

## МЕТОДИКА ИСПОЛЬЗОВАНИЯ СРЕДСТВ СТРУКТУРИЗАЦИИ УЧЕБНОГО МАТЕРИАЛА

Сулима Е.Н., Миленин В.М

**Аннотация.** *Описываются средства структуризации информационных массивов на основе использования растущих пирамидальных сетей. Приводятся примеры обработки информации на сайте ОСТРОВ ЗНАНИЙ.*

**Ключевые слова:** *структуризация естественно-языкового текста, представление и обработка знаний.*

**ACM Classification Keywords:** *1.2 ARTIFICIAL INTELLIGENCE - 1.2.4 Knowledge Representation Formalisms and Methods, K.3 COMPUTERS AND EDUCATION - K.3.1 Computer Uses in Education*

---

### Введение

---

Создание формальной структуры учебного материала выражается в определении системы отношений между понятиями: меронимии, гипонимии, исключения, дополнения, ассоциации, сходства, различия, близости, отдаленности, причины, следствия, до и после и т.д. Связи устанавливаются не только между понятиями, но и между предметами, явлениями. Основной вывод исследований по запоминанию учебного материала тривиален и состоит в том, что структурированный материал заучивается легче, чем бессвязный. Чем больше связей нового материала со старым, тем легче осуществляется процесс запоминания. Запоминание информации не следует понимать как процесс "раскладывания" ее по заранее заготовленным "полкам". Одновременно с восприятием информации осуществляется как структурирование информации, так и структурирование самой памяти под влиянием воспринимаемой и уже хранящейся информации. Воспринятая информация отображается в структуре памяти.

Какие же типы компьютерных информационных моделей наиболее близки к моделям, которые использует человек для осуществления своей жизнедеятельности? В информационно-аналитических системах для обработки сложных структур разнотипных данных и знаний используются логико-лингвистические модели [Поспелов, 1981], т.е. такие модели, в которых основными элементами является не числа и вычислительные операции, а имена и логические связи. Естественно-языковой текст также может быть адекватно представлен с помощью логико-лингвистической модели.

Одной из хорошо апробированных возможных реализаций логико-лингвистических моделей является организация памяти информационно-аналитических систем в виде *растущих пирамидальных сетей* (РПС) [Гладун, 2000].

---

## Построение РПС

---

Процесс построения растущей пирамидальной сети описан во многих публикациях [Гладун, 1994, Гладун, 1987]. Напомним основные понятия и процессы, происходящие при построении РПС. *Растущей пирамидальной сетью* называется ациклический ориентированный граф, в котором нет вершин, имеющих одну заходящую дугу. Примеры РПС приведены на рис.1,2,3. Вершины, не имеющие заходящих дуг, называются *рецепторами*, остальные вершины – *концепторами*. Подграф пирамидальной сети, включающий вершину  $a$  и все вершины, от которых имеются пути к вершине  $a$ , называется *пирамидой* вершины  $a$ . Вершины, входящие в пирамиду вершины  $a$ , образуют ее *субмножество*. Множество вершин, к которым имеются пути от вершины  $a$ , называется ее *супермножеством*.

В субмноестве и супермноестве вершины  $a$  выделяются 0-субмножество и 0-супермножество, состоящие из тех вершин, которые связаны с вершиной  $a$  непосредственно. При построении сети входной информацией служат наборы значений признаков, описывающих некоторые объекты произвольной предметной области (материалы, состояния агрегата, ситуации, термины учебного курса и т.п.). Рецепторы соответствуют значениям признаков. В различных задачах это могут быть имена свойств, отношений, состояний, действий, объектов или классов объектов. По мере ввода в сеть новых объектов формируются концепторы – специальные вершины, соответствующие пересечениям множеств рецепторов различных объектов, а также описаниям объектов в целом. В начальном состоянии сеть состоит только из рецепторов. Концепторы формируются в результате работы алгоритма построения сети. Приведем описание алгоритма построения сети[Гладун, 1994].

