

Ю.М. Щербина*, Т.В. Шестакевич, В.А. Висоцька
 *Львівський національний університет ім. І. Франка,
 кафедра дискретного аналізу та інтелектуальних систем,
 Національний університет "Львівська політехніка",
 кафедра інформаційних систем та мереж

НАУКОВИЙ НАПРЯМ ТА НАВЧАЛЬНА ДИСЦИПЛІНА "МАТЕМАТИЧНА ЛІНГВІСТИКА"

© Щербина Ю.М., Шестакевич Т.В., Висоцька В.А., 2010

Розглянуто передумови виникнення та основні положення математичної лінгвістики, проаналізовано різні погляди на її предмет. Запропоновано програму відповідного навчального курсу та розкрито її зміст.

Ключові слова: математична лінгвістика, квантитативна лінгвістика, комбінаторна лінгвістика.

This article deals with the preconditions of the appearance and main principles of the mathematical linguistics and different points of view on its subject were analyzed. The program of corresponding educational course was proposed. The context of the program has been revealed.

Key words: mathematical linguistics, computational linguistics, combinational linguistics.

Вступ

Розвиток обчислювальної техніки спростив та прискорив дослідження практичного застосування накопичених лінгвістичних знань, що стало поштовхом до розвитку такої галузі мовознавства, як прикладна лінгвістика. Об'єктом прикладної лінгвістики є розроблення методів та засобів застосування у соціальній практиці людини результатів фундаментальних лінгвістичних досліджень, спрямованих на пізнання онтологічних властивостей мови [10]. Для виконання завдань прикладної лінгвістики застосовуються методи комп'ютерної (машинної, обчислювальної, інженерної) лінгвістики, математичної лінгвістики, лінгводидактики, термінознавства, перекладознавства, контрастивної лінгвістики, структурної лінгвістики тощо.

Комп'ютерна лінгвістика (машинна, обчислювальна, інженерна лінгвістика) займається застосуванням комп'ютера (технологій та програм опрацювання даних) для моделювання функціонування мови в певних умовах та виконання лінгвістичних завдань, а також розробляє лінгвістичні аспекти комп'ютеризації. У широкому розумінні до комп'ютерної лінгвістики зараховують усе, що пов'язане з використанням комп'ютерів у мовознавстві.

Математична лінгвістика (виникла в 50-ті роки ХХ ст. [11]) – галузь науки на межі мовознавства та математики, яка вивчає можливості застосування математичних методів для опису та дослідження природних і деяких штучних мов, для пояснення лінгвістичних подій. Підґрунтям її появи стала необхідність в уточненні основних лінгвістичних понять, потреба в уведенні точніших та об'єктивніших методів для аналізу та синтезу мови та тексту, поява міжпредметних зв'язків з іншими галузями, що вимагають спілкування мовою математики. З розвитком можливостей комп'ютерних технологій також виникла потреба, зокрема, у машинному перекладі та автоматизованому інформаційному пошуку. Саме **математичну лінгвістику** вважають теоретичним підґрунтям прикладної лінгвістики. Підкреслюючи спільність поняттєвого апарату, математичну лінгвістику іноді розділяють на галузі мовознавства та математики, а також зазначають, що в частині **використання** розроблених математичних моделей для опису будови природних мов математична лінгвістика належить до такої галузі досліджень, як *штучний інтелект*.

Постановка проблеми

Застосування математичних методів у лінгвістиці зумовлене двома причинами.

По-перше, розвиток теорії і практики мовознавства вимагає точніших і об'єктивніших методів для аналізу мови і тексту. Одночасне використання математичних прийомів у систематизації, вимірюванні та інтерпретації лінгвістичного матеріалу у поєднанні з якісною інтерпретацією результатів дає змогу мовознавцям глибше проникнути у таємниці побудови мови і утворення тексту [17].

По-друге, контакти мовознавства з іншими науками (акустикою, фізіологією вищої нервової діяльності, кібернетикою та обчислювальною технікою) постійно розширюються, і можуть розвиватись лише з використанням математичної мови, яка має високий ступінь загальності та універсальності для різних гілок знань [11].

