量子计算(二)——量子态与量子门

一、量子态与局部观测

1.量子态与计算基态

我们知道一个量子位对应的**叠加态**可以表示为两个**本征态**的叠加，即$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$；对于两个量子位，则表示为$\alpha|00\rangle + \beta|01\rangle + \theta|10\rangle + \delta|11\rangle$；扩展为$n$位时，则量子态可以表示为：

$$|\psi\rangle = \sum_{i=0}^{2^n-1} a_i|i\rangle$$

在这里，符号$|i\rangle$中的i通常会展开写为一个长度为n的二进制数，而它又代表一个元素个数为2^n、但只在第i位为1的列向量，我们将$|i\rangle$这样的态称为**计算基态**。各计算基态本身是**单位的**(即模长为1)，而各计算基态之间是**正交的**，也就是说：

$$\langle i|j\rangle = \langle i||j\rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

> 复习：左右矢互为共轭转置。当同维的左矢乘上右矢时，相当于普通的向量乘法，将会得到一个数。对于计算基态来说，自然只在$|i\rangle, |j\rangle$相等时，能够使$\langle i|j\rangle = 1$。

量子态本身需要符合**归一化条件**，也就是各计算基态出现的概率之和必须为1，这可以写成以下两种形式：

$$\sum_{i=0}^{2^n-1} |a_i|^2 = 1 \quad , \quad \text{或者}$$

$$||\psi\rangle|^2 = \langle\psi|\psi\rangle = 1$$

2.局部观测

现在假设分别有两个量子位$|\psi\rangle = \alpha_1|0\rangle + \beta_1|1\rangle, |\varphi\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$，那么这两个量子位共同组成的系统可以用张量积表示为：

$$
\begin{aligned}
|\psi\rangle|\varphi\rangle &= |\psi\rangle \otimes |\varphi\rangle \\
&= (\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) \\
&= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle
\end{aligned}
$$

那么对这个系统整体进行观测时，$00, 01, 10, 11$状态出现的概率分别是$|\alpha_1\alpha_2|^2, |\alpha_1\beta_2|^2, |\beta_1\alpha_2|^2, |\beta_1\beta_2|^2$。但我们考虑一种情况：假如只对其中一个量子位观测，那么结果又将如何？或者说，仅针对其中某一个量子位，观测到其结果为0或1的概率分别为多少？

您可能会说这很简单呀！例如只观测第一个量子位，其状态为1的概率很显然就是$10, 11$状态出现的概率之和$|\beta_1\alpha_2|^2 + |\beta_1\beta_2|^2$嘛！这确实是一种方法，但一种更系统的做法是，将量子态重写为$|0\rangle|\psi_0\rangle + |1\rangle|\psi_1\rangle$的形式，那么状态$0, 1$的概率就分别是$|\psi_0\rangle, |\psi_1\rangle$的**模的平方**了！以上述为例，将$|\psi\rangle|\varphi\rangle$重写为：

$$
|0\rangle(\alpha_1\alpha_2|0\rangle + \alpha_1\beta_2|1\rangle) + |1\rangle(\beta_1\alpha_2|0\rangle + \beta_1\beta_2|1\rangle)
$$

所以第一个量子位的状态为0的概率就是$|\alpha_1\alpha_2|0\rangle + \alpha_1\beta_2|1\rangle|^2 = |\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2$了。

这可以扩展到多个量子位的情况。假设有n位量子位，要观测第i位量子位，那么就将第i位为$0, 1$的计算基态分别提取出来并写成$|0\rangle|\psi_0\rangle + |1\rangle|\psi_1\rangle$的形式，再分别计算$|\psi_0\rangle, |\psi_1\rangle$的**模的平方**即可。

到这里还尚未结束！必须强调，当第i位被观测之后，这个量子位就已经“**永久地**”落入到确定状态了，此时的量子态就不能按原来的形式写了！因为在第一节中已经强调，量子态需要符合**归一化条件**，我们必须要继续写出观测之后的量子态。

还是以上文的二量子位系统为例，第一量子位被观测后，假如观测结果是0，那么剩余系统的叠加态就应当是$|\psi_0\rangle = \alpha_1\alpha_2|0\rangle + \alpha_1\beta_2|1\rangle$；然而我们知道$|\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2 \neq 1$，因此需要将这个态重整**归一化**，这就是说剩余系统的叠加态要写成：

$$
\frac{\alpha_1\alpha_2|0\rangle + \alpha_1\beta_2|1\rangle}{\sqrt{|\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2}} = \frac{\alpha_1\alpha_2}{\sqrt{|\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2}}|0\rangle + \frac{\alpha_1\beta_2}{\sqrt{|\alpha_1\alpha_2|^2 + |\alpha_1\beta_2|^2}}|1\rangle
$$

