

Method	Learn	Fill-in
BC	Policy $\pi(a s)$ directly	A
DMP	Trajectory (attractor sys.)	C
IRL	Reward $R(s)$	B

Behavioral Cloning

Supervised learning: $\pi_\theta(s) \approx \pi^*(s)$.

► **Compounding error** [PS3 Q5a]: small mistakes \rightarrow unseen states \rightarrow more errors \rightarrow crash. Error grows $O(\varepsilon T^2)$.

► **DAgger**: run learned π , query expert at visited states, re-train.

BC vs IRL [PS3 Q5b: D, Q5c: C]

BC copies actions directly (supervised). Assumes expert optimal.

IRL infers reward function, then optimizes policy via RL. IRL more robust, transfers better.

► Q5c: BC assumes expert actions are optimal; IRL relaxes this.

DMPs (Dynamic Movement Primitives)

$\tau \dot{v} = K(g-x) - Dv + (g-x_0)f(s)$, $\tau \dot{s} = -\alpha s$ ($s: 1 \rightarrow 0$).

$f(s) = \sum w_i \psi_i(s)s$. As $s \rightarrow 0$: $f \rightarrow 0 \rightarrow$ spring-damper converges to g .

Weights learned via linear regression. Temporal/spatial invariance.

Imitation Challenges [Prac Q3c — 4 pts]

1. Compounding error / distribution shift (see above).

2. Multimodal demonstrations: different demos show different strategies for same state (e.g., go left OR right around obstacle). Unimodal policy (Gaussian) \rightarrow averages modes \rightarrow dangerous (goes straight into obstacle).

Solutions: GMMS, Diffusion Policy (generative, handles multiple modes).

Monte Carlo [Prac Q2b fill-in]

Estimate via random sampling (E) + statistical analysis (F).

CONVOLUTION [Prac Q1c]

Slide kernel over image, element-wise multiply, sum \rightarrow 1 output pixel.

$$h[m, n] = \sum_{k, l} g[k, l] f[m+k, n+l]. \quad \text{Output:}$$

$(W-K+1) \times (H-K+1)$.

Box $\frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightarrow$ blur. Sobel V: $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & -1 \end{pmatrix}$ detects vertical edges.

Sobel H: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ -1 & -2 & -1 \end{pmatrix}$ detects horizontal edges.

Sharpening: 2x original \rightarrow blurred \rightarrow amplifies edges.

Gaussian: $G_\sigma = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$, smoother blur than box.

► Prac Q1c kernel: negative top, positive bottom \rightarrow horizontal edge detect (Sobel-H variant) \rightarrow answer shows horizontal edges.

Conv ex: 3x3 image with Sobel-H center pixel:

$$\sigma = (-1)(a_{00}) + (-2)(a_{01}) + (-1)(a_{02}) + (1)(a_{20}) + (2)(a_{21}) + (1)(a_{22})$$

Large $|\sigma| \rightarrow$ strong horizontal edge. $\sigma > 0 \rightarrow$ dark-to-light going down.

VI vs PI COMPARISON [PS3 Q4f]

	Value Iteration	Policy Iteration
Updates	V values only	π and V
Per-iter	$O(S ^2 A)$	$O(S ^3 + S ^2 A)$
# Iters	Many (slow conv.)	Few (fast conv.)
Converges to	V^* directly	π^* directly
Simplicity	Simpler	More complex
In practice	—	Often faster
Early stop	Yes (monitor V)	No (need full eval)

► VI: one Bellman backup per iteration (approx.).
PI: exact policy eval + greedy improvement, guaranteed to converge in finite steps (finite # policies).

ROTATION & TRIG REFERENCE

2D Rotation Matrix

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{Rotates point by } \theta \text{ CCW.}$$

$$R^{-1} = R^T = R(-\theta).$$

3D Elementary Rotation Matrices

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad R_y(\beta) = \begin{bmatrix} c\beta & 0 & s\beta \\ 0 & 1 & 0 \\ -s\beta & 0 & c\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} c\gamma & -s\gamma & 0 \\ s\gamma & c\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

► R_y has $+s\beta$ top-right, $-s\beta$ bottom-left (opposite of R_x, R_z !).

