



OSTIS-2011

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:

СЕМАНТИЧЕСКАЯ ТЕХНОЛОГИЯ КОМПОНЕНТНОГО ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

С.А. Самодумкин (samodumkin@bsuir.by), С.И. Сорока, А.И. Махина, А.С. Глазунов
*Белорусский государственный университет информатики и радиоэлектроники,
г.Минск, Республика Беларусь*

В работе рассмотрена технология проектирования интеллектуальных геоинформационных систем на основе открытой семантической технологии компонентного проектирования интеллектуальных систем.

Ключевые слова: интеллектуальная геоинформационная система, семантическая технология.

Введение

В индустрии разработки геоинформационных систем (ГИС) в настоящее время обозначилась потребность в их интеллектуализации, т.е. решение задач традиционно относящихся к геоинформатике с применением методов искусственного интеллекта, и в первую очередь – интеллектуального поиска. Существующие инструментальные ГИС [Абламейко, 2000], [Крючков, 2006] являющиеся средством разработки прикладных ГИС, задачи интеллектуального поиска не решают по ряду причин:

- практически все они построены на внутренних (закрытых) форматах представления пространственных данных, а обменные открытые форматы служат лишь средством обмена картографическими данными между различными инструментальными ГИС;
- тематические данные сопоставлены с конкретными пространственными объектами и нет дополнительной возможности установления связей и отношений между такими данными;
- реализация прикладных задач геоинформатики осуществляется на внутренних языках программирования, тем самым упрощая лишь доступ к пространственным данным, а карта служит только средством визуализации.

С другой стороны, известные технологии проектирования интеллектуальных систем используют картографические материалы, как правило, в виде растрового изображения, т.е. отсутствует возможность рассматривать карту как совокупность географических объектов с заданными топологическими и предметно-ориентированными (в зависимости от типа карты) отношениями. В последнее время благодаря развитию технологии «Semantic Web», ключевым элементом которой являются онтологии, появилась возможность в ГИС акцентировать внимание на семантике предметных знаний [Schevers, 2005], осуществлять интеграцию и слияние различных наборов данных в смежных областях, устанавливать правила предметной области и их запись с использованием RDF (Resource Description Framework), что безусловно расширяет возможности ГИС-технологий. Данные обстоятельства указывают на необходимость интеграции пространственных данных и предметных знаний. В работе в качестве языкового средства для представления знаний используется универсальный абстрактный язык семантических сетей, обеспечивающий унифицированное кодирование семантических сетей [Голенков и др., 2001].

Целью данной работы является создание в рамках открытой семантической технологии компонентного проектирования интеллектуальных систем (ОСТИС) [OSTIS, 2010] технологии проектирования интеллектуальных геоинформационных систем, что позволит разрабатывать новый класс систем – интеллектуальные геоинформационные системы.

Поставленная цель определяет следующие задачи:

- разработка модели пространственной информации совместимой с моделью представления знаний проекта ОСТИС;
- разработка способа задания карт антропоморфного способу записи исходного текста базы знаний на псевдоестественном языке;
- разработка средств трансляции карт в базу знаний интеллектуальной системы;
- разработка системы операций, реализующей пространственные запросы, и вопросно-ответный режим прикладной геоинформационной системы;
- разработка средств визуализации карт.

Компоненты семантической технологии проектирования интеллектуальных геоинформационных систем

В соответствие с общими принципами проектирования интеллектуальных систем по технологии ОСТИС выделяются компоненты баз знаний, машина обработки знаний и пользовательский интерфейс. Спецификой рассматриваемого класса систем является, во-первых, способ кодирования исходной картографической информации и ее интеграция с предметными базами знаний, во-вторых, особый вид операций, реализующих пространственные запросы, и, в-третьих, новый вид пользовательского интерфейса – картографический интерфейс. На рисунке 1 представлены компоненты семантической технологии проектирования интеллектуальных геоинформационных систем.

