Towards Capturing Order-Independent P

Neil Immerman

College of Computer and Information Sciences
University of Massachusetts, Amherst
Amherst, MA, USA

people.cs.umass.edu/~immerman
Restrict attention to the complexity of computing individual bits of the output, i.e., decision problems.

How hard is it to check if input has property $S$?

How rich a language do we need to express property $S$?

There is a computable isomorphism between these two approaches.
Restrict attention to the complexity of computing individual bits of the output, i.e., **decision problems**.
Restrict attention to the complexity of computing individual bits of the output, i.e., decision problems.

How hard is it to check if input has property $S$?
Restrict attention to the complexity of computing individual bits of the output, i.e., decision problems.

How hard is it to check if input has property $S$?

How rich a language do we need to express property $S$?
Restrict attention to the complexity of computing individual bits of the output, i.e., decision problems.

How hard is it to check if input has property $S$?

How rich a language do we need to express property $S$?

There is a computable isomorphism between these two approaches.
REACH = \{ G, s, t \mid s \xrightarrow{*} t \}
Inductive Definitions and Least Fixed Point

\[
\text{REACH} = \{ G, s, t \mid s \overset{*}{\rightarrow} t \}
\]

\[
\text{REACH} \not\in \text{FO}
\]
Inductive Definitions and Least Fixed Point

\[ E^*(x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z (E^*(x, z) \land E^*(z, y)) \]

\[ \text{REACH} = \{ G, s, t \mid s \overset{*}{\rightarrow} t \} \]

\[ \text{REACH} \notin \text{FO} \]
Inductive Definitions and Least Fixed Point

\[ E^*(x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z(E^*(x, z) \land E^*(z, y)) \]

\[ \varphi_{tc}(R, x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z(R(x, z) \land R(z, y)) \]

\[
\begin{align*}
\text{REACH} &= \{ G, s, t \mid s \xrightarrow{*} t \} \\
\text{REACH} &\notin \text{FO}
\end{align*}
\]
Inductive Definitions and Least Fixed Point

\[ E^*(x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z(E^*(x, z) \land E^*(z, y)) \]

\[ \varphi_{tc}(R, x, y) \equiv x = y \lor E(x, y) \lor \exists z(R(x, z) \land R(z, y)) \]

\[ \varphi_{tc}^G : \text{binRel}(G) \rightarrow \text{binRel}(G) \]

monotone \[ R \subseteq S \Rightarrow \varphi_{tc}^G(R) \subseteq \varphi_{tc}^G(S) \]

\[ \text{REACH} = \{ G, s, t \mid s \rightarrow^* t \} \]

\[ \text{REACH} \notin \text{FO} \]
Inductive Definitions and Least Fixed Point

\[ E^*(x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z(E^*(x, z) \land E^*(z, y)) \]

\[ \varphi_{tc}(R, x, y) \equiv x = y \lor E(x, y) \lor \exists z(R(x, z) \land R(z, y)) \]

\[ \varphi^G_{tc} : \text{binRel}(G) \rightarrow \text{binRel}(G) \]

monotone \[ R \subseteq S \Rightarrow \varphi^G_{tc}(R) \subseteq \varphi^G_{tc}(S) \]

\[ E^* = (\text{LFP}\varphi_{tc}) \]

\[ \text{REACH} = \{ G, s, t \mid s \xrightarrow{\ast} t \} \]

\[ \text{REACH} \not\in \text{FO} \]
Inductive Definitions and Least Fixed Point

\[ E^*(x, y) \overset{\text{def}}{=} x = y \lor E(x, y) \lor \exists z(E^*(x, z) \land E^*(z, y)) \]

\[ \varphi_{tc}(R, x, y) \equiv x = y \lor E(x, y) \lor \exists z(R(x, z) \land R(z, y)) \]

\[ \varphi^G_{tc} : \text{binRel}(G) \rightarrow \text{binRel}(G) \]

**monotone** \[ R \subseteq S \Rightarrow \varphi^G_{tc}(R) \subseteq \varphi^G_{tc}(S) \]

\[ G \in \text{REACH} \iff G \models (\text{LFP}\varphi_{tc})(s, t) \]

\[ E^* = (\text{LFP}\varphi_{tc}) \]

\[ \text{REACH} = \{ G, s, t \mid s \xrightarrow{*} t \} \]

\[ \text{REACH} \not\in \text{FO} \]
**Thm.** \( P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \)

\( \text{FO}[n^{O(1)}] \) means for graphs with \( n \) vertices, the formula \( \varphi_n \) expressing the property has \( n^{O(1)} \) quantifiers, but only a **fixed number** of requantified **variables**, \( x_1, \ldots, x_k \), i.e., \( \varphi_n \in \mathcal{L}^k \).
Thm. \[ P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \]

Graphs are completely general structures, i.e., any structure can be encoded as a graph.

