

UDC 004.4'41: 004.942

Information Interaction as a Mechanism of Semantic Gap Elimination

Victor Y. Tsvetkov

Moscow State University of Geodesy and Cartography, Russia
Dr. (Engineering), Dr. (Economics), Professor
E-mail: cvj2@mail.ru

Abstract. The article studies semantic gap as an objective phenomenon, shows that semantic gap occurs both in parallel computing and in other areas. Semantic description of the content is revealed as a set of different descriptions. Causes of semantic gap are described. The content of information exchange is explained in the article. Information interaction in the semantic field is interpreted as a mechanism to lessen the gap.

Keywords: formalization; modeling; semantic description; semantic analysis; semantic gap; formal grammar; Chomsky hierarchy; information interaction.

Введение. Семантический разрыв характеризует различие между описаниями объектов различных формальных представлений, например, языковое и символьное представление, языковое и формульное представление [1]. В компьютерной науке, семантический разрыв рассматривается как ситуация, когда обычная деятельность человека: наблюдения и задачи - передаются в вычислительное представление с потерей семантической содержательности. «Разрыв» означает разницу между многозначностью формулировки контекстуальных знания естественного языка и формализаций воспроизводимых и вычислительной среде в формальных языках, например, в языках программирования. «Семантический» в данном случае означает уменьшение или существенное исключение смысловой содержательности. Уменьшению семантического разрыва посвящены многие работы, особенно в области параллельных вычислений. Однако семантический разрыв имеет место всегда в других областях, когда происходит кодирование информации или построение модели объекта с уменьшением семантического содержания. Поэтому вопросы устранения или уменьшения семантического разрыва актуальны во многих областях.

Семантическое описание. Семантическое описание включает совокупность описаний. Внешняя семантика может быть определена как семантика модели внешнего мира. Формальная семантика – это семантика модели внешнего мира сформулированная на естественном языке.

Математическая семантика – это семантика, определения всех объектов, сформулированная на том же математическом языке. Смысл конструкций при этом описывается самими конструкциями. Машинная семантика – это семантика, определения всех объектов и действий, происходящие внутри компьютера. Архитектурная компьютерная абстракция - это некий внутренний уровень с формальной семантикой. Она часто представляется графом. Таким образом, формальной машинной семантике присваивается внешняя семантика с помощью человеческой интерпретации.

Формальная грамматика определяет формальный язык, который является множеством конечной длины последовательности символов. Например, последовательность строк, которые могут быть построены с применением правил продукции в другую последовательность символов, которая изначально содержит только начальный символ.

Согласно иерархии Хомского [2] формальные языки и формальные грамматики делятся на 4 типа по их условной сложности: тип 0 — неограниченные; тип 1 — контекстно-зависимые; тип 2 — контекстно-свободные; тип 3 — регулярные. К третьему типу относятся регулярные (автоматные) — самые простые из формальных грамматик. Они являются контекстно-свободными, но с ограниченными возможностями.

Для описания синтаксиса языков программирования нужны более мощные грамматики, чем регулярные. Обычно для этого используют укорачивающие контекстно-свободные грамматики (УКС-грамматики) [3]. Грамматики этого класса, с одной стороны,

позволяют достаточно полно описать синтаксическую структуру реальных языков программирования; с другой стороны, для разных подклассов УКС-грамматик существуют достаточно эффективные алгоритмы разбора.

Рассмотрение языков в иерархии Хомского показывает, что автоматизированный способ перевода с одного языка на другой невозможен выше определенного уровня. Следовательно, семантические разрывы возникают всегда при языковых (модельных) переходах и требуют дополнительных действий.

В практической деятельности любое представление задач реального мира требует перевода контекстной знаний эксперта (приложений высокого уровня) в воспроизводимые операции вычислительной машины (низкий уровень) или в модели обработки информации. Многообразие естественного языка позволяет описать задачи, которые невозможно адекватно описать на формальном языке. Поэтому для уменьшения семантического разрыва применяют различные средства. Одним из таких средств является информационное взаимодействие.

