

А.М.Лях

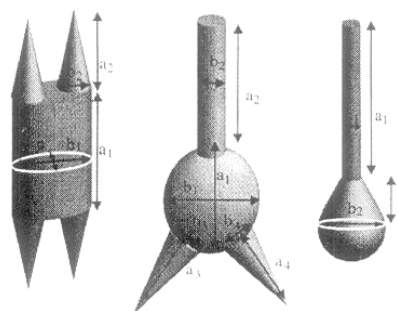
*Институт биологии южных морей НАН Украины, г.Севастополь*

### ОПИСАНИЕ ФОРМЫ ОДНОКЛЕТОЧНЫХ ВОДОРΟΣЛЕЙ И АЛГОРИТМОВ РАСЧЕТА ИХ ОБЪЕМА И ПЛОЩАДИ ПОВЕРХНОСТИ НА ЯЗЫКЕ XML

Предложена схема описания формы одноклеточных водорослей на языке XML. Описание состоит из трех блоков, хранящих информацию о простых геометрических телах и их комбинациях, правилах вывода формул расчета их объемов и площадей поверхности, и списка видов микроводорослей с указанием соответствующих им геометрических тел. По описанию можно автоматически восстановить алгоритмы расчета объема и площади поверхности микроводорослей. Использование подобных XML-описаний при обмене информацией о форме микроводорослей позволит разработать стандартные наборы геометрических фигур и стандартизировать процедуру определения биомассы геометрическим способом.

Живая масса или биомасса клеток фитопланктона – важнейшая экологическая характеристика их состояния. Массу любого организма можно вычислить по его объему и средней плотности. Обычно плотность одноклеточных водорослей принимается равной единице, поэтому, значение массы клеток будет совпадать со значением их объема. Другой, не менее важной характеристикой одноклеточных водорослей является их общая площадь поверхности, а точнее отношение площади поверхности к объему. Оно определяет интенсивность процессов обмена вещества и энергии через поверхность клетки, скорости потребления биогенов, утилизации органических веществ и т.п. Таким образом, при исследовании одноклеточных водорослей важно знать их геометрические характеристики – объем и площадь поверхности.

Традиционным и наиболее распространенным методом определения геометрических характеристик микроводорослей является метод истинных объемов – метод представления формы клеток в виде одного или нескольких геометрических тел [1]. В настоящее время существует несколько десятков наборов геометрических фигур, используемых для аппроксимации формы клеток фитопланктона (рис.1), причем разные авторы для имитации формы одного и того же вида микроводоросли используют отличающиеся наборы геометрических фигур, что приводит к различиям в оценке массы клетки [2 – 4]. Разработка стандартных аппроксимирующих фигур позволит унифицировать методы расчета объема и площади поверхности микроводорослей.



Р и с . 1 . Геометрические фигуры, аппроксимирующие форму клеток фитопланктона, состоят из простых примитивов: цилиндров, конусов, сфер (из [4]).

дорослей, что позволит получать массивы сопоставимых биологических данных. Для разработки стандартных фигур необходимо сотрудничество ученых из разных стран, которых, в первую очередь, следует обеспечить быстрым динамическим способом обмена информацией. Сейчас наиболее быстрым и безграничным средством передачи данных является сеть Интернет. Информация, передаваемая по сети, должна быть описана в едином формате, понятном всем участникам обмена. Это означает, что данные должны быть описаны на специальном языке, которым в настоящий момент является универсальный язык описания упорядоченных данных – *XML (eXtensible Markup Language)* [5].

Язык *XML* широко используется в естественных науках – биологии, океанологии, гидробиологии, примером чего служат созданные и разрабатываемые спецификации, такие как: *MarineXML* – язык описания океанологических данных [6], *Darwin Core* – язык описания биологических коллекций [7], *OBIS Shema* – язык описания биогеографической информации [8] и многие другие. На языке *MathML* реализуются математические вычисления [9]; стандарт *VRML* применяется для описания трехмерных сцен [10]. Однако к настоящему моменту не разработано спецификации *XML*, пригодной для описания простых геометрических тел (шара, цилиндра, конуса) и их комбинаций, которые используются при моделировании формы живых организмов.

