

Б.А.Кауров, канд. биол. наук

(ФГУ Российский геронтологический научно-клинический центр, Москва)

ЭЛЕМЕНТЫ СОЗДАНИЯ СПЕЦИАЛИЗИРОВАННОГО ЯЗЫКА GERON ДЛЯ ФОРМАЛИЗОВАННОГО ОПИСАНИЯ И МОДЕЛИРОВАНИЯ ВОЗРАСТНЫХ ИЗМЕНЕНИЙ ОРГАНИЗМА ЧЕЛОВЕКА

Рассмотрена возможность создания специализированного языка GERON для формализованного описания и моделирования возрастных изменений организма человека.

Ключевые слова: формализованный язык описания, возрастные изменения, старость, человек.

В связи с интенсивным и постоянным накоплением новой разноуровневой биологической информации, необходимостью ее своевременной обработки и нахождения взаимосвязей с предшествующей информацией встает острая проблема ее логической обработки. Для решения данной проблемы создано целое научное направление – анализ знаний (Data Minig) [1]. Однако, несмотря на то, что оно в разной форме существует уже несколько десятилетий, значительного прорыва в этой области пока не наблюдается. На то есть объективные причины. Прежде всего это сложность автоматической структуризации, формализации и семантического анализа специализированной текстовой информации, требующей огромной предварительной работы по формализации исходной предметной области. Такая формализация невозможна без создания соответствующего языка описания данных.

Применительно к геронтологии это означает создание специфического и формализованного языка описания и моделирования различных молекулярно-клеточных и морфофизиологических процессов старения. В сообщении [2] рассмотрены некоторые возможные варианты таких языков. В связи с существенно возросшим интересом к этой теме стали появляться специализированные языки, основанные на графических подходах. Самым последним амбициозным проектом в данной области стало создание языка системной биологии SBGN [3] для графического описания биохимических и клеточных процессов.

Однако SBGN, как и другие графические языки [4, 6], имеет существенное ограничение в плане визуальной наглядности нарисованных с его помощью схем при большом количестве узлов и дуг, отображающих разные процессы. В этом можно убедиться на примере моей схемы старения человека [5], построенной с помощью системного графического языка Стар Tools [6]. Кроме того, они не позволяют осуществлять динамическое моделирование рассматриваемых процессов.

На мой взгляд, более перспективным является описание системных биологических процессов на основе формализованных языков с другой идеологией построения, близкой к объектно-ориентированной с элементами искусственного интеллекта, к которым относится и разрабатываемый мной язык GERON. Попытки создания подобного языка мной предпринимались и раньше [7]. С этой целью по-

сле анализа соответствующих медицинских и биологических текстов был выделен ряд основных классов объектов, формализованы их свойства и возможные значения, сформулированы правила их поведения. Ведущим понятием в языке GERON является объект. Поведение объектов внутри класса подчиняется определенным правилам с частично детерминированной логикой. Объект обладает такими параметрами как принадлежность к определенному классу (шаблону) объектов на основании одного или нескольких объединяющих их признаков; специфическими в пределах своего класса свойствами (включая способность к определенным взаимоотношениям с другими объектами, в том числе из других классов) и их значениями (постоянными или переменными); способностью к сохранению и изменению (действиям) своего состояния с помощью определенных методов в ответ на определенные события (сигналы к действиям); уровнем вложенности (организации) в объекты других классов (были выделены следующие основные уровни: генетический, молекулярно-клеточный, органно-тканевый, организменный и популяционный с соответствующими характеристиками); наследованием свойств при своей репликации. Все перечисленные параметры объекта зависят от времени. Внутри каждого класса объектов при необходимости выделяются свои подклассы, обладающие дополнительным специфическим набором свойств и их значений. Например, внутри класса объектов "клетки" выделены подклассы объектов, соответствующих разным типам дифференцированных клеток. Учитывая необходимость широкого использования разных баз данных, в настоящее время делается попытка реализовать данный язык на платформе СУБД VFP, поддерживающей как модульное, так и визуальное объектно-ориентированное программирование высокого уровня, а также связь с другими приложениями.

