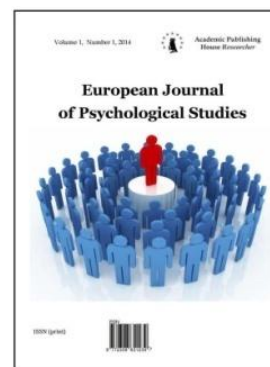


Copyright © 2015 by Academic Publishing House *Researcher*



Published in the Russian Federation  
European Journal of Psychological Studies  
Has been issued since 2014.  
ISSN: 2312-0363  
E-ISSN: 2409-3297  
Vol. 5, Is. 1, pp. 37-44, 2015

DOI: 10.13187/ejps.2015.5-37  
[www.ejournal12.com](http://www.ejournal12.com)



UDC 159.09.7

## Cognitive Science of Information Retrieval

Victor Ya. Tsvetkov

Moscow State Institute of Electronics and Mathematics, National Research University "Higher School of Economics", Russian Federation  
Doctor of Technical Sciences, Professor  
E-mail: cvj2@mail.ru

### Abstract

This article describes the process of information retrieval as cognitological procedure. For information retrieval used auxiliary concepts: information design, information environment, information relations. This article describes the basic concepts of search: concept and pattern. The symbolic modeling as a procedure for constructing patterns is presented. Also, the philosopher cognitive science is given.

**Keywords:** psychology; cognitive science; linguistics; cognitive linguistics; and information relations.

### Введение

Когнитологию часто рассматривают как метод нового понимания теории информации. Близкое к ней понятие, которое рассматривают как синоним – когнитивистика. Ее трактуют как «cognitive science [1] — когнитивистика, наука о мышлении» или «наука, изучающая и моделирующая принципы организации и работы естественных и искусственных интеллектуальных систем» [2]. Когнитологию связывают с психологией и анализом знания [3], но более точно, по мнению автора, связывать ее с анализом познания [4] и обеспечением развития знания. Когнитология связана с понятиями неявного знания [5, 6]. В самом широком смысле когнитологию можно связать с познанием, философией, психологией, информатикой, семантикой, передачей информации, лингвистикой.

Аспект философского рассмотрения когнитологии позволяет говорить о философии когнитологии. Можно отметить две проблемы философии когнитологии: философские проблемы, сформулированные в терминах когнитологии и философии; проблемы когнитологии, объектом которых является деятельность по изучению когнитологии как науки. Можно выделить две группы проблем, связанных с когнитологией: собственно проблемы когнитологии, философские проблемы когнитологии.

Информационные проблемы когнитологии связаны с исследованием формализованных информационных языков [7, 8, 9], формальных информационных конструкций [10], информационных единиц [11, 12], их интерпретаций [13] и информационным моделированием.

Когнитологию разделяют на прикладную и теоретическую. Прикладная когнитология исследует предметные области в аспекте решения практических задач. Она тесно связана с

предметным миром. Прикладная когнитология исследует практические вопросы преобразования неявного знания в явное.

Теоретическая когнитология занимается исследованием корректных процедур анализа и развитием теоретических положений, которые служат основой для прикладной когнитологии. Теоретическая когнитология исследует теоретические вопросы преобразования неявного знания в явное, в частности занимается построением моделей таких переходов.

Достаточно много исследований связано с применением когнитологии в сфере управления. В данной работе остановимся на применении когнитологии в области информационного поиска.

**Информационный поиск.** Теория в области информационного поиска тесно связана с когнитивной лингвистикой. Когнитивная лингвистика как самостоятельная область современной лингвистической науки, выделилась из когнитивной науки [14]. Когнитивная лингвистика исследует сознание на материале языка. Она исследует когнитивные процессы и анализирует типы ментальных репрезентаций в сознании человека на основе применения к языку имеющихся в распоряжении лингвистики собственно лингвистических методов анализа с последующей когнитивной интерпретацией результатов исследования. В области информационного поиска существует свой язык, что делает возможным и удобным применение методов когнитивной лингвистики для анализа этой области.

В процессе информационного поиска субъект создает информационные конструкции, для которых существует информационное окружение. При поиске субъект использует информационные отношения между образами в информационном поле. На основе созданных информационных конструкций выполняется поиск с использованием информационных отношений и получается результат поиска.

