3.2: Equivalence and Correctness of Regular Expressions

In this section, we:

- say what it means for regular expressions to be equivalent;
- show a series of results about regular expression equivalence;
- show how regular expressions can be synthesized and proved correct.

Equivalence of Regular Expressions

We say that regular expressions α and β are *equivalent* iff $L(\alpha) = L(\beta)$.

Equivalence of Regular Expressions

We say that regular expressions α and β are *equivalent* iff $L(\alpha) = L(\beta)$.

We define a relation $\approx$ on **Reg** by: $\alpha \approx \beta$ iff α and β are equivalent.

Equivalence of Regular Expressions

We say that regular expressions α and β are *equivalent* iff $L(\alpha) = L(\beta)$.

We define a relation $\approx$ on **Reg** by: $\alpha \approx \beta$ iff α and β are equivalent.

For example, $L((00)^* + \%) = L((00)^*)$, and thus $(00)^* + \% \approx (00)^*$.

Equivalence of Regular Expressions

We say that regular expressions α and β are *equivalent* iff $L(\alpha) = L(\beta)$.

We define a relation $\approx$ on **Reg** by: $\alpha \approx \beta$ iff α and β are equivalent.

For example, $L((00)^* + \%) = L((00)^*)$, and thus $(00)^* + \% \approx (00)^*$.

One approach to showing that $\alpha \approx \beta$ is to show that $L(\alpha) \subseteq L(\beta)$ and $L(\beta) \subseteq L(\alpha)$. The following proposition is useful for showing language inclusions, not just ones involving regular languages.

Language Inclusions

Proposition 3.2.1

(1) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 \cup A_2 \subseteq B_1 \cup B_2$.

(2) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 \cap A_2 \subseteq B_1 \cap B_2$.

(3) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $B_2 \subseteq A_2$, *then* $A_1 - A_2 \subseteq B_1 - B_2$.

(4) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 A_2 \subseteq B_1 B_2$.

(5) *For all* $A, B \in \mathbf{Lan}$ *and* $n \in \mathbb{N}$, *if* $A \subseteq B$, *then* $A^n \subseteq B^n$.

(6) *For all* $A, B \in \mathbf{Lan}$, *if* $A \subseteq B$, *then* $A^* \subseteq B^*$.

Language Inclusions

Proposition 3.2.1

(1) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 \cup A_2 \subseteq B_1 \cup B_2$.

(2) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 \cap A_2 \subseteq B_1 \cap B_2$.

(3) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $B_2 \subseteq A_2$, *then* $A_1 - A_2 \subseteq B_1 - B_2$.

(4) *For all* $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, *if* $A_1 \subseteq B_1$ *and* $A_2 \subseteq B_2$, *then* $A_1 A_2 \subseteq B_1 B_2$.

(5) *For all* $A, B \in \mathbf{Lan}$ *and* $n \in \mathbb{N}$, *if* $A \subseteq B$, *then* $A^n \subseteq B^n$.

(6) *For all* $A, B \in \mathbf{Lan}$, *if* $A \subseteq B$, *then* $A^* \subseteq B^*$.

Language Inclusions (Cont.)

Proof. (1) and (2) are straightforward. We show (3) as an example, below. (4) is easy. (5) is proved by mathematical induction, using (4). (6) is proved using (5).

For (3), suppose that $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, $A_1 \subseteq B_1$ and $B_2 \subseteq A_2$. To show that $A_1 - A_2 \subseteq B_1 - B_2$, suppose $w \in A_1 - A_2$. We must show that $w \in B_1 - B_2$. It will suffice to show that

□

Language Inclusions (Cont.)

Proof. (1) and (2) are straightforward. We show (3) as an example, below. (4) is easy. (5) is proved by mathematical induction, using (4). (6) is proved using (5).

For (3), suppose that $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, $A_1 \subseteq B_1$ and $B_2 \subseteq A_2$. To show that $A_1 - A_2 \subseteq B_1 - B_2$, suppose $w \in A_1 - A_2$. We must show that $w \in B_1 - B_2$. It will suffice to show that $w \in B_1$ and $w \notin B_2$.

□

Language Inclusions (Cont.)

Proof. (1) and (2) are straightforward. We show (3) as an example, below. (4) is easy. (5) is proved by mathematical induction, using (4). (6) is proved using (5).

For (3), suppose that $A_1, A_2, B_1, B_2 \in \textbf{Lan}$, $A_1 \subseteq B_1$ and $B_2 \subseteq A_2$. To show that $A_1 - A_2 \subseteq B_1 - B_2$, suppose $w \in A_1 - A_2$. We must show that $w \in B_1 - B_2$. It will suffice to show that $w \in B_1$ and $w \notin B_2$.

Since $w \in A_1 - A_2$, we have that

□

Language Inclusions (Cont.)

Proof. (1) and (2) are straightforward. We show (3) as an example, below. (4) is easy. (5) is proved by mathematical induction, using (4). (6) is proved using (5).

For (3), suppose that $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, $A_1 \subseteq B_1$ and $B_2 \subseteq A_2$. To show that $A_1 - A_2 \subseteq B_1 - B_2$, suppose $w \in A_1 - A_2$. We must show that $w \in B_1 - B_2$. It will suffice to show that $w \in B_1$ and $w \notin B_2$.

Since $w \in A_1 - A_2$, we have that $w \in A_1$ and $w \notin A_2$. Since $A_1 \subseteq B_1$, it follows that $w \in B_1$. Thus, it remains to show that $w \notin B_2$.

□

Language Inclusions (Cont.)

Proof. (1) and (2) are straightforward. We show (3) as an example, below. (4) is easy. (5) is proved by mathematical induction, using (4). (6) is proved using (5).

