Finding useful questions in a natural environment

Jonathan D. Nelson (jnelson@cogsci.ucsd.edu)
Cognitive Science Dept., University of California at San Diego
9500 Gilman Dr., Dept. 0515, La Jolla, CA 92093-0515 USA

Identifying useful questions (tests, experiments, or queries) is important for a host of situations, including scientific reasoning, word learning, and vision. If a probabilistic belief model is used to describe an inquirer's knowledge, then each question's usefulness may be calculated using an explicit sampling norm (utility). Prominent sampling norms in psychological literature include Bayesian diagnosticity and log diagnosticity (Good, 1950), information gain (mutual information or Kullback-Liebler distance: Oaksford & Chater, 1994, 1996), probability gain (minimal error: Baron, 1985), and impact (absolute difference: Klayman & Ha, 1987). Strong claims about both the psychological reality and normative basis of particular norms have been made, in papers that calculate only a single sampling norm. Yet a literature review produced no treatment of when the sampling norms disagree with each other, and whether there are theoretical or empirical reasons to prefer a particular norm.

Skov & Sherman (1986) provided an early probabilistic study of information gathering. Participants were told (for instance) that on the planet Vuma 50% of creatures are gloms and 50% are fizos; that 28% of gloms and 32% of fizos wear a hula hoop; and that 10% of gloms and 50% of fizos smoke maple leaves. Given the goal of finding out whether a novel Vumian was a glom or fizo by asking either whether they wear a hula hoop or whether they smoke maple leaves, most participants asked about maple leaves. Skov & Sherman took this as evidence that people are sensitive to diagnosticity. Unfortunately, this result does not show what sampling norm is closest to people's intuitions, because diagnosticity, log diagnosticity, information gain, Kullback-Liebler distance, probability gain, and impact make the same prediction. Other studies have also made claims about particular sampling norms' psychological reality or normative preeminence, without considering other sampling norms.

One frequent task in daily life is to visually ascertain a person's gender. A simplified version of this task (which negates low-resolution information available from outside the center of gaze) is to learn a person's gender by viewing one feature at a time. This task is formally equivalent to the Vuma task. We collected statistics of the gender and features of interest of about 500 passerby, 51% of whom were male, in one natural environment (Table 1). Goals were to determine (1) whether different sampling norms make contradictory claims about what features are most useful, and (2) what sampling norms would best serve in this task.

Results showed that asking about hair length maximizes information gain, Kullback-Liebler distance, probability of correctly identifying the gender, and impact (absolute change in beliefs). Skirt and beard, however, have infinite diagnosticity and log diagnosticity. This is because in the rare event that a person is wearing a skirt or dress, or has a beard or other facial hair, their gender is known with certainty. Using diagnosticity or log diagnosticity to select questions would be inefficient in this environment.

Future work will examine what sampling norms' predictions best match human questions, and whether human questions are sensitive to symmetries and other class-conditional feature dependencies of natural objects.

Table 1: Features' distribution and usefulness.

	Skirt/ dress		Glasses (s=sun)			Beard		Earring		Short hair	
	n	y	n	s	y	y	n	y	n	y	n
% males	100	0	67	6	27	16	84	2	98	93	7
females	98	2	83	3	14	0	100	47	53	7	93
diag.	infinite		1.412			infinite		7.056		13.296	
$\log_{10}$ d.	infinite		0.093			infinite		0.532		1.123	
info.	0.010		0.025			0.084		0.235		0.634	
prob.	0.010		0.065			0.062		0.220		0.420	
impact	0.010		0.080			0.080		0.225		0.430	

Acknowledgment

J. D. N. was funded by NIMH grant 5T32MH020002-05 to the Salk Institute for Neural Computation, UCSD.

References

Baron, J. (1985). *Rationality and Intelligence*. Cambridge: Cambridge University Press.

Good, I. J. (1950). *Probability and the Weighing of Evidence*. New York: Charles Griffin

Klayman, J. & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information. *Psychological Review, 94*, 211-228

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608-631.

Oaksford, M., & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review, 103(2)*, 381-391.

Skov, R. B. & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Psychology, 22*, 93-121.