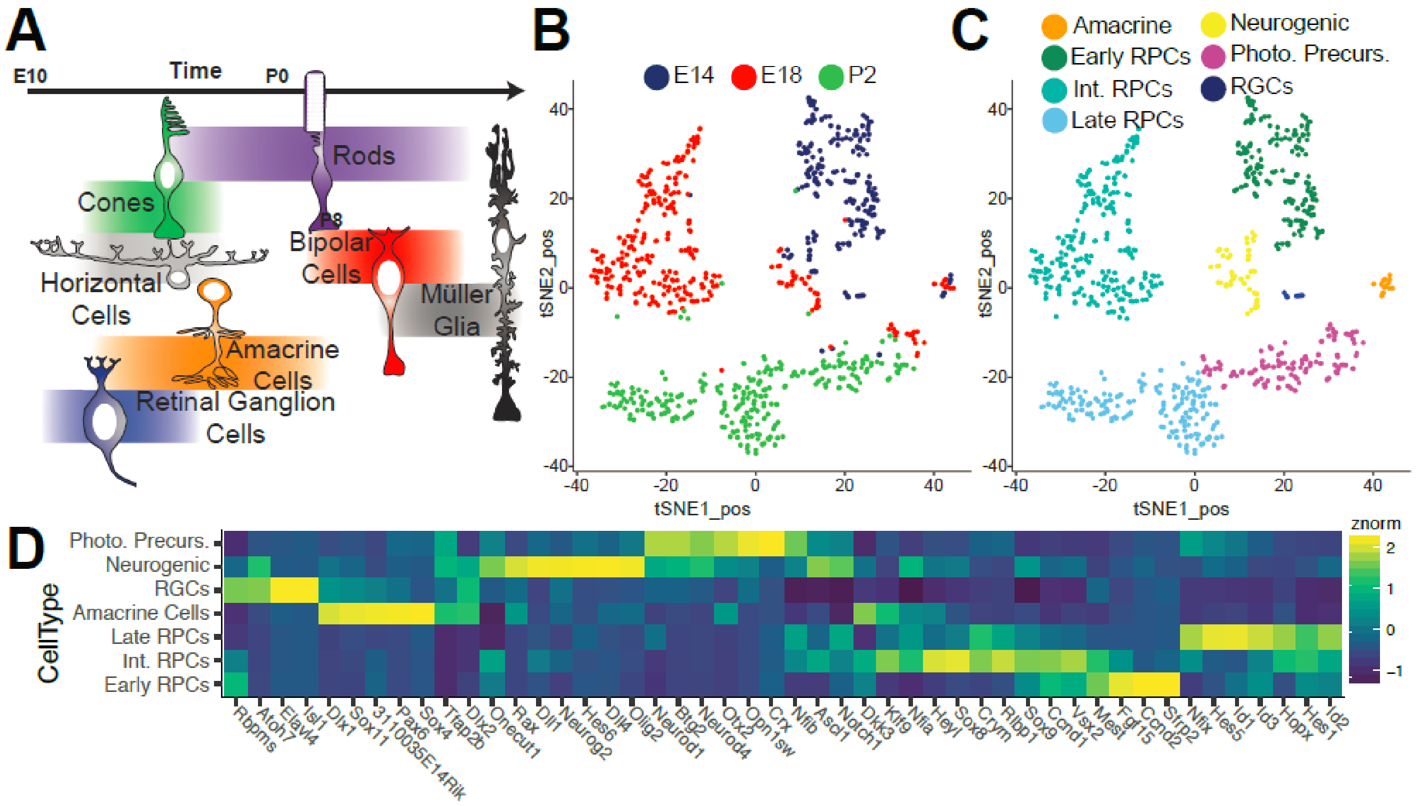
Generative AI Models
ECE 598 LV – Lecture 18

Lav R. Varshney

29 March 2022

Information Lattice Learning: Learning laws of neurogenesis



[B. Clark, et al., "Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification," *Neuron*, June 2019.]

- Single-cell RNA sequence data analysis for understanding the rules that govern pattern formation in neurodevelopment

[Yu, Varshney, Stein-O'Brien, 2019]

