

Clase #10 de 27

Especificación Formal Máquinas de Estado & Expresiones Regulares

Mayo 22, Lunes

Repaso Clase Anterior

- [Glosario](#)
- [Tareas](#)

Agenda para esta clase

- Especificación Formal de Máquinas de Estado
- Expresiones Regulares

Máquinas de Estado

Parte III

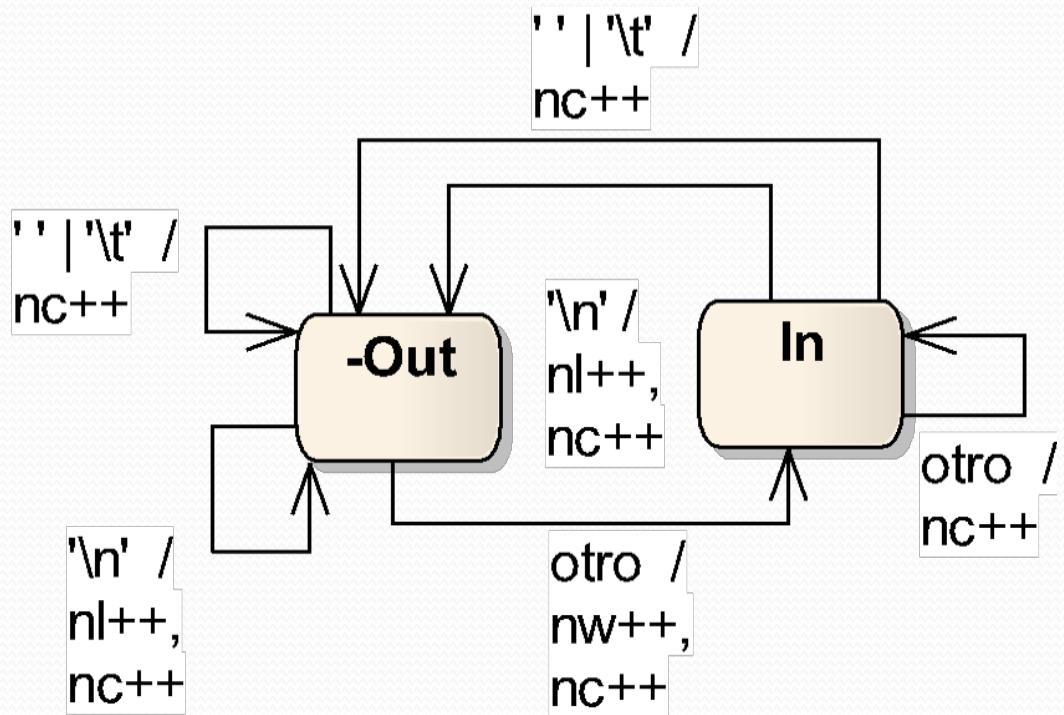
Especificación Formal

Máquinas de Estado para especificar comportamiento

- ¿Cuál es el comportamiento?
- ¿Cómo describirlo?
 - ¿Prosa?
 - ¿Caso de uso?
 - ¿Función?
- Dominio e Imagen
 - $F : \Sigma_i^* \rightarrow \Sigma_o^*$
- ¿Como representar la función?
 - Diagrama de Transiciones.

Formalización Matemática para Mealy

- ¿Es suficiente un Digrafo?
- ¿Es suficiente una Función?
 - $T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o$
- Par ordenando (2-upla)
 - $(T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o, \text{ inicial})$
- 5-upla
 - Elementos
 - Estados
 - Estado inicial
 - Entradas
 - Salidas
 - Función
 - Definición
 - $M = (Q, e_0, \Sigma_i, \Sigma_o, T)$
 - $T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o$
- 6-upla
 - Función Acción separada
 - $M = (Q, e_0, \Sigma_i, \Sigma_o, T, A)$
 - $T : Q \times \Sigma_i \rightarrow Q$
 - $A : Q \times \Sigma_i \rightarrow \Sigma_o$.

