# Expressibility of simple unary generalized quantifier

Shohei Okisaka

Mathematical Institute, Tohoku University

September 5, 2014

# Outline

Many theorems in model theory fail if we restrict to *finite structures*.

- Compactness
  Let $T = \{\varphi_{\geq n} \mid n \geq 1\}$ where $\varphi_{\geq n}$ means "there are at least $n$ elements", then, any finite subset of T is satisfiable in finite structures but $T$ is not.
- Completeness

**Theorem (Trakhtenbrot(1950))**

*The halting problem can be reducible to finitely satisfiability problem. i.e for any TM M, we can construct FO-sentence $\varphi_M$ which satisfying:*
*$M(< M >)$ halts iff $\varphi_M$ is satisfiable by finite structure.*

R. Fagin show the first *descriptive complexity* result.

**Theorem (Fagin(1974))**

*Let $K$ be a class of finite structures, then*

$$K \text{ is } \Sigma^1_1 \text{ definable} \Leftrightarrow K \text{ is NP-computable}$$

Rmk $K$ is NP-computable means if finite structure $\mathcal{A}$ is given, then it is NP-computable to decide whether $\mathcal{A} \in K$.

e.g) (undirected) graph $\mathcal{G}$ is 3-colorable iff $\mathcal{G}$ satisfies

$$\exists C_1 \exists C_2 \exists C_3 ((\forall x (C_1(x) \vee C_2(x) \vee C_3(x))) \wedge (\forall x \forall y (E(x,y) \rightarrow \bigwedge \neg (C_i(x) \wedge C_i(y)))))$$

Other complexity classes are also characterized
if we restrict to *ordered structures*.

| complexity | logic |
|---|---|
| $AC^0$ | $FO(\leq, +, \times)(Immerman, 88)$ |
| $AC^0(m)$ | $FO{+}D_m(\leq, +, \times)$ |
| $TC^0$ | $FO{+}M(\leq, +, \times)$ |
| NL | $FO{+}TC\ operator(\leq)(Immerman, 83)$ |
| P | FO+*least fixpoint operator*$(\leq)$ (Immerman, Vardi 82) |
| PSPACE | $FO{+}partial\ fixpoint\ operator(\leq)(Vardi, 82)$ |

$$\mathcal{A} \models D_m x\varphi(x) :\Leftrightarrow \#\{a \in A \mid \mathcal{A} \models \varphi(a)\} \equiv 0 \mod m$$
$$\mathcal{A} \models Mx\varphi(x) :\Leftrightarrow \#\{a \in A \mid \mathcal{A} \models \varphi(a)\} \geq \#A/2$$

How can you show ”class $K$ is not definable in logic $\mathcal{L}$?”

How can you show "class $K$ is not definable in logic $\mathcal{L}$?"
$\rightarrow$Ehrenfeucht-Frásse game is a tool to show such undefinability.

Let $\tau$ be *finite relational vocabulary*,
$\mathcal{A}, \mathcal{B}$ be $\tau$-str, $k, m \geq 0$, $\bar{a} \in A^k, \bar{b} \in B^k$
$m$-round EF-game $G_m((\mathcal{A}, \bar{a}), (\mathcal{B}, \bar{b}))$ is defined as follows.

- There are two players (I and II)
- This game consists of $m$-rounds
- $i$-th round (FO-move)
  - I choose $\mathcal{A}$ or $\mathcal{B}$, (assume choose $\mathcal{A}$,) I choose $c_i \in A$
  - Then II choose $d_i \in B$ (similarly when I choose $\mathcal{B}$)
- After $m$-th round,
  II win iff $\bar{a}c_1 \cdots c_m \mapsto \bar{b}d_1 \cdots d_m$ is *partial isomorphism*.

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | |
| $\mathbb{Z}$ | | |

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | |
| $\mathbb{Z}$ | $a$ | |

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | |
| $\mathbb{Z}$ | $a$ | $a-1$ |

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | ? |
| $\mathbb{Z}$ | $a$ | $a - 1$ |

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | ? |
| $\mathbb{Z}$ | $a$ | $a-1$ |

For any element $x$, $0x \mapsto a(a-1)$ is not partial isomorphism.

