Finite and Algorithmic Model Theory

Lecture 5 (Dresden 09.11.22, Short version with Errors)

Lecturer: Bartosz "Bart" Bednarczyk

Technische Universität Dresden & Uniwersytet Wrocławski

European Research Council
Established by the European Commission

Today's agenda

Goal: Prove that Ehrenfeucht-Fraïssé games works + Simplification of E-F games with Hanf's locality

1. Recap of Ehrenfeucht-Fraïssé games.

2. Back-and-Forth Equivalence with threshold m. Notation: ($\mathfrak{A} \simeq_m \mathfrak{B}$).

$\mathfrak{A} \simeq_m \mathfrak{B}$ iff Duplic∀or has winning strategy in m-round E-F games on $\mathfrak{A}$ and $\mathfrak{B}$.

3. Hintikka formulae, i.e. describing the m-isomorphism type of a τ-structure $\mathfrak{A}$ with an $\mathrm{FO}_m[\tau]$ formula.

$$\mathfrak{A} \simeq_m \mathfrak{B} \text{ iff } \mathfrak{B} \models \varphi^{\mathfrak{A},m}_{\text{Hintikka}}.$$

4. Gaifman Graphs and r-neighbourhoods

5. Examples of Hanf(r, t)-equivalent structures.

6. Hanf's theorem + applications to inexpressivity in FO.

7. Proof of Hanf's theorem.

Lecture based on

Chapter 3.5 of [Libkin's Book]

Slides 29–33, 43–51 of [Montanari]

19:23-24:32 of lecture by [Anuj Dawar]

Slides 80–110 by [Diego Figueira]

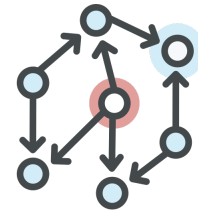
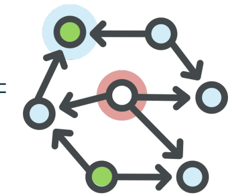
Feel free to ask questions and interrupt me!

Don't be shy! If needed send me an email (bartosz.bednarczyk@cs.uni.wroc.pl) or approach me after the lecture!

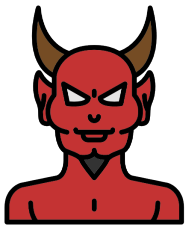
Reminder: this is an advanced lecture. Target: people that had fun learning logic during BSc studies!

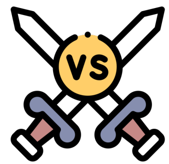
Recap of Ehrenfeucht-Fraïssé games

- Duration: m rounds.
- Playground: two τ-structures $\mathfrak{A}$ and $\mathfrak{B}$.

$\mathfrak{A} :=$  $\mathfrak{B} :=$ 

- Two players: Spoil∃r (D∃vil/∃loise/∃ve/Player I) vs Duplic∀tor (∀ngel/∀belard/∀dam/Player II)





Goal of $\forall$: $\mathfrak{A}, \mathfrak{B}$ "look the same".

Goal of $\exists$: pinpoint the difference.

- During the i-th round:

1. $\exists$ selects a structure (say $\mathfrak{A}$) and picks an element (say $a_i \in A$)

2. $\forall$ replies with an element (say $b_i \in B$) in the other structure (in this case $\mathfrak{B}$)

so that $(a_1 \mapsto b_1, \ldots, a_i \mapsto b_i)$ is a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

- $\exists$ wins if $\forall$ cannot reply with a suitable element. $\forall$ wins if he survives m rounds.

Theorem (Fraïssé 1954 & Ehrenfeucht 1961)

$\forall$ has a winning strategy in m-round Ehrenfeucht-Fraïssé game on τ-structures $\mathfrak{A}$ and $\mathfrak{B}$ iff $\mathfrak{A} \equiv_m^\tau \mathfrak{B}$.

Back and Forth Equivalence (a.k.a. Bisimulations)

We define an FO-m-bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$ as the relation $\mathcal{Z} \subseteq \bigcup_{i=0}^{m} A^i \times B^i$ with $(\varepsilon, \varepsilon) \in \mathcal{Z}$ fulfilling:

- **(atomic harmony):** $\mathfrak{A}\restriction_{\overline{a}} \cong \mathfrak{B}\restriction_{\overline{b}}$
- **(forth):** if $|\overline{a}| < m$, then for all $c \in A$, there is $d \in B$ such that $(\overline{a}c, \overline{b}d) \in \mathcal{Z}$.
- **(back):** if $|\overline{b}| < m$, then for all $d \in B$, there is $c \in A$ such that $(\overline{a}c, \overline{b}d) \in \mathcal{Z}$.

From m-round E-F Games to m-bisimulations

Take $\mathcal{Z} := \Big\{(\overline{a}_{1\ldots i}, \overline{b}_{1\ldots i}) \mid 1 \leq i \leq m, \text{ and } (\overline{a}, \overline{b}) \text{ is a history of the winning play of } \forall \text{ in } m\text{-round E-F game}\Big\}$.

From m-bisimulations to m-round E-F Games

Play as Duplicator, employing witnesses guaranteed by **(forth)** and **(back)** conditions.

Bisimulation as a more general concept

- One can define bisimulations $\simeq^{\mathsf{L}}_{\omega}$ (for ω rounds) for any logic L, e.g. Modal/Descr./Temporal logics.
- An abstract categorical and comonadic approaches: [Joyal et al.'1994] and [Abramsky'2022].
- Van-Benthem Theorems for L $\subseteq$ FO: φ is preserved under $\simeq^{\mathsf{L}}_{\omega}$ iff φ is equiv. to some $\psi \in \mathcal{L}$.

m-Hintikka formulae

Goal: describe the m-isomorphism type of a τ-structure $\mathfrak{A}$ with an $\mathrm{FO}_m[\tau]$ formula.

Fix a structure $\mathfrak{A}$, a k-tuple $\overline{a}$ from A, and a k-tuple of variables $\overline{x}$. Define $\varphi^k_{(\mathfrak{A},\overline{a})}(\overline{x})$ inductively as

- **(Base):** $\varphi^0_{(\mathfrak{A},\overline{a})}(\overline{x}) := \underbrace{\bigwedge_{\text{atomic } \lambda(\overline{x}),\, \mathfrak{A}\models\lambda(\overline{a})} \lambda(\overline{x}) \quad \wedge \bigwedge_{\text{atomic } \lambda(\overline{x}),\, \mathfrak{A}\not\models\lambda(\overline{a})} \neg\lambda(\overline{x})}_{\text{atomic harmony}}$

- **(Step):** $\varphi^k_{(\mathfrak{A},\overline{a})}(\overline{x}) := \underbrace{\bigwedge_{c\in A} \exists x_k \; \varphi^{k-1}_{(\mathfrak{A},\overline{a}c)}(\overline{x}, x_k)}_{\text{forth: responses for challenges in } \mathfrak{A}} \quad \wedge \quad \underbrace{\forall x_k \bigvee_{c\in A} \varphi^{k-1}_{(\mathfrak{A},\overline{a}c)}(\overline{x}, x_k)}_{\text{back: responses for challenges in } \mathfrak{B}}$

Call $\varphi^m_{(\mathfrak{A},\varepsilon)}$ the m-Hintikka formula. Goal: $\mathfrak{B} \models \varphi^m_{(\mathfrak{A},\varepsilon)}$ iff there is an m-bisimulation $\mathcal{Z}$ between $\mathfrak{A}$ and $\mathfrak{B}$.

