Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2005 Proceedings

Americas Conference on Information Systems (AMCIS)

2005

Visualization and Bayesian Inference

Vince Kellen
DePaul University, vkellen@depaul.edu

Visualization and Bayesian Inference

Vince Kellen
DePaul University
vkellen@depaul.edu

ABSTRACT

The apparent difficulty people have with making Bayesian inferences has been researched heavily over the past 25 years, with conflicting explanations regarding the causes of and the cures for this inadequacy. Some researchers have improved Bayesian reasoning by representing the problem visually, but usually as a tool to teach Bayesian reasoning skills. This research examines facilitating reasoning performance in naïve Bayesian subjects without attempting to teach Bayesian reasoning skills. This approach is more relevant for everyday decision support situations where subjects do not or need not possess knowledge of Bayes theorem (naïve subjects). Several different visual representations (VRs) will be examined to determine which visualization technique generates the best decision performance. For this specific problem, certain visualization representations (VRs) may reveal the problem structure better than others, improving decision making, regardless of the whether number is represented as a natural frequency or a probability. VRs should be stable with regard to different base rates and reference class sizes. Using dual processing theories of cognition, this research will explain other aspects of this judgment task, including how users create and choose their strategies in solving this task and why subjects may have low levels of confidence in their results yet exhibit high task performance. Hopefully this research will help paint a clearer picture of the best ways for decision support systems to represent information in Bayesian inference tasks to naïve subjects and how VRs can enhance naïve subject performance in a variety of judgment and decision making tasks.

Keywords

Information visualization, Bayesian reasoning, judgment, natural frequencies, decision making, uncertainty, base-rate neglect, dual processing theories

INTRODUCTION

Information visualization techniques are frequently used in decision support systems, under the notion, usually presumed, that visualization improves decision making. Research from other disciplines is painting a more complex and subtle relationship between how a problem is represented and how people make judgments and decisions. In this research, we focus on VRs for a specific and well-researched problem, the "mammography problem," which asks subjects to infer a probability. The problem text, (from Gigerenzer & Hoffrage, 1995) reads as follows:

> The probability of breast cancer is 1% for women at age forty who participate in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer?

The normative answer requires applying a form of Bayes' theorem to calculate the posterior probability that a woman with positive results has cancer. The symbols *H* and –*H* denote the two hypotheses (cancer and no cancer, respectively) and *D* denotes the data representing the positive test results. The formula for computing *p(H | D)* is:

$$p(H \mid D) = \frac{p(H)\,p(D \mid H)}{p(H)\,p(D \mid H) + p(-H)\,p(D \mid -H)} = \frac{(.01)(.80)}{(.01)(.80) + (.99)(.096)} = .078$$

Prior research indicates subject accuracy is typically poor for the text-only probability version of the problem (Gigerenzer & Edwards, 2003; Gigerenzer & Hoffrage, 1995; Girotto & Gonzalez, 2001; Kurzenhauser & Hoffrage, 2002; Sedlmeier, 2000; Sloman, Over, Slovak, & Stibel, 2003).

PRIOR RESEARCH GOALS

Most of the research on this particular problem has had one of two goals: to improve subjects Bayesian reasoning skills through training or to understand how subjects make Bayesian, or equivalent, judgments. A few studies treat the visual representation as an interactive tool, typically to teach Bayesian reasoning (Cole, 1989; Sedlmeier, 2000), while others use it as a non-interactive tool in conjunction with a textual version of the problem (Sloman et al., 2003). Cole (1989) is perhaps the only study that compares different visual representations, but his was focused on teaching Bayesian reasoning, not in improving performance in naïve subjects.

The list of textual and visual representations discussed or tested in the body of research on this problem includes: contingency tables (Cole, 1989), trees (Martignon & Wassner, 2002), frequency bars or "beam cut" (Gigerenzer & Hoffrage, 1995), Euler circles (Sloman et al., 2003), signal detection, probability curves (Cole, 1989; Meuser, Cowan & Meuser, 1999), frequency grids (Cole, 1989; Gigerenzer & Hoffrage, 1995) and Bayesian boxes (Burns, 2004). Figures 1 through 6 show some of the VRs this research will examine.

NATURAL FREQUENCIES AND BAYESIAN INFERENCE

Gigerenzer and Hoffrage (1995) have argued that representing this problem as *natural frequencies* rather than as probabilities improves subjects' performance. Natural frequencies (Figure 7) arise from natural sampling and are not normalized with respect to base rates (Gigerenzer & Hoffrage, 1999). Natural sampling is the process of encountering instances in a population sequentially – in other words, as human beings normally encounter observations in their lives. Citing Cosmides & Tooby (1996), Gigerenzer and Hoffrage (1995) propose that evolution has given us minds that can reason with natural frequencies better than probabilities.

