

# ОЦЕНКА КОГНИТИВНОЙ СЛОЖНОСТИ МОДЕЛЕЙ

Григорьев А.В.

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

## Abstract

*Grigoriev A.V. Estimation cognitive complexity of models. Definitions of complexity in the various mathematical theories are considered. A method of an estimation cognitive complexity of models, generalizing line of definitions of complexity and oriented on peculiarity of offered conceptual model of subject area is offered.*

## Введение

К специфике работ, посвященных средствам оценки когнитивной сложности (КС) моделей [1], можно отнести наличие: концептуальной модели предметной области (КМ ПО) различной сложности; языка для описания моделей; меры для оценки и ограничения КС, основанной на интерпретации понятия структурной (системной) сложности. Анализ недостатков данных работ позволяет сделать следующие выводы: концептуальная модель предметной области (КМ ПО), используемая для расчета КС должна иметь обобщенный вид; мера КС моделей объектов, должна коррелировать не только с отдельными показателями системной сложности, но и системной сложностью вообще; мера КС должна иметь комплексный характер, т.е. строится на основе не только структурной (системной) сложности, но и учитывать положения других известных теорий определения сложности. Автором предложена более общая КМ ПО [8,9], чем используемые ранее [1]. Таким образом, актуальным представляется решение задачи разработки мер КС моделей объектов, основанные на положениях всех известных теорий сложности и ориентированные на предлагаемую КМ ПО.

## 1 Обзор известных теорий сложности

С целью обобщения понятия КС моделей рассмотрим существующие теории определения сложности. 1. Аксиоматическая концепция системной сложности [4]. Определяет сложность как неотрицательное число в соответствии с пятью аксиомами сложности: А) Подсистема не может быть сложнее системы; В) Сложность системы, являющейся параллельным соединением подсистем, равна сложности максимально сложной подсистемы; С) Сложность системы, являющейся последовательным соединением подсистем, не превышает суммы сложности этих подсистем; D) Сложность системы, состоящей из двух последовательных подсистем, охваченных обратной связью, не превышает суммы сложностей этих подсистем и обратной связи; E) Среди систем, удовлетворяющих аксиомам А-D, существуют системы минимальной (нулевой) сложности. Может использоваться для ограничения сложности процесса восприятия моделей человеком. 2. Теория сложности конечных автоматов и полугрупп [5]. Определяет сложность как наименьшее число групп, требуемое для декомпозиции полугруппы преобразований, основанной на таком представлении автомата, при котором каждый входной сигнал автомата индуцирует перестановку элементов в множестве его состояний. Существует аналогия между полугрупповой сложностью и сложностью САУ (см. ниже). 3. Теория сложности систем автоматического управления (ТС САУ) [6]. Цель теории - задание модели САПР САУ. Оперирует понятиями алфавита и операцией спивания как средства для создания сложных слов. Определяет шкалы сложности для технических заданий и описаний соответствующих

