ОЦЕНКА СЛОЖНОСТИ ДАННЫХВ ПРОЦЕССЕ ЧИСЛЕННО-АНАЛИТИЧЕСКОГО РЕШЕНИЯ ЗАДАЧ

Демиденко М.И., аспирант., Полтавский национальный технический университет им Юрия Кондратюка. , г.Полтава
maxx@pntu.poltava.ua

Вступление. В настоящее время входные языки широко используемых систем компьютерной алгебры (СКА) разрабатываются на основе представлений о данных в процессе решения прикладной задачи как о совокупности отделенных выражений языка представления данных [2]. При такой реализации языка отношения между данными, отражающие суть условия задачи и более глубокие закономерности исследуемой предметной области, становятся доступными для анализа и обработки решающей системой только после выполнения всех необходимых подстановок. Такие подстановки нередко сопровождаются взрывным увеличением объема промежуточных данных, что существенно затрудняет или, даже, делает практически невозможным как интерактивный, так и автоматический анализ свойств промежуточных данных.

Все современные СКА обладают аппаратом управления преобразованиями [3]. Вместе с тем, их общим свойством является и отсутствие в этом аппарате средств для автоматической выработки критериев такого управления. Таким образом, одним из важных и актуальных аспектов проблемы интеллектуализации программного обеспечения в современной компьютерной алгебре является разработка таких средств и методик их применении.

Важным шагом на пути решения этой задачи является уточнение понятия «сложность» таким образом, чтобы это давало возможность непосредственно в ходе решения оценивать сложность данных и среди множества альтернатив выбирать такие преобразования или такое представление данных, которые приводят к наиболее простому результату.

О сложности задачи. Наиболее в настоящее время разработаны представления о «сложности» вычислений, что сводится к анализу алгоритмов: временная сложность, емкостная сложность и т.д. [4,5].

Для решения задач средствами ЧАМ такие подходы не применимы, так как в процессе числено-аналитического решения кроме числовых и символьных преобразований данных программа выполняет и интеллектуальные функции – распознавание свойств промежуточных данных и выбор необходимого преобразования. Таким образом, добавляется новый аспект сложности – сложности свойств данных.

Известны и феноменологические уточнения понятия сложности задачи [6,7,8], которые позволили выделить проблему интеллектуализации в современной компьютерной алгебры, однако не содержат необходимой конструктивной основы для вычисления сложности.

В работах [9, 10] предложена теоретико-множественная модель задачи как объекта интеллекта и введено понятие функциональных и структурных свойств данных. С точки зрения таких представлений сложность данных можно оценивать количеством состояний такого объекта.

В данной работе рассматривается сложность одного выражения.

Количество информации выражения

Процесс численно-аналитического решения можно представит в виде цепи алфавитных отображений.

$$A_0 \xrightarrow{F_1} A_1 \xrightarrow{F_2} A_2 \xrightarrow{F_3} \ldots \xrightarrow{F_{n-1}} A_{n-1} \xrightarrow{F_n} A_n$$

где An – математические выражения, описывающие данные на каждом шаге решения,

Fn - преобразования этих данных

Поставим в соответствие каждой операции, выполняемой при вычислении значения выражения A, множество их операндов, не принимая во внимание их значений, и будем рассматривать это выражение как систему мультимножеств, упорядоченных очередностью выполнения операций, выполняемых при вычислении значения выражения A. Мультимножество образуют элементы, некоторые из которых одинаковы в том смысле, что являются операндами одной и той же операции, другие считаем отличными от них. Операции рассматриваем, как алфавитные отображения, множества выражений в себя. Таким образом, унарные операции здесь не рассматриваются, т.к. они определены не на мультимножествах, а по своей сути осуществляют вычисление значения своего операнда. Для всех операций считаем, что арность $K \geq K_i \geq 2$, где i – номер операции.

1) i = 1. Пусть эта операция K_1 – арная определяется в мультимножестве

$$a_1 : \overbrace{\underbrace{o_{11} o_{11} \ldots o_{11}}_{K_1} o_{12} o_{13} \ldots o_{1K-K_1+1}}^{K},$$

где элементы одинаковы в указанном выше смысле.

Т.к. $K_1 \leq K$ возможное число состояний выражения A, соответствующих первой операции, определяется возможными значениями арности K_1 и вычисляется как мультиномиальный коэффициент

$$C_1 = \sum_{K_1=2}^{K} C_K^{K_1}.$$

В частности, если $K_1 = K$, то эта операция применяется к множеству операндов

$$a_1 : o_{11} \ldots o_{1K}$$

и выражение A описывается характеристической схемой $\otimes_1$ отношения в виде именной формы

$$A_1 : \otimes_1 \left(o_{11} \ldots o_{1K} \right).$$

Например, если $\otimes_1$- операция агрегации, то A – вектор размерности K

$$A_1 : \left(o_{11}, \ldots, o_{1K} \right),$$

а последовательность операций в A содержит I =1 членов.

2) i = 2. Пусть эта операция K_2 – арная. Она определена в мультимножестве

$$a_2 : \underbrace{o_{21}, \ldots, o_{21}}_{K_2} o_{22} o_{23} \ldots o_{2K-(K_1+K_2)+2},$$

один из этих $K - K_1 + 1$ элементов имеет вид именной формы

$$o : \otimes_1 \left(o_{11} \ldots o_{1K_1} \right),$$

описывающей характеристическою схему отношения, соответствующего второй операции.

