

полиномов по соответствующему аргументу и соответствующей степени

$$V_{rl}(\omega, T_i) = \sum_{j=0}^m a_j \omega^j \sum_{p=0}^n b_p T_i^p.$$

Выводы. В итоге

$$Z(\Omega, \omega, T_i) = S(\Omega, \omega, T_i) \sum_{j=0}^m a_j \omega^j \sum_{p=0}^n b_p T_i^p$$

В функции $V_{rl}(\omega, T_i)$ присутствуют частоты Ω , как низкочастотные колебания поверхности, которую описывает $V_{rl}(\omega, T_i)$. Решение задачи определения параметров a_j, b_p решается методом наименьших квадратов в виде

$$\sigma^2 = [Z(\Omega, \omega, T_i) - S(\Omega, \omega, T_i) \sum_{j=0}^m a_j \omega^j \sum_{p=0}^n b_p T_i^p]^2$$

при этом задача решается методом покоординатного спуска, при котором параметры a_j, b_p меняются последовательно по правилу

$$a_i \pm \Delta a_i, \quad b_p \pm \Delta b_p$$

минимизируя σ_g^2 . Степени полиномов последовательно увеличиваются, уменьшая минимумы σ_g^2 , где g - номер итерации повышения степени.

Библиографические ссылки

1. Карпов О.Н. Компьютерные технологии распознавания речевых сигналов. Монография. / О.Н. Карпов, А.Г. Габович, Б.Г. Марченко, В.А. Хорошко, Л.Н. Щербак. –К., 2005. – 138 с.
2. Карпов О.Н. Описание спектрально-временного представления речевых сигналов в классе производных функций Гаусса второго порядка. / О.Н. Карпов, Г.В. Зирнеева // Питання прикладної математики та мат. моделювання. – 2004.–С.88–97
3. Карпов О.Н. Сравнение свойств колебательных функций в задаче анализа спектров речевых сигналов. / О.Н. Карпов, Г.В. Зирнеева // Актуальні проблеми автоматизації та інформаційних технологій. – Т.8.–2004. С. 14–19.

Надійшла до редколегії 20.08.10

УДК 534.4: 621.391

О.І. Лучинкіна, П.Ю. Аксьоненко, О.М. Карпов

Дніпропетровський національний університет імені Олеся Гончара

ТРАНСКРИПТОР ЯК ОДИН З ВУЗЛІВ ЗАГАЛЬНОЇ ЕКСПЕРТНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ МОВИ

Проводиться аналіз структури транскриптору та пропонується схема побудови транскриптору як експертної системи.

Ключові слова: експертна система, надійність, розпізнавання мови, транскриптор

Проводится анализ структуры транскриптора и предлагается схема построения транскриптора как экспертной системы.

Ключевые слова: экспертная система, надежность, распознавание речи, транскриптор

The analysis of the structure of transcriptor holds and the scheme of transcriptor as an expert system is proposed.

Keywords: expert system, reliability, speech recognition, transcriptor

Постановка проблеми. На сьогоднішній день, під поняттям «розпізнавання мови» приховується ціла сфера наукової та інженерної діяльності. Загалом, кожна задача розпізнавання мовлення зводиться до того, щоб виділити, класифікувати, і, відповідним чином, відреагувати на людську мову з вхідного звукового потоку. Це може бути і виконання певної дії на команду людини, і виділення певного слова-маркера з великого масиву телефонних переговорів, і системи для голосового вводу тексту.

Аналіз досліджень та постановка задачі. Згідно з опублікованими на сьогоднішній день даними надійність розпізнавання мови для систем, які працюють з однослівними командами, сягає 99,5 %, командами, що складаються з двох слів – 97,5 %, з трьох слів – 92,5 %, і чотирьох слів – 91,9 %. Але при проектуванні системи розпізнавання злитного мовлення перед розробниками постає проблема розпізнавання не лише мовного сигналу, але й лінгвістичного змісту. Тобто необхідно розробити методи підвищення надійності розпізнавання ланцюжка з великої кількості слів.

Системи розпізнавання мовлення класифікують за такими ознаками [1; 2; 3; 4]:

– тип речі (злитий або ізольований);

- тип елементів словника (фонеми, склади, слова);
- залежність від диктора;
- степінь деталізації;
- розмір словника (здатність обробляти велику кількість слів як загальної, так і спеціальної категорії з різних областей з метою збільшення точності та ефективності систем розпізнавання).

