Dynamic Consistency and the Value of Information

by Stephen Morris

Department of Economics, 
University of Pennsylvania, 
McNeil Building, 
3718 Locust Walk, 
Philadelphia PA 1910-6297

June 1992

Abstract: Expected utility maximizers make dynamically consistent choices if and only if they update beliefs using Bayes rule. For agents who cannot fully anticipate future contingencies, dynamic consistency may be too strong a requirement. Information is valuable if, for every decision problem, ex ante utility is higher if the information is observed. A necessary and sufficient condition for information to be valuable is that beliefs are noisy Bayes updated. Agents must revise their beliefs in the correct direction, but not necessarily as much as implied by Bayes updating. This work builds on Geanakoplos (1989) who showed that the valuable information condition is key to no speculation results, and gave conditions, when agents' information is represented (non-partition) possibility correspondences and agents update by Bayes rule, for information to be valuable. I show here that possibility correspondences are valuable, for some beliefs (not necessarily Bayes updated), if and only if they satisfy two properties, non-delusion and knowing that you know, while choices satisfy dynamic consistency, for some beliefs, if and only if possibility correspondences are partitions.

* This paper (together with companion paper, Morris (1992b)) incorporates material from chapters III and IV of my dissertation at Yale University (Morris (1991)). I thank my advisor John Geanakoplos for invaluable advice and support. This work arose out of a conversation with him about the relation between Blackwell's theorem and a result in Geanakoplos (1989). I am grateful for comments from participants at the Harvard Decision Sciences Colloquium and the financial support of an Anderson Prize Fellowship from the Cowles Foundation and an University of Pennsylvania Research Foundation grant.
1. Introduction

Expected utility theory provides one account of how rational people should make choices under uncertainty. Expected utility theory is static. Rational people "should" behave as if they are maximizing the expected value of some utility function under some probability distribution, but expected utility theory has nothing to say about how those probabilities "should" be revised over time (for example, following the arrival of new information). Many critics, empirical and theoretical, argue that expected utility theory is too restrictive. Yet without further dynamic restrictions on beliefs, expected utility theory is surely too weak for most economic applications.

In practice, rational people are assumed to use Bayes rule in updating beliefs, although the implicit restrictions on dynamic choice implied by Bayes rule are not always well understood. For expected utility maximizers, Bayes rule is equivalent to assuming that agents' preferences satisfy dynamic consistency. If you are going to observe some information, and you could commit to some information-contingent plan of action after observing the information, you would not want to revise that plan after you had actually observed the information. This result, which has appeared in various guises in philosophical and economic literatures, is proved in section 3.

Dynamic consistency is plausible enough as an axiom of rationality when the structure of the information is well understood, so that you are able to imagine all the possible signals you might observe in the future or might have observed in the past (equivalently, all possible states of your belief in the past and future). But can you imagine all possible signals? The axioms of (static) expected utility maximization seem modest in comparison. I may well have beliefs about some set of payoff-relevant events, and maximize expected utility given those beliefs, at every date and contingency, without knowing what my beliefs would be at another time or another contingency.

But if dynamic consistency is dropped, are no restrictions to be imposed on how beliefs are revised over time? A weaker restriction than dynamic consistency is a valuable information property, applied in a related context (discussed below) by Geanakoplos (1989). A decision problem is a set of actions available to an agent, and a function assigning utility to each action in each payoff-relevant state of the world. Given ex ante beliefs about payoff-relevant events, the uninformed value of a decision problem is the maximum expected utility attainable without information. But the agent will be observing a signal, and will have some ex post beliefs conditional on each signal (not necessarily derived by Bayes rule). A decision rule maps signals to actions, and is ex post optimal if conditional expected utility is

3. See references in footnote 11.
maximized for each signal. If we (unlike the agent) know the true distribution of signals, conditional on payoff-relevant events, we can calculate the informed value of the decision problem: the ex ante expected utility of following an ex post optimal decision rule, after observing information. Following Blackwell (1951), information is said to be valuable if the informed value of a decision problem is at least as great as the uninformed value, for every decision problem. Note that if conditional beliefs are generated by Bayes rule on the true distribution, information is always valuable. This confirms that dynamic consistency is a strictly stronger requirement than valuable information.

The valuable information property has both normative and positive significance. If we were thinking about designing an artificially intelligent system to update beliefs, and it was impossible to build in dynamic consistency because it was impossible to specify all possible information signals in advance, valuable information would be a natural property to try and build in, guaranteeing as it does that ex ante utility is increasing in information. Such monotonicity implies that convergence to true beliefs will eventually occur, given enough information. For the same reasons, thinking of economic agents as boundedly rational information processors, valuable information is a natural normative property to impose.

But the valuable information property is also of positive importance in economic models with asymmetric information. The non-existence of speculation is equivalent to agents’ beliefs satisfying the valuable information property [Geanakoplos (1989)]. This is discussed further in section 5.

In theorem 1 of section 3, it is shown that a condition called noisy Bayes updating is a necessary and sufficient condition for valuable information. Beliefs satisfy noisy Bayes updating if conditional beliefs are a weighted sum of the “true” Bayes updated conditional beliefs, and beliefs conditional on other signals. Agent beliefs are generated by adding “noise” to true Bayes beliefs. To see why, note that if the agent always ignored all information, and acted as if he had observed a unique signal with his ex ante beliefs, he could not be made worse off for any decision problem. As noted above, he also cannot be made worse off if he updates his beliefs with Bayes rule. Theorem 1 essentially states that linear combinations of these updating rules guarantee that information cannot make the agent worse off in any decision problem.

The above results about revising beliefs can be translated into results about revising knowledge in the following way. Suppose that in each payoff-relevant state of the world, there is a unique signal which is observed. Thus the distribution of signals conditional on the payoff-relevant states of the world is degenerate. If \( \omega \) is the true state, the possibility set \( P(\omega) \) is the set of states assigned strictly positive probability under that unique signal observed at state \( \omega \). Now we can ask what implications Bayes and noisy Bayes updating have for the possibility correspondence \( P \), mapping states of the world to states of the world believed possible. By looking at the possibility correspondence, we are throwing away information in the distribution.

In the case of the state structure, there exist communication and coordination problems of informational incompleteness and information asymmetry. Building on the framework of agents can be represented by a partition. So you can either impose the structure on the agents to assume, or you can let an agent knows something about the other agent's beliefs satisfying noisy Bayes updating, knowing that the other agent knows that you know that he knows, and so on. But which was it in section 3 which was it in section 5? In the valuative consideration the valuative considerations for both conditional beliefs, an agent updating beliefs and knowing that the other agent has the belief. Because, by definition, noisy Bayes updating is consistent with the restriction.

In summary, beliefs are

5. Hintikka
If we (unlike the agent) know the true distribution of signals, conditional on which we can calculate the informed value of the decision problem: the ex ante expected utility of the best optimal decision rule, after observing information. Following Blackwell (1958), information is valuable if the informed value of a decision problem is at least as great as its expected value of every decision problem. Note that if conditional beliefs are generated by an optimal decision rule, information is always valuable. This confirms that dynamic consistency is a stronger requirement than valuable information.

The information property has both normative and positive significance. If we were to build an artificial intelligence system to update beliefs, and it was impossible to build such a system that was impossible to specify all possible information signals in advance, it would be a natural property to try and build in, guaranteeing as it does that ex ante expected utility is maximized. Such monotonicity implies that convergence to true beliefs will be achieved. For the same reasons, thinking of economic agents as rational processors, valuable information is a natural normative property to have.

The information property is also of positive importance in economic models with non-existence of speculation is equivalent to agents' beliefs satisfying the single portfolio hypothesis (J. Geanakoplos (1989)). This is discussed further in section 5.

In section 3, it is shown that a condition called noisy Bayes updating is a necessary condition for having valuable information. Beliefs satisfy noisy Bayes updating if conditional on all information, and acted as if the agent had observed a unique signal with ex ante expected utility worse off for any decision problem. As noted above, he also cannot have beliefs that are consistent with Bayes rule. Theorem 1 essentially states that linear decision rules guarantee that information cannot make the agent worse off in any state.

In order to revise beliefs can be translated into results about revising knowledge. Suppose that in each payoff-relevant state of the world, there is a unique signal that indicates the distribution of signals conditional on the payoff-relevant states of the world. In the state, the possibility set P(ω) is the set of states assigned strictly positive probability to a signal observed at state ω. Now we can ask what implications Bayes and the possibility correspondence p, mapping states of the world to states of information: we are now concerned only with the support of beliefs (in a finite state space), not with the distribution.