Information lattice learning: decompose and recompose

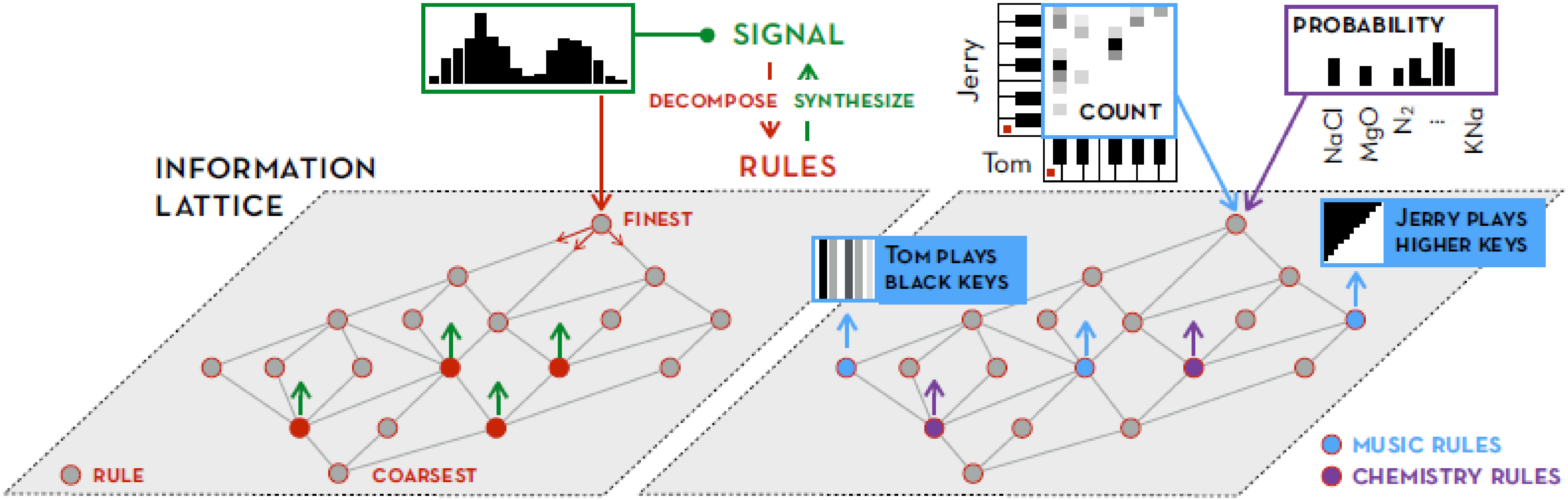
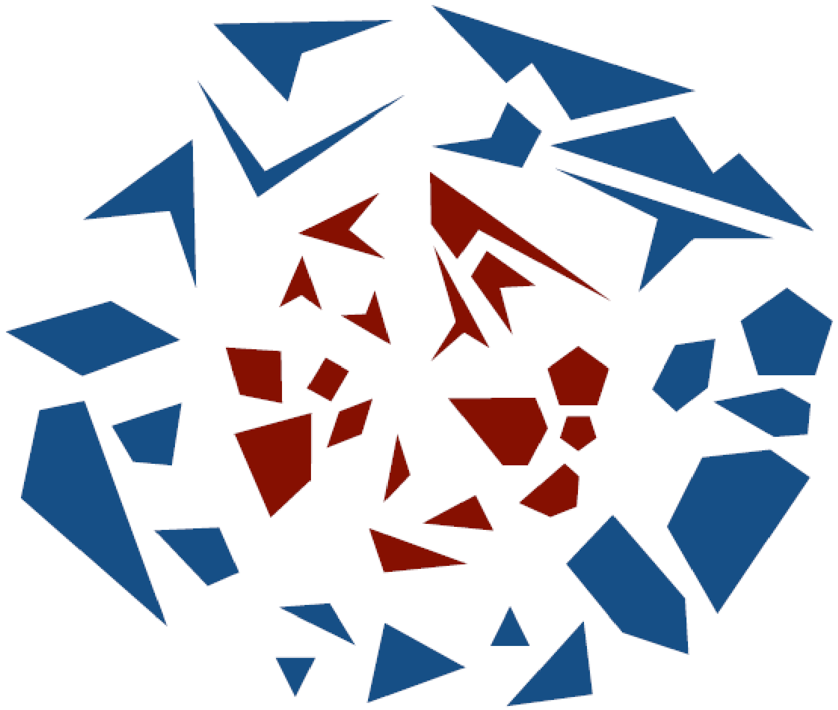
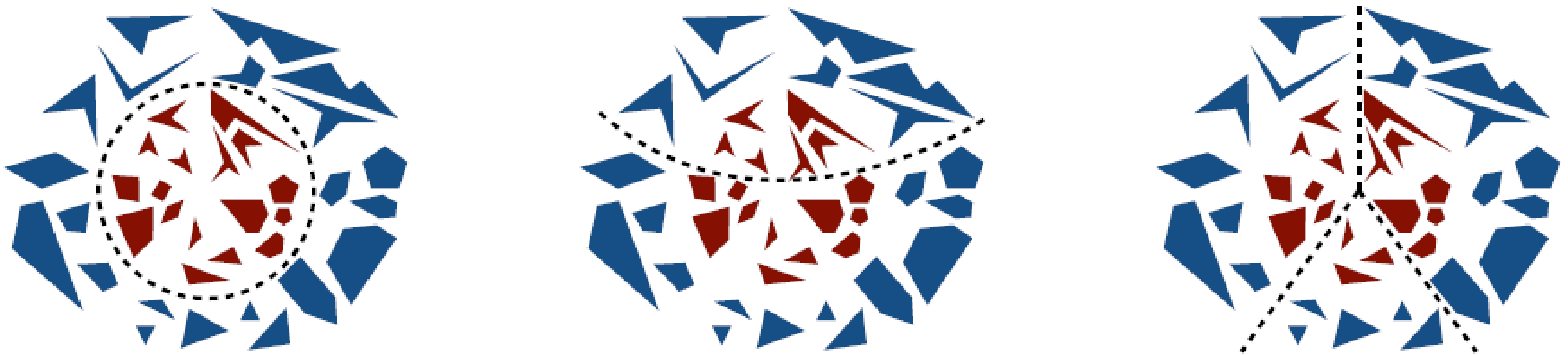


Figure 1: ILL's main idea: decompose the signal into rules that are individually simple but collectively expressive. A lattice is first constructed regardless of the signal (prior-driven), yet the same lattice may be later used to learn rules (data-driven) of signals from different topics, e.g. music and chemistry.