При вводе признакового описания объекта, рецепторы, соответствующие значениям признаков, входящим в описание, переводятся в *состояние возбуждения*. Процесс возбуждения распространяется по сети. Концептор переводится в состояние возбуждения, если возбуждены все вершины его 0-субмноества. Рецепторы и концепторы сохраняют состояние возбуждения в течение выполнения всех операций достройки сети.

Пусть при вводе описания объекта  $F_a$  – подмножество возбужденных вершин 0-субмноества вершины  $a$ ;  $G$  – множество возбужденных вершин сети, не имеющих других возбужденных вершин в своих супермноествах.

Ввод новых вершин производится по следующим правилам.

### Правило А1.

Если вершина  $a$  не возбуждена, и множество  $F_a$  содержит более одного элемента, то дуги, соединяющие вершины из множества  $F_a$  с вершиной  $a$ , ликвидируются, и в сеть вводится новый концептор, который соединяется заходящими дугами с вершинами множества  $F_a$  и исходящей дугой с вершиной  $a$ . Новая вершина находится в состоянии возбуждения.

Выполнение правила A1 иллюстрируется рис. 1 (а, б, в). Сеть рис. 1 (б) возникает после возбуждения в сети рис. 1 (а) рецепторов 4, 5, а сеть рис. 1 (в) возникает после дополнительного возбуждения в сети рис. 1 (б) рецепторов 2, 3.

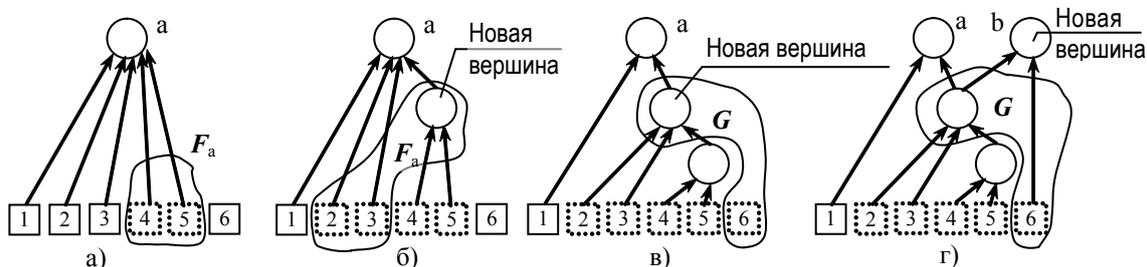


Рис. 1.

Как следует из правила A1, условием ввода в сеть новой вершины является ситуация, когда некоторая вершина сети оказывается не полностью возбужденной (возбуждаются не все, но не менее двух вершин ее 0-субмножества). Новые вершины вводятся в субмножества не полностью возбужденных вершин. Они представляют в сети пересечения описаний объектов.

После введения новых вершин во все участки, где удовлетворяется условие правила A1, выполняются правила A2 или A3, завершающие построение пирамиды объекта.

#### Правило A2.

Если множество  $G$  содержит более одного элемента и не включает вершины, помеченной именем введенного объекта, к сети присоединяется новый концептор, который соединяется заходящими дугами со всеми вершинами множества  $G$ . Новая вершина находится в возбужденном состоянии.

Выполнение правила A2 иллюстрируется рис.1 (в, г). Сеть рис.1 (г) возникает после возбуждения в сети рис.1 (в) рецепторов 2,3,4,5,6.

### **Особенности РПС**

Далее остановимся на особенностях РПС, которые позволяют эффективно их использовать для решения задач. В РПС структурирование памяти осуществляется путем построения иерархической сетевой структуры одновременно с восприятием информации. Сформированная структура отображает особенности решаемой задачи. Алгоритм построения сети обеспечивает автоматическое установление ассоциативной близости между объектами на основе общих элементов из их описаний. Таким образом, при построении сети формируются конъюнктивные классы объектов, т.е. осуществляется классификация без учителя. При добавлении новых объектов нет необходимости выполнять операции поиска и преобразования сети по всем введенным объектам – формирование концепторов ограничено небольшими фрагментами сети и происходит только в пирамидах тех объектов, рецепторы которых совпадают с рецепторами нового объекта. Принципиальное решение проблемы сокращения времени

поиска в РПС состоит в отнесении операций по установлению сходства объектов на этап ввода информации об объектах в память, который выполняется только один раз при построении сети.

