

Pedro Páramo

Con Analizador Lexico y Sintactico en FLEX

VINCULACIÓN INTERNACIONAL COIL-UAEH-UDEA

EQUIPO:

JUAN PABLO PÉREZ

LUISA MARTÍNEZ

LINARES CARRADA JASIEL

LÓPEZ LÓPEZ FERNANDO

MARTÍNEZ ÁNGELES MARIO RODRIGO

ROCHA HERNANDEZ ROGELIO

Índice

Introducción	2
Marco teórico.....	3
Objetivos.....	5
Desarrollo	6
Resultados	8
Conclusiones	9
Referencias	10

Introducción

El presente proyecto, desarrollado en el marco de la vinculación internacional COIL-UAEH-UDEA, tiene como propósito integrar los conocimientos de lenguajes formales, autómatas, compiladores y procesamiento del lenguaje natural (PLN) para construir una herramienta de análisis léxico y sintáctico aplicada al texto literario Pedro Páramo de Juan Rulfo.

El trabajo surge del interés por fomentar la inclusión educativa a través del uso de tecnologías que faciliten la comprensión lingüística entre hablantes de diferentes variantes del español. Para ello, se implementó un sistema capaz de reconocer, clasificar y adaptar regionalismos colombianos, traduciendo el texto a un lenguaje más accesible para estudiantes iberoamericanos.

El proyecto combina el uso de Lex y Yacc para el procesamiento computacional del lenguaje con técnicas de minería de texto y análisis emocional, permitiendo no solo la detección de tokens y estructuras gramaticales, sino también la generación de nubes de palabras y la interpretación afectiva del discurso.

Además, la colaboración interdisciplinaria entre los equipos de México y Colombia permitió consolidar una experiencia de aprendizaje internacional, reforzando competencias en prog

Marco teórico

Sobre Objetivo de Desarrollo Sostenible 4...

Sobre **Autómatas y Compiladores**...[puede ser lo que ya les puse en el documento que describe el proyecto]

Un **Autómata Finito Determinista (AFD)** se define como una tupla:

$$AFD = (\Sigma, Q, \delta, q_0, F)$$

- Σ es el alfabeto de entrada
- Q es el conjunto finito y no vacío de los estados
- δ es la función de transición, $\delta: Q \times \Sigma \rightarrow Q$
- $q_0 \in Q$ es el estado inicial
- $F \subset Q$ es un conjunto de estado finales de aceptación ($F \neq \emptyset$)

Sea $A = (\Sigma, Q, \delta, q_0, F)$ un AFD y sea $w = w_1 w_2 \dots w_n$ una cadena de símbolos donde $w_i \in \Sigma$.

Entonces, **A acepta w** si existe una secuencia de estados $r_0, r_1, \dots, r_n \in Q$ con tres condiciones:

1. $r_0 = q_0$
2. $\delta(r_i, w_i + 1) = r_{i+1}$ para $i = 0, 1, \dots, n-1$
3. $r_n \in F$

La condición 1 establece que el AFD comienza en el estado inicial.

La condición 2 establece que el AFD cambia desde un estado hacia otro estado de acuerdo con la función de transición.

La condición 3 establece que el AFD acepta la cadena de entrada si termina en un estado de aceptación. Entonces, **A reconoce el lenguaje L** si $L = \{w | A \text{ acepta } w\}$.

Una **tabla de transición de estados** es una tabla que muestra a qué estado se moverá un autómata finito, basándose en el estado actual y las entradas. Una tabla de estados es esencialmente una tabla de verdad en la cual algunas de las entradas son el estado actual, y las salidas incluyen el siguiente estado, junto con otras salidas.

Generalmente se asocia con cada autómata un **diagrama de transición de estados**. Cada nodo del diagrama corresponde a un estado. El estado inicial se indica mediante una flecha que no tiene nodo origen. Los estados finales se representan con un círculo doble. Si existe una transición del estado e_i al estado e_j para un símbolo de entrada a , existe entonces un arco rotulado con a desde el nodo e_i al nodo e_j ; es decir que $\delta(e_i, a) = e_j$, se representa en el diagrama como se muestra en la Figura 1.

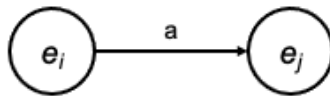


Figura 1. Diagrama de transición.

Sobre la extracción de información de textos, minería de textos...(procesamiento del lenguaje natural (PLN))

Preprocesamiento del texto

Tokenización

Eliminación de palabras vacías (stopwords)

Normalización

Extracción de palabras clave

Nubes de palabras (Word Clouds)

Bibliotecas de PLN en Python: NLTK, spaCy, scikit-learn.

Análisis de sentimientos

Clasificación de texto

Sobre educación diferenciada y la importancia del lenguaje empleado para la correcta comprensión de textos.

Objetivos

Promover la inclusión educativa mediante el diseño de experiencias de aprendizaje diferenciadas que consideren las jergas lingüísticas de los estudiantes, apoyándose en herramientas tecnológicas de reconocimiento y análisis léxico computacional.

1. Construir el documento fuente siguiendo las especificaciones para la producción del texto.
2. Implementar un analizador léxico y sintáctico utilizando Lex y Yacc, capaz de identificar, clasificar y validar los regionalismos colombianos y las expresiones emocionales presentes en un documento fuente.
3. Desarrollar un sistema de procesamiento lingüístico que traduzca el documento fuente a un lenguaje comprensible para estudiantes iberoamericanos, reconociendo automáticamente las emociones predominantes y generando una nube de palabras representativa.
4. Construir un prototipo funcional que integre las herramientas tecnológicas desarrolladas, mostrando la ejecución, los resultados y la documentación del código, en un entorno visual aplicable a contextos educativos.
5. Esbozar la implementación de los resultados para favorecer estrategias de atención diferenciada en ambientes educativos.

Desarrollo

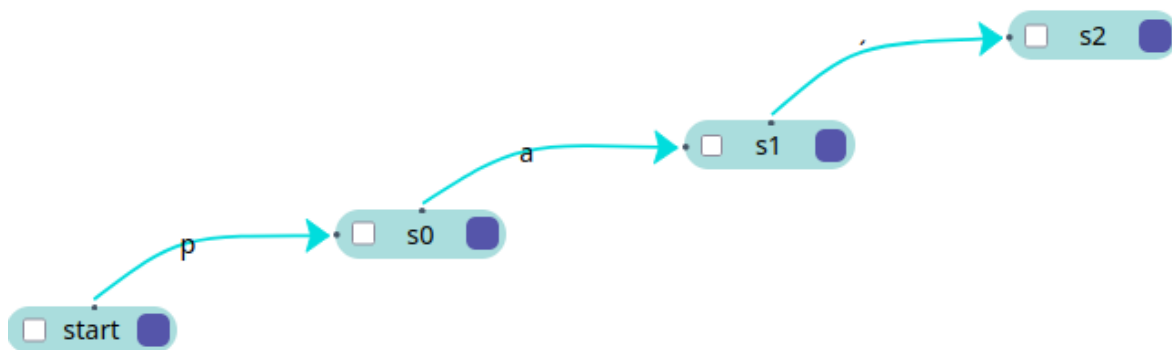
1. **Análisis Léxico: Autómatas empleados** en el para reconocer los tokens requeridos para las resoluciones.

$$AFD = (\Sigma, Q, \delta, q_0, F)$$

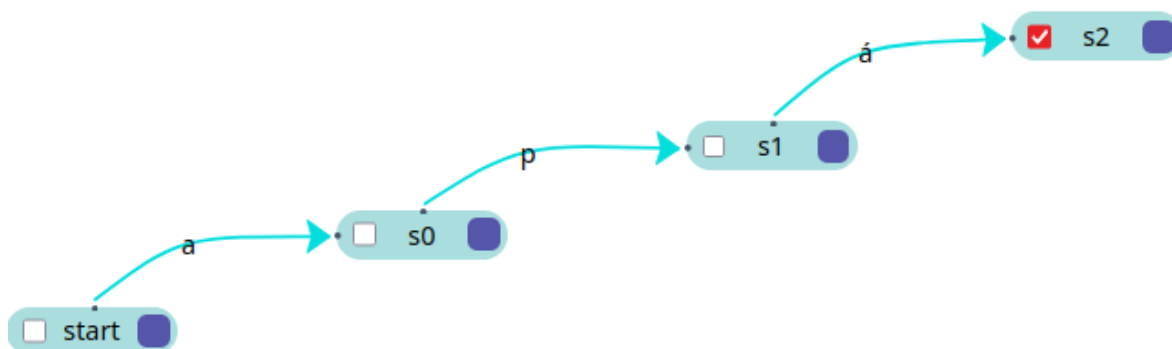
- Σ es el alfabeto de entrada
- Q es el conjunto finito y no vacío de los estados
- δ es la función de transición, $\delta: Q \times \Sigma \rightarrow Q$
- $q_0 \in Q$ es el estado inicial
- $F \subset Q$ es un conjunto de estado finales de aceptación ($F \neq \emptyset$)

Sea $A = (\Sigma, Q, \delta, q_0, F)$ un AFD y sea $w = w_1 w_2 \dots w_n$ una cadena de símbolos donde $w_i \in \Sigma$.

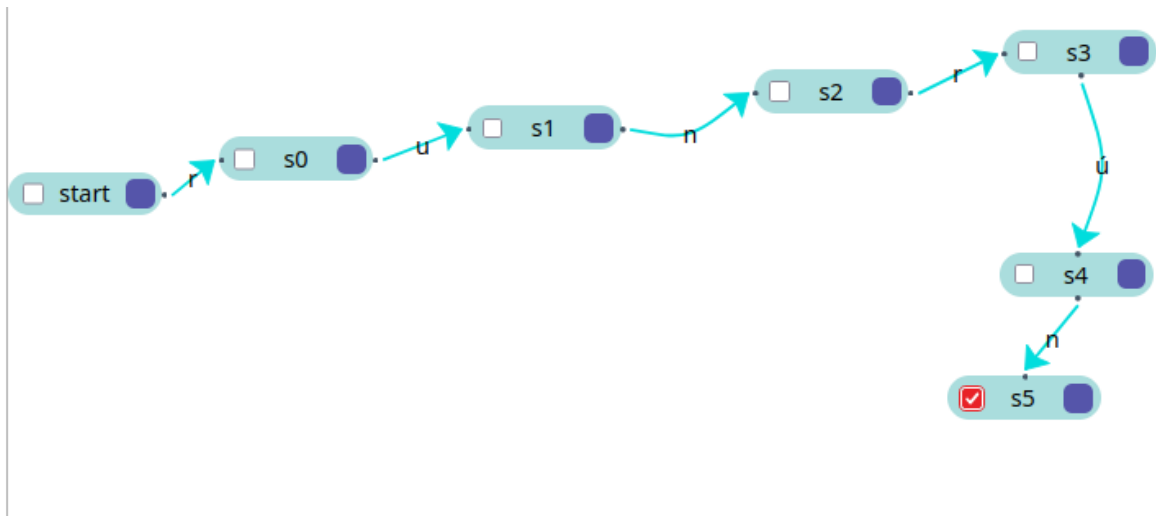
pa'



