

CO6: Introduction to Computational Neuroscience

<http://iec-lnc.ens.fr/group-for-neural-theory/teaching-260/article/co6-course>

Lecturer: Johann Lussange
Ecole Normale Supérieure
29 rue d'Ulm, 2nd floor,
e-mail: johann.lussange@ens.fr

Tutor: Gregory Dumont
Ecole Normale Supérieure
29 rue d'Ulm, ground floor, room 72
e-mail: gregory.dumont@ens.fr

Solutions to the 1st exercise sheet

On the following pages, you will find example solutions to the written exercises. Although these (hopefully) all are correct, they are not necessarily the only possible solutions. In particular, you are free to name the variables you use - as long as you define (= tell me explicitly) what you mean by which letter.

After all, variables are just name tags, and the same letter in different contexts may mean two very different things. Or you might want to represent the same thing by two different letters, depending on the situation in which you use it. But when it is clear from the context and from your definitions what exactly you are referring to, it is often convenient to use a very simple notation. If you are talking about the probability of making a particular choice, for example, all of the following *can* mean the same thing - if you *tell* me that this is what you want:

$$p(c = 1), \ p(1), \ p_1, \ \hat{P}_1, \ \Pi(1), \ \dots$$

If you have any questions regarding these solutions or the exercises in general, feel free to contact by e-mail (johann.lussange@ens.fr or boris.gutkin@ens.fr).

Math reminder

Although this week's exercise mostly relied on basic math, there were some concepts which may be less commonplace to people with a non-mathematical background. Quite certainly, you had to deal with the multiplication of exponential functions:

$$e^a e^b = e^{a+b}$$

$$(\dots)^{-a} = \frac{1}{(\dots)^a}$$

You may also have encountered their inverse - (natural) logarithms...

$$e^a = b \Rightarrow a = \log b$$

...as well as the capital-sigma notation for sums:

$$\sum_{k=1}^N f(k) = f(1) + f(2) + \dots + f(N)$$

For example:

$$\sum_{x=1}^3 \frac{x^2}{x+1} = \frac{1}{1+1} + \frac{4}{2+1} + \frac{9}{3+1} = \frac{49}{12}$$

If you don't yet feel comfortable working with these, don't worry. If you let us know by e-mail or via the 'anonymous comment box' on our website, I will cover them again during the next seminar.

If you want to do some reading to refresh your memory of the kind of math required for this class, there are many books out there which you may find helpful. In particular, the following one seems to cover all of the concepts you would have required to obtain full score on the initial math test, from sums and logarithms to integrals and eigenvalues:

*Introduction to Mathematics
for Life Scientists*
Edward Batschelet, Springer 1979

This being said, I haven't actually read it yet, so I can't tell you how well-written this book really is. Judge for yourselves.

Attribution of points

For each exercise sheet, you can obtain up to 20 points, plus a number of bonus points for working on the exercises labeled as 'advanced'.

This time, points were attributed as follows:

exercise	points
1a	1
1b	1
1c	6
1d	5
1e	4
1f	3
1g	+ 4

While the amount of equations I have included in the solutions corresponds approximately what I expect when grading your exercises, you certainly won't need as much text as I produced. This being said, it doesn't hurt to add a sentence or two to make sure I understand exactly what you're doing!

Bees and flowers

Exercise 1a

Using the definition of a softmax policy given on the exercise sheet, this was an easy one:

$$\begin{aligned} \sum_{c=1}^2 p(c) &= p(c=1) + p(c=2) \\ &= \frac{e^{\beta m_1}}{e^{\beta m_1} + e^{\beta m_2}} + \frac{e^{\beta m_2}}{e^{\beta m_1} + e^{\beta m_2}} \\ &= \frac{e^{\beta m_1} + e^{\beta m_2}}{e^{\beta m_1} + e^{\beta m_2}} = 1 \end{aligned}$$

Do you remember why the softmax policy looks exactly this way?

