

УДК 004.89, 004.942

DOI 10.51691/2541-8327_2021_8_2

**О ПОДХОДЕ К ИНТЕРВЬЮИРОВАНИЮ ЭКСПЕРТОВ ДЛЯ
РЕАЛИЗАЦИИ ИДЕАЛИЗИРОВАННОГО ЭКСПЕРИМЕНТА ПРИ
ФОРМИРОВАНИИ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ СЛОЖНЫХ
ОБЪЕКТОВ**

Столбов А.Б.*к.т.н., младший научный сотрудник**ФГБУН Институт динамики систем и теории управления имени В.М.**Матросова СО РАН**Россия, г. Иркутск***Лемперт А.А.***к.ф.-м.н., ведущий научный сотрудник**ФГБУН Институт динамики систем и теории управления имени В.М.**Матросова СО РАН**Россия, г. Иркутск***Аннотация**

В статье рассматривается подход к решению задачи интервьюирования экспертов на основе общего алгоритма формирования математических моделей сложных объектов. Использование интервью позволяет извлекать недостающие в условиях неопределенности данные и знания для выполнения специализированных методик идентификации коэффициентов математических моделей на основе различных вариантов идеализированного эксперимента с применением экспертных оценок. Вопросы интервью генерируются автоматически на основе заранее заданных процедур, содержательно связанных с этапами общего алгоритма формирования моделей. При этом для хранения информации используется предметная и проблемная концептуальные модели в

форме онтологий. Программная реализация подхода выполнена в виде компонента платформы создания систем, основанных на знаниях. Представлена модель данных компонента, архитектура модулей, описана их функциональность. Приведены иллюстративные примеры. Отличительной особенностью предложенного компонента является явная связь элементов модели интервью с элементами онтологии исследования.

Ключевые слова: интеллектуальная поддержка процесса математического моделирования, сложный объект, концептуальное моделирование, интервьюирование эксперта, программное обеспечение.

***TOWARDS AN APPROACH TO INTERVIEWING EXPERTS FOR THE
IMPLEMENTATION OF AN IDEALIZED EXPERIMENT DURING THE
CREATION OF MATHEMATICAL MODELS OF COMPLEX OBJECTS***

Stolbov A.B.

candidate of technical sciences, junior researcher

Matrosov Institute for System Dynamics and Control Theory SB RAS

Russia, Irkutsk

Lempert A.A.

candidate of physical and mathematical sciences, leading researcher

Matrosov Institute for System Dynamics and Control Theory SB RAS

Russia, Irkutsk

Annotation

The article considers an approach to solving the problem of interviewing experts based on a general algorithm for forming mathematical models of complex objects. The application of interviews allows us to extract the missing data and knowledge in the uncertainty conditions to perform specialized methods for identifying the coefficients of mathematical models based on various variants of an idealized

experiment using expert evaluation. Interview questions are generated automatically based on predefined procedures that are meaningfully related to the stages of the general model formation algorithm. At the same time, domain and problem conceptual models in the form of ontologies store desired information. The software implementation is made in the form of a component of the platform for creating knowledge-based systems. The data model of the component, its architecture, and modules functionality are described. Illustrative examples are given. A distinctive feature of the proposed component is the explicit connection of the elements of the interview model with the elements of the research ontology.

Key words: intellectual support of the mathematical modeling process, complex object, conceptual modeling, expert interviewing, software.

В последние годы в научных исследованиях по различными направлениями наметился устойчивый тренд на всё более комплексное изучение процессов и объектов [13, 14]. Данный процесс связан как с достижениями в отдельных научных дисциплинах, так и с общими тенденциями технического прогресса: увеличение вычислительных мощностей, повышение уровня коммуникаций между учеными, публикация программ и данных в открытых источниках, семантическое описание данных. В этих условиях востребованы исследования, рассматривающие общесистемные проблемы комплексного изучения сложных объектов, основанные на процедурах системного анализа, принципах междисциплинарного подхода, методах математического моделирования. При этом математические модели могут выступать не только как готовое средство обработки существующих данных для получения количественных оценок, но и как исследовательский инструмент предварительного этапа, определяя процесс изучение сложного объекта в условиях неопределенности и задавая процедуры по поиску и

извлечению новых данных и знаний, изначально в исследовании отсутствующих или представленных не в полной мере.