Взаимодействие и информационное взаимодействие. Взаимодействие как процесс известно во многих науках. Взаимодействие — объективная и универсальная форма движения и развития, которая определяет существование и структурную организацию материальной или нематериальной системы. Взаимодействие включает процессы воздействия объектов друг на друга, процессы воздействия внешней среды на объекты. Взаимодействие позволяет выделить среди множества объектов — объекты взаимодействия и не взаимодействующие объекты.

Информационное взаимодействие [4] относительно новый вид взаимодействия. Оно осуществляется в информационном поле и является развивающимся феноменом [5]. Информационное взаимодействие имеет разновидности. Рассмотрим в данной работе прямое взаимодействие (Trusted Path). Это такое информационное взаимодействие, при котором передаваемая информация не подвергается перехвату или искажению. Взаимодействие отражает отношение и связь между объектами взаимодействия. Среди взаимодействующих объектов в информационном поле можно выделить исходные и порожденные и относительно независимые взаимодействующие объекты. В аспекте семантики можно говорить об исходном и порожденном объекте. Для таких объектов информационное взаимодействие может приводить к следующим явлениям:

- изменению состояния одного или нескольких взаимодействующих объектов;
- изменению или появлению свойств у объектов взаимодействия;
- изменению качественных и количественных характеристик объектов взаимодействия;
- изменению отношений между объектами
- изменению уровня информированности объектов взаимодействия
- появление или уменьшение степени обусловленности между объектами

Информационное взаимодействие (Information interaction) в широком смысле — процесс взаимодействия информационных и неинформационных объектов друг на друга через информационную среду, информационные модели и информационные технологии. Одной из особенностей информационного взаимодействия является возможность информационного копирования. Это означает, что при информационном взаимодействии возможна передача части свойств и признаков одного (передающего информацию) объекта в другой (принимающий информацию) путём их копирования или путём полного перемещения. При этом информационные свойства передающего объекта не изменяются.

Информационное взаимодействие, в отличие от других видов взаимодействия, может нарушать закон Хиральной чистоты (Л. Пастер, 1860) и допускает полное клонирование объекта. Клонирование означает полное соответствие между оригиналом и копией. Оно возможно в информационном поле и нарушает Хиральную чистоту.

Информационное взаимодействие как инструмент преодоления семантического разрыва — процесс взаимодействия субъектов и объектов, основным содержанием которого является достижение необходимой цели семантического соответствия между исходным и порожденным объектами на основе использования информационных ресурсов (объемов, потенциалов, структур, качественных и количественных признаков [6]).

Процесс информационного взаимодействия в семантическом поле.

Основная часть процессов информационного взаимодействия связана с двухсторонним или односторонним обменом информации. По существу информационное взаимодействие структурно повторяет процесс управления с обратной связью. При формализации описания один объект является исходным (эталон) второй порожденным (копией). Это определяет качественное неравенство между ними. Исходный объект имеет большую полноту описания. При таком взаимодействии участвует субъект, поэтому в информационном поле необходимо рассматривать триаду «эталон – субъект – копия». При информационном взаимодействии субъект принимает информацию о текущем описании порожденного объекта. Если текущее описание не адекватно эталону, то формируется новое формальное описание в рамках того языка, на котором оно выполнено. Этот цикл повторяется, пока не будут исчерпаны возможности языка формального описания порожденного объекта. Процесс информационного взаимодействия приведен на рис. 1.