Information lattice learning for knowledge discovery

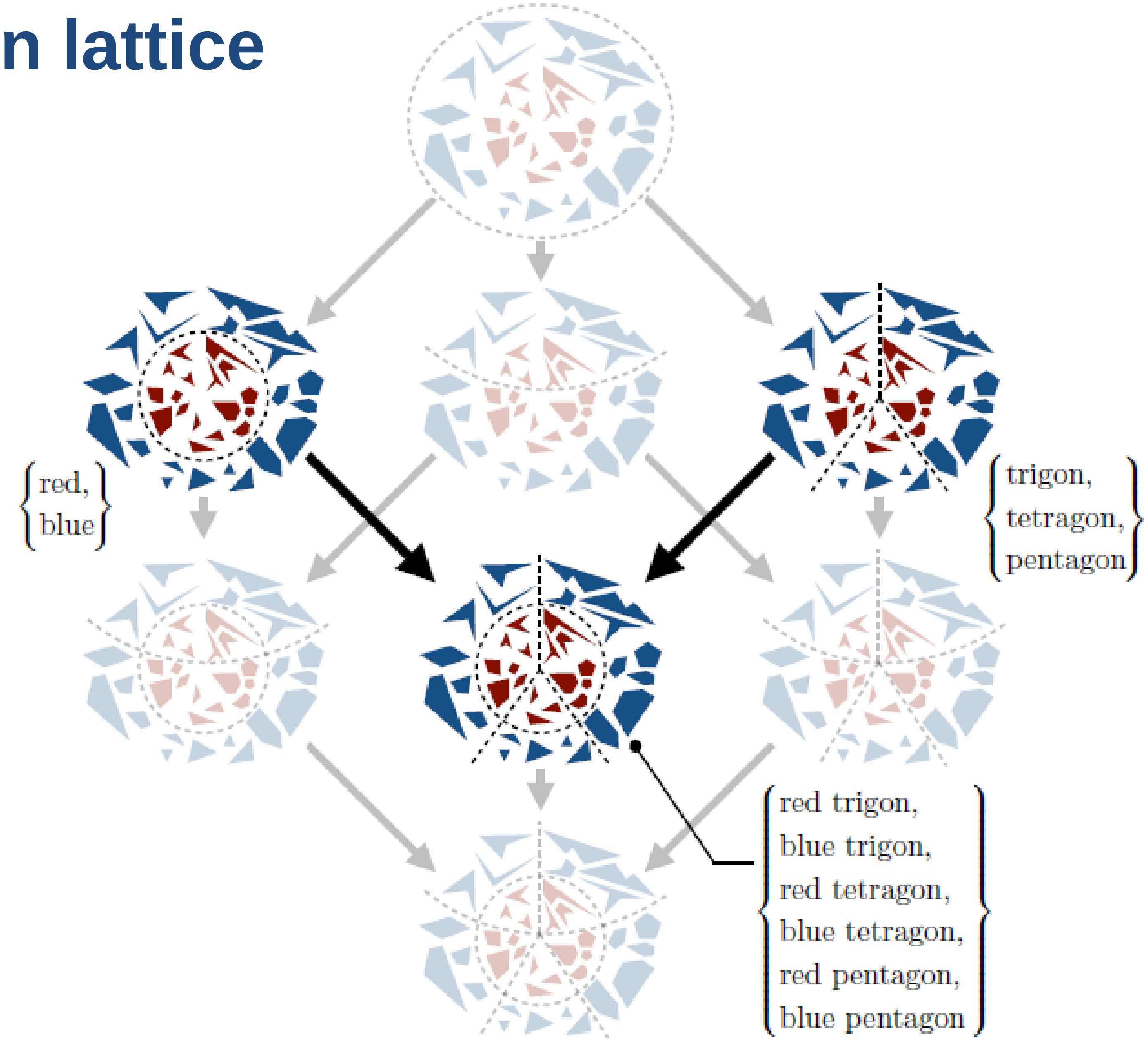


Information lattice learning for knowledge discovery



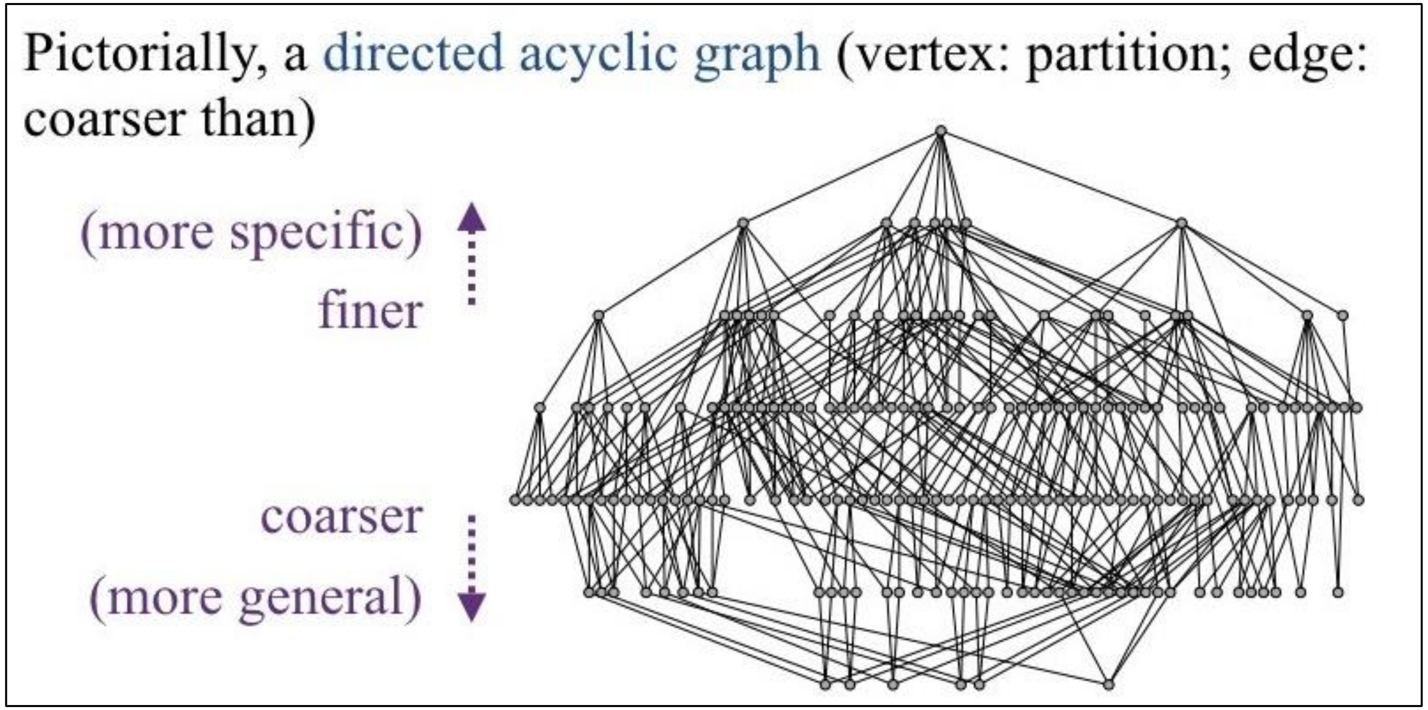
{red, blue} {convex, concave} {trigon, tetragon, pentagon}

Information lattice



Abstraction universe as partition lattice

- A set X can have multiple partitions (Bell number $B_{|X|}$)
- Let $\mathfrak{B}_X^*$ denote the family of all partitions of a set X, so $|\mathfrak{B}_X^*| = B_{|X|}$
- Compare partitions of a set by a partial order on $\mathfrak{B}_X^*$
    - Partial order yields a *partition lattice*, a hierarchical representation of a family of partitions



Information theoretic algorithm for rule learning

Learning is achieved by statistical inference on a partition lattice

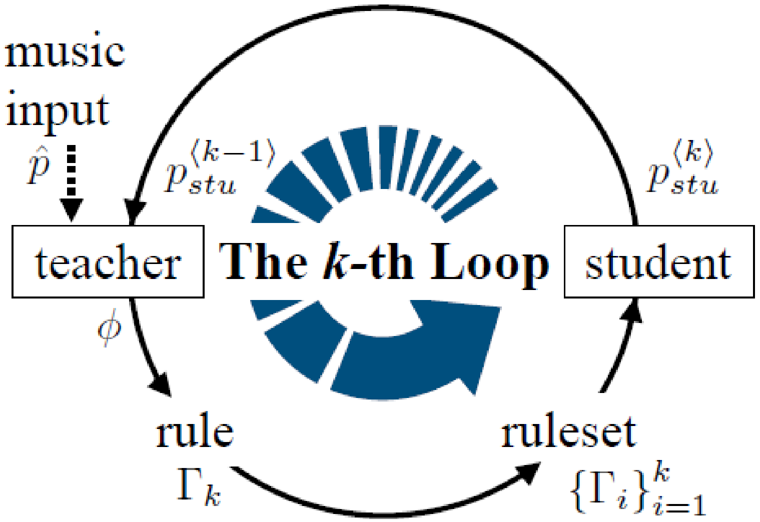
The iterative cooperation between a discriminator (teacher) and a generator (student).