РПС удобны для выполнения операций ассоциативного поиска [Гладун, 1994, Гладун, 2000]. Например, можно выбрать все объекты, которые включают заданное объединение значений признаков, прослеживая пути, выходящие из вершины сети, которая отвечает этому объединению. Для выбора всех объектов, описания которых пересекаются с описанием заданного объекта, достаточно проследить пути, которые выходят из вершин пирамиды этого объекта. Все поисковые операции пирамидальной сети ограничиваются сравнительно малым участком сети, которая включает пирамиду объекта и вершины, непосредственно связанные с ней. Переход от конвергированных представлений объектов (концепторов) к развернутым - наборам рецепторов осуществляется путем просмотра пирамид объектов в различных направлениях (сверху вниз и снизу вверх). В результате появляется возможность решать практические задачи на основе анализа больших объемов данных.

Важным свойством пирамидальных сетей является их иерархичность, которая позволяет естественным образом отображать структуру составных объектов, родовидовые связи, связи объект-свойство. Обучение РПС состоит в формировании в них структур, представляющих понятие [Гладун, 1994]. Понятие – элемент системы знаний, представляющий собой обобщенную логическую признаковую модель класса объектов, с помощью которой реализуются процессы распознавания и генерации моделей конкретных объектов. После построения пирамидальной сети создается структура, являющаяся представлением описаний объектов. Каждая вершина сети определяет конъюнктивный класс объектов. РПС позволяет построить понятие в форме логического выражения, включающего имена рецепторов и логические операции дизъюнкции, конъюнкции и отрицания [Gladun, 2008]. Построенное понятие включает все наиболее важные признаки, характеризующие класс, и отображает характерные для этого класса логические связи между этими признаками. Представление понятия в виде логического выражения является наглядным, хорошо интерпретируемым и может быть использовано с целью более глубокого понимания закономерностей, присущих предметной области. Важной особенностью метода формирования понятий в РПС является возможность введения в понятие так называемых разделительных признаков, которые не принадлежат объектам исследуемого класса. В результате сформированные понятия имеют более компактную логическую структуру. РПС реализованы в программной системе “Конфор” [Гладун, 2002]. Система “Конфор” не является предметно-ориентированной и может быть использована для анализа данных в любой предметной области при условии, что данные представлены как множество описаний объектов в виде набора значений признаков.

---

### **Структуризация учебного материала с использованием РПС**

---

Содержательная (семантическая) структура учебного материала предполагает смысловое, а не механическое запоминание. Такая структура отражает объективные зависимости, существующие в

природе, обществе и между людьми. При механическом запоминании материал заучивается таким, каким он предлагается, при смысловом - происходит организация и упорядочение сообщаемой информации, ее фильтрация. Современный школьник должен уметь в достаточной степени систематизировать и обобщать свои знания, знать соотношения и связи между структурными элементами системы знаний. Автоматизированная структуризация учебного материала позволяет в первую очередь в рамках электронного учебного ресурса связать контент ресурса в единую структуру, проследить связи между понятиями, используемыми в различных информационных источниках, которыми постоянно пополняется учебный ресурс.

Методика использования средств автоматизированной структуризации учебного материала включает следующие этапы:

- 1) автоматизированный поверхностный семантический анализ текста [Величко, 2009];
- 2) выделение терминов из текста на основе проведенного анализа [Величко, 2009];
- 3) выделение отношений между терминами на основе проведенного анализа;
- 4) для каждого проанализированного текста – формирование файла с выделенными терминами и отношениями между ними;
- 5) построение растущей пирамидальной сети, объединяющей термины и связывающей отдельные учебные материалы в единое информационное пространство.

