Generative AI Models
ECE 598 LV – Lecture 18

Lav R. Varshney
29 March 2022
Information Lattice Learning: Learning laws of neurogenesis

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

[Yu, Varshney, Stein-O’Brien, 2019]
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

\{\text{red, blue}\}
\{\text{convex, concave}\}
\{\text{trigon, tetragon, pentagon}\}
Information lattice

\{\text{red, blue}\}
\{\text{trigon, tetragon, pentagon}\}
\{\text{red trigon, blue trigon, red tetragon, blue tetragon, red pentagon, blue pentagon}\}
Abstraction universe as partition lattice

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

Pictorially, a directed acyclic graph (vertex: partition; edge: coarser than)

(more specific)  
\[\text{finer}\]

(coarser)  
\[\text{(more general)}\]
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:

\[
\begin{align*}
\text{maximize} & \quad D(p^{(k-1)}_{\phi,stu} \parallel \hat{p}_\phi) \\
\text{subject to} & \quad \phi \in \Phi \setminus \Phi^{(k-1)} \\
\text{(discrete optimization)} & \\
\end{align*}
\]

max. Bayesian surprise

The student solves:

\[
\begin{align*}
\text{maximize} & \quad S_q(p^{(k)}_{stu}) \\
\text{subject to} & \quad p^{(k)}_{stu} \in \Gamma_1 \\
& \quad \ldots \\
& \quad p^{(k)}_{stu} \in \Gamma_k \\
\text{(linear least-squares)} & \\
\end{align*}
\]

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

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

We then define the Kronecker product of two graphs simply as the Kronecker product of their corresponding adjacency matrices.
(d) Adjacency matrix of $K_1$

(e) Adjacency matrix of $K_2 = K_1 \otimes K_1$
(a) $K_3$ adjacency matrix ($27 \times 27$)

(b) $K_4$ adjacency matrix ($81 \times 81$)
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 $\mathcal{P}^{[k]}_1 = \mathcal{P}_k$, 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 $p_{uv}$, $p_{uv} \in \mathcal{P}_k$. 
https://playgameoflife.com/
https://www.wolframscience.com/nks/p170--cellular-automata/
Create a next-state rule set, or select a preset.

<table>
<thead>
<tr>
<th>Rule 30</th>
<th>Rule 90</th>
<th>Rule 110</th>
<th>Rule 114</th>
<th>Rule 184</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>110</td>
<td>101</td>
<td>100</td>
<td>011</td>
<td>010</td>
</tr>
<tr>
<td>001</td>
<td>000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

30

Select a starting condition:

- **Impulse**
- **Left**
- **Center**
- **Right**

- **25%**
- **50%**
- **75%**
- **Random**

[Start] [Pause] [Scroll continuously]
Information Lattice Learning: Learning laws of neurogenesis

- 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/