

МОДЕЛЬ ФОРМАЛИЗАЦИИ ТРЕБОВАНИЙ ПРИ КОНЦЕПТУАЛЬНОМ ПРОЕКТИРОВАНИИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ

В статье рассматриваются вопросы, актуальные для обеспечения экономической эффективности концептуального проектирования автоматизированных систем (АС). Начальный этап проектирования состоит в выявлении, отборе и систематизации требований заказчика к АС, которые формулируются в текстовой форме на естественно-профессиональном языке, допускающем неоднозначное истолкование. Для разработки программного обеспечения требуется исключение неопределенности. В работе представлена эффективная модель формализации требований заказчиков, основанная на логико-лингвистическом подходе.

Предметная область. Результаты исследований и разработок, представленных ниже, относятся к предметной области концептуального проектирования автоматизированных систем (АС), интенсивно использующих программное обеспечение. Такое проектирование начинается с выявления, отбора и систематизации требований к АС, в процессах коллективного обсуждения, интервью, мозгового штурма, размышления и других видов рассудочной деятельности.

Независимо от специфики источника требований к разрабатываемой АС первоначальная регистрация проводится в текстовой форме на естественно-профессиональном языке. Такая форма обычно многозначна и содержит различного рода неопределенности, которые способны привести к различного рода ошибкам. Поэтому для перехода от требований к их спецификации необходима специальная обработка, уточняющая содержание их значений. Поэтому от исходной информации о требованиях на естественном языке следует перейти к другим, более строгим представлениям на формальном или полуформальном языке.

В практике таких переходов используются различные языки спецификаций, как сокращенная форма естественного языка, обладающая свойствами структурированности, абстрактности и ясности, с возможностями включения в синтаксис конструкции языков программирования. Кроме того, для широкого

ряда языков спецификаций (так называемые исполнимые спецификации) существуют программные средства для автоматической трансляции формальных спецификаций в программу на языке высокого уровня, а также проверки их правильности.

Различают два подхода в разработке формальных спецификаций на специализированном языке: алгебраический подход, при котором система описывается в терминах операций и их отношений; и модельно-ориентированный подход, при котором модель системы строится с использованием математических объектов, таких как множества и функции. Таким образом, модельно-ориентированные языки лучше подходят для описания систем, базирующихся на состояниях. В свою очередь, языки алгебраической нотации оптимальны в спецификации абстрактных типов данных¹. В работе² в качестве средств разработки формальных спецификаций выбираются следующие: псевдокод, деревья решений, диаграммы

¹ *Соммервилл И.* Инженерия программного обеспечения.: пер. с англ. / И. Соммервилл. М.: Изд. дом «Вильямс», 2002. 624 с.

² *Леффингуэлл Д.* Принципы работы с требованиями к программному обеспечению. Унифицированный подход: пер. англ. / Д. Леффингуэлл, Д. Уидриг. М.: Изд. дом «Вильямс», 2002. 448 с.

деятельности, различные поведенческие модели: модели сущности связей, объектно-ориентированные модели, модели потоков данных, модели конечных автоматов.

Наиболее известные языки алгебраических спецификаций – **LOTOS, OBJ, LARCH** – формальные языки, использующие алгебраический подход, разработанный для спецификации поведения информационных систем распределенной и параллельной обработки и для описания интерфейсов подсистем, внутри больших распределенных систем.

Модельно-ориентированные языки применяются для описания состояний систем программного обеспечения. Наиболее известные **Z, VDM, B**.

Спецификации в этих языках состоят из математических моделей, сформированных из множеств, списков и отображений наряду с операциями, которые изменяют состояние модели. Например, базовым модулем языка **Z** является схема, состоящая из имени, описания переменных и предиката³.

В результате обязательного использования в языках формальных спецификаций математического аппарата формальные спецификации, особенно если записаны в виде сложных математических формул, сложны для чтения их неспециалистами в области математики. Этот факт создал предпосылки к разработкам языков спецификаций, отвечающих требованиям простоты и применимости к специфике решаемой задачи формальных спецификаций в процессах разработки информационных систем.