им прототипов. Слова, более длинные по размеру или полученные сшиванием более простых слов, определяются как менее весомые. В случае несравнимости слов порядок может быть установлен на основе экспертных оценок (метод диаграмм Хассе). Использует критерий сложности как средство оптимизации соотношения сложности технических заданий и прототипов. Вводит меру сложности проектирования, т.е. затрат на разработку, и рассматривает ее соотношение с типами задач проектирования и сложностью технических заданий. Использует понятие лексикографической сложности (см. ниже) для упорядочивания шкал сложности проектирования. Задаёт сложность процесса проектирования как лексикографическую сложность «траектории» движения САПР САУ, проходящей через множества прототипов. Позволяет проводить аналогию между сложностью САУ и полугрупповой сложностью. Удовлетворяет всем аксиомам концепции системной сложности, при этом отдельные аксиомы системной сложности рассматриваются как частные случаи требований к САУ. Данная теория может рассматриваться как обобщение понятия КС моделей до понятия *сложности процесса их проектирования*. 4. Сложность задач структурного синтеза (СЗСС) [7]. Определяет трудности формализации задач синтеза решений. Предполагает такие уровни: А) наличие готового решения с неизвестным набором параметров; Б) выбор структуры из конечного числа вариантов путем перебора; В) выбор варианта из конечного, но слишком большого для простого перебора множества; Г) выбор варианта из множества, мощность которого неизвестна; Д) поиск решения, не известного ранее. Теория СЗСС может рассматриваться как обобщение некоторых понятий ТС САУ. 5. Теория сложности алгоритмов и вычислений (ТСА и В) [8]. Опиерирует прямыми и двойственными, статическими и динамическими мерами сложности алгоритмов. Прямые меры сложности определяют сложность как затраты определенных ресурсов в случае обработки заданных данных. Пример прямых статических мер сложности: количество элементарных операций в алгоритме. Пример прямых динамических мер сложности: время работы с данным набором данных. Двойственная мера сложности показывает минимальные затраты ресурсов, необходимых для достижения нужного результата. Пример: минимальное время, необходимое для достижения результата. Вводится абстрактная мера сложности на множестве алгоритмов как семейство частично рекурсивных функций. Пример: временная мера сложности, определенная как взвешенная сумма шагов, если различные операции требуют различного времени выполнения. Рассматриваются асимптотически оптимальные двойственные меры для заданных прямых мер. Сопоставляя алгоритмам устройства для их реализации (машина Тьюринга) позволяет рассматривать обобщенную сложность Колмогорова как форму двойственной меры сложности. Данная теория может рассматриваться как обобщение понятия КС моделей до понятия *сложности процесса их моделирования*. 6. Теория сложности решения задач (ТСРЗ) [7]. Определяет сложность решения задачи, задаваемую в форме функции  $Q(N)$  для подсчета числа требуемых операций  $Q$  от размера задачи  $N$ . Предполагает такие классы задач: А) неразрешимые задачи ( $N=10^{93}$ ); Б) класс  $P$ , полиномиальной сложности; В) класс  $NP$ , позволяющий за полиномиальное время проверить правильность решения; Г) класс  $NP$  – полных задач, для которых неизвестен полиномиальный алгоритм точного решения, но любые две задачи могут быть сведены друг к другу за полиномиальное время. ТС примыкает как частный случай к ТСА и В. 7. Теория синтаксического упорядочивания (ТСУ) [9]. Существуют следующие пути для решения задачи построения частичного упорядочивания на множестве термов: рекурсивный, лексикографический, улучшенная рекурсивная декомпозиция, улучшенная лексикографическая декомпозиция, упорядочивание Кнута-Бендикса. Все указанные упорядочивания могут быть расширены введением статуса или сигнатур, содержащих АС-символы. Рассмотрим их

подробнее. *Лексикографическое отношение порядка* на двух последовательностях термов соответствует отношению порядка первой неэквивалентной пары термов, имеющих одинаковые номера в последовательностях. Если имеет место полная эквивалентность для одной из последовательностей, то более короткая цепочка предполагается и меньшей по порядку. *Отношение порядка на мультимножествах*, т.е. упорядоченных последовательностях, допускающих многократное вхождение элементов. Отличается от лексикографического отношения порядка тем, отношение эквивалентности должно проверяться для всех пар термов с произвольными номерами в последовательностях. *Рекурсивное упорядочивание* задает отношение для двух термов, являющихся значениями двух отличных функций, заданных над некоторыми последовательностями аргументов. Порядок в этом случае определяется упорядочиванием на множестве идентификаторов функций и множествах последовательностей аргументов. *Упорядочивание на множестве функций* может проводиться по Кнуту-Бендиксу. В этом случае на множестве функциональных символов задается мера веса, определяющая отношение порядка. *Упорядочивание со статусом* предполагает, что на множестве функциональных символов определяется функция, задающая тип упорядочивания: мультимножественный или лексикографический. *Упорядочивание с АС-термами* предполагает, что на множестве функциональных символов выполняются аксиомы ассоциативности и коммутативности. Все методы ТСУ являются средствами для определения сложности проектирования или сложности моделирования объектов.

