

# СЕМАНТИЧЕСКОЕ ПРОТОКОЛИРОВАНИЕ ПРОЦЕССОВ ОБРАБОТКИ ЗНАНИЙ

Ивашенко В. П.

Кафедра интеллектуальных информационных технологий, Факультет информационных технологий и управления, Белорусский государственный университет информатики и радиоэлектроники  
Минск, Республика Беларусь  
E-mail: ivashenko@bsuir.by

*В статье рассматриваются виды процессов обработки знаний, подходы, структуры и алгоритмы протоколирования событий в этих процессах*

## ВВЕДЕНИЕ

Важным качеством интеллектуальных систем, управляемых знаниями, является способность объяснять своё решение, на основании каких знаний оно было принято. Способность объяснить своё решение, наличие знания о нём и его предпосылках, соответствующие интроспективные способности являются одними из существенных отличительных признаков систем, обладающих признаками искусственного сознания, от систем искусственного интеллекта [1]. С другой стороны решение задач в интеллектуальных системах, способных к обучению, требует интеграции различных моделей и методов решения задач. Таким образом, в интеллектуальных системах сочетание способностей управляться знаниями и обучаться требует универсальных механизмов, которые позволили бы обеспечить возможность объяснения решений, получаемых конкретной интеллектуальной системой.

## Задачи СЕМАНТИЧЕСКОГО ПРОТОКОЛИРОВАНИЯ

Сложность задачи семантического протоколирования (СП) определяется сложностью процессов обработки знаний (ПОЗ), протекающих в интеллектуальной системе. ПОЗ могут циклическими и ациклическими. Ациклические процессы могут быть: последовательно-линейными, древовидными, (ациклическими) сетевыми. Циклические процессы могут быть: последовательно-петлевыми, древовидно-петлевыми, (циклическими) сетевыми. При наличии возможности учёта времени сложность ПОЗ может выражаться временной сложностью. Время здесь понимается в обобщённом смысле как порядковый номер состояния процесса (обработки знаний). Таким образом, для учёта времени, требуется иметь некоторое вполне упорядоченное множество или решётку «номеров», (взаимно) однозначно сопоставленных состояниям процесса. Если «номера» являются частью соответствующих им состояний, то время является собственным временем процесса, иначе – несобственным. В случае наличия собственного или несобственного времени СП может осуществляться с заданной частотой (плотностью)

на временном интервале. Также, в зависимости от наблюдения событий и доступности вычислительных ресурсов, среди протоколируемых в рамках процессов событий будем различать: последовательные, логически независимые, физически независимые. Для протоколируемых ПОЗ рассматриваются следующие задачи: построение протокола процесса для нового добавляемого события, итерирование исходных и предшествующих событий, выявление предшествования между парой событий, выявление наибольшего множества общих предшествующих событий множества событий и другие.



Рис. 1 – Пример sc-цепочки и sc-мультицепочки

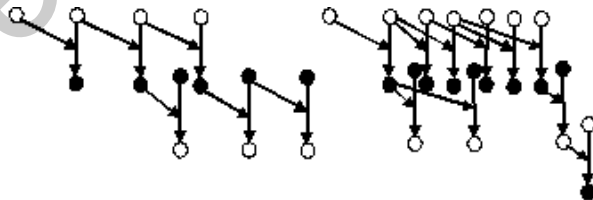


Рис. 2 – Пример sc-метацепочки и sc-метамультицепочки

В зависимости от вида и сложности ПОЗ протокол формируется в виде структур различной сложности, которые в модели унифицированного семантического представления знаний могут быть представлены ключевыми узлами и (бинарными) отношениями или такими конструкциями как sc-мультицепочки и sc-цепочки (структурное кодирование (см. рис. 1 и 2)). К таким структурам относятся: триплетно-свободный биномиальный иерархический стек (для событий последовательно-линейных процессов), триплетно-свободное биномиальное иерархическое дерево (для ациклических процессов с последовательными событиями), «пирамидальная структура» (для древовидных процессов с логически независимыми событиями (см. рис 3 и 4)) и «пирамидальный граф» (для ациклических сетевых процессов с логически независимыми событиями), основан-

ные на приближениях треугольника Серпинского (каждый треугольник в которых соответствует вершине), гиперкуб (для процессов с физически независимыми событиями) и их аналоги для циклических процессов.