Let's play $G_2((\mathbb{N}, \leq), (\mathbb{Z}, \leq))$

| str | 1 | 2 |
|---|---|---|
| $\mathbb{N}$ | 0 | ? |
| $\mathbb{Z}$ | $a$ | $a-1$ |

For any element $x$, $0x \mapsto a(a-1)$ is not partial isomorphism.
In fact, $\mathbb{N} \models \exists x \forall y (x \leq y)$ & $\mathbb{Z} \not\models \exists x \forall y (x \leq y)$

The *quantifier rank* $\mathrm{qr}(\varphi)$ of FO formula $\varphi$ is defined as follows.
$\varphi$:atomic $\Rightarrow \mathrm{qr}(\varphi)=0$, $\mathrm{qr}(\neg\varphi)=\mathrm{qr}(\varphi)$,
$\mathrm{qr}(\varphi \vee \psi)=\max\{\mathrm{qr}(\varphi), \mathrm{qr}(\psi)\}$, $\mathrm{qr}(\exists x \varphi)=\mathrm{qr}(\varphi)+1$

**Theorem**

*The followings are equivalent.*

1. *II has winning strategy in* $G_m((\mathcal{A}, \bar{a}), (\mathcal{B}, \bar{b}))$
2. $(\mathcal{A}, \bar{a}) \equiv_m (\mathcal{B}, \bar{b})$

$$(\mathcal{A}, \bar{a}) \equiv_m (\mathcal{B}, \bar{b}) :\Leftrightarrow \forall\varphi\, (\mathrm{qr}(\varphi) \leq m \Rightarrow \mathcal{A} \models \varphi(\bar{a}) \textit{ iff } \mathcal{B} \models \varphi(\bar{b}))$$

If we want to show the statement "$K$ is not definable in FO", it's enough to show

$$\forall n \in \mathbb{N}, \exists \mathcal{A} \in K \ \& \ \exists \mathcal{B} \notin K \ s.t \ \mathcal{A} \equiv_n \mathcal{B}$$

Using EF-game, we can show FO can not define the following classes.

- $\{(A, P^{\mathcal{A}}) \mid \#P \equiv 0 \mod m\}$
- $\{(A, \leq) \mid \#A \text{ is even.}\}$

If we want to show the statement "$K$ is not definable in FO", it's enough to show

$$\forall n \in \mathbb{N}, \exists \mathcal{A} \in K \ \& \ \exists \mathcal{B} \notin K \ s.t \ \mathcal{A} \equiv_n \mathcal{B}$$

Using EF-game, we can show FO can not define the following classes.

- $\{(A, P^{\mathcal{A}}) \mid \#P \equiv 0 \mod m\}$
- $\{(A, \leq) \mid \#A \text{ is even.}\}$

The expressibility of FO is so limited.
We consider to extend FO by adding new quantifier.

- First-order formula cannot describe such as
” there are finitely many ...” or ” there are uncountably many...”

- First-order formula cannot describe such as
” there are finitely many ...” or ” there are uncountably many...”

- Mostowski introduced *generalized quantifier* to express such sentence in 1957.

- First-order formula cannot describe such as
  " there are finitely many ..." or " there are uncountably many..."
- Mostowski introduced *generalized quantifier* to express such sentence in 1957.
- Lindström extended the concept in 1966, which is also called *Lindström quantifier*.

.

Let $\tau := \{R_1, \cdots, R_m\}$ be finite relational vocabulary and $K$ a class of finite $\tau$-str.

## Definition

generalized quantifier $Q_K$ given by $K$ is defined as follows:
for any finite str $\mathcal{A}$,
$\mathcal{A} \models Q_K \bar{x}_1, \cdots, \bar{x}_m(\varphi_1(\bar{x}_1), \cdots, \varphi_m(\bar{x}_m)) \Leftrightarrow (A, \varphi_1^{\mathcal{A}}, \cdots, \varphi_m^{\mathcal{A}}) \in K$
where $\bar{x}_k$ is seq of variables which length is equal to the arity of $R_k$
and $\varphi_k^{\mathcal{A}} := \{\bar{a} \mid \mathcal{A} \models \varphi_k(\bar{a})\}$

We denote the extension of FO equipped with generalized quantifier $Q_K$ by FO($Q_K$).

$Q_K$ is called *simple* if $\tau$ has only one relation symbol and *unary* if $\tau$ has only unary symbols.

・Examples・

Let $P, Q$ be unary relation symbols.