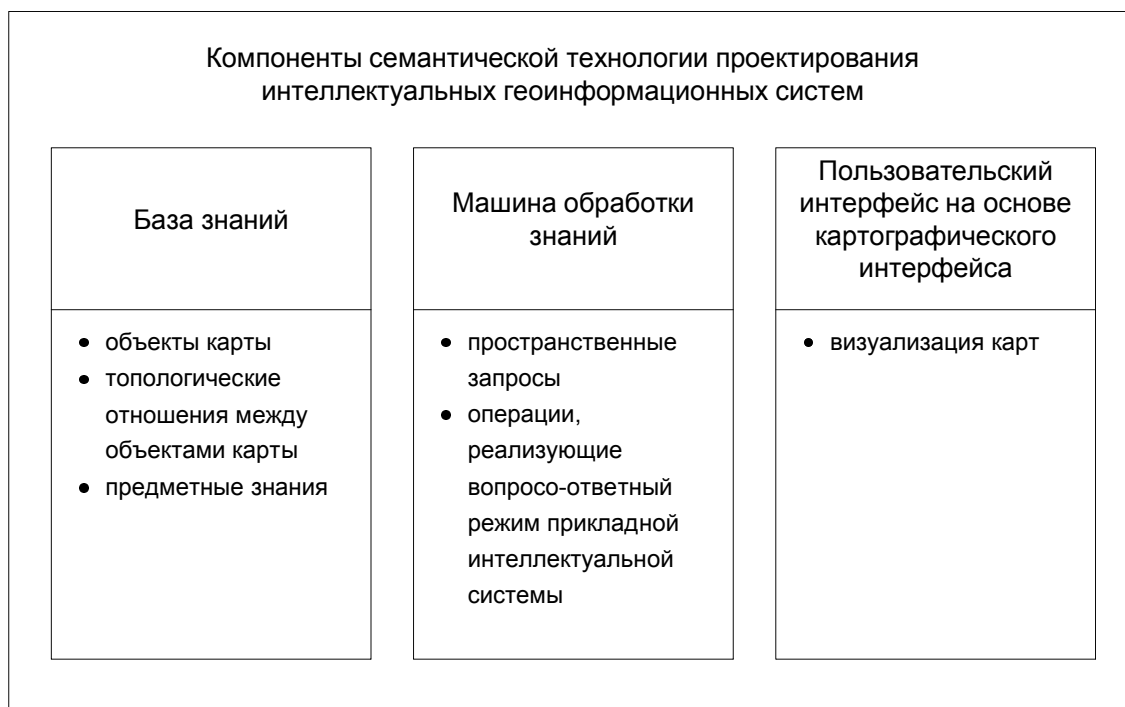


Рисунок 1. – Компоненты семантической технологии проектирования интеллектуальных геоинформационных систем.

База знаний интеллектуальной геоинформационной системы может быть получена поэтапно. В качестве исходной информации для формирования базы знаний выступают электронная карта, в которой хранится совокупность географических объектов, классификатор топографической информации, в котором описаны типы объектов, установлены правила задания объектов и их

свойств, и внешние предметно-ориентированные базы знаний (рисунок 2). На первом этапе трансляции происходит семантическое сопоставление географических объектов карты и установление топологических отношений между объектами карты. В результате формируется первая версия базы знаний, в которой находятся объекты карты и установлены топологические отношения. Таким образом, в результате трансляции уже появляется возможность в получении новых знаний.

Вторая версия базы знаний является результатом интеграции базы знаний, полученная на первом этапе и внешних предметно-ориентированных баз знаний. Такая интеграция позволяет нарастить базу знаний новыми видами знаний, а также устранить омонимию географических объектов. Так, например, использование базы знаний населенных пунктов с кодами по СОАТО (системе обозначений объектов административно-территориального деления и населенных пунктов) позволяет взаимно однозначным образом идентифицировать населенные пункты с одинаковым названием, но принадлежащим различным административно-территориальным единицам.



Рисунок 2. – Компоненты баз знаний и этапы их получения

Представление географических объектов в базе знаний потребовало разработки модели пространственной информации. Главными критериями, предъявляемыми к данной модели являются:

- интегрируемость с моделью знаний, которая лежит в основе построенных на основе технологии ОСТИС интеллектуальных систем;
- кодирование объектов карт и их признаков (характеристик) в соответствии с классификатором, на основе которого создана цифровая карта.

В соответствии с первым критерием модель пространственной информации относится к классу семантических моделей, а карта, основанная на данной модели, в базе знаний представляется семантической сетью. Второй критерий накладывает ограничения на исходные цифровые карты не по способу их представления (во внутреннем или обменном формате она представлена), а на семантическое содержание объектов карты, которое задается с помощью классификатора топографической информации, отображаемой на карте. Цифровые карты в Республике Беларусь создаются на основе классификатора ОКРБ 012-2007 «Цифровые карты местности. Топографическая информация, отображаемая на топографических картах и планах городов» [ОКРБ, 2007], который положен в основу онтологии карты.

Каждый объект карты однозначно задается классификационным кодом объекта цифровой карты (ЦК), имеет восемь ступеней классификации и состоит из кода класса, кода подкласса, кода группы, кода подгруппы, кода отряда, кода подотряда, кода вида, кода подвида. Правило задания классификационного кода объекта карты в семантической модели пространственной информации имеет вид, представленный на рисунке 3.