For (3), suppose that $A_1, A_2, B_1, B_2 \in \mathbf{Lan}$, $A_1 \subseteq B_1$ and $B_2 \subseteq A_2$. To show that $A_1 - A_2 \subseteq B_1 - B_2$, suppose $w \in A_1 - A_2$. We must show that $w \in B_1 - B_2$. It will suffice to show that $w \in B_1$ and $w \notin B_2$.

Since $w \in A_1 - A_2$, we have that $w \in A_1$ and $w \notin A_2$. Since $A_1 \subseteq B_1$, it follows that $w \in B_1$. Thus, it remains to show that $w \notin B_2$.

Suppose, toward a contradiction, that $w \in B_2$. Since $B_2 \subseteq A_2$, it follows that $w \in A_2$—contradiction. Thus we have that $w \notin B_2$.
□

Basic Equivalences

Proposition 3.2.2

(1) $\approx$ *is reflexive on* **Reg**, *symmetric and transitive.*

(2) *For all* $\alpha, \beta \in$ **Reg**, *if* $\alpha \approx \beta$, *then* $\alpha^* \approx \beta^*$.

(3) *For all* $\alpha_1, \alpha_2, \beta_1, \beta_2 \in$ **Reg**, *if* $\alpha_1 \approx \beta_1$ *and* $\alpha_2 \approx \beta_2$, *then* $\alpha_1\alpha_2 \approx \beta_1\beta_2$.

(4) *For all* $\alpha_1, \alpha_2, \beta_1, \beta_2 \in$ **Reg**, *if* $\alpha_1 \approx \beta_1$ *and* $\alpha_2 \approx \beta_2$, *then* $\alpha_1 + \alpha_2 \approx \beta_1 + \beta_2$.

Proof. Follows from the properties of $=$. As an example, we show Part (4).

Basic Equivalences (Cont.)

Proof (cont.). Suppose $\alpha_1, \alpha_2, \beta_1, \beta_2 \in$ **Reg**, and assume that $\alpha_1 \approx \beta_1$ and $\alpha_2 \approx \beta_2$. Then $L(\alpha_1) = L(\beta_1)$ and $L(\alpha_2) = L(\beta_2)$, so that

$$\begin{aligned} L(\alpha_1 + \alpha_2) &= L(\alpha_1) \cup L(\alpha_2) = L(\beta_1) \cup L(\beta_2) \\ &= L(\beta_1 + \beta_2). \end{aligned}$$

Thus $\alpha_1 + \alpha_2 \approx \beta_1 + \beta_2$. □

Basic Equivalences (Cont.)

Proposition 3.2.3
*Suppose $\alpha, \beta, \beta' \in$ **Reg**, $\beta \approx \beta'$, $\mathit{pat} \in$ **Path** is valid for α, and β is the subtree of α at position pat. Let α' be the result of replacing the subtree at position pat in α by β'. Then $\alpha \approx \alpha'$.*

Proof. By induction on α. □

Equivalences for Union

Proposition 3.2.4

(1) *For all* $\alpha, \beta \in \mathbf{Reg}$, $\alpha + \beta \approx \beta + \alpha$.

(2) *For all* $\alpha, \beta, \gamma \in \mathbf{Reg}$, $(\alpha + \beta) + \gamma \approx \alpha + (\beta + \gamma)$.

(3) *For all* $\alpha \in \mathbf{Reg}$, $\$ + \alpha \approx$

(4) *For all* $\alpha \in \mathbf{Reg}$, $\alpha + \alpha \approx$

(5) *If* $L(\alpha) \subseteq L(\beta)$, *then* $\alpha + \beta \approx$

Proof.

(1) Follows from the commutativity of $\cup$.

(2) Follows from the associativity of $\cup$.

□

Equivalences for Union

Proposition 3.2.4

(1) *For all* $\alpha, \beta \in$ **Reg**, $\alpha + \beta \approx \beta + \alpha$.

(2) *For all* $\alpha, \beta, \gamma \in$ **Reg**, $(\alpha + \beta) + \gamma \approx \alpha + (\beta + \gamma)$.

(3) *For all* $\alpha \in$ **Reg**, $\$ + \alpha \approx \alpha$.

(4) *For all* $\alpha \in$ **Reg**, $\alpha + \alpha \approx$

(5) *If* $L(\alpha) \subseteq L(\beta)$, *then* $\alpha + \beta \approx$

Proof.

(1) Follows from the commutativity of $\cup$.

(2) Follows from the associativity of $\cup$.

(3) Follows since $\emptyset$ is the identity for $\cup$.

□

Equivalences for Union

Proposition 3.2.4

(1) *For all* $\alpha, \beta \in \mathbf{Reg}$, $\alpha + \beta \approx \beta + \alpha$.

(2) *For all* $\alpha, \beta, \gamma \in \mathbf{Reg}$, $(\alpha + \beta) + \gamma \approx \alpha + (\beta + \gamma)$.

(3) *For all* $\alpha \in \mathbf{Reg}$, $\$ + \alpha \approx \alpha$.

(4) *For all* $\alpha \in \mathbf{Reg}$, $\alpha + \alpha \approx \alpha$.

(5) *If* $L(\alpha) \subseteq L(\beta)$, *then* $\alpha + \beta \approx$

Proof.

(1) Follows from the commutativity of $\cup$.

(2) Follows from the associativity of $\cup$.

(3) Follows since $\emptyset$ is the identity for $\cup$.

(4) Follows since $\cup$ is idempotent: $A \cup A = A$, for all sets A.

□

Equivalences for Union

Proposition 3.2.4

(1) *For all* $\alpha, \beta \in \textbf{Reg}$, $\alpha + \beta \approx \beta + \alpha$.

(2) *For all* $\alpha, \beta, \gamma \in \textbf{Reg}$, $(\alpha + \beta) + \gamma \approx \alpha + (\beta + \gamma)$.

(3) *For all* $\alpha \in \textbf{Reg}$, $\$ + \alpha \approx \alpha$.