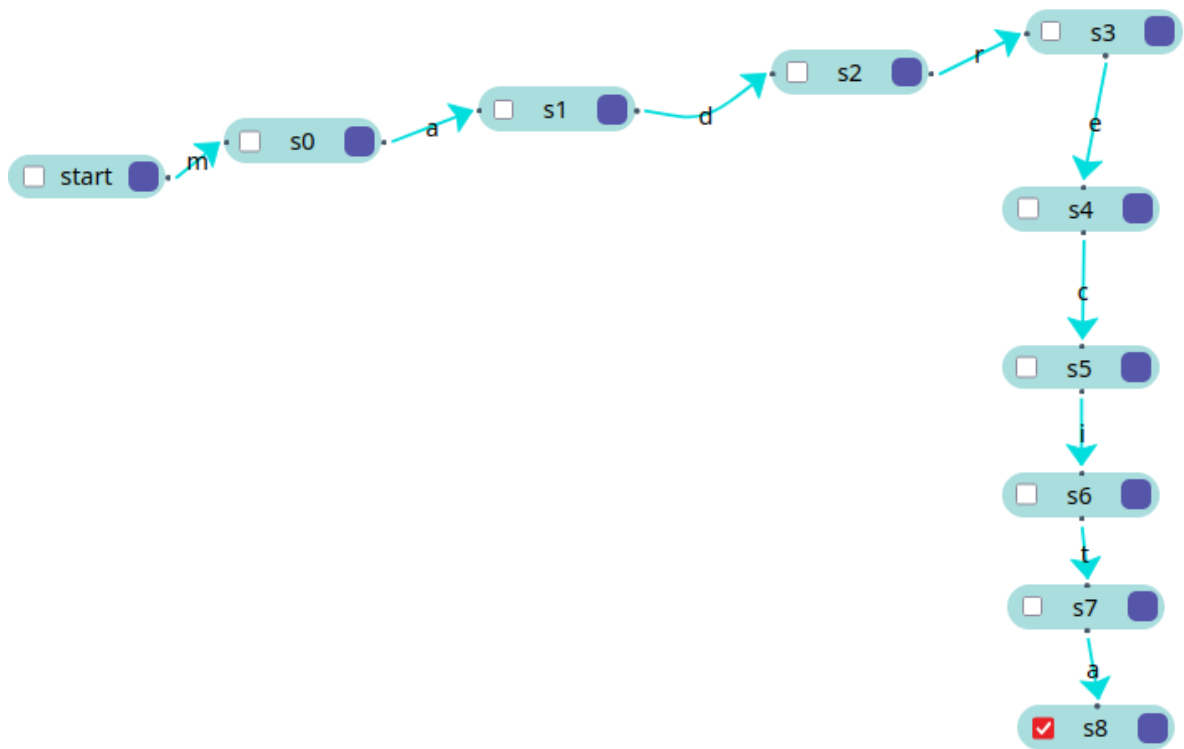
Apá



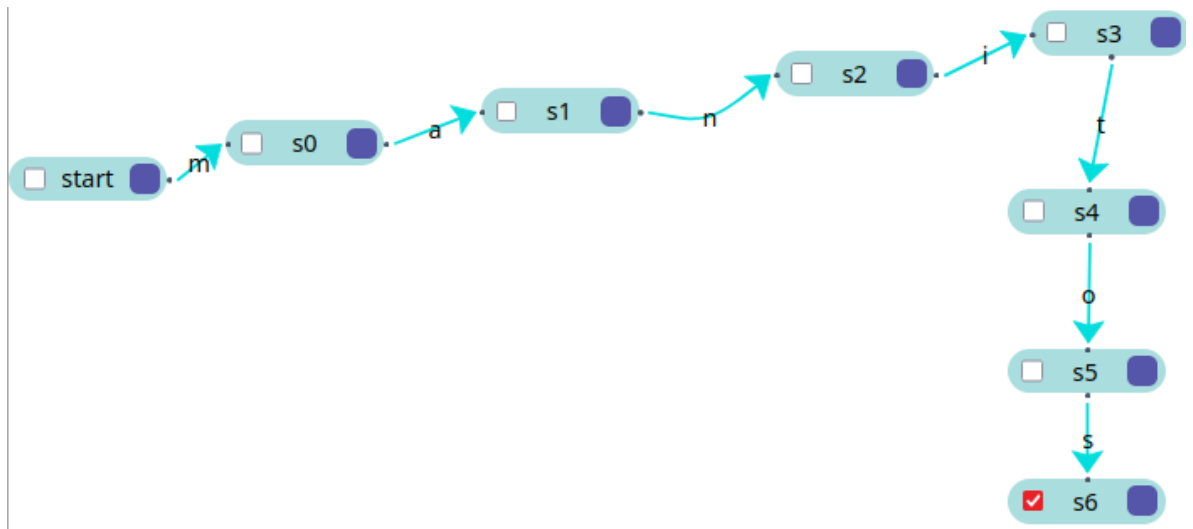
runrún



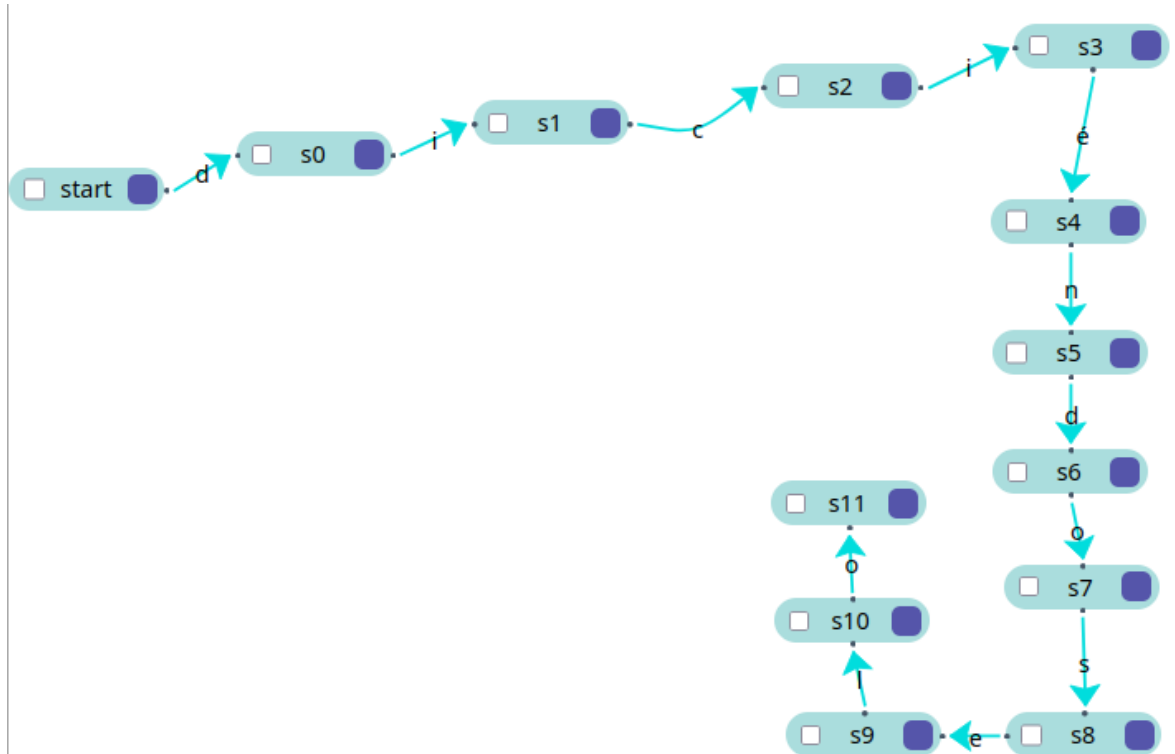
madrecita



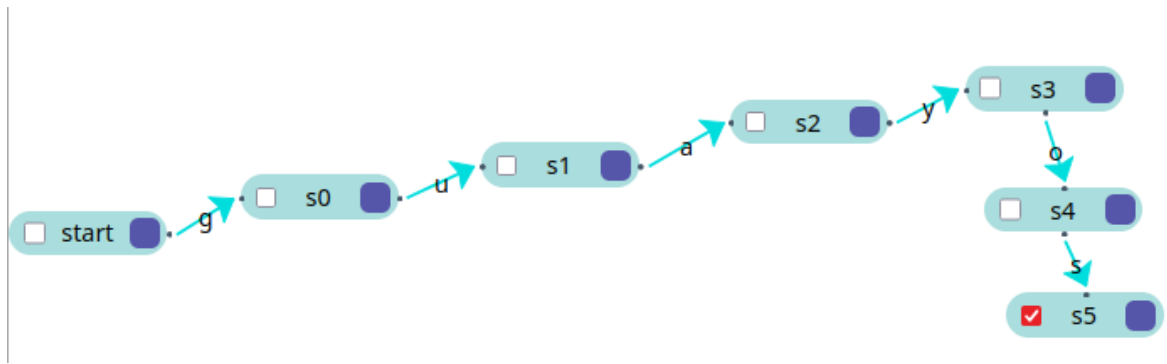
manitos



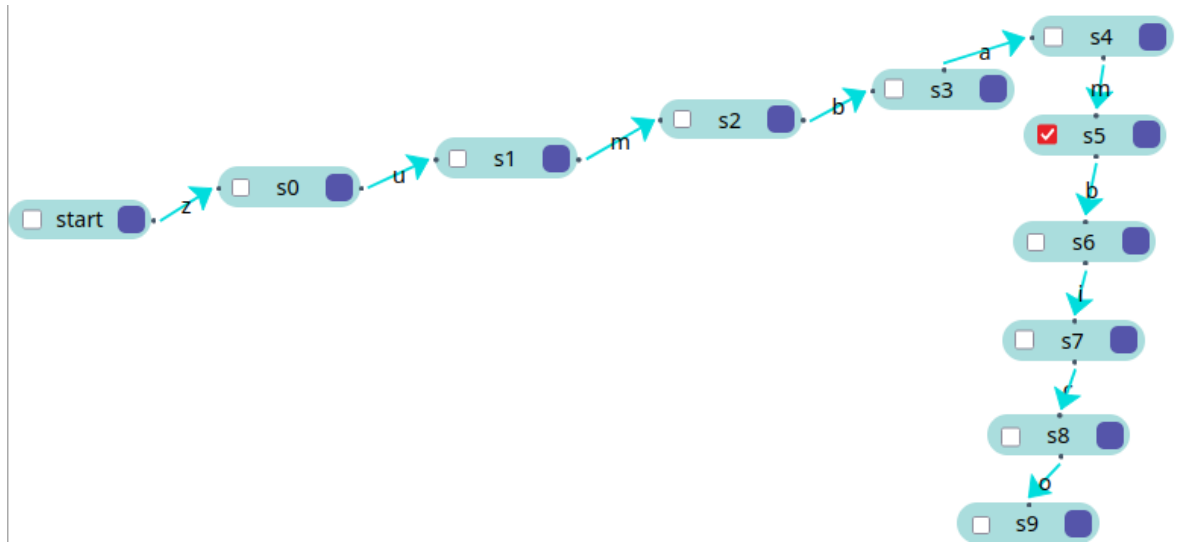
dicíendoselo



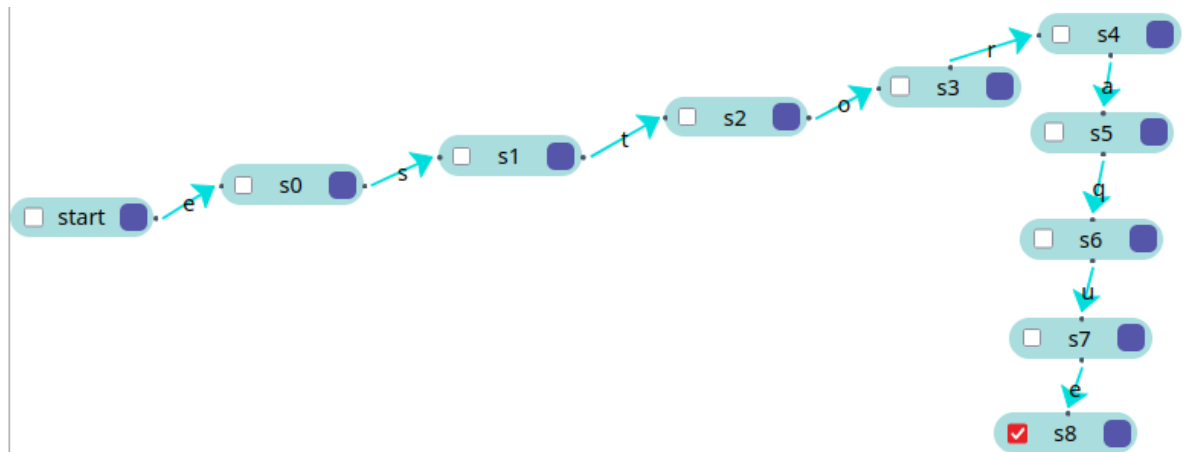
guayos



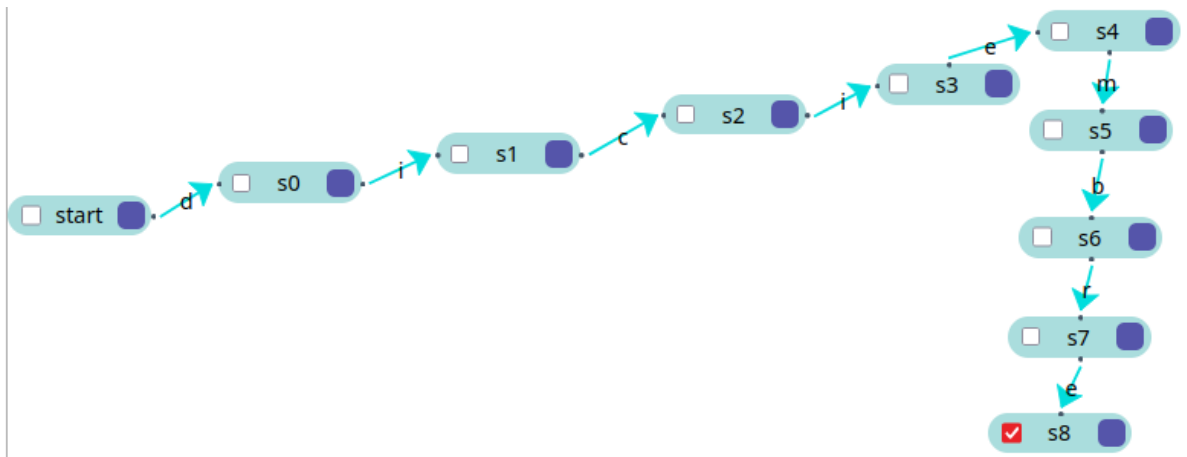
zumbambico



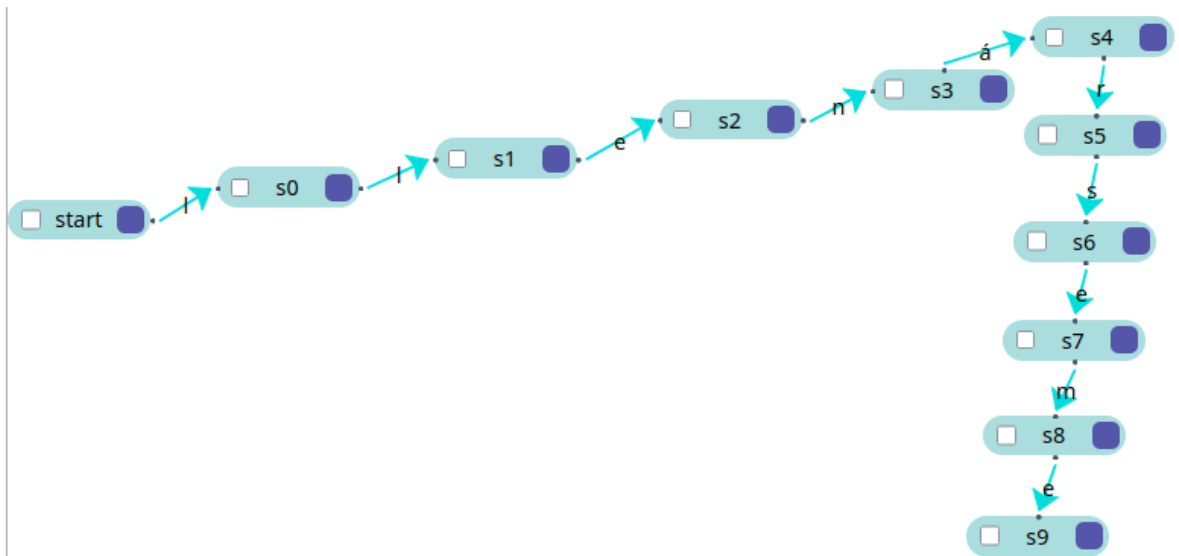
estoraque



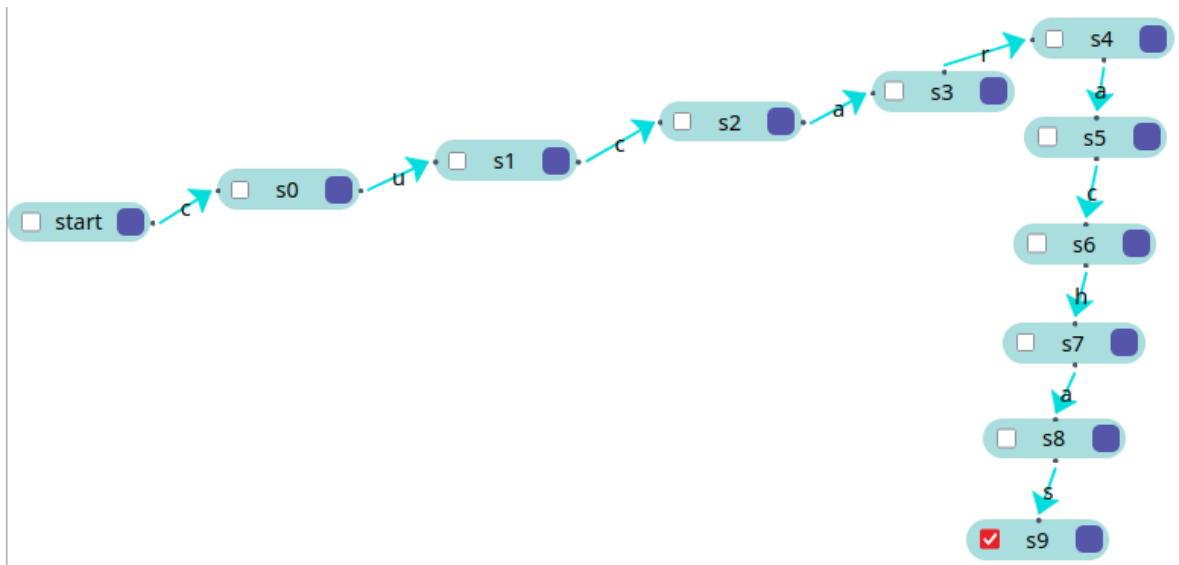
diciembre



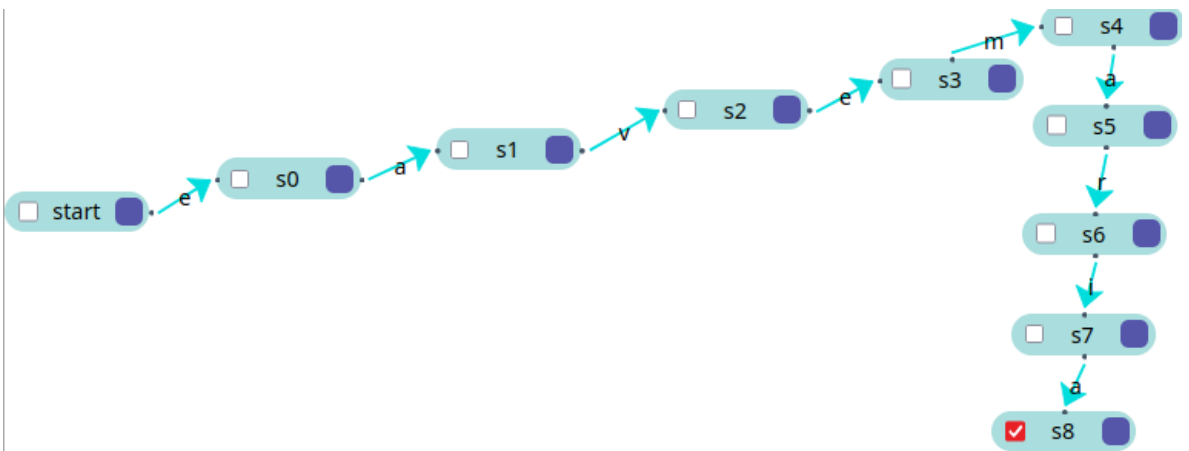
llenarse