才是正确的。因此扩展到n位的情况，观测第i位之后剩余量子位的叠加态应是：

$$
\begin{cases}
\dfrac{1}{||\psi_0\rangle|}|\psi_0\rangle = \dfrac{1}{\sqrt{\langle\psi_0|\psi_0\rangle}}|\psi_0\rangle \quad , \quad \text{若观测结果为}0 \\
\dfrac{1}{||\psi_1\rangle|}|\psi_1\rangle = \dfrac{1}{\sqrt{\langle\psi_1|\psi_1\rangle}}|\psi_1\rangle \quad , \quad \text{若观测结果为}1
\end{cases}
$$

3.纠缠态

对于某些特殊的双量子位量子态，也就是：

$$\alpha|00\rangle+\beta|11\rangle \quad 或 \quad \alpha|01\rangle+\beta|10\rangle$$

对其进行**局部观测**时，我们发现剩下的一个量子位的量子态总会只剩下$|0\rangle$或$|1\rangle$，这相当于它也落入了一个百分百概率确定的状态，这时我们称这种量子态为**纠缠态**。纠缠态是一种特殊的叠加态。

纠缠态中最重要的四种又是**贝尔纠缠态**(也称**贝尔基**)，又称为EPR对，这在量子计算中极常用到：

$$|\Phi_{\pm}\rangle=\frac{1}{\sqrt{2}}(|00\rangle\pm|11\rangle)$$

$$|\Psi_{\pm}\rangle=\frac{1}{\sqrt{2}}(|01\rangle\pm|10\rangle)$$

> 为什么还有贝尔基的称呼方式？您可以验证这四种纠缠态相互之间也是正交的。这意味着双量子位系统不仅可以表示为计算基态的线性和，也能表示为这四个纠缠态的线性和(还记得上一篇文章中光子与偏振片的例子吗？只要分解方向之间是垂直正交的，那就是合法的)。虽然这种表示方法不常用，但在一些量子算法中，按贝尔基分解时会揭示一些奇妙的性质。

4.纯态与混合态

在以上所述的各情况中，量子系统都可以只用一个量子叠加态$|\psi\rangle$来描述——虽然系统要落入哪个本征态是不确定的，但它的叠加态却是可以确定的，可以**只用一个右矢**的态矢量来表示，这时我们称之为**纯态**。

然而还有一些量子系统连是什么叠加态都不能确定，它具有n个不同的叠加态$|\psi_1\rangle,...,|\psi_i\rangle,...,|\psi_n\rangle$，处于这些叠加态的概率分别是$p_1,..,p_i,...,p_n$，称之为**混合态**。为了描述混合态我们不能单纯地只使用态矢量来表示，例如下面这种简单的线性加权和就是不行的：

$$\sum_{i=1}^{n}p_i|\psi_i\rangle$$

因为最终各本征态出现的概率和不会为1。我们定义**密度矩阵**：

$$\rho_{mix}=\sum_{i=1}^{n}p_i|\psi_i\rangle\langle\psi_i|$$

来描述混合态。当然纯态也可以用密度矩阵来描述，它以概率1出现在量子态$|\psi\rangle$，因而其密度矩阵就是$\rho=|\psi\rangle\langle\psi|$。

> 一个量子系统所处的状态称为**量子态**，它**不是纯态就是混合态**。**叠加态和计算基态**只存在于纯态中。也就是说各种态的性质可以归纳如下：

$$\text{量子态}\begin{cases}\text{纯态}\begin{cases}\text{叠加态}\\ \text{计算基态}\end{cases}\\ \text{混合态}\end{cases}$$

当然在上一篇文章中也说过，如果您认为计算基态也是一种特殊的叠加态，这并无不妥；但我们说叠加态时，会侧重表现其**叠加**的性质。

5.密度矩阵的性质

在给定密度矩阵的时候我们可以通过某些特征确定矩阵描述的是纯态还是混合态。首先注意到**纯态的密度矩阵的任意次幂都是相等的**，因为：

$$\rho^2 = |\psi\rangle\langle\psi||\psi\rangle\langle\psi| = |\psi\rangle(\langle\psi|\psi\rangle)\langle\psi| = |\psi\rangle\langle\psi| = \rho$$

而混合态是**不等的**：

$$\rho_{mix}^2 = \sum_i \sum_j p_i p_j |\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j| \neq \rho_{mix}$$

