

Н. К. Тимофієва, д. т. н., с.н.с. (Україна, Київ)

СПОСОБИ ВСТАНОВЛЕННЯ СУТІ ОБ'ЄКТА В ЗАДАЧАХ ІЗ СЕМАНТИКИ

До семантики відносяться задачі із штучного інтелекту: розпізнавання мовлення, в яких встановлюється суть слова, речення чи зміст зв'язного тексту, клінічна діагностика, порівняння текстів на плагіат, автоматичний переклад текстів з однієї мови на іншу. Для встановлення суті об'єкта проводиться пошук еталона в бібліотеці з наступним їхнім порівнянням. Цей пошук виконується двома способами: за ознаками, які задаються як вхідні дані, та за вхідним об'єктом, для якого шляхом порівняння знаходиться певний еталон.

Аналіз останніх досліджень та публікацій за темою. Теорії інформаційного пошуку, до якої відноситься проблема семантики, присвячено дуже багато літератури. При моделюванні інформаційних потоків вивчаються структурні зв'язки між масивами документів, які в них входять. На сьогоднішній день для моделювання цих зв'язків використовують фрактальний аналіз, методи нелінійної динаміки, теорію клітинних автоматів, методи самоорганізації тощо [1].

Постановка задачі. Для розв'язання задач із семантики необхідно проводити порівняння шуканого об'єкта з еталоном. Як правило, еталон знаходиться в базі даних. В роботі показано, що його пошук проводиться за ознаками, які задаються як вхідні дані, та за вхідним об'єктом. Також існують задачі, в яких вхідна інформація містить і об'єкт, який необхідно розпізнати, і еталон, з яким він порівнюється.

Основна частина. Задачі із семантики, як правило, відносяться до задач розпізнавання і для встановлення суті певного об'єкта потребують знаходження його еталону в базі даних. Це – розпізнавання мовлення, розпізнавання дитячого, жіночого, чоловічого голосів, задача клінічної діагностики, порівняння текстів на плагіат, автоматичний переклад текстів з однієї мови на іншу тощо. Для встановлення суті предмета використовують еталонну інформацію, яка міститься в базах даних, що та являє собою структуровану сукупність взаємопов'язаних даних певної предметної області. Для швидкого знаходження в ній необхідної інформації ця база має бути відповідно структурована. Наприклад, для зручності знаходження потрібної книги в бібліотеці розроблено різні каталоги, які структуровані певним чином. Аналіз порівняння вхідної інформації та еталону показує, що пошук певного еталону проводиться двома способами: за первинними ознаками, які описують шуканий об'єкт, і за заданим об'єктом. При моделюванні вхідних даних для пошуку інформації за першим способом має місце покриття певними ознаками заданих об'єктів. Ознаки розділяються на такі, які характеризують лише заданий об'єкт, за якими досить просто його визначити в базі даних. В цьому разі задача є розв'язною. Якщо однакові ознаки описують різні об'єкти, але за допомогою диференціального аналізу можна знайти потрібний об'єкт, то така задача є частково розв'язною. Якщо одні і ті ж ознаки характеризують різні об'єкти і за ними не можна ідентифікувати шуканий, то виникає ситуація невизначеності. При другому способі попереднє покриття певними ознаками еталону і об'єкта не проводиться.

Існують задачі, які для розпізнавання об'єктів не потребують бібліотеки еталонів. В них вхідні дані розділяються на сегменти з наступним визначенням подібності отриманих частин. Вхідна інформація в цьому разі містить і об'єкт, який необхідно розпізнати, і еталон, з яким він порівнюється. В деяких задачах за еталон приймають або вираз, за яким визначають суть вхідної інформації, або задають умови, за якими можна розпізнати певний об'єкт. Для розв'язання таких задач бібліотека еталонів не використовується.

Висновок. Отже, пошук еталону в бібліотеці еталонів проводиться двома способами: за первинними ознаками, які описують шуканий об'єкт, і за заданим об'єктом. Також існують задачі з розпізнавання вхідні дані яких розділяються на сегменти з наступним визначенням подібності отриманих частин. Вхідна інформація в цьому разі містить і об'єкт, який необхідно розпізнати, і еталон, з яким він порівнюється. В деяких задачах за еталон приймають або вираз, за яким визначають суть вхідної інформації, або задають умови, за якими можна розпізнати заданий об'єкт. Розв'язання таких задач проводиться без бібліотеки еталонів.

Література.

1. Ланде Д.В., Фурашев В.М. *Основи інформаційного і соціально-правового моделювання*: монографія. К.: ТОВ "ПанТот", 2012. 144 с.