У зв'язку із поширенням інтелектуальних інформаційних систем опрацювання інформації, систем перекладу з однієї природної мови на іншу, систем керування людина-машина-людина зростає необхідність підготовки інженерних кадрів, які володіють точними методами у мовознавстві та можуть кваліфіковано і свідомо використовувати увесь спектр сучасного програмного забезпечення у цій галузі.

Однією із перших в Україні фахівців з прикладної лінгвістики почала готувати кафедра прикладної лінгвістики Національного університету "Львівська політехніка".

Аналіз сучасних досліджень і публікацій

Лінгвістичні задачі, інтуїтивно сформульовані чи такі, що не мають повного розв'язку, із застосуванням математичних методів доцільно замінити однією або декількома простішими, логічно сформульованими математичними задачами, які мають *алгоритмічний* розв'язок. Таке розділення складної лінгвістичної проблеми на простіші задачі, які алгоритмізуються, називатимемо *математичною експлікацією* лінгвістичного об'єкта або явища [18].

Такий підхід до розв'язування завдань мовознавства дає змогу математичній лінгвістиці брати участь у розв'язанні таких прикладних задач [1, 10, 17]:

- розроблення формальних моделей природних та штучних мов;
- вирішення питань практичної транскрипції та транслітерації;
- дешифрування невідомих писемностей;
- усний та письмовий переклад, розроблення систем автоматичного послівного і пооборотного машинного перекладу, семантичний переклад тексту;
- авторська та часова атрибуція твору;
- створення систем стенографії, систем письма для сліпих;
- завдання судової та кримінальної лінгвістики;
- завдання лінгводидактики:
 - навчання рідній та іноземній мові;
 - розроблення відповідних лінгвістичних методик (навчання дітей та дорослих, навчання емігрантів тощо);
- розроблення раціональної та стабільної орфографії;
- автоматичне розпізнавання та синтез мови;
- розроблення автоматизованих систем опрацювання текстової інформації:
 - автоматичне зняття полісемії;
 - автоматичне анотування;
 - індексування та реферування;
 - синтаксичний аналіз;
 - компресія текстів;
 - інтелектуальний аналіз тексту з метою виділення неявного змісту ("читання між рядками");
- розроблення лінгвістичного забезпечення автоматизованих систем управління;

- комп'ютеризація навчання та автоматизація викладання за допомогою комп'ютера;
 - створення автоматизованих системи навчання;
 - створення системи дистанційного навчання;
 - створення автоматизованого робочого місце перекладача та викладача;
- розв'язання задач прикладного термінознавства;
 - систематизація та автоматизація терміносистем певних галузей;
 - укладання термінологічних словників;
 - стандартизація й уніфікація науково-технічної термінології;
- створення автоматизованих редакційно-видавничих систем;
- створення машинних фондів мов;
- вивчення процесів та створення правил утворення нових назв товарів, виробів тощо.

Також математична лінгвістика задіяна у вирішенні (за допомогою взаємодії відповідних математичних методів) таких специфічних завдань мовознавства та прикладної лінгвістики, зокрема [1, 11, 18]:

- побудовані за допомогою обернених тригонометричних функцій моделі можна використовувати для опису структурних зсувів (наприклад, поява або зникнення лінгвістичних одиниць) у галузі лексикології, морфології, фонології, синтаксису та стилістики;
- у роботах з діахронічної лінгвістики часто наводяться таблиці, що містять кількісні дані про використання досліджуваного явища на різних етапах його історії. Такі таблиці показують динаміку лінгвістичного процесу, поглибленіше аналізувати який можна за допомогою елементарних функцій, графічно, а також з використанням аналітичних моделей досліджуваного процесу;
- класична глотохронологія має на меті наближено датувати процес розходження діалектів та споріднених мов, а також кількісно оцінити ступінь їх спорідненості, і використовує для цього статистичні методи;
- методи математичної статистики також допомагають у створенні словників (зокрема частотних та статистичних), автоматичних словників, тезаурусів, у вирішенні завдань стилеметрії та інформаційного пошуку;
- постійне розширення сфери діяльності сприяє неухильному збільшенню термінологічного словника людини, і для лексикографічної практики, а особливо під час побудови систем машинного перекладу та реферування, важливо мати прогноз кількісного росту термінології у різних галузях знань. Таке прогнозування дало б можливість свідомо планувати випуск та ефективно поповнювати електронні політехнічні та галузеві словники та довідники;
- для визначення швидкості зміни в мові та мовленні, а також їх кількісних оцінок, використовується поняття похідної;
- тригонометричні функції та поняття границі дають змогу моделювати циклічний і стрибкоподібний характер лінгвістичних процесів;
- можна також описати процеси приросту та накопичення нових мовних елементів, використовуючи теорію рядів та інтегрування;
- приблизні оцінки вимірів обмежень, що накладаються на вживання лінгвістичних одиниць системою та нормою мови, можна здійснювати за допомогою комбінаторики;
- лінгвістичні дослідження в частині підготовки систем машинного перекладу та інформаційного пошуку вимагають, зокрема, знання об'єму вибірки, необхідного для забезпечення із заданою ймовірністю появи певної лінгвістичної події хоча б раз. Знаходження достатнього об'єму вибірки необхідне для визначення достовірності отриманих висновків.

Формулювання цілей статті

Математична лінгвістика бере участь у вирішенні завдань, які є актуальними не лише для мовознавства, тому уміння їх розв'язувати є вимогою часу для лінгвістів. Необхідним є володіння такими розділами математики, як комбінаторика [9], теорія ймовірностей [4], математична

статистика [4, 11], теорія інформації [16], теорія формальних породжувальних граматики [2, 3, 9], теорія автоматів [9], мережі Петрі [12]. Опанований математичний апарат студенти зможуть застосовувати для набуття навичок у розв'язуванні таких задач мовознавства, як прогнозування появи в уривку певної кількості складових тексту з відповідних класів, визначення об'єму даних та кількості інформації, визначення надлишковості тексту, виконання авторської, стилевої чи навіть мовної атрибуції уривку, застосування лінгвістичних схем у криптолінгвістиці, побудова компіляторів на лексичному та синтаксичному етапах, дослідження логічних правил побудови зв'язного тексту тощо. Викладені міркування стали основою програми освітньої дисципліни “Математична лінгвістика” для студентів, що здобувають фах “Прикладна лінгвістика”.

Аналіз отриманих наукових результатів

Структуру курсу “Математична лінгвістика” можна подати у вигляді схеми, зображеної на рис. 1.



Рис. 1. Схематична структура курсу “Математична лінгвістика”

У викладенні матеріалу до курсу відображено поділ математичної лінгвістики на два напрями: квантитативну та комбінаторну лінгвістику. Кожна з двох частин курсу містить тематичний опис математичних методів, необхідних для ґрунтовно висвітлення фундаментальних засад математичної лінгвістики. Для закріплення знань з теми запропоновані контрольні запитання та тестові завдання, а також тридцять варіантів завдань трьох рівнів складності для самостійної роботи. Методологію вирішення завдань мовознавства за допомогою відповідних математичних методів викладено у покроковому аналітичному розв'язуванні відповідного завдання, організувати його у вигляді інформаційної системи студентам пропонується самостійно. Усі практичні завдання організовано так, щоб студенти могли також застосувати знання з різноманітних розділів мовознавства.

Квантитативна лінгвістика

Перша частина курсу “Математична лінгвістика” – це квантитативна лінгвістика (обчислювальна лінгвістика), яка вивчає та описує лінгвістичні явища за допомогою методів “кількісної” математики (теорія ймовірностей, математична статистика тощо). Схема зв'язків між методами розділів математики та розділами квантитативної лінгвістики курсу “Математична лінгвістика” наведена на рис. 2.

