UNM Physics 452/581: Introduction to Quantum Information, Solution Set 1, Fall 2007

1.1 Pgate mania

- (a) **Matrix Representation of Irreversible Pgates** A gate just represents a function that takes in n pbits and spits out m pbits, $f : \mathbb{B}^n \to \mathbb{B}^m$ which in turn can be represented as a $2^m \times 2^n$ matrix. Since the irreversible versions of AND and OR take in $n = 2$ pbits and spit out $m = 1$ pbit, the representative matrices must be 2×4. Similarly, the FANOUT gate takes in $n = 1$ pbit and spits out $m = 2$ pbits, indicating a 4×2 matrix.

  - AND
  $$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
  - OR
  $$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$
  - FANOUT
  $$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

  Reversibility Conditions In order for a gate to be reversible, we know that we must be able to uniquely determine the input given an output. This further implies that a reversible pgate on n pbits must output n pbits, otherwise we would not have enough labels to uniquely map the outputs to the inputs. Pgates which satisfy these constraints are *permutations*. As mentioned in Lecture 1, permutation matrices have a single 1 in each row and column. Many of you were on the right trail, mentioning that the matrix representation must have an inverse that is a stochastic matrix. However, naming permutations explicitly was the desired answer (and was discussed in lecture 1).

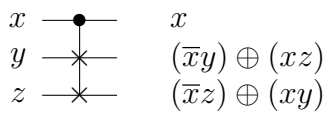
- (b) **Fredkin and Toffoli Gates** I will use the following notation for this question. The product of two boolean variables (xy), represents the logical AND operation. The sum of two boolean variables ($x + y$) represents the logical OR operation. The NOT of a variable is represented as $\overline{x}$. XOR, or addition modulo 2, is still represented as $x \oplus y$. The truth table for the Fredkin gate is

x	y	z	x'	y'	z'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	1	0
1	1	0	1	0	1
1	1	1	1	1	1

The truth Table for the Toffoli gate is given in the lecture notes. As maps, we see that the Fredkin map is $\text{FRED}(x, y, z) \mapsto (x, (\overline{x}y) \oplus (xz), (\overline{x}z) \oplus (xy))$ and the Toffoli map is $\text{TOFF}(x, y, z) \mapsto (x, y, z \oplus (xy))$.

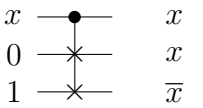
In both cases, the minimum number of gates required to simulate the other is 4. A proof of such a minimum requires exhaustively demonstrating that 1, 2 or 3 gates is insufficient to achieve the simulation. Such a proof is omitted here and was not required for credit. Unfortunately, intuition seems to be the best alternative to finding the minimum number of gates and the related circuit.

– i. Fredkin simulates Toffoli. We first explore some of the features of the Fredkin gate. Generally, we write the gate as

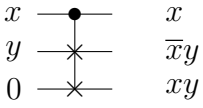


For particular choices of ancilla, we find

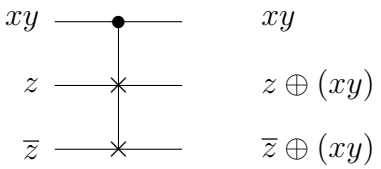
1. FANOUT/NOT



2. AND

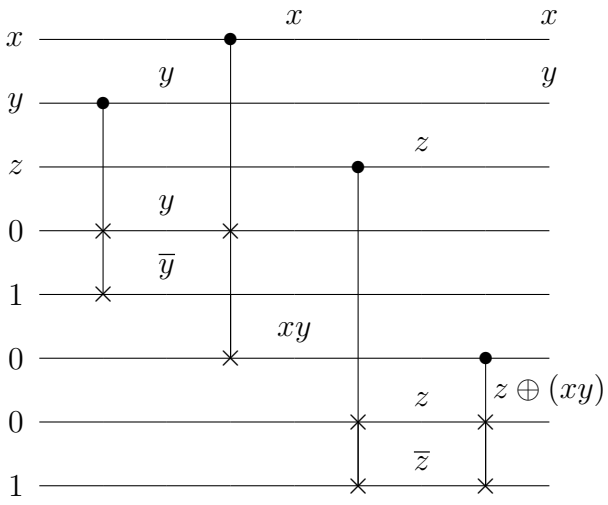


In terms of our desired output, we see that the first two inputs are trivially mapped (aside from any possibly FANOUTs needed). The third output requires combining xy with z:

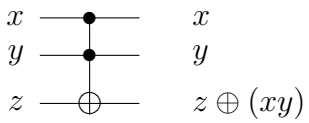


Combining these steps gives the following circuit, (with relevant intermediary

outputs labeled):