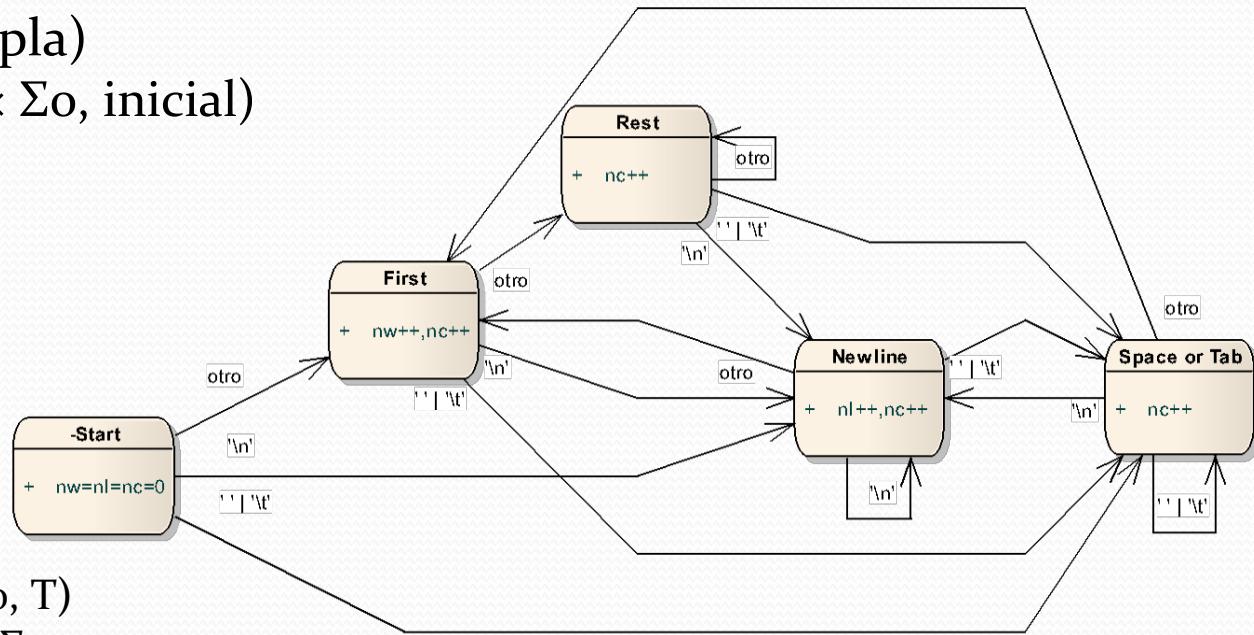


Diferencias Moore y Mealy

- Funciones
 - Mealy
 - $A : Q \times \Sigma_i \rightarrow \Sigma_o$
 - Moore
 - $A : Q \rightarrow \Sigma_o$
 - General
 - $T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o$
- Algoritmos particulares
 - Mealy
 - Seleccionar Estado
 - Seleccionar Carácter
 - Accionar
 - Actualizar Estado
 - Moore
 - Seleccionar Estado
 - Accionar
 - Seleccionar Carácter
 - Actualizar Estado
- Algoritmos generales
 - General
 - Seleccionar Estado
 - Accionar
 - Seleccionar Carácter
 - Accionar
 - Actualizar Estado
 - Un algoritmo un poco más abstracto
 - while there's another character
 - DoMoore(s)
 - DoMealy(s, c)
 - $s = \text{GetNextState}(s, c)$