再来关注密度矩阵及其平方的**迹**(对角线元素之和)。对于纯态来说，其密度矩阵的迹显然是1，因为$|\psi\rangle$本身是**单位的**，密度矩阵上对角线元素就是原右矢对应位置元素的模平方，代表了对应本征态(计算基态)出现的概率，其和自然就是1。因此对于纯态存在：

$$\rho = \rho^2 \quad \Longrightarrow \quad tr(\rho) = tr(\rho^2) = 1$$

为了更加数学化地说明，首先展开纯态的密度矩阵$\rho = |\psi\rangle\langle\psi|$。而纯态的右矢可以进一步展开为$\sum a_i|i\rangle$，于是将纯态密度矩阵继续展开：

$$\begin{aligned}|\psi\rangle\langle\psi| &= \sum a_i|i\rangle \cdot \sum \bar{a}_j\langle i| \\ &= \sum_i \sum_j a_i\bar{a}_j|i\rangle\langle j|\end{aligned}$$

上式虽然不是一个对角阵，但其对角线上的元素是$\{|a_1|^2, |a_2|^2, ..., |a_n|^2\}$，这其中的物理含义就已非常显然了——各列上的对角线元素，就是在测量后这一列对应计算基态出现的概率。根据归一化条件，结论$tr(\rho) = tr(\rho^2) = 1$不言自明。

对于混合态密度矩阵，它的迹$tr(\rho_{mix})$也是1。这有两种理解方法：其一，因为$|\psi_i\rangle\langle\psi_i|$上的对角线元素代表了这个量子态下对应本征态(计算基态)出现的概率，其和为1，但之后加入到密度矩阵时还需要以

p_i加权；而$\sum p_i = 1$，因此迹就是各量子态出现的概率之和，即为1。其二，混合态密度矩阵上的对角线元素也代表了其对应**本征态出现的概率**，是各p_i的线性组合，和也必然为1。

> 现在再来看混合态密度矩阵$\rho_{mix} = \sum p_i|\psi_i\rangle\langle\psi_i|$。按照上述，每一个$|\psi_i\rangle\langle\psi_i|$上的对角线元素都是对应计算基态出现的概率，从而$\rho_{mix}$中的对角线元素也可以轻松写出：
>
> $$\left\{p_1\sum|a_j|^2, p_2\sum|a_j|^2, ..., p_n\sum|a_j|^2\right\}$$
>
> 从而对于混合态密度矩阵来说，各列上的对角线元素依然是在测量后对应计算基态出现的概率！因而根据归一化条件，混合态密度矩阵的迹也是1。结合稍早之前的讨论，我们获得结论：任意密度矩阵的迹都是1。

但是**混合态密度矩阵的平方的迹**必然小于1，因为根据上述的计算和讨论，显然有$tr(\rho_{mix}^2) = \sum p_i^2 < (\sum p_i)^2 = 1$。这些结论及其体现的性质使得我们具有了区分纯态与混合态的能力，只要给出量子态的密度矩阵ρ，下述两种方法任取其一：

- 计算ρ^2，对比ρ；如果$\rho = \rho^2$，那么ρ表示一个纯态，否则为混合态；
- 计算ρ^2，并计算其对角线元素和，即这个矩阵的迹$tr(\rho^2)$。如果$tr(\rho^2) = 1$，那么ρ表示一个纯态；如果$tr(\rho^2) < 1$，那么ρ表示一个混合态。

> 让我们进一步研究。在数学上，矩阵的迹还同时是**特征值之和**。由前面的讨论，密度矩阵的特征值之和也自然为1，这说明密度矩阵的特征值也是一种概率！但是特征值表示的概率**不是**对应计算基态的概率。考虑一个简单的情形，即单量子位纯态$\alpha|0\rangle + \beta|1\rangle$的密度矩阵：
>
> $$\begin{pmatrix}|\alpha|^2 & \alpha\bar{\beta}\\ \bar{\alpha}\beta & |\beta|^2\end{pmatrix}$$
>
> 容易得到它的特征值是$1, 0$，分别对应的特征向量是：
>
> $$\begin{pmatrix}\alpha & \bar{\beta}\\ \beta & -\bar{\alpha}\end{pmatrix}$$
>
> 这是在说，对于这个单量子位而言，必然处于纯态$\alpha|0\rangle + \beta|1\rangle$，而绝不可能处于$\bar{\beta}|0\rangle - \bar{\alpha}|1\rangle$，而这两个纯态是正交的。因此，将密度矩阵对角化，实际是将原量子态从计算基态张成的空间变换到另一个正交纯态张成的空间，而特征值则是量子态处于对应纯态的概率。这一点可以推广到任意量子位纯态以及混合态。因此，由于概率的非负性，任意密度矩阵是**半正定的**。

