Summary 6

Debarpita Dash

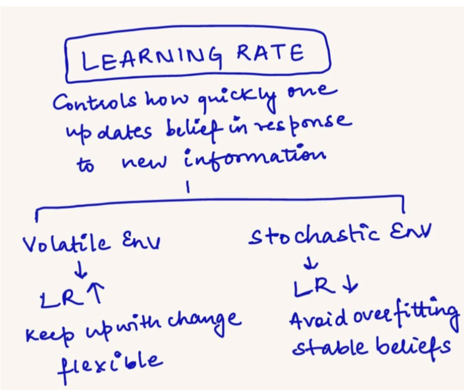
Volatility and Stochasticity

Imagine you're interacting with a cat whose mood affects whether it will scratch you. If the cat is in a good mood, there's still a 10% chance it might scratch you—this unpredictability is stochasticity. It reflects randomness within a stable underlying state: the cat's mood hasn't changed, but outcomes still vary due to noise or inherent unpredictability in behavior. In contrast, volatility refers to a deeper kind of uncertainty: the cat's mood itself can shift between good and bad. These mood changes alter the rules of the game—suddenly, the likelihood of being scratched may spike to 80% because the environment (i.e., the cat's internal state) has fundamentally changed.

People with anxiety or depression may misattribute a single bad outcome to a fundamental shift in the environment rather than random noise.This can make the world feel more threatening, fueling fear and avoidance.

Learning Through Interaction: Reinforcement Learning and the Rescorla-Wagner Rule

In reinforcement learning, an agent learns by interacting with its environment, receiving rewards based on its actions. The agent uses these rewards to gradually adjust its expectations about the value of actions. A classic model that captures this process is the Rescorla-Wagner update rule, expressed as:



$$\textbf{Reward}(t+1) = \textbf{Reward}(t) + \alpha(\textbf{Outcome}(t) - \textbf{Reward}(t))$$

This rule says that the expected reward at the next time point (t+1) is adjusted based on the current estimate (Reward(t)), the actual outcome (Outcome(t)), and the prediction error, which is the difference between the expected and actual outcome.Here α is the learning rate.

Prediction Error and the Role of Surprise in Learning

An equivalent update rule is:

$$\textbf{Reward}(t+1) = (1-\alpha) \times \textbf{Reward}(t) + \alpha \times \textbf{Outcome}(t)$$

This shows that learning is a weighted average of past belief and new evidence. Learning is driven by surprise, the size of the prediction error, not whether the outcome is good or bad. Bigger surprises are equal to bigger updates, helping the agent adapt when its current model no longer fits reality.

Brain and Reward Prediction Error(RPE)

If you expected a cookie but got two, your brain registers a positive RPE. If you expected a cookie but got none, that’s a negative RPE.

The dopamine RPE hypothesis suggests that dopaminergic neurons, particularly those in the ventral tegmental area (VTA) and substantia nigra, encode this RPE signal. Here’s how:

🔺 Positive RPE (better-than-expected outcome): Dopamine neurons fire more.
🔻 Negative RPE (worse-than-expected outcome): Dopamine neurons pause or fire less.
➖ Zero RPE (as expected): Dopamine firing stays at baseline.
This burst or drop in dopamine acts as a teaching signal, updating value estimates in the brain and guiding future behavior. Above experiment was tested in monkeys as well

Bayesian Learning: Adaptive Belief Updating in Uncertain and Volatile Environments

Bayesian learners update their beliefs using probabilistic reasoning, offering a more flexible alternative to traditional reinforcement learning models with fixed learning rates. Instead of treating all information equally, they adjust learning based on uncertainty and environmental volatility. By tracking outcomes (y_i), prediction confidence (SD_i), and volatility (k_μ), the agent maintains a mean belief (μ_i) and its variance (v*μ_i).

One-Arm Bandit and Learning Under Uncertainty

Researchers used a one-arm bandit task to study learning in uncertain environments. Participants chose between blue and green boxes for points. In the stable setting, the blue box rewarded 75% of the time. In the volatile setting, the green box offered an 80% reward rate, but the probabilities switched every 30–40 trials, requiring continual relearning.This forced participants to constantly relearn which option was better.

Brain Adaptation: mPFC and Volatility

The medial prefrontal cortex (mPFC) dynamically adjusted its activity based on how volatile the environment was. When the brain sensed the environment was rapidly changing, it increased the learning rate essentially, it trusted new evidence more. This behavior aligns with Bayesian predictions, which suggest we should weigh new information more heavily in volatile situations.

Adaptive Gain Theory (AGT)

It explains how the brain balances focus and exploration based on environmental demands, using a chemical called norepinephrine (NE). If NE levels are too low, a person may feel drowsy or inattentive. The optimal level of NE results in sharp focus and peak performance. But too much NE makes someone jittery and easily distracted. This balance is described by the Yerkes-Dodson law, which states that performance is best at moderate levels of arousal.

