Artificial Intelligence

If you are taking the **Breadth Exam:**
Answer any four (4) out of the six Basic questions B1-B6.

If you are taking the **Depth Exam:**
Answer any four (4) out of the six Basic questions B1-B6 and
answer any four (4) out of the 12 Advanced Depth questions D1-D12.

Breadth (Basic) Questions

B1. In **best-first search**, one uses a heuristic estimate, h', of the actual least cost, h, of getting from a node to a goal. If h' never overestimates h, the best-first search (or $A*$) algorithm is guaranteed to produce an optimal path to a goal. Explain why this is true, and give an explicit finite graph where h' does overestimate h and the algorithm produces a path which is not optimal.

B2. Explain the use of context-free grammars in **parsing natural languages**. In particular, give such a grammar that produces two distinct parses for each of the following two sentences:
Paint the house with green paint.
Paint the house with the attached garage
For each sentence, explain which parses correspond to a meaningful semantic interpretation. Explain why the two different uses of the word "paint" in the first sentence does *not* produce any ambiguities here.

B3. Convert the following three sentences to predicate logic, and then use **resolution** to show that the third one follows from the first two:
- If the AI exam is easy, some students are happy
- If the AI exam contains a Skolemization question, no students are happy.
- If the AI exam contains a Skolemization question, then the AI exam is not easy.

B4. Recent developments in artificial intelligence research have tended away from earlier dependence on the Von Neumann computer architecture as a feasible basis on which to build **models of intelligence**. There are at least three problems raised with respect to this architecture, viz., biological plausibility, serial processing, and the requirement for explicit programming. For each of these problems, explain why it is a problem and show how the alternative architecture (or architectures) being considered purport to overcome it.

B5. One of the acclaimed strengths of **connectionism** is that it can use distributed representations. Describe the concept of a distributed representation. Is there anything analogous in symbolic approaches to knowledge representation?

Occasionally, distributed *output* representations are used when training a neural network with Rumelhart's backpropagation algorithm. Propose and briefly justify a distributed output representation for a system that learns to classify (isolated) handwritten digits. How would the digit "1" be represented in your output representation?

B6. Consider formalizing, using the **STRIPS representation** scheme, a simple robot-controlled warehouse. Assume the warehouse stores two kinds of boxes: strong and weak. Strong boxes can support other boxes, while weak ones cannot. There can be any number of robots in the warehouse, though some may be broken at any given time. Finally, for simplicity, assume there always is free space on the warehouse floor.

Write STRIPS operators for *store-on-floor(?robot, ?box)*, *store-on-box(?robot, ?box)*, and *retrieve(?robot, ?box)*. (*?robot* and *?box* are variables.)

- For *store-on-floor*, *?robot* should move *?box* from the receiving dock to the warehouse floor.
- For *store-on-box*, *?robot* should move *?box* from the receiving dock to some stable location on top of another box (if possible).
- For *retrieve*, *?robot* should bring *?box* to the shipping dock.

Verbally describe (but do not formalize) any other operators you deem necessary. Also, briefly define any predicates you introduce.

Briefly state one aspect of a "real world", robot-controlled warehouse that is not easy to capture in the STRIPS formalization.

Advanced (Depth) Questions

D1. **Object recognition** in computer vision can be formulated as a search problem where the search space is either of local feature correspondences or of poses (i.e., positions and orientations).

(a) One difference between these two spaces is that one is discrete while the other is continuous. What implications does this have for the search process? Give your answer in terms of advantages and disadvantages of searching each of these two spaces, given this difference.

(b) The Generalized Hough Transform could be considered as performing search over both the correspondence space and the pose space. Briefly explain why after outlining how the technique works.

(c) Say we use the Hough transform as a preprocessing step before searching the correspondence space using some other method. Explain how and why this might be a good or bad idea.

D2. One method for three-dimensional shape representation in computer vision is the **aspect graph**.

(a) How is this graph defined for a polyhedral object? Include a description of what the nodes and edges correspond to, and how the pose of an object is related to this structure.

(b) Show the aspect graph for a torus (i.e., a donut-like shape containing a smooth surface, not a polyhedral approximation).

(c) Briefly describe how the aspect graph could be used in a parallel, object recognition algorithm. Assume you have the aspect graph of a single object and the goal is to detect if that object exists in a given image. Does your algorithm determine the exact pose of the object it detects? Explain how or why not.