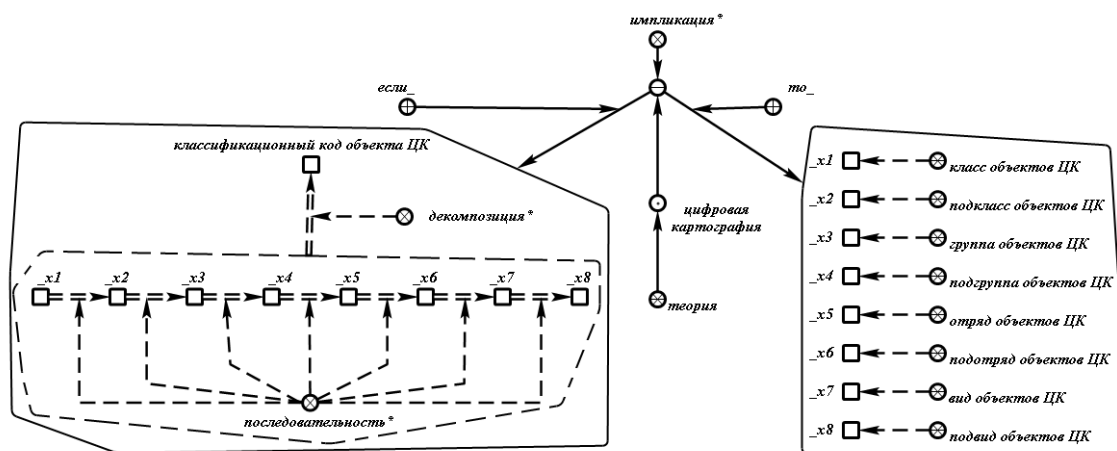


Рисунок 3. Правило формирования классификационного кода объекта карты

Семантическая модель пространственной информации включает две связанные интегрируемые семантические подсети:

- 1) семантическую сеть, однозначно определяющую объекты местности и содержащую сведения об основных признаках и постоянных свойствах этих объектов;
- 2) семантическую сеть с информацией о переменных свойствах, характеризующих объект карты и его отношение к другим объектам.

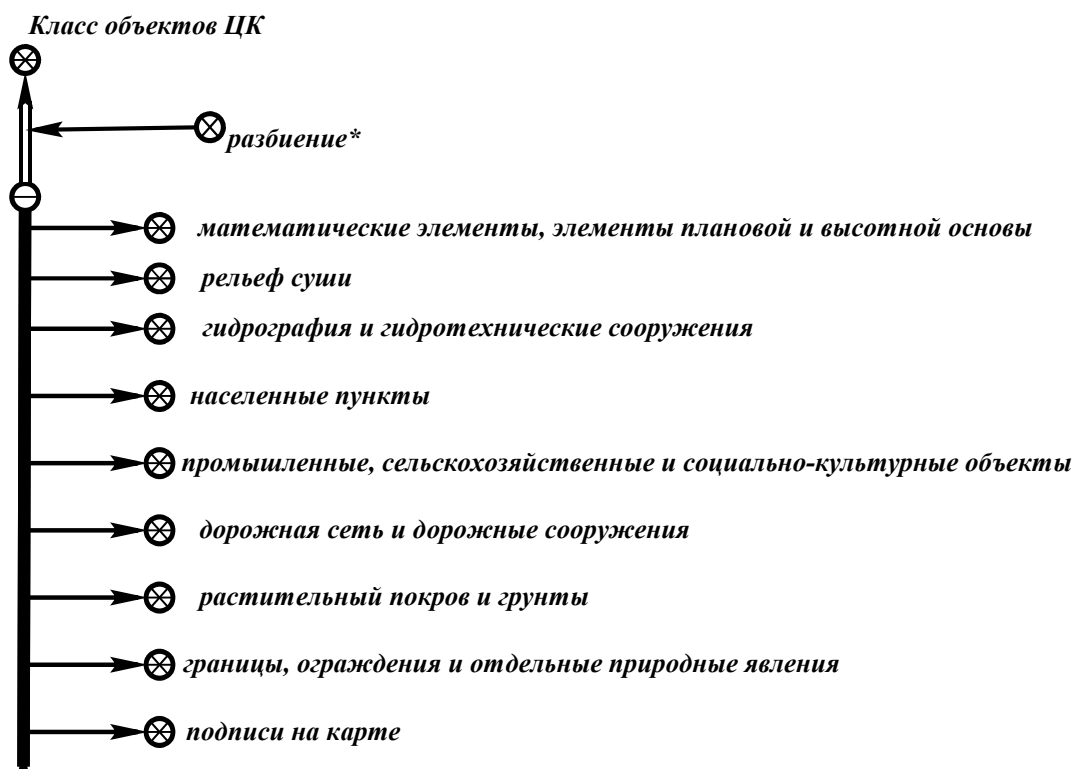
Семантическая сеть, однозначно определяющая объекты местности и содержащая сведения об основных признаках и постоянных свойствах этих объектов представляет собой иерархическую структуру. На верхнем уровне иерархии выделено восемь классов объектов цифровой карты, соответствующих элементам содержания карты