Рассмотрим пример применения описанной технологии к обработке информации сайта «ОСТРОВ ЗНАНИЙ» (<http://shkola.ostriv.in.ua>). Проанализируем две статьи из раздела «Школа - Навчання - Додаткові матеріали»: «Будова Сонця» (Рис. 2), опубликована 03 Травня 2007 и «Класифікація зірок» (Рис. 3) опубликована 17 Січня 2007. Фрагменты списков терминов, автоматически построенных на основе лингвистического анализа статей приведены на Рис.4 и Рис.5. На Рис.6. приведен фрагмент построенной пирамидальной сети, объединяющей содержание двух учебных материалов. Используя сетевое представление терминологического содержания текстов легко проследить связи между терминами, связать в единое целое учебный материал, опубликованный в различное время. Графическое представление связей между понятиями улучшает качество восприятия учебного материала, позволяет определить место научного факта в обобщенной научной картине мира, обеспечивает формирование структурно-логических знаний учащихся. Навигация по сети позволяет переходить от одного факта к другому, находить места исходного текста, связанные с упоминанием того или иного факта. Автоматически сформированные термины позволяют уточнить списки ключевых слов, использующихся в описании статьи для последующего поиска по ключевым словам и рубрикации статей.

## Будова Сонця

Зовнішній шар Сонця складається з фотосфери і хромосфери. Фотосфера, що є видимою поверхнею Сонця, має товщину приблизно 500 км і температуру близько 6000°K.

У фотосфері виявляється помітна активність в першу чергу у формі темних областей, так званих сонячних плям. За ними спостерігали ще в античні часи, але, не дивлячись на це, їх справжня природа не відома.

У пізніші часи Галілео Галілей "наново відкрив" сонячні плями, хоча це "відкриття" довго оспорювалося Крістофером Шейнером. Спостереження за сонячними плямами серед іншого, привело до краху аристотелівсько-птолемєєвської моделі Всесвіту, згідно якої всі зірки є ідеально неподільними сферами. Систематичні спостереження за сонячними плямами почалися приблизно в середині XVIII століття.

Сонячні плями здаються темними, але не тому, що вони дійсно чорні. Просто вони холодніші за фотосферу, що іоточує. Навколо найтемнішої зони плями, так званої тіні, є проміжна світла зона, так звана півтінь. Температура тіні рівна 4300-4800°K, тобто на 1000-1500°K нижча за температуру фотосфери. А ось температура півтіні рівна 5400-5500°K. Інтенсивність свічення в тіні складає приблизно 32% від інтенсивності свічення фотосфери, а півтіні - 80%. Схоже, що зниження температури усередині плям зв'язане сильними магнітними полями, відкритими Джорджем Еллері Хейлом в 1908 році. Такі поля заважають регулярному конвективному

Рис.2. Фрагмент учебного материала «Будова Сонця».

## Класифікація зірок

При досягненні в надрах зірок високої густини і температури (біля 10-12 млн. K) починаються термоядерні реакції синтезу елементів - основне джерело енергії більшості зірок. Атмосфери зір - поверхневі шари зірок, у яких формується спектр спостерігаючого випромінювання зірок. В атмосфері зірок звичайно виділяють самий зовнішній шар: протяжну корону, потім хромосферу і розташовану ще глибше фотосферу. Маса зірок (M) знаходяться в межах від 0,04 до ~ 60 M<sub>☉</sub>, світності (L) від 0,5 до сотень тис. L<sub>☉</sub>.

Зірки класифікують по світності, масі, температурі поверхні, хімічному складу, особливостям спектру. На визначених етапах зіркової еволюції ряд зірок проходить через стадію нестационарності. У залежності від маси зірки наприкінці еволюції стають або **білими карликами**, або **нейтронними зірками**, або **чорними дірами**. Існують зоряні каталоги змінних зірок, подвійних зірок, пульсарів. Поділ зірок за їхньою світністю здійснюється на групи: **надгіганти**, **яскраві і слабкі гіганти**, **субгіганти**, **карлики**, **субкарлики і білі карлики**. Розходження світимості в надгігантів і білих карликів понад 20 зоряних величини (більш ніж у 108 разів). Діаграма Герцшпрунга-Рессела виражає зв'язок між світністю і температурою (спектральним класом чи показником кольору) зірок. На цій діаграмі близькі за фізичними властивостями зірки займають відособлені області: головну послідовність зірок, послідовності надгігантів, яскравих і слабких гігантів, субгігантів, субкарликів, білих карликів і ін. Головна послідовність діаграми Герцшпрунга-Рессела, вузька смуга на цій діаграмі, у межах якої знаходиться переважна більшість зірок, перетинає діаграму по діагоналі (від високих до низьких світимостей і температур). Зірки головної послідовності (до них, зокрема, відноситься Сонце) мають однакове джерело енергії термоядерної реакції кисневого циклу. Зірки знаходяться на головній послідовності протягом приблизно 90% усього часу зоряної еволюції. Цим пояснюється переважна концентрація зірок в області головної послідовності.