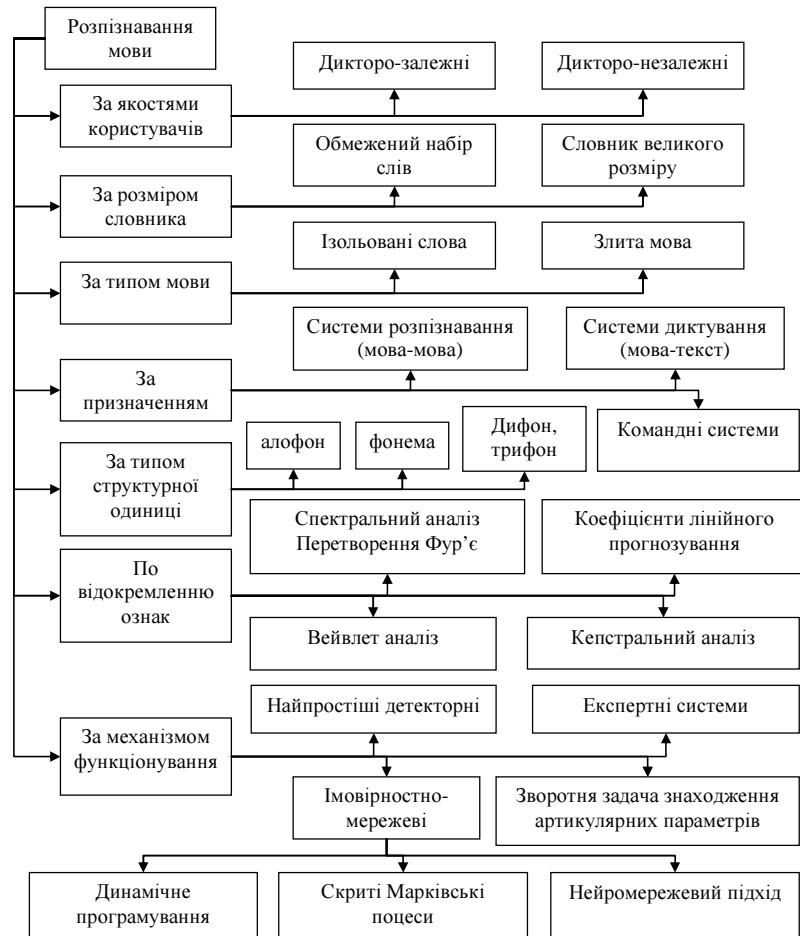


Рис 1. Класифікація систем розпізнавання мови

Кожна така система характеризується своїми методами. Розглянемо основні ознаки, за якими можна класифікувати системи

розпізнавання людської мови і те, як ця ознака може впливати на роботу системи. (рис. 1).

Класифікація систем розпізнавання мови за механізмом функціонування. У сучасних системах широко використовуються різні підходи до механізму функціонування розпізнавальних систем. Імовірно-мережевий підхід полягає в тому, що мовний сигнал розбивається на певні частини (кадри, або за фонетичною ознакою), після чого відбувається імовірнісна оцінка того, до якого саме елементу словника має відношення дана частина і (або) весь вхідний сигнал. Підхід, заснований на рішенні оберненої задачі синтезу звуку, полягає в тому, що з вхідного сигналу визначається характер руху артикуляторів мовного тракту і, за спеціальним словником відбувається визначення вимовлених фонем.

Як правило, розпізнавання мови поділяється на два етапи:

- розпізнавання голосових міток;
- розпізнавання лексичних елементів

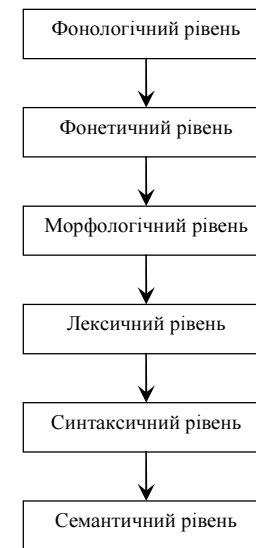


Рис. 2 Ієрархія лінгвістичних рівнів

Перший підхід припускає розпізнавання фрагментів мови за задалегідь написаним зразком. Цей підхід широко використовується у відносно простих системах, призначених для виконання задалегідь записаних речових команд.