*Информационная конструкция* — это информационно определенный обобщенный образ отражения: реальных объектов, ситуаций, явлений или процессов [10]. Обязательным для модели "информационная конструкция" является структура и элементы структуры. Это сближает информационную конструкцию как модель со сложной системой. Данное условие позволяет применять методы системного анализа для исследования этой обобщенной информационной модели и применять методы системного анализа при информационном поиске и в когнитологии.

Отражением внешней среды, в которую погружены объекты и процессы, является модель «информационной среды». *Информационная среда* — это информационно определенный обобщенный образ отражения части пространства без информационных конструкций. Целью информационного описания является информационное воспроизведение свойств отражаемого реального пространства или объекта в информационную среду и информационную конструкцию.

*Информационные отношения* — отношения, которые являются отражением и информационным описанием реально существующих отношений (информационные отношения отражения) между объектами внешнего мира, а также дополнительными отношениями (дополнительные или когнитивные информационные отношения) в искусственном информационном пространстве, которое создает человек. Это искусственное информационное пространство включает когнитивную область, то есть отражает когнитивные и психологические процессы взаимодействия человека с поисковыми системами и поисковыми образами. Поэтому в широком смысле информационные отношения описывают психологические процессы тоже.

В узком смысле информационные отношения (реального мира) можно рассматривать как специфическое отображение объективно существующих отношений реального мира. Однако это отображение несет элемент субъективности и ограниченности поскольку:

- осуществляется в рамках доступных для исследования человеку;
- осуществляется инструментами, которые может применить человек;
- осуществляется в соответствии с методиками и методами, которые изобрел человек;
- оно интерпретируется человеком в соответствии с его точкой зрения, навыками интерпретации и личными пристрастиями.

Это создает неопределенность и неоднозначность поиска информации. Когнитивные информационные отношения снимают неопределенность, содержащуюся в информационных конструкциях и реальных отношениях за счет включения психологического и когнитивных механизмов.

**Паттерн и концепт.** Основными понятиями в области информационного поиска являются паттерн и концепт. Концепты, образующие концептосферу [14], по отдельным своим признакам вступают в системные отношения сходства, различия и иерархии с другими концептами. Эта общность задает область поиска.

Для концептов существуют такие лингвокогнитивные понятия, как: номинативное поле концепта, рекуррентность концепта, лексикографическое и психолингвистическое значение, когнитивные и перцептивные образы, зоны интерпретационного поля, когнитивный классификационный признак и когнитивный дифференциальный признак [14].

Концептосфера человека как упорядоченная совокупность ментальных единиц, отражающих познannую субъектом действительность составляет информационную базу сознания и мышления человека. Концепт – принадлежность сознания человека, может быть рассмотрен как информационная единица мыслительной деятельности. Концепт – может быть рассмотрен как единица семантического пространства языка. При поиске концепт выражается языковыми средствами поисковой системы.

Паттерн есть часть концепта как информационной единицы, закрепленная символически в целях коммуникации и поиска. Паттерн – форма, образец, который субъект создает при организации поиска. Отличие паттерна от поискового слова состоит в повторяемости паттерна. Поисковое слово разовый инструмент. Паттерн – инструмент обобщения и накопления опыта. Связь концепт – паттерн показана на рис. 1.

