Monadic Quantifiers Reognized by Deterministic Pushdown Automata

Makoto Kanazawa

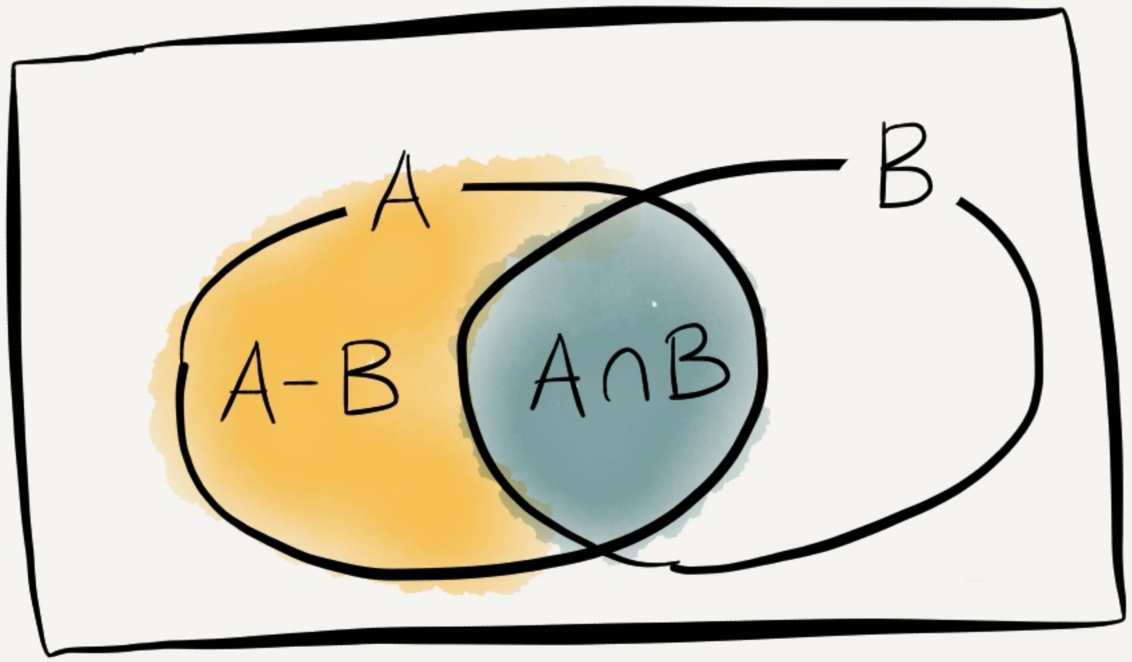
NII, Tokyo, Japan

- Mathematical Linguistics
  - Generalized Quantifiers
    - Semantic Automata

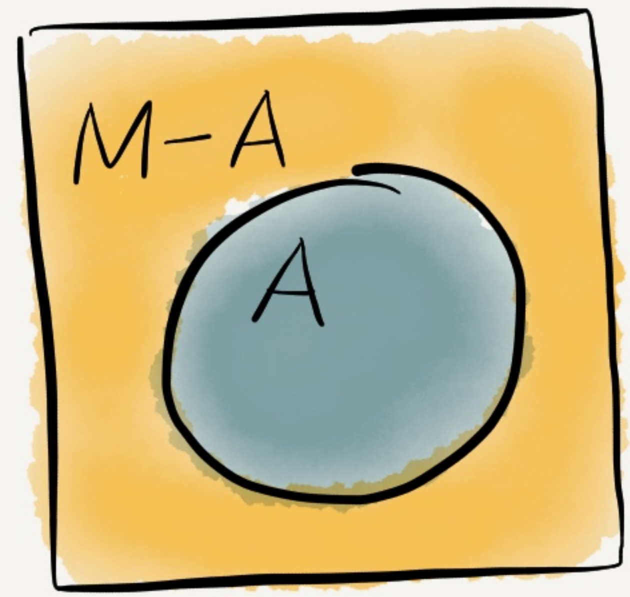
Classification of quantifiers in terms of automata recognizing them