## 2 Оценка когнитивной сложности моделей в инструментальной оболочке для построения интеллектуальных САПР

### 2.1 Особенности предлагаемой КМ ПО

Унифицированная КМ ПО (УКМ ПО) инструментальной оболочки для построения интеллектуальных САПР описана в работах [2,3]. УКМ ПО строится в форме семиотической модели [10] и включает множества: базовых элементов, правил построения синтаксически правильных формул, аксиом и правил вывода семантически правильных формул. Особенности УКМ ПО можно разделить на две группы, исходя из того, влияют они на решение задачи определения КС моделей или не влияют. Перечислим особенности, оказывающие влияние на решение задачи определения КС моделей: в состав модели входят структурная составляющая, а так же макро- и микромоделные функциональные составляющие; модель объекта есть совокупность интенциональных (именованных реляционных [10]) отношений различных типов, задающих иерархию отношений; базовыми отношениями иерархии типов отношений являются отношения принадлежности значений свойствам; модель объекта есть совокупность отношений высшего уровня иерархии - замкнутых ограниченных по сложности моделей внутренних сред прототипов; совокупность сред образует модель посредством отношений, связывающих между собой внутренние и внешние среды отдельных прототипов; по составу элементов модель счетная и ограниченная; модель прототипа принадлежит некоторому уровню недоопределенности семиотической модели; уровень определяется значениями трех атрибутов модели, задающими подуровни ее «грубости», неполноты и альтернативности.

Исходя из структуры семиотической модели, можно выделить в ее составе такие подмножества:  $F$  – подмножество функциональных символов, входящих в состав множества базовых элементов;  $S$  – подмножество синтаксически правильных формул;  $A$  – множество аксиом;  $V$  – подмножество семантически правильных формул. При этом выполняется:  $A \subseteq V \subseteq S$ . Определение системной (структурной) сложности моделей эквивалентно заданию отношения порядка на множестве  $S$ .

## 2.2 Мера сложности на множестве отношений

Рассмотрим совокупность отношений с точки зрения их градации по «уровню глубины», в соответствии с подходом [11]. Применим данный подход с учетом специфики предлагаемой КМ ПО. Рассмотрим множество функциональных символов или множество типов интенциональных отношений

$$F = \{f_j\}_j, j = \overline{1, N}. \quad (1)$$

В множестве синтаксически правильных формул  $S$  можно выделить подмножества, задающие внутреннюю структуру любого типа функционального символа, построенную на основе «нижележащих» типов функциональных символов:

$$T = \{T_j\}_j, j = \overline{1, N}. \quad (2)$$

В составе  $T$  имеют место подмножества:

1) Базовые типы отношений

$$T_b = \{T_j\}_j, j \leq N_b; \quad (3)$$

2) Небазовые типы отношений

$$T_n = \{T_j\}_j, N_b < j \leq N. \quad (4)$$

В нашем случае  $T_b$  содержит единственный тип, определяемый функциональным символом  $f_1$  - «отношение принадлежности значений типу данных».

$T_n$  определяется всеми прочими производными типами отношений (границы блоков, связи и т.д.). Небазовые типы есть сложные слова, построенные с помощью реляционных операций: «&» (одновременное существование во времени); «∨» (альтернативное существование во времени); «.» - произведения (отношение двух цепочек слов, имеющих некоторое общее подмножество слов); «-1» - обращения (изменение порядка следования двух слов на обратный); операций, производных от перечисленных. В частности: отношения принадлежности значения свойству есть логическое отношение «∨» над значениями; связь есть множество, составленное отношением произведения «.» над парой свойств, принадлежащей двум блокам; среда есть объединение связей в множество через отношение логического «&» и т.д. Каждому типу  $T_i$  соответствует грамматическое правило, задающее все возможные идентификаторы структур отношений, составляющих данный тип (список прототипов):

$$T_i = \bigvee_{k=1}^{K_i} \{O_{ik}\}. \quad (5)$$

Размерность множества прототипов соответствуют следующим уровням теории СЗСС: А) наличие готового решения с неизвестным набором параметров; Б) выбор структуры из конечного числа вариантов путем перебора; В) выбор варианта из конечного, но слишком большого для простого перебора множества.