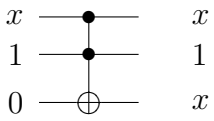


– ii. Toffoli simulates Fredkin. We write the Toffoli gate as

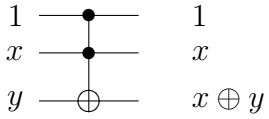


Relevant properties of the Toffoli gate are

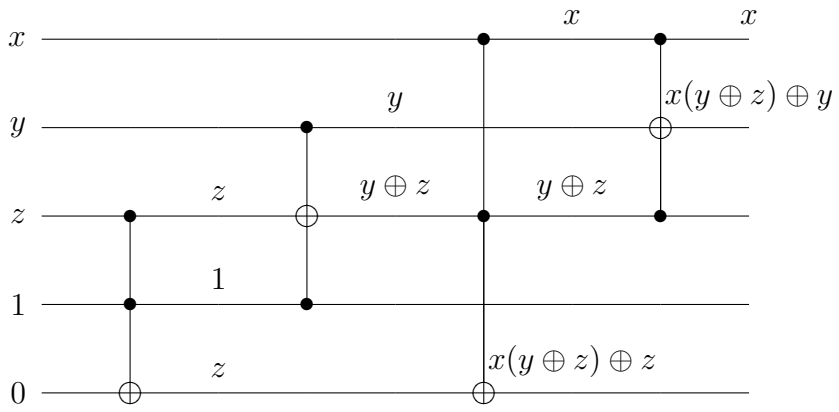
1. FANOUT



2. XOR/CNOT



We also need to recognize that the Fredkin map can be equivalently written as $\text{FRED}(x, y, z) \mapsto (x, x(y \oplus z) \oplus y), x(y \oplus z) \oplus z$. This can be intuitively seen by noticing that if the control bit x is not set, the outputs are unchanged. If x is set, the outputs are swapped using the fabled XOR swap trick. Combining these results, we write the complete circuit (with relevant intermediary ouputs labeled):



- (c) **(i) Pdits and Pdit Gates** Let $\mathbb{D} = \{0, 1, \ldots, d-1\}$ represent the d values a pdit takes on (analogous to $\mathbb{B}$ for pbits). For gates that act on two qubits, the basis is labeled by $\mathbb{D} \otimes \mathbb{D} = \{0 \otimes 0, 0 \otimes 1, 0 \otimes 2 \ldots, 1 \otimes 0, \ldots (d-1) \otimes (d-1)\}$, which gives d^2 distinct labels. Therefore, CSUM, which takes 2 pdits, will be represented by a matrix of size $d^2 \times d^2$. However, under the map $(i, j) \mapsto (i, i \oplus j)$, i remains unchanged, indicating a block diagonal structure for the matrix. Let C_i represent the i-th block of the entire CSUM matrix:

$$\begin{bmatrix} \mathrm{C}_1 & 0 & \cdots & 0 \\ 0 & \mathrm{C}_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \mathrm{C}_d \end{bmatrix}$$

Within the C_0 block, we see the output is simply the identity. In the C_1 block, the output is $1 \oplus j$, so that $\{0 \mapsto 1, 1 \mapsto 2, \ldots, (d-2) \mapsto d, (d-1) \mapsto 0$. This is just a single permutation or "cycle" of the basis states. Further consideration reveals that each successive block is the next cycle. Explictly:

  - CSUM, $d = 3$.

$$\mathrm{C}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathrm{C}_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathrm{C}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

  - CSUM, $d = 4$.

$$\mathrm{C}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathrm{C}_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathrm{C}_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\mathrm{C}_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

(ii) Inverses Since the matrix is block diagonal, each block can be inverted separately. The inverse of a cycle is just cycling the other way! We see then that

- $d = 3$,

$$\mathrm{CSUM}^{-1} = \begin{bmatrix} \mathrm{C}_0 & 0 & 0 \\ 0 & \mathrm{C}_2 & 0 \\ 0 & 0 & \mathrm{C}_1 \end{bmatrix}$$

- $d = 4$,

$$\mathrm{CSUM}^{-1} = \begin{bmatrix} \mathrm{C}_0 & 0 & 0 & 0 \\ 0 & \mathrm{C}_3 & 0 & 0 \\ 0 & 0 & \mathrm{C}_2 & 0 \\ 0 & 0 & 0 & \mathrm{C}_1 \end{bmatrix}$$

1.2 Something Is Rotten in the State of Denmark

Let $p(b)$ denote the probability that Hamlet is drawing from the pocket containing coins with bias $b \in \{H, T\}$. Let $p(b|i)$ denote the probability that Hamlet decides the pocket has bias b given that his flip yields outcome i. Let $p(i|b)$ denote the probability that outcome i will be observed if he flips the coin with bias b. Let $\pi(b)$ denote Hamlet's prior probability assesment that he has chosen the pocket with bias b.

