Automata Theory

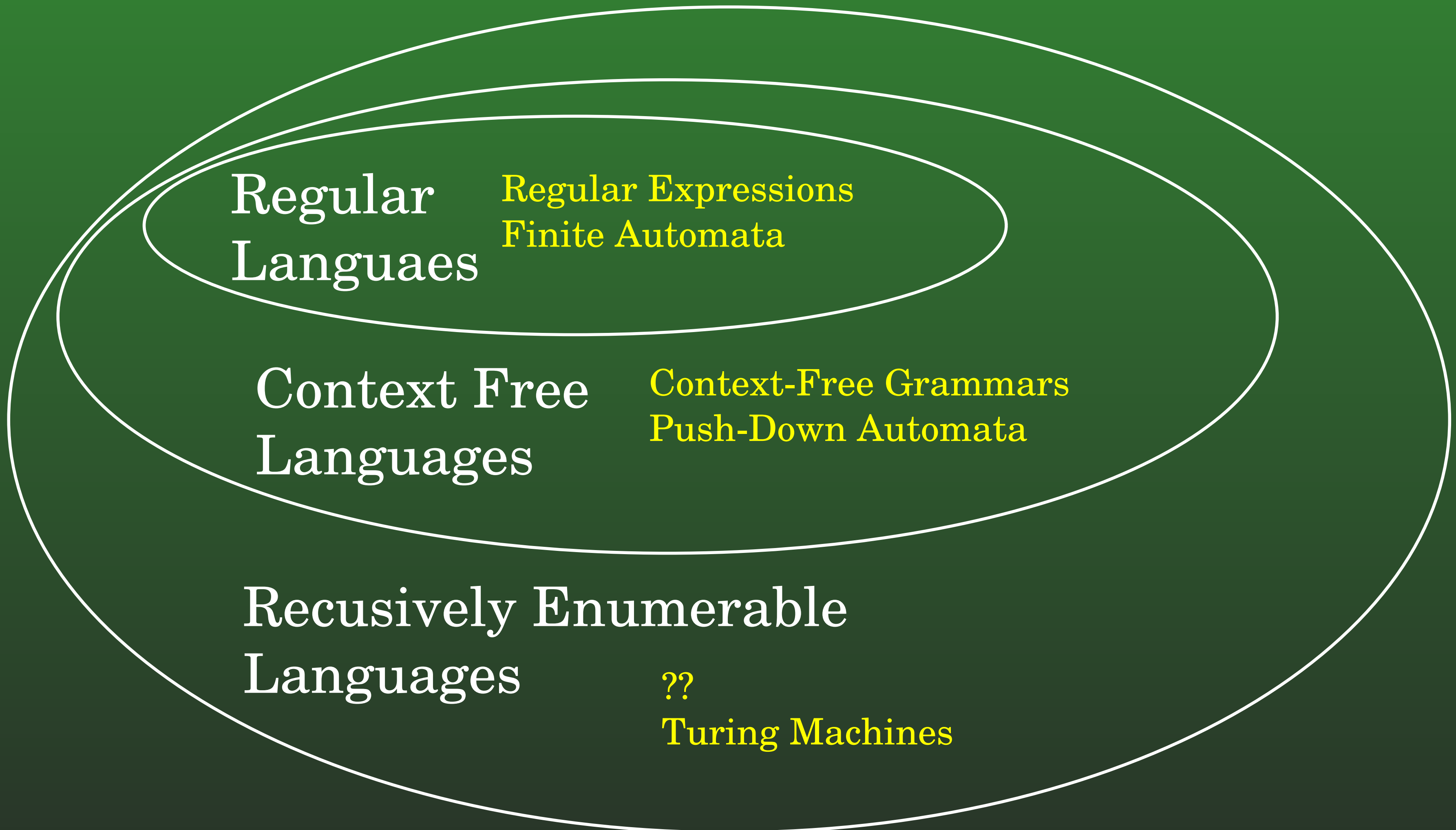
CS411-2015F-13

Unrestricted Grammars

David Galles

Department of Computer Science
University of San Francisco

13-0: Language Hierarchy



13-1: CFG Review

$G = (V, \Sigma, R, S)$

- V = Set of symbols, both terminals & non-terminals
- $\Sigma \subset V$ set of terminals (alphabet for the language being described)
- $R \subset ((V - \Sigma) \times V^*)$ Set of rules
- $S \in (V - \Sigma)$ Start symbol

13-2: Unrestricted Grammars

$G = (V, \Sigma, R, S)$

- V = Set of symbols, both terminals & non-terminals
- $\Sigma \subset V$ set of terminals (alphabet for the language being described)
- $R \subset (V^*(V - \Sigma)V^* \times V^*)$ Set of rules
- $S \in (V - \Sigma)$ Start symbol

13-3: Unrestricted Grammars

- $R \subset (V^*(V - \Sigma)V^* \times V^*)$ Set of rules
- In an Unrestricted Grammar, the left-hand side of a rule contains a string of terminals and non-terminals (at least one of which must be a non-terminal)
- Rules are applied just like CFGs:
  - Find a substring that matches the LHS of some rule
  - Replace with the RHS of the rule

13-4: Unrestricted Grammars

- To generate a string with an Unrestricted Grammar:
  - Start with the initial symbol
  - While the string contains at least one non-terminal:
    - Find a substring that matches the LHS of some rule
    - Replace that substring with the RHS of the rule

13-5: Unrestricted Grammars

- Example: Grammar for $L = \{a^n b^n c^n : n > 0\}$
  - First, generate $(ABC)^*$
  - Next, non-deterministically rearrange string
  - Finally, convert to terminals ($A \rightarrow a$, $B \rightarrow b$, etc.), ensuring that string was reordered to form $a^*b^*c^*$

13-6: Unrestricted Grammars

- Example: Grammar for $L = \{a^n b^n c^n : n > 0\}$

$$
\begin{aligned}
S &\rightarrow ABCS \\
S &\rightarrow T_C \\
CA &\rightarrow AC \\
BA &\rightarrow AB \\
CB &\rightarrow BC \\
CT_C &\rightarrow T_C c \\
T_C &\rightarrow T_B \\
BT_B &\rightarrow T_B b \\
T_B &\rightarrow T_A \\
AT_A &\rightarrow T_A a \\
T_A &\rightarrow \epsilon
\end{aligned}
$$

13-7: Unrestricted Grammars

$$
\begin{aligned}
S &\Rightarrow ABCS \\
&\Rightarrow ABCABCS \\
&\Rightarrow ABACBCS \\
&\Rightarrow AABCBCS \\
&\Rightarrow AABBCCS \\
&\Rightarrow AABBCCT_C \\
&\Rightarrow AABBCT_Cc \\
&\Rightarrow AABBT_Ccc \\
&\Rightarrow AABBT_Bcc \\
&\Rightarrow AABT_Bbcc \\
&\Rightarrow AAT_Bbbcc \\
&\Rightarrow AAT_Abbcc \\
&\Rightarrow AT_Aabbcc \\
&\Rightarrow T_Aaabbcc \\
&\Rightarrow aabbcc
\end{aligned}
$$

13-8: Unrestricted Grammars

$$
\begin{aligned}
S &\Rightarrow ABCS & &\Rightarrow AAABBBBCCCT_C \\
&\Rightarrow ABCABCS & &\Rightarrow AAABBBCCT_Cc \\
&\Rightarrow ABCABCABCS & &\Rightarrow AAABBBCT_Ccc \\
&\Rightarrow ABACBCABCS & &\Rightarrow AAABBBT_Cccc \\
&\Rightarrow AABCBCABCS & &\Rightarrow AAABBBT_Bccc \\
&\Rightarrow AABCBACBCS & &\Rightarrow AAABBT_Bbccc \\
&\Rightarrow AABCABCBCS & &\Rightarrow AAABT_Bbbccc \\
&\Rightarrow AABACBCBCS & &\Rightarrow AAAT_Bbbbccc \\
&\Rightarrow AAABCBCBCS & &\Rightarrow AAAT_Abbbccc \\
&\Rightarrow AAABBCCBCS & &\Rightarrow AAT_Aabbbccc \\
&\Rightarrow AAABBCBCCS & &\Rightarrow AT_Aaabbbccc \\
&\Rightarrow AAABBBCCCS & &\Rightarrow T_Aaaabbbccc \Rightarrow aaabbbccc
\end{aligned}
$$

13-9: Unrestricted Grammars

- Example: Grammar for $L = \{ww : w \in a, b^*\}$

13-10: Unrestricted Grammars

- Example: Grammar for $L = \{ww : w \in a, b^*\}$
- Hints:
  - What if we created a string, and then rearranged it (like $(abc)^* \rightarrow a^n b^n c^n$)

13-11: Unrestricted Grammars

- Example: Grammar for $L = \{ww : w \in a, b^*\}$
- Hints:
  - What if we created a string, and then rearranged it (like $(abc)^* \rightarrow a^n b^n c^n$)
  - What about trying ww^R ...

13-12: Unrestricted Grammars

- $L = \{ww : w \in a, b^*\}$

$$
\begin{aligned}
S &\rightarrow S'Z \\
S' &\rightarrow aS'A \\
S' &\rightarrow bS'B \\
S' &\rightarrow \epsilon
\end{aligned}
$$

$$
\begin{aligned}
AZ &\rightarrow XZ & \quad BZ &\rightarrow YZ \\
AX &\rightarrow XA & \quad AY &\rightarrow YA \\
BX &\rightarrow XB & \quad BY &\rightarrow YB \\
aX &\rightarrow aa & \quad aY &\rightarrow ab \\
bX &\rightarrow ba & \quad bY &\rightarrow bb
\end{aligned}
$$

Unrestricted Grammars

- L_{UG} is the set of languages that can be described by an Unrestricted Grammar:
  - $L_{UG} = \{L : \exists$ Unrestricted Grammar $G, L[G] = L\}$
- Claim: $L_{UG} = L_{re}$
- To Prove:
  - Prove $L_{UG} \subseteq L_{re}$
  - Prove $L_{re} \subseteq L_{UG}$

13-14: $L_{UG} \subseteq L_{re}$

- Given any Unrestricted Grammar G, we can create a Turing Machine M that semi-decides $L[G]$

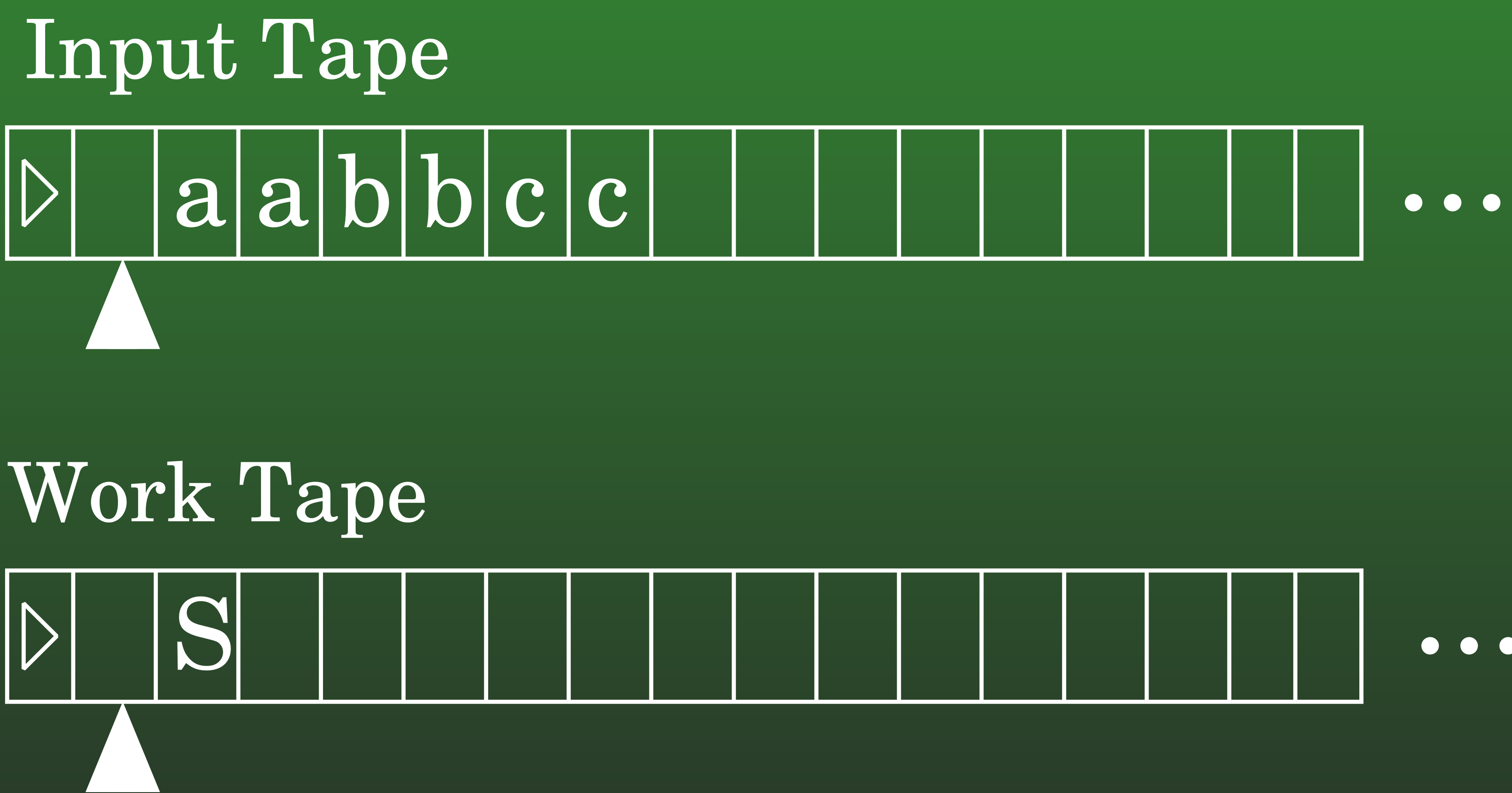
13-15: $L_{UG} \subseteq L_{re}$

- Given any Unrestricted Grammar G, we can create a Turing Machine M that semi-decides $L[G]$
- Two tape machine:
  - One tape stores the input, unchanged
  - Second tape implements the derivation
  - Check to see if the derived string matches the input, if so accept, if not run forever

13-16: $L_{UG} \subseteq L_{re}$

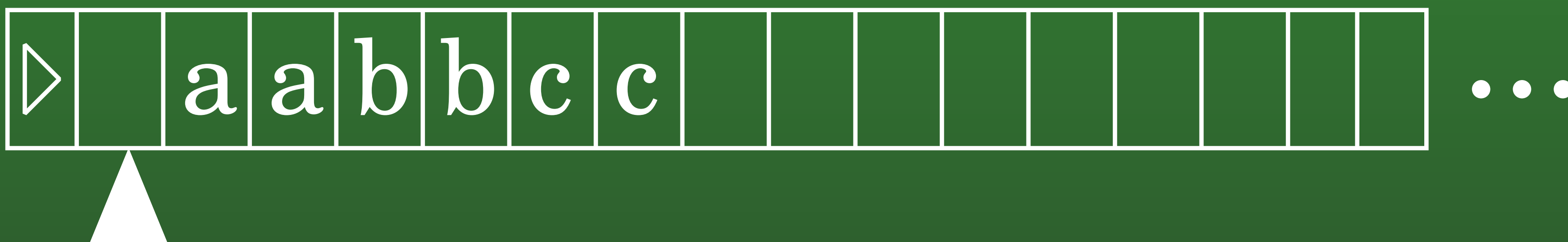
- To implement the derivation on the second tape:
  - Write the initial symbol on the second tape
  - Non-deterministically move the read/write head to somewhere on the tape
  - Non-deterministically decide which rule to apply
  - Scan the current position of the read/write head, to make sure the LHS of the rule is at that location
  - Remove the LHS of the rule from the tape, and splice in the RHS