Настоящая работа находится в контексте исследований, связанных с направлением по изучению сложных объектов, начатое В.И. Гурманом и развиваемое его учениками. К ключевым публикациям, заложившим основу методологии, можно отнести следующие [1, 2, 5, 6]. Далее рассмотрим применение этой методологии на примере реализации одного из вариантов получения линеаризованного варианта математической модели сложного объекта в условиях неопределенности, включая этапы параметрической идентификации с использованием идеализированного эксперимента на основе экспертных оценок. Более формально и подробно этот алгоритм представлен в [7]. Здесь же представлена сокращенная версия этого алгоритма (первые 6 этапов) с заменой математической формализации на содержательные пояснения и с отдельным вниманием к вопросам, связанным с обработкой знаний.

Этап 1. Абстрактная модель.

На этом этапе осуществляется определение состава эндогенных и экзогенных переменных некой абстрактной динамической модели функционирования сложного объекта. При этом о математической природе этой модели делается единственное предположение – переменные связаны между собой в форме автономной системы обыкновенных дифференциальных уравнений (ОДУ), структура которой неизвестна.

Этап 2. Формирование «опорных» точек.

В пространстве эндогенных и экзогенных переменных происходит определение набора «опорных» точек, в окрестности которых обосновано проводить линеаризацию абстрактной динамической модели. Поэтому для включения точки в набор предъявляется содержательное требование: её значения должны отражать некоторое состояние близкое в каком-то смысле к стационарному и устойчивому. Как правило, такие значения определяются в Дневник науки | www.dnevniknauki.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

результате обработки длительных наблюдений за объектом исследования или других объектов аналогичных всему объекту или его частям. При определении набора «опорных» точек используются и экспертные оценки, основанные как на знании обще-предметных закономерностей, так и особенностей функционирования изучаемого объекта. При этом для определения по координатным значениям «опорной» точки могут использоваться различные способы по различным массивам данных. В результате набор «опорных» точек зависит от абстрактной модели только по номенклатуре, в то время как значения известны независимо от модели. Примерами «опорных» точек являются: среднее многолетнее значение показателя, очищенного от выбросов в данных; равновесное состояние; фоновый (частично фоновый) уровень показателя, т.е. значения показателя при отсутствии всех (части) внешних воздействий.

Этап 3. Выбор «опорной» точки из набора.

При анализе реальных систем получение единой модели как результат линеаризация около единственной «опорной» точки как правило невозможно осуществить. Поэтому используется серия приближений, где каждой области линеаризации соответствует множество содержательных допущений: например, сезонность, территориальная локализация и др.

Этап 4. Линеаризация абстрактной модели методом Тейлора.

При указанных поведении системы вблизи выбранной опорной точки осуществляется линеаризация абстрактной модели методом Тейлора. В этом случае параметрами модели становятся коэффициенты матриц, каждый из которых связывает две переменные и может быть содержательно проинтерпретирован следующим образом: они характеризуют реакцию системы к изменениям показателей состояния и внешних воздействий относительно равновесия (если единицы варьирования времени выбраны малыми в пределах возмущений).

Этап 5. Реализация идеализированного эксперимента.

Как отмечается в монографии [1] под абстрактным экспериментом понимается любая процедура, позволяющая восстановить искомый оператор. Это может быть реальный эксперимент; наблюдение в некоторых естественных условиях, приближающихся к требуемым условиям эксперимента; анализ литературных данных по определенной системе, вытекающей из заданных требований. Комплекс мероприятий, связанных с получением необходимой информации для реализации абстрактного эксперимента и определения конкретных коэффициентов взаимодействия между отдельными компонентами моделируемого объекта называется идеализированным экспериментом.

В данной работе идеализированный эксперимент реализуется в форме интервьюирования экспертов. Для этого разработана специальная методика связи процесса интервью с концептуальной моделью, и в рамках оригинальной платформы работы со знаниями реализован программный компонент интервьюирования. Результатом данного этапа являются экспертные оценки поведения частей моделируемого сложного объекта в различных задаваемых условиях.

Этап 6. Полученные экспертные оценки используются совместно для вычисления коэффициентов модели по специальным формулам компонентного эксперимента [1].

Последующие этапы. После определения параметров модели производятся тестовые расчёты для верификации модели. В случае провала тестов предыдущие этапы могут повторяться или вынесено заключение о невозможности адекватного моделирования на основе имеющегося набора информации по данной методике. В случае успешного прохождения тестов выполняются сценарные расчёты с их последующим анализом и взаимным сравнением. Делаются выводы о поведении изучаемого объекта в различных условиях.

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

Таким образом, результатом данного исследования в контексте поддержки процесса моделирования сложного объекта является решения по сохранению информации о содержательной интерпретации данных и предметно-ориентированных допущениях, задаваемых алгоритмом, и применение этой информации для организации процесса получения необходимых данных в ходе электронного текстового интервью.

