

# Mental Models

Antti Halme

Online  
15 September 2020

# Agenda

- Hello (18:45)
- **Introduction** (19:00)
- Housekeeping
- **Human Judgement, Human Nature**
- **Abstract Thinking**  
~break~
- **Physics, Chemistry, Biology**
- **Systems, Numeracy**
- **Economics, War**
- **Discussion**
- Finish (21:30)



*Most of this material is pulled from the Farnam Street blog. See their [Mental Models](#) page for much more on the topic.*

# Introduction

- **Mental Models**

- Thinking tools, concepts, and representations
- Understanding, making sense of the world
- Real world analogies for every occasion
- New connections, new opportunities
- Simplify complexity
- Relevance, saliency
- Reasoning, *better thinking*



# Human Judgement 1 / 2



- **Representativeness Heuristic**
  - We fail to account for base rates
  - We think in stereotypes
  - We rely on false conjunctions: extra detail makes a lie
- **Availability Heuristic**
  - We recall the important, the frequent, and the recent
- **First Conclusion Bias**
  - We settle on the first plausible answer

# Human Judgement 2 / 2



- **Overgeneralisation**
  - We form categories from small samples
- **Commitment / Consistency Bias**
  - We are not great at changing our mind or habits
- **Hindsight Bias**
  - We convince ourselves that we knew it all along
- **Confirmation / Falsification Bias**
  - We believe what we want to see, choose to see

# Human Nature 1 / 4



- **Trust**
  - We choose to trust others outside our family
- **Denial**
  - We cope and survive through “behavioural inertia”
- **Envy, Jealousy**
  - We compare our share with those of others
- **Love, Hate**
  - We are blinded by our passion for ideas, things and people
- **Stress**
  - We rely on our instincts, when low on energy; bias amplification



- **Incentive Bias**
  - We are controlled by our desires
- **Pavlovian Association**
  - We learn to respond to indirect incentives
- **Safety in Numbers**
  - We are team players and socially conditioned
- **Sense of Justice**
  - We are careful arbiters of fairness and propriety



- **Curiosity Instinct**
  - We ask questions and find answers
- **Narrative Instinct**
  - We tell stories to construct and seek meaning
- **Language Instinct**
  - We are hardwired to receive and transmit meaning





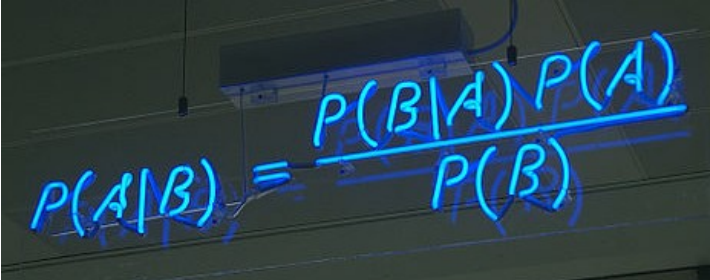
- **Fundamental Attribution Error**
  - We are surprised by how “inconsistently” others behave
- **Boredom Syndrome**
  - We act, even when actions are not needed
- **Relative Satisfaction/Misery**
  - We are happy relative to our past self or our peers

# Abstract Thinking 1 / 4

- **Circle of Competence**
  - Knowing the limits of our knowledge helps us make better decisions about decision-making
- **First Principles**
  - Reverse-engineer complexity by breaking down
  - Build new results up from from essentials
- **Thought Experiments**
  - Using imagination to examine ideas on the outset

# Abstract Thinking 2 / 4

- **Second-order Thinking**
  - Causes have causes, effects have effects
- **Probabilistic Thinking**
  - Estimation methods for identifying likely outcomes
- **Fat-tailed processes**
  - Abnormal black swan events are surprisingly common
- **Bayesian updating**
  - Using data to update beliefs



A photograph of a whiteboard with the Bayesian formula  $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$  written in blue marker. The formula is written on a whiteboard with a dark background, and the text is in a casual, handwritten style.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

# Abstract Thinking 3 / 4

- **Inversion**
  - Flip the problem around and think backwards
- **Occam's Razor**
  - Simpler explanations are more likely to be true
- **Hanlon's Razor**
  - Assume the least amount of intent in events
  - “Don't attribute to malice what can be explained by stupidity”

# Abstract Thinking 4 / 4

- **The Map is Not the Territory**
  - The map of reality is not reality
  - Maps are reductions
  - Maps are imperfect representations
  - Maps are useful because they don't show everything
  - Maps are point-in-time, perishable



