

УПРАВЛЕНИЕ КОМБИНИРОВАННЫМ ВЫВОДОМ В ДИСКРЕТНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ

А. Я. Фридман^а, доктор техн. наук, профессор

^аИнститут информатики и математического моделирования технологических процессов КНЦ РАН, Апатиты, РФ

Постановка проблемы: из-за больших объемов информации, хранящейся в базах данных и знаний интеллектуальных систем, степень отличия механизма обработки этой информации от полного перебора считается мерой интеллектуальности подобных систем. На сокращение перебора, в частности, направлены различные способы управления выводом, в том числе ранее предложенный автором метод управления прямым выводом в интеллектуальных системах с дискретными доменами переменных путем анализа эвристических индикаторов хода вывода, использующих некоторые подмножества этих доменов. Однако этот метод однонаправленный, что не позволяет ускорить вывод за счет своевременного изменения его направления и эффективного разрешения конфликта. Цель настоящей работы состоит в распространении того же теоретико-множественного подхода на задачи управления комбинированным выводом и разрешением конфликта для сокращения перебора вариантов продолжения вывода. **Результаты:** разработана динамическая стратегия управления комбинированным детерминированным и вероятностным выводом в дискретных продукционных интеллектуальных системах, анализирующая структуру следствий из входящих в конфликтное множество продукций и сравнивающая ее со структурой предпосылок цели вывода, чтобы выбрать продукцию, применение которой способно в максимальной степени подтвердить истинность текущей цели. Известно, что динамические стратегии обеспечивают большую гибкость, чем встроенные, в которых ход выбора предопределен априорно. Предложенный подход к представлению и анализу информации в интеллектуальных системах отличается от существующих тем, что в нем оперативно учитывается внутренняя структура данных и знаний интеллектуальной системы. Это позволяет повысить скорость вывода в дискретной интеллектуальной системе. **Практическая значимость:** представленные в работе правила управления применением продукций позволяют ускорить процесс достижения цели вывода при детерминированном, вероятностном и нечетком представлении информации в интеллектуальной системе.

Ключевые слова — дискретная интеллектуальная система, комбинированный вывод, управление применением продукций.

Введение

Задача в интеллектуальных системах (ИнС) в самом общем виде определяется как пара (S, G), где S — исходная информация (source — что дано), а G — целевая информация (goal — что найти, получить). В качестве S и G могут служить совокупности известных фактов, ситуаций и состояний. Для конкретности далее исследуются ИнС с продукционной базой знаний (БЗ).

Информация о задаче и ходе ее решения хранится в базе данных (БД), а о способах изменения данных — в БЗ. Каждое правило (продукция) состоит из двух частей. В левой описывается группа фактов или знания о текущей ситуации, а в правой содержатся действия, которые должны быть выполнены, если в БД найдется информация, соответствующая левой части правила. Действия правила могут состоять в модификации набора фактов в БД, например, в добавлении нового факта. Эти новые факты в свою очередь могут быть в дальнейшем использованы для сопоставления с левыми частями других правил. Отсюда следует, что применение правила состоит из двух этапов: сопоставления левой части правила с текущим состоянием БД и (в случае успешного сопоставления) выполнения действий, указанных в правой части правила.

Процесс применения правил порождает цепочку вывода. Существуют два основных способа ее порождения. Один из них получил название прямого вывода, а другой — обратного. При прямом выводе поиск новой информации происходит в направлении стрелок (импликаций), разделяющих левые и правые части правил. Решатель ИнС использует информацию из левых частей, чтобы вывести информацию, содержащуюся в правых частях. Поэтому иногда прямой вывод называют поиском в пространстве состояний ИнС.

Обратный вывод основан на идее сведения задачи к набору более мелких подзадач, решение которых в совокупности эквивалентно решению исходной задачи. Они декомпозируются дальше, и т. д., в результате первоначальная цель реализуется последовательностью элементарных действий, например запросами информации об истинности или ложности исходных данных, нужных для достижения цели вывода.