Для хранения значимой информации о методах моделирования (проблемно-ориентированные знания) и об объекте моделирования (предметно-ориентированные знания) используется адаптация стандартной структуры онтологической модели [10, 11] в следующей форме: понятие, атрибут, отношение, экземпляр. Более подробно о конкретных онтологиях для поддержки процесса моделирования излагается, например, в [8, 9, 12], в т.ч. и в публикации авторов [3].

С учетом указанных предыдущих наработок предлагается хранить информацию о модельных допущениях (согласно алгоритму) в следующем виде:

- условие на наличие понятия и/или атрибута понятия и/или отношения между понятиями;
- ограничения на значения (точные или интервальные, количественные или качественные) экземпляров.

В качестве иллюстрации: если модель заболеваемости населения рассматривается для зимнего периода, то в онтологии необходимо наличие отношения типа «является» между понятием «Период моделирования» и понятиями «Ноябрь», «Декабрь», «Январь», «Февраль», «Март». Отметим, что данное содержательное условие (в т.ч. другие последующие примеры) может концептуально моделироваться иначе, например, через свойства понятий или значения экземпляра. Ключевым здесь является то, как эта информация обрабатывается и как на её основе происходит генерация вопросов интервью.

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

Информацию о физическом смысле параметров линеаризованной модели предлагается хранить в виде экземпляров (обозначения условны):

- экземпляры понятия «переменная 1» и «переменная 2» (x и y) добавлены в номенклатуру концептуальной модели – либо через наличие отношения (на уровне понятий), либо как атрибуты (на уровне экземпляров);

- в экземпляре понятия «Реакция относительно равновесия» значения атрибутов «Источник» и «Приёмник» равны x и y;

- в экземпляре понятия «Реакция относительно равновесия» значения атрибутов «Единица измерения источника» и «Единица измерения приёмника» равны допустимому (установлены и действительно являются единицами измерений, что можно проверить, например, через иерархию наследования);

- в экземпляре понятия «Реакция относительно равновесия» значения атрибута «Фоновый уровень» уставлено.

Данная информация позволит автоматически сформировать вопрос респонденту-эксперту для оценки эффекта влияния.

Программная реализация идеализированного эксперимента при формировании математических моделей сложных объектов с применением интервьюирования экспертов выполнена на базе оригинальной веб-ориентированной платформы по созданию систем, основанных на знаниях (далее упоминается как платформа создания СОЗ или платформа) [15]. В этой платформе уже присутствует вся функциональность, связанная с концептуальным моделированием в форме онтологий, разработкой продукционных баз знаний, хранением данных числового и строкового типа. Дополнительные возможности, связанные с организацией процесса интервьюирования, выполнены в форме дополнительного компонента интервьюирования со следующей архитектурой:

- модуль формирования вопросов интервью;
- модуль генерации интервью.

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

Компонент имеет клиент-серверную архитектуру. Для хранения данных используется СУБД PostgreSQL. Компонент поддерживает собственную модель интервью, полученную по результатам анализа известных средств генерации опросов (Moodle, Google Forms, SurveyGizmo и др.) и с учетом требований текущего исследования.

В интервью может быть два типа структур для хранения вопросов:

- обычный/простой тип – содержит список ответов;
- составной тип – содержит список вопросов.

При этом каждый тип вопроса содержит информацию:

- уникальный идентификатор вопроса;
- максимальное время ответа на вопрос;
- текст вопроса;
- максимальное количество ответов на вопрос;
- список тематик, к которой можно отнести вопрос;
- список элементов онтологии (понятие, атрибут, отношение, экземпляр) с

которым связан вопрос.

На текущий момент модель интервью поддерживает несколько типов ответов. Каждый тип ответа содержит следующую информацию:

- уникальный идентификатор;
- идентификаторе вопроса в методе оценки ответа (если есть);
- список элементов онтологии (понятие, атрибут, отношение, экземпляр) с

которым связан вопрос.

- момент времени указания ответа (выбор в элементе управления «checkbox» или «radio»).

В модели имеется 3 типа ответов.

1. Респонденту предоставляется текст ответа, который ему необходимо выбрать. В этом случае ответ имеет два значения: выбран или не выбран. При

формировании результатов интервью в значение ответа заносится номер ответа. Текст каждого ответа формируется разработчиком интервью.