- $K_\exists = \{(A, P^{\mathcal{A}}) \mid P^{\mathcal{A}} \neq \emptyset\}$,
  $\mathcal{A} \models Q_{K_\exists} x\varphi(x) \Leftrightarrow \varphi^{\mathcal{A}} \neq \emptyset \Leftrightarrow \mathcal{A} \models \exists x\varphi(x)$.
- $D_3 = \{(A, P^{\mathcal{A}}) \mid \#P^{\mathcal{A}} \equiv 0 \mod 3\}$,
  $\mathcal{A} \models Q_{D_3} x\varphi(x) \Leftrightarrow \#\varphi^{\mathcal{A}} \equiv 0 \mod 3 \Leftrightarrow \mathcal{A} \models D_3 x\varphi(x)$.
- $M = \{(A, P^{\mathcal{A}}) \mid \#P^{\mathcal{A}} \geq \#A/2\}$,
  $\mathcal{A} \models Q_M x\varphi(x) \Leftrightarrow \#\varphi^{\mathcal{A}} \geq \#A/2 \Leftrightarrow \mathcal{A} \models Mx\varphi(x)$.
- $I = \{(A, P^{\mathcal{A}}, Q^{\mathcal{A}}) \mid \#P^{\mathcal{A}} = \#Q^{\mathcal{A}}\}$,
  $\mathcal{A} \models Q_I x, y(\varphi(x), \psi(y)) \Leftrightarrow \#\varphi^{\mathcal{A}} = \#\psi^{\mathcal{A}}$.

Using generalized quantifiers, we can restate the characterization of some complexity classes.

| complexity class | logic |
|---|---|
| $AC^0(m)$ | $FO(Q_{D_m})(\leq, +, \times)$ |
| $TC^0$ | $FO(Q_M)(\leq, +, \times)$ |

Using generalized quantifiers, we can restate the characterization of some complexity classes.

| complexity class | logic |
|---|---|
| $AC^0(m)$ | $FO(Q_{D_m})(\leq, +, \times)$ |
| $TC^0$ | $FO(Q_M)(\leq, +, \times)$ |

Can we also characterize other classes like P or NP in terms of generalized quantifier ??

To capture P in terms of generalized quantifier,
we need more definition.

For $\tau = \{R_1, \cdots, R_m\}$, $k > 0$, let $\tau(k) = \{R_1^k, \cdots, R_m^k\}$
where if $R_i$ is l-ary relation symbol, $R_i^k$ is kl-ary relation symbol.

## Definition

Let $K$ be a class of $\tau$-str. *k-th vectorization* of K is class of $\tau(k)$-str defined as follows:

$$K^k := \{(A, (R_1^k)^{\mathcal{A}}, \cdots, (R_m^k)^{\mathcal{A}}) \mid (A^k, (R_1^k)^{\mathcal{A}}, \cdots, (R_m^k)^{\mathcal{A}}) \in K\}$$

Rmk: If $(R_i^k)^{\mathcal{A}}$ is kl-ary relation over $A$,
we can see $(R_i^k)^{\mathcal{A}}$ as l-ary relation over $A^k$.

We denote the logic FO($\{Q_{K^l} \mid l > 0\}$) by FO+$K$.

· Examples ·

- $K_\exists = \{(A, P^{\mathcal{A}}) \mid P^{\mathcal{A}} \neq \emptyset\}$,
  $\mathcal{A} \models Q_{K_\exists^3} x_1 x_2 x_3 \varphi(x_1, x_2, x_3) \Leftrightarrow \mathcal{A} \models \exists x_1 \exists x_2 \exists x_3 \varphi(x_1, x_2, x_3)$.
- $D_3 = \{(A, P^{\mathcal{A}}) \mid \# P^{\mathcal{A}} \equiv 0 \mod 3\}$,
  $\mathcal{A} \models Q_{D_3^2} xy\varphi(x, y) \Leftrightarrow \#\{(a, b) \in A^2 \mid \mathcal{A} \models \varphi(a, b)\} \equiv 0$ mod 3.
- $M = \{(A, P^{\mathcal{A}}) \mid \# P^{\mathcal{A}} \geq \# A/2\}$,
  $\mathcal{A} \models Q_{M^2} xy\varphi(x, y) \Leftrightarrow \#\{(a, b) \in A^2 \mid \mathcal{A} \models \varphi(a, b)\} \geq \# A^2/2$.

We can define a class of finite structures which captures P, i.e.