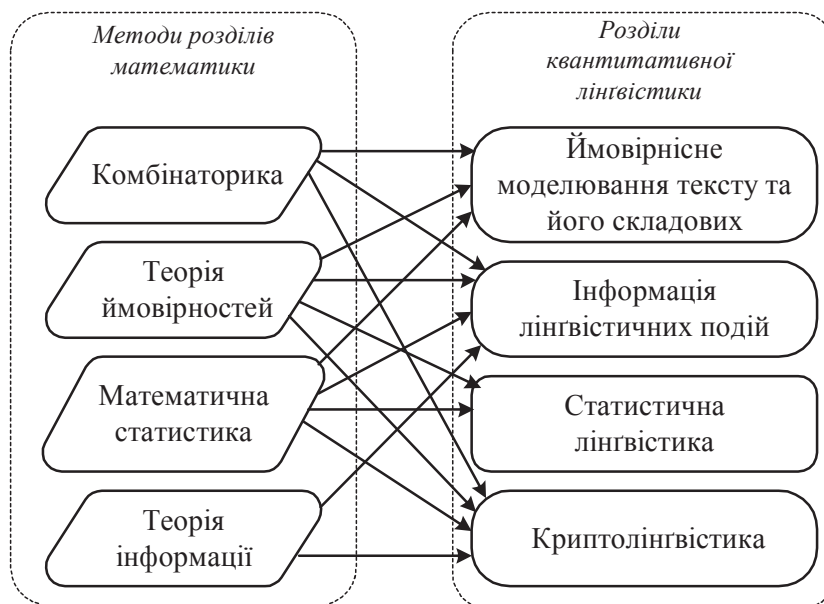


Рис. 2. Схема зв'язків між методами математики та розділами квантитативної лінгвістики

У розділі "Ймовірнісне моделювання тексту та його складових" досліджуються механізми породження тексту, причому вважається, що результати окремого лінгвістичного випробування (наприклад, вибір з тексту окремих його одиниць) не мають важливого значення, через те необхідно організувати багаторазове виконання серії експериментів, коди одна і та сама лінгвістична подія повторюється багато разів. Від того, як побудоване ймовірнісно-лінгвістичне випробування та організовано вибір з тексту окремих його одиниць, залежить вибір тієї чи іншої моделі опису тексту.

У лінгвістичних дослідженнях і особливо під час реалізації алгоритмів машинного послідовного перекладу та інформаційного пошуку постійно виникають завдання, пов'язані з прогнозуванням появи в сегменті заданої довжини певної кількості словоформ чи словосполучень, що належать певним класам. Ймовірнісне моделювання тексту та складів, словосполучень, граматичних класів тощо дає змогу вирішити це завдання та визначати об'єм вибірки, необхідної для забезпечення із заданою ймовірністю появи хоча б один раз відповідної лінгвістичної одиниці.

Вивчення функціонування мови та мовлення за допомогою ймовірнісного моделювання тексту спирається на моделі теорії ймовірності та комбінаторики.

Для ілюстрації процесу ймовірнісного моделювання тексту розв'язують типове завдання, яке для опрацювання тексту вимагає також знань. Наприклад, для кількох стилів української мови необхідно знайти відносну частоту появи лінгвістичної одиниці (наприклад, іменника). Використовуючи отримані дані, обчислити ймовірність того, що серед навмання взятих кількох слів буде відповідна кількість (0, 1, 2, ...) іменників.

Розділ "Інформація лінгвістичних подій" аналізує формально-структурні характеристики інформації: об'єм даних, кількість інформації та коефіцієнт інформативності. Поняття ентропії як міри невизначеності лінгвістичного дослідження дає змогу вимірювати кількість інформації як зміну (зменшення) невизначеності стану системи, і, відповідно, приріст інформації. Для визначення кількості інформації застосовують комбінаторний та ймовірнісний підходи.

Комбінаторні оцінки інформації використовують для аналізу "гнучкості мови", тобто дають змогу вимірювати розгалуженість продовження тексту для заданого словника та заданих правил побудови речень. Також обчислення кількості інформації дає змогу приблизно оцінити виміри обмежень, що накладаються на використання лінгвістичних одиниць системою та нормою мови [11].

Зацікавлять студентів матеріали, що стосуються інформаційних вимірів кодування інформації в частині інформаційних втрат у каналах із шумом, розрахунку швидкості передавання інформації та пропускну здатності інформаційних каналів зв'язку.

Обчислення формально-структурних характеристик інформації вимагає знань комбінаторики, засад теорії ймовірності та основ теорії інформації.