Наиболее гибкий способ вывода — комбинированный, когда режимы прямого и обратного вывода чередуются по мере исчерпания возможности получения новых фактов в текущем направлении вывода. Комбинированный вывод реализован во многих ИнС, самая известная из них — экспертная система PROSPECTOR [1].

Различные способы управления выводом способствуют сокращению перебора, в том числе предложенный авторами [2] метод управления прямым выводом в интеллектуальных системах с дискретными доменами переменных путем анализа эвристических индикаторов хода вывода, использующих некоторые подмножества этих доменов. Однако из-за однонаправленности этого метода невозможно ускорить вывод за счет своевременного изменения его направления и эффективного разрешения конфликта. В настоящей работе предлагается сократить перебор вариантов продолжения вывода на основе того же теоретико-множественного подхода к задачам управления комбинированным выводом и разрешением конфликта.

Конфликт в ИнС диагностируется решателем (интерпретатором), когда имеется несколько вариантов продолжения вывода и нужно выбрать наиболее перспективное из них [3]. Чаще всего конфликт возникает, если на некотором шаге вывода $v = 1, 2, \dots$ появляется несколько продукций, предпосылки которых подтверждаются текущим множеством фактов, хранящихся в БД. Эти продукции называют *конфликтным множеством* (CS_v — conflict set).

Существуют следующие основные способы разрешения конфликта.

1. Правилам заранее приписываются приоритеты, и между ними устанавливается отношение общего порядка. Из CS_v выбирается правило с максимальным приоритетом.

2. На множестве правил устанавливается частичный порядок, и выбирается подмножество правил с максимальным приоритетом. Например, правило P_j предпочтительнее правила P_l тогда и только тогда, когда левая часть правила P_l (представленная как множество фактов) является подмножеством левой части правила P_j . Первым применяется правило с более длинным списком фактов в левой части. Такой способ разрешения конфликта получил название «специальный случай».

3. На фактах в БД задается отношение полного порядка. Из CS_v выбирается та продукция, в левой части которой используются данные с максимальным приоритетом. Примером такого подхода служит приписывание возраста фактам в БД, тогда из CS_v выбирается продукция, использующая наиболее «молодые» факты.

4. На данных из БД задается отношение частичного порядка, с учетом которого разрешается конфликт, как и в способе 3.

5. Случайный выбор используют в случаях, когда после применения других способов разрешения конфликта не происходит выбора ни одного правила из CS_v .

Общий недостаток перечисленных способов состоит в том, что они не учитывают внутреннюю

структуру заложенных в ИнС знаний. Поэтому ниже предлагается способ управления выводом, анализирующий структуру следствий из входящих в CS_v продукций и сравнивающий ее со структурой предпосылок цели вывода, чтобы выбрать продукцию, применение которой позволяет ускорить процесс достижения этой цели.

Формализация задачи и алгоритм комбинированного детерминированного вывода

Пусть набор переменных x_i , которые могут быть результатами вывода, задан в универсуме $U = X_1 \times X_2 \times \dots \times X_n$, где X_i — дискретные или дискретизированные домены (множества допустимых значений) соответствующих переменных.

Тогда любой факт, хранящийся в БД, можно представить в виде свидетельства «за» некоторые допустимые значения переменной:

$$x_i \in \Phi_i^+, \Phi_i^+ \subseteq X_i, \quad (1)$$

либо в виде свидетельства «против» некоторых допустимых значений:

$$x_i \notin \Phi_i^-, \Phi_i^- \subseteq X_i. \quad (2)$$

Появление факта вида (2) при детерминированном выводе можно интерпретировать как «сужение» домена, приняв

$$X_i^{v+1} = X_i^v \setminus \Phi_i^- \quad (3)$$

и $X_i^0 = X_i$, поэтому в следующем разделе принято, что в БД есть только факты вида (1).

Продукции в БЗ ИнС обычно хранятся в формате «условие — действие» [2]:

$$A_j \supset B_j, \quad (4)$$

где A_j и B_j — конъюнкции утверждений вида (1) или (2). Для простоты будем считать, что каждое из этих утверждений есть высказывание. Цель вывода G задается как такая же конъюнкция.