Язык спецификаций **SIMPLE** применяется для верификации программ. В качестве основного достоинства разработчики отмечают легкость изучения и простоту конструкций взамен сложности синтаксических конструкций, трудную для понимания семантику, требующую существенных познаний в логике и других сложностей типичных языков

³ Butler M. Introductory Notes on Specification with Z / M. Butler. Режим доступа: <http://www.ecs.soton.ac.uk>.

⁴ Земцов Н.А. Формальное моделирование спецификаций процессов для композиционного проектирования потоков работ / Н.А. Земцов, С.А. Ступников. Режим доступа: <http://synthesis.ipi.ac.ru>

спецификаций, применяемых в существующих системах верификации⁴.

В лаборатории методов композиционного проектирования ИС ИПИ РАН в качестве канонической информационной модели, предназначенной для унифицированного представления спецификаций компонентов, применяют язык **СИНТЕЗ**, ориентированный на семантическую интероперабельность и композиционное проектирование информационных систем в широком диапазоне существующих неоднородных информационных ресурсов. Этот язык обладает гибридными возможностями, обеспечивающими интеграцию как структурированных, так и слабо структурированных моделей данных⁵.

В исследовании принято решение использовать язык спецификаций, подобный прологовскому языку. Это позволит проводить автоматизированную предикативную обработку текстовых представлений, регистрирующих содержание требований, что способно снять большую часть ошибок на уровне семантики. Разработанные методы и средства встроены в специализированную технологию концептуального проектирования, в которой используется вопросно-ответное моделирование рассуждений в группе разработчиков **АС**, что позволило заимствовать многие решения для работы с вопросами и ответами, а также перенести их на действия по предикативной обработке требований.

Для демонстрации сущности задачи предикативной обработки перейдем к лингвистической терминологии.

Общие принципы. В основе системы автоматизированной формализации требований лежат принципы машинного перевода. В частности, поиск в тексте ключевых признаков логической и семантической связи между простыми предикативными единицами (предикативная единица – пара подлежащее-сказуемое) осуществляется на основании информации, заложенной в лингвистическую базу. В ней содержится словарь естественного языка, таблицы, которые содержат информацию о словах и словоформах языка

⁵ Ануреев И.С. Язык спецификаций SIMPLE / И.С. Ануреев. Режим доступа: <http://www.ict.nsc.ru/ws/Lyap2001/2201>.

и используются в морфологическом, синтаксическом и логико-семантическом анализе текстовых единиц.

По завершении этапа предсинтаксического анализа каждое предложение преобразуется в последовательность базовых текстовых единиц, которые обычно соответствуют отдельным словам. Каждая такая единица характеризуется отношением к группе индикаторов связи, положением в предложении (в начале предложения, перед знаком препинания, после знака препинания), отношением к той или иной группе частей речи.

В требованиях особого типа, выраженных мультипредикативными предложениями, образуется система отношений между простыми требованиями или частями таких систем. Для формирования приблизительной картины логических связей в многосоставных предложениях необходимо выявить простые предикативные единицы с применением метода анализа позиции исследуемого слова относительно остальных чаще всего соседних слов или ключевых признаков.

Для разрешения задачи определения вида синтаксической и, следовательно, логической связи между двумя простыми предикативными единицами необходимо располагать информацией о присутствии в сложном предложении следующих ключевых признаков:

- 1) сочинительных союзов, образующих связь между однородными членами и простыми предложениями;
- 2) подчинительных союзов и союзных слов, образующих подчинительную связь между простыми предложениями;
- 3) соотносительных или указательных слов, выполняющих роль определяющего признака местоименно-соотносительного типа предложения;
- 4) лингвистических средств выражения модальности-возможности в предложении (частицы *бы, б, ли*);
- 5) лингвистических средств выражения отрицания (частица *не*), общности (*все, для всех, каждый, для каждого, ни один, любой*) и существования (*некоторый, существует*).
- 6) лексического выражения тождественности одной предикативной единицы другой (*такой ... как, так ... как, точ-*

но, словно).

