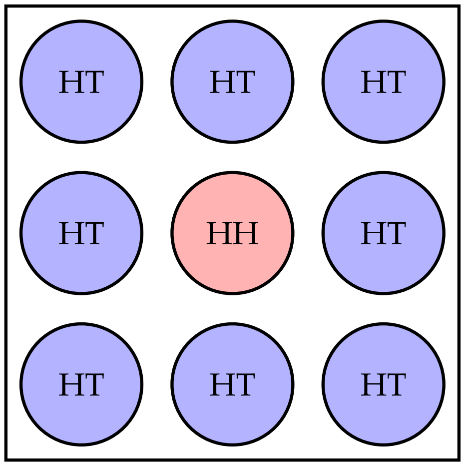
Assignment 2 Submission deadline: 23rd Mar 2026, 11:30 pm

1. **A coin box with strange coins**: Li Lei has a coin box that contains 9 coins. 8 of them are fair, standard coins (heads and tails) and 1 coin has heads on both sides.



1.1) **(2 points)** If Li Lei randomly chooses a coin and flip it, what is the probability of getting a head?

1.2) **(3 points)** Li Lei randomly chose a coin without looking at it. He flipped it four times, and his friend Han Meimei told him that he got four heads. Now, if Li Lei flip the coin again, what is the probability of getting a head?

2. **Virus detection kits**: It is known that the ***prevalence*** of a virus is 0.01. There are two detection kits for this particular virus used by many hospitals. They are produced by different manufacturers and based on completely different assays. **Kit A** has a ***sensitivity*** of 0.9 and a ***specificity*** of 0.95. **Kit B** has a ***sensitivity*** of 0.95 and a ***specificity*** of 0.9. A random person goes to a hospital to get tested by both kits **A** and **B**. Let events $\boldsymbol{E_1}$ = { Kit **A** shows a positive result }, $\boldsymbol{E_2}$ = { Kit **B** shows a positive result } and $\boldsymbol{E_3}$ = { The person carries the virus }.

2.1) **(5 points)** Calculate $\mathbb{P}(\boldsymbol{E_1})$.

2.2) **(5 points)** Calculate $\mathbb{P}(\boldsymbol{E_2})$.

2.3) **(7½ points)** Are events $\boldsymbol{E_1}$ and $\boldsymbol{E_2}$ independent? Why or why not? You can use the mathematical formula to prove your conclusion, or you can also use your own words to describe your intuition about this question.

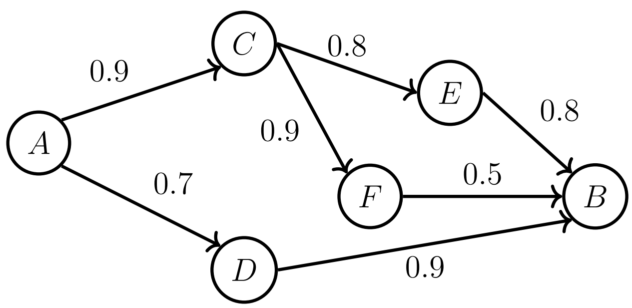
2.4) **(7½ points)** Calculate $\mathbb{P}(\boldsymbol{E_3} \mid \boldsymbol{E_1} \cap \boldsymbol{E_2})$.

3. **Five-sided die**: You have a fair five-sided die[1]. The sides of the die are numbered from 1 to 5. Each die roll is independent of all others, and all faces are equally likely to come out on top when the die is rolled. Suppose you roll the die twice.

   3.1) **(5 points)** Let event ***A*** = { the total of two rolls is 10 }, event ***B*** = { at least one roll resulted in 5 } and event ***C*** = { at least one roll resulted in 1 }. Are events ***A*** and ***B*** independent? Are events ***A*** and ***C*** independent?

   3.2) **(5 points)** Let event ***D*** = { the total of two rolls is 7 }, event ***E*** = { the difference between the two roll outcomes is exactly 1 } and event ***F*** = { the second roll resulted in a higher number than the first roll }. Are events ***E*** and ***F*** independent? Are events ***E*** and ***F*** independent given that event ***D*** has occurred?

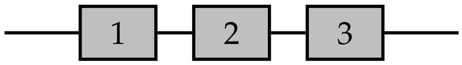
4. **(9 points)** **Network reliability and signalling pathway**: Let's start with a simple computer network connecting two nodes from A to B through intermediate nodes C, D, E, F as shown in the picture below:



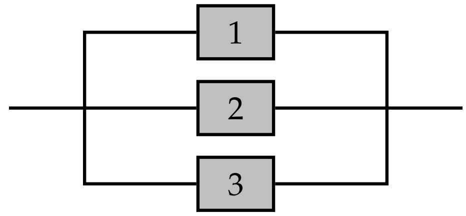
For every pair of directly connected nodes, say i and j, there is a given probability p_{ij} that the link from i to j is up. It is generally of interest to calculate the probability that there is a path connecting A and B in which all links are up.