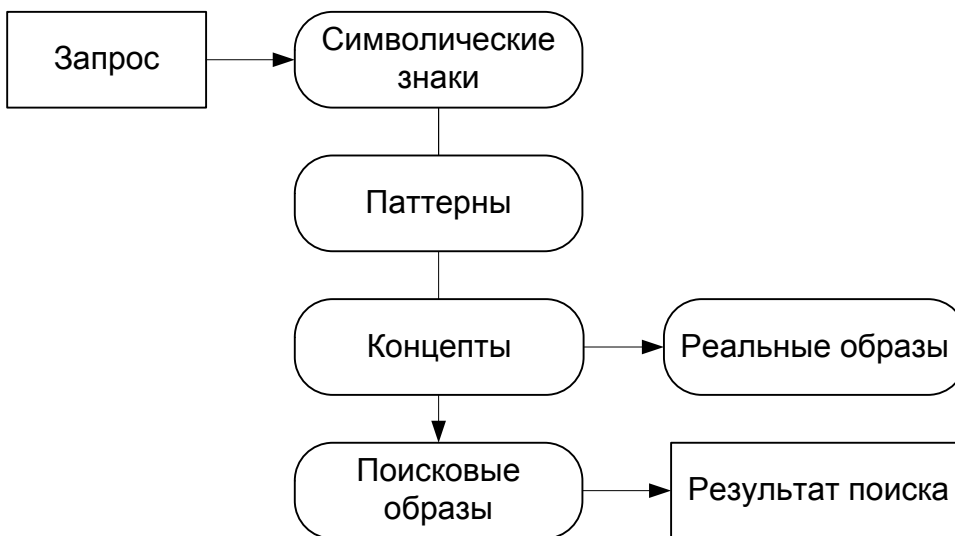


Рис. 1. Концепт, паттерн и результат поиска

Эффективность поиска оценивается по степени соответствия между реальными образами и результатом поиска (рис.1). Поисковые образы строятся как комбинации паттернов на основе связанных с ними концептов. концепт задает содержание, паттерн задает форму.

Задачи, которые необходимо решить при информационном поиске включают логический анализ выборки, учитывающий конкретные особенности концепта, и статистическое кодирование паттерна. Объектами прикладного исследования паттерна являются длина слова и его различимость.

Особенность информационного поиска в человеко-машинной процедуре состоит в том, что человек задает паттерн включая смысловую составляющую, а компьютер ее исключает и работает с совокупностью символов. Это обстоятельство должен учитывать субъект при

формировании поискового запроса и конструировании паттерна. Поэтому процедуру перехода от смысла к символу решает человек и это является психологической задачей перехода. Для отображения концепта могут быть использованы множество паттернов близких по смысловому значению.

**Естественное мышление и логический анализ.** Проблема организации поиска и выбора паттернов связана с философией логики. Философия логики [15] включает разделение логических процедур, совершаемых при помощи формализованных языков, и естественного мышления, не пользующегося формализованными языками.

Применение естественного мышления для построения паттернов возможно с помощью перечисления их феноменологических характеристик, либо путь построения их психологических моделей. Феноменологические характеристики мышления могут служить эмпирическим базисом при проверке логических моделей на адекватность.

Например, на основании феноменологических характеристик может строиться логическая модель мышления, из которой выводятся «эмпирические» следствия. Эти следствия сравниваются с заранее определенными феноменологическими характеристиками мышления. Степень толерантности будет служить оценкой адекватности предложенной модели. «Эмпирические» феноменологические характеристики мышления всегда предположительны. Они являются обобщением опыта исследования мышления, а потому зависят от принимаемых теоретических предпосылок, от тех стереотипных психологических моделей, в рамках которых они интерпретируются. Когда речь идет о естественном мышлении, особое значение приобретает когнитивность выделяемых характеристик. В понятие естественного мышления включаются стереотипы, которые встречаются в практике, науке и других областях человеческой деятельности. Иначе говоря, естественное мышление рассматривается как обычный «стереотипный» процесс

**Особенности паттерна информационного поиска.** Любой паттерн информационного поиска представляет собой формализованное описание решения часто встречающейся задачи поиска. Кроме того, паттерн информационного поиска не обязательно имеет общеупотребительное понятие. И наоборот, редкость и необычность паттерна является основой успешного поиска компьютером необходимой информации. Правильно сформулированный паттерн информационного поиска позволяет, отыскав однажды удачное решение, пользоваться им снова.

Для модели информационного поиска главным критерием является уникальность и различимость поискового образа. Область применимости паттерна может принадлежать совокупности областей, в которых исследуются объекты.

Важным начальным этапом при работе с паттернами является адекватное моделирование концептов в рассматриваемой предметной области. Рассмотрим некоторые общие принципы формирования паттернов информационного поиска объектов. В таблице 1 приведены проблемы и их решения с помощью паттернов.

Таблица 1. Решение проблем поиска объектов с помощью паттернов информационного поиска

<b>Проблема</b>	Найти группы атомарных объектов, реализующих одну функцию
<b>Решение</b>	Создать паттерн с доминирующим признаком функция и второстепенным признаком атомарный объект
<b>Проблема</b>	Найти группы атомарных объектов, реализующих одну стратегию
<b>Решение</b>	Создать паттерн с доминирующим признаком стратегия и второстепенным признаком атомарный объект
<b>Проблема</b>	Найти абстракцию по реализации.