Рис.3. Фрагмент учебного материала «Класифікація зірок».

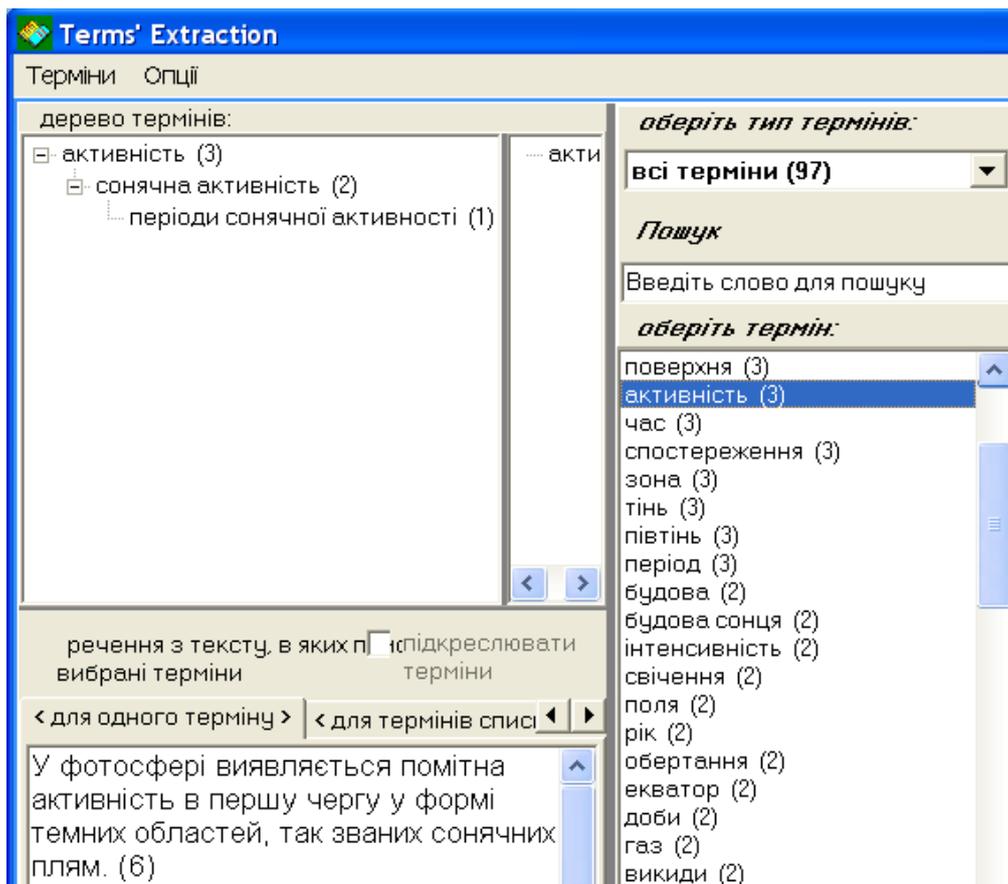


Рис.4. Фрагмент списка терминов, выделенных при анализе учебного материала «Будова Сонця».