Другий підхід складніше. При його реалізації з потоку мовлення виділяються окремі лексичні елементи – фонемі і алофони, які потім об'єднуються у склади і морфемі. Строго кажучи, саме цей підхід і використовується «справжніми» системами розпізнавання мови.

При розпізнаванні мови на рівні лексичних елементів доцільно розглянути ієрархію лінгвістичних рівнів (рис. 2)

Кожний з рівнів, що представлені на рис.2, характеризується своїми правилами, та законами з різних галузей наук – акустики, лінгвістики, фізики, математики, тощо. Система розпізнавання злитного мовлення потребує великого словника та текстової бази для навчання. Тому одним з важливих рівнів роботи системи є її навчання. Окремим етапом цього є прямий транскриптор, що дозволяє навчити систему.

Фонетична транскрипція. Фонетична транскрипція [5; 6; 7] – це особливий вид запису мови, який використовується для фіксації на листі особливостей її звучання.

Російська мова влаштована так, що одна літера на письмі може передавати два звуки (ёж) або, навпаки, дві літери – один звук (грузчик).

У транскрипції ж кожному звуку завжди відповідає свій особливий знак: [jóш], [гр'уш':ьк]. У письмовій формі після м'яких приголосних звуків замість а, о, у, э пишуться букви я, ё, ю та е, м'якість приголосного на кінці слова позначається спеціальною буквою ь, «м'який знак» (ять). У транскрипції м'якість приголосного завжди позначається однаково – знаком апострофа поруч з м'яким приголосним: [м'ат']. М'якість непарних м'яких приголосних [г'] та [ш':] також завжди позначається в транскрипції. Винятком є тільки позначення в транскрипційному запису палатальний (і тому за визначенням м'якого) приголосного [j] – при ньому не прийнято ставити апостроф. Графічно буква, підкоряючись правилам правопису (орфографії), часто не передає той звук, який вимовляється у слові (наприклад, ёж). Для транскрипції знака існує тільки одне правило – як можна точніше зафіксувати вимову звуку з урахуванням його відмінності від усіх інших звуків: [jóш].

На письмі існує спеціальне позначення для ненаголошених голосних звуків. У словах «карандаш», «багаж», «вблизи», наприклад, пишуться однакові літери для позначення голосних як ударних, так і ненаголошених, хоча звуки вимовляються по-різному: у ненаголошених складах слабше і коротше, а також у деяких випадках і зовсім по-іншому, ніж в ударному. У транскрипції необхідно підкреслити цю різницю або різними позначеннями і знаком наголосу

([каранд`аш]), або тільки знаком наголосу і відсутністю такого, так як ненаголошені [а] та [и] в двоскладових словах не зазнали якісних змін: [баг`аш], [вбл'из`и].

Якщо у письмений формі знак наголосу ставиться лише у спеціальних виданнях (словниках, підручниках для іноземців, дитячої літератури), то в транскрипції наголос відзначається обов'язково, коли у слові більше, ніж один, склад.

Звуки, що зазнали кількісну редуцію, можуть бути представлені тими ж знаками, що й ударні звуки, але без знаку наголосу, проте для позначення зредукованих голосних, які втратили внаслідок цього свою особливу якість, використовуються не тільки знаки, прийняті для позначення останніх гласних, а й деякі спеціальні знаки: [ь] (ер) і [ь] (ерь). Вони лише візуально збігаються з буквами, які в графіку взагалі не позначають звуків, але мають інші функції.

На письмі довгота тих приголосних звуків, які в російській мові бувають тільки довгими, позначається спеціальним чином: один із двох галасливих приголосних, які завжди є м'якими і довгими, позначається літерою щ (звук [ш':]), другий звук, [ж ':], не має спеціальної букви для свого позначення. Довгі звуки, що утворюються в російській мові при збігу двох однакових приголосних, позначаються двома однаковими літерами (касса); проте явище, при якому збіг двох різних приголосних дає один довгий звук (счет), ніяк не відбивається на листі.

У транскрипції тривалість приголосних звуків позначається одноманітно: як правило, двокрапкою праворуч від звуку ([ж':], [ш':], [до`ас:а]). У підручниках можна зустріти інше позначення довготи звуку: горизонтальну риску над відповідним транскрипційним знаком чи два однакових знака.