# BREAK



# Physics & Chemistry 1 / 2

- **Thermodynamics**
  - Energy cannot be created or destroyed
- **Inertia**
  - Things in motion continue where they are going
- **Friction / Viscosity**
  - Resistance at the interface drains momentum
- **Relativity**
  - It is difficult observe the system you are in
- **Leverage**
  - Great output from a small input



# Physics & Chemistry 2 / 2

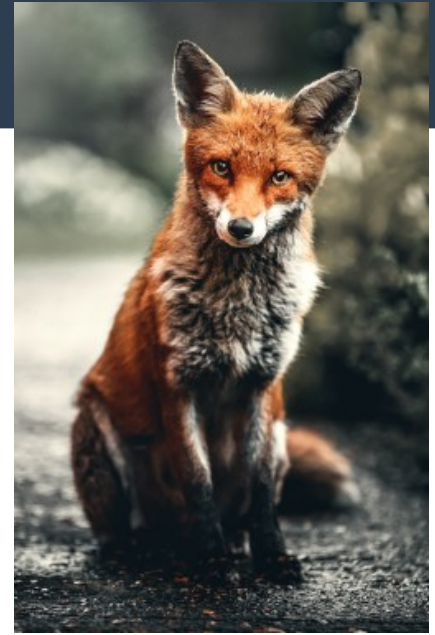
- **Velocity**
  - Motion has a direction, a vector nature
- **Reciprocity**
  - Actions have consequences in equal measure
- **Activation Energy**
  - The reaction starts only when a critical limit is reached
- **Catalysts**
  - The extra ingredient that makes the soup work
- **Alloying**
  - Combinations as more than the sum of their components





# Biology 1 / 2

- **Natural Selection and Extinction**
  - As conditions change, the fittest survive
- **Adaptation, The Red Queen Principle**
  - Have to keep running to stay in the game
- **Ecosystems**
  - Life is diverse, there are many ways to play the game
- **Niches**
  - Competition for resources leads to specialisation
- **Self-Preservation Instinct**
  - The real game is to keep on playing the game





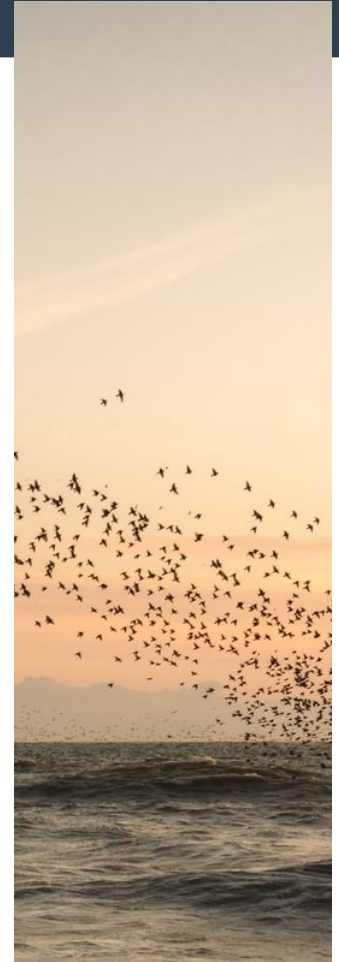
- **Replication**
  - Life is built on high-fidelity copying
- **Laziness**
  - Efficient energy use is a competitive advantage
- **Cooperation**
  - From mutual benefit to ever greater levels of organisation
- **Hierarchical Organisation**
  - Yearning for authority, leadership, and structure
- **Incentives**
  - Reward to get more of the same, to keep playing the game

# Systems 1 / 2

- **Feedback Loops**
  - Cyclic structures lead to complex flows
- **Bottlenecks, Constraints**
  - Limitations on flow change the dynamics of the system
- **Homeostasis, Equilibrium**
  - Self-regulation keeps control values within acceptable range
- **Scale**
  - Some processes and behaviours are sensitive to scale
- **Law of Diminishing Returns**
  - Positive trends often fade away in terms of incremental value
- **Churn**
  - To be keep the game alive, what is lost must be replaced

# Systems 2 / 2

- **Irreducibility**
  - There are fundamental limits to scalability and leverage
- **Cumulative Advantage (Preferential Attachment)**
  - Leaders reap the rewards; network effect, momentum
- **Margin of Safety, Backups**
  - Over-resource, over-allocate to allow for error and decay
- **Criticality**
  - Moments before a phase shift are loaded with expectation
- **Algorithms**
  - Processes can be thought of in terms of discrete steps
- **Emergence**
  - Higher-level behaviour from the interaction of lower level elements