Monadic Quantifiers

Type $\langle 1,1\rangle$

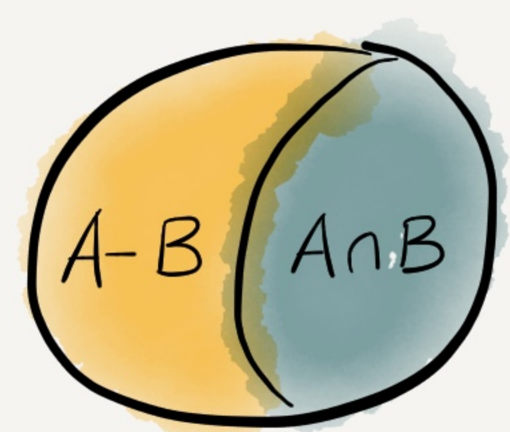


Type $\langle 1\rangle$



CONS $\quad Q_M AB \Leftrightarrow Q_M A(A\cap B)$

EXT $\quad Q_M AB \Leftrightarrow Q_{A\cup B} AB$



Q: type $\langle 1,1 \rangle$, CONS, EXT ISOM

$$V_Q = \{ (|A-B|, |A \cap B|) \mid QAB \}$$

$$\subseteq \mathbb{N}^2$$

$$W_Q = \{ w \in \{a,b\}^* \mid (\#_a(w), \#_b(w)) \in V_Q \}$$

commutative subset of $\{a,b\}^*$

$$\text{every } A\, B \iff A \subseteq B$$

$$V_{every} = \{0\} \times \mathbb{N}$$

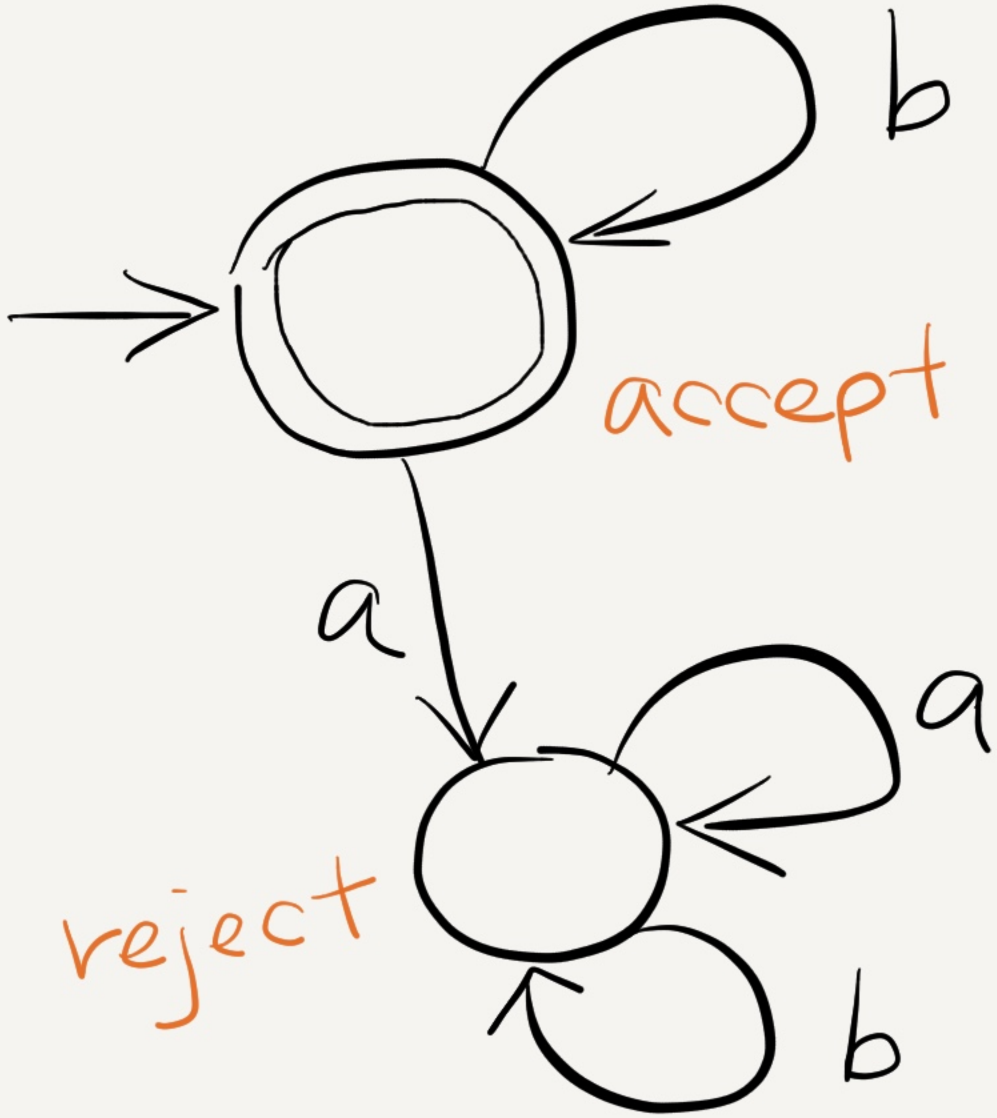
$$W_{every} = b^*$$

$$\text{most } A\, B \iff |A - B| < |A \cap B|$$