The teacher solves:

$$\begin{aligned} &\text{maximize} && D\left(p_{\phi,stu}^{\langle k-1\rangle} \,\|\, \hat{p}_\phi\right) \\ &\text{subject to} && \phi \in \Phi \backslash \Phi^{\langle k-1\rangle} \end{aligned}$$

(discrete optimization)

max. Bayesian surprise



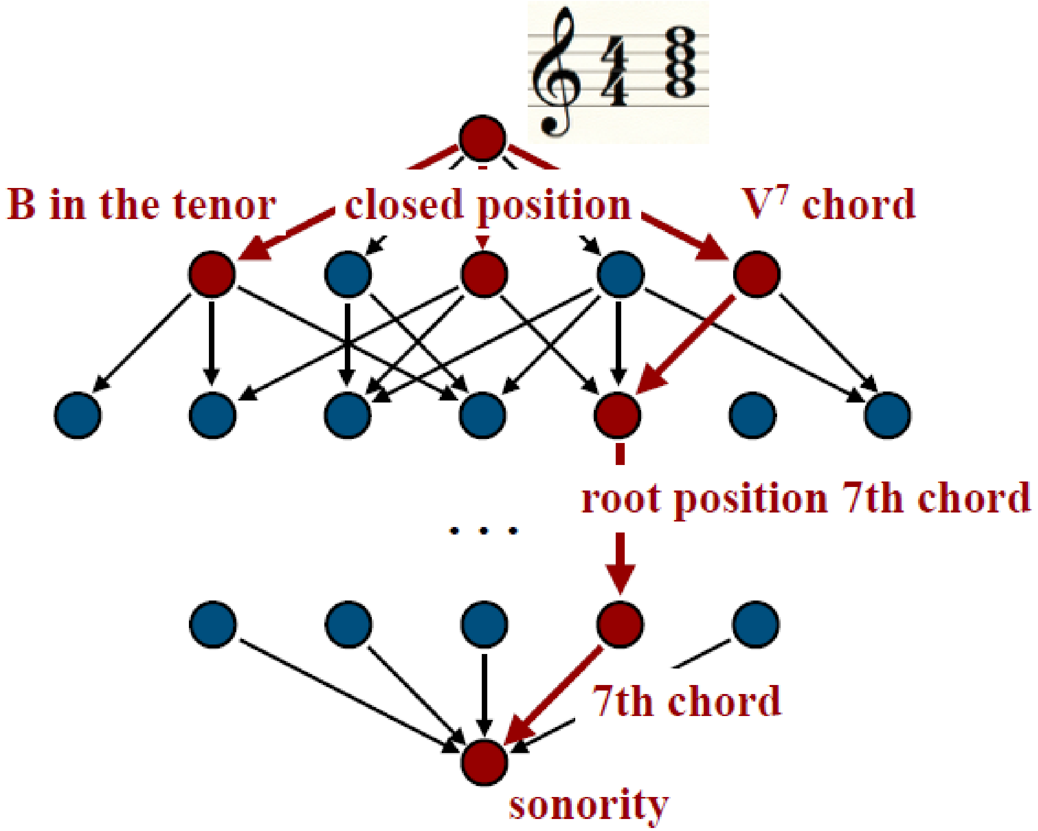
The student solves:

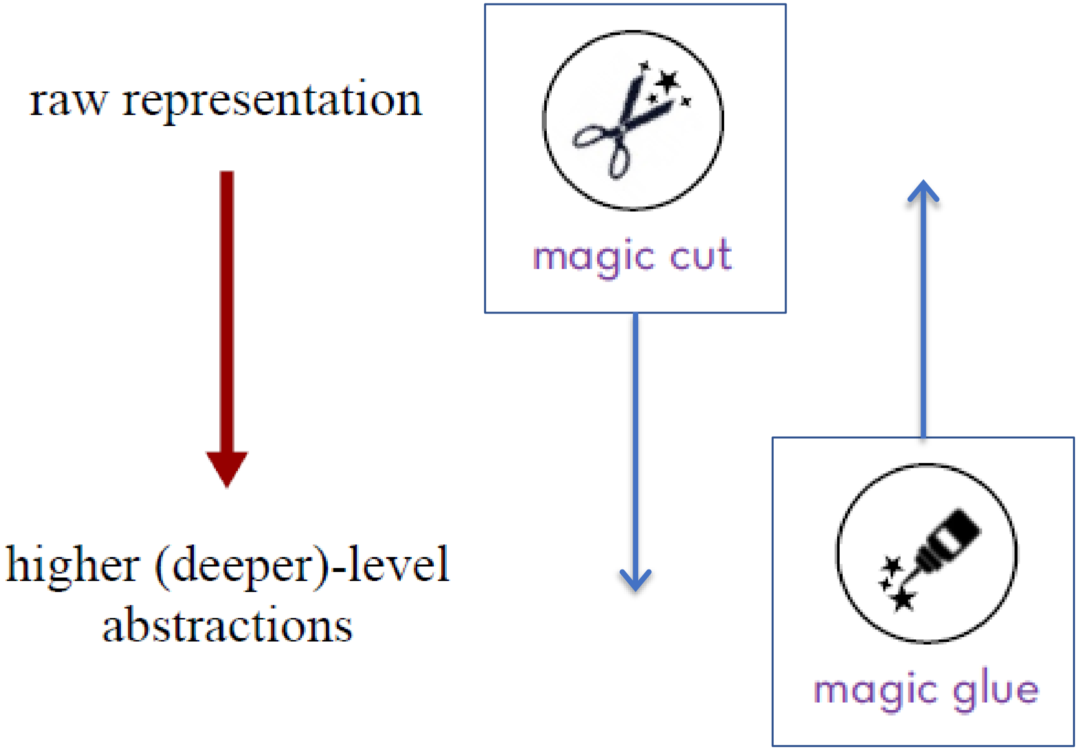
$$\begin{aligned} &\text{maximize} && S_q\left(p_{stu}^{\langle k\rangle}\right) \\ &\text{subject to} && p_{stu}^{\langle k\rangle} \in \Gamma_1 \\ & && \cdots \\ & && p_{stu}^{\langle k\rangle} \in \Gamma_k \end{aligned}$$

(linear least-squares)