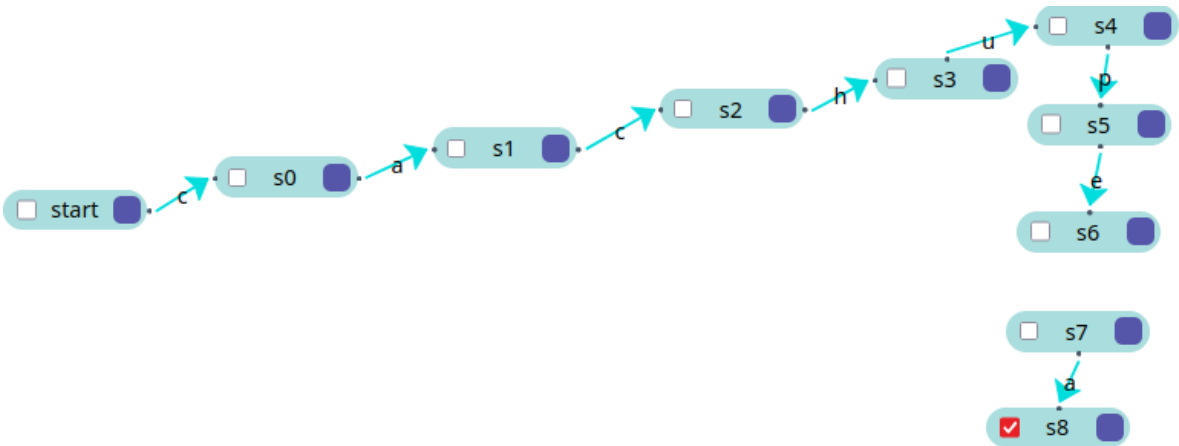
cucarachas



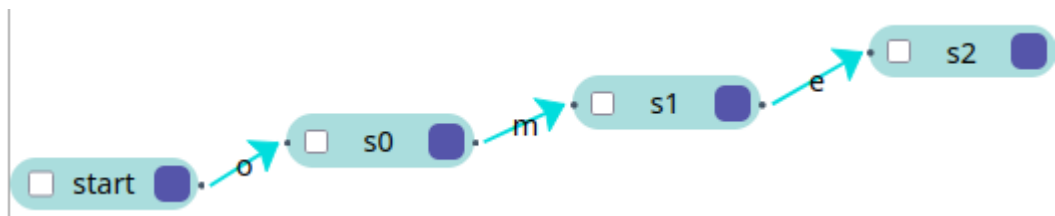
eavemaria



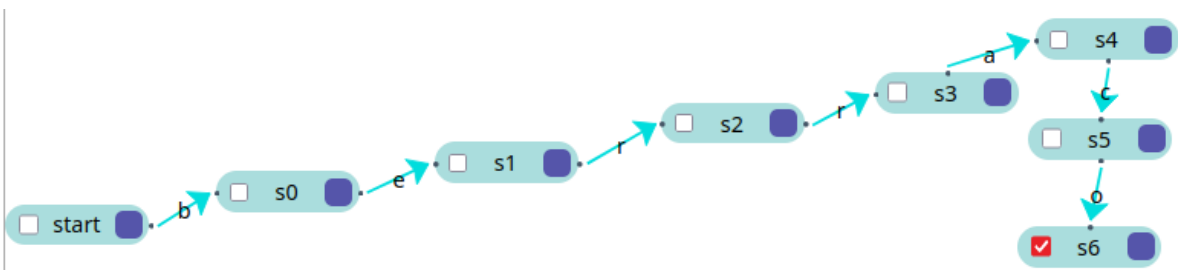
cachupe



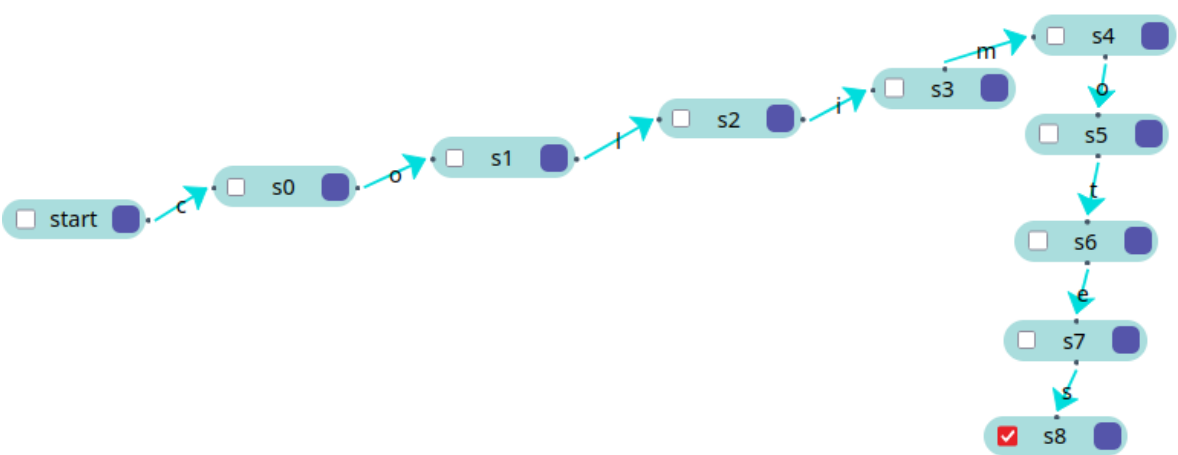
ome



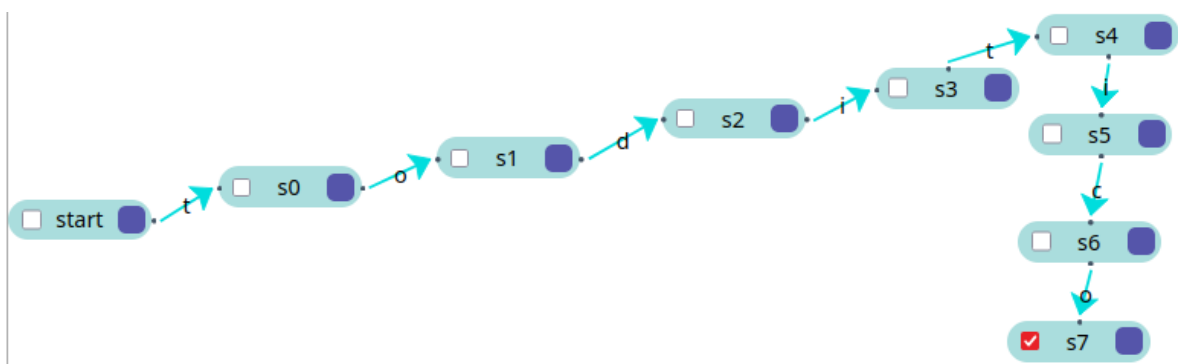
berraco



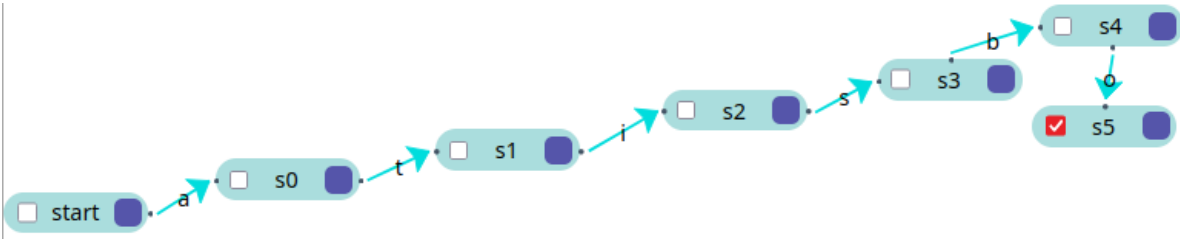
Colimotes



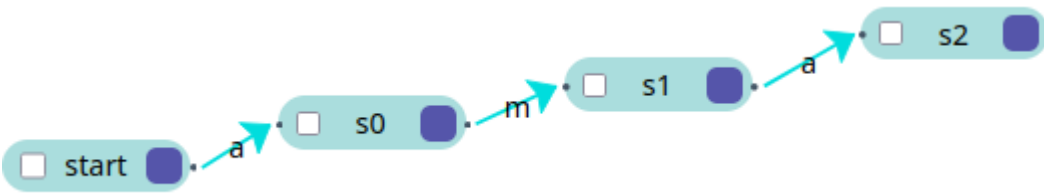
toditico



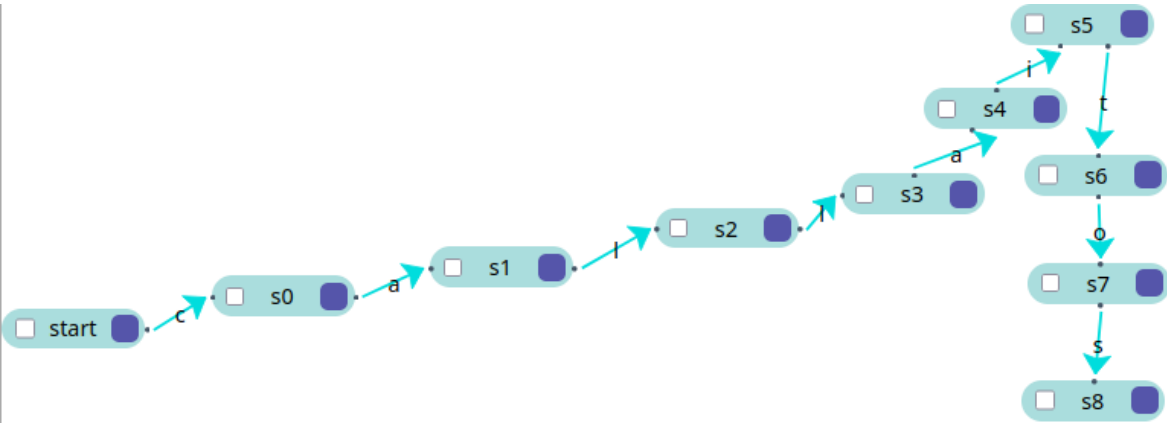
atisbo



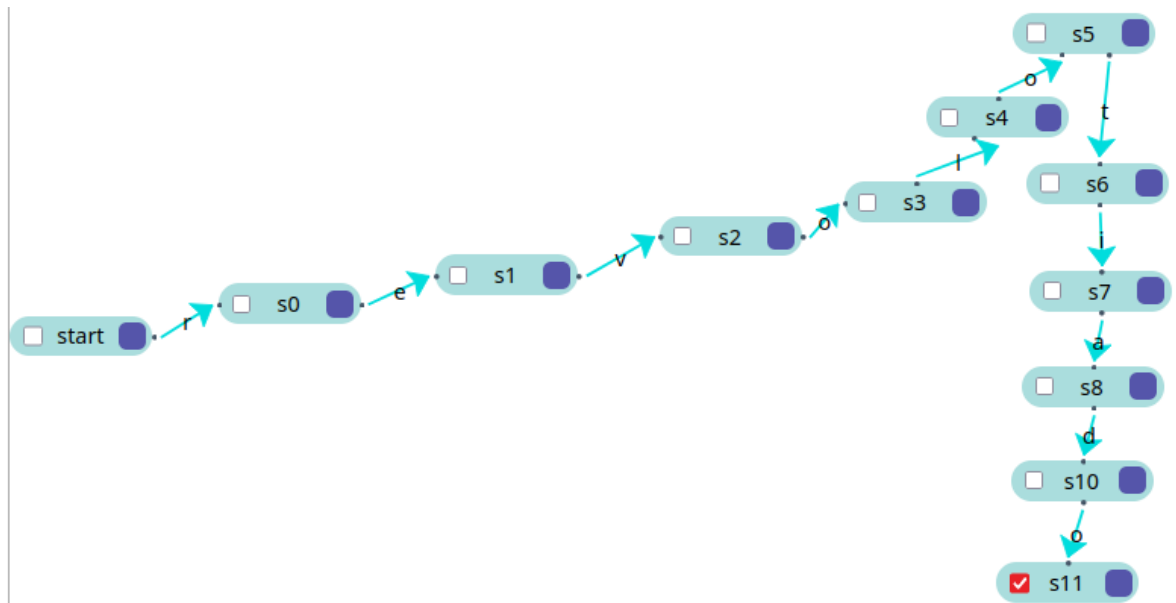
ama



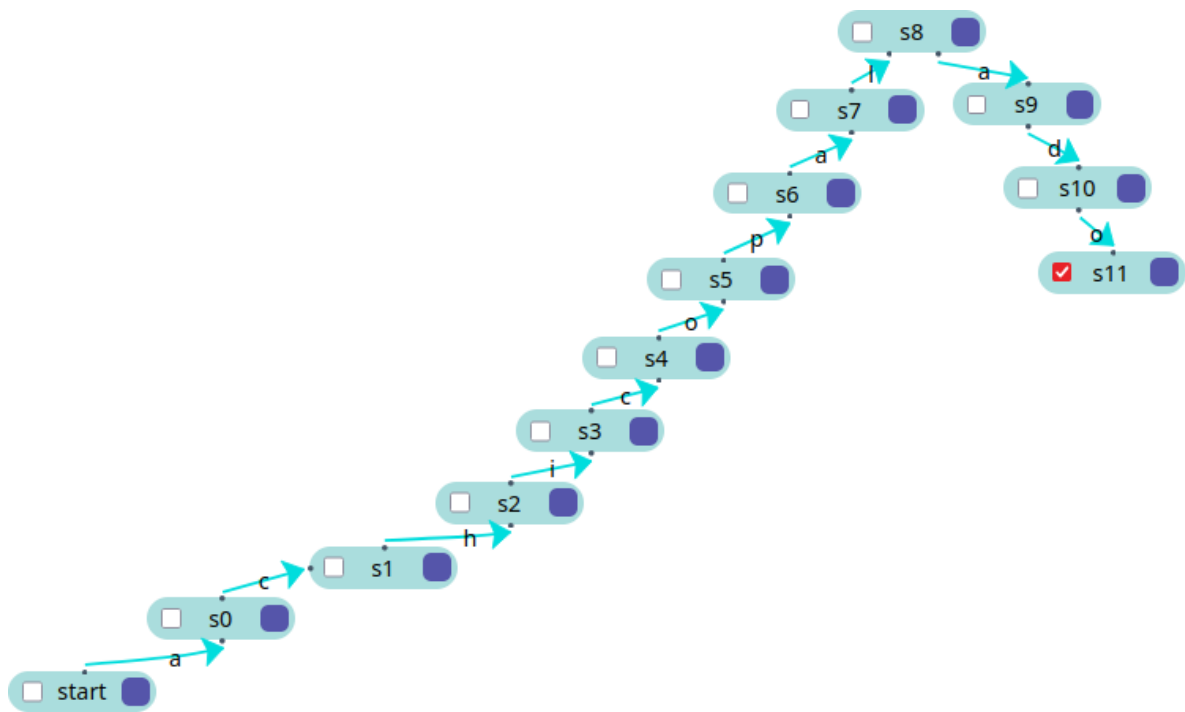
callaitos



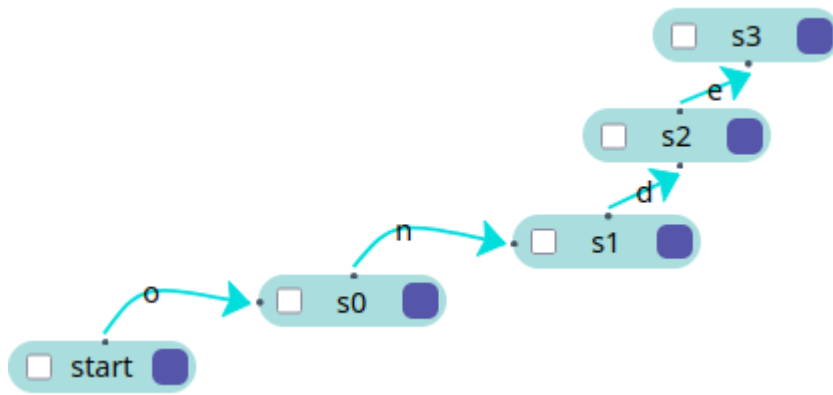
revolotiado



