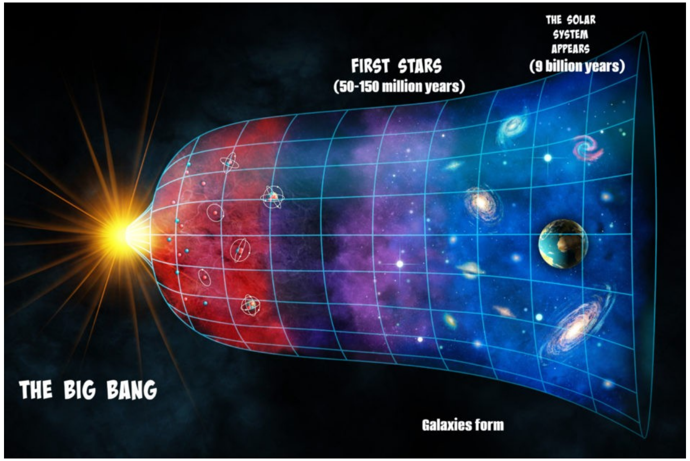
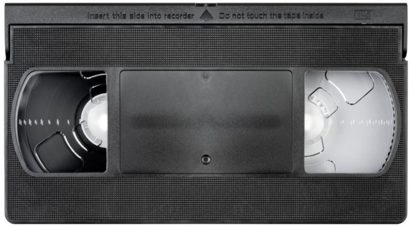
Information Dynamics and the Arrow of Time

Aram Ebtekar

Background

- The Universe consists of events ("data") on a 4D manifold
  - Events are addressed by spacetime coordinates, like data on a VHS tape
- The Universe is fully determined by:
  - Local dynamics (i.e., laws of physics)
  - Initial conditions (i.e., the Big Bang)
- The dynamics relate events across time
  - Thus, it's possible for characters living on the manifold to be aware of a past and future





Background

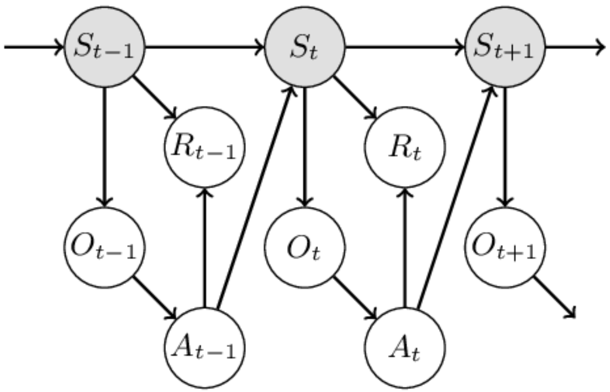
- What distinguishes cause from effect?
  - Key to understanding growth, decay, memory, learning, planning, etc.

Background

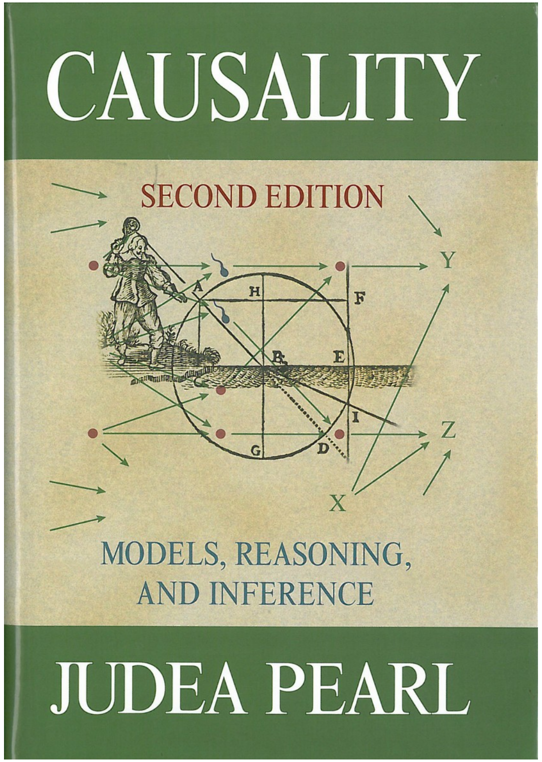
- Judea Pearl mathematized causality in terms of probabilistic graphical models
  - Example from AI: we can represent the POMDP

