Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((\text{len}, \text{offset})\)
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\), or a string reference in form \((len, offset)\)
- Example encoding:

  \[
  \text{abbaabbbabab} \rightarrow \begin{array}{c}
  a = (0,1, a) \\
  b = (0,1, b) \\
  a = (0,1, a) \\
  b = (3,4) \\
  a = (0,0)
\end{array}
  \]

  \[
  \text{abbaabbbabab} \Rightarrow (0,4, \text{abba})(3,4)(2,5)(3,2)(0,0)
  \]

- Decoding:

  \[
  \text{abbaabbaabbabbaabbbaabbaabbbabab} \\
  \text{abbaabbbabab}
  \]

©2016 D. S. Hirschberg
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\), or a string reference in form \((\text{len}, \text{offset})\)
- Example

  **encoding:**
  
  \[
  \text{abbaabbbabab} \rightarrow a = (0,1,a)
  \]

  \[
  \text{abbaabbbabab} \rightarrow b = (0,1,b)
  \]
  \[
  \text{abbaabbbabab} \rightarrow b = (0,1,b)
  \]
  \[
  \text{abbaabbbabab} \rightarrow a = (0,1,a)
  \]
  \[
  \text{abbaabbbabab} \rightarrow \text{EOF} = (0,0)
  \]

  \[
  \text{abbaabbbabab} \Rightarrow (0,4,\text{abba}) (3,4) (2,5) (3,2) (0,0)
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character (C), or a string reference in form (len, offset)
- Example encoding:
  - |abbaabbbabab → a = (0,1,a)
  - a|bbaabbbabab

"LZ-4" text
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \( (C) \),
  or a string reference in form \( (\text{len}, \text{offset}) \)
- Example
  encoding:

  \[
  \begin{align*}
  |\text{abbaabbbabab} & \rightarrow a = (0,1,\text{a}) \\
  \text{a|bbabab} & \rightarrow b = (0,1,\text{b})
  \end{align*}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)
- Example encoding:

  \[
  \begin{array}{l}
  |abbaabbbabab \rightarrow a = (0,1,a) \\
  a|bbaabbbabab \rightarrow b = (0,1,b) \\
  ab|baabbbbabab \\
  \end{array}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((\text{len}, \text{offset})\)
- Example encoding:

  \[
  \begin{align*}
  |abbaabbbabab & \rightarrow a = (0,1,a) \\
  a|bbaabbbabab & \rightarrow b = (0,1,b) \\
  ab|baabbbabab & \rightarrow b = (0,1,b) \\
  \end{align*}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)
- **Example encoding:**

  | \(\text{abbaabbbabab} \rightarrow a = (0,1,a)\) |
  | \(\text{a|bbabbbabab} \rightarrow b = (0,1,b)\) |
  | \(\text{ab|baabbbabab} \rightarrow b = (0,1,b)\) |
  | \(\text{abb|aabbbabab} \rightarrow \text{EOF} = (0,0)\) |

  \(\text{abbaabbbabab} \Rightarrow (0,4,\text{abba})(3,4)(2,5)(3,2)(0,0)\)
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)
- Example encoding:

<table>
<thead>
<tr>
<th>Original String</th>
<th>Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>abbaabbbabab</td>
<td>a = (0,1,a)</td>
</tr>
<tr>
<td>a bbaabbbabab</td>
<td>b = (0,1,b)</td>
</tr>
<tr>
<td>ab baabbbabab</td>
<td>b = (0,1,b)</td>
</tr>
<tr>
<td>abb aabbbabab</td>
<td>a = (0,1,a)</td>
</tr>
</tbody>
</table>

  \(abbaabbbabab\) \(\Rightarrow\) \((0,4,abba),(3,4),(2,5),(3,2),(0,0)\)

Decoding:

```
abbaabbaabbabbaabbbaabbaabbbabab
```
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)
- Example

**encoding:**

| abbaabbbabab → a = (0,1,a) |
| abbaabbbabab → b = (0,1,b) |
| abbaabbbabab → b = (0,1,b) |
| abbaabbbabab → a = (0,1,a) |
| abbaabbbabab |
| abbaabbbabab |
| abbaabbbabab |

©2016 D. S. Hirschberg
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\), or a string reference in form \((\text{len, offset})\)
- Example

