

## ГЕОРЕФЕРЕНЦИЯ КАК ИНСТРУМЕНТ АНАЛИЗА И ПОЛУЧЕНИЯ ЗНАНИЙ

### GEOREFERENCING AS THE ANALYSIS AND KNOWLEDGE TOOL



**Цветков В.Я. / Tsvetkov V. Ja.**

Доктор технических наук, профессор Московского государственного университета геодезии и картографии, Заслуженный деятель науки и образования Российской Федерации / Doctor of Tech.Sci., professor of The State University of geodesy and cartography, Honored worker of science and education of Russian Federation.

e-mail: [cvj2@mail.ru](mailto:cvj2@mail.ru)

**Аннотация.** В статье описан метод получения новых знаний на основе геореференции. Как понятие оно соотносится с науками о Земле. Как механизм получения знания геореференция относится к области структурной лингвистики и искусственного интеллекта. Раскрываются особенности этого механизма в сфере геоинформатики.

**Ключевые слова:** Геоинформатика, знание, моделирование.

**Abstract.** In paper the method of reception of new knowledge on a basis геореференции is described. As the concept it corresponds with sciences about the Earth. As the mechanism of reception of knowledge геореференция concerns area of structural linguistics and an artificial intellect. Features of this mechanism in geoinformatics sphere reveal. #

**Keywords:** Geoinformatics, knowledge, modelling.

Геореференция – термин, который широко применяется за рубежом [1, 2] при организации поиска пространственных объектов или при получении знаний. Формальное определение геореференции следующее.

*Геореференция* – соотнесение информации с пространственным (географическим) фактором. Геореференция используется в двух аспектах: пространственном и лингвистическом. Пространственный аспект связан с одной из главных задач геоинформатики [3] — поиском и использованием пространственных отношений. Лингвистический аспект, переходящий в настоящее время в онтологический аспект, связан с соотношением понятий и терминов; с извлечением и получением знаний. За рубежом главным образом рассматрива-

ют лингвистический аспект построения и применения геореференции.

На самом деле референция, или информация, построенная на ссылках, используется очень часто в реальной практической деятельности. Например, Декартова система координат определяется (в зарубежных источниках) как: «Система для локализации точки на основе отношения (референции *by reference*) ее расстояния от осей, пересекающихся под прямым углом, часто представляемых как сетка на карте» [1]

При лингвистическом подходе геореференция рассматривается как частный случай общего понятия референции. Референция (англ. *Referens* - относящийся, сопоставляющий) определяется как отнесённость включённых в текст или документ имён, именных групп или их эквивалентов к объ-

ектам действительности (референтам, денотатам). В первичном понимании именно географические названия и отношения между ними определяют геореференцию в этом аспекте. По мере возникновения и развития геоинформатики в значение термина «геореференция» менялось.

В данной статье под термином «геореференция» будем понимать пространственно-лингвистический аспект трактовки этого понятия.

В отличие от лингвистики, в которой референция рассматривается как имя или термин, в геоинформатике геореференция может представлять выражение. В тоже время лингвистические свойства в новом понятии геореференции сохраняются.

В лингвистическом аспекте геореференция должна отвечать трем принципам: замещения, предметности и однозначности.

*Принцип замещения* заключается в том, что выражения, обозначающие один и тот же объект, могут взаимозаменяться. Например, имена «Москва» и «Столица России» могут взаимозаменяться. Соответственно, если любые имена  $x$  и  $y$  имеют одинаковый денотат, то утверждения  $A(x)$  и  $A(y)$  также имеют одинаковые значения при разных именах  $X$  и  $Y$ .

*Принцип предметности* заключается в том, что сложное выражение, или имя, должно выражать связи между предметами, а не между именами, составляющими сложное имя. Например, геореференция «кадастровый номер» отражает связи между элементами кадастровой системы, а не между названиями этих элементов

*Принцип однозначности* заключается в том, что выражение, или имя, должно обозначать только один предмет. Например, геореференция «почтовый адрес» отражает определенный объект или субъект.