13-17: $L_{UG} \subseteq L_{re}$

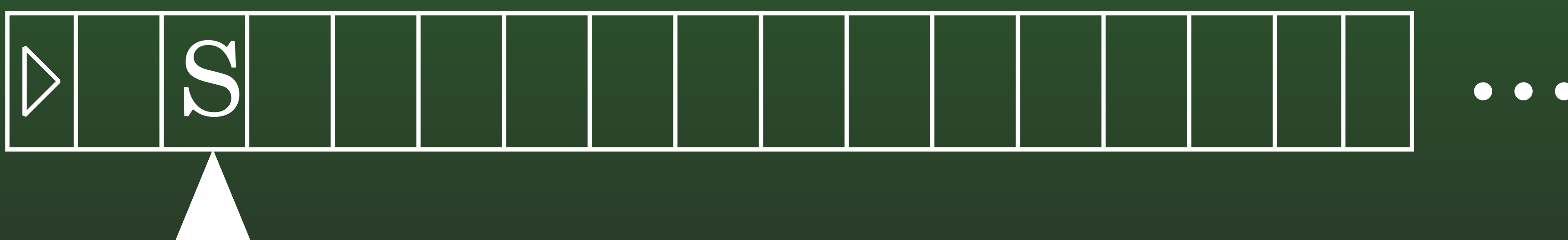


13-18: $L_{UG} \subseteq L_{re}$

Input Tape

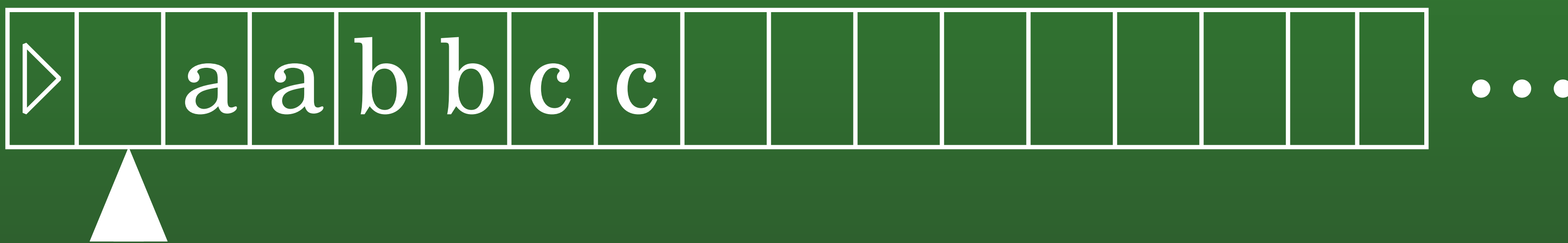


Work Tape

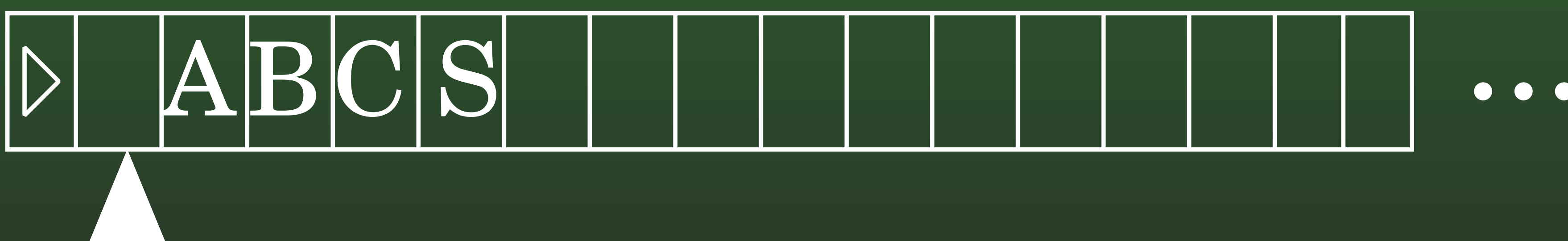


13-19: $L_{UG} \subseteq L_{re}$

Input Tape

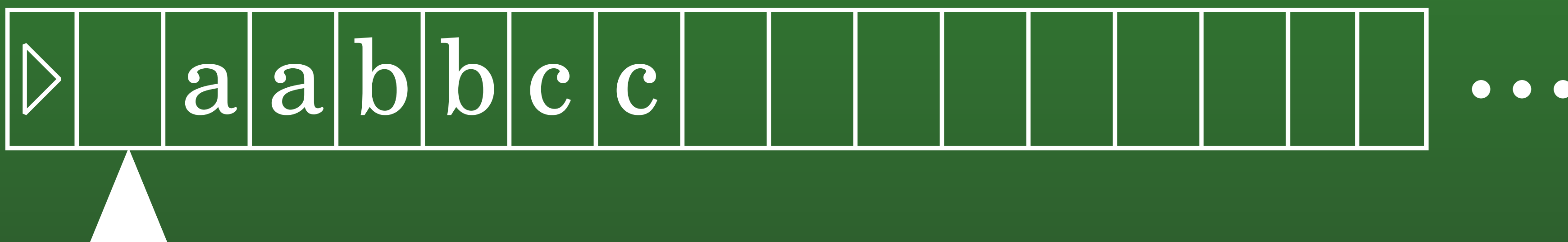


Work Tape

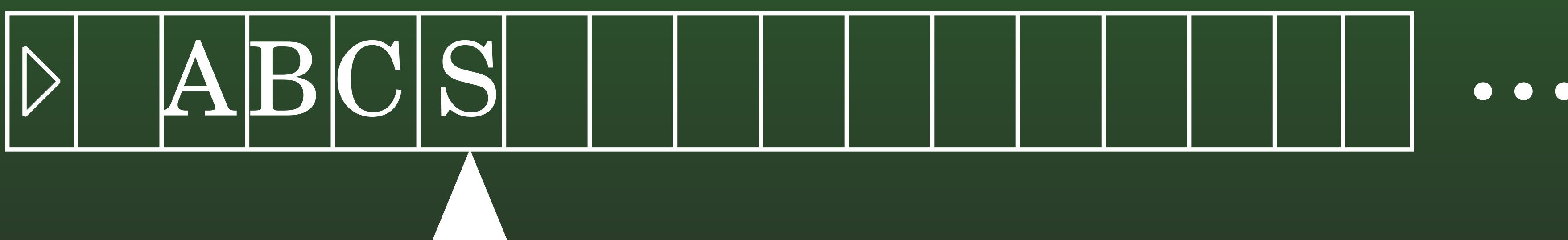


13-20: $L_{UG} \subseteq L_{re}$

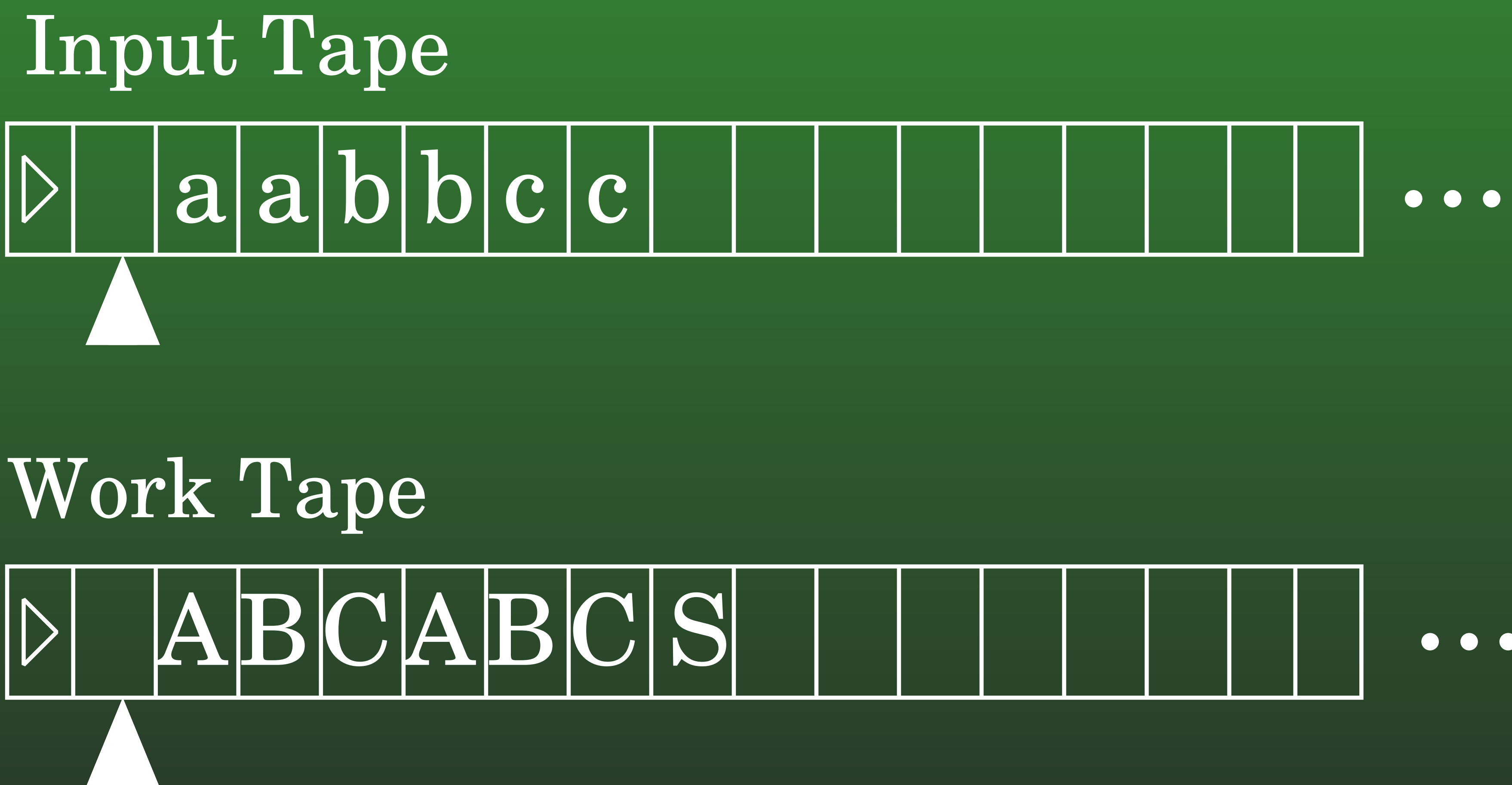
Input Tape



Work Tape

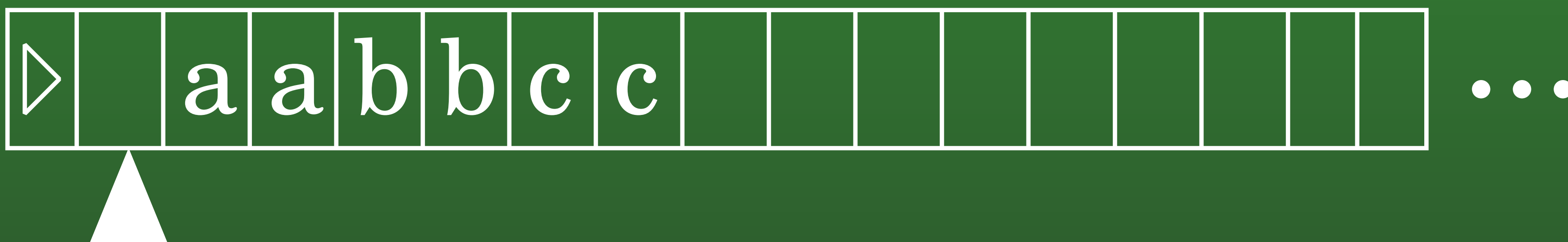


13-21: $L_{UG} \subseteq L_{re}$



13-22: $L_{UG} \subseteq L_{re}$

