Name: _____   Date: _____

Answer Key: 8th Grade AI Logic Breach Quiz

Deconstruct 10 complex scenarios involving algorithmic bias, heuristic search architectures, and the ethical dilemmas of reinforcement learning.

---

1. In the context of the 'Trolley Problem' applied to autonomous vehicles, which ethical framework focuses on minimizing total harm by prioritizing the lives of the many over the few?

Answer: C) Utilitarianism

> Utilitarianism is an ethical theory that determines right from wrong by focusing on outcomes, specifically aiming for the 'greatest good for the greatest number.'

2. Large Language Models (LLMs) possess 'Artificial General Intelligence' (AGI) because they can reason across multiple domains with human-level consciousness.

Answer: B) False

> Current LLMs are 'Narrow AI'; they predict tokens based on statistical patterns and do not possess actual consciousness or General Intelligence (AGI).

3. Which term describes the 'invisible' process where an AI model identifies patterns in data that humans cannot explain or trace, often leading to transparency issues?

Answer: B) Black Box Effect

> The 'Black Box' problem refers to AI systems where the internal workings or decision-making processes are not interpretable by humans.

4. An AI is trained to recognize pneumonia in X-rays but fails when used at a different hospital because it learned to identify a specific hospital's watermark instead of medical features. This is an example of:

Answer: A) Overfitting

> Overfitting occurs when a model learns the 'noise' or specific details of the training data (like a watermark) rather than the actual concept it is meant to identify.

5. In Reinforcement Learning, the system learns through a process of 'Trial and Error' by maximizing which specific variable?

Answer: B) The Reward Signal

Reinforcement Learning agents are programmed to take actions in an environment to maximize a numerical reward signal over time.

6. If an AI hiring tool systematically rejects resumes with feminine-coded words because it was trained on 20 years of male-dominated hiring data, the model is suffering from _____.

Answer: B) Algorithmic Bias

Algorithmic bias occurs when human prejudices or historical inequities are reflected in the training data and then automated by the AI.

7. A 'Generative Adversarial Network' (GAN) consists of two neural networks—a generator and a discriminator—that compete against each other to create realistic data.

Answer: A) True

GANs use a competitive architecture where the generator creates data and the discriminator tries to tell if it is real or fake, improving the quality of the output.

8. Which of the following is the most significant challenge regarding the 'alignment problem' in AI development?

Answer: B) Ensuring an AI's goals match human values and intentions

The alignment problem is the difficult task of ensuring that an AI system's objectives and behaviors are perfectly aligned with human morality and goals.

9. When an AI hallucination occurs, what is actually happening within the large language model?

Answer: C) The model generates a statistically likely but factually wrong output

Hallucinations occur because AI models process probability rather than truth; they generate the most likely 'next word' even if that word is factually incorrect.

10. Symbolic AI (GOFAI) relies on pre-programmed 'if-then' rules rather than learning from massive datasets like modern Machine Learning.

Answer: A) True

Symbolic AI, or 'Good Old Fashioned AI,' is based on high-level human-readable representations and logic rather than statistical processing of big data.