This is a typical problem of assessing the reliability of a system consisting of components that can fail ***independently***. You will encounter this type of system very often in real life. Such a system can often be divided into subsystems, where each subsystem consists in turn of several components that are connected either in series or in parallel, like shown below:

[1]Can you imagine what a fair five-sided die looks like in your head?



Series connection



Parallel connection

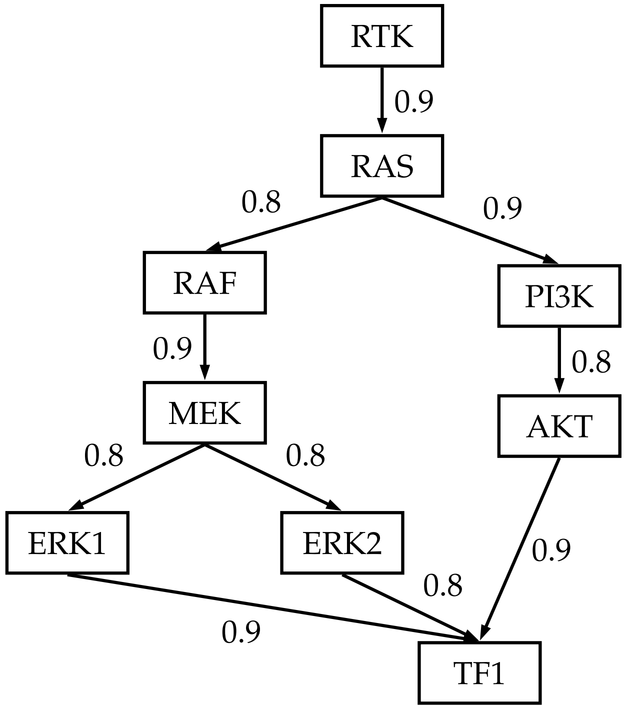
Let a subsystem consist of components $1, 2, 3, ..., m$, and let p_i be the probability that component i is up ("succeeds"). Then a series subsystem succeeds if ***all*** of its components are up, so its probability of success is the product of the probabilities of success of the corresponding components, $i.e.$

$$\mathbb{P}\,(\text{series subsystem succeeds}) = \prod_{i=1}^{m} p_i = p_1\, p_2\, p_3 \cdots p_m$$

A parallel subsystem succeeds if ***any one*** of its components succeeds, so its probability of failure is the product of the probabilities of failure of the corresponding components, $i.e.$

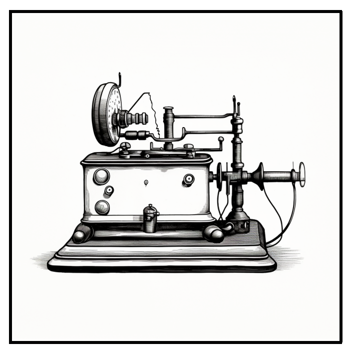
$$\begin{aligned}\mathbb{P}\,(\text{parallel subsystem succeeds}) &= 1 - \mathbb{P}\,(\text{parallel subsystem fails}) \\ &= 1 - \prod_{i=1}^{m}(1 - p_i) \\ &= 1 - (1 - p_1)(1 - p_2)(1 - p_3)\cdots(1 - p_m)\end{aligned}$$

Now let's look at a signalling pathway inside a cell, which is very important for the functions of a cell. The signalling pathway consists of a cascade of protein phosphorylation events. Many proteins are only activated when they are phosphorylated. In the following diagram depicting a simplified version of the MAP and PI3 kinase pathways, "$A \to B$" means "when A is activated, it has a probability of phosphorylating B", and the probability of the phosphorylation is indicated on the edge:



Assume all phosphorylation events are independent and all proteins shown in the above picture must be activated first before they can phosphorylate their downstream proteins to activate them. Now, the protein RTK is activated, what is the probability that the protein TF1 is activated ?

5. **Communication Through A Noisy Channel**: A source transmits a message (a string of symbols) through a noisy communication channel. Each symbol is 0 or 1 with probability p and $1-p$, respectively, and is received incorrectly with probability ϵ_0 and ϵ_1, respectively. Errors in different symbol transmissions are independent.



5.1) **(3 points)** What is the probability that the k-th symbol is received correctly?

5.2) **(3 points)** What is the probability that the string of symbols 1011 is received correctly?

5.3) **(4 points)** In an effort to improve reliability, each symbol is transmitted three times and the received string is decoded by the majority rule. In other words, a 0 (or 1) is transmitted as 000 (or 111, respectively), and it is decoded at the receiver as a 0 (or 1) if and only if the received three-symbol string contains at least two 0s (or 1s, respectively). What is the probability that a 0 is correctly decoded?

