

Лекционный материал по дисциплине
(заполняется по каждому виду учебного материала)

ДИСЦИПЛИНА Системный анализ данных в СППР
(полное наименование дисциплины без сокращений)

ИНСТИТУТ ИТ

КАФЕДРА Вычислительной техники
полное наименование кафедры

ВИД УЧЕБНОГО МАТЕРИАЛА Лекция
(в соответствии с пп.1-11)

ПРЕПОДАВАТЕЛЬ Сорокин А.Б.
(фамилия, имя, отчество)

СЕМЕСТР пятый
(указать семестр обучения, учебный год)

1. ЛЕКЦИЯ. Онтологии и онтологические системы

К XX в. сложились две формы культуры — естественнонаучная и гуманитарная, различающиеся методами познания. Произошло разделение на людей, мыслящих предпочтительно формально или, напротив, — гуманитарно. Эти формы мышления связывали с различными полушариями головного мозга.

Гуманитарное познание формирует образ, целостность, а формальное мышление обеспечивает отображение элементов и законов их взаимодействия. Гуманитарное знание связано с определением смысла, назначения, целесообразности (телеология), цели. Вершиной гуманитарного знания традиционно считается философия. *Формальное мышление* традиционно базируется на математике.

В европейской культуре более предпочтительными и развитыми оказались формальные методы. В то же время в XIX—XX вв. возник ряд проблем, труднообъяснимых с помощью формальных методов. Для их решения механическая концепция, превратившаяся в физико-математическую, оказалась недостаточной. Напротив, развивающаяся диалектическая концепция объясняла объективность противоречий и даже необходимость их для существования и развития сложного мира, но не имела формализованного аппарата, которому отдает предпочтение европейская наука.

Возникли попытки распространения на гуманитарную сферу познания физико-математических методов — физикализм. С помощью физикализма были объяснены некоторые проблемы. Однако к середине XX в. стал очевидным кризис концепции физикализма, что привело к возникновению интегральных концепций, объединяющих возможности гуманитарного и формального мышления.

Формальные методы не позволяют выявить содержание исследуемых процессов, понять их целостность, хотя и могут помочь активизировать интуицию и опыт специалистов, гуманитарное мышление для выявления содержания, отобразить законы взаимодействия компонентов, полученные эмпирически.

Анализ содержания, исследование процессов постановки задач позволили обратить внимание на особую роль человека: человек является носителем целостного восприятия, сохранения целостности при расчленении проблемы, распределении работ, носителем системы ценностей, критериев принятия решения (появился термин «лицо, принимающее решение» — ЛПР).

После осознания необходимости интеграции гуманитарного и формального мышления между философией и математикой появился спектр дисциплин, которые сочетают средства гуманитарного познания, позволяющего отобразить

содержание познаваемого объекта, и формальных методов, отражающих изученные законы строения и функционирования объектов и помогающих таким образом в развитии познания в обозримые сроки

Для выработки единых обобщающих терминов, единого языка общения представителей разных наук Норберт Винер в 1934 г. собрал в Принстоне на семинар ученых многих специальностей (нейрофизиологов, инженеров-связистов, конструкторов вычислительной техники и т.д.). Для названия новой науки об общих принципах управления в живых организмах и машинах был принят термин «кибернетика». А существующее словосочетание «кибернетический подход» может рассматриваться как первая интегральная концепция естествознания, объединяющая гуманитарное и формальное знания. Однако в последующем в связи с неоднозначной трактовкой термина «кибернетика» и употреблением его во многих работах (особенно зарубежных), связанных с разработкой технических аналогов живых организмов, этот термин стал использоваться в более узком смысле — как одно из направлений теории систем, занимающееся процессами управления техническими объектами.

Для обобщения дисциплин, связанных с исследованием и проектированием сложных объектов различной природы, возникла теория систем и системный анализ.

Основоположником теории систем считают биолога Л. Фон Берталанфи, который в 30-е гг. XX в. ввел понятие открытой системы и сформулировал основные идеи и закономерности обобщающего направления, названного теорией систем.

Важный вклад в становление системных представлений внес (еще до Берталанфи) наш соотечественник А. А. Богданов (Александр Александрович Малиновский). Однако в силу исторических причин предложенная им всеобщая организационная наука — тектология (от греч. «тектон» — строитель) не нашла распространения и практического применения

В нашей стране вначале теорию систем активно развивали философы В. Г. Афанасьев, В. Н. Садовский, В. С. Тюхтин, А. И. Уёмов.

Параллельно развивались направления, родственные теории систем: исследование операций (это направление, возникшее для исследования военных операций, применялось в различных сферах, в том числе и в экономике), имитационное моделирование, ситуационное моделирование, синергетика, информационный подход.