Алгоритм комбинированного детерминированного вывода

Шаг 1. Интеллектуальная система начинает работать в режиме обратного вывода, когда каждый из конъюнктов начальной (глобальной) цели сравнивается с конъюнктами правых частей B_j продукций (4) и, если найдена продукция, позволяющая определить истинность этого конъюнкта, вместо него в формулу цели подставляются конъюнкты, содержащиеся в левой части A_j такой продукции. Если ни одна продукция не подходит, ИнС переходит в режим прямого вывода.

Шаг 2. При прямом выводе левые части продукций (4) сравниваются с фактами вида (1), имеющимися в БД ИнС. Начальное множество фактов вводится в ИнС извне (пользователем или

автоматически). Если найдена продукция, левая часть которой A_j подтверждается текущим набором фактов (т. е. истинны все содержащиеся в ней конъюнкты), и эта продукция исполняется, то в БД добавляются все факты вида (1), содержащиеся в B_j . При появлении фактов вида (2) домены соответствующих переменных сужаются по формуле (3). Если нет подходящих продукций, то ИнС возвращается в режим обратного вывода, причем в качестве цели выбирается только часть конъюнктов, содержащихся в текущей цели вывода (иначе обратный вывод сразу остановится).

При реализации описанного выше комбинированного вывода возможны две проблемы: 1) конфликт между продукциями при прямом выводе и 2) неоднозначность выбора подмножества конъюнктов текущей цели для продолжения обратного вывода. Для разрешения подобных проблем предназначены стратегии управления выводом, обычно имеющие эвристический характер [2]. Ниже предлагается одна из таких стратегий, основанная на идеях анализа состава атрибутов и сравнения мощностей множеств, которые предложены авторами для расчета меры неопределенности следствия в рамках алгебры кортежей [4, 5].

Стратегия управления детерминированным выводом

Алгоритм разрешения конфликта

Пусть на некотором шаге v вывода конфликтное множество имеет вид

$$CS_v = \{P_1, P_2, \dots, P_j, \dots, P_N\}, \quad (5)$$

где каждая продукция P_j устроена как в (4).

Пусть также к этому шагу текущая цель задана M предпосылками:

$$G_v = G_1 \wedge G_2 \wedge \dots \wedge G_k \dots \wedge G_M, \quad (6)$$

причем структура G_k та же, что у конъюнктов, входящих в A_j и B_j , т. е. имеет вид (1) или (2).

Тогда цель разрешения конфликта состоит в выборе такой продукции из (5), применение которой в максимальной степени подтвердит истинность текущей цели.

Для этого в следствии каждой продукции из (5) выберем только множество конъюнктов с переменными, одинаковыми с конъюнктами текущей цели (6), вычтем содержащиеся в них множества значений переменных Φ_{ij}^+ из аналогичных множеств Φ_{ik}^+ в (6) и вычислим мощность декартова произведения полученных разностей, обозначив его $\delta\Phi_{ik}^v$. Эта мощность показывает количество сочетаний значений переменных, истинность которых останется неопределенной после применения данной продукции. Поэтому следует применить ту продукцию из CS_v , для которой число $\delta\Phi_{ik}^v$ минимально.

Выбор подмножества конъюнктов цели для обратного вывода

Здесь идея та же, что и при разрешении конфликта: выбрать из конъюнктов текущей цели (6) те, истинность которых легче всего выяснить с помощью имеющейся информации, но теперь «степень истинности» нужно сравнивать с фактами, содержащимися в БД. При этом уже нет необходимости анализировать множества значений переменных в Φ_k^+ , достаточно выбрать из (6) те конъюнкты, в которых имена переменных совпадают с именами переменных в фактах из БД.

С учетом изложенного получается следующий алгоритм.

Алгоритм управления комбинированным детерминированным выводом

Шаг 1. Проверка истинности всех конъюнктов глобальной цели G_v . Если все они истинны, то стоп: истинность глобальной цели доказана имеющимися фактами. Если есть хотя бы один ложный конъюнкт, то стоп: истинность глобальной цели противоречит имеющимся фактам. Ложность конъюнкта констатируется, когда после пересчета по формуле (3) получено $X_i^{v+1} = \emptyset$. Если есть конъюнкты, истинность которых еще не определена, из них формируется текущая цель (6).