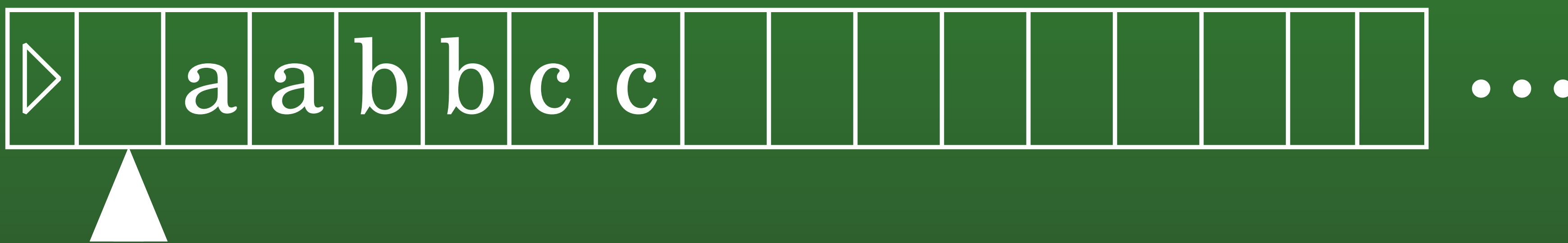
Input Tape



Work Tape

13-23: $L_{UG} \subseteq L_{re}$

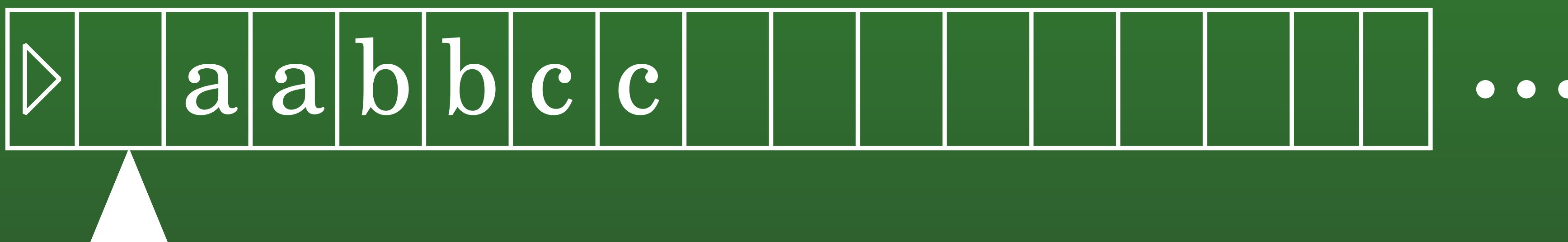
Input Tape



Work Tape

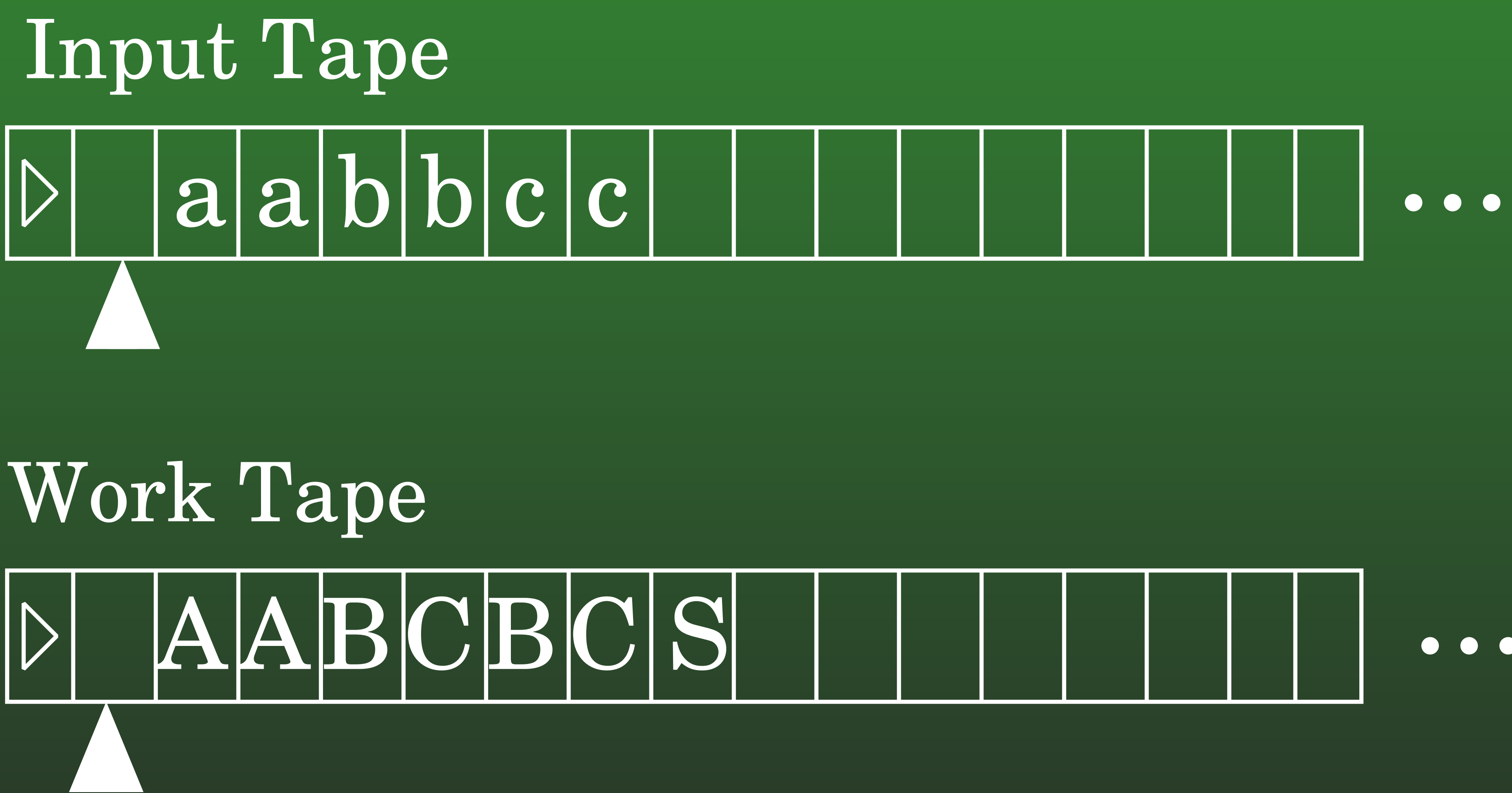
13-24: $L_{UG} \subseteq L_{re}$

Input Tape

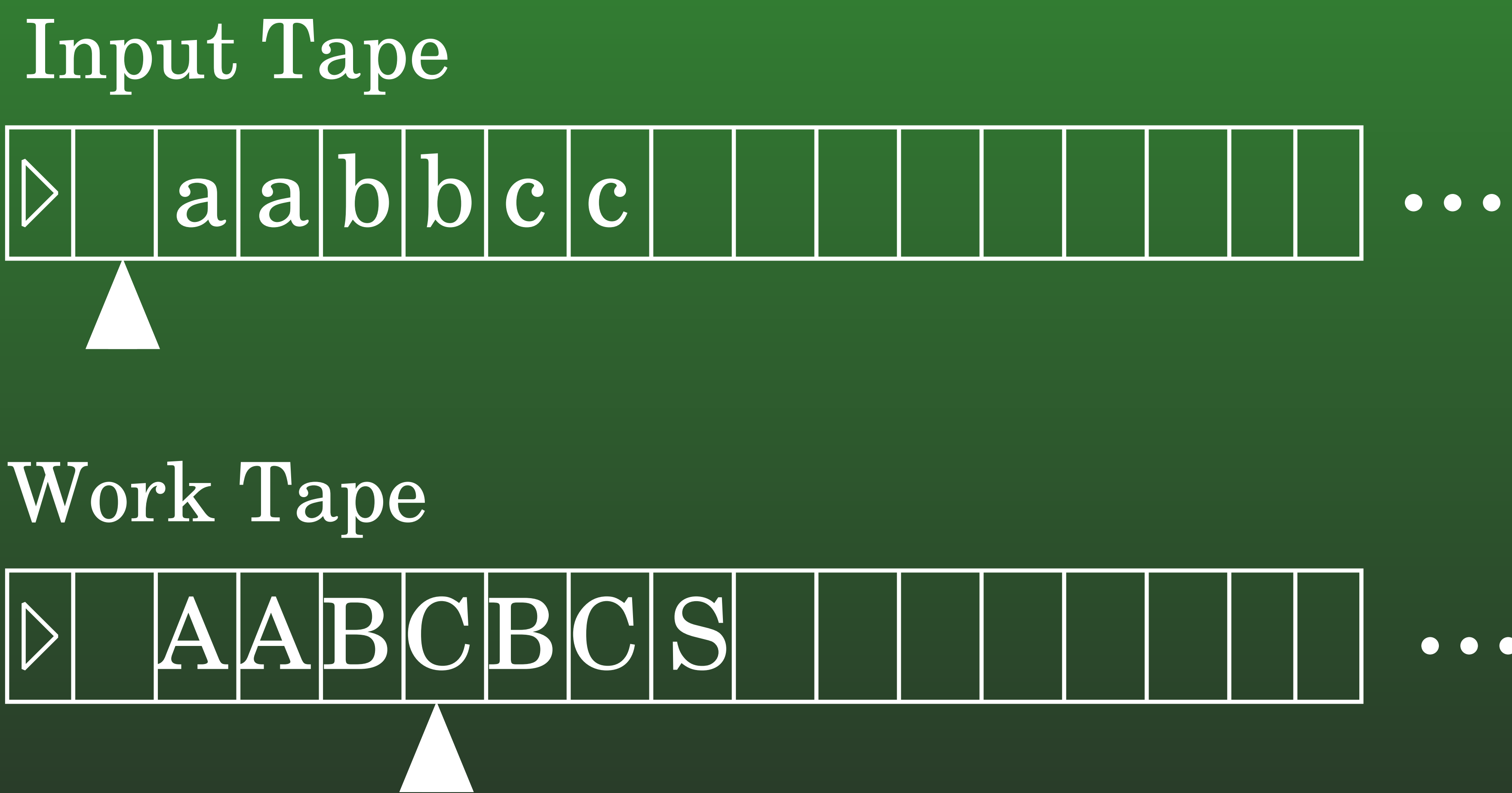


Work Tape

13-25: $L_{UG} \subseteq L_{re}$

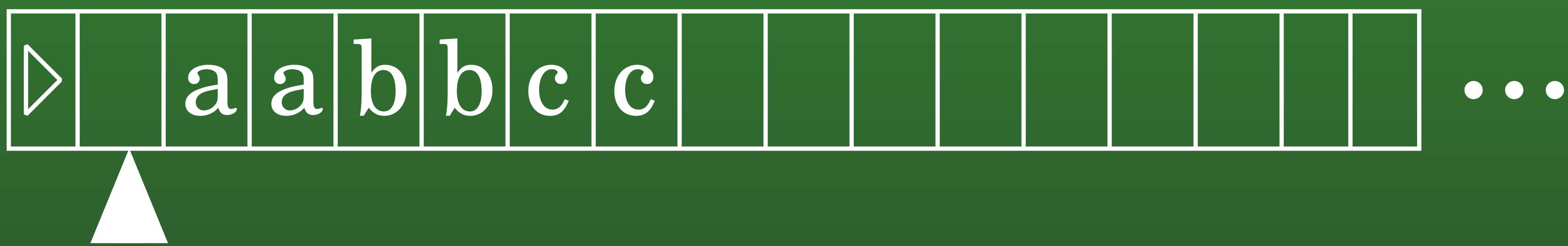


 $L_{UG} \subseteq L_{re}$



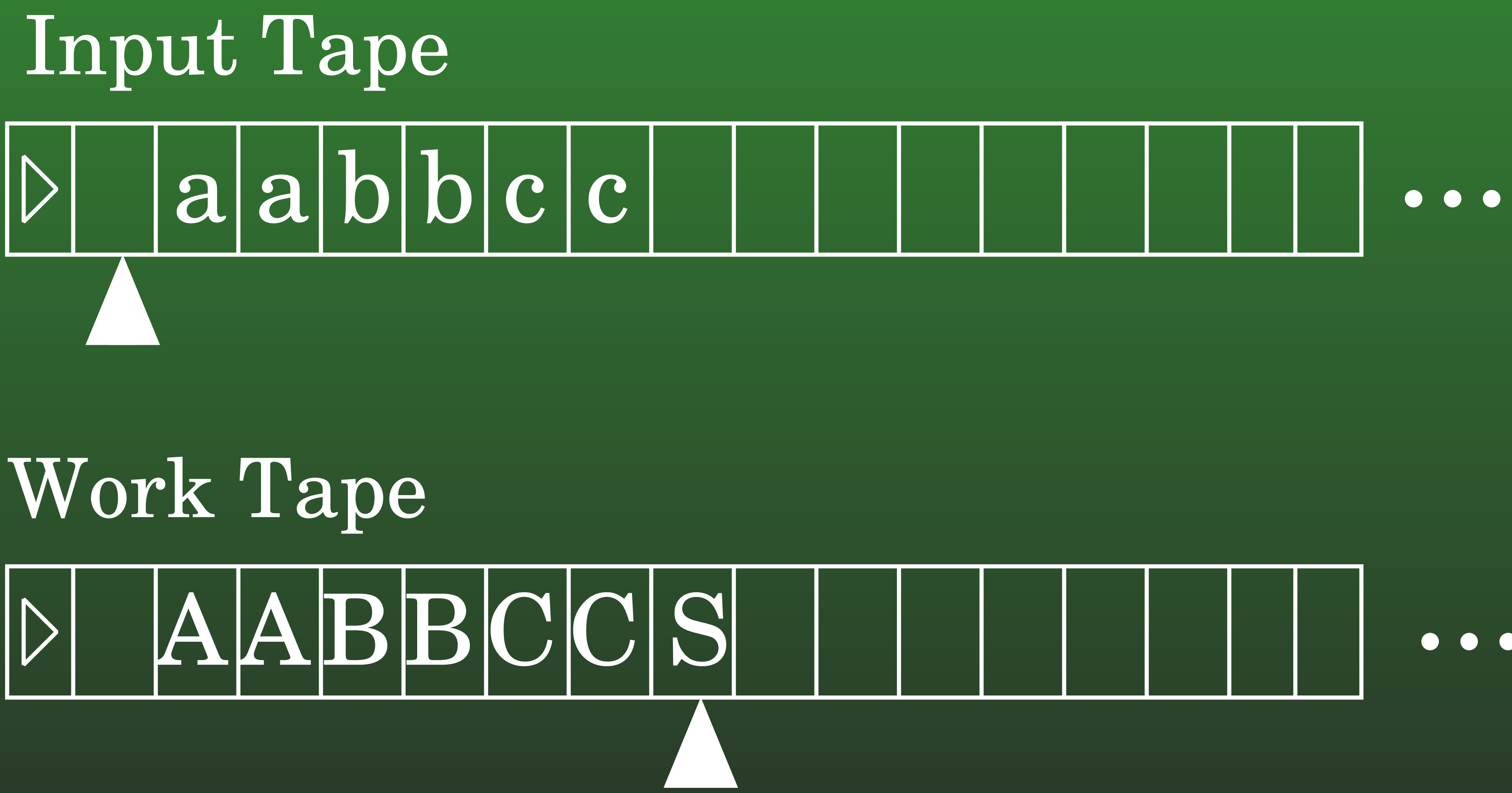
13-27: $L_{UG} \subseteq L_{re}$

Input Tape



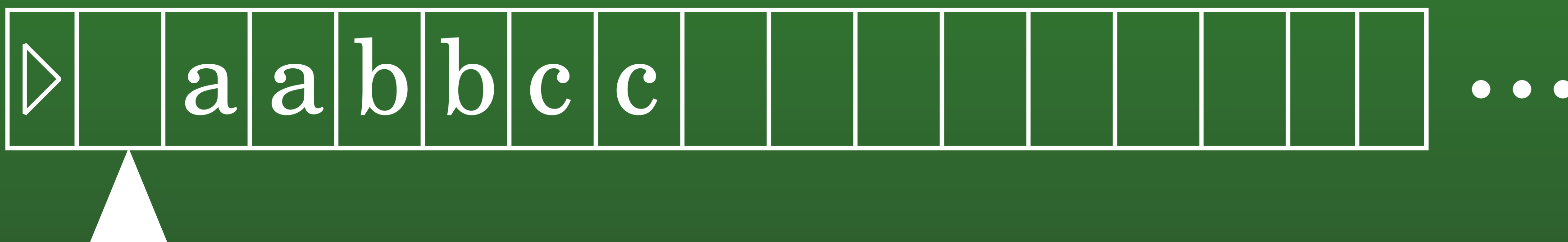
Work Tape

 $L_{UG} \subseteq L_{re}$



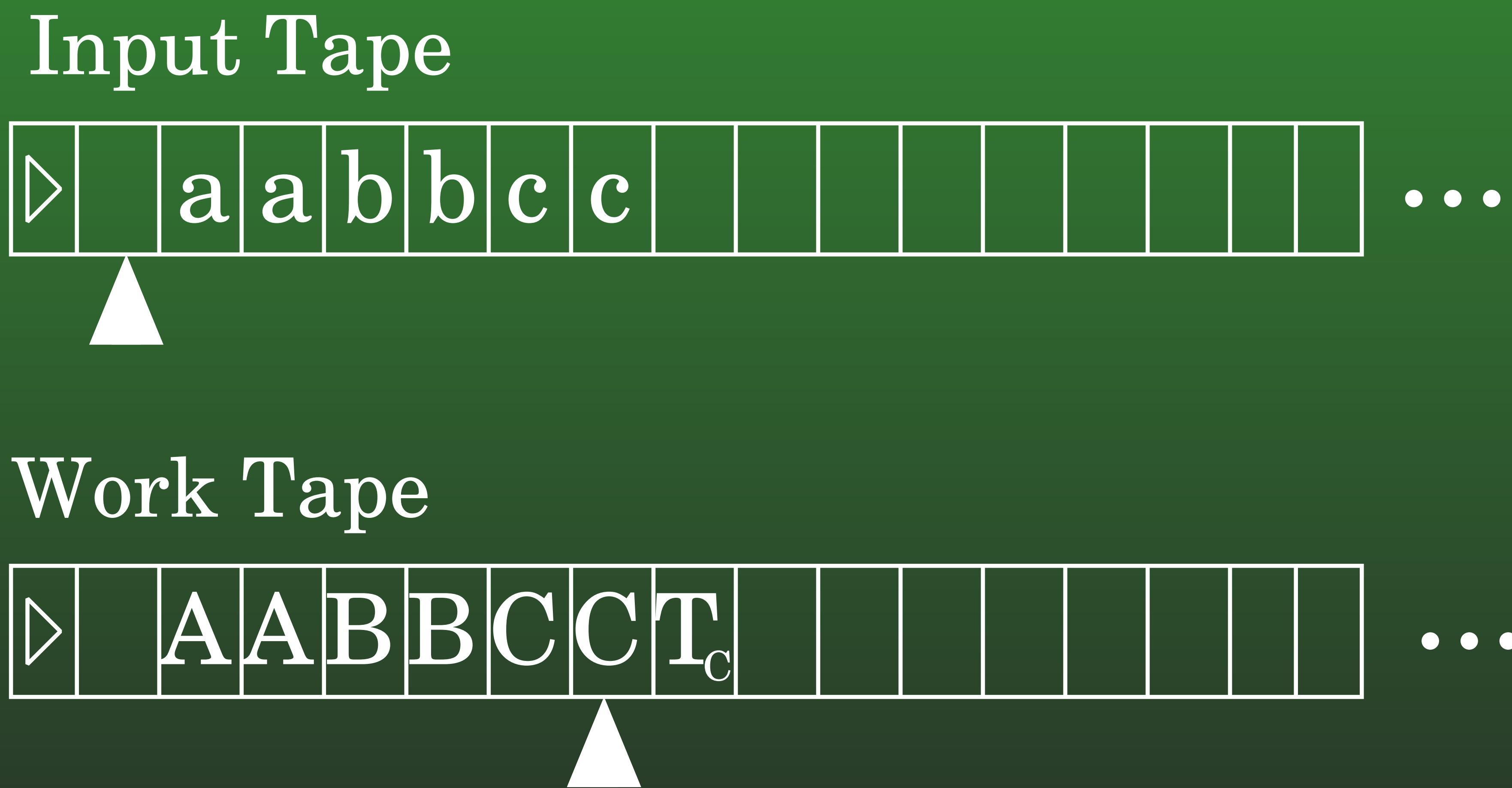
