02458 Cognitive modelling E19
2^{nd} Assignment

Adrian Atienza (s192780)
Arianna Taormina (arita@fysik.dtu.dk)
Marie Claire Capolei (macca@elektro.dtu.dk)
Maud Leclerc (s191975)

Bayes Exercise 1

Bayesian approach in the study of perception is based on some fundamental concepts as the likelihood function and the prior.

Re-constructing a 3D image with 2D information is an ill-posed problem, we just do not have enough information to get to a unique solution: there are infinite number of possible interpretations. That's the case of Necker Cube, the cube perception is highly influenced by the prospective, e.g. from-above-prospective (fap) and from-below-prospective (fbp), resulting in an unstable perception.

Although the unstable perception scenario, the human visual system experiences only two possible interpretation as compatible with the sensory information, over an infinite range. The visual system does not overlap the two interpretations, but it alternates the two prospective fap and fbp, in a reversal process determined by the observer internal fluctuations.

Likelihood and priors in the Bayesian framework give us a good interpretation of why this possibly happens.

The Likelihood captures how likely a particular scene corresponds to the perceived image, or rather, it represents how likely is an assumption to be true.

Let B be the set of observed outcomes, and A the set of parameters describing the stochastic process. The likelihood is conditioned by the B observation and is function of the A unknown parameters that need to maximize the probability of having the observation,

$L(A|B) = P(B|A)$

In fact, we proceed by maximizing the likelihood (aka minimizing the minus log likelihood; this is mainly to move the problem from maximizing the result of products to minimizing the result of sums, its more handy).

The other important factor is given by Priors: those are distributions that represent the observer prior believes (of which the observer is not necessarily aware of) before some evidences are taken into account, and provide a "principled way of formulating constraints on possible scenes that lead to unambiguous visual perception". Priors represent what we think is more plausible in the real word sort of "independently" from the current sensory information (this is not completely true as we will specify soon).

Bayes theorem:

- $P(A|B)*P(B) = P(B|A)*P(A)$
- $P(A|B) = P(B|A)*P(A) * (1/P(B))$

Where

- $P(B|A)$ = Likelihood (Not a probability distribution)

- P(A) = Prior, characterized by bias and confidence in the bias. If we model the prior in the form of normal distributions, then bias = mean of the distribution and the confidence in inversely proportional to the std standard deviation. Which mean and which std to use then? The mean can be given by the average perception, but confidence levels estimation, aka std, is trickier. Confidence in fact can be tested and constructed by changing the condition of the stimulus - lightening or orientation, e.g. increasing the contrast in an image decreases the standard deviation of the Prior that encompass the luminance effect, which then is more stable.
  Prior are not fixed entities but are affected by the stimulus characteristics.

- (1/P(B)) = corrective term that turns the Likelihood into a probability distribution, it is not necessary, and it can be omitted.

- P(A|B) = is the posterior distribution that considers both geometrical knowledge and prior weights. If we don't have a strong prior, then our final output will be mainly influenced by the likelihood, or rather, the observations.

1) Question 1:

Let $\mathbf{S}_{3x8}$ be the matrix describing the 3-dimensional structure of the real cube wireframe, and $\mathbf{I}_{3x8}$ the matrix describing the projection of $\mathbf{S}$ on the 2D plane given the $\mathbf{M}_{4x4}$ orthographic transformation matrix. $\mathbf{M}$ is function of the observer's α_{el} elevation and α_{az} azimuth angles.
We first estimate the $\mathbf{S}_{hat}$ observer's perception of the cube by minimizing the likelihood error of the estimate image,

$$\text{Argmin}[\text{ WSSE}] = \text{Argmin}[\text{ std}_N^{-2} \bullet (\mathbf{I} - \mathbf{I}_{hat})^2],$$

where WSSE stand for weighted sum of squared errors, although we are here considering only the likelihood error. $\mathbf{I}_{hat}$ is the projection of the $\mathbf{S}_{hat}$ observer's intuitive estimate on the 2D plane,

$$\mathbf{I}_{hat} = \mathbf{M} \ \mathbf{S}_{hat} \ ,$$

And std_N is the standard deviation of the noise affecting the visual perception.
Considering the observer's position respect to the cube is,

$$\alpha_{el} = 25° \ , \ \alpha_{az} = -32° \ ,$$

And the noise is normally distributed with standard deviation,

$$\text{std}_N = 2,$$

The result is shown in **Fig.1**, from the observer's position (left cube) the perceived image (in blue) tend to resemble a cube, but if we change prospective the wireframe results mashed. The Likelihood only is not enough to predict the 3D scene. It gives in fact only a good prediction of the 2D image projected on our field of vision, which inevitably breaks the moment we move around the viewpoint. Priors on the consistency of volumes, lengths, angles or luminosity can bridge the gap of information. When two Priors are in discordance, we have the switching from one interpretation to the other.

