# Some Notes on
# Finite and Hyperfinite model theory

N. Ikodinović
ikodinovic@matf.bg.ac.rs

Dubrovnik, September, 2016.

# Three seminal papers

1. Gaifman: Concerning measures on first order calculi, Israel J. Math 2 (1964), 1-18
2. Fagin: Probabilities on finite models, J. Symb. Logic 41 (1976), 50-58
3. Keisler: Hyperfinite model theory, in: R. O. Gandy, J. M. E. Hyland (eds.) Logic Colloquim 76, North-Holland (1977), 5-110

# Specifying probabilities on $\mathbf{Sent}_L$

Gaifman: A probabilty on $\mathbf{Sent}_L$ is a function $P : \mathbf{Sent}_L \to [0,1]$ such that:

- $P(\alpha \vee \beta) + P(\alpha \wedge \beta) = P(\alpha) + P(\beta)$
- $P(\neg\alpha) = 1 - P(\alpha)$
- $P(\alpha) = P(\beta)$, if $\vdash \alpha \leftrightarrow \beta$
- $P(\alpha) = 1$, if $\vdash \alpha$.

A probability on $L$ can be conceived as a non-trivial, non-negative, finitely-additive probability measure on the Lindenbaum-Tarski algebra of the sentences of $L$.

# Specifying probabilities on $\mathbf{Sent}_L$

- Let $(M, R)$ be a classical structure:

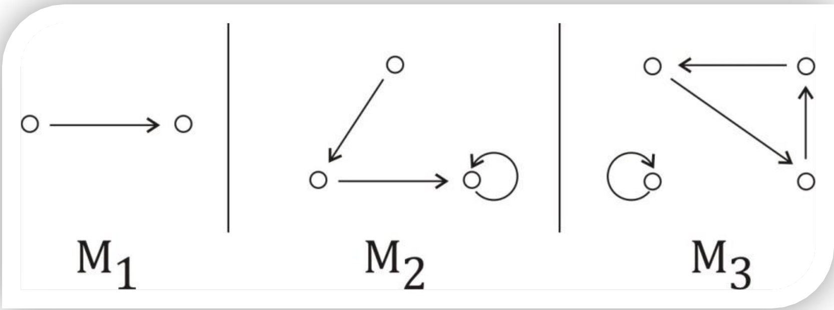
$$P_{\mathbf{M}} : \mathrm{Sent}_L \to [0,1],\ P_{\mathbf{M}}(\varphi) = \begin{cases} 1, \mathbf{M} \vDash \varphi \\ 0, \mathbf{M} \nvDash \varphi \end{cases}$$

- Let $\mathfrak{M}$ be a set of classical structures with a probability $m$ on the algebra of subsets of $\mathfrak{M}$ generated by $[\varphi]_{\mathfrak{M}} = \{\mathbf{M} \in \mathfrak{M} \mid \mathbf{M} \vDash \varphi\}$:

$$P_{\mathfrak{M}} : \mathrm{Sent}_L \to [0,1],\ P_{\mathfrak{M}}(\varphi) = \int_{\mathfrak{M}} P_{\mathbf{M}}([\varphi]_{\mathfrak{M}})\, dm(\mathbf{M})$$

# Specifying probabilities on $\mathbf{Sent}_L$

Example 1. $L = \{B\}, \mathrm{ar}(B) = 2$




$$\begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \mathbf{M}_3 \\ 5/8 & 1/4 & 1/8 \end{pmatrix}$$

$$P(\forall x \exists y\, Bxy) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}$$

$$P\big(\exists x \exists y\, (x \neq y \rightarrow Bxy)\big) = 1$$

# Probabilistic models

$$M \neq \emptyset, L(M) = \{B\} \cup \{c_a \mid a \in M\}$$

Classical model:

$B^M : M \times M \to \{0,1\}$

$c_a^M = a, a \in M$

$\mathcal{J} : \mathrm{Atomic}_{L(M)} \to \{0,1\}$

$\hat{\mathcal{J}}: \mathrm{Sent}_{L(M)} \to \{0,1\}$

Probabilistic models

$\mathcal{P} : \mathrm{Atomic}_{L(M)} \to [0,1]$

is not suficient (the values assigned to atomic sentences do not determine unique values for Boolean combinations of these sentences)!

# Probabilistic models

$M \neq \emptyset, L(M) = \{B\} \cup \{c_a \mid a \in M\}$

Classical model:

$B^M : M \times M \to \{0,1\}$

$c_a^M = a, a \in M$

$\mathcal{I} : \mathrm{Atomic}_{L(M)} \to \{0,1\}$

$\hat{\mathcal{I}}: \mathrm{Sent}_{L(M)} \to \{0,1\}$

Probabilistic models

$\mathcal{P} : \mathrm{Bool}_{L(M)} \to [0,1]$

# Gaifman's condition

Theorem. If $P : \mathrm{Sent}_{L(M)} \to [0,1]$ is a probability, then

(G) $P\big(\exists x \alpha(x)\big)$

$= \sup\{P\big(\alpha(a_1) \vee \cdots \vee \alpha(a_k)\big) \mid k \in \mathbb{N}, a_1, \ldots, a_k \in M\}$.