Разработка формализованного языка описания возрастных изменений организма GERON позволит создать необходимую методическую основу для развития нового направления в геронтологии – возрастной биоинформатики человека [8] и последующего связывания созданной базы данных его возрастных изменений с экспертной системой Ageing-Expert [2]. Все это в итоге ускорит построение практически значимой формализованной системной модели старения человека и будет способствовать более эффективному анализу геронтологических данных.

ЛИТЕРАТУРА

1. Чубукова И.А. Data Mining. – М.: БИНОМ, 2008. – 382 с.
2. Кауров Б.А. О формализованной системной модели старения человека // Информатика и системы управления. – 2009. – N4(22). – С. 26-28.
3. Le Novire N. et al. The Systems Biology Graphical Notation // Nature Biotechnology. – 2009. N27(8). – P. 735-741.
4. Буч Г., Рамбо Д., Якобсон И. Язык UML. Руководство пользователя. – М.: ДМК, 2007. – 496 с.
5. Кауров Б.А. Новая схема старения человека // VI Всероссийская научно-практ. конф. "Общество, государство и медицина для пожилых". – М., 2009. – С. 31-32.
6. Муромцев Д.И. Концептуальное моделирование знаний в системе Spar Tools. – СПб. СПб ГУ ИТМО, 2009. – 83 с.

7. Кауров Б.А. Живой организм как имитационная система // Биометрические аспекты изучения целостного организма. – М.: МГУ, 1987. – С. 82-91.
8. Кауров Б.А. Возрастная биоинформатика человека – новое направление в геронтологии // VI Всероссийская научно-практ. конф. "Общество, государство и медицина для пожилых". – М., 2009. – С. 33-34.

E-mail: bokar@mail.ru.

УДК 001.891.34:681.51.012

О.Г. Берестнева, д-р техн. наук, **Е.А. Муратова**, канд. техн. наук
(Томский политехнический университет)

ПРОБЛЕМЫ УНИФИКАЦИИ ДАННЫХ В НАУЧНЫХ ПСИХОЛОГИЧЕСКИХ И МЕДИЦИНСКИХ ИССЛЕДОВАНИЯХ¹

Представлен обзор приемов унификации разнотипных данных. Предложен эффективный алгоритм адаптивного кодирования разнотипной информации, позволяющий обрабатывать малые выборки, строить локальные диагностические шкалы, формировать новые знания на основе экспериментального материала.

Ключевые слова: информационные интеллектуальные системы, унификация данных и знаний, алгоритм адаптивного кодирования разнотипной информации.

Введение

В настоящее время использование информационных интеллектуальных систем (ИИС), в том числе и автоматизированных рабочих мест (АРМ), выполняющих роль эксперта и консультанта, бесспорно актуально. При их разработке решаются проблемы, связанные с представлением данных и знаний, выявлением закономерностей, формированием решающих правил, принятием и обоснованием решений. Выявление закономерностей из имеющихся массивов данных на начальных этапах проектирования ИИС и АРМ – один из наиболее сложных и трудоемких процессов, и он не всегда заканчивается успешно, поскольку базы данных содержат неполную, нечеткую, разнотипную и противоречивую информацию.

Например, специалисту в области медицинской психодиагностики придется оперировать следующими видами информации:

- 1) *психодиагностической* (данные психодиагностического эксперимента, измеренные в количественных, порядковых, номинальных шкалах);
- 2) *медицинской* (результаты параклинических методов исследования, имеющие количественный и интервальный характер);
- 3) *консультационно-биологической* (пол, возраст, образование, профессия, семейное положение – качественные шкалы);

¹ Работа выполнена при финансовой поддержке РФФИ (проект №08-06-00313а).