Balanced Parentheses

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.
- Examples: $[[[]]], [][[[]]][], [][][[][]]$.

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.
- Examples: $[[[]]], [][[[]]][], [][][[][]]$.
- Unbalanced: $[[[]], []][$

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.
- Examples: $[[[]]], [][[[]]][], [][][[][]]$.
- Unbalanced: $[[[]], []][$
- Grammar: $S \rightarrow [S] | SS | \epsilon$.

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.
- Examples: $[[[]]], [][[[]]][], [][][[][]]$.
- Unbalanced: $[[[]], []][$
- Grammar: $S \to [S]|SS|\epsilon$.
- Derive $[][]$: $S \to SS \to [S]S \to []S \to [][S] \to [][]$.

The Problem

- An expression over $\{[,]\}$ such that
  (i) every left parenthesis has a matching right parenthesis,
  (ii) the matched pairs are well nested.
- Examples: $[[[]]], [][[[]]][], [][][[][]]$.
- Unbalanced: $[[[]], []][$
- Grammar: $S \rightarrow [S]|SS|\epsilon$.
- Derive $[][]$: $S \rightarrow SS \rightarrow [S]S \rightarrow []S \rightarrow [][S] \rightarrow [][]$.
- Derive $[[]]$: $S \rightarrow [S] \rightarrow [[S]] \rightarrow [[]]$.

Necessary Conditions for Balance

- $L(x) := \#[(x) =$ the number of left parentheses in x.

Necessary Conditions for Balance

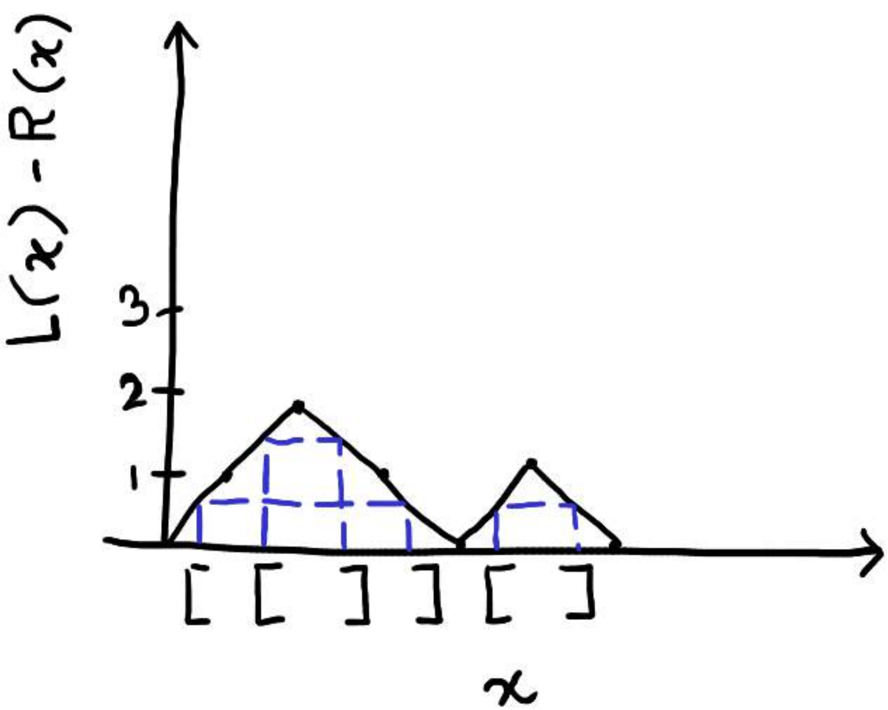
- $L(x) := \#[(x) =$ the number of left parentheses in x.
- $R(x) := \#](x) =$ the number of right parentheses in x.