Theorem. Let $(M, \mathcal{P})$ be a probabilistic model (i.e., $\mathcal{P} : \mathrm{Bool}_{L(M)} \to [0,1]$). Then there is a unique probability $\mathcal{P}^*$ which extends $\mathcal{P}$ to $\mathrm{Sent}_{L(M)}$ and satisfies (G).

Theorem. Every probability on $\mathrm{Sent}_L$ has a probabilistic model whose power is $\aleph_0 + |\mathrm{Sent}_L|$.

# Probabilistic models

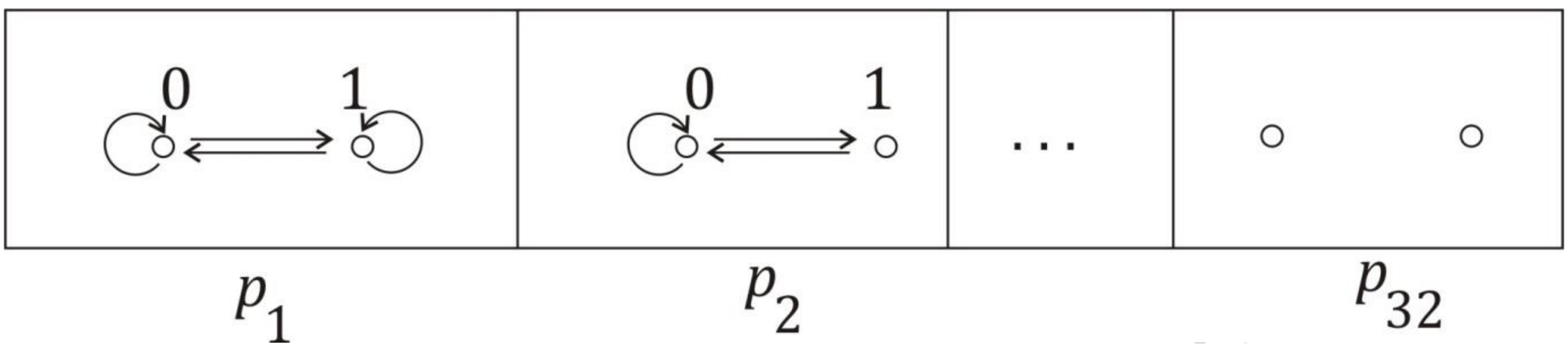
Example 2. $M = \{0,1\}$

$$\mathcal{P}(B00 \wedge B01 \wedge B10 \wedge B11) = p_1$$
$$\mathcal{P}(B00 \wedge B01 \wedge B10 \wedge \neg B11) = p_2$$
$$\vdots$$
$$\mathcal{P}(\neg B00 \wedge \neg B01 \wedge \neg B10 \wedge \neg B11) = p_{32}$$
$$p_1 + p_2 + \cdots + p_{32} = 1$$

$$\mathcal{P} : \mathrm{Sent}_{L(M)} \to [0,1]$$

# Probabilistic models

Example 2. $M = \{0,1\}$




$$p_1 + p_2 + \cdots + p_{32} = 1$$

$$\mathcal{P} : \mathrm{Sent}_{L(M)} \to [0,1]$$

# Probabilistic models

Example 3. $B : \{0,1\}^2 \to [0,1], \begin{pmatrix} B00 & B01 & B10 & B11 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$

$$B00 \wedge B01 \wedge B10 \wedge B11 \mapsto \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{6}$$

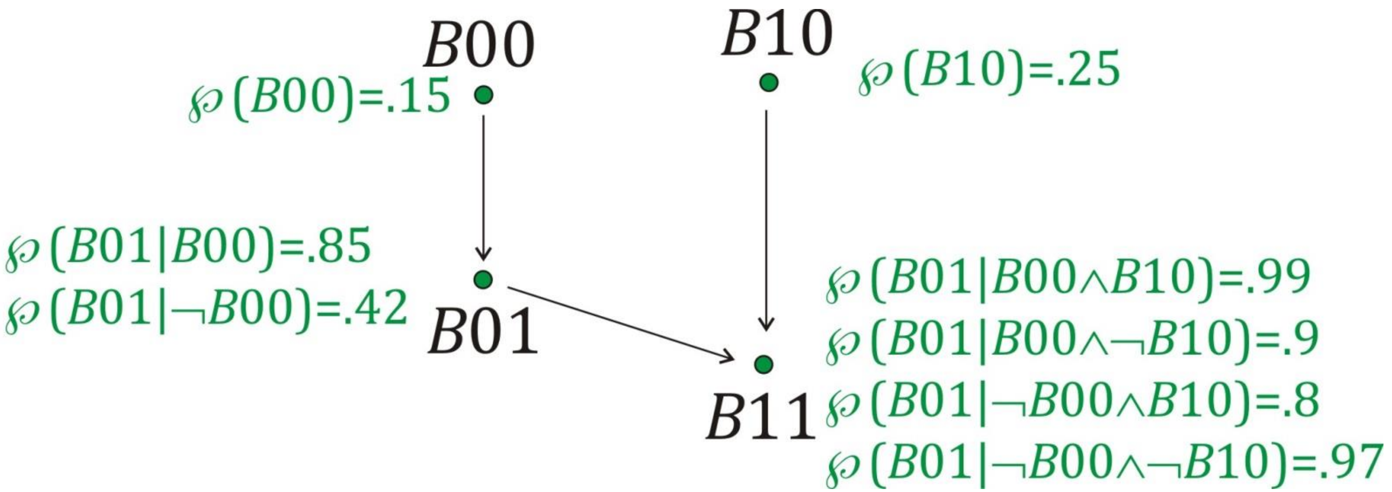
$$B00 \wedge B01 \wedge B10 \wedge \neg B11 \mapsto \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{5}{6}$$

