Information has non-negative expected utility

Richard Kennaway
School of Computing Sciences, University of East Anglia

23 March 2012

This note demonstrates that for a utility-maximising agent, the prior expected value of new information is non-negative.

First, an example to illustrate the principle. Suppose you are faced with two choices, A and B. One of them is right and one is wrong, and it's very important to make the right choice, because being right will confer some large positive utility U (you marry the princess), while the wrong choice will get you $-U$ (you are eaten by a tiger). However, you're not sure which is the right choice. From the information that you have, you compute a 51% chance that A is right, and 49% that B is right. So, you "shut up and multiply", and choose A for an expected utility of $0.02U$.

Suppose the choice does not have to be made immediately, and that you can do something to get better information about whether A or B is the right choice. Say you can make an observation giving a result a or b, such that $P(A|a) = P(B|b) = 99\%$. You make the observation and then choose A if the observation was a and B if the observation was b. In both cases, your probability that you are right is 99%, so your expected utility from choosing according to the observation is $0.98U$, an increase over not making the observation of $0.96U$.

Clearly, you should make the observation. Even though you cannot expect what it will tell you, you can expect to greatly benefit from whatever it tells you.

THEOREM 1 *Every act of observation has, before you make it, a non-negative expected utility.*

PROOF. Let the set of actions available to an agent be C. For each action c in C, the agent has a probability distribution over possible outcomes. Each outcome has a certain utility. For present purposes it is not necessary to distinguish between outcomes and their utility, so we shall consider the agent to have, for each action c, a probability distribution $P_c(u)$ over utilities u. The expectation value $\int_u uP_c(u)$ of that distribution is the prior expected utility of the choice c, and the agent's rational choice, given no other information, is to choose that c which maximises $\int_u uP_c(u)$. The resulting utility is $\max_c \int_u uP_c(u)$.

Now suppose the agent makes an observation, with result o. This gives the agent a new probability distribution for each choice c over outcomes: $P_c(u|o)$. It should choose the c that maximises $\int_u uP_c(u|o)$.

The agent also has a prior distribution of observations $P(o)$. Before making the observation, the expected distribution of utility returned by doing c after the observation is $\int_o P(o)P_c(u|o)$. This is equal to $P_c(u)$, as it should be, by the principle that your prior estimate of your posterior distribution of a variable must coincide with your prior distribution.

We therefore have the following expected utilities. If we choose the action without making the observation, the utility is

$$\max_c \int_u uP_c(u)$$

If we observe, then choose, we get

$$\int_o P(o) \max_c \int_u uP_c(u|o)$$

The second of these is always at least as large as the first. Proof:

$$\begin{aligned} & \max_c \int_u u P_c(u) \\ = & \max_c \int_u u \int_o P(o) P_c(u|o) \\ = & \max_c \int_o P(o) \int_u u P_c(u|o) \\ \leq & \int_o P(o) \max_c \int_u u P_c(u|o) \end{aligned}$$

□

Equality can easily hold. In the original example, suppose that your prior probabilities were 75% for A being right, and 25% for B. You make an additional and rather weak observation which, if it comes out one way raises your posterior probability for A to 80%, while if it comes out the other way, diminishes it to 60%. In either case you will still choose A and your expected utility (prior to actually making the observation) is unchanged. In general, the observation can only increase your expected utility if there is a non-zero probability of it telling you enough to change your mind. Below that threshold its expected value is zero. The following theorem makes this precise.

THEOREM 2 *As act of observation has, before you make it, a positive expected utility if and only if the probability of it changing your decision is positive.*

PROOF. Let c_{opt} be any value of c that maximises $\int_u u P_c(u)$. Let $f(o)$ be the increase in utility from observing o: $f(o) = \max_c \int_u u(P_c(u|o) - P_{c_{opt}}(u|o))$. It is positive exactly when observing o forces you to change your choice. Let f_s be the characteristic function of the support of f.

The proof of Theorem 1 shows that the expected increase in utility is $\int_o P(o) f(o)$. The probability of your observation changing your choice is $\int_o P(o) f_s(o)$. It is a general theorem of measure theory that either of these is positive if and only if the other is. □

The purpose of this note is to prove the above simple theorems of utility theory. It has therefore ignored all of the following possible complications of the issue which, while of interest in themselves, are not the present subject:

1. There may be a cost to obtaining extra information, and a cost to recalculating utility. These costs can be added into the utility calculation.

2. Higher-order considerations are ignored. Such considerations include dealing with enemies who may gain an advantage by knowing that you know something, cognitive biases contaminating your observations, or being driven mad by discovering things that man was not meant to know. A longer list of reasons that information can hurt you can be found in [1].

3. The question of whether people, or any other decision-making agents, are usefully modelled as maximisers of expected utility.

References

[1] Nick Bostrom. Information hazards: a typology of potential harms from knowledge. *Review of Contemporary Philosophy*, 10:44–79, 2011. `http://www.nickbostrom.com/information-hazards.pdf`.