Definability in First Order Logic and Second Order Logic

Vocabularies

We deal with (possibly many-sorted) relational structures.

Sort symbols are

$$U_\alpha : \alpha \in \mathbb{N}$$

Relation symbols are

$$R_{i,\alpha} : i \in Ar, \alpha \in \mathbb{N}$$

where Ar is a set of *arities*, i.e. of finite sequences of sort symbols.

In the case of one-sorted vocabularies, the arity is just of the form $\langle U, U, \ldots n \ldots, U \rangle$ which will denoted by n.

A **vocabulary** is a *finite* set of *finitary* **relation symbols**, usually denoted by τ, τ_i or σ.

τ-structures

Graphs: $\langle V; E \rangle$ with vertices as domain and edges as relation.
$\langle V \sqcup E, R_G \rangle$ with two sorted domain of vertices and edges and incidence relation.

Labeled Graphs: As graphs but with unary predicates for vertex labels and edge labels depending whether edges are elements or tuples.

Binary Words: $\langle V; R_<, P_0 \rangle$ with domain lineraly ordered by $R_<$ and colored by P_0, marking the zero's.

τ-structures: General relational structures.

Properties of a τ-structure

A **property** of τ-structures is a class $\mathcal{P}$ of τ-structures closed under τ-isomorphisms.

- All *finite* τ-structures.
- All $\{R_{2,0}\}$-structures where $R_{2,0}$ is interpreted as a linear order.
- Al finite 3-dimensional matchings $3DM$, i.e. all $\{R_{3,0}\}$-structures with universe A where the interpretation of $R_{3,0}$ contains a subset $M \subseteq A^3$ such that no two triples of M agree in any coordinate.
- All binary words which are palindroms.

A τ-structure $\mathcal{A}$ has property $\mathcal{P}$ iff $\mathcal{A} \in \mathcal{P}$.

First Order Logic $FOL(\tau)$:

For structures of the form $\mathcal{A} = \langle V, R_1^V, \ldots R_M^V \rangle$ and $\tau = \{R_1' \ldots, R_M\}$

Variables: $u, v, w, \ldots$ ranging over elements of the domain V.
R_j a $\rho(j)$-ary relation symbol whose interpretation is R_j^V.
Atomic formulas: $R_j(\bar{u})$, $u = v$.
Connectives: $\wedge, \vee, \neg,$
Quantifiers: $\forall v, \exists v$

Monadic Second Order Logic $MSOL(\tau)$:

Additionally we have
Variables: $X, Y, Z, \ldots$
ranging over subsets of V.
Atomic formulas: $u \in X, v \in Y, \ldots$
Quantifiers: $\forall X, \exists X$.

Theorem:[Büchi, Trakhtenbrot, 1961]
A class of binary words is:
recognizable by a finite
(non-deterministic) automaton
iff it is $MSOL$-definable
(iff it is regular).

Example: $(101 \vee 1001)^*$
101 1001 101 101 1001 1001 101........

Exercise: Find the $MSOL$-formula.

Second Order Logic $SOL^n(\tau)$ and $SOL(\tau)$:

We extend $MSOL(\tau)$ by the following features:

Variables: $X^m, Y^m, Z^m, \ldots$ for $m \leq n$

Atomic formulas: $(u_1, \ldots, u_m) \in X^m, \ldots$
Quantifiers: $\forall X^m, \exists X^m$.

$SOL = \bigcup_n SOL^n$

Clearly we have in expressing power (and syntactically)

$$MSOL(\tau) \subseteq SOL^2(\tau) \subseteq SOL(\tau)$$

In SOL^2 we can quantifier over **arbitrary sets of pairs of vertices**,

Definition 1 ($\mathcal{L}(\tau)$-Definability)

Recall that an $\mathcal{L}(\tau)$-sentence is an $\mathcal{L}(\tau)$-formula without free variables.

Given a regular logic $\mathcal{L}$ and a class of τ-structures K, we say that K is $\mathcal{L}(\tau)$-*definable* if there is a $\mathcal{L}(\tau)$-sentence θ such that for every τ-structure $\mathcal{A}$

$$\mathcal{A} \models \theta \text{ iff } \mathcal{A} \in K.$$

We write $Mod_{\mathcal{L}(\tau)}(\theta)$ for the class of τ-structures $\mathcal{A}$ such that $\mathcal{A} \models \theta$.