In theorem 2 of section 4, it is shown that if beliefs satisfy Bayes updating, the possibility correspondence P must be a partition. Conversely, if a possibility correspondence is a partition, then there exist conditional beliefs consistent with the possibility correspondence (i.e. with supports equal to the possibility sets) such that beliefs satisfy Bayes updating. A partition is the standard representation of information in economics, statistics and other fields. In some recent work in economics, however, information has been represented by possibility correspondences which are not necessarily partitions.

Building on epistemic logic, state spaces and possibility correspondences can be constructed where agents can not know something, but do not know that they do not know it (this cannot happen with partition). Such representations are natural when we get away from the idea that the agents "understand" the structure of the state space. There are two other properties of partitions, however, which it is natural to assume, even if the structure of the state space is not informally or formally understood. If an agent knows something, it is true (non-delusion) and he knows that he knows it (knowing that you know). If beliefs satisfy noisy Bayes updating, the possibility correspondence P must satisfy non-delusion and knowing that you know. Conversely, if a possibility correspondence satisfies non-delusion and knowing that you know, then there exist beliefs consistent with the possibility correspondence such that beliefs satisfy noisy Bayes updating. This result is an extension of a remarkable result in Geanakoplos (1989), which was in fact the starting point for the research in this paper. Geanakoplos (1989) considered when the valuable information property was satisfied by possibility correspondences. But rather than stating the result for any beliefs consistent with the possibility correspondences (as I do), he assumed that conditional beliefs were derived by Bayes updating on possibility sets. Because of this restriction on beliefs, an additional restriction (nestedness) on the possibility correspondences - as well as non-delusion and knowing that you know - is required to ensure that the possibility correspondence is valuable. Because, by the argument of section 3, I think of Bayes updating as a consequence of dynamic consistency - itself stronger than valuable information - I do not find it natural to impose this restriction.

In section 5, it is shown that no speculation occurs in a class of economic models if and only if beliefs are noisy Bayes updated. Section 6 concludes, with a discussion of two companion papers [Morris(1992b,c)] extending this work.

2. Example

A new test for rabies has been developed. The test has an $e_2 \in (0,1)$ false positive rate: proportion $e_2$ of patients who do not have rabies test positive. The test has an $e_3 \in (0,1)$ false negative rate: proportion $e_3$ of patients who really have rabies test negative. Before the test, a patient believed there was a probability $\pi \in (0,1)$ that he had rabies. After observing a positive test result, he believes there is a probability $\gamma_2 \in (0,1)$ that he has rabies, while after observing a negative test result, he believes there is a probability $\gamma_3 \in (0,1)$.

For a "rational" patient who understands the model, and knows $e_2$ and $e_3$, the ex ante probability $\pi$ and the conditional probabilities $\gamma_2$ and $\gamma_3$ will be linked via Bayes' rule:

$$\gamma_2^* = \frac{\pi (1-e_3)}{\pi (1-e_3) + (1-\pi) e_2} \quad \gamma_3^* = \frac{\pi e_3}{\pi e_3 + (1-\pi)(1-e_2)}$$

I assume, without loss of generality, that $e_2 + e_3 \leq 1$, so that the result labelled positive cannot make it less likely that the patient has rabies i.e. $\gamma_2^* \geq \gamma_3^*$. But perhaps the patient does not know the true values of $e_2$ and $e_3$, or knows them but fails to update correctly. The patient is nonetheless a static expected utility maximizer: he satisfies the Savage (1954) axioms at each point in time. We can ask (as external observers, knowing the true values of $e_2$ and $e_3$), whether the test is valuable to the agent, given his beliefs, $(\pi, \gamma_2, \gamma_3)$.

The patient must choose one out of a possible set of treatments (including no treatment). His utility depends on the treatment and whether, in fact, he has rabies. For a particular decision problem (a set of possible treatments), the value of the test is the difference between his ex ante utility if he takes the test, and chooses a treatment given his conditional beliefs $\gamma_2$ and $\gamma_3$, and his ex ante utility if he does not take the test, and chooses a treatment given his prior belief $\pi$. The test is said to be valuable if, for every decision problem, the value of the test is non-negative.

Theorem in section 3 states that information is valuable in this sense if and only if conditional beliefs are noisy Bayes updates. This translates, in the example, to the requirement that $\gamma_2^* \geq \gamma_2 \geq \pi \geq \gamma_3 \geq \gamma_3^*$, and $\gamma_2^* = \gamma_2 = \pi = \gamma_3 = \gamma_3^*$ if $\gamma_2 = \gamma_3$. The test is valuable as long as the patient updates his beliefs in the right direction, but no more than implied by Bayes updating.

To see why, suppose the patient must choose between a single treatment or no treatment. The treatment is painful and has a utility cost $c$ (whether or not the patient has rabies). But, if the patient has rabies and is not treated, a much higher utility cost $k$ of late treatment is incurred. An expected utility maximizer will accept the treatment if the probability of rabies is greater than $c/k$.

6. Slovic
2. Example

A test has been developed. The test has an \( \epsilon_p \in (0,1) \) false positive rate: proportion have rabies test positive. The test has an \( \epsilon_n \in (0,1) \) false negative rate: do really have rabies test negative. Before the test, a patient believed there that he had rabies. After observing a positive test result, he believes there that he has rabies, while after observing a negative test result, he believes there no rabies. Patient who understands the model, and knows \( \epsilon_p \) and \( \epsilon_n \), the ex ante probabilityabilities \( \gamma_p \) and \( \gamma_n \) will be linked via Bayes’ rule:

\[
\frac{\pi(1-\epsilon_n)}{\pi(1-\epsilon_n) + (1-\pi)\epsilon_p} \gamma_p^* = \frac{\pi \epsilon_n}{\pi \epsilon_n + (1-\pi)(1-\epsilon_p)} \gamma_n
\]

Loss of generality, that \( \epsilon_p + \epsilon_n \leq 1 \), so that the result labelled positive cannot patient has rabies i.e. \( \gamma_p^* \geq \gamma_n^* \). But perhaps the patient does not know the knows them but fails to update correctly. The patient is nonetheless a static who satisfies the Savage (1954) axioms at each point in time. We can ask (as the true values of \( \epsilon_p \) and \( \epsilon_n \)), whether the test is valuable to the agent, given choose one out of a possible set of treatments (including no treatment). His treatment and whether, in fact, he has rabies. For a particular decision problem (1), the value of the test is the difference between his ex ante utility if he takes treatment given his conditional beliefs \( \gamma_p \) and \( \gamma_n \), and his ex ante utility if he does not takes a treatment given his prior belief \( \pi \). The test is said to be valuable if, for the value of the test is non-negative.

On 3 states that information is valuable in this sense if and only if conditional rates. This translates, in the example, to the requirement that \( \gamma_p^* \geq \gamma_p \geq \pi \) \( = \pi = \gamma_n = \gamma_n^* \) if \( \gamma_p = \gamma_n \). The test is valuable as long as the patient the right direction, but no more than implied by Bayes updating.

Use the patient must choose between a single treatment or no treatment. The utility cost \( c \) (whether or not the patient has rabies). But, if the patient has much higher utility cost \( k \) of late treatment is incurred. An expected utility treatment if the probability of rabies is greater than \( c/k \).

Now suppose \( \gamma_p^* > \gamma_p > \pi > \gamma_n > \gamma_n^* \). For any value of \( c/k \), the patient cannot be made worse off in ex ante utility terms by observing the test. If a positive result is observed, the patient will make the same choice he would have made if with Bayes updated beliefs unless \( \gamma_p^* > c/k \geq \gamma_p \), when he may reject the treatment when he should have accepted it. But in this case, he would have rejected the treatment anyway if he had not observed the information. Analogously, if a negative result is observed, the patient will make the correct choice unless \( \gamma_n^* < c/k \leq \gamma_n \), when he may accept the treatment when he should have rejected it. But in this case, he would have accepted it if he had not observed the information.

Thus if beliefs satisfy noisy Bayes updating, the test must be valuable. Let us show the converse: if beliefs do not satisfy noisy Bayes updating, there exists a decision problem such that the patient is made worse off by observing the information. Suppose the patient updates beliefs in the right direction but too much: \( \gamma_p > \gamma_p^* > \pi \). Now suppose that \( \gamma_p > c/k > \gamma_p^* \). Then the patient will accept the treatment after observing a positive test result (because \( \gamma_p > c/k \)), even though he should have refused it (\( \gamma_p^* < c/k \)) and he would have refused it if he had not observed the test (\( \pi < c/k \)).