► Trick: the “1” is on the axis you rotate about ($x \rightarrow$ row/col 1, $y \rightarrow$ 2, $z \rightarrow$ 3).

Rotation Properties

$\det(R)=1$, $R^T R=I$, $R^{-1}=R^T$. Group: SO(3).

Composition: $R_{total} = R_z \cdot R_y \cdot R_x$ (right-to-left = intrinsic).

► $R_A \cdot R_B \neq R_B \cdot R_A$ — rotations don't commute!

Euler Angles (ZYX convention)

Any 3D rotation $= R = R_z(\alpha) R_y(\beta) R_x(\gamma)$ (yaw-pitch-roll).

Yaw α : rotation about z (heading).

Pitch β : rotation about y (nose up/down).

Roll γ : rotation about x (tilt left/right).

► Gimbal lock at pitch $= \pm 90^\circ$: lose 1 DoF (yaw and roll become same axis).

Axis-Angle (Rodrigues' Formula)

Rotate by angle θ about unit axis $\hat{\omega} = [\omega_x, \omega_y, \omega_z]^T$:

$$R = I + \sin \theta [\hat{\omega}]_x + (1 - \cos \theta) [\hat{\omega}]_x^2$$

where skew-symmetric: $[\hat{\omega}]_x = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$

Special cases: $\theta=0 \rightarrow R=I$. $\theta=\pi \rightarrow R=2\hat{\omega}\hat{\omega}^T - I$.

Homogeneous Transform (4x4)

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} R^T & -R^T t \\ 0 & 1 \end{bmatrix}.$$

Chain: ${}^A T_C = {}^A T_B \cdot {}^B T_C$.

Trig Values (no calculator!)

$$\theta \mid 0 \quad 30^\circ \quad 45^\circ \quad 60^\circ \quad 90^\circ$$

$$\sin \mid 0 \quad \frac{1}{2} \quad \frac{\sqrt{2}}{2} \quad \frac{\sqrt{3}}{2} \quad 1$$

$$\cos \mid 1 \quad \frac{\sqrt{3}}{2} \quad \frac{\sqrt{2}}{2} \quad \frac{1}{2} \quad 0$$

$\sin(-\theta) = -\sin \theta$, $\cos(-\theta) = \cos \theta$.

$\sin(90^\circ) = 1$, $\cos(90^\circ) = 0$, $\sin(-90^\circ) = -1$.

KEY DEFINITIONS [Fill-in Targets]

EKF [PS3 Q2b]: Recursive estimator for nonlinear systems; linearizes models via Jacobians to apply Kalman filter updates.

Markov [PS3 Q4a]: Given current state, future is independent of history.

IK challenges [Prac Q4c]: Joint limits, singularities, arm redundancy, multiple solutions, workspace limits.

Meas. model [PS3 Q1c]: $P(\text{sensor reading} | \text{robot at location})$.

Loop closure [PS3 Q2d]: Re-observing known landmark \rightarrow corrects accumulated drift.

C-space [PS3 Q3d]: Space of all robot configurations (joint angles). Obstacles mapped to C-space.

RRT step: $q_{new} = q_{near} + \varepsilon \cdot \frac{q_{rand} - q_{near}}{\|q_{rand} - q_{near}\|}$.

