

## МЕТОДИКА ТЕСТИРОВАНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ КОГНИТИВНОЙ СЛОЖНОСТИ МОДЕЛЕЙ РАЗЛИЧНЫХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ

Григорьев А.В.

Кафедра ПМИИ, ДонГТУ  
grigorie@r5.dgtu.donetsk.ua

### Abstract

*Grigoriev A.V. A technique of testing for definition cognitive complexity of models for various subject areas. Questions construction of the tests for estimation allowable cognitive complexity for the operators of computers are discussed. A technique for estimation cognitive complexity of models for various subject areas is offered.*

### Введение

В настоящее время достаточно хорошо исследованы фундаментальные и некоторые прикладные аспекты организации функционирования систем «человек-машина». Так, в работе [1] сформулирован ряд принципов построения таких систем. Наиболее полно данные принципы реализованы в концепции адаптивного компьютера [2]. Согласно данной концепции внутрикомпьютерная среда должна включать: блок зондирования характеристик оператора; математическую модель (ММ) оператора; генератор ММ оператора; ММ рабочей среды; генератор ММ рабочей среды. Блок зондирования обеспечивает получение долгосрочных и актуальных характеристик оператора. Долгосрочные характеристики определяются психофизиологическим типом человека (сангвиник и т.п.), а актуальные характеристики зависят от многочисленных внешних факторов и внутреннего состояния человека. Для определения долгосрочных характеристик применяются, например, тесты ММРІ, Айзека, Стрелая; для определения актуальных характеристик - тест Люшера, теппинг-тест, и т.п. В качестве интегральной модели актуальных характеристик оператора [2] используется показатель уровня внимания, заданный как функция времени. Для настройки данной модели на состояние конкретного пользователя используется методика управления проведением дополнительных тестов. Проблема снижения и ограничения сложности процесса взаимодействия пользователя и компьютера в системах искусственного интеллекта (СИИ) так же важна, как и в любых других системах взаимодействия «человек-машина». Системы искусственного интеллекта имеют ряд специфических этапов работы (обучение, вывод решения и т.д.), для которых необходим учет ряда специфических характеристик оператора. Спецификой обладают и конкретные подклассы СИИ, например, интеллектуальные системы автоматизации проектирования (САПР), комплексы автоматизации системного анализа (СА) и системы автоматизации научных исследований (АСНИ). Общей спецификой перечисленных подклассов является необходимость оценки и ограничения когнитивной сложности (КС) представления моделей объектов. Проанализируем существующие методы решения задачи ограничения КС на примере ряда работ. В работе [3], посвященной разработке языка структурного анализа SA, предназначенного для передачи понимания, проблема ограничения сложности взаимодействия человека и системы SA решается путем применения принципа «все, о чем есть смысл говорить, должно быть представлено не более, чем из шести частей». Такой подход основан на известном в психологии «законе  $7 \pm 2$ » [4], гласящем, что устойчивая коммуникация возможна лишь с  $7 \pm 2$

коммуникантами. В работе [5], посвященной разработке языка для описания систем взаимодействующих процессов, рассматривается корреляция КС моделей и некоторых показателей структурной сложности систем взаимодействующих процессов. С этой целью авторами данной работы проводилось тестирование различных категорий пользователей ЭВМ. Кроме того, показана возможность выявления диапазона значений показателей структурной сложности, обеспечивающего необходимый для безошибочной работы уровень «понимаемости» (>95%) системы функциональных процессов.

К специфике перечисленных работ можно отнести наличие концептуальной модели предметной области (КМ ПО) различной сложности, языка для описания моделей, а так же меры для оценки и ограничения КС, основанной на интерпретации понятия структурной (системной) сложности.

Можно сделать следующие выводы:

- интегральная характеристика оператора должна учитывать ограничения на КС моделей;
- необходимо дополнительно определить специфические для интеллектуальных САПР и АСНИ долговременные и актуальные характеристики операторов;
- данные характеристики должны отражать сложность восприятия, проектирования и контроля функционирования (моделирования) моделей объектов;
- необходимо обеспечить построение меры КС моделей для различных предметных областей и языков представления модели.

### 1 Постановка задачи тестирования

В течении нескольких лет среди студентов ДонГТУ 4-х и 5-х курсов специальностей «Программное обеспечение» (ПО) и «Менеджмент информационных систем» (МИ) проводилось тестирование на восприятие КС. Тесты проводились, соответственно, в курсах «Автоматизация проектирования сложных объектов и систем» и «Моделирование экономических и производственных процессов». Общее число тестируемых составило 214 человек.

Табл. 1 Состав тестовых примеров.

