Recapitulation:

Ehrenfeucht - Fraïssé Theorem

Duplicator wins $\underbrace{G_k((S_v,\vec{s}),(S_w,\vec{t}))}_{\text{Duplicator can respond } k\text{-rounds to choices of Spoiler}}$ iff $\underbrace{(S_v,\vec{s}) \equiv_{k,n} (S_w,\vec{t})}_{\text{cannot be distinguished by FO[<]-formulas of qd} \leq k}$.

Application:

$(aa)^*$ is <u>not</u> FO[<]-definable

=> This reasons over <u>all</u> FO[<]-formulas.

<u>Proof:</u>

If $(aa)^* = L(\varphi)$ for some φ,

then φ has $qd(\varphi) = k$ for some $k \in \mathbb{N}$.

Choose words large enough to win game:

$G_k(a^{2^k}, a^{2^k+1})$ won by duplicator.

Thus, by Ehrenfeucht - Fraïssé Theorem

a^{2^k} and a^{2^k+1} cannot be distinguished by formula of qd $\leq k$.

Thus, cannot be distinguished by φ. ↯ □

How to prove the existence of such a winning strategy?

<u>Lemma:</u>

Duplicator wins $G_k(a^{2^k}, a^{2^k+1})$.

<u>Proof:</u>

Establish stronger result:

For all k, duplicator wins $G_k(a^i, a^j)$ with $i,j \geq 2^k - 1$.

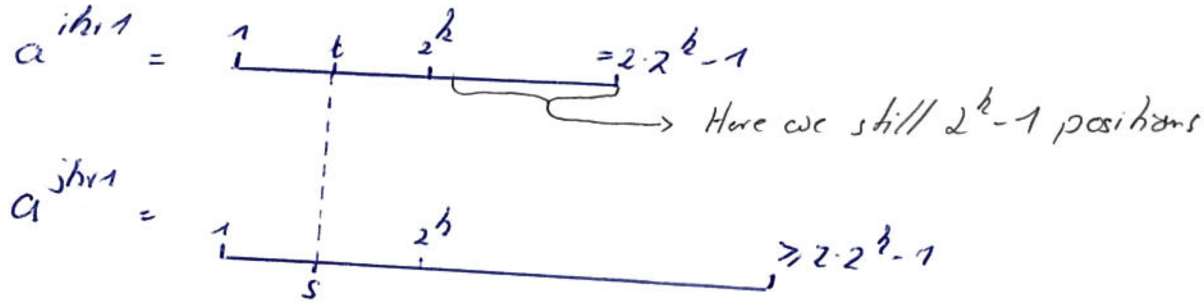
Proceed by induction on k:

<u>IB:</u> <u>$k=1$</u> Duplicator wins $G_1(a, a^n)$ with $n \geq 1$.

IS: Assume duplicator has a winning strategy for $G_k(a^{i_k}, a^{j_k})$ with $i_k, j_k \geq 2^k - 1$.
Consider

$$G_{k+1}(a^{i_{k+1}}, a^{j_{k+1}}) \text{ with } i_{k+1} = 2^{(k+1)} - 1 = 2 \cdot 2^k - 1$$
$$j_{k+1} \geq 2^{(k+1)} - 1$$

Then



Case 1:

Spoiler pics $s \leq 2^k$ in $a^{j_{k+1}}$.
Let duplicator pic $t = s \leq 2^k$ in $a^{i_{k+1}}$
Then

$G_k(a^{s-1}, a^{s-1})$ is won by duplicator as the words coincide.

Moreover,

$G_k(a^{i_{k+1}-s}, a^{j_{k+1}-s})$ won by duplicator by the induction hypothesis.

To see this,

$$\underset{(\text{Def. } i_{k+1})}{i_{k+1} - s = 2 \cdot 2^k - 1 - s} \underset{(\text{Def. } s)}{\geq} 2 \cdot 2^k - 1 - 2^k = 2^k - 1.$$

Case 2:

Spoiler pics $s > 2^k$. Similar.

$\square$

Proof (of Ehrenfeucht-Fraïssé theorem):
"=>" By induction on $k \in \mathbb{N}$.

IA: $k=0$ Duplicator wins 0 rounds iff $\vec{s} \mapsto \vec{t}$ is partial isomorphism.
This means for every atomic formula

$\varphi = P_a(x)$ and $\varphi = x < y$

we have

$$S_v, I[\vec{s}/\vec{x}] \models \varphi \text{ iff } S_w, I[\vec{t}/\vec{x}] \models \varphi.$$

By induction on structure of formulas, this carries over to arbitrary quantifier-free formulas.

<u>IS:</u> Assume a win of duplicator on $G_k((S_v, \vec{s'}), (S_w, \vec{t'}))$ entails $(S_v, \vec{s'}) \equiv_{k,n'} (S_w, \vec{t'})$.

Consider win of duplicator on $G_{k+1}((S_v, \vec{s}), (S_w, \vec{t}))$.

Show that

$$S_v, I[\vec{s}/\vec{x}] \models \varphi \text{ iff } S_w, I[\vec{t}/\vec{x}] \models \varphi$$

where φ is of $qd(\varphi) = k+1$.

<u>Case $\exists x: \psi$</u> (where $qd(\psi) = k$):

If $S_v, I[\vec{s}/\vec{x}] \models \exists x: \psi$, then there is $s \in D_v$ so that

$$S_v, I[s.\vec{s}/x.\vec{x}] \models \psi.$$

Let spoiler start with s.

By assumption that duplicator wins $G_{k+1}((S_v, \vec{s}), (S_w, \vec{t}))$, can reply by position $t \in D_w$.