(4) *For all* $\alpha \in \textbf{Reg}$, $\alpha + \alpha \approx \alpha$.

(5) *If* $L(\alpha) \subseteq L(\beta)$, *then* $\alpha + \beta \approx \beta$.

Proof.

(1) Follows from the commutativity of $\cup$.

(2) Follows from the associativity of $\cup$.

(3) Follows since $\emptyset$ is the identity for $\cup$.

(4) Follows since $\cup$ is idempotent: $A \cup A = A$, for all sets A.

(5) Follows since, if $L_1 \subseteq L_2$, then $L_1 \cup L_2 = L_2$.

□

Equivalences for Concatenation

Proposition 3.2.5

(1) *For all* $\alpha, \beta, \gamma \in$ **Reg**, $(\alpha\beta)\gamma \approx \alpha(\beta\gamma)$.

(2) *For all* $\alpha \in$ **Reg**, $\%\alpha \approx \alpha \approx \alpha\%$.

(3) *For all* $\alpha \in$ **Reg**, $\$\alpha \approx \quad \approx \alpha\$$.

Proof.

(1) Follows from the associativity of language concatenation.

(2) Follows since $\{\%\}$ is the identity for language concatenation.

□

Equivalences for Concatenation

Proposition 3.2.5

(1) *For all* $\alpha, \beta, \gamma \in \mathbf{Reg}$, $(\alpha\beta)\gamma \approx \alpha(\beta\gamma)$.

(2) *For all* $\alpha \in \mathbf{Reg}$, $\%\alpha \approx \alpha \approx \alpha\%$.

(3) *For all* $\alpha \in \mathbf{Reg}$, $\$\alpha \approx \$ \approx \alpha\$$.

Proof.

(1) Follows from the associativity of language concatenation.

(2) Follows since $\{\%\}$ is the identity for language concatenation.

(3) Follows since $\emptyset$ is the zero for language concatenation.

□

Distributivity of Concatenation Over Union

Proposition 3.2.6

(1) *For all* $L_1, L_2, L_3 \in \textbf{Lan}$, $L_1(L_2 \cup L_3) =$

(2) *For all* $L_1, L_2, L_3 \in \textbf{Lan}$, $(L_1 \cup L_2)L_3 =$

Distributivity of Concatenation Over Union

Proposition 3.2.6

(1) For all $L_1, L_2, L_3 \in \textbf{Lan}$, $L_1(L_2 \cup L_3) = L_1L_2 \cup L_1L_3$.

(2) For all $L_1, L_2, L_3 \in \textbf{Lan}$, $(L_1 \cup L_2)L_3 = L_1L_3 \cup L_2L_3$.

Proof. We show the proof of Part (1); the proof of the other part is similar. Suppose $L_1, L_2, L_3 \in \textbf{Lan}$. It will suffice to show that

Distributivity of Concatenation Over Union

Proposition 3.2.6

(1) For all $L_1, L_2, L_3 \in \mathbf{Lan}$, $L_1(L_2 \cup L_3) = L_1 L_2 \cup L_1 L_3$.

(2) For all $L_1, L_2, L_3 \in \mathbf{Lan}$, $(L_1 \cup L_2)L_3 = L_1 L_3 \cup L_2 L_3$.

Proof. We show the proof of Part (1); the proof of the other part is similar. Suppose $L_1, L_2, L_3 \in \mathbf{Lan}$. It will suffice to show that

$$L_1(L_2 \cup L_3) \subseteq L_1 L_2 \cup L_1 L_3 \subseteq L_1(L_2 \cup L_3).$$

Distributivity (Cont.)

Proof (cont.). To see that $L_1(L_2 \cup L_3) \subseteq L_1L_2 \cup L_1L_3$, suppose $w \in L_1(L_2 \cup L_3)$. We must show that $w \in L_1L_2 \cup L_1L_3$.

Distributivity (Cont.)

Proof (cont.). To see that $L_1(L_2 \cup L_3) \subseteq L_1L_2 \cup L_1L_3$, suppose $w \in L_1(L_2 \cup L_3)$. We must show that $w \in L_1L_2 \cup L_1L_3$. By our assumption, $w = xy$ for some $x \in L_1$ and $y \in L_2 \cup L_3$. There are two cases to consider.

Distributivity (Cont.)

Proof (cont.). To see that $L_1(L_2 \cup L_3) \subseteq L_1L_2 \cup L_1L_3$, suppose $w \in L_1(L_2 \cup L_3)$. We must show that $w \in L_1L_2 \cup L_1L_3$. By our assumption, $w = xy$ for some $x \in L_1$ and $y \in L_2 \cup L_3$. There are two cases to consider.

- Suppose $y \in L_2$. Then $w = xy \in L_1L_2 \subseteq L_1L_2 \cup L_1L_3$.
- Suppose $y \in L_3$. Then $w = xy \in L_1L_3 \subseteq L_1L_2 \cup L_1L_3$.

Distributivity (Cont.)

Proof (cont.). To see that $L_1 L_2 \cup L_1 L_3 \subseteq L_1(L_2 \cup L_3)$, suppose $w \in L_1 L_2 \cup L_1 L_3$. We must show that $w \in L_1(L_2 \cup L_3)$. There are two cases to consider.

□

Distributivity (Cont.)

Proof (cont.). To see that $L_1L_2 \cup L_1L_3 \subseteq L_1(L_2 \cup L_3)$, suppose $w \in L_1L_2 \cup L_1L_3$. We must show that $w \in L_1(L_2 \cup L_3)$. There are two cases to consider.

- Suppose $w \in L_1L_2$. Then $w = xy$ for some $x \in L_1$ and $y \in L_2$. Thus $y \in L_2 \cup L_3$, so that $w = xy \in L_1(L_2 \cup L_3)$.
- Suppose $w \in L_1L_3$. Then $w = xy$ for some $x \in L_1$ and $y \in L_3$. Thus $y \in L_2 \cup L_3$, so that $w = xy \in L_1(L_2 \cup L_3)$.