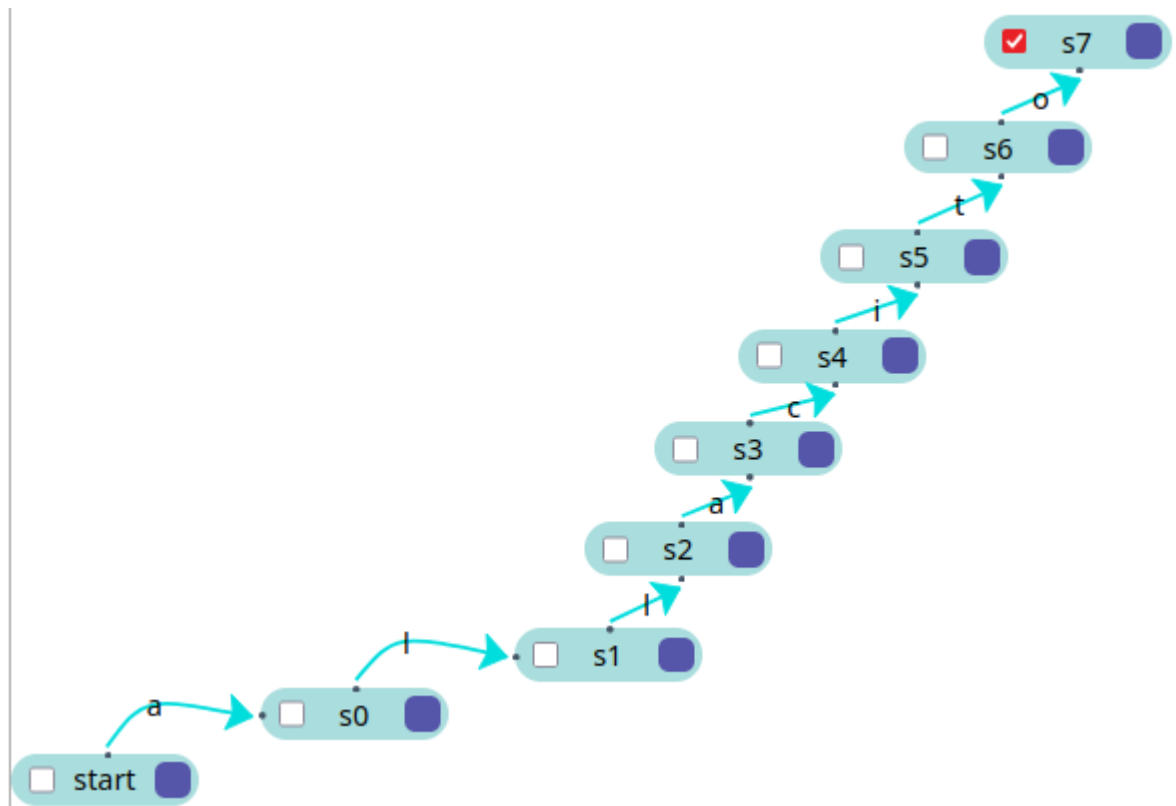
achicopalado



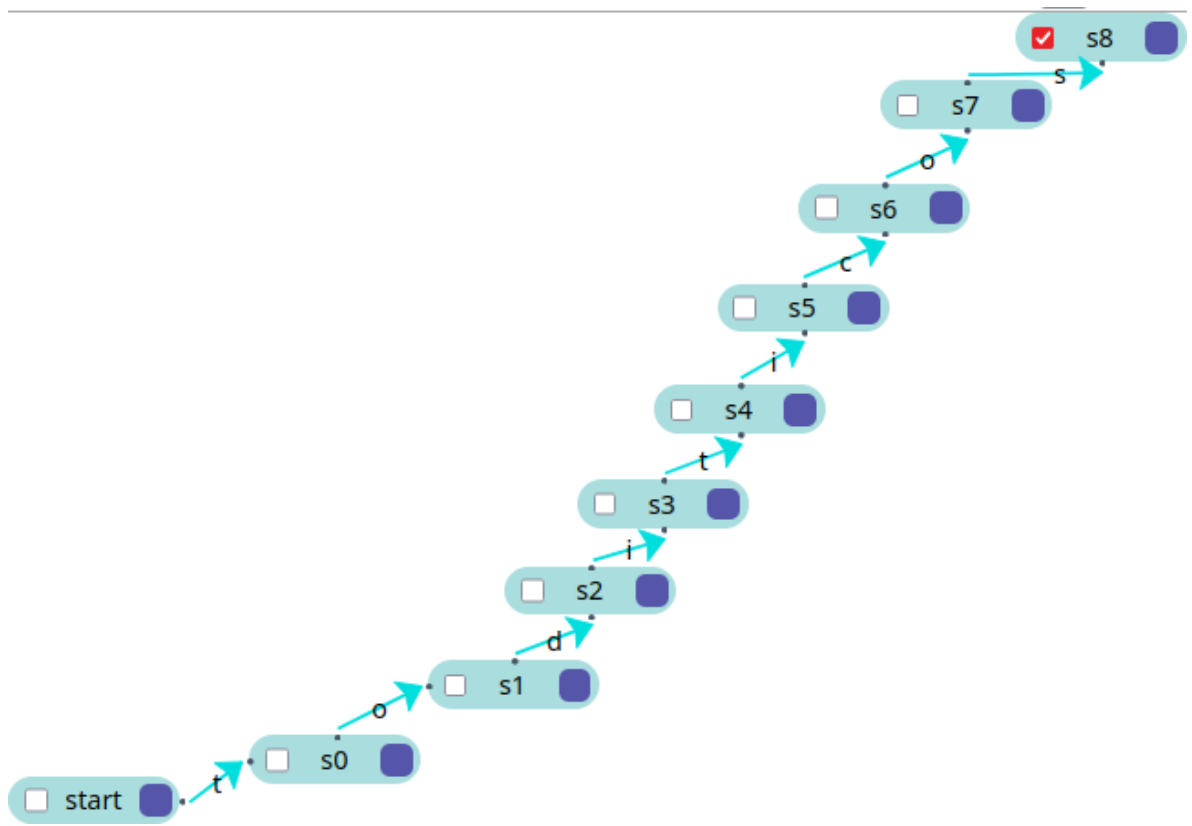
onde



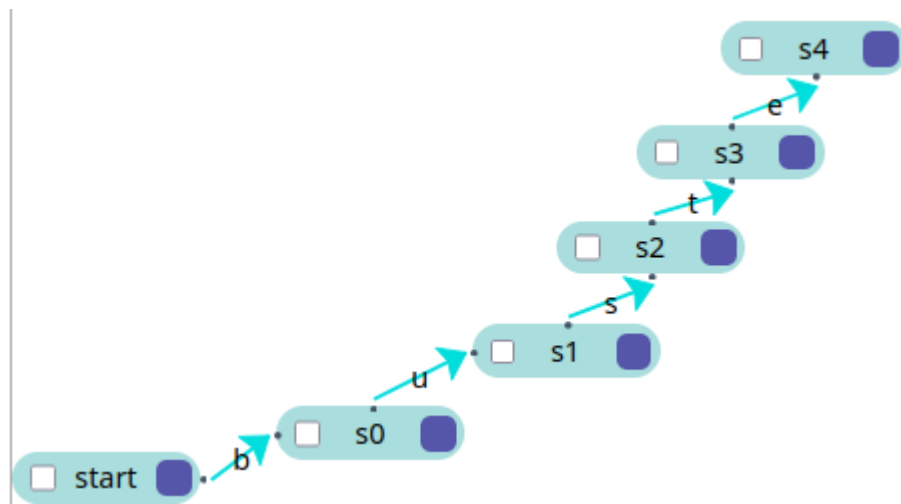
allacito



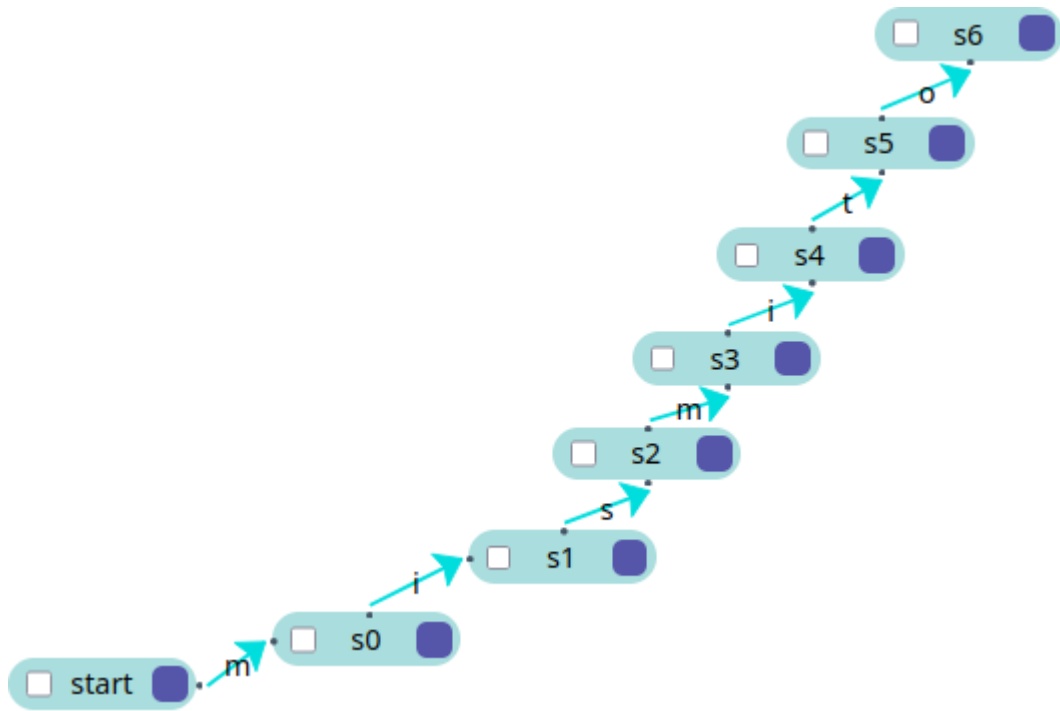
toditicos



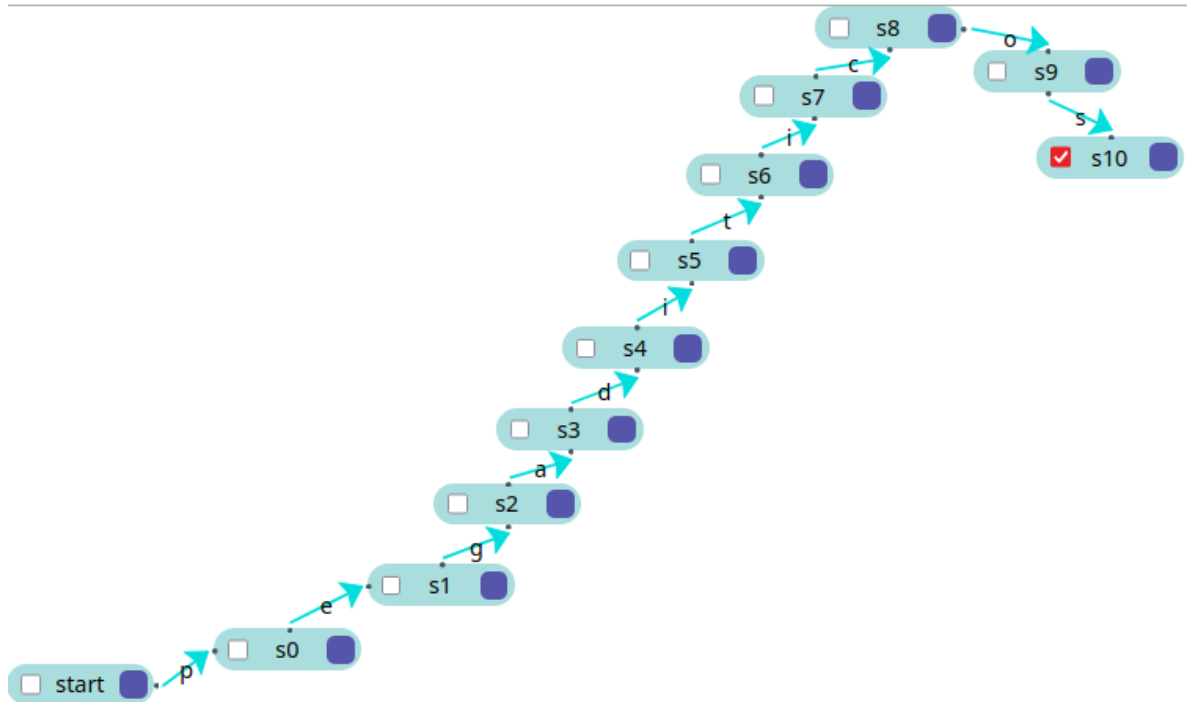
buste



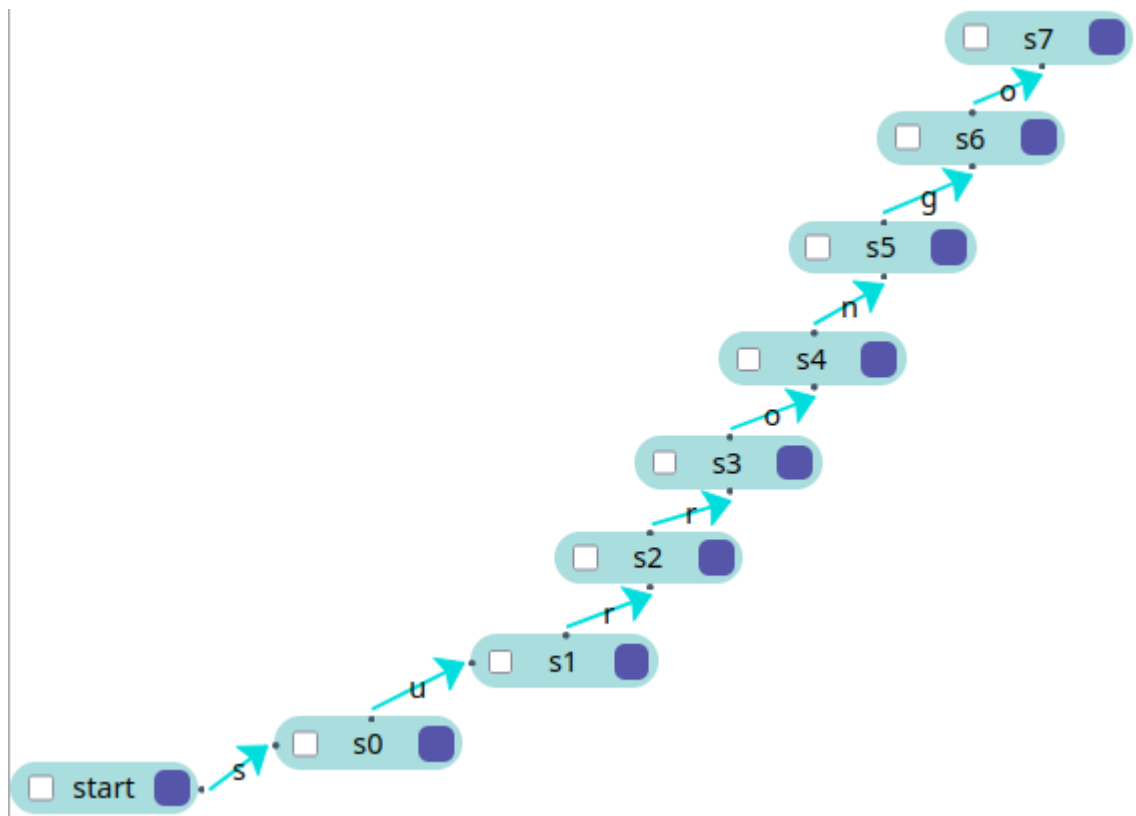
mismo



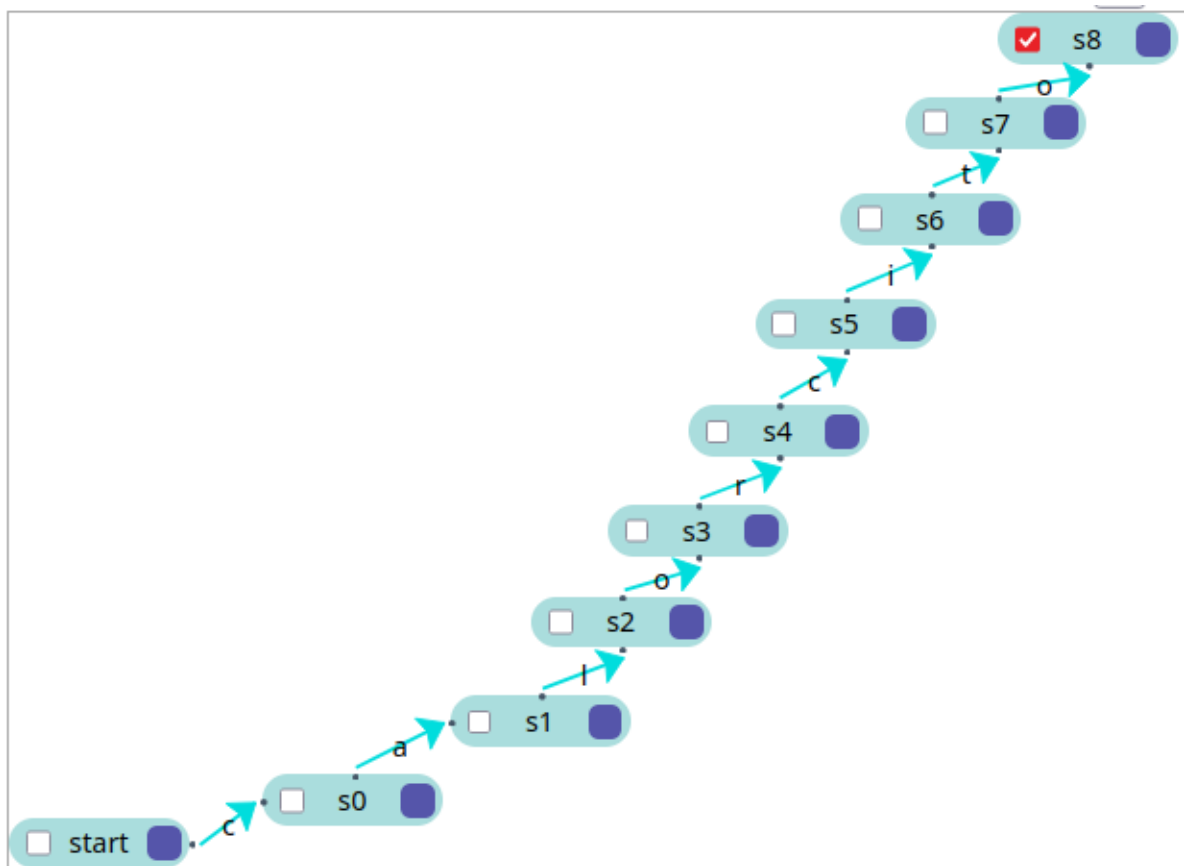
pegaditicos



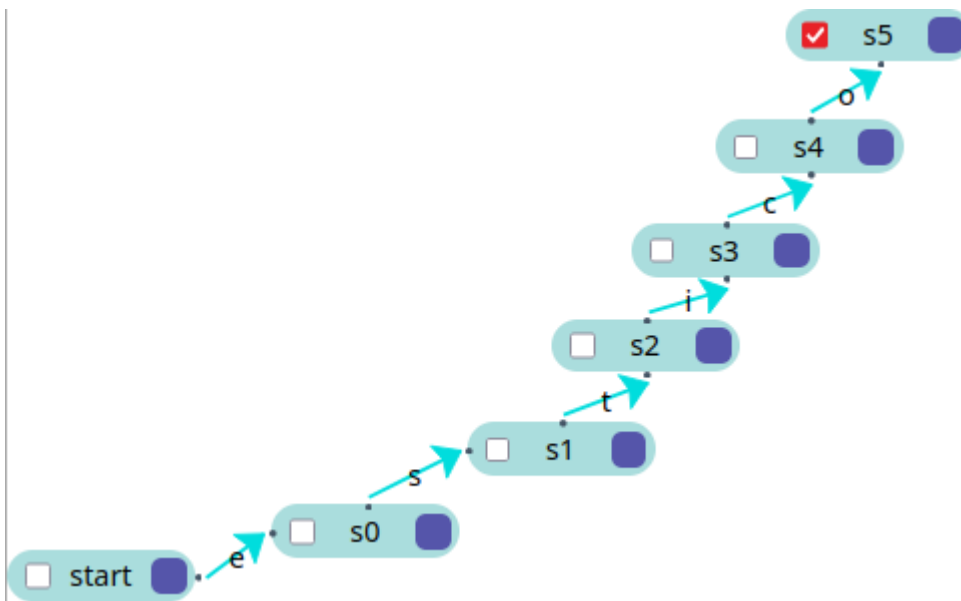
surrongo



calorcito



estico



2. Análisis Sintáctico: Gramáticas regulares utilizadas

$$G = \{\Sigma_T, \Sigma_N, S, P\}$$