 $L_{UG} \subseteq L_{re}$

Input Tape

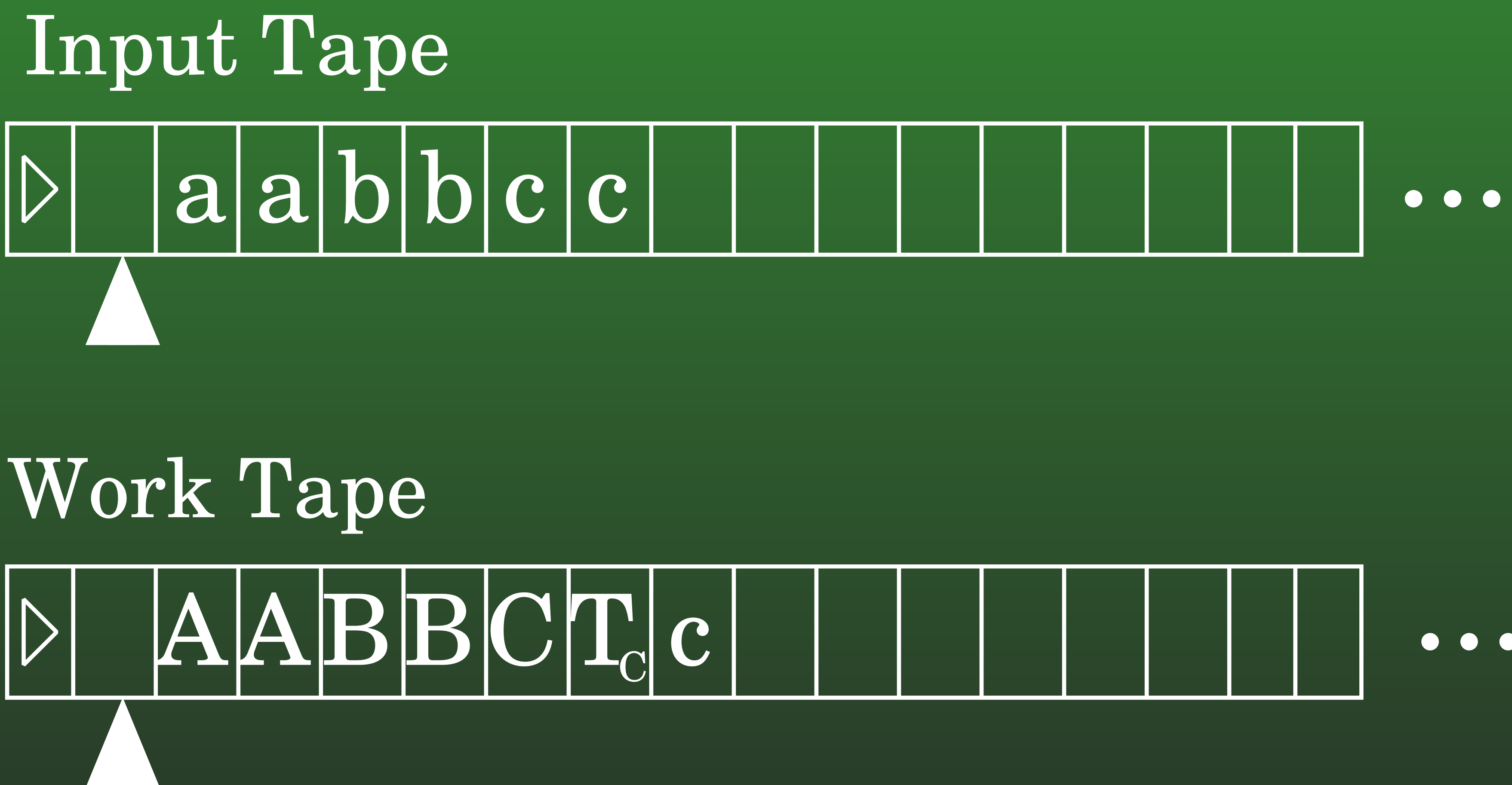


Work Tape

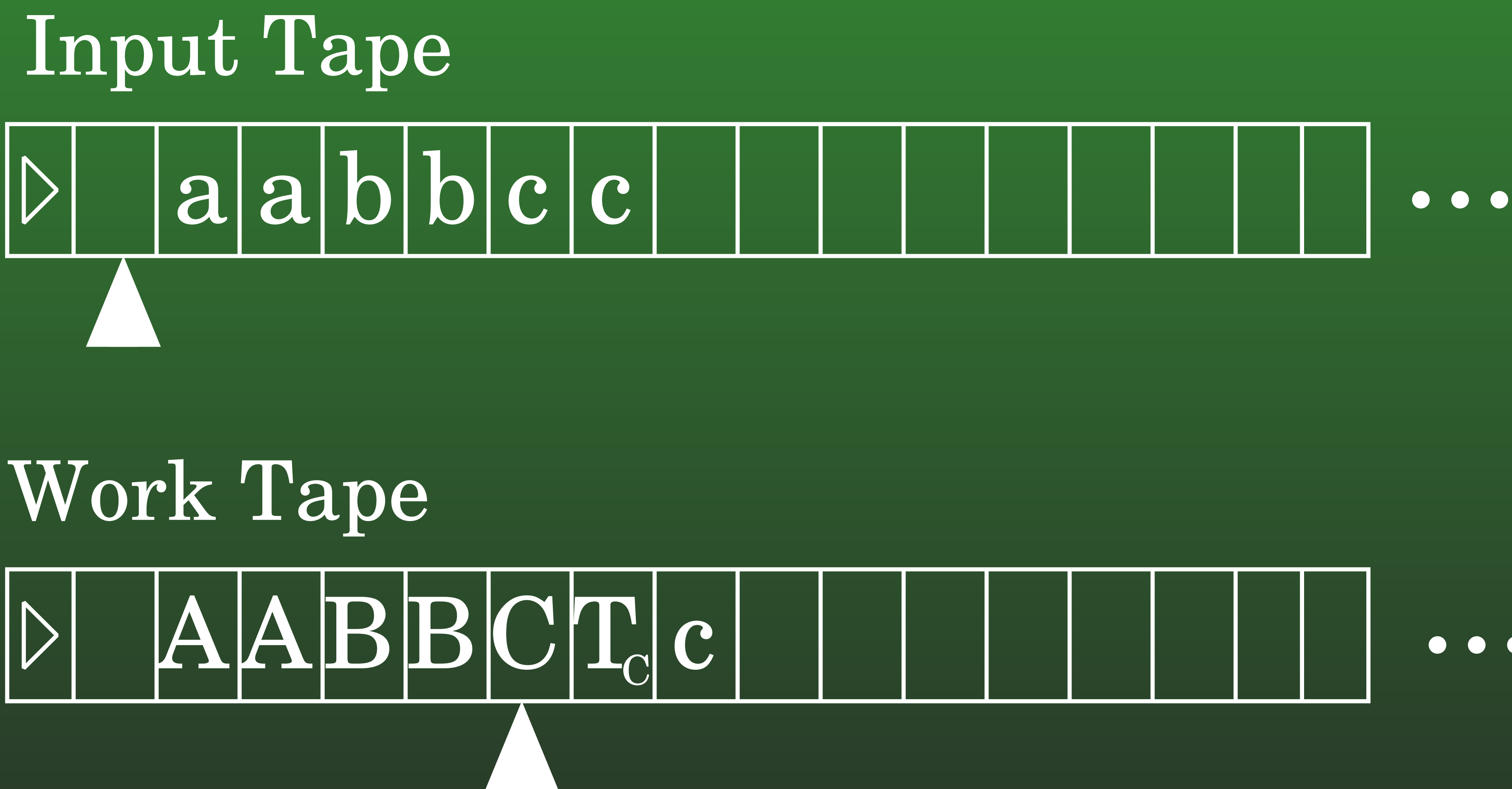
13-30: $L_{UG} \subseteq L_{re}$



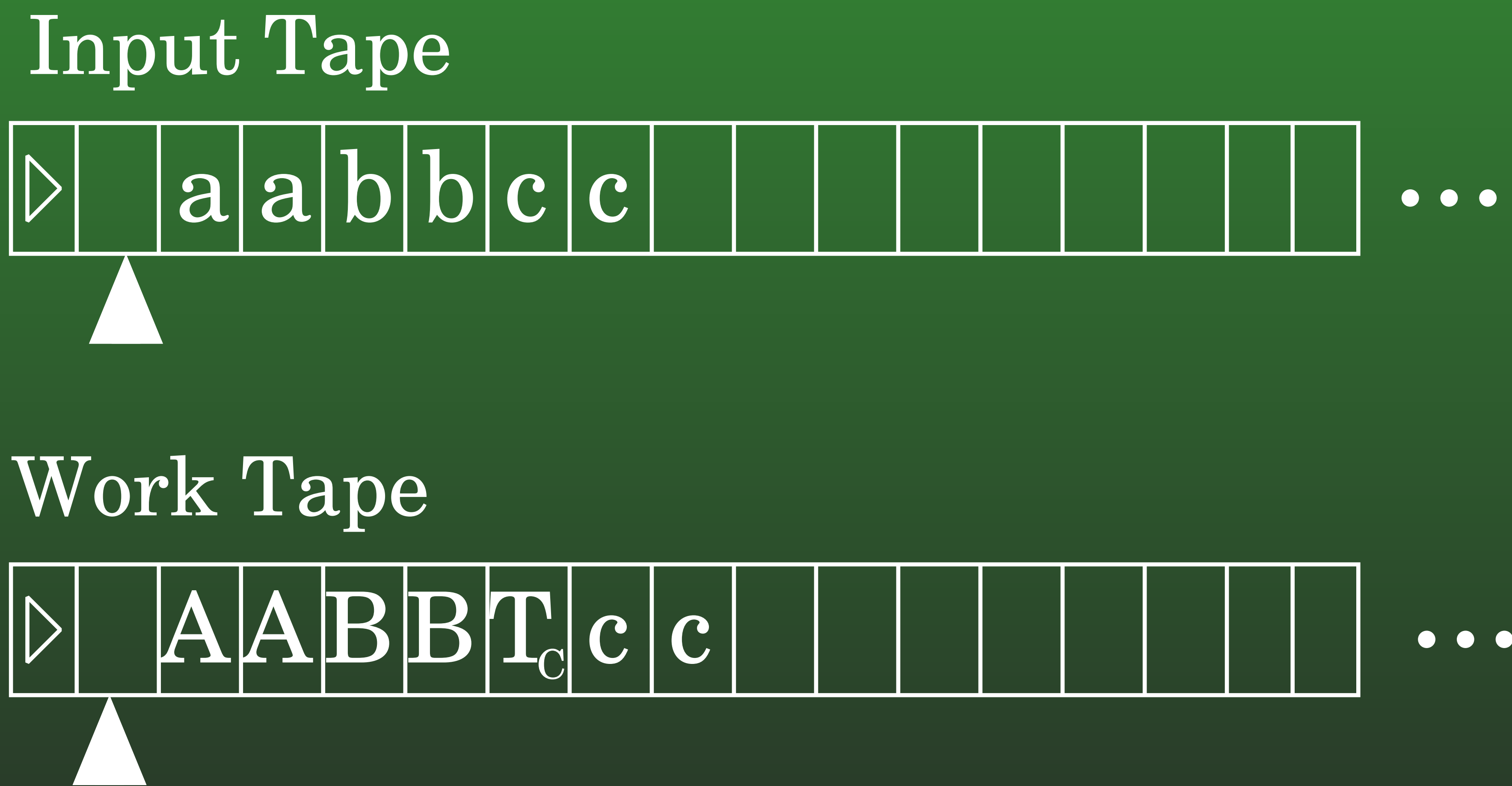
13-31: $L_{UG} \subseteq L_{re}$



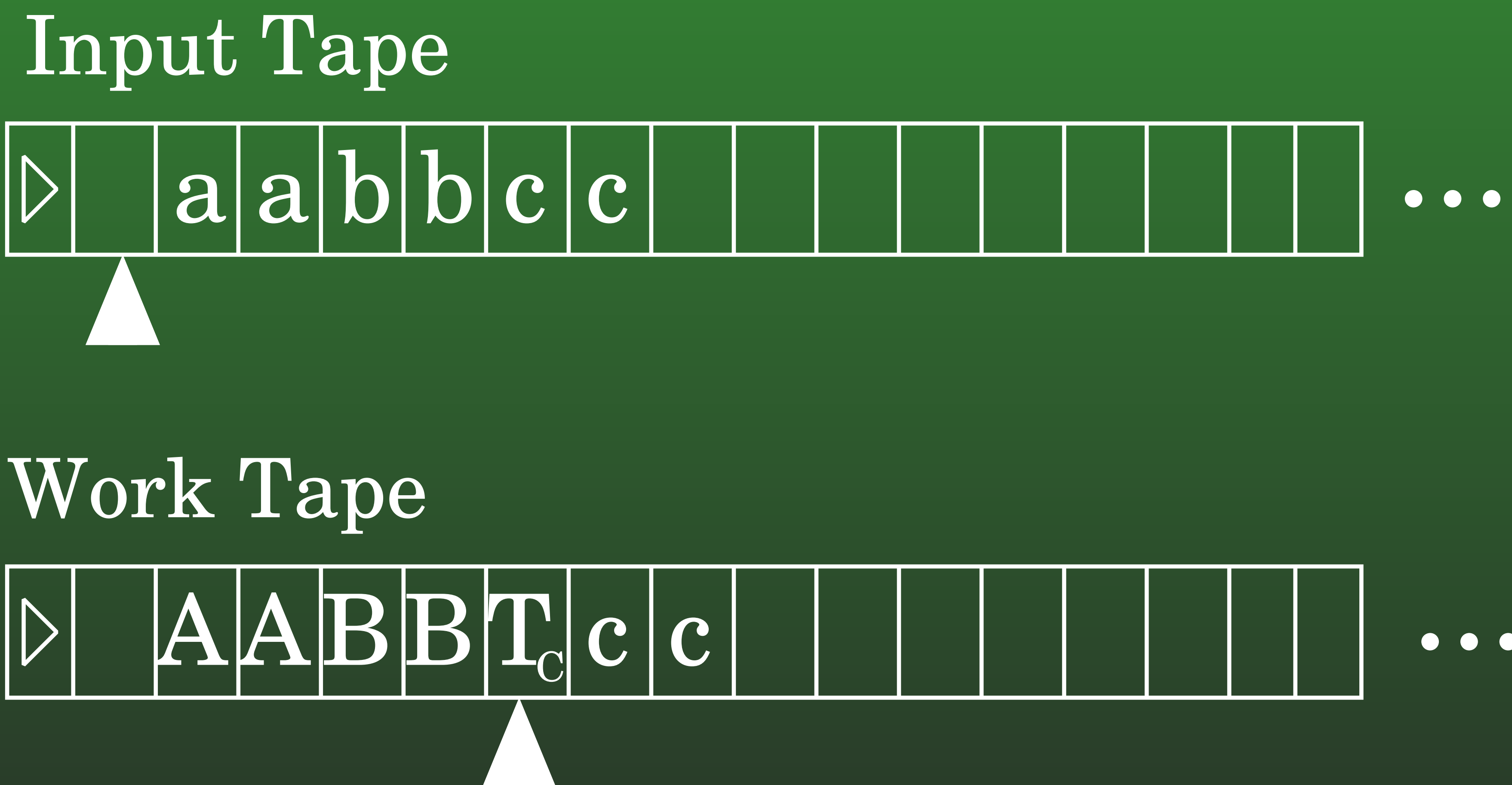
 $L_{UG} \subseteq L_{re}$



$L_{UG} \subseteq L_{re}$

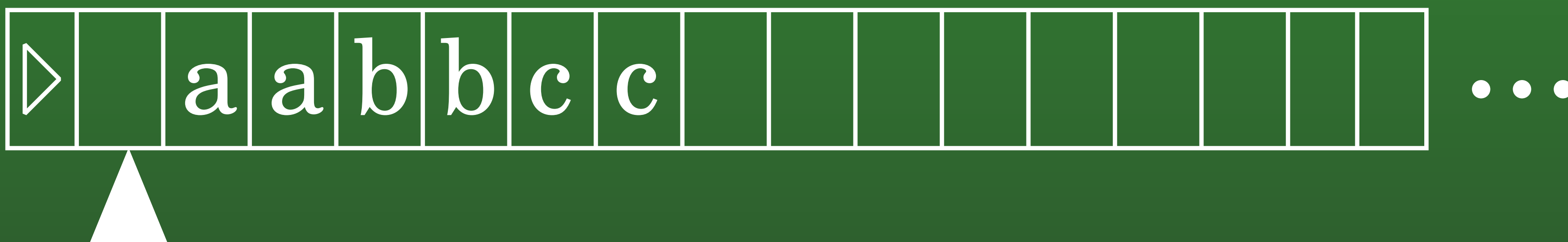


13-34: $L_{UG} \subseteq L_{re}$

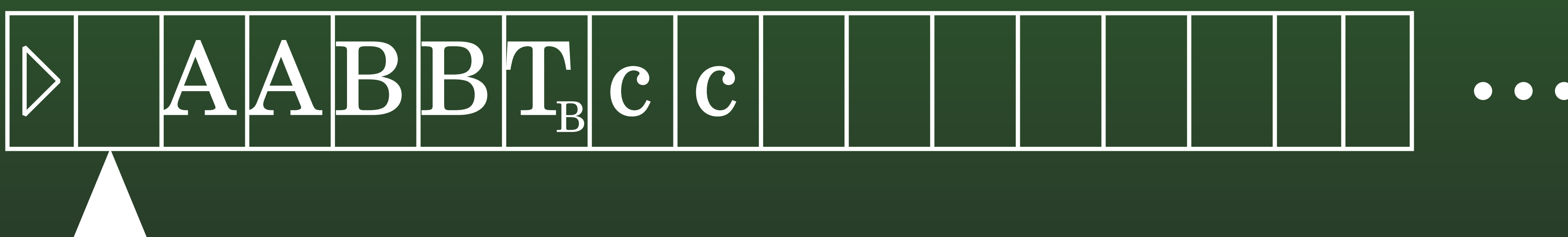


13-35: $L_{UG} \subseteq L_{re}$

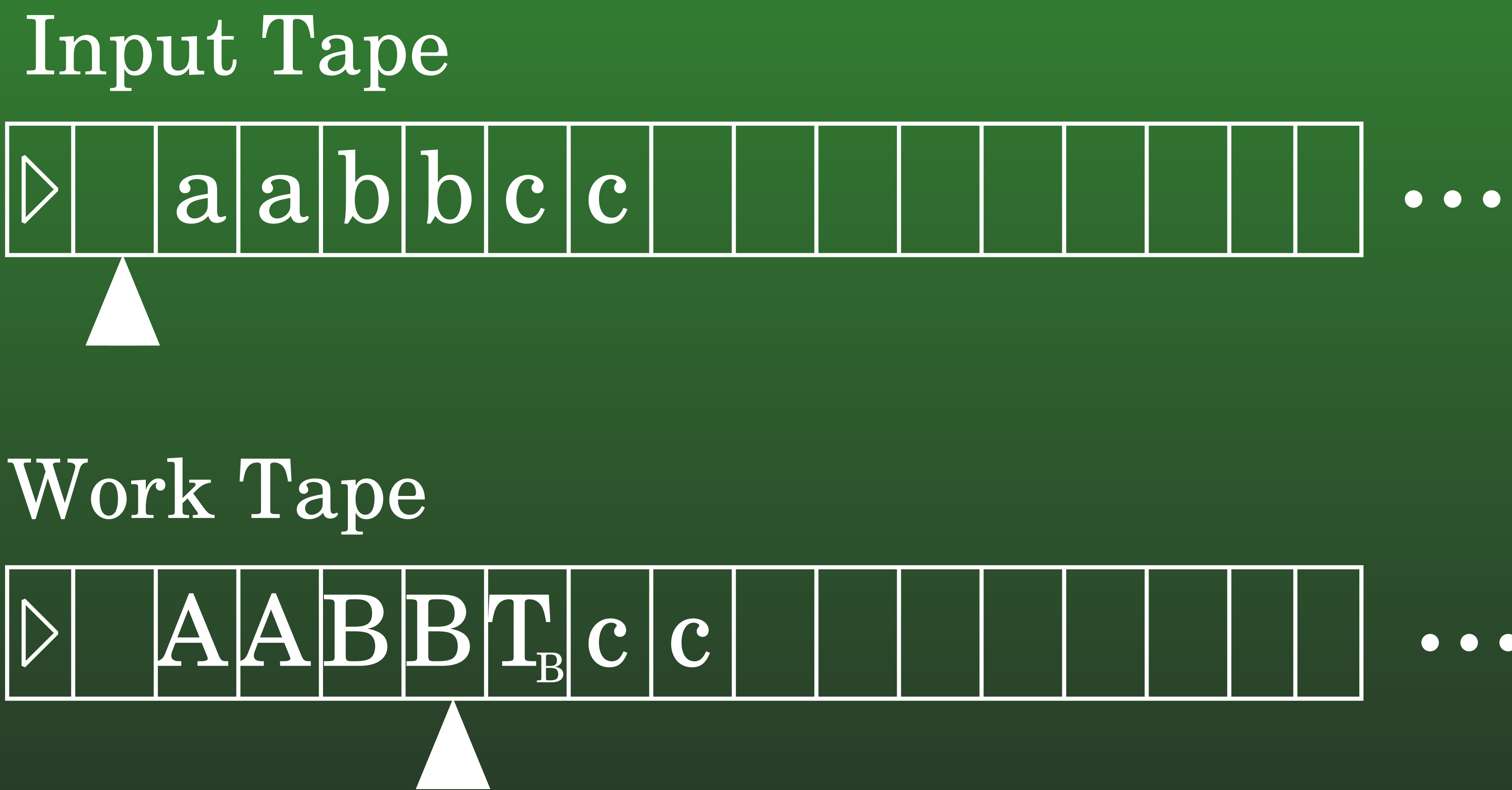
Input Tape



Work Tape