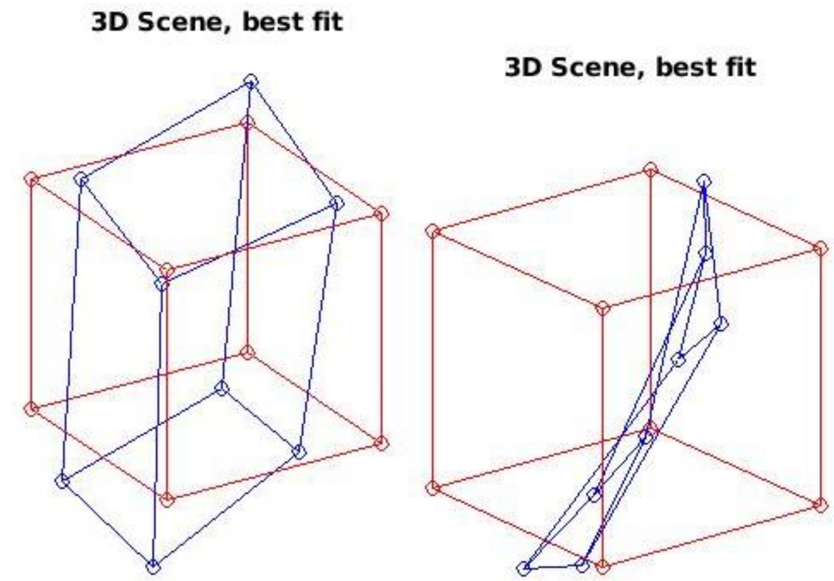

Figure 1. Necker cube perception from different point of view considering only the likelihood error. In red the real cube, in blue the perceived.

We extend the WSSE weighted sum of squared errors used for the optimization of the estimate, including first the information regarding the angles,

$$\mathrm{Argmin}[\ \mathrm{WSSE}] = \mathrm{Argmin}[\ \mathrm{std}_N^{-2} \bullet (\mathbf{I} - \mathbf{I}_{hat})^2 + \mathrm{std}_{s1}^{-2} \bullet (\mathbf{S}_1 - \mathbf{S}_{1,hat})^2],$$

Where $\mathbf{S}_1$ is a matrix of rectangular angles, and $\mathbf{S}_{1,hat}$ is the matrix describing the estimated angle, computed from the $\mathbf{S}_{hat}$ matrix.

Given the same prospective as in the previous test case, and perception noise,

$$\mathrm{std}_{s1} = 2,$$

The outcome is apparently improved (**Fig.2**), however it results not robust to changes in the observation angles (right cube).