Класс объектов ЦК

– Разбиение:

- математические элементы и элементы плановой и высотной основы
= 10000000 код ЦК
- рельеф суши
= 20000000 код ЦК
- гидрография и гидротехнические сооружения
= 30000000 код ЦК
- населенные пункты
= 40000000 код ЦК
- промышленные, сельскохозяйственные и социально-культурные объекты
= 50000000 код ЦК
- дорожная сеть и дорожные сооружения
= 60000000 код ЦК
- растительный покров и грунты
= 70000000 код ЦК
- границы, ограждения и отдельные природные явления
= 80000000 код ЦК

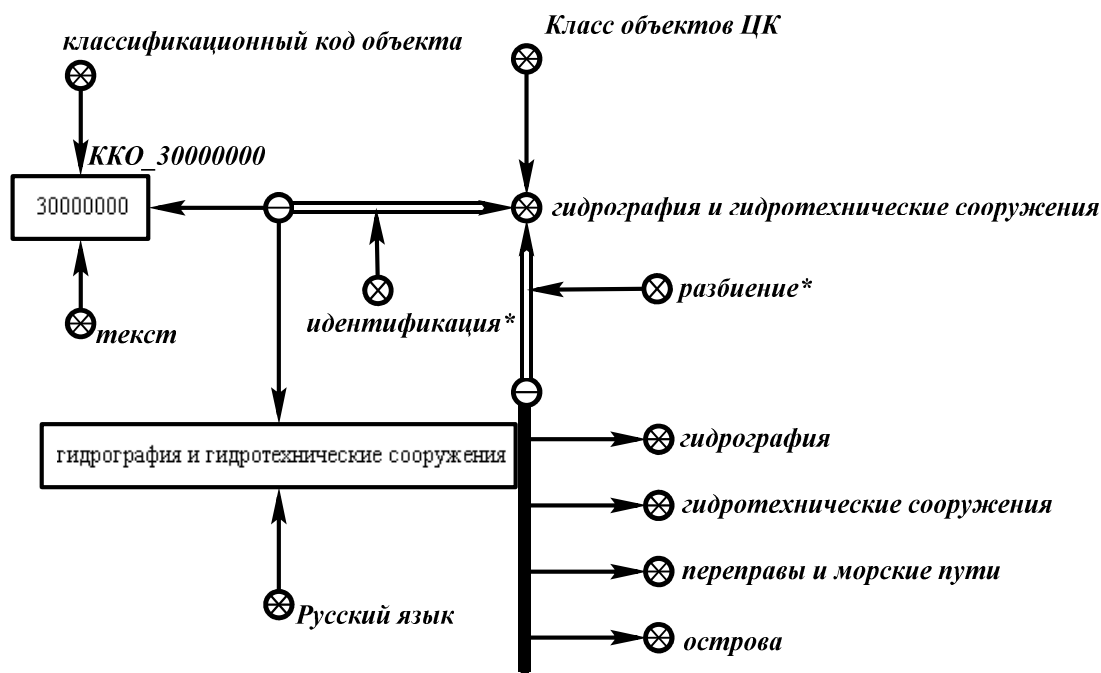
- названия и подписи
= 90000000 код ЦК



Развитие каждого класса в глубину и ширину различно и зависит от степени взаимосвязности объектов и выбранных признаков классификации. Например, разбиение класса объектов ЦК *гидрография и гидротехнические сооружения* с классификационным кодом 30000000 на осуществляется на четыре подкласса

гидрография и гидротехнические сооружения

- = Класс всевозможных объектов ЦК, относящихся к гидрографии и гидротехническим сооружениям
- = 30000000 код ЦК
- Разбиение :
 - гидрография
 - = Подкласс всевозможных объектов ЦК, относящихся к гидрографии
 - = 31000000 код ЦК
 - гидротехнические сооружения
 - = Подкласс всевозможных объектов ЦК, относящихся к гидротехническим сооружениям
 - = 32000000 код ЦК
 - переправы и морские пути
 - = Подкласс всевозможных объектов ЦК, относящихся к переправам и морским путям
 - = 33000000 код ЦК
 - острова
 - = Подкласс всевозможных объектов ЦК, относящихся к островам
 - = 34000000 код ЦК