□

Distributivity (Cont.)

Proposition 3.2.7

(1) For all $\alpha, \beta, \gamma \in$ **Reg**, $\alpha(\beta + \gamma) \approx \alpha\beta + \alpha\gamma$.

(2) For all $\alpha, \beta, \gamma \in$ **Reg**, $(\alpha + \beta)\gamma \approx \alpha\gamma + \beta\gamma$.

Proof. Follows from Proposition 3.2.6. Consider, e.g., the proof of Part (1). By Proposition 3.2.6(1), we have that

$$\begin{aligned} L(\alpha(\beta+\gamma)) &= L(\alpha)L(\beta+\gamma) \\ &= L(\alpha)(L(\beta) \cup L(\gamma)) \\ &= \\ &= \\ &= \end{aligned}$$

Thus $\alpha(\beta + \gamma) \approx \alpha\beta + \alpha\gamma$. □

Distributivity (Cont.)

Proposition 3.2.7

*(1) For all $\alpha, \beta, \gamma \in$ **Reg**, $\alpha(\beta + \gamma) \approx \alpha\beta + \alpha\gamma$.*

*(2) For all $\alpha, \beta, \gamma \in$ **Reg**, $(\alpha + \beta)\gamma \approx \alpha\gamma + \beta\gamma$.*

Proof. Follows from Proposition 3.2.6. Consider, e.g., the proof of Part (1). By Proposition 3.2.6(1), we have that

$$\begin{aligned} L(\alpha(\beta+\gamma)) &= L(\alpha)L(\beta+\gamma) \\ &= L(\alpha)(L(\beta) \cup L(\gamma)) \\ &= L(\alpha)L(\beta) \cup L(\alpha)L(\gamma) \\ &= L(\alpha\beta) \cup L(\alpha\gamma) \\ &= L(\alpha\beta + \alpha\gamma) \end{aligned}$$

Thus $\alpha(\beta + \gamma) \approx \alpha\beta + \alpha\gamma$. □

Inclusions for Kleene Closure

Proposition 3.2.8

- *For all* $L \in \textbf{Lan}$, $LL^* \subseteq L^*$.
- *For all* $L \in \textbf{Lan}$, $L^*L \subseteq L^*$.

Proof. E.g., to see that $LL^* \subseteq L^*$, suppose $w \in LL^*$. Then $w = xy$ for some $x \in L$ and $y \in L^*$. Hence $y \in L^n$ for some $n \in \mathbb{N}$. Thus $w = xy \in LL^n = L^{n+1} \subseteq L^*$. □

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* =$

(2) $\{\%\}^* =$

(3) *For all* $L \in \textbf{Lan}$, $L^* L =$

(4) *For all* $L \in \textbf{Lan}$, $L^* L^* =$

(5) *For all* $L \in \textbf{Lan}$, $(L^*)^* =$

(6) *For all* $L_1 L_2 \in \textbf{Lan}$, $(L_1 L_2)^* L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* =$

(3) *For all* $L \in \textbf{Lan}$, $L^* L =$

(4) *For all* $L \in \textbf{Lan}$, $L^* L^* =$

(5) *For all* $L \in \textbf{Lan}$, $(L^*)^* =$

(6) *For all* $L_1 L_2 \in \textbf{Lan}$, $(L_1 L_2)^* L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \mathbf{Lan}$, $L^* L =$

(4) *For all* $L \in \mathbf{Lan}$, $L^* L^* =$

(5) *For all* $L \in \mathbf{Lan}$, $(L^*)^* =$

(6) *For all* $L_1 L_2 \in \mathbf{Lan}$, $(L_1 L_2)^* L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \mathbf{Lan}$, $L^*L = LL^*$.

(4) *For all* $L \in \mathbf{Lan}$, $L^*L^* =$

(5) *For all* $L \in \mathbf{Lan}$, $(L^*)^* =$

(6) *For all* $L_1L_2 \in \mathbf{Lan}$, $(L_1L_2)^*L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \mathbf{Lan}$, $L^* L = L L^*$.

(4) *For all* $L \in \mathbf{Lan}$, $L^* L^* = L^*$.

(5) *For all* $L \in \mathbf{Lan}$, $(L^*)^* =$

(6) *For all* $L_1 L_2 \in \mathbf{Lan}$, $(L_1 L_2)^* L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \textbf{Lan}$, $L^*L = LL^*$.

(4) *For all* $L \in \textbf{Lan}$, $L^*L^* = L^*$.

(5) *For all* $L \in \textbf{Lan}$, $(L^*)^* = L^*$.

(6) *For all* $L_1L_2 \in \textbf{Lan}$, $(L_1L_2)^*L_1 =$

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \mathbf{Lan}$, $L^*L = LL^*$.

(4) *For all* $L \in \mathbf{Lan}$, $L^*L^* = L^*$.

(5) *For all* $L \in \mathbf{Lan}$, $(L^*)^* = L^*$.

(6) *For all* $L_1L_2 \in \mathbf{Lan}$, $(L_1L_2)^*L_1 = L_1(L_2L_1)^*$.

Equivalences for Kleene Closure

Proposition 3.2.9

(1) $\emptyset^* = \{\%\}$.

(2) $\{\%\}^* = \{\%\}$.

(3) *For all* $L \in \mathbf{Lan}$, $L^*L = LL^*$.

(4) *For all* $L \in \mathbf{Lan}$, $L^*L^* = L^*$.

(5) *For all* $L \in \mathbf{Lan}$, $(L^*)^* = L^*$.

(6) *For all* $L_1 L_2 \in \mathbf{Lan}$, $(L_1L_2)^*L_1 = L_1(L_2L_1)^*$.