$$\mathbb{P}(\mathbf{S},\mathbf{O},\mathbf{A},\mathbf{R}) = \mathbb{P}(S_0)\prod_{t=0}^{T}\mathbb{P}(O_t \mid S_t)\mathbb{P}(A_t \mid O_t)\mathbb{P}(R_t \mid S_t, A_t)\mathbb{P}(S_{t+1} \mid S_t, A_t)\cdots$$

    by the following graph:

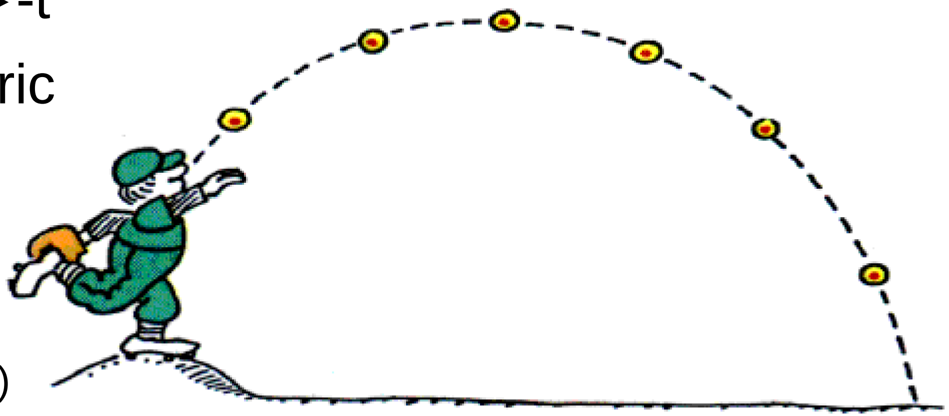


  - Decision nodes A_t are subject to optimization
    - They represent "free will", selecting among many counterfactual futures



Background

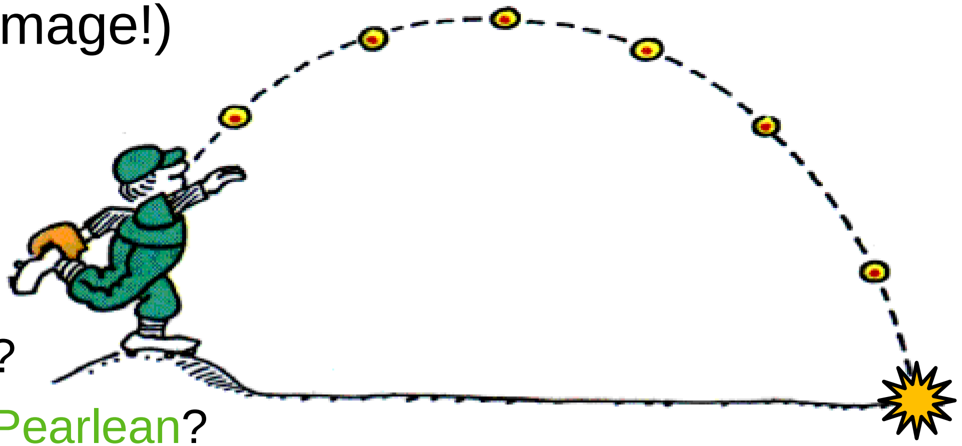
- Problem: the leading theories in physics look nothing like Pearl's models! In fact, they are symmetric under time reversal*
  - That means every rewinded movie is physically valid
  - Example: Newton's law $F = ma = m(d/dt)^2 x$ is invariant to the substitution t → -t
  - A baseball's trajectory is symmetric
    - ... until it lands!
    - We also ignored air resistance



* up to parity & charge conjugation (CPT symmetry)

Background

- Macroscopically, the landing appears to break symmetry
- Microscopically, there's no issue:
  - Kinetic energy transfers to air & ground molecules, as heat & sound
  - In rewind: air & ground molecules miraculously converge to push the ball up (and repair any impact damage!)
- Similarly, physics can unshatter a glass or unfry an egg
  - But it's *unlikely*
    - Why are the statistics asymmetric?
    - Why are the statistics specifically Pearlean?