На следующем уровне иерархии подкласс всевозможных объектов ЦК, относящихся к гидрографии разбивается на 7 групп

гидрография

= Подкласс всевозможных объектов ЦК, относящихся к гидрографии

= 31000000 код ЦК

– Р а з б и е н и е :

- водоемы

- = Группа всевозможных объектов ЦК, относящихся к водоемам

- = 31100000 код ЦК

- объекты прибрежной полосы

- = Группа всевозможных объектов ЦК, относящихся к объектам прибрежной полосы

- = 31200000 код ЦК

- рельеф дна

- = Группа всевозможных объектов ЦК, относящихся к объектам рельефа дна

- = 31300000 код ЦК

- водотоки

- = Группа всевозможных объектов ЦК, относящихся к водотокам

- = 31400000 код ЦК

- характеристики гидрографии на карте, выделяемые как самостоятельные объекты

- = Группа всевозможных объектов ЦК, относящихся к характеристикам

- гидрографии на карте, выделяемых как самостоятельные объекты

- = 31500000 код ЦК

- водоисточники

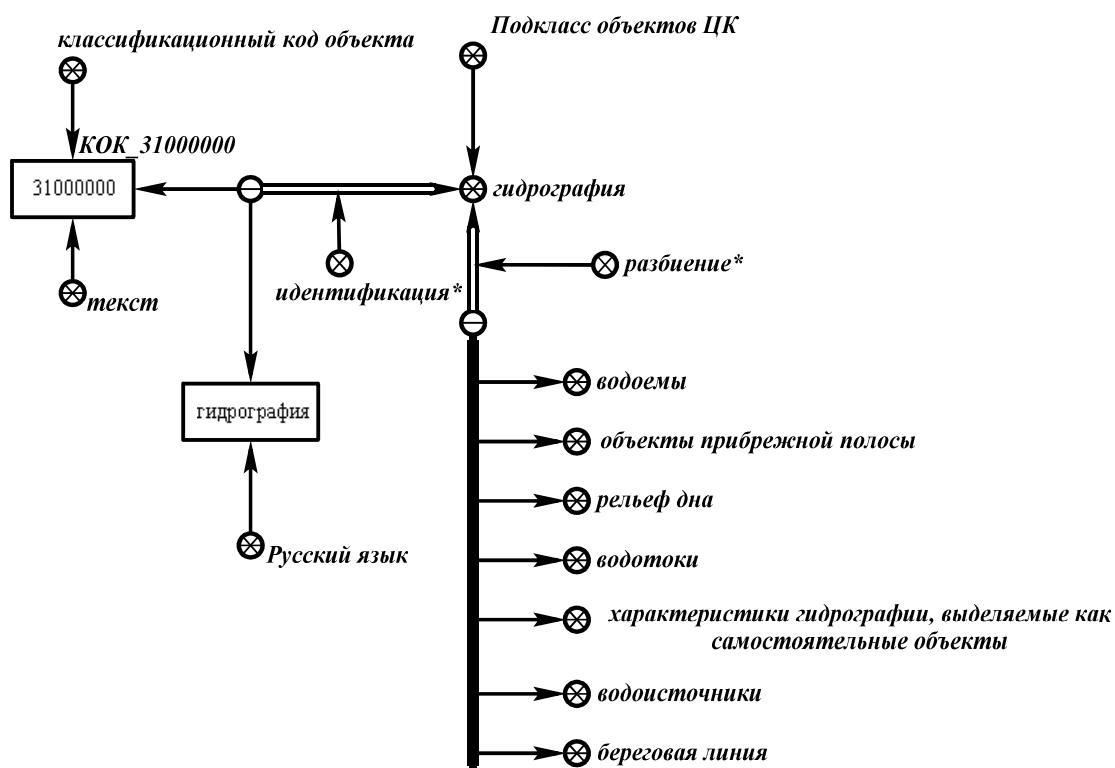
- = Группа всевозможных объектов ЦК, относящихся к водоисточникам