<b>Решение</b>	Найти более высокий класс абстракции по реализации и сформировать паттерн высокого класса по реализации.
<b>Проблема</b>	Найти атрибут множества объектов
<b>Решение</b>	Создать промежуточный атрибутивный объект - паттерн атрибут и осуществить поиск по этому паттерну

Следует отметить, что при информационном поиске возникает задача не только нахождения некоего объекта (слова), но и системы взаимосвязанных объектов (совокупность связанных слов – предложение). В таблице 2 приведены паттерны к системе.

Таблица 2. Решение проблем поиска систем объектов с помощью паттернов информационного поиска

<b>Проблема</b>	Найти части системы, чтобы изменение одних ее элементов не влияло на другие?
<b>Решение</b>	Идентифицировать точки общих влияний и определить это как общность множеств и сформировать паттерн по принципу "не фактор".
<b>Проблема</b>	Найти части системы, исключающие другие свойства.
<b>Решение</b>	Создать паттерн ограничение (П), включающий атрибуты и ли сущности по принципу "только П" и включить этот паттерн в сложный
<b>Проблема</b>	Найти взаимодействие разных объектов в системе.
<b>Решение</b>	Создать процессуальный объект - паттерн адаптер, включающий общие атрибуты взаимодействующих объектов и направить поиск по этому паттерну - адаптеру.
<b>Проблема</b>	Найти связанную систему объектов на множестве отношений
<b>Решение</b>	Создать механизм построения паттерна, используя критерий максимальной различимости.

Решение последней проблемы говорит о том, что паттерном может быть не только конструкция, но механизм его построения.

**Символьное моделирование.** Создание паттерна связано с символьным моделированием или S-моделями [16]. S-модели – это вид символьных моделей, существование которых возможно только в соответствующей информационной среде (s-среда). Символьная модель - это частный случай информационной модели [17], заданная множеством символьных конструктивных объектов (s-объектов) и системой правил, определяющих их размещение в конструкции и значения атрибутов. Для символьных моделей s-среда это окружение моделей в котором они существуют и не существуют вне его [18]. С позиций семантических информационных единиц символьная модель является моделью низшего уровня – моделью символов.

Модель информационного поиска (model of information retrieval - MIR) является развитием символьной модели и с позиций семантических информационных единиц является моделью более высокого уровня – моделью слов. Модели информационного поиска делятся на дескриптивные и прескриптивные [19]. Дескриптивные делятся на субстанциональные, атрибутивные и процедурные. Прескриптивные модели могут быть

только процедурными. По аналогии MIR модели – это вид моделей, существование которых возможно только в среде языка информационного поиска.

Человек должен выбрать паттерн обладающий максимальной символической различимостью. Компьютер анализирует сходство и различие символических конструкций, а не их смысл. Таким образом задача информационного поиска приводит к необходимости символического моделирования, в котором используются критерии символической различимости и сохранения смыслового содержания. Значительная часть символического моделирования осуществляется в когнитивной области субъекта [20]. Поддержкой когнитивного моделирования являются различные формальные методы снижающие информационную нагрузку на субъект

Рассмотрим поиск системы объектов на примере электронной библиотеки "Т-Library". Рассмотрим название публикации: "Иванов В.В. Моделирование научных исследований в автоматизации и проектировании – М. ГКНТ, 1991. Мы хотим найти ссылки на источники, которые цитируют данную публикацию. То есть мы хотим найти систему объектов, связанную с известным объектом.

Во многих поисковых системах имеются несколько полей, которые заполняются по степени информированности субъекта об объекте поиска. В нашем случае имеются поля для ФИО, год публикации и тема или слово. Паттерны ограничения в нашем случае: ФАМИЛИЯ ИНИЦИАЛЫ, ГОД.

Задавая ФИО мы исключаем (ограничиваем) других авторов подобных работ. Задавая ГОД, мы исключаем публикации, написанные в другие периоды.