It is interesting to note that this was exactly the actual updating behavior observed in early experiments on choice under uncertainty by psychologists:

"The subject is presented with the following situation: Two bookbags are filled with poker chips. One bookbag has 70 red chips and 30 blue chips, while the other bag holds 30 red chips and 70 blue chips. The subject does not know which bag is which. The experimenter flips a coin to choose one of the bags and then begins to draw chips from the chosen bag. After drawing a chip he shows it to the subject.... The subject is asked to estimate the probability that the predominantly red bag is being sampled. At the start, before the first chip is drawn, the subject is required to give a probability estimate of .5, indicating that each bag is equally likely to have been chosen.**

This problem has exactly the structure of the example I presented. The ex ante probability of a predominantly red bag is \( \pi = .5 \). The probability of drawing a blue ball from the predominantly red bag is \( \epsilon = .3 \). The probability of drawing a red ball from the predominantly blue bag is \( \epsilon = .3 \). So the Bayes updated probability that the bag contains mostly red balls, after observing a red ball, is \( \gamma_p^* = 0.7 \), and the Bayes updated probability that the bag contains mostly red balls, after observing a blue ball, is \( \gamma_n^* = 0.3 \).

The subjects' estimates of the probabilities typically satisfied the noisy Bayes updating condition, but not Bayes updating: the probability that the bag contained predominantly red balls, after observing a red ball, \( r \), was assessed higher than 0.5 but less than 0.7. This behavior was known as conservative updating in the psychology literature.

I cite this work in order to claim that noisy Bayes updating is typical, but to give a sense of why it is a natural concept in thinking about revisions of beliefs. I will be focussing on its normative significance in this paper and its companions.

More recent experimental work suggests that people may sometimes update more than implied by Bayes updating, and thus violate noisy Bayes updating. Kahneman and Tversky\(^7\) argue that people use a "representativeness" heuristic which can work in the opposite direction. Proportion \( \pi \) of the guests at a hotel are lawyers, and proportion \((1-\pi)\) engineers. You are told that one particular guest is argumentative. What is the probability that he is a lawyer? Subjects assess a probability that he is lawyer based on how "representative" argumentativeness is felt to be as a description of lawyers. This posterior probability is insensitive - excessively insensitive under Bayes Rule - to the prior probability \( \pi \). This implies a violation of noisy Bayes updating for some \( \pi \). Thus the representativeness heuristic may cause information to be overvalued relative to prior beliefs. A consequence of the results of this paper is that this kind of "mistake" in updating can be more serious in some sense than noisy Bayes updating.

Section 3: The Value of Information and Dynamic Consistency

When is information valuable? More precisely, when is it the case that, for every decision problem, depending on some finite uncertainty space \( \Omega \), that you might face, it is better to have observed a signal correlated with the outcome in \( \Omega \), than not to have observed it. It is a corollary of Blackwell's theorem that, for agents who update their beliefs using Bayes rule, such information is always valuable. But if Bayes rule is a sufficient condition on probability updating for information to be always valuable, what is a necessary condition? A weaker version of Bayes updating - here described as noisy Bayes updating - is necessary as well as sufficient. Beliefs satisfy noisy Bayes updating if beliefs about the uncertainty

More verifiable statements of noisy Bayes updating can also be made.

Suppose that for some \( a \in A \) the following holds:

\[
\text{If belief } B(a|s) \text{ and action } a \in A \text{ results in state } s, \text{ then belief } B(a'|s) \text{ and action } a \in A \text{ results in state } s'.
\]

7. Kahneman, Slovic and Tversky (1982), part II.

8. Blackwell's Theorem states that an experiment is ex ante preferred to another (for every decision problem) if and only if the latter is a "noisy" version of the former. Since any experiment is a noisy version of no experiment, the result following is indeed a corollary of Blackwell's theorem. See Blackwell (1951) for the original statement, Crémer (1982) for an alternative exposition. A companion paper (Morris (1992b)) derives a generalization of Blackwell's theorem in the general case of agents not updating according to Bayes rule.
of the probabilities typically satisfied the noisy Bayes updating condition, probability that the bag contained predominantly red balls, after observing greater than 0.5 but less than 0.7. This behavior was known as conservative literature.

In order to claim that noisy Bayes updating is typical, but to give a sense of thinking about revisions of beliefs. I will be focussing on its normative its companions.

Conceptual work suggests that people may sometimes update more than implied violate noisy Bayes updating. Kahneman and Tversky argue that people heuristic which can work in the opposite direction. Proportion \( \pi \) of the agents proportion (1-\( \pi \)) engineers. You are told that one particular guest is probability that he is a lawyer? Subjects assess a probability that he is lawyer argumentativeness is felt to be as a description of lawyers. This posterior cessively insensitive under Bayes Rule - to the prior probability \( \pi \). This yes updating for some \( \pi \). Thus the representativeness heuristic may cause relative to prior beliefs. A consequence of the results of this paper is that thinking can be more serious in some sense that noisy Bayes updating.

### The Value of Information and Dynamic Consistency

Beliefs are valuable if, for every decision problem \((A, u)\), every optimal decision rule \(f\), and every action \(a \in A\),

\[
\sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega, s) u(f(s), \omega) \geq \sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega) u(a, \omega).
\]

9. An alternative characterization would be to say that beliefs are valuable if the inequality hold for every decision problem, for some optimal decision rule \(f\), for every action \(a \in A\). This is a non-trivial distinction, since without Bayes updating, different (ex post) optimal decision rules may give different ex ante utility. But Morris (1992b) shows that as long as no two signals have the same conditional belief (i.e., \(\gamma(\omega | s) = \gamma(\omega | s')\) for all \(\omega \in \Omega\) implies \(s = s'\)), there is no difference between the two definitions. Morris (1992b) also explores the relation between a number of different notions of when information is valuable.
If beliefs satisfy dynamic consistency, the agent would not want to revise a plan of action, drawn up before observing the signals but conditional on them, after observing the signals. This criterion is too strong if the agent does not understand the true relation between signals and states.

Beliefs satisfy Bayes updating if

$$\gamma(\omega|s) = \frac{\pi(\omega,s)}{\sum_{\omega' \in \Omega} \pi(\omega',s)} \text{ for all } \omega \in \Omega, s \in S.$$  

Beliefs satisfy noisy Bayes updating if there exist $\alpha: S \to (0, 1)$ and $\theta: S^2 \to \mathbb{R}_+$, such that

$$[a] \quad \gamma(\omega|s) = \alpha(s) \frac{\pi(\omega,s)}{\sum_{\omega' \in \Omega} \pi(\omega',s)} + [1 - \alpha(s)] \sum_{s' \neq s} \theta(s,s') \gamma(\omega|s'), \text{ for all } \omega \in \Omega, s \in S.$$  

where $[b] \sum_{s' \neq s} \theta(s,s') = 1$, and $[c] \sum_{s' \neq s} \frac{\pi(\omega,s)}{\pi(s)} \geq \sum_{s' \neq s} \left\{ \frac{[1 - \alpha(s') \theta(s'|s)] \sum_{s'' \neq s'} \pi(\omega,s''|s')} {\pi(s')} \right\}$.

Note that if beliefs satisfy Bayes updating, they certainly satisfy noisy Bayes updating [set $\alpha(s) = 1 \forall s \in S$]. Noisy Bayes beliefs are an average of Bayes beliefs and beliefs conditional on other signals. Condition [b] says that $\theta$ is a Markov process. Condition [c] places restrictions on that class of Markov processes. Note that if $\theta(s,s') = k$ for all $s \neq s'$, then condition [c] is always satisfied. The intuition behind noisy Bayes updating is that agents are conservative as discussed in the example of the previous section. Some alternative characterizations of noisy Bayes updating are given in lemma 2 below.

**THEOREM 1** Beliefs are valuable if and only if they satisfy noisy Bayes updating.

Beliefs satisfy dynamic consistency if and only if they satisfy Bayes updating.

The theorem will be proved via a couple of lemmas.

