Multiple Systems – Lesson 02

Deval Deliwala

December 18, 2023

Contents

We can always choose to view multiple systems *together* as if they form a single, compound system – to which the discussion in the previous section applies.

1 Classical Information

Recognizing how mathematics works in the familiar setting of classical information is helpful in understanding why quantum information is described the way it is.

1.1 Classical States via the Cartesian Product

We will start with classical states of multiple systems. We will begin by discussing just two systems and then generalize further.

To be precise, let us suppose that X is a system whose classical state set is Σ and Y is a second system having classical state set Γ. Because these are *classical state sets*, our assumption is that Σ and Γ are both finite and nonempty. It could be that $\Sigma = \Gamma$, but this is not necessarily so.

Now imagine the two systems X and Y are placed side-by-side, with X on the left and Y on the right. We can views these two systems as if they form a single system (X, Y) or XY.

The set of classical states of XY is the *Cartesian Product* of Σ and Γ, which is the set defined as

Cartesian Product

$$\Sigma \times \Gamma = (a, b) : a \in \Sigma \text{ and } b \in \Gamma$$

The Cartesian Product is the mathematical notion of viewing an element of one set and an element of a second set together, as if they form a single set.

For more than two systems, the situation generalizes in a natural way. If we suppose that $X_1, \cdots, X_n$ are systems having classical state sets $\Sigma_1, \cdots, \Sigma_n$, respectively for any positive integer n, the classical state set of the n-tuple $X_1 X_2 \cdots X_n$, viewed as a single joint system, is the Cartesian Product

$$\Sigma_1 \times \cdots \times \Sigma_n = (a_1, \cdots, a_n) : a_1 \in \Sigma_1 \cdots a_n \in \Sigma_n$$

1.2 Representing states as strings

It is convenient to write a classical state $(a_1, \cdots, a_n)$ as a string '$a_1 \cdots a_n$'. The notion of a string, which is fundamentally important concept in CS, is formalized in mathematical terms through the cartesian product.

For example suppose that $X_1, \cdots, X_{10}$ are bits, so the classical state sets of these systems are all the same

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_{10} = 0, 1$$

The set $0, 1$ is referred to as the *binary alphabet*. There are $2^{10} = 1024$ classical states of the joint system $X_1 X_2 \cdots X_{10}$ which are elements of the set

$$\Sigma_1 \times \Sigma_2 \times \cdots \times \Sigma_{10} = 0, 1^1 0.$$

Written as strings, these classical states look like this:

$$\begin{gathered} 0000000000 \\ 0000000001 \\ 0000000010 \\ 0000000011 \\ 0000000100 \\ \vdots \\ 1111111111 \end{gathered}$$

For the classical state 0001010000 we see that X_4 and X_6 are in the state 1, while all other systems are in the state 0.

1.3 Probabilistic States

Probabilistic states associate a probability with each classical state of a system. Thus, a probabilistic state of multiple systems – viewed collectively as if they form a single system – associates a probability with each element of the Cartesian product of the classical state sets of the individual systems.

Suppose X and Y are both bits. So their corresponding classical state sets are $\Sigma = 0, 1$ and $\Gamma = 0, 1$. Here is a probabilistic state of the pair XY:

$$\begin{aligned} \Pr((X, Y) = (0, 0)) &= \frac{1}{2} \\ \Pr((X, Y) = (0, 1) &= 0 \\ \Pr((X, Y) = (1, 0) &= 0 \\ \Pr((X, Y) = (1, 1) &= \frac{1}{2} \end{aligned}$$

This probabilistic state is one in which both X and Y are random bits – each is 0 with probability 1/2 and 1 with probability 1/2 – but the classical states of the two bits always agree. This is an example of a *correlation* between these systems.

1.4 Ordering Cartesian Product State Sets

To represent a probabilistic state of multiple systems as a Cartesian product, one must decide on an ordering of the product's elements. The entries in each $n-$ tuple are viewed as being ordered by significance that *decreases from left to right*.

The Cartesian Product $1, 2, 3 \times 0, 1$ is ordered like this:

$$(1,0),(1,1),(2,0),(2,1),(3,0),(3,1)$$

The probabilistic state represented by the probability vector above

$$\begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}$$

Where from top to bottom we get the probability associated with 00, then 01, then 10, then 11.

