Comp 330 — Lec 24 — Nov 23rd

Added a video about Lecture 23:
$\varphi_M(L(M))$

Undecidable problems about CFLs

Recall DPs about CFLs:

ISIN(G, w): "Given a CFG G & $w \in \Sigma^*$, is $w \in L(G)$?" → Decidable CYKS
(S ↓ Sakai)

EMPTY-CF(G): "Given CFG G, is $L(G) = \emptyset$?" → Decidable
GEN

Many simple decision problems about CFLs

① ALL-CF(G): "Given a CFG G, is $L(G) = \Sigma^*$?"
↑ set of terminals of G $G = (V, S, \Sigma, P)$
Undecidable!

② INTERSECT-CF(G_1, G_2): "Given a pair of CFGs G_1 & G_2, is $L(G_1) \cap L(G_2) \neq \emptyset$?"
Undecidable!

To show that ALL-CF is undecidable:

$$\overline{HP} \leq_m \text{ALL-CF}$$

binary encoding G

1. Convert $I_{\overline{HP}} = \langle M, x \rangle$ to $I_{\text{ALL-CF}} = \langle G \rangle$

2. M loops on x $\iff$ $L(G) = \Sigma^*$

set of terminals of G.

What could G possibly be? It's related to the computation history of a TM M running on input x. Intuition: If M loops on x then there is no finite string which can represent its computation history.

Q Γ Σ $\rightarrow \Sigma \subseteq \Gamma$

Def Given a TM M & an input $x \in \Sigma^*$, a valid computation history is a string of the form

$$\# \alpha_0 \# \alpha_1 \# \cdots \# \alpha_N \#$$

$abac\,q_7\,cba$

1) $\#$ is a special separator token $\notin \Gamma$

2) $\forall\, i \in 0 \ldots N$, $\alpha_i \in \Delta^*$ $\quad \Delta = \Gamma \cup Q$

& α_i is an instantaneous config. interpret q_0 as the letter q_0

3) $\alpha_0 = \vdash s\, x\, \square$ $\quad$ $\alpha_N = \vdash u\, q_a\, v\, \square$

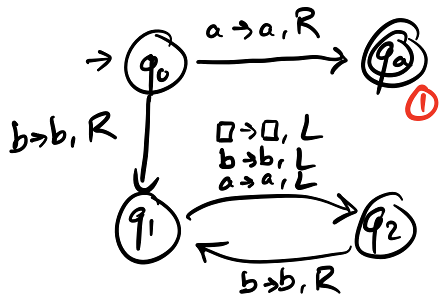
start state of M $\quad$ $= \vdash u\, q_R\, v\, \square$ $\quad$ $u, v \in \Gamma^*$

4) $\forall i \in 0 \ldots N-1$, α_i, α_{i+1} must follow

$$\alpha_i \xrightarrow[M]{1} \alpha_{i+1}$$

<u>Ex</u> Lecture 18

M



$x = abb$

Contains accept state => accepting I.C.

① $\alpha_{M,x} = \# \underbrace{\vdash q_0\, abb \square}_{\alpha_0} \# \underbrace{\vdash a\, q_a\, bb \square}_{\alpha_1 = \alpha_N} \#$

$x = ba$

② $\alpha_{M,x} = \# \underbrace{\vdash q_0\, ba \square}_{\alpha_1} \# \underbrace{\vdash b\, q_1\, a \square}_{\alpha_2} \#$

$\vdash q_2\, ba \square \# \vdash b\, q_1\, a \square \# \ldots$

infinite string

Valid computation histories are <u>finite</u> strings

=> ② is not a valid computation history.

$Q \quad \Gamma \quad \Sigma$

<u>Def</u> Given a TM M and $x \in \Sigma^*$,

$\Delta = \Gamma \cup Q \cup \{\#\}$

$\text{VALCOMPS}(M, x) = \{ \alpha \in \Delta^* :$ α is a valid computation history of M running on $x \}$

Remark Since M is a deterministic semi-infinite tape TM, there are only 2 options:

$$|\text{VALCOMPS}(M,x)| = 1 \leftrightarrow M \text{ halts on } x$$

$$|\text{VALCOMPS}(M,x)| = 0 \leftrightarrow M \text{ loops on } x$$

Ex In previous M

$$\text{VALCOMPS}(M,\ x=abb) = \{\underbrace{\text{"}\# \vdash q_0 abb \square \# \vdash a q_a bb \square \#\text{"}}_{\text{a single thing}}\}$$