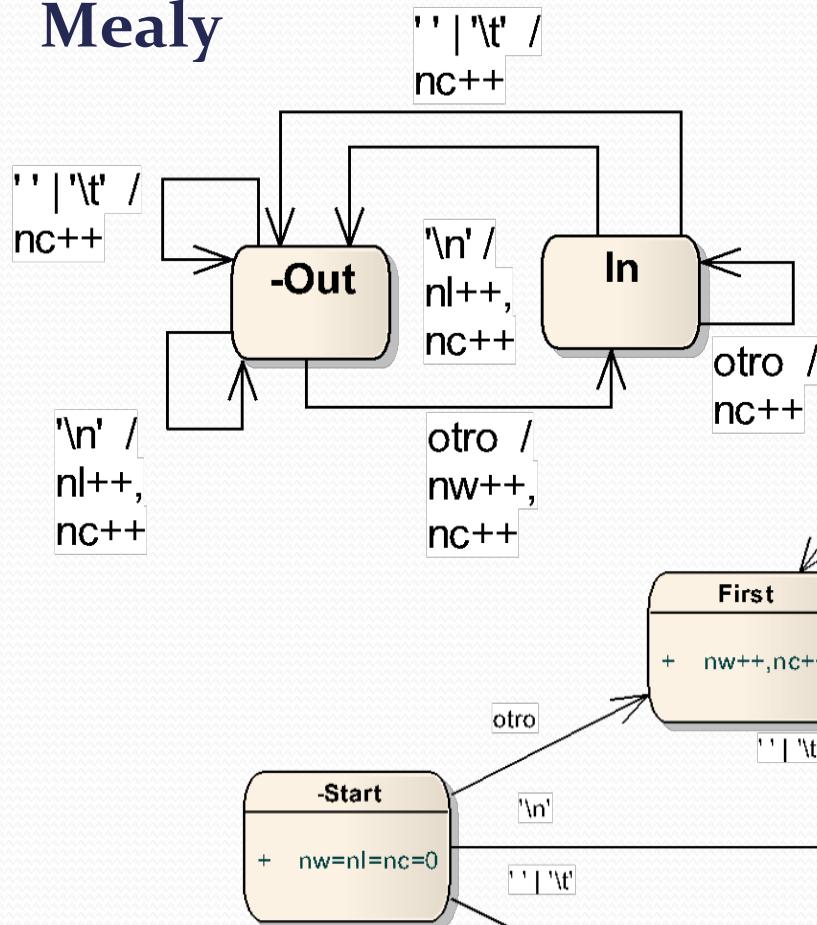
Formalización Matemática para Moore

- Par ordenando (2-upla)
 - $(T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o, \text{ inicial})$
- 5-upla
 - Elementos
 - Estados
 - Estado inicial
 - Entradas
 - Salidas
 - Función
 - Definición
 - $M = (Q, e_0, \Sigma_i, \Sigma_o, T)$
 - $T : Q \times \Sigma_i \rightarrow Q \times \Sigma_o$
- 6-upla
 - Función Acción separada
 - $M = (Q, e_0, \Sigma_i, \Sigma_o, T, A)$
 - $T : Q \times \Sigma_i \rightarrow Q$
 - $A : Q \rightarrow \Sigma_o$.