D3. Micro-electrode studies with monkeys have found individual neurons that fire selectively to different faces or other complex objects within 70 to 200 milliseconds, while it takes roughly 1 to 2 milliseconds for one neuron to fire into and excite another.

A. Succinctly describe the **primate visual system**, emphasizing its important structures (such as neurons, rods and cones, retina, cortical areas, columns, hypercolumns) and how they are linked.

B. Describe what you think is the most appropriate structure of processes to embed into this structure in order to recognize such complex objects within only 35 to 200 serial steps. Justify your choices in terms of what is known about brains where you can.

D4. You are to develop a computer **vision system** that can recognize a small number, like 5 or 10, different object-classes (e.g., chair, table, couch, lamp, vase) that will be input in 3 different ways:

A. as single frames from a black-white television camera,

B. as line drawings,

C. as either of the above, but converted into a dotted picture, but where the objects, although dots and patches, remain easily recognizable by human beings.

Describe what you think are the two most promising approaches to explore. Make clear why you chose these two approaches. Make sure to explain how you would try to handle the dotted images with each, and also the many different possible types of each object class (that is, the intra-object differences).

D5. Convert the following sentences to clauses and derive the empty clause, $\square$, using **paramodulation** and **binary resolution**.

$a \neq 0$

$\neg(0 < a * a)$

$\forall x(x = x)$

$\forall x((0 < x) \vee (x = 0) \vee (x < 0))$

$\forall x \forall y(((0 < x) \wedge (0 < y)) \Rightarrow (0 < x * y))$

$\forall x(x < 0 \Rightarrow 0 < m(x))$

$\forall x \forall y(x * y = m(x) * m(y))$

Here, x, y, are variables and $a, 0$ are constants.

D6. In **Prolog**, consider the program P:

```
p(a,b).
p(b,c).
anc(X,Y) :- p(X,Y).
anc(X,Z) :- p(X,Y), anc(Y,Z).
```

With this program, the queries, `?- anc(a,c)` and `?- not anc(b,a)` will both get `yes` answers. Explain why both $anc(a, c)$ and $\neg anc(b, a)$ are logical consequences of the Clark completion of P.

D7. Approximately 50% of English text includes **metaphors** of some sort. Discuss ways in which a natural language processing system might detect and interpret metaphors in the course of processing large quantities of text of varied subject matter.

D8. Discuss the use of syntactic and semantic features in a **natural language processing** system. Include a discussion of inheritance.

D9. What is meant by "**common sense knowledge**"? Give examples. Contrast with other kinds of knowledge. Why is it argued that expert systems can not achieve their full potential unless and until they come to "possess" common sense? In particular, specify some typical expert system deficiencies that might thereby be overcome. Briefly describe some of the research (e.g., the CYC project led by Lenat) that bears on this problem and evaluate its likelihood of success.

D10. Outline the essential capabilities of one "**knowledge representation**" language (e.g., KRL, KL-ONE, LOOPS, CLOS); contrast them with the capabilities of traditional programming languages (e.g., LISP, C); and contrast them with the capabilities of more recent "object-oriented" programming languages (e.g., Object Pascal, C++, Modula III). Briefly discuss the problems that have arisen in attempts to implement AI knowledge representation languages. How, if at all, will parallel architectures contribute to a solution of these problems?

D11. Like most AI programs, **inductive learning** algorithms search a space. Describe the space searched by (a) Quinlan's ID3 algorithm and (b) Rumelhart's backpropagation algorithm (i.e, what information characterizes a node?).

Compare and contrast the two spaces searched. Several investigators have found only minor differences in generalization performance when comparing ID3 to backpropagation on common problems. Discuss this result in light of the spaces searched.

Which search-control strategy does each method follow? Briefly describe any evaluation functions used. How would each algorithm change if beam search were used instead (or is this impossible)?

D12. Assume you have an **inductive learning** algorithm LA, along with N positive examples and M negative examples of some concept you wish to learn. Consider the following three situations.

a. You wish to estimate the future accuracy of LA on the concept represented by the training examples. Describe and justify an experimental methodology for estimating this.
b. LA has p "free" parameters (one such parameter might be the number of hidden units in a neural network). Describe and justify an experimental methodology for deciding how to set these parameters for your sample problem.
c. Someone claims algorithm LA' works better than LA. How could you empirically evaluate this claim?