

Writing Problem Sets and Exams

Voon Hui Lai and Heidi Klumpe (2018)

Activity 1:

Identify the level of Bloom's Taxonomy targeted by this problem.

Rewrite the problem to target a lower or higher level of Bloom's Taxonomy.

Activity 2:

Write the objective of the problem.

Rewrite some aspect of the problem to better meet that objective.

Activity 3:

Write your step-by-step guide for putting together a problem set or exam.

BE/Bi 103, Fall 2016: Homework 2

Due 1pm, Sunday, October 9

(c) 2016 Justin Bois. Problem 2.2 is licensed under a [Creative Commons Attribution License CC-BY 4.0](#). All code contained therein is licensed under an [MIT license](#). Problem 2.1 may not be shared.

This homework was generated from an Jupyter notebook. You can download the notebook [here](#).

Problem 2.1 (20 pts)

This problem features unpublished research, which the authors wish to remain private until their work is published. BE/Bi 103 students may access the password-protected homework [here](#).

Problem 2.2 (Exploring fish sleep data, 80 pts)

In [Tutorial 2](#), we investigated sleeping states of zebrafish larvae. We had a discussion about what are the best metrics for a sleepful states based on the one-minute interval activity data we have. I think we agree that it is far from obvious how to define a sleepful state. In this problem, you will work with your group to come up with some good ways to parametrize sleep behavior and estimate the values of these parameters.

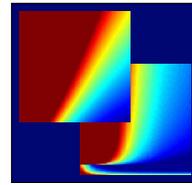
Choose two different ways to parametrize sleep behavior. You can use sleep metrics from the [Prober, et al. paper](#), ones we discussed in [Tutorial 2](#), or (preferably, and for more fun) invent your own. For each of the ways of parametrizing sleep, provide instructive plots and estimate the values of the parameters. Compare the relative strengths and weaknesses of the sleep metrics you propose and give a discussion on which parametrization(s) you prefer.

Learning From Data

Yaser Abu-Mostafa, *Caltech*

<http://work.caltech.edu/telecourse>

Self-paced version



Homework # 2

All questions have multiple-choice answers ([a], [b], [c], ...). You can collaborate with others, but do not discuss the selected or excluded choices in the answers. You can consult books and notes, but not other people's solutions. Your solutions should be based on your own work. Definitions and notation follow the lectures.

Note about the homework

- The goal of the homework is to facilitate a deeper understanding of the course material. The questions are not designed to be puzzles with catchy answers. They are meant to make you roll up your sleeves, face uncertainties, and approach the problem from different angles.
- The problems range from easy to difficult, and from practical to theoretical. Some problems require running a full experiment to arrive at the answer.
- The answer may not be obvious or numerically close to one of the choices, but one (and only one) choice will be correct if you follow the instructions precisely in each problem. You are encouraged to explore the problem further by experimenting with variations on these instructions, for the learning benefit.
- You are also encouraged to take part in the forum

<http://book.caltech.edu/bookforum>

where there are many threads about each homework set. We hope that you will contribute to the discussion as well. Please follow the forum guidelines for posting answers (see the “BEFORE posting answers” announcement at the top there).

© 2012-2015 Yaser Abu-Mostafa. All rights reserved. No redistribution in any format. No translation or derivative products without written permission.

● Hoeffding Inequality

Run a computer simulation for flipping 1,000 virtual fair coins. Flip each coin independently 10 times. Focus on 3 coins as follows: c_1 is the first coin flipped, c_{rand} is a coin chosen randomly from the 1,000, and c_{min} is the coin which had the minimum frequency of heads (pick the earlier one in case of a tie). Let ν_1 , ν_{rand} , and ν_{min} be the *fraction* of heads obtained for the 3 respective coins out of the 10 tosses.

Run the experiment 100,000 times in order to get a full distribution of ν_1 , ν_{rand} , and ν_{min} (note that c_{rand} and c_{min} will change from run to run).

