General Principles of Human and Machine Learning:

Pop Quiz #4

Name: **Date:** **Grade:**

Questions:

1. How did Carroll's first function learning experiment go beyond previous stimulus-response learning tasks? How is learning a function more than just learning stimulus-response associations?

2. What is the "bias-variance" trade-off?

3. Name one disadvantage to similarity-based theories of function learning. Which phenomena that are characteristic of human learners do they fail to capture?

4. How is function learning useful in reinforcement learning settings?

5. Name two mechanisms of human exploration in settings with large state spaces, where not all states can be explored. Provide a brief explanation

6. What is the poverty of the stimulus argument? Do you think it is valid? Provide an argument either in support or against it.

7. Explain the principle behind the DYNA algorithm. What is the difference between learning online vs offline? Can you give an example of an efficient sampling method of experiences to be replayed?

8. Display different classes of RL algorithm (MF, MB, SR) within an efficiency/flexibility diagram - where flexibility is on the y-axis and efficiency on the x-axis.

9. Why is manifold analysis a useful tool for the analysis of big data sets? Can you give an example of manifold analysis, and briefly describe its principle?

10. Imagine an experiment in which we want to differentiate between regions of the brain that respond to odors and regions of the brain that respond to touch using Representational Similarity Analysis. In that experiment, participants are placed in an fMRI and receive 20 different types of touch on the hands and on the feet and smell 20 different types of odors. The stimuli are randomly presented, and a fan gets rid of any odor in between. What is the size of the representation dissimilarity matrix in that example? Give the formula of one entry of the RDM, or explain with words how to build it.

Answers key

1. Carroll showed that people can learn continuous mappings between stimuli and responses. Rather than learning discrete S-R associations, people learn functions, which correspond to a set of rules or programs, allowing for flexible generalization through interpolation and extrapolation

2. Bias and variance are two sources of error, and they complement one another (e.g., as bias goes down, variance goes up). This is due to overfitting, where more complex models have better training accuracy and lower bias, but tend to overfit the test data due to higher variance.

3. Similarity-based theories of function learning lack strong inductive biases, and fail to capture patterns of human extrapolation (e.g., tendency to prefer linear extrapolation)

4. Function learning is the basis for value function approximation in RL, where agents learn an (implicit) value function mapping states to value expectations. This allows for more flexible generalization and adaptive behavior in settings with vast or continuous state spaces

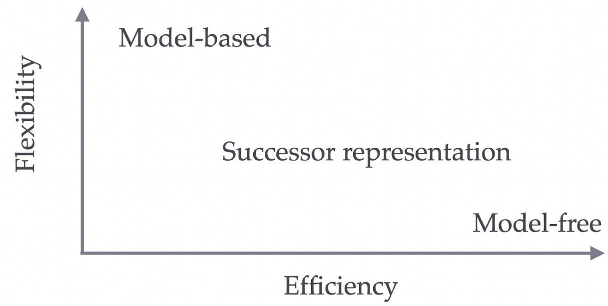
5. Mechanisms of human exploration
    a. Generalization (of past experiences onto novel settings). This is captured with a Gaussian Process model of function learning, mapping states onto expected rewards. *Optional*: the length-scale parameter lambda captures the extent to which people generalize (broadly vs. narrowly).
    b. (Uncertainty-) directed exploration. This is captured using upper confidence bound (UCB) sampling, where expected rewards are inflated by the subjective uncertainty (also can be described as a linear sum of reward expectations + beta*uncertainty). *Optional*: The exploration bonus beta captures the degree to which exploring uncertain regions of the search space is valued relative to exploiting high reward explorations
    c. Random exploration. This is captured using the temperature parameter of a softmax choice rule, allowing the agent to inject some random noise into its policy and potentially discover better actions by accident.

6. The poverty of the stimulus argument states that humans do not experience enough data to sufficiently learn language.
   a. Pros:
      i. The argument is based on there not being enough “positive” examples, which motivates the need to develop better hypotheses of how people generalize and can do knowledge transfer across contexts.
      ii. We still haven’t really developed artificial language models that can learn language with a human-level of experience
   b. Cons:
      i. Contextual learning largely negates the argument, since LSA (Landauer & Dumais) show that semantics can be learned based on the co-occurrence patterns of words
      ii. We simply haven’t figured out the more sophisticated learning mechanisms of human language learners that can account for how the degree of experience can yield fluent language. There is no need to assume something like Universal Grammar.

7. Dyna stores states, actions, future states and rewards during online experience, and uses offline periods to replay those experiences. This enables one to save ‘real experience’ in order to learn the value function/policy. Prioritised replay samples stored events based on their learning potential: typically, based on the TD error associated to the events replayed (such that events that would improve the behavior or the estimates the most would have higher chances to be selected). Other options are Importance Sampling (based on the probability distribution of the target policy compared to the behavior policy.

8. What in the SR compared to a traditional model-based approach makes it more efficient? In what aspects is it less flexible than a model-based approach?

   The SR contains cumulative information about discounted state occupancy. While a traditional model-based approach only has information about one time-step ahead in the markov chain. Thus, planning is very costly. As an SR agent has information about states that are many steps away, planning is more efficient.



9. Manifolds offer low-dimensional representation of a large dataset. One classical example is PCA, in which the principal components capture the axis of highest variance of the dataset. Thus, keeping only the main PCs enables us to cut the dimension of the dataset while preserving a high degree of variance/ separability between the points.

10. RSA allows us to look at similarity in stimulus-evoked responses. We would compute the correlation between the patterns of activity elicited by all stimuli. Thus, for areas that respond more to odors, their activity patterns should be more correlated within olfactory stimuli, and for areas that respond more to touch, their activity patterns should be more correlated within tactile stimuli.
What is the size of the representation dissimilarity matrix in that example?
It is 40*40 (number of stimuli * number of stimuli)

Give the formula of one entry of the RDM, or explain with words how to build it.
One entry of the matrix is 1-c[i,j], in which c[i,j] is the correlation between the patterns of activity evoked by the stimulus i and the patterns of activity evoked by the stimulus j.