Proof ($\Leftarrow$) [We leave ($\Rightarrow$) as an exercise.]

Induction over k. Assumption: For any $(\overline{a}, \overline{b}) \in \mathcal{Z}$ with $|\overline{a}|{=}|\overline{b}|{=}m{-}k$ we have $\mathfrak{B} \models \varphi^i_{(\mathfrak{A},\overline{a})}(\overline{b})$.

For $k = 0$ we are done by **(atomic harmony)**. For $k > 0$, take $(\overline{a}, \overline{b}) \in \mathcal{Z}$ with $|\overline{a}| = |\overline{b}| = m{-}k{-}1$.

Take any $c \in A$. By **(forth)** there is $d \in B$ so that $(\overline{a}c, \overline{b}d) \in \mathcal{Z}$. By ind. ass. $\mathfrak{B} \models \varphi^i_{(\mathfrak{A},\overline{a}c)}(\overline{b}d)$.

Thus $\mathfrak{B} \models \exists x_i \; \varphi^k_{\overline{a}c}(\overline{b}, x_i)$. By the choice of c, we conclude $\mathfrak{B} \models \bigwedge_{c\in A} \exists x_i \; \varphi^k_{\overline{a}c}(\overline{b}, x_i)$.

By reasoning similarly and employing **(back)**, we conclude the satisfaction of the RHS of $\varphi^k_{\overline{a}c}(\overline{b})$. $\square$

Main theorem about Ehrenfeucht-Fraïssé games

Lemma: For any τ-structures $\mathfrak{A}, \mathfrak{B}$ and $m \in \mathbb{N}$, the following are equivalent:

1. Duplic∀tor has the winning strategy in any m-round Ehrenfeucht-Fraïssé game played on $\mathfrak{A}$ and $\mathfrak{B}$.
2. There exists an m-bisimulation between $\mathfrak{A}$ and $\mathfrak{B}$.
3. $\mathfrak{B}$ satisfies the m-Hintikka formulae constructed from $\mathfrak{A}$.
4. $\mathfrak{A}$ and $\mathfrak{B}$ agree on all $\mathrm{FO}_m[\tau]$ sentences.

We've already seen that **(1)** $\Leftrightarrow$ **(2)** and **(2)** $\Leftrightarrow$ **(3)**. Clearly **(4)** $\Rightarrow$ **(3)**, thus it suffices to show **(2)** $\Rightarrow$ **(4)**.

Proof [**(2)** $\Rightarrow$ **(4)** by induction] Let $\mathcal{Z}$ be an m-bisimulation. The case $m = 0 \rightsquigarrow$ **(atomic harmony)**.

Note that every $\mathrm{FO}_m[\tau]$ formula is a boolean combination of formulae of the form $\exists x\ \psi$.

So it suffices to show the lemma for $\exists x\ \psi$ with $\mathrm{qr}(\varphi) \leq m{-}1$. Let $\mathfrak{A} \models \exists x\ \psi$. (Case with $\mathfrak{B}$ is symmetric).

Take $a \in A$ such that $\mathfrak{A} \models \psi(a)$. By **(forth)** we get $b \in B$ for which $(a, b) \in \mathcal{Z}$.

By ind. ass. b in $\mathfrak{B}$ satisfies the same $\mathrm{qr}(m{-}1)$-sentences as a in $\mathfrak{A}$. So $\mathfrak{B} \models \psi(b)$. Thus $\mathfrak{B} \models \exists x\ \psi$. $\square$

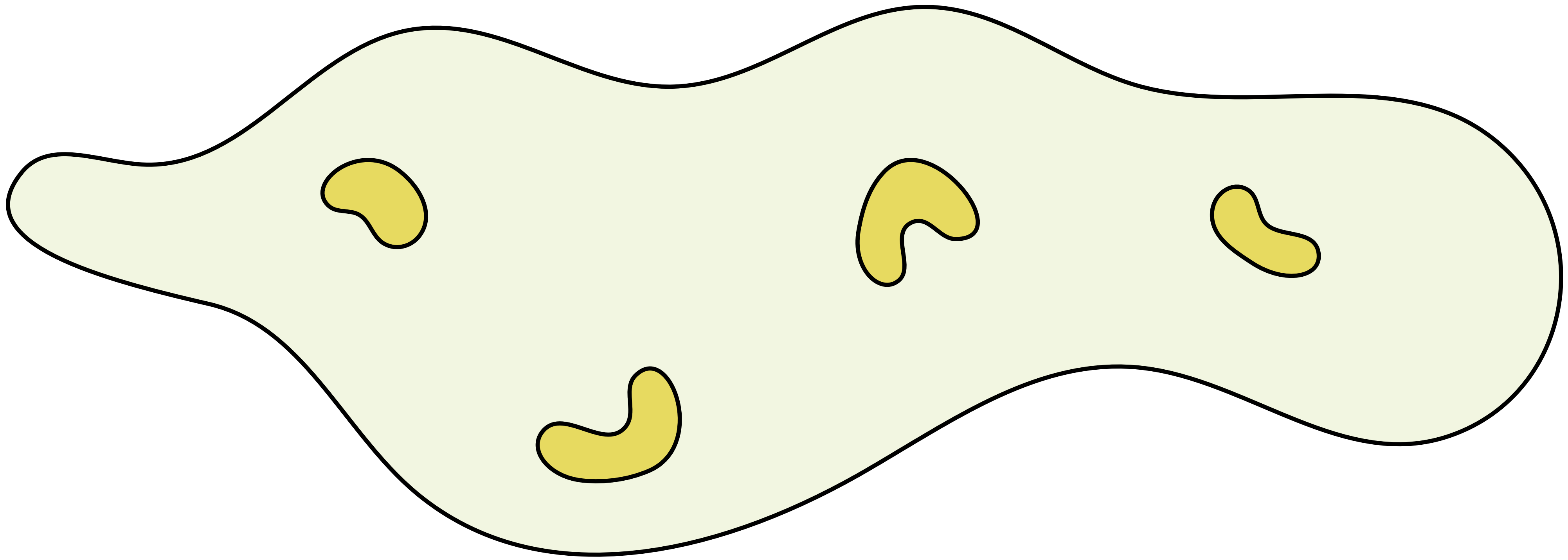
induction | atomic harmony | Simplify $\mathrm{FO}_m[\tau]$ | reduce | intro | witness | forth | ind. ass. | conclude

We will now go through slides 78–110 from ESSLI 2016 by [Diego Figueira].