Exercise 1b

Since all exponential functions are non-zero for finite arguments, this one is straightforward as well:

$$\begin{aligned} p(c=1) &= \frac{e^{\beta m_1}}{e^{\beta m_1} + e^{\beta m_2}} \\ &= \frac{e^{\beta m_1}}{e^{\beta m_1} + e^{\beta m_2}} \cdot \frac{e^{-\beta m_1}}{e^{-\beta m_1}} \\ &= \frac{e^{\beta m_1} e^{-\beta m_1}}{e^{\beta m_1} e^{-\beta m_1} + e^{\beta m_2} e^{-\beta m_1}} \\ &= \frac{e^{(\beta m_1 - \beta m_1)}}{e^{(\beta m_1 - \beta m_1)} + e^{(\beta m_2 - \beta m_1)}} \\ &= \frac{1}{1 + e^{\beta(m_2 - m_1)}} \end{aligned}$$

Exercise 1c

For $\beta = 1$, the graph of the probability function should look like the one in figure ??:

Apart from being mathematically plausible, this should intuitively make sense as well. So let's have a look at different values of $d = m_2 - m_1$. A very high d means that flower 2 is a lot more attractive than flower 1. The probability for the bee to go to flower 1 should thus be close to zero. Which it is. On the

other hand, a very low d means that flower 1 is far more attractive, and consequently, the bee should virtually always go there. The probability $p(c=1) = p_1$ should thus be close to one. Which it is.

If the difference between the two flowers is less pronounced, the bee should show a mixed behaviour of exploration and exploitation: It should go more often to the more attractive flower, yet still pay a visit to the less attractive one from time to time. Thus, there should be a smooth transition from $p_1 = 1$ to $p_1 = 0$. If both flowers are equally attractive, the probability to go to flower 1 should be at exactly 0.5. Which it is.

Exercise 1d

For different values of β , the probability distribution should look just like figures ??, ?? and ?? included at the end of this document.

Remember that β is a measure of how much the bee tends to exploit its knowledge, rather than to improve it by exploring its environment. So for a higher value of β , the bee should stick with the more attractive flower for most of the time - even if it is only slightly more attractive than the other. Thus, even for values of d which are relatively close to zero, the probability p will now be near one or near zero.

If $\beta = 0$, on the other hand, there should be no exploitation whatsoever. Which means that no matter what the bee already knows about its environment, it is just going to keep on exploring, and to go to both flowers with equal likelihood. Thus, its probability distribution will be flat - that is, stay at a constant value of 0.5, no matter how similar or how different the two flowers are.

Now what happens if β is negative? Because

$$\begin{aligned} p_1(d; \beta = -1) &= \frac{1}{1 + e^{-1(+d)}} \\ &= \frac{1}{1 + e^{+1(-d)}} \\ &= p_1(-d; \beta = 1) \end{aligned}$$

the probability distribution looks just like the one for $\beta = 1$, albeit reflected along the abscissa. But what does this mean for the behaviour of such a bee? It means that the bee will tend to go to the flower which it believes to be the *less attractive* one, and even more so the less attractive it is! This, of course, is a decidedly stupid behaviour, which would probably lead to our stupid bee starving. Thus, although we can easily deal with negative values of β within our mathematical framework, these do not make any sense biologically.

Exercise 1e

Here, we will have to expand the softmax framework we used earlier from two flowers to many flowers. In order to get there, remember why the probability distribution for two flowers looks the way it does. How did we come up with it in the first place?

Well, imagine that what we want is to find a simple model for the two-flower case. (a) We simply want it to be more likely for the bee to go to the more attractive flower than to go to the less likely one, and we want this probability to increase smoothly with the expected reward difference. (b) The second thing we want is that this increase should be steeper than just linear, that is: a flower which is twice as attractive than the other should be visited *more* than twice as often. Which is biologically plausible, because finding twice as much food is a huge advantage.