$\vdots$

$$\neg B00 \wedge \neg B01 \wedge \neg B10 \wedge \neg B11 \mapsto \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{5}{6}$$

# Probabilistic models

Example 4.




$$\mathcal{P}(B00 \wedge \neg B01 \wedge \neg B10 \wedge B11)$$
$$= \mathcal{P}(B00) \cdot \mathcal{P}(\neg B01 \mid B00) \cdot \mathcal{P}(\neg B10) \cdot \mathcal{P}(B11 | B00 \wedge \neg B01)$$
$$= 0.15 \cdot (1 - 0.85) \cdot (1 - 0.25) \cdot 0.9$$

# Random structure $M(n,p)$

Erdős, Rényi: On the Evolution of the Random graph, Mat. Kutató Int. Közl 5 (1960), 17-60

$M = \{0,1,\ldots,n-1\},\ 0 \leq p \leq 1$

Define $\mathcal{P} : \mathrm{Bool}_{L(M)} \to [0,1]$ by:

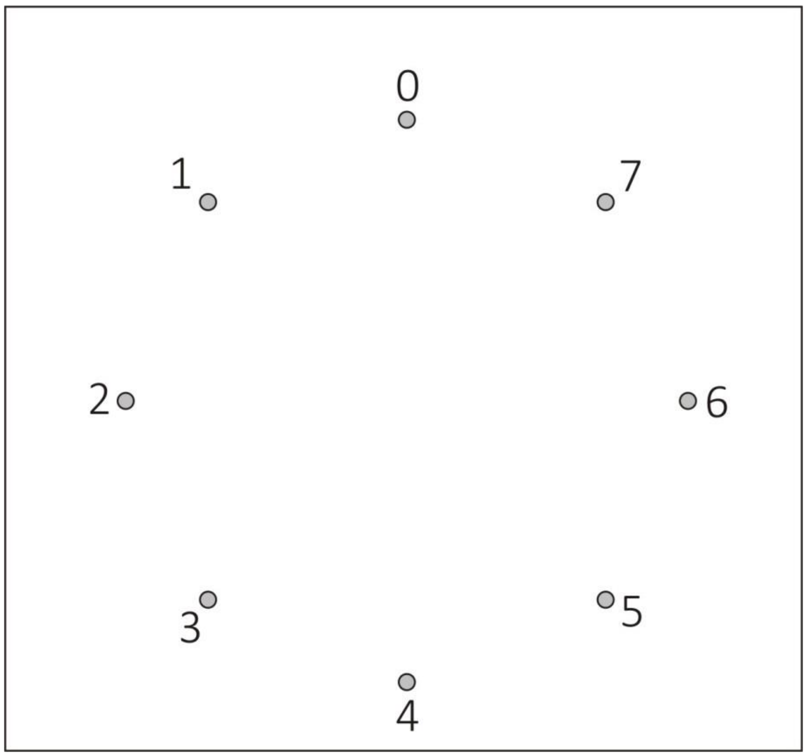
$\mathcal{P}(Bab) = p, a, b \in M$, and these events are mutually independent.

$$\mathcal{P}\left(\bigwedge_{i=1}^{k} Ba_i b_i \wedge \bigwedge_{j=k+1}^{n} \neg Ba_j b_j\right) = p^k (1-p)^{n-k},$$

where sentences $Ba_i b_i$ and $Ba_j b_j$ are not identical.