Another technique: Locality

Idea: First order logic can only express "local" properties

Local = properties of nodes which are close to one another



Hanf locality

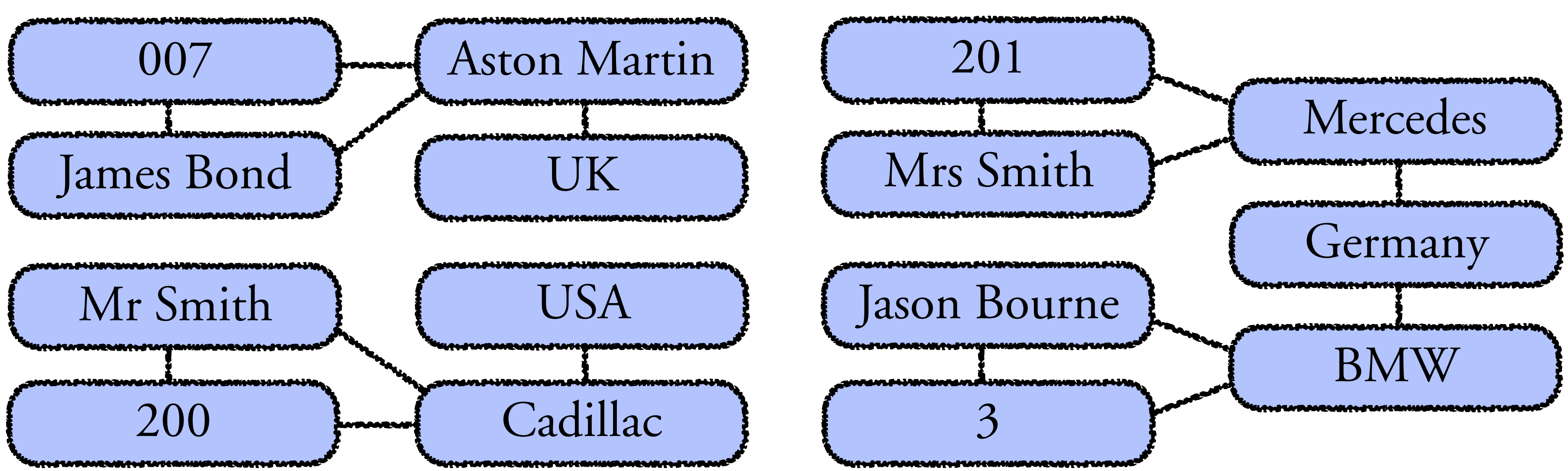
Definition. The **Gaifman graph** of a structure $S = (V, R_1, \dots, R_m)$ is the **undirected** graph

$$G_S = (V, E) \quad \text{where} \quad E = \{\, (u,v) \mid \exists\, (\dots, u, \dots, v, \dots) \in R_i \text{ for some } i \,\}$$

Agent	Name	Drives
007	James Bond	Asto
200	Mr Smith	Cadi
201	Mrs Smith	Mercedes
3	Jason Bourne	BMW

Car	Country
	UK
	USA
Mercedes	Germany
BMW	Germany

The Gaifman graph of a graph G is the underlying undirected graph.

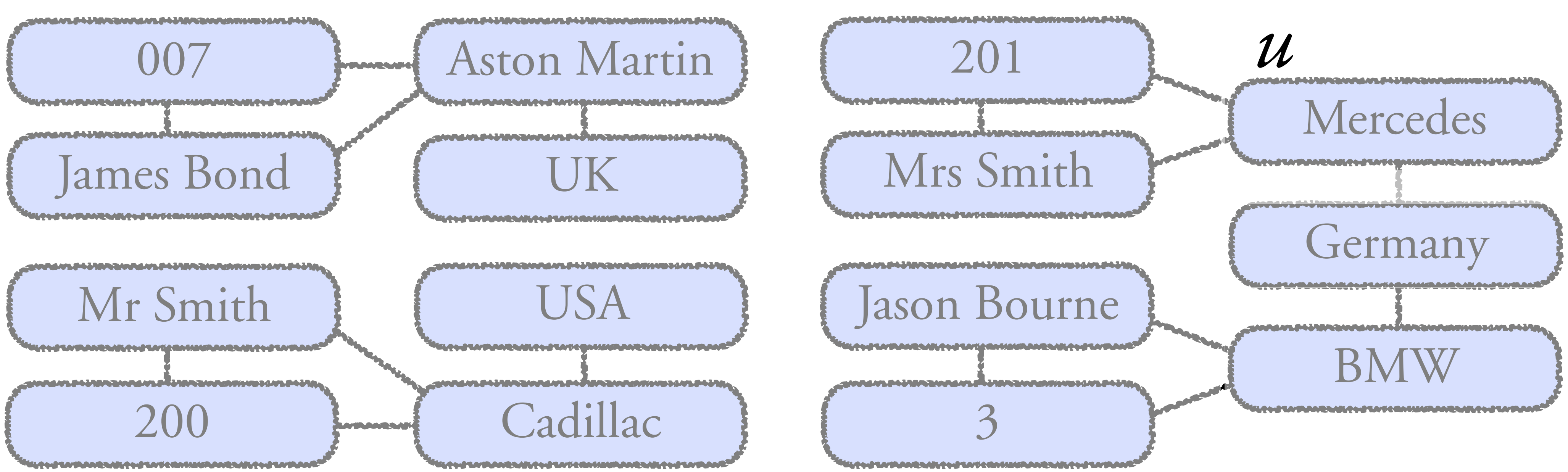


Hanf locality

- dist (u, v) = **distance** between u and v in the Gaifman graph
- $S[u, r]$ = sub-structure induced by $\{ v \mid \text{dist}(u, v) \leq r \}$ = **ball** around u of radius r

Agent	Name	Drives
007	James Bond	Aston Martin
200	Mr Smith	Cadillac
201	Mrs Smith	Mercedes u
3	Jason Bourne	BMW

Car	Country
Aston Martin	UK
Cadillac	USA
u Mercedes	Germany
BMW	Germany



Hanf locality

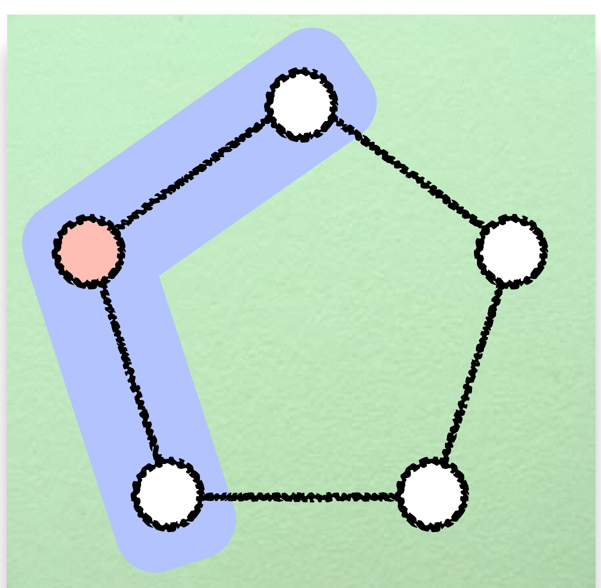
Definition. Two structures S_1 and S_2 are **Hanf(*r*, *t*) - equivalent**