Нарешті, поняття «слово» у письмовому графіку і в транскрипції – це не одне і те ж. У письмовій графіку воно – самостійна або службова частина мови (прийменник в – теж слово і пишеться окремо), у транскрипції – це фонетичне слово, тобто деяке що складається з послідовності складів єдине ціле з одним організуючим центром – ударним складом. Таким чином, приводи, частинки, союзи, який виголошується злино з іншими словами, в транскрипції пишуться також разом і з позначенням усіх змін, які відбулися з звуками, їх складовими: в школу [фшк`олу], с ним [с'н`им], спросил бы [спрас`ил'п], за рекой [з`рик`ој], под гору [п`одг`ру].

Фонетична транскрипція у задачі розпізнавання мови. Ступінь подробиці транскрипційного запису розмови залежить від того, що вона повинна зафіксувати: всі особливості звучання, включаючи

індивідуальні і зумовлені ситуацією породження мовлення (емоційні і до.), або особливості фонетичної системи мови.

Текст на природній мові являє собою впорядкований потік символів. Символи обробляються послідовно, один за одним, у порядку їх розташування в тексті.

На первинних рівнях аналізу перед нами стоїть завдання розбиття висловлювання на елементи первинного аналізу. В якості елемента первинного аналізу будемо брати фонему. Як фонетичний алфавіт будемо використовувати набір з 48 фонем: 12 – для голосних звуків (враховуючи, що кожна голосна може бути ударною і ненаголошеною) і 36 – для приголосних (враховуючи, що кожна приголосна може бути твердою і м'якою).

Таким чином, отримуємо алфавіт:

- Голосні: а а! е е! о! и и! у у! ы ы! э!.
- Приголосні: б, б' в в' г г' д д' ж з з' й к к' л л' м м' н н' п п' р р' с с' т т' ф ф' х х' ц ч ш щ.

Проектування та розробка транскриптору як вузла експертної системи. При створенні фонетичного транскриптору необхідно брати до уваги логіку транскрипції. Проблема полягає в тому, що неможливо поставити у відповідність кожній літері конкретний її звук. Так, наприклад, транскрипція слова «ёж» – [jош], а слова «кожа» – [кожа].

Звук літери міняє своє значення залежно від розташування у слові, наголосу, порядку слідування, властивостей літер, тощо.

Писати програмний продукт який буде перевіряти десятки однотипних перевірок для кожної літери при транскрибуванні не є доцільним. Тому проаналізувавши сучасні методи вирішення подібних проблем [8; 9] можна дійти висновку, що доцільно будувати таку систему як система штучного інтелекту.

Мета проектування експертних систем полягає в розробці програмних продуктів, що можуть вирішувати задачі, які є важкими для розв'язання експертами у зв'язку з необхідністю обробки великої кількості даних, або застосування великої кількості правил. За допомогою експертних систем ми отримуємо результати, які не поступаються за якістю і ефективності рішенням, до яких можуть дійти експерти. У більшості випадків експертні системи вирішують задачі, які важко формалізувати, або задачі, що не мають алгоритмічного рішення.

Розрізняють два види експертних систем:

1. Статичні. У таких системах обробка інформації йде на основі постійних і незмінних даних і правил з бази. Знання можуть додаватися до системи, але не у процесі функціонування системи.

2. Динамічні. Це системи орієнтовані на роботу з постійно мінливою інформацією, яка міститься в інформаційній базі і базі знань системи (компонентом бази знань є база даних).