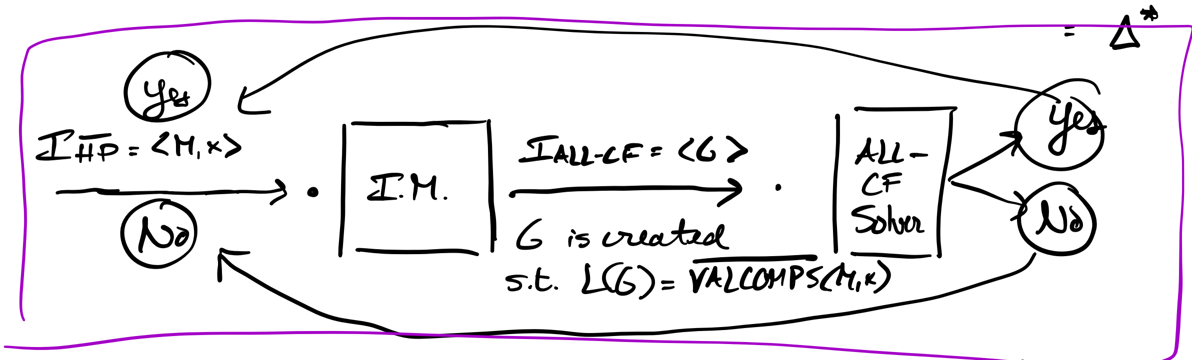
$$\text{VALCOMPS}(M,\ x=ba) = \{\ \}$$

$\overline{HP} \le_m$ ALL-CF $\rightarrow$ Given a CFG $G=(V,S,\underset{\substack{\uparrow \\ \text{set of} \\ \text{Terminals}}}{T},P)$, is $L(G) = T^*$?

Key observation:

$$M \text{ loops on } x \iff \text{VALCOMPS}(M,x) = \emptyset$$

$$\iff \overline{\text{VALCOMPS}(M,x)} = \overline{\emptyset} = \Delta^*$$



$\overline{HP}$ Solver.

$$ANS(I_{ALL\text{-}CF}) = \text{Yes} \Rightarrow L(G) = \overline{\text{VALCOMPS}(M,x)} = \Delta^* \Rightarrow$$

$\Rightarrow$ VALCOMPS$(M,x) = \phi$

$\Rightarrow$ M loops x

$\Rightarrow$ ANS$(I_{\overline{HP}})$ = Yes

ANS(I_{ALL-CF}) = No $\Rightarrow$ $\overline{\text{VALCOMPS}(M,x)} \neq \phi$

$\Rightarrow$ M halts on x

$\Rightarrow$ ANS$(I_{\overline{HP}})$ = No

If we show that $\overline{\text{VALCOMPS}(M,x)}$ is CF then $\overline{HP} \leq_m$ ALLCF $\Rightarrow$ ALLCF is undecidable.

To show that $\overline{\text{VALCOMPS}}$ is CF, we will first write the conditions $\alpha \in \Delta^*$, $\alpha \in$ VALCOMPS

$\alpha = \#\alpha_0\#\alpha_1\# \cdots \#\alpha_N$

c_1) α is a string of the form

$\#\alpha_0\# \ldots \#\alpha_N\#$

$\alpha_i \in (Q \cup \Gamma)^*$

This can be modeled via Reg Exp

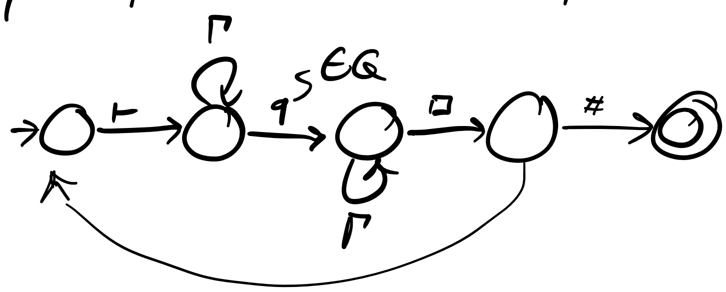
$\alpha REX = (s + q_a + q_r + \vdash + \sqcup + a + b)^*$

