Proof of Lemma II.2.1.2.

Lemma (II.2.1.2.)

1. Assume a grammar \( G \) which has only productions of the form

\[
A \rightarrow Bw \text{ or } A \rightarrow w
\]

for some \( w \in T^* \), \( A, B \in N \). Then \( L(G) = L(G') \) for some left-linear grammar \( G' \), which can be computed from \( G \).

2. Assume a grammar \( G \) which has only productions of the form

\[
A \rightarrow wB \text{ or } A \rightarrow w
\]

for some \( w \in T^* \), \( A, B \in N \). Then \( L(G) = L(G') \) for some right-linear grammar \( G' \), which can be computed from \( G \).
Proof of Lemma II.2.1.2.

In a first step we omit all transitions $A \rightarrow B$ for $A, B \in N$:

Let $G = (N, T, S, P)$ be a grammar having such transitions.

We form a grammar $G'$ having no such transitions as follows:

<table>
<thead>
<tr>
<th>grammar</th>
<th>$G'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>terminals</td>
<td>$N$</td>
</tr>
<tr>
<td>nonterminals</td>
<td>$T$</td>
</tr>
<tr>
<td>start symbol</td>
<td>$S$</td>
</tr>
</tbody>
</table>
| productions | $A \rightarrow w$ if $A \Rightarrow^*_G A' \rightarrow w$ for some $A, A' \in N$, $w \in T^*$  
$A \rightarrow wB$ if $A \Rightarrow^*_G A' \rightarrow wB$ for some $A, A', B \in N$, $w \in T^*$ |

So in $G'$ we just jump over all silent transitions $A \rightarrow B$ in $G$.

We can in fact decide whether $A \Rightarrow^* A'$, since such a derivation must have the form $A = A'$ or $A = A_1 \Rightarrow A_2 \Rightarrow \cdots \Rightarrow A_n = A$ for some $A_i \in N$.

And if such derivation exists then a derivation exists in which all $A_i$ are distinct (omit loops).

Therefore $n$ can be restricted to the number of elements in $N$, and therefore there are only finitely many possible derivations, which we can enumerate. For each of them we can check whether it is in fact a derivation, and therefore determine all possible derivations $A \Rightarrow^* A'$.

End of Proof of II.2.1.2.

We have now obtained a grammar which doesn’t contain silent productions of the form $A \rightarrow B$ for nonterminals $A, B$.

The following lemma shows that such languages are definable by left-linear or right-linear grammars.
Proof of Lemma II.2.1.3.

Lemma (II.2.1.3.)

1. Assume a grammar $G$ which has only productions of the form
   $$ A \rightarrow Bw \text{ or } A \rightarrow w' $$
   for some $w \in T^+$, $w' \in T^*$, $A, B \in N$. Then $L(G) = L(G')$ for some left-linear grammar $G'$, and $G'$ can effectively computed from $G$.

2. Assume a grammar $G$ which has only productions of the form
   $$ A \rightarrow wB \text{ or } A \rightarrow w' $$
   for some $w \in T^+$, $w' \in T^*$, $A, B \in N$. Then $L(G) = L(G')$ for some right-linear grammar $G'$, and $G'$ can effectively computed from $G$.

Theorem

(a) Let $G = (N, T, S, P)$ be a left-linear grammar, $A \in N$, $w \in (N \cup T)^*$, $A \Rightarrow^* w$.
Then the derivation of $A \Rightarrow^* w$ is

$$ A \Rightarrow A_1 a_1 \Rightarrow A_2 a_2 a_1 \Rightarrow \cdots \Rightarrow A_n a_n \cdots a_2 a_1 = w \quad (1) $$

or
$$ A \Rightarrow A_1 a_1 \Rightarrow A_2 a_2 a_1 \Rightarrow \cdots \Rightarrow A_n a_n \cdots a_2 a_1 \Rightarrow a_{n+1} a_n \cdots a_2 a_1 = w \quad (2) $$

or
$$ A \Rightarrow A_1 a_1 \Rightarrow A_2 a_2 a_1 \Rightarrow \cdots \Rightarrow A_n a_n \cdots a_2 a_1 \Rightarrow a_n \cdots a_2 a_1 = w \quad (3) $$

for productions
- $A_i \rightarrow a_{i+1} a_i \ (\text{in } (1)-(3))$,
- $A_n \rightarrow a_{n+1} \ (\text{in } (2))$
- $A_n \rightarrow \epsilon \ (\text{in } (3))$

