On individual neutrality and collective decision making

Mu Zhu\textsuperscript{1,*}, Shangsi Wang\textsuperscript{1}, Lu Xin\textsuperscript{1}

\textsuperscript{1} University of Waterloo, 200 University Ave W, Waterloo, ON, Canada N2L 3G1.

* Corresponding author; email: m3zhu@uwaterloo.ca

May 10, 2014

Abstract

We derive a simple mathematical “theory” to show that two decision-making entities can work better together only if at least one of them is occasionally willing to stay neutral. This provides a mathematical “justification” for an age-old cliché among marriage counselors.

Key words: average; Bayes theorem; probability; ROC curve; synergy; utility.

AMS classification: 62C05.

1 Introduction

Suppose that, for each decision one has to make, there are two possible courses of action, a good one ($\mathcal{G}$) and a bad one ($\mathcal{B}$). Let $h$ represent a decision-making agent, where $h(\mathcal{G}) > h(\mathcal{B})$ means $h$ favors the good action; $h(\mathcal{G}) = h(\mathcal{B})$ means $h$ is neutral; and $h(\mathcal{G}) < h(\mathcal{B})$ means $h$ favors the bad action. In addition, suppose that the payoffs associated with favoring the good action, being neutral, and favoring the bad action are $+1$, $0$, and $-1$, respectively. Let $V(h)$ denote the expected payoff for $h$. Then,

$$V(h) = (+1) \times P(h(\mathcal{G}) > h(\mathcal{B})) + 0 \times P(h(\mathcal{G}) = h(\mathcal{B})) + (-1) \times P(h(\mathcal{G}) < h(\mathcal{B})).$$

(1)
Clearly, $V(h)$ is a measure of the tendency for $h$ to favor the good action: $V(h) > 0$ means $h$ is more likely to take the good action; $V(h) < 0$ means it is more likely to take the bad one; and $V(h) = 0$ means it essentially acts randomly.

Let $h_1$ and $h_2$ be two decision-making agents, e.g., a husband and wife. In such settings, it is natural to require (e.g., for concerns of fairness) that neither $h_1$ nor $h_2$ be allowed to dominate the collective decision. This means their respective votes must be on the same scale and equally weighted. Without loss of generality, we assume that both $h_1$ and $h_2$ are binary functions, with $h_i(x) = 1$ meaning that $h_i$ voted for action $x$ and $h_i(x) = 0$ meaning that $h_i$ voted against action $x$. The collective decision is represented by their mean,

$$\bar{h} = \frac{h_1 + h_2}{2}.$$  

**Definition 1** The collective decision $\bar{h}$ is said to be synergistic if

$$V(\bar{h}) \geq \frac{V(h_1) + V(h_2)}{2};$$

it is said to have positive synergy if the above inequality is strict.

In what follows, we derive “if and only if” conditions for the collective decision $\bar{h}$ to be synergistic and examine the implications of those conditions. Before we proceed, however, it is first necessary for us to explain why our definition of synergy above should depend on the specific payoff structure $(+1, 0, -1)$ that gave rise to the expression $V(h)$ in equation (1).

### 1.1 Connection to existing literature: ROC curves

Our choice of $V(h)$ is deeply related to the so-called receiver-operating characteristic (ROC) curve (Pepe 2003), a concept widely used in many scientific fields such as signal detection and medical diagnostic tests.

In the context of medical diagnostic tests, let $h(x)$ be a measurement taken on the patient $x$. Suppose the test declares $x$ to be “sick” if $h(x) < c$ for some threshold $c$, and “healthy” if $h(x) \geq c$. The test is not perfect. For a given $c$, some healthy patients are declared “sick” (false positives), while some sick patients are declared “healthy” (false negatives).

It is easy to see that decreasing the threshold $c$ will always reduce the false positive rate but increase the false negative rate of such a test, while increasing the threshold $c$ will
always reduce its false negative rate but increase its false positive rate. For example, in
the extreme case of \( c = -\infty \), nobody is declared “sick” so there can be no false positives,
but the false negative rate reaches 100% since everyone is declared “healthy” including all
truly sick individuals. The situation is similar but reversed in the other extreme case of
\( c = +\infty \). There is an inherent trade-off between the two types of errors — reducing one
always increases the other, and the optimal choice of \( c \) depends on their relative costs.