A very simple way (mathematically speaking) to combine these two ideas is to use an exponential function. Hence, our probability to go to flower 1 might just be proportional to an exponential function which depends on the attractiveness of flower 1 (that is: the reward m_1 expected from going there):

$$p(c=1) = p(1) = A \cdot e^{\beta m_1} \quad (1)$$

Here, A is just some constant for which we haven't decided yet what we want it to look like. β is another constant which will determine just how steep the increase with increasing attractiveness is. Now, do we need to make some

more assumptions to decide on which A to use in our model? No, we don't! There is a very basic idea from probability theory which tells us exactly what A has to be: Since the bee has to do *something*, the probability to do something has to be 100%. Which means that the sum of the probabilities of all possible actions has to be equal to 1. For two flowers, there are only two possible actions, and hence:

$$p(1) + p(2) = 1 \quad (2)$$

Which, according to our simple model, is the same as:

$$\begin{aligned} & A \cdot e^{\beta m_1} + A \cdot e^{\beta m_2} = 1 \\ \Leftrightarrow & A \cdot (e^{\beta m_1} + e^{\beta m_2}) = 1 \\ \Leftrightarrow & A = \frac{1}{e^{\beta m_1} + e^{\beta m_2}} \end{aligned}$$

And now we're done! We just found the softmax policy you know from the lecture:

$$p(c) = A \cdot e^{\beta m_c} = \frac{e^{\beta m_c}}{e^{\beta m_1} + e^{\beta m_2}}$$

This has worked well for two flowers - but how are we going to make it work for N flowers? Well, the only step in which we included the total number of flowers was in equation (??). So all we have to do is to rewrite this equation for N flowers instead of 2, and we should get the model we're looking for:

$$\begin{aligned} & p(1) + p(2) + \dots + p(N) = 1 \\ \Leftrightarrow & \sum_{k=1}^N p(k) = 1 \\ \Leftrightarrow & \sum_{k=1}^N A \cdot e^{\beta m_k} = 1 \\ \Leftrightarrow & A \cdot \sum_{k=1}^N e^{\beta m_k} = 1 \\ \Leftrightarrow & A = \frac{1}{\sum_{k=1}^N e^{\beta m_k}} \end{aligned}$$

If we substitute this into equation (??) just like before, this yields our new model for N different flowers:

$$p(c) = A \cdot e^{\beta m_c} = \frac{e^{\beta m_c}}{\sum_{k=1}^N e^{\beta m_k}} \quad (3)$$

As before, the parameter β will determine how sensitively the bee reacts to differences in attractiveness. In other words: we will use higher values of β to describe a "greedy", "exploitative" bee, whereas lower values correspond to a more "exploratory" animal.

Do you see how you can plot equation (??) as a function of "reward difference" when there are N different flowers to compare?

Exercise 1f

In the above exercises, we weren't really interested in how the bee comes up with its internal estimate of the expected rewards in the first place. So, where do these m_i come from?

In the lecture, we have seen that from one visit to a flower to the next, the bee may update its internal estimate in the following way:

$$m_{i,\text{new}} = m_{i,\text{old}} + \epsilon \cdot (r_{i,\text{now}} - m_{i,\text{old}})$$

Hence, such an "online update" depends on the previous internal estimate and on the prediction error $r_{i,\text{now}} - m_{i,\text{old}}$ to which it just led.

If we don't just look at this from trial to trial, but as a continuous function of time, we will obtain the following:

$$\begin{aligned} m_i(t) &= m_i(t-T) \\ &+ \epsilon \cdot [r_i(t) - m_i(t-T)] \end{aligned} \quad (4)$$

where T is the time between two visits to a flower.

Advanced exercise: Exercise 1g

If the rewards stay at constant values r_i over time, the bee's internal estimates will eventually converge to these values. According to the above learning rule, though, the bee will only adapt its internal estimate of flower i when it actually visits this very flower. The interval between two such visits is determined by the attractiveness of the flower - and on the precise behavioural policy the bee is following.

Let us first consider a *very attractive* flower among our N different specimen. The more "exploratory" the bee behaves (low β), the more time will pass between two visits to the same flower, and the longer it will take for the bee's estimate to approach the actual value of the reward. The more the bee tends to "exploit" its knowledge (high β), the more often it will visit this attractive flower, and the sooner its corresponding estimate will converge.

For *very high* β , the bee will virtually stop exploring, and stick to one of the first flowers it has visited - if it really is the most attractive or not.

