Show Me Everything You Know About [Birds]

The Fan Effect in Knowledge Management Systems

October 20, 2020

R.J. Cordes

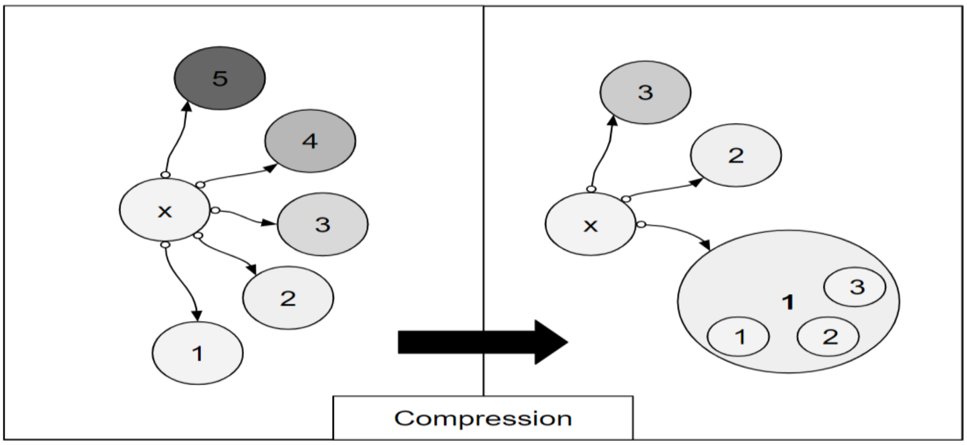
One of the primary problems that have become apparent to me while using Knowledge Management Systems (KMS) which allow for the user to traverse the paths of the network for exploration is something that, after a playful conversation with a colleague, I now refer to as the "Show me everything about [Birds]" Problem. Early into the development of the knowledge base or when a topic within it is still young, such a query can be quite useful—exposing opportunities, even in the 1st degree relationships, for meaningful retrieval and exploration, respectively: "Ah, this is the document which gives a definition of birds", "Odd, I did not know that the CIA used pigeons and ravens to spy on conversations". However, even the most obscure of topics within a knowledge base will be exposed to a titration of occasional backlinks and mentions—and will sometimes be dumped on by the occasional bursts of interactions by some user who, by happenstance, suddenly found use for the topic or felt that it was in need of some love and attention. These backlinks accumulate and soon, there is an overflow that disables, even in the 1st degree relationships, any meaningful discovery other than to know that "Wow, there is definitely a lot to know about [Birds]."

People who study the KMS in our brain (neuroscientists and pedagogists) refer to this problem, in their domain, as the *fan effect.* The fan effect describes the relationship between a cue concept and attached concepts (e.g. the cue-word math and attached algebra, calculus, and statistics): the higher the number of concepts related to a cue the higher the time requirements and rate of error of recalling the attached. This concept shows up in many disparate domains under different names such as value disambiguation in Cartography, wherein it is noted that users of map products cannot reliably differentiate between more than eight value types on a map (e.g. population density, traffic density, elevation etc.). The brain solves for this problem using 2 primary methods.

The First Method: Iterative & Dynamic Compression

The first method may be well represented by the quote by Alfred North Whitehead: “Knowledge shrinks as wisdom grows, for details are swallowed up by principles”. Albert Avey in his early 20th century philosophy survey-course textbook “Function and Forms of Thought” notes that at the start of the process of learning a new subject, **the examples are the knowledge attached to the subject**. However, as induction continues, and examples are accrued, we condense the memorable residue of each example into pattern and then the pattern becomes the knowledge of the subject. To return to our math example, we may begin, through induction to pick up that where addition, multiplication, and exponentiation are operations which induce growth and compress them into a category within operations (Operations which expand value) or conclude that addition and subtraction, multiplication and division, and exponentiation and roots each belong to their own respective categories, each a set of inverses of one another.

