

# CS 228 : Logic in Computer Science

Krishna. S

# Quantifier Rank

---

Let  $\varphi$  be a FO formula over words. Define the **quantifier rank** of  $\varphi$  denoted  $c(\varphi)$

- ▶ If  $\varphi$  is atomic ( $x = y, x < y, S(x, y), Q_a(x)$ ) then  $c(\varphi) = 0$
- ▶  $c(\neg\varphi) = c(\varphi)$
- ▶  $c(\varphi \wedge \psi) = \max(c(\varphi), c(\psi))$
- ▶  $c(\exists x\varphi) = c(\varphi) + 1$
- ▶ Quantifier free formulae written in DNF :  $C_1 \vee C_2 \vee \dots \vee C_n$
- ▶ Formulae of quantifier rank  $k + 1$  written as a disjunction of the conjunction of formulae, each formula of the form  $\exists x\varphi, \neg\exists x\varphi$  or  $\varphi$ , with  $c(\varphi) \leq k$ . Eliminate repeated disjuncts/conjuncts.  
(Think Prenex normal form)

# Number of FO formulae of rank $\leq c$

---

Let  $\mathcal{V}$  be a finite set of first order variables, and let  $c \geq 0$ . There are finitely many FO formulae in normal form (therefore, upto equivalence) with rank  $\leq c$  over  $\mathcal{V}$ .

Proof ideas in the following slides.

# Number of FO formulae of rank $\leq c$

---

Let  $\mathcal{V}$  be a finite set of first order variables. Fix a finite signature  $\tau$ . Let there be  $m$  atomic formulae over  $\tau$  having variables from  $\mathcal{V}$ .

- ▶ If  $\mathcal{V}$  has 2 variables  $x, y$ , and  $\tau$  has  $Q_a, S, <$ .
- ▶ Atomic formulae :  $\{Q_a(x), Q_a(y), S(x, y), x < y\}$
- ▶ Let  
 $G = \{Q_a(x), \neg Q_a(x), Q_a(y), \neg Q_a(y), S(x, y), \neg S(x, y), x < y, \neg(x < y)\}$   
be the closure of all atomic formulae containing all formulae and their negations.
- ▶ Each subset of  $G$  is a possible conjunct  $C_i$ .
- ▶ All possible disjuncts using each  $C_i$  : formulae in DNF of rank 0

# Number of FO formulae of rank $\leq c$

---

Let  $\mathcal{V}$  be a finite set of first order variables. Fix a finite signature  $\tau$ . Let there be  $m$  atomic formulae over  $\tau$  having variables from  $\mathcal{V}$ .

- ▶  $2m$  atomic/negated atomic formulae
- ▶ Number of conjunctions  $C_i$  possible  $\leq 2^{2m}$
- ▶ Number of formulae in DNF  $\leq 2^{2^{2m}}$  ( $c = 0$ )

# Rank 1

---

Let there be  $p$  formulae  $\varphi$  of rank 0.

- ▶  $2p$  formulae of the form  $\exists x\varphi, \neg\exists x\varphi$
- ▶  $2^{2p}$  conjunctions of rank 1
- ▶ Conjoining any one of the  $p$  formulae of rank 0 gives all conjuncts of rank 1 :  $p2^{2p}$  more
- ▶ Possible conjuncts of rank 1 is  $q = (p + 1)2^{2p}$
- ▶ Possible disjuncts of these :  $2^q$

# Some Notation

---

Given a word  $w = a_1 \dots a_n$ , and a finite set of variables  $\mathcal{V}$ , define a  $\mathcal{V}$ -enriched-word with respect to  $w$  as

- ▶  $(a_1, U_1)(a_2, U_2) \dots (a_n, U_n)$  where
- ▶  $\bigcup_i U_i = \mathcal{V}$
- ▶  $U_i \cap U_j = \emptyset$