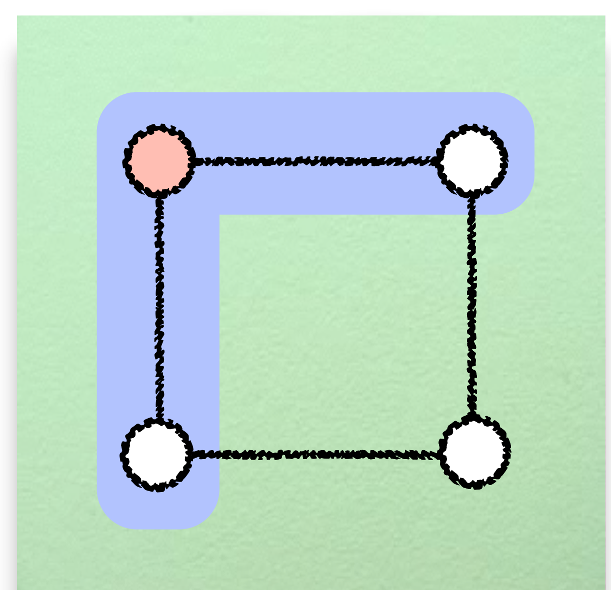
iff for each structure B, the two numbers

$$\#u \text{ s.t. } S_1[u,r] \cong B \qquad \#v \text{ s.t. } S_2[v,r] \cong B$$

are *either the same* or *both* $\geq t$.

Example. S_1, S_2 are Hanf(1, 1) - equivalent iff they have the *same balls* of radius 1





Hanf locality

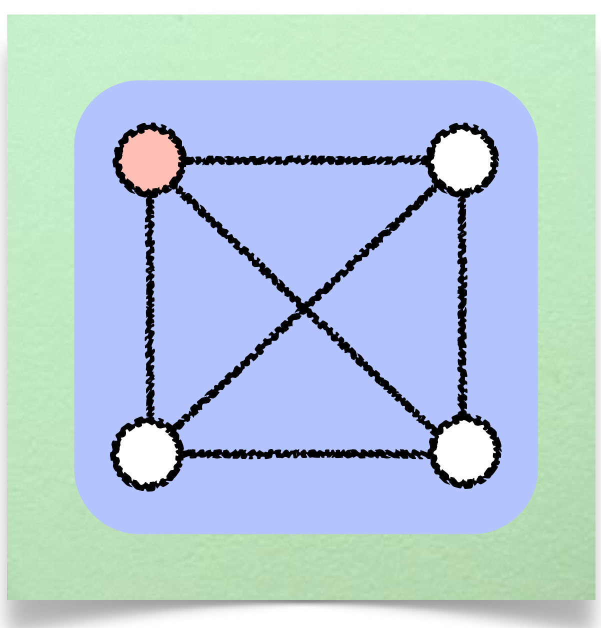
Definition. Two structures S_1 and S_2 are **Hanf(r, t)-equivalent**

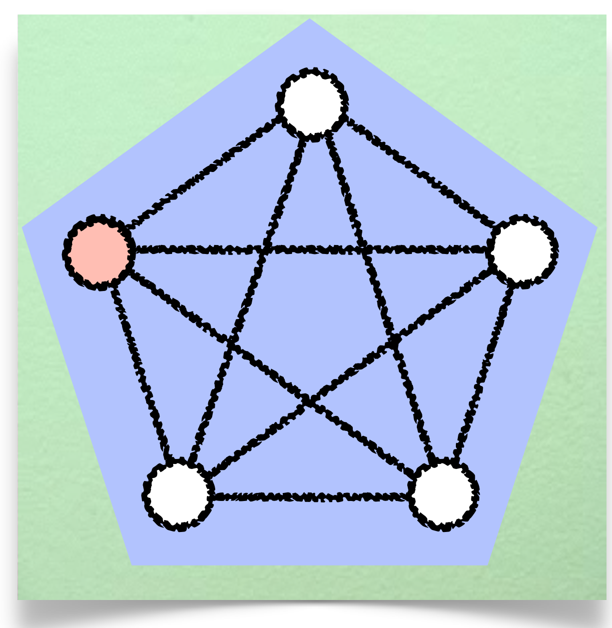
iff for each structure B, the two numbers

$$\#u \text{ s.t. } S_1[u, r] \cong B \qquad \#v \text{ s.t. } S_2[v, r] \cong B$$

are *either the same* or *both* $\geq t$.