Шаг 2. Обратный вывод: каждый из конъюнктов текущей цели (6) сравнивается с правыми частями B_j всех продукций (4), содержащихся в БЗ, и если найдена продукция, позволяющая определить истинность этого конъюнкта, вместо него в формулу цели подставляются конъюнкты, содержащиеся в левой части A_j такой продукции. Если ни одна из имеющихся в БЗ продукций не подходит, ИнС переходит в режим прямого вывода.

Шаг 3. Прямой вывод: левые части продукций (4) из БЗ сравниваются с фактами, имеющимися в БД ИнС. Если найдена продукция, левая часть которой A_j подтверждается текущим набором фактов (т. е. истинны все содержащиеся в ней конъюнкты), она добавляется в конфликтное множество CS_v . Если нет подходящих продукций, то переход к процедуре разрешения конфликта.

Шаг 4. Разрешение конфликта: вычисление $\delta\Phi_{ik}^v$ для каждой продукции из CS_v и применение той продукции, для которой число $\delta\Phi_{ik}^v$ минимально. При этом в БД добавляются факты вида (1), входящие в левую часть примененной продукции, и по формуле (3) перевычисляются домены переменных, о которых появились факты вида (2). Если (6) не имеет общих переменных со следствиями продукций из CS_v , то переход на шаг 1.

Шаг 5. Выбор подмножества конъюнктов цели для обратного вывода: выбор из (6) тех конъюнктов, в которых имена переменных совпадают с именами переменных в текущем множестве

фактов из БД. Если такие найдены, то соответствующие конъюнкты включаются в G_v и переход на шаг 1. Если совпадений нет, то стоп: вывод завершен неуспешно — истинность глобальной цели нельзя выяснить по имеющимся фактам. Ложность цели вывода констатируется, когда доказана ложность хотя бы одного конъюнкта из (6), т. е. в нем не осталось ни одного возможного значения.

Управление вероятностным выводом

При вероятностном выводе каждому значению каждой переменной x_i , т. е. каждому элементу x_{ir} множества X_i (пусть мощность этого множества равна N_i), приписывается некоторое значение текущей вероятности $p_v(x_{ir})$ того, что переменная x_i принимает значение x_{ir} . В начале вывода обычно считают, что все возможные значения переменных равновероятны, т. е.

$$p_0(x_{ir}) = 1/N_i. \quad (7)$$

Естественно, пользователь имеет возможность изменить начальное распределение вероятностей значений переменных.

Факты могут попасть в БД двумя путями: они либо вводятся пользователем по запросу ИнС, либо выводятся применением одной из продукций. В обоих случаях при вероятностном выводе каждому факту (1) или (2) приписывается некоторая вероятность, означающая, что переменная x_i принимает или не принимает значения, указанные, соответственно, во множествах Φ_i^+ или Φ_i^- . Для листовых фактов (вводимых вручную) это делает пользователь, при срабатывании какого-либо правила она вычисляется с учетом вероятностей фактов, входящих в предпосылки этого правила, и вероятности самого правила p_j , отражающей мнение экспертов о достоверности причинно-следственной связи между предпосылками и следствиями правила [6]. По понятным причинам в БЗ ИнС входят только достаточно достоверные правила, для которых p_j превышает заданное пороговое значение (например, 0,9 или выше [7]). Таким образом, для правил вероятностного вывода вместо (4) требуется формат

$$A_j \supset B_j, p_j. \quad (8)$$

Обычно [6] всем следствиям сработавшего правила приписывается вероятность, равная произведению p_j на вероятность условия этого правила, которая определяется как минимальная вероятность одного из конъюнктов, входящих в предпосылки правила, либо как произведение вероятностей таких конъюнктов.

После вычисления вероятностей следствий правила по известным формулам пересчета априорной вероятности в апостериорную [7] вычисля-

ется новое распределение вероятностей значений переменных, входящих в следствия примененного правила.