№	Предметная область	Язык представления модели	Количество тестовых примеров
1	Схемотехника (микропроцессорные системы)	Принципиальные схемы	8
2	Экономика	Таблицы (графики)	8(8)

Целью тестирования было определение корреляции восприятия КС с долговременными и актуальными характеристиками операторов для различных групп экспертов, различных ПО и форм представления моделей. Студенты тестировались на общем наборе тестовых примеров (таблица 1). В зависимости от типа ПО тестируемые условно делились на две категории: эксперты в данной предметной области и дилетанты. Каждая ПО и тип языка представления модели был представлен одинаковым числом тестовых примеров. Состав тестовых примеров характеризовался следующим образом: а) все примеры интерпретировались как структурно-функциональные модели, задающие некоторый состав отношений в соответствии с принятой унифицированной концептуальной моделью ПО [6,7]; б) в состав примеров включались как явно «необозримые» примеры, так и «слишком» простые примеры; в) с точки зрения структурной составляющей модели примеры отличались как составом типов блоков и типов данных, так и наличием, числом измерений, а так же размером регулярных составляющих модели (массивов); г) графики рассматривались как форма



задания функциональной макромоделей объектов при заданной структурной модели границы объектов; д) таблицы и принципиальные схемы (ПС) рассматривались как формы задания функциональных микромоделей при явном (для ПС) или неявном (для таблиц) задании структурной модели границ объектов.

## 2 Организация тестирования

Приведенный ниже порядок тестирования операторов задает алгоритм работы блока зондирования. Тестируемым предлагалось исследовать КС моделей сложных систем для некоторой ПО. Тест имеет три этапа.

Этап 1. Предварительно перед тестированием каждый из тестируемых задавал субъективные характеристики для себя как эксперта в данной ПО. Для этого требовалось по шкале 1-5 задать уровни следующих долговременных характеристик: компетентность в данной ПО; зрительная память; абстрактное мышление. Указывался пол эксперта: мужской или женский. Т.о. осуществлялся учет отличных характеристик экспертов. Учитывался и ряд общих характеристик экспертов. Например, общие характеристики студентов-экспертов специальности ПО: 1) возраст - 21 год; 2) уровень подготовки - профессиональные программисты - бакалавры, 4 курс; 3) уровень знакомства со схмотехникой средний, т.е. изучены курсы электроники, электротехники и САПР.

Этап 2. На данном этапе требовалось оценить по шкале оценок 1-5 определенное количество моделей на степень возможности выполнения экспертом следующего ряда операций: зрительный анализ; понимание; запоминание; возможный контроль функционирования. Все четыре перечисленные характеристики являются актуальными характеристикам оператора. Предварительно каждый эксперт ознакомился с общей трактовкой этих операций:

1) «Выполнить зрительный анализ» означало, не вникая в смысл схемы, охватить ее глазами («обозримость»);

2) «Понять схему» означало определить, что имеется на входе схемы, а что получается на выходе; иными словами, требовалось выявить порядок функционирования блоков внутри схемы;

3) «Запомнить схему» - означало для эксперта быть способным воспроизвести схему в общих чертах на бумаге;

4) «Обеспечить контроль функционирования схемы» - означало, что в ходе моделирования схемы эксперт не утрачивает возможность контроля функционирования отдельных составляющих и всей схемы в целом.

Этап 3. Тестируемым предлагалось дать экспертную оценку относительной КС (ОКС) всех моделей. Оценка проводилась в два шага.

На первом шаге предлагалось определить относительную сложность для некоторого набора пар примеров, что означало задание частичного порядка на множестве примеров.

На втором шаге тестируемые строили полный линейный порядок КС всех примеров по шкале 1-8, ставя более высокие номера более сложным моделям. При этом не допускались группы эквивалентности прототипов по сложности.

Таким образом, экспертом задавалась шкала  $Ш_0$  ОКС для набора прототипов.

Примечание. Этап 1 может быть расширен специализированными тестами для определения отличных долговременных характеристик экспертов. Однако результаты статистического анализа тестов показывают высокую степень близости совокупности актуальных оценок для экспертов, имеющих эквивалентные субъективно заданные долговременные характеристики. Это позволяет утверждать, что «этап 1» в данном виде обладает достаточным уровнем полноты и достоверности.