FORMULA QUICK REF

DH: $T = \text{Rot}(x, \alpha)\text{Trans}(x, a)\text{Trans}(z, d)\text{Rot}(z, \theta)$

IK: $c\theta_2 = \frac{x^2 + y^2 - l_1^2 - l_2^2}{2l_1l_2}; \theta_1 = \text{atan}2(y, x) - \text{atan}2(l_2s_2, l_1 + l_2c_2)$

Camera: $s[u, v, 1]^T = K[R|t][X, Y, Z, 1]^T$. Pose: $C = -R_t$

Bellman: $V^* = \max_a \sum_{s'} P(s'|s, a)[R + \gamma V^*(s')]$

Q-val: $Q(s, a) = R(s, a) + \gamma V(s')$. Self-loop: $V = R/(1-\gamma)$

PF Bayes: $P(\text{loc}|r) = \frac{P(r|\text{loc})P(\text{loc})}{P(r)}$. Meas. model = $P(r|\text{loc})$

PF motion: $x' = x + d' \cos \Theta'$, $y' = y + d' \sin \Theta'$ (turn first, then move)

SLAM: EKF $O(n^2)$. Fast $O(Mn)$. EKFs = $M \times n$. Fast wins: $M \ll n$

Stereo: $Z = fB/d$. FoV: $\text{atan}(d/2f)$. Gear: $n:1 \rightarrow \text{spd}/n, \text{trq} \times n$

Grübler: dof = $M(N-1-J) + \sum f_i$

DMP: $\tau \dot{v} = K(g-x) - Dv + (g-x_0)f(s); f = \sum \psi_i s_i$

$\tau \dot{s} = -\alpha s$

Kabsch: $H = P^T Q$, svd $\rightarrow R = V \text{diag}(1, 1, \det(VU^T)) U^T, t = \bar{q} - R\bar{p}$

Singularity: $\det(J) = l_1 l_2 \sin \theta_2 = 0$ at $\theta_2 = 0, \pi$

Conv: $h[m, n] = \sum g[k, l] f[m+k, n+l]$. Gauss: $\frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$

Common Pitfalls & Exam Tips

► DH: R joint $\rightarrow \theta$ varies; P joint $\rightarrow d$ varies. Don't mix up!

► Camera: $[R|t]$ transforms world \rightarrow camera, NOT camera pose. Camera position = $-R^T t$.

► IK: Always check $|c\theta_2| \leq 1$ (reachable?). Two solutions: elbow up/down.

► PF motion: Robot turns first (θ), then moves (x, y). Each particle gets different noise.

► Resampling: Concentrates particles, does NOT increase total count. Weights reset to $1/N$.

► MDP: V^* gives value of best action. Q^* gives value of specific action. π^* tells you which action.

► VI vs PI: VI can stop early (check ΔV). PI must finish full evaluation before checking convergence.

► Discount: γ close to 0 \rightarrow greedy/myopic. γ close to 1 \rightarrow values future. $\gamma=1 \rightarrow$ may not converge.

► SLAM: EKF-SLAM = one big joint state. FastSLAM = factored (particles + small EKFs).

► BC vs IRL: BC copies actions (fast but brittle). IRL learns why (reward), then finds new policy.

How to Solve Common Problem Types

Grübler: Count links (incl. ground), count joints, identify types \rightarrow plug in.

DH table: Assign frames (z =joint axis, x =common normal), read off α, a, d, θ .

FK: Plug DH params into T matrix for each joint, multiply ${}^0 T_1 \cdot {}^1 T_2 \dots$

IK: Use law of cosines for θ_2 , then atan2 for θ_1 . Check reachability.

Camera proj: Back-project (need depth!) \rightarrow transform frame \rightarrow project with $K \rightarrow$ divide by Z_c .

PF predict: Apply motion + noise to each particle independently.

PF update: Weight = how well particle's expected reading matches actual reading.

MDP Q-value: $Q = R + \gamma V(s')$ if deterministic. $Q = \sum P(s')[R + \gamma V(s')]$ if stochastic.

Self-loop: $V = R + \gamma V \rightarrow V = R/(1-\gamma)$. Always check for this pattern!

Kabsch: Center points, compute $H = P^T Q$, SVD, assemble R , then $t = \bar{q} - R\bar{p}$.