For a *less attractive* flower, on the other hand, these behavioural modes will have quite the opposite effect. "Exploratory" behaviour leads to an increased number of visit, and hence to a quicker convergence. An "exploitative" bee will only rarely visit such an unattractive flower, and it will take a long time for its internal estimate of the corresponding reward to converge to its actual value.

Finally, we may want to quantify just how quickly the internal estimate for a given flower converges to its real value. Or, in other words: we would like to find a mathematical expression describing how steep the increase or decrease over time of the internal estimate really is. This is where so-called *characteristic time constants* come in handy.

What is a characteristic time constant, then? It is the time it takes for the value of a function to increase or decrease by a given percentage.

For example, the most common measure for exponential decays is the time it takes for the function value to decrease from a given value to $1/e$ times this value. It is important to note that for exponential functions, this time interval is constant, no matter which initial value we're examining. So, if we look at an exponential function $f(t) = A \cdot \exp(-at)$ at some point in time t , its characteristic time constant θ is given by:

$$\begin{aligned} f(t + \theta) &= \frac{1}{e} \cdot f(t) \\ \Leftrightarrow A \cdot e^{-a(t+\theta)} &= \frac{1}{e} \cdot A \cdot e^{-at} \\ \Leftrightarrow e^{-at} e^{-a\theta} &= \frac{1}{e} \cdot e^{-at} \\ \Leftrightarrow e^{-a\theta} &= \frac{1}{e} \end{aligned}$$

So in order to determine θ , all we need is a logarithm:

$$\begin{aligned} e^{-a\theta} &= \frac{1}{e} \\ \Leftrightarrow e^{a\theta} &= e \\ \Leftrightarrow \log(e^{a\theta}) &= \log(e) \\ \Leftrightarrow a\theta &= 1 \\ \Leftrightarrow \theta &= \frac{1}{a} \end{aligned}$$

We can thus rewrite our function $f(t)$ as follows:

$$f(t) = A \cdot e^{-at} = A \cdot e^{-t/\theta}$$

Actually, characteristic time constants are often denoted by τ . Here, we use θ instead in order to avoid confusion with the time difference " τ " as in the lecture on temporal difference learning.

So which are the characteristic time constants of convergence in our bees-and-flowers problem? In order to calculate it, we have to approximate the stepwise decrease of the bee's internal estimates by a continuous function, for which we may then determine its θ value. Let us consider the case of a flower which is visited every T seconds (for example, a flower

which is very attractive and/or attracts an "exploitative" bee).

We are thus looking for a function $M_i(t) = r_i + A \cdot e^{-t/\theta}$ which at times T , $2T$, $3T$ etc. yields the same values as our previously defined learning rule $m_i(t)$. In other words, we want this function $M_i(t)$ to fulfill equation (??), and thus:

$$\begin{aligned} M_i(t) &= M_i(t - T) + \epsilon \cdot [r_i - M_i(t - T)] \\ \Leftrightarrow r_i + A \cdot e^{-t/\theta} &= r_i + A \cdot e^{-(t-T)/\theta} + \\ &\quad + \epsilon \cdot [r_i - r_i - A \cdot e^{-(t-T)/\theta}] \\ \Leftrightarrow e^{-t/\theta} &= e^{-t\theta} e^{+T/\theta} - \epsilon \cdot e^{-t\theta} e^{+T/\theta} \\ \Leftrightarrow 1 &= (1 - \epsilon) \cdot e^{T/\theta} \\ \Leftrightarrow \frac{1}{1 - \epsilon} &= e^{T/\theta} \\ \Leftrightarrow \log((1 - \epsilon)^{-1}) &= \log(e^{T/\theta}) \\ \Leftrightarrow -\log(1 - \epsilon) &= T/\theta \\ \Leftrightarrow \theta &= \frac{-T}{\log(1 - \epsilon)} \quad (> 0) \end{aligned}$$

Are you aware of why we decided to use the concept of characteristic time constants instead of simply calculating the "total time" it takes to reach r_i - for example, by setting $m_i(t_{\text{final}}) = r_i$ and calculating t_{final} from it?

PS: Time constants for other cases (less attractive flowers etc.) can be deduced in a similar way. For $\beta = 0$, the bee will visit all flowers equally often, no matter how high or low its corresponding estimates are. Then, the characteristic time constants can be calculated as above, albeit with time intervals $N \cdot T$ instead of T .

