Consider the signature \((E, <)\).

Consider structures \(G = (V, E, <)\) in which \(E\) is a graph relation and \(<\) is a linear order.

There is no first order sentence \(\gamma\) in this signature such that

\[ G \models \gamma \text{ if, and only if, } (V, E) \text{ is connected.} \]

Proof

Suppose there was such a formula \(\gamma\).

Let \(\gamma'\) be the formula obtained by replacing every occurrence of \(E(x, y)\) in \(\gamma\) by the following formula

\[
\begin{align*}
y &= x + 2 \\
(x = \max \land y = \min + 1) \lor \\
(y = \min \land x = \max - 1).
\end{align*}
\]

Then, \(\neg \gamma'\) defines evenness on linear orders!

Proof

We obtain two disjoint cycles on linear orders of even length, and a single cycle on linear orders of odd length.
**Reduction**

The above is, in fact, a *first-order definable reduction* from the problem of evenness of linear orders to the problem of connectivity of ordered graphs.

It follows from the above that there is no first order formula that can express the *transitive closure* query on graphs. *Any such formula would also work on ordered graphs.*

**Gaifman Graphs and Neighbourhoods**

On a structure $A$, define the binary relation:

$$E(a_1, a_2) \text{ if, and only if, there is some relation } R \text{ and some tuple } \mathbf{a} \text{ containing both } a_1 \text{ and } a_2 \text{ with } R(\mathbf{a}).$$

The graph $GA = (A, E)$ is called the *Gaifman graph* of $A.$

$\text{dist}(a, b)$ — the distance between $a$ and $b$ in the graph $(A, E)$.

$\text{Nbd}^A_r(a)$ — the substructure of $A$ given by the set:

$$\{b \mid \text{dist}(a, b) \leq r\}$$

**Hanf Locality Theorem**

We say $A$ and $B$ are *Hanf equivalent* with radius $r$ ($A \simeq_r B$) if, for every $a \in A$ the two sets

$$\{a' \in A \mid \text{Nbd}^A_r(a) \cong \text{Nbd}^B_r(a')\} \text{ and } \{b \in B \mid \text{Nbd}^B_r(a) \cong \text{Nbd}^B_r(b)\}$$

have the same cardinality

and, similarly for every $b \in B.$

**Theorem (Hanf)**

For every vocabulary $\sigma$ and every $p$ there is $r \leq 3^p$ such that for any $\sigma$-structures $A$ and $B$: if $A \simeq_r B$ then $A \equiv_p B.$

In other words, if $r \geq 3^p,$ the equivalence relation $\simeq_r$ is a refinement of $\equiv_p.$

**Hanf Locality**

*Duplicator*’s strategy is to maintain the following condition:

After $k$ moves, if $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ have been selected, then

$$\bigcup_i \text{Nbd}^A_{3^p-k-1}(a_i) \cong \bigcup_i \text{Nbd}^B_{3^p-k-1}(b_i)$$

If *Spoiler* plays on $a$ within distance $2 \cdot 3^p-k-1$ of a previously chosen point, play according to the isomorphism, otherwise, find $b$ such that

$$\text{Nbd}^A_{3p-k-1}(a) \cong \text{Nbd}^B_{3p-k-1}(b)$$

and $b$ is not within distance $2 \cdot 3^p-k-1$ of a previously chosen point.

Such a $b$ is guaranteed by $\simeq_r.$
Uses of Hanf locality

The Hanf locality theorem immediately yields, as special cases, the proofs of undefinability of:

- connectivity
- 2-colourability
- acyclicity
- planarity

A simple illustration can suffice.

Connectivity

To illustrate the undefinability of connectivity and 2-colourability, consider on the one hand the graph consisting of a single cycle of length $4r + 6$ and, on the other hand, a graph consisting of two disjoint cycles of length $2r + 3$.

Acyclicity

A figure illustrating that acyclicity is not first-order definable.

Planarity

A figure illustrating that planarity is not first-order definable.
**Monadic Second Order Logic**

MSO consists of those second order formulas in which all relational variables are unary.

*That is, we allow quantification over sets of elements, but not other relations.*

Any MSO formula can be put in prenex normal form with second-order quantifiers preceding first order ones.