Models of Learning: Healthy vs. Lesioned

Three models describe how the brain manages uncertainty and change. In the Healthy Model, the brain accurately tracks both randomness (stochasticity, denoted as s_t) and genuine environmental changes (volatility, v_t), allowing it to adapt learning rates effectively. In the Stochasticity Lesion Model, the brain fails to recognize randomness and mistakenly interprets random fluctuations as meaningful changes. This leads to an excessive increase in the learning rate and overreaction to noise. Conversely, in the Volatility Lesion Model, the brain overlooks real changes in the environment and treats them as noise. This causes the learning rate to remain too low, preventing proper adaptation when actual change occurs.

Sleep and Learning: The Role of N3 Sleep

Sleep, especially deep sleep (N3 stage), significantly improves learning by helping the brain distinguish between randomness and real environmental change. People with high trait anxiety (HTA) often confuse randomness (stochasticity) with actual change (volatility), leading to incorrect adjustments in learning rate. N3 sleep helps fix this by reducing the learning rate when randomness is just noise, preventing overreactions. During N3 sleep, delta waves — low-frequency, high-amplitude brain waves (greater than 75 μV) — dominate, as shown in EEG recordings. These waves are crucial for stabilizing learning processes and optimizing future decision-making.

Naturalistic Decision Making

Ethology: The scientific study of animal behavior as it occurs in natural environments, focusing on instinctive or evolved behaviors that emerge in specific contexts.
Behavioral Ecology: A related field that investigates how animal behavior changes in response to environmental variation, and how such behaviors may contribute to survival and reproduction.

Example of Naturalistic Decision Making

The White Bellbird holds the record for the loudest bird call, reaching up to 125 decibels—loud enough to cause hearing damage. This remarkable vocalization plays an adaptive role in attracting mates. Despite the potential harm, females approach closely during these calls. Interestingly, males adjust the direction of their calls to avoid hurting the females. This behavior showcases how selection can shape and modulate adaptations in a natural context.

Tinbergen's Four Questions: A Framework for Understanding Behavior

To fully understand animal behavior, Tinbergen proposed four key questions:

- Mechanism (How?): What immediate factors cause the behavior, such as neural activity, hormones, or sensory stimuli?

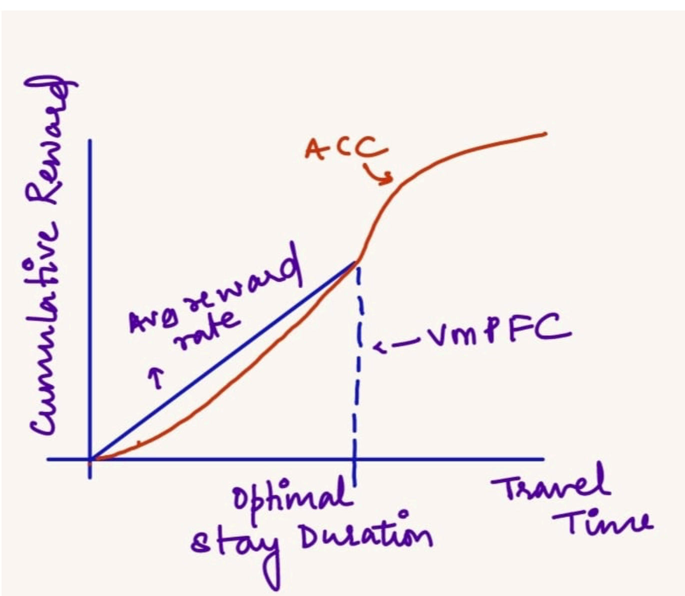
- Development (How did it develop?): How does the behavior emerge and change over an individual's lifespan?
- Function (Why?): What is the adaptive value or purpose of the behavior in promoting survival or reproduction?
- Evolution (Why did it evolve?): What is the evolutionary origin and history of the behavior across species?

This framework allows scientists to explore behavior at multiple levels—from immediate causes to evolutionary significance.

Ecological Pressures Shape Patience and Impulsivity in Primates

Species-specific ecological demands influence decision-making traits. For instance, marmosets feed on slow-releasing gum and tend to show high levels of patience, waiting for delayed rewards. This reflects low temporal discounting. On the other hand, tamarins chase fast-moving fruits and insects, which are more immediately rewarding but require quick action. They exhibit low effort discounting, showing a preference for immediate, low-effort choices. These differences demonstrate how evolutionary pressures mold cognitive traits like impulsivity and patience to match ecological niches.

Marginal Value Theorem(MVT)



MVT explains when an animal should leave a foraging patch: it should move on when the instantaneous reward rate falls below the average environmental reward rate, factoring in travel costs. Neurobiologically, the anterior cingulate cortex (ACC) tracks the reward–cost trade-off, increasing its firing rate as returns diminish until a threshold prompts disengagement. Studies show that humans and monkeys follow near-optimal foraging strategies. The posterior cingulate cortex (PCC) evaluates salience across options, while the vmPFC computes encounter value, but not search costs. This neural mechanism mirrors real-world decisions about whether to persist or seek better alternatives.