- = 31600000 код ЦК

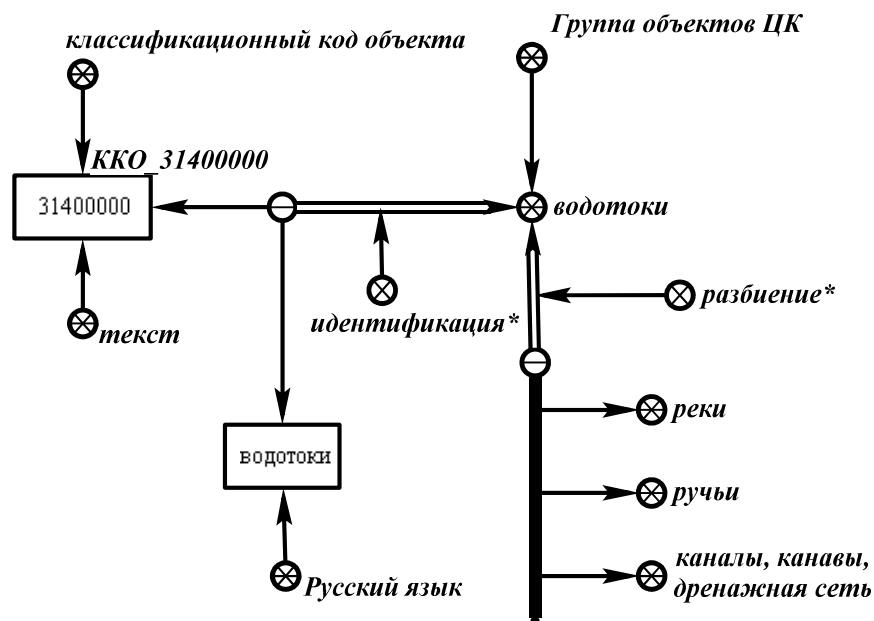
- береговая линия

- = Группа всевозможных объектов ЦК, относящихся к береговым линиям

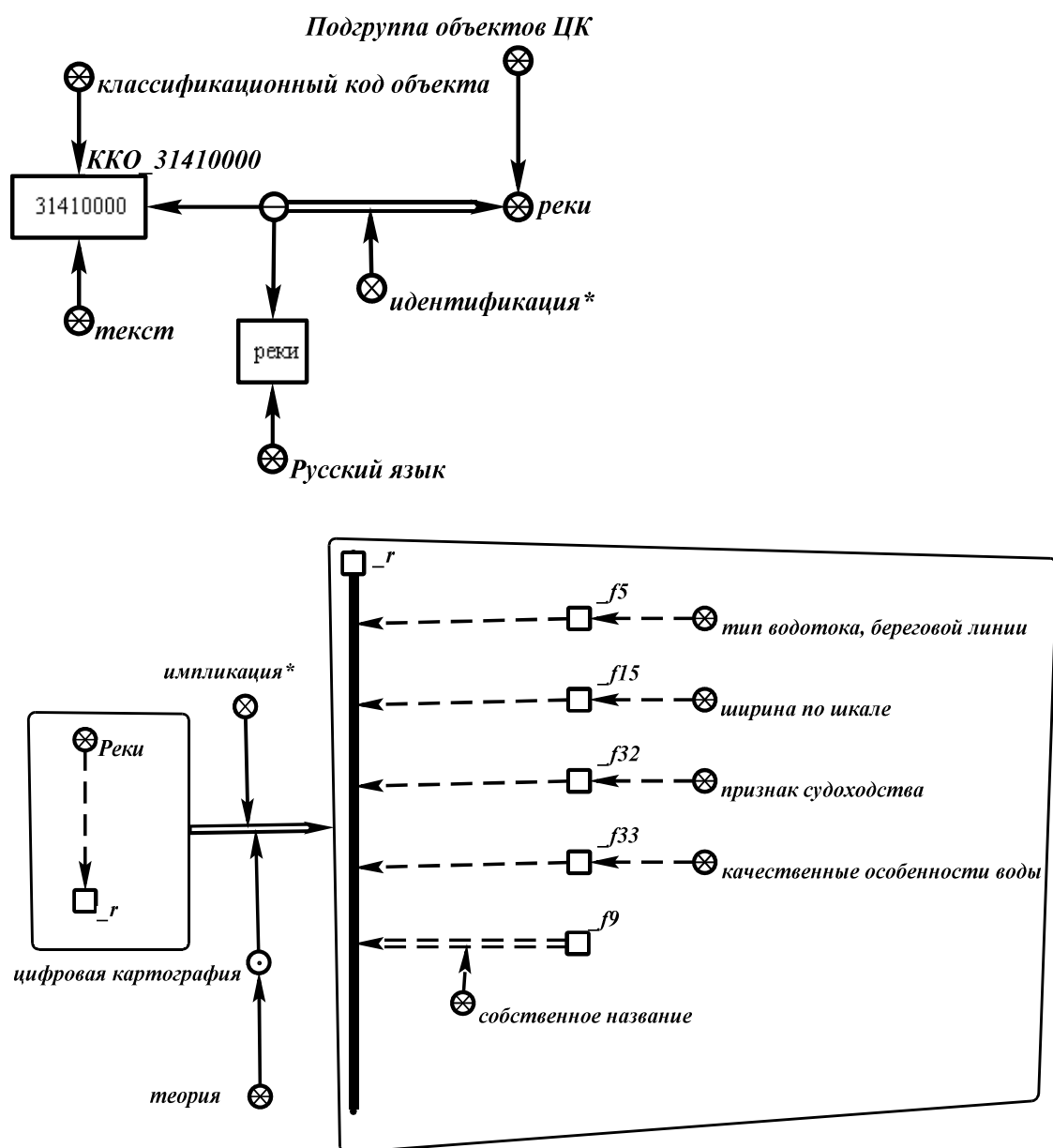
- = 31700000 код ЦК



В свою очередь группы объектов ЦК могут разбиваться на подгруппы и далее на отряды, подотряды, виды и подвиды. Например, группа всевозможных водотоков ЦК разбивается на три подгруппы

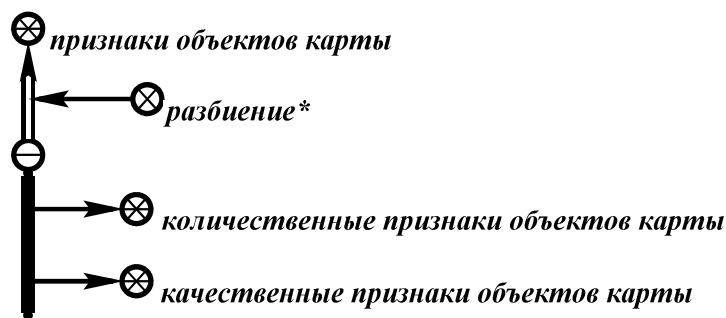


На нижней ступени классификации находится совокупность однотипных объектов карты, содержащих при необходимости, определенный набор количественных и качественных характеристик. В рассматриваемом примере для объектов карты, принадлежащих подгруппе *реки*, определен количественный