**Fact**

*There is a class of finite structures $L_P$ s.t. for any class of finite ordered structures $K$, $K$ is P-computable iff $K$ is definable in FO+$L_P$.*
*The same statement holds for L, NL, NP, PSPACE.*

Note

- It is shown that P can't be captured by the logic FO($Q_K$) for any $K$ (Hella,1992).
- Some classes like $D_m$ collapse *vectorization hierarchy*. i.e. FO+$D_m$ is equivalent to FO($Q_{D_m}$).

We investigate expressibility of the most simplest case.
Let $\tau = \{P\}$ ($P$: unary), for $S \subseteq \mathbb{N}$, we define a class of $\tau$-str $K_S$ by

$$K_S := \{(A, P^{\mathcal{A}}) \mid \#P^{\mathcal{A}} \in S\}$$

Then, the semantics of the generalized quantifier is given by

$$\mathcal{A} \models Q_{K_S} x \varphi(x) \Leftrightarrow \#\varphi^{\mathcal{A}} \in S$$

We investigate expressibility of the most simplest case.
Let $\tau = \{P\}$ ($P$: unary), for $S \subseteq \mathbb{N}$, we define a class of $\tau$-str $K_S$ by

$$K_S := \{(A, P^{\mathcal{A}}) \mid \#P^{\mathcal{A}} \in S\}$$

Then, the semantics of the generalized quantifier is given by

$$\mathcal{A} \models Q_{K_S} x \varphi(x) \Leftrightarrow \#\varphi^{\mathcal{A}} \in S$$

Question.
Given two subset $S, T \subseteq \mathbb{N}$, when is FO+$K_T$ ( or FO($Q_{K_T}$) ) more expressive than FO+$K_S$ ( FO($Q_{K_S}$) ) ??

## Definition

For any logic $\mathcal{L}, \mathcal{L}'$, we say $\mathcal{L}'$ is *more expressive than* $\mathcal{L}$ ($\mathcal{L} \leq \mathcal{L}'$) if for any $\tau$ and any $\tau$-formula $\varphi$ in $\mathcal{L}$, there exists $\tau$-formula $\psi$ in $\mathcal{L}'$ which is equivalent to $\varphi$.

## Lemma

*For two classes $K, L$,*

1. *FO($Q_K$)$\leq$ FO($Q_L$) iff $K$ is definable in FO($Q_L$)*
2. *FO+$K$ $\leq$ FO+$L$ iff $K$ is definable in FO+$L$*

From now on, $\tau = \{P\}$ ($P$: unary), and $\mathcal{A}$ is $\tau$-str.
Given $S \subseteq \mathbb{N}$, let $S + m := \{n + m \mid n \in S\}$, then for example

$$\mathcal{A} \in K_{S+1} \Leftrightarrow \mathcal{A} \models \exists y(P(y) \wedge Q_{K_S} x(x \neq y \wedge P(x)))$$

So, $\mathrm{FO}(Q_{K_{S+1}}) \leq \mathrm{FO}(Q_{K_S})$.

**Theorem (Corredor(1986))**

*For $S, T \subseteq \mathbb{N}$,*
*$FO(Q_{K_S}) \leq FO(Q_{K_T})$ iff $\exists T' \in \mathcal{B}(\{T + m \mid m \geq 0\})$ s.t $\#(S \Delta T') < \infty$*

**Corollary**

*For $m, m' > 0$,*
*$FO(Q_{D_m}) \leq FO(Q_{D_{m'}})$ iff $m \mid m'$*

(Sketch of proof.) It's enough to show left to right.
At first, note that quantifier rank of $\varphi \in$ FO+$K_T$ is defined similarly.
Furthermore, EF-game for FO+$K_T$ is also defined as FO case but add $Q_{K_T}$-move:

- I choose $\mathcal{A}$ or $\mathcal{B}$ (assume choose $\mathcal{A}$), I choose $X \subset A$ which is closed under automorphism which fixes chosen elements ,
- II choose $Y \subseteq B$ which satisfies $\#X \in T$ *iff* $\#Y \in T$
- I choose $b \in Y$, then II choose $a \in X$.