donde:

- Σ_T es un alfabeto de símbolos terminales
- Σ_N es un alfabeto de símbolos no terminales
- S es el símbolo inicial de la gramática
- P es un conjunto de producciones gramaticales

Además, se cumple:

$$S \in \Sigma_N$$

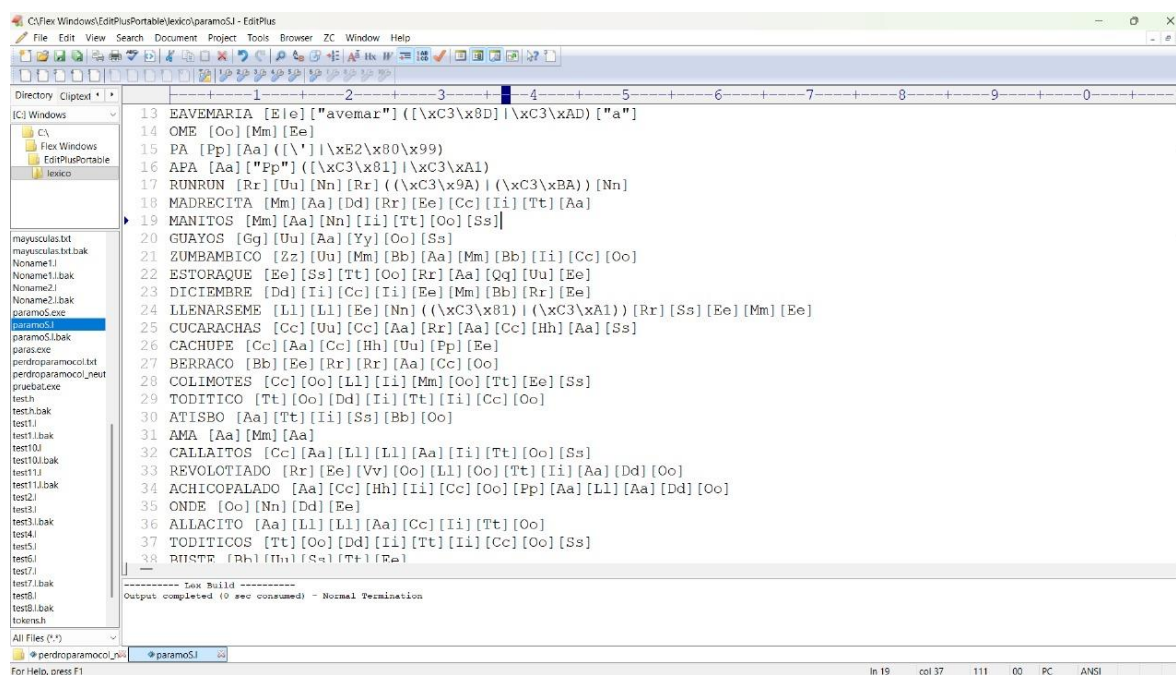
$$\Sigma_T \cap \Sigma_N = \emptyset$$

$$\Sigma = \Sigma_T \cup \Sigma_N$$

La gramática formal G permite generar un lenguaje $L = \{x \in \Sigma_T^* | S \xrightarrow{*} x\}$

Las **expresiones regulares** permiten describir con exactitud y sencillez cualquier lenguaje regular.

