COSC343: Artificial Intelligence

Lecture 25: Revision Lecture

Alistair Knott

Dept. of Computer Science, University of Otago

Other AI-related courses on offer in the dept

Next semester:

- COSC421: Neural models of language (Ali)

Next year (Semester 1):

- COSC420: Neural Networks/Machine Learning (CS Dept Allstars)
- COSC450: Computer Vision and Graphics (Steven Mills)

Format of the exam

There will be 6 questions in total, and you have to answer 5.

- 3 questions are on Lech's part
- 3 questions are on my part.

- Each question is worth **12 marks**.
- There are **60 marks in total** in the exam.

You should bring a (non-programmable) calculator, for probability questions.

You're not allowed to bring anything else.

Some hints on study technique

- There will be a mixture of technical (quantitative) questions and short answer questions.
- You should revise all the lecture notes, and all the tutorial questions.
- You should look at past exam papers. (You can ignore questions on logic, Prolog, lambda calculus, expert systems. New topics this year: kernel methods, convolutional networks, Elman networks.)

Some quick hints on exam technique:

- Read the questions properly!
- Tailor your answer to the number of marks a question is worth!
- Define any terms you use!
- Use examples to illustrate!
- Reread your answers!

Core AI concepts

- Definitions of AI; the Turing test
- Intelligent agents
- Different agent architectures (reflex agents, goal-based agents, utility-based agents, learning agents)

Embodied AI and robots

- Difficulty of making AI agents operate in the real world
- Reflex agents, Braitenburg vehicles
- Agents and their environments: complex emergent behaviours
- Brooks' subsumption architecture
- LEGO robots; multithreaded NXC

Machine learning

- The concept of a hypothesis space
- Polynomial functions
- The concept of an error surface
- Linear regression
- Training sets, testing sets, overfitting

Probabilistic reasoning

- Basic concepts: sample space, sample points, full joint probability distribution, event, random variable
- Estimating probabilities from relative frequencies
- Conditional probability: definition in terms of prior probabilities.
- The chain rule: rewriting $P(a \wedge b \ldots \wedge n)$ as a product involving a prior probability and a set of conditional probabilities
- Normalisation: $P(a|b) + P(\neg a|b) = 1$
- Independence and conditional independence
- Bayes' rule (expressed with single probabilities and whole distributions)
- Naive Bayes models

Machine learning

- Classification vs regression
- Decision trees: how to use them; how to build them
- Using information theory to decide which features carry most information about an item's class
- Learning curves: the effect of training set size

Neural networks

- Real neurons
- Simple perceptrons with threshold activation functions
- The perceptron learning rule (a.k.a. delta rule)
- Linear separability
- Encoding Boolean values using 1 and 0
- The XOR problem
- Epochs of training

Neural networks

- Sigmoid perceptrons
- Error surfaces, and the gradient descent learning rule
  (You don't need to know how to calculate partial gradients!)
- The problem of local minima for gradient descent
- Multi-layer perceptrons: extra expressiveness
- The back-propagation rule: why it's needed; where it comes from
  (You don't need to be able to derive it!)
- Convolutional neural networks

Unsupervised learning

- Supervised vs unsupervised learning methods
- Principal component analysis
- Clustering (esp. k-means clustering)
- Hierarchical clustering techniques, and dendrograms
- Autoencoder networks
- 'Deep' networks, mixing unsupervised and supervised learning

Optimisation and Genetic Algorithms

- Ways of solving optimisation problems: analytical, exhaustive search, gradient-based methods
- Hill-climbing, stochastic hill-climbing
- Simulated annealing
- Biological evolution by natural selection
- Genetic algorithms

State space search

- The concept of a state space
- Uninformed search methods: depth-first, breadth-first, uniform-cost, depth-limited, iterative deepening
- Pseudocode for different search algorithms
- Time and space complexity of these different methods
- Tree search and graph search
- Heuristic search: greedy search, A* search
- Examples of heuristics: straight-line distance; Manhattan distance
- Optimistic (=admissible) heuristics; consistent heuristics
- Relaxed problems
- Adversarial search: Minimax; alpha-beta pruning

Natural language

- Language as a kind of action
- Compositionality of language: (the meaning of a sentence is formed from the meanings of its words)
- Well-formed and ill-formed sentences
- Hierarchical structure in sentences
- Phrase structure grammars: how you can use them to build parse trees for sentences
- Agreement (between subjects and verbs; between determiners and nouns)
- Grammars using variables to capture agreement constraints
- Parsing as a kind of search
- Syntactic ambiguity: how to represent it, and how to disambiguate

Natural language

- Probabilistic language models
- Bayesian text classification
- Linear language models: *n*-grams (& the sparse data problem)
- Uses of *n*-gram models: word-sense disambiguation, speech interpretation, predictive texting
- Probabilistic grammars: how they solve the problem of spurious ambiguities in wide-coverage grammars
- How to derive a probabilistic grammar from a hand-parsed corpus; how to use it to disambiguate.
- Simple recurrent networks: their architecture and training; how they can learn grammatical categories, and word boundaries without supervision