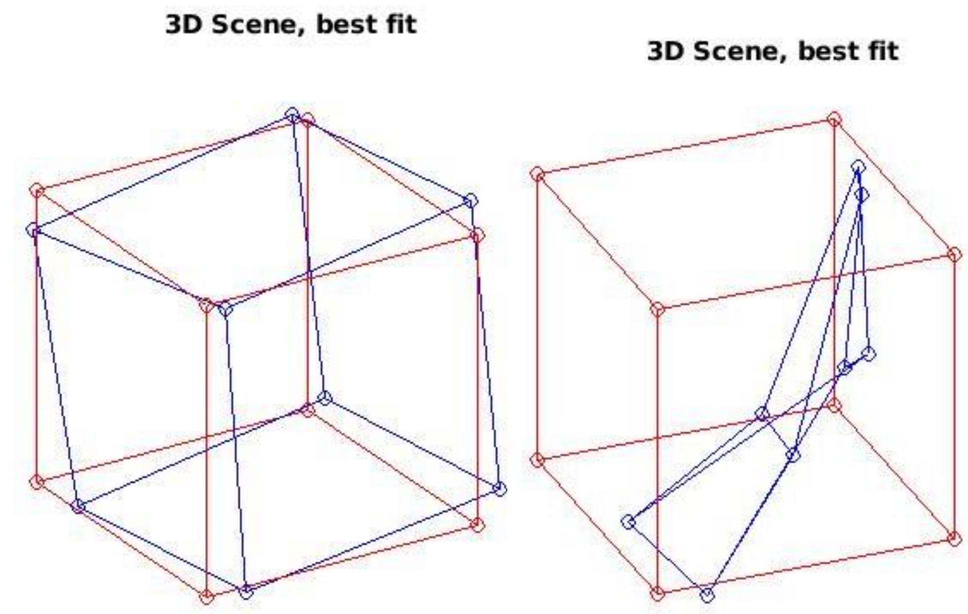

Figure 2. Necker cube perception from different point of view considering the likelihood and the perceived angles error. In red the real cube, in blue the perceived.

We consider an extra features in the estimation: the lengths of the cube,

$$\text{Argmin[WSSE]} = \text{Argmin[} std_N^{-2} \bullet (\mathbf{I} - \mathbf{I}_{hat})^2 + std_{s1}^{-2} \bullet (\mathbf{S}_1 - \mathbf{S}_{1,hat})^2 + std_{s2}^{-2} \bullet (\mathbf{S}_2 - \mathbf{S}_{2,hat})^2 \text{]},$$

Where $\mathbf{S}_2$ is a matrix of real lengths, and $\mathbf{S}_{2,hat}$ is the matrix describing the estimated lengths, computed from the $\mathbf{S}_{hat}$ matrix. The perception results highly robust to varying observer's position thanks to the larger amount of prior information exploited during the perception process (**Fig.3**).

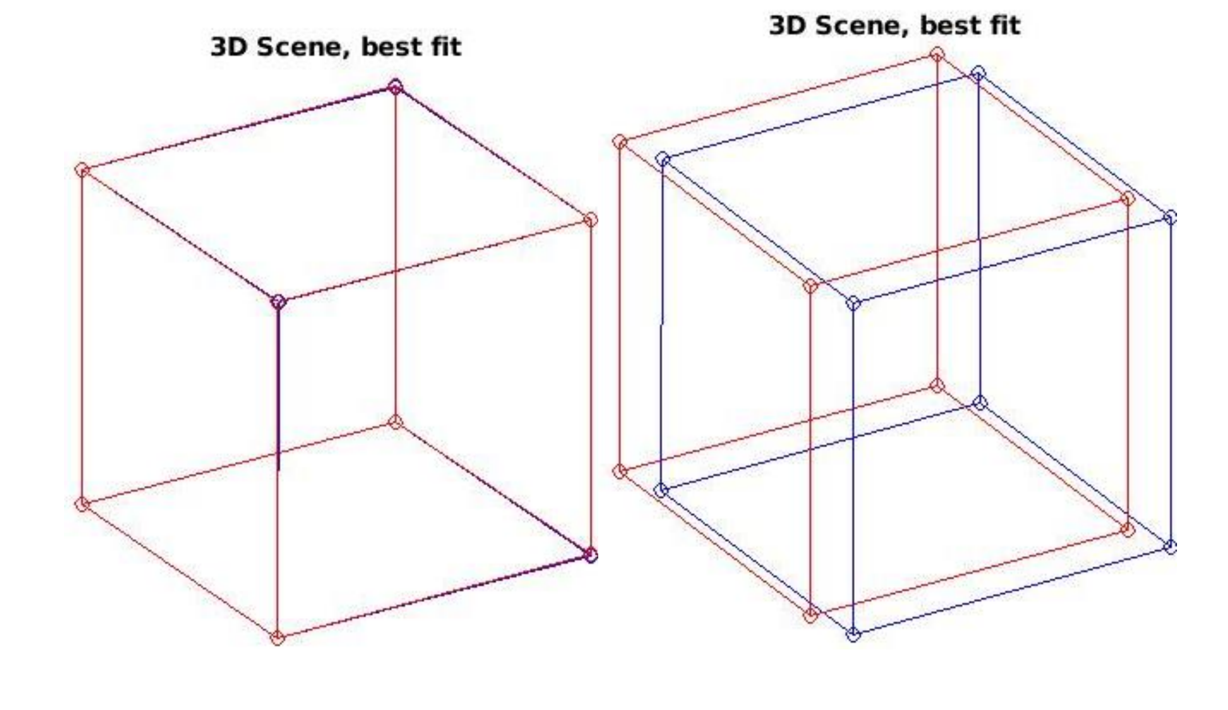

Figure 3. Necker cube perception from different point of view considering the likelihood, the perceived angles error, and the lenght. In red the real cube, in blue the perceived.

other comments?