$$V_{most} = \{(x, y) \in \mathbb{N}^2 \mid x < y\}$$

$$W_{most} = \{w \in \{a, b\}^* \mid \#_a(w) < \#_b(w)\}$$

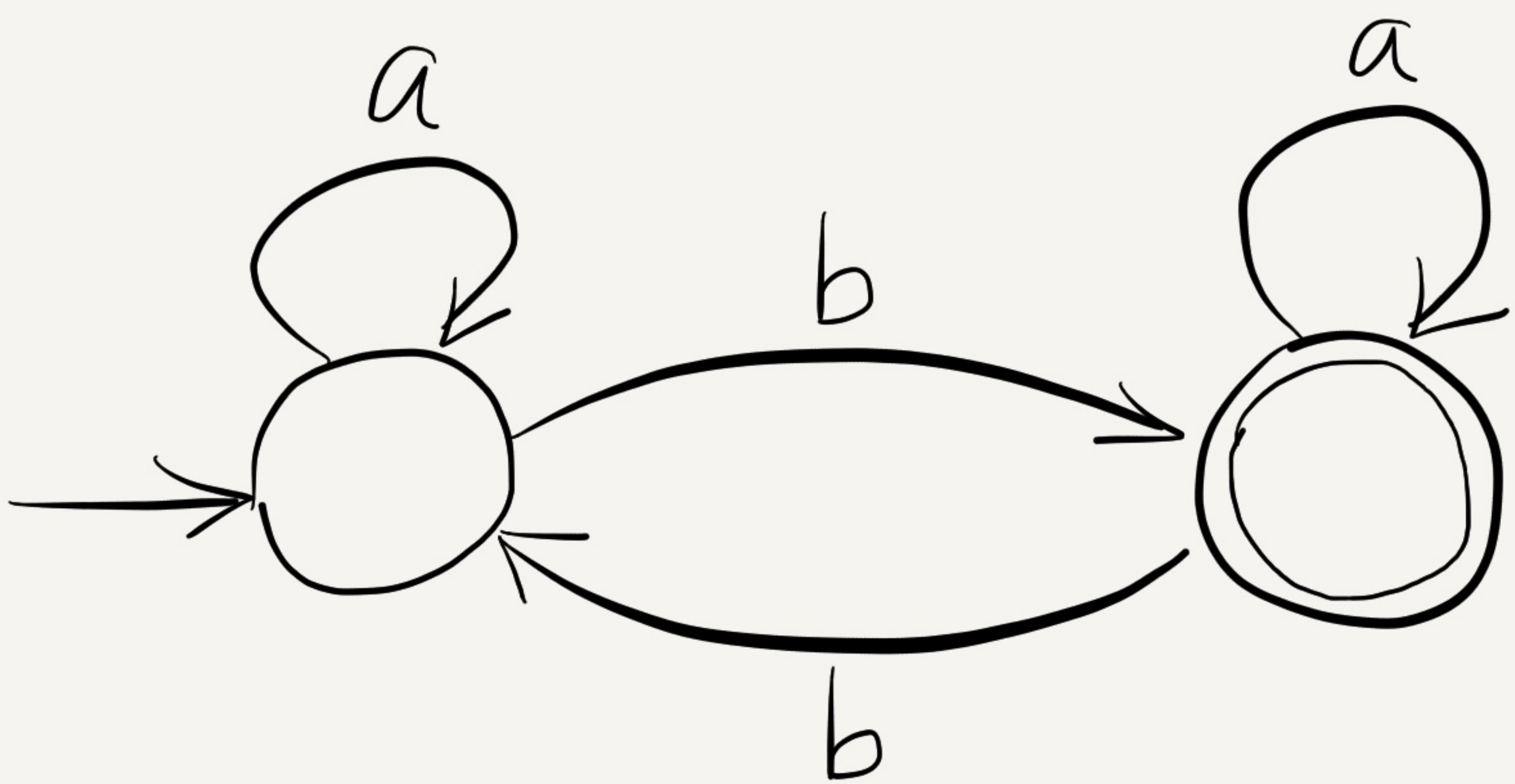
Automaton for every



DFA

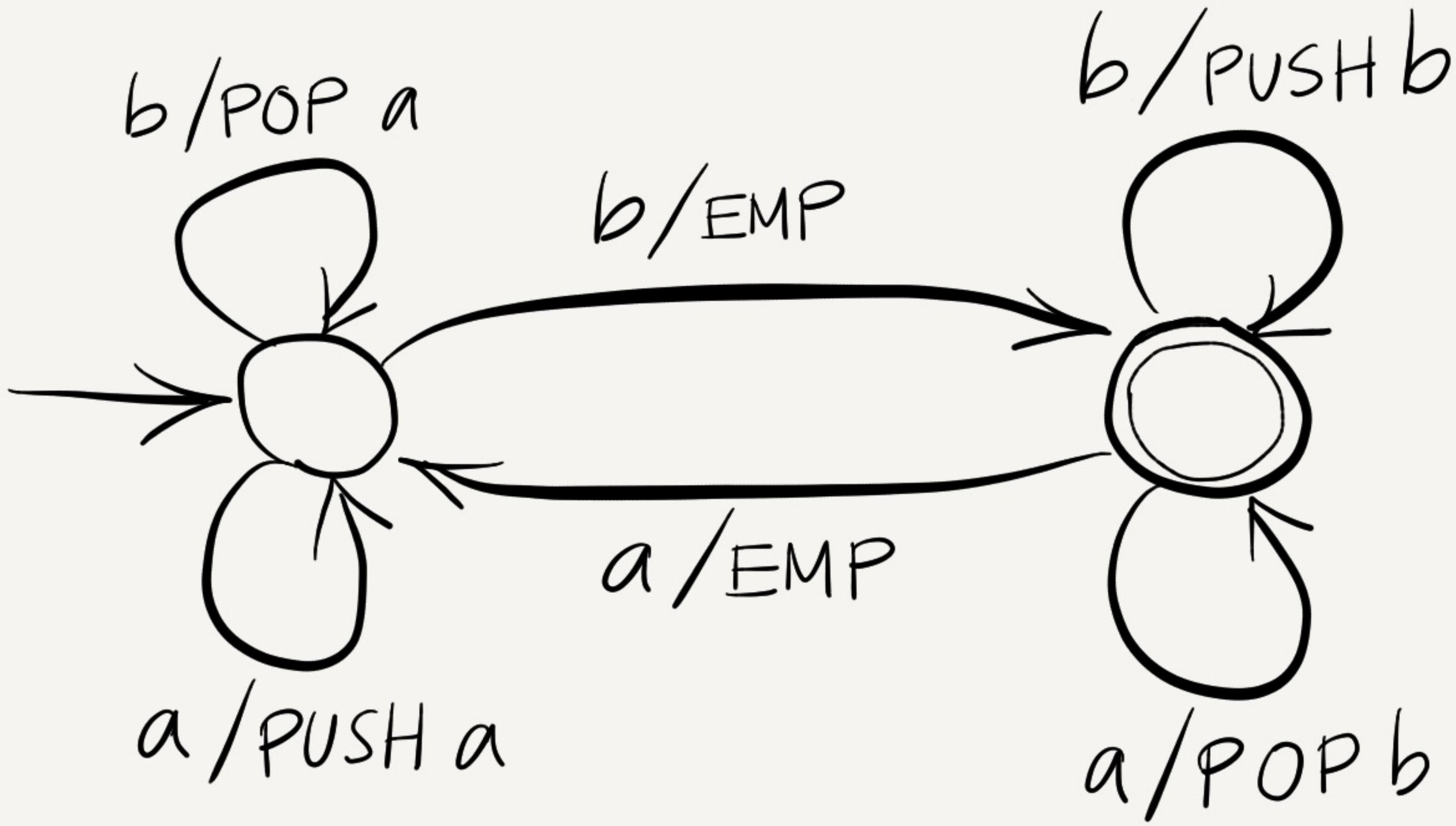
an odd number of

$|A \cap B|$ is odd

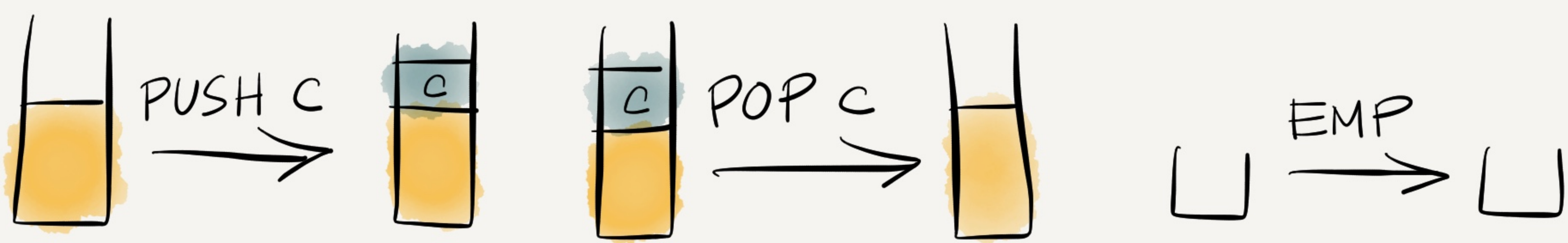


DFA

most



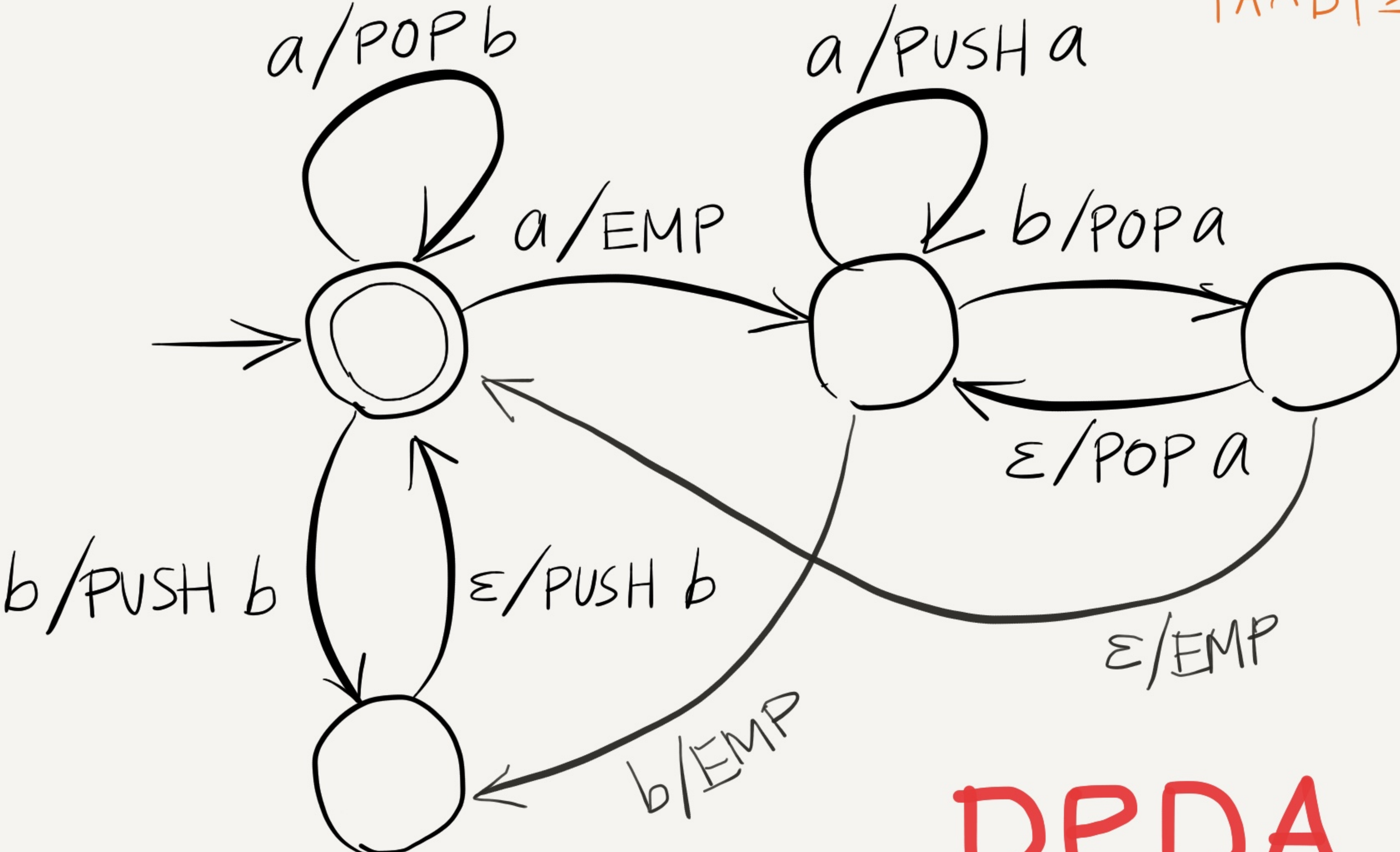
DPDA



at least one third

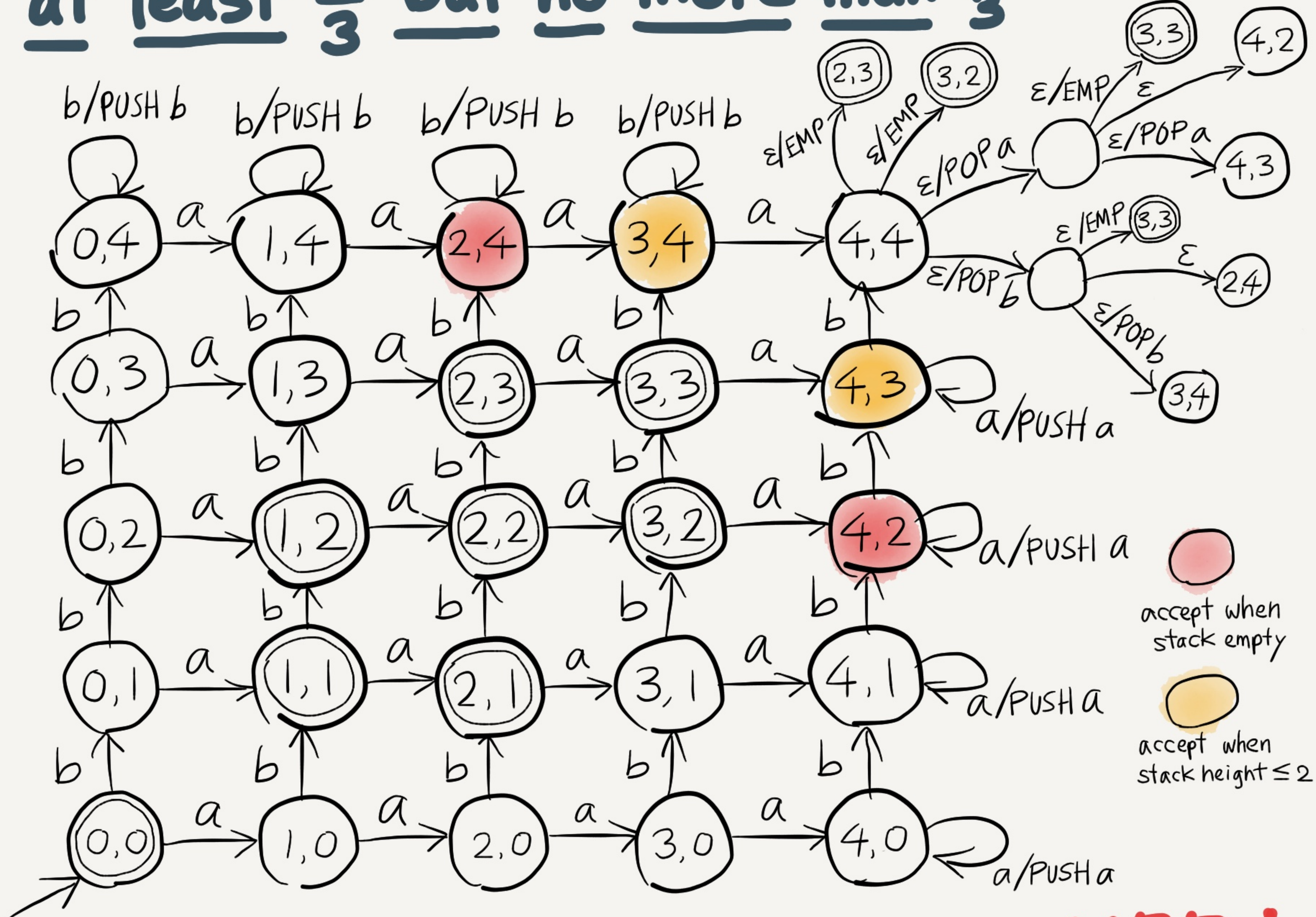
$|A \cap B| \geq \frac{1}{3}|A|$

$|A \cap B| \geq \frac{1}{2}|A-B|$



DPDA

at least $\frac{1}{3}$ but no more than $\frac{2}{3}$



NPDA

W_Q accepted by	V_Q	Q definable in
DFA $\equiv$ NFA		$FO(D_n)$
DPDA		
NPDA	semilinear	$MSO(I)$