Эти принципы определяют построение геореференции как выражения и определенной информационной единицы. Именно в геоинформатике появились семантические информационные единицы [4].

Главным при построении использовании обычной референции являются отношения. При построении геореференции используют как обычные,

так и пространственные отношения. Большая часть пространственных отношений ориентирована на пространственные топологические отношения и ассоциации, такие, как *административная иерархия места*, что ярко выражено в кадастровой иерархии.

Механизмы геореференции и референции отличаются друг от друга. В осуществлении обеих участвуют: фонд автономных единиц и их актуализаторы. К актуализаторам, оформляющим именную группу, относятся (речь идет не только о русском языке): артикли, числительные, неопределенные, отрицательные, притяжательные и указательные прилагательные. Различие в том, что к фонду автономных единиц геореференции относятся: имена собственные (но не нарицательные) и именные группы; личные, неопределенные, указательные местоимения.

Отсюда геореференция чаще всего предстает как идентифицирующая геореференция. то есть указывающая на событие, явление, процесс или геообъект. Идентифицирующая геореференция опирается на три вида отношений: *указание*, *именование* и *обозначение*. Эти три вида отношений задают три способа представления геореференции как инструмента получения знания.

*Обозначение* — отношение, которое применяют для представления геореференции в ситуации явного описания геообъекта. Например, «Столица России». Оно связано с однозначным описанием, процедурами идентификации и идентификаторами

*Именование* — отношение, которое применяют для представления геореференции при неявном описании геообъекта. Например, «Город - герой». Оно связано с ссылками и предикацией.

*Указание* — отношение, которое применяют для представления геореференции при отсутствии описания геообъекта и оно обеспечивает референцию к объектам, находящимся в каких-либо пространственных отношениях с искомым геообъектом. Например, «Северо-Запад». Оно связано с многозначным описанием и процедурами ограничения или условиями.

Исследования в данной области проводились за рубежом, в частности, они описаны в работах [2], [5], где сформулирована основная задача процесса определения геореференций — умение соотносить название географического местоположения с геообъектом, его атрибутами или связанными геообъектами. Это дает основание построить структуру.

Рассмотрим структурную модель геореференции применительно к геоинформатике. Геореференция может быть описана графом  $G(X, U)$ , который представляет собой пару множеств  $X, U$ . Здесь  $X$  - множество вершин (терминов, понятий),  $U$  - множество ребер (отношений, связей, дефиниций).

В геоинформатике, в отличие от лингвистики, геореференция опирается на три

характерные группы данных: место время тема. Отношения порядка формируют структуру графовой модели. Если последовательно брать отношения четкого порядка, то это будет структурная модель простой «линейной» геореференции. На рис.1 приведен пример простой линейной геореференции «Почтовый адрес».

Здесь  $O_1$  - страна,  $O_2$  – регион,  $O_3$  – город,  $O_4$  – улица,  $O_5$  – дом,  $O_6$  – квартира (офис),  $O_7$  – субъект или объект,  $r$  – простое отношение (референция) «содержит».

Особенностью схемы является то, что отношение  $r$  повторяется между разными вершинами модели. В общем случае между каждой вершиной может быть разное отношение.