For intermediate β , it is worth noting that the time it actually takes for the internal estimates to decrease or increase by a factor of e will also be longer than the θ above, because the internal estimate will repeatedly stall at constant values for some time if the probability to visit the flower is significantly lower than one. Still, it is possible to determine the expectation

value of the time needed for such a change by combining a learning rule and a policy describing the probabilities of possible choices (like the softmax policy above).

Figures

This week, you will find several figures attached to this document. Figure 1 illustrates exercise 1c, whereas figures 2, 3 and 4 are associated with exercise 1d.

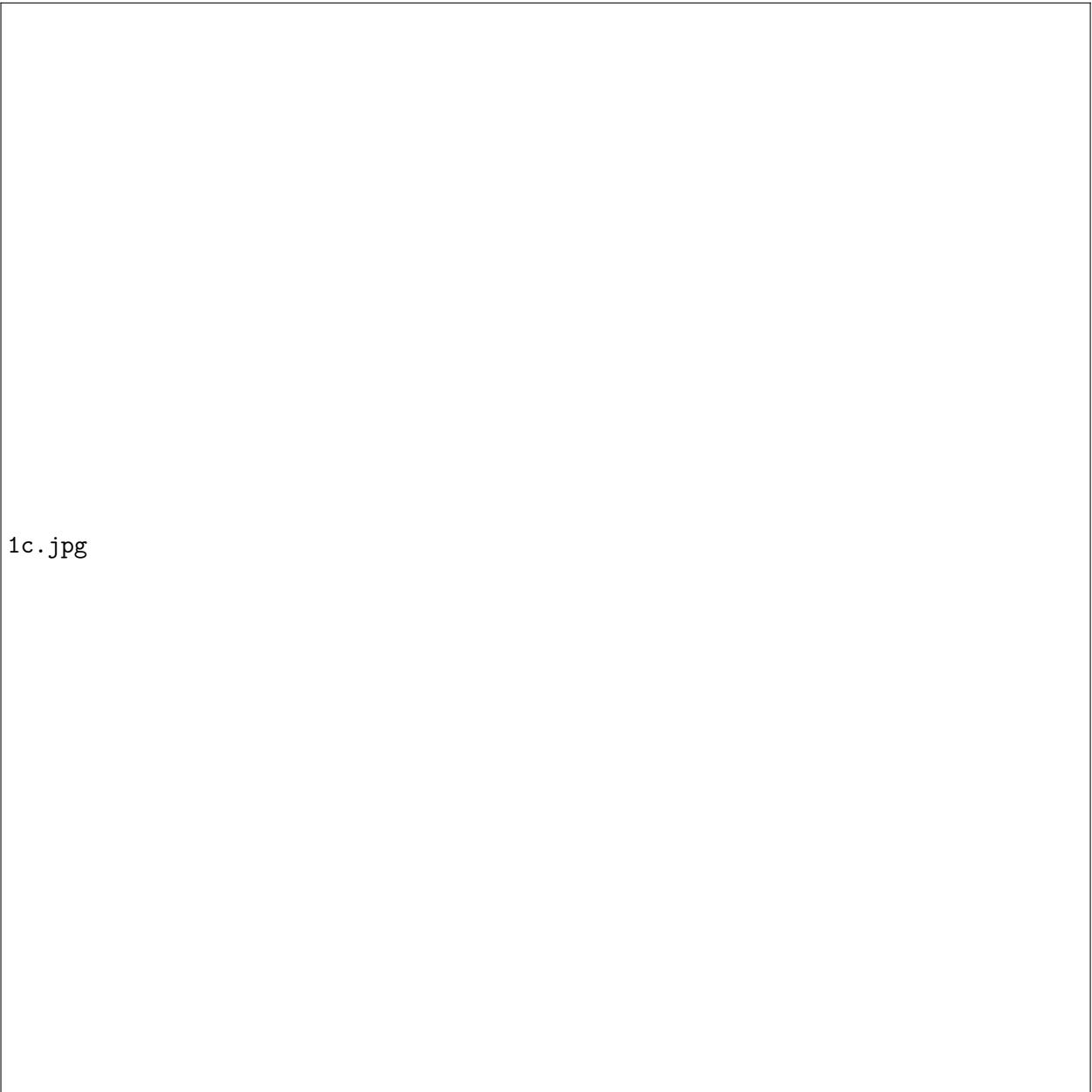
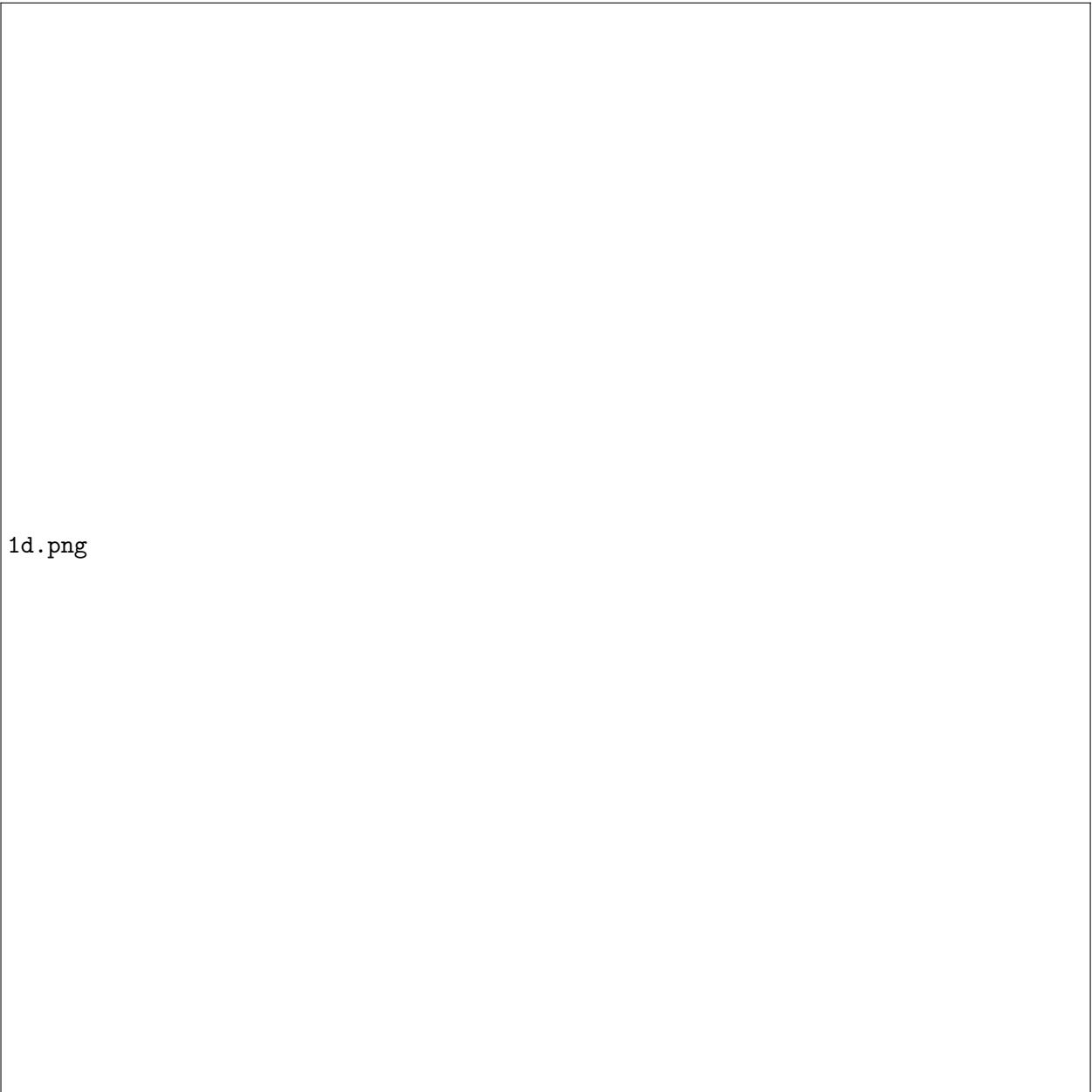
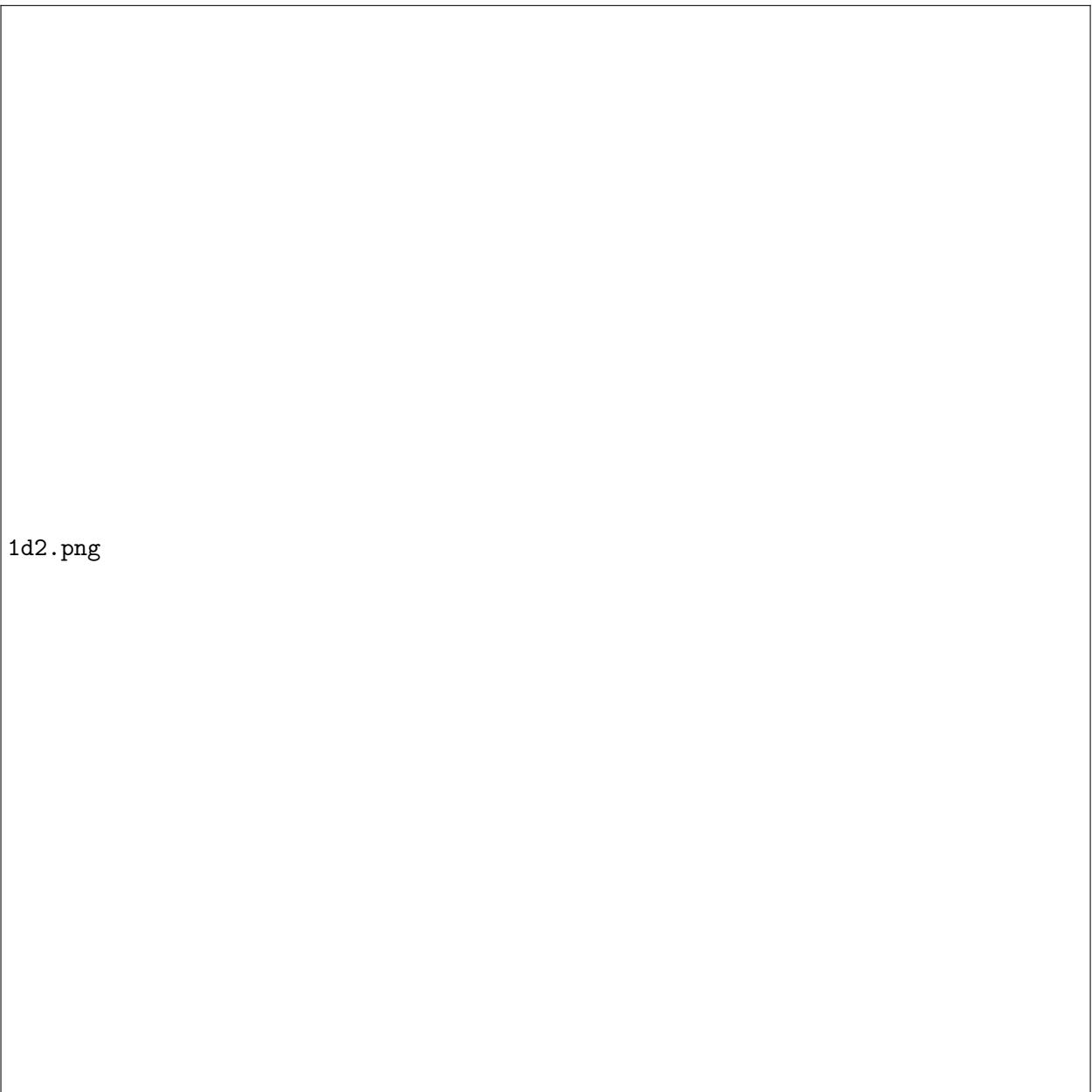