2) Question 2:

With only 8 vertexes out of 12 , we don’t get enough information from the Priors to have a good prediction. Modifying the code with the number of angles and lengths to 12 solved the issue.

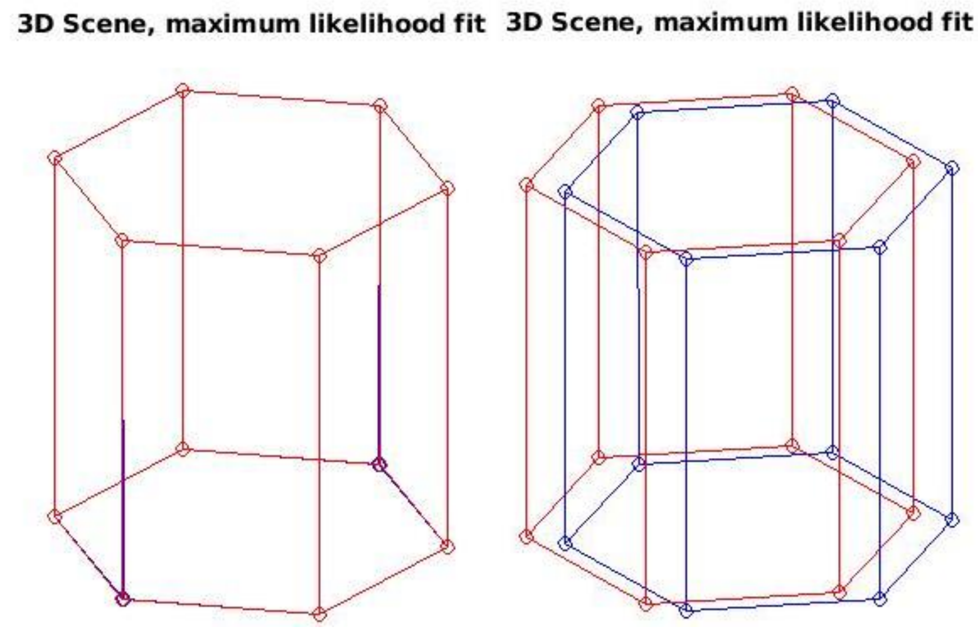

Figure 3. Hexagonal prism perception from different point of view considering the likelihood, the perceived angles error, and the lenght. In red the real prism, in blue the prism.

3) Question 3:

Binocular vision allows us to perceive a single 3D image of our surroundings, gives a wide field of vision (ca 190 of which 120 for both eyes) and, very important, the different position of the eyes in the head (defined as parallax or binocular disparity) allows our brain to get to a precise depth perception via a process called stereopsis. In the last model we have once again the real 3D Scene (S), with the real projection on 2D planes of vision of each eye (I1 and I2). Here we have hp same elevation (EL1 = EL2) for both and a different azimuth angle (AZ1 = AZ2-60 degrees). We calculate the minimum of NegLogLikelihood given by the sum of MSE for each eye (S_guess as a contribution from both viewpoints and therefore the projections on the two different planes I_guess_1 and I_guess_2).

We see that, also without any Prior on volume, lengths, angles o light, we get a perfect estimate of the original 3D scene.

Bayes Exercise 2

4) Question 1:

Ripe juju-fruits are orange with a wavelength of

$$Wave_{Juju,ripe} = 600 \pm 50 \text{ nm},$$

while not ripe juju fruits are in a range of

$$Wave_{Juju,\sim ripe} = 500 \pm 50 \text{ nm},$$

We can describe this information with two normal distribution respectively,

$$N_{w,Juju,ripe}(600,50^2),$$
$$N_{w,Juju,\sim ripe}(500,50^2).$$

The P(Juju,ripe) prior probability of juju-fruits to be ripe is,

$$P(Juju,ripe) = 0.15,$$

accordingly, the P(Juju,~ripe) prior probability of juju-fruits to not be ripe is,

$$P(Juju,\sim ripe) = 1 - P(Juju,ripe) = 0.85.$$

The posterior probability of a juju-fruit to be ripe given a wavelength between 540-550 nm is low,

$$P(Juju,ripe \mid 540\text{-}550) = \beta^{-1} \bullet P(540\text{-}550 \mid Juju,ripe) \bullet P(Juju,ripe) = \mathbf{0.1263}$$