Example. K_n, K_{n+1} are **not** Hanf$(1, 1)$-equivalent



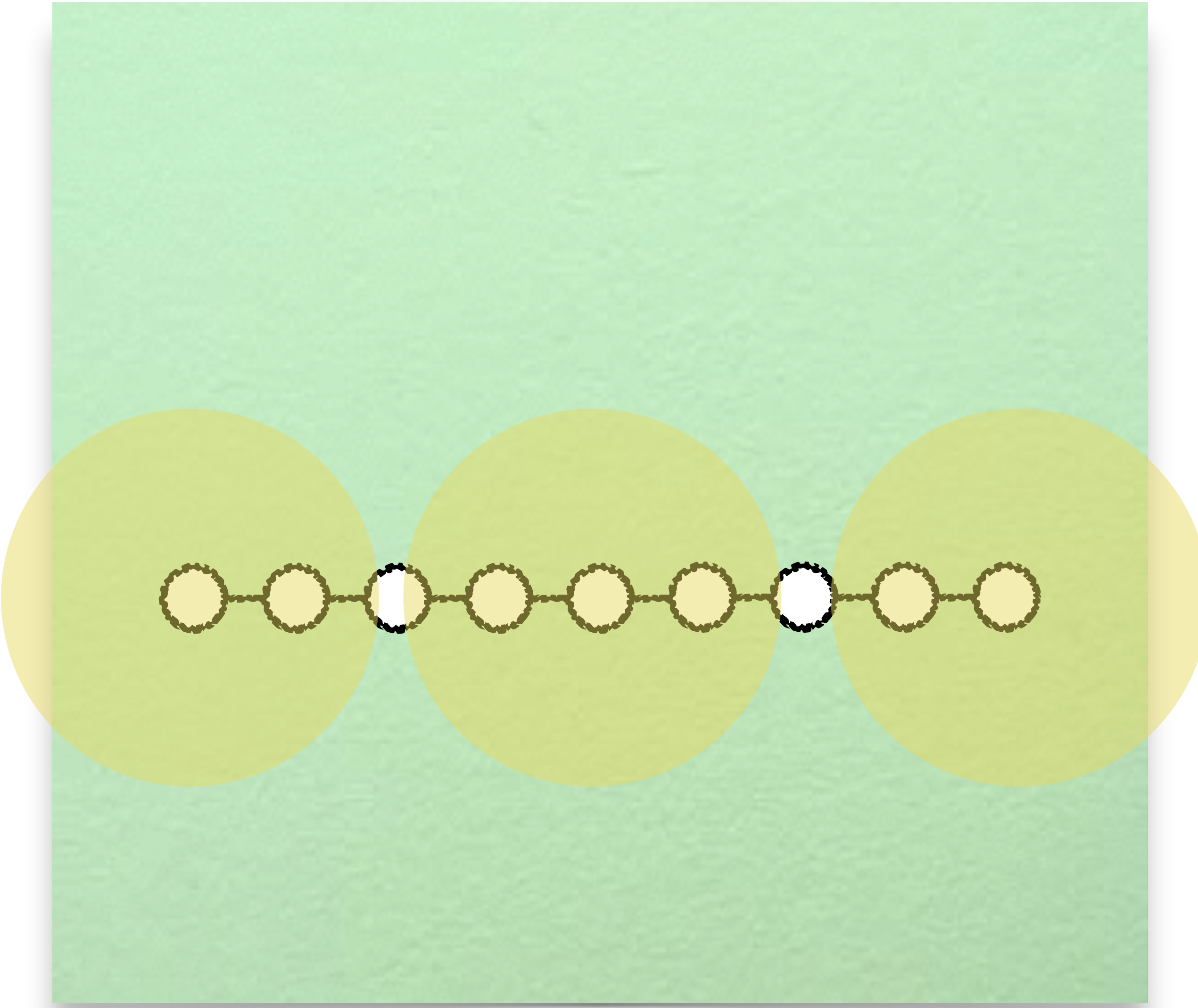


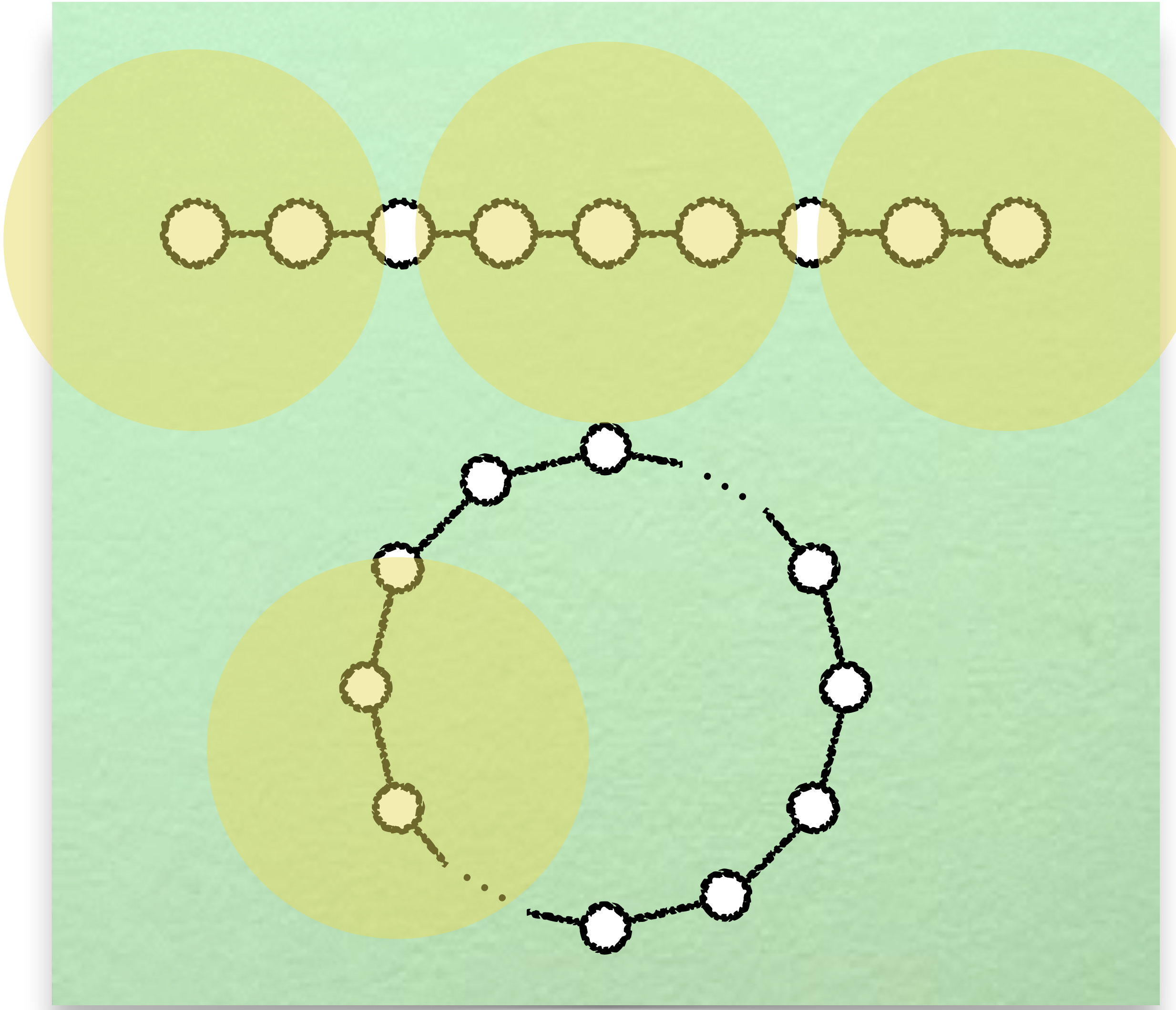
Hanf locality

Theorem. If S_1, S_2 are **Hanf(r, t) - equivalent**, with $r = 3^n$ and $t = n$
then S_1, S_2 are **n - equivalent** (they satisfy the same sentences with quantifier rank n)

[Hanf '60]

Exercise: prove that *acyclicity* is not FO-definable (on finite structures)