Background

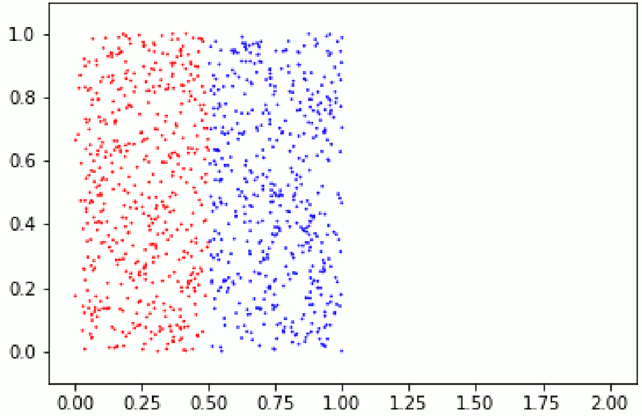
- One possibility is to discover an asymmetric fundamental law
  - It would need to explain why we see both reversibility and causality
  - We will not take this approach
- The other possibility is that the initial condition is special
  - It seems hard to tie the Big Bang's entropy to our perception of time
    - Even if this approach is correct, is causality too complex to understand from first principles? Just as psychology is not derived from quantum physics
  - Using simple models, we'll see that's not the case!
    - The arrow of time is really about the interplay between chaos and information

Agenda

- In four stages, we develop a rigorous model that's time-symmetric microscopically but not macroscopically
  - Stage 1: introduce the baker's map as a foundation
  - Stage 2: emulate a 1D random walk
  - Stage 3: emulate general Markov chains
  - Stage 4: emulate full-blown Pearlean causality
- Then, we examine its macroscopic statistics
  - The 2nd law of thermodynamics, and more
  - Consequent asymmetric phenomena: memory and agency

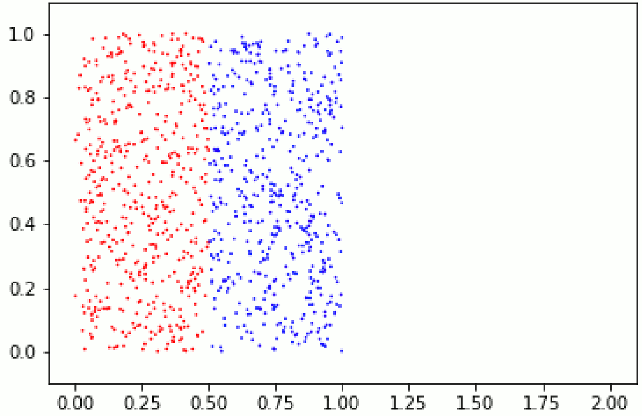
Model 1: The Baker’s Map

- A classical one-particle system’s state consists of a 3D position and 3D momentum, i.e., a point in 6D phase space
  - Make it simpler: if the particle moves in 1D, the phase space is 2D
  - Even simpler: discretize time and choose a convenient bijection for the dynamics
- The baker’s map acts on the unit square
  - Stretch horizontally, squeeze vertically, cut and glue to get back the same square
  - It’s reversible, chaotic, and area-preserving



