A homomorphism on an alphabet is a function that gives a string for each symbol in that alphabet.

Example: \( h(0) = ab, h(1) = \varepsilon. \)

Extend to strings by \( h(a_1a_2\ldots a_n) = h(a_1)h(a_2)\ldots h(a_n). \)

Example: \( h(01010) = ababab. \)
Closure Properties: Homomorphism

- If $L$ is regular, and $h$ is a homomorphism on its alphabet, then $h(L) = \{ h(w) \mid w \in L \}$ is also regular.
- **Proof:** Let $E$ be a regular expression for $L$.
  - Apply $h$ to each symbol in $E$.
  - Language of resulting RE is $h(L)$. 

Mridul Aanjaneya  
Automata Theory 3/23
Closure under Homomorphism: Example

- Let $h(0) = ab$, $h(1) = \varepsilon$.
- Let $L$ be the language of a regular expression $01^* + 10^*$.
- Then $h(L)$ is the language of regular expression $ab\varepsilon^* + \varepsilon(ab)^*$.
- **Note:** $ab\varepsilon^* + \varepsilon(ab)^* = ab + (ab)^* = (ab)^*$. 
Inverse Homomorphisms

- Let \( h \) be a homomorphism and \( L \) a language whose alphabet is the output language of \( h \).
- \( h^{-1}(L) = \{ w \mid h(w) \in L \} \).
Inverse Homomorphisms: Example

- Let $h(0) = ab$, $h(1) = \varepsilon$.
- Let $L = \{abab, baba\}$.
- $h^{-1}(L) =$ the language with two 0's and any number of 1's $= L(1^*01^*01^*)$.
- **Note:** No string maps to $baba$, any string with exactly two 0's maps to $abab$. 
• Start with a DFA $A$ for $L$.
• Construct a DFA $B$ for $h^{-1}(L)$ with:
  • The same set of states.
  • The same start state.
  • The same final states.
  • Input alphabet = the symbols to which the homomorphism $h$ applies.
• The transitions for $B$ are computed by applying $h$ to an input symbol $a$ and seeing where $A$ would go on sequence of input symbols $h(a)$.
• Formally, $\delta_B(q,a) = \delta_A(q,h(a))$. 
• $h(0) = ab$, $h(1) = \varepsilon$. 
Closure Proof for Inverse Homomorphism

- **Induction** on $|w|$ shows that $\delta_B(q_0,w) = \delta_A(q_0,h(w))$.

- **Basis:** $w = \varepsilon$.
  \[
  \delta_B(q_0,\varepsilon) = q_0, \text{ and } \delta_A(q_0,h(\varepsilon)) = \delta_A(q_0,\varepsilon) = q_0.
  \]

- **Inductive Step:** Let $w = xa$, assume **IH** for $x$.

  - $\delta_B(q_0,w) = \delta_B(\delta_B(q_0,x),a) = \delta_B(\delta_A(q_0,h(x)),a)$ (from **IH**)
  - $= \delta_A(\delta_A(q_0,h(x)),h(a))$ (by definition of $B$)
  - $= \delta_A(q_0,h(x)h(a))$ (by definition of extended $\delta_A$)
  - $= \delta_A(q_0,h(w))$ (by definition of $h$)
Decision Property: Equivalence

- Given regular languages $L$ and $M$, is $L = M$?
- Algorithm involves constructing the product DFA $P$ from DFA’s for $L$ and $M$.
- Make the final states of $P$ be those states $[q,r]$ such that exactly one of $q$ and $r$ is a final state of its own DFA.
- Thus, $P$ accepts $w$ iff $w$ is in exactly one of $L$ and $M$.
- The product DFA’s language is empty iff $L = M$.

**Note:** We already have a better algorithm to test emptiness.
Minimum State DFA

- In principle, since we can test for equivalence of DFA’s we can, given a DFA A find the DFA with the fewest states accepting $L(A)$.
- Test all smaller DFA’s for equivalence with A.
- But thats a terrible algorithm.
Efficient State Minimization

- Construct a table with all pairs of states.
- If you find a string that distinguishes two states (takes exactly one to an accepting state), mark that pair.
- Algorithm is a recursion on the length of the shortest distinguishing string.
Efficient State Minimization

- **Basis:** Mark a pair if exactly one is a final state.
- **Induction:** Mark \([q,r]\) if there is some input symbol \(a\) such that \([\delta(q,a),\delta(r,a)]\) is marked.
- After no more marks are possible, the unmarked states are equivalent and can be merged into one state.
Suppose $q_1, \ldots, q_k$ are indistinguishable states.
Replace them by one state $q$.
Then $\delta(q_1, a), \ldots, \delta(q_k, a)$ are all indistinguishable states, otherwise, we should have marked at least one more pair.
Let $\delta(q, a) =$ the representative state for that group.
Eliminating Indistinguishable States

- Unfortunately, combining indistinguishable states could leave us with *unreachable* states in the *minimum-state* DFA.
- Thus, before or after, *remove* states that are not reachable from the start state.
Example: State Minimization

A B
\[0 \quad 1 \quad 0,1\]

C D
\[1 \quad 1 \quad 0\]

\[[A,C] \quad [B,C] \quad [B,D] \quad [A,D]\]
Proposition

If state $p$ is indistinguishable from $q$, and $q$ is indistinguishable from $r$, then $p$ is indistinguishable from $r$.

- **Proof:** The outcome (accept or don’t) of $p$ and $q$ on input $w$ is the same, and the outcome of $q$ and $r$ on $w$ is the same, then likewise the outcome of $p$ and $r$. 
• We have combined states of the given DFA wherever possible.
• Could there be another, completely unrelated DFA with fewer states?
• No. The proof involves minimizing the DFA we derived with the hypothetical better DFA.
Proof: No Unrelated, Smaller DFA

- Let $A$ be our minimized DFA; let $B$ be a smaller equivalent.
- Consider an automaton with the states of $A$ and $B$ combined.
- Use distinguishable in its contrapositive form:
  - If states $q$ and $p$ are indistinguishable, so are $\delta(q,a)$ and $\delta(p,a)$. 

Mridul Aanjaneya
Inductive Hypothesis

Hypothesis

Every state \( q \) of \( A \) is indistinguishable from some state of \( B \).

- Induction is on the length of the shortest string taking you from the start state of \( A \) to \( q \).
• **Basis:** Start states of $A$ and $B$ are indistinguishable, because $L(A) = L(B)$.

• **Induction:** Suppose $w = xa$ is a shortest string getting $A$ to state $q$.

• By the **IH**, $x$ gets $A$ to some state $r$ that is indistinguishable from some state $p$ of $B$.

• Then $\delta(r,a) = q$ is indistinguishable from $\delta(p,a)$. 
• **Key idea:** Two states of A **cannot** be indistinguishable from the same state of B, or they would be indistinguishable from each other.
  - But A is already a **minimum** state DFA!
• Thus, B has **at least** as many states as A.
Fibonacci Numbers and the Golden Ratio

\[ F_n = F_{n-1} + F_{n-2} \]

\[ \Rightarrow F_n = \frac{1}{\sqrt{5}} \left\{ \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right\} \]