Operationalization of Meme Selection Criteria: Methodologies to Empirically Test Memetic Predictions

Klaas Chielens
Ecco,Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
klaas@chielens.net

Francis Heylighen
Ecco,Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
fheyligh@vub.ac.be

Abstract

This paper reviews a number of recent approaches to put memetics to the test of quantitative measurability. The focus is on the selection criteria for the spreading of memes put forward by Heylighen (1997), which include utility, novelty, simplicity, coherence, authority and proselytism. The general hypothesis is that memes scoring higher on these criteria will survive longer and be more prevalent than others. This can be tested by checking which story elements best survive a chain of person-to-person transmissions ("Chinese whispers" game), by simulating the cognitive and social processes that determine this differential survival and spread, and by correlating the score on the selection criteria with the actual frequency with which a meme is encountered. In a pilot study using an Internet survey, this method was applied specifically to virus hoaxes, which can be seen as paradigmatic examples of clearly delimited, self-reproducing messages.

1 Introduction

In 1976 Dawkins coined the term ‘meme’ to denote the cultural equivalent of the biological gene, i.e. an information pattern that is being copied from person to person. Examples of memes are jokes, ideas, traditions, rumors, fashions and chain letters. Each of these information systems spreads by means of communication from one to several carriers. Thus, a successful meme can be compared to a cultural virus that "infects" a growing group of hosts. Over the past decade, an increasing number of publications has been devoted to memetics (e.g. Blackmore, 2000 & Aunger, 2001), proposing explanations for phenomena from viral marketing to consciousness and religion.

However, the memetic approach has been criticized by many authors (Aunger 2001). Two major shortcomings can be pointed out: 1) it is hard to define what exactly a meme is; 2) the theoretical statements of memetics are as yet too vague to be empirically verifiable or falsifiable (Edmonds, 2002). The present paper proposes a broad methodology to address these problems. We will argue that: a) a memetic perspective can suggest concrete and non-trivial predictions; b) given a suitable memetic unit of investigation, these predictions can be tested empirically. This should establish a firm operational footing for memetics, allowing a comparison of the strengths and weaknesses of different models, and thus transforming memetics from a collection of suggestive hypotheses into a true scientific discipline.

2 Meme Selection Criteria

The core idea of memetics is that the popularity or success of a meme is determined by natural selection. At any moment, several memes are in competition for the attention of potential hosts and only those memes will spread that are well-adapted to the socio-cultural environment formed by these hosts and the network of their interactions; the others will become extinct. This leads to the generic prediction that “fitter” (i.e. better adapted) memes will become more widespread than less fit ones. To operationalize this as yet very abstract (and to some degree tautological) idea, concrete selection criteria need to be formulated that specify the degree to which a meme is adapted to its environment.

Several authors have tried to formulate principles that govern the spread of information. For example, Dawkins (1976), generalizing from the characteristics of biological evolution, listed the following three characteristics for any successful replicator, and thus for a meme: copying-fidelity, fecundity (number of copies made per time unit), and longevity (duration that any copy will survive). Working from a viral marketing perspective, Godin (2002) introduced the concept of the velocity with which ‘Idea Viruses’ spread from person to person. The social psychologists Schaller, Conway & Tanchuk (2002) focused on the communicability of a cultural trait. However, these characterizations of memetic fitness remain very broad and vague: what is it that makes a meme more communicable, fecund, or faster in spreading? They therefore offer little guidance in making non-trivial predictions.