Т.к. $K_2 \le K - K_1 + 1$, возможное число состояний выражения A, соответствующих второй операции, определяется возможными значениями арности K_2 и вычисляется как мультиномиальный коэффициент

$$C_2 = \sum_{K_2=2}^{K-K_1+1} C_{K-K_1+1}^{K_2}.$$

В частности, если $K_2 = K - K_1 + 1$, то эта операция применяется к множеству операндов

$$a_2 : o_{21} \ldots o_{2K-K_1+1}$$

и выражение A описывается характеристической схемой $\otimes_1$ отношения в виде именной формы

$$A_2 : \otimes_2 \left(o_{21} \ldots o_{2K-K_1+1} \right).$$

Методом индукции доказаны:

Утверждение 1.

Операция с номером i определена в мультимножестве

$$a_i : \underbrace{o_{i1}, \ldots, o_{i1}}_{K_i} o_{i2} o_{i3} \ldots o_{iK - \sum_{m=1}^{i-1} K_m + i - 1}, \tag{1}$$

а число возможных состояний выражения A, соответствующих этой операции, определяется возможными значениями арности этой операции и равно

$$C_i = \sum_{K_i=2}^{K - \sum_{m=1}^{i-1} K_m + i - 1} C^{K_i}_{K - \sum_{m=1}^{i-1} K_m + i - 1}. \tag{2}$$

Утверждение 2.

Множество состояний выражения A, как K-арного отношения, значение которого вычисляется последовательным выполнением $I \geq 1$ операций арности $K \geq K_i \geq 2$, равно

$$C = \prod_{i=1}^{I} C_i ,$$

(3)

где C_i – число состояний мультимножества – области определения соответствующей операции.

Все возможные состояния выражения A, как структуры равновозможны. Это дает возможность для количественной оценки сложности выражения используя известную формулу Р.Хартли

$$I = \log_2 N$$

(4)

с (3) следует, что если сложность выражения A возростает

$$K - 1 = \sum_{j=1}^{I} k_j ,$$

(5)

где $k_j = K_j - 1 \geq 1$.

соотношение (5) дает возможность найти количество последовательностей операций на множестве из K операндов, их их арности.

Результаты. На основе разработанных представлений о сложности можно ввести конструктивное определение «упрощения» или усложнения выражений:

Упрощающим преобразованием называется преобразование, которое уменьшает количество операндов выражения.

Усложняющим преобразованием называется преобразование, которое увеличивает количество операндов выражения

Преобразование, которое не изменяет количество операндов, будем называть единичным преобразованием

Примерами упрощающего преобразования является *операция приведения подобных, замена переменных*. Усложняющим преобразованием есть операция *подстановки.* Упрощающее преобразование будет уменьшать количество информации в выражении, соответственно усложняющее преобразование увеличивает количество информации. Данные примеры не являются тривиальными, поскольку предоставляется возможность оценить степень усложнения или упрощения результата.

Вывод. Таким образом, управление представляем себе так: процедура «прикидывает» количество информации при каждом из возможных преобразований из некоторого списка. Выбирает то, которое ведет к результату с наименьшей информацией.

На основании предложенного представления о сложности не только получили взрыв данных, хорошо известное явление, но и количественно описали это явление. Исходя из критериев сложности, предложенных в данной работе установлено, что в процессе преобразований выражений не избежать «взрывного» роста данных.

Литература

1. Колмогоров А.Н. Теория информации и теория алгоритмов,- М.,Наука, 1987.-304с.
2. Клименко В.П., Ляхов А.Л., Фишман Ю.С. Основные тенденции развития языков систем компьютерной алгебры // Математические машины и системы.– 2002. - № 2. – С. 29-64.
3. Ляхов А.Л., Демиденко М.И. Современные проблемы информатизации в моделировании и программировании: Сб. трудов XI международной открытой научной конференции. – Воронеж: Научная книга, 2006. – С. 293-295.
4. Гери М., Джонсон Д., Вычислительные машины и труднорешаемые задачи,- М., Мир, 1982. – 418 с
5. Ахо А., Хопкрофт Дж.,Ульман Дж., Построение и анализ вычислительных алгоритмов,-М., Мир, 1979.-536 с.
6. Ляхов О.Л. Деякі сучасні проблеми застосування чисельно-аналітичних методів// Математичні машини і системи. – 2003. - № 2. – С.54-63.
7. Глушков В.М., Брановицкий В.И., Довгялло А.М., Рабинович З.Л. Стогний А.А. Человек и вычислительная техника, - К., «Наукова думка» - 1971. -294 с.
8. Успенский В.А, Семенов А.Л.Теория алгоритмов. Основные открытия и приложения, М., Наука , 1987.- 288 с.

9. Клименко В.П., Ляхов А.Л. Прикладная математическая задача как объект компьютерной алгебры// Математические машины и системы. – 2003. – № 3-4. – С. 103-123.
10. Ляхов О.Л. Інтелектуалізація розв’язування наукових і прикладних задач на основі методів комп’ютерної алгебри// Автореферат дисертації на здобуття наукового ступеня доктора технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем, Київ-2004. – 35 с.