\text{FO}[n^{O(1)}] \text{ means for graphs with } n \text{ vertices, the formula } \varphi_n \text{ expressing the property has } n^{O(1)} \text{ quantifiers, but only a \textbf{fixed number} of requantified variables, } x_1, \ldots, x_k, \text{ i.e, } \varphi_n \in \mathcal{L}^k.
Thm. \[ P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \]

Graphs are completely general structures, i.e., any structure can be encoded as a graph. **Restrict to graphs.**

\text{FO}[n^{O(1)}] \text{ means for graphs with } n \text{ vertices, the formula } \varphi_n \text{ expressing the property has } n^{O(1)} \text{ quantifiers, but only a fixed number of requantified variables, } x_1, \ldots, x_k, \text{ i.e, } \varphi_n \in \mathcal{L}^k.
LFP is a Polynomial Iteration Operator

**Thm.**  \[ P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \]

Graphs are completely general structures, i.e., any structure can be encoded as a graph. **Restrict to graphs.**

\( \text{FO}[n^{O(1)}] \) means for graphs with \( n \) vertices, the formula \( \varphi_n \) expressing the property has \( n^{O(1)} \) quantifiers, but only a **fixed number** of requantified **variables** \( x_1, \ldots, x_k \), i.e., \( \varphi_n \in \mathcal{L}^k \).

Above **Thm** requires **ordering relation**, \( \leq \).
**Thm.** \( P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \)

Graphs are completely general structures, i.e., any structure can be encoded as a graph. **Restrict to graphs.**

\( \text{FO}[n^{O(1)}] \) means for graphs with \( n \) vertices, the formula \( \varphi_n \) expressing the property has \( n^{O(1)} \) quantifiers, but only a **fixed number** of requantified **variables**, \( x_1, \ldots, x_k \), i.e, \( \varphi_n \in \mathcal{L}^k \).

Above **Thm** requires **ordering relation**, \( \leq \).

Necessary for encoding computation – **inputs to computers are ordered**.
**Thm.** \[ P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}] \]

Graphs are completely general structures, i.e., any structure can be encoded as a graph. **Restrict to graphs.**

\( \text{FO}[n^{O(1)}] \) means for graphs with \( n \) vertices, the formula \( \varphi_n \) expressing the property has \( n^{O(1)} \) quantifiers, but only a **fixed number** of requantified variables, \( x_1, \ldots, x_k \), i.e, \( \varphi_n \in \mathcal{L}^k \).

Above **Thm** requires **ordering relation**, \( \leq \).

Necessary for encoding computation – **inputs to computers are ordered.**

Unnatural for graphs – the ordering of the vertices is **irrelevant.**
Thm. \quad P = \text{FO}(\text{LFP}) = \text{FO}[n^{O(1)}]

Graphs are completely general structures, i.e., any structure can be encoded as a graph. \quad \textbf{Restrict to graphs.}

\text{FO}[n^{O(1)}] \text{ means for graphs with } n \text{ vertices, the formula } \varphi_n \text{ expressing the property has } n^{O(1)} \text{ quantifiers, but only a \textbf{fixed number} of requantified \textit{variables}, } x_1, \ldots, x_k, \text{ i.e, } \varphi_n \in \mathcal{L}^k.

Above \textbf{Thm} requires \textbf{ordering relation}, \leq.

Necessary for encoding computation – \textbf{inputs to computers are ordered}.

Unnatural for graphs – the ordering of the vertices is \textbf{irrelevant}.

\textbf{Wanted:} a language capturing Order-Independent P (OIP).
Want to Capture Order-Indepdendent P (OIP)

\[ \text{FO}(\text{LFP}) = P \]

\[ \text{FO}(\text{wo} \leq)\text{(LFP)} \subseteq \text{OIP} \]
Want to Capture Order-Independent P (OIP)

\[ \text{FO}(LFP) = \text{P} \]

\[ \text{FO}(\text{wo} \leq)(LFP) \subseteq \text{OIP} \]

\[ \text{EVEN} \overset{\text{def}}{=} \{ G \mid |V^G| \equiv 0 \pmod{2} \} \]
Want to Capture Order-Independent P (OIP)

\[ \text{FO} \left( \text{LFP} \right) = P \]

\[ \text{FO} \left( \text{wo} \leq \right) \left( \text{LFP} \right) \subseteq \text{OIP} \]

\[ \text{EVEN} \overset{\text{def}}{=} \{ G \mid |V^G| \equiv 0 \pmod{2} \} \]

\[ \text{EVEN} \in \text{OIP} - \text{FO} \left( \text{wo} \leq \right) \left( \text{LFP} \right). \]
Want to Capture Order-Independent $P$ (OIP)

$$
\text{FO}(\text{LFP}) = P
$$

$$
\text{FO}(\text{wo} \leq) \text{(LFP)} \subseteq \text{OIP}
$$

$$
\text{EVEN} \overset{\text{def}}{=} \{ G \mid |V^G| \equiv 0 \pmod{2} \}
$$

$$
\text{EVEN} \in \text{OIP} - \text{FO}(\text{wo} \leq) \text{(LFP)}.
$$

Thus, $\text{FO}(\text{wo} \leq) \text{(LFP)} \subset \not\subseteq \text{OIP}$
Want to Capture Order-Independent P (OIP)

\[
\text{FO}(\text{LFP}) = \text{P}
\]

\[
\text{FO}(\text{wo}\leq)(\text{LFP}) \subseteq \text{OIP}
\]

\[
\text{EVEN} \overset{\text{def}}{=} \{ G \mid |V^G| \equiv 0 \pmod{2} \}
\]

\[
\text{EVEN} \in \text{OIP} - \text{FO}(\text{wo}\leq)(\text{LFP}).
\]

Thus,

\[
\text{FO}(\text{wo}\leq)(\text{LFP}) \not\subseteq \text{OIP}
\]

How do we prove \(
\text{EVEN} \not\in \text{FO}(\text{wo}\leq)(\text{LFP})
\)?
Ehrenfeucht-Fraïssé Game

$g^k_m(G, H)$  $m$ moves,  $k$ pebbles,  2 players

Samson: show a difference.

