Quantum computation and the CHSH game

Mathcamp 2024

Glenn Sun and Jennifer Gao

These notes may contain errors,
and should not considered an authoritative source.

Contents

1 The classical world **2**

2 The linear algebra world **3**

3 The quantum world **5**

4 What changed? **6**

1 The classical world

Consider the following game, in which Alice and Bob cannot communicate:

1. Scooter picks $x, y \in \{0, 1\}$ uniformly at random.
2. Scooter gives Alice x and gives Bob y.
3. Alice replies with $a \in \{0, 1\}$, and Bob replies with $b \in \{0, 1\}$.
4. Alice and Bob win if $x \wedge y = a \oplus b$. (That is, x AND $y = a$ XOR b. Alternatively, when $x = y = 1$, they win by outputting different bits, otherwise they win by outputting the same bit.)

A simple strategy that wins with 75% probability is for Alice and Bob to always output 1.

Theorem 1.1. *In the classical world, Alice and Bob cannot do better than 75%.* ⌟

Proof. Note that if Alice and Bob use a deterministic strategy, then the only way they do better than 75% is to get 100%. A deterministic strategy is a function $A : \{0, 1\} \to \{0, 1\}$ for Alice and another function $B : \{0, 1\} \to \{0, 1\}$ for Bob. You can either list out every combination to see that none get 100%, or cleverly observe that always winning means

$$\begin{aligned} A(0) \oplus B(0) &= 0 \\ A(0) \oplus B(1) &= 0 \\ A(1) \oplus B(0) &= 0 \\ A(1) \oplus B(1) &= 1, \end{aligned}$$

but XOR of the LHS is 0, while the XOR of the RHS is 1.

If Alice and Bob decide to use randomness as well, they cannot do better than deterministic. Think of randomness as flipping coins. If flipping coins allowed the overall strategy to succeed more than 75% of the time, then there exists at least one sequence of coin flips that leads to success more than 75% of the time. But if they just agreed to always use those coin flip results from the start, that would be a determinstic algorithm that succeeds more than 75% of the time. □

In general, for the purposes of analogy, a classical system with randomness can be thought of as follows:

1. **State:** A vector $\vec{x}$ representing probabilities of each configuration, with $\sum_i x_i = 1$ and all $x_i \geq 0$.
2. **Actions:** A *stochastic* matrix A, meaning that whenever $\vec{x}$ is a state, then $A\vec{x}$ is also a state. This is equivalent to every column being a probability vector.
3. **Observation:** Sample a configuration of $\vec{x}$ according to its probabilities.

Example 1.2. Take the system consisting of 2 coins. There are four possible configurations: HH, HT, TH, and TT, so the state is a vector of length 4. A starting configuration could be $\vec{x} = (1, 0, 0, 0)$, meaning that both coins are heads for certain.

A possible action would be flipping the second coin. This is described by the matrix:

$$A = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}.$$

In particular, this expresses how HH and HT both become uniformly random between HH and HT, and similarly for TH and TT.

If we applied A to $\vec{x}$ defined above, we would get $(0.5, 0.5, 0, 0)$. After the action, we don't know whether the system is HH or HT until we look at it ("superposition"). However, when we do look at the system, we find that it is HH with probability 0.5 and HT with probability 0.5. ⌟

2 The linear algebra world

Sticking with the classical world in mind for now, one of the basic things that we need to do is describe what happens when you consider two independent systems as one bigger system. For example, we might want to take two coins which can exist separately and independently, but consider them as one system.

Definition 2.1. The *tensor product* of $\vec{x} = (x_1, \dots, x_n)$ and $\vec{y} = (y_1, \dots, y_m)$ is the vector of length mn,

$$\vec{x} \otimes \vec{y} = (x_1\vec{y}, \dots, x_n\vec{y}) = (x_1y_1, x_1y_2, \dots, x_ny_{m-1}, x_ny_m).$$ ⌟

Example 2.2. Consider a system with 2 coins in state HT (i.e. $(0, 1, 0, 0)$), along with a system with 1 coin which is H or T with equal probability (i.e. $(0.5, 0.5)$). Together, the system is 3 coins which is HTH and HTT with equal probability. If we enumerated the states with the same pattern as before, as

HHH, HHT, HTH, HTT, THH, THT, TTH, TTT,

then the system together is in state $(0, 0, 0.5, 0.5, 0, 0, 0, 0)$.