$Q = \{s, q_a, q_r\}$

$\Gamma = \{\vdash, \sqcup, a, b\}$

$r_{c_1} = \#\alpha REX\#(\alpha REX\#)^*\alpha REX\#$

$c_2)$ $\forall\ \alpha_i \in (\Gamma \cup Q)^*$, α_i contains exactly one $q \in Q$, start with $\vdash$, end with a $\square$



$c_3)$ α_0 should have the following form

$$r_{c_3} = \vdash s \underbrace{\Gamma^*}_{(a+b+\vdash+\square)^*} \square$$

$c_4)$ α_N must be a halting config

$$r_{c_4} = (\vdash \Gamma^* q_a \Gamma^* \square) + (\vdash \Gamma^* q_r \Gamma^* \square)$$

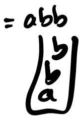
$c_5)$ $\forall\ i \in 0 \ldots N-1$, α_i, α_{i+1} must follow

$\alpha_i \xrightarrow[M]{1} \alpha_{i+1}$

α_i $\qquad$ α_{i+1}

$\# \ u a q b v \ \# \ u a b p v \ \#$ $\qquad$ $\{ww : w \in \{a,b\}^*\}$ NOT CF

$\delta(q,b) = (p,b,R)$ $\qquad$ $w = abb$



Cannot model c_5 even with CFG / PDA.

$$\text{VALCOMPS}(M,x) = \{\alpha \in \Delta^* : c_1 \wedge c_2 \wedge c_3 \wedge c_4 \wedge c_5\}$$
$$= \{\alpha \in \Delta^* : c_1\} \cap \{\alpha \in \Delta^* : c_2\} \cap \ldots \cap \{\alpha \in \Delta^* : c_5\}$$

$$= \underset{\text{REG}}{L_1} \cap \underset{\text{REG}}{L_2} \cap \underset{\text{REG}}{L_3} \cap \underset{\text{REG}}{L_4} \cap \underset{\text{NOT CF}}{L_5}$$

$$\overline{\text{VALCOMPS}} = \underbrace{\underbrace{\underbrace{\overline{L_1}}_{\text{REG}} \cup \overline{L_2} \cup \overline{L_3} \cup \underbrace{\overline{L_4}}_{\text{REG}}}_{\text{REG}} \cup \boxed{\underbrace{\overline{L_5}}_{\text{is CF!}}}}_{\text{CF}}$$

→ Show that $\overline{L_5}$ is CF.

Goal: Show $\overline{L_5}$ is CF. Design a PDA P s.t. $L(P) = \overline{L_5}$

$$\alpha \in \Delta^* , \quad \alpha \in \overline{L_5} \Rightarrow \alpha \notin L_5$$

$$\Rightarrow \exists\ \alpha_i, \alpha_{i+1}$$

$$\alpha_i \underset{M}{\overset{1}{\not\to}} \alpha_{i+1}$$

PDA will search for such a pair of α_i, α_{i+1}. If it finds this pair then it accepts. If it can't find such a pair, it will never get to accept ⇒ reject.

PDA $\boxed{P}$

α_i, α_{i+1}

$\alpha_i \underset{M}{\overset{1}{\not\to}} \alpha_{i+1}$

verify this guess

$$\alpha = \# \alpha_0 \# \alpha_1 \# \dots \# \alpha_i \# \alpha_{i+1} \# \dots \# \alpha_N \#$$

guesses that α_i is the first of the inconsistent α_i, α_{i+1}

What cases lead to a violation?

Case 1 δ of M is not respected

$$\overbrace{\# \;\ldots\; bq_1a \;\ldots\;}^{\alpha_i} \# \;\overbrace{\ldots\; q_2bb^{\times} \;\ldots\;}^{\alpha_{i+1}} \#$$

violation b/c b is written instead of a

$$\delta(q_1, a) = (q_2, a, L)$$

Case 2 Miscopying

$$\ldots \; \overbrace{\#\underbrace{\quad}_{10} a \quad bq_1a}^{\alpha_i} \qquad \overbrace{\#\underbrace{\quad}_{10} b \quad q_2ba}^{\alpha_{i+1}} \qquad \#$$

miscopied letter

How can a PDA detect and verify case 1 / case 2 ?