Delilah: preserve isomorphism.

For all $m$, $D$ wins $g^2_m(G, H)$; but $S$ wins $g^3_3(G, H)$. 
Ehrenfeucht-Fraïssé Game

$g^k_m(G, H)$  $m$ moves,  $k$ pebbles,  2 players

**Samson**: show a difference.
Ehrenfeucht-Fraïssé Game

$g^k_m(G, H)$\  $m$ moves, \  $k$ pebbles, \  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

$g_m^k(G, H)$  $m$ moves,  $k$ pebbles,  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

$G^k_m(G, H)$  $m$ moves,  $k$ pebbles,  2 players

Samson: show a difference.  Delilah: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

$g_m^k(G, H)$  $m$ moves,  $k$ pebbles,  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

$G_m^k(G, H) \quad m \text{ moves, } \quad k \text{ pebbles, } \quad 2 \text{ players}$

**Samson**: show a difference.  **Delilah**: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

$g^k_m(G, H)$  $m$ moves, $k$ pebbles, 2 players

**Samson:** show a difference.  **Delilah:** preserve isomorphism.
Ehrenfeucht-Fraïssé Game

\[ g^k_m(G, H) \] \( m \) moves, \( k \) pebbles, 2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.
Ehrenfeucht-Fraïssé Game

\( \mathcal{G}_m^k(G, H) \) \( m \) moves, \( k \) pebbles, 2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.

For all \( m \), \( D \) wins \( \mathcal{G}_m^2(G, H) \);
Ehrenfeucht-Fraïssé Game

$$G^k_m(G, H) \quad m \text{ moves, } \quad k \text{ pebbles, } \quad 2 \text{ players}$$

**Samson**: show a difference.  **Delilah**: preserve isomorphism.

For all $$m$$, D wins $$G^2_m(G, H)$$;
Ehrenfeucht-Fraïssé Game

$G^k_m(G, H)$  $m$ moves,  $k$ pebbles,  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.

For all $m$, $D$ wins $G^2_m(G, H)$;
Ehrenfeucht-Fraïssé Game

$\mathcal{G}_m^k(G, H)$  $m$ moves,  $k$ pebbles,  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.

For all $m$, $D$ wins $\mathcal{G}_m^2(G, H)$;

![Graphs G and H with their respective vertices and edges labeled with variables $x_1, x_2, x_3$]
Ehrenfeucht-Fraïssé Game

\( g^k_m(G, H) \)  \( m \) moves,  \( k \) pebbles,  2 players

**Samson**: show a difference.  **Delilah**: preserve isomorphism.  
For all \( m \), \( D \) wins \( g^2_m(G, H) \);  \( \) but \( S \) wins \( g^3_3(G, H) \).
Notation: \( G \sim^k_m H \) means that Delilah has a winning strategy for \( G^k_m(G, H) \).
**Notation:** \( G \sim_m^k H \) means that Delilah has a winning strategy for \( G^k_m(G, H) \).

**Thm.** D has a winning strategy on the \( m \)-move, \( k \)-pebble game on \( G, H \) iff \( G \) and \( H \) agree on all formulas using \( k \) variables and quantifier depth \( m \).

\[ G \sim_m^k H \iff G \equiv_m^k H \]
**Thm.** \( \text{EVEN} \) requires \( n + 1 \) variables without ordering. Thus \( \text{EVEN} \not\in \text{FO}(\text{wo}\leq)(\text{LFP}) \).
Thm. \textbf{EVEN} requires \( n + 1 \) variables without ordering. Thus \textbf{EVEN} \( \not\in \text{FO}(\text{wo}\leq)(\text{LFP}) \).

**proof:**

\[
\begin{array}{cccc}
\text{\( G_{2m} \)} & \vdots & \text{\( H_{2m+1} \)} \\
g_1 & \vdots & h_1 \\
g_2 & \vdots & h_2 \\
g_{2m} & \vdots & h_{2m} \\
\end{array}
\]
Thm. **EVEN** requires $n + 1$ variables without ordering. Thus **EVEN** $\not\in \text{FO(wo} \leq \text{)}(\text{LFP})$.

**proof:**

\[ g_1 \quad x_1 \quad g_2 \quad \ldots \quad g_{2m} \]

\[ H_{2m+1} \quad \vdots \quad H_{2m+1} \]

\[ g_1 \quad h_1 \quad g_2 \quad h_2 \quad \ldots \quad g_{2m} \quad h_{2m} \quad h_{2m+1} \]
**Thm.** \textbf{EVEN} requires $n + 1$ variables without ordering. Thus \textbf{EVEN} $\not\in$ FO(wo$\leq$)(LFP).