Expresiones regulares en flex



C:\Flex Windows\EditPlusPortable\lexico\paramoSI - EditPlus

File Edit View Search Document Project Tools Browser ZC Window Help

Directory | ClipText | 1 2 3 4 5 6 7 8 9 0

13 EAVEMARIA [E][o] [{"avemar"}] ({\xC3\x8D}{\xC3\xAD} [{"a"}]
14 OME [Oo][Mm][Ee]
15 PA [Pp][Aa] ({\x2E\x80\x99})
16 APA [Aa] [{"Pp"}] ({\xC3\x81}{\xC3\xA1})
17 RUNRUN [Rr][Uu][Nn][Rr] ({\xC3\x9A}{\xC3\xBA}) [Nn]
18 MADRECITA [Mm][Aa][Dd][Rr][Ee][Cc][Ii][Tt][Aa]
19 MANITOS [Mm][Aa][Nn][Ii][Tt][Oo][Ss]
20 GUAYOS [Gg][Uu][Aa][Yy][Oo][Ss]
21 ZUMBAMBICO [Zz][Uu][Mm][Bb][Aa][Mm][Bb][Ii][Cc][Oo]
22 ESTORAQUE [Ee][Ss][Tt][Oo][Rr][Aa][Qq][Uu][Ee]
23 DICIEMBRE [Dd][Ii][Cc][Ii][Ee][Mm][Bb][Rr][Ee]
24 LLENARSEME [Ll][Ll][Ee][Nn] ({\xC3\x81}{\xC3\xA1}) [Rr][Ss][Ee][Mm][Ee]
25 CUCARACHAS [Cc][Uu][Cc][Aa][Rr][Aa][Cc][Hh][Aa][Ss]
26 CACHUPE [Cc][Aa][Cc][Hh][Uu][Pp][Ee]
27 BERRACO [Bb][Ee][Rr][Rr][Aa][Cc][Oo]
28 COLIMOTES [Cc][Oo][Ll][Ii][Mm][Oo][Tt][Ee][Ss]
29 TODITICO [Tt][Oo][Dd][Ii][Tt][Ii][Cc][Oo]
30 ATISBO [Aa][Tt][Ii][Ss][Bb][Oo]
31 AMA [Aa][Mm][Aa]
32 CALLAITOS [Cc][Aa][Ll][Ll][Aa][Ii][Tt][Oo][Ss]
33 REVOLOTIADO [Rr][Ee][Vv][Oo][Ll][Oo][Tt][Ii][Aa][Dd][Oo]
34 ACHICOPALADO [Aa][Cc][Hh][Ii][Cc][Oo][Pp][Aa][Ll][Aa][Dd][Oo]
35 ONDE [Oo][Nn][Dd][Ee]
36 ALLACITO [Aa][Ll][Ll][Aa][Cc][Ii][Tt][Oo]
37 TODITICOS [Tt][Oo][Dd][Ii][Tt][Ii][Cc][Oo][Ss]
38 BUSTE [Bb][U][S][T][E]

----- Text Build -----
Output completed (0 sec consumed) - Normal Termination

All Files (*)

paramoSI

For Help, press F1

In 19 col 37 111 00 PC ANSI

C:\Flex Windows\EditPlusPortable\lexico\paramoSI - EditPlus

File Edit View Search Document Project Tools Browser ZC Window Help

Directory | ClipText | 1 2 3 4 5 6 7 8 9 0

55 {MADRECITA} { fputs("madre", yyout); }
56 {MANITOS} { fputs("manos", yyout); }
57 {GUAYOS} { fputs("zapatos", yyout); }
58 {ZUMBAMBICO} { fputs("tonto", yyout); }
59 {ESTORAQUE} { fputs("tipo", yyout); }
60 {DICIEMBRE} { fputs("diciembre", yyout); }
61 {LLENARSEME} { fputs("llenarseme", yyout); }
62 {CUCARACHAS} { fputs("cucarachas", yyout); }
63 {CACHUPE} { fputs("olor fuerte", yyout); }
64 {BERRACO} { fputs("muy valiente", yyout); }
65 {COLIMOTES} { fputs("cerros", yyout); }
66 {TODITICO} { fputs("todo", yyout); }
67 {ATISBO} { fputs("mirada", yyout); }
68
69 {CALLAITOS} { fputs("callados", yyout); }
70 {REVOLOTIADO} { fputs("alborotado", yyout); }
71 {ACHICOPALADO} { fputs("desanimado", yyout); }
72 {ONDE} { fputs("donde", yyout); }
73 {ALLACITO} { fputs("allá", yyout); }
74 {TODITICOS} { fputs("todos", yyout); }
75 {BUSTE} { fputs("usted", yyout); }
76 {MISMITO} { fputs("mismo", yyout); }
77 {PEGADITICOS} { fputs("pegados", yyout); }
78 {SURRONGO} { fputs("hijo", yyout); }
79 {CALORCITO} { fputs("calor", yyout); }
80 {BUSTE} { fputs("buste", yyout); }

----- Text Build -----
Output completed (0 sec consumed) - Normal Termination

All Files (*)

paramoSI

For Help, press F1

In 19 col 37 111 00 PC ANSI

```

76 {MISMITO}      { fputs("mismo", yyout); }
77 {PEGADITICOS} { fputs("pegados", yyout); }
78 {SURRONGO}     { fputs("hijo", yyout); }
79 {CALORCITO}    { fputs("calor", yyout); }
80 {ESTICO}       { fputs("este", yyout); }
81
82 .\n           { fputs(yytext[0], yyout); }
83
84 %%
85
86 int main(void) {
87     FILE *file;
88     char name[] = "perdroparamocol.txt";
89     char outname[] = "perdroparamocol_neutro.txt";
90     file = fopen(name, "r");
91     if (!file) {
92         printf("Error al abrir el archivo %s.\n", name);
93         return 1;
94     }
95
96     yyout = fopen(outname, "w");
97     if (!yyout) {
98         printf("Error al crear el archivo %s.\n", outname);
99         fclose(file);
100        return 1;
101    }
102 }

```

Output completed (0 sec consumed) - Normal Termination

```

91 if (!file) {
92     printf("Error al abrir el archivo %s.\n", name);
93     return 1;
94 }
95
96 yyout = fopen(outname, "w");
97 if (!yyout) {
98     printf("Error al crear el archivo %s.\n", outname);
99     fclose(file);
100    return 1;
101 }
102
103 yyin = file;
104 yylex();
105 fclose(file);
106 fclose(yyout);
107
108 printf("Texto con regionalismos reemplazados guardado en: %s\n", outname);
109 return 0;
110 }
111

```

Output completed (0 sec consumed) - Normal Termination

3. **Notación Extended Bakus Naur Form:** Expresar las producciones con (EBNF) antes del código de Yacc.

4. Traducción

Programa de LEX para el análisis léxico del documento fuente (tokenización)

Evidencias de ejecución desde línea de comandos

```
Microsoft Windows [Versión 10.0.26100.7171]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Users\m9060>cd C:\Flex Windows\EditPlusPortable\lexico
C:\Flex Windows\EditPlusPortable\lexico>lex contadorparamo.l
C:\Flex Windows\EditPlusPortable\lexico>cc -w lex.yy.c -o 1234
C:\Flex Windows\EditPlusPortable\lexico>1234.exe
Encontrado: PA
Encontrado: RUNRUN
Encontrado: APA
Encontrado: MADRECITA
Encontrado: OME
Encontrado: MANITOS
Encontrado: PA
Encontrado: OME
Encontrado: AMA
Encontrado: MANITOS
Encontrado: GUAYOS
Encontrado: ZUMBAMBICO
Encontrado: ESTORAQUE
Encontrado: DICIEMBRE
Encontrado: PA
Encontrado: LLENARSEME
Encontrado: CUCARACHAS
Encontrado: AMA
Encontrado: PA
Encontrado: CACHEUPE
Encontrado: AMA
Encontrado: OME
Encontrado: BERRACO
Encontrado: ACHICOPALADO
Encontrado: PA
Encontrado: COLIMOTES
Encontrado: AMA
Encontrado: TODITICO
Encontrado: ATISBO
Encontrado: MISHITO
Encontrado: AMA
Encontrado: CALLAITOS
Encontrado: REVOLOTIADO
Encontrado: ONDE
Encontrado: ALLACITO
Encontrado: AMA
Encontrado: AMA
Encontrado: TODITICOS
Encontrado: ONDE
Encontrado: ONDE
Encontrado: BUSTE
Encontrado: MISHITO
Encontrado: PEGADITICOS
Encontrado: SURRONGO
Encontrado: CALORCITO
Encontrado: ESTICO
=== CONTEOS FINALES ===
Total palabras encontradas: 46

Total palabras encontradas: 46
Detalles:
pa': 5
ap8: 1
runr:n: 1
madrecita: 1
manitos: 2
guayos: 1
zumbambico: 1
estoraque: 1
diciembre: 1
llen8rseme: 1
cucarachas: 1
cachupe: 1
eavemaria: 3
ome: 3
berraco: 1
colimotes: 1
toditico: 1
atisbo: 1
ama: 7
callaitos: 1
revolotiado: 1
achicopalado: 1
onde: 3
allacito: 1
toditicos: 1
buste: 1
mishito: 2
madrecita: 1
manitos: 2
guayos: 1
zumbambico: 1
estoraque: 1
diciembre: 1
llen8rseme: 1
cucarachas: 1
cachupe: 1
eavemaria: 3
ome: 3
berraco: 1
colimotes: 1
achicopalado: 1
onde: 3
allacito: 1
toditicos: 1
buste: 1
mishito: 2
pegaditicos: 1
surrongo: 1
calorcito: 1
estico: 1
C:\Flex Windows\EditPlusPortable\lexico>
```

5. Eliminación de palabras vacías (stopwords)

Programa de LEX para el análisis léxico del documento fuente

```
C:\Flex Windows\EditPlusPortable\lexico>quitar_preposiciones.l
C:\Flex Windows\EditPlusPortable\lexico>cc lex.yy.c
C:\Flex Windows\EditPlusPortable\lexico>
C:\Flex Windows\EditPlusPortable\lexico>quitar_preposiciones.exe
Archivo generado sin preposiciones: sinpreposiciones.txt
C:\Flex Windows\EditPlusPortable\lexico>
```



```
sinpreposiciones.txt
Archivo  Editar  Ver  H1  B  I  A
Vea pues, vine Comala porque llegó rumor acá podí dar papá, tal Pedro Pá ramo. madre fue dijo, yo prometí vendrí buscarlo apenas ella
muriera. apreté manos viera sí, porque ya estabaú ltimas yo, vea, prometíé ndole todo pedi.
" No dejes ir verlo- dijo-. llama así asa. Estoy segura va dar mucho gusto verlo. "
claro, ¿ yo qué hací? dije sí, iba ir, tanto repetí seguí dicié ndoselo ya costó soltar esas manos frí as quedaron tiesas.
Antes colgar zapatos dijo:
- "í ste tonto, no vas pedir nada ese tipo, nada má s nos toca. Usted sabe gustan cuentas claras chocolate espeso. ese olvido nos tuvo, có
brase lo caro esperá, total todo marrano llega diciembre. "
- Así será vieja
Pero vea, qué digo mentiras, pensaba cumplirle. hace poco empezaron dar vueltas sueñ os, llen@ @ rseme cabeza cucarachas. así, dar me
cuenta, fui armando todo mundo alrededor eseñ or Pedro Pá ramo, marido ama. eso fue agarré camino Comala.
es llegó solazo eavemaria... Má s caliente boca suegra poco venteaba traí olor olor fuerte
camino era ida venida: " Sube baja segú n va viene. va, sube; é l viene, baja."
- ¿ Có mo decí s vos llama pedazo ve allá abajo?
- Comala, patró n.
- ¿ Está seguro ya es Comala?
- Eavemaria ome, mas seguro muy valiente
- ¿ qué ve esto tan desanimado?
- Son tiempos, patró n.
Yo imaginaba distinto, vea, recordaba vieja cuentos: nostalgia suspiros. Pero, eso tan bueno no dan tanto. Igual ella siempre vivió
embohada Comala, volver; pero jamá s atrevió. Ahora yo vengo lugar. Traigo mismos ojos ella miró estas tierras, porque dejó ver: " Ahí,
pasando puerto cerros, hay vista muy bonita, mijo, llanura... "
verde, algo amarilla maí z todo punto. ese lugar mirada comala, blanqueando tierra, iluminando durante noche. voz era toda secreta, casi
apagada, hablara mismo... ama.
¿ va usted comala, puede saber? i preguntaban.
voy ver apa conteste.
i ahl dijo
Ln 46, Col 1 3,327 caracteres.  Texto sin formato  100%  Windows (CRLF)  ANSI
```

6. Extracción de palabras clave

Programa de LEX para el análisis léxico del documento fuente (normalización y obtención de palabras para Word Cloud)

```
C:\Flex Windows\EditPlusPortable\lexico>lex paramoS.l
C:\Flex Windows\EditPlusPortable\lexico>cc -w lex.yy.c -o paras2
C:\Flex Windows\EditPlusPortable\lexico>paras2.exe
Texto con regionalismos reemplazados guardado en: perdroparamocol_neutro.txt
C:\Flex Windows\EditPlusPortable\lexico>paras2.exe
Texto con regionalismos reemplazados guardado en: perdroparamocol_neutro.txt
C:\Flex Windows\EditPlusPortable\lexico>
```

```
File Edit Selection View Go Run ...  base de datos
perdroparamocol.txt  perdroparamocol_neutro.txt X
C:\Flex Windows\EditPlusPortable\lexico> perdroparamocol_neutro.txt
1 Vea pues, me vine para Comala porque me llegó un rumor de que por acá podía dar con mi papá, un tal Pedro Páramo. Mi madre fue la que me di
2 "No dejes de ir a verlo -me dijo-. Se llama así y así. Estoy segura de que le va a dar mucho gusto verlo."
3 Y claro, ¿yo qué hacía? Le dije que sí, que iba a ir, y tanto se lo repetí que seguí diciéndoselo hasta cuando ya me costó soltar esas mano
4 Antes de colgar los zapatos me dijo:
5 - "Oíste tonto, no le vas a pedir nada a ese tipo, nada más lo que nos toca. Usted sabe que a mí me gustan las cuentas claras y el chocolat
6 - Así será Vieja
7 Pero vea, para qué le digo mentiras, ni pensaba cumplirle. Hasta que hace poco me empezaron a dar vueltas unos sueños, a llen@rseme la cab
8 Y como es que llegó en un solazo que eavemaria. Más caliente que boca e suegra y lo poco que venteaba traía un olor a olor fuerte
9 El camino era de ida y venida: "Sube o baja según se va o se viene. Para el que va, sube; para él que viene, baja."
10 - ¿Cómo decís vos que se llama el pedazo que se ve allá abajo?
11 -Comala, patrón.
12 - ¿Está seguro de que ya es Comala?
13 -Eavemaria ome, mas seguro que un muy valiente
14 - ¿Y por qué se ve esto tan desanimado?
15 -Son los tiempos, patrón.
16 Yo me lo imaginaba distinto, vea, como lo recordaba mi vieja en sus cuentos: Con nostalgia y suspiros. Pero, de eso tan bueno no dan tanto.
17 verde, algo amarilla por el maíz en todo su punto. desde ese lugar mirada a comala, blanqueando la tierra, iluminando durante la noche. y s
18 ¿y a que va usted a comala, si se puede saber? oí que me preguntaban.
19 voy a ver a mi apa conteste.
20 ¡ahl dijo el
21 y nos quedamos callados.
22 caminábamos loma abajo. oyendo el paso alborotado de los burros. los ojos reventados por el sueño, en la canícula de agosto.
23 buena la fiesta le va montar- volví a oír la voz del que iba allí a mi lado-. se pondrá a saltar en un pie después de tantos años que nad
24 luego añadió.
25 sea usted quien sea, se alegrará en verlo.
26 en la reverberación del sol, la llanura parecía un charco transparente, deshecha en vapores por donde se veía el cielo gris. y mas allá, u
27 y qué cosas tiene tu padre, si se puede saber
28 ni le he visto- le dije- solo que disque se llama Pedro Páramo.
29 ¡ah conste!
30 sí, me dijeron que así se llamaba.
31 oí otra vez el ah del arriero.
```

7. Programa de YACC para el análisis sintáctico del documento fuente

1. Marco Conceptual de Traducción Adaptativa

❖ Fundamentos de Adaptación Lingüística

La transposición lingüística se sustenta en principios de lingüística contrastiva y sociolingüística aplicada. El proceso implica una reestructuración profunda que trasciende la traducción literal, requiriendo un análisis pragmático-discursivo del texto fuente (Baker, 1992). La adecuación dialectal opera mediante mapeo de variantes léxicas y reconstrucción de patrones sintácticos según normas del español estándar iberoamericano (Chesterman, 1998).