# Random structure $M(8,0.5)$

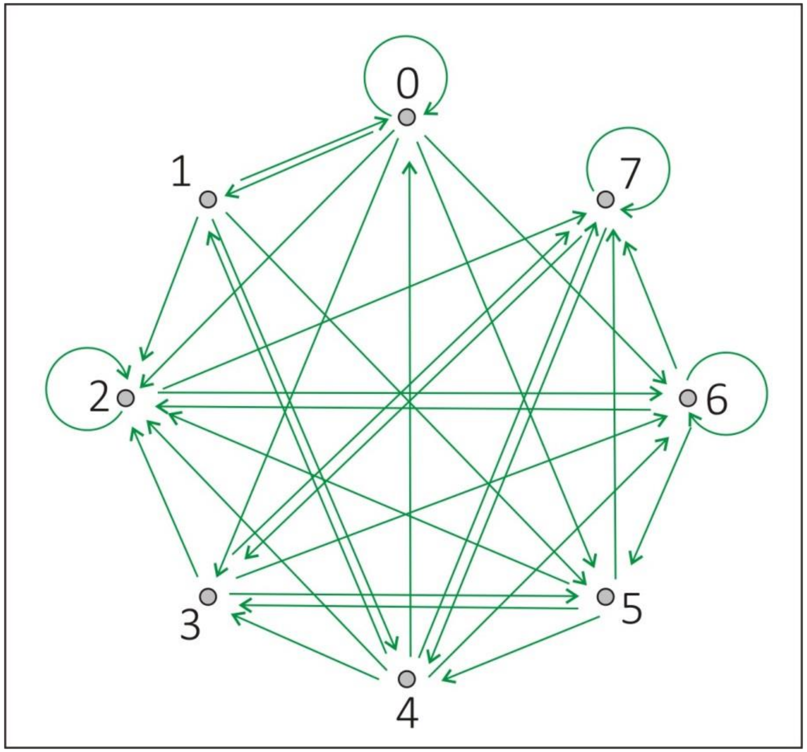
Example 5.




| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | H | H | H | H | T | H | H | T |
| 1 | H | T | H | T | H | H | T | T |
| 2 | T | T | H | T | T | T | H | T |
| 3 | T | T | H | T | T | H | H | H |
| 4 | H | H | H | H | T | T | H | T |
| 5 | T | T | T | H | H | T | T | H |
| 6 | T | T | H | T | T | H | H | H |
| 7 | T | T | T | H | H | T | T | H |

# Random structure $M(8,0.5)$

Example 5.




| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | H | H | H | H | T | H | H | T |
| 1 | H | T | H | T | H | H | T | T |
| 2 | T | T | H | T | T | T | H | T |
| 3 | T | T | H | T | T | H | H | H |
| 4 | H | H | H | H | T | T | H | T |
| 5 | T | T | T | H | H | T | T | H |
| 6 | T | T | H | T | T | H | H | H |
| 7 | T | T | T | H | H | T | T | H |

# Gaifman's simple example

Example 6: Let $M$ an infinite set, and $p$ a fixed number, $0 < p < 1$.

Define $\mathcal{P} : \mathrm{Bool}_{L(M)} \to [0,1]$ by:

- $\mathcal{P}(Bab) = p$, $a, b \in M$,
- $\mathcal{P}\left(\bigwedge_{i=1}^{k} Ba_i b_i \wedge \bigwedge_{j=k+1}^{n} \neg Ba_j b_j\right) = p^k (1-p)^{n-k}$,

where sentences $Ba_i b_i$ and $Ba_j b_j$ are not identical,

- $\mathcal{P}(a = b) = 0$, $a \neq b$.

Let $\mathcal{P}^* : \mathrm{Sent}_{L(M)} \to [0,1]$ be the extension of $\mathcal{P}$ satisfying **(G)**.

# Gaifman's simple example

Lemma. If $\alpha, \beta \in \mathrm{Sent}_{L(M)}$ and no constant occurs both in $\alpha$ and $\beta$ then

$$\mathcal{P}^*(\alpha \wedge \beta) = \mathcal{P}^*(\alpha) \cdot \mathcal{P}^*(\beta).$$

In particular, for every $\sigma \in \mathrm{Sent}_L$

$$\mathcal{P}^*(\sigma) = \mathcal{P}^*(\sigma \wedge \sigma) = (\mathcal{P}^*(\sigma))^2,$$

and hence $\mathcal{P}^*(\sigma)$ is either 0 or 1.

$(M, \mathcal{P})$ determines a complete theory!

# The complete theory

Let $V_n = \{v_0, v_1, \dots, v_{n-1}\}$ be a set of distinct variables.

A complete diagram $\sigma(V_n)$ is a conjunction s.t. for every pair $(v_i, v_j) \in V_n \times V_n$ either $Bv_iv_j$ or $\neg Bv_iv_j$ is a conjunct.

$$\sigma(V_3) = Bv_0v_0 \wedge \neg Bv_0v_1 \wedge Bv_0v_2 \wedge$$
$$\wedge \neg Bv_1v_0 \wedge Bv_1v_1 \wedge Bv_1v_2 \wedge$$
$$\wedge Bv_2v_0 \wedge \neg Bv_2v_1 \wedge \neg Bv_2v_2$$

# The complete theory

A complete diagram $\sigma(V_{n+1})$ extends another complete diagram $\sigma'(V_n)$ if $\vdash \sigma'(V_n) \to \sigma(V_{n+1})$.

$$\sigma(V_3) = Bv_0v_0 \wedge \neg Bv_0v_1 \wedge Bv_0v_2 \wedge$$
$$\wedge \neg Bv_1v_0 \wedge Bv_1v_1 \wedge Bv_1v_2 \wedge$$
$$\wedge Bv_2v_0 \wedge \neg Bv_2v_1 \wedge \neg Bv_2v_2$$

extends

$$\sigma'(V_2) = Bv_0v_0 \wedge \neg Bv_0v_1 \wedge$$
$$\wedge \neg Bv_1v_0 \wedge Bv_1v_1$$

# The complete theory

Let $E$ be the set of all sentences $\varepsilon_n$:

$$\forall v_0 \dots v_{n-1} \left( \bigwedge_{i \neq j} v_i \neq v_j \wedge \sigma(V_n) \right.$$

$$\left. \rightarrow \exists v_n \bigwedge_i v_i \neq v_n \wedge \sigma'(V_{n+1}) \right)$$

Theorem. $E$ is consistent and complete.