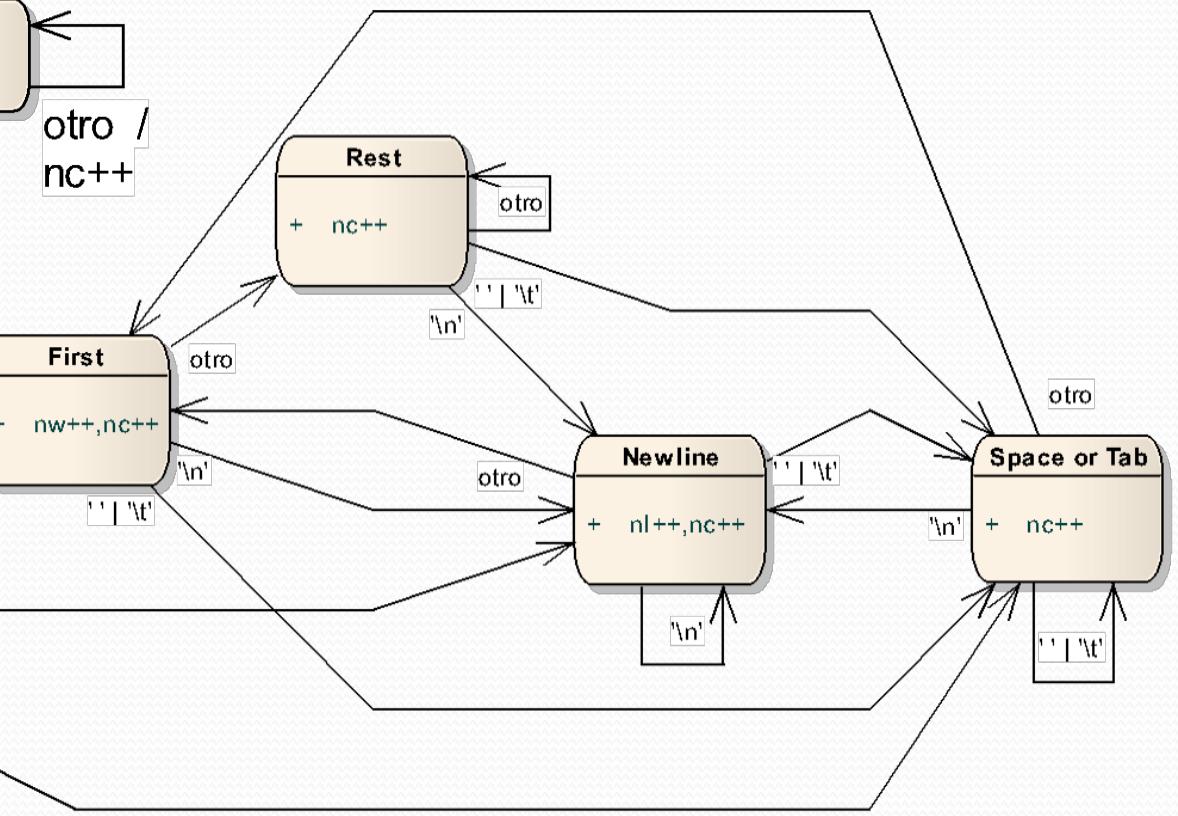


Moore y Mealy – Modelos Equivalentes

Mealy



Moore



Expresiones Regulares

Definición de los lenguajes de las Categorías Léxicas

- Palabras reservadas
- Puntuación
- Identificadores
- Constantes
 - Entero
 - Flotante
 - Carácter
 - Enumeración
- Literales de cadena.

Ejemplo – Enteros decimales

- $\{1,2,4,3,5,6,7,8,9\} \{0,1,2,3,4,5,6,7,8,9\}^* \{u,U,l,L,\varepsilon\}$
- Necesitamos una forma más compacta
- "Demasiadas llaves y comas".

Definición formal de una ER

- Sintaxis

- $a \in \Sigma$
- E y F representan cualquier ER
- \emptyset es una ER
- ε es una ER
- a es una ER
- $E+F$ es una ER
- EF es una ER
- E^* es una ER
- (E) es una ER

- Semántica

- $L(x)$ significa el LF asociado con x
- $L(\emptyset) = \{\} = \emptyset$
- $L(\varepsilon) = \{\varepsilon\}$
- $L(a) = \{a\}$
- $L(E+F) = L(E) \cup L(F)$
- $L(EF) = L(E) \times L(F)$
- $L(E^*) = L(E)^*$
- $L((E)) = (L(E))$

Aplicaciones

- Buscar patrones
- Procesar texto
- Herramientas
 - Grep
 - Clases Regex
- Sintaxis Léxica
 - Generación de Analizadores Léxicos.

Términos de la clase #10

Definir cada término con la bibliografía

- Especificación Formal de Máquinas de Estado
 - Comportamiento que expresan las máquinas de estado
 - Función Transición
 - Máquinas de estado como 2-upla
 - Máquinas de estado como 5-upla
 - Máquinas de estado como 6-upla
 - Máquinas de Estado de Mealy
 - Máquinas de Estado de Moore
 - Conjunto de Estados
 - Estado inicial
 - Alfabeto de Entrada
 - Alfabeto de Salida
 - Función Accionar o Acción
- Expresiones Regulares
 - Operaciones con Conjuntos
- Lenguaje Universal
- Clausura de Kleene
- Sintaxis de las ER
- Semántica de las ER
- ER Universal
- Aplicaciones de las ER

Tareas para la próxima clase

1. Continuar con el trabajo #3.



¿Consultas?



