions. Along this development it seems that in particular the initial task of pest control and the counter argument of the impact on the rest of the (forest) ecosystem contend with each other. This contest seems mutually beneficial to the improvement of the solution. As above mentioned, the various pest control ideas had been made public and were changed based on the input of the whole community – in this case – the aforementioned impact on the forest ecosystem.

Discussion

The goal of this paper was to contribute to the study of idea improvement, a core component of knowledge building. Assessing idea improvement is of complex nature. One aspect that has been neglected so far is to

determine idea improvement by assessing the relevance of ideas in the community. The current results show that the way of data preparation and analysis can be a powerful tool to complement the summative appraisal of learners' discourse. We revealed that the group had broadened their discussion and improved their ideas in the sense of understanding the complexity of the initial problem by relating it to adjacent issues (e.g. *forest ecosystem*), which here were the fittest in the end.

In sum, we found that the few memes which survived the discourse were also more complex ones. So in this case, the fitness of the memes was related to their quality in terms of complexity. However, in respect to the potential of the described framework, there are also some limitations: first, our analyses zoomed in on just one aspect that could help assess idea improvement. In order to gain a complete understanding of what determines idea improvement, various facets need to be taken into account. Hence, future research should focus on combining several aspects (e.g. such as specificness, see Zhang at al., 2007) when assessing idea improvement. A second limitation lies in the reductionist approach taken: evolutionary processes only allow a very limited view on ideas; possibly it even neglects important aspects of ideas that are characteristic for idea development. We do not suggest that our framework alone is sufficient, but it can help to focus the questions for deeper analysis. Overall, we conclude that our framework provides a way to more directly research idea improvement as a process in which the community is a driving force.

References

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: L. Erlbaum Associates.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Hong, H.-Y., & Sullivan, F. R. (2009). Towards an idea-centered, principle-based design approach to support learning as knowledge creation. *Educational Technology Research and Development*, *57*, 613-627.

Heylighen, F. (1999). What makes a meme successful? Selection criteria for cultural evolution. *Proceedings of the 15th International Congress on Cybernetics, Namu*r, p. 418- 423.

Mayring, P. (2000). Qualitative content analysis. Forum: Qualitative Social Research, 1(2). Retrieved August 25, 2011, from http://www.qualitative-research.net/fqs-texte/2-00/02-00mayring-e.htm.

Niu, H., & van Aalst, J. (2009). Participation in Knowledge-Building Discourse: An analysis of online discussions in mainstream and honors social studies courses. *Canadian Journal of Learning and Technology 35*(1), http://www.cjlt.ca/index.php/cjlt/article/viewArticle/515/245.

Nye, B. D. (2011). *Modeling Memes: A memetic view of affordance learning*. Publicly accessible Penn Dissertations, paper 336. Retrieved August 25, from http://repository.upenn.edu/edissertations/336.

Scardamalia, M. (2003). Knowledge building environment: Extending the limits of the possible in education and knowledge work. In. A. DiStefano, K. E. Rudestam, & R. Silverman (Eds.), *Encyclopedia of distributed learning* (pp. 269 - 272). Thousand Oaks, CA: Sage Publications.

Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In Encyclopedia of education, second edition. (pp. 1370-1373). New York: Macmillan Reference, USA

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: theory, pedagogy, and technology. In R. K. Sawyer (Eds.), *Cambridge handbook of the learning sciences.* (p. 97-115). New York: Cambridge University Press.

Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer Supported Collaborative Learning*, *5*(1), 5-42.

van Aalst, J. (2006). Rethinking the nature of online work in asynchronous learning networks. *British Journal of Educational Technology, 37*, 279-288.

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Educational Technology Research & Development, 55*, 117-145.

Acknowledgments

This study resulted from an exploration the authors began during the Summer School "Making Sense of Social Media", co-sponsored by the Knowledge Media Research Center and Tuebingen ScienceCampus, August 1-4, 2011. The data were originally collected as part of a New Economy Collaborative Research grant from the Social Sciences and Humanities Research Council of Canada to Marlene Scardamalia (OISE/University of Toronto). The authors thank Hui Niu (Simon Fraser University) for permission to reanalyze the Knowledge Forum® database.