- ▶ A  $\mathcal{V}$ -enriched-word is over the alphabet  $\Sigma \times 2^{\mathcal{V}}$
- ▶  $(a, \{x\})(b, \{y, z\})(c, \emptyset)(d, \{u, v\})$  is a  $\{x, y, z, u, v\}$ -enriched word with respect to the word  $abcd$ .
- ▶ We will refer to  $\mathcal{V}$ -enriched-word structures as  $\mathcal{V}$ -structures from here on

# Notational Semantics

---

Given a  $\mathcal{V}$ -structure  $w = (a_1, S_1) \dots (a_n, S_n)$ ,

- ▶  $w \models Q_a(x)$  iff there exists  $j$  such that  $a_j = a$  and  $x \in S_j$ 
  - ▶  $(a, \{y\})(b, \{u, v\})(a, \{x\}) \models Q_a(x)$
- ▶  $w \models (x = y)$  iff there exists  $j$  such that  $x, y \in S_j$ 
  - ▶  $(a, \{x\})(b, \{y, z\})(c, \emptyset) \not\models (x = y)$
- ▶  $w \models x < y$  iff there exists  $i < j$  such that  $x \in S_i, y \in S_j$ 
  - ▶  $(a, \{x\})(b, \{y, z\})(c, \emptyset) \models x < y$
- ▶  $w \models \exists x Q_a(x)$  iff there exists  $i$  such that  $(a_1, S_1) \dots (a_i, S_i \cup \{x\}) \dots (a_n, S_n) \models Q_a(x)$ 
  - ▶  $(b, \{y, z\})(a, \{u\})(c, \emptyset) \models \exists x Q_a(x)$  since  
 $(b, \{y, z\})(a, \{x, u\})(c, \emptyset) \models Q_a(x)$



# Notational Semantics

- ▶  $(a, \emptyset)(a, \emptyset)(b, \emptyset) \models \forall x \exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])$  iff
- ▶  $(a, \emptyset)(a, \emptyset)(b, \emptyset) \models \neg \exists x \neg [\exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])]$  iff
- ▶  $(a, \emptyset)(a, \emptyset)(b, \emptyset) \not\models \exists x \neg [\exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])]$  iff
  - ▶  $(a, \{x\})(a, \emptyset)(b, \emptyset) \not\models \neg [\exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])]$  and
  - ▶  $(a, \emptyset)(a, \{x\})(b, \emptyset) \not\models \neg [\exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])]$  and
  - ▶  $(a, \emptyset)(a, \emptyset)(b, \{x\}) \not\models \neg [\exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])]$
- ▶  $(a, \{x\})(a, \emptyset)(b, \emptyset) \models \exists y (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])$  iff
- ▶  $(a, \{x\})(a, \emptyset)(b, \{y\}) \models (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])$

Similarly,  $(a, \emptyset)(a, \{x\})(b, \{y\}) \models (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])$  and  
 $(a, \emptyset)(a, \emptyset)(b, \{x, y\}) \models (Q_a(x) \rightarrow [(x < y) \wedge Q_b(y)])$

# Logical Equivalence

---

- ▶ Let  $w_1, w_2$  be two  $\mathcal{V}$ -structures and let  $r \geq 0$ .
- ▶ Write  $w_1 \sim_r w_2$  iff  $w_1, w_2$  satisfy the same set of FO formulae of rank  $\leq r$ .
- ▶  $(a, \emptyset)(b, \emptyset) \sim_0 (a, \emptyset)(b, \emptyset)(a, \emptyset)$
- ▶  $(a, \emptyset)(b, \emptyset) \approx_2 (a, \emptyset)(b, \emptyset)(a, \emptyset)$
- ▶  $(a, \emptyset)(b, \emptyset) \sim_1 (a, \emptyset)(b, \emptyset)(a, \emptyset)$ ?
- ▶  $\sim_r$  is an equivalence relation
- ▶ **Finitely** many equivalence classes : each class consists of words that behave the same way on formulae of rank  $\leq r$