Процес обчислення інформації лінгвістичних подій демонструється прикладами, у яких необхідно визначити триграму (*Триграма*), з котрої починається прізвище студента, транслітероване англійською мовою. За допомогою словника обчислюють відносні частоти появи голосних букв після ланцюжка *ПробілТриграма* в англійській мові. Необхідно визначити кількість інформації для кожного з ланцюжків *ПробілТриграмаГолосна* та її середнє значення; обчислити кількість інформації, що одержується з досліду з п'ятьма рівноймовірними результатами та порівняти отримане значення зі значенням середньої кількості інформації.

У розділі "Статистична лінгвістика" розглядаються завдання визначення статистичної структури тексту із використанням методів математичної статистики та теорії ймовірностей. Розділ охоплює три підрозділи прикладної лінгвістики та суміжних із нею наук: лінгвометрії, стилеметрії та глотохронології.

Лінгвометрія – галузь прикладної лінгвістики, що виявляє, вимірює та аналізує кількісні характеристики одиниць різних рівнів мови чи мовлення [1]. Застосовуючи апарат математичної статистики, лінгвометрія бере участь у вирішенні таких завдань мовознавства, як створення та порівняння словників (зокрема частотних та статистичних), створення автоматичних словників, тезаурусів, створення систем стенографії, автоматичне визначення мови, інформаційний пошук тощо.

Завдання автоматичного визначення мови можна сформулювати так: відформатовані фрагменти текстів, де літери розташовані за зниженням частот їх появи в уривку, необхідно проаналізувати, наприклад, за допомогою дослідження частотності вживання літер та визначити мову фрагментів.

Інший тип завдань, що належать до лінгвометрії, полягає в оцінці фрагмента тексту за допомогою коефіцієнтів різноманітності та зв'язності мовлення. Необхідно прочитати та письмово переказати уривок тексту (уривки пропонує викладач, завдання доцільно розв'язувати під час заняття). Знайти коефіцієнти зв'язності мовлення, лексичної різноманітності, синтаксичної складності, індексів концентрації та винятковості для уривку та для переказу та порівняти результати. Наступним етапом аналізу фрагментів тексту є дослідження його внутрішньої "динаміки".

Типологія, діагностика, атрибуція, класифікація, реконструкція, датування тощо елементів функціонального чи індивідуального стилю – ось неповний перелік завдань, які розв'язує стилеметрія [1, 7].

За допомогою наукового апарату атрибуції текстів стилеметрія бере участь у розв'язанні трьох основних груп практичних завдань [1, 8]: дослідження літератури й історії (наприклад, "Шекспірівське питання", авторство "Тихого Дону" тощо), завдання у сфері освіти та психології (проблема плагіату особливо загострилась з розвитком глобальної мережі Інтернет [6]), а також судова практика (об'єктами досліджень можуть бути письмові свідчення очевидців чи заповіти, анонімні листи тощо). Перспективою у розвідках авторської атрибуції текстів є, наприклад, дослідження збереження авторського стилю у перекладах текстів [15].

Розв'язування завдання з авторської атрибуції текстів дають змогу оцінити, для якого із досліджуваних відформатованих уривків (слововживання розташовані за зниженням частот їх появи в уривку, власні назви видалені з тексту) імовірність належати до еталону є найбільшою.

Глотохронологія (з *gp. glotta* – мова, *chronos* – час і *logos* – слово, вчення; хронологія слів) – це підрозділ порівняльно-історичного та статистичного мовознавства, предметом якого є виявлення швидкості мовних змін та визначення на основі цього часу розділення (дивергенції) споріднених мов та ступеня близькості між ними [1, 5, 14]. Для досліджень використовують, зокрема, методи математичної статистики.

Пізнавальним завданням для студентів стане аналіз літературних надбань американських та англійських письменників, що творили упродовж певного відрізка часу. Дослідження полягатиме у визначенні тенденції до збільшення (зменшення) кількості спільних слів ядерної лексики в роботах англійських та американських авторів.

Дослідження характеристик мов (частоти появи певних літер, комбінацій літер, загальні шаблони тощо) для застосування отриманих знань у криптології аналізуються в розділі "Криптолінгвістика".