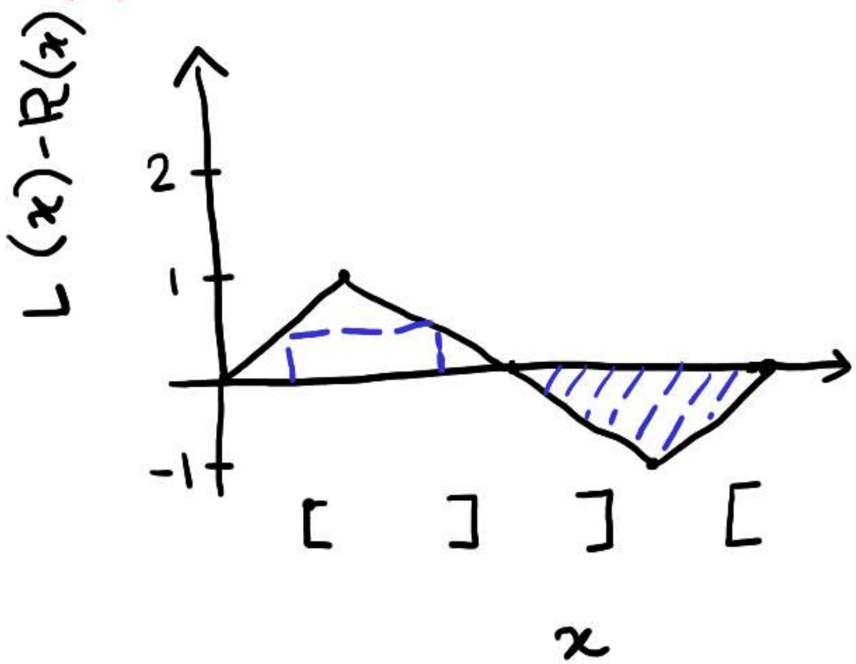
Necessary Conditions for Balance

- $L(x) := \#[(x) =$ the number of left parentheses in x.
- $R(x) := \#](x) =$ the number of right parentheses in x.
- Necessary conditions: A string x of parentheses is *balanced* iff:
  (i) $L(x) = R(x)$,
  (ii) for all prefixes y of x, $L(y) \geq R(y)$. - A right parenthesis can only match to a left parenthesis to its left.

Sufficient Conditions for Balance

- The above conditions are sufficient:
  Look at the graph of $L(x) - R(x)$ v x.





Production $S \rightarrow [S]|SS|\epsilon$

- Need to show that the given grammar $S \rightarrow [S]|SS|\epsilon$ generates exactly the set of strings satisfying the 2 balanced parantheses conditions.

Production $S \rightarrow [S]|SS|\epsilon$

- Need to show that the given grammar $S \rightarrow [S]|SS|\epsilon$ generates exactly the set of strings satisfying the 2 balanced parantheses conditions.
- Proof:
  ($\Rightarrow$) If $S \rightarrow_G^* x$ then x satisfies (i) and (ii).

Production $S \rightarrow [S]|SS|\epsilon$

- Need to show that the given grammar $S \rightarrow [S]|SS|\epsilon$ generates exactly the set of strings satisfying the 2 balanced parantheses conditions.
- Proof:
  ($\Rightarrow$) If $S \rightarrow_G^* x$ then x satisfies (i) and (ii).
- Induction on length of the derivation of x.
  In fact, we show that for any $\alpha \in (N \cup \Sigma)^*$, if $S \rightarrow_G^* \alpha$, then α satisfies (i) and (ii).
  In fact, induction on length of derivation of α.

- Base case: $S \to_G^0 \alpha$, so $\alpha = S$ and the two conditions are trivially satisfied.

- Base case: $S \to_G^0 \alpha$, so $\alpha = S$ and the two conditions are trivially satisfied.
- Induction step: $S \to_G^n \beta \to_G^1 \alpha$.

- Base case: $S \rightarrow^0_G \alpha$, so $\alpha = S$ and the two conditions are trivially satisfied.
- Induction step: $S \rightarrow^n_G \beta \rightarrow^1_G \alpha$.
- By IH, β satisfies (i) and (ii).

- $\beta \rightarrow^1_G \alpha$ can happen due to three types of productions:

- $\beta \to_G^1 \alpha$ can happen due to three types of productions:
- $S \to \epsilon$. So $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1\beta_2$: No change in order of parentheses and α satisfies (i) and (ii) iff β satisfies them.

- $\beta \to_G^1 \alpha$ can happen due to three types of productions:
- $S \to \epsilon$. So $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1\beta_2$: No change in order of parentheses and α satisfies (i) and (ii) iff β satisfies them.
- Similar argument for $S \to SS$.

- $\beta \to_G^1 \alpha$ can happen due to three types of productions:
- $S \to \epsilon$. So $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1 \beta_2$: No change in order of parentheses and α satisfies (i) and (ii) iff β satisfies them.
- Similar argument for $S \to SS$.
- $S \to [S]$: Then $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1 [S] \beta_2$.