Для обобщения дисциплин, связанных с исследованием и проектированием сложных систем, используется термин «системные исследования», а иногда

сохраняется термин «системный подход», который широко применялся в первые годы становления теории систем в двух смыслах — в смысле методологического направления философии, и в прикладном аспекте, как синоним понятия «комплексный подход».

Относительно философского представления в термине «системность» заключается два смысла:

- первый составляет отождествление системности с объективным, независимым от человека свойством действительности, что делает ее онтологическим, объективно-диалектическим свойством всего сущего;

***Онтология** — раздел философии, изучающий фундаментальные принципы устройства бытия, его начала, сущностные формы, свойства и категориальные распределения.*

- другой под системностью подразумевает накопленные людьми представления о самом свойстве, т.е. она представляет собой гносеологическое явление, некоторые знания о системах различной природы.

***Гносеология**, или учение о познании — это раздел философии, изучающий возможности познания мира человеком, структуру познавательной деятельности, формы знания в его отношении к действительности, критерии истинности и достоверности знания, его природу и границы.*

Для изучения этих двух смыслов Вам предлагается:

- Онтологическое представление изучить в рамках первой практической работы и первой лекции.

- Гносеологическое представление изучить в рамках остальных лекций.

1.1. Понятие онтологии

Возникновение онтологий и их стремительное развитие связано с проявлением в нашей реальности следующих новых факторов:

- колоссальный рост объемов информации, предъявляемых для обработки (анализа, использования) специалистам самых различных областей деятельности;

- чрезвычайная зашумленность этих потоков (повторы, противоречивость, разноуровневость, и т.п.);

- острая необходимость в использовании одних и тех же знаний разными специалистами в разных целях;

- всеобщая интернетизация нашей жизни и острая необходимость в структуризации информации для её представления пользователям и более эффективного поиска;

- необходимость сокращения времени на поиск нужной информации и повышения качества информационных услуг в Интернете.

Появление онтологий стало *ответом* ряда наук, связанных с информационными технологиями и системами искусственного интеллекта на перечисленные проблемы. Именно они обеспечили возможность их перехода на новый качественный уровень обработки и поиска информации, в том числе в *Web*-системах. Рассмотрим это понятие подробнее.

Онтология (от др.-греч. *онтос* – сущее, *логос* – учение, понятие) – термин, определяющий учение о бытии, о сущем, в отличие от *гносеологии* – учения о познании. Автор термина «онтология» является Христиан Вольф (1679–1754). Ее предмет – изучение абстрактных и общих философских категорий, таких как бытие, субстанция, причина, действие, явление и т.д., а сама онтология претендовала на полное объяснение причин всех явлений.

В информационных технологиях и интеллектуальных системах этот термин был принят для обозначения представлений о моделях мира (бытии, сущностей, причинных связей и т.п.).

Наиболее распространенными стали следующие определения.

Онтология – это точная спецификация концептуализации.

Онтология – это формальная точная спецификация совместно используемой концептуализации.

Онтологии – это базы знаний специального типа, которые могут *читаться* и *пониматься*, *отчуждаться* от разработчика *и/или* физически *разделяться* их пользователями.

В этих определениях ключевыми являются понятия: концептуализация (совместно используемая); спецификация (формальная, точная). Поясним смысл этих терминов.

Под *концептуализацией* понимается абстрактная модель явлений (процессов) в мире, составленная посредством определения существенных для описания данных явлений понятий, т.е. концептов.

Точность подразумевает, что типы используемых понятий и ограничения на область применения данных понятий явно определены.

Формальность означает, что онтология должна быть ориентирована на компьютерное представление, что исключает использование естественных языков в полной мере, в связи с их неоднозначностью и сложностью.

Совместное использование отражает понятие того, что онтология описывает всеобщие знания, т.е. не персональные знания одного человека, а знания, принятые в группе, сообществе.

Для обеспечения совместного использования вводится ещё понятие *онтологического соглашения* – объема словаря, достаточного для описания представляемых абстрактной моделью знаний, предназначенных для совместного использования. Другими словами, онтология ссылается на конкретные понятия, взятые из специального словаря, используемого для описания некоторой реальности, и на множество толкований предполагаемых значений слов, входящих в словарь.

Еще одно определение онтологии, используется в рамках технологии агентов и соответствующее спецификации *FIPA (Foundation for Intelligent Physical Agents)*.

Онтология – это подробная спецификация структуры определенной проблемной области, включающей: словарь (список) логических констант и предикатных символов для описания предметной области и набор логических высказываний, формулирующих существующие в данной проблемной области ограничения и определяющих интерпретацию словаря.