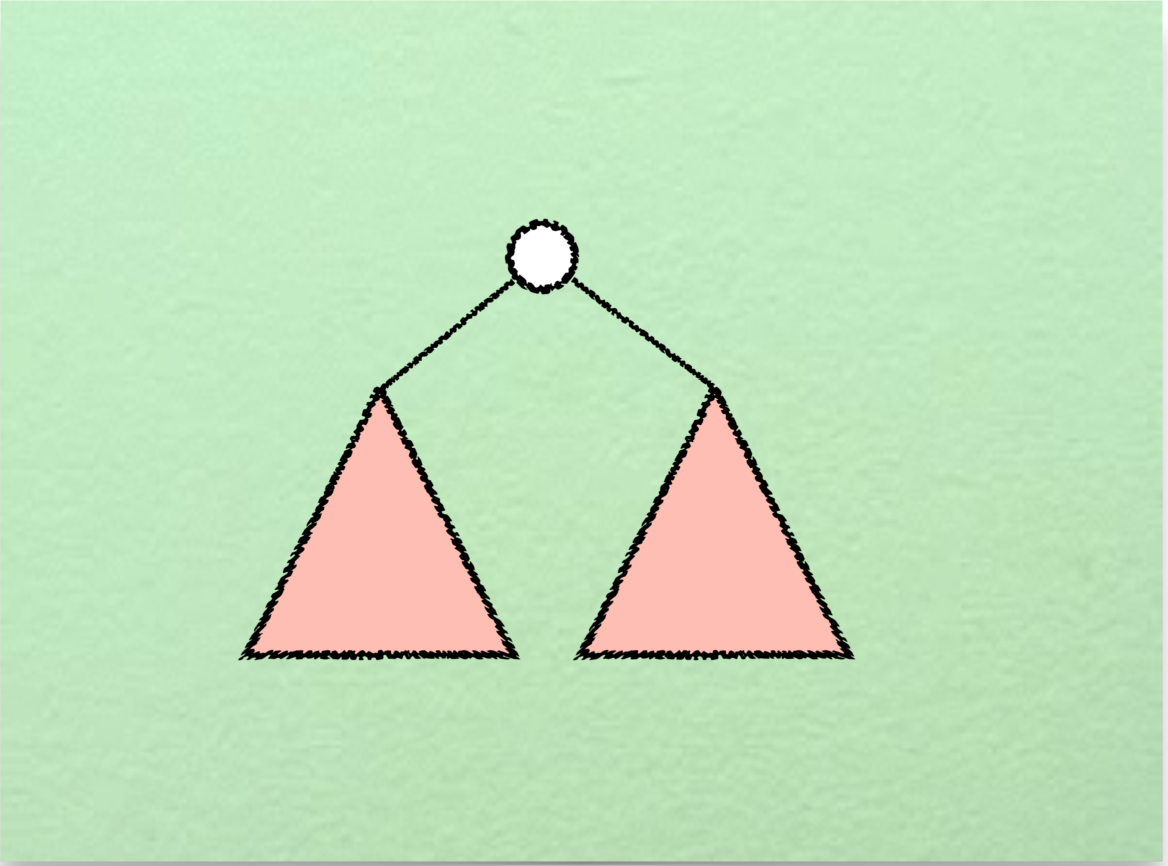


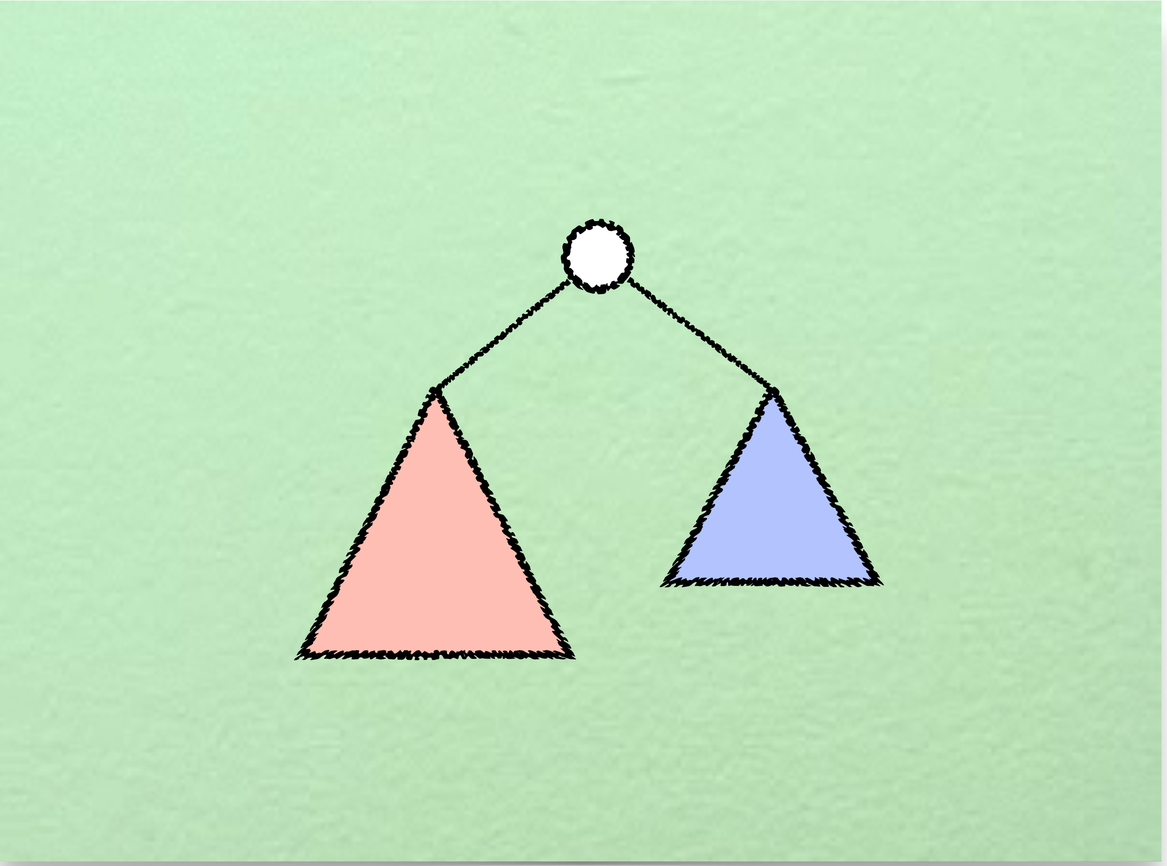


Hanf locality

Theorem. S_1, S_2 are n-equivalent (they satisfy the same sentences with quantifier rank n)
whenever S_1, S_2 are Hanf(r,t)-equivalent, with $r = 3^n$ and $t = n$.
[Hanf '60]

Exercise: prove that testing whether a binary tree is *complete* is not FO-definable





Hanf locality

Theorem. S_1, S_2 are n-equivalent (they satisfy the same sentences with quantifier rank n)

whenever S_1, S_2 are Hanf(r,t)-equivalent, with $r = 3^n$ and $t = n$.