Отдельные прототипы  $O_{ik}$  предполагают одновременное использование ряда функциональных символов - «подтипов»  $T_{ik} = \big\&_{j=1}^{K_{ij}} \{f_{ik}^j\}$ . Каждый подтип представлен подмножеством своих прототипов (объединение по «&»), выбранным из состава возможных прототипов  $O_{ik}^{jl}$  (задается объединением по «∨»). В пределах заданной структуры  $O_{ik}$  будем различать «экземпляры отношений», т.е. отдельные отношения, возможно принадлежащие одному и тому же подтипу и прототипу, но отличающиеся дополнительной идентификацией. Т.о. можно записать соотношение:

$$O_{jk} \leftrightarrow (T_{jk} = \{T_{jk}^i\}_i, i < j, i \in I_{jk}, I_{jk} \subset I_j = \{1, 2, \dots, j\}, T_{jk}^i = \{O_{jk}^{il}\}_l, l \in L_{jk}^i). \quad (6)$$

Формально к множеству  $T$  можно подойти следующим образом: отдельный прототип есть И-дерево; тип как совокупность прототипов есть И-ИЛИ-дерево. Системная сложность для различных отношений внутри прототипа или типа может задаваться, исходя из следующей трактовки аксиом системной сложности: А') Тип сложнее, чем подтип или прототип; В') Объединение значений по « $\vee$ » имеет сложность, равную максимальной сложности простых выражений; С') Объединенных значений операциями « $\&$ » и « $\cdot$ », имеет сложность, не превышающую суммы сложности простых выражений; D') Сложность выражения, состоящего из обращения простого отношения («-1»), не превышает суммы сложностей простого отношения и отношения обращения; D'') Сложность выражения, состоящего из двух простых выражений, объединенных по И (« $\&$ ») или по ИЛИ (« $\vee$ »), при том, что одно из простых понятий есть исходное сложное выражение (рекурсия), не превышает суммы сложностей этих понятий и отношения рекурсии; E') Символы значений типов свойств есть минимальное по сложности выражение. Таким образом, для оценки сложности на множестве отношений, составляющих модель в данной форме представления, возможно применить аддитивную меру сложности.

### 3 Построение меры абсолютной когнитивной сложности моделей

#### 3.1 Общее описание метода

Для упорядочивания по сложности совокупности отношений в пределах отдельно взятого прототипа способом, позволяющим удовлетворить данную аксиоматику, можно предложить следующую модификацию метода диаграмм Хассе [6]. Суть метода диаграмм Хассе заключается в упорядочивании сложных слов по двум основным принципам: 1) более длинное слово проще короткого; 2) слово, включающее в себя другое слово, определяется как более простое. В случае наличия несравнимых слов привлекаются эксперты. Положение более простых слов в сложном слове безразлично, важен только состав слов. Допускается вхождение более простых слов в любое количество более сложных слов. Существует алфавит простейших слов, среди которых имеется слово  $K0$ , большее любого из слов. Отдельный узел диаграммы или сложное слово первоначально имеет ряд направленных дуг, определяющих отношения порядка сложного слова со словами, входящими в его состав. В результате упорядочивания диаграмма разбивается на ряд уровней, отличающихся определенной длиной слов. На каждом уровне все слова упорядочены по возрастанию сложности на основе задаваемых впрямую экспертных оценок. Т.о. определяется полный или линейный порядок на множестве слов, т.е. шкала  $Ш_1$ .