**proof:**

\[
\begin{align*}
G_{2m} & 
\vdots \\
g_1 & x_1 \\
g_2 & \\
g_{2m} & \\
H_{2m+1} & \\
\vdots & \\
h_1 & x_1 \\
h_2 & \\
h_{2m} & \\
h_{2m+1} & 
\end{align*}
\]
Thm. **EVEN** requires \( n + 1 \) variables without ordering. Thus **EVEN** \( \not\in \text{FO}(\text{wo}\leq)(\text{LFP}) \).

**proof:**

\[
\begin{align*}
G_{2m} & : g_1 & x_1 \\
g_2 & : & \\
g_{2m} & : & \\
\vdots & : & \\
\end{align*}
\]

\[
\begin{align*}
H_{2m+1} & : h_1 & x_1 \\
h_2 & : & x_2 \\
h_{2m} & : & \\
h_{2m+1} & : &
\end{align*}
\]
Thm. **EVEN** requires $n + 1$ variables without ordering. Thus **EVEN** $\not\in \text{FO}(\text{wo} \leq)\text{(LFP)}$.

**proof:**

\[
\begin{align*}
  &g_1 \quad x_1 \\
  &g_2 \quad x_2 \\
  &\vdots \\
  &g_{2m} \\
  &G_{2m} \\
  &h_1 \quad x_1 \\
  &h_2 \quad x_2 \\
  &\vdots \\
  &h_{2m} \\
  &h_{2m+1} \\
  &H_{2m+1}
\end{align*}
\]
Thm. **EVEN** requires \( n + 1 \) variables without ordering. Thus **EVEN** \( \not\in \text{FO(wo}\leq\text{)}(\text{LFP}) \).

**proof:**

\[
\begin{align*}
G_{2m} & \quad : \\
g_1 & \quad x_1 \\
g_2 & \quad x_2 \\
g_{2m} & \quad x_{2m} \\
H_{2m+1} & \quad : \\
h_1 & \quad x_1 \\
h_2 & \quad x_2 \\
h_{2m} & \\
h_{2m+1}
\end{align*}
\]
Thm. **EVEN** requires \( n + 1 \) variables without ordering. Thus **EVEN** \( \not\in \text{ FO(wo}\leq\text{)}(\text{LFP}) \).

**proof:**

\[
G_{2m} : \\
g_1 x_1 \\
g_2 x_2 \\
g_{2m} x_{2m} \\
\\
h_1 x_1 \\
h_2 x_2 \\
h_{2m} x_{2m} \\
h_{2m+1}
\]

\[
l = 2m + 1\\n\therefore H_{2m+1}
\]
Thm. **EVEN** requires $n + 1$ variables without ordering. Thus **EVEN** $\not\in \text{FO(wo}\leq)(\text{LFP})$.

**proof:**

\[
\begin{align*}
G_{2m} & \sim^m H_{2m+1}
\end{align*}
\]
Add Counting to FO Logic

Two sorts: **Numbers**: \{0, 1, \ldots, n\}, \leq, Plus, Times and **Vertices**: \{v_1, \ldots, v_n\}, E, C_1, C_2 \ldots
Two sorts: **Numbers**: \{0, 1, \ldots, n\}, \leq, Plus, Times and **Vertices**: \{v_1, \ldots, v_n\}, \ E, C_1, C_2 \ldots

Combine with counting terms: \#x(\varphi(x)).
Add Counting to FO Logic

Two sorts: **Numbers**: \{0, 1, \ldots, n\}, \leq, Plus, Times and
**Vertices**: \{v_1, \ldots, v_n\}, E, C_1, C_2 \ldots

Combine with counting terms:  \#x(\varphi(x)).

\[
\text{EVEN} \equiv \exists i (\text{Plus}(i, i, \#x(x = x)))
\]
Two sorts: **Numbers**: \{0, 1, \ldots, n\}, \leq, Plus, Times and
**Vertices**: \{v_1, \ldots, v_n\}, E, C_1, C_2 \ldots

Combine with counting terms: \( \#x(\varphi(x)) \).

\[
\text{EVEN} \equiv \exists i \left( \text{Plus}(i, i, \#x(x = x)) \right)
\]

Let \( C^k \overset{\text{def}}{=} \text{FO}^k(\text{COUNT}) \); \( \text{FPC} \overset{\text{def}}{=} \text{FO}(\text{LFP}, \text{COUNT}) \).
Two sorts: **Numbers**: \{0, 1, \ldots, n\}, \leq, Plus, Times and **Vertices**: \{v_1, \ldots, v_n\}, E, C_1, C_2 \ldots

Combine with counting terms: \#x(\varphi(x)).

\[
\text{EVEN} \equiv \exists i \ (\text{Plus}(i, i, \#x(x = x)))
\]

Let \( C^k \overset{\text{def}}{=} \text{FO}^k(\text{COUNT}) \); \ FPC \overset{\text{def}}{=} \text{FO}(\text{LFP}, \text{COUNT}).

\[
\text{FO(wo}\leq)(\text{LFP}) \subsetneq \ FPC \subseteq \text{OIP}
\]
Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.
Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.
Stable Coloring of Vertices

Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.
Stable Coloring of Vertices

Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.
Stable Coloring of Vertices

Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.
Stable Coloring of Vertices

Start with a colored graph, and repeatedly color each vertex by how many neighbors it has of each color.