The resulting game

$$G_k((S_v, s.\vec{s}), (S_w, t.\vec{t}))$$

is won by duplicator.

By the induction hypothesis ($\vec{s'} = s.\vec{s}$, $\vec{t'} = t.\vec{t}$),

$$(S_v, s.\vec{s}) \equiv_{k,n+1} (S_w, t.\vec{t})$$

As ψ has $qd(\psi) = k$, we conclude

$$S_w, I[t.\vec{t}/x.\vec{x}] \models \psi.$$

This means, there is $t \in D_w$:

$$S_w, I[\vec{t}/\vec{x}] \models \exists x: \psi.$$

Remaining cases again by induction on structure of formulas of $qd = k+1$.

"$\Leftarrow$" Let $(S_v, \vec{s}) \equiv_{k,n} (S_w, \vec{t})$.

Construct a formula $\varphi^k_{v,\vec{s}}$ of qd k that characterises win of $G_k((S_v, \vec{s}), -)$ from duplicators point of view:

Let $\vec{x} = (x_1, \dots, x_n)$

(Δ) iff Duplicator wins $G_k((S_v, \vec{s}), (S_u, \vec{r}))$

$$S_u, I[\vec{r}/\vec{x}] \models \varphi^k_{v,\vec{s}},$$

for arbitrary $(S_u, \vec{r})$.

Assume we constructed this formula (see below how it works).

Then duplicator wins

$$G_k((S_v, \vec{s}), (S_v, \vec{s}))$$

by copying spoiler moves.

By construction of $\varphi^k_{v,\vec{s}}$ (Δ), this means

$$S_v, I[\vec{s}/\vec{x}] \models \varphi^k_{v,\vec{s}}.$$

We assume that $(S_v, \vec{s}) \equiv_{k,n} (S_w, \vec{t})$.

By definition of k-equivalence and the fact that

$$qd(\varphi^k_{v,\vec{s}}) = k,$$

we get

$$S_w, I[\vec{t}/\vec{x}] \models \varphi^k_{v,\vec{s}}.$$

Again with equivalence (Δ) for $\varphi^k_{v,\vec{s}}$ we get

$$G_k((S_v, \vec{s}), (S_w, \vec{t}))$$

as required.

<u>Key tool in construction of $\varphi^k_{v,\vec{s}}$</u>:

Lemma on winning situations for duplicator.

Construction of $\varphi^k_{v,\vec{s}}$ by induction on k:

IA:
$k = 0$ $\quad \varphi^0_{v,\vec{s}} := \bigwedge_{\substack{s_i \in \vec{s} \\ v(s_i) = a}} P_a(x_i) \wedge \bigwedge_{\substack{s_i, s_j \in \vec{s} \\ s_i < s_j}} x_i < x_j$

IS: Assume we already constructed $\varphi^k_{v,\vec{s'}}$ of $qd = k$ for $G_k((S_v, \vec{s'}), -)$.

Then for $G_{k+1}((S_v, \vec{s}), -)$ we set

$$\varphi^{k+1}_{v,\vec{s}} := \underbrace{\bigwedge_{s \in D_v} \exists x : \varphi^k_{v,s.\vec{s}}(x, \vec{x})}_{(a)} \wedge \underbrace{\forall x : \bigvee_{s \in D_v} \varphi^k_{v,s.\vec{s}}(x, \vec{x})}_{(b)}.$$

By winning lemma, this yields a winning strategy for duplicator

(a) If spoiler selects from v, then there is a position x in u with which duplicator can reply (2a in lemma)

(b) Whatever spoiler chooses from u, there is a position s in v with which duplicator can respond (2b in lemma)

Quantifier depth is $k+1$.

Note:

- Selects in v by Boolean connectives
- Selects in u by quantification.

3.2 Closure properties

- Find "subclass" of regular languages that characterises FO[<]-definable languages
- Algebraic characterisation (as opposed to logical)
  ↳ Focus on closure properties.

Definition (Star-free languages):
Let Σ an alphabet. The class of star-free languages (over Σ), denoted by SF_Σ, is the smallest class so that

(1) $\emptyset \in SF_\Sigma$, $\{\varepsilon\} \in SF_\Sigma$, $\{a\} \in SF_\Sigma$ f.a. $a \in \Sigma$.

(2) If $L_1, L_2 \in SF_\Sigma$ then $L_1 . L_2$, $L_1 \cup L_2$, $\overline{L_1} \in SF_\Sigma$.

Example:

Star-free does not mean that every description of the language avoids Kleene-star:

(a) Σ^* is star-free by $\Sigma^* = \overline{\emptyset}$.

(b) If L_1, L_2 are star-free, so are $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$ and $L_1 \setminus L_2 = L_1 \cap \overline{L_2}$

(c) Let $D \subseteq \Sigma$. Then D^* is star-free by

$$D^* = \Sigma^* \setminus (\Sigma^* . \underbrace{(\Sigma \setminus D)}_{\text{There is a non-}D\text{ letter.}} . \Sigma^*)$$

(d) Language $(a.b)^*$ is star-free as

$$(a.b)^* = \Sigma^* \setminus \underbrace{b.\Sigma^*}_{\substack{\text{Remove words}\\ \text{starting with } b}} \setminus \underbrace{\Sigma^*.aa\Sigma^*}_{\substack{\text{Remove words}\\ \text{with consecutive}\\ a\text{s}}} \setminus \Sigma^*.bb\Sigma^* \setminus \Sigma^*.a$$

<u>Goal:</u> Show that star-free languages are precisely the FO[<]-definable languages.

REG = WMSO-definable

$\boxed{\text{SF = FO[<]-definable}}$

$(aa)^*$