Модификация метода состоит в следующем: 1) Предполагается изменение знака отношения сложности на противоположный, т.е. отношение  $O1$  считается более сложным чем  $O2$ , если  $O1$  включает в себя  $O2$ , а не наоборот. Соответственно,  $K0$  есть наименьшее по сложности слово и определяется экспертным путем, максимального по сложности слова не существует. 2) Сложность слов оценивается посредством аддитивной меры структурной сложности, т.е. по взвешенному числу составляющих их функциональных символов различных уровней иерархии. Аддитивная мера структурной сложности обеспечивает выполнение всех аксиом системной сложности в приведенной выше трактовке. 3) Порядок сложности для ряда несравнимых слов первоначально задается системой неизвестных коэффициентов. 4) Сложность каждого отдельно взятого слова из разряда несравнимых рассчитывается по совокупности данных об общей системной сложности для ряда прототипов. 5) В качестве величин сложности прототипов могут использоваться оценки их когнитивной сложности, что делает возможным использование предлагаемой оценки системной сложности в качестве меры КС моделей; когнитивная сложность прототипов определяется

экспертным путем; б) В случае наличия в модели недоопределенностей (NIL), порядок их сложности определяется следующими принципами, определенным ранее в [3]: А) каждый NIL, заданный в составе множества элементов некоторого отношения декомпозиции, связывающего различные уровни модели, имеет минимальный вес в пределах данного множества элементов; Б) множество NIL одного уровня упорядочиваются между собой как ряд несравнимых символов; С) минимальный по сложности NIL нижнего уровня имеет минимальный (нулевой) вес среди цепочек длины «1», т.е. есть является символом «K0»; Д) полная цепочка «минимальных» NIL имеет минимальный (нулевой) вес среди цепочек данной длины; Е) любая полностью определенная цепочка функциональных символов имеет вес больший, чем любая недоопределенная цепочка. Пример упорядочивания приведен на рис. 2.