Proof. The six parts can be proven in order using Proposition 3.2.1. All parts but (2), (5) and (6) can be proved without using induction.

As an example, we show the proof of Part (5). To show that $(L^*)^* = L^*$, it will suffice to show that $(L^*)^* \subseteq L^* \subseteq (L^*)^*$.

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$,

□

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$, $(L^*)^n \subseteq L^*$.

□

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$, $(L^*)^n \subseteq L^*$.

- **(Basis Step)** We have that $(L^*)^0 = \{\%\} = L^0 \subseteq L^*$.

□

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$, $(L^*)^n \subseteq L^*$.

- **(Basis Step)** We have that $(L^*)^0 = \{\%\} = L^0 \subseteq L^*$.
- **(Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $(L^*)^n \subseteq L^*$. We must show that $(L^*)^{n+1} \subseteq L^*$. By the inductive hypothesis, Proposition 3.2.1(4) and Part (4), we have that $(L^*)^{n+1} = L^*(L^*)^n \subseteq L^*L^* = L^*$.

□

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$, $(L^*)^n \subseteq L^*$.

- **(Basis Step)** We have that $(L^*)^0 = \{\%\} = L^0 \subseteq L^*$.
- **(Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $(L^*)^n \subseteq L^*$. We must show that $(L^*)^{n+1} \subseteq L^*$. By the inductive hypothesis, Proposition 3.2.1(4) and Part (4), we have that $(L^*)^{n+1} = L^*(L^*)^n \subseteq L^*L^* = L^*$.

Now, we use the result of the induction to prove that $(L^*)^* \subseteq L^*$. Suppose $w \in (L^*)^*$. We must show that $w \in L^*$. Since $w \in (L^*)^*$, we have that $w \in (L^*)^n$ for some $n \in \mathbb{N}$. Thus, by the result of the induction, $w \in (L^*)^n \subseteq L^*$.

□

Equivalences for Kleene Closure (Cont.)

Proof (cont.). To see that $(L^*)^* \subseteq L^*$, we use mathematical induction to show that, for all $n \in \mathbb{N}$, $(L^*)^n \subseteq L^*$.

- **(Basis Step)** We have that $(L^*)^0 = \{\%\} = L^0 \subseteq L^*$.
- **(Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $(L^*)^n \subseteq L^*$. We must show that $(L^*)^{n+1} \subseteq L^*$. By the inductive hypothesis, Proposition 3.2.1(4) and Part (4), we have that $(L^*)^{n+1} = L^*(L^*)^n \subseteq L^*L^* = L^*$.

Now, we use the result of the induction to prove that $(L^*)^* \subseteq L^*$. Suppose $w \in (L^*)^*$. We must show that $w \in L^*$. Since $w \in (L^*)^*$, we have that $w \in (L^*)^n$ for some $n \in \mathbb{N}$. Thus, by the result of the induction, $w \in (L^*)^n \subseteq L^*$.

For the other inclusion, we have that $L^* = (L^*)^1 \subseteq (L^*)^*$. □

Equivalences for Kleene Closure (Cont.)

Proposition 3.2.11

(1) $\mathsf{\$}^* \approx \mathsf{\%}$.

(2) $\mathsf{\%}^* \approx \mathsf{\%}$.

(3) *For all* $\alpha \in \textbf{Reg}$, $\alpha^*\alpha \approx \alpha\,\alpha^*$.

(4) *For all* $\alpha \in \textbf{Reg}$, $\alpha^*\alpha^* \approx \alpha^*$.

(5) *For all* $\alpha \in \textbf{Reg}$, $(\alpha^*)^* \approx \alpha^*$.

(6) *For all* $\alpha, \beta \in \textbf{Reg}$, $(\alpha\beta)^*\alpha \approx \alpha(\beta\alpha)^*$.

Equivalences for Kleene Closure (Cont.)

Proposition 3.2.11

(1) $\$^* \approx \%$.

(2) $\%^* \approx \%$.

(3) *For all* $\alpha \in \textbf{Reg}$, $\alpha^*\alpha \approx \alpha\,\alpha^*$.

(4) *For all* $\alpha \in \textbf{Reg}$, $\alpha^*\alpha^* \approx \alpha^*$.

(5) *For all* $\alpha \in \textbf{Reg}$, $(\alpha^*)^* \approx \alpha^*$.

(6) *For all* $\alpha, \beta \in \textbf{Reg}$, $(\alpha\beta)^*\alpha \approx \alpha(\beta\alpha)^*$.

Proof. Follows from Proposition 3.2.9. Consider, e.g., the proof of Part (5). By Proposition 3.2.9(5), we have that

$$L((\alpha^*)^*) = L(\alpha^*)^* = (L(\alpha)^*)^* = L(\alpha)^* = L(\alpha^*).$$

Thus $(\alpha^*)^* \approx \alpha^*$. □

Proving the Correctness of Regular Expressions

We look at the harder of two regular expression synthesis and proof of correctness examples.

Define

$$\begin{aligned} &A = \{001, 011, 101, 111\}, \text{ and} \\ &B = \{\, w \in \{0,1\}^* \mid \text{for all } x, y \in \{0,1\}^*, \text{ if } w = x0y, \\ &\qquad\qquad \text{then there is a } z \in A \text{ such that } z \text{ is a prefix of } y \,\}. \end{aligned}$$

Proving the Correctness of Regular Expressions

We look at the harder of two regular expression synthesis and proof of correctness examples.

Define

$$A = \{001, 011, 101, 111\}, \text{ and}$$
$$B = \{\, w \in \{0,1\}^* \mid \text{for all } x, y \in \{0,1\}^*, \text{ if } w = x0y, \text{ then there is a } z \in A \text{ such that } z \text{ is a prefix of } y \,\}.$$

So B consists of those strings of 0's and 1's in which every occurrence of 0 is

Proving the Correctness of Regular Expressions

We look at the harder of two regular expression synthesis and proof of correctness examples.