## Non-Expressibility in FO : The Game Begins

# Come, Lets Play

---

- ▶ Given two  $\mathcal{V}$ -structures  $w_1, w_2$ , lets play a game on the pair of words  $w_1, w_2$
- ▶ There are 2 players : Spoiler and Duplicator
- ▶ Play for  $r$ -rounds,  $r \geq 0$
- ▶ Spoiler wants to show that  $w_1, w_2$  are different ( $w_1 \not\sim_r w_2$ )
- ▶ Duplicator wants to show that they are same ( $w_1 \sim_r w_2$ )
- ▶ Each player has  $r$  pebbles  $z_1, \dots, z_r$

# Moves of the Game

---

- ▶ At the start of each round, spoiler chooses a structure.
- ▶ Duplicator gets the other structure
- ▶ Spoiler places his pebble say  $z_i$  on one of the positions of his chosen word
- ▶ Duplicator must keep the pebble  $z_i$  on one of the positions of her word
- ▶ A pebble once placed, cannot be removed
- ▶ The game ends after  $r$  rounds, when both players have used all their pebbles

# A Play

---

- ▶  $w_1 = (a, \emptyset)(b, \emptyset)$  and  $w_2 = (a, \emptyset)(b, \emptyset)(a, \emptyset)$
- ▶ 2 rounds, so 2 pebbles :  $z_1, z_2$
- ▶ Spoiler picks  $w_2$ , duplicator picks  $w_1$
- ▶ Round 1:
  - ▶ Spoiler :  $(a, \{z_1\})(b, \emptyset)(a, \emptyset)$
  - ▶ Duplicator :  $(a, \{z_1\})(b, \emptyset)$
  - ▶ After round 1, we have two  $\{z_1\}$  structures  $(w'_1, w'_2)$
- ▶ Round 2:
  - ▶ Spoiler continues on the structure  $w'_2$
  - ▶ Duplicator gets  $w'_1$  to play
  - ▶ Spoiler :  $(a, \{z_1\})(b, \emptyset)(a, \{z_2\})$
  - ▶ Duplicator :  $(a, \{z_1, z_2\})(b, \emptyset)$  or  $(a, \{z_1\})(b, \{z_2\})$

# Winner

---

- ▶ Start with two  $\emptyset$  structures  $(w_1, w_2)$
- ▶  $r$ -round game, pebble set  $\mathcal{V} = \{z_1, \dots, z_r\}$
- ▶ Each round changes the structures
- ▶ At the end of  $r$ -rounds, we have two  $\mathcal{V}$ -structures  $(w'_1, w'_2)$
- ▶ Duplicator wins iff for every atomic formula  $\alpha$ ,  
 $w'_1 \models \alpha$  iff  $w'_2 \models \alpha$
- ▶ That is,  $w'_1 \sim_0 w'_2$
- ▶ Spoiler wins otherwise.

# Winner

---

Given two word structures  $(w_1, w_2)$ , duplicator wins on  $(w_1, w_2)$  if for every atomic formula  $\alpha$ ,  $w_1 \models \alpha$  iff  $w_2 \models \alpha$



# Play continues

---

- ▶ Who won in the earlier play?
- ▶ We had
  - ▶  $(a, \{z_1\})(b, \emptyset)(a, \{z_2\})$  and  $(a, \{z_1, z_2\})(b, \emptyset)$
  - ▶  $(a, \{z_1\})(b, \emptyset)(a, \{z_2\}) \models (z_1 < z_2)$
  - ▶  $(a, \{z_1, z_2\})(b, \emptyset) \not\models (z_1 < z_2)$  or
  - ▶  $(a, \{z_1\})(b, \emptyset)(a, \{z_2\})$  and  $(a, \{z_1\})(b, \{z_2\})$
  - ▶  $(a, \{z_1\})(b, \emptyset)(a, \{z_2\}) \models Q_a(z_2)$
  - ▶  $(a, \{z_1\})(b, \{z_2\}) \not\models Q_a(z_2)$
- ▶ Spoiler wins in two rounds
- ▶ If the game was played only for one round, who will win?