2. Предопределенный список ответов на основе некоторого повторно используемого шаблона. Используется в случае стандартных вариантов ответа, например, об образовании или поле. В этом случае необходимо предварительно создать шаблон с вариантами ответов. При формировании результатов в значение ответа заносится номер ответ. При создании интервью разработчик выбирает шаблон, а текст ответа изменять не может.

3. Открытый ответ с использованием текстовой формы ввода.

Основная функции модуля формирования вопросов интервью – ввод и редактирование информации по следующим сущностями модели данных интервью: профиль, опрос, профильный опрос, тема, вопрос, вопрос по теме, вопрос интервью. Предполагается, что ввод данных сможет осуществлять эксперт-предметник или математик-модельер. Модуль состоит из следующих блоков:

1. Блок основной страницы – хранит информацию о структуре страницы , обеспечивает инициализацию других блоков модуля.

2. Контроллер-представление в форме дерева – используется для выбора группы экранных форм для редактирование данных, связанных с интервью, вопросом, профилем интервью, темой вопроса.

3. Клиентское представление для обработки (ввода и редактирования) данных в одну таблицу БД. Имеет следующие элементы управления: таблица (используется библиотека <http://tabulator.info/>), текстовое поле, область текста, переключатель, кнопки (добавить, удалить, изменить запись). Параметры настройки источников данных: имя базы данных, имя таблицы, имя полей таблицы, элемент управления для поля, параметры элементы управления (возможность редактирования, видимость элемента для пользователя).

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

4. Клиентское представление для обработки (ввода и редактирования) данных в связанные таблицы БД, реализующих связь многие ко многим. Элементы управления и параметры настройки источников данных аналогичны блоку 3 с добавлением информации о связях таблиц БД: имя первой и второй таблиц с данными и имена соответствующих полей-идентификаторов, имя таблиц связи, имена полей-идентификаторов внешних ключей, а также ссылка на связанный блок 3. Идентификатор одной из сущностей связи многие ко многим определяется на основе текущей активной записи связанного блока 3, другая сущность выбирается в этом блоке.

5. Обработчик запросов на сервере. Добавить, удалить, обновить записи в таблицах. Запрос формируется блоками 3 или 4.

Модуль генерации интервью состоит из следующих блоков:

1. Генератор пользовательского интерфейса на основе модели опроса.

2. Блок генерация спецификации интервью содержит код создания модели интервью по информации из концептуальной модели и данных, подготовленных модулем формирования вопросов интервью в соответствии с выбранным профилем.

3. Блок генерации пользовательского интерфейса интервью осуществляет запуск генератора интервью с различными параметрами: использовать таймер или нет, показывать вопросы по одному или сразу весь список, тип поведения при исчерпании отведенного на вопрос времени (отключение возможности ответа на вопрос, кнопка перехода доступна; автоматический переход; ничего не происходит), возможность перехода на следующий вопрос без ответа на текущий.

Таким образом, компонент интервьюирования является компонентом общего назначения и позволяет реализовать интервью как в привязке к концептуальной модели, так и независимо от неё. По умолчанию реализуется простой алгоритм взаимодействия с респондентом, где все вопросы задаются

Дневник науки | www.dnevnikaui.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

последовательно в соответствии с порядковым номером. При этом при наличии в справочных целях может выдаваться дополнительная информация, связанная с текущим онтологическим контекстом. Сама же эта информация по умолчанию не влияет на алгоритм прохождения интервью.

Для реализации идеализированного эксперимента в генератор встроены специализированные процедуры формирования модели интервью соответствующие этапам алгоритма 1, 2, 5 (задаются как ключевое слово в названии связанных с интервью профилей). В качестве иллюстрации: для генерации вопросов по содержательной интерпретации коэффициентов линеаризованной модели используется либо шаблоны по умолчанию, либо вопрос с тематиками «Влияние эндогенной переменной на эндогенную» или «Влияние экзогенной переменной на эндогенную». В этом случае для каждой текущей переменной математической модели (связь ищется в явно указанных ассоциированных элементах онтологии, связанной с вопросом, или транзитивно, или запрашивается у респондента в случае отсутствия информации или дублирования) генерируется свой вопрос. Например, «укажите величину увеличение показателя 1 при увеличении показателя 2 на x единиц в отсутствии других влияний и относительно фонового уровня, где показатель 1 – частота заболеваний верхних дыхательных путей, показатель 2 – концентрация пыли, x единиц – 10 мкг/м³, фоновый уровень – обращений на 1000 человек».

