



ОСНОВИ РОЗРОБКИ ТРАНСЛЯТОРІВ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

| Рівень вищої освіти | Перший (бакалаврський) |
|---|---|
| Галузь знань | 12 Інформаційні технології |
| Спеціальність | 121 Інженерія програмного забезпечення |
| Освітня програма | Інженерія програмного забезпечення інтелектуальних кібер-фізичних систем в енергетиці |
| Статус дисципліни | Нормативна |
| Форма навчання | заочна |
| Рік підготовки, семестр | 3 курс осінній семестр |
| Обсяг дисципліни | 7,5 кредитів ECTS /225 годин (16 годин лекцій, 12 годин практичних занять, 197 годин — самостійна робота студента) |
| Семестровий контроль/ контрольні заходи | Екзамен, МКР, РГР, представлення робіт комп'ютерного практикуму |
| Розклад занять | http://rozklad.kpi.ua |
| Мова викладання | Українська |
| Інформація про керівника курсу / викладачів | Лектор: к.т.н., доцент, Статівка Юрій Іванович, statyvka-yu@ill.kpi.ua Практичні: к.т.н., доцент, Статівка Юрій Іванович, statyvka-yu@ill.kpi.ua |
| Розміщення курсу | https://ecampus.kpi.ua |

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

В курсі розглядаються як фундаментальні принципи створення і імплементації мов програмування загального призначення, так і практичні аспекти автоматизації розробки трансляторів.

Значна увага приділяється проектуванню мов програмування, їх специфікації, математичним основам представлення та опрацювання формальних мов.

Розглядаються загальні та спеціалізовані засоби побудови компіляторів та інтерпретаторів: алгоритми лексичного, синтаксичного та семантичного аналізу, генерування коду, засобів виконання коду цільової мови.

Метою дисципліни є опанування студентами теоретичних знань та набуття практичного досвіду проєктування і імплементації імперативних мов програмування .

Предмет дисципліни — методи та алгоритми побудови, специфікації та імплементації мов програмування високого рівня.

Завдання. В результаті вивчення дисципліни у студентів повинні сформуватися наступні компетентності:

фахові:

— здатність брати участь у проєктуванні програмного забезпечення, включаючи проведення моделювання (формальний опис) його структури, поведінки та процесів функціонування. (ФК 02);

— здатність розробляти архітектури, модулі та компоненти програмних систем (ФК 03);

— здатність застосовувати транслятори мов програмування при реалізації програмних систем (ФК 19).

Після засвоєння навчальної дисципліни студенти мають продемонструвати такі програмні результати навчання:

— аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідникові ресурси і знання з урахуванням сучасних досягнень науки і техніки (ПРН 01);

— знати і застосовувати відповідні математичні поняття, методи доменного, системного і об'єктно-орієнтованого аналізу та математичного моделювання для розробки програмного забезпечення (ПРН 05);

— знати і застосовувати на практиці фундаментальні концепції, парадигми і основні принципи функціонування мовних, інструментальних і обчислювальних засобів інженерії програмного забезпечення (ПРН 07);

— знати і застосовувати методи розробки алгоритмів, конструювання програмного забезпечення та структур даних і знань (ПРН 13);

— знати і вміти використовувати фундаментальний математичний інструментарій при побудові алгоритмів та розробленні сучасного програмного забезпечення (ПРН 25);

— застосовувати на практиці фундаментальні концепції, парадигми і основні принципи функціонування мовних, інструментальних і обчислювальних засобів інженерії програмного забезпечення (ПРН 32).

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Пререквізити дисципліни. Знання, отримані при вивченні дисциплін: «Компоненти програмної інженерії», «Основи програмування», «Комп'ютерна дискретна математика», «Алгоритми та структури даних».

Постреквізити дисципліни. Компетенції, отримані студентами в процесі вивчення дисципліни можуть бути використані при виконанні дипломної роботи.

3. Зміст навчальної дисципліни

Розділ 1. Основи теорії формальних мов

Тема 1.1. Вступ. Формальні мови. Компіляція. Фази компіляції.