Fin de la clase

Clase #11 de 27

Lenguajes Formales y sus Componentes & Formalización de Autómatas Finitos

Mayo 29, Lunes

Repaso Clase Anterior

- Glosario
- Tareas

Agenda para esta clase

- Carácter, Alfabeto, y String
- Lenguajes Formales
- Intervalo
- Formalización de Autómatas Finitos
- Cardinalidad de Lenguajes y Resumen



Símbolo, Alfabeto & Cadena

Símbolo o Carácter

- Es Indivisible
- Es cualquier elemento que se defina como símbolo
- Pertenece a un Alfabeto (Σ)
- Ejemplo [Mvíci]3: El alfabeto $\Sigma = \{0, 1\}$ proporciona los caracteres para construir los números binarios
- Ejemplo: El $\Sigma = \{\text{if}, \text{for}\}$ tiene dos caracteres
- Ejercicio [Mvíci]1: Escriba el Σ para construir el conjunto de los números enteros con signo en base 10
- Operaciones
 - Concatenar dos caracteres
 - a con $b = ab$
 - Potencia7
 - $a^7 = aaaaaaa$
 - $\text{for}^7 = ?$
 - ¿Cerradas?

Alfabeto Σ

- Es un Conjunto, cuyos elementos son símbolos o caracteres
- Restricciones básicas
 - Finito
 - No vacío
- Ejemplos
 - $\{\} = \emptyset \Rightarrow$ no es alfabeto, por ser vacío
 - Naturales \Rightarrow no es alfabeto, por ser infinito
 - ASCII
 - Dígitos $= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
 - Meses $= \{\text{Lunes, Martes, Miércoles, Jueves, Viernes, Sábado}\}$
- Restricción avanzada
 - Sus elementos no pueden formarse por yuxtaposición de otros elementos
- Ejemplos
 - $\{0,1, 1001\}$
 - $\{\text{if, for}\}$
 - $\{\text{a, b, c, ch, d, e, f, g, h, i, j, k, l, ll, m, n, ñ, o, p, q, r, rr, s, t, u, v, w, x, y, z}\}$.

Caracteres y Alfabetos – Implementación en ANSI C

- Alfabeto tipo de dato char
- Carácter valor y variable
 - A
 - Literal 'A'
 - Literal 65
 - $c \in \Sigma$
 - char c;
- Concatenar dos caracteres
 - ¿Uso?
 - void DosCaracteres(int, int, char *);
- Potencia 7
 - ¿Uso?
 - void SieteVeces(int, char *);

Cadena o String

- Secuencia finita
 - $\Sigma=\{a, b\}$
 - aab
 - Pares ordenados y Producto Cartesiano
 - ((a, a), b)
 - N-Tupla
 - (a, a, b)
 - Ejercicio 3: Dado el $\Sigma=\{\text{if}, \text{for}\}$, construya una cadena que tenga cuatro caracteres
- Algunas Operaciones
 - Concatenación
 - Longitud
 - Potenciación
 - Inversa
 - Instancias de un carácter
 - Contiene tal carácter
 - ¿Cuáles son cerradas?

Cadenas especiales

- Secuencia vacía
 - ()
 -
- Símbolos
 - Épsilon ϵ ó
 - Lambda λ
- ¿A qué alfabeto pertenece?
 - (Evitar la respuesta obvia)
- En ANSI C
 - """
- Un solo carácter
 - a (carácter)
 - En ANSI C
 - 'a'
 - a (string)
 - En ANSI C
 - "a"

Tipo de dato de una cadena

- Valores, Variables, y Tipo de Dato
 - ¿ Tipo de Dato Matemático ?
 - a
 - aa
 - aaaaaaaaa
 - ϵ
 - ¿ Tipo de Dato C?
 - "a"
 - "aa"
 - "aaaaaaaa"
 - ""
- ¿Tipo de dato de las cadenas en el Lenguaje Pascal?
- ¿Tipo de dato de las cadenas en el Lenguaje C?