Other authors have started listing more concrete and detailed criteria that together determine the fitness of a meme. For example, Castelfranchi's criteria (2001) focus on the social and cultural mechanisms of cultural transmission. A different list of criteria (Heylighen, 1997, 1998) focuses on the ways memes adapt to their hosts. In this work, four general criteria families are distinguished: objective, subjective, inter-subjective and meme-centered, depending on whether the selection depends on outside, objective reality, the individual subject or host of the meme, the process of transmission between subjects, or the internal properties of the meme itself. Heylighen (1998) proposes a four-stage model for memetic replication: 1) assimilation of a meme by a host; 2) retention within the host's memory; 3) expression by the host through behavior, language or some other medium; 4) transmission of the expression to one or more other hosts. At each stage there is selection, in the sense that some memes will be successfully assimilated, retained, expressed or transmitted, while others will not. A fit meme must pass all stages. The different selection criteria are typically active at different stages of this replication process.

The following is a selection of the most important criteria of this model[Heylighen, 1997, 1998], that can be easily operationalized:

- utility (the meme contains useful or valuable information)
- novelty (the meme is sufficiently different from already known memes)
- coherence (the meme is consistent with the knowledge that the hosts already have)
- simplicity (since complex memes difficult to process, less important details tend to be left out)
- formality (the less context or background communicating hosts share, the more important it is to express the meme explicitly)
- expressivity (the meme is easily expressible in the available languages or media)
- authority (the source is recognized as being trustworthy)
- conformity (the majority of hosts agree on the meme)
- proselytism (the meme explicitly incites its hosts to spread it further)

The first four of these are subjective and therefore depend on the host: what is useful or novel for one person may not be so for another one. The next four are intersubjective: they depend on the relations and forms of communication between hosts, and thus on the structure of the socio-cultural system. The last one, proselytism, is an example of a meme-centered criterion, that depends only on the meme itself. Simple examples of such self-promoting memes are viral sentences that contain a copy instruction, such as ‘Copy me’ or ‘say me’ (Hofstadter, 1996).

The general prediction that can be derived from this model is that, all other things being equal, if one meme scores higher on one of these criteria than another meme, it will also be fitter, and therefore spread more far and wide. For example, of two otherwise equivalent injunctions the one that is backed up by an authority (such as the pope), or by the majority of the population is likely to make more converts than the one that is not; the one that is novel will attract more attention and therefore spread faster; the one that fits in with people's existing ideas is more likely to be understood and believed and therefore to be memorized and expressed, etc. Moreover, the more criteria a meme fulfils the greater its overall fitness. Thus, the criteria, if valid, would provide us with a set of guidelines for how to recognize and design successful memes.

3 Methodologies for testing the selection criteria

3.1 Creating a memetic transmission chain

Different paradigms exist to study the spreading of memes. Perhaps the most direct, interactive one is the old game of "telephone" or "Chinese whispers", in which one person tells a story to another one, who then tells what (s)he remembers of it to the next person in line, who passes it on to the next one, and so on. At the end of the transmission chain, the final version is compared to the original story. To the amusement of the participants, the differences generally make the end story almost unrecognizable from the begin story.

From a memetic perspective, the different elements of such a story can be seen as individual memes. Some of these memes will be fitter, in the sense that they survive the many omissions and variations during the consecutive transmission better than others. Thus, the results of such a game may show what distinguishes good memes from poor ones.

An elegant example of this approach can be found in the psychological experiments of Lyons & Kashima (2001, 2003). In their game, the first participant read a made-up story about a non-existent tribe, the Jamayans. This participant 1 would retell the story to participant 2, who would retell it to 3, and 3 to 4, who told the final version to the experimenters. Before the experiment started, all participants had received background information about what kind of people the Jamayans were supposed to be, and what opinion the other participants had about that. The story consisted of consecutive elements (e.g. "a Jamayan boy encounters a bear", "he climbs in a tree", "he throws a branch at the bear", etc.). Some of these elements fit with the background knowledge (e.g. climbing in a tree is consistent with the Jamayans being fearful), others did not (e.g. throwing a rock is inconsistent with Jamayans being peaceful).