Тема 1.2. Елементи теорії формальних мов.

Тема 1.3. Скінченні автомати.

Розділ 2. Лексичний, синтаксичний та семантичний аналіз програм

Тема 2.1. Лексичний аналіз програм.

Тема 2.2. Автомат з магазинною пам'яттю (МП-автомат).

Тема 2.3. Призначення та види синтаксичного аналізу. Метод рекурсивного спуску.

Тема 2.4. Семантичний аналіз програм.

Тема 2.5 Приклад реалізації семантичного аналізатора.

Розділ 3. Трансляція та виконання програм

Тема 3.1. Проміжні форми подання програм.

Тема 3.2. Трансляція у ПОЛІЗ і виконання ПОЛІЗ-програм.

Тема 3.3. Віртуальні стекові машини. CLR. JVM.

Тема 3.4. Переклад на CIL та Java-асемблер.

Розділ 4. Автоматизація розробки трансляторів.

Тема 4.1. Засоби автоматизації розробки трансляторів

Тема 4.2. ANTLR4 як компілятор компіляторів.

Тема 4.3. Варіанти використання ANTLR4.

Розділ 5. Окремі питання розробки трансляторів.

Тема 5.1. Генерування коду для реєстрових машин.

Тема 5.2. Оптимізація цільового коду.

Тема 5.3. Розробка трансляторів. Підсумки.

4. Навчальні матеріали та ресурси

Основна література

1. Стативка Ю.І. Формальні мови: Основні концепти і представлення [Текст]: навч. посіб. / Ю. І. Стативка. – Київ: КПІ ім. Ігоря Сікорського, 2023. – 87 с.
2. Медведєва В.М. Транслятори: лексичний та синтаксичний аналізатори [Текст] : навч. посіб. / В.М. Медведєва, В.А. Третяк. – К. : НТУУ «КПІ», 2012. – 148с.
3. Медведєва В.М. Транслятори: внутрішнє подання програм та інтерпретація [Текст] : навч. посіб. / В.М. Медведєва, В.А. Третяк. . – К. : Текст, 2015. – 144 с.
4. Aho, Alfred, Lam, Monica, Sethi, Ravi, Ullman, Jeffrey Compilers: Principles, Techniques, and Tools, 2nd edition. - Addison Wesley, 2006. - 1040 p.
5. Jason Bock CIL Programming: Under the Hood of .NET. - Apress, 2008. 355 p.
6. Joshua Engel Programming for the Java™ Virtual Machine. - Addison Wesley, 1999. 512 p.

Додаткова література

7. Hopcroft J. E., Motwani R., Ullman J. D. Introduction to Automata Theory, Languages, and Computation (3rd ed.). Pearson, 2013. 560 p.
8. A. V. Aho and J. D. Ullman, The Theory of Parsing, Translation, and Compiling, Vol. 1, Parsing. Prentice Hall, 1972. 1030 p.
9. Winskel G. The formal semantics of programming languages: an introduction. Cambridge, Massachusetts, London: MIT Press, 1993. 384 p.

Інформаційні ресурси

10. <https://www.antlr.org/>
11. <https://github.com/antlr/antlr4/blob/master/doc/getting-started.md>
12. <https://github.com/antlr/antlr4/blob/4.6/doc/index.md>
13. [EBNF Visualizer: веб-сайт. URL: http://dotnet.jku.at/applications/visualizer/](http://dotnet.jku.at/applications/visualizer/)

14. [EBNF 2 RailRoad: веб-сайт. URL: https://github.com/matthijsgroen/ebnf2railroad](https://github.com/matthijsgroen/ebnf2railroad)
15. [Railroad Diagram Generator: веб-сайт. URL: https://bottlecaps.de/rr/ui](https://bottlecaps.de/rr/ui)
16. [JFLAP Version 7.1 RELEASED July 27, 2018: веб-сайт. URL: https://www.jflap.org/.](https://www.jflap.org/)
17. [Rodger S. H., Finley T. W. JFLAP – An Interactive Formal Languages and Automata Package Jones & Bartlett Publishers, Sudbury, 2005. 212 p. – URL: https://www2.cs.duke.edu/csed/jflap/jflapbook/jflapbook2006.pdf.](https://www2.cs.duke.edu/csed/jflap/jflapbook/jflapbook2006.pdf)
18. <https://learn.microsoft.com/en-us/dotnet/framework/tools/ilasm-exe-il-assembler>
19. <https://learn.microsoft.com/en-us/dotnet/framework/tools/ildasm-exe-il-disassembler>
20. <https://docs.oracle.com/javase/specs/jvms/se7/html/>
21. <https://github.com/drstrng/Krakatau-noff>