признак – *собственное название*, и пять качественных признаков: *тип водотока, береговой линии*; *ширина по шкале*; *признак судоходства*, *качественные особенности воды*; *характер расположения объекта относительно земной (водной) поверхности*



Вторая подсеть задает переменные свойства, характеризующие объект карты и его отношение к другим объектам и строится по принципу фасетного метода классификации, т.е. отнесение классов географических объектов, определенных в первой подсети, к разным признакам.

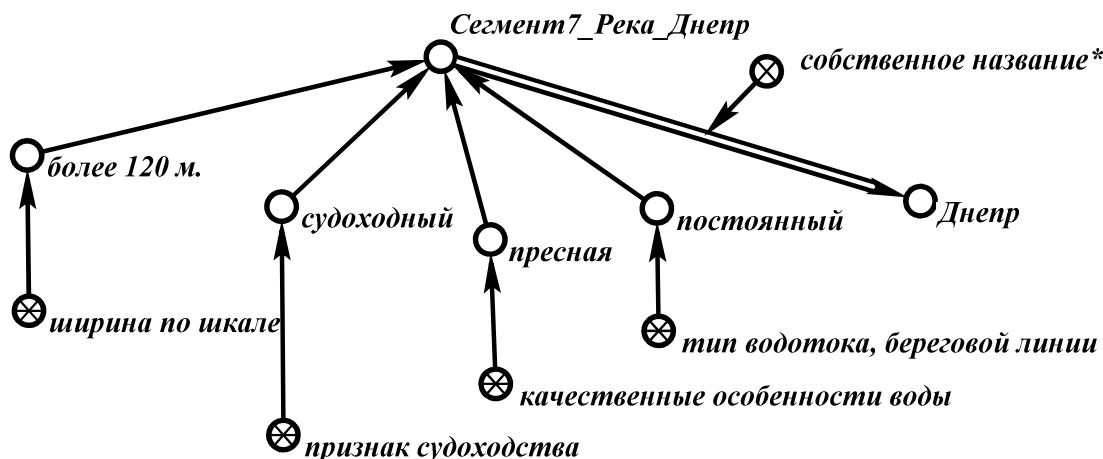
Все признаки объектов карты разбиваются на два подмножества признаков (*количественные признаки объектов карты* и *качественные признаки объектов карты*)