# Unique Winner

---

Given structures  $w_1, w_2$ , and a number of rounds  $r$ , exactly one of the players win.

- ▶ Think of a tree of all possible sequences of plays.
- ▶ The root is  $(w_1, w_2)$ , and nodes are all possible pairs  $(w'_1, w'_2)$  of obtainable structures.
- ▶ Mark a leaf node as S or D depending on whether Spoiler won or Duplicator won.
- ▶ An interior node corresponding to a move of Duplicator is labeled D if there any children labeled D; likewise for Spoiler

# Logical Equivalence and Winning

---

Let  $w_1, w_2$  be  $\mathcal{V}$ -structures and let  $r \geq 0$ . Then  $w_1 \sim_r w_2$  iff Duplicator has a winning strategy in the  $r$ -round game on  $(w_1, w_2)$ .

# Logical Equivalence and Winning

---

Assume  $w_1 \sim_r w_2$ , and induct on  $r$

- ▶ Base :  $r = 0$  and  $w_1 \sim_0 w_2$ . Duplicator wins, since by assumption,  $w_1, w_2$  agree on all atomic formulae.
- ▶ Assume for  $r - 1$  :  $w_1 \sim_{r-1} w_2 \Rightarrow$  Duplicator has a winning strategy in a  $r - 1$  round game

# Logical Equivalence and Winning

---

- ▶ Now, let  $w_1 \sim_r w_2$ , and assume spoiler wins the  $r$ -round game on  $(w_1, w_2)$ .
  - ▶ Assume spoiler starts on  $w_1$ , places a pebble  $z_1$  somewhere on  $w_1$
  - ▶ The resultant structure is  $w'_1$
  - ▶ In response, duplicator places her pebble somewhere on  $w_2$
  - ▶ The resultant structure is  $w'_2$
  - ▶ By assumption, spoiler wins the  $r - 1$  round game on  $(w'_1, w'_2)$
  - ▶ By inductive hypothesis,  $w'_1 \sim_{r-1} w'_2$
  - ▶ Let  $\psi$  be the conjunction of all formulae of rank  $r - 1$  in normal form that are satisfied by  $w'_1$
  - ▶ Then  $w'_1 \models \psi$ ,  $w'_2 \not\models \psi$
  - ▶ We thus have

$$w_1 \models \exists z_1 \psi, w_2 \not\models \exists z_1 \psi$$

contradicting  $w_1 \sim_r w_2$

# Logical Equivalence and Winning : Converse

---

Assume Duplicator wins  $r$ -round games on  $(w_1, w_2)$  and induct on  $r$

- ▶ Base :  $r = 0$  and Duplicator wins. Then  $w_1, w_2$  agree on all atomic formulae, and hence  $w_1 \sim_0 w_2$
- ▶ Assume for  $r - 1$  : Duplicator has a winning strategy in a  $r - 1$  round game  $\Rightarrow w_1 \sim_{r-1} w_2$

# Logical Equivalence and Winning : Converse

---

- ▶ Now, let duplicator win in the  $r$  round game, but  $w_1 \approx_r w_2$ .
  - ▶  $w_1 \approx_r w_2 \Rightarrow$  there is some formula  $\psi$ ,  $c(\psi) = r$  such that  $w_1 \models \psi$ ,  $w_2 \not\models \psi$
  - ▶ Assume  $\psi = \exists z_1 \varphi$ . Then  $c(\varphi) = r - 1$
  - ▶ Since  $w_1 \models \exists z_1 \varphi$ , spoiler can keep pebble  $z_1$  somewhere in  $w_1$  obtaining  $w'_1$  satisfying  $\varphi$
  - ▶ In reply, duplicator keeps pebble  $z_1$  on  $w_2$  obtaining  $w'_2$
  - ▶ By assumption,  $w'_2 \not\models \varphi$
  - ▶ Also, by assumption, duplicator wins the  $r - 1$  round game on  $(w'_1, w'_2)$  : this by inductive hypothesis says that  $w'_1 \sim_{r-1} w'_2$
  - ▶ That is, either both  $w'_1, w'_2$  satisfy  $\varphi$ , or both don't, a contradiction.