# The complete theory

The theory $E$ has exactly one, up to isomorphism, countable model – the Rado graph.

That model is `most general’ in the sense that:

- every possible finite model is realized there as a submodel, and
- for every finite submodel every possible finite extension of it is realized.

# Probabilities on finite models

What is the probability that a certain sentence holds for a randomly chosen finite structure?

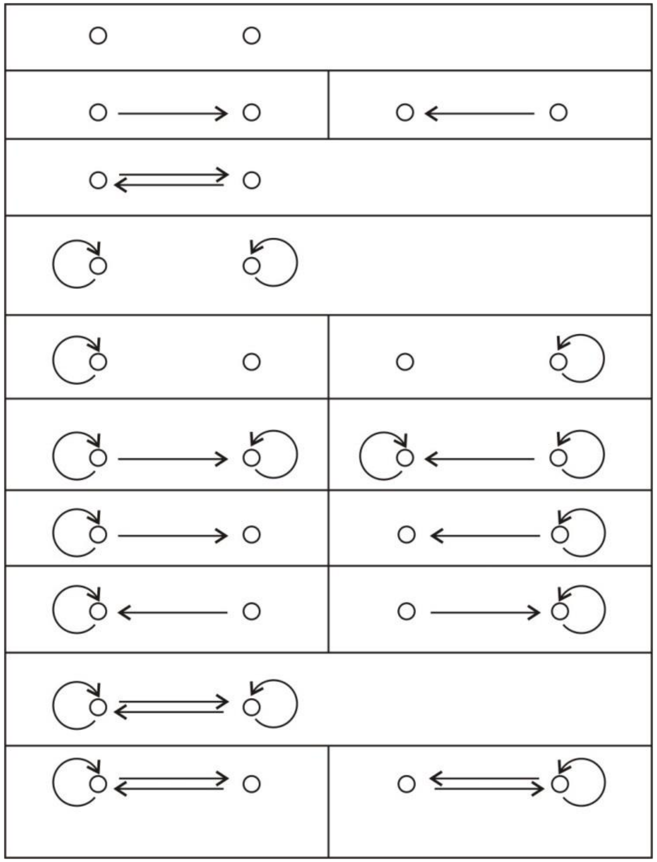
$\mathfrak{M}_n$ the class of all $n$-structures (on $\{0,1,\dots,n-1\}$).

$\mathfrak{M}_n[\sigma] = \{\mathbf{M} \in \mathfrak{M}_n \mid \mathbf{M} \vDash \sigma\}$

$$P_n(\sigma) = \frac{|\mathfrak{M}_n[\sigma]|}{|\mathfrak{M}_n|}$$

# Specifying probabilities on $\mathbf{Sent}_L$

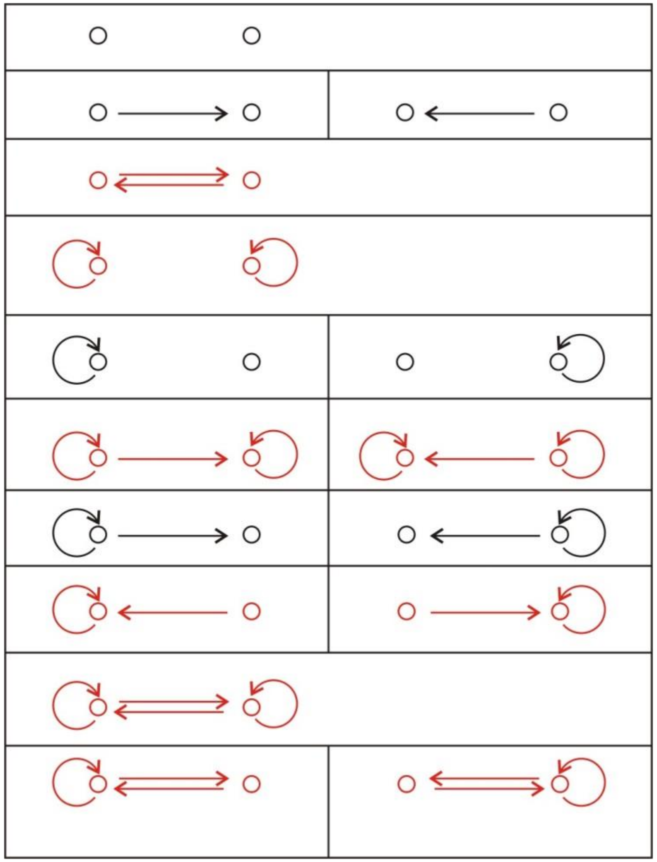
Example 7.



