

A dynamic memory management system based on forgetting and recalling.

Michal PTASZYNSKI [†], Koichi SAYAMA [†], Rafal RZEPKA ^{††}, Kenji ARAKI ^{††}
 Otaru University of Commerce, Department of Information and Management Science [†]
 Hokkaido University, Graduate School of Information Science and Technology ^{††}

Abstract:

Thinking of how to improve naturalness and adequacy of speech in conversation software, authors propose a dynamic database management system. The system borrows some features of the forgetting mechanisms in humans. The core of the system, forgetting and recalling algorithms, depends on frequency of usage of context units and its emotive value derived from linguistic- and communication sciences based-analysis.

1. Introduction.

Forgetting is a constant and chronological process (Ebbinghaus, 1885) in which parts of knowledge become rearranged, inaccessible or inactive (Markovitch and Scott, 1988). Memory has also much in common with emotions, what was proved in psychology (Zitman, 2001) as well as in language acquisition research (Luk, 2002). Memories with stronger emotive affection are harder to forget (Wolfe, 2006).

We find forgetting a disadvantage or a defect in the human mind. However, scientists from different fields state that forgetting is a useful process facilitating effective knowledge acquisition (Markovitch and Scott, 1988), and actually helping people organize memory and remember about important things (Kahn et al., 2007).

Thinking of how to improve naturalness and adequacy of speech in conversation software, authors thought of a dynamic database management system based on the statements that forgetting is: a) a crucial element in the process of learning; b) dependant on frequency of used connections and c) emotiveness of connections; d) beneficial in organizing knowledge.

2. Description of the proposed system.

In a conversation program supported with an expanding database (DB, see Figure 1) former conversations will be archived just like in an internet messenger. When a user encounters the program a new conversation will start. The program will query the database searching for adequate context keywords (n-grams, conversational procedures) to produce an utterance. If it does not find any, it is considerable to let the program query internet for finding adequate knowledge (associations). Every entry after a query will be treated as a separate Context Unit (CU). For better distinction we can propose a division into Context Units in the form of a dialogue with a user (dCU) and those in the form of an association list gathered from the Internet (aCU). For the program though it will not make a difference what kind of CU is saved in the archive.

However, since the database will expand with every new CU, very soon the program would have to process enormous number of entries. As states Araki (2004), this is one of the difficult problems in processing context-bound sentences. To avoid it authors propose a system with an implemented algorithm of forgetting and recalling.

2.1 Algorithm of forgetting.

Forgetting means here that, as stated in the previous paragraphs, parts of knowledge will become inaccessible or inactive. Whether to “shade” a CU will be determined on two conditions: frequency of usage and emotive value.

Since there are context connections in the DB used more often and those used rarely, it is easy to base the forgetting process on the frequency of using the connections. That is, if a connection will not be in use for a specified amount of time, an event would be marked as inaccessible.

As mentioned above, forgetting in humans is based also on the emotive strength of neuronal connections. This mechanism will be transferred into the algorithm by adding emotive values to the archived items during the process of gathering information. This value will be derived from linguistic- and communication sciences based-analysis. Every utterance will be analyzed concerning differences in code elaboration and language usage. This states for such parts of speech like exclamations, interjections, syllable prolongations, exclamation marks, mimetic words (*gitaigo*), emotive lexicon, hypocoristics, and code sophistication (*suru/da* vs *masu/desu*) (Ptaszyński, 2006). The general emotive value of a CU will be equal to the approximation of emotive values of each utterance in the CU.

The contents to be forgotten by the program should be selected by taking into account those values along with the frequency of the links in use. The final decision whether to shade a CU would be based on those two conditions.

2.2 Algorithm of recalling.

Although there were proposals of creating an algorithm of forgetting before with some good ideas, it seems that the attention was always

concentrated only on forgetting during the learning process. Markovitch and Scott (1988) as well as later Ishikawa (1990) by forgetting meant that the unused links of large knowledge amount were ought to be forgotten in the meaning of deleting them from database. This way the notion of recalling information from faded database was ignored.

In the proposed algorithm the recalling process is nothing more than looping back to the process of gaining knowledge. When a user utters the program, a new conversation starts and the program searches through its database for an appropriate context data to respond for the dialog. If this ends in success, the connection to the used context is renewed. If not, the program starts the process of gaining knowledge (Internet). When a new associative context unit is gained, it is compared to the inactive forgotten archives. If the new gained CU covers or links to the inactive (forgotten) parts of knowledge in the database, those parts would be activated and their connections renewed. If the CU contains a new data, it is added and archived in the database as an active CU.