❖ Estratificación del Proceso Adaptativo

- Nivel Macrotextual: Reorganización de la estructura discursiva según convenciones retóricas de la región objetivo (James, 1980)
- Nivel Mesotextual: Reformulación de conectores lógicos y mecanismos de cohesión (Baker, 1992)
- Nivel Microtextual: Sustitución de unidades léxicas y reconfiguración de estructuras sintácticas (Chesterman, 1998)

2. Teoría de Análisis Emocional Textual

❖ Modelización Multidimensional Afectiva

El análisis emocional se fundamenta en teorías de psicología del lenguaje y lingüística computacional afectiva (Scherer, 2005). Se conceptualiza como un proceso de descomposición en capas sucesivas:

- Capa Proposicional: Identificación de actitudes proposicionales mediante marcadores modales y evidenciales (Osgood et al., 1957)
- Capa Expresiva: Detección de marcadores de subjetividad y evaluación (Mohammad & Turney, 2013)
- Capa Interpersonal: Análisis de posicionamiento dialógico y estrategias de face-work (Scherer, 2005)

❖ Taxonomía Emocional Aplicada

La clasificación emocional integra modelos dimensionales y categoriales, estableciendo correspondencias entre (Mohammad & Turney, 2013):

- Indicadores léxicos directos (términos emocionalmente cargados)
- Marcadores gramaticales (estructuras exclamativas, interrogativas retóricas)
- Patrones discursivos (narración de eventos, construcción de climax dramáticos)

3. Teoría de Representación Semiótica Visual

❖ Fundamentos de la Visualización Lingüística

La generación de nubes de palabras se conceptualiza como un proceso de semiosis visual donde (Bertin, 1967):

- La frecuencia léxica se traduce en prominencia visual
- Las relaciones semánticas se manifiestan mediante proximidad espacial
- Los dominios conceptuales se codifican a través de esquemas cromáticos

❖ Principios de Diseño Information-Foraging

La disposición visual sigue principios de (Tufte, 1983; Lohmann et al., 2017):

- Jerarquización por relevancia cognitiva
- Agrupamiento por campos semánticos
- Contrastación por polaridad afectiva

4. Arquitectura Teórica de Procesamiento Integrado

❖ Modelo de Pipeline Lingüístico-Cognitivo

El sistema se conceptualiza como una secuencia de transformaciones abstractas (Jurafsky & Martin, 2021):

- Fase de Descomposición: Segmentación en unidades procesables (Manning & Schütze, 1999)
- Fase de Análisis: Aplicación paralela de modelos especializados (Jurafsky & Martin, 2021)
- Fase de Síntesis: Integración de representaciones intermedias (Manning & Schütze, 1999)
- Fase de Realización: Generación de outputs multimodal (Jurafsky & Martin, 2021)

❖ Teoría de Retroalimentación Adaptativa

El mecanismo de ajuste continuo se basa en (Manning & Schütze, 1999):

- Monitoreo de métricas de comprensión lectora
- Actualización dinámica de bases de conocimiento
- Recalibración de parámetros de procesamiento

5. Modelado Gramatical Extendido

❖ Teoría de Gramáticas Formales Aplicadas

La especificación EBNF establece los fundamentos formales para el procesamiento automático, definiendo (ISO/IEC 14977:1996):

- Conjunto de categorías gramaticales relevantes

- Inventario de reglas de reescritura aplicables
 - Esquema de derivación canónica para el dominio específico
- ❖ Jerarquía de Constructos Lingüísticos
- La formalización distingue entre (Aho et al., 2006):
 - Unidades primarias (constituyentes inmediatos)
 - Estructuras complejas (patrones recursivos)
 - Relaciones de dependencia (restricciones de co-ocurrencia)

6. Modelo de Componentes Interrelacionados

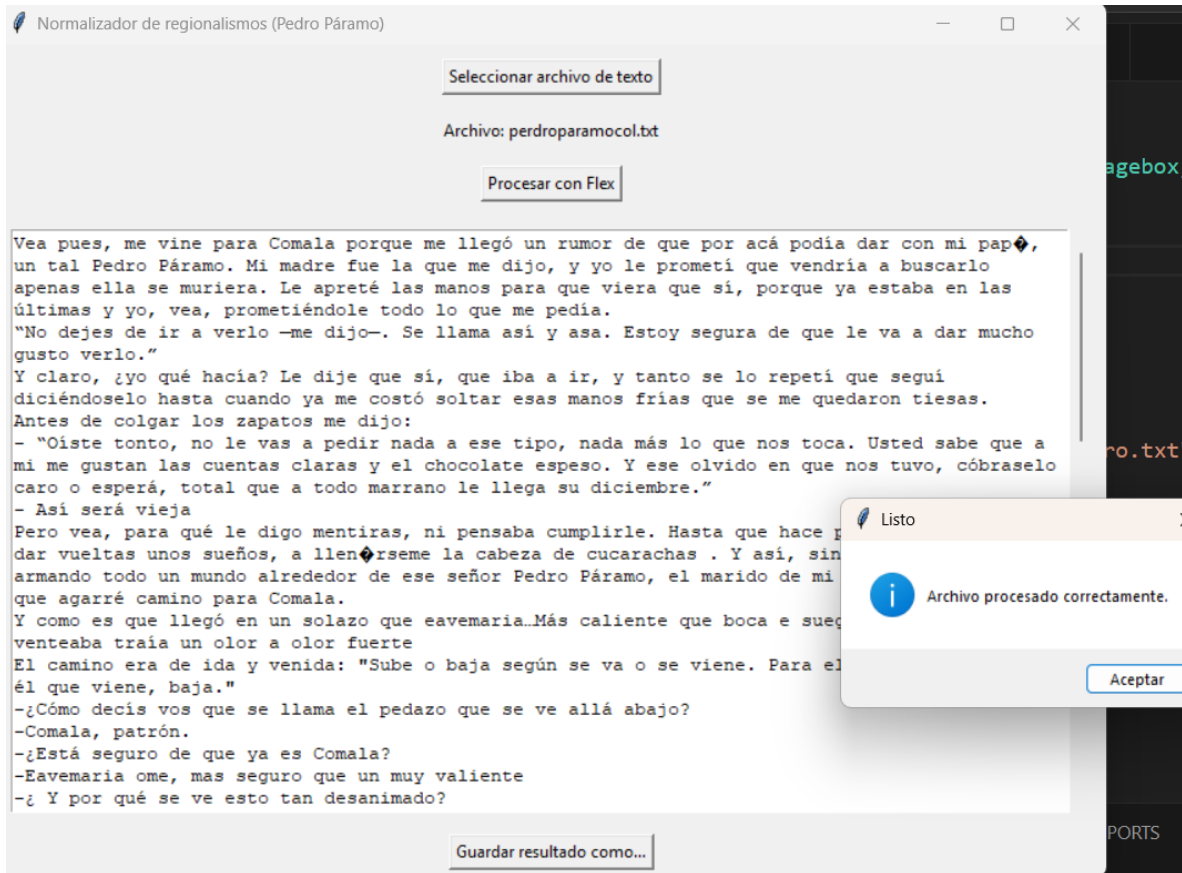
- ❖ Interfaz entre Subsistemas
- La arquitectura teórica define las interfaces conceptuales entre (Jurafsky & Martin, 2021):
 - Módulo de procesamiento lingüístico basal
 - Subsistema de análisis afectivo
 - Componente de visualización semántica

- ❖ Teoría de Flujo de Información

El modelo especifica las transformaciones sucesivas de representaciones intermedias mediante (Manning & Schütze, 1999):

- Operaciones de mapeo entre niveles de abstracción
- Transformaciones de formato entre componentes
- Proyecciones entre dominios representacionales

8. Desarrollo de la interfase gráfica



9. Interacción de Equipo México/Colombia



Integración 🇨🇴

Elmario7u7, ~ F, Roger, +57 301 6723911, +57 305 2772857, Tú



Info. del grupo

Añadir miembros

30/10/2025



Elmario7u7



Descripcion_de_palabras .docx

DOCX • 8 kB

Hola compañeros buenas tardes, nos podrían apoyar con estas descripciones de estos regionalismos es para el proyecto porfis

12:10 p. m.

Descargar



Porfa 12:22 p. m. ✓✓



12:22 p. m. ✓✓



Elmario7u7



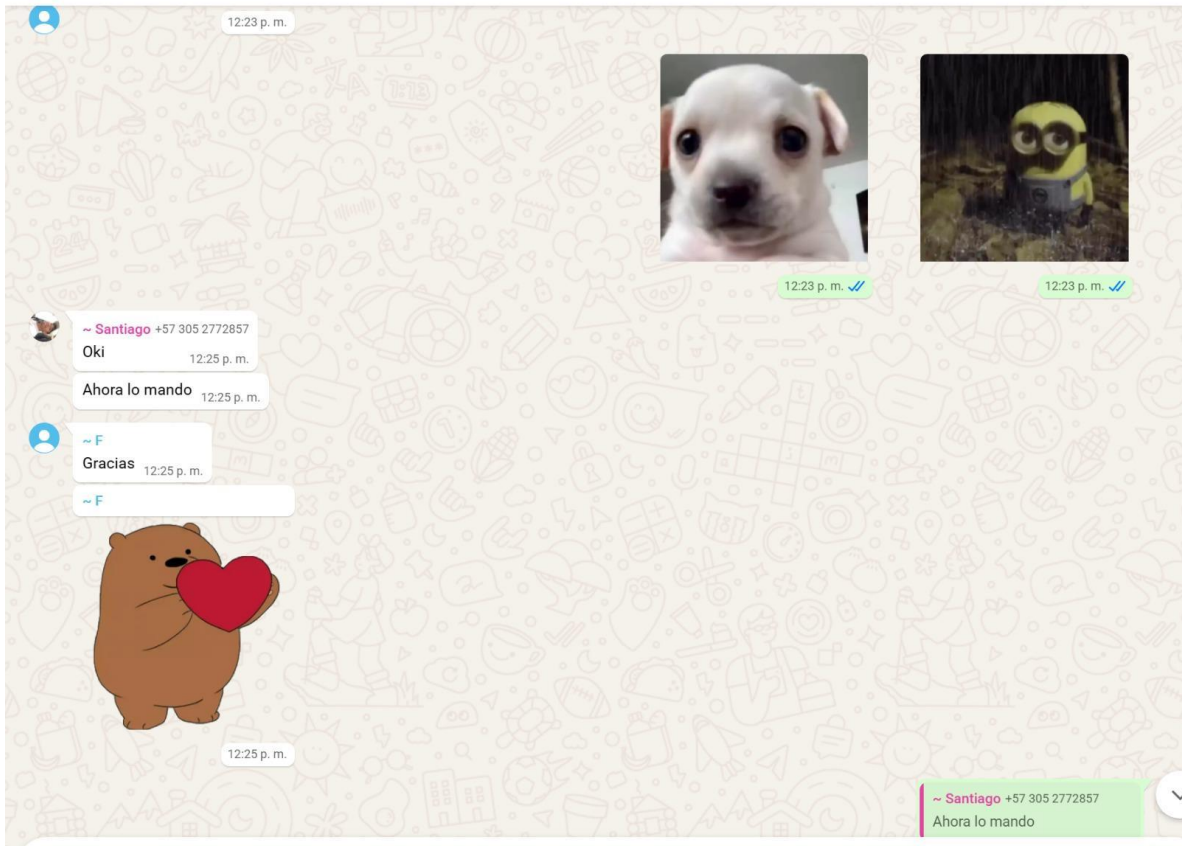
Escribe un mensaje





Integración 🇲🇽🇵🇪

Elmario7u7, ~ F, Roger, +57 301 6723911, +57 305 2772857, Tú





Integración 🇲🇽🇵🇷

Elmario7u7, ~ F, Roger, +57 301 6723911, +57 305 2772857, Tú



30/10/2025

12:27 p. m. ✓✓



~ Mariana~ 🦋 +57 301 6723911

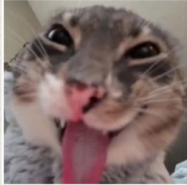
Elmario7u7

📄 Hola compañeros buenas tardes, nos podrían apoyar con estas descripciones de estos regionalismos es para el proyecto porfis

No

12:33 p. m.

~ Mariana~ 🦋 +57 301 6723911 ✓



12:33 p. m.



Roger



12:33 p. m.



~ F



Escribe un mensaje





Integración 🇨🇺

Elmario7u7, ~ F, Roger, +57 301 6723911, +57 305 2772857, Tú



12:46 p. m.

~ Mariana 🦋 +57 301 6723911

~ F

Si pueden para hoy por favor si para mañana

Claro que sí, con mucho gusto les ayudamos

1:04 p. m.

~ Mariana 🦋 +57 301 6723911



1:04 p. m.

Mentiras, hágale qué a la noche o a más tardar mañana procuramos enviarlo

1:04 p. m.



Roger

~ Mariana 🦋 +57 301 6723911

Claro que sí, con mucho gusto les ayudamos

Gracias

1:05 p. m.

Roger



Escribe un mensaje





Integración 🇲🇽🇵🇷

Elmario7u7, ~ Mariana, +57 301 6723911, +57 305 2772857, Tú



11:08 a. m.

5/11/2025



~ Mariana 🦋 +57 301 6723911



Descripcion_de_palabras 2.docx

DOCX • 2 MB

10:04 a. m.



Descargar

Hola compas, que pena la demora. Espero les sirva

10:05 a. m.