After several such experiments under varying conditions, a statistical analysis of the story elements that remained at the end of the game found a number of systematic effects that appear to confirm four of the above criteria: 1) coherence: elements inconsistent with the background information were more likely to be left out; 2) novelty: elements that the participants assumed were already known by the others were more likely to be left out; 3) simplicity: details or embellishments that did not affect the story line tended to be left out; 4) conformity: when the participants were told that the majority of them believed that the Jamayans were, e.g., peaceful, they were more likely to leave out elements inconsistent with this fact than if they thought that this was only a minority opinion.

3.2 Simulating meme evolution

A second paradigm for quantitative memetic investigation is simulation. There have been many agent-based simulations of how cultural replicators can spread through a population (e.g. Best, 1997), of which the first one to explicitly speak about memes may well be Gabora (1995). However, the agents and the memes used in these simulations are generally too simple to be used as models for the higher cognitive, emotional and social dynamics that govern meme transmission among humans. One of the only selection criterion to emerge (i.e. without being imposed by the programmer) from such simulations is conformity: the more agents already host a meme, the higher the probability that the other agents will be infected as well (cf. Boyd & Richerson, 1985).

Van Overwalle, Heylighen & Heath have started to investigate more realistic models in which agents do not just copy a message (with or without errors), but actively "reinterpret" messages, based on their own subjective experience with other agents and messages. To achieve this, agents are represented by simple neural networks that learn from experience. A message then corresponds to a pattern of activation over the nodes in such a network, and communication to the spread of that activation from agent to agent via variable inter-agent connections. The strength of the connection between two agents represents the degree of trust of the one in the information received from the other. This trust is learned on the basis of the degree to which information from that agent is confirmed by own knowledge and other sources.

This approach may allow the selection criteria to be derived from the dynamics of such a distributed connectionist network, rather than have them posited to some degree ad hoc. A preliminary simulation (Van Overwalle, Heylighen & Heath, 2004) indeed suggests that this can be achieved. For example, the reinforcement of inter-agent links through the increase of trust builds authority for the sending agents, and tells them which information the receiving agents are likely to already know and agree with, making it less important for them to transmit detailed, explicit reports (novelty and formality). Moreover, spread of activation along existing connections will automatically attenuate inconsistent (coherence) or complex (simplicity) signals, while amplifying signals that are confirmed by many different sources (conformity) or that activate in-built rewards or punishments (utility). As a first test, this simulation (Van Overwalle et al., 2004) has been able to replicate the most important quantitative results from the aforementioned study of Lyons & Kashima (2001) concerning the probability with which inconsistent or novel story elements are replicated in their "Chinese whispers" game.

3.3 Analyzing existing meme frequencies

A different paradigm for memetic investigations is the collection of existing memes (e.g. urban legends), together with an estimate of their success (e.g. the actual frequency with which a given legend is encountered on the web, or the likeliness that a person is to pass on the story to someone else). The study can then look for correlations between actual or apparent success rates and different criteria to test in how far high scores on the criteria predict memetic fitness.

Heath, Bell & Sternberg (2001) used this method to investigate a number of properties that fall under the general heading of "utility". Utility is a very broad category that includes any estimate of the importance or value of the information contained in a meme. Some of these estimates will be made rationally, e.g. by considering the plausibility of a meme; others will be made more intuitively or emotionally, e.g. by reacting with pleasure to an implied opportunity or fear to an implied danger. From the emotional components of this value judgment, Heath et al. focused on disgust because this is a relatively simple emotion whose strength is easy to measure. When comparing different urban legends that contained an element of disgust (e.g. the story of a man who discovers a dead rat in the cola bottle he has just been drinking from), they found that the more disgusting variations typically were more likely to be spread than the less disgusting ones. The same applied to plausibility, thus confirming two components of a broader utility criterion.

4 Virus hoaxes as paradigmatic memes

A shortcoming of the previous studies is that they work with rather vague and variable memetic units: "story elements", "traits" or "patterns of activation". As such they do not satisfy Dawkins' requirement of copying-fidelity or the general criticism that memes

lack a clear definition and are difficult to analyze. The last operational approach tackles this problem by looking at a very clear-cut example of a cultural replicator: a virus hoax (Sophos).