Define

$$A = \{001, 011, 101, 111\}, \text{ and}$$
$$B = \{\, w \in \{0,1\}^* \mid \text{for all } x, y \in \{0,1\}^*, \text{ if } w = x0y, \text{ then there is a } z \in A \text{ such that } z \text{ is a prefix of } y \,\}.$$

So B consists of those strings of 0's and 1's in which every occurrence of 0 is immediately followed by an element of A.

Proving the Correctness of Regular Expressions

We look at the harder of two regular expression synthesis and proof of correctness examples.

Define

$$A = \{001, 011, 101, 111\}, \text{ and}$$
$$B = \{\, w \in \{0,1\}^* \mid \text{for all } x, y \in \{0,1\}^*, \text{ if } w = x0y,$$
$$\text{then there is a } z \in A \text{ such that } z \text{ is a prefix of } y \,\}.$$

So B consists of those strings of 0's and 1's in which every occurrence of 0 is immediately followed by an element of A.

We will find a regular expression that generates B, and prove it correct.

Synthesis

E.g.:

- %

Synthesis

E.g.:

- % is in B;
- 00111

Synthesis

E.g.:

- % is in B;
- 00111 is in B;
- 0000111

Synthesis

E.g.:

- % is in *B*;
- 00111 is in *B*;
- 0000111 is not in *B*; and
- 011

Synthesis

E.g.:

- % is in B;
- 00111 is in B;
- 0000111 is not in B; and
- 011 is not in B.

Synthesis

E.g.:

- $\%$ is in B;
- 00111 is in B;
- 0000111 is not in B; and
- 011 is not in B.

Note that, for all $x, y \in B$, xy

Synthesis

E.g.:

- $\%$ is in B;
- 00111 is in B;
- 0000111 is not in B; and
- 011 is not in B.

Note that, for all $x, y \in B$, $xy \in B$, i.e., $BB \subseteq B$.

Synthesis

E.g.:

- % is in B;
- 00111 is in B;
- 0000111 is not in B; and
- 011 is not in B.

Note that, for all $x, y \in B$, $xy \in B$, i.e., $BB \subseteq B$.

Furthermore, for all strings x, y, if $xy \in B$, then

Synthesis

E.g.:

- % is in B;
- 00111 is in B;
- 0000111 is not in B; and
- 011 is not in B.

Note that, for all $x, y \in B$, $xy \in B$, i.e., $BB \subseteq B$.

Furthermore, for all strings x, y, if $xy \in B$, then y is in B.

Synthesis (Cont.)

How should we go about finding a regular expression α such that $L(\alpha) = B$?

Synthesis (Cont.)

How should we go about finding a regular expression α such that $L(\alpha) = B$?

Because

- $\% \in B$,
- for all $x, y \in B$, $xy \in B$,
- for all strings x, y, if $xy \in B$ then $y \in B$,

our regular expression can have the form

Synthesis (Cont.)

How should we go about finding a regular expression α such that $L(\alpha) = B$?

Because

- $\% \in B$,
- for all $x, y \in B$, $xy \in B$,
- for all strings x, y, if $xy \in B$ then $y \in B$,

our regular expression can have the form β^*, where β generates all the strings that are

Synthesis (Cont.)

How should we go about finding a regular expression α such that $L(\alpha) = B$?
Because

- $\% \in B$,
- for all $x, y \in B$, $xy \in B$,
- for all strings x, y, if $xy \in B$ then $y \in B$,

our regular expression can have the form β^*, where β generates all the strings that are *basic* in the sense that they are nonempty elements of B with no non-empty proper prefixes that are in B.

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and $0000x$ is never basic.

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

0001

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

00011

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

001

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and $0000x$ is never basic.

000111 is the only basic string beginning with 000.

0011

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 0101

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000*x* is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111, 0101

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111, 010101

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and 0000x is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111, 01010111

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and $0000x$ is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111, 01010111, 0101010111, etc.

Synthesis (Cont.)

Clearly, 1 is basic, so there are no more basic strings that begin with 1.

But what about the basic strings beginning with 0?

No sequence of 0's is basic, and $0000x$ is never basic.

000111 is the only basic string beginning with 000.

00111 is the only basic string beginning with 001.

But what about the basic strings beginning with 01?

We have 0111, 010111, 01010111, 0101010111, etc.

Fortunately, there is a simple pattern here: we have all strings of the form $0(10)^n111$ for $n \in \mathbb{N}$.

Synthesis (Cont.)

By the above considerations, it seems that we can let our regular expression be

$$(1 + 0(10)^*111 + 00111 + 000111)^*.$$

Synthesis (Cont.)

By the above considerations, it seems that we can let our regular expression be

$$(1 + 0(10)^*111 + 00111 + 000111)^*.$$

But, using some of the equivalences we learned about above, we can turn this regular expression into

$$(1 + 0(0 + 00 + (10)^*)111)^*,$$

which we take as our α. Now, we prove that $L(\alpha) = B$.

Correctness Proof

Let

$$X = \{0\} \cup \{00\} \cup \{10\}^* \quad \text{and} \quad Y = \{1\} \cup \{0\}X\{111\}.$$

Then, we have that

$$\begin{aligned} X &= L(0 + 00 + (10)^*), \\ Y &= L(1 + 0(0 + 00 + (10)^*)111), \text{ and} \\ Y^* &= L((1 + 0(0 + 00 + (10)^*)111)^*) = L(\alpha). \end{aligned}$$

Correctness Proof

Let

$$X = \{0\} \cup \{00\} \cup \{10\}^* \quad \text{and} \quad Y = \{1\} \cup \{0\}X\{111\}.$$

Then, we have that

$$\begin{aligned} X &= L(0 + 00 + (10)^*), \\ Y &= L(1 + 0(0 + 00 + (10)^*)111), \text{ and} \\ Y^* &= L((1 + 0(0 + 00 + (10)^*)111)^*) = L(\alpha). \end{aligned}$$

Thus, it will suffice to show that $Y^* = B$. We will show that $Y^* \subseteq B \subseteq Y^*$.