max. creativity

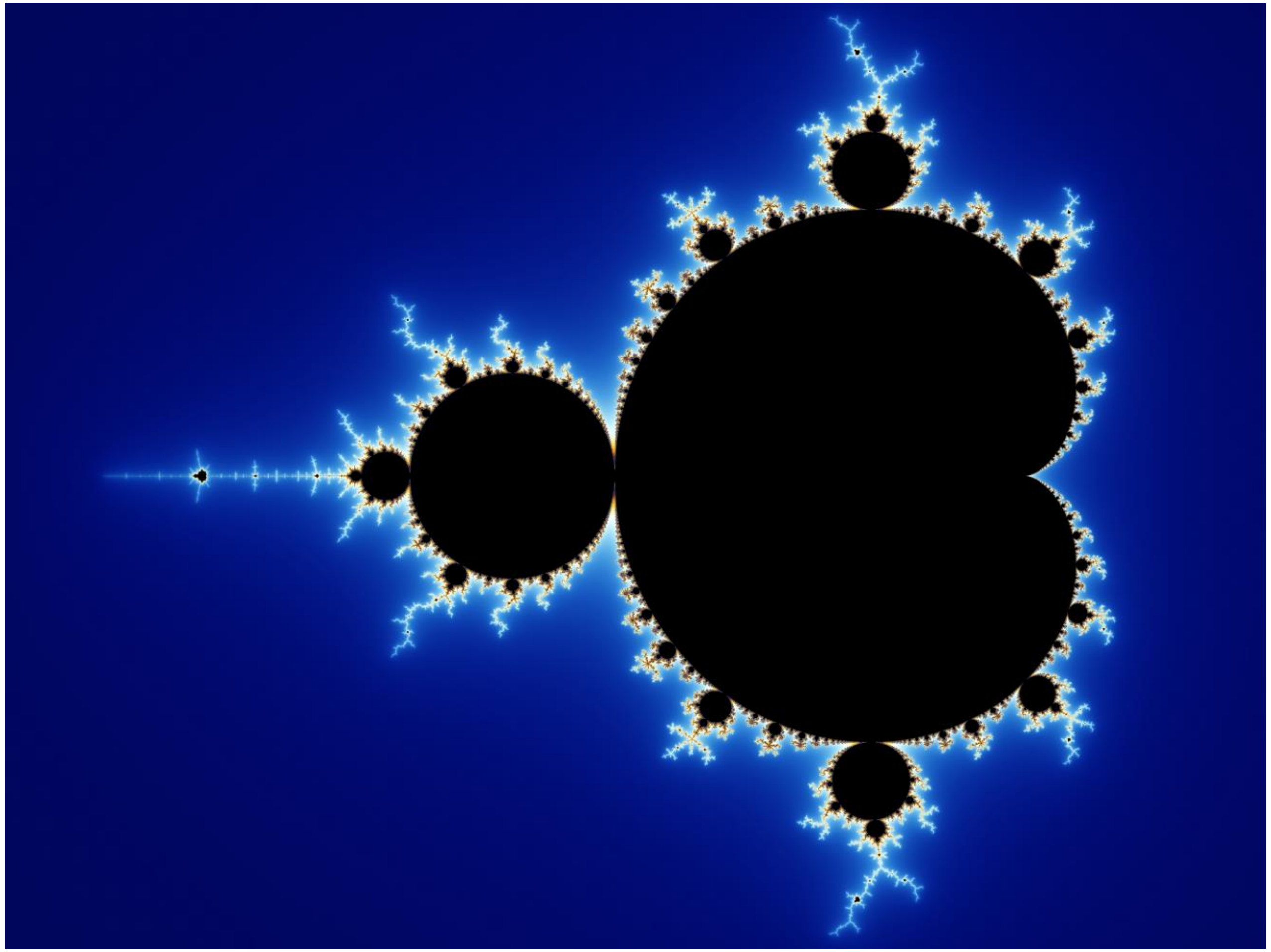
Magic cuts and magic glue involve moving up and down ILL





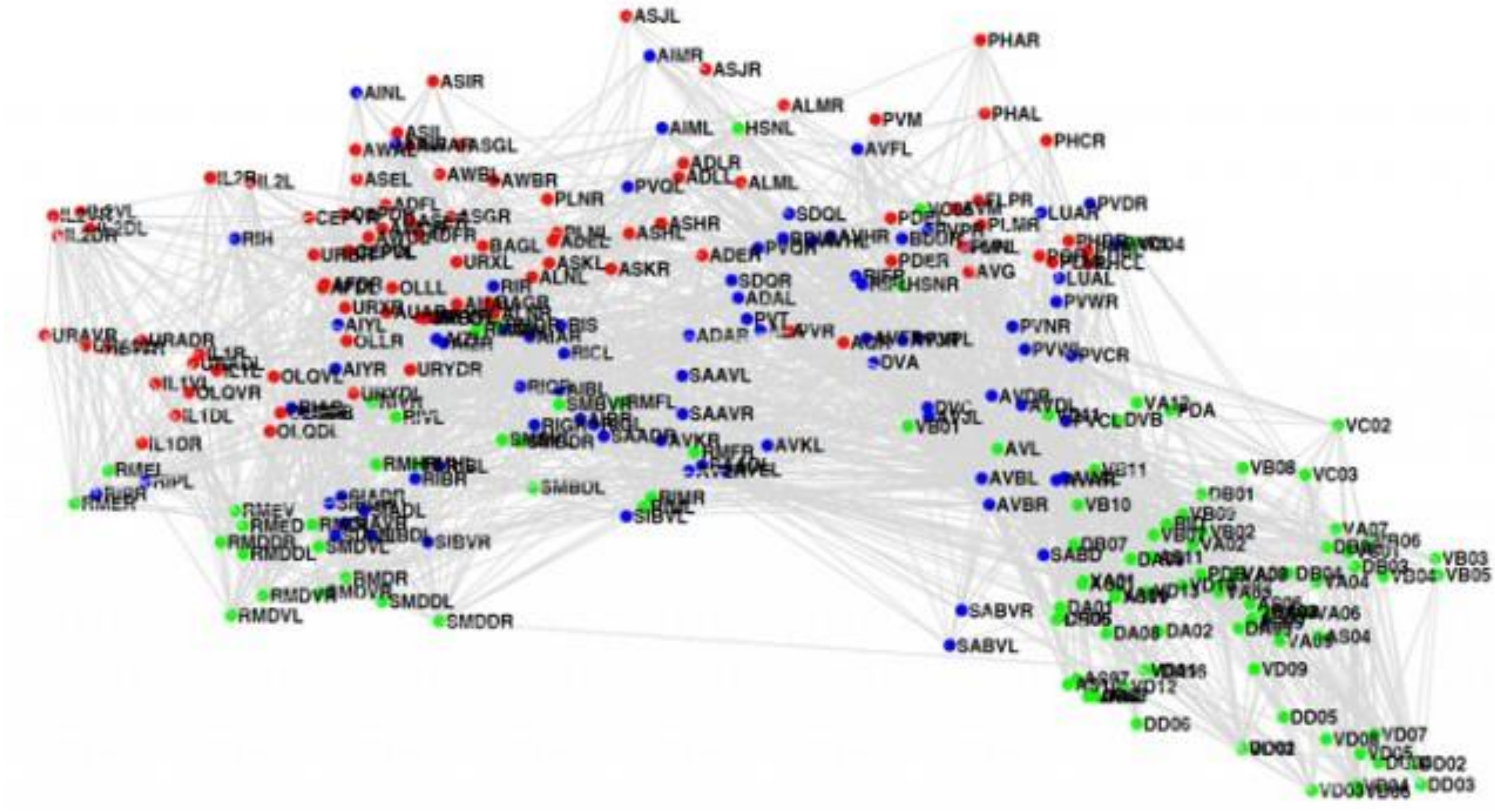
Generative Algorithms based on Rules

Fractals



https://upload.wikimedia.org/wikipedia/commons/a/a4/Mandelbrot_sequence_new.gif

https://en.wikipedia.org/wiki/Julia_set#Quadratic_polynomials



[Varshney et al., 2011]

Definition 1 (Kronecker product of matrices) *Given two matrices* $\mathbf{A} = [a_{i,j}]$ *and* $\mathbf{B}$ *of sizes* $n \times m$ *and* $n' \times m'$ *respectively, the Kronecker product matrix* $\mathbf{C}$ *of dimensions* $(n \cdot n') \times (m \cdot m')$ *is given by*