3 letter windows (3 l.w.)

$$\# \quad \alpha_i \quad \# \quad \alpha_{i+1} \quad \#$$

Idea 1 : $\# \, u_1 \, x \, u_2 \, \# \, w_1 \, y \, w_2 \, \#$ $x \,\&\, y \Rightarrow$ Miscopying / δ not respected

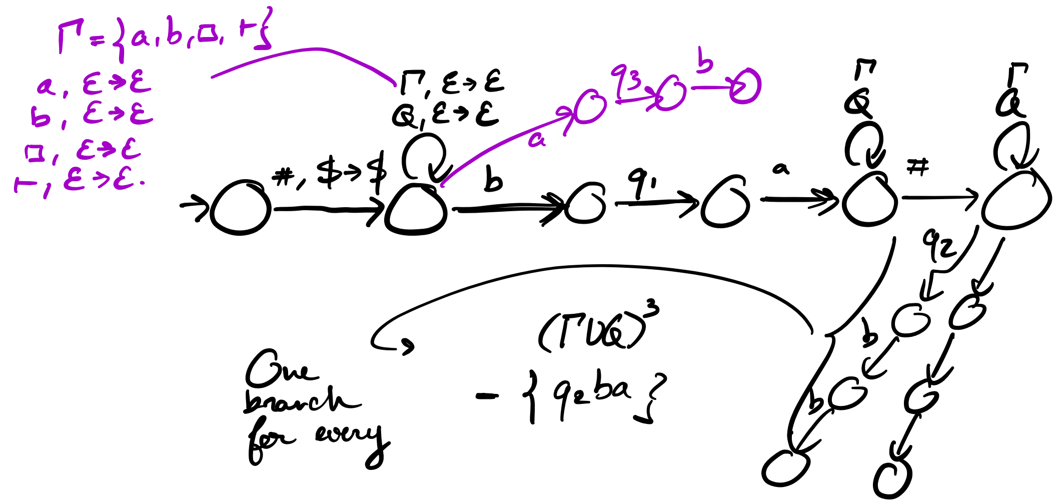
$|u_1| = |w_1|$ ...

Related to checking a mismatch b/w x & y

$\Rightarrow$ Since Γ, Q & δ are finite, the PDA can hardcode every possible mismatch

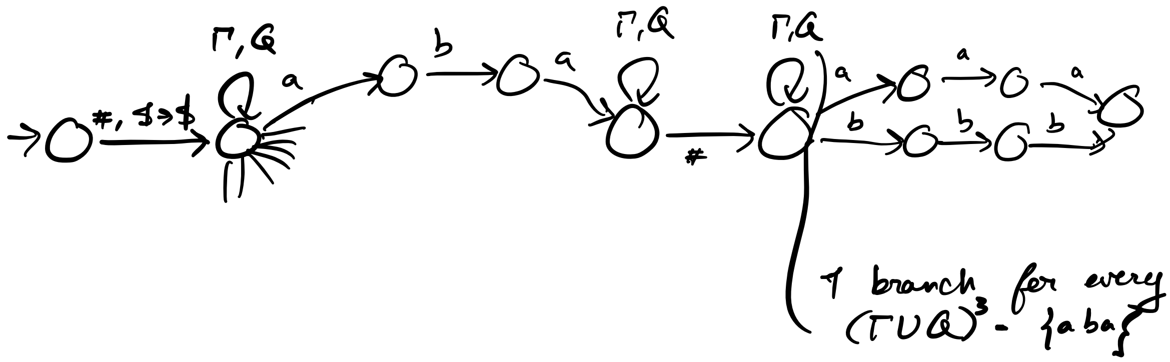
$$\delta(q_1, a) = (q_2, a, L)$$

Case 1 $\# \ldots \boxed{bq_1a} \ldots \# \ldots \boxed{q_2bb^{\times}} \ldots \#$