Theorem (van Benthem 1986)

W_Q is accepted by an NPDA

$\iff$ V_Q is semilinear

$S \subseteq \mathbb{N}^\ell$ is **semilinear** $\overset{def}{\iff}$

S is a finite union of linear sets

$S \subseteq \mathbb{N}^\ell$ is **linear** $\overset{\text{def}}{\iff}$

$$S = L(\vec{u}; \{\vec{v}_1, \dots, \vec{v}_n\})$$

for some $\vec{u}, \vec{v}_1, \dots, \vec{v}_n \in \mathbb{N}^\ell$

generators

$$L(\vec{u}; \{\vec{v}_1, \dots, \vec{v}_n\}) \overset{\text{def}}{=}$$

$$\{\vec{u} + k_1\vec{v}_1 + \cdots + k_n\vec{v}_n \mid k_i \in \mathbb{N}\}$$

offset

Write $L(O; G)$ for $\bigcup_{\vec{u} \in O} L(\vec{u}; G)$

$S \subseteq \mathbb{N}^\ell$ is semilinear $\Longleftrightarrow$ Ginsburg & Spanier 1966

S is definable in <u>Presburger arithmetic</u>

$$\{(x,y)\in\mathbb{N}^2 \mid \tfrac{1}{3}(x+y) \le y \le \tfrac{2}{3}(x+y)\}$$

$$= \{(x,y)\in\mathbb{N}^2 \mid \tfrac{1}{2}x \le y \le 2x\}$$

$$= L(\{(0,0),(1,1),(2,2)\};\{(1,2),(2,1)\})$$

W_Q accepted by	V_Q	Q definable in
DFA $\equiv$ NFA	?	$FO(D_n)$
DPDA	?	
NPDA	semilinear	$MSO(I)$

cf. Mostowski 1998

Theorem W_Q is accepted by a DFA

$$\Updownarrow$$

$$
\begin{aligned}
V_Q = {} & L(O_1; \emptyset) \cup \\
& L(O_2; \{(k,0)\}) \cup \\
& L(O_3; \{(0,\ell)\}) \cup \\
& L(O_4; \{(k,0),(0,\ell)\})
\end{aligned}
$$

