ÁREA 2. DESARROLLO DE SOFTWARE DE BASE

SUBÁREA 2.2 COMPILADORES

BIBLIOGRAFÍA DE LA GUÍA:

- Aho, Alfred. (2008). Compiladores, principios, técnicas y herramientas. Prentice Hall México.
- Giró, Juan et al. (2015) Lenguajes formales y teoría de autómatas. Alfaomega.
- Hopcroft, John E., Motwani, Rajeev y Ullman., Jeffrey D. (2007). Introducción a la teoría de autómatas, lenguajes y computación. 3a ed. Pearson, Addison Wesley.
- Kelley, Dean. (1995). Teoría de Autómatas y Lenguajes Formales. España: Prentice Hall.
- Rendell, Paul. (2016). Turing Machine Universality of the Game of Life. Springer International Publishing.

Acerca de la bibliografía: El libro más utilizado para el estudio de compiladores es el Alfred Aho. Sin embargo, este libro es muy extenso y abarca muchos temas. Es poco probable que en el examen pregunten cosas demasiado específicas y especializadas. Por lo tanto, el libro se debe usar como referencia general, y no tanto para estudiar en detalle todo.

TEMAS IMPORTANTES:

Conceptos generales de compiladores e intérpretes:

El primer capítulo del libro de Aho es recomendable leerlo, pero también se pueden apoyar de recursos como las páginas:

https://towardsdatascience.com/understanding-compilers-for-humans-version-2-157f0edb02dd

https://medium.com/@thelukaswils/understanding-compilers-for-humans-ba970e045877

Análisis Léxico:

Además de que tienen que entender la función de este tipo de análisis, este tema también involucra el uso de **Expresiones Regulares** y **Autómatas**, que forman parte del tema de **Lenguajes Regulares** (que se estudiaron en la materia de Algoritmos y Complejidad).

Lenguajes Regulares:

Este tipo de lenguajes son representados o definidos por las Expresiones Regulares y los Autómatas. Es el tipo de lenguaje más simple que un modelo computacional puede representar.

Pueden consultar los siguientes videos de la materia de Algoritmos y Complejidad acerca de este tema, los autómatas y las expresiones regulares:

https://www.youtube.com/watch?v=R74u8seZpHA (Introducción a los Lenguajes Regulares)
https://www.youtube.com/watch?v=SzLwybf0crc (Diseño de Autómatas)
https://www.youtube.com/watch?v=0NyVzmpCjp4 (Autómatas no Deterministas)
https://www.youtube.com/watch?v=HEYUUVqhfOY (Expresiones Regulares)

Expresiones Regulares:

Es importante saber cómo se utilizan y sus operadores para poder representar lenguajes regulares. En el caso de los compiladores se usan principalmente para representar nombres de variables, expresiones aritméticas, entre otras cosas. Se les recomienda usar el libro usado en la materia de Algoritmos y Complejidad: Sipser, Michael. *Introduction to the Theory of Computation*. Cengage Learning, 3rd Edition, 2012. Este tema también viene en el libro de Aho.

Autómatas:

Los autómatas son equivalentes a las expresiones regulares, por lo que también se suelen usar para el estudio y diseño del analizador léxico. Este tema se puede subdividir en varios subtemas importantes:

- DFAs (autómatas finitos deterministas)
- NFAs (autómatas finitos NO deterministas)
- Conversión entre NFAs -> DFAs

Al igual que las expresiones regulares, se recomienda el uso del libro de Sipser para su estudio; sin dejar de consultar el libro de Aho para ver los temas desde el punto de vista de un compilador, ya que el libro de Sipser es más general.

Análisis Sintáctico:

Después del analizador léxico toca el turno del analizador sintáctico. Es importante entender el proceso, sus elementos, y cómo funciona. Uno de los temas que se derivan de este es el de las **Gramáticas Libres de Contexto** y los **Autómatas de Pila** (pushdown automata); que a su vez forman parte del tema de **Lenguajes Libres de Contexto** (context-free languages).

Lenguajes Libres de Contexto:

Este tipo de lenguajes son más poderosos que los regulares, ya que permiten representar lenguajes más complejos. Se usan las Gramáticas Libres de Contexto y los Autómatas de Pila para definirlos y generarlos. Pueden consultar los siguientes videos de la materia de Algoritmos y Complejidad acerca de estos temas:

https://www.youtube.com/watch?v=8AvPDJLmEx4 (Gramáticas Libres de Contexto)
https://www.youtube.com/watch?v=mdsCjnowXJY (Autómatas de Pila)

Gramáticas Libres de Contexto:

Es importante saber cómo se definen y sus reglas de operación para poder representar lenguajes libres de contexto. En el caso de los compiladores se usan principalmente para definir la gramática del lenguaje de programación que se desea compilar. Por lo que es de suma importancia este tema. Conceptos importantes son los de **producción**, **derivación**, **elementos terminales** y **no terminales**, **árbol de análisis sintáctico**, entre otros. Se les recomienda usar el libro usado en la materia de Algoritmos y Complejidad, como introducción: Sipser, Michael. *Introduction to the Theory of Computation*. Cengage Learning, 3rd Edition, 2012. También leer este tema en el libro de Aho.

Autómatas de Pila:

En inglés son los pushdown automatas, y son equivalentes a las gramáticas libres de contexto. Son parecidos, pero NO iguales, a los autómatas deterministas finitos. Se recomienda el uso del libro de Sipser para su estudio; sin dejar de consultar el libro de Aho.

Máquinas de Turing:

Las máquinas de Turing son un modelo computacional (en la categoría de los autómatas, gramáticas, expresiones regulares y autómatas de pila), pero mucho más poderoso. Normalmente, estas máquinas no se estudian en el tema de Compiladores, pero la bibliografía de la guía contiene un libro de este tema, por lo que es posible que lo pregunten. Se les recomienda usar el libro usado en la materia de Algoritmos y Complejidad: Sipser, Michael. *Introduction to the Theory of Computation*. Cengage Learning, 3rd Edition, 2012. Además, pueden consultar los siguientes videos de la misma materia:

https://www.youtube.com/watch?v=XhkxF3u62_U
https://www.youtube.com/watch?v=Qj16AeF7ecU

Generación de Código Intermedio:

Es importante entender el proceso, que sucede después del parser (análisis sintáctico), y antes del generador de código. Para este proceso se pueden utilizar variantes del árbol sintáctico y el código de tres direcciones, que sería importante conocer y entender en general a qué se refieren. Este tema se encuentra en el libro de Aho.