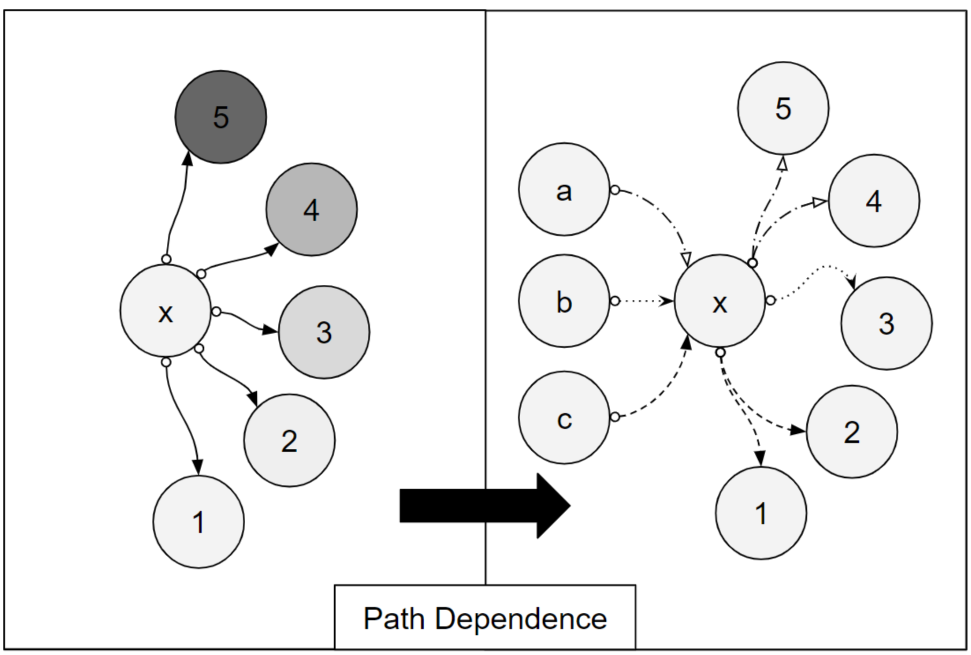
Famed pedagogists Jean Piaget and Lev Vygotsky, for all their disagreements, do not disagree on this notion of induction of examples into **schema**. Vygotsky suggests that, most important to this condensation or compression of the residue of knowledge, is the presence of a More Knowledgeable Other, who may point out the most helpful examples, who may light the path and correct misunderstandings to expedite the constant reduction of examples into pattern and the deduction of pattern into principle. Piaget on the other hand stresses play and active engagement, the opportunity to generate many examples of specific instances in order to expedite the process. Modern pedagogists might argue both are equally important.



Aaron S. Benjamin, in a highly comprehensive meta-analysis, "Memory is More than just Remembering: Strategic Control of Encoding, Accessing Memory, and Making Decisions", notes that the competence of computer memory in facilitating its host's actions is its ability to stay static whereas the competence of human memory in facilitating its host's actions is its inability to stay static. Every time a memory is recalled, it changes, its context changes, its relationship to other memories changes. Over time, this change becomes more minute, but it always changes. Even where memory and meaning appear to be static, this is often illusory, such is the case with the nature of phobia, where patients fail to make meaningful progress. A research project at the University of Amsterdam showed that it isn't that the memory is unchanged it's that the "long term potentiation" (the brain's perceived importance) which triggered the crystallization of the memory and meaning of the events which caused the trauma, is so high and so well reinforced by the number of times it was revisited that no amount of voluntary exposure in a day can be done such that it doesn't resolidify the trauma during sleep.

The Second Method: Path Dependence

The second method by which the brain solves for the fan effect is something which can be referred to as "path dependence" or, more simply, context. If we include not just a cue word, but a prefix and a cue word, or a prefix, a cue word, and a suffix we create pattern in identification patterns that identify patterns (see "fun"). This is just a variation on compression, in which where we were before helps identify where we go after we take our next step. We may now return to our math example, where we may say "rate of change" prior to saying our cue-word "math" and know that this means we are speaking about calculus, or "the moments described" prior to saying "within math" and know we are speaking of statistics.



Looking Forward

Current enterprise-grade KMS frameworks such as those implemented by Palantir or LegalZoom generally rely on querying the system prior to traversing the network in order to solve for the fan effect. Learning Management Systems literature and standard development working groups within the U.S. Army Training Doctrine and Command (TRADOC), the Department of Defense, and the IEEE all indicate a similar reliance. Non-enterprise note-taking frameworks that offer KMS features like Obsidian.md offer far less. Knowledge Management Systems which are characterized by their query-affordances, such as Google, generally rely on word-association in queries and generalizations about your background based on past searches rather than focusing on your current understanding of a subject and the current path you may be on in the process of developing that understanding. The Computer Science, Biology, and Artificial Intelligence communities have already translated and integrated research on topics like Communication of Sequential Processes (CSP), Pedagogy, Fuzzy Logics, Markov Blankets, and Neuroscience, the producers of next-generation Knowledge Management Systems technology should consider doing the same.