This is exactly the tensor product:

$$(0, 1, 0, 0) \otimes (0.5, 0.5) = (0, 0, 0.5, 0.5, 0, 0, 0, 0).$$ ⌟

Definition 2.3. A classical state is *entangled* if it cannot be written as the tensor product of 2-configuration states. ⌟

Example 2.4. Take 2 coins that are both heads-up. This is not entangled because we can write

$$(1, 0) \otimes (1, 0) = (1, 0, 0, 0).$$

Now tape them together, put them in a box, and shake the box. The state is HH and TT with equal probability, i.e. $(0.5, 0, 0, 0.5)$. This is an entangled state! To prove it, suppose there exist a, b, c, d such that

$$(a, b) \otimes (c, d) = (ac, ad, bc, bd) = (0.5, 0, 0, 0.5).$$

But $ad = 0$ implies $a = 0$, which contradicts $ac = 0.5$, or $d = 0$, which contradicts $bd = 0.5$.

To use this to debunk a "quantum paradox", simply remove the tape from the coins, and give one to Alice and the other to Bob, still hidden from view. Now, when one person looks at their coin, they immediately know their other person's coin, no matter how far away they are. Faster-than-light communication! (/s) ⌟

If all of this sounds familiar, that might be because in the classical world, these definitions are exactly equal to some things in probability theory—marginal and joint distributions and independence. We use different words because physics like using different words, and they are a little different in the quantum world.

Now that we've talked about how to combine independent states into one state, the last thing in the classical world I want to talk about is how to combine independent actions into one action.

Definition 2.5. The tensor product of matrices is the block matrix given by

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix}.$$

It is the unique matrix satisfying

$$(A \otimes B)(\vec{x} \otimes \vec{y}) = (A\vec{x}) \otimes (B\vec{y}).$$

(This can be easily proven by looking at basis vectors.) ⌟

The equation at the end there is really the way to think about tensor product. If you want to act on $\vec{x}$ with A and act on $\vec{y}$ with B, it is entirely equivalent to acting on the joint system $\vec{x} \otimes \vec{y}$ with $A \otimes B$.

Example 2.6. Suppose you created the state $(0.5, 0, 0, 0.5)$ and separated two entangled coins. You can still perform actions on the entangled coins individually, such as tossing them, or turning them upside-down. What is the matrix that then expresses turning the second coin upside-down?

By the above, we get

$$A = I \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

If you do a sanity check and multiply this with basis vectors, this is indeed what you need to turn the second coin upside down.

Importantly, when we want to express an action, the matrix doesn't care about whether the state it acts on is entangled or not. If it's not entangled, then applying separate actions to each piece won't entangle it. Meanwhile, being entangled does not preclude us from applying separate actions to each piece, the result might just remain entangled. ⌟

3 The quantum world

There is just one main difference* in the quantum world. If you think about a classical state as a vector with unit L_1 norm (a.k.a. taxicab, Manhattan, etc.), then you know that the L_2 norm (a.k.a. standard distance) is often more mathematically interesting thing to study. So the main difference is that quantum states have unit L_2 norm.

1. **State:** A vector $\vec{x}$ with unit norm, i.e. $\sum_i |x_i|^2 = 1$.
2. **Actions:** A *unitary* matrix A, meaning that whenever $\vec{x}$ has unit norm, then $A\vec{x}$ also has unit norm.
3. **Observation:** Sample a configuration of $\vec{x}$, where configuration i is selected with probability $|x_i|^2$.

Everything else about the quantum world is pretty much the same as the classical world. We can put together independent systems using tensor product in the same way, since $|x_i|^2|y_j|^2 = |x_i y_j|^2$ means the probability justification we gave continues to hold.

In quantum mechanics, there are several equivalents to "coins" that can form the basic combinatorial object of study. One option is the electron, where heads/tails is the direction of a quantity called electron spin, which can be up/down upon observation. A system with n electrons will be represented as a vector of length 2^n, just like coins.

Theorem 3.1. *There is a quantum strategy where Alice and Bob win approximately 85% of the time.* ⌟

Proof. Define the rotation matrix to be:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

which rotates by θ counterclockwise. Rotations are rigid transformations, and are hence clearly unitary.

The strategy is as follows:

1. Alice and Bob construct a two-electron system with $(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}})$ before the game, then take their individual electron with them.
2. Alice does nothing if she receives $x = 0$, and applies $R_{\pi/4}$ is she receives $x = 1$. Alice then looks at her electron and replies with what she sees.
3. Bob applies $R_{\pi/8}$ if he receives $y = 0$, and applies $R_{-\pi/8}$ if he receives $y = 1$. Bob then looks at his electron and replies with what he sees.