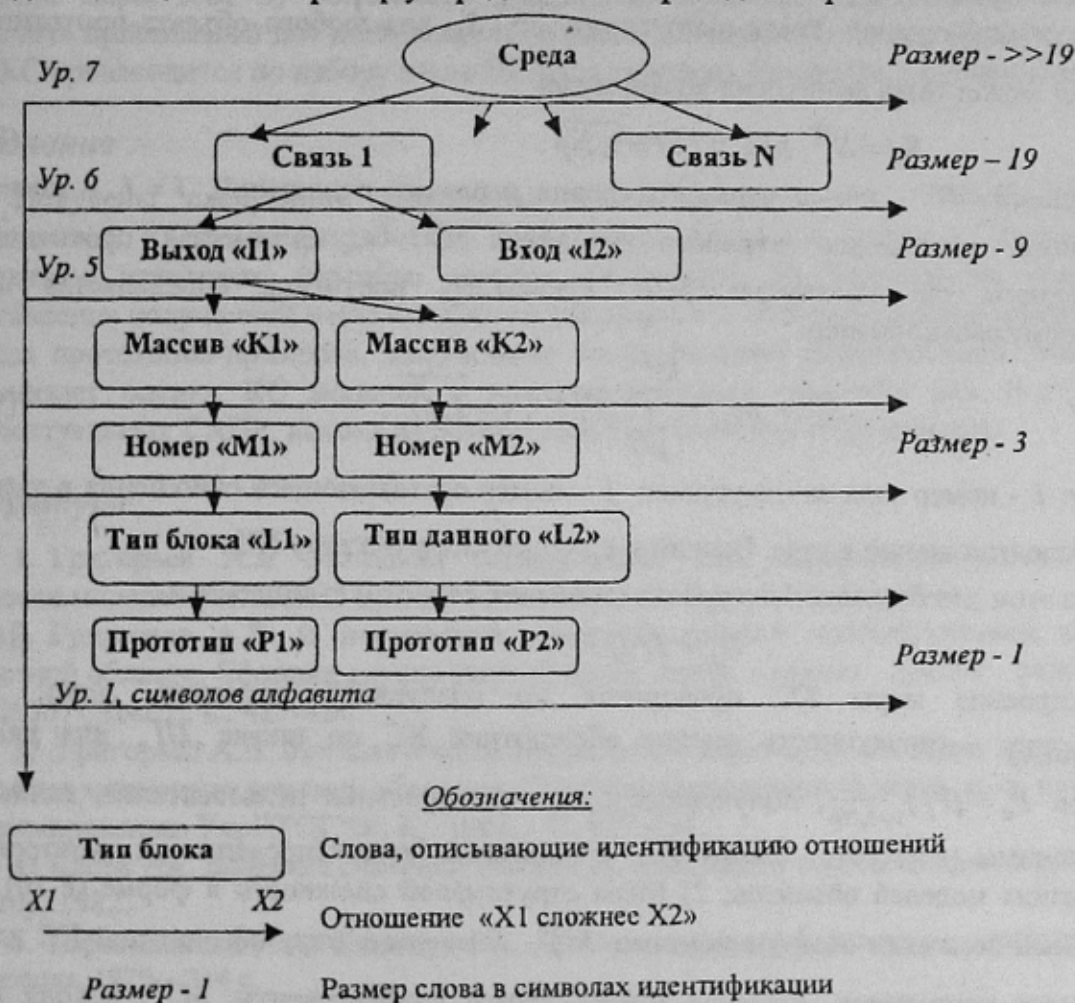


Рис. 2 Пример построения меры сложности на упрощенном описании внутренней среды блока-прототипа.

Предлагаемый подход предполагает упорядочивание по весу на множестве функциональных символов и есть вариант упорядочивания по Кнуту-Бендиксу. Взаимное упорядочивания двух несравнимых сложных слов в случае наличия у них общей «нижележащей» ветви выполняется, фактически, лексикографическим способом, а при наличии в их составе ряда непересекающихся нижележащих ветвей упорядочивание выполняется фактически мультимножественным способом. Таким образом, на функциональных символах неявно определяется функция статуса упорядочивания, задающая тип упорядочивания: мультимножественный или лексикографический.

### 3.2 Детальное изложение метода

Пусть дана некоторая замкнутая ограниченная модель внутренней среды прототипа  $P$ , представленная множеством экземпляров отношений различных уровней общности. Данные отношения представлены таблично и могут быть упорядочены по взаимному включению в соответствии со шкалой  $Ш_1$ . Вес несравнимых слов, задан системой неизвестных коэффициентов:

$$M = \{M_i^m\}_{i=1, Im}^{m=1, W}. \quad (7)$$

Здесь:  $W$  - число уровней типов отношений, имеющих место в прототипе;  $m$  - номер уровня иерархии отношений,  $i$  - номер отношения в пределах уровня;  $Im$  - число отношений на уровне. Общее число отношений в диаграмме будем обозначать как  $I$ . Тогда структурная (системная) сложность  $S_l$  для любого объекта-прототипа  $P_l, l = \overline{1, Np}$  может быть вычислена по формуле:

$$S_l = S_1^W \text{ для } \forall l: l = \overline{1, Np}. \quad (8)$$

Здесь:  $m = W$  - номер верхнего уровня иерархии отношений,  $i = I$  - номер единственного отношения верхнего уровня;  $S_1^W$  - сложность прототипа, представленного как замкнутая среда. Сложность прототипа определяется по рекуррентному соотношению:

$$S_i^m = M_i^{m*} \sum_{j=1}^{K_i^{m-1}} S_j^{m-1}, i = \overline{1, I_i^m}. \quad (9)$$

Тут:  $i$  - номер узла  $m$ -го уровня;  $j$  - номер составляющего отношения в узле;  $K_i^{m-1}$  - число отношений в узле. Всякий узел содержит структуру NIL.