Tipo de dato de cualquier cadena

$$\Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots \cup \Sigma^{255}$$

$$\bigcup_{i=0}^n \Sigma^i$$

$$\Sigma^* = \bigcup_{i=0}^{\infty} \Sigma^i$$

Lenguajes Formales

Lenguajes Naturales & Formales

- Lenguajes Naturales
 - Para comunicarse entre personas
 - Evolución continua
 - Reglas gramaticales y sintaxis surgen después
 - El pragmatismo se puede obtener aunque la sintaxis o la semántica no sea la correcta
 - Preciso
- Lenguajes Formales
 - Para comunicarse entre sistemas o para formalizar conceptos
 - Evolución discreta
 - Reglas gramaticales y sintaxis surgen primero
 - El pragmatismo se obtiene solo si la sintaxis y la semántica son correctas
 - Ambiguo

Lenguajes Formales

- Conjunto de cadenas
- Sus elementos se llaman palabras
- Definición
 - Por extensión
 - Por comprensión

Ejemplo 27

Sea el siguiente Lenguaje Formal descripto por EXTENSIÓN: $L = \{101, 1001, 10001, 100001\}$. Utilizando el operador "supraíndice", este lenguaje puede ser descripto por COMPRENSIÓN, en forma más compacta, así: $L = \{10^n1 / 1 \leq n \leq 4\}$. Entonces, la cadena 1001 ("uno-cero-cero-uno") es una palabra del lenguaje L , mientras que 1100 ("uno-uno-cero-cero") es una cadena construida con caracteres del mismo alfabeto pero no es una palabra de L .

- Diagramas sintácticos
- Expresiones regulares
- Máquinas de estado como Autómatas Finitos
- Ejemplo [Mv1c1] 27.

Ejemplo 31 [Mv1c1]

- $L = \{(abc)^n / 0 \leq n \leq 3\}$
 $L = \{\epsilon, abc, abcabc, abcabca\}.$
- ¿Alfabeto?
- ¿ $|abc| = 3$?
- ¿Concatenación o potencia cerrada?
 - La palabra vacía es un miembro L
 - La concatenación de las palabras abc y abcabc produce otra palabra de este lenguaje: abcabca.
 - En cambio, la concatenación de la palabra abcabc consigo misma produce la cadena abcabcaabc, que no es una palabra de este lenguaje.
 - La potencia $(abc)^2$ es una palabra del lenguaje
- Definir L por medio de LN
 - La cadena abc hasta tres veces o ninguna vez

Ejemplo 32 [Mv1c1]

- $L = \{a^{2n+1} / 0 \leq n \leq 200\}$
¿Concatenación cerrada?
- Ejercicio 15: Definir L por medio de LN

Cardinalidad de un Lenguaje Formal

- Ejemplos
 - $L_1 = \{a, ab, aab\}$
 - $L_2 = \{\} = \emptyset$
 - $L_3 = \{\varepsilon\}$

Sublenguajes

- Sublenguajes de un Lenguaje
 - \emptyset es sublenguaje de todo lenguaje
- ¿Cómo se expresa "Todos los Sublenguajes de un Lenguaje"?
 - Conjunto potencia: $P(L)=2^L$
- Ejercicio. Completar:
 - $L = \{a, b\}$
 - $2^L = \{$
 - $|2^L| = |2|^{|L|} =$

Lenguajes Formales – El Lenguaje Universal sobre un Alfabeto – Definición formal de Σ^*

- $\Sigma = \{a, b\}$ Alfabeto
- $\Sigma^0 = \{\epsilon\}$
Lenguaje de palabras de longitud cero
 - $\Sigma^0 \neq \emptyset$
 - $\emptyset = \{\}$
- $\Sigma = \Sigma^1 = \{a, b\}$
Lenguaje de palabras de longitud uno
- $\Sigma \times \Sigma =$
 $\Sigma^2 = \{aa, bb, ab, ba\}$
Lenguaje de palabras de longitud dos
- $\Sigma \times \Sigma \times \Sigma =$
 $\Sigma^2 \times \Sigma =$
 $\Sigma^3 = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$
Lenguaje de palabras de longitud tres
- $\Sigma^* =$
 $\Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots =$
$$\Sigma^* = \bigcup_{i=0}^{\infty} \Sigma^i$$

Lenguajes Formales – Dado un Σ

¿Cuántos lenguajes?