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} \doteq \begin{pmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \dots & a_{1,m}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \dots & a_{2,m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}\mathbf{B} & a_{n,2}\mathbf{B} & \dots & a_{n,m}\mathbf{B} \end{pmatrix}.$$

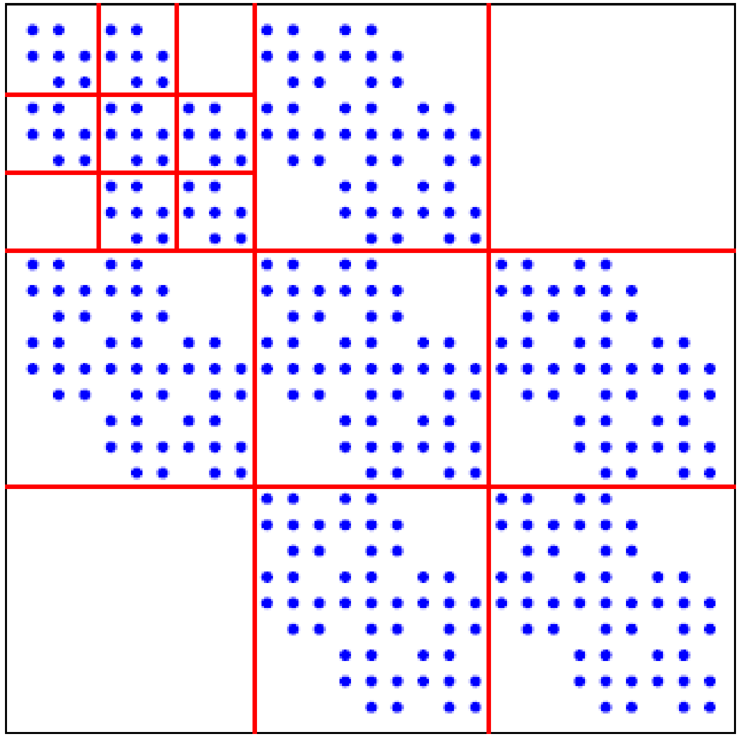
We then define the Kronecker product of two graphs simply as the Kronecker product of their corresponding adjacency matrices.

1	1	0
1	1	1
0	1	1

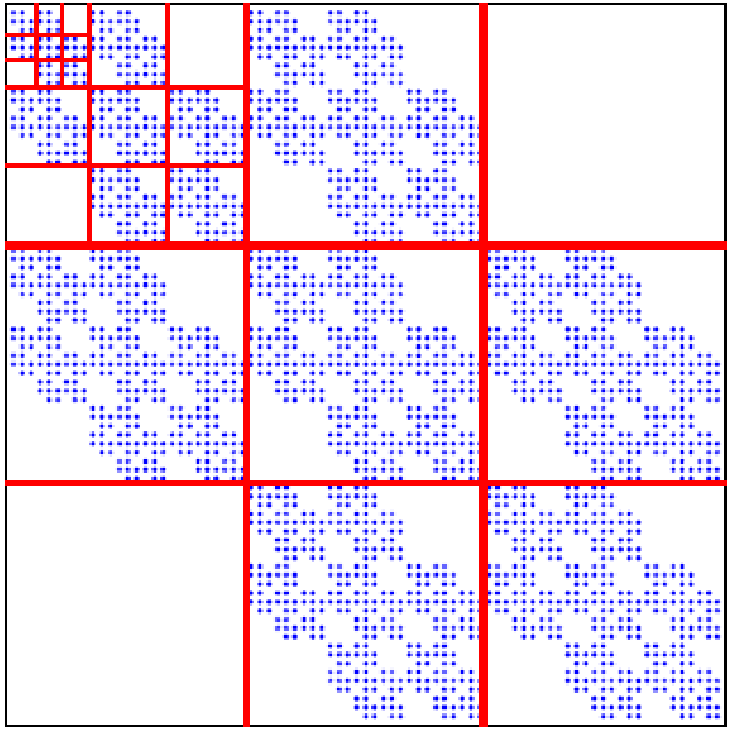
(d) Adjacency matrix of K_1

K_1	K_1	0
K_1	K_1	K_1
0	K_1	K_1

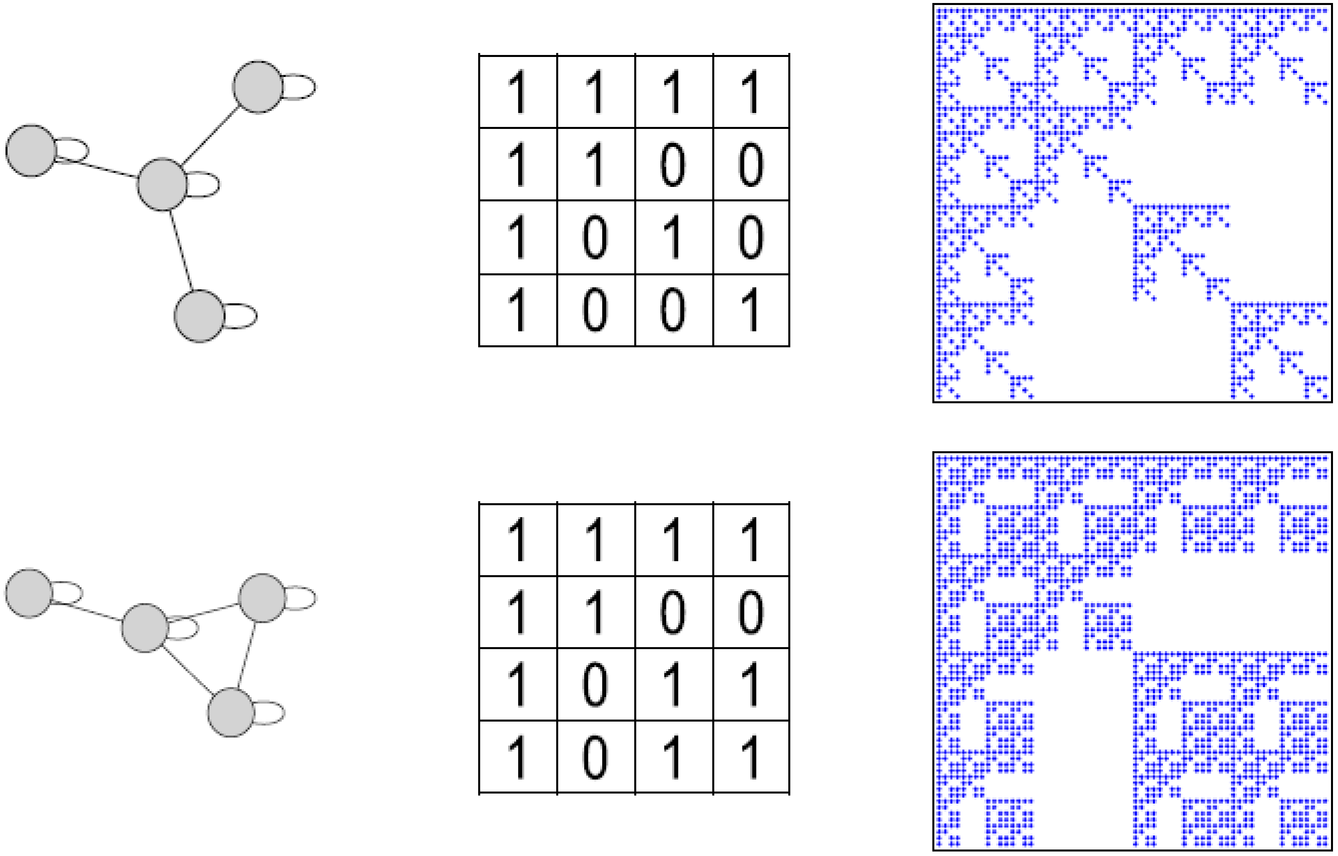
(e) Adjacency matrix of $K_2 = K_1 \otimes K_1$