Model 1: The Baker's Map

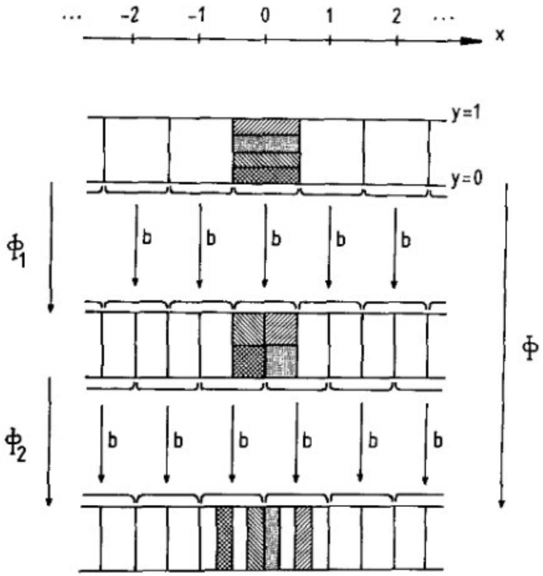
- Writing the coordinates in binary, the baker's map becomes
  - $(0.x_0x_1x_2\ldots,\ 0.x_{-1}x_{-2}x_{-3}\ldots) \to (0.x_1x_2x_3\ldots,\ 0.x_0x_{-1}x_{-2}\ldots)$
- Suppose the initial distribution is continuous, but we only see a fixed number of the most significant bits
  - Then, for large *i*, the x_i are uniform & i.i.d.
  - Eventually, we *only* see large indices *i*
  - Therefore, the state appears to fully mix!
    - The sequence of random digits provides an infinite reserve of "hidden entropy" to extract



Model 2: Multibaker Chain

- Let's extend the baker's map to yield more interesting dynamics
- Gaspard (1992) emulated a random walk on $\mathbb{Z}$, by identifying each "macrostate" with a "microscopic" unit square
  - Overall state space is $\mathbb{Z} \times [0,1) \times [0,1) \simeq \mathbb{R} \times [0,1)$
  - Apply two successive chains of baker's maps
  - When y-coordinate starts uniformly distributed, the result is a random walk with probabilities:

$$p(s,\, s') = \begin{cases} 1/2 & \text{if } s' = s, \\ 1/4 & \text{if } s' = s \pm 1, \\ 0 & \text{otherwise.} \end{cases}$$

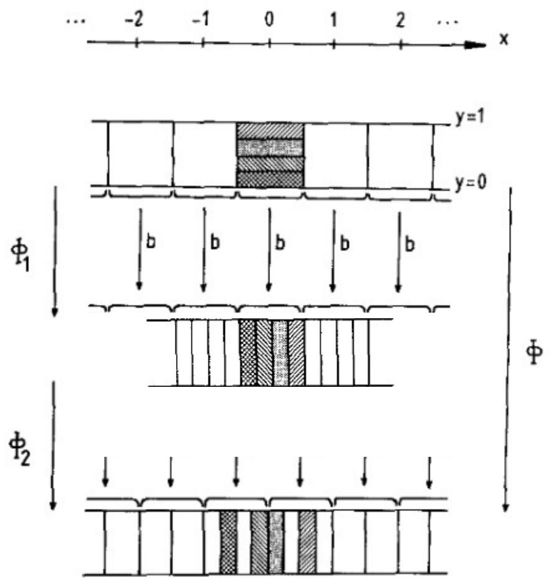


Model 2: Multibaker Chain

- Consider an alternative decomposition of the multibaker mapping
  - Write the state's coordinates in base *m* = 4
  - First, act "microscopically" by shifting **x**
  - Then, act "macroscopically" by permuting the columns, which are given by (s, x_0)
    - Permuting the columns is equivalent to applying a bijection $T : (s, x_0) \mapsto (s', x_0')$, as follows:

$$(s.x_{-1}x_{-2}x_{-3}\ldots,\ 0.x_0x_1x_2\ldots)$$
$$\xrightarrow{\text{shift}} (s.x_0x_{-1}x_{-2}\ldots,\ 0.x_1x_2x_3\ldots)$$
$$\xrightarrow{\text{permute}} (s'.x_0'x_{-1}x_{-2}\ldots,\ 0.x_1x_2x_3\ldots)$$

$$T(s,\,0) = (s-1,\,3)$$
$$T(s,\,1) = (s,\,1)$$
$$T(s,\,2) = (s,\,2)$$
$$T(s,\,3) = (s+1,\,0)$$



