COL 352 Introduction to Automata and Theory of Computation

Nikhil Balaji

Bharti 420
Indian Institute of Technology, Delhi
nbalaji@cse.iitd.ac.in

January 12, 2023

Lecture 5: Nondeterminism: Subset Construction

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language.

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

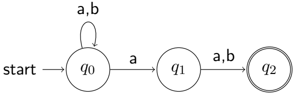
We will work it out for an example.

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

We will work it out for an example.

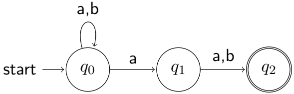


Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

We will work it out for an example.

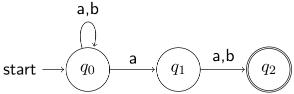


Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

We will work it out for an example.



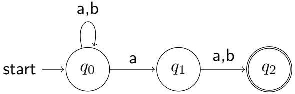
$$\varnothing \quad \{0\} \quad \{1\} \quad \{2\} \quad \{0,1\} \quad \{0,2\} \quad \{1,2\} \quad \{0,1,2\}$$

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

We will work it out for an example.



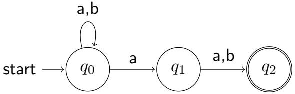
	$\varnothing$	$\{0\}$	$\{1\}$	$\{2\}$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$	$\{0,1,2\}$
a	$\varnothing$	$\{0,1\}$	$\{2\}$	$\varnothing$	$\{0,1,2\}$	$\{0,1\}$	$\{2\}$	$\{0,1,2\}$

Power of NFAs

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

We will work it out for an example.



	$\varnothing$	$\{0\}$	$\{1\}$	$\{2\}$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$	$\{0,1,2\}$
a	$\varnothing$	$\{0,1\}$	$\{2\}$	$\varnothing$	$\{0,1,2\}$	$\{0,1\}$	$\{2\}$	$\{0,1,2\}$
b	$\varnothing$	$\{0\}$	$\{2\}$	$\varnothing$	$\{0,2\}$	$\{0\}$	$\{2\}$	$\{0,2\}$

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language.

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Proof.

Let $A = (Q, \Sigma, q_0, F, \delta)$. We will construct a DFA $A' = (Q', \Sigma, q_0', F', \Delta)$ such that $L(A') = L(A)$.

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Proof.

Let $A = (Q, \Sigma, q_0, F, \delta)$. We will construct a DFA $A' = (Q', \Sigma, q'_0, F', \Delta)$ such that $L(A') = L(A)$.

Subset construction

$$Q' = 2^Q$$

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Proof.

Let $A = (Q, \Sigma, q_0, F, \delta)$. We will construct a DFA $A' = (Q', \Sigma, q_0', F', \Delta)$ such that $L(A') = L(A)$.

Subset construction

$$Q' = 2^Q,$$

$$q_0' = \{q_0\}$$

Subset construction

Lemma

Let A be an NFA. Then $L(A)$ is a regular language. That is, NFA and DFA accept the same set of languages.

Proof.

Let $A = (Q, \Sigma, q_0, F, \delta)$. We will construct a DFA $A' = (Q', \Sigma, q'_0, F', \Delta)$ such that $L(A') = L(A)$.

Subset construction

$Q' = 2^Q$,

$q'_0 = \{q_0\}$,

$F' = \{S \subseteq Q \mid S \cap F \neq \varnothing\}$.

For each $S \subseteq Q, \Delta(S, a) = \bigcup_{p \in S} \delta(p, a)$.

$\square$

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Definition

Let $A = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
Let $\hat{\delta} : Q \times \Sigma^* \subseteq Q$ be defined as follows:

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Definition

Let $A = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
Let $\hat{\delta} : Q \times \Sigma^* \subseteq Q$ be defined as follows:
Let $q \in Q$

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Definition

Let $A = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
Let $\hat{\delta} : Q \times \Sigma^* \subseteq Q$ be defined as follows:
Let $q \in Q$

$$\hat{\delta}(q, \epsilon) := q$$

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Definition

Let $A = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
Let $\hat{\delta} : Q \times \Sigma^* \subseteq Q$ be defined as follows:
Let $q \in Q$

$$\hat{\delta}(q, \epsilon) := q$$

$$\hat{\delta}(q, xa) := \{p \mid \exists q' \in \hat{\delta}(q, x) \text{ s.t. } p \in \delta(q', a)\}$$

Correctness of subset construction

We will prove the following first.