1. The average value of ν_{min} is closest to:

- [a] 0
- [b] 0.01
- [c] 0.1
- [d] 0.5
- [e] 0.67

2. Which coin(s) has a distribution of ν that satisfies the (single-bin) Hoeffding Inequality?

- [a] c_1 only
- [b] c_{rand} only
- [c] c_{min} only
- [d] c_1 and c_{rand}
- [e] c_{min} and c_{rand}

● Error and Noise

Consider the bin model for a hypothesis h that makes an error with probability μ in approximating a deterministic target function f (both h and f are binary functions). If we use the same h to approximate a noisy version of f given by:

$$P(y \mid \mathbf{x}) = \begin{cases} \lambda & y = f(x) \\ 1 - \lambda & y \neq f(x) \end{cases}$$

3. What is the probability of error that h makes in approximating y ? *Hint: Two wrongs can make a right!*

- [a] μ
- [b] λ
- [c] $1-\mu$
- [d] $(1-\lambda) * \mu + \lambda * (1-\mu)$
- [e] $(1-\lambda) * (1-\mu) + \lambda * \mu$

4. At what value of λ will the performance of h be independent of μ ?

- [a] 0
- [b] 0.5
- [c] $1/\sqrt{2}$
- [d] 1
- [e] No values of λ

● Linear Regression

In these problems, we will explore how Linear Regression for classification works. As with the Perceptron Learning Algorithm in Homework # 1, you will create your own target function f and data set \mathcal{D} . Take $d = 2$ so you can visualize the problem, and assume $\mathcal{X} = [-1, 1] \times [-1, 1]$ with uniform probability of picking each $\mathbf{x} \in \mathcal{X}$. In each run, choose a random line in the plane as your target function f (do this by taking two random, uniformly distributed points in $[-1, 1] \times [-1, 1]$ and taking the line passing through them), where one side of the line maps to $+1$ and the other maps to -1 . Choose the inputs \mathbf{x}_n of the data set as random points (uniformly in \mathcal{X}), and evaluate the target function on each \mathbf{x}_n to get the corresponding output y_n .

5. Take $N = 100$. Use Linear Regression to find g and evaluate E_{in} , the fraction of in-sample points which got classified incorrectly. Repeat the experiment 1000 times and take the average (keep the g 's as they will be used again in Problem 6). Which of the following values is closest to the average E_{in} ? (*Closest* is the option that makes the expression $|\text{your answer} - \text{given option}|$ closest to 0. Use this definition of *closest* here and throughout.)

- [a] 0
- [b] 0.001
- [c] 0.01
- [d] 0.1
- [e] 0.5

ChE101 Homework 4 Winter Term 2018

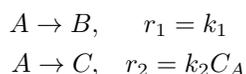
Distributed 24 January 2018

Due 2 February 2018 by 17:00 in class to Kari's lockbox (outside SPL 331)

For this homework, make sure you've read **Ch. 4** in Schmidt, as well as any material covered in previous homeworks. Please list all assumptions made to solve the problems and cite any sources used in solving the problems. Please show all your work, have your name on each page, write neatly, make your final answer clear, and start each problem on a new sheet of paper.

1) (25 points) **Selectivity and yield for parallel reactions** Adapted from Schmidt 4.31

The reactions



give 67% selectivity to the desired product B at 75% conversion of A in a CSTR with a residence time of 1 min and $C_{A0} = 2 \frac{\text{mol}}{\text{L}}$ in the feed.

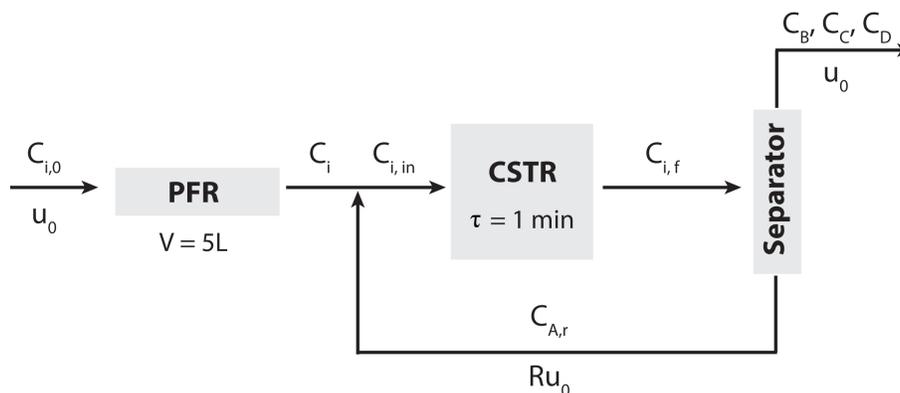
- (a) ***Before consulting your notes, or solving the problem***, indicate whether the CSTR or PFR (both operating at the conditions specified above) has higher:

Conversion of A
Selectivity for B
Yield of B

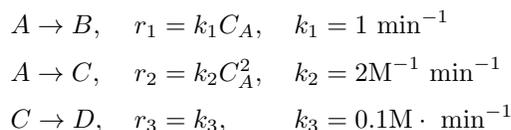
It's okay if your answer is incorrect! Just make a guess and defend it. (No credit if you do not defend your guess.)