Проблема однозначности определения вида синтаксической связи в сложном предложении снимается, если связь между простыми частями образуется либо сочинительными союзами (в сложносочиненном предложении), либо посредством семантических союзов (в сложноподчиненном предложении), которые служат не только для формального прикрепления придаточного предложения к главному, но и для выражения тех или иных смысловых отношений. Например, причинно-следственные отношения между двумя переменными совершенно определено выражаются союзами «*вследствие того что*», «*из-за того что*» «*по причине*», «*потому что*» и т.п. Прологоподобное представление предложения $P_{r_a} :- P_{r_m}$ или $P_{r_m} :- P_{r_a}$ также в случае предполагаемой причины $PSBL(P_{r_a}) :- PSBL(P_{r_m})$, где P_{r_m} , P_{r_a} – предикаты главного и придаточного предложений соответственно; $PSBL$ – предикат модальности «возможность» в логике предикатов первого порядка.

В предложениях времени отношения следования событий или их одновременности выражаются явно с помощью союзных средств или не явно, посредством наречий (*раньше, позже, во время*). В прологоподобном представлении устанавливается прямой порядок предикатов (P_{r_m}, P_{r_a}), в случае предшествовании событий главного предложения событиям придаточного (*перед тем как*), и обратной (P_{r_a}, P_{r_m}) в противоположном случае (*после того как*). В случае, когда события главного и придаточного предложений происходят одновременно (*в то время как, в тот момент когда*), соответствующие предикаты собираются в одной формуле ($P :- P_{r_a}, P_{r_m}$).

Таким образом, в рассмотренных примерах полностью восстановить картину формальной организации предложения возможно, если поставить в соответствие союзам, описанным в лингвистической базе, определенный смысл, который определит вид логической связи между связующими элементами.

В предложениях, где предикативные единицы связываются функциональными союзными средствами (*кто, где, когда, как*),

которые выполняют чисто синтаксическую роль выражения зависимости, придаточного предложения от главного, невозможно, опираясь только на семантический анализ, сделать определенный вывод о виде логической связи. Однако в силу того, что в нескольких случаях для разных типов сложных предложений разработанные формулы сходны, то разброс в вариантах синтеза пары предложений в формулу сокращается до четырех.

Способ 1. В прикомпаративных (здесь и далее классификация сложноподчиненных предложений⁶) предложениях ввести дополнительное сказуемое – компаратив, обозначающий отношения между предикатами главного и придаточного предложений. В стандартном случае:

$$CMPR [Pr_m(Sb_m;Ob_m)Ac_m; Pr_a(Sb_a;Ob_a)Ac_a]$$

В случае отсутствия в придаточном предложении сказуемого:

$$CMPR [Pr_m(Sb_m;Ob_m)Ac_m; Pr_m(Sb_a;Ob_a)Ac_a]$$

где *CMPR* – предикат-компаратив. В предложениях соответствия (связанной структуры) отношения между простыми предикатами прослеживаются причинно-следственные отношения:

$$CMPR_a [Pr_a(Sb_a;Ob_a)Ac_a] - CMPR_m [Pr_m(Sb_m;Ob_m)Ac_m]$$

где *CMPR_m*, *CMPR_a* – предикаты-компаративы главного и придаточного предложений, соответственно.

В отождествительных (с предметным и определительным значением коррелятивных слов) и в сравнительных предложениях ввести дополнительное сказуемое-тождество:

$$Identity (Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_a;Ob_a)Ac_a);$$

$$Identity (Pr_m(Sb_m;Ob_m)Ac_m, PSBL(Pr_a(Sb_a;Ob_a)Ac_a))$$

где *Identity* – предикат-тождество.