Lemma

For any NFA A, there is an equivalent DFA A' such that $L(A') = L(A)$.

Definition

Let $A = (Q, \Sigma, \delta, q_0, F)$ be an NFA.
Let $\hat{\delta} : Q \times \Sigma^* \subseteq Q$ be defined as follows:
Let $q \in Q$

$$\hat{\delta}(q, \epsilon) := q$$

$$\hat{\delta}(q, xa) := \{p \mid \exists q' \in \hat{\delta}(q, x) \text{ s.t. } p \in \delta(q', a)\}$$

Proof of Correctness

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$
Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$

Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

- $\hat{\delta}(q_0, x) = \hat{\Delta}(\{q_0\}, x) = S$ (say) (By Induction Hypothesis)

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$
Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

- $\hat{\delta}(q_0, x) = \hat{\Delta}(\{q_0\}, x) = S$ (say) (By Induction Hypothesis)
- $\hat{\delta}(q_0, xa) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\delta}$).

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$
Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

- $\hat{\delta}(q_0, x) = \hat{\Delta}(\{q_0\}, x) = S$ (say) (By Induction Hypothesis)
- $\hat{\delta}(q_0, xa) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\delta}$).
- $\hat{\Delta}(\{q_0\}, xa) = \Delta(S, a) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\Delta}$).

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$
Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

- $\hat{\delta}(q_0, x) = \hat{\Delta}(\{q_0\}, x) = S$ (say) (By Induction Hypothesis)
- $\hat{\delta}(q_0, xa) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\delta}$).
- $\hat{\Delta}(\{q_0\}, xa) = \Delta(S, a) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\Delta}$).
- Therefore, $\hat{\delta}(q_0, xa) = \hat{\Delta}(\{q_0\}, xa)$.

Proof of correctness

Lemma

For all $w \in \Sigma^*$, $\hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$.

Proof.

Base case: Let $w = \epsilon$. $\hat{\delta}(\{q_0\}, \epsilon) = q_0 = \Delta(\{q_0\}, \epsilon) = \hat{\Delta}(\{q_0\}, \epsilon)$
Induction step: Let $w = x \cdot a$ where $w \in \Sigma^*$, $a \in \Sigma$

- $\hat{\delta}(q_0, x) = \hat{\Delta}(\{q_0\}, x) = S$ (say) (By Induction Hypothesis)
- $\hat{\delta}(q_0, xa) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\delta}$).
- $\hat{\Delta}(\{q_0\}, xa) = \Delta(S, a) = \bigcup_{q \in S} \delta(q, a)$ (By definition of $\hat{\Delta}$).
- Therefore, $\hat{\delta}(q_0, xa) = \hat{\Delta}(\{q_0\}, xa)$.

$\square$

From Lemma, we have for all $w \in \Sigma^*, \hat{\delta}(q_0, w) = \hat{\Delta}(\{q_0\}, w)$. Hence,for all $w, \hat{\delta}(q_0, w) \cap F \neq \varnothing$ iff $\hat{\Delta}(\{q_0\}, w) \in F'$.

Complexity of Subset Construction

- If NFA has n states, then DFA has 2^n states.

Complexity of Subset Construction

- If NFA has n states, then DFA has 2^n states.
- But not all these states need to be reachable!

Complexity of Subset Construction

- If NFA has n states, then DFA has 2^n states.
- But not all these states need to be reachable!

Two questions

- Does this blowup really occur when only considering reachable states?
- On examples where it does not occur can we have a subset construction that is efficient?

Worst case blowup of Subset Construction

- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.

Worst case blowup of Subset Construction

- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.
- For each example in family show that if a smaller DFA than 2^n exists, there is a contradiction.

Worst case blowup of Subset Construction

- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.
- For each example in family show that if a smaller DFA than 2^n exists, there is a contradiction.