  **encoding:**

  \[
  \begin{align*}
  |abbaabbbabab & \rightarrow a = (0,1,a) \\
  a|bbaabbbabab & \rightarrow b = (0,1,b) \\
  ab|baabbbabab & \rightarrow b = (0,1,b) \\
  abb|aabbbabab & \rightarrow a = (0,1,a) \\
  abba|abbbabab & \rightarrow abb = (3,4) \\
  \end{align*}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)

- **Example encoding:**
  
  \[
  \begin{align*}
  |abbaabbbabab & \rightarrow a = (0,1,a) \\
  a|bbbaabbbabab & \rightarrow b = (0,1,b) \\
  ab|baabbbbabab & \rightarrow b = (0,1,b) \\
  abb|aabbbabab & \rightarrow a = (0,1,a) \\
  abba|abbbabab & \rightarrow abb = (3,4) \\
  abbaabb|babab &
  \end{align*}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\), or a string reference in form \((len, offset)\)
- Example encoding:

  \[
  \begin{align*}
  \mid abbaabbbabab & \rightarrow a = (0,1,a) \\
  a\mid bbaabbbbabab & \rightarrow b = (0,1,b) \\
  ab\mid baabbbbabab & \rightarrow b = (0,1,b) \\
  abb\mid aabbbbabab & \rightarrow a = (0,1,a) \\
  abba\mid abbbbabab & \rightarrow abb = (3,4) \\
  abbaabb\mid babab & \rightarrow ba = (2,5) 
  \end{align*}
  \]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\), or a string reference in form \((len, offset)\)
- Example

**encoding:**

| abbaabbbabab | → a = (0,1,a) |
| a|bbabab | → b = (0,1,b) |
| ab|baabbbabab | → b = (0,1,b) |
| abb|aabbbabab | → a = (0,1,a) |
| abba|abbbabab | → abb = (3,4) |
| abbaabb|babab | → ba = (2,5) |

| abbaabbbba|bab |
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \( (C) \),
  or a string reference in form \( (len, offset) \)

**Example encoding:**

\[
\begin{align*}
|abbaabbbabab \rightarrow a &= (0,1,a) \\
|abaabbbabab \rightarrow b &= (0,1,b) \\
|aabbabab \rightarrow b &= (0,1,b) \\
|ababbbabab \rightarrow a &= (0,1,a) \\
|abbaabbbabab \rightarrow abb &= (3,4) \\
|abbaabbabab \rightarrow ba &= (2,5) \\
|abbaabbbba \rightarrow bab &= (3,2)
\end{align*}
\]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \( (C) \), or a string reference in form \((\text{len}, \text{offset})\)
- Example

  encoding:

  |   | abbaabbbabab       | a = (0,1,a) |
  | ab | bbaabbbabab       | b = (0,1,b) |
  | ab | baabbbabab       | b = (0,1,b) |
  | ab | bbabbbabab       | a = (0,1,a) |
  | abba| bbbabab        | abb = (3,4) |
  | abba| aabbab          | ba = (2,5) |
  | abbaa| bbbab          | bab = (3,2) |
  | abbaabbbabab|                | EOF = (0,0) |
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \((C)\),
  or a string reference in form \((len, offset)\)

**Example**

**encoding:**

\[
\begin{array}{l}
\text{abbaabbbabab} \rightarrow a = (0,1,a) \\
\text{a|bbabab} \rightarrow b = (0,1,b) \\
\text{aba|abbabab} \rightarrow b = (0,1,b) \\
\text{abb|aabbabab} \rightarrow a = (0,1,a) \\
\text{abba|abbbabab} \rightarrow abb = (3,4) \\
\text{abbaabb|babab} \rightarrow ba = (2,5) \\
\text{abbaabbb|bab} \rightarrow bab = (3,2) \\
\text{abbaabbbbabab} \Rightarrow (0,4,\text{abba})(3,4)(2,5)(3,2)(0,0) \\
\end{array}
\]
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character (C), or a string reference in form \( (len, offset) \)