- [Mv1c1] Ejercicio 10:
Dado el LF = {Argentina, Holanda, Brasil}
Indique el Σ (alfabeto) mínimo
- Dado el Σ anterior, indique cuántos lenguajes de cada cardinalidad existen
 - 0
 - 1
 - 2
 - n
 - Infinita

Intervalo

20 minutos

Formalización de Automátas Finitos

Autómatas Finitos

- Repaso

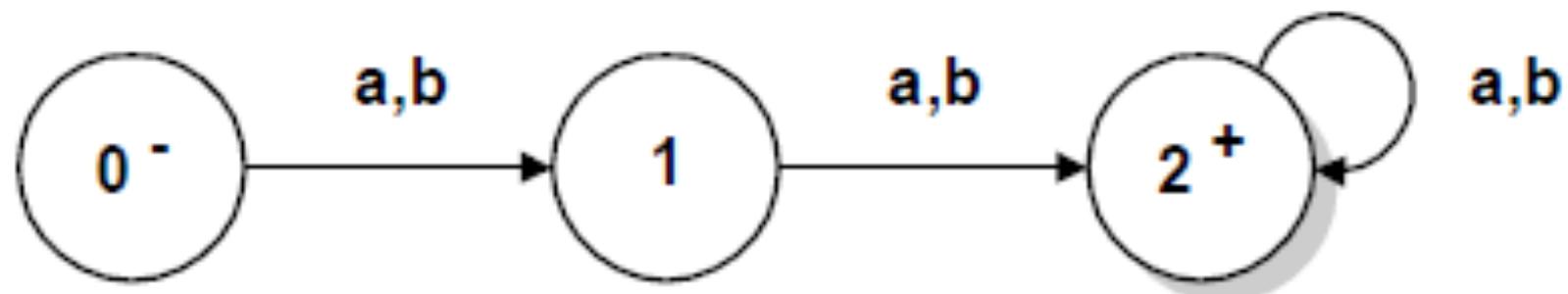
Autómatas Finitos – SSL vol 2

- $M = (Q, \Sigma, T, q_0, F)$
- $T : Q \times \Sigma \rightarrow Q / T(q, x) = z$
 - Representaciones
 - $((q, x), z)$
 - (q, x, z)
 - $q \Rightarrow x \Rightarrow z$
 - Digrafo
- Tabla de transiciones: Cuadro de doble entrada
 - $Q \times \Sigma \rightarrow Q$
 - Inicial
 - Finales

Ejemplo 14

- $M = (Q, \Sigma, T, q_0, F)$
- $Q = \{0, 1, 2, 3\}$
- $\Sigma = \{a, b\}$
- $q_0 = 0$
- $F = \{2, 3\}$
- $T = \{0 \Rightarrow a \Rightarrow 1, 1 \Rightarrow a \Rightarrow 2, 1 \Rightarrow b \Rightarrow 3, 3 \Rightarrow a \Rightarrow 3\}$
- Armar tabla
- Armar digrafo

Ejemplo 15



- ¿Cuántas transiciones?
- Particiones del conjunto Σ

Tablas completas

- Sin completar
 - a b
 - 0- 1 -
 - 1 2 3
 - 2+ - -
 - 3+ 3 -
- "Huecos"
- Función parcial versus completa
- Función sobreyectiva
- Completa
 - a b
 - 0- 1 4
 - 1 2 3
 - 2+ 4 4
 - 3+ 3 4
 - 4 4 4
- Estado de rechazo
- Autómatas equivalentes

Autómata Finito No Determinístico

- $M = (Q, \Sigma, T, q_0, F)$
- $T : Q \times (\Sigma \cup \{\epsilon\}) \rightarrow 2^Q / T(q, x) = Z$
- SSL vol3, Ejemplo 5, página 11
 - $a(a+b)^*b$
- Ejemplo 8
 - $ab+a^*ba$

Cardinalidad & Resumen

Cardinalidad de los Conjuntos

- Σ
 - conjunto de símbolos, ó alfabeto
 - Finito, sí se puede numerar
- Σ
 - conjunto de cadenas de longitud uno, ó lenguaje de palabras de longitud uno
 - Finito, sí se puede numerar
- Σ^n
 - conjunto de cadenas de cierta longitud, ó lenguaje de palabras cierta longitud
 - Finito, sí se puede numerar
- $\sum_{i \leq n}$
 - conjunto de cadenas de hasta cierta longitud, ó lenguaje de palabras de hasta cierta longitud
 - Finito, sí se puede numerar
- Σ^*
 - conjunto de cadenas de cualquier longitud, ó lenguaje de palabras de cualquier longitud
 - Infinito
 - ¿Se puede numerar?
 - ¿Está "Hamlet" y "El Quijote" en este conjunto? ¿Qué más?
- $2^{\Sigma^*} = P(\Sigma^*)$
 - conjunto con todos los lenguajes que se puede formar con Σ
 - "Más" infinito todavía
 - ¿Se puede numerar?
 - ¿Está "Jeringoso" en este conjunto? ¿Está C? ¿Y Smalltalk?.