for some $k, \ell \in \mathbb{N}$ and

finite sets $O_1, O_2, O_3, O_4 \subseteq \mathbb{N}^2$

Lemma $W \subseteq \Sigma^*$ regular $\Rightarrow$

$$\exists p \in \mathbb{N}\ \forall z \in \Sigma^* \big(|z| \geq p \Rightarrow$$

$$\exists x_1 x_2 x_3 \in \Sigma^*$$

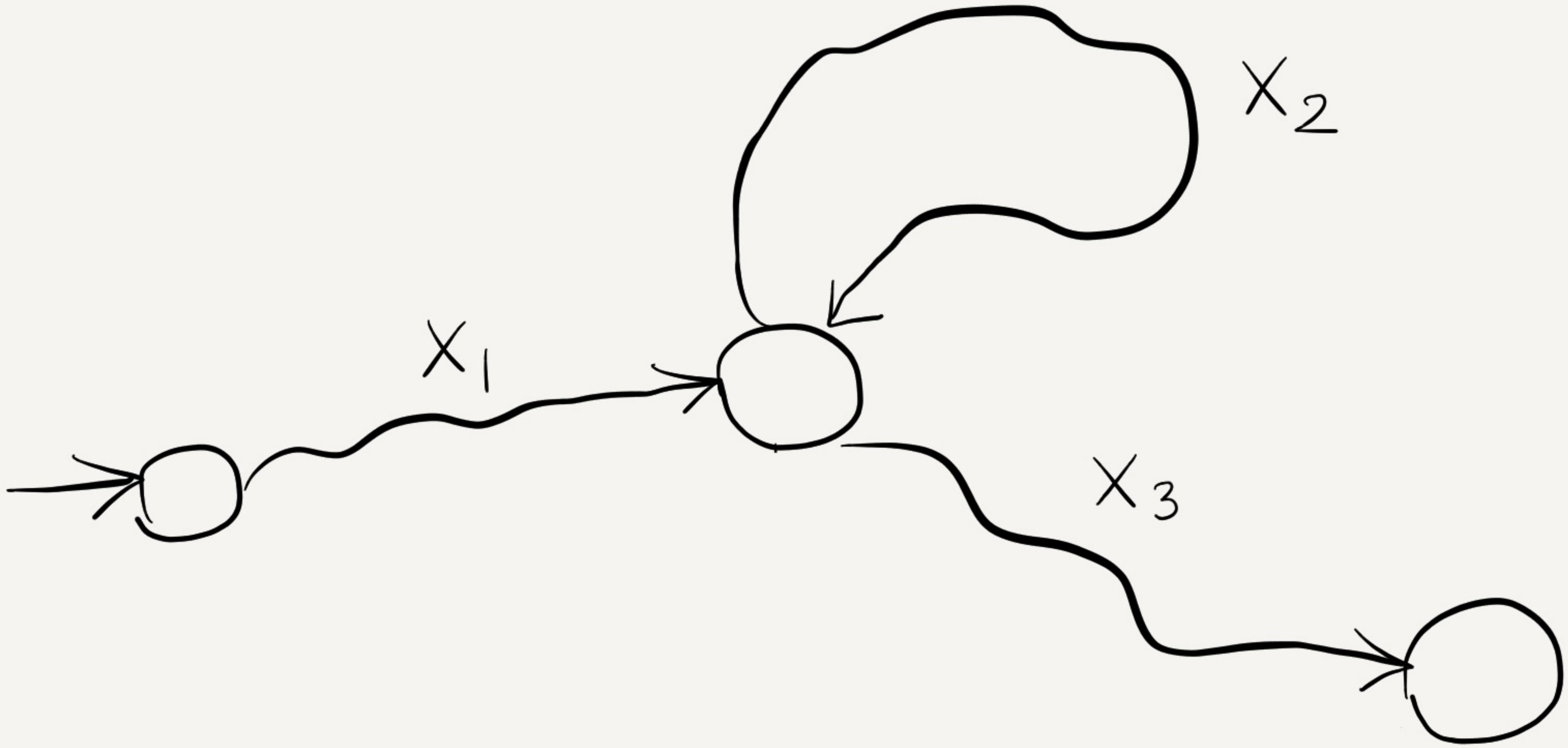
$$\big(z = x_1 x_2 x_3 \ \wedge$$

$$x_2 \neq \varepsilon \ \wedge$$

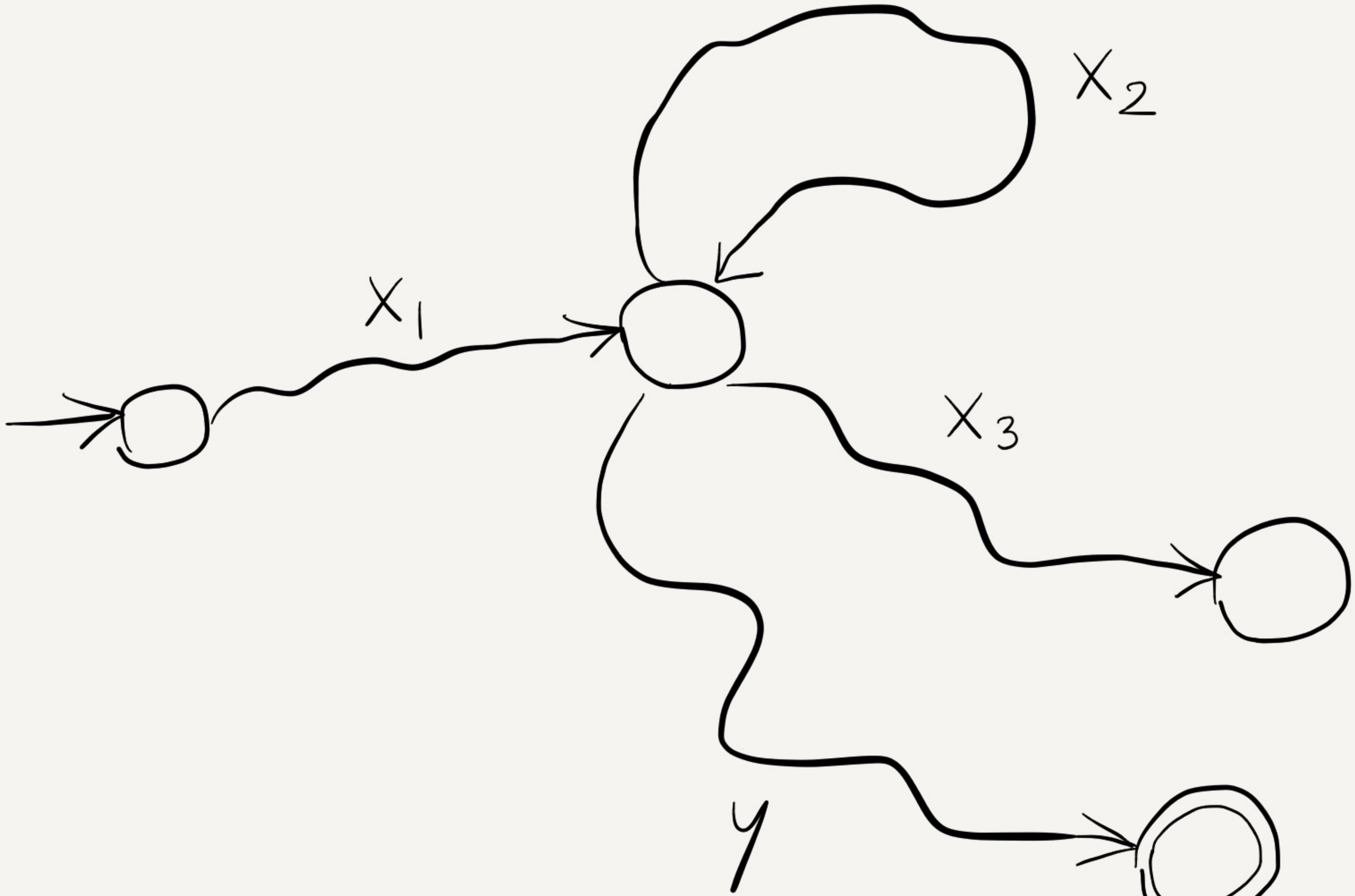
$$|x_1 x_2| \leq p \ \wedge$$

$$\forall y \in \Sigma^*\ \forall j \in \mathbb{N}$$

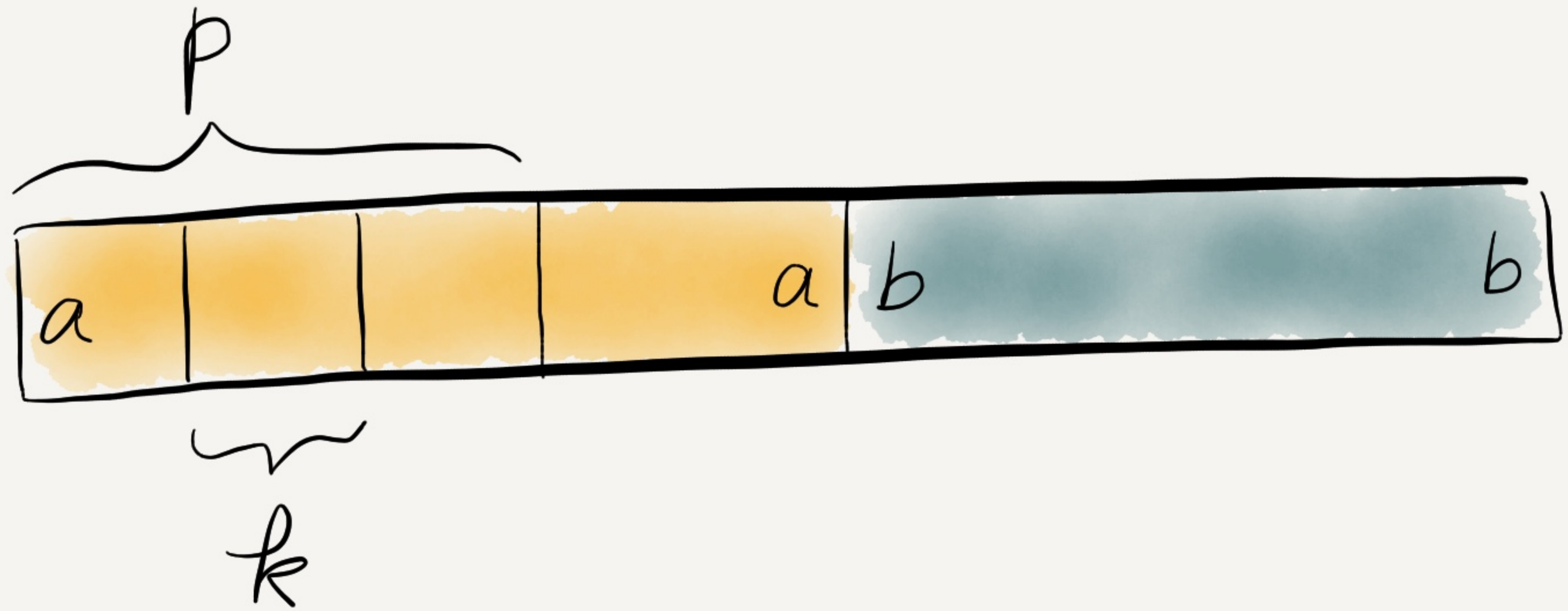
$$(x_1 x_2 y \in W \Leftrightarrow x_1 x_2^j y \in W)\big)\big)$$