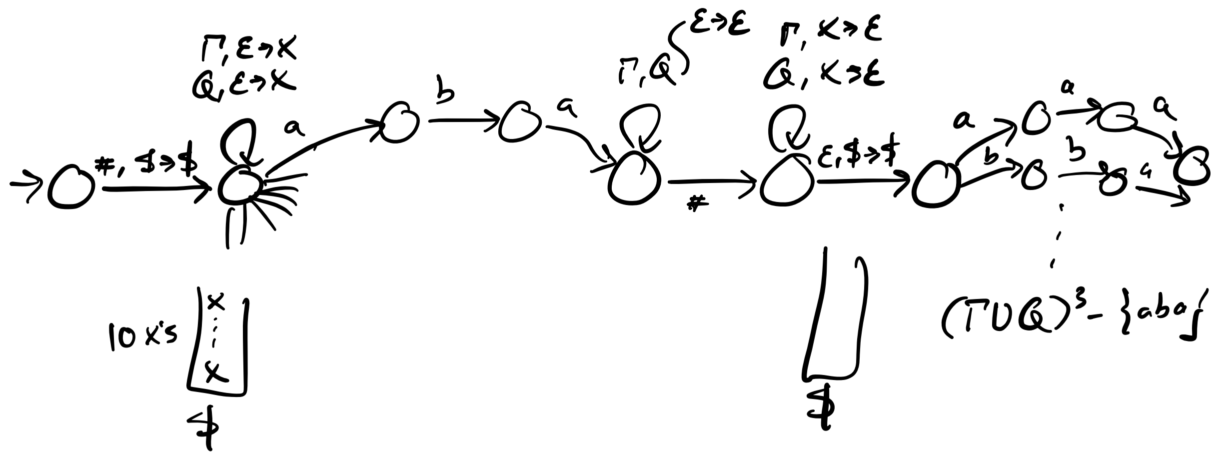


$\delta(q_1, a) = (q_2, a, L)$

Case 2 # aba … bq_1a … — # aaa … q_2ba … #



Idea 2: Related to guessing and verifying x & y at the same distance from their respective #. ⇒ Use the stack to push until the PDA guesses that x is the inconsistent 3.l.w. and then pop starting from the next # verifying the 3l.w y with x once the stack is empty.



What's the point? PDA P accepts $\overline{L_5}$

$\Rightarrow \overline{L_5}$ is CF $\Rightarrow \exists$ CFG G s.t.

$$L(G) = \overline{L_1} \cup \ldots \cup \overline{L_4} \cup \overline{L_5}$$
$$= \overline{\text{VALCOMPS}(M,x)}$$

Given a $\overset{I_{\overline{HP}}}{\overset{\shortparallel}{\langle M, x\rangle}}$, we can create a CFG G s.t. $L(G) = \overline{\text{VALCOMPS}}$ and check $\overline{\text{VALCOMPS}} = \Delta^*$

$\text{VALCOMPS} = \emptyset$

This acts as an $\overline{HP}$ solver $\Rightarrow$ ALL-CF is undecidable. $\square$

ALL-CF is undecidable. Is it CE or co-CE?

If answer to ALL-CF $\{G = (V, S, T, P)\}$ is no then $\exists$ $w \in T^*$ s.t. $w \notin L(G)$. If I know this is true a priori, I can check/verify.

for $w \in T^*$ do

1. Run CYKS algorithm on G & w

   ↳ Halts & gives a Y/N answer in finite # steps

2. If CYKS says No then accept

ALL-CF is co-CE!

Ex Show that CONCAT-CF is undecidable

CONCAT-CF(G_1, G_2, G_3) : " Given CFGs G_1, G_2, G_3, is $L(G_1) \cdot L(G_2) = L(G_3)$?"

$$\text{ALL-CF} \leq_m \text{CONCAT-CF}$$

1. $I_{\text{ALL-CF}} = \langle G \rangle \rightarrow I_{\text{CONCAT-CF}} = \langle G_1, G_2, G_3 \rangle$

   $G_1 := G$

   $G_2 := S \to \varepsilon \qquad L(G_2) = \{\varepsilon\}$

   $G_3 := S \to aS \mid bS \mid \varepsilon \qquad L(G_3) = \Sigma^*$

2. yes $\rightsquigarrow L(G) = \Sigma^* \implies L(G_1) = L(G) = \Sigma^*$

   $\underset{\Sigma^*}{L(G_1)} \cdot \underset{\{\varepsilon\}}{L(G_2)} = \underset{\Sigma^*}{L(G_3)}$ ↳ yes