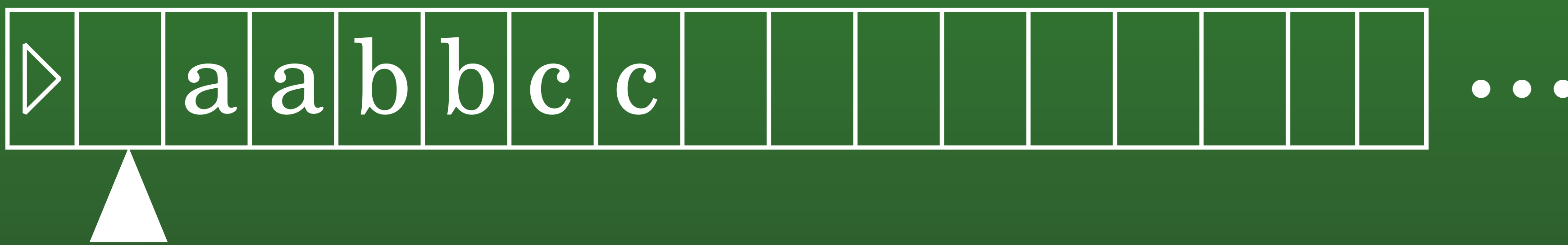


13-36: $L_{UG} \subseteq L_{re}$

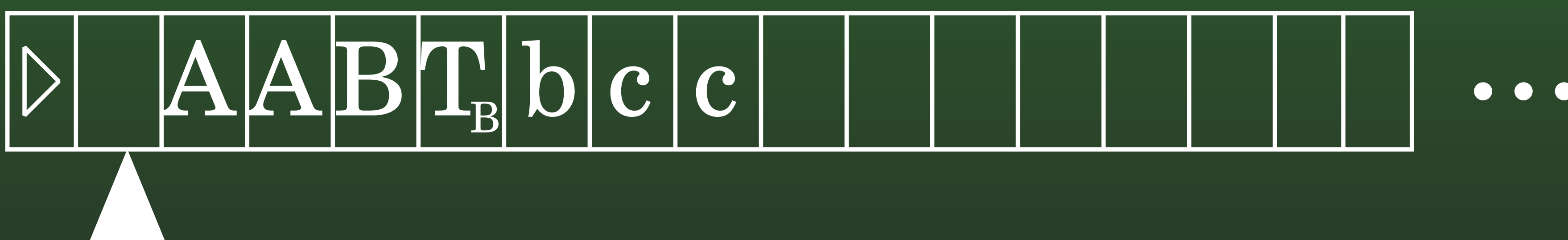


13-37: $L_{UG} \subseteq L_{re}$

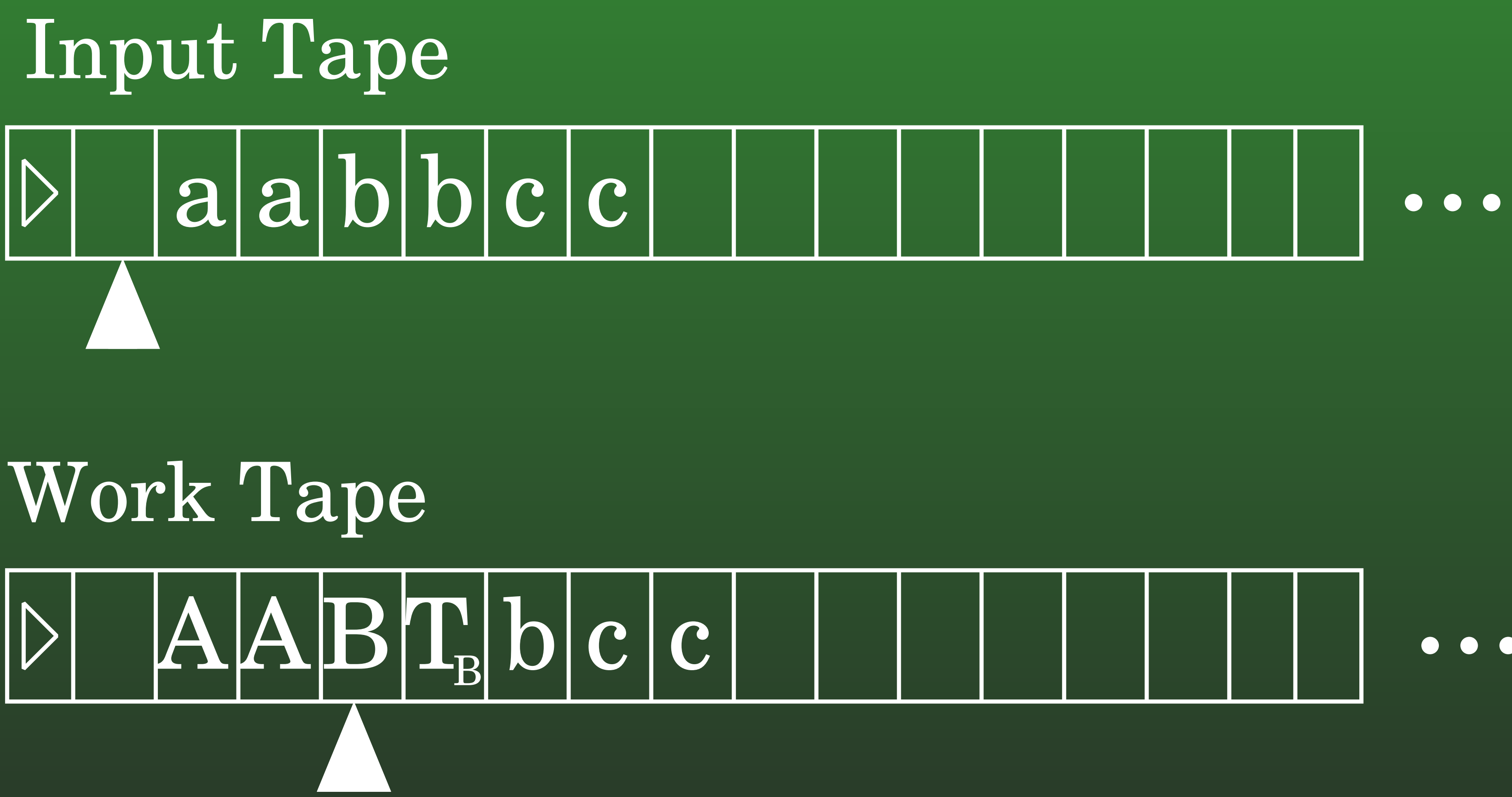
Input Tape



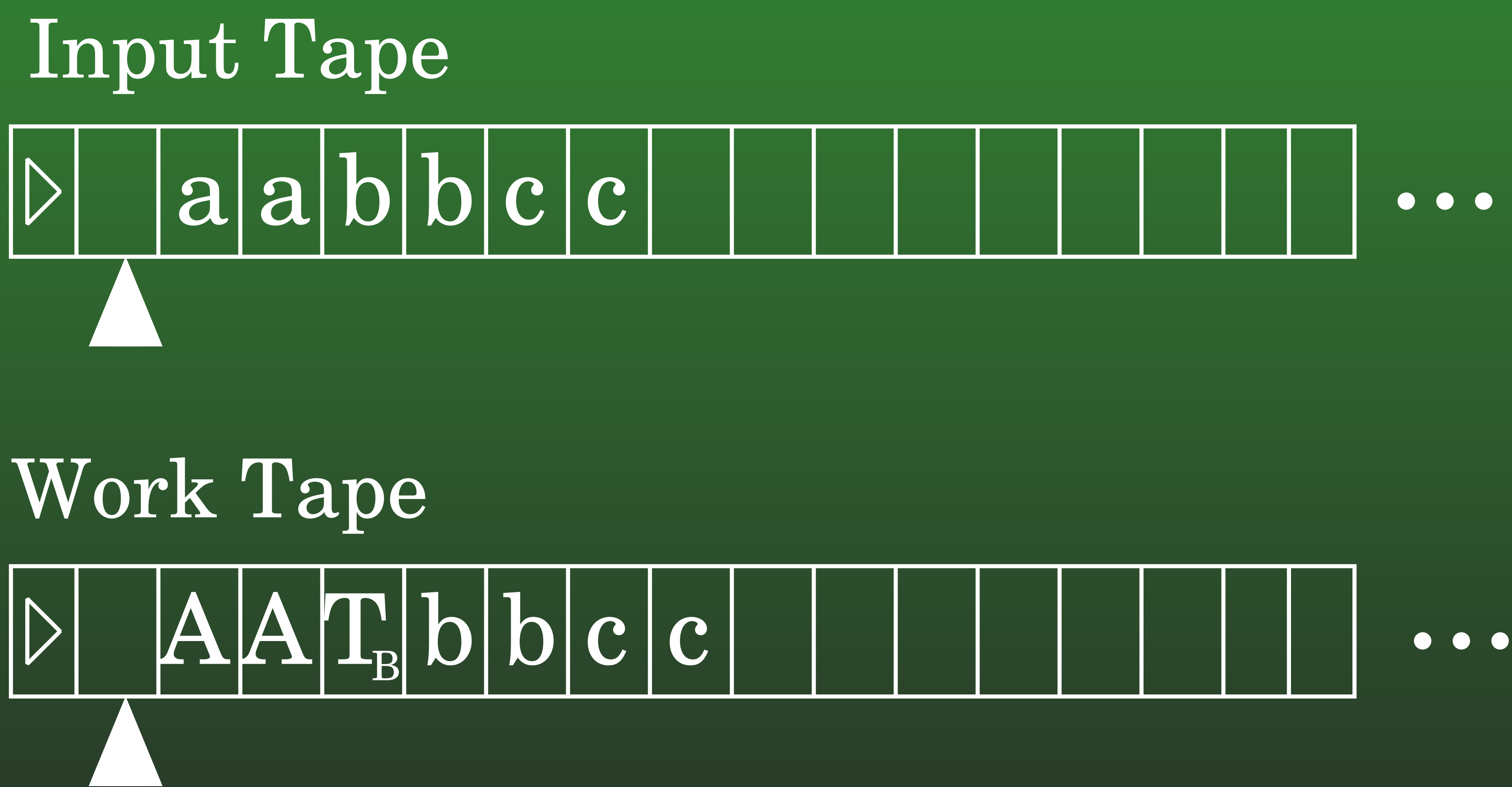
Work Tape



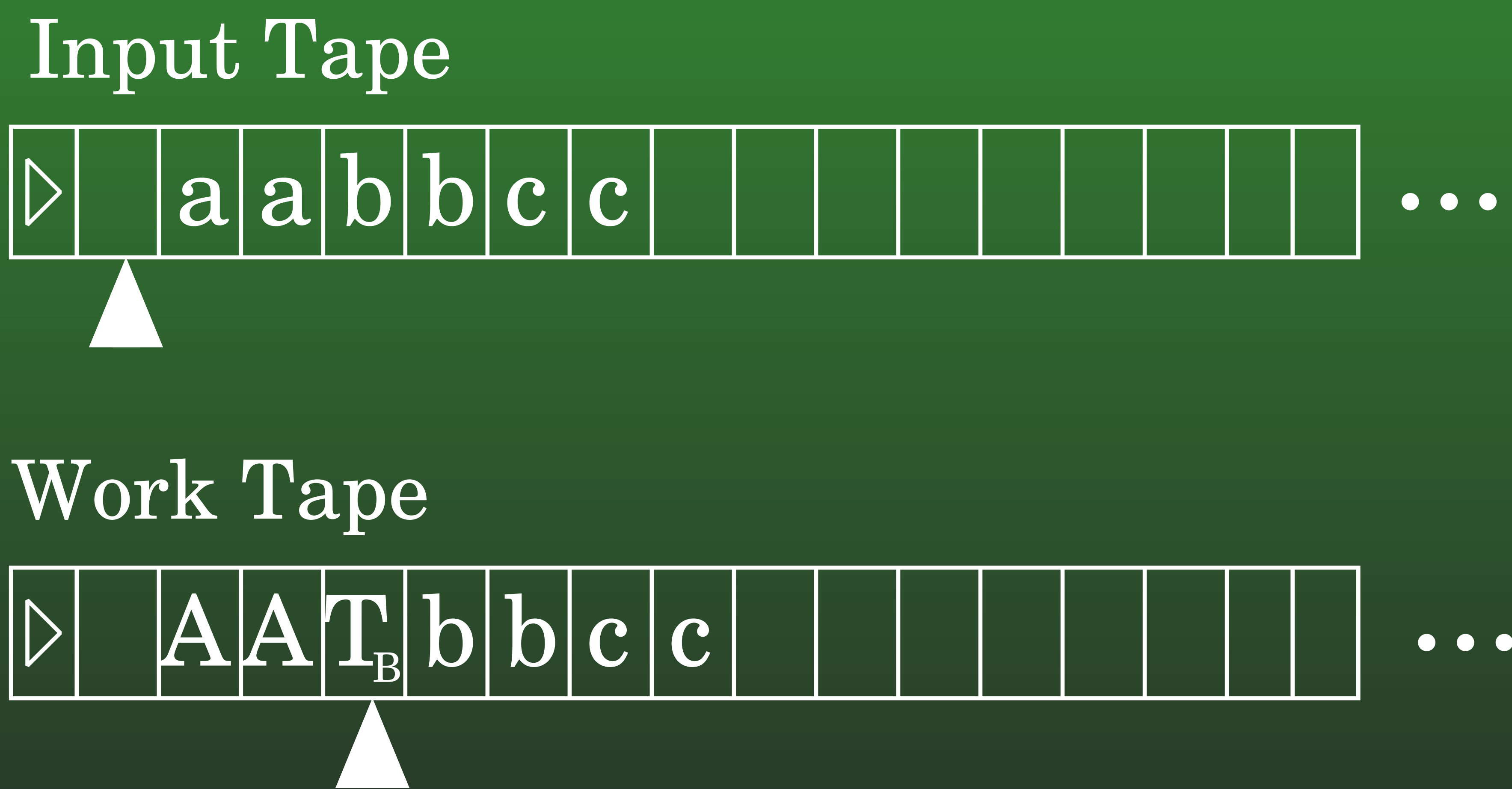
 $L_{UG} \subseteq L_{re}$



13-39: $L_{UG} \subseteq L_{re}$

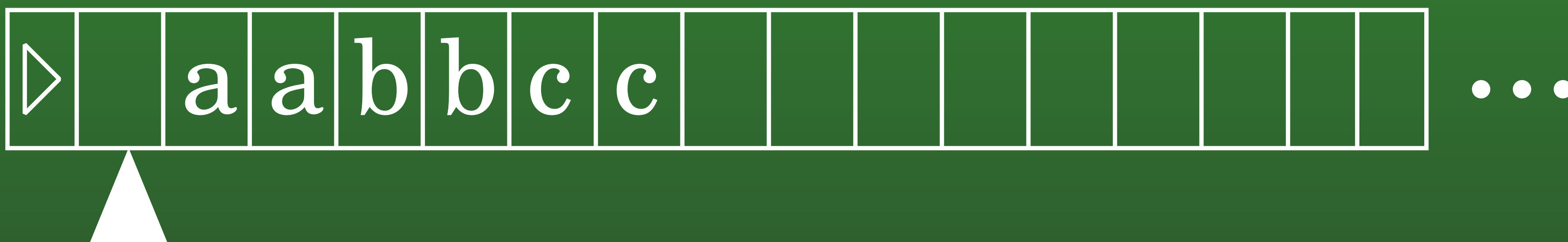


13-40: $L_{UG} \subseteq L_{re}$

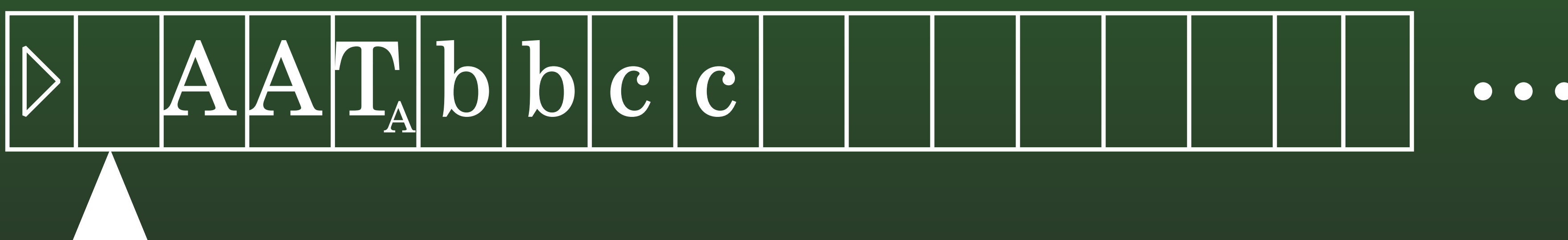


 $L_{UG} \subseteq L_{re}$

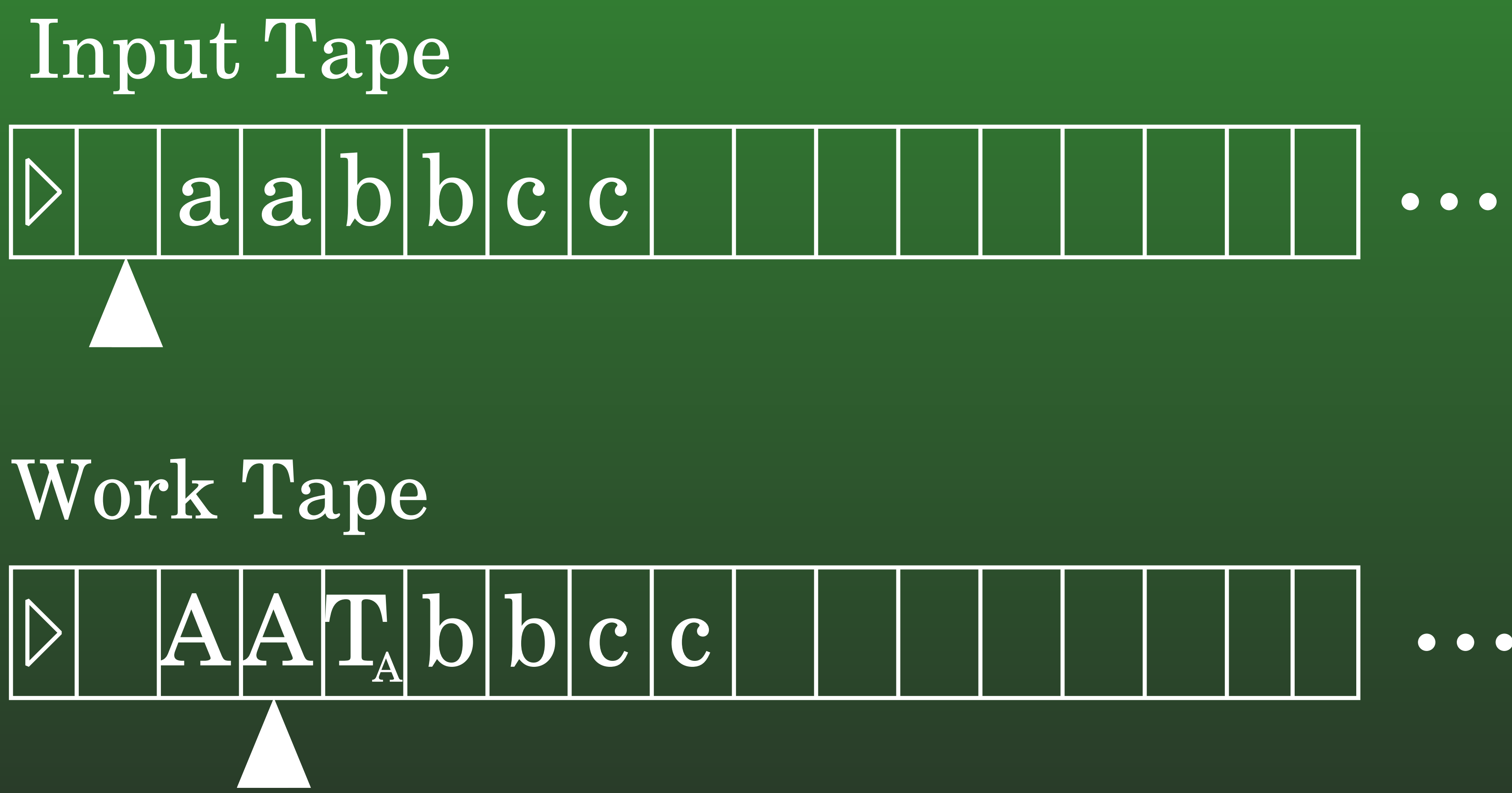
Input Tape



Work Tape

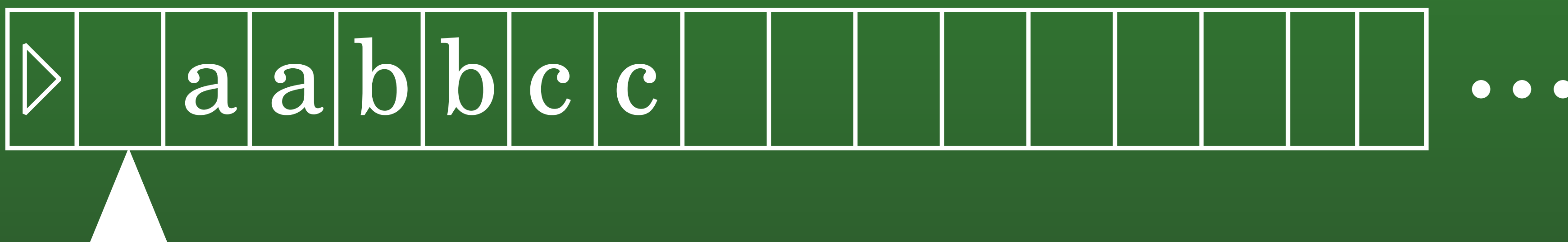