Mon. $\Sigma^1_1$ — MSO formulas with only existential second-order quantifiers in prenex normal form.

Mon. $\Pi^1_1$ — MSO formulas with only universal second-order quantifiers in prenex normal form.

**Undefinability in MSO**

The method of games and *locality* can also be used to show *inexpressibility* results in MSO.

In particular, there is a Mon.$\Sigma^1_1$ query that is not definable in Mon.$\Pi^1_1$ (Fagin 1974)

*Note:* A similar result without the *monadic* restriction would imply that NP $\neq$ co-NP and therefore that P $\neq$ NP.

**Connectivity**

Recall that *connectivity* of graphs can be defined by a Mon.$\Pi^1_1$ sentence.

$$\forall S(\exists x Sx \land (\forall x \forall y (Sx \land Exy) \rightarrow Sy)) \rightarrow \forall x Sx$$

and by a $\Sigma^1_1$ sentence (simply because it is in NP).

We now aim to show that *connectivity* is not definable by a Mon.$\Sigma^1_1$ sentence.

**MSO Game**

The *m*-round monadic Ehrenfeucht game on structures $\mathcal{A}$ and $\mathcal{B}$ proceeds as follows:

- At the *i*-th round, *Spoiler* chooses one of the structures (say $\mathcal{B}$) and plays either a point move or a set move.
  
  In a point move, he chooses one of the elements of the chosen structure (say $b_i$) — *Duplicator* must respond with an element of the other structure (say $a_i$).

  In a set move, he chooses a subset of the universe of the chosen structure (say $S_i$) — *Duplicator* must respond with a subset of the other structure (say $R_i$).
MSO Game

• If, after $m$ rounds, the map
  \[ a_i \mapsto b_i \]
  is a partial isomorphism between
  \((A, R_1, \ldots, R_q)\) and \((B, S_1, \ldots, S_q)\)
  then \textit{Duplicator} has won the game, otherwise \textit{Spoiler} has won.

Existential Game

The $m,p$-move existential game on \((A, B)\):

• First \textit{Spoiler} makes $m$ set moves on $A$, and \textit{Duplicator} replies on $B$.
• This is followed by an Ehrenfeucht game with $p$ point moves.

If \textit{Duplicator} has a winning strategy, then for every \textit{Mon.$\Sigma_1^1$} sentence:
  \[ \phi \equiv \exists R_1 \ldots \exists R_m \psi \]
  with $\text{qr}(\psi) = p$,
  if $A \models \phi$ then $B \models \phi$

Variation

To show that a Boolean query $Q$ is not \textit{Mon.$\Sigma_1^1$} definable, find for each $m$ and $p$

• $A \in Q$; and
• $B \not\in P$; such that
• \textit{Duplicator} wins the $m,p$ move game on \((A, B)\).

Or,
• \textit{Duplicator} chooses $A$.
• \textit{Spoiler} colours $A$ (with $2^m$ colours).
• \textit{Duplicator} chooses $B$ and colours it.
• They play a $p$-round Ehrenfeucht game.

If we define the \textit{quantifier rank} of an MSO formula by adding the following inductive rule to those for a formula of FO:

if $\phi = \exists S \psi$ or $\phi = \forall S \psi$ then $\text{qr}(\phi) = \text{qr}(\psi) + 1$

then, we have

\textit{Duplicator} has a winning strategy in the $m$-round monadic Ehrenfeucht game on structures $A$ and $B$ if, and only if, for every sentence $\phi$ of MSO with $\text{qr}(\phi) \leq m$

\[ A \models \phi \quad \text{if, and only if}, \quad B \models \phi \]
Application

Write $C_n$ for the graph that is a simple cycle of length $n$.

For $n$ sufficiently large, and any *colouring* of $C_n$, we can find an $n' < n$ and a colouring of

$C_{n'} \oplus C_{n-n'}$ the disjoint union of two cycles—one of length $n'$, the other of length $n - n'$

So that the graphs $C_n$ and $C_{n'} \oplus C_{n-n'}$ are $\simeq_r$ equivalent.

Taking $n > (2r + 1)^{2^m + 2}$ suffices.

Reading List for this Handout

1. Ebbinghaus and Flum. Section 2.4.
2. Libkin. Chapter 4.
3. Grädel et al. Section 2.3 and 2.5