В изъяснительных, сравнительных, фразеологического типа предложениях ввести дополнительный предикат *возможно*, для выражения модальности *возможность* в логике предикатов первого порядка:

- сравнительные

$$Pr_m(Sb_m;Ob_m)Ac_m :- PSBL (Pr_a(Sb_a;Ob_a)Ac_a);$$

- фразеологические

$$PSBL(Pr_a(Sb_a;Ob_a)Ac_a) :- Pr_m(Sb_m;Ob_m)Ac_m;$$

- изъяснительные

$$PSBL (Pr_a(Sb_a;Ob_a)Ac_a).$$

Способ 2. В качественно-определительных, выделительно-определительных предложениях включить характеризующее слово в предикат придаточного предложения в качестве субъекта, объекта или ассоциативной составляющей:

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_a;[характеризующее слово])Ac_a;$$

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a([характеризующее слово];Ob_a)Ac_a;$$

где «характеризующее слово» – опорное слово, к которому «присоединяется» придаточное.

Способ 3. В относительно-распространительных, распространительных, выделительно-определительных, фразеологического типа, вмещающих предложениях включить субъект, объект или ассоциативную составляющую главного предложения в предикат придаточного предложения. В результате получаем следующие прологоподобные представления формул предложений.

- в придаточном предложении нет субъекта (подлежащего):

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_m;Ob_a)Ac_a$$

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_m;Ob_a)Ac_a + Ac_m$$

- в придаточном предложении нет ни субъекта (подлежащего), ни объекта (дополнения):

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_m;Ob_m)Ac_a + Ac_m;$$

- в придаточном предложении нет объекта (дополнения):

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_a;Ob_m)Ac_a + Ac_m;$$

- в придаточном предложении обстоятельства совпадают с главным предложением:

$$P :- Pr_m(Sb_m;Ob_m)Ac_m, Pr_a(Sb_a;Ob_a)Ac_a + Ac_m.$$

Способ 4. В относительно распространительных, отождествительных (с предметным значением коррелятивных слов, обстоятельственно-пространственным значением коррелятивных слов) предложениях

⁶Русская грамматика /АН СССР. Ин-т рус.яз. М., 1980. Режим доступа: <http://rusgram.narod.ru/>

включить союзное слово в предикат придаточного предложения в качестве субъекта или объекта. Прологоподобные представления формул предложений:

$P: - \text{Pr}_m(Sb_m; Ob_m)AC_m, \text{Pr}_a(SConj; Ob_a)AC_a$

$P: - \text{Pr}_m(Sb_m; Ob_m)AC_m, \text{Pr}_a(Sb_m; SConj)AC_a,$

где *SConj* – подчинительный союз, союзное слово.

В отождествительных (с предметным значением коррелятивных слов, с обстоятельственно-пространственным значением коррелятивных слов) предложениях включить соотносительное слово в предикат придаточного предложения в качестве субъекта, объекта или ассоциативной составляющей:

$P: - \text{Pr}_m(Sb_m; Ob_m)CWord, \text{Pr}_a(Sb_a; Ob_a)SConj$

$P: - \text{Pr}_a(Sb_a; Ob_a)CWord, \text{Pr}_m(Sb_m; Ob_m)SConj,$

$P: - \text{Pr}_m(Sb_m; Ob_m)AC_m, \text{Pr}_a(CWord; Ob_a)AC_a$

где *CWord* – соотносительное или опорное слово в местоименно-соотносительных предложениях.

Специфика логико-семантического анализа для формирования прологоподобных представлений в английском языке, а также модель функционирования системы формирования формальных описания рассмотрены автором⁷.

Разработанный инструментарий формализации в его приложениях к работе с требованиями подтвердил эффективность в снятии неопределенности в построении

⁷ *Валюх В.В.* Система перевода технических текстов в формальное представление/ В.В. Валюх // КЛИН- 2004: тр. междунар. конф. Ульяновск: УлГТУ, 2004. Т.1. С.53-54.

Valyukh V.V. Methods of machine translation of texts from natural into predicate logic-based language / V.V.Valyukh // Interactive systems and technologies : Collect. scientific papers Intern. conf. Ulyanovsk: UISTU. С. 173 – 178.