Bayes' rule tells us that

$$p(b|i) = \frac{\pi(b)p(i|b)}{p(i)}. \tag{1}$$

Hamlet's decision procedure (called the Bayes' decision method in the literature) has the feature that when outcome i is obtained, the probability of a correct decision is the maximum of $p(H|i)$ and $p(T|i)$. In other words, his probability of error is

$$P_e = \sum_i p(i)\,(1 - \max\{p(H|i), p(T|i)\}) \tag{2}$$

$$= \sum_i p(i)\,\min\{p(H|i), p(T|i)\} \tag{3}$$

$$= \sum_i \min\{\pi(H)p(i|H), \pi(T)p(i|T)\}, \tag{4}$$

where Bayes' rule was used in the last step.

From the details of the problem, we have

$$\pi(H) = \pi(T) = 1/2 \tag{5}$$

and

$$p(i|H) = p_i, \qquad p(i|T) = q_i, \tag{6}$$

where

$$\vec{\mathbf{p}} := \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}, \qquad \vec{\mathbf{q}} := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{7}$$

in part (a) and

$$\vec{\mathbf{p}} := \begin{bmatrix} 0.96 \\ 0.04 \end{bmatrix}, \qquad \vec{\mathbf{q}} := \begin{bmatrix} 0.04 \\ 0.96 \end{bmatrix} \tag{8}$$

in part (b).

Plugging in these values we get for part (a),

$$P_e(1 \text{ flip}) = \frac{1}{2}\min\{0.9, 0\} + \frac{1}{2}\min\{0.1, 1\} = 0.05. \tag{9}$$

and for part (b),

$$P_e(1 \text{ flip}) = \frac{1}{2}\min\{0.96, 0.04\} + \frac{1}{2}\min\{0.04, 0.96\} = 0.04. \tag{10}$$

Therefore, according to this measure, the pbits in part (b) are more distinguishable than the pbits in part (a).

The extension to a two-flip test follows in the same framework, only now there are four outcomes to consider. For part (c) we have

$$\begin{aligned} P_e(2 \text{ flips}) &= \frac{1}{2}\min\{0.9 \times 0.9, 0 \times 0\} + \frac{1}{2}\min\{0.9 \times 0.1, 0 \times 1\} && (11) \\ &\quad + \frac{1}{2}\min\{0.1 \times 0.9, 1 \times 0\} + \frac{1}{2}\min\{0.1 \times 0.1, 1 \times 1\} && (12) \\ &= 0.005, && (13) \end{aligned}$$

and for part (d) we have

$$\begin{aligned} P_e(2 \text{ flips}) &= \frac{1}{2}\min\{0.96 \times 0.96, 0.04 \times 0.04\} + \frac{1}{2}\min\{0.96 \times 0.04, 0.04 \times 0.96\} && (14) \\ &\quad + \frac{1}{2}\min\{0.04 \times 0.96, 0.96 \times 0.04\} + \frac{1}{2}\min\{0.04 \times 0.04, 0.96 \times 0.96\} && (15) \\ &= 0.04. && (16) \end{aligned}$$

After this two-coin flip test, the pair of coins from part (a) are ten times more distinguishable according to this measure, yet the coins from part (b) are no more distinguishable than before. Evidently, this measure is not one that amplifies well under further data acquisition. So while this measure has a good operational motivation, it's not so great at assessing how distinguishable pbits $\vec{\mathbf{p}}$ and $\vec{\mathbf{q}}$ are without specifying an exact number of samplings.

A (possibly) deeper understanding is provided by noting that the part (b) coins are symmetric in probability, whereas those in part (a) are not. Since Hamlet's prior does not

prefer one coin type to another, additional flips of the symmetric coins do not help amplify the results of a single flip, because the outcome probabilities for multiple flips maintain the symmetry. When summing over the possible flip outcomes i, we add together "symmetric probability pairs" (so flipping tails-tails with the heads-biased coin contributes the same amount to the probability of error as flipping heads-heads with the tails-biased coin). In short, the coins of part (b) result in errors in the same way, whereas those in part (a) do not.