(a) K_3 adjacency matrix (27×27)



(b) K_4 adjacency matrix (81×81)



Initiator K_1 | K_1 adjacency matrix | K_3 adjacency matrix

Theorem 5 (Multinomial degree distribution) *Kronecker graphs have multinomial degree distributions, for both in- and out-degrees.*

Theorem 6 (Multinomial eigenvalue distribution) *The Kronecker graph K_k has a multinomial distribution for its eigenvalues.*

Theorem 7 (Multinomial eigenvector distribution) *The components of each eigenvector of the Kronecker graph K_k follow a multinomial distribution.*

Theorem 12 *If K_1 has diameter D and a self-loop on every node, then for every k, the graph K_k also has diameter D.*

Definition 14 (Stochastic Kronecker graph) *Let $\mathcal{P}_1$ be a $N_1 \times N_1$* probability matrix*: the value $\theta_{ij} \in \mathcal{P}_1$ denotes the probability that edge (i, j) is present, $\theta_{ij} \in [0, 1]$.*

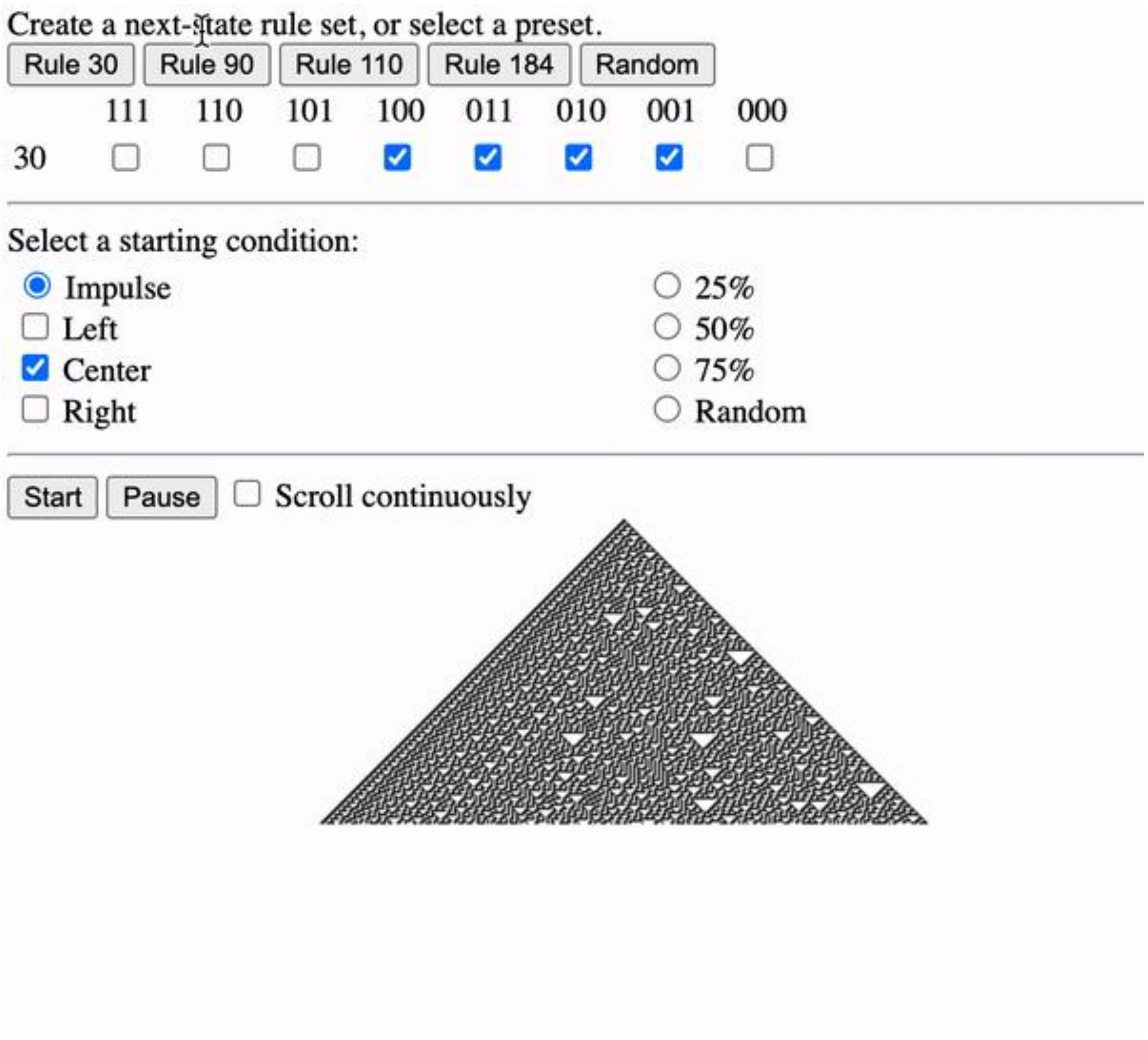
Then k^{th} Kronecker power $\mathcal{P}_1^{[k]} = \mathcal{P}_k$, where each entry $p_{uv} \in \mathcal{P}_k$ encodes the probability of an edge (u, v).

