EF Games #2: Inexpressibility Proofs

Christoph Koch

Inexpressibility proofs

We use the methodology theorem:

Theorem (Methodology theorem)

Given a Boolean query Q. There is **no** *FO sentence that expresses Q if and only if there are, for each k, structures $\mathcal{A}_k$, $\mathcal{B}_k$ such that*

- $\mathcal{A}_k \models Q$,
- $\mathcal{B}_k \not\models Q$ *and*
- $\mathcal{A}_k \sim_k \mathcal{B}_k$.

To prove inexpressibility, we only have to

- construct suitable structures $\mathcal{A}_k$ and $\mathcal{B}_k$ and
- prove that $\mathcal{A}_k \sim_k \mathcal{B}_k$. (This is usually the difficult part.)

Example: Inexpressibility of the parity query

Definition (parity query)

Given a structure $\mathcal{A}$ with empty schema (i.e., only $|\mathcal{A}|$ is given).
Question: Does $|\mathcal{A}|$ have an even number of elements?

- Construction of the structures $\mathcal{A}_n$ and $\mathcal{B}_n$ for arbitrary n:

$$|\mathcal{A}_n| := \{a_1, \ldots, a_n\} \qquad |\mathcal{B}_n| := \{b_1, \ldots, b_{n+1}\}$$

Lemma

$\mathcal{A}_n \sim_k \mathcal{B}_n$ for all $k \leq n$.

(This is shown on the next slide.)

- On the other hand, $\mathcal{A}_n \models \text{Parity}$ if and only if $\mathcal{B}_n \not\models \text{Parity}$.
- It thus follows from the methodology theorem that parity is not expressible in FO.

Example: Inexpressibility of the parity query

Lemma

$\mathcal{A}_n \sim_k \mathcal{B}_n$ for all $k \leq n$.

Proof.

We construct a winning strategy for Duplicator. This time no strategy trees are explicitly shown, but a general construction is given.
We handle the case in which Spoiler plays on $\mathcal{A}_n$. The other direction is analogous. If $S_i \mapsto a$ then

- $D_i \mapsto b$ where b is a new element of $|\mathcal{B}_n|$ if a has not been played on yet (=no token was put on it);
- If, for some $j < i$, $S_j \mapsto a, D_j \mapsto b'$ or $S_j \mapsto b', D_j \mapsto a$ was played then $D_i \mapsto b'$.

Over k moves, we only construct partial isomorphisms in this way and obtain a winning strategy for Duplicator. □

Eulerian graphs

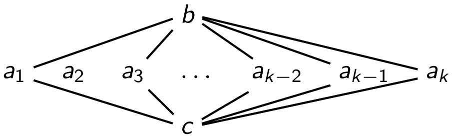
Definition

Eulerian graph: a graph that has a Eulerian cycle, i.e., a round trip that visits each edge of the graph exactly once.

Theorem

The Boolean query "Eulerian Graph" is not expressible in FO.

Proof sketch: Graph $\mathcal{A}_k$:



Graph $\mathcal{B}_k := \mathcal{A}_{k+1}$.
For all k: $\mathcal{A}_k \sim_k \mathcal{B}_k$. $\mathcal{A}_k$ is Eulerian if and only if k is even, i.e., iff $\mathcal{B}_k$ is not Eulerian.

Undirected Paths

$$
\begin{array}{ll}
L_n & a_1 - a_2 - a_3 - \dots - a_{i-1} - a_i - a_{i+1} - \dots - a_n \\
L_n^{<a_i} & a_1 - a_2 - a_3 - \dots - a_{i-1} \\
L_n^{>a_i} & \qquad\qquad\qquad\qquad\qquad\qquad a_{i+1} - \dots - a_n
\end{array}
$$

(Nodes a_{i-1}, a_{i+1} are labeled A_i, as adjacent to a_i in L_n).

Lemma (composition lemma for paths)

$L_m \sim_{k+1} L_n$ *if and only if*

$$
\begin{array}{lll}
(1) & \forall a\, \exists b & L_m^{<a} \sim_k L_n^{<b} \;\wedge\; L_m^{>a} \sim_k L_n^{>b} \text{ and} \\
(2) & \forall b\, \exists a & L_m^{<a} \sim_k L_n^{<b} \;\wedge\; L_m^{>a} \sim_k L_n^{>b}
\end{array}
$$

Undirected Paths

Lemma (composition lemma for paths)

$$\left.\begin{array}{lll} (1) & \forall a\,\exists b & L_m^{<a} \sim_k L_n^{<b} \ \wedge\ L_m^{>a} \sim_k L_n^{>b} \\ (2) & \forall b\,\exists a & L_m^{<a} \sim_k L_n^{<b} \ \wedge\ L_m^{>a} \sim_k L_n^{>b} \end{array}\right\} \Leftrightarrow L_m \sim_{k+1} L_n$$

Proof.

