**Сліпченко Сергій Віталійович. Обробка числової і символьної інформації на основі розподілених представлень в задачах штучного інтелекту. : Дис... канд. наук: 05.13.23 – 2006**

|  |  |
| --- | --- |
| |  | | --- | | **Сліпченко С.В. Обробка числової і символьної інформації на основі розподілених представлень в задачах штучного інтелекту. – Рукопис.**  Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – Cистеми та засоби штучного інтелекту. – Інститут проблем математичних машин та систем НАН України, Київ, 2006.  Дисертація присвячена розробці й дослідженню методів розподіленого представлення та обробки числової й символьної структурованої інформації. Проведено теоретичне та експериментальне дослідження методів формування бінарних розріджених розподілених представлень інформації та методів її обробки. Розроблено методи формування бінарних представлень для вирішення задач пошуку, відображення та виводу за аналогією. Запропоновані методи дозволили покращити повноту від 11 до 22% та точність пошуку в 3-4 рази. Методи відображення та виводу за аналогією дали результати, що відповідають результатам психологічних тестів та існуючих систем, але на відміну від останніх запропоновані методи спроможні обробляти не тільки символьну інформацію. Розроблено методи представлення числових векторів і більш складних структур за допомогою бінарних розріджених розподілених представлень. В задачі прогнозування існування хімічних сполук запропоновані методи показали результат на рівні кращих систем – 94,8%-99,91%. Отримано аналітичні характеристики методу грубого кодування *Prager* (щільність кодів, перекриття, роздільна здатність та інші), які дозволяють ефективно вибирати параметри методу при вирішенні практичних задач. Алгоритми й методи реалізовані у вигляді програмних та інструментально-технологічних засобів штучного інтелекту, що пов’язані із обробкою знань та класифікацією. Виконано експериментальне дослідження на реальних і тестових даних. | |
| |  | | --- | | Дисертаційна робота присвячена розробці та дослідженню нових методів розподіленого представлення й обробки числової та символьної структурованої інформації. Аналітично та експериментально досліджено характеристики розроблених бінарних розріджених розподілених представлень інформації й методів їх обробки. Розвинені представлення та методи підвищили ефективність пошуку і відображення аналогів у базах знань, виводу за аналогією, класифікації числових даних. Розроблені методи реалізовані у вигляді програмних та інструментально-технологічних засобів для систем штучного інтелекту, які пов’язані з обробкою знань та класифікацією, і перевірені на реальних і тестових даних.  За результатами проведеного дослідження зроблено такі висновки:  1. Розроблені мова *Slang* та її інтерпретатор за рахунок поєднання символьних і розподілених представлень інформації дозволили формалізувати та суттєво спростити моделювання пошуку, відображення й виводу за аналогією.  2. Розроблений та реалізований засобами *Analogу Toolbox* метод пошуку аналогів за рахунок вдосконалення методів їх представлення за допомогою бінарних розріджених розподілених представлень дозволив покращити повноту пошуку на 11-22% і точність в 3-4 рази в порівнянні з кращим відомим символьним методом *MAC*/*FAC* при перевірці на розширеній базі аналогів *ThinkNet*, а результати запропонованих методів відображення й виводу за аналогією співпадають з результатами кращих відомих символьних методів *SME* та *ACME*.  3. Відповідність результатів експериментальних досліджень розроблених моделей пошуку, відображення та виводу за аналогією результатам відомих психологічних експериментів підтвердила адекватність запропонованих моделей. Розроблені методи представлення, пошуку, відображення та виводу за аналогією за рахунок низької обчислювальної складності та паралельності обробки інформації дозволяють забезпечити їх обчислювально-ефективну реалізацію.  4. Розроблені методи формування бінарних багатовимірних розріджених представлень для числових векторів за рахунок зв’язування розподілених представлень ознак та їх значень дозволили зменшити обчислювальні потреби та об’єми пам’яті пропорційно до розмірності вхідного вектора. Запропоновані методи відновлення числових векторів за їх розподіленими представленнями відкривають можливості інтеграції систем, що мають за основу такі представлення, з класичними архітектурами. Отримані експериментальні залежності кількості помилок декодування й відстані між вхідним і декодованим векторами дозволяють оцінювати якість відновлення для широкого діапазону параметрів та обирати ефективні значення параметрів при вирішенні практичних задач.  5. Отримані аналітичні характеристики розподілу розмірності полів, щільності, перекриття, роздільної здатності бінарних представлень числових векторів для методів кодування *Prager* дозволяють створювати ефективну конфігурацію експлуатаційних характеристик прикладних систем, що підтверджено експериментами на тестових та реальних даних баз *Elena*, *DGEN* та інших.  6. Розроблені методи кодування, реалізовані у програмному засобі *Coding* *Toolbox,*при прогнозуванні існування хімічних сполук забезпечують результат на рівні кращих досягнень у цій області: 94,8% – 99,91% (залежно від кількості компонент у сполуці) порівняно із 94,8% – 99,68% системи *АНАЛОГІЯ* та 89,11% – 99,32% системи *DISCOVERY*, що підтверджує доцільність використання запропонованих методів кодування для вирішення задач зі складною структурою даних.  7. Розроблені методи реалізовані у вигляді інструментально-технологічних та програмних засобів, які є компонентами інформаційних технологій, пов’язаних з класифікацією, пошуком і виводом за аналогією, а також можуть бути використані для впровадження моделей когнітивних процесів у системи штучного інтелекту. | |