По сравнению с другими определениями, это является более формальным, представляя онтологию в виде формальной системы, в основе которой лежит словарь терминов, с отношениями между ними, задаваемыми в виде логических высказываний.

Таким образом, *онтология* – это тоже модель предметной области, использующая все доступные средства представления знаний, релевантные для данной области, но ориентированная на стандартизацию и унификацию форм представления знаний.

1.2. Формальная модель онтологии

Примем за основу, что понятие *онтологии* предполагает определение и использование взаимосвязанной и взаимосогласованной совокупности трех компонент: словаря терминов, определений терминов и правил их обработки. Тогда, формальная модель онтологии (O) может быть определена как упорядоченная тройка вида:

$$O = \langle X, R, F \rangle, \text{ где}$$

X – конечное множество концептов (понятий, терминов) предметной области, которую представляет онтология; R – конечное множество отношений между концептами (понятиями, терминами) заданной предметной области; F – конечное множество функций интерпретации (аксиоматизации), заданных на концептах и/или отношениях онтологии.

Естественным ограничением, накладываемым на множество X , является его

конечность и не пустота. Иначе дело обстоит с компонентами F и R в определении онтологии. Очевидно, что F и R должны быть конечными множествами.

Рассмотрим граничные случаи, связанные с пустотой и не пустотой F и R .

1. Пусть $R = \emptyset$ и $F = \emptyset$. Тогда онтология $O = V = \langle X, \{ \}, \{ \} \rangle$, что соответствует простому словарю (V). Очевидно, что это случай вырожденной онтологии. Она может быть полезна для спецификации, пополнения и поддержки словарей предметной области. Онтологии-словари имеют ограниченное применение, поскольку, когда используемые термины принадлежат очень узкому (например, техническому) словарю, их смыслы уже заранее хорошо согласованы в пределах определенного (например, научного) сообщества. Например, онтологию этого типа образуют индексы машин поиска информации в Интернете.

2. $R = \emptyset$, $F \neq \emptyset$, $O = V = \langle X, \{ \}, F \rangle$. Тогда каждому элементу множества терминов из X ставится в соответствие функция (аксиома) интерпретации (одна или несколько) $f \in F$. Здесь могут быть варианты.

Допустим, что $X = X_1 \cup X_2$, причем $X_1 \cap X_2 = \emptyset$, где X_1 – множество интерпретируемых терминов, X_2 – множество не интерпретируемых терминов. При этом $\exists (x \in X_1, y^1, y^2, \dots, y^k \in X_2)$ такие, что $x = f(y^1, y^2, \dots, y^k)$, $f \in F$.

Введение в рассмотрение функции k аргументов обеспечивает более полную интерпретацию $x \in X$. Вид отображения $f \in F$ определяет выразительную мощность и практическую полезность этого вида онтологии.

Так, если предположить, что функция интерпретации задается оператором присваивания значений ($X_1 := X_2$, где X_1 – имя интерпретации X_2), то онтология трансформируется в пассивный словарь $O = V^p = \langle X_1 \cup X_2, \{ \}, \{ := \} \rangle$, так как все определения терминов из X_1 берутся из уже существующего и фиксированного множества X_2 . Практическая ценность данного словаря выше, чем простого, но он явно недостаточен, например, для представления знаний в задачах обработки информации в Интернете в силу динамического характера этой среды. Смысл таких терминов «вычисляется» каждый раз при их интерпретации. Ценность словаря для задач обработки информации, например в Интернет среде, выше, чем у предыдущей модели, но еще недостаточна, так как интерпретируемые элементы из X_1 никак не связаны между собой и, следовательно, играют лишь роль ключей входа в онтологию.

3. $R \neq \emptyset$, $F = \emptyset$, $O = V = \langle X, R, \{ \} \rangle$, т.е. множество F – пусто, множество отношений на концептах онтологии – не пусто и включает одно из отношений типа иерархии: $R = IS_A$ (*быть элементом класса*) ИЛИ POW (*part of while – часть - целое*) ИЛИ AKO (*a kind of – вид - род*) ИЛИ $contained_in$ (содержится в) ИЛИ see_also (смотрите также) и др.

Этот случай соответствует специальному подклассу онтологий, называемых простыми таксономиями:

$$O = T^O = \langle X, \{IS_A\}, \{ \} \rangle, O = T^O = \langle X, \{POW\}, \{ \} \rangle, O = T^O = \langle X, \{AKO\}, \{ \} \rangle \text{ и др.}$$

Под *таксономической структурой* понимается иерархическая система понятий, связанных между собой одним из отношений IS_A , POW , AKO и т.п. Поскольку перечисленные отношения имеют заранее фиксированную семантику, это позволяет организовывать структуру понятий онтологии в виде дерева. Такой подход имеет свои преимущества и недостатки, но в общем случае является адекватным и удобным для представления иерархий понятий.

