Teacher Answer Key

Case Study

1. The chatbot was trained primarily on data from younger patients, making it reliable for common presentations in that demographic but unreliable for Mr. Chen, an older patient where the same symptoms have different statistical implications. Deploying the system on all incoming patients regardless of age put it outside its effective training domain, which is exactly the risk the reading warns about with narrow AI.

2. Narrow AI generates confident outputs based on statistical patterns in its training data. The chatbot had seen many cases of musculoskeletal chest pain and learned to recognize that pattern confidently — but it had not been trained on sufficient data to recognize when the same symptoms in an older patient carry a different risk profile. Confident output indicates pattern-matching, not genuine clinical understanding.

Multiple Choice

1. D
2. C
3. A

Matching

1-E, 2-A, 3-D, 4-C, 5-B

Fill-in-the-Blank

1) chess, 2) training, 3) scale

Short Answer (Sample Responses)

Q1: AlphaGo is impressive because it applied massive computational resources to a single well-defined task and achieved a level of performance no human can match. It is narrow because all of that capability is confined to Go — the system has no ability to transfer its strategic reasoning to any other problem. The concentration of resources on one task is both the source of its strength and the boundary of its usefulness.

Q2: Words like 'reasoning' and 'understanding' imply cognitive processes similar to human thinking, but the accurate description is 'statistical pattern matching on training data.' The reading suggests these systems generate plausible outputs by identifying correlations in the text they were trained on, not by building internal

models of reality the way humans do. The marketing language gap between 'reasoning' and 'pattern matching' misleads users into trusting AI outputs more than warranted.

Q3: A medical AI tool should only be trusted within the specific population and condition types its training data covered. A careful user would ask: what patient population was used to train this tool, does that match my demographics, what conditions was it validated for, and has it been clinically tested against outcomes in patients like me. These questions matter because, as Mr. Chen's case shows, a system can be reliable for one population and dangerous for another.