X_1
X_2
X_3



$$x_1 x_2 y \in W \iff x_1 x_2^j y \in W$$

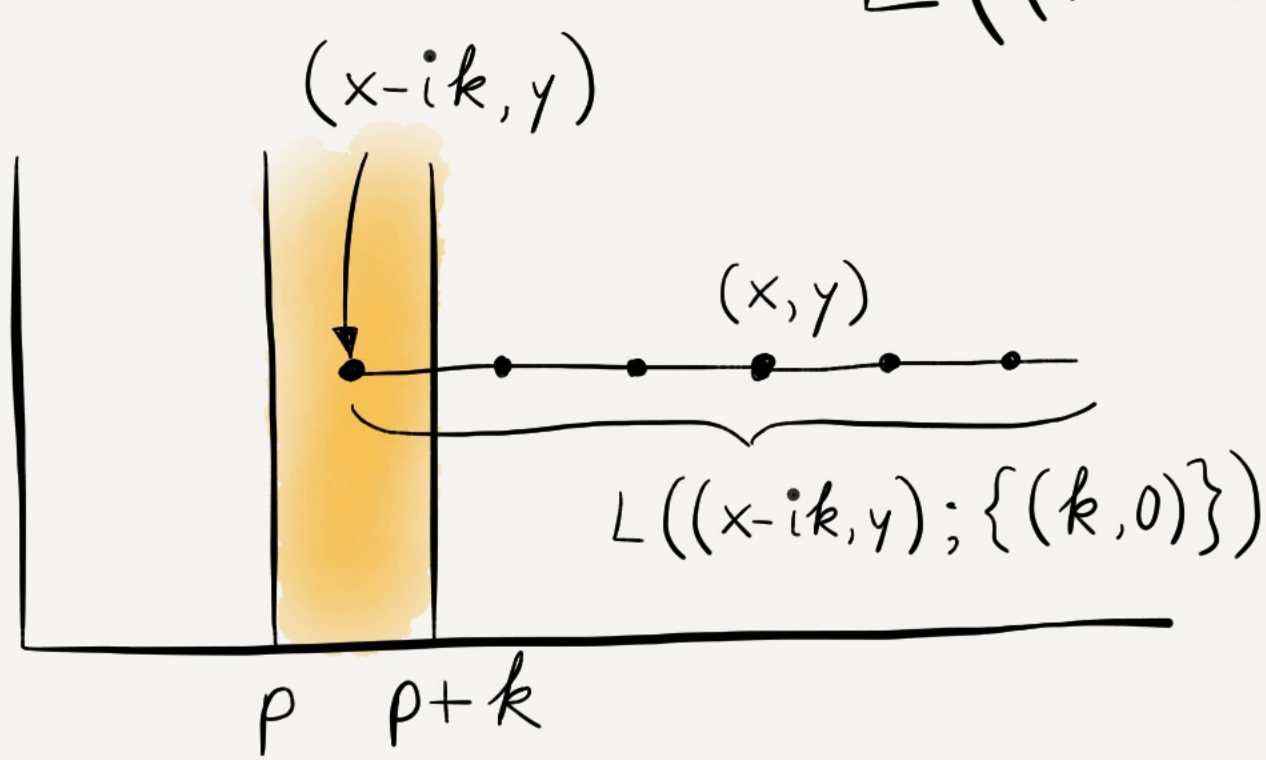


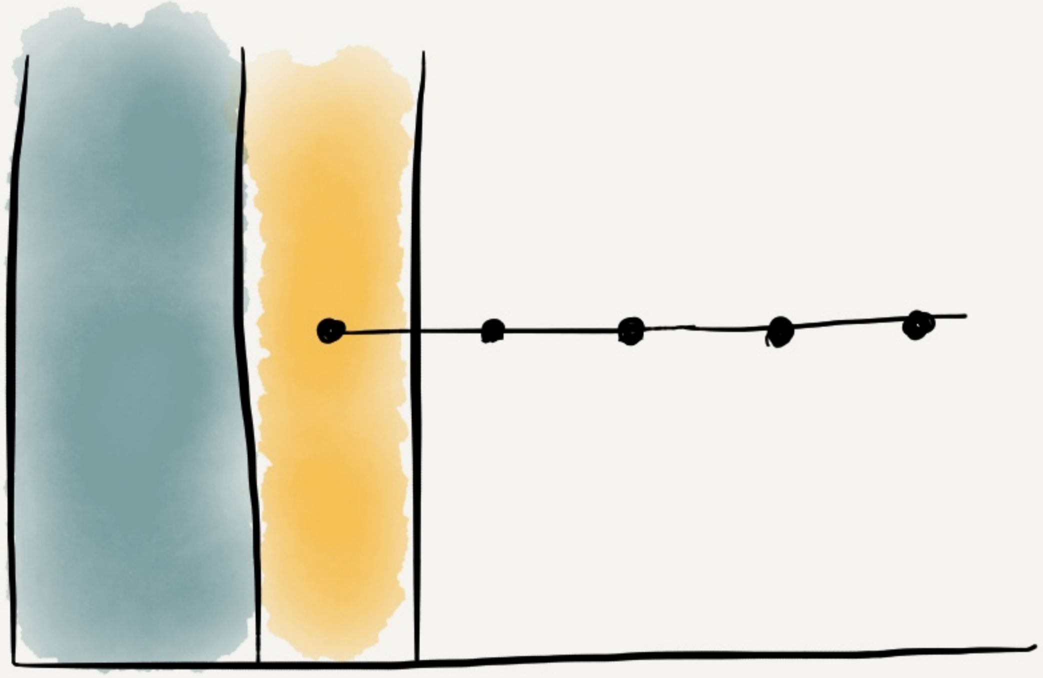
$$a^x b^y \in W \implies a^{x+(j-1)k} b^y \in W$$

$$\text{if } x \geq p$$

Lemma W_Q regular $\Rightarrow$

$$\exists p k \in \mathbb{N} \Big(1 \le k \le p \wedge \forall x y \in \mathbb{N} \Big((x,y) \in V_Q \wedge x \ge p \Rightarrow \exists i \big(p \le x - ik < p + k \wedge L((x-ik); \{(k,0)\}) \subseteq V_Q \big) \Big) \Big)$$





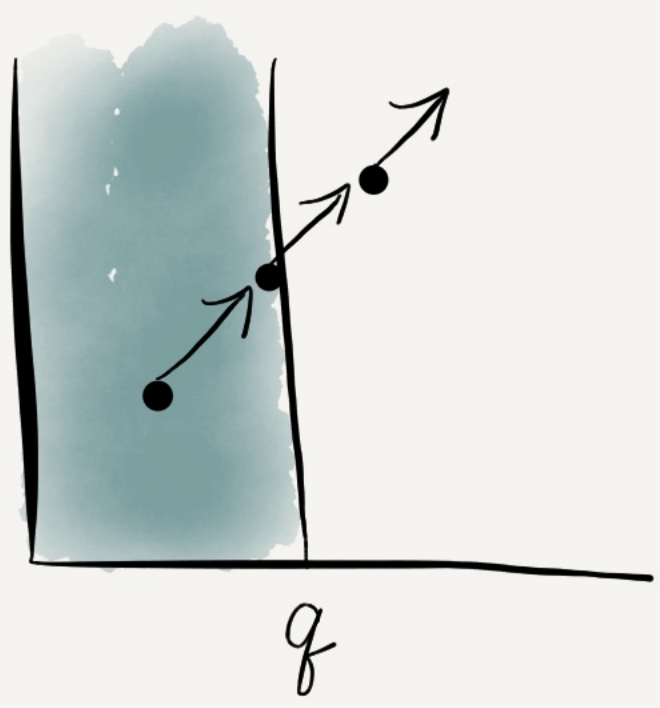
$$V_Q = V_Q \cap ([0, p-1] \times \mathbb{N}) \cup$$
$$L(V_Q \cap ([p, p+k-1] \times \mathbb{N}); \{(k, 0)\})$$

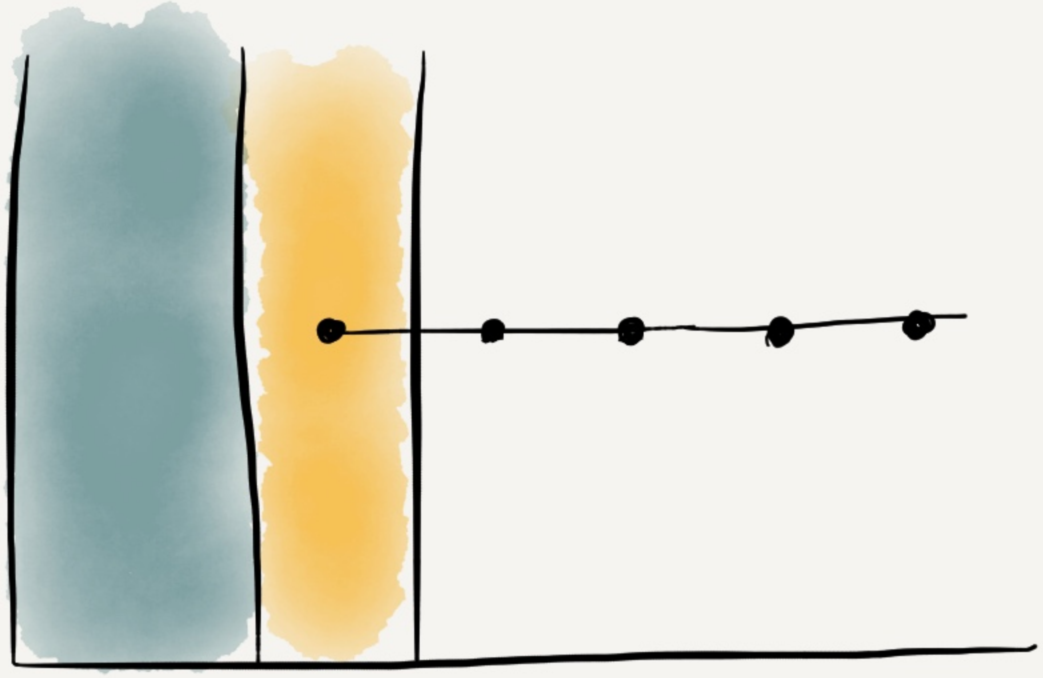
Lemma

$S \subseteq [0,q] \times \mathbb{N}$ is semilinear $\Rightarrow$

$$S = L(O_1 ; \emptyset) \cup L(O_2 ; \{(0,\ell)\})$$

for some $\ell \in \mathbb{N}$ and finite sets $O_1, O_2 \subseteq \mathbb{N}^2$





$$V_Q = V_Q \cap ([0, p-1] \times \mathbb{N}) \cup$$
$$L(V_Q \cap ([p, p+k-1] \times \mathbb{N}); \{(k,0)\})$$
$$= L(O_1; \emptyset) \cup L(O_2; \{(0,\ell)\}) \cup$$
$$L(L(O_3; \emptyset) \cup L(O_4; \{(0,\ell)\}); \{(k,0)\})$$