$P_2(\forall x \exists y\ Bxy) = ?$

# Specifying probabilities on $\mathbf{Sent}_L$

Example 7.

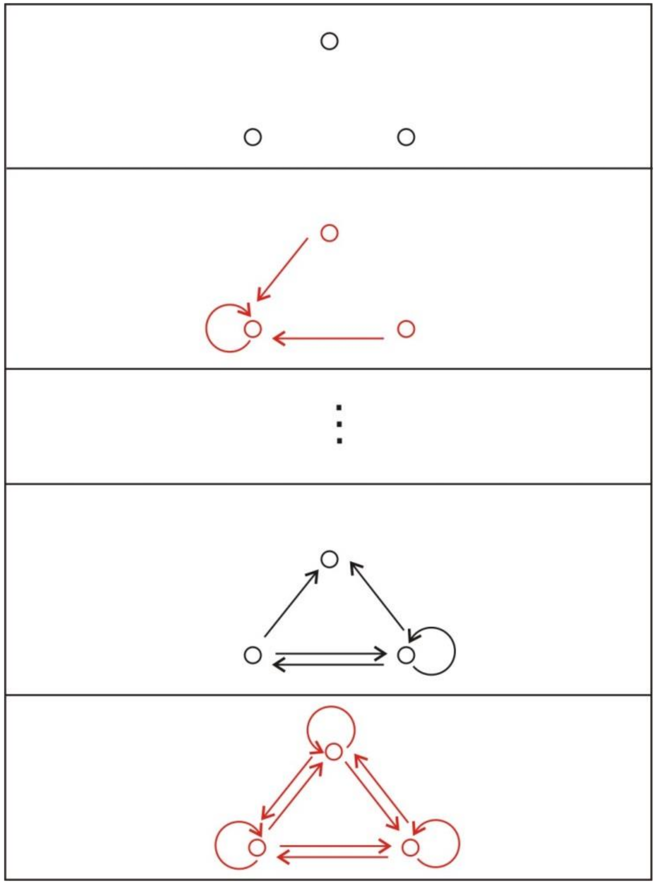


$$P_2(\forall x \exists y\, Bxy) = ?$$

$$P_2(\forall x \exists y\, Bxy) = \frac{9}{16} \approx 0.56$$

# Specifying probabilities on $\mathbf{Sent}_L$

Example 7.



$$P_3(\forall x \exists y\, Bxy) = ?$$

# 0-1 law

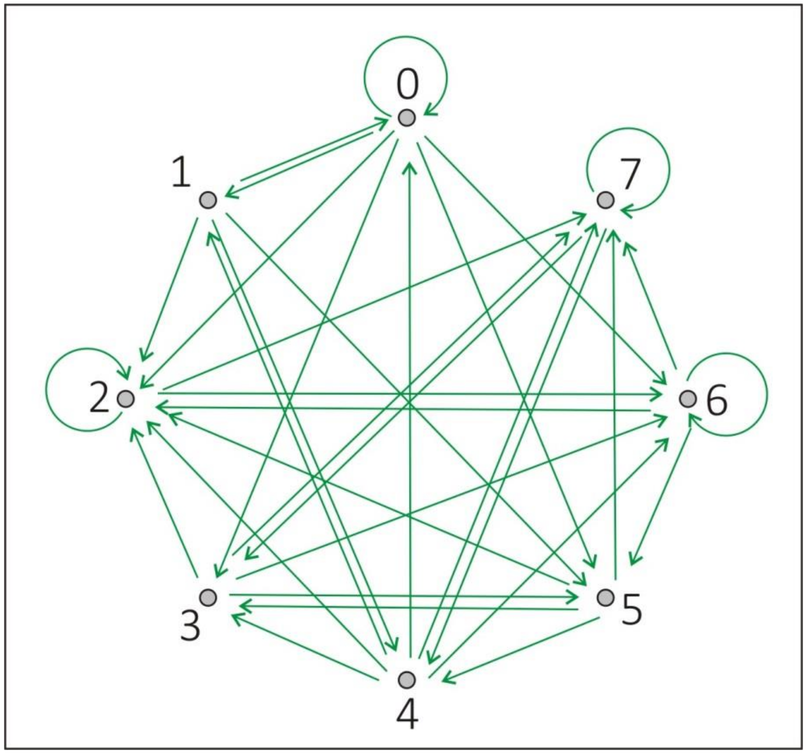
$$P_n(\sigma) = \frac{|\mathfrak{M}_n[\sigma]|}{|\mathfrak{M}_n|} \rightarrow ?, n \rightarrow \infty$$

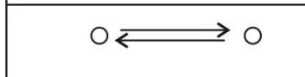
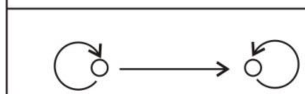
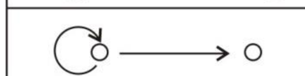
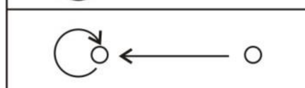
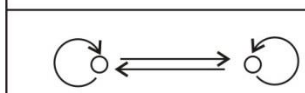
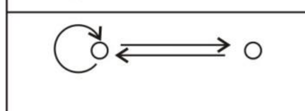
The zero-one law.