1.5 Independence of Two Systems

Intuitively speaking, two systems are independent if learning the classical state of either system has no effect on the probabilities associated with the other.

Suppose X and Y are systems having classical state sets Σ and Γ respectively. They are said to be *independent* if it is the case that

$$\Pr((X,Y) = (a,b)) = \Pr(X = a)\Pr(Y = b)$$

for every $a \in \Sigma$ and $b \in \Gamma$.

To express this in terms of probability vectors, assume the probabilistic state of (X, Y) is described as a probability vector written in Dirac notation as

$$\sum_{(a,b)\in\Sigma\times\Gamma} p_{ab}|ab\rangle$$

The above condition is equivalent to the existence of two probability vectors

$$|\phi\rangle = \sum_{a\in\Sigma} q_a|a\rangle \quad \text{and} \quad |\psi\rangle = \sum_{b\in\Gamma} r_b|b\rangle$$

representing the probabilities associated with the classical states of X and Y such that

$$p_{ab} = q_a r_b$$

For example, the probabilistic state of a pair of bits (X, Y) represented by the vector

$$\frac{1}{6}|00\rangle + \frac{1}{12}|01\rangle + \frac{1}{2}|10\rangle + \frac{1}{4}|11\rangle$$

is one in which X and Y are independent. Specifically the condition required for independence is true for the probability vectors

$$|\phi\rangle = \frac{1}{4}|0\rangle + \frac{3}{4}|1\rangle \quad \text{and} \quad |\psi\rangle = \frac{2}{3}|0\rangle + \frac{1}{3}|1\rangle$$

For example, to match the 00 entry we need $\frac{1}{6} = \frac{1}{4} \times \frac{2}{3}$. The other entries can be verified in a similar manner.

On the other hand, the probabilistic state

$$\frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle$$

does not represent independence between X and Y.

Suppose that there did exist probability vectors $|\phi\rangle$ and ψ for which $p_a b = q_a r_b$ is satisfied for every choice a and b. It would then necessarily be that

$$q_0 r_1 = \Pr((X, Y) = (0, 1)) = 0$$

This implies that either $q_0 = 0$ or $r_1 = 0$. Which means that $q_0 r_0 = 0$ or $q_1 r_1 = 0$ which isn't the case. Hence, there do not exist vectors ϕ and $|\psi\rangle$ satisfying the property required for independence.

We can now define correlation precisely as a *lack of independence.*

1.6 Tensor Products of Vectors

The condition for independence just described can be expressed more succinctly through the notion of a *tensor product.* In concrete terms, given two vectors

$$|\phi\rangle = \sum_{a \in \Sigma} \alpha_a |a\rangle \quad \text{and} \quad |\psi\rangle = \sum_{b \in \Gamma} \beta_b |b\rangle$$

the tensor product $|\phi\rangle \otimes |\psi\rangle$ is a new vector over the joint set $\Sigma \times \Gamma$ defined as

Tensor Product

$$|\phi\rangle \otimes |\psi\rangle = \sum_{(a,b)\in\Sigma\times\Gamma} \alpha_a \beta_b |ab\rangle$$

Equivalently the vector $|\pi\rangle = |\phi\rangle \otimes |\psi\rangle$ is defined by the equation

$$\langle ab|\pi\rangle = \langle a|\phi\rangle\langle b|\psi\rangle$$

being true for every $a \in \Sigma$ and $b \in \Gamma$
In this situation it is said that $|\pi\rangle$ is a *product state* or *product vector*. We often just write $|\phi\rangle|\psi\rangle$ instead of $|\phi\rangle \otimes |\psi\rangle$. The tensor product for two column vectors is defined as follows:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 \\ \vdots \\ \alpha_1\beta_k \\ \alpha_2\beta_1 \\ \vdots \\ \alpha_2\beta_k \\ \vdots \\ \alpha_m\beta_1 \\ \vdots \\ \alpha_m\beta_k \end{pmatrix}$$

For standard basis vectors:

$$|a\rangle \otimes |b\rangle = |ab\rangle = |a, b\rangle$$

The tensor product of two vectors has the important property that it is *bilinear*, which means that it is linear in each of the two arguments separately, assuming that the other argument is fixed.