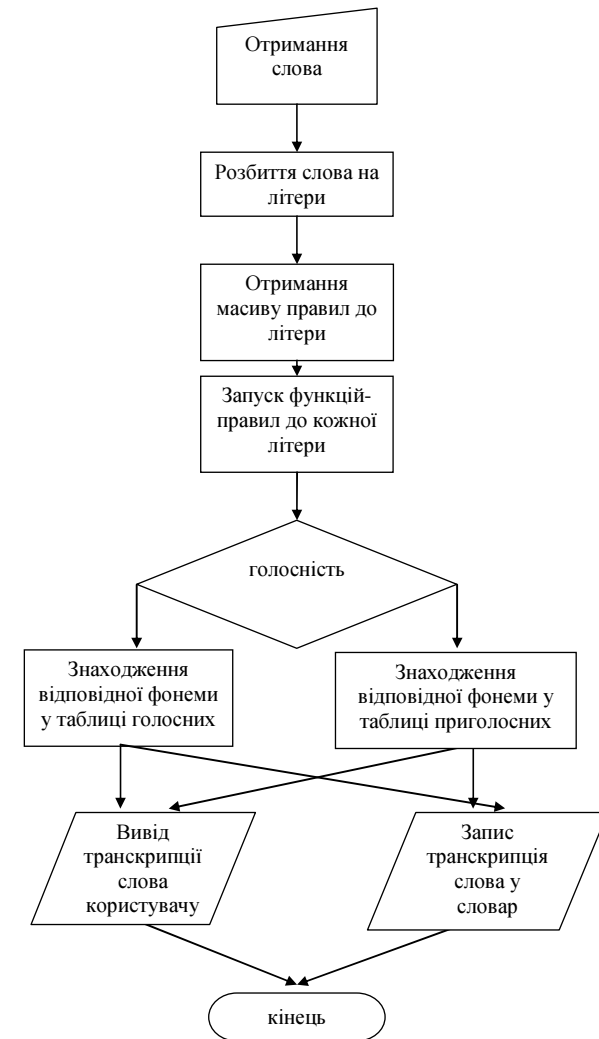


Рис 3. Блок-схема роботи експертної системи прямого фонетичного транскриптору

При розробці транскриптору доцільно використовувати статичну експертну систему з чітко заданими правилами та заздалегідь відомими елементами.

Для того щоб вирішити які саме правила потрібно використовувати для транскрибування потрібно:

1. Виділити елементи системи – звуки.
2. Виділити правила транскрипції.
3. Розставити пріоритет для кожного з правил.
4. Записати правила у формалізованому вигляді.

Роботу експертної системи прямого фонетичного транскриптору описує блок-схема на рис.3.

Пряма транскрибування вхідного виразу будується за наступними етапами (рис. 3):

- На вході система (діалоговий компонент) отримує слово яке необхідно транскрибувати.
- Розбиття вхідного слова на літери.
- Експертна система знаходить масив правил , які слід використовувати до кожної конкретно літери.
- Перевірка виконання відповідних правил до кожної літери.
- Знаходження у відповідній таблиці (голосні, приголосні) фонему яка відповідає відповідній літері за відповідним правилом.
- Видання результату користувачу.
- Запис результату у словник фонетичних транскрипцій.

Розглянемо конкретні приклади:

Слово «мороженное» – м|о|р|о`|ж|е|н|н|о|е

До кожної літери використовується відповідне правило за пріоритетом. При виконанні умов правила N N+1 правило не буде використане. Наголошеність – обов'язкова умова для транскрибування.

Після аналізу правил для кожної літери отримується поле, в якому , у відповідній таблиці для голосних і для приголосних, знаходиться фонема:

- М – проста відповідність;
- О – не наголошена;
- Р – проста відповідність;
- О` – наголошена;
- Ж– проста відповідність;
- Е – не наголошена, після приголосного;
- Н – проста відповідність;

- Н – подвоєння;
- О – не наголошена, після приголосного;
- Е – не наголошена, після приголосного.

Експертна система видає наступну відповідність:

- «М» – «М»
- «О» – «А»
- «Р» – «Р»
- «О`» – «О»
- «Ж» – «Ж:»
- «Е» – «'И»
- «Н» – «Н»
- «Н» – «:»
- «О» – «А»
- «Е» – «ЙЭ»

Наведемо приклад фрейму, який описує правила, за якими будується транскрипція голосних наголошених звуків.