Proving definability

The class of τ-structures of finite even cardinality, $EVEN(\tau)$, is definable in Second Order Logic:

- Let $\tau_1 = \{R, S, P\}$ with R, S binary and P unary, none of them in τ.
- We write a $FOL(\tau_1)$-formula $\phi_{bij}(R, P)$ which says that R is a bijection between P and its complement.
- We write a $FOL(\tau_1)$-formula $\psi_{inj}(S)$ which says that S is a proper injection of the domain into itself.
- Now the required formula is

$$\exists R \exists P \phi_{bij}(R, P) \wedge \forall S \neg \psi_{inj}(S)$$

Cographs graphs

A graph G is a **cograph** if and only if there is no induced subgraph of G isomorphic to a P_4.

P_4-sparse Graphs

A G is P_4-**sparse** if no set of 5 vertices induced more than one P_4 in G.

Cliques and Cographs are P_4-sparse.

Recall

Example 2 (3 Colorability)

The class of 3-colorable graphs is definable by a formula of Monadic Second Order Logic:

$$\exists X_1, X_2, X_3 \phi_{partition}(X_1, X_2, X_3) \wedge \bigwedge_{i=1}^{3} \phi_{color}(X_i)$$

where

- $\phi_{partition}(X_1, X_2, X_3)$ says that X_1, X_2, X_3 form a partition of the vertices and
- $\phi_{color}(X_i)$ says that there are no edges between two vertices in X_i.

Note that all the second order variables are unary and $\phi_{partition}$ and ϕ_{color} are first order formulas over $\tau = \{E, X_1, X_2, X_3\}$.

Expressibility: $FOL(\tau)$

The following are FOL-definable on graphs:

- Cographs
- P_4-sparse graphs
- Existence of prescribed (induced) subgraph H.
- Non-Existence of prescribed (induced) subgraph H.

Expressibility: $FOL(\tau)$ vs. $MSOL(\tau)$

- A graph has bounded degree $\leq k$ is FOL-expressible.
- A graph is regular of degree 17 is FOL-expressible.
- A graph is connected is not FOL expressible, but $MSOL$ expressible by

  $$\neg \exists X(closed(X) \wedge \exists v(v \notin X))$$

  with

  $closed(X) = \forall v, w((v \in X \wedge E(v, w)) \rightarrow w \in X)$

Expressibility: $MSOL(\tau)$- but not $FOL(\tau)$-expressible

- Chordal graphs are $MSOL$-definable
- Trees are cycle-free graphs.
  Trees are $MSOL$-definable
- Bipartite graphs are 2-colorable graphs.
  Bipartite graphs are $MSOL$-definable
- 3-Colorability is $MSOL$-expressible:
- There are vertex disjoint paths between the k pairs $(x_1, y_1), \ldots, (x_k, y_k)$ is $MSOL$-expressible.

We shall see in the next lecture that non of these are FOLdefinable.

Planar Graphs, revisited

A **subdivision of an edge** is a path with now branching.
A **subdivion of a graph** is a graph obtained by replacing all the edges by some subdivions thereof.

Theorem [Kuratowski 1930]:
A graph G is planar if and only if there is no induced subgraph of G isomorphic to a subdivision of K_5 or $K_{3,3}$.

Use this to show:

Proposition:
Planarity is $MSOL$ definable.

For graphs G_2 of the form $\langle V \sqcup E, R \rangle$ with V set of vertices, E set of edges, and $R \subseteq V \times E$ a binary relation expressing that v lies on the edge e.

Monadic Second Order Logic MS_2:

For graphs G_1 of the form $\langle V, E, \rangle$ with V set of vertices, $E \subseteq V^2$ set of edges (as a binary relation).

Monadic Second Order Logic MS_1:

MS_1 vs MS_2 Definable Graph Properties

- Connectivity (MS_1)
- Planarity (MS_1)
- Perfect matching (MS_2 but not MS_1)
- Hamiltonian cycle (MS_2 but not MS_1)

The following are not even MS_2-definable:

- existence of a clique of size at least $\frac{n}{2}$
- Eulerian graphs

Expressibility: $MSOL(\tau)$ vs. $SOL(\tau)$

- The existence of a hamiltonian circuit is **not** MS_1-expressible but SOL^2-expressible for graphs $G = (V, E)$.

$$\exists X^2 (Edges(X^2) \wedge SimpleCircuit(X^2) \wedge AllVerticesIn(X^2))$$

- The existence of a clique which is at least half the size of the graph (HALF-CLIQUE) is **not** $MSOL$-expressible but SOL^2-expressible.
- The class of graphs which are disjoint unions of two isomorphic components is not $MSOL$-definable, but SOL-definable