**Thm.** Stable Coloring of Vertices $= C^2$ type.

Round $m$ of stable coloring is quantifier depth of $C^2$ formula.
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.

Thus, for almost all graphs, there is a linear time algorithm to canonize the graph, i.e., sort the vertices by their $C^2$ type, so that two graphs are isomorphic iff their canonical forms are equal.
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.

Thus, for almost all graphs, there is a linear time algorithm to canonize the graph, i.e., sort the vertices by their $C^2$ type, so that two graphs are isomorphic iff their canonical forms are equal.

With high probability, $G \cong H$ iff $G \equiv_4^2 H$. 
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.

Thus, for almost all graphs, there is a linear time algorithm to canonize the graph, i.e., sort the vertices by their $C^2$ type, so that two graphs are isomorphic iff their canonical forms are equal.

With high probability, $G \cong H$ iff $G \equiv_4^2 H$.

Thus, Graph Isomorphism (GI) is linear time for random graphs.
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.

Thus, for almost all graphs, there is a linear time algorithm to canonize the graph, i.e., sort the vertices by their $C^2$ type, so that two graphs are isomorphic iff their canonical forms are equal.

With high probability, $G \cong H$ iff $G \equiv_4^2 H$.

Thus, Graph Isomorphism (GI) is linear time for random graphs.

In general the complexity of GI is unknown.
**Thm.** [Babai, Erdos, Selkow] With high probability, after four iterations of stable coloring, each vertex of a random graph has a unique color, i.e., the $C_4^2$-type of each vertex is unique.

Thus, for almost all graphs, there is a linear time algorithm to canonize the graph, i.e., sort the vertices by their $C^2$ type, so that two graphs are isomorphic iff their canonical forms are equal.

With high probability, $G \cong H$ iff $G \equiv^2_4 H$.

Thus, Graph Isomorphism (GI) is linear time for random graphs.

In general the complexity of GI is unknown.

**Thm.** [Babai, 2015] GI $\in$ DTIME[$n^{\log_7 n}$]. (Before this it was only known that GI $\in$ DTIME[$n^{\sqrt{n}}$].)
Def. Language $\mathcal{L}$ characterizes a graph $G$ iff for all graphs $H$,

$$G \equiv_\mathcal{L} H \iff G \cong H.$$
Logics Characterizing Graphs

**Def.** Language $\mathcal{L}$ **characterizes** a graph $G$ iff for all graphs $H$,

$$G \equiv_{\mathcal{L}} H \iff G \cong H.$$  

- $C^2$ characterizes almost all random graphs.
- $C^2$ characterizes all trees.
- $C^3$ characterizes all graphs of color class size 3.
**Def.** Language $\mathcal{L}$ characterizes a graph $G$ iff for all graphs $H$,

$$G \equiv_{\mathcal{L}} H \iff G \cong H.$$  

- $C^2$ characterizes almost all random graphs.
- $C^2$ characterizes all trees.
- $C^3$ characterizes all graphs of color class size 3.

**Thm.** We can test if $G \equiv_{C^k} H$ in FPC and DTIME[$n^k \log n$].
**Def.** Language $\mathcal{L}$ characterizes a graph $G$ iff for all graphs $H$,

\[ G \equiv_{\mathcal{L}} H \iff G \cong H. \]

- $C^2$ characterizes almost all random graphs.
- $C^2$ characterizes all trees.
- $C^3$ characterizes all graphs of color class size 3.

**Thm.** We can test if $G \equiv_{C^k} H$ in FPC and DTIME[$n^k \log n$].

**Cor.** If $C^k$ characterizes all graphs in a class of graphs $\mathcal{G}$ that is closed under particularizing, then $\mathcal{G}$ admits $C^k$ canonization, and thus FPC captures OIP over $\mathcal{G}$.
Logics Characterizing Graphs

**Def.** Language $\mathcal{L}$ **characterizes** a graph $G$ iff for all graphs $H$,

$$G \equiv_{\mathcal{L}} H \iff G \cong H.$$ 

- $C^2$ characterizes almost all random graphs.
- $C^2$ characterizes all trees.
- $C^3$ characterizes all graphs of color class size 3.

**Thm.** We can test if $G \equiv_{C^k} H$ in FPC and DTIME[$n^k \log n$].

**Cor.** If $C^k$ characterizes all graphs in a class of graphs $\mathcal{G}$ that is closed under particularizing, then $\mathcal{G}$ admits $C^k$ canonization, and thus FPC captures **OIP** over $\mathcal{G}$.

**proof:** Apply arbitrary FO(LFP) formula to the canonical form of the input graph.
Is FPC Equal to OIP?

[1] Is FPC Equal to OIP?
Is FPC Equal to OIP?

- Is FPC Equal to OIP?
- Does $C^4$ characterize all graphs?
Is FPC Equal to OIP?

- Is FPC Equal to **OIP**?

- Does \( C^4 \) characterize all graphs?