13-42: $L_{UG} \subseteq L_{re}$

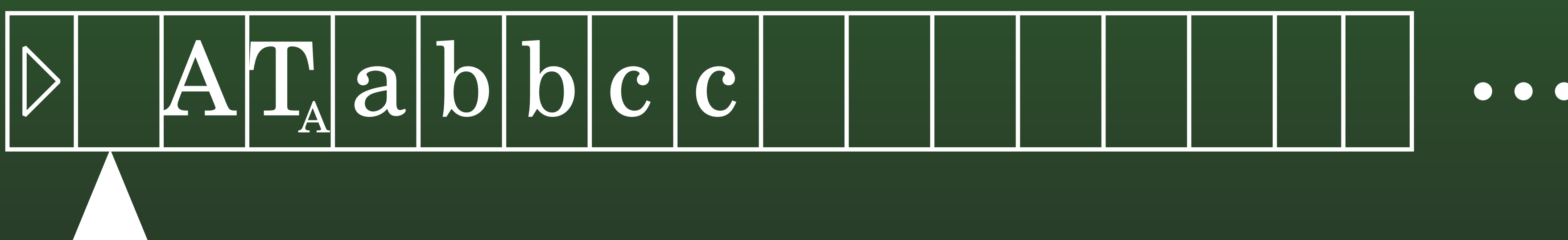


13-43: $L_{UG} \subseteq L_{re}$

Input Tape

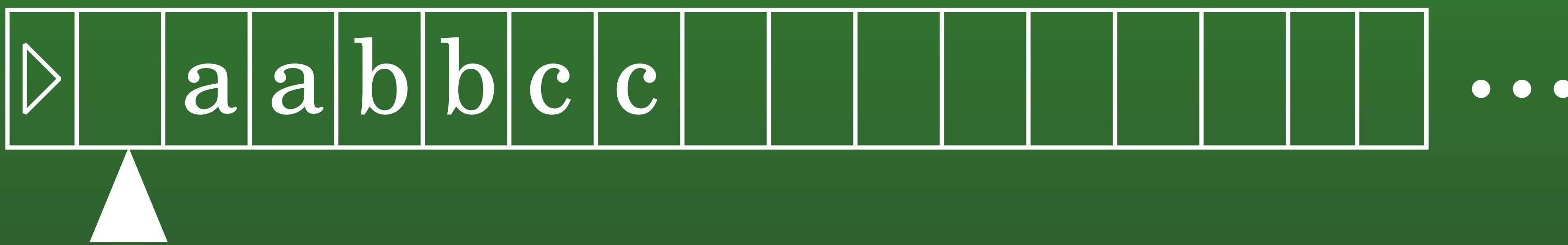


Work Tape

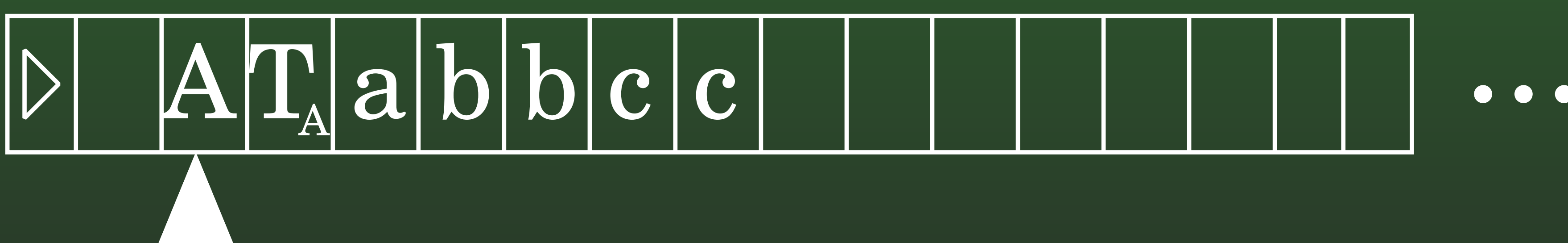


13-44: $L_{UG} \subseteq L_{re}$

Input Tape

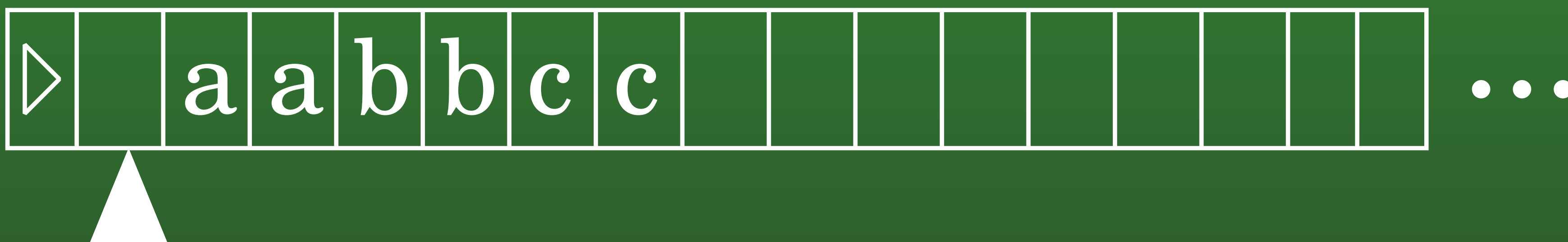


Work Tape

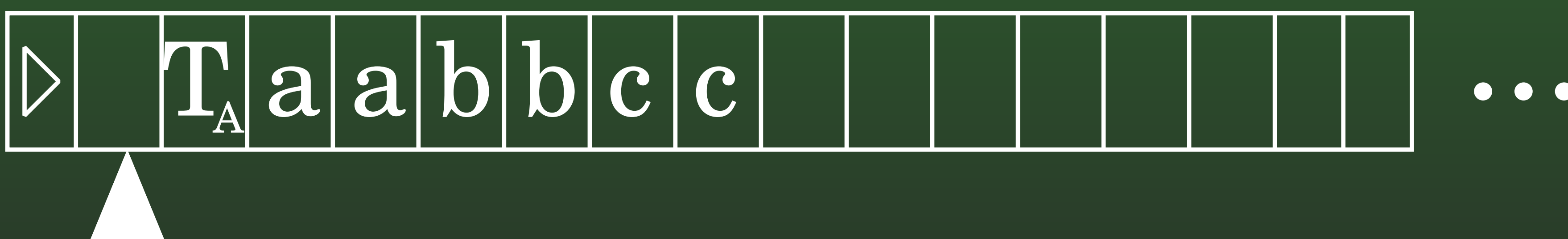


 $L_{UG} \subseteq L_{re}$

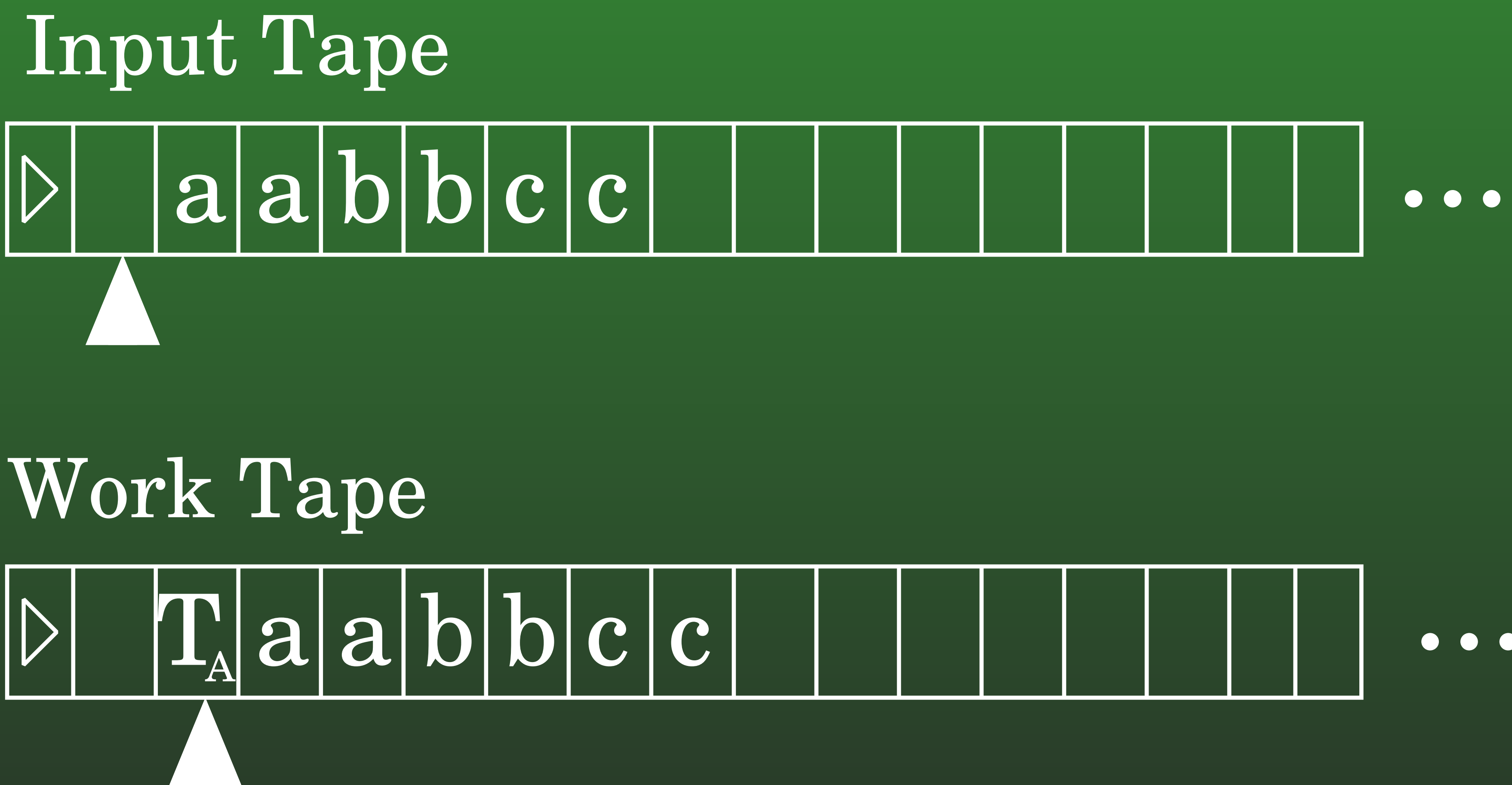
Input Tape



Work Tape

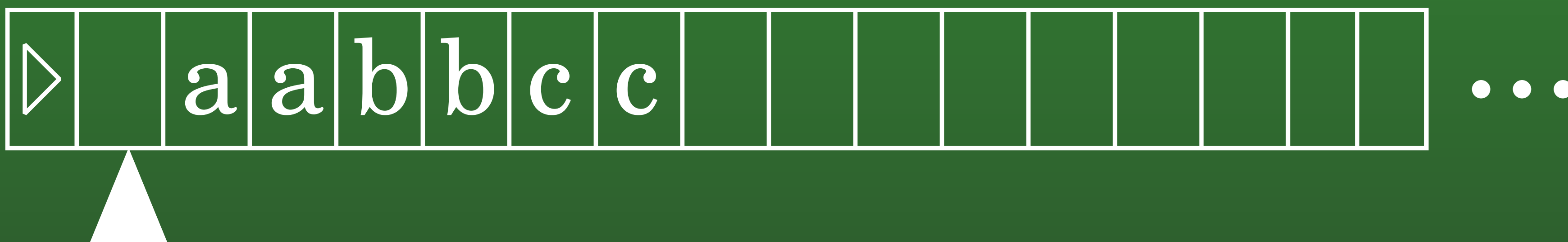


13-46: $L_{UG} \subseteq L_{re}$

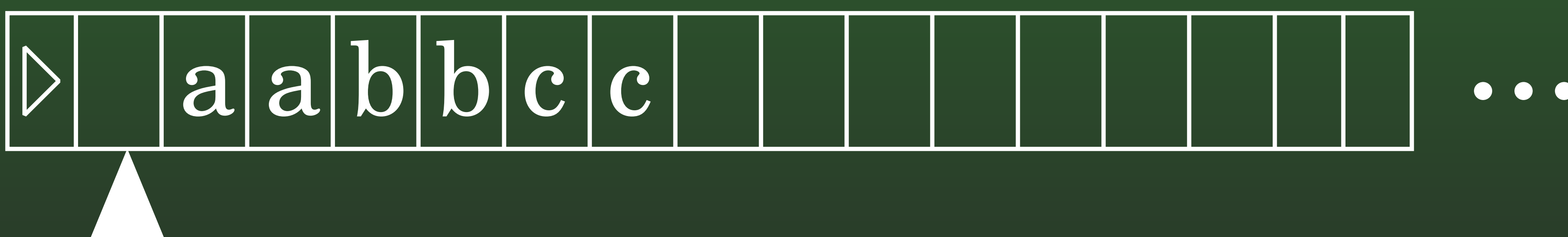