- $\beta \to^1_G \alpha$ can happen due to three types of productions:
- $S \to \epsilon$. So $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1 \beta_2$: No change in order of parentheses and α satisfies (i) and (ii) iff β satisfies them.
- Similar argument for $S \to SS$.
- $S \to [S]$: Then $\beta = \beta_1 S \beta_2$ and $\alpha = \beta_1 [S] \beta_2$.
- Condition (i): $L(\alpha) = L(\beta) + 1$
  $= R(\beta) + 1$ (IH on β and (i))
  $= R(\alpha)$

- Condition (ii): Want to show that for any prefix γ of $\alpha = \beta_1[S]\beta_2$, $L(\gamma) \geq R(\gamma)$.

- Condition (ii): Want to show that for any prefix γ of $\alpha = \beta_1[S]\beta_2$, $L(\gamma) \geq R(\gamma)$.
- If γ is a prefix of β_1, then it is a prefix of β - so done by IH.

- Condition (ii): Want to show that for any prefix γ of $\alpha = \beta_1[S]\beta_2$, $L(\gamma) \geq R(\gamma)$.
- If γ is a prefix of β_1, then it is a prefix of β - so done by IH.
- If γ is a prefix of $\beta_1[S$ but not β_1, then
  $L(\gamma) = L(\beta_1) + 1$
  $\geq R(\beta_1) + 1$ (IH as β_1 is a prefix of β)
  $\geq R(\beta_1)$
  $= R(\gamma)$.

- Condition (ii): Want to show that for any prefix γ of $\alpha = \beta_1[S]\beta_2$, $L(\gamma) \geq R(\gamma)$.
- If γ is a prefix of β_1, then it is a prefix of β - so done by IH.
- If γ is a prefix of $\beta_1[S$ but not β_1, then
  $L(\gamma) = L(\beta_1) + 1$
  $\geq R(\beta_1) + 1$ (IH as β_1 is a prefix of β)
  $\geq R(\beta_1)$
  $= R(\gamma)$.
- If $\gamma = \beta_1[S]\delta$ where δ is a prefix of β_2, then
  $L(\gamma) = L(\beta_1 S\delta) + 1$
  $\geq R(\beta_1 S\delta) + 1$ (IH and definition) $= R(\gamma)$

- Condition (ii): Want to show that for any prefix γ of $\alpha = \beta_1[S]\beta_2$, $L(\gamma) \geq R(\gamma)$.
- If γ is a prefix of β_1, then it is a prefix of β - so done by IH.
- If γ is a prefix of $\beta_1[S$ but not β_1, then
  $L(\gamma) = L(\beta_1) + 1$
  $\geq R(\beta_1) + 1$ (IH as β_1 is a prefix of β)
  $\geq R(\beta_1)$
  $= R(\gamma)$.
- If $\gamma = \beta_1[S]\delta$ where δ is a prefix of β_2, then
  $L(\gamma) = L(\beta_1 S \delta) + 1$
  $\geq R(\beta_1 S \delta) + 1$ (IH and definition) $= R(\gamma)$
- Thus (ii) also holds for α and this concludes the proof of ($\Rightarrow$): If $S \rightarrow_G^* \alpha$, then α is balanced [In particular, when α is a sentence x].

- ($\Leftarrow$) If x is balanced, then $S \rightarrow_G^* x$.

- ($\Leftarrow$) If x is balanced, then $S \to_G^* x$.
- Induction on $|x|$. By assumption, x satisfies (i) and (ii).

- ($\Leftarrow$) If x is balanced, then $S \rightarrow_G^* x$.
- Induction on $|x|$. By assumption, x satisfies (i) and (ii).
- Base case: If $|x| = 0$, then $x = \epsilon$. Then $S \rightarrow \epsilon$ is already a production.

- ($\Leftarrow$) If x is balanced, then $S \rightarrow_G^* x$.
- Induction on $|x|$. By assumption, x satisfies (i) and (ii).
- Base case: If $|x| = 0$, then $x = \epsilon$. Then $S \rightarrow \epsilon$ is already a production.
- IH: If $|x| > 0$, then
  (a) Either there exists a proper prefix y of x satisfying (i), (ii)
  (b) Or no such prefix exists.

