Final Exam Review

CMPUT 366: Intelligent Systems

Weeks 1-13

Lecture Structure

1. Exam structure and details
2. Learning objectives walkthrough
   - **Clarifying questions** are the point of this class
3. Other questions, clarifications

USRI Survey

- USRI surveys close **today at midnight**
- About half of the class has filled them in
  - **Thanks**! *whoever you are...*
- Survey link: https://usri.srv.ualberta.ca/etw/ets/et.asp?nxappid=WCQ&nxmid=start

Final Exam Details

- The final exam is **Wednesday, April 24** at **2pm** in **CSC B-2**
    - Regular classroom, **one hour earlier** than usual
- There will be **3 hours** available for the exam
- You may bring a **single**, handwritten **cheat sheet** if you wish
- You may bring a non-programmable calculator if you wish
- **All course material** is included
    - Weeks 8-13 will be more heavily weighted than weeks 1-7

Final Exam Structure

- There will be **120 marks** total
- There will be **15 short answer** questions with 1-2 sentence answers
  - The rest will be more in-depth
- There will be **no coding** questions
  - But you may be asked to **execute a few steps** of an algorithm
- Every question will be based on the **learning objectives** that we are about to walk through

Introduction to AI

- characterize simplifying assumptions made in building AI systems
- determine what simplifying assumptions particular AI systems are making
- suggest what assumptions to lift to build a more intelligent system than an existing one
- define the major representational dimensions
- classify problem statements by representational dimensions

Search

- define a directed graph
- represent a problem as a state-space graph
- explain how a generic searching algorithm works

Search (2)

- demonstrate how depth-first search will work on a graph
- demonstrate how breadth-first search will work on a graph
- demonstrate how iterative deepening DFS will work
- demonstrate how least cost first search will work on a graph
- predict the space and time requirements for depth-first and breadth-first searches

Search (3)

- devise a useful heuristic function for a problem
- demonstrate how best-first and A* search will work on a graph
- predict the space and time requirements for best-first and A* search
- justify why and when depth-bounded search is useful
- demonstrate how iterative-deepening works for a particular problem
- demonstrate how depth-first branch-and-bound works for a particular problem

Search (4)

- define hill climbing, random step, random restart
- explain why hill climbining is not complete
- explain why adding random restarts to hill climbing makes it complete
- justify when local search is appropriate for a given problem

Search (5)

- list the elements of a local search problem
- recognize a local search problem
- explain how the generic local search algorithm works
- define hill climbing and stochastic local search
- trace an execution of hill-climbing and stochastic local search
- define improving step, random step, and random restart
- explain the benefits of random steps and random restarts

Uncertainty

- define a belief network
- build a belief network for a domain
- build a correct belief network for a given joint distribution
- compute marginal and conditional probabilities from a joint distribution

Uncertainty (2)

- define a random variable
- describe the semantics of probability
- apply the chain rules
- apply Bayes' theorem

Uncertainty (3)

- define the factor objects and factor operations used in variable elimination
- explain the origins of the efficiency improvements of variable elimination
- define the high-level steps of variable elimination
- trace an execution of variable elimination

Uncertainty (4)

- justify why a belief network is a correct encoding of a joint distribution
- identify the factorization of a joint distribution encoded by a belief network
- answer queries about independence based on a belief network
- answer queries about independence based on a joint distribution

Causality

- define observational and causal query
  - explain the difference
- explain why causal queries on observational distributions can go wrong
- construct the post-intervention distribution for a causal query from an observational distribution
- evaluate a causal query given an observational distribution
- justify whether a causal model is valid

Causality (2)

- define a back-door path
- identify a back-door path
- define the back-door criterion
- identify whether a causal query is identifiable from a partially-observable causal model

Supervised Learning

- define supervised learning task, classification, regression, loss function
- represent categorical target values in multiple ways (indicator variables, indexes)
- identify an appropriate loss function for different tasks
- explain why a separate test set estimates generalization performance
- define 0/1 error, absolute error, (log-)likelihood loss, mean squared error, worst-case error

Supervised Learning (2)

- define generalization performance
- construct a decision tree using given features, splitting conditions, and stopping conditions
- define overfitting
- explain how to avoid overfitting

Supervised Learning (3)

- explain how to use the Beta and Bernoulli distributions for Bayesian learning
- derive the posterior probability of a model using Bayes' rule
- define conjugate prior
- demonstrate model averaging

Supervised Learning (4)

- estimate expectations from a finite sample
- apply Hoeffding's inequality to derive PAC bounds for given quantities
- demonstrate the use of rejection sampling and importance sampling

Deep Learning

- define an activation
- define a rectified linear unit and give an expression for its value
- describe how the units in a feedforward network are connected
- give an expression in matrix notation for a layer of a feedforward network
- explain at a high level what the Universal Approximation Theorem means
- explain at a high level how feedforward neural networks are trained
- identify the parameters of a feedforward neural network

Deep Learning (2)

- define sparse interactions and parameter sharing
- define the convolution operation
- define the pooling operation
- explain why convolutional networks are more efficient to train
- describe how the units/layers in a convolutional neural network are connected

Deep Learning (3)

- demonstrate unfolding a recurrent expression
- explain the problems with handling sequence input using dense or convolutional multi-layer neural networks
- describe the high-level idea behind recurrent neural networks

Reinforcement Learning

- define a Markov decision process and a policy
- define and give expressions for the **state-value function** and the **action-value function**
- state the Bellman optimality equations
- define returns and give expressions for episodic and discounted continuing returns
- represent a problem as a Markov decision process

Reinforcement Learning (2)

- justify why a policy is weakly better than another
- trace an execution of iterative policy evaluation
- state the **Policy Improvement Theorem** and explain **why it is important**
- trace an execution of the Value Iteration algorithm

Reinforcement Learning (3)

- explain how Monte Carlo estimation for state values works
- trace an execution of the first-visit Monte Carlo Prediction algorithm
- explain the difference between prediction and control

Reinforcement Learning (4)

- define on-policy vs. off-policy learning
- define a behaviour policy
- explain what exploring starts are and why they are necessary
- define an ε-soft policy
- explain when and why ε-soft policies are useful

Reinforcement Learning (5)

- trace an execution of the TD(0) algorithm
- trace an execution of the Q-learning algorithm
- trace an execution of the Sarsa algorithm
- define bootstrapping
- explain why bootstrapping is useful

Reinforcement Learning (6)

- explain why function approximation is useful
- explain the difference between action-value and policy gradient methods
- trace an execution of the REINFORCE algorithm

Multiagent Systems

- define **best response** and **Nash equilibrium**
- define Pareto dominance and Pareto optimality
- explain the difference between pure strategy and mixed strategy Nash equilibria

Multiagent Systems (2)

- trace an execution of backward induction
- explain the difference between imperfect information and perfect information extensive form games
- define an information set
- identify the pure strategies in an extensive form game

Multiagent Systems (3)

- define a **maxmin strategy**
- identify a maxmin strategy in an extensive form game
- define a zero-sum game
- state the Minimax Theorem and **explain its implications**
- trace an execution of the alpha-beta search algorithm
- define a multiagent mechanism

Questions?