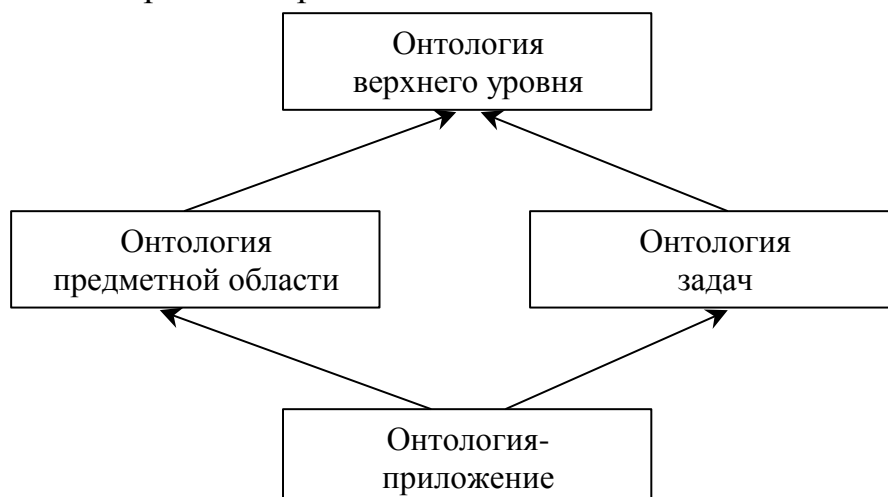
4. $R \neq \emptyset$, $F \neq \emptyset$ и $O = V = \langle X, R^I, F \rangle$, где R^I – это какое-либо из отношений иерархии (IS_A , POW , AKO , и т.п.). В этом случае строится иерархия терминов, в которой каждому из них ставится определенная процедура или аксиома интерпретация. Это, например, онтологии, в которой термины раскрываются по матрёшечному принципу.

2.4. $R \neq \emptyset$, $F \neq \emptyset$, $O = V = \langle X, R^I + R^C, F \rangle$ и $R = R^I + R^C$, где R^I – это какое-либо из отношений иерархии, R^C – специфические отношения универсального толкования типа: временные, пространственные каузальные, сравнения и т.п. Таким образом, в данном случае речь идет об иерархических структурах, на которых вводимые сущности имеют соответствующие толкования и на них определены отношения, в том числе, часто с нечетко фиксируемой семантикой. Это приводит к значительным трудностям по обработке таких структур, в связи с необходимостью учитывать свойства вводимых отношений: рефлексивность (нерефлексивность, антирефлексивность) симметричность (несимметричность, антисимметричность), транзитивность (нетранзитивность, антитранзитивность). Именно по этой причине это самые сложные онтологии и их практически очень мало, да и те, которые есть, являются очень упрощенными. Онтологии данного типа соответствуют модели расширяемой онтологии, достаточно мощной для спецификации процессов формирования пространств знаний.

1.3. Виды онтологий

С точки зрения концептуализации онтологии делятся на:

- онтологии представления, определяющие концептуализацию, которая лежит в основе формализма представления знаний. Понятия и отношения, определенные в других видах онтологий, считаются конкретизацией понятий онтологии данного вида. Предполагается, что они не зависят от конкретной ПрО. Онтологии данного вида могут использоваться для поддержки нескольких альтернативных теорий, относящихся к одной и той же предметной области;
- общие (родовые) онтологии, затрагивающие общие, фундаментальные аспекты концептуализации такие, как «часть», «причина», «участие», «представление»;
- промежуточные онтологии, содержащие общие понятия и отношения, характерные для конкретной предметной области. В идеальном случае, они используются в качестве интерфейса между онтологиями предметных областей и общими онтологиями, но могут выступать как онтологии верхнего уровня для описания знаний конкретной ПрО;
- онтологии верхнего уровня, являющиеся конкретным назначением понятий общих и промежуточных онтологий. Это уникальный модуль, находящийся над онтологиями предметной области, или являющийся самостоятельной, независимой от предметной области теорией;
- онтологии предметных областей, содержащие понятия определенной области знаний или входящих в нее областей;
- онтологии задач, описывающие определенные задачи или деятельность, используя словарь, введенный в общих, промежуточных онтологиях и онтологиях предметных областей;
- онтологии-приложения, являющиеся специализацией онтологий предметных областей и онтологий задач, и опирающиеся на определения, характерные для конкретного приложения.