- If yes, then \( \text{FPC} = \text{OIP} \) and for all graphs, \( G \cong H \iff G \cong_{C^4} H \).

  Thus, GI would be in \( \text{DTIME}[n^4 \log n] \).
Is FPC Equal to OIP?

- Is FPC Equal to \textbf{OIP}?

- Does \( C^4 \) characterize all graphs?

- If yes, then \( \text{FPC} = \text{OIP} \) and for all graphs,
  \[ G \cong H \iff G \equiv_{C^4} H. \]

  Thus, GI would be in \( \text{DTIME}[n^4 \log n] \).

**Thm.** [CFI] No!

A simple graph property (now called the CFI property) checkable in \( \text{DTIME}[n] \), requires \( v = \Omega(n) \) variables to express in \( C^v \). Thus, \( \text{CFI} \in \text{OIP} - \text{FPC} \)
Proof of CFI Thm

Each $m_i$ adjacent to an even number of $a_j$'s.

Automorphisms of $X$: switch an even number of $(a_i b_i)$ pairs.

Automorphism: $(a_2 b_2)(a_3 b_3)(m_1 m_2)(m_3 m_4)$
CFI Gadget $X$: Each $m_i$ adjacent to an even number of $a_j$'s.
Proof of CFI Thm

CFI Gadget $X$: Each $m_i$ adjacent to an even number of $a_j$’s.
Automorphisms of $X$: switch an **even number** of $(a_i,b_i)$ pairs.
Proof of CFI Thm

CFI Gadget $X$: Each $m_i$ adjacent to an even number of $a_j$’s.

Automorphisms of $X$: switch an **even number** of $(a_i b_i)$ pairs.

Automorphism: \[(a_2 b_2)(a_3 b_3)(m_1 m_2)(m_3 m_4)\]
Proof of CFI Thm

CFI Gadget $X$: Each $m_i$ adjacent to an even number of $a_j$’s.
Automorphisms of $X$: switch an even number of $(a_i b_i)$ pairs.

Automorphism: $(a_1 b_1)(a_2 b_2)(m_1 m_4)(m_2 m_3)$
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$. 

Thus $X(G_n)$ has color class size 4.
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$.

If we remove any $n$ vertices from $G_n$, it still has a connected component with more than $|V^{G_n}|/2$ vertices.
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$.

If we remove any $n$ vertices from $G_n$, it still has a connected component with more than $|V^{G_n}|/2$ vertices.

Such regular degree 3 graphs with linear-size separators exist.
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$.

If we remove any $n$ vertices from $G_n$, it still has a connected component with more than $|V^{G_n}|/2$ vertices.

Such regular degree 3 graphs with linear-size separators exist.

Color class size 1 means every vertex of $G_n$ has a unique color.
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$.

If we remove any $n$ vertices from $G_n$, it still has a connected component with more than $|V^{G_n}|/2$ vertices.

Such regular degree 3 graphs with linear-size separators exist.

Color class size 1 means every vertex of $G_n$ has a unique color.

Let $X(G_n)$ be the result of replacing each vertex $v \in V^{G_n}$ by a copy of $X$ of $v$’s color.
Let $G_n$ be a regular, degree 3 graph with $O(n)$ vertices, color class size 1 and separator size $n$.

If we remove any $n$ vertices from $G_n$, it still has a connected component with more than $|V^{G_n}|/2$ vertices.

Such regular degree 3 graphs with linear-size separators exist.

Color class size 1 means every vertex of $G_n$ has a unique color.

Let $X(G_n)$ be the result of replacing each vertex $v \in V^{G_n}$ by a copy of $X$ of $v$’s color.

Thus $X(G_n)$ has color class size 4.
$X(G_n)$: replace each vertex $v \in V^{G_n}$ by a copy of $X$ of $v$’s color, connecting $a$ to $a$ and $b$ to $b$. 
$\tilde{X}(G)$ is $X(G)$ with any one edge pair flipped.
$\tilde{X}(G)$ is $X(G)$ with any one edge pair flipped.
Prop. Let $X'(G_n)$ be $X(G_n)$ with some number, $m$, of the magenta edges flipped.

Then $X'(G_n) \cong X(G_n)$ iff $m$ is even and $X'(G_n) \cong \tilde{X}(G_n)$ iff $m$ is odd.
Prop. Let $X'(G_n)$ be $X(G_n)$ with some number, $m$, of the magenta edges flipped.

Then $X'(G_n) \cong X(G_n)$ iff $m$ is even and

$$X'(G_n) \cong \tilde{X}(G_n)$$

iff $m$ is odd.

proof: Using the automorphisms of $X$, we can move any two flips towards each other until they eliminate each other.
Is it $X(G_n)$ or $\tilde{X}(G_n)$?

**Prop.** Let $X'(G_n)$ be $X(G_n)$ with some number, $m$, of the magenta edges flipped.

Then $X'(G_n) \cong X(G_n)$ iff $m$ is even and

$$X'(G_n) \not\cong \tilde{X}(G_n)$$

iff $m$ is odd.
**Prop.** Let $X'(G_n)$ be $X(G_n)$ with some number, $m$, of the magenta edges flipped.