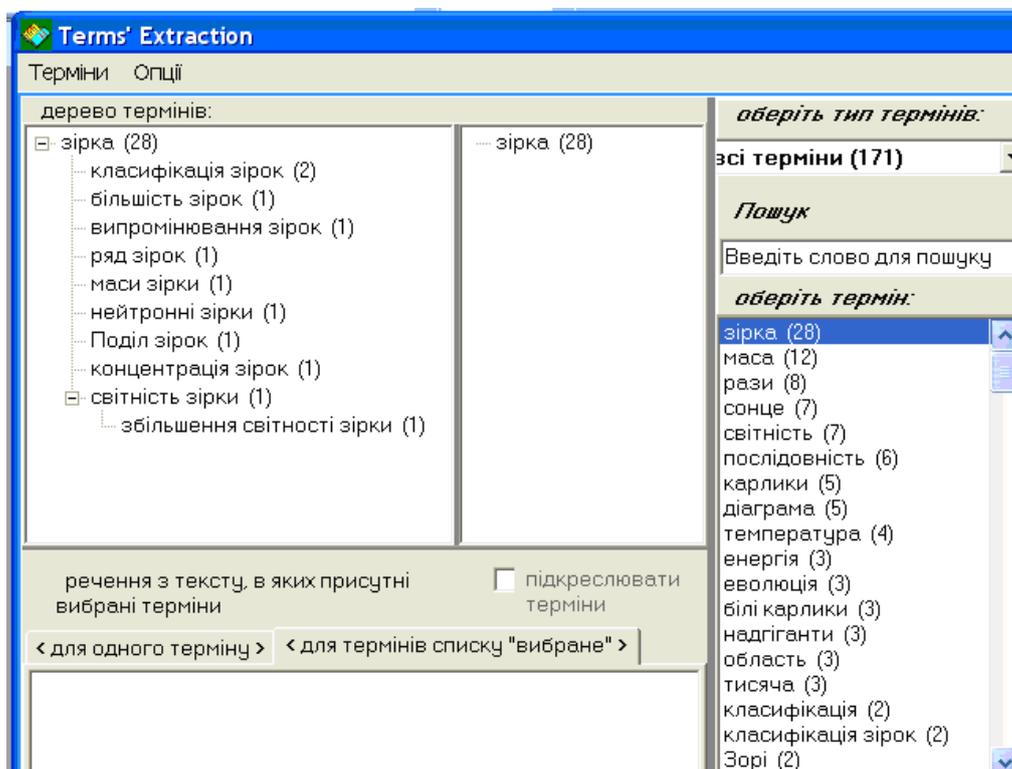


Рис.5. Фрагмент списка терминов, выделенных при анализе учебного материала «Класифікація зірок».