Virus hoaxes are email messages warning the recipients for a non-existent computer virus, and urging them to forward this warning to as many other people as possible. As such, a virus hoax is an illustration of a self-replicating message, that parasitizes the attention and computational resources of its recipients in order to maximally multiply itself. The ever wider expansion of electronic communication points us at the possible dangers of these virus hoaxes, which are threefold:

1) Virus hoaxes often propose methods of “protection” that are actually harmful (such as erasing essential program files).

2) They can create panic among naïve computer users by making them falsely believe that their computer is showing symptoms of a virus.

3) They produce economic damage by making their readers focus on the hoax instead of other activities, which results in a loss of time, energy, bandwidth and other resources.

Thus the study of how virus hoaxes spread is not only scientifically interesting, but it has direct social and economic applications. Moreover, these parasitic email messages are clearly delimited, normally undergo replication without variation, and, being pieces of text, are easy to analyze.

To test this memetics hypothesis, the statistical correlation between the score of a hoax on one of the criteria and an estimate of the degree of spreading of this hoax can be determined. It is important to make sure that enough different hoaxes are analyzed in order to obtain statistical significance. To be able to measure the degree of spreading (and thus the success) of a hoax, it is necessary to determine the exact content of the hoax text. Hoaxes are available in a number of specialized databases maintained by different organizations, such as Symantec or McAfee, on the internet. By comparing the different sources it is not only possible to find the most prevalent form but also to compare the strength of different mutations of the hoax. This could be used to recreate the evolutionary path that the hoax has followed, making a taxonomy of its different mutations (Bennett 2003).

Given the canonical form of a common variation, two or three distinguishing strings in the hoax’s text can be found that determine a unique "signature" of that text. Entering these signature strings in a search engine such as Google or AltaVista will not only find documents that contain this signature, but tell us how often these strings appear together on the internet, both on webpages or in newsgroups. This determines the number of copies of the hoax that still reside on the net.

The scoring of the selection criteria can happen in two ways: objective and subjective. Certain criteria can be measured objectively by applying linguistic techniques directly on the hoax text. Simplicity, for example, can be measured with the aid of Flesch Kincaid or Gunning-Fog readability tests, or the average sentence or word length. Other criteria can only be measured subjectively, by holding a survey in which participants are asked to indicate how strongly a hoax satisfies a certain criterion. To obtain a statistically significant score, the same hoax can be evaluated by a large number of people, after which the scores are averaged. As an extra controlling factor, the same criteria can also be scored by a group of experts.

Certain demographical data can also be gathered through the surveys, such as the participant's level of schooling or degree of knowledge of the language of the hoax. A possible hypothesis is that respondents with different levels of background knowledge accord a different order of importance to the criteria. For example, a hoax containing grammatical errors may lose credibility with a native speaker, while these errors may not be noticed by a foreign language speaker.

5 A pilot study of virus hoaxes

To test this general methodology, a small pilot study was performed in which 6 hoaxes were scored on 6 criteria by 195 participants (Chielens, 2003).

As this particular topic is closely linked to the Internet, an online survey was chosen to collect the data. One of the advantages of surveys over interviews is that there is less risk of answers being biased by social expectations. Using computer display makes it possible to represent the hoax as it would appear in a participant’s mailbox, including the capitalization and grammatical or spelling errors. Moreover, as there is no time-pressure in a survey, the participants can read and re-read the questions and the hoaxes as needed.