To obtain a graph, an instance *(or* realization*), $K = R(\mathcal{P}_k)$ we include edge (u, v) in K with probability p_{uv}, $p_{uv} \in \mathcal{P}_k$.*

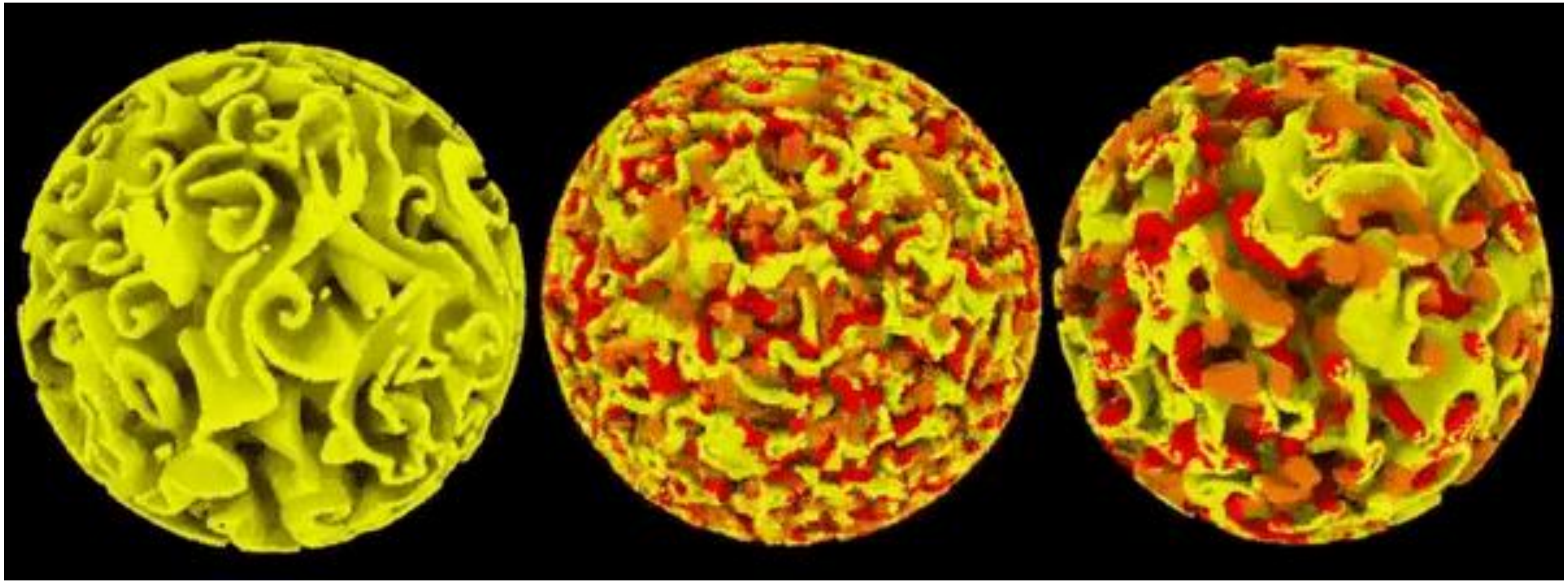
Cellular Automata

https://playgameoflife.com/

https://www.wolframscience.com/nks/p170--cellular-automata/

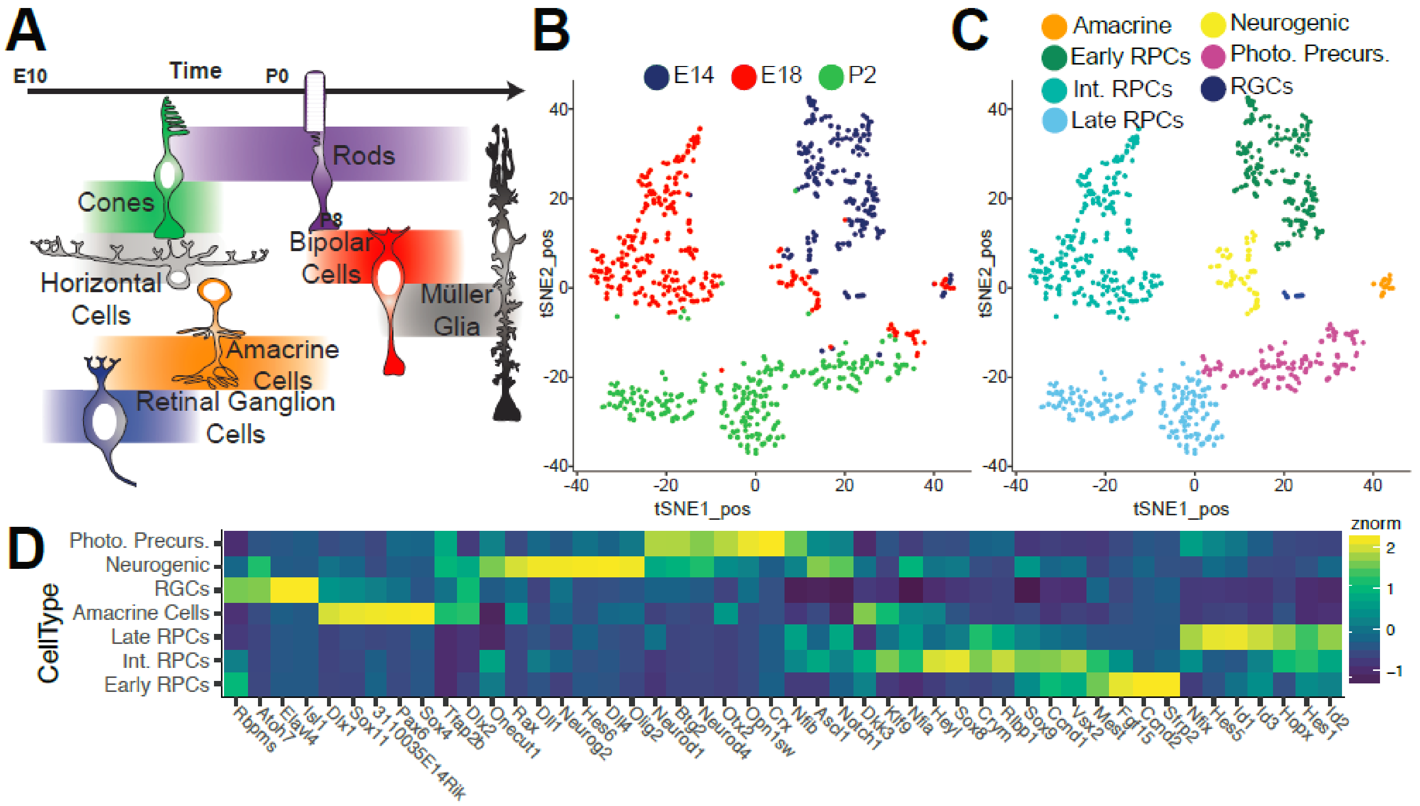


https://towardsdatascience.com/neural-cellular-automata-for-art-recreation-6d9fb61afb37



https://towardsdatascience.com/neural-cellular-automata-for-art-recreation-6d9fb61afb37

Information Lattice Learning: Learning laws of neurogenesis



[B. Clark, et al., "Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification," *Neuron*, June 2019.]

- Single-cell RNA sequence data analysis for understanding the rules that govern pattern formation in neurodevelopment

[Yu, Varshney, Stein-O'Brien, 2019]

Neural Cellular Automata

https://distill.pub/2020/growing-ca/