How We Can Explain Simple Empirical Memory Rules

Francisco Zapata[1], Olga Kosheleva[2], and Vladik Kreinovich[3]
[1]Department of Industrial, Manufacturing, and Systems Engineering
[2]Department of Teacher Education
[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
fazg74@gmail.com, olgak@utep.edu, vladik@utep.edu

Abstract

Researchers have found out that normally, we remember about 30% of the information; however, if immediately after reading, we get a test, the rate increases to 45%. In this paper, we show that Zipf law can explain this empirical dependence.

1 Formulation of the Problem

Empirical facts about memory. Researchers have consistently observed that:

- Very fast, we forget about 70% of what we have heard or read and remember only 30%; see, e.g., [1, 3]. This is the percentage of correct answers that we get if we test the students a few days after they read the material.
- Interestingly, if, immediately after reading, the students take a test on what they just read, they retain 50% more information, i.e., they remember 45% of the original material; see, e.g., [1, 5].

How can we explain these empirical facts? To best of our knowledge, there are no quantitative explanations for the above empirical facts.

What we do in this paper. In this paper, we provide such an explanation.

2 Our Explanation

Main idea: Zipf's law. The main idea behind our explanation is to use Zipf's law; see, e.g., [2, 4]. This law was first observed in linguistics: if we sort all the words from a language in the decreasing order of their frequency, then the frequency with which the k-th word appears in the texts is approximately equal to c/k, for some constant c.

The same dependence was observed in many other situations, e.g., when we sort people by wealth or sort facts by importance.

Let us apply Zipf's law to our situation. Suppose that we have read or heard N different pieces of information. According to Zipf's law, the relative importance of the k-th piece of information is approximately equal to

$$\frac{c}{k}.$$

In particular, the least important piece of information has importance c/N. This is already close to the noise level, so it is reasonable to assume that the standard deviation σ of the corresponding noise is

$$\sigma \approx \frac{c}{N}.$$

Which pieces of information does it make sense to remember? Only those about which we are absolutely sure that this is not noise, that this information is indeed true. Usually, in applications of statistics, we use the "three sigma" rule: we believe in a certain fact if its deviation from the mean exceeds three times the standard deviation; see, e.g., [6]. This rule corresponds to 99.9% confidence: if we follow this rule, we will get erroneous signal only in 0.1% of the cases.

This explains 30%. So, we remember only the pieces for which the importance is larger than or equal to

$$3\sigma \approx \frac{3c}{N},$$

i.e., for which

$$\frac{c}{k} \geq \frac{3c}{N}.$$

This inequality is equivalent to

$$k \leq \frac{N}{3}.$$

Thus, out of the original N pieces of information, we remember one third. This is very close to the empirical 30%.

The fact that we actually remember slightly less than 1/3 can be explained by the imperfection of memory mechanisms.

This also explains 45%. Indeed, testing means, in effect, that the students encounter the same information twice. It is known that when the signal is repeated t times, averaging decrease the noise by a factor of $\sqrt{t}$; see, e.g., [6]. In particular, for $t = 2$, the level of noise decreases from

$$\sigma \approx \frac{c}{N}$$

to

$$\sigma' = \frac{\sigma}{\sqrt{2}} \approx \frac{c}{\sqrt{2} \cdot N}.$$

Thus, the three sigma threshold determining which pieces of information to remember is now

$$3\sigma' = \frac{3c}{\sqrt{2} \cdot N}.$$

Thus, only pieces for which

$$\frac{c}{k} \geq 3\sigma' = \frac{3c}{\sqrt{2} \cdot N}$$

are recalled. This inequality is equivalent to

$$k \leq \frac{\sqrt{2}}{3} \cdot N \approx 0.47 \cdot N.$$

This is very close to the empirical 45%. Thus, this number is also explained.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] P. C. Brown, H. L. Roedinger III, and M. A. McDaniel, *Make it Stick*, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, and London, UK, 2014.

[2] D. Cervantes, O. Kosheleva, and V. Kreinovich, "Why Zipf's law: a symmetry-based explanation", *International Mathematical Forum*, 2018, Vol. 13, No. 6, pp. 255–258.

[3] H. Ebbinghaus, *A Contribution to Experimental Psychology*, Dover, New York, 1964.

[4] B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, California, 1983.

[5] H. L. Roediger and J. D. Karpicke, "Test-enhanced learning: taking memory tests improves long-term retention", *Psychological Science*, 2006, Vol. 17, pp. 249–255.

[6] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Boca Raton, Florida: Chapman and Hall/CRC, 2011.