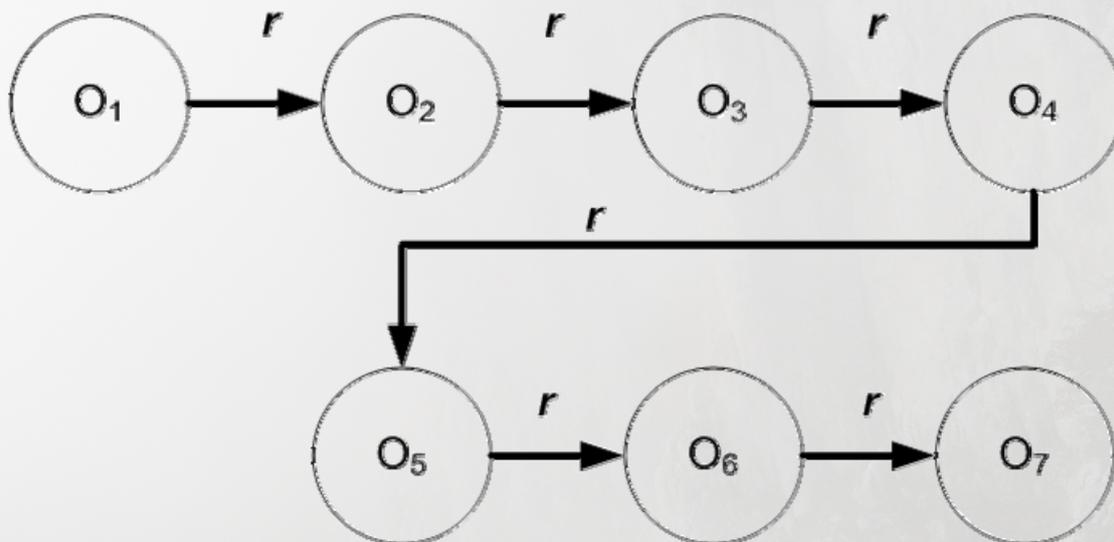


Рис.1. Линейная геореференция

По существу такая модель является обобщенным описанием уровней иерархической модели. Такая линейная геореференция ( $LGR$ ) может быть записана в общем виде как

$$LGR = (O_1, r_1; O_2, r_2; \dots O_n, r_n)$$

В этом выражении отношения  $r_i$  различны, что отражают разные индексы.

На практике редко встречаются отношения строгого порядка из-за нечеткости реальной информации, а главное, из-за многозначности описания. Это означает, что информационные элементы в  $LGR$  могут

повторяться, а связи - дублироваться. Если существует неоднозначность в отношениях, то структурная модель представляется как мультиграф. Такая структурная модель геореференции является усложнением линейной. Назовем ее «сложной» геореференцией. Сложная геореференция включает, как минимум, три связанных цепочки «место» (M), «время» (B) и «тема» (T) рис.2. Сам случай, или прецедент (Pr) является основой геореференции.

Например, рассмотрим сложную геореференцию — «Авария на участке железной дороги между Новороссийском и

Сочи». Условно она представлена схемой на рис.2.

Геореференция, представленная на рисунке 2 не показывает все элементы схемы, а выделяет различие между ними. Она включает цепочку «место»: страна – регион – дорога – участок дороги-

координаты места события. На рисунке эта цепочка отображается идентификаторами «М».

Геореференция включает цепочку «время»: год- месяц- день недели - число- время события. На рисунке эта цепочка отображается идентификаторами «В».