**LEMMA 1** Beliefs are valuable if and only if there exists $\lambda: S \to \mathbb{R}_+$ and $\mu: S^2 \to \mathbb{R}_+$ such that

$$\pi(\omega,s) = \lambda(s) \gamma(\omega|s) + \sum_{s' \neq s} [\mu(s,s') \gamma(\omega|s) - \mu(s',s) \gamma(\omega|s')], \text{ for all } \omega \in \Omega, s \in S.$$  

**LEMMA 2**

[1] \quad $\pi(\omega,s) = \lambda(s) \gamma(\omega|s) + \sum_{s' \neq s} [\mu(s,s') \gamma(\omega|s) - \mu(s',s) \gamma(\omega|s')].$

[2] \quad $\pi(\omega).$
Proof of lemma 1 An experiment is valuable if there does not exist decision problem \([A,u]\) and optimal decision rule \(f\) such that
\[
\sum_{\omega \in \Omega} \gamma(\omega|s)u(f(s),\omega) \geq \sum_{\omega \in \Omega} \gamma(\omega|s)u(a,\omega), \quad \forall a \in A
\]
\[
\sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega,s)u(f(s),\omega) < \sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega,s)u(a,\omega), \text{ for some } a \in A.
\]
But now by letting
\[
x(\omega,s) = u(f(s),\omega) - u(a^*,\omega), \text{ where } a^* \in \arg\max_{a \in A} \sum_{s \in S} \pi(\omega,s)u(a,\omega)
\]
we see that this is equivalent to the requirement that there does not exist \(x : \Omega \times S \rightarrow \mathbb{R}\) such that:
\[
\sum_{\omega \in \Omega} \gamma(\omega|s)x(\omega,s) \geq \sum_{\omega \in \Omega} \gamma(\omega|s)x(\omega,s'), \quad \forall s,s' \in S
\]
\[
\sum_{\omega \in \Omega} \gamma(\omega|s)x(\omega,s) > 0, \quad \forall s \in S.
\]
\[
\sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega,s)x(\omega,s) < 0.
\]
By Farkas' Lemma, there do not exist such \(x\) if and only if there exist \(\lambda : \Omega \rightarrow \mathbb{R}\), and \(\mu : S^2 \rightarrow \mathbb{R}\), such that
\[
\pi(\omega,s) = \lambda(s)\gamma(\omega|s) + \sum_{s' \in S} [\mu(s,s')\gamma(\omega|s') - \mu(s',s)\gamma(\omega|s')], \text{ for all } \omega \in \Omega, \quad s \in S.
\]

By dynamic consistency if and only if they satisfy noisy Bayes updating.

Lemma 2 establishes that the condition of lemma 1 is equivalent noisy Bayes updating, as well as using matrix notation to give some alternative characterizations of noisy Bayes updating. Enumerate \(\Omega = \{\omega_1, \ldots, \omega_n\}\) and \(S = \{s_1, \ldots, s_m\}\) and write \(\Gamma\) for the \(n \times m\) matrix with \((i,j)\)th co-ordinate \(\gamma(\omega_i|s_j)\) and \(\Pi\) for the \(n \times m\) matrix with \((i,j)\)th co-ordinate \(\pi(\omega_i,s_j)\).

**LEMMA 2** The following are equivalent:

1. Beliefs satisfy noisy Bayes updating
2. There exists \(\lambda : \Omega \rightarrow \mathbb{R}\), and \(\mu : S^2 \rightarrow \mathbb{R}\), such that
\[
\pi(\omega,s) = \lambda(s)\gamma(\omega|s) + \sum_{s' \in S} [\mu(s,s')\gamma(\omega|s') - \mu(s',s)\gamma(\omega|s')], \text{ for all } \omega \in \Omega, \quad s \in S.
\]
[3] There exist non-negative diagonal matrices $L$ and $D$, and Markov matrix $\Theta$ such that

$$\Pi = \Gamma [L + D[t-I(\cdot)]]$$

[4] There exists matrix $M$ such that (i) $\Pi = \Gamma M$ (ii) $MI \geq 0$ and (iii) $M$ has non-positive off diagonal elements

Note that a matrix of the form of $M$ in [3] can be shown to have a non-negative inverse, so that it is necessary (but not sufficient) for noisy Bayes updating that there exists a non-negative matrix $N$ such that $\Gamma = IIN$.

Proof of lemma 2

[1] is equivalent to [2]

Suppose condition [2] holds. Then $\pi(\omega, s) > 0$ for some $\omega$, for all $s = \lambda(s) + \Sigma_{\nu \cdot \mu}(s,s') > 0$ for all $s$, so [2] can be re-written as

$$\gamma(\omega|s) = \frac{\pi(\omega, s) + \sum_{s'}^\mu(\cdot|s')\gamma(\cdot|s')}{\lambda(s) + \sum_{s'}^\mu(\cdot|s')}, \text{ for all } \omega \in \Omega, s \in S.$$ 

Now substitute $(\alpha, \theta)$, satisfying the conditions below, for $(\lambda, \mu)$, in the expression above

$$\alpha(s) = \sum_{\omega \in \Omega} \pi(\omega, s) \lambda(s) + \sum_{s' \in S} \mu(s, s'), \quad \theta(s,s') = \frac{\mu(s,s')}{\sum_{s' \in S} \mu(s', s')}, \text{ if } \sum_{s' \in S} \mu(s', s) > 0$$

Now noisy Bayes updating condition (a) holds by substitution, (b) by definition, and (c) is equivalent to the restriction that $\lambda \geq 0$.

[2] is equivalent to [3]

If condition [2] holds, let $L$ be matrix with $(i,i)$th element $\lambda(s)$, $D$ be matrix with $(i,i)$th element $\Sigma_{\nu \cdot \mu}(s,s')$ and $\Theta$ be matrix with $(i,j)$th co-ordinate $= \lambda(s,s'')/\lambda(s,s)$. $L$ and $D$ are non-negative diagonal matrices, $\Theta$ is Markov, and condition of [3] holds. A similar converse argument holds.

10. Nikaido (1968) chapter II.
negative diagonal matrices L and D, and Markov matrix $\Theta$ such that
$[\Theta, D + \Theta]$

Fix $M$ such that
(i) $M = \Gamma M$ (ii) $M1 = 0$ and (iii) $M$ has non-positive off

the form of $M$ in [3] can be shown to have a non-negative inverse, so
sufficient for noisy Bayes updating that there exists a non-negative matrix

Then $\pi(\omega, s) > 0$ for some $\omega$, for all $s = \lambda(s) + \sum_{s' \in s} \mu(s', s) > 0$ for all

$$
\frac{\pi(\omega, s)}{\lambda(s) + \sum_{s' \in s} \mu(s', s)} > 0, \text{ for all } \omega \in \Omega, s \in S.
$$

using the conditions below, for $(\lambda, \mu)$, in the expression above

$$
\theta(s, s') = \frac{\mu(s', s)}{\sum_{s' \in s} \mu(s', s)} \text{ if } \sum_{s' \in s} \mu(s', s) > 0
$$

Condition (a) holds by substitution, (b) by definition, and (c) is equivalent to

be matrix with $(i, i)$th element $\lambda(s)$, $D$ be matrix with $(i, i)$th element
with $(i, j)$th co-ordinate $= \mu(s, s') / d(s, s)$. $L$ and $D$ are non-negative

$\Theta$ is equivalent to [4]

If condition [2] holds, then $M = L + D(1-\Theta)$ satisfies (i), (ii) and (iii) in condition [3]. If condition [3]
holds, then let the $(i, j)$th element of diagonal matrix $L$ equal the $i$th component of vector $M1$. Let the
$(i, i)$th element of diagonal matrix $D$ be the $i$th element of $M - L$. Finally let $\Theta = -D^{-1}(M - L)$.
Because $M$ satisfies (i), (ii) and (iii), $L$ and $D$ are non-negative and $\Theta$ is Markov $\blacksquare$

Proof of theorem 1 Lemmas 1 and 2 imply the first part of the theorem. For dynamic consistency, note that
the problem is separable for each signal, i.e. beliefs satisfy dynamic consistency if and only if for
every decision problem $(A, u)$, and every signal $s \in S$,

$$
\sum_{\omega \in \Theta} \gamma(\omega | s) u[a, \omega] \geq \sum_{\omega \in \Theta} \gamma(\omega | s) u[a, \omega], \text{ for all } a \in A.
$$

$$
\sum_{\omega \in \Theta} \pi(\omega, s) u[a, \omega] \geq \sum_{\omega \in \Theta} \pi(\omega, s) u[a, \omega], \text{ for all } a \in A.
$$