Представленные на рис. онтологии рассматриваются следующим образом.

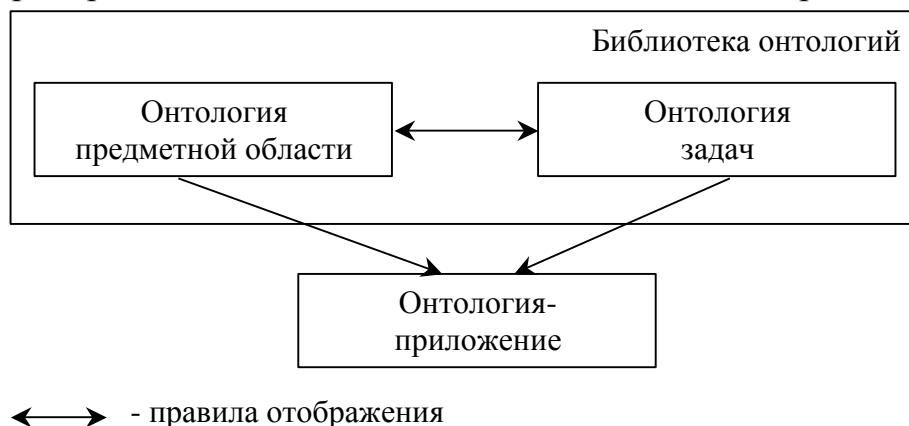
Онтологии верхнего уровня описывают общие понятия такие, как пространство, время, материя, объект, событие, действие и т.п., которые не зависят от конкретной задачи в ПрО.

Онтологии предметной области состоят из объектов и связей между ними, описанных в терминологии конкретной предметной области.

Онтологии задач ориентированы на решение конкретных проблем и включают все понятия, необходимые для описания процесса логического вывода, от самых абстрактных понятий, относящихся к схеме вывода до более специальных, характерных для отдельных методов.

Онтологии-приложения описывают понятия, зависящие как от конкретной проблемной области, так и от задачи. Эти понятия часто характерны для обеих соединяемых онтологий и соответствуют ролям, играемым объектами предметной области при выполнении определенной деятельности.

Онтология предметной области (или ее часть) может образовывать концептуальные связи с несколькими ПрО. Одно и то же множество объектов с одним и тем же набором свойств может использоваться в различных ПрО, для каждой ПрО характерна определенная модель поведения объектов предметной области. Знания предметной области вместе с моделью, описывающей их поведение в рамках определенных в онтологии задач функций, образуют частную модель ПрО. Программная реализация данной модели является онтологией-приложением. Онтологию-приложение можно рассматривать как усеченную онтологию ПрО применительно к включенным в онтологию-приложение задачам.



1.4. Операции над онтологиями

Операции над онтологиями ориентированы на реализацию как процессов: разработка онтологий; так и их поддержки и организации повторного их использования.

Для осуществления этих процессов используется множество операций над

онтологиями, которые по своему функциональному назначению можно разбить на 6 групп:

- **операции по редактированию;**

В данную группу операций входят операции, связанные с разработкой и редактированием онтологий на этапе их построения.

Создание (creating) онтологии состоит из четырех основных этапов:

- *составление спецификации*, определяющей допустимую область применения описываемых онтологией знаний,
- *определение задач*, в решении которых используются рассматриваемые знания;
- *разработка концептуализации*, то есть определение объектов предметной области, составляющих описываемую область знаний, и отношений между ними;
- *формализация* представляемых знаний.

Возможны два способа построения новой онтологии:

- (1) создание новой онтологии с нуля;
- (2) создание новой онтологии на основе имеющихся.

Если создание новой онтологии ведётся на уже имеющихся (ранее разработанных) онтологиях, то выполняются операции поиска и получения доступа к онтологиям, их кодирование и интеграция.

На стадии *поддержки (maintenance)* разработанной онтологии используются операции, обеспечивающие *организацию доступа* к онтологии, а также ее *просмотр* и *модификацию*.

Модификация онтологии, в свою очередь, включает в себя операции по *добавлению, удалению, перемещению и редактированию* элементов онтологии и их определений. *Редактирование*, в свою очередь, использует операции по *проверке онтологии на согласованность и просмотр онтологии (browsing)*. Возможности *просмотра* предоставляются редакторами онтологий и серверами, обеспечивающими хранение онтологий и доступ к библиотекам онтологий.

Организация доступа (accessing) – обеспечение доступа к просмотру и редактированию онтологий, хранящихся на сервере онтологий (сетевой компьютер, обеспечивающий доступ пользователей к хранящимся на нем онтологиям), из удаленных приложений или через Интернет.