The criteria that were chosen to be included in the survey needed to be easily understood by the average participant. From the list above, the following criteria were selected: novelty, simplicity, utility, authority and proselytism. In the introduction to the survey each of these criteria was described so as to clarify its meaning. The short descriptions of the criteria were repeated with every question in the survey, as were the values (on a five-point scale) that could be entered for the criterion. For example, simplicity was tested with the following question: “How easy is it to understand this message? Is it hard to grasp or is it pretty clear and simple? (1: Very Hard / 5: Very Easy)”

The criterion of novelty was renamed to originality, in order to avoid a confusion with the idea that the hoax should be objectively ‘new’. Authority probed in how far the presumable source of the information (e.g. "This dangerous virus was first announced by IBM and Microsoft") appeared trust-

worthy. Utility was split up in a negative component, danger, and a positive one, benefit, since these hoaxes always warn of the great danger that may befall the ignorant recipient of a virus, but more rarely also mention the positive measures that can be taken to protect against the virus. Another reason for this split is that negative information normally produces a stronger mental reaction than positive information, a phenomenon called "negativity bias" (Ito et al., 1998). The criterion of proselytism (called "replication pressure" in the survey) is a particularly salient characteristic of virus hoaxes, which typically urge recipients to pass on the warning to all their friends and acquaintances.

After the participants had scored each of the criteria on a scale from one to five, the average scores were calculated, and correlated with the frequency with which the hoaxes appeared on the web or on newsgroups. One of the strongest correlations was found with the novelty criterion. This fits in with Godin's idea of the "filled vacuum" (2002): a meme can diffuse most easily in a niche where no similar memes are present yet. Specifically for hoaxes, a possible explanation for this correlation is that when a new type of hoax appears, it is not immediately recognized as a fake, whereas a hoax similar to older hoaxes will be found out more quickly. Another strong correlation was found for the criterion of benefit. Proposing a solution to a potential danger may help the hoax to spread as it gives the recipient a feeling of control, while it can indirectly confirm the false threat, as when the recipient carries out the hoax's instructions for tackling the problem and finds that, indeed, the specified file exists on his or her hard drive. Hoaxes that carried a warning with a possible ‘solution’ were indeed considered to have a higher benefit rating than hoaxes which only carried a warning.

The other correlations were too weak to be significant. This is probably due to the lack of data, as it is difficult to find reliable correlations when there are only 6 elements to compare.

However, another plausible explanation for the lack of correlation may be that the hoaxes used were by definition rather successful, since they otherwise wouldn't have appeared in hoax databases. This would mean that they were already close to the optimal score for the most critical criteria, so that a significant further increase in the score would be too much of a good thing, damaging the hoax's credibility. For example, the warning that a virus will erase your hard disk and damage your computer is already frightening enough; adding that it moreover may make you blind and put your house on fire would make the hoax lose its credibility. Similarly, it is likely that a too high proselytism score will not lead to a higher replication rate but to a ridicule of the hoax. A hoax that consists merely of ‘please pass me on’ phrases will not be passed on due to the lack of content, because people simply do not take it seriously (Hofstadter, 1996). A similar effect was found by Heath et al. (2002) in their investigation of disgusting urban legends: for the most successful legends, they found that it was impossible to create a more disgusting version, and the only plausible variations scored lower in disgust.

If most hoaxes in the sample would cluster around the peak value for a criterion, this would erase any clear correlation. To tackle this problem, further research would either need to use a more fine-grained statistical method than correlation coefficients to determine the relation between frequency and criterion scores, or artificially vary the score of a hoax to see whether it would lose in virulence, as Heath et al. (2002) did with some of their urban legends. An explanation for the fact that benefit and novelty still produced good correlations may be that these are less critical properties for virus hoaxes, unlike danger or proselytism, so that a typical hoax still has "room for improvement" on these dimensions.

6 Conclusion

Probably the most serious criticism of memetics is that it has not as yet produced any empirically verifiable predictions (Edmonds, 2002). Reviewing a number of partial and preliminary studies, using data about real memes or simulations of the social and psychological processes that govern their transmission, this paper has shown how memetic theories can be operationalized. This allows us to produce to a number of concrete, non-trivial and testable predictions, with immediate applications in domains such as viral marketing, the spread of rumors, or of parasitic email messages. It is our hope that this general approach will provide inspiration for other researchers to build more realistic and sophisticated memetic models, and to gather the detailed empirical evidence that will be necessary to convince other scientists of the value of the memetic perspective.