$P_n(\sigma)$ converges to either $0$ or $1$.

# Random structure $M = (8,0.5)$

Example 8.

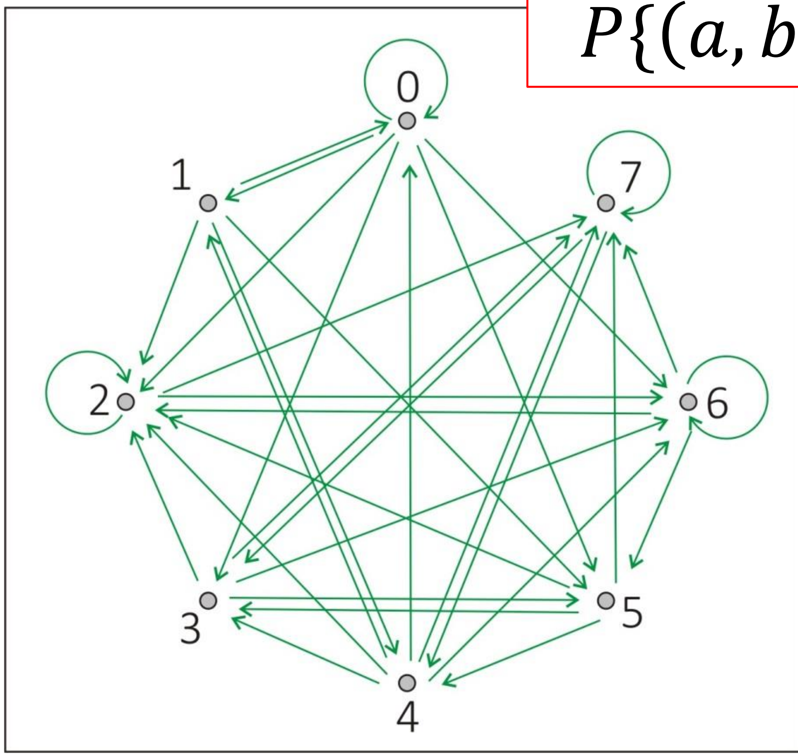



| | |
|---|---|
|  | 13 |
|  | 15,34, 45 |
|  | 14,35 |
|  | 07 |
|  | 16,17 |
|  | 02,27,67 |
|  | 03,05,25,56,57 |
|  | 04,12,23,24,36,46 |
|  | 26 |
|  | 01,37,47 |

# Random graph

$\sigma(v_0, v_1) = (Bv_0v_0 \vee Bv_0v_1) \wedge (Bv_1v_0 \vee Bv_1v_1)$

$P\{(a, b) \in M \times M \mid \mathbf{M} \vDash \sigma[a, b]\} = ?$



# Random graph

$$\sigma(v_0, v_1) = (Bv_0v_0 \vee Bv_0v_1) \wedge (Bv_1v_0 \vee Bv_1v_1)$$

$$P\{(a,b) \in M \times M \mid \mathbf{M} \vDash \sigma[a,b]\} =?$$

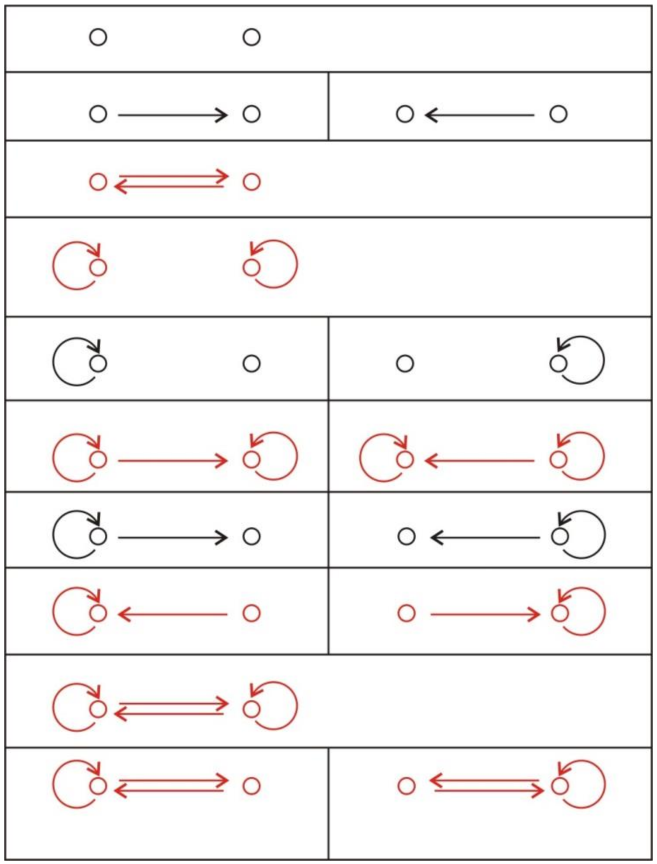
| | |
|---|---|
|  | 13 |
|  | 15,34, 45 |
|  | 14,35 |
|  | 07 |
|  | 16,17 |
|  | 02,27,67 |
|  | 03,05,25,56,57 |
|  | 04,12,23,24,36,46 |
|  | 26 |
|  | 01,37,47 |