Вместо названия используют только одно слово, так как программа считает пробелы и несоответствие пробелов в запросе пробелам в оригинале не даст нужный результат. Это психологический процесс. Поэтому необходимо выбрать одно слово. Это логическая и психологическая задача для субъекта. Выбор слова направлен на поиск его уникальности и неповторяемости. При этом в русском языке появляется дополнительная возможность форма падежного отношения. Слова в предложении имеют не каноническую сигнификативную форму, а форму выражающую падежное отношение. Это повышает неповторяемость слова. Поэтому целесообразно выбирать именно такую форму. В рассматриваемом примере, когда субъект выбрал паттерн "моделирование" он получил 189 ссылок, среди которых большинство были ссылки на другие работы.

При выборе паттерна "исследований" было получено 54 ссылки включающие другие работы. При выборе паттерна "проектировании" было получено 26 ссылок на искомую работу. Поиск окончен.

Таким образом задача информационного поиска лежит в области когнитологии. Использование паттернов и концептов позволяет обобщать теорию информационного поиска и накапливать опыт в этом направлении.

### **Заключение**

Когнитология информационного поиска является специфическим разделом когнитологии, в котором человек "вынужден моделировать восприятие" машины или интеллектуальной системы. Человек вынужден моделировать символическое восприятие вводя в символическую модель семантическое содержание. Когнитология позволяет за счет группировки фактов и феноменов, за счет обобщения их по различным характеристикам, за счет включения когнитивной области знания в формальную область обозначений – находить связи, структуры и функции информационных конструкций. Она позволяет уменьшать информационную неопределенность, создаваемую формальными моделями, за счет когнитивных моделей и процессов. В информационном поиске этот подход позволяет "думать за компьютер", что повышает эффективность поиска и расширяет его границы. Механизм паттернов и концептов служит основой такого поиска.

### **Примечания<sup>^</sup>**

1. Johnson-Laird P. N. Mental models in cognitive science //Cognitive science. 1980. Т. 4. № 1. p.71-115.

2. Борковский А.Б. Англо-русский словарь по программированию и информатике. М.: Русский язык, 1987. 335 с.
3. Stich S.P. From folk psychology to cognitive science: The case against belief. – the MIT press, 1983.
4. Новикова Т.В. Системная когнитология как способ познания // Вестник Томского государственного университета. Философия. Социология. Политология. 2010. № 2.
5. Цветков В.Я. Неявное знание и его разновидности // Вестник Мордовского университета. 2014. Т. 24. № 3. с. 199-205.
6. Kimble, C. Knowledge management, codification and tacit knowledge // Information Research. 2013. 18(2). P. 577.
7. Кузнецов А.М. Когнитология, «антропоцентризм», «языковая картина мира» и проблемы исследования лексической семантики // Этнокультурная специфика речевой деятельности. 2000. С. 8-18.
8. Цветков В.Я. Язык информатики // Успехи современного естествознания. 2014. №7. с.129-133.
9. Дридзе Т.М. Язык информации и язык реципиента как факторы информированности / Речевое воздействие. Проблемы прикладной психолингвистики. М.: Наука. 1972. С. 34-80.
10. Tsvetkov V. Ya. Information Constructions // European Journal of Technology and Design. 2014. Vol.(5), № 3. p. 147-152.
11. Tsvetkov V. Ya. Information Units as the Elements of Complex Models // Nanotechnology Research and Practice. 2014. Vol.(1), № 1. p. 57-64.
12. Du Preez P. Units of information in the acquisition of language // Language and Speech. 1974. Т. 17. №. 4. p. 369-376.
13. Reinhart T. Anaphora and semantic interpretation. – London: Croom Helm, 1983.
14. Попова З.Д., Стернин И.А. Когнитивная лингвистика. М.: АСТ: «Восток-Запад», 2007. 227 с.
15. Брюшинкин В.Н. Логика, мышление, информация. Л.: Изд-во Ленингр. ун-та, 1988. 152 с.
16. Ильин А.В., Ильин В.Д. S-моделирование объектов информатизации. М.: ИПИ РАН, 2010. 412 с.
17. Цветков В.Я. Модели и моделирование. М.: Госинформобр, 2006. 94 с.
18. Ильин В.Д., Соколов И. А. Символьное моделирование в человеко-машинной среде: основы концепции // Информационные технологии и вычислительные системы. 2008. № 1. С. 51-60.
19. Цветков В.Я., Воинов А.И. Управление и антропоэнтропия // Современные наукоёмкие технологии. 2008. №5. с. 39-41.
20. Баксанский О.Е., Кучерлена А.И. Образ мира: когнитивный подход. М.: Альтекс, 2000. 107 с.