При этом для базового 1-го уровня (значений свойств) выполняется

$$S_j^1 = 1; \forall j. \quad (10)$$

Построение меры КС проводится по следующей методике. Дано: 1)  $A = \{A_l\}_{l=1, Np}$  - совокупность оценок абсолютной КС по шкале  $Ш_a$  для ряда прототипов  $P_o = \{P_l\}_{l=1, Np}$ , полученные путем тестирования пользователей; данные оценки должны отражать сложность восприятия, проектирования и контроля моделирования моделей объектов; 2) Мера структурной сложности в форме (8-10) с неизвестными весовыми коэффициентами  $M_i^m$ . Получить: меру абсолютной КС для произвольных прототипов. Решение задачи. Ищем коэффициенты  $M_i^m$ , исходя из системы уравнений:

$$A_l = S_l^{m^l} \text{ для } \forall l: l = \overline{1, Np}. \quad (11)$$

По числу примеров в наборе тестов возможны следующие случаи: 1)  $Np \geq Nf$  - возможен полный расчет  $M$  при условии совпадения числа примеров с общим количеством прототипов всех возможных типов отношений, входящих в состав набора примеров. Данное условие практически не выполнимо. 2)  $Np = W$  - возможен частичный расчет  $M$ , при этом вес всех отношений одного размера (типа) эквивалентен между собой:

$$M = \{M_m\}_{m=1, W}; M_m^l = M_m, \forall i = \overline{1, Im}; \quad (12)$$

В этом случае система уравнений (11) вырождается в линейную и не требует сложных методов решения. Соответствует обычному подходу к структурной сложности, отличие

состоит только в составе и порядке относительной сложности типов отношений. Предлагается принять случай 2 за основу упрощенной меры АКС. 3)  $Np < W$  - - возможен частичный расчет  $M$ , при этом часть коэффициентов нижележащих типов отношений из случая 2 определяются как константы (0 или 1), тогда

$$M = \{M_m\}_{m=1, Np}; M_m^i = 1, i > Np. \quad (13)$$

На основе случаев 2 и 3 разрабатывается программная система, способная: 1) обеспечивать расчет совокупности коэффициентов когнитивной сложности типов отношений, составляющих модель прототипов в инструментальной оболочке по построению интеллектуальных САПР; 2) оценить КС любого прототипа на основе известной меры КС; 3) обеспечить упрощение представления модели с целью обеспечить приемлемый для пользователя уровень КС моделей. Расчет коэффициентов меры КС производится по набору весов КС ряда тестовых примеров, упомянутых в [1].

### Заключение

Получены следующие результаты: 1) В соответствии с особенностями предложенной УКМ ПО разработана новая мера сложности моделей, являющаяся обобщением известных способов оценки сложности; 2) Предложена методика представления полученной меры сложности как меры КС на основе оценок абсолютной КС ряда прототипов-примеров, полученных тестированием пользователей. Мера КС обеспечивает оценку КС моделей в инструментальной оболочке для построения интеллектуальных САПР, исходя из результатов тестирования пользователя.

### Литература

1. Григорьев А.В. Методика тестирования для определения когнитивной сложности моделей различных предметных областей. В этом же сборнике.
2. 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.
3. 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.
4. Касти Дж. Большие системы: связность, сложность и катастрофы: Пер. с англ. - М.: Мир, 1982.
5. Первозванский А.А., Гайцгори В.Г. Декомпозиция, агрегирование и приближенная оптимизация, 1979. - 344 с.
6. Солодовников В.В., Тумаркин В.И. Теория сложности и проектирование систем управления. - М.: Наука. Гл. ред. физ.-мат. лит., 1990. - 186 с.
7. Норенков И.П. Разработка систем автоматизации проектирования. М.: МГТУ им. Э.Н.Баумана, 1994. - 207 с.
8. Бургин М.С. Меры сложности в аксиоматической теории алгоритмов. //Методы проектирования интеллектуальных прикладных программных систем. К.: ИК, 1992. - С. 60-67.
9. Коваль Л.В., Коваль Ю.В. Методы упорядочивания термов в задачах пополнения. // Там же. - С. 67-78.
10. Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. - 288 с.
11. Хинтикка Я. Логико эпистемологические исследования. // Под ред. Садовникова В.Н. и Смирнова В.А. М.: Прогресс, 1980. - 450 с.