1.3 So teach, when's the homework due?

For this question, assume that our pdit is given by

$$\vec{p} := \begin{bmatrix} p_F \\ p_P \\ p_N \end{bmatrix}$$

- (a) Let $\vec{p}_n$ represent your state of knowledge at week n. The probablity of having a full assignment for week n is $P(F_n) = (p_F)_n$, the first entry of the vector. In order to update our state of knowledge for week $n+1$, we have $P(F_{n+1}) = P(F_{n+1}|F_n)P(F_n) + P(F_{n+1}|P_n)P(P_n) + P(F_{n+1}|N_n)P(N_n)$. We can write similar equations for N_{n+1} and P_{n+1}. The graph given in the assignment represents the conditional probabilities in the equations, where each inward arrow represents the probablity to be where the arrow's head is at week $n+1$ given that you were at the tail location at week n. The full P matrix applied to $\vec{p}_n$ updates to $\vec{p}_{n+1}$ using these rules and each row in P represents the inward arrows for that state in the graph.

  $$P = \begin{bmatrix} 0.40 & 0.45 & 0.80 \\ 0.35 & 0.40 & 0.15 \\ 0.25 & 0.15 & 0.05 \end{bmatrix}$$

  You can easily verify that this is a stochastic matrix

- (b) To find the probabilities two weeks from now, we apply the P matrix in series:

  $$\begin{aligned} \vec{p}_2 &= P^2\vec{p}_0 \\ &= \begin{bmatrix} 0.4 & 0.45 & 0.8 \\ 0.35 & 0.4 & 0.15 \\ 0.25 & 0.15 & 0.05 \end{bmatrix}^2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &\approx \begin{bmatrix} 0.43 \\ 0.35 \\ 0.22 \end{bmatrix} \end{aligned}$$

  There is a 43% chance of having a full assignment in two weeks, given that there is no assignment this week.

- (c) Considering the results of part (b), we see that the matrix P^2 maps pdits at week 0 to pdits at week 2. As P_{13} represents the conditional probability $P(F_1|\ N_0)$, P^2_{13}

represents the conditional probability $P(F_2|\ N_0)$. We will now use these conditional probabilities between weeks 2 and 0 with Bayes rule to find $P(N_0|F_2)$. Since we have no information about the past, we take our vector of priors to be $\vec{p}_0 = [1/3,\ 1/3,\ 1/3]^T$.

$$P(N_0|\ F_2) = \frac{P(F_2|\ N_0)P(N_0)}{P(F_2)} = \frac{P(F_2|\ N_0)P(N_0)}{P(F_2|N_0)P(N_0) + P(F_2|P_0)P(P_0) + P(F_2|N_0)P(N_0)} \approx 0.30$$

For the remaining parts of the question, it will be useful to diagonalize the P matrix. Recall from linear algebra that the process of diagonalization involves finding a matrix V such that $D = V^{-1}PV$ is diagonal. In this representation, V is a matrix whose columns are the eigenvectors of P and $D_{ii} = \lambda_i$, where λ_i is the i-th eigenvalue. Using MATLAB's `eig` command, we find that P has eigenvalues $1, -0.26, 0.11$ and a corresponding V matrix:

$$\begin{pmatrix} -0.7922 & -0.8099 & 0.5445 \\ -0.5354 & 0.3153 & -0.7992 \\ -0.2930 & 0.4946 & 0.2547 \end{pmatrix}$$

where each column in V corresponds to the the first, second and third eigenvalue respectively. Using the diagonal form of a matrix is useful for computer powers of that matrix, as

$$P^n = PPP\cdots PP = PVV^{-1}PVV^{-1}PV\cdots V^{-1}P = V(V^{-1}PV)^nV^{-1} = VD^nV^{-1}$$

and calculating D^n requires simply raising each diagonal entry to the power n.

- (d) Using the above, we calculate $\vec{p}_{16} = VD^{16}V^{-1}\vec{p}_0$ and find $\vec{p}_{16} = [.49\ .33\ .18]^T$, so the probability of a partial assignment in 16 weeks is 0.33.
- (e) An invariant pdit state $\vec{\pi}$ satisfies $\vec{\pi} = P\vec{\pi}$ and is thus simply the $+1$ eigenvalue of P. Looking at our results above, we see that this is the first column in V. Since eigenvalues are defined up to a constant, we can renormalize that column so that it represents a valid pdit state to find $\vec{\pi} = [.49\ .33\ .18]^T$.
- (f) The results of parts (d) and (e) suggest that $\vec{\pi}$, the $+1$ eigenvector, is related to the convergent gate. Using the diagonal form, we have that $(D^n)_{ii} = (D_{ii})^n$. For $\lambda_i < 1$, $(\lambda_i)^n \to 0$ as $n \to \infty$, but $1^n = 1$ for all n. So as n goes to infinity, D just becomes the matrix

  $$D^\infty = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

  So that

  $$P^\infty = VD^\infty V^{-1} = \begin{pmatrix} .49 & .49 & .49 \\ .33 & .33 & .33 \\ .18 & .18 & .18 \end{pmatrix}$$