Детальное рассмотрение вопросов вычисления текущих распределений вероятностей значений переменных универсума U выходит за рамки настоящей работы, поскольку здесь нас интересуют отличия стратегии управления вероятностным выводом по сравнению с представленным выше детерминированным выводом. Они состоят в следующем.

1. «Сужение» текущих доменов переменных универсума аналогично (3) обычно производится все же для ускорения хода вывода, но из списка возможных значений некоторой переменной исключаются не все значения, указанные в фактах (2) «против», а только те, текущая апостериорная вероятность которых стала нулевой либо ниже заданной минимально допустимой величины (например, 0,1 или 0,2 [5, 6]).

2. В конфликтное множество (5) включаются только продукции, в условии которых перечисляются достаточно вероятные факты (например, минимальная вероятность конъюнктов равна 0,9 или 0,8 [6, 7]). Процедура разрешения конфликта строится аналогично детерминированному выводу, но при сопоставлении следствий правил (5) и конъюнктов цели (6) также учитываются только достаточно вероятные подмножества Φ_i^+ или Φ_i^- доменов переменных.

3. При выборе конъюнктов цели для обратного вывода анализируются не только имена входящих в них переменных, но и подмножества их значений Φ_i^+ или Φ_i^- с учетом текущих вероятностей составляющих эти подмножества значений переменных. Для доказательства ложности цели выбираются подмножества с наиболее низкими вероятностями, для доказательства истинности — с наиболее высокими.

4. Алгоритм управления комбинированным вероятностным выводом включает те же шаги, что и алгоритм управления детерминированным выводом, но единственным критерием преждевременного завершения вывода здесь является ложность одного из конъюнктов цели (6), в противном случае вывод продолжается, пока в БЗ есть применимые продукции. Когда таковых не остается, реализуется дополнительный шаг алгоритма, ответственный за интерпретацию результата вывода.

5. В недетерминированных (вероятностных, нечетких и др.) стратегиях управления выводом проблема интерпретации результата намного сложнее, чем в детерминированных стратегиях. В общем случае ее решение связано с созданием специальных процедур интерпретации, в основном эвристических и существенно зависящих от предметной области ИнС. Поэтому в настоящей работе они не рассматриваются.

Заключение

Разработанная стратегия управления выводом относится к динамическим (настраиваемым [3, 8]) стратегиям, так как параметры выбора на каждом шаге вывода зависят от результатов предыдущего шага. Такие стратегии обеспечивают большую гибкость, чем встроенные [3, 8] стратегии, в которых ход выбора предопреде-

лен априорно. Предложенный теоретико-множественный подход к представлению и анализу информации в ИнС применим как к детерминированному выводу, так и к вероятностному или нечеткому.

Работа частично поддержана грантами РФФИ (проекты № 13-07-00318-а, 14-07-00256-а, 14-07-00257-а, 14-07-00205-а, 15-07-04760-а, 15-07-02757-а).

Литература

1. Уотермен Д. Руководство по экспертным системам: пер. с англ. — М.: Мир, 1989. — 388 с.
2. Фридман А. Я., Кулик Б. А., Курбанов В. Г. Метод управления прямым логическим выводом в интеллектуальных системах с дискретными доменами переменных // Информационно-управляющие системы. 2015. № 1. С. 29–34. doi:10.15217/issn1684-8853.2015.1.29
3. Лорьер Ж.-Л. Системы искусственного интеллекта: пер. с франц. — М.: Мир, 1991. — 568 с.
4. Кулик Б. А., Зуенко А. А., Фридман А. Я. Алгебраический подход к интеллектуальной обработке данных и знаний. — СПб.: Изд-во Политехн. ун-та, 2010. — 235 с.
5. Зуенко А. А., Кулик Б. А., Фридман А. Я. Новые аспекты логического вывода: рост неопределенности и индуктивное обобщение // Шестая Всероссийская конференция по проблемам управления, 30 сентября–5 октября 2013 г.: материалы конференции. В 4 т. — Ростов-на-Дону: Изд-во Южного федерального ун-та, 2013. Т. 1. С. 34–39.
6. Марселлус Д. Программирование экспертных систем на Турбо Прологе. — М.: Финансы и статистика, 1994. — 256 с.
7. Darwiche A. Modeling and Reasoning with Bayesian Networks. — Cambridge University Press, 2009. — 526 p.
8. Рассел С., Норvig П. Искусственный интеллект: современный подход (AIMA): пер. с англ. 2-е изд. — М.: Вильямс, 2005. — 1424 с.