Where β is the ratio,

$$\beta = (P(540\text{-}550 \mid Juju,ripe) \bullet P(Juju,ripe) + P(540\text{-}550 \mid Juju,\sim ripe) \bullet P(Juju,\sim ripe)$$

and P(540-550| Juju,ripe) is the likelihood of having that wavelength interval if the juju-fruit is ripe, or rather,

$$P(540\text{-}550 \mid N_{w,Juju,ripe}(600,50^2)) = \mathbf{L}(N_{w,Juju,ripe}(600,50^2) \mid 550) - \mathbf{L}(N_{w,Juju,ripe}(600,50^2) \mid 540),$$

Same argument is valid for the P(540-550| Juju,~ripe) likelihood of not being ripe given that wavelength interval.

This low probability is highly compromised by the low prior P(Juju,ripe) to have ripe juju-fruit, for instance, if we change the prior to,

$$P(Juju,ripe) = 0.85,$$

The posterior probability of having a ripe juju-fruit given that wavelength range drastically increase,

$$P(Juju,ripe \mid 540\text{-}550) = 0.8228.$$

5) Question 2:

The jungle offers to the monkeys several type of fruits, statistics show that among all the edible fruits only a small percentage are juju-fruits, or rather the prior of being a juju-fruit is,

$$P(Juju) = 0.10,$$

While half of the fruits are mongo berries,

$$P(mongo\ berries) = 0.50,$$

and the remaining are chakavas,

$$P(chakavas) = 1 - P(Juju) - P(mongo\ berries) = 0.40.$$

The reflected light of ripe and not ripe mongo berries and chakavas can describe as normal distribution,

$$N_{w,mongo,ripe}(580,20^2),$$
$$N_{w,mongo,\sim ripe}(520,20^2),$$

for the mongo berries, and

$$N_{w,chakavas,ripe}(400,100^2),$$
$$N_{w,\ chakavas,\sim ripe}(550,100^2),$$

for the chakavas respectively. The P(mongos,ripe) and P(chakavas,ripe) prior probability of mongos berries and chakavas to be ripe are,

$$P(mongos,ripe) = 0.80,$$

$$P(chakavas,ripe) = 0.10,$$

accordingly, the P(mongos,~ripe), P(chakavas,~ripe) prior probability of mongos berries and chakavas to not be ripe is,

$$P(mongos,\sim ripe) = 1 - P(mongos,ripe) = 0.20,$$

$$P(chakavas,\sim ripe) = 1 - P(chakavas,ripe) = 0.90,$$

The probability that a random fruit is ripe given a reflecting wavelength between 540-550 is,

$$P(ripe|540\text{-}550) = \beta^{-1} \bullet \alpha = \mathbf{0.4018},$$

Where α is the sum of the likelihoods multiplied the priors of being ripe for each fruit, times the probability of existence of that type of fruit,

$$\alpha = \alpha_1 \bullet P(Juju) + \alpha_2 \bullet P(mongos) + \alpha_3 \bullet P(chakavas)$$

$$\alpha_1 = P(540\text{-}550|\ Juju,ripe) \bullet P(Juju,ripe),$$

$$\alpha_2 = P(540\text{-}550|\ mongos,ripe) \bullet P(mongos,ripe),$$

$$\alpha_3 = P(540\text{-}550|\ chakavas,ripe) \bullet P(chakavas,ripe),$$

and β is the ratio considering both the probability of being or not ripe,

$$\beta = \beta_1 \bullet P(Juju) + \beta_2 \bullet P(mongos) + \beta_3 \bullet P(chakavas)$$

$$\beta_1 = P(540\text{-}550|\ Juju,ripe) \bullet P(Juju,ripe) + P(540\text{-}550|\ Juju,\sim ripe) \bullet P(Juju,\sim ripe),$$

$$\beta_2 = P(540\text{-}550|\ mongos,ripe) \bullet P(mongos,ripe) + P(540\text{-}550|\ mongos,\sim ripe) \bullet P(mongos,\sim ripe),$$

$$\beta_3 = P(540\text{-}550|\ chakavas,ripe) \bullet P(chakavas,ripe) + P(540\text{-}550|\ chakavas,\sim ripe) \bullet P(chakavas,\sim ripe).$$

The overall probability of having a ripe fruit is largely influenced by the juju-fruits and chakavas prior of being ripe, together they represent the 50% of the fruit available in the jungle and their prior probability of being ripe is incredibly low, moreover the considered range can be out of the ripe wavelength for the mongos berries.
The event considering that the monkey is enjoying eating the fruit could be seeing as a signal that the fruit is ripe increasing the prior of that specific fruit, and consequentially influencing the future observations.