We define the winning strategy for $k+1$ moves as follows:

- W.l.o.g., Spoiler Spoiler chooses node a of structure L_m in the first move.
- Because of (1), there is a b in L_n such that Duplicator wins in k moves on $L_m^{<a}$, $L_n^{<b}$ and on $L_m^{>a}$, $L_n^{>b}$.
- We can combine the two winning strategies into one combined strategy:
  - If Spoiler chooses a node $\leq a$ in L_m in the i-th move, then Duplicator answers according to the winning strategy for $L_m^{<a}$ and $L_n^{<b}$, not counting the moves that were played in the other pair of structures.
  - If Spoiler chooses a node $\geq a$, we answer analogously using Duplicator's winning strategy for $L_m^{>a}$, $L_n^{>b}$. □

Undirected Paths

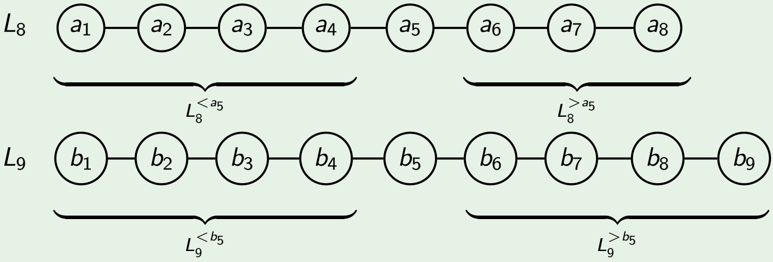
It follows:

Theorem

$L_m \sim_k L_n$ if and only if $m = n$ or $m, n \geq 2^k - 1$.

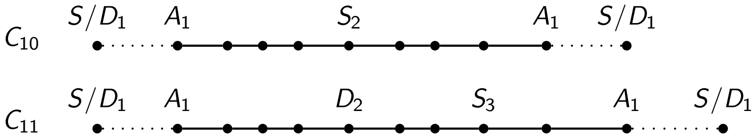
So for $n < 2^k - 1$, $L_n \nsim_k L_{n+1}$; for $n \geq 2^k - 1$, $L_n \sim_k L_{n+1}$.

Example ($L_8 \sim_3 L_9$)



Cycles

- (Isolated) directed cycles C_n: Graphs with nodes $\{v_1, \ldots, v_n\}$ and edges $\{(v_1, v_2), (v_2, v_3), \ldots, (v_{n-1}, v_n), (v_n, v_1)\}$.
- There is an analogous composition lemma for (directed or undirected) cycles.
- After the first move, there is one distinguished node in the cycle, the one with token S_1 or D_1 on it.
- We can treat this cycle like a path obtained by cutting the cycle at the distinguished node.



- Theorem. If $n \geq 2^k$, then $C_n \sim_k C_{n+1}$.

2-colorability

Definition

2-colorability: Given a graph, is there a function that maps each node to either "red" or "green" such that no two adjacent nodes have the same color?

Theorem

2-colorability is not expressible in FO.

Proof Sketch.

For each k,

- $\mathcal{A}_k$: C_{2^k}, the cycle of length 2^k.
- $\mathcal{B}_k$: C_{2^k+1}, the cycle of length $2^k + 1$.
- $\mathcal{A}_k \sim_k \mathcal{B}_k$.
- However, a cycle C_n of length n is 2-colorable iff n is even.

Inexpressibility follows from the EF methodology theorem. □

Acyclicity

From now on, "very long/large" means simply 2^k.

Theorem

Acyclicity is not expressible in FO.

Proof Sketch.

- $\mathcal{A}_k$: a very long path.
- $\mathcal{B}_k$: a very long path plus (disconnected from it) a very large cycle.
- $\mathcal{A}_k \sim_k \mathcal{B}_k$.

□

Graph reachability

Theorem

Graph reachability from a to b is not expressible in FO.

a, b are constants or are given by an additional unary relation with two entries.

Proof Sketch.

- $\mathcal{A}_k$: a very large cycle in which the nodes a and b are maximally distant.
- $\mathcal{B}_k$: two very large cycles; a is a node of the first cycle and b a node of the second.
- $\mathcal{A}_k \sim_k \mathcal{B}_k$.

□