- (b) Compute the selectivity and conversion for this reaction and feed in a PFR with $\tau = 1$ min.
- (c) What are the conversion and selectivity for the same PFR (i.e. same rate expressions and τ) if we change the feed concentration to $C_{A0} = 1 \frac{\text{mol}}{\text{L}}$?
- (d) What τ and conversion will give 99% selectivity in a CSTR? (Assume $C_{A0} = 1 \frac{\text{mol}}{\text{L}}$.)
- (e) ***After you complete the problem***, look back at your guess from Part (a).
- If any of your guesses were incorrect, write a short explanation of why and defend the correct answer.
 - If all your guesses were correct, devise a reaction scheme which reverses one of these trends (i.e. a CSTR is higher than a PFR, or vice versa).

2) (20 points) **Reactors in series, with undesired side reactions**



You are asked to operate the plant sketched above. The following reactions occur:



All reactions are in the aqueous phase. A pure feed stream of A ($C_{A0} = 1 \frac{\text{mol}}{\text{L}}$) enters a 5 L PFR at a volumetric flow rate of $u_0 = 10 \frac{\text{L}}{\text{min}}$. The effluent of the PFR is then fed to a CSTR. The CSTR has a recycle stream after the product stream is purged in a separator that is 100% efficient for separating B , C , and D out of the CSTR outlet stream.

- Find the conversion of A in the PFR and the selectivities and yields for B , C , and D in the PFR.
- Given a CSTR residence time of $\tau = 1 \text{ min}$ and a CSTR outlet concentration of $C_{A,f} = 0.15 \frac{\text{mol}}{\text{L}}$, find the recycle ration R , the recycle stream concentration $C_{A,r}$, and the CSTR volume V_{CSTR} .