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Лекційні заняття

Розділ 1. Основи теорії формальних мов

Лекція 1. Вступ. Формальні мови. Компіляція. Фази компіляції.

Місце та призначення компілятора в програмному забезпеченні комп'ютера. Основні блоки та їх взаємодія. Структура компілятора.

Лекція 2. Елементи теорії формальних мов.

Основні поняття: мова, граматика, виведення, сентенційна форма, рекурсивні правила.

Приведення граматик. Трансформація граматик. Усунення ϵ -продукцій. Усунення ланцюгових правил. Усунення марних символів. Факторизація граматик. Нормальна форма Чомскі.

Форми подання граматик мов програмування: БНФ, РБНФ. Синтаксичне дерево і абстрактне синтаксичне дерево. Синтаксичні діаграми.

Лекція 3. Скінченні автомати.

Поняття розпізнавача. Види розпізнавачів. Скінченні автомати. Визначення, правила функціонування, форми подання. Детерміновані скінченні автомати (ДСА). Недетерміновані скінченні автомати (НСА). Мінімізація ДСА. Автоматні мови. Еквівалентність множини регулярних і автоматних мов.

Розділ 2. Лексичний, синтаксичний та семантичний аналіз програм

Лекція 4. Лексичний аналіз програм.

Лексичний аналізатор. Призначення, вхід, вихід. Діаграма станів. Лексичний аналіз програм на основі діаграми станів (переходів). Приклад реалізації лексичного аналізатора на основі діаграми станів.

Лекція 5. Призначення та види синтаксичного аналізу. Рекурсивний спуск

Призначення, види синтаксичного розбору. Низхідний розбір. Алгоритм рекурсивного спуску. Перетворення граматика для реалізації розбору без повернень та зациклювання.

Лекція 6. Семантичний аналіз програм.

Семантика мови і поведінка програми. Аналіз відповідності програми семантиці, декларованій у специфікації мови. Приклад реалізації семантичного аналізатора.

Розділ 3. Трансляція та виконання програм

Лекція 7. Проміжні форми подання програм і генерування коду.

Призначення проміжних форм, їх види. Польський інверсний запис. Алгоритм Дейкстри побудови ПОЛІЗ. Пріоритети. Віртуальна стекова машина.

Розділ 4. Автоматизація розробки трансляторів

Лекція 8. Засоби автоматизації розробки трансляторів

Генератор парсерів: Yacc, Bison, Lex, Flex, Narry, BNFC, ANTLR4.

ANTLR4: Структура, встановлення, використання. Засоби проектування та часу виконання.

Нотація. Приклад використання

Практичні заняття

Практичне заняття 1. Елементи теорії формальних мов.

Формальні мови. Ланцюжки, операції над ними. Формальні граматики. Виведення. Побудова граматики за описом мови. Побудова опису мови за граматиною.

Практичне заняття 2. Специфікація формальних мов.

Структура специфікації — лексика, синтаксис, семантика. Ревю специфікацій.

Практичне заняття 3. Скінченні автомати і лексичний аналіз

Скінченні автомати. Діаграма станів. Лексичний аналіз.

Практичне заняття 4. Синтаксичний і семантичний аналіз

Метод рекурсивного спуску. Реалізація. Семантика мови і поведінка програми.

Практичне заняття 5. Генерування коду і виконання програм.

Генерування postfix- та il-коду. Postfix Stack Machine. CLR.

Практичне заняття 6. МКР.

Виконання модульної контрольної роботи.