The ROC curve essentially traces the two types of errors over the entire range of \( c \), and
the area under the ROC curve, or simply “area under the curve” (AUC), is a commonly used
performance metric for evaluating these diagnostic tests. The AUC has the advantage of not
depending on the relative costs of the two types of errors. In addition, it has the following
interesting probabilistic interpretation \([\text{Hanley and McNeil 1982; Pepe 2003}]\): Suppose \( \mathcal{H} \)
denotes the set of all healthy patients, and \( \mathcal{S} \) denotes the set of all sick patients. Then,

\[
\text{AUC}(h) \equiv 1 \times P(h(\mathcal{G}) > h(\mathcal{B})) + \frac{1}{2} \times P(h(\mathcal{G}) = h(\mathcal{B})) + 0 \times P(h(\mathcal{G}) < h(\mathcal{B}))
\]

for any randomly chosen \( \mathcal{G} \in \mathcal{H} \) and \( \mathcal{B} \in \mathcal{S} \). If \( P(h(\mathcal{G}) = h(\mathcal{B})) = 0 \), meaning \( h \) does not
produce ties between healthy and sick patients, this is simply the “folklore” statement that
AUC\( (h) \) is the probability that the test \( h \) correctly orders healthy patients ahead of sick
ones. The factor “1/2” means ties are broken at random.

It is easy to see from (1) and (3) that

\[
V(h) = 2\text{AUC}(h) - 1.
\]

In other words, our expected payoff \( V(h) \), as defined in (1), is merely a linearly transformed
version of the widely-used AUC for evaluating the effectiveness of medical diagnostic tests.
One can think of \( V(h) \) as measuring the ability of \( h \) to “diagnose” the good action when
faced with a decision.

2 Results

We now state two lemmas for respectively the independent and dependent cases. Together,
they establish necessary and sufficient conditions for the collective decision \( \bar{h} \) to be syner-
gistic. Proofs of the two lemmas are given in the appendices. However, the punchline of the
paper, which we state in Section 3, is a consequence of these two lemmas, rather than the two lemmas themselves.

2.1 Independent case

First, suppose \( h_1 \) and \( h_2 \) act independently. For \( i = 1, 2 \), let

\[
\begin{align*}
    a_i &= P(h_i(\mathcal{G}) > h_i(\mathcal{B})), \\
    b_i &= P(h_i(\mathcal{G}) = h_i(\mathcal{B})), \\
    c_i &= P(h_i(\mathcal{G}) < h_i(\mathcal{B})).
\end{align*}
\]

In other words, \( a_i \) is the probability that \( h_i \) favors the good action; \( b_i \) is the probability that \( h_i \) is neutral; and \( c_i \) is the probability that \( h_i \) favors the bad action.

**Lemma 1** Suppose \( h_1 \) and \( h_2 \) are statistically independent. Then, \( \bar{h} \) is synergistic if and only if

\[
b_1(a_2 - c_2) + b_2(a_1 - c_1) \geq 0.
\]  

(4)

**Corollary 1** For \( \bar{h} \) to be synergistic, it is sufficient (but not necessary) for both \( a_1 \geq c_1 \) and \( a_2 \geq c_2 \).

Since \( V(h_1) = a_1 - c_1 \) by (1) and likewise for \( h_2 \), the conclusion here is simply this: the collective decision is synergistic as long as both decision makers are “no worse than random”, having a slightly higher chance of favoring the good action rather than the bad one.

2.1.1 Connection to existing literature: Weak learners

The notion of being “no worse than random” is analogous to that of a “weak learner” (Schapire 1990), an important concept in PAC learning theory (Valiant 1984) and the theory of boosting (Freund and Schapire 1996).

2.2 Dependent case

More generally, suppose that \( h_1 \) and \( h_2 \) are not independent. For clarity, we make a small change in the notation to describe this case, using lowercase letters \( a, b, c \) for the marginal
probabilities of \( h_1 \), and uppercase letters \( A, B, C \) for those of \( h_2 \):

\[
a = P(h_1(G) > h_1(B)), \quad A = P(h_2(G) > h_2(B));
\]
\[
b = P(h_1(G) = h_1(B)), \quad B = P(h_2(G) = h_2(B));
\]
\[
c = P(h_1(G) < h_1(B)), \quad C = P(h_2(G) < h_2(B)).
\]

Furthermore, since \( h_1 \) and \( h_2 \) are dependent, we use \( P_{xX} \) to refer to the conditional probabilities of \( h_2 \) given \( h_1 \), and \( Q_{Yy} \) to refer to the conditional probabilities of \( h_1 \) given \( h_2 \), as displayed in Table 1. These notations for \( P \) and \( Q \) are analogous to those typically used to denote transition probabilities for Markov chains (Ross 1997).

Of course, \( P \) and \( Q \) are related by Bayes theorem (Bayes 1763). For example,

\[
P(h_2(G) = h_2(B)|h_1(G) < h_1(B)) = \ldots
\]
\[
= \frac{P(h_1(G) < h_1(B)|h_2(G) = h_2(B)) \times P(h_2(G) = h_2(B))}{P(h_1(G) < h_1(B))} \tag{5}
\]
or, using the compact notations of Table 1

\[
P_{cB} = \frac{Q_{Bc} \times B}{c}, \quad \text{or} \quad cP_{cB} = BQ_{Bc}.
\]

This leads to the following proposition, which is needed in order to prove Lemma 2 that follows.