3) (15 points) **Theory of reaction mechanisms**

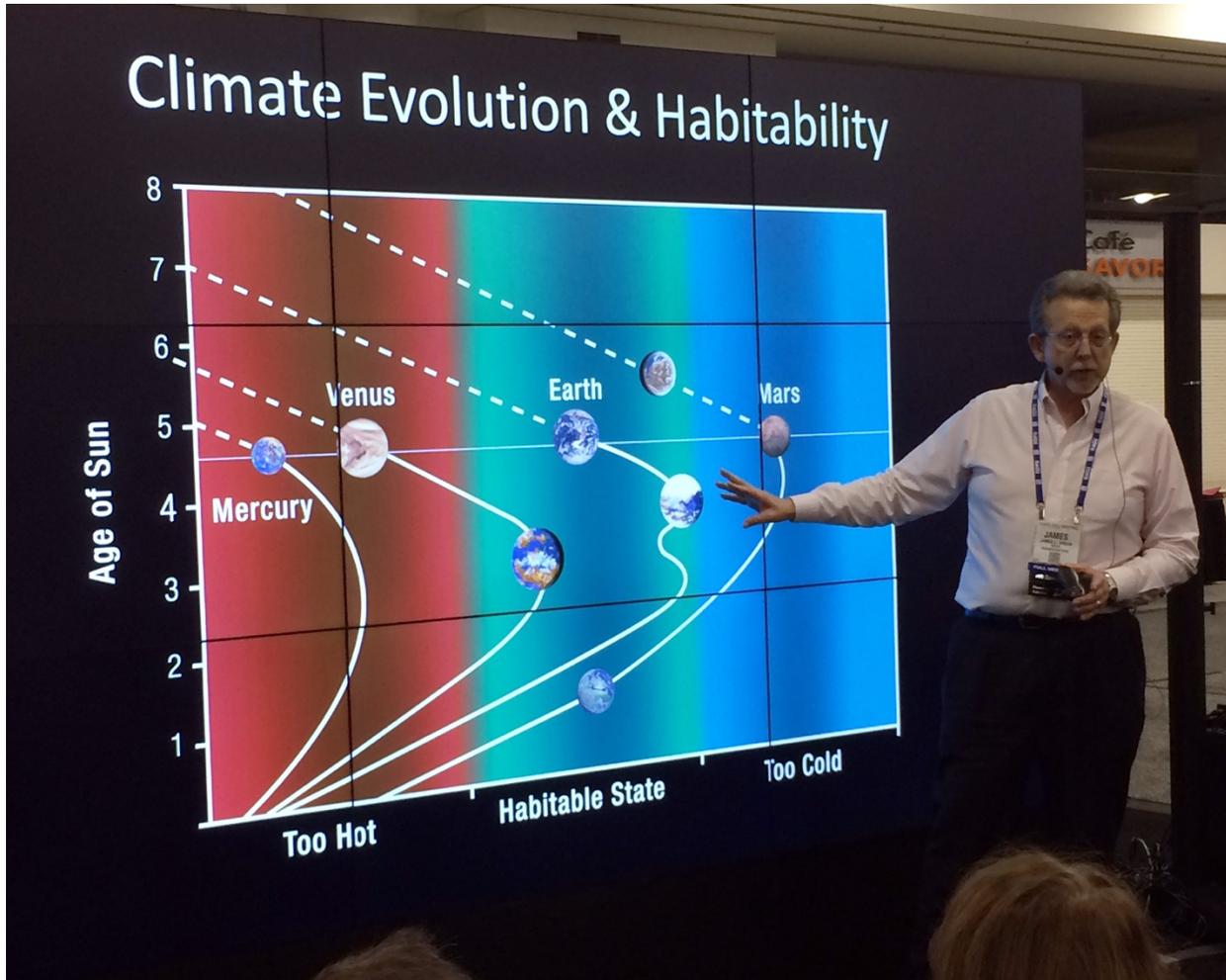
Explain (with at least one complete sentence) your intuition for each of the following. Feel free to include longer explanations or mathematical expressions if it is useful, but the goal is to make things we take for granted feel more obvious, by reminding ourselves where they come from.

- Why reaction rates depend on concentrations.
- Why we raise the concentration to the power of the stoichiometric coefficient in elementary reactions.
- Why the Arrhenius rate has an exponential dependence on activation energy.
- How unimolecular reactions occur, since a single molecule is unlikely to interact energetically with itself.

4) (15 points) **Mass action kinetics for other ODE models**

The possibility of life on Mars is a very exciting one for many astrobiologists (and most humans). Researchers found that methane in the Martian atmosphere varied seasonally.* This was surprising because (in the absence of microbes) the expected half life of methane in the atmosphere is on the order of

2. Habitability Through Time. Mike saw NASA administrator Jim Green give a talk about planetary habitability at the American Geophysical Meeting (AGU) last December. Check out the photo below.



- In general, all of the terrestrial planets start too hot for life, cool off, and then get warmer. **Briefly explain the physical phenomena that drive this trend.**
- Look at Mars's curve. About 1 billion years in to the Solar System's history, this graph places Mars firmly in a "habitable state." **Briefly state the geological evidence that contributes to this conclusion.**
- Earth's curve experiences wiggles that bring it to the brink of exiting a "habitable state" on the cold side. **What do these wiggles represent, and by what process did Earth fight off these cool states?**
- Today Mars hoards abundant water ice in its polar caps and subsurface. In the graph, Mars returns to a "habitable state" about 1 billion years into the future because this ice

is presumed to melt and sit stably on the surface of the globe. **Describe any caveats to this scenario that you can think of.**

- e) Although Venus is not tidally locked, its extremely slow rotation rate ($P_{rot} \sim 117$ Earth days) means its dayside faces the sun for months at a time. As shown in the graph, Venus can make a brief jaunt into a “habitable state” if it has reflective water clouds that pile up on the dayside, keeping the surface from overheating. However, this assumes that Venus still has any water by $\sim 2\text{--}3$ billion years into its history. **Briefly explain all the mechanisms by which Venus might’ve lost its water, taking note of the fact that there is no evidence that Venus ever had a planetary dynamo.**