Model 3: Markov-baker Chain

- Now, let's generalize to any countable state space, any base *m*, and any transition function $T : (s, x_0) \mapsto (s', x'_0)$
  - If x_0 is uniformly distributed on {0, 1,..., m-1}, then the macroscopic transition probabilities *p*(*s*, *s*') are multiples of 1/*m*
  - If **x** is uniformly i.i.d., the macroscopic dynamics *p* is homogeneous
  - If *T* is bijective, every state *s* has exactly *m* images and *m* preimages; therefore, *p* is doubly stochastic

$$
\begin{aligned}
&(s.x_{-1}x_{-2}x_{-3}\ldots,\ 0.x_0x_1x_2\ldots) \\
\xrightarrow{\text{shift}}\ &(s.x_0x_{-1}x_{-2}\ldots,\ 0.x_1x_2x_3\ldots) \\
\xrightarrow{\text{permute}}\ &(s'.x'_0x_{-1}x_{-2}\ldots,\ 0.x_1x_2x_3\ldots)
\end{aligned}
$$

Model 3: Markov-baker Chain

- Conversely, let's emulate an arbitrary Markov chain described by:
  - A countable state space, WLOG taken to be $\mathbb{Z}^{\geq 0} := \{0,\, 1,\, 2,\, \ldots\}$
  - A doubly stochastic matrix *p* whose entries are multiples of 1/*m*
- Let *T* act bijectively, on pairs of state and base-*m* digit, by

$$T\left(s,\; i + m\sum_{r=0}^{s'-1} p(s,\, r)\right) := \left(s',\; i + m\sum_{r=0}^{s-1} p(r,\, s')\right) \quad \forall s, s', i \in \mathbb{Z}^{\geq 0},\; i < m \cdot p(s, s')$$

  - For a uniformly random digit *X*, we verify the transition probabilities:

$$\mathbb{P}\left(\exists y,\, T(s,\, X) = (s',\, y)\right) = \frac{1}{m} \cdot m \cdot p(s,\, s') = p(s,\, s')$$

Model 3: Markov-baker Chain

- Let’s review what we have so far
  - By endogenizing randomness into the state, *every* doubly stochastic homogeneous Markov chain can be made deterministic & reversible
    - Conversely, we see why macroscopic systems tend to be Markovian
  - In this representation, symmetry is broken *only* by the initial condition
    - Initially (but not later), the digits are uniformly and independently distributed
  - The 2nd law of thermodynamics is a known property of Markov chains
    - Therefore, it's also a property of our reversible systems!
  - Already, this model is powerful enough to study some open questions
    - E.g., what sorts of initial conditions suffice to get the 2nd law?

Motivating Model 4

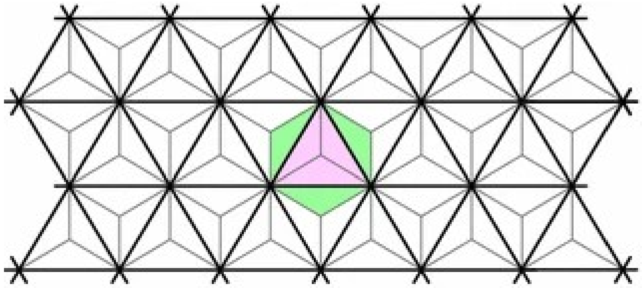
- Let's go beyond Markov chains: Pearl notes that causality is more readily inferred in the presence of colliders

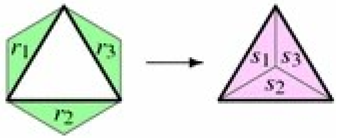
$$
\begin{array}{ccccccccccc}
\cdots \to & A_0 & \to & A_1 & \to & A_2 & \to & A_3 & \to \cdots & & \\
 & & & & \searrow & & & & & & \\
\cdots \to & B_0 & \to & B_1 & \to & B_2 & \to & B_3 & \to \cdots & &
\end{array}
$$

  - Consider an interaction, where system A causally influences system B
  - Example 1: B is a memory that records an observation of A
  - Example 2: A is an agent that intervenes to manipulate B
- In either case, information from A enters future states of B