### 3 Обработка результатов тестирования

Обработка результатов тестирования состояла в построении модели оператора. Были определены следующие цели:

- 1) упорядочить полную систему оценок в соответствии с частичным порядком на подмножестве оценок, связанных с ОКС прототипов, т.е. построить шкалу абсолютной КС (АКС) прототипов  $Ш_A$ ;
- 2) ввести на шкале  $Ш_A$  интервал допустимых значений КС;
- 3) исходя из граничных оценок на  $Ш_A$ , определить истинность условий ограничения для некоторого прототипа, имеющего соответствующую оценку;
- 4) обеспечить возможность расчета абсолютной КС прототипов на основе модифицированной меры структурной (системной) сложности.

#### 3.1 Построение шкалы весомости операций для определения КС.

Были выделены группы экспертов по ПО, эквивалентных по значениям общих и отличных долговременных характеристик.

Для каждой группы средствами EXCEL была рассчитана корреляция между ОКС прототипа (КСП) по шкале  $Ш_0$  и значениями оценок для прототипов в различных группах экспертов. Ее усредненные значения представлены в таб. 2.

Табл. 2 Зависимость между уровнем компетентности и корреляцией сложности выполнения различных операций с ОКС.

Компетентность	Корреляция сложности операций над прототипами и ОКС			
	Обозримость	Контроль	Запоминание	Понимание
1	0,95	0,94	0,90	0,87
2	0,85	0,87	0,83	0,64
3	0,78	0,90	0,83	0,79
4	0,77	0,87	0,89	0,66

Анализ зависимостей позволил сделать следующий ряд выводов:

1) Существуют следующие зависимости между КСП, оценками сложности и уровнем компетентности экспертов: для наиболее компетентных экспертов КСП определяется сложностью запоминания; для средне компетентных экспертов КСП определяется сложностью контроля функционирования; для мало компетентных экспертов КСП определяется сложностью охвата глазами, т.е. обозримости.

2) На границе между группами наблюдается равенство по важности смежных в данном списке параметров сложности.

3) Для наиболее подготовленных соблюдается такой порядок важности параметров: запоминание - контроль функционирования - обозримость. Для наименее подготовленных соблюдается обратный порядок, т.е.: обозримость - контроль - запоминание.

4) «Понимание» в данной его формулировке и месте в порядке определения относительной КСП менее всего влияет на величину КСП.

5) Пол практически не влияет на восприятие сложности.

На рис. 1 приведены «гладкие» EXCEL-графики, задающие зависимости между центрированными относительно математического ожидания показателями корреляции сложности выполнения отдельных операций над прототипами и относительной КСП для различных уровней компетентности экспертов.

Т.о. на основе шкалы  $Ш_0$  для каждой группы экспертов оказалось возможным ввести линейный порядок на множестве операций, задающий их «важность» для определения КСП, т.е. шкалу  $Ш_П$ .



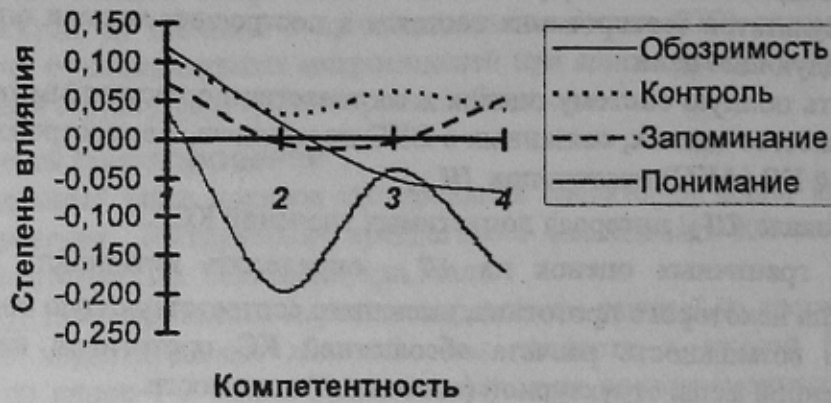


Рис. 1 Влияние уровня компетентности на корреляцию между сложностью отдельных операций над прототипами и ОКС.

### 3.2 Построение шкалы абсолютной КСП.

Анализ результатов тестирования показывает, что возможны ситуации, когда из-за субъективного фактора система оценок и значения ОКС противоречат как друг - другу так и условиям тестирования. Например: 1) совокупность явно больших по значению оценок имеет более низкую ОКС (или наоборот); 2) имеются различные значения ОКС для одинаковых систем оценок (или наоборот). Был сделан вывод, что более достоверной мерой КС в этом случае является ОКС, как интуитивная оценка, обладающая большей полнотой по отношению к системе заданных характеристик.

Задача построения шкалы АКС прототипов  $\mathcal{Ш}_A$  решалась следующим образом.