13-47: $L_{UG} \subseteq L_{re}$

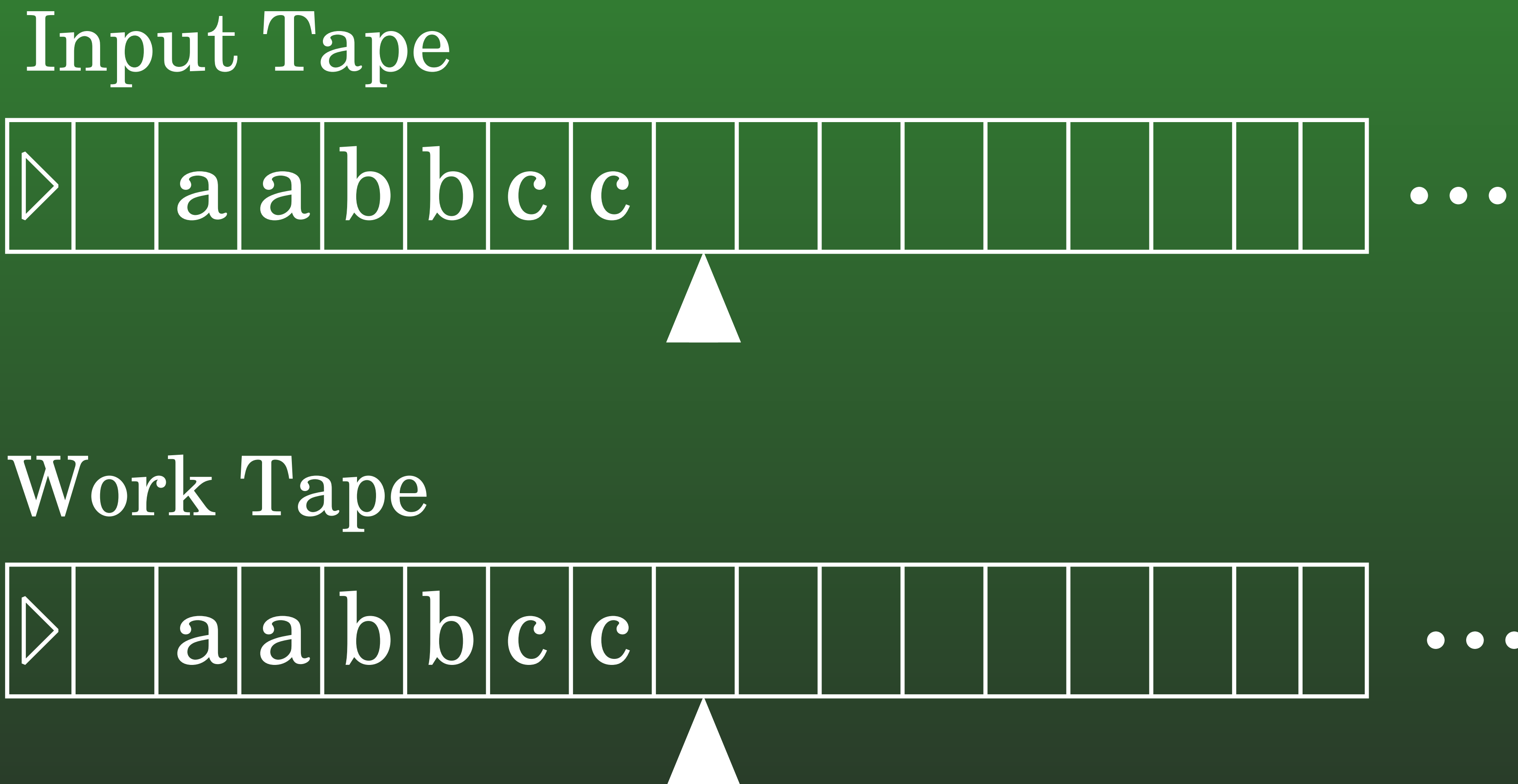
Input Tape



Work Tape



13-48: $L_{UG} \subseteq L_{re}$



13-49: $L_{re} \subseteq L_{UG}$

- Given any Turing Machine M that semi-decides the language L, we can create an Unrestricted Grammar G such that $L[G] = L$

13-50: $L_{re} \subseteq L_{UG}$

- Given any Turing Machine M that semi-decides the language L, we can create an Unrestricted Grammar G such that $L[G] = L$
  - Will assume that all Turing Machines accept in the same configuration: $(h, \triangleright\underline{\sqcup})$
  - Not a major restriction – why?

