Rationality and the Dutch Book argument (i.a. Halpern 2.2.1)

- If belief is a quantified probability, then it is important to explain what the numbers represent, where they come from, and why finite additivity is appropriate.

- The "Dutch Book" argument (DBA) tracing back to independent work by F.Ramsey (1926) and B.deFinetti (1937), offers prudential grounds for action in conformity with **personal probability**.

- DBA represent the possibility of a new kind of justification for epistemological principles (Kolmogorov's axioms of probability).

- A DBA relies on some descriptive or normative assumptions to connect degrees of belief with willingness to wager -- for example, a person with degree of belief p in sentence S is assumed to be willing to pay up to and including \$$p$ for a **unit wager on *S*** (i.e., a wager that pays \$1 if S is true) and is willing to sell such a wager for any price equal to or greater than \$$p$. A rational person is assumed to be equally willing to buy or sell such a wager when the price is exactly \$$p$.

- DBAs can be used to check the (in) consistency of probability judgements.

A *Dutch Book* is a combination of wagers which, on the basis of deductive logic alone, can be shown to entail a sure loss.

Example 1

Suppose that agent's degrees of belief in S and $\sim S$ (written $bel(S)$ and $bel(\sim S)$) are each .51, and, thus that their sum 1.02 (greater than one). On the behavioral interpretation of degrees of belief introduced above, the agent would be willing to pay $bel(S) \times \$1$ for a unit wager on S and $bel(\sim S) \times \$1$ for a unit wager on $\sim S$. If a bookie B sells both wagers to our agent for a total of \$1.02, the combination would be a *synchronic Dutch Book* -- synchronic because the wagers could both be entered into at the same time, and a *Dutch Book* because the agent would have paid \$1.02 on a combination of wagers guaranteed to pay exactly \$1. Thus, the agent would have a guaranteed net loss of \$.02

Let $U \subseteq W$ be an event (set of possibilities)

Then a **bet *b* on *U*** is a triple $b = [U, S, \alpha]$, where $S \geq 0$ and $0 \leq \alpha \leq 1$.

S is called the stake of b
α is called the betting quotient of b (and $\alpha/(1-\alpha)$ the odds of b)

	U	**~U**
bet on U	win $(1-\alpha)S$	lose αS
bet against U	lose $(1-\alpha)S$	win αS
abstain from betting	status quo	status quo

The bet becomes less and less attractive as α gets larger!

Example 2
Suppose you wager 10$ on the complete outsider *Born Loser* with scores 19:1 at the bookmakers. Then the **stak**e of your bet is 200 (=19x10+1x10). The **odds** are 1:19 and the **betting quotient** is 0.05.

Remark: the bet $\sim b = [\sim U, S, 1-\alpha]$ is called the complementary bet for b.

Definition "value of a bet"

Let $b = [U, S, \alpha]$ be a bet and let E be an U-specific event (i.e. $E \subseteq U$ or $E \subseteq \sim U$), then the value $\|b\|_E$ of the bet b at the event U is

$$\|b\|_E = \begin{cases} (1-\alpha)\mathrm{S} & \text{if } E \subseteq U \\ -\alpha\mathrm{S} & \text{if } E \subseteq \sim U \end{cases}$$

Definition "book" and "value of a book"

For a given algebra of events, a **book B** is a finite set of bets on certain events of the algebra such that $[U, \mathrm{S}, \alpha] \in \mathbf{B}$, $[U, \mathrm{S}, \alpha'] \in \mathbf{B} \Rightarrow \alpha = \alpha'$.
The value of a book with regard to an book-specific event E, $\|\mathbf{B}\|_E$, is the sum of the values of the bets contained in the book.
[Note: regarding a book **B**, E is a book-specific event iff E is an U-specific event for all bets $[U, S, \alpha] \in \mathbf{B}$].

Definition “Dutch book”

A book **B** with regard to a given algebra of events is called a **Dutch book** if for every book-specific event E, $\|\mathbf{B}\|_E < 0$. (Hence, the agent will have a guaranteed net loss!).