Сохранение (saving) – операция, обеспечивающая возможность хранения промежуточных версий онтологии в ходе ее разработки.

Выпуск (publishing) – помещение онтологии на сервер онтологий с целью предоставления возможности ее повторного использования. Онтология, которая

должна быть помещена на сервер, предварительно оценивается экспертами, устанавливающими соответствие разработанной онтологии ее назначению.

- **алгебра онтологий;**

Алгебра онтологий разработана для систем, основанных на знаниях, с целью построения более крупных онтологических систем, использующих знания из нескольких проблемных областей. В основе построения таких систем лежит *объединение* информации из нескольких баз знаний. Основные действия алгебры являются простыми операциями, они предоставляют возможности *выбора объектов* из исходной предметной области и *помещения* их в новые. Композиция информации из нескольких предметных областей осуществляется в соответствии со следующей схемой.

1. В онтологиях предметных областей осуществляется *поиск терминов*, представляющих интерес для конкретной прикладной задачи.

2. Для найденных терминов на основании правил *строятся отображения для извлечения и соединения контекстов*. *Контекстом* считается часть более общей онтологии, являющаяся внутренне согласованной, достоверной самой по себе. В рамках каждой предметной области *формируются* контексты на основании отображений для извлечения. Выполняется *реструктуризация* для представления данных во внутреннем формате с внесением при необходимости семантических исправлений и уточнений. Сформированные контексты образуют срезы онтологий, значимые для рассматриваемой прикладной задачи, сами по себе также являющиеся онтологиями и называемые *соединяемыми онтологиями*.

3. В данных (соединяемых) онтологиях на основании отображений для соединения *определяются общие* для них контексты. В пределах этих контекстов соединяемые онтологии являются *согласованными*.

4. Контексты последовательно *объединяются* в результирующую онтологию, описывающую собранные из разных областей знания применительно к рассматриваемой задаче.

Алгебра онтологий состоит из трех основных операций.

Пересечение (intersection) – формирование из двух соединяемых контекстов нового, включающего в себя термины, используемые в обоих контекстах. Операция включает в себя задачу *определения минимального пересечения*.

Объединение (union) – формирование новой непротиворечивой онтологии, включающей термины, используемые несколькими контекстами (результат операции пересечения), и оставшиеся термины из соединяемых контекстов, не являющиеся совместно используемыми.

Вычитание (difference) – создание несимметричных подмножеств – наборов

терминов, которые включены в один контекст, но не используются в другом. Термины, совместно используемые несколькими контекстами, удаляются из результирующей онтологии. В результате остаются только уникальные термины.

Алгебра онтологий обладает следующими свойствами: операции могут *комбинироваться*, *переупорядочиваться*, альтернативные переупорядочивания могут *оцениваться*, возможна *оптимизация* порядка выполнения операций, результаты предыдущих операций могут *сохраняться и повторно использоваться*.

- **операции по интеграции онтологий:**

Для её выполнения применяются следующие операции.

Выборка (slicing, extraction) – предназначена для формирования внутренне согласованной части исходной онтологии или ее среза, который будет включен в новую онтологию. В алгебре онтологий такой срез назывался *контекстом*. В качестве начальных условий принимается, что часть онтологии, которая будет использоваться в новой онтологии, известна. Относительно этой части формируются *максимальный* и *оптимальный* срезы.

Срез называется *максимальным*, если любой использующий данную часть онтологии логический вывод, который выполняется в исходной онтологии, может быть выполнен, используя сформированный срез. По таксономическим связям классов определяется ближайший общий надкласс для выбранных классов, который включается в срез вместе со значимыми классами и классами, стоящими на промежуточных уровнях таксономии между найденным общим классом и значимыми классами. В самом общем случае максимальная выборка равна исходной онтологии.

Отсечение (pruning) – служит для удаления из максимального среза классов, которые являются несущественными для новой онтологии.

Выделяются: *простая интеграция* (интеграция нескольких онтологий (или частей) в одну онтологию); *соединение (merging)* онтологий как интеграция нескольких независимо разработанных онтологий. Исходные онтологии должны удовлетворять требованиям: совместимости формализмов и языков представления интегрируемых онтологий; отсутствия противоречий в структурных и семантических представлениях описываемых онтологиями знаний. Примером простой интеграции служит операция «*приращение*» (*augment*). Результатом операции является приращение некоторой онтологии проблемной области посредством извлечения и включения выбранных частей из независимо разработанных онтологий.