For each $n \in \mathbb{N}$, let

$$L_n = \{w \in \{0,1\}^* \mid n\text{-th symbol from the end is a } 1\}$$

Worst case blowup of Subset Construction

- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.
- For each example in family show that if a smaller DFA than 2^n exists, there is a contradiction.

For each $n \in \mathbb{N}$, let

$$L_n = \{w \in \{0,1\}^* \mid n\text{-th symbol from the end is a } 1\}$$

- Exercise: give an NFA for An which accepts L_n. How many states does it have?
- Can L_n be accepted by a DFA with less than 2^n states?

Worst case blowup of Subset Construction

- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.
- For each example in family show that if a smaller DFA than 2^n exists, there is a contradiction.

For each $n \in \mathbb{N}$, let

$$L_n = \{w \in \{0,1\}^* \mid n\text{-th symbol from the end is a } 1\}$$

- Exercise: give an NFA for An which accepts L_n. How many states does it have?
- Can L_n be accepted by a DFA with less than 2^n states?

$$L = \{x \in \{a\}^* \mid |x| \text{ is divisible by } 3 \text{ or } 5\}$$

Worst case blowup of Subset Construction

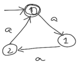
- Is one NFA enough? A family of examples such that for each n there is an example NFA parametrized by n.
- For each example in family show that if a smaller DFA than 2^n exists, there is a contradiction.

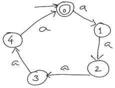
For each $n \in \mathbb{N}$, let

$$L_n = \{w \in \{0,1\}^* \mid n\text{-th symbol from the end is a } 1\}$$

- Exercise: give an NFA for An which accepts L_n. How many states does it have?
- Can L_n be accepted by a DFA with less than 2^n states?

divisible by 3 or $5\}$





Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \ldots a_n$ and $b_1 \ldots b_n$ that must both end in the same state q.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \ldots a_n$ and $b_1 \ldots b_n$ that must both end in the same state q.
- Suppose a_i and b_i are different and wlog let $a_i = 1, b_i = 0$.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \dots a_n$ and $b_1 \dots b_n$ that must both end in the same state q.
- Suppose a_i and b_i are different and wlog let $a_i = 1, b_i = 0$.
- Consider words

$$w = a_1 \dots a_i \dots a_n 0^{i-1} \in L_n$$

$$w' = b_1 \dots b_i \dots b_n 0^{i-1} \notin L_n$$

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \dots a_n$ and $b_1 \dots b_n$ that must both end in the same state q.
- Suppose a_i and b_i are different and wlog let $a_i = 1, b_i = 0$.
- Consider words

$$w = a_1 \dots a_i \dots a_n 0^{i-1} \in L_n$$

$$w' = b_1 \dots b_i \dots b_n 0^{i-1} \notin L_n$$

- But A is a DFA, so starting from q on reading 0^{i-1} we reach the same state.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \dots a_n$ and $b_1 \dots b_n$ that must both end in the same state q.
- Suppose a_i and b_i are different and wlog let $a_i = 1, b_i = 0$.
- Consider words

$$w = a_1 \dots a_i \dots a_n 0^{i-1} \in L_n$$

$$w' = b_1 \dots b_i \dots b_n 0^{i-1} \notin L_n$$

- But A is a DFA, so starting from q on reading 0^{i-1} we reach the same state. Hence A will either accept both w and w' or reject both.

Exponential Blowup of Subset Construction

Theorem

For all $n \in \mathbb{N}$, every DFA accepting L_n has at least 2^n states.

Proof.

- Suppose not. For some $n \in \mathbb{N}$ let L_n be accepted by DFA A with less than 2^n states.
- Then, there exist $a_1 \dots a_n$ and $b_1 \dots b_n$ that must both end in the same state q.
- Suppose a_i and b_i are different and wlog let $a_i = 1, b_i = 0$.
- Consider words

$$w = a_1 \dots a_i \dots a_n 0^{i-1} \in L_n$$

$$w' = b_1 \dots b_i \dots b_n 0^{i-1} \notin L_n$$

- But A is a DFA, so starting from q on reading 0^{i-1} we reach the same state. Hence A will either accept both w and w' or reject both. Contradiction!