Clearly this is true if beliefs satisfy Bayes updating. If beliefs do not satisfy Bayes updating,
there exists $\omega' \in \Omega, s' \in S$ and $k \in (0, 1)$ such that

$$
\gamma(\omega' | s') > k > \frac{\pi(\omega', s')}{\sum_{\omega \in \Omega} \pi(\omega, s')}
$$

Now consider decision problem where $A = (a, a_0)$, $u(a_0, \omega) = 0$ for all $\omega \in \Omega$, $u(a_0, \omega) = 1-k$ and $u(a_0, \omega') = -k$, for all $\omega \neq \omega'$.

$$
\sum_{\omega \in \Theta} \gamma(\omega | s') u[a, \omega] > \sum_{\omega \in \Theta} \gamma(\omega | s') u[a_0, \omega].
$$

but

$$
\sum_{\omega \in \Theta} \pi(\omega, s') u[a_0, \omega] = 0 > \sum_{\omega \in \Theta} \pi(\omega, s') u[a_0, \omega]. \blacksquare
$$

The dynamic consistency result provides a decision theoretic justification for Bayes Rule. The
axioms of expected utility theory provide a static account of choices under uncertainty. They represent
the requirement that agents' choices are coherent or consistent. The axioms do not relate to the question
of whether the agent processes information correctly, so they can hardly imply Bayes rule, as is often
loosely assumed. Hacking (1967) noted that an additional dynamic consistency assumption is implicit in
many interpretations of subjective expected utility theory: choices after observing information are the same as the conditional choices that would be made before observing the information\textsuperscript{11}.  

Section 4: Possibility Correspondences

In the previous section, a distinction was made between the "payoff-relevant" states of the world \( \Omega \) and the signals \( S \) that the agent observes. The content of this distinction is that we look at decision problems which depend only on the payoff-relevant states, and not on the signals. We can imagine looking at richer classes of decision problems as we expand \( \Omega \) to include more complicated descriptions of the world (with the true probabilities \( \pi \) and \( \pi \) post beliefs \( \gamma \) expanded correspondingly). In this section, we consider only cases where the space of payoff-relevant events is sufficiently rich, relative to the set of signals, so that, for each payoff-relevant state, there is only one possible signal, i.e. for each \( \omega \in \Omega \), there is a unique \( s \in S \) such that \( \pi(\omega, s) > 0 \).  

The advantage of pursuing this approach - requiring the class of decision problems to be sufficiently rich - is that it is possible to relate the results of section 3 to work in epistemic logic which considers (explicitly or implicitly) a state space which includes a description of agents' knowledge as part of the description of the state. In particular, Geanakoplos (1989) has derived important results (discussed below) on the value of information in such a setting. In this section, I briefly report some results, discussed in more detail in companion papers Morris (1992b) and Morris (1992c), on the relation between this paper and Geanakoplos (1989).

Consider a finite state space \( \Omega \), where each state can be thought of as a complete description of relevant aspects of the world. In each state \( \omega \in \Omega \), there is a non-empty subset of states \( P(\omega) \subseteq \Omega \) which are thought possible. Now information is represented by a possibility correspondence, \( P : \Omega \rightarrow 2^\Omega \).


12. If we think of the payoff-relevant state space \( \Omega \) as representing possible truth valuations of some set of propositions in some language, then the assumption that there is a unique signal for every state of the world is equivalent to the requirement there is a proposition in the language representing every state of belief. Or, as Skyrms (1990) puts it, "does knowledge always come nicely packaged as a proposition in one's subjective probability space?" (page 92).
Section 4: Possibility Correspondences

On, a distinction was made between the "payoff-relevant" states of the world agent observes. The content of this distinction is that we look at decisions on the payoff-relevant states, and not on the signals. We can imagine decision problems as we expand \( \Omega \) to include more complicated descriptions (probabilities \( \pi \) and \( \varepsilon \) post beliefs \( \gamma \) expanded correspondingly). In this case where the space of payoff-relevant events is sufficiently rich, relative to each payoff-relevant state, there is only one possible signal, i.e. for each \( S \) such that \( \pi(\omega, s) > 0 \).

Pursuing this approach - requiring the class of decision problems to be possible to relate the results of section 3 to work in epistemic logic which (sic) a state space which includes a description of agents' knowledge as part of. In particular, Geanakoplos (1989) has derived important results (discussed in formation in such a setting. In this section, I briefly report some results, from companion papers Morris (1992b) and Morris (1992c), on the relation between 1989).

The space \( \Omega \), where each state can be thought of an a complete description of. In each state \( \omega \in \Omega \), there is a non-empty subset of states \( P(\omega) \subset \Omega \) which information is represented by a possibility correspondence, \( P: \Omega \to 2^\Omega \).

Let \( P \subset 2^\Omega \) be the range of \( P \). This representation of information has been used by a number of recent papers in the economics literature.

An agent's choices will depend not only on information (which states are possible), but also on beliefs. Suppose there is a "true", strictly positive, probability distribution \( \pi \) on \( \Omega \). Beliefs \( \gamma: \Omega \times P \to R \), are conditional beliefs if \( C = \{ \omega \in \Omega \mid \gamma(\omega, C) > 0 \} \), \( \forall C \in P \). This possibility correspondence framework can now be seen as a special case of the framework in the previous section, with \( S = P \), and \( \pi(\omega, C) = \pi(\omega) \) if \( C = \{ \omega \} \), 0 otherwise.

Define a decision problem as above: a finite action set \( A \) and a function \( u: A \times \Omega \to R \). A decision rule, \( f: P \to A \), is optimal, given conditional beliefs \( \gamma \), if, for every action \( a \in A \),

\[
\sum_{\omega \in C} \gamma(\omega, C) u[f(C), \omega] \geq \sum_{\omega \in C} \gamma(\omega, C) u[a, \omega]
\]

is valued if there exist conditional beliefs \( \gamma \) such that for every decision problem, every optimal decision rule \( f \), and every action \( a \in A \),

\[
\sum_{\omega \in \Omega} \pi(\omega) u[f(P(\omega)), \omega] \geq \sum_{\omega \in \Omega} \pi(\omega) u[a, \omega]
\]

\( P \) satisfies dynamic consistency if there exist conditional beliefs \( \gamma \) such that, for every decision problem and every optimal decision rule \( f \), and every decision rule \( g \),

\[
\sum_{\omega \in \Omega} \pi(\omega) u[f(P(\omega)), \omega] \geq \sum_{\omega \in \Omega} \pi(\omega) u[g(P(\omega)), \omega]
\]

The possibility correspondence theorem will depend on some critical properties of the possibility correspondences, properties which have a long history of discussion among philosophers before their importance relevance in decision theory was realized. \( P \) satisfies non-delusion if \( \omega \in P(\omega) \) for all \( \omega \in \Omega \). The true state of the world is always thought possible. \( P \) satisfies knowing that you know if \( \omega' \in P(\omega) \Rightarrow P(\omega') \subset P(\omega) \). If state \( \omega' \) is thought possible when the true state is \( \omega \), then any state thought possible at \( \omega' \) is also thought possible at \( \omega \). \( P \) is a partition if \( P(\omega) = \{ \omega \in \Omega \mid P(\omega') = P(\omega) \} \). This

---


13. Brandenburger, Dake and Geanakoplos (1989) called these structures possibility correspondences: they have also been variously described as generalized information structures and generalized partitions; see references in note 4 above.

14. The equivalence between the possibility correspondence representation and a "signal space" or type space representation was shown (in a more general, many agent, context) in Brandenburger, Dake and Geanakoplos (1989). That result is discussed further below.

15. See Halpern (1986) and references in note 16.
is the standard economists' representation of information, which is strictly stronger than non-delusion and knowing that you know combined.

The significance of these conditions is clearest if we translate them from this possibility correspondence framework into a knowledge operator representation\(^{16}\). You know that an event A is true if every state which you think possible is an element of A. Thus let \( K_p(A) = \{ \omega \in \Omega \mid P(\omega) \subseteq A \} \). Now the interpretation is that \( K_p(A) \) is the set of states where you "know" event A. Then \( P \) satisfies non-delusion if and only if \( K_p(A) \subseteq A \), for all \( A \subseteq \Omega \). If you know A, then A must be true (i.e. the true state is contained in A). \( P \) satisfies knowing that you know if and only if \( K_p(A) \subseteq K_p(K_p(A)) \), for all \( A \subseteq \Omega \). If you know A, then you know that you know it. \( P \) is a partition if, and only if, in addition to satisfying non-delusion and knowing that you know. \( -K_p(A) \subseteq K_p(-K_p(A)) \), for all \( A \subseteq \Omega \). If you don't know A, then you know that you don't know it\(^{19}\).