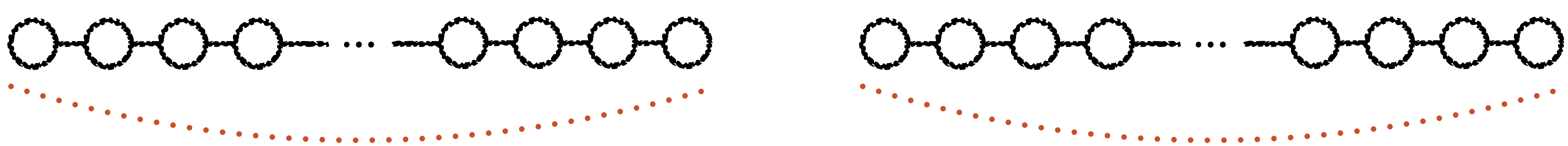
[Hanf '60]

Why so **BIG**?

Remember $\phi_k(x,y)$ = "there is a path of length 2^k from x to y"

$$\phi_0(x, y) = E(x, y), \text{ and}$$
$$\phi_k(x,y) = \exists z\, (\, \phi_{k-1}(x, z) \wedge \phi_{k-1}(z, y)\,)$$

$qr(\phi_k) = k$



$2 \cdot 2^n + 1$ $\qquad$ $2 \cdot 2^n$

Not (n+2)-equivalent yet they have the same $2^n - 1$ balls.

Hanf's theorem proof: Part I

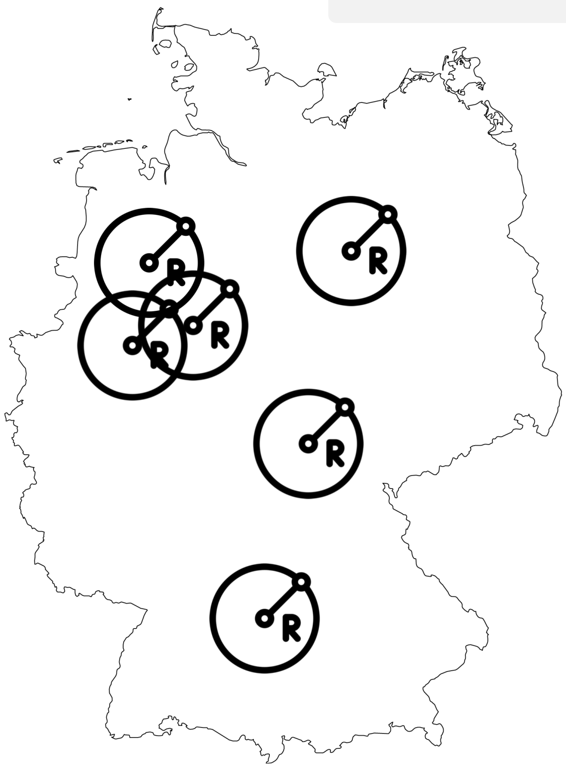
If $\mathfrak{A}$ and $\mathfrak{B}$ are Hanf$(3^n, n)$-equivalent then $\mathfrak{A} \equiv_m \mathfrak{B}$.

Proof

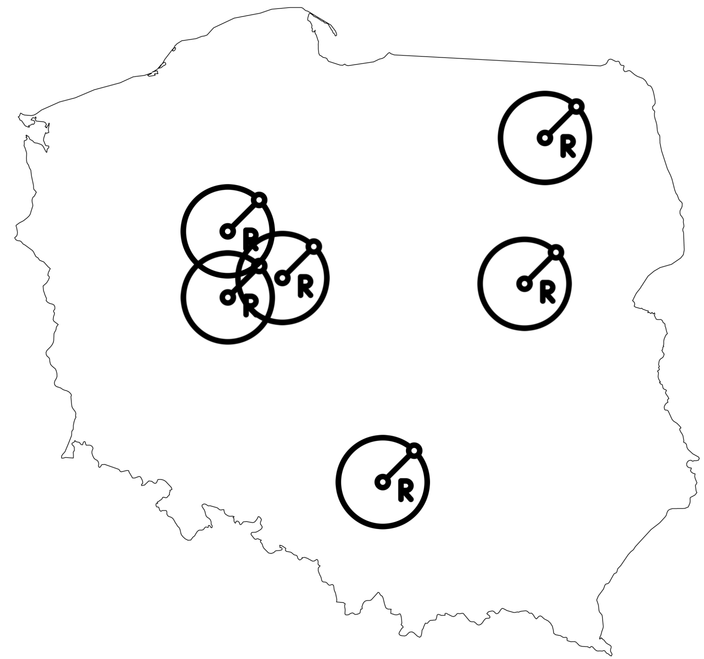
Let $a_1, a_2, \ldots, a_k \in A$ and $b_1, b_2, \ldots, b_k \in B$ be the history of the play after k rounds.

(Invariant): $\bigcup_{i=1}^{k} \mathfrak{A}[a_i, 3^{n-k}] \cong \bigcup_{i=1}^{k} \mathfrak{B}[b_i, 3^{n-k}]$

$\mathfrak{A} :=$



$\mathfrak{B} :=$



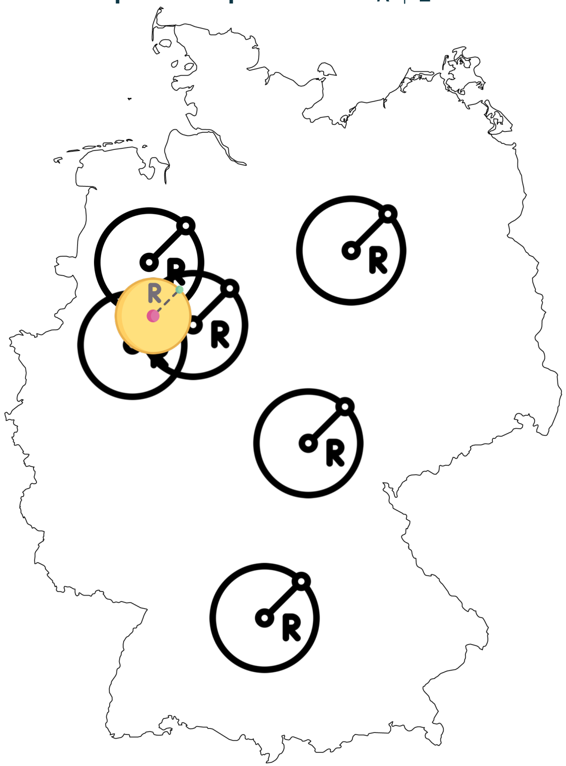
Hanf's theorem proof: Part II

Let $a_1, a_2, \ldots, a_k \in A$ and $b_1, b_2, \ldots, b_k \in B$ be the history of the play after k rounds.