Перелік робіт комп'ютерного практикуму

1. Лексичний аналіз програм.
2. Синтаксичний аналіз програм.
3. Семантичний аналіз програм.
4. Трансляція у ПОЛІЗ та виконання ПОЛІЗ-програм.
5. Трансляція для віртуальної стекової машини (CLR/JVM).
6. Автоматизація побудови транслятора засобами ANTLR4.

Розрахунково-графічна робота

Індивідуальне завдання на проектування та специфікацію простої імперативної мови програмування.

Контрольна робота

Метою контрольної роботи є закріплення та перевірка теоретичних знань із освітнього компонента, набуття студентами практичних навичок самостійного вирішення задач.

Модульна контрольна робота (МКР) виконується після вивчення Розділів 1-2. Кожен студент отримує завдання одного з варіантів, відповідно до якого необхідно виконати побудову граматики, скінченного або магазинного автомата, та виконати приведення граматики.

Самостійна робота студента

Теми для самостійного опрацювання:

Розділ 1. Основи теорії формальних мов

Тема 1. Робота з граматами у JFLAP.

[1, розд. 1.5; 15, розд. 3]

Тема 2. Класифікація граматик Чомські.

[1, розд. 1.3.4; 2, розд. 2.1.2]

Тема 3. Приведення граматик.

Трансформація граматик. Усунення ϵ -продукцій. Усунення ланцюгових правил. Усунення лівої рекурсії. Усунення марних символів. Факторизація граматик. Нормальна форма Чомські.

[1, розділ 1.3.5]

Тема 4. Форми подання граматик.

Нотація Бекуса-Наура. Розширена нотація Бекуса-Наура. Синтаксичне дерево. Абстрактне синтаксичне дерево. Синтаксичні діаграми. Однозначність граматики. Варіанти граматик арифметичного виразу. Ebnf-visualizer (та інші інструменти) для генерування синтаксичних діаграм.

[1, розділ 1.3.5; 1, розд. 1.2.4; 2, розділ 2.1.4]

Тема 5. Регулярні вирази та мови.

Визначення. Метамова представлення регулярних виразів. Граф регулярного виразу. Еквівалентність множини регулярних і автоматних мов. Робота з регулярними виразами у JFLAP. [1, розділ 1.2; 2, розділ 3.5; 15, розд. 4]

Тема 6. Скінченні автомати.

Поняття розпізнавача. Види розпізнавачів. Скінченні автомати. Визначення, правила функціонування, форми подання. Детерміновані скінченні автомати (ДСА). Недетерміновані скінченні автомати (НСА). Мінімізація ДСА. Автоматні мови. Еквівалентність множини регулярних і автоматних мов. Робота з скінченними автоматами у JFLAP.

[1, розділ 1.4.1; 2, розділ 3.4; 15, розд. 4]

Розділ 2. Лексичний, синтаксичний та семантичний аналіз програм

Тема 7. Лексичний аналіз програм.

Лексичний аналізатор. Призначення, вхід, вихід. Діаграма станів. Лексичний аналіз програм на основі діаграми станів.

[2, розділи 3.1-3.3; 4 розділ 3.4]

Тема 8. Приклад реалізації лексичного аналізатора на основі діаграми станів.

Постановка завдання. Приклад виконання.

Див. матеріали до КП 2

Тема 9. Автомат з магазинною пам'яттю (МП-автомат).

Визначення, правила функціонування, форми подання. МП-автомат і контекстно вільні граматики. Робота з МП-автоматами у JFLAP.

[1, розділ 1.4.2; 2, розділ 4.3; 15, розд. 4]

Тема 10. Призначення та види синтаксичного аналізу. Рекурсивний спуск

[2, розділи 4.1-4.2; 4, розділи 2.4 та 4.4]

Тема 11. Реалізація синтаксичного аналізатора на основі алгоритму рекурсивного спуску.

Див. матеріали до КП 3, частина 1.

Тема 12. Використання автомата з магазинною пам'яттю для синтаксичного розбору.

Нерекурсивний синтаксичний аналіз. Низхідний та висхідний синтаксичний аналіз.