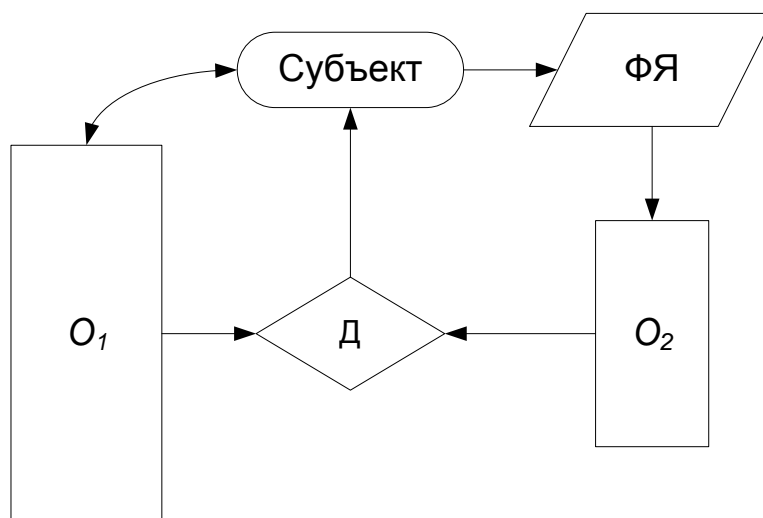


Рис. 1. Процесс информационного взаимодействия

Исходный объект O_1 служит основой описания объекта O_2 . Субъект использует O_1 как эталон и с помощью формального языка ФЯ формирует объект O_2 . При отсутствии информационного взаимодействия процесс формирования заканчивается. При наличии информационного взаимодействия в информационном поле осуществляется сравнение информационных потоков I_1 и I_2 . Эти потоки содержат описание объектов на естественном языке и на формальном. В дискриминаторе Д осуществляется сравнение описаний и результат сравнения направляется субъекту. Субъект осуществляет формальный и главное семантический анализ. Семантический анализ служит для выявления семантического разрыва и формирования дополнительных описаний для его преодоления или уменьшения. Такой процесс повторяется, пока не будет достигнут баланс возможностей формального языка и описания объекта O_2 .

В этом процессе информационное взаимодействие направлено на изменение информационных ресурсов порожденного объекта с целью их уменьшения семантического разрыва. Поэтому процессы информационного взаимодействия можно рассматривать как процессы семантического развития порожденного объекта. Они протекают в информационном поле «субъект – объекты».

Если выделить информационное взаимодействие только в информационной области, то можно выделить дополнительную специфику этих процессов. В информационной области оперируют с информационными объектами, информационными ситуациями, информационными позициями [7]. В процессе информационного взаимодействия (как развития) происходят следующие изменения информационных объектов:

- увеличение сложности порожденных объектов (возрастание коэффициента эмергентности);

- модификация структуры порожденных объектов;
- повышение адаптивности порожденных объектов к внешней среде;
- интеграция информационных объектов;
- количественное изменение характеристик порожденных объектов;

Важной характеристикой такого информационного взаимодействия является инкрементальность [8] и ресурсность [9].

Инкрементальность взаимодействия выражается в том, что развитие предшествующей стадии включается в последующую с наполнением целевой функции или накоплением ресурса. Ресурсность взаимодействия выражается в том, что субъект в процессе взаимодействия накапливает опыт. Который затем можно применить как информационный ресурс при формировании следующих порожденных объектов.

Информационное взаимодействие субъекта и объектов состоит из этапов. Применительно к преодолению семантических разрывов взаимодействию оно имеет вид:

1. Выявление семантического разрыва;
2. Определение (цели) способа формализации для уменьшения разрыва;
3. Определение последовательности действий;
4. Осуществление последовательности действий;
5. Восприятие нового состояния порожденного объекта;
6. Интерпретация нового состояния;
7. Оценка результата информационного взаимодействия.

Этап оценки результата информационного взаимодействия включает оценки, на которых субъект пытается понять: достиг он своей цели полностью или частично? Насколько валидны полученные результаты?

Если цель не достигнута, происходит возвращение первому этапу. Это определяет цикличность [10] и инкрементальность [8] такого информационного взаимодействия. Таким образом, информационное взаимодействие в сложной совокупности «человек – объекты» может быть описано следующей последовательностью:

«ЧТО = КАК = ДЕЛАТЬ = ОЦЕНКА».

Эту последовательность действий необходимо использовать как технологическую парадигму при информационном взаимодействии для уменьшения или устранения семантического разрыва.

Выводы. Семантические разрывы являются объективно существующим явлением, связанным с использованием языковых средств разных уровней и ситуациями неадекватного описания моделей по отношению к свойствам реального мира. Следовательно, эта проблема существует и будет проявляться в новых формах.

Информационное взаимодействие является одним из методов уменьшения семантических разрывов. Преимущество информационного взаимодействия в допустимости клонирования информационных объектов и возможности нарушения закона Хиральной чистоты. Это обеспечивает наибольшее соответствие между порождающим и порожденным объектом в семантическом поле.