Для операции *соединения* онтологий характерно *объединение* одинаковых

по содержанию, но возможно отличающихся по форме представления элементов онтологии, в один элемент. Новая онтология строится на базе *соединения* классов, атрибутов, аксиом и других элементов онтологий, касающихся той же самой проблемы, для которой строится новая онтология. Для того, чтобы обеспечить единую форму представления соединяемых элементов онтологий, операции по переводу и преобразованию осуществляются до операции соединения. Онтология, полученная в результате соединения онтологий, *унифицирует* знания, хранящиеся в других онтологиях. В зависимости от количества изменений, вносимых в новую онтологию (полученную в результате соединения), могут быть выделены различные уровни интеграции: *установка соответствий*, *частичная согласованность* и *унификация*.

Установка соответствий (alignment) – операция, задающая отношения между онтологиями посредством установки связей между эквивалентными понятиями в онтологиях. Такой вид соединения сохраняет структуру исходных онтологий, не требует переименования терминов, изменений в аксиомах, определениях, доказательствах и т.п., т.е. является наиболее «слабой» формой соединения.

Частичная согласованность – операция, устанавливающая между онтологиями соответствия, поддерживающие эквивалентный логический вывод и вычисления, выполняемые для всех эквивалентных понятий и отношений.

Унификация – операция, результатом которой является новая универсальная онтология. Унификация определена как частичная согласованность онтологий, которая может быть

- **операции по агрегированию и декомпозиции;**

Эти операции предназначены для *разбиения* (декомпозиции) онтологий на модули и *сбора* онтологий из модульных компонент.

Модуляризация (modularization) или *декомпозиция (decomposition)* – операция по разбиению онтологии на модули, которые могут служить основой для разработки новых онтологий. Возможны подходы к разбиению онтологий на модули, основанные на:

- области применения: упорядочивание базы знаний, описываемой онтологией, по темам, заданным инженером по знаниям;
- решаемой задаче: упорядочивание базы знаний по правилам логического вывода, применяемых для ее решения, а также по используемым в логическом выводе объектам;
- микро-теориях: микро-теория состоит из набора высказываний и ограничена областью применения (для каких объектов, при каких условиях и т.п.

данная теория может быть использована);

- контекстах: выявление разобобщенных контекстов и помещение их в самостоятельные модули. Для выявления разобобщенных контекстов используются, как правило, специальные методы анализа контекстов.

Помещенные в библиотеку модули могут повторно использоваться при построении онтологий для других предметных областей.

Составление (assembling, composing) – операция конструирования новой онтологии из модулей. Новая онтология строится посредством выбора подходящих модулей из библиотеки и их соединения.

- **операции по преобразованию;**

Эти операции ориентированы на приведение онтологий к общему виду или выражения формы представления одной онтологии в форме другой с целью предоставления возможностей их повторного использования в операциях по интеграции.

Переформулирование (reformulation) – операция по преобразованию представления входной теории. Например, данная операция может быть использована для преобразования функциональных предикатов в отношения.

Построение таксономии (taxonomy design) – операция по формированию иерархической структуры, основанной на отношении *IS_A*, в которой свойства более общих понятий наследуются подчиненными им понятиями.

Перевод (translation) – операция по переводу онтологий, разработанных в одном формализме и реализованных с помощью одних языков, в другие формализмы и языки. Реализация операции связана с разработкой средств, способных осуществлять взаимные переводы между несколькими форматами.

- **операции по сравнению, проверке и оценке.**

Сравнение (matching) – операция, устанавливающая степень соответствия онтологий. Данная операция включает с себя операции по *сравнению концептуальных структур* и операции, *определяющие степень переводимости одной онтологии в другую*. Сравнение концептуальных структур заключается в *определении степени совместимости понятий*, используемых в разных онтологиях. Концептуальная структура определяется областью применения описываемых знаний. Для сравнения выбираются несколько онтологий, описывающих одну и ту же прикладную проблему. Степень совместимости оценивается в соответствии с классификацией: *согласие / соответствие / конфликт / противоречие*. В соответствии с указанной классификацией концептуальные структуры могут находиться в следующих отношениях:

- согласующиеся (consensus): разные онтологии имеют одинаковые

структуру и терминологию;

- соответствующие (correspondence): в разных онтологиях одинаковые структурные элементы имеют разные имена (несовпадения в терминологии);
- конфликтующие (conflict): в разных онтологиях разные структурные элементы названы одинаково;
- противоречивые (contrast): в разных онтологиях нет совпадений ни в структурах, ни в терминологии.

Проверка (validation) – операция, проверяющая корректность и непротиворечивость (внутреннюю согласованность) онтологии. В ходе операции производится семантический и синтаксический анализ онтологии.