Correctness Proof (Cont.)

Lemma 3.2.17
For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq B$.

Proof. We proceed by mathematical induction.

- **(Basis Step)** We have that $0111 \in B$. Hence $\{0\}\{10\}^0\{111\} = \{0\}\{\%\}\{111\} = \{0\}\{111\} = \{0111\} \subseteq B$.

Correctness Proof (Cont.)

Lemma 3.2.17

For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq B$.

Proof. We proceed by mathematical induction.

- **(Basis Step)** We have that $0111 \in B$. Hence $\{0\}\{10\}^0\{111\} = \{0\}\{\%\}\{111\} = \{0\}\{111\} = \{0111\} \subseteq B$.
- **Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $\{0\}\{10\}^n\{111\} \subseteq B$. We must show that $\{0\}\{10\}^{n+1}\{111\} \subseteq B$. Since

  $$\{0\}\{10\}^{n+1}\{111\} = \{0\}\{10\}\{10\}^n\{111\}$$

Correctness Proof (Cont.)

Lemma 3.2.17
For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq B$.

Proof. We proceed by mathematical induction.

- **(Basis Step)** We have that $0111 \in B$. Hence $\{0\}\{10\}^0\{111\} = \{0\}\{\%\}\{111\} = \{0\}\{111\} = \{0111\} \subseteq B$.
- **Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $\{0\}\{10\}^n\{111\} \subseteq B$. We must show that $\{0\}\{10\}^{n+1}\{111\} \subseteq B$. Since

$$\begin{aligned}\{0\}\{10\}^{n+1}\{111\} &= \{0\}\{10\}\{10\}^n\{111\} \\ &= \{01\}\{0\}\{10\}^n\{111\}\end{aligned}$$

Correctness Proof (Cont.)

Lemma 3.2.17

For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq B$.

Proof. We proceed by mathematical induction.

- **(Basis Step)** We have that $0111 \in B$. Hence $\{0\}\{10\}^0\{111\} = \{0\}\{\%\}\{111\} = \{0\}\{111\} = \{0111\} \subseteq B$.
- **Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $\{0\}\{10\}^n\{111\} \subseteq B$. We must show that $\{0\}\{10\}^{n+1}\{111\} \subseteq B$. Since

$$\begin{aligned}\{0\}\{10\}^{n+1}\{111\} &= \{0\}\{10\}\{10\}^n\{111\} \\ &= \{01\}\{0\}\{10\}^n\{111\} \\ &\subseteq \{01\}B && \text{(inductive hypothesis)},\end{aligned}$$

it will suffice to show that $\{01\}B \subseteq B$.

Correctness Proof (Cont.)

Lemma 3.2.17

For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq B$.

Proof. We proceed by mathematical induction.

- **(Basis Step)** We have that $0111 \in B$. Hence $\{0\}\{10\}^0\{111\} = \{0\}\{\%\}\{111\} = \{0\}\{111\} = \{0111\} \subseteq B$.
- **Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $\{0\}\{10\}^n\{111\} \subseteq B$. We must show that $\{0\}\{10\}^{n+1}\{111\} \subseteq B$. Since

$$\begin{aligned}\{0\}\{10\}^{n+1}\{111\} &= \{0\}\{10\}\{10\}^n\{111\} \\ &= \{01\}\{0\}\{10\}^n\{111\} \\ &\subseteq \{01\}B \qquad \text{(inductive hypothesis),}\end{aligned}$$

it will suffice to show that $\{01\}B \subseteq B$. But this is false!

Correctness Proof (Cont.)

Let

$$C = \{\, w \in B \mid 01 \text{ is a prefix of } w \,\}.$$

Lemma 3.2.17
For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq C$.

Correctness Proof (Cont.)

Let

$$C = \{\, w \in B \mid 01 \text{ is a prefix of } w \,\}.$$

Lemma 3.2.17

For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq C$.

Proof. . . . It will suffice to show that $\{01\}C \subseteq C$. Suppose $w \in \{01\}C$. We must show that $w \in C$. We have that $w = 01x$ for some $x \in C$.

Correctness Proof (Cont.)

Let

$$C = \{\, w \in B \mid 01 \text{ is a prefix of } w \,\}.$$

Lemma 3.2.17

For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq C$.

Proof. . . . It will suffice to show that $\{01\}C \subseteq C$. Suppose $w \in \{01\}C$. We must show that $w \in C$. We have that $w = 01x$ for some $x \in C$. Thus w begins with 01. It remains to show that $w \in B$. Since $x \in C$, we have that x begins with 01.

Correctness Proof (Cont.)

Let

$$C = \{\, w \in B \mid 01 \text{ is a prefix of } w \,\}.$$

Lemma 3.2.17
For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq C$.

Proof. . . . It will suffice to show that $\{01\}C \subseteq C$. Suppose $w \in \{01\}C$. We must show that $w \in C$. We have that $w = 01x$ for some $x \in C$. Thus w begins with 01. It remains to show that $w \in B$. Since $x \in C$, we have that x begins with 01. Thus the first occurrence of 0 in $w = 01x$ is followed by $101 \in A$. Furthermore, any other occurrence of 0 in $w = 01x$ is within x,

Correctness Proof (Cont.)

Let

$$C = \{\, w \in B \mid 01 \text{ is a prefix of } w \,\}.$$

Lemma 3.2.17
For all $n \in \mathbb{N}$, $\{0\}\{10\}^n\{111\} \subseteq C$.