### References

1. Johnson-Laird P. N. Mental models in cognitive science // Cognitive science. 1980. Т. 4. № 1. p. 71-115.
2. Borkovskii A.B. Anglo-russkii slovar' po programmirovaniyu i informatike. М.: Russkii yazyk, 1987. 335 s.
3. Stich S.P. From folk psychology to cognitive science: The case against belief. – the MIT press, 1983.
4. Novikova T.V. Sistemnaya kognitologiya kak sposob poznaniya // Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya. 2010. № 2.
5. Tsvetkov V.Ya. Neyavnoe znanie i ego raznovidnosti // Vestnik Mordovskogo universiteta. 2014. Т. 24. № 3. s. 199-205.
6. Kimble, C. Knowledge management, codification and tacit knowledge // Information Research. 2013. 18(2). P. 577.

7. Kuznetsov A.M. Kognitologiya, «antropotsentrizm», «yazykovaya kartina mira» i problemy issledovaniya leksicheskoi semantiki //Etnokul'turnaya spetsifika rechevoi deyatel'nosti. 2000. S. 8-18.
8. Tsvetkov V.Ya. Yazyk informatiki // Uspekhi sovremennogo estestvoznaniya. 2014. №7. s.129-133.
9. Dridze T.M. Yazyk informatsii i yazyk retsipienta kak faktory informirovannosti /Rechevoe vozdeistvie. Problemy prikladnoi psikholingvistiki. M.: Nauka. 1972. S. 34-80.
10. Tsvetkov V. Ya. Information Constructions // European Journal of Technology and Design. 2014. Vol.(5), № 3. p. 147-152.
11. Tsvetkov V. Ya. Information Units as the Elements of Complex Models // Nanotechnology Research and Practice. 2014. Vol.(1), № 1. r. 57-64.
12. Du Preez P. Units of information in the acquisition of language //Language and Speech. 1974. T. 17. №. 4. p. 369-376.
13. Reinhart T. Anaphora and semantic interpretation. – London: Croom Helm, 1983.
14. Popova Z.D., Sternin I.A. Kognitivnaya lingvistika. M.: AST: «Vostok-Zapad», 2007. 227 s.
15. Bryushinkin V.N. Logika, myshlenie, informatsiya. L.: Izd-vo Leningr. un-ta, 1988. 152 s.
16. Il'in A.V., Il'in V.D. S-modelirovanie ob"ektov informatizatsii. M.: IPI RAN, 2010. 412 s.
17. Tsvetkov V.Ya. Modeli i modelirovanie. M.: Gosinformobr, 2006. 94 s.
18. Il'in V.D., Sokolov I. A. Simvol'noe modelirovanie v cheloveko-mashinnoi srede: osnovy kontseptsii //Informatsionnye tekhnologii i vychislitel'nye sistemy. 2008. № 1. S. 51-60.
19. Tsvetkov V.Ya., Voinov A.I. Upravlenie i antropoentropiya // Sovremennye naukoemkie tekhnologii. 2008. №5. s. 39-41.
20. Baksanskii O.E., Kucherlena A.I. Obraz mira: kognitivnyi podkhod. M.: Al'teks, 2000. 107 s.

УДК 159.09.7

### **Когнитология информационного поиска**

Виктор Яковлевич Цветков

Московский институт электроники и математики Национального исследовательского университета «Высшая школа экономики», Российская Федерация  
Доктор технических наук, профессор  
E-mail: cvj2@mail.ru

**Аннотация.** Статья описывает процедуру информационного поиска как когнитологическую процедуру. В качестве вспомогательных понятий поиска применяют понятия: информационная конструкция, информационное окружение информационное отношение. В качестве основных понятий поиска используют понятия концепт и паттерн. Рассмотрено символическое моделирование как процедура построения паттернов. Рассмотрена философия когнитологии.

**Ключевые слова:** психология; когнитология; лингвистика; когнитивная лингвистика; информационные отношения.