Примечания:

1. Dorai C, Venkatesh S. Bridging the Semantic Gap with Computational Media Aesthetics. IEEE MultiMedia 2003;10(2):15-17.
2. Хомский Н., Дж. Миллер Введение в формальный анализ естественных языков // Кибернетический сборник / Под ред. А.А.Ляпунова и О.Б.Лупанова. М.: Мир, 1965.
3. Кулагин В.П. Алгебра сетевых моделей для описания параллельных вычислительных систем // Автоматизация и современные технологии. 1993. №2. С. 25-30.
4. Иванников А.Д., Тихонов А.Н., Мордвинов В.А. Получение знаний методами информатики и геоинформатики // Вестник Московского государственного областного университета. 2012. №3. С. 140-142.
5. Оболяева Н.М. Устранение информационной асимметрии как инструмент повышения качества образования // Геодезия и аэрофотосъемка. 2012. №6. С. 123-124.

6. Оболяева Н.М. Информационное управление качеством образования// Дистанционное и виртуальное обучение. 2012. № 12. С. 58-62.
7. Железняков В.А. Интеллектуальное обновление информации в банке геоданных // Инженерные изыскания. 2012. № 5. С. 58-61.
8. Цветков В.Я., Железняков В.А. Инкрементальный метод проектирования электронных карт. // Инженерные изыскания. 2011. № 1 январь. С. 66-68.
9. Цветков В.Я. Информационные модели и информационные ресурсы // Геодезия и аэрофотосъемка. 2005. №3. С. 85-91.
10. Цветков В.Я., Корнаков А.Н. Циклическая модель информационного управления // Современные наукоемкие технологии. 2010. №2. С. 129-131.

References:

1. Khomskii N., Dzh. Miller Vvedenie v formal'nyi analiz estestvennykh yazykov // Kiberneticheskiy sbornik / Pod red. A.A.Lyapunova i O.B.Lupanova. M.: Mir, 1965.
2. Kulagin V.P. Algebra setevykh modelei dlya opisaniya parallel'nykh vychislitel'nykh sistem // Avtomatizatsiya i sovremennye tekhnologii. 1993. №2. S. 25-30.
3. Ivannikov A.D., Tikhonov A.N., Mordvinov V.A. Poluchenie znaniy metodami informatiki i geoinformatiki // Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. 2012. №3. S. 140-142.
4. Obolyaeva N.M. Ustranenie informatsionnoi asimmetrii kak instrument povysheniya kachestva obrazovaniya // Геодезия и аэрофотосъемка. 2012. №6. S. 123-124.
5. Obolyaeva N.M. Informatsionnoe upravlenie kachestvom obrazovaniya// Distsionnoe i virtual'noe obuchenie. 2012. № 12. S. 58-62.
6. Zheleznyakov V.A. Intellektual'noe obnovlenie informatsii v banke geodannykh // Inzhenernye izyskaniya. 2012. № 5. S. 58-61.
7. Tsvetkov V.Ya., Zheleznyakov V.A. Inkremental'nyi metod proektirovaniya elektronnykh kart. // Inzhenernye izyskaniya. 2011. № 1 yanvar'. S. 66-68.
8. Tsvetkov V.Ya. Informatsionnye modeli i informatsionnye resursy // Геодезия и аэрофотосъемка. 2005. №3. S. 85-91.
9. Tsvetkov V.Ya., Kornakov A.N. Tsiklicheskaya model' informatsionnogo upravleniya // Sovremennye naukoemkie tekhnologii. 2010. №2. S. 129-131.

УДК 004.4'41: 004.942

Информационное взаимодействие как механизм устранения семантического разрыва

Виктор Яковлевич Цветков

Московский государственный университет геодезии и картографии, Россия
Доктор технических наук, доктор экономических наук, профессор
E-mail: cvj2@mail.ru

Аннотация. В статье описан семантический разрыв как объективное явление. Показано, что семантический разрыв встречается не только при параллельных вычислениях, но и в других областях. Раскрывается содержание семантического описания как совокупности описаний. Описаны причины появления семантического разрыва. В статье раскрывается содержание информационного взаимодействия. Описан процесс информационного взаимодействия в семантическом поле и механизм уменьшения разрыва.

Ключевые слова: формализация; моделирование; семантическое описание; семантический анализ; семантический разрыв; формальные грамматики; иерархия Хомского; информационное взаимодействие.