References

Aunger, R., (Ed.) (2001). Darwinizing Culture: The Status of Memetics As aScience Oxford/New York, Oxford University Press.

Best, M. (1997). Models for Interacting Populations of Memes. Journal of Memetics - Evolutionary Models of Information Transmission, 1. Retrieved October 31, 2004, from http://jomemit.cfpm.org/1997/vol1/best_ml.html

Blackmore, S. (2000). The Meme Machine Oxford/New York, Oxford University Press (Original work published 1999)

Boyd, R., & Richerson, P. J. (1985). Culture and the Evolutionary Proces Chicago, Chicago University Press.

Castelfranchi, C. (2001). Towards a Cognitive Memetics: Socio-Cognitive Mechanisms for Memes Selection and Spreading. Journal of Memetics - Evolutionary Models of Information Transmission. Retrieved October 31, 2004, from http://jom-emit.cfpm.org/2001/vol5/castelfranchi_c.html

Chielens, K.. The Viral Aspects of Language: A Quantitative Research of Memetic Selection Criteria (unpublished thesis, Vrije Universiteit Brussel, 2003)

Dawkins, R. (1989). The Selfish Gene (2nd ed.) Oxford/New York, Oxford University Press (Original work published 1976)

Edmonds, B. (2002). Three Challenges for the Survival of Memetics. Journal of Memetics - Evolutionary Models of Information Transmission. Retrieved October 31, 2004, from http://jom-emit.cfpm.org/2002/vol6/edmonds_b_letter.html

Gabora, L. (1993). Meme and variations:: A computer model of cultural evolution. In . L. Nadel, & D. Stein (Ed.), Lectures in Complex Systems New York, Addison-Wesley.

Godin, S. (2000). Unleashing the Ideavirus London, Simon & Schuster.

Heath C., Bell C., & Sternberg E. (2001). Emotional Selection in Memes: The CAse of Urban Legends. Journal of Personality and Social Psychology, 81, 1028-1041.

Heylighen, F. (1997). Objective, subjective and intersubjective selectors of knowledge. Evolution and Cognition, 3, 63-67.

Heylighen, F. (1998). What makes a meme succesful?: Selection Criteria for Cultural Evolution.. In , Proc. 16th Int. Congress on Cybernetcis Namur, Association Internat. de Cybernétique.

Hofstadter, D. R. (1985). Metamagical Themas: Questing for the Essence of Mind and Pattern New York, Basic Books.

Ito T. A., Larsen J. T., & Cacioppo J. T. (1998). Negative information weighs more heavily on the Brain: The negativity bias in evaluative categorizations. Journal of Personality and Social Psychology, 887-900.

Lyons A., & Kashima Y. (2001). The reproduction of culture: communication processes tend to maintain cultural stereotypes. Social Cognition, 19, 372-394.

Lyons A., & Kashima Y. (2003). How are stereotypes maintained through communication?: The influence of Stereotype scharedness. Journal of Personality and Social Psychology, 989-1005.

Ming, L., Bin, M., & Bennet C. (2003). Chain Letters & Evolutionary Histories. Scientific American, 76-81.

Schaller M., Conway L. G., & Tanchuk T. L. (2002). Selective pressures on the once and future contents of ethnic stereotypes: effects of the 'communicability' of traits. Journal of Personality and Social Psychology, 861-877.

Van Overwalle, F., Heylighen, F. & Heath, M. (2004). From Communication between individuals to collective beliefs. Retrieved October 31, 2004, from http://pcp.vub.ac.be/Papers/SiennaSimulation.pdf