**Proposition 1** For \( x = a, b, c \) and \( Y = A, B, C \), \( xP_{xY} = YQ_{Yx} \).

**Lemma 2** Suppose \( h_1 \) and \( h_2 \) are statistically dependent, with conditional probabilities given by Table 1. Then, \( h \) is synergistic if and only if

\[
b (P_{bA} - P_{bC}) + B (Q_{Ba} - Q_{Bc}) \geq 0. \tag{6}
\]

**Corollary 2** For \( h \) to be synergistic, it is sufficient (but not necessary) for both \( P_{bA} \geq P_{bC} \) and \( Q_{Ba} \geq Q_{Bc} \).

Notice the similarity and symmetry of the two “if and only if” conditions, (4) and (6). They can both be interpreted as follows: “Whenever one decision maker is neutral, the other one is more likely to favor the good action rather than the bad one.” Clearly, this is a highly intuitive characterization of the notion of synergy.
Table 1: Conditional probabilities. (I) $P_{xX}$ denotes various conditional probabilities of $h_2$ given $h_1$. (II) $Q_{Yy}$ denotes various conditional probabilities of $h_1$ given $h_2$. For example, $P_{aB} = P(h_2(G) = h_2(B)|h_1(G) > h_1(B))$, $Q_{Bc} = P(h_1(G) < h_1(B)|h_2(G) = h_2(B))$, etc.

(II)

|       | $h_1(G) > h_1(B)$ | $h_1(G) = h_1(B)$ | $h_1(G) < h_1(B)$ |
|-------|------------------|------------------|------------------|
| $h_2(G) > h_2(B)$ | $P_{aA}$ | $P_{aB}$ | $P_{aC}$ |
| $h_2(G) = h_2(B)$ | $P_{bA}$ | $P_{bB}$ | $P_{bC}$ |
| $h_2(G) < h_2(B)$ | $P_{cA}$ | $P_{cB}$ | $P_{cC}$ |

3 Conclusion

It is easy to see that the inequality (4) will become an equality when $b_1 = b_2 = 0$. The same can be said for inequality (6) when $b = B = 0$. These cases correspond to the situation where neither $h_1$ nor $h_2$ is ever neutral about a decision.

Definition 2 A decision-making agent $h$ is said to be opinion-loaded if its probability of being neutral is zero, i.e., if $P(h(G) = h(B)) = 0$.

Theorem 1 There can be no positive synergy in the collective decision $\bar{h}$ if both decision-making agents are opinion-loaded.

Our results, therefore, imply that there can be no positive synergy in the collective decision without individual neutrality. In other words, the willingness to compromise is not an option; it is a necessity!

Intuitively, this is because opinion-loaded decision-makers are never willing to admit that they may sometimes have a hard time making a good decision. As a result, they don’t give the other decision-maker a chance to take over the decision when it can be beneficial to do so; this explains why there can be no positive synergy. In order for there to be positive


synergy at all, it is necessary for at least one agent to have a strictly positive probability of remaining neutral.

For centuries, marriage counselors must have been giving such advice to couples all over the world, but they may not be aware of the mathematical justification for their age-old practice.

A Proof of Lemma 1

Since both \( h_1 \) and \( h_2 \) are binary functions, \( \tilde{h}(G) > \tilde{h}(B) \) if and only if

(i) \( h_1(G) > h_1(B) \) and \( h_2(G) > h_2(B) \); or

(ii) \( h_1(G) > h_1(B) \) and \( h_2(G) = h_2(B) \); or

(iii) \( h_1(G) = h_1(B) \) and \( h_2(G) > h_2(B) \).

That \( h_1 \) and \( h_2 \) are statistically independent means

\[
P(\tilde{h}(G) > \tilde{h}(B)) = a_1 a_2 + a_1 b_2 + a_2 b_1.
\]

Likewise,

\[
P(\tilde{h}(G) < \tilde{h}(B)) = c_1 c_2 + c_1 b_2 + c_2 b_1.
\]

Using the definition (1) and (7)-(8), we get

\[
V(\tilde{h}) = \frac{V(h_1) + V(h_2)}{2} \nonumber
\]

\[
= \left[\left(\frac{a_1 a_2 + a_1 b_2 + a_2 b_1}{2}\right) - (c_1 c_2 + c_1 b_2 + c_2 b_1)\right] - \frac{(a_1 - c_1) + (a_2 - c_2)}{2} \nonumber
\]

\[
= a_1 a_2 - c_1 c_2 + \left(\frac{b_2 - 1}{2}\right)(a_1 - c_1) + \left(\frac{b_1 - 1}{2}\right)(a_2 - c_2). \tag{9}
\]