[2, розділи 4.3.2-4.3.4; 4, розділ 4.4.4]

Тема 13. Приклад реалізації семантичного аналізатора.

Див. матеріали до КП 3, частина 2.

Розділ 3. Трансляція та виконання програм

Тема 14. Трансляція у ПОЛІЗ.

Переклад у ПОЛІЗ конструкцій мови високого рівня.

[3, розділи 1.1, розділ 4.4]

Тема 15. Виконання ПОЛІЗ-програм: Postfix State Machine (PSM).

Див. матеріали до КП 4, частина 2.

Тема 16. Common Language Runtime (CLR). Переклад на Common Intermediate Language (CIL).

Віртуальна стекова машина .Net. ILASM. ILDASM. Набір інструкцій CLR. Асемблювання і виконання il-програм.

[5, розділ 3; 16; 17]

Тема 17. Java Virtual Machine (JVM). Переклад на Java-асемблер.

Віртуальна стекова машина Java. Набір інструкцій JVM. Засоби асемблювання та дизасемблювання для JVM. Асемблери та дизасемблери третіх сторін: Krakatau.

Асемблювання і виконання j-програм.

[6, розділи 2,3; 20; 21; 18]

Розділ 4. Автоматизація розробки трансляторів

Тема 18. Засоби автоматизації розробки трансляторів

ANTLR4: Структура, встановлення, використання. Засоби проектування та часу виконання. Особливості представлення граматики.

[10; 12]

Приблизний розподіл часу СРС

| № з/п | Вид самостійної роботи | Кількість | Кількість годин СРС |
|--------------|--|-----------|---------------------|
| 1 | Виконання робіт комп'ютерного практикуму | 6 | 36 |
| 2 | Виконання РГР | 1 | 12 |
| 3 | Підготовка до практичних занять | 6 | 6 |
| 4 | Опрацювання тем, винесених на СРС | 18 | 107 |
| 5 | Підготовка до МКР | 1 | 6 |
| 6 | Підготовка до екзамену | 1 | 30 |
| Разом | | | 197 |

Політика та контроль

6. Політика навчальної дисципліни (освітнього компонента)

Система вимог при вивченні дисципліни:

- правила відвідування занять: заборонено оцінювати присутність або відсутність здобувача на аудиторному занятті, в тому числі нараховувати заохочувальні або штрафні бали. Відповідно до РСО даної дисципліни бали нараховують за відповідні види навчальної активності на лекційних та практичних заняттях.
- правила поведінки на заняттях: студент має можливість отримувати бали за відповідні види навчальної активності на лекційних та практичних заняттях, передбачені РСО дисципліни. Використання засобів зв'язку для пошуку інформації на гугл-диску викладача, в інтернеті, в дистанційному курсі на платформі Сікорський здійснюється за умови вказівки викладача;
- політика дедлайнів та перескладань: якщо студент не проходив або не з'явився на МКР (без поважної причини), його результат оцінюється у 0 балів. Перескладання результатів МКР не передбачено;
- політика щодо академічної доброчесності: Кодекс честі Національного технічного університету України «Київський політехнічний інститут» <https://kpi.ua/files/honorcode.pdf> встановлює загальні моральні принципи, правила етичної поведінки осіб та передбачає політику академічної доброчесності для осіб, що працюють і навчаються в університеті, якими вони мають керуватись у своїй діяльності, в тому числі при вивченні та складанні контрольних заходів з дисципліни «Основи розробки трансляторів»;
- при використанні цифрових засобів зв'язку з викладачем (мобільний зв'язок, електронна пошта, переписка на форумах та у соцмережах тощо) необхідно

дотримуватись загальноприйнятих етичних норм, зокрема бути ввічливим та обмежувати спілкування робочим часом викладача.

7. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Система рейтингових (вагових) балів та критерії оцінювання

Поточний контроль: активність — питання і відповіді на лекційних заняттях, питання і відповіді та участь у ревію на практичних заняттях; вправи на практичних заняттях; МКР; виконання завдань до практичних занять; виконання та представлення робіт комп'ютерного практикуму, виконання та представлення РГР.