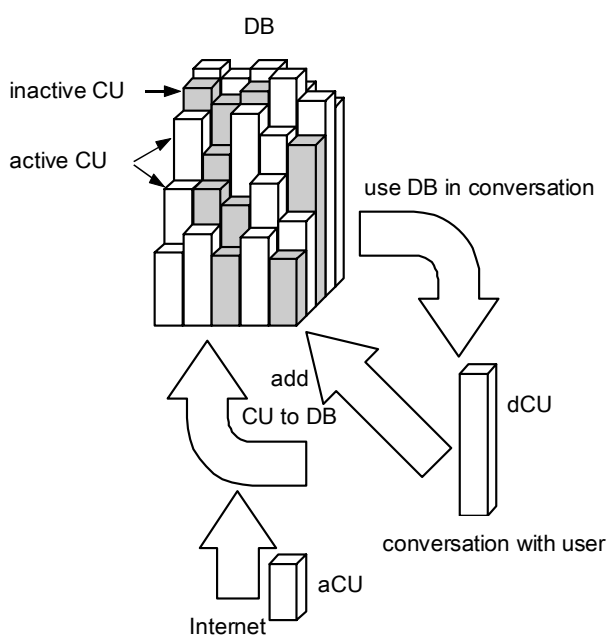


Figure 1 Simplified diagram of the system

3. Plan of implementation.

Although the system is still in its' theoretical phase, some of elements of it might be thought of as already done. To the preliminary testing phase there can be used already existing conversation software and searching engines. The decision in these matters will be taken after thorough review and comparing of the present state of knowledge in the field.

For the part of emotive analysis authors plan to work out and propose an original system based on the linguistic and communication sciences approach towards emotions and feelings recognition in the text (Ptaszyński, 2006). After combining together all parts of dynamic

database management system there will be a preliminary test phase, where efficiency and performance of the system with forgetting-recalling algorithm activated and deactivated will be compared. As the results will be found promising, an "in practice" test phase will start. Here there are two considerable ways of performing the test. In the first one, based on the idea of a Turing test, "naturalness of speech" will be tested in a standard way – a user will encounter the program and evaluate its' performance on the spot. In the second idea of the test, which authors thought of, a program will encounter another conversation software and the record of this dialogue will be evaluated by readers.

4. Conclusion.

Fading the unused context connections would improve the speed and performance of the program in processing large databases containing context-bound sentences. Since the database will be renewed according to the present needs, the program will operate on an actual and appropriate database all the time. In the final effect, a program working on such system, should have comparable ability to process and produce language with those in humans.

References:

- [1] Araki K. (2004) *Shizengengoshori kotohajime - kotoba wo oboe kaiwa no dekiru konpyūta* -, Tokyo: Morikita Publishers
- [2] Ebbinghaus, H. (1885) *Memory A Contribution to Experimental Psychology*, (transl.: Utsugi T. (1978) *Kioku nituite: jikken shinrigaku he no kōken*, Tokyo: Seishinsho Shobō)
- [3] Ishikawa M. (1990) "A Structural Connectionist Learning Algorithm with Forgetting" (*Journal of Japanese Society for Artificial Intelligence* 5(5), JSAI)
- [4] Kahn, I., et al. (2007) "Decreased demands on cognitive control reveal the neural processing benefits of forgetting", (*Nature Neuroscience*, 10, 908-914)
- [5] Markovitch, S. and Scott, P. D. (1988) "The Role of Forgetting in Learning", (*Proceedings of The Fifth International Conference on Machine Learning*, Ann Arbor, MI: Morgan Kaufmann)
- [6] Nurray Luk (2002) "The Role of Emotions in Language Teaching", (in: *The Journal of the Imagination in Language and Teaching*, vol. VII, 2002.03)
- [7] Ptaszynski, M. (2006) *Moeru gengo - Intānetto keijiban no ue no nihongo kaiwa ni okeru kanjōhyōgen no kōzō to kigōrontekikinō no bunseki - "2channeru" denshikeijiban o rei toshite* -, Poznan: UAM
- [8] Wolfe, P. (2006) "The role of meaning and emotion in learning" (*New Directions for Adult and Continuing Education*, vol. 2006, Issue 110, pp. 35 - 41)
- [9] Zitman, F. G. (2001) *Emotion and memory in mood- anxiety disorders*, Schreuder, BJN