Several authors have suggested, that for problems like this one, certain representations reveal the problem's hidden set structure making the solution easier to obtain (Evans, Handley, Perham, Over, & Thompson, 2000; Girotto & Gonzalez, 2001; Mellers & McGraw, 1999). This particular problem has a three-part nested set structure of interest – women who are regularly screened, women who test positive for the disease and women who have the disease. Once the set structure becomes clear, subjects can assign values to the appropriate subset and compute a correct answer.

A well-known natural frequency version of the problem, from Gigerenzer & Hoffrage (1995), is as follows:

> 10 out of every 1,000 woman at age forty who participate in routine screening have breast cancer. 8 out of every 10 women with breast cancer will also get a positive mammograhy. 95 out of every 990 women without breast cancer will also get a positive mammograhy. Here is a new representative sample of 100 women at age forty who got a positive mammography in a routine screening. How many of these women do you expect to actually have breast cancer?

When presented with natural frequencies, subjects can more easily compute the correct answer, shown below:

$$p(H \mid D) = \frac{d \,\&\, h}{d \,\&\, h + d \,\&\, -h} = \frac{8}{8+95} = 0.078$$

The terms *d & h* are the number of cases with the symptom and the disease and *d & -h* is the number of cases having the symptoms but lacking the disease. Gigerenzer & Hoffrage, (1995) identify three relevant reasons natural frequencies simplify this problem: 1) natural frequencies are computationally simpler (as we just saw); 2) natural frequencies reduce attentional demands since only two pieces of information need to be considered, *d&h* and *d&-h*; 3) base rates (e.g., 10 out of 1,000) can be ignored, since the natural frequency information preserves the base rate information within the frequencies.

The frequency versus probability debate

Over the past decade, a rather vigorous debate concerning the effect natural frequencies have on performance of Bayesian inference problems has ensued. The points of the debate relevant to visual representations include properly defining what is a natural frequency; single-case versus aggregate cases representation; identification of the nested sets; common versus rare events; the size of the reference class and the how much natural frequencies facilitate Bayesian inference.

Some of the debate has identified different interpretations of the term *natural frequencies*. Simply putting a numeric value in terms of frequencies versus probabilities (e.g., 3 out of 100 versus a 0.03 probability) does not make it a natural frequency nor does it give it the benefits of natural frequencies (Gigerenzer & Hoffrage, 1999). Evans et al. (2000) and Sloman et al. (2003) argue that it is by revealing nested set relationships that frequency formats yield better performance. Gigerenzer &

Edwards (2003) state that natural frequencies (aggregate cases) make the reference class clear where as single event probability representations (single cases) leave it open to interpretation.

Natural frequencies may not always yield better performance. Mellers and McGraw (1999) indicates that frequency formats facilitate Bayesian reasoning only with rare events. Brase (2002) points out that prior research has suggested that when making judgments about populations of people, subjects are sensitive to population sizes, with subjects giving more normative answers for small reference class sizes (less than 100). Judgment might be better when subjects deal with scales that align with typical human experience.

Both sides of the aisle seem to agree that once the reference class and the nested sets are clear, the problem is more easily solved. This research expects that visual representations do better than natural frequency text representations in making the problem structure clear (reference class and nested set relationships) and hypothesizes that when good VRs are used, whether the numbers are a natural frequency or probability form should not matter. Since base rates and reference class sizes are encoded visually as proportions, VRs should be stable regardless of the reference class size and rarity of the event.

DUAL PROCESSING AND JUDGMENT AND DECISION MAKING

According to dual-processing theory, the human mind is comprised of two distinct cognitive systems underlying thought. Evans et al. (2003) consider system 1 as comprising implicit, domain-specific neural networks reflecting the learning history of the individual, but also including innate or automatic modules. System 2, the explicit system, is uniquely human and linked to language and reflective consciousness (Evans et al, 2003). Stanovich & West (2000), and others, including Kahneman (2002), Sloman (2002) and Hogarth (2002), use dual-process theories of reasoning as possible interpretations for explaining base rate neglect and other errors in Bayesian inference. While details of the dual process theories explained by various researchers are not identical, some agreement exists. Table 1 describes the two systems (Stanovich & West, 2000).