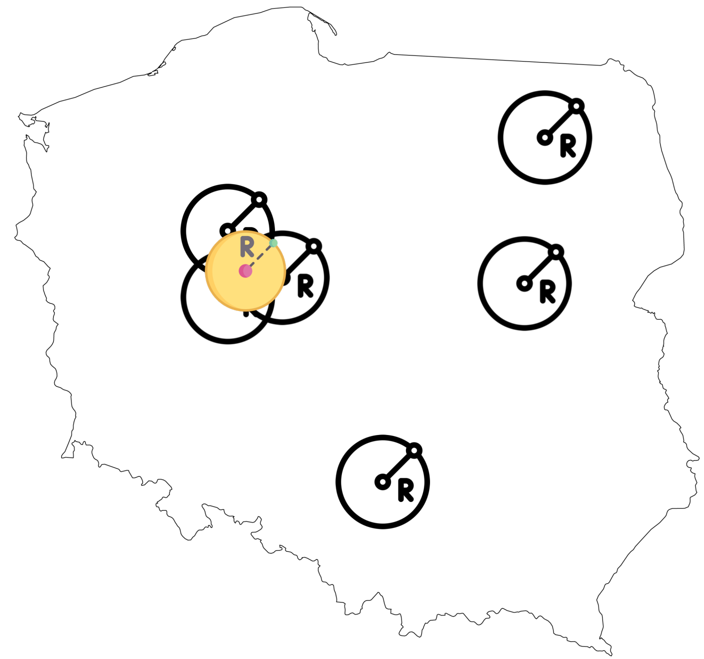
$$\textbf{(Invariant):} \; \bigcup_{i=1}^{k} \mathfrak{A}[a_i, 3^{n-k}] \cong \bigcup_{i=1}^{k} \mathfrak{B}[b_i, 3^{n-k}]$$

Suppose that Spoiler picked $a_{k+1} \in A$ such that $\mathrm{dist}(a_{k+1}, a_i) \leq 2 \cdot 3^{n-k}$ holds for some a_i.

$\mathfrak{A} :=$



$\mathfrak{B} :=$



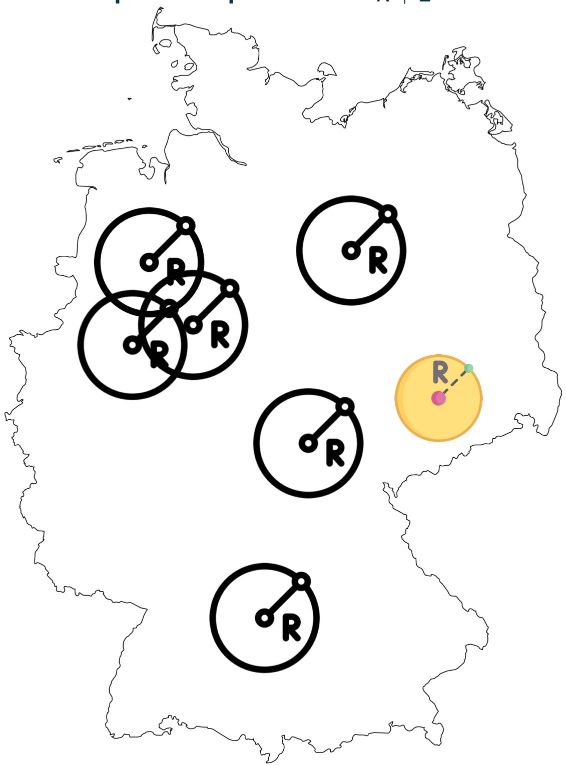
We know how to reply since $\mathfrak{A}[a_{k+1}, 3^{n-k-1}]$ is fully contained in some previously selected balls.

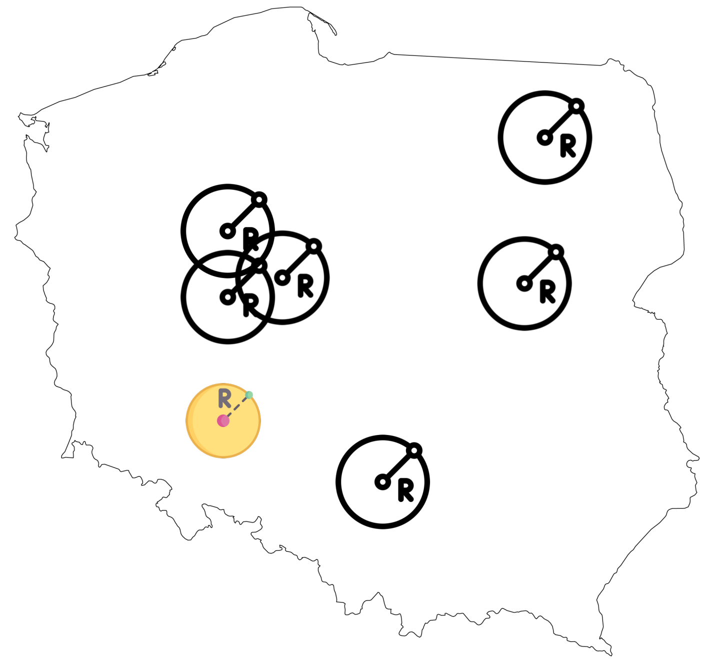
Hanf's theorem proof: Part II

Let $a_1, a_2, \ldots, a_k \in A$ and $b_1, b_2, \ldots, b_k \in B$ be the history of the play after k rounds.

$$\textbf{(Invariant):} \bigcup_{i=1}^{k} \mathfrak{A}[a_i, 3^{n-k}] \cong \bigcup_{i=1}^{k} \mathfrak{B}[b_i, 3^{n-k}]$$

Suppose that Spoiler picked $a_{k+1} \in A$ such that $\mathrm{dist}(a_{k+1}, a_i) > 2 \cdot 3^{n-k}$ holds for some a_i.

$\mathfrak{A} :=$ 

$\mathfrak{B} :=$ 

We know how to reply since we have sufficiently many realisations of $\mathfrak{A}[a_{k+1}, 3^{n-k-1}]$ in $\mathfrak{B}$.

Copyright of used icons, pictures and slides

1. Universities/DeciGUT/ERC logos downloaded from the corresponding institutional webpages.
2. Idea icon created by Vectors Market — Flaticon `flaticon.com/free-icons/idea`.
3. Head icons created by Eucalyp — Flaticon `flaticon.com/free-icons/head`
4. Graph icons created by SBTS2018 — Flaticon `flaticon.com/free-icons/graph`
5. Angel icons created by Freepik — Flaticon `flaticon.com/free-icons/angel`
6. Devil icons created by Freepik and Pixel perfect — Flaticon `flaticon.com/free-icons/devil`
7. VS icons created by Freepik — Flaticon `flaticon.com/free-icons/vs`
8. Robot icon created by Eucalyp — Flaticon `flaticon.com/free-icons/robot`.
9. Warning icon created by Freepik - Flaticon `flaticon.com/free-icons/warning`.
10. Slides 78–110 from ESSLI 2016 by [`Diego Figueira`]
11. German and Poland maps by Vemaps.com `https://vemaps.com/europe`.
12. Radius icons created by Freepik - Flaticon `flaticon.com/free-icons/radius`.