**THEOREM 2 (Possibility Correspondence)**  
\( P \) is valuable if and only if it satisfies non-delusion and knowing that you know. \( P \) satisfies dynamic consistency if and only if \( P \) is a partition.

The theorem can be proved in different ways. In a companion paper, Morris (1992c), a constructive proof is given, which in fact generalizes beyond the expected utility setting of this paper: if non-delusion or knowing that you know fail, a decision problem can be constructed where ex ante utility is strictly lower if information is observed; if non-delusion and knowing that you know hold, then the conditional beliefs ensuring the theorem is true can be explicitly constructed.

Here, we are comparing a possibility correspondence with no information. In Morris (1992b), possibility correspondences are compared with each other, and constructive arguments are not possible. A double application of linear programming methods can be used to characterize properties of possibility correspondences such that one is more valuable than another. First, conditions on beliefs are derived by the dual methods as in the previous section. Then, the existence of beliefs consistent with a given knowledge structure and satisfying more valuable restrictions constitutes another linear programming problem. Theorem 2 is a special case of results in that paper.

---

16. See Aumann (1989), Binmore and Brandenburger (1990) and Geanakoplos (1992) for detailed discussions of these relationships.
17. This property has various names in the literature: non-delusion [Geanakoplos (1989)], the axiom of knowledge [Binmore and Brandenburger (1990)], the axiom of correctness [Ferrante (1991)].
presentation of information, which is strictly stronger than non-delusion and
sincere.

These conditions are clearest if we translate them from this possibility
into a knowledge operator representation. You know that an event A is
think possible is an element of A. Thus let \( K_\alpha(A) = \{ \omega \in \Omega \mid P(\omega) \subset A \} \). \( K_\alpha(A) \) is the set of states where you "know" event A. Then P satisfies non-
C A, for all A \( \subset \Omega \). If you know A, then A must be true (i.e. the true
satisfies knowing that you know if and only if \( K_\alpha(A) \subset K_\alpha(K_\alpha(A)) \), for all
when you know that you know it. P is a partition if, and only if, in addition
knowing that you know, \( \neg K_\alpha(A) \subset K_\alpha(\neg K_\alpha(A)) \), for all A \( \subset \Omega \). If you
now that you don't know it.

Theorem 1 \[ P \text{ is valuable if and only if it satisfies non-delusion and }
\] satisfies dynamic consistency if and only if P is a partition.

proved in different ways. In a companion paper, Morris (1992c), a
which in fact generalizes beyond the expected utility setting of this paper:
that you know fail, a decision problem can be constructed where ex ante
information is observed; if non-delusion and knowing that you know, then
the theorem is true can be explicitly constructed.

Having a possibility correspondence with no information. In Morris (1992b),
were compared with each other, and constructive arguments are not possible.
programming methods can be used to characterize properties of possibility
is more valuable than another. First, conditions on beliefs are derived by
previous section. Then, the existence of beliefs consistent with a given
satisfying more valuable restrictions constitutes another linear programming
factual case of results in that paper.

Geanakoplos (1989) first considered the question of whether a possibility correspondence is
valuable in the setting described above. He assumed, however, that beliefs after observing the information were generated by what I will call naive Bayes updating: \( \gamma(\omega \mid C) = \pi(\omega) / \sum_{\omega' \in C} \pi(\omega') \). The rationale for this assumption is discussed below. This additional restriction on beliefs results in an
additional restriction on possibility correspondences: P satisfies nestedness if P(\( \omega \)) \( \cap \) P(\( \omega' \)) = \( \emptyset \), P(\( \omega \)) \( \cup \) P(\( \omega' \)) = \( \Omega \). There is a very natural interpretation of non-delusion, knowing that you know and
interpretation.

**Theorem 2 [Geanakoplos (1989)]** Suppose conditional beliefs \( \gamma \) satisfy naive Bayes updating. Then
P is valuable if and only if it satisfies non-delusion, knowing that you know and nestedness.

The theorem was proved by induction on the size of \( \Omega \), but can also be proved by the two
methods outlined above (see Morris (1991) chapter 3). An example from Geanakoplos (1989) illustrates
the significance of nestedness and the difference between theorem 2 and the original result.

**Example**

\( \Omega = \{ a, b, c \} \)
\( \pi = (2/7, 3/7, 2/7) \)
\( P(a) = \{ a, b \} \)
\( P(b) = \{ b \} \)
\( P(c) = \{ b, c \} \)

P satisfies non-delusion, knowing that you know, but not nestedness. Naive Bayes updated beliefs
are \( \gamma(a \mid \{ a, b \}) = \gamma(c \mid \{ b, c \}) = 2/5; \gamma(b \mid \{ a, b \}) = \gamma(b \mid \{ b, c \}) = 3/5; \) and \( \gamma(b \mid \{ b \}) = 1 \).

Suppose A = {B,N} ("bet" or "don't bet"). Let u(N,a) = u(N,b) = u(N,c) = 0. Let u(B,a) = u(B,c) = -1, while u(B,b) = 1. Consider decision rule f(\( \omega \)) = B, \( \forall \omega \in \Omega \): f is optimal given Bayes updated beliefs, but this decision rule gives expected payoff of -1/7, strictly less than 0 which could be
achieved by always not betting. So observing P makes the agent worse off in terms of ex ante utility.

On the other hand, consider arbitrary set of conditional beliefs, \( \gamma \), consistent with P. Let \( \gamma(a \mid \{ a, b \}) = \alpha \) and \( \gamma(c \mid \{ b, c \}) = \beta \), where \( \alpha, \beta \in (0,1) \). Thus \( \gamma(b \mid \{ a, b \}) = 1-\alpha \) and \( \gamma(b \mid \{ b, c \}) = 1-\beta \); of course, \( \gamma(c \mid \{ c \}) = 1 \). Now it is a consequence of theorem 2 that there exists beliefs \( \gamma \), consistent with P, such that P is valuable. Specifically, P is valuable if and only if \( 1/\alpha + 1/\beta \leq 7/2 \). It may be helpful to
understanding the theorem to see how the linear algebra drives this result. Consider a decision
problem [A,u], and optimal decision rule f. For every action d \( \in A \), optimality of f implies the following
five inequalities hold:-

1. \( u(d) \geq u(f(d)) \) for all d \( \in A \).
2. \( u(f(a)) \geq u(f(b)) \) for all a, b \( \in A \) with a \( \neq b \).
3. \( u(f(a \cdot b)) \geq u(f(a) \cdot f(b)) \) for all a, b \( \in A \).
4. \( u(f(a \cdot b)) \geq u(f(a) \cdot f(b)) \) for all a, b \( \in A \) with a \( \neq b \).
5. \( u(f(a \cdot b)) \geq u(f(a) \cdot f(b)) \) for all a, b \( \in A \) with a \( \neq b \).

These inequalities are consequences of the axioms of rationality and consistency.
\[ \alpha u[f(a,b),a] + (1 - \alpha)u[f(a,b),b] \geq \alpha u[d,a] + (1 - \alpha)u[d,b], \]
\[ u[f(b),b] \geq u[d,b], \]
\[ \beta u[f((b,c),c)] + (1 - \beta)u[f((b,c),b)] \geq \beta u[d,c] + (1 - \beta)u[d,b]. \]
\[ u[f((b,c),b)] \geq u[f(a,b),b], \]
\[ u[f(b),b] \geq u[f((b,c),c),b]. \]
\[ \frac{2}{7} u(f(P(a),a) + \frac{2}{7} u(f(P(b),b) + \frac{2}{7} u(f(P(c),c) + \frac{2}{7} u(d,a) + \frac{2}{7} u(d,b) + \frac{2}{7} u(d,c) \]

Multiplying these five inequalities by the constants on the right, and summing, gives the inequality on the bottom - as long as the constants are non-negative, i.e. \(1/\alpha + 1/\beta \leq 7/2\). We can check that naive Bayes updating (\(\alpha = \beta = 2/5\), \(1/\alpha + 1/\beta = 5\)) fails this condition.