5.4) **(4 points)** For what values of ϵ_0 is there an improvement in the probability of correct decoding of a 0 when the scheme of **5.3)** is used?

5.5) **(4 points)** Suppose that the scheme of **5.3)** is used. What is the probability that a symbol was 0 given that the received string is 101?

6. **(5 points)** **Another Prisoner's Dilemma**: The *Prisoner's Dilemma* is a classic game analysed in the game theory. Today, we have another prisoner, and he is in a different dilemma. There are only three prisoners (including him) in the prison. The release of two of them has been announced, but their identity is kept secret. He considers asking a friendly guard to tell him who is the prisoner other than himself that will be released, but hesitates based on the following rationale: at his present state of knowledge, the probability of being released is $2/3$, but after he knows the answer, the probability of being released will become $1/2$, since there will be two prisoners (including him) whose fate is unknown and exactly one of the two will be released. Apparently, he does not have a good understanding of conditional probabilities and independence. Let event **A** = { He is one of the chosen prisoner to be released } and event **B** = { The friendly guard tells him the identify of one of the prisoner other than himself that will be released }. Show him that A and B are independent.

7. **Independence of event complements**: if events $\boldsymbol{A}$ and $\boldsymbol{B}$ are independent, are the following events independent or not? If they are independent, prove your answers. If they are not independent, please provide an example.

   7.1) **(1 point)** $\boldsymbol{A}^C$ and $\boldsymbol{B}$

   7.2) **(1 point)** $\boldsymbol{A}$ and $\boldsymbol{B}^C$

   7.3) **(1 point)** $\boldsymbol{A}^C$ and $\boldsymbol{B}^C$

8. **(4 points)** **Will It Rain In The Weekend?**: The weather forecaster says that the probability of rain on Saturday is 25% and that the probability of rain on Sunday is 25%. Li Lei uses the additivity axiom to claim that the probability of rain during the weekend is 25% + 25% = 50%. Why is he wrong? What is the actual probability of rain during the weekend assuming rain on Saturday and rain on Sunday are independent?

9. **(5 points)** **The Bayesian Search Theory**: Recall from what we talked about during the lecture. We divide the search area into a grid of cells:

$i = 1$	2	3	4	5
6	7	$\cdots$		

We index those cells as i, where $i = 1, 2, 3, \cdots$. Then we gather experts, discuss, use simulations $etc.$ to come up with a set of ***prior probabilities*** about where the lost object is. We denote them as π_i. For example, π_1 is the prior probability of the lost object is in cell 1. The next step is to consider our equipment, the complexity of each cell and other factors to decide the probability of finding the lost object in cell i given that the lost object is in cell i. We denote the probability as p_i. For example, p_3 is the probability of finding the lost object in cell 3 given that it is indeed in cell 3. Once we finish the setup, the probability of finding the lost object in cell i is basically $\pi_i p_i$. We simply start from the cell with the highest $\pi_i p_i$.

Now we search cell i but fail to find the lost object in cell i. Given that has occurred, calculate the ***posterior probabilities*** that the lost object is in this cell i and other cell j ($j \neq i$). Compare them to the prior probabilities π_i or π_j, respectively. Are the posterior probabilities bigger or smaller than their corresponding prior probabilities? Are your results consistent with your intuition?

Optional: If you have time and know how to code, you can spend some of your spare time writing a simulation about the Bayesian search on a 5×5 grid of cells with different values of π_i, p_i and vary the true location of the lost object. Run the simulation many times to get a sense about the Bayesian search method. I suppose this is a good practice for Java that your learnt in your first year.

10. **Finding the dog**: Li Lei has lost his dog in either forest A (with a priori probability 0.4) or forest B (with a priori probability 0.6). On any given day, if the dog is in A and Li Lei spends a day searching for it in A, the conditional probability that he will find the dog that day is 0.25. Similarly, if the dog is in B and Li Lei spends a day looking for it there, the conditional probability that he will find the dog that day is 0.15. The dog cannot go from one forest to the other. The forests are dangerous at night, so Li Lei can search only in daytime and have to go back home during nighttime. Answer the following question:

10.1) **(4 points)** In which forest should Li Lei look to maximise the probability he finds his dog on the first day of the search?

10.2) **(4 points)** Given that Li Lei looked in A on the first day but did not find his dog, what is the probability that the dog is in A?

10.3) **(4 points)** If Li Lei flips a fair coin to determine where to look on the first day and successfully finds the dog on the first day, what is the probability that he looked in A?

10.4) **(4 points)** If the dog is alive and not found by the N^{th} day of the search, it will die that evening with probability $\frac{N}{N+2}$. Li Lei has decided to look in A for the first two days. What is the probability that he will find a live dog for the first time on the second day?