Should we write something else?

6) Question 3:

We tested the picking capabilities of the monkey per 1000 trials.

We first initialized a multinomial probability distribution including all the possible scenario of P(fruit type, ripe) and P(fruit type, ~ripe),

$$\mathbf{Probabilities}_{fruit} = [P(Juju,ripe) \bullet P(Juju)\ ,\ P(Juju,\sim ripe) \bullet P(Juju)\ ,\ P(mongos,ripe) \bullet P(mongos)\ ,\ P(mongos,\sim ripe) \bullet P(mongos),\ P(chakavas,ripe) \bullet P(chakavas)\ ,\ P(chakavas,\sim ripe) \bullet P(chakavas)\ ]$$

We used this distribution to create a vector $\mathbf{Fruits}_{1xtrials}$ of trials = 1000 random numbers corresponding to the ripe or not ripe fruits.

Per 1000 times we picked a value from the vector (fruit,(not) ripe) and the corresponding m_i mean and std_i standard deviation of the *i-th* wavelengths distribution, we used this value to calculate the random interval,

$$\mathbf{interval}_i = [x-5,x+5]$$

that will be used in to compute the likelihoods,

$$P(\mathbf{interval}_i\ |\ N_{w,fruit,ripe}(m_i,\ std_i^{\ 2})),$$

$$P(\mathbf{interval}_i\ |\ N_{w,fruit,\sim ripe}(m_i,\ std_i^{\ 2})).$$

As in question 2, for the I-th trial the probability that the I-th fruit is ripe given a reflecting wavelength between the interval $\mathbf{interval}_i$ is,

$$P(ripe|\ \mathbf{interval}_i) = \beta^{-1} \bullet \alpha\ ,$$

Where α is the sum of the likelihoods multiplied the priors of being ripe for each fruit, times the probability of existence of that type of fruit,

$$\alpha = \alpha_1 \bullet P(\text{Juju}) + \alpha_2 \bullet P(\text{mongos}) + \alpha_3 \bullet P(\text{chakavas})$$

$$\alpha_1 = P(\textbf{interval}_i \mid \text{Juju,ripe}) \bullet P(\text{Juju,ripe}),$$

$$\alpha_2 = P(\textbf{interval}_i \mid \text{mongos,ripe}) \bullet P(\text{mongos,ripe}),$$

$$\alpha_3 = P(\textbf{interval}_i \mid \text{chakavas,ripe}) \bullet P(\text{chakavas,ripe}),$$

and β is the ratio considering both the probability of being or not ripe,

$$\beta = \beta_1 \bullet P(\text{Juju}) + \beta_2 \bullet P(\text{mongos}) + \beta_3 \bullet P(\text{chakavas})$$

$$\beta_1 = P(\textbf{interval}_i \mid \text{Juju,ripe}) \bullet P(\text{Juju,ripe}) + P(\textbf{interval}_i \mid \text{Juju,\textasciitilde ripe}) \bullet P(\text{Juju,\textasciitilde ripe}),$$

$$\beta_2 = P(\textbf{interval}_i \mid \text{mongos,ripe}) \bullet P(\text{mongos,ripe}) + P(\textbf{interval}_i \mid \text{mongos,\textasciitilde ripe}) \bullet P(\text{mongos,\textasciitilde ripe}),$$

$$\beta_3 = P(\textbf{interval}_i \mid \text{chakavas,ripe}) \bullet P(\text{chakavas,ripe}) + P(\textbf{interval}_i \mid \text{chakavas,\textasciitilde ripe}) \bullet P(\text{chakavas,\textasciitilde ripe}).$$

If the I-th $P(\text{ripe} \mid \textbf{interval}_i)$ probability is higher than 0.5, so is ripe, and the picked fruit is actually ripe than we increase a counter. The accuracy of the monkey will be,

Accuracy = sum of correct response / trials = 0.8010

Exercise 3

See code