Example 1, continued

$\mathbf{B} = \{[S, \$1, 0.51], [\sim S, \$1, 0.51]\}$ is a Dutch book. It is simply to show that for each event E that either is contained in S or in $\sim S$, the value of **B** is –0.02:

case 1: $E \subseteq S$, then $\|\mathbf{B}\|_E = 0.49 \text{ x } \$1 - 0.51 \text{ x } \$1 = -\$0.02$

case 2: $E \subseteq \sim S$, then $\|\mathbf{B}\|_E = -0.51 \text{ x } \$1 + 0.49 \text{ x } \$1 = -\$0.02$

The aim is to give a general characterization of the bets a rational agent will accept.

Definition "acceptance set"

For a given algebra of events, the *acceptance set* of an rational agent X is a set Acc_X such that

1. If $[U, S, \alpha] \in \mathrm{Acc}_X$ and $\lambda > 0$, then $[U, \lambda S, \alpha] \in \mathrm{Acc}_X$
2. If $[U, S, \alpha] \in \mathrm{Acc}_X$ and $0 \leq \alpha' \leq \alpha$, then $[U, S, \alpha'] \in \mathrm{Acc}_X$
3. For each event $U \subseteq W$ there is a unique $0 \leq \alpha \leq 1$ such that $[U, 1, \alpha]$ and $[\sim U, 1, 1\text{-}\alpha] \in \mathrm{Acc}_X$.

- The first condition on acceptance sets implies that the acceptability of a bet should not depend on the stake, but only on the event and the betting quotient (or odds).
- If a bet is accepted, then the second condition requires that bets on the same event and with the same stake, but with a more favourable betting quotient should be accepted.
- The third condition requires that for each event there is a unique breaking point, that is, a betting quotient at which one is indifferent as to which side of the bet one takes (either betting on the event, or against the event at the reverse odds).

- The unique α mentioned above is called the X's degree of belief in U, written $\mu_X(U)$
- An acceptance set is completely determined by its belief function $\mu_X(U)$
- The conditions on acceptance sets do not rule out the possibility that its belief function violates some or all of the Kolmogorov axioms for probabilities.

Definition "coherence"

A acceptance set Acc_X and its belief function μ_X are called coherent iff Acc_X does not contain a Dutch book with regard to the algebra of events under discussion.

- Big question: What conditions are satisfied by coherent belief functions?

Example 3

Assume that U and V are events such that $U \cap V = \varnothing$. Further assume that

(i) $\mu_X(U) = 0.3$,
(ii) $\mu_X(V) = 0.2$,
(iii) $\mu_X(U \cup V) = 0.6$.

Now assume that Acc_X is determined by a belief function that satisfies the conditions (i)-(iii). Then the following is a Dutch book contained in Acc_X:

$\mathbf{B} = \{[\sim U, 1, 0.7], [\sim V, 1, 0.8], [U \cup V, 1, 0.6]\}$. To prove this consider the following three possibilities for a *B*-specific event *E*:

1. $E \subseteq \sim U \cap V$, then $\|\mathbf{B}\|_E = (1\text{-}0.7) - 0.8 + (1\text{-}0.6) = -0.1$
2. $E \subseteq U \cap \sim V$, then $\|\mathbf{B}\|_E = -0.7 + (1\text{-}0.8) + (1\text{-}0.6) = -0.1$
3. $E \subseteq \sim U \cap \sim V$, then $\|\mathbf{B}\|_E = (1\text{-}0.7) + (1\text{-}0.8) - 0.6 = -0.1$

Example 3 illustrates an instance of the Dutch book theorem

Theorem

For any algebra of events: Acc_X is coherent iff $\mu_X(U)$ is a probability function (satisfying the Kolmogorov axioms)

Remark:
There is a large body of literature on the Dutch book argument. (e.g. see Frans Voorbraak http://staff.science.uva.nl/~fransv/):

"Although it has been criticised by many authors, it remains a strong, intuitive appealing, argument for using probabilities as degrees of belief of an ideally rational agent. The argument is not airtight, but it is reasonable to demand that any proposal for an alternative theory should be accompanied by an explanation why the Dutch book argument does not disqualify the proposed theory."