Theorem W_Q is accepted by a DFA

$$\Updownarrow$$

$\Uparrow$ counting modulo

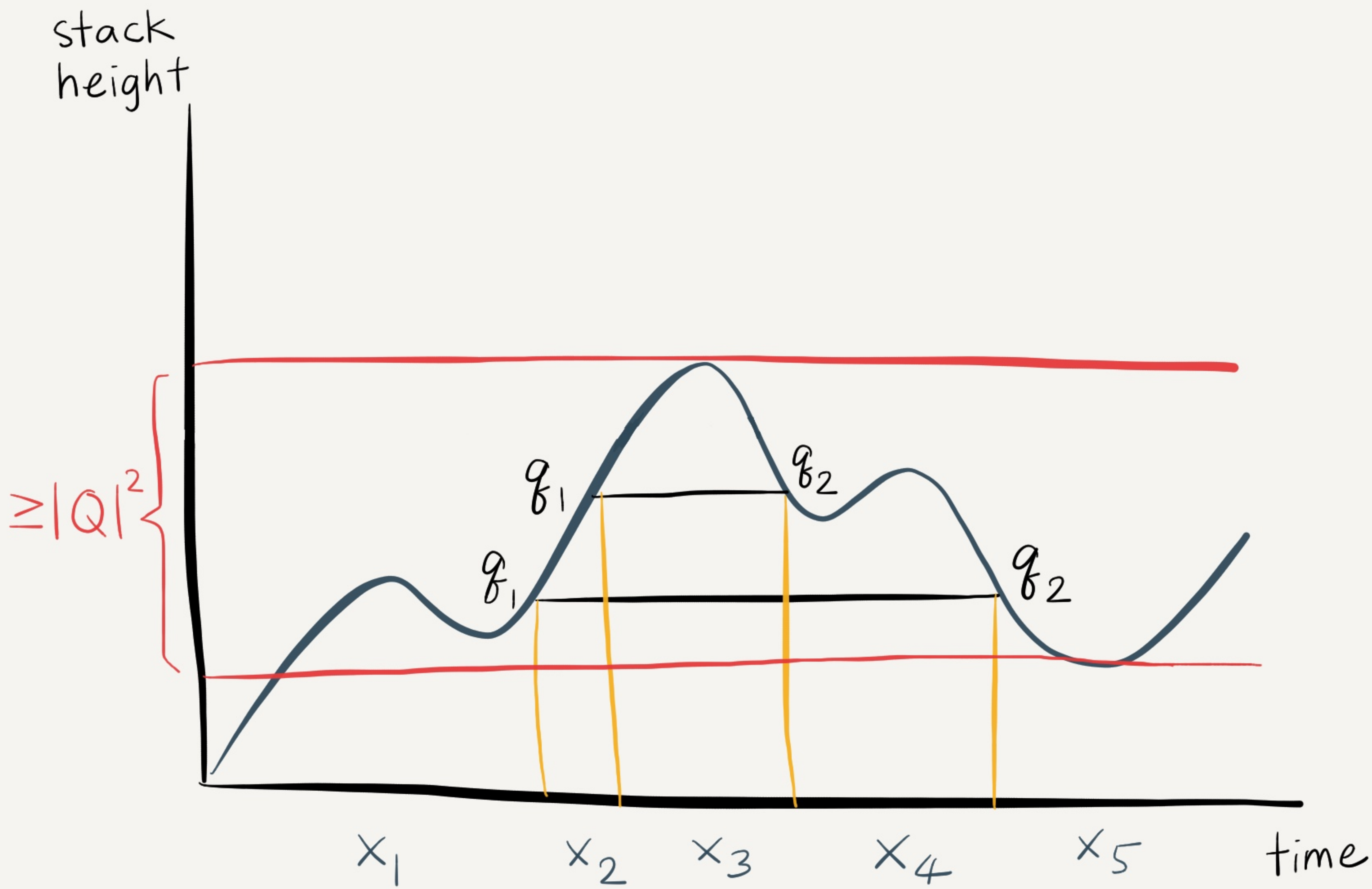
$$
\begin{aligned}
V_Q = {} & L(O_1; \emptyset) \cup \\
& L(O_2; \{(k,0)\}) \cup \\
& L(O_3; \{(0,\ell)\}) \cup \\
& L(O_4; \{(k,0),(0,\ell)\})
\end{aligned}
$$

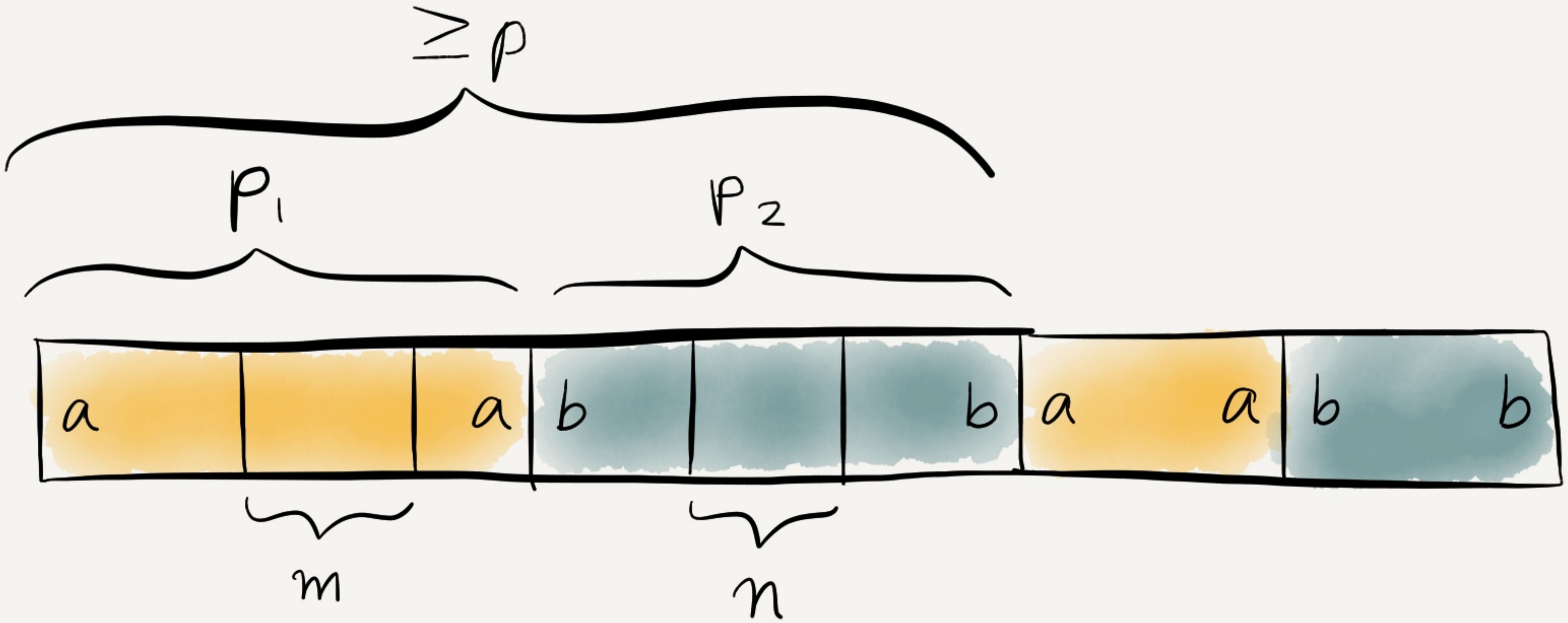
for some $k, \ell \in \mathbb{N}$ and finite sets $O_1, O_2, O_3, O_4 \subseteq \mathbb{N}^2$

W_Q accepted by	V_Q	Q definable in
DFA $\equiv$ NFA	generators $\subseteq \{(k,0),(0,\ell)\}$	$FO(D_n)$
DPDA	?	
NPDA	semilinear	$MSO(I)$