Оценка (evaluating) – операция, в результате которой выносится заключение о соответствии онтологии ее назначению. Операция выполняется экспертами перед *выпуском* онтологии. Онтология оценивается по ее соответствию: требованиям, спецификации, реальному миру, степени компетентности.

Приведенный перечень операций может варьироваться в различных системах представления по названиям, по определениям, модифицироваться, но, в принципе, является типовым для разработки и поддержки онтологий любого класса.

1.5. Инструменты инженерии онтологий

Рассмотрим кратко наиболее известные инструменты инженерии онтологий, поддерживающие процессы создания онтологий и их трансляции.

- ***Ontolingua.*** Среда разработки, обеспечивающая коллективное использование, предоставляющая набор средств для создания онтологий и библиотеку модулей онтологий. Среда доступна через Интернет. Имеющиеся средства ориентированы на создание онтологий посредством сбора и расширения онтологий, хранящихся в библиотеке. Данная среда позволяет переименование понятий из различных составляющих онтологий, устраняет неоднозначности в ссылках на понятия в процессе ввода/вывода, поддерживает циклическое включение одной онтологии в другую, предоставляет пользователям возможности для расширения включенных онтологий такими способами, как добавление упрощающих допущений или ограничений на область допустимых значений и расширение областей допустимых значений полиморфных операторов. Для многократного использования существующих онтологий поддерживаются операция *включения* (включение всего содержимого внешней онтологии без изменений), *полиморфного уточнения* (расширение включаемых аксиом посредством уточнения их операторов) и *ограничения* (ограничения, накладываемые на включаемые аксиомы).

- **OntoEdit.** Среда разработки для проектирования, приспособления и импорта моделей знаний для прикладных систем. Среда поддерживает многоязычную разработку онтологий и множественное наследование. Поддерживается выполнение следующих функций: определение понятий в иерархии понятий, множественное наследование, определение локальных и глобальных отношений, определение экземпляров классов, многоязычность, определение базовых аксиом таких, как разрозненные понятия, симметричные отношений, транзитивные отношений и т.п..

- **OilEd.** Редактор онтологий, позволяющий пользователю строить онтологию, используя *OIL (The Ontology Inference Layer* – онтологический уровень логического вывода). *OIL* предлагается для WWW-представления онтологий и использования уровня логического вывода для онтологий, который объединяет широко используемые примитивы моделирования, заимствованные из языков, основанных на фреймовом представлении, с формальной семантикой и методами логики первого порядка, поддерживающими механизм вывода. Среда *OilEd* не предназначена для построения онтологий большого размера, интеграции онтологий и поддержки других видов деятельности, направленных на создание онтологий, она является разновидностью «текстового редактора» для редактирования онтологий, обладает достаточной функциональностью, чтобы предоставить пользователю возможность построения онтологий и продемонстрировать, как используется механизм вывода системы для проверки разработанных онтологий на согласованность.

- **Protégé.** Система предоставляет интегрированную среду редактирования баз знаний и наращиваемую архитектуру для создания основанных на знаниях систем пользователя. *Protégé* позволяет строить онтологию предметной области (определять классы и их иерархию, слоты, задавать ограничения на значения слотов, устанавливать связи между классами и определять их свойства); генерировать изменяемые по желанию пользователя формы на приобретение знаний (вначале генерируются заданные по умолчанию формы для получения экземпляров классов, опираясь на заданные пользователем слоты, затем сгенерированные формы могут быть изменены посредством переупорядочивания полей на экране, изменения размеров, меток и других свойств); вводить знания предметной области (экземпляры классов в онтологию могут быть добавлены разработчиком). Система является платформой для доступа к другим приложениям с внедренными системами, основанными на знаниях. *Protégé* – это библиотека, предоставляющая доступ другим приложениям для просмотра баз знаний.

- **Web-DESO.** Система, предназначенная для создания онтологий, описывающих некоторую ПрО. Онтологии предметных областей соединяются в одну результирующую онтологию предметной области. Такая же операция выполняется для онтологий задач, представленных в виде дерева декомпозиции структуры задача-метод. Две результирующие онтологии помещаются в библиотеку и интегрируются в онтологию-приложение посредством перевода их представления во внутреннее представление данной онтологии, которое определено формальной теорией онтологии верхнего уровня. Онтология-приложение хранится в библиотеке и включает в себя, кроме перечисленных онтологий, онтологии источников знаний. Источники знаний выявляются в ходе работы системы, после чего их представление переводится в нотацию онтологии верхнего уровня. Многократный доступ к знаниям осуществляется через онтологию-приложение.