# Numeracy 1 / 2

- **Distribution**
  - Processes produce values in a characteristic pattern
- **Compounding**
  - A regular multiplier manifests as interest on interest
- **Law of Large Numbers**
  - With more event instances, results converge to the expectation
- **Multiplying by Zero**
  - In a product of many components, the critical value dominates
- **Algebraic Equivalence**
  - Symbolic manipulation opens the door to numeric analogy



- **Fooled by Randomness**
  - We tend to pick up patterns in a non-sequential, unordered world
- **Regression to the Mean**
  - Long deviations from the average normally tend to disappear
- **Surface Area**
  - More surface area, more exposure to the environment
- **Global and Local Maxima**
  - Some peaks are higher than others, but it's hard to see up close

# Economics 1 / 3



- **Creative Destruction**
  - Out with the old, in with the new
- **Opportunity Cost**
  - Doing one thing means not being able to do another
- **Comparative Advantage**
  - Trading is beneficial, even if one party is better at everything
- **Specialisation (Pin Factory)**
  - The system benefits when everyone doesn't do everything
- **Seizing the Middle**
  - There are more moves to be made from the middle of the board

- **Trademarks, Patents, Copyright**
  - Limits on free distribution protect the creative establishment
- **Double-Entry Bookkeeping**
  - Every transaction has two parties; error detection by invariance
- **Bribery**
  - Removing the enforcer may be easier than playing by the rules
- **Arbitrage**
  - Two markets price the same thing in profitably different ways





- **Utility (Marginal, Diminishing, Increasing)**
  - Usefulness of additional units tends to vary with scale
- **Supply and Demand**
  - Limited supply leads to competition and price discovery
- **Scarcity**
  - Constrained supply leads to fierce competition and high prices
- **Mr. Market**
  - The market is like a moody neighbour, with his ups and downs

# War



- **Two-Front War**
  - Splitting resources weakens the overall position
- **Seeing the Front**
  - First-hand experience of trenches helps with decisions higher up
- **Asymmetric Warfare**
  - Each side plays by their own rules; underdogs are unpredictable
- **Counterinsurgency**
  - Asymmetry requires a deliberate strategy from the overdogs
- **Mutually Assured Destruction**
  - As opponents get stronger the *less* likely they are to engage

# Discussion

- **Summary**

- Human Judgement, Human Nature :: Heuristics, biases, instincts
- Abstract Thinking :: A critical view of the knowledge enterprise
- Physics, Chemistry, Biology :: Phenomena from the natural world
- Systems, Numeracy :: Complex behaviour from simple rules
- Economics, War :: Competition, rule breaking, tactics, strategy

- **Questions**

- What makes a good mental model?
- What is a tool? What does it mean for a concept to be a tool?
- What is easy, what is hard – for human and machine?

# Epilogue

*“The best language [today] seems to be more colorless and limp than some of the language of [centuries past]. There's a vividness, a willingness to use metaphor and literary flourishes that you are less likely to see today. [...] It may be that because we have so many technical terms available to us, that we don't **reach for the metaphor** and that will drain prose of some of its vitality, even though it kind of makes it [easier] **to convey abstract ideas.**”*

- Steven Pinker on contemporary language, **during a Q&A** for a talk based on his book *The Sense of Style* (2014)

Thanks!

# Photos

- SHAKESPEARE GHOST TOWN by Lachlan Donald on Unsplash
- Woman in black crew neck series by Maria Lysenko on Unsplash
- WANDERKARTE by niklas\_hamann on Unsplash
- [Bayes Theorem](#), Wikimedia Commons
- Coffee mug and biscuit by Daria Nepriakhina on Unsplash
- Water Droplet by Sourav Mishra on Pexels
- Laboratory Glasses by Rodolfo Clix on Pexels
- Fox sitting on ground by Alex Andrews on Pexels
- Five parrots perched on brown wooden surface by Magda Ehlers on Pexels
- Murmurations over Brighton by Rhys Kentish on Unsplash
- [Answer to Life 42](#), Wikimedia Commons
- Heavy Industry by Ant Rozetsky on Unsplash
- Fruit stall at supermarket by Fikri Rasyid on Unsplash
- Otis Air National Guard Base Re-enacters by Suzy Brooks on Unsplash