Elmario7u7

~ Mariana 🦋 +57 301 6723911

Hola compas, que pena la demora. Espero les sirva

Muchas gracias

10:17 a. m.

Elmario7u7



10:18 a. m.

~ Mariana 🦋 +57 301 6723911

Descripcion_de_palabras 2.docx



Resultados

1. Documento objeto (desde la terminal después de ejecutados en conjunto los programas de Lex y Yacc, con la “traducción” ya sin jerga colombiana)
2. Herramienta tecnológica desarrollada (desde donde se ejecuta automáticamente los programas desarrollados, en un entorno visual) útil para su uso en entornos educativos (que permita descargar un pdf del documento objeto).
3. Pantallas de la herramienta y evidencias de ejecución de la traducción
4. Pantallas de la herramienta y evidencias de ejecución de la extracción de las Stop Words
5. Pantallas de la herramienta y evidencias de ejecución de las palabras clave (cuáles y cuántas, puede ser desde la línea de comandos)
6. Herramienta tecnológica desarrollada (desde donde se ejecuta automáticamente los programas desarrollados, en un entorno visual) útil para su uso en entornos educativos.
7. Pantallas de la herramienta y evidencias del Análisis de sentimientos y Word cloud
8. Esbozo básico del planteamiento didáctico para emplear los resultados obtenidos.
9. Descripción general de la aplicación práctica en estrategias de atención diferenciada.

NOTA: Documentar los resultados mediante capturas de pantalla y descripciones de estas.

Conclusiones

Reflexiones sobre la colaboración interdisciplinaria y aprendizaje obtenido (elaborado con aportaciones de todos los estudiantes)

El desarrollo de este proyecto COIL permitió constatar la relevancia del análisis léxico-sintáctico como herramienta educativa para promover la comprensión intercultural del lenguaje. La implementación de los analizadores mediante Lex y Yacc evidenció la aplicabilidad de los autómatas finitos y las gramáticas formales en la resolución de problemas lingüísticos reales.

Desde la perspectiva técnica, se logró diseñar un prototipo funcional que integra la tokenización, eliminación de palabras vacías, extracción de palabras clave y análisis emocional, fortaleciendo las competencias en procesamiento del lenguaje natural y visualización semántica.

En el plano académico, la colaboración entre los equipos de México y Colombia propició un intercambio de saberes que enriqueció la comprensión del fenómeno lingüístico regional y la importancia de adaptar los recursos tecnológicos al contexto educativo.

Finalmente, se concluye que este tipo de experiencias fomenta una formación más integral, en la que la tecnología se convierte en un medio para la inclusión, la empatía y la innovación educativa, reafirmando el papel del lenguaje como puente entre culturas.

Referencias

[En formato APA 7ª. Ed]

Aho, A. V., Ravi, S. A. V., Ullman, J. D. (1998). *Compiladores: Principios, técnicas y herramientas*. Addison Wesley Longman.

Giró, J., Vázquez, J., Meloni, B., Constable, L. (2015). *Lenguajes formales y teoría de autómatas*. Editorial Alfaomega. Argentina.

Gutú, O. (2013). Primer curso en teoría de autómatas y lenguajes formales. Pearson Educación. México.

Hopcroft, J. E. Motwani, R. & Ullman, J. D. (2007). Introducción a la teoría de autómatas, lenguajes y computación. Pearson Educación. México.

Anexo A. Documento fuente

1. Fragmento

VINE a Comala porque me dijeron que acá vivía mi padre, un tal Pedro Páramo. Mi madre me lo dijo. Y yo le prometí que vendría a verlo en cuanto ella muriera. Le apreté sus manos en señal de que lo haría, pues ella estaba por morir y yo en un plan de prometerlo todo. "No dejes de ir a visitarlo -me recomendó. Se llama de este modo y de este otro. Estoy segura de que le dar gusto conocerlo." Entonces no pude hacer otra cosa sino decirle que así lo haría, y de tanto decírselo se lo seguí diciendo aun después de que a mis manos les costó trabajo zafarse de sus manos muertas. Todavía antes me había dicho: -No vayas a pedirle nada. Exígele lo nuestro. Lo que estuvo obligado a darme y nunca me dio... El olvido en que nos tuvo, mi hijo, cóbraselo caro. -Así lo haré, madre. Pero no pensé cumplir mi promesa. Hasta que ahora pronto comencé a llenarme de sueños, a darle vuelo a las ilusiones. Y de este modo se me fue formando un mundo alrededor de la esperanza que era aquel señor llamado Pedro Páramo, el marido de mi madre. Por eso vine a Comala. Era ese tiempo de la canícula, cuando el aire de agosto sopla caliente, envenenado por el olor podrido de la saponarias. El camino subía y bajaba: "Sube o baja según se va o se viene. Para el que va, sube; para él que viene, baja." -¿Cómo dice usted que se llama el pueblo que se ve allá abajo? -Comala, señor. -¿Está seguro de que ya es Comala? -Seguro, señor. -¿Y por qué se ve esto tan triste? -Son los tiempos, señor. Yo imaginaba ver aquello a través de los recuerdos de mi madre; de su nostalgia, entre retazos de suspiros. Siempre vivió ella suspirando por Comala, por el retorno; pero jamás volvió. Ahora yo vengo en su lugar. Traigo los ojos con que ella miró estas cosas, porque me dio sus ojos para ver: "Hay allí, pasando el puerto de Los Colimotes, la vista muy hermosa de una llanura verde, algo amarilla por el maíz maduro. Desde ese lugar se ve Comala, blanqueando la tierra, iluminándola durante la noche." Y su voz era secreta, casi apagada, como si hablara consigo misma... Mi madre. -¿Y a qué va usted a Comala, si se puede saber? -oí que me preguntaban. -Voy a ver a mi padre contesté. -¡Ah! -

dijo él. Y volvimos al silencio. Caminábamos cuesta abajo, oyendo el trote rebotado de los burros. Los ojos reventados por el sopor del sueño, en la canícula de agosto. -Bonita fiesta le va a armar -volví a oír la voz del que iba allí a mi lado-. Se pondrá contento de ver a alguien después de tantos años que nadie viene por aquí. Luego añadió: -Sea usted quien sea, se alegrará de verlo. En la reverberación del sol, la llanura parecía una laguna transparente, deshecha en vapores por donde se traslucía un horizonte gris. Y más allá, una línea de montañas. Y todavía te; la más remota lejanía. -¿Y qué trazas tiene su padre, si se puede saber? -No lo conozco -le dije-. Sólo sé que se llama Pedro Páramo. -¡Ah!, vaya. -Sí, así me dijeron que se llamaba. Oí otra vez el "¡ah!" del arriero.

Me había topado con él en Los Encuentros, donde se cruzaban varios caminos. Me estuve allí esperando, hasta que al fin apareció este hombre. -¿A dónde va usted? -le pregunté. -Voy para abajo, señor. -¿Conoce un lugar llamado Comala? -Para allá mismo voy. Y lo seguí. Fui tras él tratando de emparejarme a su paso, hasta que pareció darse cuenta de que lo seguía disminuyó la prisa de su carrera. Después los dos íbamos tan pegados que casi nos tocábamos los hombros. -Yo también soy hijo de Pedro Páramo -me dijo. Una bandada de cuervos pasó cruzando el cielo vacío, haciendo cuar, cuar, cuar. Después de trastumbar los cerros, bajamos cada vez más. Habíamos dejado el aire caliente allá arriba y nos íbamos hundiendo en el puro calor sin aire. Todo parecía estar como en espera de algo. -Hace calor aquí -dije. -Sí, y esto no es nada me contestó el otro-. Cállese. Ya lo sentirá más fuerte cuando lleguemos a Comala. Aquello está sobre las brasas de la tierra, en la mera boca del infierno. Con decirle que muchos de los que allí se mueren, al llegar al infierno regresan por su cobija. -¿Conoce usted a Pedro Páramo? -le pregunté. Me atreví a hacerlo porque vi en sus ojos una gota de confianza. -¿Quién es? -volví a preguntar. -Un rencor vivo -me contestó él.

2. Traducción al coloquialismo colombiano

Vea pues, me vine **pa'** Comala porque me llegó un **runrún** de que por acá podía dar con mi **apá**, un tal Pedro Páramo. Mi **madrecita** fue la que me dijo, y yo le prometí que vendría a buscarlo apenas ella se muriera. Le apreté las manitos **pa'** que viera que sí, porque ya estaba en las últimas y yo, vea, prometiéndole todo lo que me pedía.

“No dejes de ir a verlo —me dijo—. Se llama así y asa. Estoy segura de que le va a dar mucho gusto verlo.”

Y claro, ¿yo qué hacía? Le dije que sí, que iba a ir, y tanto se lo repetí que seguí diciéndoselo hasta cuando ya me costó soltar esas **manitos** frías que se me quedaron tíasas.

Antes de colgar los **guayos** me dijo:

- “Oíste **zumbambico**, no le vas a pedir nada a ese **estoraque**, nada más lo que nos toca. Usted sabe que a mi me gustan las cuentas claras y el chocolate espeso. Y ese olvido en que nos tuvo, cóbraselo caro o esperá, total que a todo marrano le llega su **diciembre**.”

- Así será vieja

Pero vea, **pa'** qué le digo mentiras, ni pensaba cumplirle. Hasta que hace poco me empezaron a dar vueltas unos sueños, a **llenárseme** la cabeza de **cucarachas**. Y así, sin darme cuenta, fui armando todo un mundo alrededor de ese señor Pedro Páramo, el marido de mi ama. Por eso fue que agarré camino **pa'** Comala.

Y como es que llegó en un solazo que **eavemaria**... Más caliente que boca e suegra y lo poco que venteaba traía un olor a **cachupe**

El camino era de ida y venida: "Sube o baja según se va o se viene. Para el que va, sube; para él que viene, baja."

-¿Cómo decís vos que se llama el pedazo que se ve allá abajo?

-Comala, patrón.

-¿Está seguro de que ya es Comala?

-Eavemaria ome, mas seguro que un berraco Yo me lo imaginaba distinto, vea, como lo recordaba mi vieja en sus cuentos: Con nostalgia y suspiros. Pero, de eso tan bueno no dan tanto. Igual ella siempre vivió embobada por Comala, por volver; pero jamás se atrevió. Ahora yo vengo en su lugar. Traigo los mismos ojos con que ella miró estas tierras, porque me los dejó pa' ver: "Ahí, pasando el puerto de Los Colimotes, hay una vista muy bonita, mijo, de una llanura..."

verde, algo amarilla por el maíz en toditico su punto. desde ese lugar atisbo a comala, blanqueando la tierra, iluminando durante la noche. y su voz era toda secreta, casi apagada, como si hablara con el mismito... Mi ama.

¿y a que va usted a comala, si se puede saber? oí que me preguntaban.

voy a ver a mi apa conteste.

¡ah! dijo el

y nos quedamos callaitos.

caminábamos loma abajo. oyendo el paso revolotiado de los burros. los ojos reventados por el sueño, en la canícula de agosto.

buenota la fiesta le va montar- volví a oír la voz del que iba allí a mi lado-. se pondrá a saltar en un pie después de tantos años que nadie se asoma por aquí.

luego añadió.

-¿ Y por qué se ve esto tan achicopalado?

-Son los tiempos, patrón.

sea usted quien sea, se alegrará en verlo.

en la reverberación del sol, la llanura parecía un charco transparente, deshecha en vapores por **onde** se veía el cielo gris. y mas **allacito**, una línea de montañas todavia la mas lejana de las lejanas

y qué cosas tiene tu padre, si se puede saber

ni le he visto- le dije- solo que disque se llama Pedro Páramo.

¡ah conste!

si, me dijeron que así se llamaba.

oi otra vez el ah del arriero.

me lo había cruzado en **toditicos** los encuentros, donde se metía varios caminos. Me estuve allí esperando, hasta que por fin apareció tal hombre.

¿**Pa'** donde va usted?- le pregunté.

¿De casualidad **buste** conoce comala?

pa allí **mismito** voy.

y lo seguí, fui tras él tratando de seguirle el paso, hasta como que cayo en cuenta de que estaba detrás de el bajo el paso, después los dos íbamos tan **pegaditicos** que casi nos chocamos los hombros.

Yo también soy un **surrongo** de Pedro Páramo.

un vuelo de cuervos paso por el cielo vacío, haciendo cuar, cuar, cuar.

después de rodar la montaña, bajamos cada vez más. habíamos dejado el aire caliente alla arriba y nos íbamos hundiendo en el bochorno sin aire. Todo parecia estar como en espera de algo.

Hace como **calorcito**-dije.

si, y **estico** no es nada me contestó el otro-. cálmese. ya lo sentirá más fuerte cuando lleguemos a comala. aquello entre las brasas de la tierra, en la mera boca del pozo ardiente.

¿Conoce usted disque a Pedro Páramo?

me atreví a hacerle esa pregunta porque vi en sus ojos una gota de buena fe.

¿Quién es? volví a preguntar

un rencor vivo me contestó.

Palabra coloquial	Descripción /
-------------------	---------------

	Significado
pa'	Amigo/para
apá	Padre
runrún	Rumor
madrecita	Madre
manitos	Amigo/Compañero
diciéndoselo	Decir/Contar
guayos	Zapatos deportivos

zumbambico	Niño malcriado
estoraque	Persona pesada/insopora ble
diciembre	Mes del año donde se celebra el nacimiento de Jesús (Navidad)
llenárseme	Llenar
cucarachas	Insecto

eavemaria	<p>Expresión paísa que denota</p> <p>admiración, sorpresa, alegría o decepción</p>
cachupe	<p>Olor desagradable</p>
ome	<p>Muletilla que expresa</p> <p>sorpresa, admiración.</p> <p>Tambien se utiliza para dirigirse a alguien</p>
berraco	<p>Persona valiente/ cualidad de</p> <p>algo difícil o exagerado</p>

Colimotes	Gentilicio que nombra a los nacidos en Colima
toditico	Diminutivo de todo
atisbo	Indicio o sospecha
ama	Madre/Mamá
callaitos	Acción de callar

<p>revolotiado</p>	<p>participio pasado del verbo</p> <p>"revolotear" utilizada comúnmente para describir el</p> <p>movimiento de una persona alrededor de otra o un movimiento</p> <p>continuo y rápido.</p>
<p>achicopalado</p>	<p>Triste</p>
<p>onde</p>	<p>Donde</p>
<p>allacito</p>	<p>Indica lugar/allá</p>

toditicos	Diminutivo de todo
buste	Usted
mismo	Diminutivo de mismo/ puede indicar una acción que deba hacerse ya
pegaditicos	Diminutivo de pegado/estrech o
surrongo	persona impertinente, necia o cansona

calorcito	Dimutivo de calor
estico	Dentro de poco