We assume that $\forall T' \in \mathcal{B}(\{T + m \mid m \geq 0\})\ \#(S \Delta T') = \infty$, and show for any $n \in \mathbb{N}$, there exists $\mathcal{A} \in K_S$, $\mathcal{B} \notin K_S$ s.t
$\forall \varphi \in$ FO+$K_T$ $\mathrm{qr}(\varphi) \leq n \Rightarrow \mathcal{A} \models \varphi$ *iff* $\mathcal{B} \models \varphi$

we fix $n \in \mathbb{N}$,

**Lemma**

*there exists* $u \in S$ *&* $v \notin S$ *s.t*

- $u, v > n$
- *for any* $m < n$, $u \in T + m$ *iff* $v \in T + m$

Let $\mathcal{A} = (A, A)$, $\mathcal{B} = (B, B)$ where $\#A = u$, $\#B = v$.
Then $\mathcal{A} \in K_S$ & $\mathcal{B} \notin K_S$.

we fix $n \in \mathbb{N}$,

**Lemma**

*there exists* $u \in S$ *&* $v \notin S$ *s.t*

- $u, v > n$
- *for any* $m < n$*,* $u \in T + m$ *iff* $v \in T + m$

Let $\mathcal{A} = (A, A)$, $\mathcal{B} = (B, B)$ where $\#A = u$, $\#B = v$.
Then $\mathcal{A} \in K_S$ & $\mathcal{B} \notin K_S$.
We need to check II win in EF-game for FO+$K_T$ between $\mathcal{A}$ and $\mathcal{B}$.

In $i$-th move,

- If I choose FO-move and $a \in A$, II can choose $b \in B$ since $u, v > n$.
- If I choose $Q_{K_T}$-move and $X \subseteq A$,
    - if $X$ does not contain unchosen element, II choose $Y$ as set of correspondings (in this case $\#X = \#Y$).
    - if $X$ contains unchosen element, then $X$ contain all of such elements. II choose $Y$ as set of unchosen elements and correspondings in $X$.
      In this case, $\#X = u - m$ & $\#Y = v - m$ $(m < n)$,

  So any case, $\#X \in T$ *iff* $\#Y \in T$

□

How about ordered case? For example,

$$\mathcal{A} \in D_4 \Leftrightarrow \mathcal{A} \models Q_{D_2} x P(x) \wedge Q_{D_2} x (P(x) \wedge Q_{D_2} y (P(y) \wedge y \leq x))$$

So, FO($Q_{D_4}$) $\leq$ FO($Q_{D_2}$) on ordered.

**Theorem (Nurmonen(2000))**

*For $m, k > 0$, FO($Q_{D_{m^k}}$) $\leq$ FO($Q_{D_m}$) on ordered.*

**Corollary**

*For $m, m' > 0$,*
*FO($Q_{D_m}$) $\leq$ FO($Q_{D_{m'}}$) on ordered iff $\forall p$:prime, $p \mid m \Rightarrow p \mid m'$*

How about vectorized case?

$$\mathcal{A} \in K_S \Leftrightarrow \mathcal{A} \models \exists z_1 \exists z_2((z_1 \neq z_2) \wedge Q_{K^2_{2S}} xy((x = z_1 \vee x = z_2) \wedge P(y)))$$

$$\mathcal{A} \in K_S \Leftrightarrow \mathcal{A} \models Q_{K^2_{S^2}} xy(P(x) \wedge P(y))$$

where $2S := \{2n \mid n \in S\}$, $S^2 := \{n^2 \mid n \in S\}$.
So, FO+$K_S \leq$ FO+$K_{2S}$, FO+$K_{S^2}$

**Theorem**

*For $S, T \subseteq \mathbb{N}$,*
*FO+$K_S \leq$ FO+$K_T$*
*iff $\exists T' \in \mathcal{B}(\{f^{-1}(T) \mid f \in \mathbb{Z}[x]^+\})$ s.t $\#(S \Delta T') < \infty$*

$$f \in \mathbb{Z}[x]^+ \Leftrightarrow f = \sum_{k=0}^{n} a_k x^k \text{ where } a_k \in \mathbb{Z} \text{ \& } a_n > 0$$

- H.D. Ebbinghaus. J. Flum. “*Finite Model Theory*”. Springer. 1995
- G. Kolaitis. J. Väänänen. “Generalized quantifiers and pebble games on finite structures.” *Annals of Pure and Applied Logic* 74. pp. 23-75. 1995.
- L. J. Corredor. “El reticulo de las logicas de primer orden con cuantificadores cardinales identification.” *Revista Colombiana de Mathemáticas* 20. pp. 1-26. 1986.
- J. Nurmonen. “Counting modulo quantifiers on finite structures.” *Information and Computation* 160.1. pp. 62-87. 2000.
- K. Luosto. “On vectorizations of unary generalized quantifiers.” *Archive for Mathematical Logic* 51.3-4. pp. 241-255. 2012.