No $\nearrow$ $L(G) \neq \Sigma^*$ $\Rightarrow$ $L(G_1) = L(G) \neq \Sigma^*$ $\rightsquigarrow$ No

$$\underbrace{L(G_1)}_{\neq\ \Sigma^*} \cdot \underbrace{L(G_2)}_{=\ \{\varepsilon\}} = \underbrace{L(G_3)}_{=\ \Sigma^*}$$

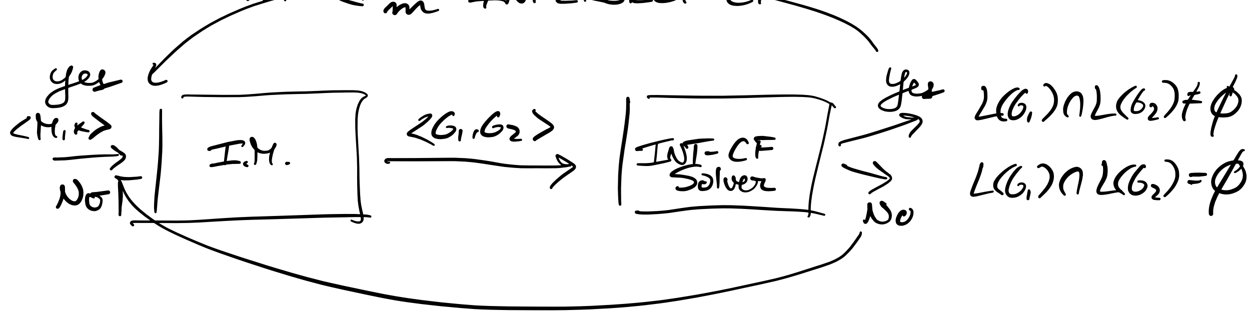
$$\underbrace{\Sigma^* \cdot \{\varepsilon\}}_{\neq\ \Sigma^*}$$

This shows that CONCAT-CF is undecidable

INTERSECT-CF $\langle G_1, G_2\rangle$ is undecidable

"Given CFGs G_1 & G_2, $L(G_1) \cap L(G_2) \neq \phi$?"

HP $\leq_m$ INTERSECT-CF



Design G_1 & G_2 s.t. $L(G_1) \cap L(G_2) =$ VALCOMPS2(M, x)

if N is even, α_N
if N is odd, α_N^R

VALCOMPS2$(M,x) = \{ \#\alpha_0\#\alpha_1^R\#\alpha_2\#\alpha_3^R\#\alpha_4\# \cdots \#\alpha_N\# :$
$\exists\, \alpha \in \Delta^*,\ \alpha = \#\alpha_0\#\alpha_1\# .. \#\alpha_N$
is a valid computation history of
M seen on x $\}$

$(xy)^R = y^R x^R$

$\# \alpha_0 \# \alpha_1^R \# \alpha_2 \# \alpha_3^R \# \; \alpha_i \# \; \dots \# \alpha_N \# :$

$\alpha_0 \qquad\qquad \alpha_1^R \qquad \alpha_1 = u\,a\,b\,q_1\,v$

PDA $\quad P_1 \quad \# u\, a q_0 b\, v \# \; v^R q_1\, ba\, u^R \#$

$\hookrightarrow$ Match $\alpha_0 \leftrightarrow \alpha_1^R$

$\alpha_2 \leftrightarrow \alpha_3^R$

$\vdots$

PDA $\quad P_2$

$\hookrightarrow$ Match $\alpha_1^R \leftrightarrow \alpha_2$

$\alpha_3^R \leftrightarrow \alpha_4$

$\vdots$

$P_1 \to G_1 \quad \& \quad P_2 \to G_2$

$\alpha_i \xrightarrow{M} \alpha_{i+1}$

i even

i odd

Yes

$L(G_1) \cap L(G_2) = \text{VALCOMPZ}(M, x)$

$\neq \emptyset$

$\Rightarrow$ M halts on x **Yes**

No

$L(G_1) \cap L(G_2) = \text{VALCOMPS2}(M, x) = \emptyset$

$\Rightarrow \text{VALCOMPS}(M, x) = \emptyset$

$\Rightarrow$ M loops x **No**

$HP \leq_m$ INT-CF

$\uparrow$

undecidable

but CE