Proof. . . . It will suffice to show that $\{01\}C \subseteq C$. Suppose $w \in \{01\}C$. We must show that $w \in C$. We have that $w = 01x$ for some $x \in C$. Thus w begins with 01. It remains to show that $w \in B$. Since $x \in C$, we have that x begins with 01. Thus the first occurrence of 0 in $w = 01x$ is followed by $101 \in A$. Furthermore, any other occurrence of 0 in $w = 01x$ is within x, and so is followed by an element of A because $x \in C \subseteq B$. Thus $w \in B$. □

Correctness Proof (Cont.)

Lemma 3.2.18

$Y \subseteq B$.

Proof. Uses Lemma 3.2.17. □

Correctness Proof (Cont.)

Lemma 3.2.18
$Y \subseteq B$.

Proof. Uses Lemma 3.2.17. □

Lemma 3.2.19
$Y^* \subseteq B$.

Proof. It will suffice to show that, for all $n \in \mathbb{N}$, $Y^n \subseteq B$, and we proceed by mathematical induction.

- **(Basis Step)** Since $\% \in B$, we have that $Y^0 = \{\%\} \subseteq B$.
- **(Inductive Step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $Y^n \subseteq B$. Then $Y^{n+1} = YY^n \subseteq BB \subseteq B$, by Lemma 3.2.18 and the inductive hypothesis.

□

Correctness Proof (Cont.)

Lemma 3.2.20
$B \subseteq Y^*$.

Proof. Since $B \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

Correctness Proof (Cont.)

Lemma 3.2.20

$B \subseteq Y^*$.

Proof. Since $B \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

We proceed by strong string induction. Suppose $w \in \{0,1\}^*$, and assume the inductive hypothesis: for all $x \in \{0,1\}^*$, if x is a proper substring of w, then

Correctness Proof (Cont.)

Lemma 3.2.20
$B \subseteq Y^*$.

Proof. Since $B \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

We proceed by strong string induction. Suppose $w \in \{0,1\}^*$, and assume the inductive hypothesis: for all $x \in \{0,1\}^*$, if x is a proper substring of w, then

$$\text{if } x \in B, \text{ then } x \in Y^*.$$

Correctness Proof (Cont.)

Lemma 3.2.20
$B \subseteq Y^*$.

Proof. Since $B \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

We proceed by strong string induction. Suppose $w \in \{0,1\}^*$, and assume the inductive hypothesis: for all $x \in \{0,1\}^*$, if x is a proper substring of w, then

$$\text{if } x \in B, \text{ then } x \in Y^*.$$

We must show that

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

Suppose $w \in B$. We must show that $w \in Y^*$.

Correctness Proof (Cont.)

Lemma 3.2.20
$B \subseteq Y^*$.

Proof. Since $B \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

We proceed by strong string induction. Suppose $w \in \{0,1\}^*$, and assume the inductive hypothesis: for all $x \in \{0,1\}^*$, if x is a proper substring of w, then

$$\text{if } x \in B, \text{ then } x \in Y^*.$$

We must show that

$$\text{if } w \in B, \text{ then } w \in Y^*.$$

Suppose $w \in B$. We must show that $w \in Y^*$. There are three main cases to consider. (See the book for more details.)

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000$

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000111t =$

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000111t = ((0)(00)(111))t \in YY^* \subseteq Y^*$, by the inductive hypothesis.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000111t = ((0)(00)(111))t \in YY^* \subseteq Y^*$, by the inductive hypothesis.
    - Suppose $y = 1z$ for some z, so $w = 001z$.

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000111t = ((0)(00)(111))t \in YY^* \subseteq Y^*$, by the inductive hypothesis.
    - Suppose $y = 1z$ for some z, so $w = 001z$. Thus there is a v such that $w = 001$

Correctness Proof (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w \in Y^0 \subseteq Y^*$.
- Suppose $w = 0x$ for some x.
  - Suppose $x = 0y$ for some y, so $w = 00y$.
    - Suppose $y = 0z$ for some z, so $w = 000z$. Thus, there is a t such that $w = 000111t = ((0)(00)(111))t \in YY^* \subseteq Y^*$, by the inductive hypothesis.
    - Suppose $y = 1z$ for some z, so $w = 001z$. Thus there is a v such that $w = 00111v = ((0)(0)(111))v \in YY^* \subseteq Y^*$, by the inductive hypothesis.

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$.

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010.

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
    - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u$

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u1$

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u111r = ((0)(10u)(111))r \in YY^* \subseteq Y^*$, by the inductive hypothesis.

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
    - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u111r = ((0)(10u)(111))r \in YY^* \subseteq Y^*$, by the inductive hypothesis.
    - Suppose $y = 1z$ for some z, so that $w = 011z$. Thus,

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u111r = ((0)(10u)(111))r \in YY^* \subseteq Y^*$, by the inductive hypothesis.
  - Suppose $y = 1z$ for some z, so that $w = 011z$. Thus, there is a u such that $w = 0111u = ((0)(\%)(111))u \in YY^* \subseteq Y^*$, by the inductive hypothesis.

□

Correctness Proof (Cont.)

Proof (cont.).

Suppose $w = 0x$ for some x. (Cont.)

- Suppose $x = 1y$ for some y, so $w = 01y$.
  - Suppose $y = 0z$ for some z, so that $w = 010z$. Let u be longest prefix of z in $\{10\}^*$, and v be such that $z = uv$. Thus $w = 010uv$, and $010u$ ends with 010. Thus, there is an r such that $w = 010u111r = ((0)(10u)(111))r \in YY^* \subseteq Y^*$, by the inductive hypothesis.
  - Suppose $y = 1z$ for some z, so that $w = 011z$. Thus, there is a u such that $w = 0111u = ((0)(\%)(111))u \in YY^* \subseteq Y^*$, by the inductive hypothesis.
- Suppose $w = 1x$ for some x. Then, $w = 1x \in YY^* \subseteq Y^*$, by the inductive hypothesis.

□