So should we assume naive Bayes updating? As argued in the introduction, the results of section 3 provide one rationale for not assuming naive Bayes updating. The decision theoretic justification for Bayes updating is dynamic consistency, a condition which is stronger than the information is valuable criterion. On the other hand, an argument in Brandenburger, Dekel and Geanakoplos (1989) appears to give a rationale for naive Bayes updating. Consider state space \(\Omega\), possibility correspondence \(P\) (with range \(P\)), prior \(\pi\) and conditional beliefs, \(\gamma\), not necessarily derived by Bayes updating. Now fix a decision problem \((A,u)\) depending on \(\Omega\). We can construct an equivalent representation of this problem where beliefs are Bayes updated and information is represented by a partition. Let \(\Omega' = \Omega \times P\). Define possibility correspondence \(P'\) on \(\Omega'\) by \(P'((\omega,C)) = \{ (\omega',C') \mid C' = C \}\). Define prior \(\pi'\) on \(\Omega'\) by \(\pi'((\omega,C)) = \pi(\omega)\) if \(C = P(\omega), 0\) otherwise. Define another prior \(\gamma'\) on \(\Omega'\) by \(\gamma'((\omega,C)) = \lambda(C) \gamma(\omega|C)\), for some strictly positive distribution \(\lambda\) on \(P\). Finally, define \(u^*\): \(A \times \Omega' \to R\) by \(u^*[a, (\omega,C)] = u[a, \omega]\), for all \(a \in A, \omega \in \Omega, C \in P\). Now if we derive conditional beliefs after observing \(P\) by naive Bayes updating, the initial decision problem \([\Omega, P, \pi, \gamma, A, u]\) is equivalent to the decision problem \([\Omega^*, P', \pi', \gamma', A, u^*]\) with naive Bayes updated conditional beliefs. Note that we cannot in general guarantee that \(\pi^* = \gamma^*\). Does this mean that the assumption of naive Bayes updating is without loss of generality in some sense? This is true if \(we are only interested in the one decision problem (A,u). But properties like dynamic consistency and valuable information are not about a particular decision problem, but depend on the class of decision problems contingent on the uncertainty space. By expanding the state space from \(\Omega\) to \(\Omega^*\), we have changed the problem. For this reason, the assumption of naive Bayes updating is not
\[ u(f(b)), b) \geq u(f(b)), b) \geq u(f(b)), b) \geq u(f(b)), b) \geq u(f(b)), b) \geq u(f(b)), b) \]
\[ 1 - \frac{1}{1 + \beta} \]

\[ \gamma(w, s) > 0 \}

An analogue of theorem 2 is also given without the restriction that only one signal is observed at each payoff-relevant state of the world. In (Morris 1992c), the static expected utility assumption is dropped, and theorem 2 is derived directly from axioms (which do not imply expected utility maximization) on preferences.

Section 5: Speculation

Geanakoplos (1989) showed that the property that information is valuable underlies various no speculation results. In this section, I show that an analogue to that result for the framework of section 3. The following notion of speculation is adapted from the definition in Geanakoplos (1989). Suppose there are |A| agents, \{1, ..., |A|\}, a finite uncertainty space \( \Omega \), and each agent observes a signal, \( s_i \in S_i \). Write \( S = S_1 x ... x S_{|A|} \) with generic element \( s \), and \( S_i = S_1 x ... x S_{|A|} \), with generic element \( s_i \). Suppose the "true" probability distribution over \( \Omega \times S \) is \( \pi \). Suppose that after observing signal \( s_i \), agent \( i \) has beliefs \( \gamma(w, s_i | s_i) \) about \( \Omega x S \).

A game consists of a finite set of actions for each agent \( i \), \( A_i \), and a payoff function for each agent \( i \), \( u_i : \Omega x S_i \rightarrow \mathbb{R} \), where \( A = A_1 x ... x A_i \). Each agent \( i \) chooses an action rule, \( e_i : S_i \rightarrow A_i \). Agent \( i \)'s interim payoff (after observing signal \( s_i \)) if he chooses action \( a_i \) and other agents use rules \( e_j = (e_1, e_2, ..., e_{i-1}, e_i, e_{i+1}, ..., e_{|A|}) \) is

\[ v_i(e_i, s, a_i) = \sum_{\omega \in \Omega} \sum_{x \in S_i} \gamma(\omega, s, a_i) u_i(\{a_i, e_i(\omega, s, a_i)\}, \omega) \]

Profile of action rules \( f \) is a Nash equilibrium if \( v_i(e_i, s, a_i) \geq v_i(f, s, a_i) \) for all \( i, a_i \in A_i \) and \( s_i \in S_i \). Define agent \( i \)'s maxmin action \( d_i \) and maxmin payoff \( m_i \) by

\[ d_i \in \text{argmax}_{a_i \in A_i} \left\{ \min_{\omega \in \Omega} \sum_{x \in S_i} \pi(\omega, s) u_i(\{a_i, e_i(\omega, s, a_i)\}, \omega) \right\} \]

\[ m_i = \min_{f_i} \sum_{\omega \in \Omega} \sum_{x \in S_i} \pi(\omega, s) u_i(\{d_i, f_i(x)\}, \omega) \]
Thus agent i could choose to play action $a_i$ always (whatever his signal) and guaranteeing himself $m_i$. Say that profile of action rules f is speculative if $\Sigma_{\omega \in \Omega} \Sigma_{s \in S_i} \pi(\omega, s) u_i(f(s), s, \omega) < m_i$ for some i. Thus in a speculative Nash equilibrium some agent i could have achieved the guaranteed ex ante utility $m_i$ by choosing the maxmin action $a_i$, but takes a poor gamble instead.

**THEOREM 3:** There does not exist a speculative Nash equilibrium in any game if and only if each agent's conditional beliefs $\gamma_i$ are noisy Bayes updates of $\pi$.

**Proof of theorem 3** Suppose $f$ is a Nash equilibrium. Then for all $i, s_i \in S_i, a_i \in A_i$

$$\sum_{\omega \in \Omega} \sum_{s \in S_i} \gamma(\omega, s, f) u_i(f(s), s, \omega) < \sum_{\omega \in \Omega} \sum_{s \in S_i} \gamma(\omega, s, f) u_i(a_i(f(s), s), \omega)$$

If, for each i, $\gamma_i$ is a noisy Bayes update of $\pi$, then by theorem 1, for each i,

$$\sum_{\omega \in \Omega} \sum_{s \in S_i} \pi(\omega, s) u_i(f(s), \omega) < \sum_{\omega \in \Omega} \sum_{s \in S_i} \pi(\omega, s) u_i(d_i, \omega) \geq m_i$$

Conversely, if $\gamma_i$ is not a Bayes update of $\pi$ for any i, then by theorem 1 there exists finite set $A_i, y: A_i \times \Omega \times S_i \rightarrow \mathbb{R}$ and $f_i: S_i \rightarrow A_i$ such that

$$\sum_{\omega \in \Omega} \sum_{s \in S_i} \gamma(\omega, s, f) y(s, \omega) \geq \sum_{\omega \in \Omega} \sum_{s \in S_i} \gamma(\omega, s, f) y(a_i^*, s, \omega), \text{ for all } s \in S_i, a_i \in A_i$$

but

$$\sum_{\omega \in \Omega} \sum_{s \in S_i} \pi(\omega, s) y(f(s), s, \omega) < \sum_{\omega \in \Omega} \sum_{s \in S_i} \pi(\omega, s) y(a_i^*, s, \omega), \text{ for some } a_i^* \in A_i$$

Now for each $j \neq i$, let $A_j = S_j, u_j(\omega, s) = 0$ for all $\omega \in \Omega$ and $f_j(s) = s$ for all $s \in S_j$. Let $w_i[a(s, \omega)] = y(s, \omega) - y(a_i^*, s, \omega)$ for all $a_i \in A_i, s_i \in S_i$, and $\omega \in \Omega$. Maxmin payoffs in this game are at least 0 for each agent (action $a_i^*$ guarantees agent i utility 0). But profile f is a Nash equilibrium with $\Sigma_{\omega \in \Omega} \Sigma_{s \in S_i} \pi(\omega, s) u_i(f(s), \omega) < 0$.