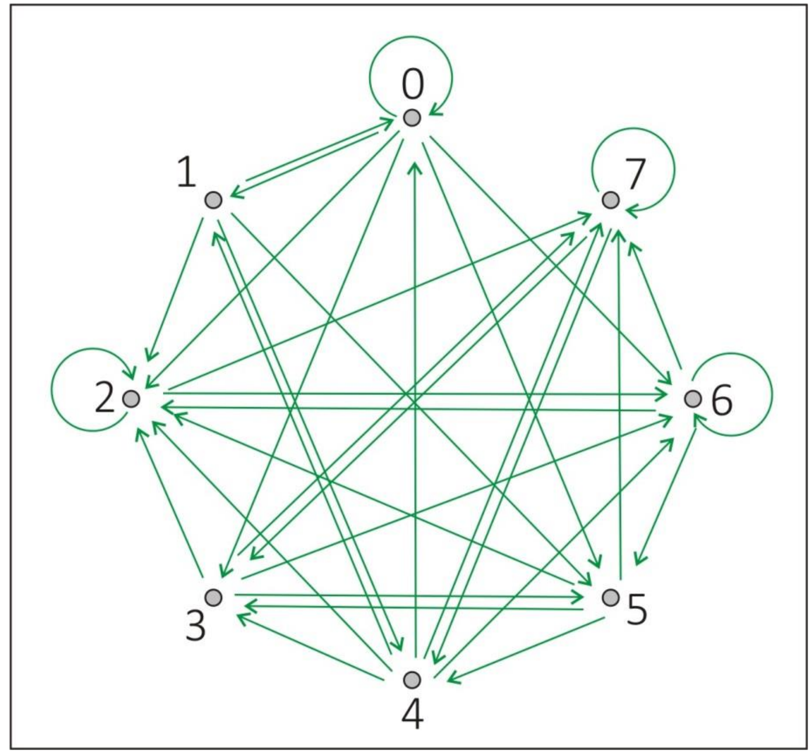
$$\ldots \vDash \forall x \exists y\, Bxy$$

$$= \frac{16}{\binom{8}{2}} = \frac{16}{28} \approx 0.57$$

Count models!

Count pairs!






$P_2(\forall x \exists y\ Bxy) \approx 0.56$

$P\{\sigma(v_0, v_1)\} \approx 0.57$

# Count models vs. count tuples

Intuition: If $H \gg n$, then $(H, 0.5)$ contains (almost) all $n$-structures as its substructures. Moreover, the distribution of $n$-structures inside $(H, 0.5)$ is (almost) uniform.

# Hyperfinite model theory

Keisler: The purpose of hyperfinite model theory is to study and classify a type of models which arises in applied mathematics.

··· [These] models have usually been either countable sequences of finite models or structures built upon the real numbers. Hyperfinite models provide a better source of infinite models which closely approximate large finite phenomena.

··· Hyperfinite models deal with limiting behavior of finite models.

# Hyperfinite model theory

Definition. A hyperfinite probability space is a pair $(M, \mu)$ where $A$ is a nonempty hyperfinite set and $\mu$ is an internal function $\mu : M \rightarrow {}^*[0,1]$ such that $\sum_{a \in M} \mu(a) = 1$.

$\langle \mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3, \dots \rangle_{\mathcal{U}}$ - a hyperfinite set

$\langle \mathfrak{M}_1[\sigma], \mathfrak{M}_2[\sigma], \mathfrak{M}_3[\sigma], \dots \rangle_{\mathcal{U}} \subseteq \langle \mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3, \dots \rangle_{\mathcal{U}}$

$\left\langle \frac{|\mathfrak{M}_1[\sigma]|}{|\mathfrak{M}_1|}, \frac{|\mathfrak{M}_2[\sigma]|}{|\mathfrak{M}_2|}, \frac{|\mathfrak{M}_2[\sigma]|}{|\mathfrak{M}_2|}, \dots \right\rangle_{\mathcal{U}} \in {}^*[0,1]$

# Probability logic $L_{\omega P}$

Probability logic $L_{\omega P}$ is like first-order logic, but instead of ∀ and ∃ it has probability quantifiers $(P\vec{x} > r)$, $r \in \mathbb{Q}$.

$$\mathbf{M} \vDash (Px > r)\varphi(x) \text{ iff } \mu\{x \in M \mid \mathbf{M} \vDash \varphi(\vec{x})\} > r$$

# Logic and Probability

| First-order logic | Probability logic | Probability theory |
|---|---|---|
| Downward Löwenheim- Skolem theorem | Elementary submodel theorem | Weak law of large numbers |
| The zero-one law | Elementary subsequence theorem | Strong law of large numbers |

Thank you for your attention