13-51: $L_{re} \subseteq L_{UG}$

- Given any Turing Machine M that semi-decides the language L, we can create an Unrestricted Grammar G such that $L[G] = L$
  - Will assume that all Turing Machines accept in the same configuration: $(h, \triangleright\underline{\sqcup})$
  - Not a major restriction – why?
  - Add a “tape erasing” machine right before the accepting state, that erases the tape, leaving the read/write head at the beginning of the tape

13-52: $L_{re} \subseteq L_{UG}$

- Given any Turing Machine M that semi-decides the language L, we can create an Unrestricted Grammar G such that $L[G] = L$
  - Grammar: Generates a string
  - Turing Machine: Works from string to accept state
- Two formalisms work in different directions
- Simulating Turing Machine with a Grammar can be difficult ..

13-53: $L_{re} \subseteq L_{UG}$

- Two formalisms work in different directions
  - Simulate a Turing Machine – in reverse!
  - Each partial derivation represents a configuration
  - Each rule represents a *backwards* step in Turing Machine computation

13-54: $L_{re} \subseteq L_{UG}$

- Given a TM M, we create a Grammar G:
  - Language for G:
    - Everything in Σ_M
    - Everything in K_M
    - Start symbol S
    - Symbols $\triangleright$ and $\triangleleft$

13-55: $L_{re} \subseteq L_{UG}$

- Configuration $(Q, \triangleright u\underline{a}w)$ represented by the string: $\triangleright uaQw\triangleleft$

For example, $(Q, \triangleright \sqcup ab\underline{c} \sqcup a)$ is represented by the string $\triangleright \sqcup abcQ \sqcup a\triangleleft$

13-56: $L_{re} \subseteq L_{UG}$

- For each element in δ_M of the form:
  - $((Q_1, a), (Q_2, b))$
- Add the rule:
  - $bQ_2 \rightarrow aQ_1$
- Remember, simulating *backwards* computation

13-57: $L_{re} \subseteq L_{UG}$

- For each element in δ_M of the form:
  - $((Q_1, a), (Q_2, \leftarrow))$
- Add the rule:
  - $Q_2 a \rightarrow a Q_1$

 $L_{re} \subseteq L_{UG}$

- For each element in δ_M of the form:
  - $((Q_1, \sqcup), (Q_2, \leftarrow))$
- Add the rule
  - $Q_2 \triangleleft \rightarrow \sqcup Q_1 \triangleleft$
- (undoing erasing extra blanks)

13-59: $L_{re} \subseteq L_{UG}$

- For each element in δ_M of the form:
  - $((Q_1, a), (Q_2, \rightarrow))$
- Add the rule
  - $abQ_2 \rightarrow aQ_1b$
- For all $b \in \Sigma$

13-60: $L_{re} \subseteq L_{UG}$

- For each element in δ_M of the form:
  - $((Q_1, a), (Q_2, \rightarrow))$
- Add the rule
  - $a \sqcup Q_2 \triangleleft \rightarrow aQ_1 \triangleleft$
- (undoing moving to the right onto unused tape)

13-61: $L_{re} \subseteq L_{UG}$

- Finally, add the rules:
  - $S \rightarrow \triangleright \sqcup h \triangleleft$
  - $\triangleright \sqcup Q_s \rightarrow \epsilon$
  - $\triangleleft \rightarrow \epsilon$

13-62: $L_{re} \subseteq L_{UG}$

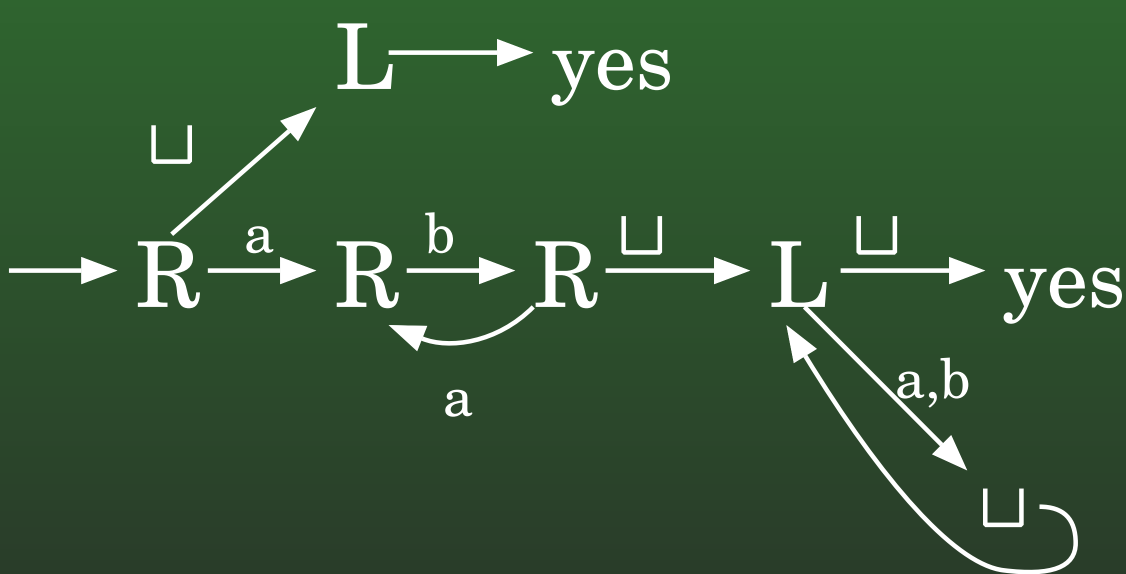
- If the Turing machine can move from
  - $\triangleright\underline{\sqcup}w$ to $\triangleright h\underline{\sqcup}$
- Then the Grammar can transform
  - $\triangleright \sqcup Q_h \triangleleft$ to $\triangleright \sqcup Q_s w \triangleleft$
- Then, remove $\triangleright \sqcup Q_s$ and $\triangleleft$ to leave w

13-63: $L_{re} \subseteq L_{UG}$

- Example:
  - Create a Turing Machine that accepts $(ab)^*$, halting in the configuration $(h, \triangleright\underline{\sqcup})$
  - (assume tape starts out as $\triangleright\underline{\sqcup}w$)

13-64: $L_{re} \subseteq L_{UG}$

- Example:
  - Create a Turing Machine that accepts $(ab)^*$, halting in the configuration $(h, \triangleright\underline{\sqcup})$



13-65: $L_{re} \subseteq L_{UG}$

	a	b	$\sqcup$
q_0	$(q_1, \rightarrow)$	$(q_1, \rightarrow)$	$(q_1, \rightarrow)$
q_1	$(q_2, \rightarrow)$		$(q_h, \leftarrow)$
q_2		$(q_3, \rightarrow)$	
q_3	$(q_2, \rightarrow)$		$(q_4, \leftarrow)$
q_4	$(q_5, \sqcup)$	$(q_5, \sqcup)$	$(q_h, \sqcup)$
q_5			$(q_4, \leftarrow)$

13-66: $L_{re} \subseteq L_{UG}$

- $((q_0, a), (q_1, \rightarrow))$
  - $aaQ_1 \rightarrow aQ_0a$
  - $abQ_1 \rightarrow aQ_0b$
  - $a \sqcup Q_1 \rightarrow aQ_0\sqcup$
  - $a \sqcup Q_1\triangleleft \rightarrow aQ_0\triangleleft$

13-67: $L_{re} \subseteq L_{UG}$

- $((q_0, b), (q_1, \rightarrow))$
  - $baQ_1 \rightarrow bQ_0a$
  - $bbQ_1 \rightarrow bQ_0b$
  - $b \sqcup Q_1 \rightarrow bQ_0\sqcup$
  - $b \sqcup Q_1\triangleleft \rightarrow bQ_0\triangleleft$

13-68: $L_{re} \subseteq L_{UG}$

- $((q_0, \sqcup), (q_1, \rightarrow))$
  - $\sqcup a Q_1 \rightarrow \sqcup Q_0 a$
  - $\sqcup b Q_1 \rightarrow \sqcup Q_0 b$
  - $\sqcup \sqcup Q_1 \rightarrow \sqcup Q_0 \sqcup$
  - $\sqcup \sqcup Q_1 \triangleleft \rightarrow \sqcup Q_0 \triangleleft$

13-69: $L_{re} \subseteq L_{UG}$

- $((q_1, a), (q_2, \rightarrow))$
  - $aaQ_2 \rightarrow aQ_1a$
  - $abQ_2 \rightarrow aQ_1b$
  - $a \sqcup Q_2 \rightarrow aQ_1\sqcup$
  - $a \sqcup Q_2\triangleleft \rightarrow aQ_1\triangleleft$

13-70: $L_{re} \subseteq L_{UG}$

- $((q_1, \sqcup), (q_h, \leftarrow))$
  - $h\sqcup \rightarrow \sqcup Q_1$

13-71: $L_{re} \subseteq L_{UG}$

- $((q_2, b), (q_3, \to))$
  - $baQ_3 \to bQ_2a$
  - $bbQ_3 \to bQ_2b$
  - $b \sqcup Q_3 \to bQ_2\sqcup$
  - $b \sqcup Q_3\triangleleft \to bQ_2\triangleleft$

13-72: $L_{re} \subseteq L_{UG}$

- $((q_3, a), (q_4, \rightarrow))$
  - $aaQ_4 \rightarrow aQ_3a$
  - $abQ_4 \rightarrow aQ_3b$
  - $a \sqcup Q_4 \rightarrow aQ_3\sqcup$
  - $a \sqcup Q_4 \triangleleft \rightarrow aQ_3 \triangleleft$

13-73: $L_{re} \subseteq L_{UG}$

- $((q_4, a), (q_5, \sqcup))$
  - $\sqcup Q_5 \rightarrow aQ_4$
- $((q_4, b), (q_5, \sqcup))$
  - $\sqcup Q_5 \rightarrow bQ_4$
- $((q_4, \sqcup), (q_h, \sqcup))$
  - $\sqcup h \rightarrow \sqcup Q_4$
- $((q_5, \sqcup), (q_4, \leftarrow))$
  - $Q_4 \sqcup \rightarrow \sqcup Q_5$

13-74: $L_{re} \subseteq L_{UG}$

$S \rightarrow \triangleright \sqcup h \triangleleft$

$\triangleright \sqcup Q_0 \rightarrow \epsilon$

$\triangleleft \rightarrow \epsilon$

$aaQ_1 \rightarrow aQ_0a$

$abQ_1 \rightarrow aQ_0b$

$a \sqcup Q_1 \rightarrow aQ_0\sqcup$

$a \sqcup Q_1\triangleleft \rightarrow aQ_0\triangleleft$

$baQ_1 \rightarrow bQ_0a$

$bbQ_1 \rightarrow bQ_0b$

$b \sqcup Q_1 \rightarrow bQ_0\sqcup$

$b \sqcup Q_1\triangleleft \rightarrow bQ_0\triangleleft$

$\sqcup aQ_1 \rightarrow \sqcup Q_0a$

$\sqcup bQ_1 \rightarrow \sqcup Q_0b$

$\sqcup \sqcup Q_1 \rightarrow \sqcup Q_0\sqcup$

$\sqcup \sqcup Q_1\triangleleft \rightarrow \sqcup Q_0\triangleleft$

$aaQ_2 \rightarrow aQ_1a$

$abQ_2 \rightarrow aQ_1b$

$a \sqcup Q_2 \rightarrow aQ_1\sqcup$

$a \sqcup Q_2\triangleleft \rightarrow aQ_1\triangleleft$

$h\sqcup \rightarrow \sqcup Q_1$

$baQ_3 \rightarrow bQ_2a$

$bbQ_3 \rightarrow bQ_2b$

$b \sqcup Q_3 \rightarrow bQ_2\sqcup$

$b \sqcup Q_3\triangleleft \rightarrow bQ_2\triangleleft$

$aaQ_4 \rightarrow aQ_3a$

$abQ_4 \rightarrow aQ_3b$

$a \sqcup Q_4 \rightarrow aQ_3\sqcup$

$a \sqcup Q_4\triangleleft \rightarrow aQ_3\triangleleft$

$\sqcup Q_5 \rightarrow aQ_4$

$\sqcup Q_5 \rightarrow bQ_4$

$\sqcup h \rightarrow \sqcup Q_4$

$Q_4\sqcup \rightarrow \sqcup Q_5$

13-75: $L_{re} \subseteq L_{UG}$

- Generating $abab$

$$
\begin{array}{rcl}
S & \Rightarrow & \underline{\triangleright \sqcup h \triangleleft} \\
\triangleright \underline{\sqcup h} \triangleleft & \Rightarrow & \triangleright \underline{\sqcup Q_4} \triangleleft \\
\triangleright \sqcup \underline{Q_4 \triangleleft} & \Rightarrow & \triangleright \sqcup \underline{\sqcup Q_5 \triangleleft} \\
\triangleright \sqcup \underline{\sqcup Q_5} \triangleleft & \Rightarrow & \triangleright \sqcup \underline{a Q_4} \triangleleft \\
\triangleright \sqcup a \underline{Q_4 \triangleleft} & \Rightarrow & \triangleright \sqcup a \underline{\sqcup Q_5 \triangleleft} \\
\triangleright \sqcup a \underline{\sqcup Q_5} \triangleleft & \Rightarrow & \triangleright \sqcup a \underline{b Q_4} \triangleleft \\
\triangleright \sqcup ab \underline{Q_4 \triangleleft} & \Rightarrow & \triangleright \sqcup ab \underline{\sqcup Q_5 \triangleleft} \\
\triangleright \sqcup ab \underline{\sqcup Q_5} \triangleleft & \Rightarrow & \triangleright \sqcup ab \underline{a Q_4} \triangleleft \\
\triangleright \sqcup aba \underline{Q_4 \triangleleft} & \Rightarrow & \triangleright \sqcup aba \underline{\sqcup Q_5 \triangleleft} \\
\triangleright \sqcup aba \underline{\sqcup Q_5} \triangleleft & \Rightarrow & \triangleright \sqcup aba \underline{b Q_4} \triangleleft
\end{array}
$$

13-76: $L_{re} \subseteq L_{UG}$

- Generating $abab$

$$
\begin{aligned}
\rhd \sqcup abab\underline{Q_4 \lhd} &\Rightarrow \rhd \sqcup abab\underline{\sqcup Q_3 \lhd} \\
\rhd \sqcup aba\underline{b \sqcup Q_3 \lhd} &\Rightarrow \rhd \sqcup aba\underline{b Q_2 \lhd} \\
\rhd \sqcup ab\underline{ab Q_2} \lhd &\Rightarrow \rhd \sqcup ab\underline{a Q_3 b} \lhd \\
\rhd \sqcup a\underline{ba Q_3} b \lhd &\Rightarrow \rhd \sqcup a\underline{b Q_2 a} b \lhd \\
\rhd \sqcup \underline{ab Q_2} ab \lhd &\Rightarrow \rhd \sqcup \underline{a Q_1 b} ab \lhd \\
\rhd \underline{\sqcup a Q_1} bab \lhd &\Rightarrow \rhd \underline{\sqcup Q_0 a} bab \lhd \\
\underline{\rhd \sqcup Q_0} abab \lhd &\Rightarrow abab \rhd \\
abab\underline{\lhd} &\Rightarrow abab
\end{aligned}
$$