Лінгвістичні схеми в розв'язуванні завдань криптології передбачають використання методів математичної статистики, теорії ймовірностей, комбінаторики. Скінченні автомати, що розглядаються в комбінаторній частині математичної лінгвістики, є дуже зручним наочним інструментом ілюстрації процесу кодування та декодування.

Для моделювання процесу дешифрування як приклад пропонується порівняти частоти букв в українській, англійській, російській мовах та у зашифрованому тексті. Для цього необхідно ознайомитись з частотністю літер відповідних мов та обчислити частоти появи символів та їх сполучень у зашифрованому тексті. Спираючись на отримані дані, потрібно визначити мову тексту та розшифрувати його.

Комбінаторна лінгвістика

Другою частиною математичної лінгвістики є комбінаторна лінгвістика (рис. 3), яка вивчає та описує лінгвістичні явища за допомогою методів "не кількісної" математики (теорія формальних граматики, мережі Петрі тощо) [11, 18].

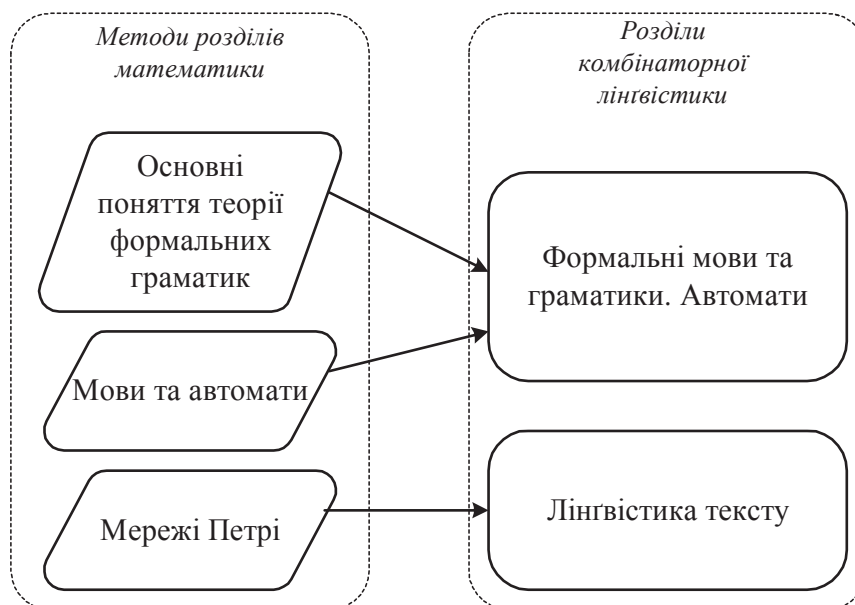


Рис. 3. Схема зв'язків між темою та розділами комбінаторної лінгвістики в курсі "Математична лінгвістика"

У розділі "Формальні мови та граматики. Автомати" розглядається процес моделювання за допомогою формальних граматики та автоматів лексичного етапу роботи компілятора, коли під час аналізу вхідного тексту у ньому виділяються найпростіші конструкції мови – ідентифікатори, символи, числа, символи операцій, службові слова тощо. Також розглядається другий (синтаксичний) етап роботи компілятора, де тіло синтаксичного аналізатора – це діаграма переходів відповідного детермінованого скінченного автомата. Для задання синтаксису алгоритмічної мови використовуються формальні граматики, які Н. Хомський ввів як формалізацію одного із розділів лінгвістики – синтаксису.

Основою для методів формалізації мови є теорія формальних граматики. Скінченні автомати розглядатимемо як абстрактні моделі найпростіших пристроїв обробки даних.

Комплексне завдання для закріплення знань, отриманих у цьому розділі, полягає у реалізації лексичного аналізатора мови програмування. Теоретична частина організовується у формі скінченного автомата, який розрізняє такі класи лексем: числа, рядкові та символні константи, коментарі, зарезервовані слова, оператори, розділові знаки, ідентифікатори.