(b) Let $G = (N, T, S, P)$ be a right-linear grammar, $A \in N$, $w \in (N \cup T)^*$, $A \Rightarrow^* w$.
Then the derivation of $A \Rightarrow^* w$ is

$$ A \Rightarrow a_1 A_1 \Rightarrow a_1 a_2 A_2 \Rightarrow \cdots \Rightarrow a_1 a_2 \cdots a_n A_n = w \quad (1) $$

or
$$ A \Rightarrow a_1 A_1 \Rightarrow a_1 a_2 A_2 \Rightarrow \cdots \Rightarrow a_1 a_2 \cdots a_n A_n \Rightarrow a_1 a_2 \cdots a_n A_{n+1} = w \quad (2) $$

or
$$ A \Rightarrow a_1 A_1 \Rightarrow a_1 a_2 A_2 \Rightarrow \cdots \Rightarrow a_1 a_2 \cdots a_n A_n \Rightarrow a_1 a_2 \cdots a_n = w \quad (3) $$

for productions
- $A_i \rightarrow a_{i} a_i+1 \ (\text{in } (1)-(3))$
- $A_n \rightarrow a_{n+1} \ (\text{in } (2))$
- $A_n \rightarrow \epsilon \ (\text{in } (3))$. 

Derivations in Regular Grammars
II.2.1. Regular Languages (12.2)

Proof

The above are the only derivations possible.

II.2.2. Regular Expressions (13.8)

Proof of Lemma II.2.2.1.

Lemma (II.2.2.1.)

Let \( G, G' \) be both left-linear grammars or both right-linear grammars. Then we can define a left-linear or right-linear grammars \( G_i \) s.t.

1. \( L(G_1) = L(G) \cup L(G') \),
2. \( L(G_2) = L(G).L(G') \),
3. \( L(G_3) = L(G)^* \).

These grammars can be computed from \( G \) and \( G' \).

Assume in 1./2./3.

\[ G = (T, N, S, P), \quad G' = (T', N', S', P') \]

After renaming of nonterminals we can assume \( N \cap N' = \emptyset \).
Let \( S'' \) be a new symbol not in \( N \cup N' \cup T \cup T' \).
We define multi-step left/right-linear grammars with those properties, from which one can construct ordinary (one-step) left/right-linear grammars with those properties.
We only carry out the proof for right-linear grammars.
II.2.2. Regular Expressions (13.8)

Proof of 1.

We define $G_1$ as follows:

<table>
<thead>
<tr>
<th>grammar</th>
<th>$G_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>terminals</td>
<td>$T \cup T'$</td>
</tr>
<tr>
<td>nonterminals</td>
<td>$N \cup N' \cup {S''}$</td>
</tr>
<tr>
<td>start symbol</td>
<td>$S''$</td>
</tr>
</tbody>
</table>
| productions | $S'' \rightarrow S$
  $S'' \rightarrow S'$
  $P$
  $P'$ |

So $G_1$ has the productions from $G$ and $G'$ plus

$$S'' \rightarrow S \text{ and } S'' \rightarrow S'. $$

Derivations in $G_1$ have the form

$$S'' \Rightarrow S \Rightarrow^* w$$

and

$$S'' \Rightarrow S' \Rightarrow^* w'$$

for derivations

$$S \Rightarrow^*_G w$$

and

$$S' \Rightarrow^*_G w'$$

So for $w'' \in (T \cup T')^*$ we have

$$S'' \Rightarrow^*_G w'' \text{ iff } S \Rightarrow^*_G w' \text{ or } S' \Rightarrow^*_G w',$$

so $L(G'') = L(G) \cup L(G')$.

Proof of 2.

We define $G_2$ as follows:

<table>
<thead>
<tr>
<th>grammar</th>
<th>$G_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>terminals</td>
<td>$T \cup T'$</td>
</tr>
<tr>
<td>nonterminals</td>
<td>$N \cup N'$</td>
</tr>
<tr>
<td>start symbol</td>
<td>$S$</td>
</tr>
</tbody>
</table>
| productions | $A \rightarrow aA'$ for $A \rightarrow aA' \in P$ ($A, A' \in N$, $a \in T$)
  $A \rightarrow aS'$ for $A \rightarrow a \in P$ ($A \in N$, $a \in T$)
  $P'$ |

So $G_2$ has

- the productions from $G'$,
- the productions of the form $A \rightarrow aA$ from $G$ and
- productions $A \rightarrow aS'$, if $A \rightarrow a$ is a production from $G$.