**Example**

**encoding:**

- \( a.baabbbabab \rightarrow a = (0,1,a) \)
- \( ab.baabbbabab \rightarrow b = (0,1,b) \)
- \( ab.baabbbabab \rightarrow b = (0,1,b) \)
- \( abb.aabbbabab \rightarrow a = (0,1,a) \)
- \( abba.abbbabab \rightarrow abb = (3,4) \)
- \( abbaabbbabab \rightarrow ba = (2,5) \)
- \( abbaabbbabab \rightarrow bab = (3,2) \)
- \( abbaabbbabab \rightarrow EOF = (0,0) \)
- \( abbaabbbabab \Rightarrow (0,4,abba)(3,4)(2,5)(3,2)(0,0) \)

**decoding:**

- abba
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character (C), or a string reference in form (len, offset)
- Example

**encoding:**

<table>
<thead>
<tr>
<th>abbaabbbabab</th>
<th>$a = (0,1,a)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>bbababab</td>
</tr>
<tr>
<td>ab</td>
<td>baabbbabab</td>
</tr>
<tr>
<td>ab</td>
<td>baabbbabab</td>
</tr>
<tr>
<td>ab</td>
<td>baabbbabab</td>
</tr>
<tr>
<td>abba</td>
<td>abbbabab</td>
</tr>
<tr>
<td>abbaabb</td>
<td>babab</td>
</tr>
<tr>
<td>abbaabbbbabab</td>
<td>$EOF = (0,0)$</td>
</tr>
<tr>
<td>abbaabbbbabab</td>
<td>$\Rightarrow (0,4,abba)(3,4)(2,5)(3,2)(0,0)$</td>
</tr>
</tbody>
</table>

**decoding:**

```
abba
abbaabb
```
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character ($C$), or a string reference in form ($len$, $offset$)

**Example encoding:**

- $abbaabbbabab \rightarrow a = (0,1,a)$
- $abaabbbabab \rightarrow b = (0,1,b)$
- $abbaabbbabab \rightarrow b = (0,1,b)$
- $abbabbbabab \rightarrow a = (0,1,a)$
- $abbaabbbabab \rightarrow abb = (3,4)$
- $abbaabbbabab \rightarrow ba = (2,5)$
- $abbaabbbbaabab \rightarrow bab = (3,2)$
- $abbaabbbbabab \rightarrow EOF = (0,0)$
- $abbaabbbbabab \Rightarrow (0,4,abba)(3,4)(2,5)(3,2)(0,0)$

**Decoding:**

- abba
- abbaabb
- abbaaabbba
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character \( C \), or a string reference in form \((\text{len}, \text{offset})\)

**Example**

**encoding:**

<table>
<thead>
<tr>
<th>abbaabbbabab → a = (0,1,a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>abaabbbabab → b = (0,1,b)</td>
</tr>
<tr>
<td>abbaabbbabab → b = (0,1,b)</td>
</tr>
<tr>
<td>aabbaabbbabab → a = (0,1,a)</td>
</tr>
<tr>
<td>abbaabbbabab → abb = (3,4)</td>
</tr>
<tr>
<td>abbaabbbabab → ba = (2,5)</td>
</tr>
<tr>
<td>abbaabbbabab → bab = (3,2)</td>
</tr>
<tr>
<td>abbaabbbabab → bab = (3,2)</td>
</tr>
<tr>
<td>abbaabbbabab ⇒ (0,4,abba)(3,4)(2,5)(3,2)(0,0)</td>
</tr>
</tbody>
</table>

**decoding:**

abba
abbaaabb
abbaabbbba
abbaabbbbabab
Lempel-Ziv sliding window

- Search window to find longest match with prefix of lookahead buffer
- Output at each step is either a character (C), or a string reference in form (len, offset)
- Example

  **encoding:**

  \[
  \begin{align*}
  \text{abbaabbbabab} & \rightarrow a = (0,1,a) \\
  a\text{bbbaabbbabab} & \rightarrow b = (0,1,b) \\
  ab\text{baabbbabab} & \rightarrow b = (0,1,b) \\
  \text{abbababbbabab} & \rightarrow a = (0,1,a) \\
  \text{abbaabbbabab} & \rightarrow \text{abb} = (3,4) \\
  \text{abbaabbabab} & \rightarrow \text{ba} = (2,5) \\
  \text{abbaabbbabab} & \rightarrow \text{bab} = (3,2) \\
  \text{abbaabbbabab} & \Rightarrow (0,4,\text{abba})(3,4)(2,5)(3,2)(0,0)
  \end{align*}
  \]

  **decoding:**

  - abba
  - abbaabb
  - abbaabbbba
  - abbaabbbabab