Figure 1: *Softmax policy with $\beta = 1$ (as in exercise 1c).*



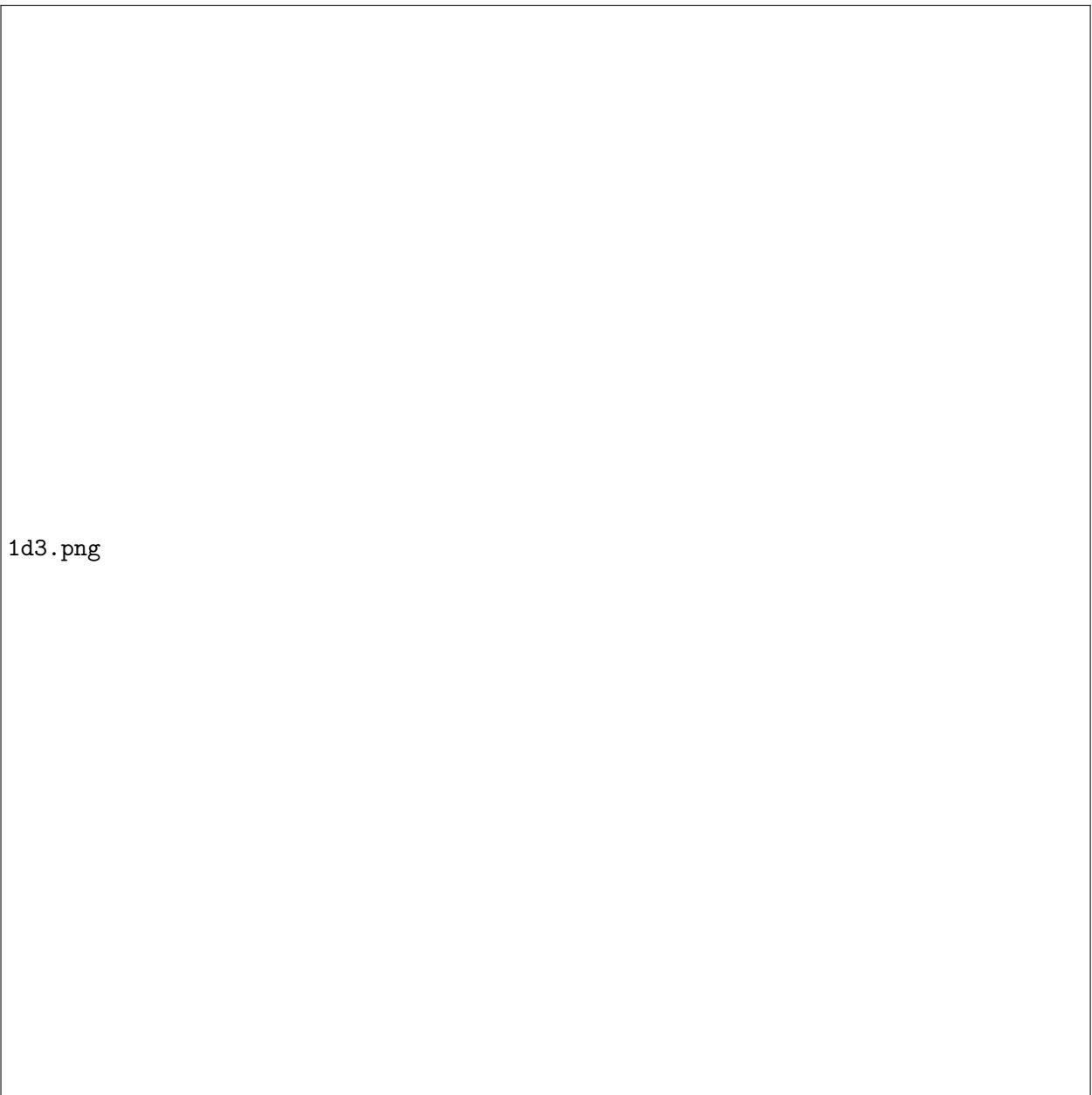
1d.png

Figure 2: *Softmax policy with $\beta = 10$ (as in exercise 1d).*



1d2.png

Figure 3: *Softmax policy with $\beta = 0$ (as in exercise 1d).*



1d3.png

Figure 4: *Softmax policy with $\beta = -1$ (as in exercise 1d).*