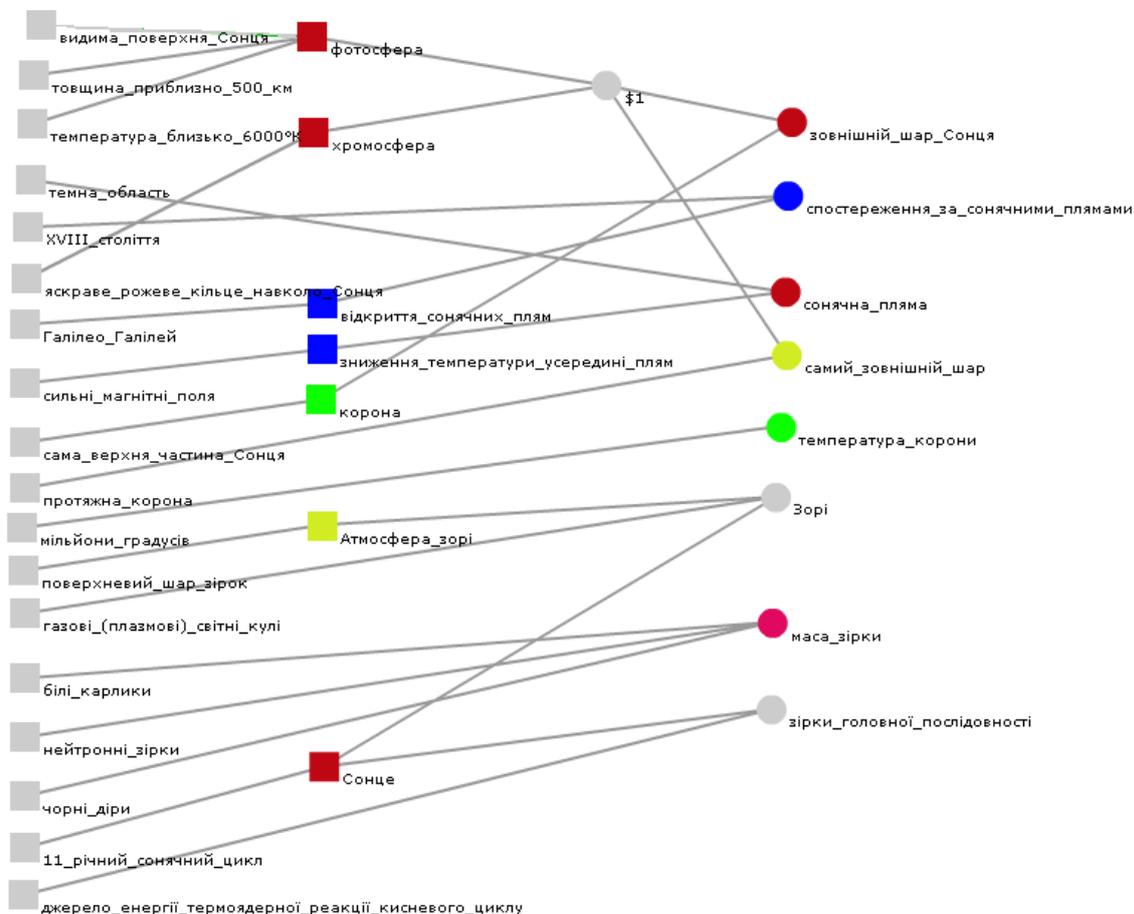


Рис.6. Фрагмент построенной пирамидальной сети, представляющей содержание двух учебных материалов

## ЗАКЛЮЧЕНИЕ

Растущая пирамидальная сеть является сетевой памятью, самонастраивающейся на структуру входной информации. Такие качества, как простота внесения изменений, совмещение процессов ввода информации с ее классификацией, обобщением и выделением существенных признаков, высокая ассоциативность, делают растущие пирамидальные сети важным компонентом интеллектуальных систем. Одной из главных задач логической структуризации учебного материала, как средства систематизации и обобщения знаний учащихся, является формирование не формальных, а осознанных знаний, позволяющих их активное применение в творчестве, решении задач анализа, синтеза, обобщения, поиска новых знаний. Современные средства автоматизированной структуризации учебного материала помогают решать задачу формирования активных знаний у учащихся, используя возможности системно-структурного метода для управления познавательной деятельностью и выделяя ориентиры для самостоятельного обновления и пополнения знаний.

## Литература

---

- [Поспелов, 1981] Поспелов Д.А. Логико-лингвистические модели в системах управления.-Москва: Энергоиздат.- 1981.
- [Гладун, 2000] Гладун В.П. Партнерство с компьютером. Человеко-машинные целеустремленные системы.-Киев: Port-Royal. – 2000.
- [Гладун, 1994] Гладун В. П. Процессы формирования новых знаний. – София: СД «Педагог 6». – 1994.
- [Гладун, 1987] Гладун В.П. Планирование решений.-Киев: Наукова думка.– 1987.
- [Gladun, 2008] Victor Gladun, Vitaly Velichko, Yurii Ivaskiv. Selfstructured Systems. International Journal "Information Theories & Applications". FOI ITHEA, Sofia. – Volume 15 – 2008, Number 1 – P. 5-13.
- [Гладун, 2002] Гладун В.П., Ващенко Н.Д., Величко В.Ю. Программный комплекс решения аналитических задач методами правдоподобного вывода // Искусственный интеллект.- 2002.-№2.- с.430-437.
- [Величко, 2009] Виталий Величко, Павел Волошин, Светлана Свитла Автоматизированное создание тезауруса терминов предметной области для локальных поисковых систем. "Knowledge – Dialogue – Solution" International Book Series "INFORMATION SCIENCE & COMPUTING", Number 15. – FOI ITHEA Sofia, Bulgaria. - 2009. – pp.24-31.