Since \( a_i + b_i + c_i = 1 \) for \( i = 1, 2 \), (9) becomes

\[
a_1 a_2 - c_1 c_2 + \left(\frac{1}{2} - a_2 - c_2\right)(a_1 - c_1) + \left(\frac{1}{2} - a_1 - c_1\right)(a_2 - c_2) \nonumber
\]

\[
= \left(a_1 a_2 - c_1 c_2\right) + \frac{a_2 - c_2}{2} - (a_2 + c_2)(a_1 - c_1) + \frac{a_1 - c_1}{2} - (a_1 + c_1)(a_2 - c_2) \nonumber
\]

\[
= \frac{a_2 - c_2}{2} + \frac{a_1 - c_1}{2} - (a_1 a_2 - c_1 c_2) \nonumber
\]

\[
= \frac{1}{2} \left[(a_2 - c_2) + (a_1 - c_1) - 2(a_1 a_2 - c_1 c_2)\right], \tag{*}
\]
where

\[
(*) = \left[ (a_2 - c_2) - (a_1 a_2 - c_1 c_2) \right] + \left[ (a_1 - c_1) - (a_1 a_2 - c_1 c_2) \right] \\
= \left[ (1 - a_1) a_2 - (1 - c_1) c_2 \right] + \left[ (1 - a_2) a_1 - (1 - c_2) c_1 \right].
\] (10)

Finally, using the fact that \(a_i + b_i + c_i = 1\) for \(i = 1, 2\) again, we see that (10) is equal to

\[
\begin{align*}
[(b_1 + c_1) a_2 - (a_1 + b_1) c_2] &+ [(b_2 + c_2) a_1 - (a_2 + b_2) c_1] \\
&= b_1 a_2 - b_1 c_2 + b_2 a_1 - b_2 c_1 \\
&= b_1 (a_2 - c_2) + b_2 (a_1 - c_1),
\end{align*}
\]

which proves the lemma. □

**B Proof of Lemma 2**

By the same argument used to start the proof of Lemma 1

\[
V(h) - \frac{V(h_1) + V(h_2)}{2} \\
= \left[ (aP_A + aP_B + bP_A) - (cP_C + cP_B + bP_C) \right] - \frac{(a - c) + (A - C)}{2}.
\] (11)

But

\[
A = aP_A + bP_A + cP_C \quad \text{and} \quad C = aP_C + bP_C + cP_C,
\]

so (11) is equal to

\[
(aP_A + aP_B + bP_A) - (cP_C + cP_B + bP_C) - ... \\
\ldots - \frac{(a - c)}{2} - \frac{(aP_A + bP_A + cP_C) - (aP_C + bP_C + cP_C)}{2}.
\]

Gathering terms multiplying \(a, b,\) and \(c,\) respectively, this becomes

\[
a \left( P_A + P_B - \frac{1 + P_A - P_A}{2} \right) + b \left( P_B - P_C - \frac{P_B - P_C}{2} \right) - ... \\
\ldots - c \left( P_C + P_C - \frac{1 - P_C + P_C}{2} \right).
\] (12)

Using the fact that

\[
P_A + P_B + P_A = 1 \quad \text{and} \quad P_C + P_C + P_C = 1,
\]
expression (12) can be simplified to
\[ a \frac{P_{aB}}{2} + b \left( \frac{P_{bA} - P_{bC}}{2} \right) - c \frac{P_{cB}}{2}. \] (13)

However, by Proposition 1 we have
\[ a P_{aB} = B Q_{Ba} \quad \text{and} \quad c P_{cB} = B Q_{Bc}. \]

Substituting this into (13) leads to the conclusion that (11) ≥ 0 if and only if
\[ b (P_{bA} - P_{bC}) + B (Q_{Ba} - Q_{Bc}) \geq 0. \quad \Box \]

Acknowledgments

This work is partially supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. Philosophical Transactions of the Royal Society, 53, 370–418.

Freund, Y. and Schapire, R. (1996). Experiments with a new boosting algorithm. In Proceedings of the 13th International Conference on Machine Learning, pages 148–156. Morgan Kaufmann, San Francisco, CA, USA.

Hanley, J. A. and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, 143, 29–36.

Pepe, M. S. (2003). The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press.

Ross, S. M. (1997). Introduction to Probability Models. Academic Press, 6th edition.

Schapire, R. (1990). Strength of weak learnability. Machine Learning, 5, 197–227.

Valiant, L. (1984). A theory of the learnable. Communications of the ACM, 27(11), 1134–1142.