# FO-definable languages

---

Assume  $L$  is FO-definable, and  $L = L(\varphi)$  with rank of  $\varphi$  being  $k$ .

- ▶ Let  $L = \{v_1, v_2, v_3, \dots\}$  and  $\bar{L} = \{w_1, w_2, w_3, \dots\}$
- ▶ Play a  $k$  round game on  $v_i \in L$  and  $w_j \notin L$ . Let  $\psi_{v_i, w_j}$  be the formula of rank  $k$  that distinguishes the two words.
- ▶ Consider the formula

$$[\psi_{v_1, w_1} \wedge \psi_{v_1, w_2} \wedge \dots \wedge \psi_{v_1, w_n} \wedge \dots]$$

$$\vee$$

$$[\psi_{v_2, w_1} \wedge \psi_{v_2, w_2} \wedge \dots \wedge \psi_{v_2, w_n} \wedge \dots]$$

$$\vee$$
$$\vdots$$



# FO-definable languages

---

$$\psi_L = \bigvee_{v \in L} \bigwedge_{w \notin L} \psi_{vw}$$

- ▶ Each  $\psi_{vw}$  has rank at most  $k$
- ▶ Up to equivalence, there are finitely many formulae of rank  $k$
- ▶ Hence the disjunction and conjunction are finite
- ▶  $\psi_L$  is a proper formula (of finite size)
- ▶  $\psi_L$  captures  $L$  since each  $v \in L$  satisfies  $\bigwedge_{w \notin L} \psi_{vw}$  while none of the  $w \notin L$  satisfy  $\bigwedge_{w \notin L} \psi_{vw}$

# FO-definable languages

---

Given a property  $\mathcal{K}$ , if for any pair  $v \in \mathcal{K}$  and  $w \notin \mathcal{K}$ , spoiler has a winning strategy in the  $k$ -round EF game on  $v$  and  $w$ , then there is a rank  $k$  FO formula  $\varphi_{\mathcal{K}}$  that defines the property  $\mathcal{K}$ .

$$\varphi_{\mathcal{K}} = \bigvee_{v \in \mathcal{K}} \bigwedge_{w \notin \mathcal{K}} \psi_{vw}$$

where  $\psi_{vw}$  is as explained in the previous slide.

- Note that  $k$  is fixed in the above, and is independent of the choices of the words.

# Implications of the Game on FO definability

---

## FO Definability

$L$  is FO definable  $\Rightarrow$  there exists an  $r$  such that for every  $(w_1, w_2)$  pair, such that  $w_1 \in L$ ,  $w_2 \notin L$ , spoiler wins in  $r$  rounds

## Non FO Definability

For all  $r \geq 0$ , there exists a  $(w_1, w_2)$  pair with  $w_1 \notin L$ ,  $w_2 \in L$ , duplicator wins in  $r$  rounds  $\Rightarrow L$  is not FO definable

# $(aa)^*$ is not $FO[<]$ Definable

---

- ▶ Assume that there is a sentence  $\varphi$  that defines words of even length, with  $c(\varphi) = r$ .
- ▶ Then,  $a^i \models \varphi$  iff  $i$  is even
- ▶ Show that for all  $r > 0$ ,  $a^{2^r} \sim_r a^{2^r-1}$

# $(aa)^*$ is not $FO[<]$ Definable

---

- ▶ Base case :  $(a, \emptyset)(a, \emptyset)$  and  $(a, \emptyset)$  for  $r = 1$
- ▶ In one round, duplicator wins on  $(a, \emptyset)(a, \emptyset)$  and  $(a, \emptyset)$
- ▶ Consider  $(aaaa, aaa)$  for  $r = 3$ . Who wins?
- ▶ Consider  $(aaaa, aaa)$  for  $r = 2$ . Who wins?