1) Активність

Активність (частота, змістовність) участі студента у процесі обговорення відповідних тем на заняттях — питання і відповіді на лекційних заняттях, питання і відповіді та участь у ревію на практичних заняттях, виконання вправ на практичних заняттях оцінюється, максимум, 10 балами, які може отримати кожен студент за семестр.

2) Роботи комп'ютерного практикуму

Максимальна кількість балів за усі виконані роботи комп'ютерного практикуму дорівнює 36 балів. Розподіл балів за виконання робіт практикуму:

| № з/п | Роботи комп'ютерного практикуму | Кількість балів |
|--------|--|-----------------|
| 1 | Лексичний аналіз програм | 6 |
| 2 | Синтаксичний аналіз програм | 6 |
| 3 | Семантичний аналіз програм | 6 |
| 4 | Трансляція у ПОЛІЗ та виконання ПОЛІЗ-програм | 6 |
| 5 | Трансляція для CLR/JVM | 6 |
| 6 | Побудова синтаксичного аналізатора засобами ANTLR4 | 6 |
| Разом: | | 36 |

Критерії оцінювання:

Виконання робіт комп'ютерного практикуму:

- правильне функціонування розробленого модуля є необхідною умовою зарахування виконаної роботи;
- виконана у повному обсязі і представлена робота – максимальна кількість балів (6= 3 + 3);
- виконана і представлена частково (не представлена) — відповідна частка від максимальної кількості балів.

3) Модульна контрольна робота

Максимальна кількість балів за модульну контрольну роботу дорівнює 10 балів.

Якість виконання роботи:

- виконана у повному обсязі з необхідними текстовими поясненнями дій та результатів – максимальна кількість балів;
- виконана частково з поясненнями — відповідна частка від максимальної кількості балів;
- виконана без текстових пояснень дій та результатів – не більше трьох балів;

Семестровий контроль: екзамен.

Максимальний ваговий бал $\Gamma_{\text{ісп}} = 20$

На усному екзамені студент обирає білет, який містить одне теоретичне питання та одне практичне завдання. Теоретичне питання оцінюється максимально до 8 балів, практичне — до 12 балів.

Умови допуску до семестрового контролю: виконані роботи 1-4 комп'ютерного практикуму та РГР, стартовий рейтинг (R_c) — не менше 40 балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

| Бали | Оцінка |
|---|--------------|
| 95 - 100 | Відмінно |
| 85 - 94 | Дуже добре |
| 75 - 84 | Добре |
| 65 - 74 | Задовільно |
| 60 - 64 | Достатньо |
| Менше 60 | Незадовільно |
| $R < 40$ або не виконана одна з робіт № 1-4 комп'ютерного практикуму, або не виконана РГР | Не допущено |

Розрахунок шкали (R) рейтингу:

Сума вагових балів контрольних заходів протягом семестру (шкала рейтингу) складає:

$$R = \Gamma_{\text{КП}} + \Gamma_{\text{РГР}} + \Gamma_{\text{МКР}} + \Gamma_{\text{активн}} + \Gamma_{\text{ісп}} = 36 + 19 + 15 + 10 + 20 = 100 \text{ балів.}$$

Максимальний стартовий рейтинг становить $R_c = \Gamma_{\text{КП}} + \Gamma_{\text{РГР}} + \Gamma_{\text{МКР}} + \Gamma_{\text{активн}} = 80$ балів.

Рейтинг іспиту дорівнює 20 балів. Мінімальний рейтинг допуску до іспиту становить 40 балів.

Таким чином, рейтингова шкала з кредитного модуля складає

$$R = 80 + 20 = 100 \text{ балів.}$$

Робочу програму навчальної дисципліни (силабус) «Основи розробки трансляторів»:

Склав доцент кафедри ІПЗЕ, к.т.н., доц. Стативка Юрій Іванович

Ухвалено кафедрою ІПЗЕ (протокол № 34 від 10.05.2024 р)

Погоджено Методичною комісією НН ІАТЕ ім.Ігоря Сікорського (протокол № 9 від 31.05.2024 р.)