A derivation in $G_2$ starts with a derivation

$$S \Rightarrow a_1A_1 \Rightarrow a_1a_2A_2 \Rightarrow a_1a_2a_3A_3 \Rightarrow \cdots \Rightarrow a_1a_2 \cdots a_{n-1}A_{n-1} \Rightarrow a_1a_2 \cdots a_nS'$$

for derivations in $G$ of the form

$$S \Rightarrow a_1A_1 \Rightarrow a_1a_2A_2 \Rightarrow a_1a_2a_3A_3 \Rightarrow \cdots \Rightarrow a_1a_2 \cdots a_{n-1}A_{n-1} \Rightarrow a_1a_2 \cdots a_n.$$
II.2.2. Regular Expressions (13.8)

Proof of 2.

Then this is followed by a derivation

\[ a_1a_2 \cdots a_nS' \Rightarrow a_1a_2 \cdots a_nb_1B_1 \Rightarrow a_1a_2 \cdots a_nb_1b_2B_2 \Rightarrow \cdots \]
\[ \Rightarrow a_1a_2 \cdots a_nb_1b_2 \cdots b_{m-1}B_{m-1} \Rightarrow a_1a_2 \cdots a_nb_1b_2 \cdots b_m \]

for a derivation in \( G' \) of the form

\[ S' \Rightarrow b_1B_1 \Rightarrow b_1b_2B_2 \Rightarrow \cdots \]
\[ \Rightarrow b_1b_2 \cdots b_{m-1}B_{m-1} \Rightarrow b_1b_2 \cdots b_m \]

Therefore \( S \Rightarrow^* w \) for some \( w \in (T \cup T')^* \) if and only if \( S \Rightarrow^* w' \) and \( S' \Rightarrow^* w'' \) for some \( w', w'' \) s.t. \( w = ww'' \).

So \( L(G_2) = L(G)L(G') \).

II.2.2. Regular Expressions (13.8)

Proof of 3.

We define \( G_3 \) as follows:

- **grammar** \( G_3 \)
- **terminals** \( T \)
- **nonterminals** \( N \)
- **start symbol** \( S \)
- **productions**
  - \( S \rightarrow \epsilon \),
  - \( A \rightarrow aA' \) for \( A \rightarrow aA' \in P (A, A' \in N, a \in T) \)
  - \( A \rightarrow aS \) for \( A \rightarrow a \in P (A \in N, a \in T) \)

Derivations in \( G_3 \) are \( S \Rightarrow \epsilon \) or they start similarly as for concatenation with

\[ S \Rightarrow^* wS \]

for a derivation in \( G \)

\[ S \Rightarrow^* w \]

and \( w \in N^+ \). In the latter case it can continue either (using \( S \rightarrow \epsilon \) ) with

\[ wS \Rightarrow w \]

or with

\[ wS \Rightarrow^* ww'S \]

for a derivation in \( G \)

\[ S \Rightarrow^* w' \]

Again in the latter case we can continue (using \( S \rightarrow \epsilon \) ) with

\[ ww'S \rightarrow ww' \]

or with

\[ ww'S \Rightarrow^* ww'w''S \]

for a derivation in \( G \)

\[ S \Rightarrow^* w'' \]

We obtain that in \( G_3 \) we have

\[ S \Rightarrow^* w \]

if there exist derivations in \( G \) of

- \( S \Rightarrow^* w_1 \)
- \( S \Rightarrow^* w_2 \)
- \( \ldots \)
- \( S \Rightarrow^* w_n \)

s.t. \( w = w_1w_2 \ldots w_n \). So we get

\[ L(G_3) = \{ w_1w_2 \ldots w_n \mid n \geq 0, w_1, \ldots, w_n \in L(G) \} = L(G)^* \]
Proof of Lemma II.2.2.2.

Lemma (II.2.2.2.)
Let $E$ be a regular Expression. Then there exist both left-linear and right-linear grammars $G$, $G'$ s.t.

$$L(E) = L(G) = L(G')$$

$G$ and $G'$ can be computed from $L$.

Proof: By Lemma II.2.2.1, and the fact that the finite languages $\emptyset$, $\{\epsilon\}$ and $\{a\}$ are regular.

Induction on the definition of regular expressions.

**Case 1:** $L = \emptyset$, $\epsilon$, $a$ (where $a \in T$). Then $L$ is finite, therefore definable by a left/right-linear grammar.

**Case 2:** $L = (L_1) \mid (L_2)$ or $L = (L_1)L_2$ or $L = (L_1)^*$. By IH $L_i$ are defined by left/right-linear grammars $G_i$. By Lemma II.2.2.1. it follows that $L$ can be defined by a left/right-linear grammar.