# $(aa)^*$ is not $FO[<]$ Definable

---

- ▶ Show that for all  $k \geq 2^r - 1$ , duplicator has a winning strategy for the  $r$ -round game in  $(a^k, a^{k+1})$ , for all  $r \geq 0$
- ▶ Induct on  $r$
- ▶ If  $r = 1$ , then on  $(a, aa)$  duplicator wins in one round
- ▶ Assume now that the claim is true for  $\leq r - 1$

# $(aa)^*$ is not $FO[<]$ Definable

---

- ▶ Let  $k \geq 2^r - 1$ , and consider the structures

$$(a^k, a^{k+1})$$

- ▶ Spoiler puts pebble  $z_1$  in one of the words obtaining

$$(a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^t$$

- ▶  $s \leq \frac{k-1}{2}$  or  $t \leq \frac{k-1}{2}$

# $(aa)^*$ is not $FO[<]$ Definable

---

- ▶ Assume  $s \leq \frac{k-1}{2}$ . Duplicator puts her pebble  $z_1$  on the  $(s+1)$ th letter of the other word obtaining

$$(a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^{t'}$$

where  $t' = t + 1$  or  $t' = t - 1$ .

- ▶ The structures after round 1 are thus

$$(a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^t, (a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^{t'}$$

- ▶ We have  $2^r - 1 \leq k = \min(t, t') + s + 1 \leq \min(t, t') + \frac{k-1}{2} + 1$
- ▶ Hence  $\min(t, t') \geq \frac{k-1}{2} \geq 2^{r-1} - 1$
- ▶ By inductive hypothesis, duplicator has a winning strategy for the  $r - 1$  round game on  $(a^t, a^{t'})$ .



# Duplicator's Win

---

- ▶ Use the duplicator's winning strategy for the  $r - 1$  round game on  $(a^t, a^{t'})$ , to obtain a winning strategy in  $r - 1$  rounds on

$$(a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^t, (a, \emptyset)^s(a, \{z_1\})(a, \emptyset)^{t'}$$

- ▶ Whenever spoiler plays on a structure on letter  $i \leq s + 1$ , duplicator plays on the same position on the other structure
- ▶ When spoiler plays at a position  $i > s + 1$  in either word, duplicator plays in the part of the other word  $> s + 1$  using her winning strategy in  $(a^t, a^{t'})$

# Duplicator's Win

---

- ▶ At the end of  $r$  rounds, we have structures  $w'_1, w'_2$ .
- ▶ For  $i \leq s + 1$ , pebble  $z_j$  appears at position  $i$  of  $w'_1$  iff pebble  $z_j$  appears at position  $i$  of  $w'_2$
- ▶ Let's erase the first  $s + 1$  letters in  $w'_1, w'_2$ , obtaining  $v'_1, v'_2$
- ▶  $v'_1, v'_2$  are the words that result after  $r' \leq r - 1$  rounds of play on  $(a^t, a^{t'})$ . Recall that duplicator won this.
- ▶ Show that  $w'_1, w'_2$  satisfy the same atomic formulae

# Duplicator's Win

---

- ▶ Atomic Formulae :  $Q_a(z_j)$  : Both  $w'_1, w'_2$  satisfy this.
- ▶  $w'_1 \models z_i < z_j$ . If  $z_i, z_j$  are in the first  $s + 1$  letters, then  $w'_2 \models z_i < z_j$ .
- ▶ If  $z_i, z_j$  occur in the last  $|w'_1| - s - 1$  positions, then  $v'_1 \models z_i < z_j$ .  
By duplicator's win in  $(a^t, a^{t'})$ ,  $v'_2 \models z_i < z_j$
- ▶ If  $z_i$  appears among the first  $s + 1$  letters and  $z_j$  after the first  $s + 1$  letters of  $w'_1$ , same is true in  $w'_2$ .

# Historically Speaking

---

The games that we saw are due to Ehrenfeucht and Fraïssé

Reference: Finite Automata, Formal Logic and Circuit Complexity, by Howard Straubing.