System 1	**System 2**
Associative	Rule-based
Holistic	Analytic
Automatic	Controlled
Relatively undemanding of cognitive capacity	Demanding of cognitive capacity
Relatively fast	Relatively slow
Highly contextualized task construal	Decontextualized task construal
Conversational and socialized task construal	Asocial task construal

Table 1. Dual processing and judgment and decision making

In problem solving situations, the two systems cue different responses and most likely interact or cooperate with each other in generating a response (Stanovich & West, 2000; Sloman, 2002). In McElroy & Seta (2003, 2004) studies that examined framing effects and dual-processing theory, they confirmed some aspects of dual processing theory: a) when task relevancy is high, system 2 is better engaged leading to more normative answers; b) users with an extremely analytic style as measured by the Zenhausen (1978) preference test generate more normative answers; c) hemispheres of the brain corresponding to system 1 and 2 influence the normative answer and can be manipulated to interfere with normative answers generated by system 2.

This research proposes that the processing of VRs in Bayesian inference problems involves cooperation between these two systems. VRs may facilitate improved judgment performance by engaging the context-specific, pre-attentive system 1 with visual cues that help subjects identify the reference class and the nested set relationships. VRs can give stronger system 2 cues (through system 1 pre-attentive processing) than textual natural frequencies facilitating performance. Task relevancy may also have much to do in explaining the impact VRs have on decision making. VRs should have a greater impact in improving judgment accuracy for tasks with lower relevancy. Because VRs may engage system 1 in stronger ways than do text representations, VRs may inhibit generation of multiple strategies from which subjects select, reducing problem solving time. In keeping with the information processing approach to judgment and decision making and its relationship to dual processing theory (Payne & Bettman, 2004), by engaging system 1 in the judgment task, VRs may reduce cognitive effort by reducing or eliminating strategy selection. When VRs are used, whether natural frequencies or probabilities are used should not matter. The visual cues are used by System 1 to constrain generation of System 2 strategies, surpassing the benefit natural frequencies may provide.

PROPOSED RESEARCH AGENDA

While this particular problem and others like it have been heavily researched, more needs to be understood. This research will examine the effectiveness of several visual representations in enhancing performance on this problem for naïve Bayesian subjects. These VRs include:

- Euler circles
- frequency bar
- frequency grid
- treemaps

This research will compare text and visual representations of the mammography problem. Results from this study can be compared with prior studies and the impact VRs have on task performance will be compared with the textual manipulations of the problem. This research will use process tracing techniques in a laboratory setting to determine which strategies subjects select from and use to solve this Bayesian inference problem and whether frequency formats or VRs have the greater influence on judgment speed and accuracy.

Based on dual processing theories and prior research into Bayesian inference problems, effective VRs should

- Make the reference class and nested set relationships pre-attentively transparent thereby helping subjects establish their problem solving strategy with fewer strategy shifts detected
- Increase judgment speed and accuracy regardless of the numeric format (frequency versus probability)
- Be robust enough to work well regardless of the rarity of the event (different base rates) and the size of the reference class

VRs may facilitate other classic judgment and decision problems, including mitigating framing effects and conjunctive probability errors; overcoming violations of stochastic dominance and other non-Bayesian selection of gambles; and reducing overconfidence in planning. The author argues that dual processing theories provide a better explanation for how VRs contribute to Bayesian inference in naïve subjects than do natural frequency, nested sets, information processing, or mental model explanations. Information and decision support systems may need to expand their repertoire of visualization techniques to include ones that can reduce judgment and decision bias through proper engagement of System 1.