Then $X'(G_n) \cong X(G_n)$ iff $m$ is even and 

$$X'(G_n) \cong \tilde{X}(G_n)$$

iff $m$ is odd.

**proof:** Using the automorphisms of $X$, we can move any two flips towards each other until they eliminate each other.  \( \square \)
\[ \tilde{X}(G_n) \]
$\tilde{X}(G_n)$
Every one of the $m_i$’s is connected to an even number of $a_j$’s.
Every one of the $m_i$'s is connected to an odd number of $a_j$'s.
Def. \( \text{CFI} = \{ (X'(G) \mid X'(G) \cong X(G) \} \) for \( G \) is connected, reg. deg. 3, \( cc(G) = 1 \).
The CFI Problem

**Def.** CFI = \{ (X'(G) \mid X'(G) \cong X(G) ) \} for G is connected, reg. deg. 3, cc(G) = 1.

**Prop.** CFI ∈ DTIME[\eta].
The CFI Problem

Def. \( \text{CFI} = \left\{ (X'(G) \mid X'(G) \cong X(G) \right\} \) for \( G \) is connected, reg. deg. 3, \( \text{cc}(G) = 1 \).

Prop. \( \text{CFI} \in \text{DTIME}[n] \).

proof Use the ordering to label boundary pairs \( a_i, b_i \) when \( a_i \leq b_i \). Then count the number, \( m \), of flips of vertices and edges mod 2. \( X'(G) \in \text{CFI} \) iff \( m \) is even.
\( \tilde{X}(G_n) \)
Thm. \[ \text{CFI} \in \text{OIP} - \text{FPC}. \]
Thm. $\text{CFI} \in \text{OIP} - \text{FPC}$.  

proof We show that $X(G_n) \equiv_{C^n} \tilde{X}(G_n)$.  

Counting doesn't help since $c(G_n) = 4$. Suffices to show that $X(G_n) \sim_n \tilde{X}(G_n)$.  

Initially no pebbles on the board, Samson places $x_1$ on $X(v)$ in one of the two graphs. Note that the largest connected component $C_1$ of $G - \{v\}$ includes over half the vertices of $G$.  

Delilah moves the flip into $C_1$. If she removes the flip, then the two graphs are isomorphic. Delilah answers according to this isomorphism.  

Inductively, after step $m$, Delilah has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.
**Thm.** CFI ∈ OIP − FPC.

**proof** We show that $X(G_n) \equiv^n \tilde{X}(G_n)$.

Counting doesn’t help since $cc(X(G_n)) = 4$. Suffices to show that $X(G_n) \sim^n \tilde{X}(G_n)$. 

Initially no pebbles on the board, Samson places $x_1$ on $X(v)$ in one of the two graphs. Note that the largest connected component $C_1$ of $G - \{v\}$ includes over half the vertices of $G$. Delilah moves the flip into $C_1$. If she removes the flip, then the two graphs are isomorphic. Delilah answers according to this isomorphism. Inductively, after step $m$, Delilah has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.
Thm. \( \text{CFI} \in \text{OIP} - \text{FPC} \).

**proof** We show that \( X(G_n) \equiv_{C^n} \tilde{X}(G_n) \).

Counting doesn’t help since \( \text{cc}(X(G_n)) = 4 \). Suffices to show that \( X(G_n) \sim^n \tilde{X}(G_n) \).

Initially no pebbles on the board, **Samson** places \( x_1 \) on \( X(\nu) \) in one of the two graphs. Note that the largest connected component \( C_1 \) of \( G - \{\nu\} \) includes over half the vertices of \( G \). **Delilah** moves the flip into \( C_1 \). If she removes the flip, then the two graphs are isomorphic. **Delilah** answers according to this isomorphism.
Thm. \( \text{CFI} \in \text{OIP} - \text{FPC} \).

**proof** We show that \( X(G_n) \equiv_{C^n} \tilde{X}(G_n) \).

Counting doesn’t help since \( \text{cc}(X(G_n)) = 4 \). Suffices to show that \( X(G_n) \sim^n \tilde{X}(G_n) \).

Initially no pebbles on the board, **Samson** places \( x_1 \) on \( X(\nu) \) in one of the two graphs. Note that the largest connected component \( C_1 \) of \( G - \{\nu\} \) includes over half the vertices of \( G \). **Delilah** moves the flip into \( C_1 \). If she removes the flip, then the two graphs are isomorphic. **Delilah** answers according to this isomorphism.

Inductively, after step \( m \), **Delilah** has not yet lost, so there is an isomorphism from chosen points in \( X(G_n) \) to chosen points in \( \tilde{X}(G_n) \) which extends to an isomorphism of the whole graphs in which a flip in \( \tilde{G}_n \) in \( C_m \) has been removed.
Inductively, after step $m$, Delilah has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.
Inductively, after step $m$, **Delilah** has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.

**Samson** picks up the $x_i$ pebbles and places one on some $X(v)$. Note that $C_m$ and $C_{m+1}$ both contain over half the vertices of $G_n$.

Thus they have some vertex $w \in C_m \cap C_{m+1}$. 
Inductively, after step $m$, **Delilah** has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.

