360 (Newcomb's paradox) There are two boxes, one transparent and one opaque. Inside the transparent box there is a visible $1,000. The player can choose to take the contents of both boxes, or just the opaque box. The amount in the opaque box was determined by a predictor, who predicted, with probability p of being correct, whether the player will take both boxes or just the opaque box. If the predictor predicted that the player will take both boxes, then the opaque box contains nothing. If the predictor predicted that the player will take just the opaque box, then the opaque box contains $1,000,000. The player knows, when making the choice, what the two possible amounts are, and that the amount was determined by the predictor with probability p of being correct. What should the player do?

After trying the question, scroll down to the solution.

§ To decrease the number of 0s to be written, the amounts of money will be in units of \$1,000 (so \$1,000 = 1 unit and \$1,000,000 = 1,000 units). Let binary variable a be the predictor's prediction: $a=0$ predicts both boxes will be chosen; $a=1$ predicts that only the opaque box will be chosen. Let binary variable b be the choice made by the player: $b=0$ says the player chose both boxes; $b=1$ says the player chose only the opaque box. Let m be the amount of money placed in the opaque box. The actions for a player who chooses both boxes, followed by the amount of money received by the player, are as follows.

$$
\begin{array}{l}
(\ (\ \ \textbf{if } 1/2 \textbf{ then } a:=0 \textbf{ else } a:=1 \textbf{ fi}. \\
\qquad \textbf{if } a \textbf{ then } m:=1000 \textbf{ else } m:=0 \textbf{ fi}. \\
\qquad b:=0 \qquad\qquad\qquad\qquad\qquad\qquad) \\
\quad \| \ \ \textbf{if } p \textbf{ then } a'=b' \textbf{ else } a'\neq b' \textbf{ fi} \qquad\qquad) . \\
m+1
\end{array}
$$

In the top line, the predictor predicts, but the player doesn't know how, so to the player it looks random. On the second line, the money in the opaque box is determined. On the next line, the player chooses both boxes. Those first three lines are what happens, and at the same time, we know that with probability p the prediction is correct, and that's the next line. The parallel lines $\|$ are defined as

$$A \| B \ = \ A \times B \ / \ (\Sigma a, b, m \cdot \ A \times B)$$

We know both A and B, so they are multiplied, but that may not result in a distribution, so we divide by the sum to make it a distribution. The final line is the money received by the player. Now we calculate. A quick use of some laws yields

$$= \quad \textbf{if } p \textbf{ then } 0+1 \textbf{ else } 1000+1 \textbf{ fi}$$

A slower calculation, translating each line, yields

$$
\begin{array}{ll}
= & (\ (\ \ (a'=0)\times(b'=b)\times(m'=m)/2 + (a'=1)\times(b'=b)\times(m'=m)/2. \\
& \qquad a\times(a'=a)\times(b'=b)\times(m'=1000) + (1-a)\times(a'=a)\times(b'=b)\times(m'=0). \\
& \qquad (a'=a)\times(b'=0)\times(m'=m) \qquad\qquad\qquad\qquad\qquad) \\
& \quad \| \ \ p\times(a'=b') + (1-p)\times(a'\neq b') \qquad\qquad\qquad\qquad\qquad) . \\
& m+1
\end{array}
$$

$$
\begin{array}{ll}
= & (\ (\qquad (a'=0)\times(b'=0)\times(m'=0)/2 \\
& \qquad + (a'=1)\times(b'=0)\times(m'=1000)/2 \ \ ) \\
& \quad \| \ \ p\times(a'=b') + (1-p)\times(a'\neq b') \qquad) . \\
& m+1
\end{array}
$$

$$
\begin{array}{ll}
= & (\quad p\times(a'=0)\times(b'=0)\times(m'=0) \\
& + \ \ (1-p)\times(a'=1)\times(b'=0)\times(m'=1000) \ \ ). \\
& m+1
\end{array}
$$

$$= \quad p + (1-p)\times 1001$$

$$= \quad 1001 - 1000\times p$$

That is the average number of units of money received by the player, if the player chooses both boxes. Now let's look at the other choice. The actions for a player who chooses only the opaque box, followed by the amount of money received by the player, are as follows.

$$\Sigma m \cdot \left(\left(\begin{array}{l} \textbf{if } 1/2 \textbf{ then } a:=0 \textbf{ else } a:=1 \textbf{ fi}. \\ \textbf{if } a \textbf{ then } m:=1000 \textbf{ else } m:=0 \textbf{ fi}. \\ b:=1 \end{array} \right) \;\|\; \textbf{if } p \textbf{ then } a'=b' \textbf{ else } a'\neq b' \textbf{ fi} \right).$$

By the quick route,

$$= \quad \textbf{if } p \textbf{ then } 1000 \textbf{ else } 0 \textbf{ fi}$$

By the slow route,

$$= \quad \Sigma m \cdot \left(\left(\begin{array}{l} (a'=0)\times(b'=b)\times(m'=m)/2 + (a'=1)\times(b'=b)\times(m'=m)/2. \\ a\times(a'=a)\times(b'=b)\times(m'=1000) + (1-a)\times(a'=a)\times(b'=b)\times(m'=0). \\ (a'=a)\times(b'=1)\times(m'=m) \end{array} \right) \;\|\; p\times(a'=b') + (1-p)\times(a'\neq b') \right).$$

$$= \quad \Sigma m \cdot \left(\left(\begin{array}{l} \;\;(a'=0)\times(b'=1)\times(m'=0)/2 \\ + (a'=1)\times(b'=1)\times(m'=1000)/2 \end{array} \right) \;\|\; p\times(a'=b') + (1-p)\times(a'\neq b') \right).$$

$$= \quad \Sigma m \cdot \left(\begin{array}{l} \;\;p\times(a'=1)\times(b'=1)\times(m'=1000) \\ + \; (1-p)\times(a'=0)\times(b'=1)\times(m'=0) \end{array} \right).$$

$$= \quad 1000\times p$$

That is the average number of units of money received by the player, if the player chooses only the opaque box.

Choosing both boxes beats choosing only the opaque box when

$$1001 - 1000\times p > 1000\times p$$
$$= \quad p < 1001/2000$$
$$= \quad p < 0.5005$$

and the average amount received by the player is

$$1001 - 1000\times p \text{ units}$$
$$= \quad 1001000 - 1000000\times p \text{ dollars}$$

Choosing only the the opaque box beats choosing both boxes when

$$p > 0.5005$$

and the average amount received by the player is

$$1000\times p \text{ units}$$
$$= \quad 1000000\times p \text{ dollars}$$

When $p = 0.5005$ it doesn't matter which the player chooses, and the average amount received by the player is

$$500.5 \text{ units}$$
$$= \quad \$500500$$

If the predictor definitely predicts correctly, $p=1$, the player should choose only the opaque box, and the average amount received is 1,000 units or $1,000,000 .