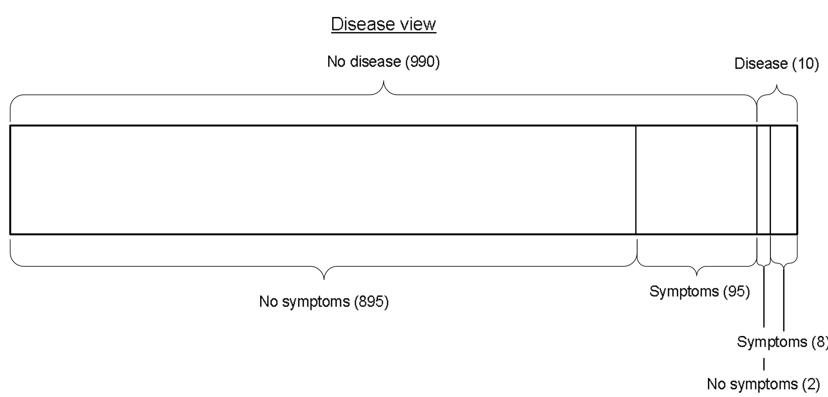

Figure 1. Frequency bar

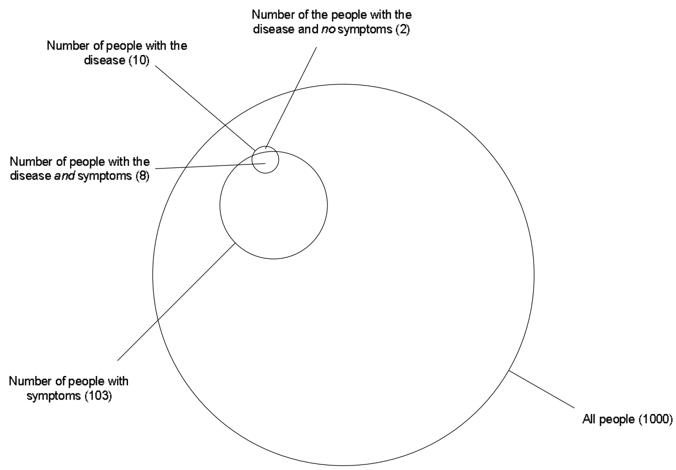

Figure 2. Euler circle

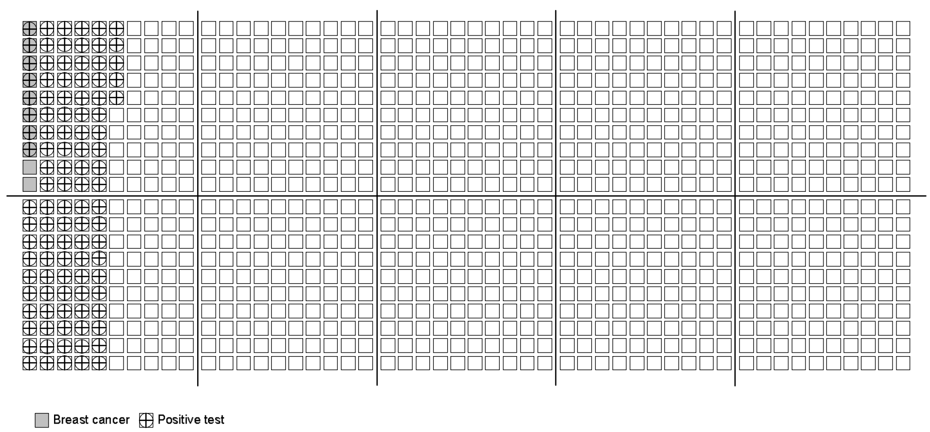

Figure 3. Frequency grid

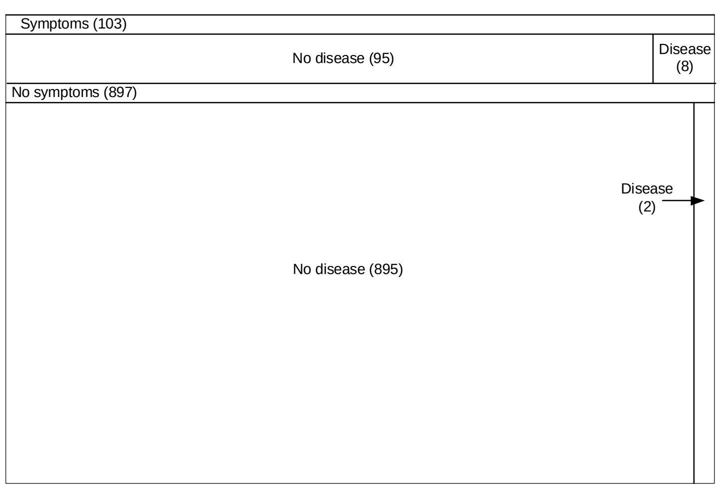

Figure 4. Treemap

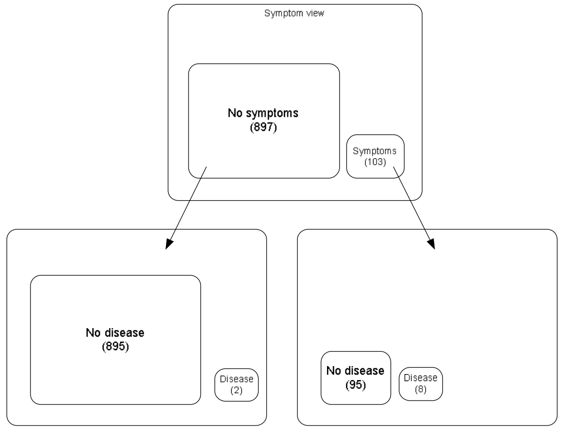

Figure 5. Visual Net hierarchy (treemap)