The intuition is that their starting state is similar to the taped coins, so they know that they're starting with the same thing. So in the beginning,

*The second difference is that the quantum world uses complex numbers. Once negative numbers are in the picture, the reason we need complex numbers is for continuity: matrices over the real numbers don't always have square roots, and you need square roots to do half of an action. The reason we need our actions to be given by matrices is that in 1998, Abrams and Lloyd proved that if quantum mechanics were nonlinear, then you can solve NP problems in polynomial time.

without acting on their electrons, they always win the "output same" cases but always fail the "output different" case. After these rotations, they are $\frac{\pi}{8}$ away from always winning in every case. But because of the way circles work, this is actually much better than always failing one case.

To calculate precisely, take first the case of $x = y = 0$. Then the resulting state is

$$I \otimes R_{\pi/8} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{8} & -\sin\frac{\pi}{8} & 0 & 0 \\ \sin\frac{\pi}{8} & \cos\frac{\pi}{8} & 0 & 0 \\ 0 & 0 & \cos\frac{\pi}{8} & -\sin\frac{\pi}{8} \\ 0 & 0 & \sin\frac{\pi}{8} & \cos\frac{\pi}{8} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\frac{\pi}{8} \\ \sin\frac{\pi}{8} \\ -\sin\frac{\pi}{8} \\ \cos\frac{\pi}{8} \end{bmatrix}.$$

Now, Alice and Bob look at their electrons. They are still entangled, so we need to look at the full state vector. The probability that they both output the same spin is then

$$\left(\frac{1}{\sqrt{2}}\cos\frac{\pi}{8}\right)^2 + \left(\frac{1}{\sqrt{2}}\cos\frac{\pi}{8}\right)^2 = \cos^2\frac{\pi}{8} \approx 85\%.$$

A nearly identical calculation (with maybe some more trigonometry) shows the same bound for the other cases. □

This game is known as the CHSH game, after John Clauser, Michael Horne, Abner Shimony, and Richard Holt who stated and proved this result in 1969. Then in the 1980s, a team of experimental physicists were able to try it out, and in fact succeeded in winning the game 84% of the time!

4 What changed?

Here's a question: what part of the proof of the classical lower bound fails when we are in a quantum world? It's not so obvious!

The answer is that pulling our coin flips to the start of the strategy is always valid in the classical model, but this is not always valid in the quantum model. We had to pull our coin flips to the start in order create a deterministic strategy from a randomized one with coin tosses in the middle. This is not possible in the quantum world.

More formally, actions and observations commute in the classical world in the following way. Suppose you have the following state and action:

$$\begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where the first component represents the state H and the second line represents the state T.

Let H and T denote the event that you see H and T respectively when you observe after the action. In the classical world, if you acted first and then observed, just multiply the matrix and vector to find that

$$\Pr[H] = ax + by$$
$$\Pr[T] = cx + dy.$$

Meanwhile, if you observe first before acting, letting H' and T' denote the observation events before the action, you will find that

$$\begin{aligned}\Pr[H'] &= x\\ \Pr[T'] &= y.\end{aligned}$$

Now, the state is $(1, 0)$ in the first case and $(0, 1)$ in the second case, so that after the action, multiplication tells us that

$$\begin{aligned}\Pr[H \mid H'] &= a\\ \Pr[H \mid T'] &= b\\ \Pr[T \mid H'] &= c\\ \Pr[T \mid T'] &= d.\end{aligned}$$

From this we conclude that

$$\begin{aligned}\Pr[H] &= \Pr[H \mid H']\Pr[H'] + \Pr[H \mid T']\Pr[T'] = ax + by\\ \Pr[T] &= \Pr[T \mid H']\Pr[H'] + \Pr[T \mid T']\Pr[T'] = cx + dy,\end{aligned}$$

which is the same as if we acted before observing.

Now, in the quantum case, if we act before observing, we end up with

$$\begin{aligned}\Pr[H] &= |ax + by|^2\\ \Pr[T] &= |cx + dy|^2.\end{aligned}$$

One can easily check with the same calculations as above that if we observe before we act, we end up with

$$\begin{aligned}\Pr[H] &= |a|^2|x|^2 + |b|^2|y|^2\\ \Pr[T] &= |c|^2|x|^2 + |d|^2|y|^2,\end{aligned}$$

which is different!

Physicists call this phenomenon “wavefunction collapse”. It means that unlike the classical case, just by looking at a quantum particle, you fundamentally change how it behaves. The wording here is important: the interesting part of quantum mechanics is that observation changes the effect of future actions. The interesting part is not that the superposition collapses into one of its possible values, as often quoted in pop media, since that happens classically too.