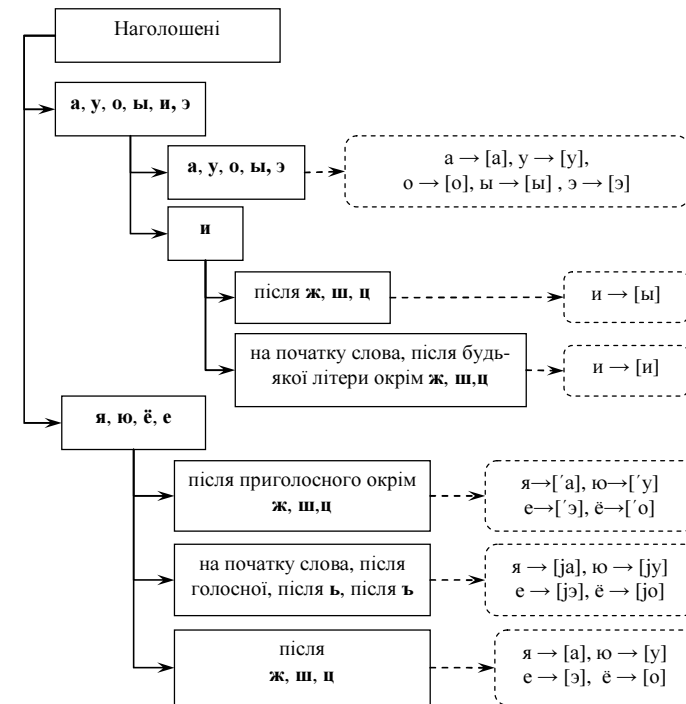


Рис 3. Граф-схема фрейму опису транскрипції ненаголошених голосних

Таким чином, у системі відокремлюються літери за деякою істотною для транскрипції ознакою. Наведемо декілька прикладів:

- приголосні ж, ш та ц відокремлюються до окремої групи твердих непарних приголосних;
- ь та ъ – літери, за допомогою яких позначається твердість та м'якість приголосних;
- я, ю, ё, э, а, у, о, ы, и – голосні;
- я, ю, ё, е – голосні, які можуть бути представлені двома звуками;
- э, а, у, о, ы, и – голосні, які представляються одним звуком;
- и – голосна, яка має властивість пом'якшувати приголосні.

Подальший аналіз допомагає нам виділити серед голосних шиплячі, тверді, дзвінки, парні, непарні, тверді, м'які.

Окрім перетворення окремих літер при побудові транскриптору ми стикаємося із задачею позиційного спрощення приголосних, асиміляції приголосних, тощо. Тобто, виникає необхідність не просто аналізу окремих літер, а аналізу деяких їх сполучень.

На рис.3 представлена граф-схема одного з підрівнів системи, на якому вирішується задача прийняття рішення відносно наголошених голосних. Подібною схемою описується вся система транскрипції. Її вузлами можуть бути не лише об'єкти та їх властивості, але і функції.

Висновки. Складність системи, велика кількість даних, які необхідно обробити, та необхідність прийняття експертного рішення підводить нас до розробки транскриптору як експертної системи.

Первісною задачею при побудові системи є добір суттєвих ознак об'єктів системи, виділення їх в окремі групи та побудова правил їх взаємодії. Подібна система є складною, тобто деякі її вузли являють собою процедури, що вирішують деякі під задачі перевірки та аналізу правил.

Бібліографічні посилання

1. **Карпов О.М.** Вступ до курсу «Технологія проектування пристроїв розпізнавання промови» / О.М. Карпов // Навч. посібн. – Д., 2000. – 64 с.
2. **Карпов О.Н.** Компьютерные технологии распознавания речевых сигналов. Монография. / О.Н. Карпов, А.Г. Габович, Б.Г. Марченко, В.А. Хорошко, Л.Н. Щербак // – К., 2005. – 138 с.
3. **Карпов О.М.** Методи та алгоритми оцінки ситуативних відхилень параметрів мови людини. Навч. пос. / О.М. Карпов, Г.В. Зирнєєва, А.О. Чугай, В.О. Асадулін. – Д., 2007 – 64 с.

4. **Карпов О.Н.** Технология построения устройств распознавания речи. Монография. / О.Н. Карпов – Д., 2001. – 184 с.

5. http://ru.wikipedia.org/wiki/Фонетическая_транскрипция

6. **Буланин Л.Л.** Фонетика современного русского языка. / ЛЛ. Буланин. – М., 1970

7. **Шанский Н.М.** Современный русский язык. В 3-х частях. Ч.1. / Н.М. Шанский, В.В. Иванов. – М., 1987.

8. **Люггер Джордж Ф.** Искусственный интеллект: стратегии и методы решения сложных проблем. Четвертое издание. / Ф. Джордж Люггер – М., 2005 – 864 с.

9. **Рассел Стюарт.** Искусственный интеллект: современный подход, 2-е издание. / Стюарт Рассел, Питер Норвиг – Вильямс, 2007 – 1408 с.

Надійшла до редколегії 12.10.10