Lemma $W \subseteq \Sigma^*$ is a DCFL $\Rightarrow$

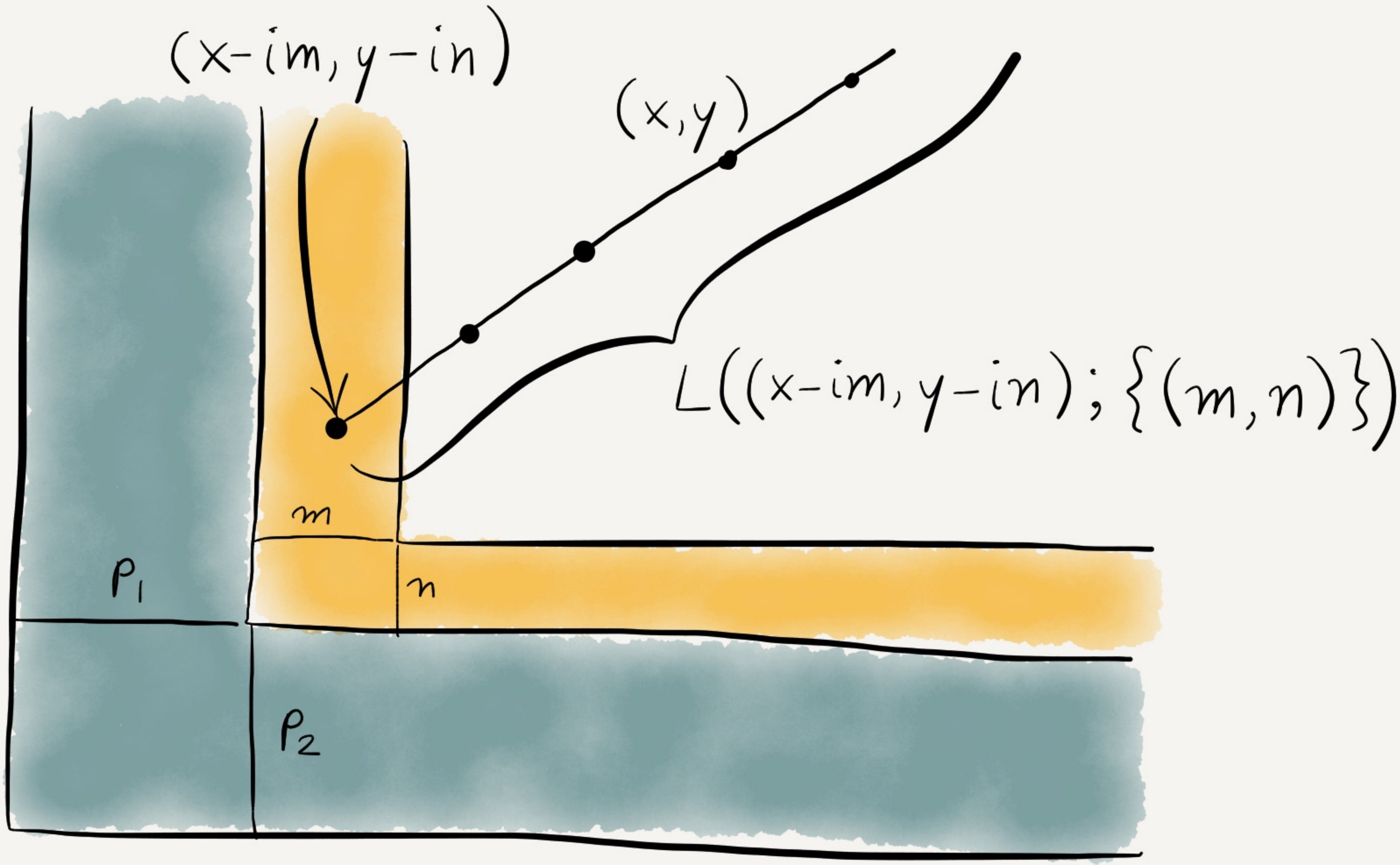
$$\exists p \in \mathbb{N}\ \forall z \in W \Big(|z| \geq p \Rightarrow$$
$$\exists x_1 x_2 x_3 x_4 x_5 \in \Sigma^*$$
$$\big(z = x_1 x_2 x_3 x_4 x_5 \wedge$$
$$x_2 x_4 \neq \varepsilon \wedge$$
$$\forall y \in \Sigma^* \forall j \in \mathbb{N}$$
$$(x_1 x_2 x_3 x_4 y \in W \Leftrightarrow x_1 x_2^j x_3 x_4^j y \in W)\big)\Big)$$


stack height
≥|Q|²
q₁
q₁
q₂
q₂
x₁
x₂
x₃
x₄
x₅
time



$$a^{x} b^{y} \in W \Longrightarrow a^{x+(j-1)m}\, b^{y+(j-1)n} \in W$$

$$\text{if } x \geq p_1 ,\ y \geq p_2$$



(x-im, y-in)
(x, y)
L((x-im, y-in); {(m, n)})
m
n
P1
P2

Theorem W_Q is accepted by a DPDA

$$\Updownarrow$$

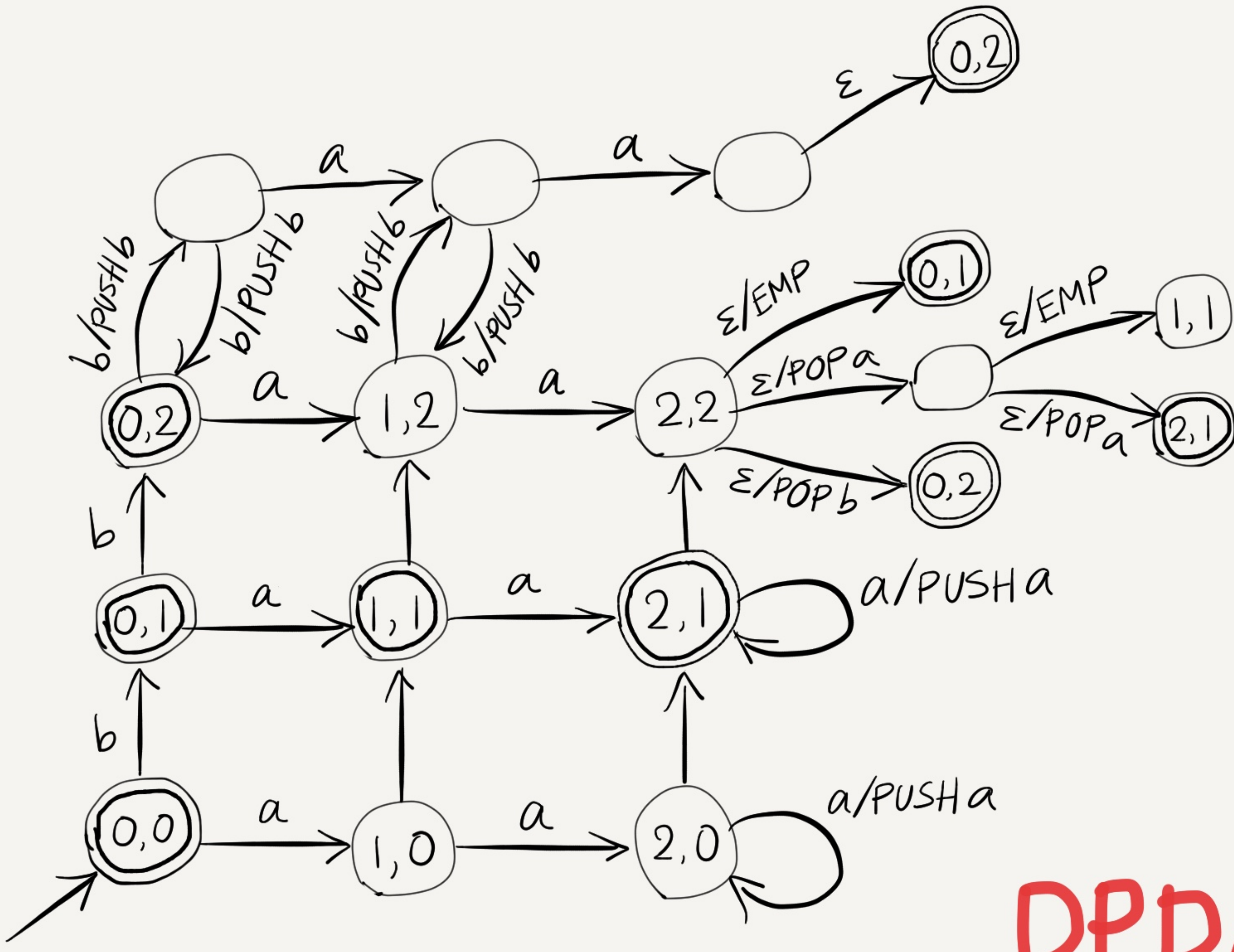
$\Uparrow$ matching + counting modulo

$$
\begin{aligned}
V_Q = \; & L(O_1; \emptyset) \cup \\
& L(O_2; \{(k,0)\}) \cup \\
& L(O_3; \{(0,\ell)\}) \cup \\
& L(O_4; \{(m,n)\}) \cup \\
& L(O_5; \{(k,0),(m,n)\}) \cup \\
& L(O_6; \{(0,\ell),(m,n)\})
\end{aligned}
$$

for some $k, \ell, m, n \in \mathbb{N}$ and

finite sets $O_1, O_2, O_3, O_4, O_5, O_6 \subseteq \mathbb{N}^2$

$L((0,1);\{(1,0),(2,1)\}) \cup L((0,0);\{(0,2),(2,1)\})$



DPDA

W_Q accepted by	V_Q	Q definable in
DFA $\equiv$ NFA	generators $\subseteq \{(k,0),(0,\ell)\}$	$FO(D_n)$
DPDA	generators $\subseteq \{(k,0),(0,\ell),(m,n)\}$	
NPDA	semilinear	$MSO(I)$

Some observations

- One counter suffices for both NPDA and DPDA.

  stack $\longrightarrow$ counter

- DPDA $\equiv$ real-time 1-CM

- NPDA-acceptable $\equiv$

  Boolean combination of DPDA-recognizable