

1. What is AI? (L0 p.17; Ch1 p.2–9)

AI $\supset$ ML $\supset$ DL $\supset$ GenAI ($\supset$ Agentic AI). AI $\neq$ learning: search & expert systems are AI with no training. L0 p.17

4 approaches = {thinking, acting} $\times$ {humanly, rationally}. Textbook adopts acting rationally (rational agent). “AI” coined 1955. p.2

Turing test: interrogator can't tell human vs. computer from written answers. Needs NLP, knowledge rep., reasoning, ML (+ vision, robotics). p.3

Thinking humanly = cognitive modeling: introspection, psych experiments, brain imaging $\rightarrow$ cognitive science. p.4

Thinking rationally (logistic). 2 obstacles: informal knowledge hard to formalize; solvable “in principle” $\neq$ solvable in practice. p.5

Why acting rationally: more general than laws of thought (inference = 1 of many ways); more amenable to scientific development. p.6

Foundations: philosophy, math, economics (maximize payoff), neuroscience, psychology, comp. eng., control theory, linguistics. p.7–9

2. History of AI (Ch1 p.10–26)

Turing Awards: '69/'71 reasoning; '75 symbolic; '94 expert systems; '11 probabilistic (Pearl); '19 deep learning; '24 RL (Barto, Sutton). p.10

1943 McCulloch & Pitts: artificial neuron; any computable function via neuron networks. 1949 Hebb: weight-update rule. 1950 first NN computer. p.11

1956 Dartmouth (McCarthy, Minsky, Shannon, Rochester) = birth of AI; neuron nets, self-improvement, creativity. p.12–13

GPS (Newell & Simon): probably the first program to embody “thinking humanly”. p.14

Physical symbol system hypothesis (1976): any intelligent system must operate by manipulating symbol data structures. p.14

Microworlds (Minsky): limited domains that seem to require intelligence; most famous = blocks world. p.16–17

1966–73 dose of reality: overconfidence from toy examples; MT failure (“vodka is good but meat is rotten”); root cause = scalability. p.18

Scalability: finding a solution “in principle” $\neq$ having the mechanisms to find it in practice. p.18

Weak methods: general but don't scale $\rightarrow$ expert systems use powerful domain-specific knowledge in narrow areas. p.19

Expert systems (1969–86): DENDRAL (molecular structure); MYCIN (~450 rules, better than junior doctors). p.19–20

AI Winter: industry few \$M (1980) $\rightarrow$ billions (1988); companies failed to deliver on extravagant promises. p.21

Back-propagation: 1969 (Bryson & Ho), reinvented mid-1980s by ≥ 4 groups. Connectionist vs. symbolic (Newell & Simon) vs. logicist (McCarthy). p.22

1987–: probability not Boolean logic; ML not hand-coding; experiments not philosophy; shared benchmarks. p.23

HMMs dominated speech recognition: rigorous math, trained on large corpora; no claim humans use HMMs. p.23–24

Bayesian networks (Pearl 1988): formalism for uncertain knowledge + probabilistic reasoning. p.24

RL: learn by trial and error from rewards; Sutton (1988) linked RL to Markov decision processes. p.24

Big data (2001–): ImageNet sparked CV revolution; IBM Watson won Jeopardy! (2011). p.25

Deep learning: ML with multiple layers of simple adjustable units. 2012 ImageNet (Hinton): big data + GPU + new training techniques. p.26

CNN: small learned filters slide over the image; edges $\rightarrow$ parts $\rightarrow$ objects; handwritten digits (1990s). AlphaGo (2015–16) = RL + DL. p.26

3. Risks & Benefits (Ch1 p.30)

★★ Benefits: free humans from repetitive work, boost production. Risks: lethal autonomous weapons; surveillance & persuasion; biased decision making; impact on employment; safety-critical applications; cybersecurity. p.30

4. Intelligent Agents (Ch2 p.2–12, 17)

Agent: anything perceiving its environment through sensors and acting upon it through actuators. p.2

Percept: input at an instant. Percept sequence: complete history. Agent function: maps any percept sequence $\rightarrow$ action. p.3

Vacuum world: squares A, B; Left, Right, Suck, NoOp. Simple agent: if dirty then Suck else move. p.3–4

Table-driven agent: doomed to fail — table grows with sequence length. p.17

Performance measure evaluates a sequence of environment states; design by what you want in the environment, not how the agent should act. p.5

Rational agent: for each percept sequence, choose the action expected to maximize the performance measure, given percepts + built-in knowledge. p.6

Rationality maximizes expected performance; perfection maximizes actual. Rationality $\neq$ omniscience. p.7

Autonomy: learn what it can to compensate for partial or incorrect prior knowledge. p.8

PEAS = Performance, Environment, Actuators, Sensors. Taxi: P safe/fast/cheap, E roads, A steering, S cameras. p.9–10

Fully observable: sensors detect all aspects relevant to the choice of action (chess); taxi = partially. p.11

Multiagent if B maximizes a performance measure depending on A's behavior; chess = competitive. p.11

Deterministic: next state fully determined by current state + action; else stochastic (taxi). p.12

DRY RUN — not the real exam sheet. Purpose: check the “A4 行” column can be used as-is. W1 alone fills $\approx \frac{1}{2}$ page (measured: 1 page, content ends at 54% height); 18 weeks will NOT fit, so the real sheet must filter by 我 懂了嗎 $\neq$  and 考試訊號 $\in \{\text{🌀}, \text{★}, \text{★}, \text{★}\}$. Grey lines = no exam signal. Whether an A4 sheet is allowed at all is **unconfirmed**.