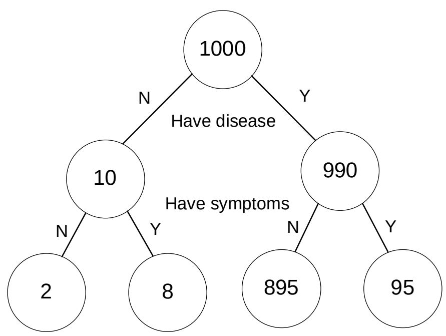

Figure 6. Frequency tree

Natural frequencies

	H	Not H
D	8	99
Not D	2	891
	10	990

Normalized frequencies

	H	Not H
D	800	100
Not D	200	900
	1000	1000

Figure 7. Contingency tables

REFERENCES

1. Brase, G. L. (2002). Which statistical formats facilitate what decisions? The perception and influence of different statistical information formats. *Journal of Behavioral Decision Making*. 15, 381-401.
2. Burns, K. (2004). Painting pictures to augment advice. *AVI '04, May 25-28.ACM.*344-349.
3. Cole, William G. Understanding Bayesian reasoning via graphical displays. *CHI '89 Proceedings, May 1989. ACM.* 381-386.
4. Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition. 58, 1-73.
5. Evans, J., Handley, S. J., Perham, N. Over, D.E., & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. *Cognition*. 77, 197-213.
6. Evans, J., Over, D. E., & Handley, S. J. (2003). A Theory of Hypothetical Thinking. In D. Hardman & L. Macchi (Eds), *Thinking: Psychological perspectives on reasoning, judgment and decision making.* John Wiley & Sons, Ltd.
7. Gigerenzer, G. & Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. *BMJ.* 327, 741-744.
8. Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review.* 102, 684-701.
9. Gigerenzer G. & Hoffrage, U. (1999). Overcoming difficulties in Bayesian Reasoning: A reply to Lewis & Keren and Mellers & McGraw. *Psychological Review.* 106, 425-430.
10. Girotto, V. & Gonzalez, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition.* 78, 247-276.
11. Hogarth, R. (2002). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. October. *UPF Economics and Business Working Paper No. 654*. http://ssrn.com/abstract=394920.
12. Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Prize Lecture, December 8, 2002. http:www.nobelprize.org*
13. Kurzenhauser, S. & Hoffrage, U. Teaching Bayesian reasoning: An evaluation of a classroom tutorial for medical students. *Medical Teacher.* 24, 516-521.
14. Lagnado, D.A. & Sloman, S. A. (2004). Inside and Outside Probability Judgment. In D. J. Koehler, & N. Harvey, N. (Eds), *Blackwell Handbook of Judgment and Decision Making.* Blackwell Publishing.
15. Martignon, L. & Wassner, C. Teaching decision making and statistical thinking with natural frequencies. *ICOTS6 2002.*
16. McElroy T. & Seta, J.J. (2003). Framing effects: An analytic-holistic perspective. *Journal of Experimental Social Psychology.* 39, 610-617.
17. McElroy, T. & Seta, J. J. (2004). On the other hand am I rational? Hemispheric activation and the framing effect. *Brain and Cognition.* 55, 572-580.
18. Mellers, B. A. & McGraw, A. P. (1999). How to improve Bayesian reasoning: Comment on Gigerenzer and Hoffrage (1995). *Psychological Review.* 106, 417-424.
19. Mueser, P. R., Cowan, N. & Mueser, K. T. (1999). A generalized signal detection model to predict rational variation in base rate use. *Cognition.* 69, 267-312.
20. Payne, J.W. & Bettman, J. R. (2004). The information-processing approach to decision-making. In D. J. Koehler, & N. Harvey, N. (Eds), *Blackwell Handbook of Judgment and Decision Making.* Blackwell Publishing.
21. Sedlmeier, P. (2000). How to improve statistical thinking: Choose the task wisely and learn by doing. *Instructional Science*. 28, 227-262.
22. Sloman, S. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin & D. Kahneman (Eds), *Heuristics and biases: The psychology of intuitive judgment.* (pp. 379-396). Cambridge, England: Cambridge University Press.
23. Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes.* 91, 296-301.
24. Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*. 23, 645-726.
25. Zenhausern R. (1978). Imagery, cerebral dominance, and style of thinking: A unified field model. *Bulletin of the Psychonomic Society*. 12, 381-384.