- ($\Leftarrow$) If x is balanced, then $S \rightarrow_G^* x$.
- Induction on $|x|$. By assumption, x satisfies (i) and (ii).
- Base case: If $|x| = 0$, then $x = \epsilon$. Then $S \rightarrow \epsilon$ is already a production.
- IH: If $|x| > 0$, then
  (a) Either there exists a proper prefix y of x satisfying (i), (ii)
  (b) Or no such prefix exists.
- Case (a): $x = yz$ where $z \neq \epsilon$.
  If y and x satisfy (i) and (ii), then so does z. (Check for yourself)

- ($\Leftarrow$) If x is balanced, then $S \rightarrow^*_G x$.
- Induction on $|x|$. By assumption, x satisfies (i) and (ii).
- Base case: If $|x| = 0$, then $x = \epsilon$. Then $S \rightarrow \epsilon$ is already a production.
- IH: If $|x| > 0$, then
  (a) Either there exists a proper prefix y of x satisfying (i), (ii)
  (b) Or no such prefix exists.
- Case (a): $x = yz$ where $z \neq \epsilon$.
  If y and x satisfy (i) and (ii), then so does z. (Check for yourself)
- By IH $S \rightarrow^*_G y$ and $S \rightarrow^*_G z$. Then
  $S \rightarrow^1_G SS \rightarrow^*_G yS \rightarrow^*_G yz = x$.

- Case (b): No such y exists. Then it must be that $x = [z]$.

- Case (b): No such y exists. Then it must be that $x = [z]$.
- If x satisfies (i) and (ii), then so does z:
  z satisfies (i) from its definition (Check for yourself).

- Case (b): No such y exists. Then it must be that $x = [z]$.
- If x satisfies (i) and (ii), then so does z:
  z satisfies (i) from its definition (Check for yourself).
- z satisfies (ii) because for all non-null prefixes u of z,
  $L(u) - R(u) = L([u) - 1 - R([u) \geq 0$.
  (Case (b): $L([u) - R([u) \geq 1$, o/w $[u$ is a proper prefix of x satisfying (i) and (ii).)

- Case (b): No such y exists. Then it must be that $x = [z]$.
- If x satisfies (i) and (ii), then so does z:
  z satisfies (i) from its definition (Check for yourself).
- z satisfies (ii) because for all non-null prefixes u of z,
  $L(u) - R(u) = L([u) - 1 - R([u) \geq 0$.
  (Case (b): $L([u) - R([u) \geq 1$, o/w $[u$ is a proper prefix of x satisfying (i) and (ii).)
- By IH $S \rightarrow_G^* z$. Then
  $S \rightarrow_G^1 [S] \rightarrow_G^* [z] = x$.

- Case (b): No such y exists. Then it must be that $x = [z]$.
- If x satisfies (i) and (ii), then so does z:
  z satisfies (i) from its definition (Check for yourself).
- z satisfies (ii) because for all non-null prefixes u of z,
  $L(u) - R(u) = L([u) - 1 - R([u) \geq 0$.
  (Case (b): $L([u) - R([u) \geq 1$, o/w $[u$ is a proper prefix of x satisfying (i) and (ii).)
- By IH $S \rightarrow_G^* z$. Then
  $S \rightarrow_G^1 [S] \rightarrow_G^* [z] = x$.
- Thus ($\Leftarrow$) is done: Every string satisfying (i) and (ii) can be derived.

- Case (b): No such y exists. Then it must be that $x = [z]$.
- If x satisfies (i) and (ii), then so does z:
  z satisfies (i) from its definition (Check for yourself).
- z satisfies (ii) because for all non-null prefixes u of z,
  $L(u) - R(u) = L([u) - 1 - R([u) \geq 0$.
  (Case (b): $L([u) - R([u) \geq 1$, o/w $[u$ is a proper prefix of x satisfying (i) and (ii).)
- By IH $S \rightarrow_G^* z$. Then
  $S \rightarrow_G^1 [S] \rightarrow_G^* [z] = x$.
- Thus ($\Leftarrow$) is done: Every string satisfying (i) and (ii) can be derived.
- Thus, grammar $S \rightarrow [S]|SS|\epsilon$ generates exactly the set of strings satisfying the 2 balanced parentheses conditions.