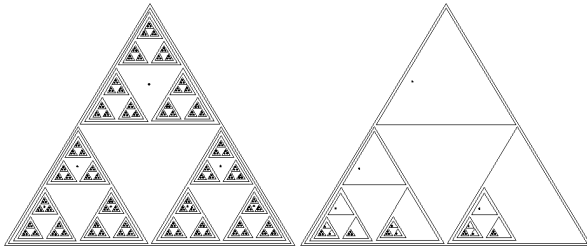


Рис. 3 – Полная пирамидальная структура и хранение в ней разреженного (фибоначчиева) дерева событий (высота  $\ln(n)/\ln(\phi)$ )

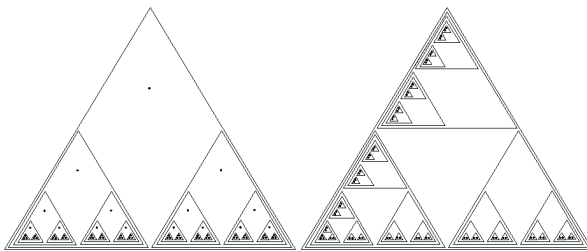


Рис. 4 – Хранение сбалансированного дерева и облиственной ветви событий в пирамидальной структуре (высоты  $\ln(n)/\ln(2)$ )

Триплетно-свободный биномиальный иерархический стек — структура, состоящая из звеньев. Каждое звено хранит коэффициент (1 или 2) увеличения количества звеньев-потомков по сравнению с предыдущим звеном, наименьший индекс события, хранимого в звеньях-потомках или в нём самом, а также ссылается на протоколируемое событие или корневое звено биномиального иерархического (под)стека потомков, на родительские, предыдущее и следующее звенья (если таковые есть). Индекс события — имя (счётчик) в системе именования, (возможно) соответствующее некоторому онтологическому базису [2]. Для каждого запротоколированного события имеется связка, связывающая его со звеном протокола, в котором оно находится.

Алгоритм процедуры построения протокола последовательно-линейного процесса для нового добавляемого события сводится к следующим шагам.

1. Получить новое событие, его индекс и первое звено (вершину) стека.
2. Если стек непустой и если в паре первых двух звеньев, последующее за первым звено имеет коэффициент 1, то требуется их замена.
3. Если требуется замена первых двух звеньев, то начиная со следующей за ними пары звеньев, пока между парой из соседних нечётного и чётного звеньев коэффициент увеличения равен единице (в чётном звене), заменить каждые такие звенья каждой па-

ры на объединённое звено, в котором: коэффициент увеличения равен 2, индекс равен индексу в подлежащем замене чётном звене, корневым звеном потомков является копия корневого звена потомков подлежащего замене нечётного звена, копия звена подлежащего замене чётного звена следует за последним звеном, являющимся копией последнего звена в последовательности копий всех звеньев последовательности звеньев, начинающейся от корневого звена потомков подлежащего замене нечётного звена, родительским звеном в копии корневого звена потомков указывается объединённое звено, предыдущим звеном во всех копиях звеньев — соответствующая копия предыдущего звена. Следующее звено за заменёнными парами звеньев, если такое есть, заменяется на звено с коэффициентом 1.

4. Если требуется замена первых двух звеньев, то они заменяются на звено с коэффициентом 2, в котором: индекс равен индексу второго звена, корневым звеном потомков становится копия первого звена (родителем в которой становится это новое звено, а следующим элементом — копия второго звена, предыдущим звеном в которой в свою очередь становится копия первого), следующим звеном становится третье звено (если таковое было), либо то звено, на которое оно было заменено.
5. Добавить в стек в качестве вершины новое звено с коэффициентом 1, указанным индексом и протоколируемым событием, следующим за вершиной звеном является прежняя вершина стека (если таковая была), или звено, на которое она была заменена.

Выявление исходных, предшествующих и итерирование событий осуществляется последовательно по индексам или бинарным поиском дихотомически в протоколе (в случае последовательно-линейного процесса это требует полилогарифмического времени, если пренебречь временем доступа к одному элементу).

## ЗАКЛЮЧЕНИЕ

Предлагаемые механизмы СП призваны обеспечить интеллектуальную систему интерфейсом и универсальными средствами объяснения запротоколированных ПОЗ при построении решений задач.

1. McDermott, D. Mind and Mechanism. — Cambridge (Mass), MIT Press. 2001.
2. Ивашенко, В. П. Модели обработки информации в интеллектуальных системах, основанных на семантических технологиях / В. П. Ивашенко, А. С. Бельчиков, А. П. Еремеев // Информационные технологии и системы 2016 (ИТС 2016) : материалы международной научной конференции (БГУИР, Минск, Беларусь, 26 октября 2016) / редкол. : Л. Ю. Шилин [и др.]. — Минск: БГУИР, 2016. — С. 106—107.