Пусть дано:

- шкала,  $\mathcal{Ш}_O$  т.е. ОКС для ряда прототипов;
- множество совокупностей значений оценок  $O'$ , т.е. декартово произведение

$$O' = O_1 \times \dots \times O_n; \quad (1)$$

- частичное отображение

$$F_o : \mathcal{Ш}_O \leftrightarrow O' \in O', \text{ при этом выполняется } \forall o_i'' < o_j'' \text{ для } i < j. \quad (2)$$

*Решение.* Строим функцию  $P_o$ , задающую отношение порядка на  $O'$  таким образом, что удовлетворяется (2). Регрессионный анализ для построения  $P_o$  в данном случае не возможен по причине того, что для  $O'$  имеются только неравенства, задающие отношения порядка в соответствии со шкалой  $\mathcal{Ш}_O$ . Исходя из этого, предлагается следующая эмпирическая форма построения функции  $P_o$ :

$$P_o(O_1, \dots, O_n) = \sum_{i=1}^n O_i * K_i. \quad (3)$$

Тут:  $K_i$  - уровни корреляции  $K_i(O_i, ОКС)$ , рассчитанные для ряда прототипов. Шкала  $\mathcal{Ш}_П$  в этом случае служит для задания сложности на множестве операций. Предложенная эмпирическая формула имеет высокий коэффициент корреляции ( $>0,94$ ) с ОКС, что оправдывает ее применимость.

### 3.3 Определение границ допустимой КСП

Определение на шкале  $\mathcal{Ш}_A$  границ допустимой КСП включает два этапа:

1). Определение нижней границы допустимой КСП:

- каждая оценка  $O_i$ , имеющая значение

$$O_i < 3, \quad (4)$$

определяет «слишком» сложный прототип;

- если в составе системы оценок хотя бы для одной оценки выполняется условие (4), то прототип объявляется «слишком» сложным;

- наибольшая по номеру в  $\Pi_A$  совокупность значений оценок, т.е. класс эквивалентности  $\mathcal{E}_2$ , для которого выполняется условие (4), есть точная нижняя грань допустимой КСП на  $\Pi_A$ ;

- прототип, принадлежащий классу эквивалентности  $\mathcal{E}_2$  и имеющий наибольшую относительную когнитивную сложность, объявляется нижней точной гранью множества допустимых КСП.

2). Определение верхней границы допустимой сложности:

- класс эквивалентности  $\mathcal{E}_5$ , такой, что верно

$$\forall O_i = 5, \quad (5)$$

может быть объявлен как «слишком» простой и является точной верхней гранью  $\Pi_A$ ;

- прототип, принадлежащий классу эквивалентности  $\mathcal{E}_5$  и имеющий наименьшую относительную сложность, объявляется верхней точной гранью КСП.

Классы эквивалентности  $\mathcal{E}_2$  и  $\mathcal{E}_5$  могут варьироваться, исходя из текущего состояния актуальных характеристик операторов.

### Заключение

На основе концепции адаптивного компьютера разработана методика тестирования пользователей и метод определения шкалы и допустимых границ абсолютной КСП прототипов, специфичных для данного оператора. Полученный результат обеспечивает возможность построения меры абсолютной КСП, определенной на модели предметной области, для произвольных прототипов.

### Литература

1. Павлов В.В. Системы человек - машина: проблемы и синтез. К.: Вища школа. Головное издательство. 1987. - 55 с.
2. Кочетенко Е.М. Побудова та алгоритми внутрішньо-комп'ютерної системи знань про людину-оператора в рамках концепції адаптивного комп'ютера. Сб. трудов первой межд. н.-п. конф. по программированию УкрПРОГ'98. К., 1998. - С. 487-492.
3. Росс. Д. Структурный анализ (SA): язык для передачи понимания. //Требования и спецификации в разработке программ. - М.: Мир, 1984. - С. 240-284.
4. Шиян А.А. Типы людей и их взаимодействие с компьютером. Сборник трудов первой межд. н.-п. конф. по программированию УкрПРОГ'98. К., 1998. - С.482-486.
5. Балашов К.В. Когнитивная сложность систем взаимодействующих процессов. Сб. научных трудов конференции по искусственному интеллекту «КИИ-94». Рыбинск. - 1994. Том 1, С. 154-158.
6. Григорьев А.В. О построении унифицированной концептуальной модели предметной области. Сборник трудов шестой межд. конф. «Знания - диалог - решения». Ялта, 1997. - Том.1., С. 427-434.
7. Григорьев А.В. Методы и средства работы с недоопределенными моделями в технологии мета-эвристических оболочек. Сборник трудов первой межд. н.-п. конф. по программированию УкрПРОГ'98. К., 1998. - С. 427-434.