Простежити зв'язок між мовами, граматиками та автоматами допомагають такі задачі. Задано тип ідентифікатора – це ланцюжок довжини 5, який може складатись із цифр, великих або малих букв українського алфавіту. Побудуйте скінченний автомат з виходом, що розпізнає ланцюжок заданого типу, і лише його, з певними обмеженнями (наприклад, символи з парними номерами –

цифри). Необхідно побудувати граматику, що породжує такий і лише такий ланцюжок, знайти можливу кількість комбінацій елементів ідентифікатора такого типу.

У розділі "Лінгвістика тексту" вивчається структура, категорії та складові тексту як автономної конструкції, розглядаються правила побудови зв'язного тексту. Реалізація таких завдань передбачає застосування методів лінгвістичного аналізу для попереднього розбору багатьох текстів та порівняння результатів. Об'єктом лінгвістичного аналізу тексту може бути, наприклад, науковий, офіційно-діловий, публіцистичний чи художній тексти.

Встановлені під час лінгвістичного аналізу логічні правила побудови зв'язного тексту можна використати в розв'язуванні задач опрацювання та синтезу природної мови, яка є важливою проблемою у створенні систем штучного інтелекту. Для моделювання асинхронних інформаційних потоків в системах обробки даних використовують модифікацію мережевої моделі – розфарбовану мережу Петрі. Введення поняття кольору дає змогу ввести нове правило взаємодії між складовими мережі (дуга, розфарбована деяким кольором, може пропустити тільки фішки відповідного кольору), коли із кожним кольором пов'язується конкретна семантична інтерпретація.

Застосування лінгвістичного аналізу дає змогу виділити правила побудови зв'язного тексту на основі аналізу текстів різного стилю. У машинній побудові зв'язних текстів, а також у реалізації машинного синтезу мови застосовують мережі Петрі, пов'язані із моделюванням причинно-наслідкових зв'язків.

Розфарбовані мережі Петрі отримують з класичних за допомогою певної модифікації.

Для моделювання етапів організації зв'язного тексту необхідно організувати розфарбовану мережу Петрі як апарат побудови тексту казки [13]. На підставі кількох відомих казок необхідно сформулювати набір параметрів мережі: основні типи персонажів та базові ситуації. Продемонструвати роботу розфарбованої мережі Петрі, одержані кілька варіантів казки, що відображають один сюжет.

Основні методи та розділи математичної лінгвістики викладені у відповідній освітній дисципліні так, щоб врахувати тісний зв'язок між кількісними та якісними характеристиками мови та підкреслити необхідність спільного застосування математичного апарату квантитативної та комбінаторної лінгвістики.

Висновки

Запропонована схема викладення матеріалу ґрунтується на практичному досвіді. Апробацію пройшли методичні вказівки до виконання лабораторних робіт з курсу "Математична лінгвістика" для студентів базового напрямку "Філологія" (2006–2007) [19, 20], укладені авторами статті. Описана структура курсу "Математична лінгвістика" дає змогу охопити та проілюструвати розділи математичної лінгвістики, що необхідні для здобуття фаху в галузі прикладної лінгвістики. Запропонований додатковий матеріал для поглибленого вивчення розділів допоможе скерувати зацікавленого студента у власних наукових дослідженнях та практичних розробках з використанням сучасних інформаційних технологій. Курс "Математична лінгвістика" призначений для студентів, що навчаються за напрямами галузей знань "Гуманітарні науки" (напрямок 6.020303 "Філологія") і "Системні науки та кібернетика" (напрямок 6.040303 "Системний аналіз").

1. Бук С. Н. *Основи статистичної лінгвістики: Навч.-метод. посіб.* / Відп. ред. проф. Ф.С. Бацевич / С.Н. Бук. – Л. : Видавничий центр ЛНУ ім. І. Франка, 2008. – 124 с. 2. Гладкий А.В. *Лекции по математической лингвистике* / А.В. Гладкий. – Новосибирский госуниверситет, 1966. 3. Гладкий А.В. *Формальные грамматики и языки* / А.В. Гладкий. – Наука, 1973. – 368 с. 4. Жлуктенко В. І. *Теорія ймовірностей і математична статистика: Навч.-метод. посіб.: У 2-х ч.* — Ч. II. *Математична статистика* / В. І. Жлуктенко, С. І. Наконечний, С. С. Савіна. — К.: КНЕУ, 2001. — 336 с. 5. Иванов В.В. *Глоттохронология // Лингвистический энциклопедический словарь* / Гл. ред. В. Н. Ярцева. — М. : Сов. энциклопедия, 1990. — С. 109-110. 6. Ланде Д. В. *Підхід до рішення проблем пошуку двомовного плагіату [Електронний ресурс]* / Д.В. Ланде, В.В. Жигало. – *Проблеми інформатизації та управління: Збірник наукових праць.* – К. : НАУ, 2008. – Вип. 2 (24). – С. 125–129. — Режим доступу до журн : www.iit.nau.edu.ua/content/archive/125-129.pdf. 7. Мартыненко Г.Я.