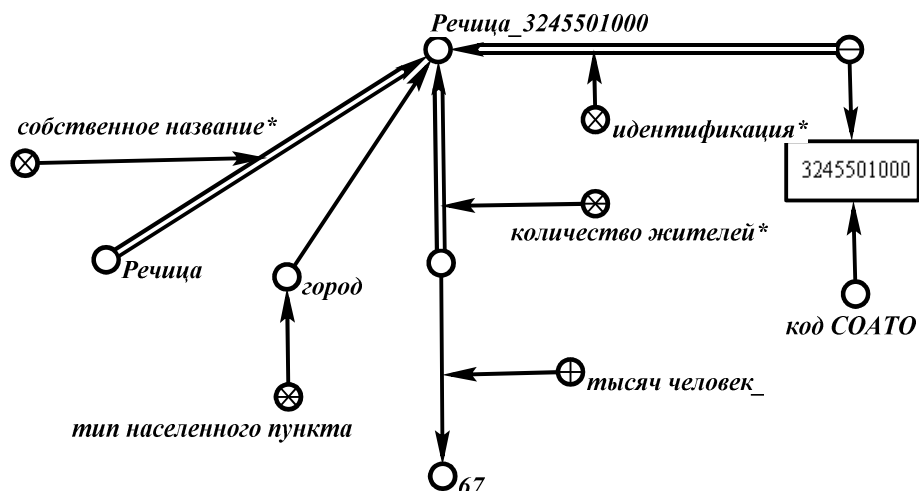
Для каждого признака в базе знаний заданы его характеристики. Так количественные признаки задаются бинарными ориентированными отношениями, связующие географический объект с значением количественного признака, а качественные признаки задаются дугами принадлежности географических объектов к возможному значению из множества всевозможных значений качественного признака.

В качестве исходных картографических данных предложено использование карт, подготовленных в открытом обменном формате mid/mif [2], что соответствует идеологии проектирования интеллектуальных систем по технологии OSTIS. Карта в формате mid/mif представляется как совокупность слоев, каждый из которых описывает набор родственных, схожих, принадлежащих одному классу объектов, но слои карты имеют связность за счет одной системы координат. Mif-файл содержит служебную информацию, задает структуру объектов слоя, их атрибуты, содержит координаты объектов. В mid-файле определены атрибуты объектов, которые объявлены в mif-файле.

В результате трансляции карты в базу знаний формируются знания о географическом объекте. Например, сегмент реки Днепр около города Речица будет описан следующей семантической сетью:



Интеграция с внешними базами знаний, где содержится предметная информация о географических объектах, позволяет расширить базу знаний новыми видами знаний. Так для населенных пунктов будет добавлена информация о территориальной принадлежности, количество жителей и т.п. Рассмотрим на примере населенного пункта – город Речица. В исходной карте содержится собственное название, тип населенного пункта (*город*), в базе знаний административно-территориального деления Республики Беларусь – информация о количестве жителей и уникальный код по системе обозначения административно-территориальных объектов (3245501000):



В текущей версии проекта используется машина обработки знаний со стандартным набором операций и начата реализация операций, осуществляющих пространственные запросы.

Интерфейс пользователя поддержан картографическим редактором, совместимым с системой координат исходного набора картографической информации.

Приведенный в работе подход к интеграции геоинформационных данных в памяти интеллектуальных справочных систем, реализованных в соответствии с технологией OSTIS, позволяет использовать разработанные компоненты как непосредственно для проектирования геоинформационных систем с интеллектуальным выводом, так и использовать на уровне многократно используемых компонентов карты местности. Скомпонованная система протестирована на примере местности – Республика Беларусь, а разработанные инструментальные средства позволяют загружать в память интеллектуальных систем карты, соответствующие классификатору [ОКРБ, 2007] в обменном формате mid/mif.

Работа выполнена при поддержке молодежного гранта БГУИР «Научно-методическое обеспечение дисциплин специализации «Интеллектуальные геоинформационные системы».

Библиографический список

- [Абламейко, 2000] Абламейко, С. В. Географические информационные системы. Создание цифровых карт / С. В. Абламейко, Г. П. Апарин, А. Н. Крючков. – Минск : Ин-т техн. кибернетики НАН Беларуси, 2000.
- [Голенков и др, 2001] Голенков, В.В. Представление и обработка знаний в графодинамических ассоциативных машинах /В. В. Голенков[и др.] – Мн. : БГУИР, 2001.
- [Крючков, 2006] Крючков, А.Н. Интеллектуальные технологии в геоинформационных системах / А.Н. Крючков, С.А. Самодумкин, М.Д. Степанова, Н.А. Гулякина; Под науч. ред. В.В.Голенкова. – Мн.: БГУИР, 2006.
- [ОКРБ, 2007] Цифровые карты местности. Топографическая информация, отображаемая на топографических картах и планах городов / ОКРБ 012-2007.
- [Schervers, 2005] Schervers, H.A. Semantic Web for an Integrated Urban Software System / H.A. Schervers, R.M. Drogemuller //MODSIM 05, P.2040-2046.