Bilinearity

1. Linearity in the first argument:

$$(|\phi_1\rangle + |\phi_2\rangle) \otimes |\psi\rangle = |\phi_1\rangle \otimes |\psi\rangle + |\phi_2\rangle \otimes |\psi\rangle$$
$$(\alpha|\phi\rangle) \otimes |\psi\rangle = \alpha(|\phi\rangle \otimes |\psi\rangle)$$

2. Linearity in the second argument:

$$|\phi\rangle \otimes (|\psi_1\rangle + |\psi_2\rangle) = |\phi\rangle \otimes |\psi_1\rangle + |\phi\rangle \otimes |\psi_2\rangle$$
$$|\phi\rangle \otimes (\alpha|\psi\rangle) = \alpha(|\phi\rangle \otimes |\psi\rangle)$$

Scalars “float freely” within tensor products.

1.7 Independence and Tensor Products for 3 or More Systems

If $X_1, \cdots, X_n$ are systems with classical state sets $\Sigma_1, \cdots, \Sigma_n$, then a probabilistic state of the combined system $X_1 X_2 \cdots X_n$ is a *product state*

$$|\psi\rangle = |\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle$$

for probability vectors $|\phi_1\rangle \cdots |\phi_n\rangle$ describing probabilistic states $X_1, \cdots, X_n$

The definition of the tensor product generalizes as the following:

$$|\psi\rangle = |\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle$$

is defined by the equation

Tensor Product Definition

$$\langle a_1 \cdots a_n|\psi\rangle = \langle a_1|\phi_1\rangle \cdots \langle a_n|\phi_n\rangle$$

for every $a_1 \in \Sigma_1, \cdots a_n \in \Sigma_n$.

Similar to the tensor product for two vectors, the tensor product for 3 or more vectors is linear in each of the arguments individually. We say the tensor product of 3 or more vectors is *multilinear*.

For multiple systems is $X_1, ...X_n$ are *independenet*, when they are in a product state, they are actually *mutually independent*. There is also other notions of independence for 3 or more systems such as *pairwise independence*.

For any positive integer n and any classical states $a_1, ..., a_n$ we have

$$|a_1\rangle \otimes \cdot \otimes |a_n\rangle = |a_1 \cdot a_n\rangle$$

1.8 Measurements of Probabilistic States

If two bits (X, Y) are described by the following probability vector

$$\frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle,$$

then the outcome measurement of 0 for X and 0 for Y is obtained with probability 1/2 and the outcome 11 is also obtained with probability 1/2.

Suppose however, that we choose not to measure *every* system, just some *proper subset* of the systems. This will result in a measurement outcome for each system that gets measured, and will also affect our knowledge of the remaining systems.

Suppose that X is a system having classical state set Σ and Y is a system having classical state set Γ, and the two systems together are in some probabilistic state. We will consider what happens when we just measure X and do nothing with Y. The situation vice versa is handled symmetrically.

We know that the probability to observe a particular classical state $a \in \Sigma$ when just X is measured must be consistent with the probabilities we would obtain under the assumption that Y was also measured. That is, we must have

$$\Pr(X = a) = \sum_{b \in \Gamma} \Pr((X, Y) = (a, b))$$

Intuitively, if this formula is wrong, it would mean that the probabilities for X measurements are influenced simply by whether or not Y is also measured, irrespective of the outcome of Y. If Y happened to be in a distant location, say, another galaxy, this would allow for faster-than-light signaling, which we obviously reject.

Given the assumption that only X is measured, there still exists uncertainty over the classical state of Y. We must update our description so this uncertainty of Y is properly reflected.

The following *conditional probability* formula reflects this uncertainty.

$$\Pr(Y = b \,|\, X = a) = \frac{\Pr((X, Y) = (a, b))}{\Pr(X = a)}$$

Here, the expression $\Pr(Y = b \,|\, X = a)$ denotes the probability that $Y = b$ *conditioned* on $X = a$.