How does this notion of speculation relate to no trade results? In chapter IV of Morris (1991), it is shown that for risk neutral agents, there exists a speculative Nash equilibrium of some game, for some group of agents $\{1, \ldots, i\}$ and "true" probability distribution $\pi$ if and only if there exists a feasible incentive compatible trade between the agents and an (i+1)th agent ("nature") with beliefs $\pi$ and no information. One implication of this relation is that conditions on each agent's beliefs to preclude speculation can serve as a standard of in-advance belief updating. On the one hand, if there is no incentive consistent belief updating belief set $\mathcal{A}$, the corresponding correspondence $\mathcal{B}$, and the results of static expected utility analysis can be valuable information for knowing that all economic assumptions are satisfied.

The above discussions of speculation in computer science and cybernetics are not strong. Our "common" assumption of rationality" is that agents know that there is incentive compatibility, they would have no incentive to speculate.

---

20. See, for example, Spohn (1987).
21. A number of recent contributions have used computer science to generate results. Spohn (1987).
play action \( a \) always (whatever his signal) and guaranteeing himself \( m_1 \), as \( f \) is speculative if \( \sum_{s \in S} \sum_{i \in I} \pi(\omega, s) u_i[f(s), \omega] < m_1 \) for some \( i \). Thus in a some agent \( i \) could have achieved the guaranteed ex ante utility \( m_1 \) by \( \omega \), but takes a poor gamble instead.

not exist a speculative Nash equilibrium in any game if and only if each are noisy Bayes updates of \( \pi \). f is a Nash equilibrium. Then for all \( i, s_1, s_2, s_3 \in S \), \( s_1, s_2, s_3 \in A \)

\[
|\pi(s_3) \pi(s_2) \pi(s_1)| \geq \sum_{\omega \in \Omega} \sum_{s_3, s_2, s_1} \gamma(\omega, s_3, s_2, s_1) u_i[\pi(s_3, s_2, s_1), \omega]
\]

ays update of \( \pi \), then by theorem 1, for each \( i, \omega, s \in S \), \( s \in A \)

\[
\sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega, s) u_i[\pi(s), \omega] \geq m_i
\]

ot a Bayes update of \( \pi \) for any \( i \), then by theorem 1 there exists finite set \( S \rightarrow A \) such that

\[
\sum_{\omega \in \Omega} \sum_{s \in S} \gamma(\omega, s, x, \omega) y(a_s, x, \omega), \text{for all } x \in S \), \( s \in A \)
\]

\[
\sum_{\omega \in \Omega} \sum_{s \in S} \pi(\omega, s) y(a_s^+, x, \omega), \text{for some } \omega \in S \), \( \omega \in A \)
\]

\[
A_1 = S \), \( u_1(s, \omega) = 0 \text{ for all } s \in A \), \( \omega \in \Omega \) and \( f(s) = s \) for all \( s \in S \). Let \( \pi \) for all \( s \in A \), \( s \in S \) and \( \omega \in \Omega \). Maximin payoffs in this game are on \( a^* \) guarantees agent i utility 0. But profile f is a Nash equilibrium with

speculation can be characterized independently of each other (as in theorem 3). There is some absolute standard of information processing error allowed. In particular, there can be speculation with only one agent, or if there are more than one agent observing identical signals and with identical (mistaken) updating. On the other hand, conditions on beliefs which preclude trade [Morris (1992b)] are relative. If there is only one agent, or if all agents observe the same signals and make the same mistakes in updating beliefs, there is still no trade, whatever the mistake.

Chapter IV of Morris (1991) also shows how this result relates to possibility correspondence results. Geanakoplos (1989) showed that there is no speculation among a group of agents with possibility correspondences, and naive Bayes updated beliefs, if and only if those possibility correspondences all satisfy non-delusion, knowing that you know and nestingness. The results of section 3 can be used to show that there exist conditional beliefs such that there is no speculation if and only if those possibility correspondences satisfy non-delusion and knowing that you know.

Section 6: Conclusion

The main contribution of this paper has been to show the correspondence, under the assumption of static expected utility maximization, between dynamic axioms on preferences (dynamic consistency and valuable information), rules for updating beliefs (Bayes and noisy Bayes updating) and representations of new information (as a partition and as a possibility correspondence satisfying non-delusion and knowing that you know). Thus the agenda of Savage (1954) of reducing properties of (static) beliefs to assumptions on preferences is extended to rules for updating beliefs and knowledge.

These relationships are of particular interest in the light of an apparent dichotomy between discussions of belief and knowledge in the literature. There is widespread consensus among philosophers, computer scientists and others engaged in representing knowledge that the partition representation is too strong. On the other hand, Bayes updating is largely unquestioned. Yet the same kind of "bounded rationality" motivates dropping the partition representation and Bayes updating: people do not necessarily know that they do not know something they do not know; and they do not take into account what they would have believed if they had observed a different signal. This paper gave a precise sense in which

---

20. See, for example, Halpern (1986) and subsequent volumes of Theoretical Aspects of Reasoning about Knowledge.
21. A number of authors have pointed out that with consensus on rules for updating beliefs, it ought to be possible to generate analogous rules for knowledge, at least if knowledge is equated with beliefs with probability one [e.g. Spohn (1988)].
partition information is the knowledge analogue of Bayes updating, while non-deluded possibility correspondences satisfying knowing that you know are the knowledge analogue of naive Bayes updating.

The importance of these issues for economic problems was illustrated by giving an analogue of a result of Geanakoplos (1989) on the equivalence between no speculation in economic environments and the valuable information property. More generally, the methods of representing dynamic choice under uncertainty allow us to model a particular kind of unforeseen contingencies: agents are assumed to understand the structure of payoff-relevant events (otherwise expected utility assumptions would not make sense), but not necessarily anticipate information they might receive or might have received⁹.

These results are extended in two directions in companion papers. In this paper, we ask when information is more valuable than no information. More generally, following Blackwell’s Theorem but without assuming Bayes updating, we can ask when one “experiment” is more valuable than another experiment. This problem is solved in Morris (1992b). An experiment is more valuable than another if and only if it is sufficient in differences for the other. This result specializes to Blackwell’s Theorem when Bayes updating is used, and to theorem 1 of this paper if the latter experiment is the “empty experiment” with no information. This theorem can be used to generate results about knowledge in a possibility correspondence framework and in a less restrictive framework where more than one signal can be generated in a given payoff-relevant state of the world. Suppose information at time 1 is represented by non-deluded possibility correspondence \( P_1 \), and information at time 2 is represented by possibility correspondence \( P_2 \). Then the ex ante utility of choosing at time 2 is higher than the ex ante utility of choosing at time 1, for some conditional beliefs, for every decision problem, if and only if \( P_2 \) satisfies non-delusion and knowing that you know and, for every event \( A \) known (not known) at time 1 in state \( \omega \), it is known at time 2 in state \( \omega \) that \( A \) was known (not known) at time 1. Morris (1992c) examines the questions raised in this paper, without the static expected utility assumption, by primitive restrictions on preferences. Knowledge is defined directly as a property of preferences, without the mediation of beliefs.

---

22. Kreps (1992) provides a model of desire for flexibility in the face of unforeseen contingencies, where those unforeseen contingencies are payoff-relevant. Here, by contrast, it is information signals (equivalently, changes in beliefs), not payoff-relevant events, which are not anticipated.
knowledge analogue of Bayes updating, while non-deluded possibility knowing that you know are the knowledge analogue of naive Bayes updating. These issues for economic problems was illustrated by giving an analogue of (89) on the equivalence between no speculation in economic environments and... More generally, the methods of representing dynamic choice under risk a particular kind of unforeseen contingencies: agents are assumed to payoff-relevant events (otherwise expected utility assumptions would not make anticipate information they might receive or might have received. It is extended in two directions in companion papers. In this paper, we ask when there is more information. More generally, following Blackwell’s Theorem but updating, we can ask when one “experiment” is more valuable than another. This is solved in Morris (1992b). An experiment is more valuable than another in differences for the other. This result specializes to Blackwell’s Theorem and, to theorem 1 of this paper if the latter experiment is the “empty” experiment. This theorem can be used to generate results about knowledge in a framework and in a less restrictive framework where more than one signal can off-relevant state of the world. Suppose information at time 1 is represented by correspondence P₁, and information at time 2 is represented by possibility he ex ante utility of choosing at time 2 is higher than the ex ante utility of the conditional beliefs, for every decision problem if and only if P₂ satisfies that you know and, for every event A known (not known) at time 1 in state st at A was known (not known) at time 1. Morris (1992c) examines paper, without the static expected utility assumption, by primitive restrictions which are defined directly as a property of preferences, without the mediation of

References


Hacking I (1967) Slightly more realistic personal probability. Philosophy of Science 34: 311–325.