UDC 519.673:004.9

doi:10.15217/issn1684-8853.2015.3.35

Control of Combined Inference in Discrete Intelligence Systems

Fridman A. Ya.^a, Dr. Sc., Tech., Professor, fridman@iimm.ru

^aInstitute for Informatics and Mathematical Modelling of Technological Processes of KSC RAS, 24A, Fersman St., 184209, Apatity, Murmansk Region, Russian Federation

Purpose: Due to ample quantities of information stored in data and knowledge bases of intelligence systems, the degree of their intelligence can be measured as the difference from the complete search. In order to shorten the search, researchers develop various algorithms to control the inference chaining. In particular, the author has previously proposed a technique to control the course of forward chaining in intelligence systems with discrete domains of variables by means of heuristic indicators using certain subsets of these domains. However, it was a unidirectional chaining which made it hard to accelerate the inference by proper control of its direction and conflict resolution. In this paper, a similar set-theory approach is extended to organize combined inference and conflict resolution with the view of shortening the variant search for inference chaining. **Results:** A dynamic control strategy is developed for combined deterministic and probabilistic inference in discrete intelligence production systems, analyzing the structure of consequences from the productions included in the conflict set and comparing it with the structure of goal premises in order to find a production able to confirm as solidly as possible that the current goal is true. Dynamic strategies are known to provide more flexibility than built-in ones in which the course of choosing is predefined a priori. The proposed approach to representing and analyzing the information in intelligent systems differs from the existing ones, as it rapidly takes into account the internal structure of the system data and knowledge. This helps to increase the inference speed in a discrete intellectual system. **Practical relevance:** In flexible discrete systems, the introduced rules of control over production application can help to provide that inference goals are reached faster, whether the information in the system is determinate, stochastic or fuzzy.

Keywords — Discrete Intelligence System, Combined Inference, Control Over Production Application.

References

1. Waterman D. *A Guide to Expert Systems*. Addison-Wesley Publ., 1985. 419 p.
2. Fridman A. Ya., Kulik B. A., Kurbanov V. G. Course Indicators for Forward Logical Inference upon List Variables]. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 1, pp. 29–34 (In Russian). doi:10.15217/issn1684-8853.2015.1.29
3. Lauriere J.-L. *Intelligence Artificielle: Resolution de Probleme par l'Homme et la Machine*. Paris, Editons Eyrolles, 1987. 473 p. (In French).
4. Kulik B. A., Zuenko A. A., Fridman A. Ya. *Algebraicheskii podkhod k intellektual'noi obrabotke dannykh i znanii* [Algebraic Approach to Intelligent Processing of Data and Knowledge]. Saint-Petersburg, Politekhnikheskii universitet Publ., 2010. 235 p. (In Russian).
5. Kulik B. A., Zuenko A. A., Fridman A. Ya. New Aspects of Logical Inference: Growth of Uncertainty and Inductive Generalization. *Shestaia Vserossiiskaia mul'tikonferentsiia po problemam upravleniia*, Rostov-na-Donu, IuFU Publ., 2013, vol. 1, pp. 34–39 (In Russian).
6. Marsellus D. *Expert Systems Programming in Turbo Prolog*. Prentice Hall Publ., 1988. 266 p.
7. Darwiche A. *Modeling and Reasoning with Bayesian Networks*. Cambridge University Press, 2009. 526 p.
8. *Artificial Intelligence. A Modern Approach*. Second ed. S. Russel, P. Norvig (eds.). Prentice Hall Publ., 2003. 1081 p.