To express these formulas in terms of probability vector, consider a probability vector $|\psi\rangle$ describing the joint state of (X, Y).

$$|\psi\rangle = \sum_{(a,b) \in \Sigma \times \Gamma} p_{ab} |ab\rangle$$

Measuring X alone yields each possible outcome with probabilities

$$\Pr(X = a) = \sum_{b \in \Gamma} p_{ab}$$

Thus the vector representing the probabilistic state of X alone is given by

$$\sum_{a\in\Sigma}\left(\sum_{c\in\Gamma}p_{ac}\right)|a\rangle$$

The updated probabilistic state of Y having obtained a particular outcome $a \in \Sigma$ of the measurement of X, is represented by the probability vector:

$$|\pi_a\rangle = \frac{\sum_{b\in\Gamma} p_{ab}|b\rangle}{\sum_{c\in\Gamma} p_{ac}}$$

If X results in a classical state a, we update our description of the probabilistic state of the joint system (X, Y) to $|a\rangle \otimes |\pi_a\rangle$.

The definition of $|\pi_a\rangle$ is the *normalization* of the vector $\sum_{b\in\Gamma} p_{ab}|b\rangle$ where we divide the sum of the entries in this vector to obtain the probability vector.

1.9 Example of Classical State Measurements

For example, suppose the classical state of X is $\Sigma = 0, 1$, the classical state of Y is $\Gamma = 1, 2, 3$, and the probabilistic state of (X, Y) is

$$|\psi\rangle = \frac{1}{2}|0,1\rangle + \frac{1}{12}|0,3\rangle + \frac{1}{12}|1,1\rangle + \frac{1}{6}|1,2\rangle + \frac{1}{6}|1,3\rangle$$

Our goal is to determine probabilities of two possible outcomes – 0 and 1, and to calculate what the resulting probabilistic state of Y is for the two outcomes, assuming X is measured.

Using the bilinearity of the tensor product, we may rewrite $|\psi\rangle$ as

$$|\psi\rangle = |0\rangle \otimes \left(\frac{1}{2}|1\rangle + \frac{1}{12}|3\rangle\right) + |1\rangle \otimes \left(\frac{1}{12}|1\rangle + \frac{1}{6}|2\rangle + \frac{1}{6}|3\rangle\right)$$

We have isolated the distinct standard basis vectors for the system being measured. This is always possible, regardless of what vector we start with.

Thus the probabilities of the two outcomes are given by

$$\Pr(X = 0) = \frac{1}{2} + \frac{1}{12} = \frac{7}{12}$$
$$\Pr(X = 1) = \frac{1}{12} + \frac{1}{6} + \frac{1}{6} = \frac{5}{12}$$

The probabilistic state of Y, conditioned on each possible outcome can also be quickly inferred by normalizing the vectors in parenthesis so that these vectors become probability vectors. That is, conditioned on X being 0, the probabilistic state of Y becomes

$$\frac{\frac{1}{2}|1\rangle + \frac{1}{12}|3\rangle}{\frac{7}{12}} = \frac{6}{7}|1\rangle + \frac{1}{7}|3\rangle$$

and conditioned on the measurement of X being 1, the probabilistic state of Y becomes

$$\frac{\frac{1}{12}|1\rangle + \frac{1}{6}|2\rangle + \frac{1}{6}|3\rangle}{\frac{5}{12}} = \frac{1}{5}|1\rangle + \frac{2}{5}|2\rangle + \frac{2}{5}|3\rangle$$

1.10 Operations on Probabilistic States

We will consider *operations* on multiple systems in probabilistic states. Returning to the typical setup where we have two systems X and Y, let us consider classical operations on the compound system (X, Y). Any operation is represented by a stochastic matrix whose rows and columns are indexed by the Cartesian Product $\Sigma \times \Gamma$.

For example, suppose X and Y are bits, and consider the operation:

> If $X = 1$, then perform a NOT operation on Y.
> Otherwise do nothing.

This is a deterministic operation known as a *controlled-NOT* or CNOT operation, where X is the *control* bit that determines whether or not a NOT operation gets applied to the *target* bit Y. Here is the matrix representation of this operation:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Its action on the standard basis states:

$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle \\ |11\rangle &\mapsto |10\rangle \end{aligned}$$

Another example is the operation having this description

Perform one of the following two operations, each with probability $\frac{1}{2}$:

- Set Y to be equal to X
- Set X to be equal to Y

The matrix representation is as follows:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

The action of this operation on the standard basis vectors:

$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto \frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle \\ |10\rangle &\mapsto \frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle \\ |11\rangle &\mapsto |11\rangle \end{aligned}$$