**Samson** picks up the $x_i$ pebbles and places one on some $X(v)$. Note that $C_m$ and $C_{m+1}$ both contain over half the vertices of $G_n$.

Thus they have some vertex $w \in C_m \cap C_{m+1}$.

**Delilah** mentally moves the flip to $X(w)$. She then answers according to the isomorphism from $X(G_n)$ to $\tilde{X}(G_n)$ where that flip in $X(w)$ has been removed.
Inductively, after step $m$, Delilah has not yet lost, so there is an isomorphism from chosen points in $X(G_n)$ to chosen points in $\tilde{X}(G_n)$ which extends to an isomorphism of the whole graphs in which a flip in $\tilde{G}_n$ in $C_m$ has been removed.

Samson picks up the $x_i$ pebbles and places one on some $X(v)$. Note that $C_m$ and $C_{m+1}$ both contain over half the vertices of $G_n$.

Thus they have some vertex $w \in C_m \cap C_{m+1}$.

Delilah mentally moves the flip to $X(w)$. She then answers according to the isomorphism from $X(G_n)$ to $\tilde{X}(G_n)$ where that flip in $X(w)$ has been removed.

Thus Delilah never loses.\qed
Recap

We have shown that the linear-time CFI problem is in \( \mathcal{OIP} - \mathcal{FPC} \).

**Cor.** \( \Omega(n) \) variables are needed to characterize graphs.
Martin Grohe has shown that many classes of graphs are characterized by $C^k$ for some $k$. This includes planer graphs, graphs of bounded genus, graphs of bounded tree width and culminating in

Thm. [Anderson, Dawar and Holm] Linear Programming is in FPC.
Martin Grohe has shown that many classes of graphs are characterized by \( C^k \) for some \( k \). This includes planer graphs, graphs of bounded genus, graphs of bounded tree width and culminating in

**Thm. [Grohe]** Any class \( \mathcal{G} \) of graphs that excludes some minor is characterized by \( C^k \) for some fixed \( k \).
Martin Grohe has shown that many classes of graphs are characterized by $C^k$ for some $k$. This includes planer graphs, graphs of bounded genus, graphs of bounded tree width and culminating in

**Thm.** [Grohe] Any class $\mathcal{G}$ of graphs that excludes some minor is characterized by $C^k$ for some fixed $k$. Thus,

- FPC captures **OIP** on $\mathcal{G}$. Thus, for graphs from $\mathcal{G}$, graph isomorphism and canonization are in P.

- For $G, H \in \mathcal{G}$, $G \cong H$ iff $G \equiv_{C^k} H$. 

**Thm.** [Anderson, Dawar and Holm] Linear Programming is in FPC.
Martin Grohe has shown that many classes of graphs are characterized by $C^k$ for some $k$. This includes planer graphs, graphs of bounded genus, graphs of bounded tree width and culminating in

**Thm.** [Grohe] Any class $\mathcal{G}$ of graphs that excludes some minor is characterized by $C^k$ for some fixed $k$. Thus,

- FPC captures $OIP$ on $\mathcal{G}$. Thus, for graphs from $\mathcal{G}$, graph isomorphism and canonization are in P.

- For $G, H \in \mathcal{G}$, $G \cong H$ iff $G \equiv_{C^k} H$.

**Thm.** [Anderson, Dawar and Holm] Linear Programming is in FPC.
Two other languages are candidates for capturing $O(\log n)$ variables.
Going Beyond FPC

Two other languages are candidates for capturing **OIP**:

- **Choiceless Polynomial Time (CPT)** [Blass and Gurevich]
  Compute using sets of sets of sets, etc., where instead of choosing the first vertex, we consider the set of all such choices, keeping the total size of all sets polynomial.
Two other languages are candidates for capturing \textbf{OIP}:

- **Choiceless Polynomial Time (CPT) [Blass and Gurevich]**
  Compute using sets of sets of sets, etc., where instead of choosing the first vertex, we consider the set of all such choices, keeping the total size of all sets polynomial.

- **Rank Logic [Dawar, Grohe, Holm, and Laubner]**
  Compute the rank of matrices expressed in an unordered setting.
Two other languages are candidates for capturing OIP:

- **Choiceless Polynomial Time (CPT) [Blass and Gurevich]** Compute using sets of sets of sets, etc., where instead of choosing the first vertex, we consider the set of all such choices, keeping the total size of all sets polynomial.

- **Rank Logic [Dawar, Grohe, Holm, and Laubner]** Compute the rank of matrices expressed in an unordered setting.

CFI is expressible in CPT and in Rank Logic, thus these are strict extensions of FPC.
Two other languages are candidates for capturing **OIP**:

- **Choiceless Polynomial Time (CPT) [Blass and Gurevich]**
  Compute using sets of sets of sets, etc., where instead of choosing the first vertex, we consider the set of all such choices, keeping the total size of all sets polynomial.

- **Rank Logic [Dawar, Grohe, Holm, and Laubner]**
  Compute the rank of matrices expressed in an unordered setting.

CFI is expressible in CPT and in Rank Logic, thus these are strict extensions of **FPC**.

What I want: more natural extension to **FPC** that adds group theory and characterizes graphs using $O(\log n)$ variables.