Новые информационные технологии систематизации и исследования художественных текстов (на материале русского рассказа XX века) [Электронный ресурс] / Г.Я. Мартыненко // Российские электронные библиотеки 2001 – Т. 4 – вып. 5. — Режим доступа до журн.: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2001/part5/martynenko>. 8. Мецераков Р.В. Модели определения авторства текста [Электронный ресурс] / Р.В. Мецераков, Н.С. Васюков. — Режим доступа до журн.: http://db.biyisk.secna.ru/conference/conference.conference.doc_download?id_thesis_dl=427. 9. Нікольський Ю. В. Дискретна математика: підручн. для студ. ВНЗ / Ю.В. Нікольський, В.В.Пасічник, Ю.М.Щербина. — К. : Видавнича група ВНУ, 2007. — 368 с. 10. Перебийніс В. С. Математична лінгвістика // Українська мова: Енциклопедія / Редкол.: В.М. Русанівський та інші. — К. : Українська енциклопедія, 2000, — С. 302–287. 11. Пиотровский Р.Г. Математическая лингвистика: Учебное пособие / Пиотровский Р. Г., Бектаев К. Б., Пиотровская А. А. — М. : Высшая школа, 1977. — 384 с. 12. Питерсон Дж. Теория сетей Петри и моделирование систем: Пер. с англ. / Дж. Питерсон. — М.: Мир, 1984.— 264 с. 13. Пропп В.Я. Морфология Волиебной сказки. Исторические корни волиебной сказки [Электронный ресурс]: [сб. трудов В.Я. Проппа] / научн. редакция, текстологический комментарий И.В. Пешкова. / В.Я. Пропп. — М. : Лабиринт, 1998. — 512 с. — Режим доступа до журн. : http://www.zipsites.ru/books/morfologiya_volsh_skazki/ 14. Старостин С.А. Сравнительно-историческое языкознание и лексикостатистика // Лингвистическая реконструкция и древнейшая история Востока. Ч. 1. [Электронный ресурс] / С.А. Старостин. — М., 1989. — Режим доступа до журн. : http://ru.wikipedia.org/wiki/Старостин,_Сергей_Анатольевич. 15 Хмелев Д. Как определить писателя? [Электронный ресурс] / Д. Хмелев. — Режим доступа до журн. : <http://offline.computerra.ru/2000/338/3010/>. 16. Хэмминг Р.В. Теория кодирования и теория информации: Пер. с англ. / Р.В. Хэмминг. — М. : Радио и связь, 1983. 17. Шульгіна В. І. Прикладна лінгвістика як навчальна та наукова спеціальність [Електронний ресурс]. — Режим доступа до статті: http://www.bdpu.org/scientific_published/akt_prob1_sl_filol-12/1.doc. 18. Щербина Ю. М. Предмет математичної лінгвістики / Ю. М. Щербина // Вісник Нац. ун-ту «Львівська політехніка». — Львів, 2002. — № 464 : Інформаційні системи та мережі. — С. 340–349. 19. Щербина Ю.М. Методичні вказівки до виконання лабораторних робіт з курсу "Математична лінгвістика" / Ю.М. Щербина, В.А. Висоцька, Т.В. Шестакевич // Львів, 2006. — № 1–5. 20. Щербина Ю.М. Методичні вказівки до виконання лабораторних робіт з курсу "Математична лінгвістика" / Ю.М. Щербина, В.А. Висоцька, Т.В. Шестакевич // Львів, 2007. — № 1–8.