根据这个讨论，我们又获得了判断给定矩阵是否是一个密度矩阵的能力。如果您在使用上述第二个方法判定密度矩阵所表示的的量子态时，出现了$tr(\rho^2) > 1$的情况，就应当使用下述性质检查密度矩阵是否

有误：

- 对于给定矩阵ρ，当且仅当$tr(\rho)=1$且半正定时，是密度矩阵。

二、量子门

现在我们来研究量子位的变换。在上篇文章说到，我们可以通过酉矩阵来使叠加态发生变换，即输入一个叠加态、则输出一个量子态，这显然是一种门电路。但与经典情况中的不同，对于经典门电路，是输入两个经典位、而只输出一个经典位，这个过程是熵增的、不可逆的——也就是说，**不能由输出推知输入**。

而量子门不会造成信息的丢失。量子门的本质是**酉矩阵**，酉矩阵总存在$UU^{\dagger}=U^{\dagger}U=I$，只要得到这个量子门的共轭转置矩阵就能恢复原来的量子态。所有的**量子门**都必须对应一个酉矩阵。可逆性是量子算法与经典算法的本质区别之一。

1.由输入输出构造量子门

假如有一个量子门，但我们并不知道它的具体数学形式，那么这个数学形式应当如何获知呢？一种方法是遍历它的输入与输出，关键在于利用各计算基态间正交的特性。我们如下操作：对于一个n位的量子门，将2^n个计算基态的输入输出遍历一遍：

$$|i\rangle \xrightarrow{\text{计算基态经过量子门后得到}} |result_i\rangle$$

那么量子门就可以写为：

$$U=\sum_{i=0}^{2^n-1}|result_i\rangle\langle i|$$

这是因为，对于任意输入的计算基态$|j\rangle$，对任意$i\neq j$都存在$\langle i|j\rangle=0$，因此：

$$U|j\rangle=\sum_{i=0}^{2^n-1}|result_i\rangle\langle i|j\rangle=|result_j\rangle\langle j|j\rangle=|result_j\rangle$$

那么对于任意叠加态，即任意计算基态的线性组合，其输出也必然是各计算基态对应输出的线性组合，并且**权重相等**：

$$U\sum_{j=0}^{2^n-1}a_j|j\rangle=\sum_{j=0}^{2^n-1}a_j|result_j\rangle$$

但我们尚且没有保证这样构造出的量子门是酉矩阵。然而证明是简单的。由于叠加态是归一的，所以$|result_i\rangle$也必然是**单位的**；同时，正交的输入必然得到**正交的输出**。为此可以做以下计算：

$$U^\dagger = \sum_{i=0}^{2^n-1} |i\rangle\langle result_i|$$

$$\begin{aligned} U^\dagger U &= \sum_{i=0}^{2^n-1}\sum_{j=0}^{2^n-1} |i\rangle\langle result_i|result_j\rangle\langle j| \\ &= \sum_{i=0}^{2^n-1} |i\rangle\langle result_i|result_i\rangle\langle i| \\ &= \sum_{i=0}^{2^n-1} |i\rangle\langle i| \\ &= I \end{aligned}$$

另一方面：

$$\begin{aligned} UU^\dagger &= \sum_{i=0}^{2^n-1}\sum_{j=0}^{2^n-1} |result_i\rangle\langle i|j\rangle\langle result_j| \\ &= \sum_{i=0}^{2^n-1} |result_i\rangle\langle i|i\rangle\langle result_i| \\ &= \sum_{i=0}^{2^n-1} |result_i\rangle\langle result_i| \\ &= I \end{aligned}$$

因此U是个**酉矩阵**。这样构造量子门是正确的。

2.简单的单量子门重顾

我们还提到过几个比较经典与重要的**单量子门**。首先是NOT门，习惯上我们更喜欢称为X门，它反转了量子比特。假设量子态$|\psi\rangle = (\alpha,\beta)^T = \alpha|0\rangle + \beta|1\rangle$，那么：

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad X|\psi\rangle = \beta|0\rangle + \alpha|1\rangle$$

然后是Z门，它将相位翻转：

$$Z=\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix},\quad Z|i\rangle=(-1)^i|i\rangle,\quad Z|\psi\rangle=\alpha|0\rangle-\beta|1\rangle$$

相位门R_θ，它将相位顺时针旋转θ角，为此我们应将量子态改写为$|\psi\rangle=cos\varphi|0\rangle+sin\varphi|1\rangle$：

$$R_\theta=\begin{pmatrix}cos\theta & -sin\theta\\sin\theta & cos\theta\end{pmatrix},\quad R_\theta|\psi\rangle=cos(\varphi+\theta)|0\rangle+sin(\varphi+\theta)|1\rangle$$