Carácter, Alfabeto, Cadena & Lenguaje – Elementos y Conjuntos

- Carácter
 - $c_x = a$
 - Conjunto finito de Caracteres → Alfabeto (Σ)
 - $\Sigma = \{a, b, c\}$
 - Tipo de dato de un carácter
 - $c_x \in \Sigma$
 - Secuencia finita de Caracteres → String
 - $s = bcaacccc$
 - Tipo de dato de un String
 - $s \in \Sigma^*$
 - Conjunto de Strings → Lenguaje
 - $L = \{cc, bcb, abbb, ccab, abcaba, cb, aa\}$
 - String que pertenece a un Lenguaje → Palabra
 - $w = ccab$
 - $w \in L$
 - Conjunto de Lenguajes → Familia de Lenguajes
 - $L \in P(\Sigma^*) = 2^{\Sigma^*}$
 - Tipo de un Lenguaje
 - $L \in T_{0,1,2,3}$
- "Una **cadena de caracteres**, o simplemente **cadena** o **string**, es una secuencia finita, posiblemente vacía, de **símbolos** o **caracteres** de un **alfabeto**, que es un **conjunto finito no vacío**"
 - "**Un lenguaje formal** es un **conjunto de cadenas**; a esas **cadenas** se las llama **palabras** de un **lenguaje**".

Términos de la clase #11

Definir cada término con la bibliografía

- Carácter, Alfabeto, y String
 - Símbolo o carácter en LF
 - Concatenar caracteres
 - Potencia de un carácter
 - Alfabeto Σ en LF
 - Símbolo o carácter en ANSI C
 - Alfabeto en ANSI C
 - Cadena o String en LF
 - Operaciones con cadenas
 - Cadenas especiales
 - Cadena vacía en LF
 - Cadena o String en ANSI C
 - Cadena vacía en ANSI C
 - Cadena de longitud uno en LF
 - Cadena de longitud uno en ANSI C
 - Tipo de una cadena en LF
 - Tipo de una cadena en ANSI C
 - Tipo de dato de cualquier cadena en LF
 - Tipo de dato de cualquier cadena en ANSI C
 - Definición de Σ^*
- Lenguajes Formales
 - Lenguaje Natural
 - Lenguaje Formal
 - Evolución
 - Reglas gramaticales
 - Reglas sintácticas
 - Alfabeto de un LF
 - Cadena que pertenece a un LF
 - Palabra
 - Definiciones y Especificaciones de LF
 - Por Lenguaje Natural
 - Por Conjuntos por Extensión
 - Por Conjuntos por Comprensión con operaciones sobre Caracteres, Strings o Lenguajes (conjuntos)
 - Lenguajes de Programación como Lenguajes Formales
 - Palabra de un LP
 - Sublenguajes de un Lenguaje
 - Conjunto potencia: $P(L)=2^L$
 - El Lenguaje Universal sobre un Alfabeto: Σ^*
- Lenguaje Universal
- Formalización de Autómatas Finitos
 - Tabla de transición completa
 - Autómata Finito No Determinístico
 - Función Transición de un Autómata Finito no Determinístico
- Cardinalidad de Conjuntos y Resumen
 - Conjunto de todas las palabras
 - Conjunto de todos los lenguajes.

Tareas para la próxima clase

1. Entregar el trabajo #3.



¿Consultas?

Clase #12 de 27

Resolución Trabajo #3 y Repaso Examen #1

Junio 5, Lunes

Repaso Clase Anterior

- Glosario
- Tareas

Agenda para esta clase

- Resolución trabajo #3
- Repaso Examen #1

Tareas para la próxima clase

1. Estudiar para el exámen.



¿Consultas?