В заключении отметим, что предложенный подход позволяет автоматизировать процесс приобретения информации от экспертов в удобной и понятной форме интервью. При этом респондент-эксперт может не быть знаком с используемой при создании математической модели методологией [1, 2, 5-7] и лишь частично знаком с изучаемой предметной областью относительно всего сложного объекта (что типично для междисциплинарных исследований). В отличие от известных систем генерации интервью в компоненте Дневник науки | www.dnevniknauki.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

интервьюирования платформы создания СОЗ осуществлена явная связь элементов модели интервью с элементами онтологии, задающих концептуальную модель исследования. Это обстоятельство позволяет эксперту в случае необходимости понимать и дополнительно исследовать контекст вопроса. Архитектура предложенного компонента обладает достаточной общностью для дальнейшего расширения новыми типами вопросов, ответов, элементов управления и т.п. Важным направлением развития компонента является задача интеграция с ещё одним компонентом платформы – создания потока работ (workflow). В области приложений планируется применение предложенного подхода и компонента для математического моделирования водных ресурсов на основе балансовых моделей [4].

Исследование выполнено при финансовой поддержке РФФИ и Правительства Иркутской области в рамках научного проекта № 20-47-380001.

Библиографический список

1. Гурман В.И. Эколого-экономические системы: Модели, информация, эксперимент / В.И. Гурман, В.А. Дыхта, Н.Ф. Кашина и др. – Новосибирск: Наука, 1987. – 216 с.
2. Гурман В.И. Эколого-экономическая стратегия развития региона: Математическое моделирование и системный анализ на примере Байкальского региона / В.И. Гурман, В.Е. Викулов, Е.В. Данилина и др. – Новосибирск: Наука, 1990. – 184 с.
3. Лемперт А.А. Использование онтологического подхода для интеллектуальной поддержки сценарного моделирования водных объектов Приангарья / А.А. Лемперт, А.Б. Столбов, В.В. Козлов // Информационные и математические технологии в науке и управлении. – 2020. - №3(19). – С.14-24.
4. Михайлов В.Н. Гидрология устьев рек / В.Н. Михайлов. – М.: Изд-во МГУ, 1998. – 176 с.

ЭЛЕКТРОННЫЙ НАУЧНЫЙ ЖУРНАЛ «ДНЕВНИК НАУКИ»

5. Модели управления природными ресурсами / под. редакцией В. И. Гурмана. – М.: Наука, 1981. – 264 с.
6. Моделирование и оценка состояния медико-эколого-экономических систем / под ред. В. А. Батурина. – Новосибирск: Наука, 2005. – 249 с.
7. Столбов А.Б. Intellectual technologies for medical-ecological-economic systems research support / А.Б. Столбов, В.А. Батурина // Восьмая Международная конференция «Системный анализ и информационные технологии» (САИТ – 2019). Труды конференции, 8 – 14 июля 2019 г., Иркутск - Листвянка, Россия. – М.: ФИЦ ИУ РАН, 2019. – С.528-533.
8. Dunlap R. Earth system curator: Metadata infrastructure for climate modeling / R. Dunlap, L. Mark, S. Rugaber, V. Balaji, J. Chastang, L. Cinquini, C. DeLuca, D. Middle-ton, S. Murphy // Earth Science Informatics. – 2008. - №1. – С.131-149.
9. Elag M. An ontology for component-based models of water resource systems / M. Elag, J.L. Goodall // Water Resour. Res. – 2013. - №49. – С.5077-5091.
10. Gangemi A. Ontology Design Patterns / A. Gangemi, V. Presutti // Handbook on Ontologies / edited by S. Staab, R. Studer. 2nd Edition.Springer Verlag. – 2009. – Pp.221-244.
11. Gruber T.A. Translation Approach to Portable Ontology Specifications / T.A. Gruber // Knowledge Acquisition. – 1993. - №5(2). – С.199-220.
12. Harpham Q. Introductory Overview: The OpenMI 2.0 Standard for Integrating Numerical Models / Q. Harpham, A. Hughes, R.V. Moore // Environmental Modelling & Software. – 2019. -Vol.122. 104549. – С.1-13.
13. Ladyman J. What is a complex system? / J. Ladyman, J. Lambert, K. Wiesner // Euro Jnl Phil Sci. – 2013. - №3. – С.33-67.
14. Moore R., Hughes A. Integrated environmental modelling: Achieving the vision / R Moore., A. Hughes // Geological Society, London, Special Publications. – 2016. - Vol.408. – С.17-34.

15. Nikolaychuk O.A The software platform architecture for the component-oriented development of knowledge-based systems / O.A. Nikolaychuk, A.I. Pavlov, A.B. Stolbov // Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). – 2018. – Pp. 1234-1239

Оригинальность 92%