Model 4: Partitioned Cellular Automaton

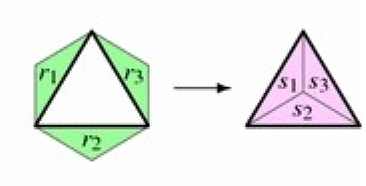
- In order to model causal separation between systems, add a discrete (i.e., cellular) spatial structure
  - Give each cell its own copy of the Markov-baker state space
    - First, apply the mapping to every cell simultaneously
    - Then, reversibly move data between adjacent cells
  - Details don’t matter: main results apply to a variety of cellular structures





Model 4: Partitioned Cellular Automaton

- The macroscopic view is equivalent to a Pearlean model that is:
  - Homogeneous in space and time
  - Full of colliders



- Therefore, Pearl's *d*-separation criterion applies
  - In particular, correlated events must have a common cause in the past
    - Can we say anything quantitative?
- We need a local version of the 2nd law of thermodynamics
  - Recall that local energy conservation is stated as a continuity equation, accounting for flux of energy that enters or leaves the system
    - For closed systems, flux = 0

$$E(A_t) = E(A_u) + \text{flux}$$

Model 4: Partitioned Cellular Automaton

- Entropy can be created, but not destroyed
  - So instead of equations, we obtain continuity *inequalities*
- Let *H* denote entropy, *I* denote mutual information, $t < u$
  - Resource law: (2nd law of thermodynamics) $H(A_t) \leq H(A_u) + \text{flux}$
  - Memory law: for *Y* not in the future of A_t, $I(A_t;\, Y) \geq I(A_u;\, Y) + \text{flux}$
  - Consequently, for disjoint systems A & B, $I(A_t;\, B_t) \geq I(A_u;\, B_u) + \text{flux}$
- Thus, mutual information can increase *only* via flux
  - In particular, correlated systems must have interacted in the past
  - Spontaneous *decrease* is possible, but thermodynamically costly

Summary

- These cellular automata serve as an *existence proof*, demonstrating how chaotic reversible dynamics can yield:
  - A thermodynamic arrow of time, i.e., the Resource law
  - A psychological arrow of time, i.e., the Memory law
  - A causal arrow of time, i.e., Pearl's *d*-separation criterion
- They can also serve as a useful tool in other lines of research
  - Some questions that appear too difficult in the context of real physics, yield clear and plausible answers within these automata
  - Please see the paper's Applications section for examples!

Epilogue: Boltzmann Brains

- In closing, I’d like to leave you with a question: how do we know anything about the world?
  - We usually talk about “measurement” as if the observer acts outside the physical system, with free will and direct observations
  - Using our automata, we endogenize the observer as a physical entity
    - Observations must be reversibly placed onto some physical memory
  - If we start with a uniform Bayesian prior, the Universe is at max entropy
    - By the Resource law, it stays at max entropy; equivalent to heat death!
    - By the Shannon identity $J(A,B) = J(A) + J(B) + I(A;B)$, where J is *lack* of entropy, our memory’s mutual information about the outside world is also zero
  - Thus, we should be unable to know anything! Unless, Solomonoff prior?

Image Credits

- Eviatar Bach https://en.wikipedia.org/wiki/Baker%27s_map
- Evan-Amos https://en.wikipedia.org/wiki/VHS
- Lucasfilm’s Star Wars: Episodes III & V
- Huang et al https://ieeexplore.ieee.org/document/8755551
- Doug Davis https://www.ux1.eiu.edu/~cfadd/1350/09Mom/CoM.html
- Emma Vanstone https://www.science-sparks.com/what-is-the-big-bang/
- Pierre Gaspard https://link.springer.com/article/10.1007/BF01048873
- Kenichi Morita https://link.springer.com/article/10.1007/s11047-017-9655-9

Resources

- Thank you! Questions, comments?
- My paper:
  - *“Information Dynamics & the Arrow of Time”* arxiv.org/abs/2109.09709
  - Contains more references, and my comments on them
- Related talk by Sean Carroll:
  - *“The Arrow of Time in Causal Networks”* youtu.be/6slug9rjaIQ