单位门I，对量子态不做操作：

$$I=\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},\quad I|\psi\rangle=|\psi\rangle$$

哈达玛变换H，使计算基态进入叠加态：

$$H=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix},\quad H|i\rangle=\frac{1}{\sqrt{2}}(|0\rangle+(-1)^i|1\rangle)$$

$$H|0\rangle=|+\rangle=\frac{1}{\sqrt{2}}|0\rangle+\frac{1}{\sqrt{2}}|1\rangle$$

$$H|1\rangle=|-\rangle=\frac{1}{\sqrt{2}}|0\rangle-\frac{1}{\sqrt{2}}|1\rangle$$

3.由单量子门扩展多量子门

单量子门扩展多量子门的数学基础是**张量积**的性质：

$$(A\otimes B)(C\otimes D)=(AC)\otimes(BD)$$

这条数学公式的物理意义是：当A,B为单量子门，C,D为单量子右矢时，分别对量子态C和D**同时分别**做酉变换A和B，相当于对量子态$C\otimes D$**整体**做酉变换$A\otimes B$。甚至，上述描述根本不需要“单量子”的限制，对于任意多位的量子位都是成立的。例如下列矩阵：

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 & 0 & 0\\1 & -1 & 0 & 0\\0 & 0 & 1 & 1\\0 & 0 & 1 & -1\end{pmatrix}$$

它很显然是单位矩阵与哈达玛变换的张量积$I\otimes H$，它就相当于在一个双量子位的系统中只对第二个量子位作哈达玛变换。

在对n量子位系统的每一量子位都做**相同**酉变换U时，我们用记号$U^{\otimes n}$来表示。例如，$H^{\otimes n}|0^n\rangle=(H|0\rangle)^{\otimes n}=\frac{1}{\sqrt{2^n}}(|0\rangle+|1\rangle)^{\otimes n}$，结果是一个所有元素均为$\frac{1}{\sqrt{2^n}}$的$2^n$阶列向量。

4.几个多量子门

由单量子门拓展多量子门是简单的，但也有许多多量子门是不能通过上述方法得到的。比如**受控非门**$CNOT$门，这是一个双量子门。当第一个量子位为0时，不对第二量子位操作；否则，将第二量子位**取非操作**。也就是对应如下变换：

$$|00\rangle \to |00\rangle \quad , \quad |01\rangle \to |01\rangle$$

$$|10\rangle \to |11\rangle \quad , \quad |11\rangle \to |10\rangle$$

应用上文由输入输出构造量子门的方法，容易得到这个量子门是：

$$\begin{aligned} CNOT &= |00\rangle\langle 00| + |10\rangle\langle 10| + |11\rangle\langle 10| + |10\rangle\langle 11| \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

您会发现这个量子门没法表示为两个单量子门的张量积。受控非门对双量子位的作用，比较简便好写的写法是$CNOT|xy\rangle = |x\rangle|x \oplus y\rangle$。$\oplus$是**按位异或**，也就是按位二进制**无进位**加法。

从$CNOT$门出发，可以得到$Toffoli$门等任意多位的控制非门。例如$Toffoli$门是一个**三量子位门**，具有两个控制位。当两个控制位都为1的时候，才对受控位取非。也就是：

$$Toffoli|xyz\rangle = |xy\rangle|xy \oplus z\rangle$$

对于一个受控U门——首先是一位的控制位，**之后**接上任意位的受控位，只有当控制位为1的时候才对受控位做变换U——这个量子门的矩阵形式是不能用单量子门张量积扩展得到的。虽然也能用输入输出构造，但显然过于繁琐。这里我们说明一个简便的构造方法。既然当控制位为0时U不作用，反之为1时作用，因此：

$$\begin{aligned} CU &= (|0\rangle\langle 0| \otimes I) + (|1\rangle\langle 1| \otimes U) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes I + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes U \\ &= \begin{bmatrix} I & O \\ O & U \end{bmatrix} \end{aligned}$$

也就是，这种情况下的受控U门，只需要简单地代入上式即可得到矩阵形式了。例如，最简单的$CNOT$门就符合这种构造方法。同理，您也可以非常容易地得到，只有控制位为0时U门才对受控位作用的受控门的矩阵形式是：

$$\begin{bmatrix} U & O \\ O & I \end{bmatrix}$$

按照同样的方法，同样可以扩展到控制位有任意位的情况。其矩阵形式依然是相同的。