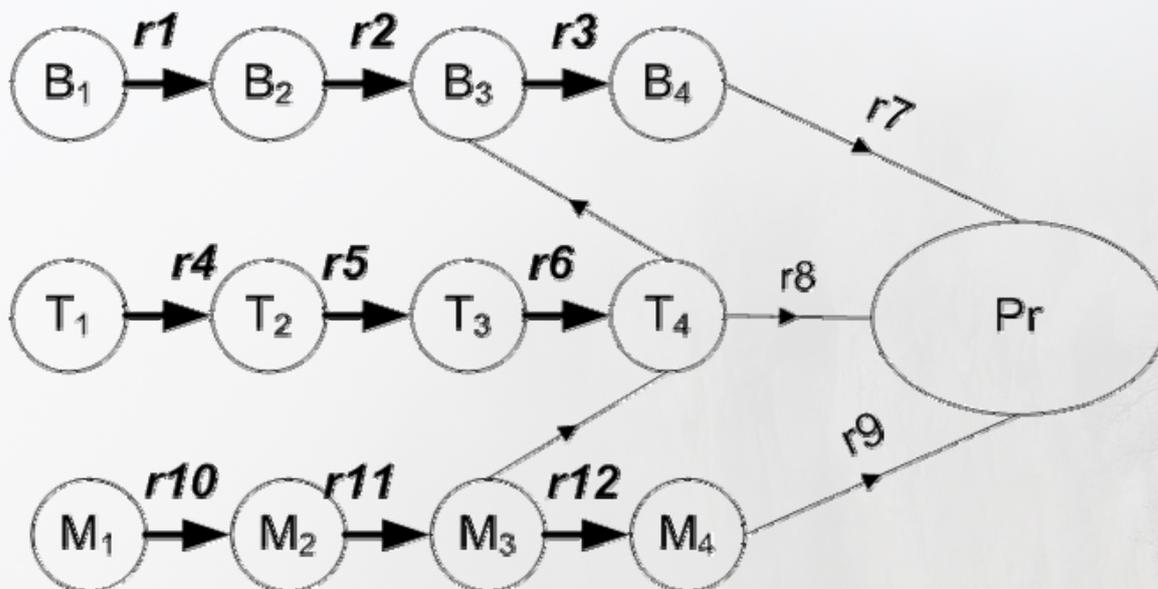


Рис.2. Структурная модель сложной геореференции

Геореференция включает цепочку «тема»: поезд- бригада- машинист- период работы машиниста- освещенность дороги – солнечная активность - сейсмическая активность и т.д. На рисунке эта цепочка отображается идентификаторами «Т». Все цепочки объединяет применяемый в теории объектно-ориентированного проектирования «прецедент» (*Pr*).

В нашем случае прецедент это авария. Он связан с возможными причинами, климатические условия, сейсмические условия, человеческий фактор, террористический акт и т.д.

Таким образом, структурно сложная геореференция представляет собой **модель события**, в которой связаны все факторы, имеющие отношение к этому событию. Доминирующей цепочкой в сложной геореференции является цепочка «место».

Рассмотрев подобные геореференции в других точках земного шара, например — «Авария на участке железной дороги в Южной Африке», «Авария на участке

железной дороги в Чили», мы можем сопоставить эти прецеденты (*Pr*). Сопоставление прецедентов по атрибутам позволяет выявить скрытые общие причины аварий или влияющие факторы, например:

машинист находился длительное время без отдыха,  
освещенность была ниже определенного уровня или  
активность на солнце имела высокое значение  
охрана участка железной дороги была недостаточна и пр.

Такие сложные геореференции дают *знание* о скрытых факторах, которые при анализе единичного случая не выявляются. Это знание помогает устранять причины негативных или новых факторов и по новому организовывать деятельность, с учетом выявленных скрытых факторов.

Аналогичные геореференции можно создавать в сфере других исследований, особенно в тех случаях, когда имеет место не стереотипная ситуация. Анализ

геореференций позволяет находить скрытые факторы, объясняющие не стереотипность, и накапливать положительный опыт в решении новых задач и получать новые знания. С позиций искусственного интеллекта важным следует считать вывод о том, что построение структурных моделей геореференции [6] позволяет получать новое знание.

### Литература

1. Heywood I., Cornelius S., Carver St. An introduction to Geographical Information Systems / Third Edition/ - Pearson Education Limited, 2006 – 426 p.
2. Hill Linda L. Georeferencing: The Geographic Associations of Information – MIT Press Cambridge, Massachusetts, London, England- 2009, - 272 p.
3. Геодезия, картография, геоинформатика и кадастр /Энциклопедия. Т.І – М.: Геодезиздат. 2008. -496 с.
4. Цветков В. Я. Семантика геоинформационных единиц / Материалы 4-й Международной конференции «Геопространственные технологии и сферы их применения» Москва 12-13 марта - М.: Информационное агентство «Гром», 2008 - с.69-71
5. Gangeni A., Pisanelli D.M., Steve G. An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies // Data & Knowledge Engineering, 1999.V.31. Pp. 183-220
6. Савиных В.П., Цветков В.Я. Развитие методов искусственного интеллекта в геоинформатике // Транспорт Российской Федерации. – 2010. –№ 5. – с.41-43.

(с) Цветков В.Я., 2011