Цель статьи – разработать структуру *XML*-документа для хранения информации о форме живых организмов. Задачи статьи: разработать формат описания геометрических тел и их комбинаций, разработать алгоритм построения формул расчета объемов и площадей поверхности геометрических тел по их описанию, определить основные структурные блоки *XML*-документа, представить концепцию обмена геометрической информацией через сеть Интернет.

**Методика.** Описание геометрических тел должно позволять автоматически, по информации, хранящейся в описании, восстанавливать формулы расчета их объемов и площадей поверхности. Это, во-первых, избавит ученых от необходимости рутинного вывода расчетных формул, во-вторых, позволит сразу использовать геометрическую информацию в расчетных программах.

Совокупность геометрических фигур разделим на простые или базовые геометрические тела – шар, конус, цилиндр и т.п., – формулы расчета геометрических характеристик которых непосредственно хранятся в описании, и комбинации базовых фигур или композитные фигуры, для которых расчетные формулы выводятся по определенным правилам.

**Описание базовых геометрических фигур.** Рассмотрим произвольное геометрическое тело  $F$  и будем считать, что функция  $volume(.)$  вычисляет его объем, а функция  $area(.)$  – его площадь поверхности. Пусть общую поверхность тела  $S$  можно разделить на множество мелких частей  $\{s_i\}$ , например, на грани, основания и боковые поверхности, так, что площадь поверхности каждой части можно вычислить по значениям некоторых параметров. Например, конус можно разделить на основание и боковую поверхность, цилиндр на боковую поверхность и два основания, а параллелепипед на шесть граней. Понятно, что общая площадь поверхности тела  $F$  будет равна сумме площадей каждой его части:

$$area(F) = \sum_i area(s_i). \quad (1)$$

Обозначим все плоские части поверхности символом  $f$  (от англ. *flat* – плоский), а неплоские символом  $c$  (от англ. *convex* – выпуклый), тогда

$$S = \{f_1, f_2, \dots, f_k, c_1, c_2, \dots, c_n\}. \quad (2)$$

Будем рассматривать только такие геометрические тела, форма которых полностью определяется конечным множеством параметров  $P$  или линейными размерами тела:

$$P = \{p_1, p_2, \dots, p_m\}. \quad (3)$$

По значению параметров можно вычислить объем и общую площадь поверхности тела и площадь каждой его части, а также по определенным правилам построить геометрическую фигуру:

$$P = \{p_1, p_2, \dots, p_m \mid volume(F) = volume(p_1, p_2, \dots, p_m), \\ area(F) = area(p_1, p_2, \dots, p_m)\}. \quad (4)$$

Назовем такие тела параметрическими. Пусть каждой части тела  $s_i$  соответствует некоторое подмножество параметров  $P_i$ , по которым можно вычислить ее площадь:

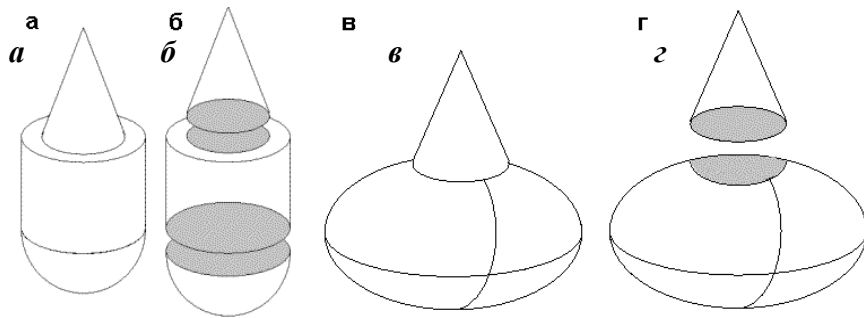
$$P_i = \{p_{i1}, p_{i2}, \dots, p_{ik} \mid area(s_i) = area(p_{i1}, p_{i2}, \dots, p_{ik})\}. \quad (5)$$

Тогда описание базовой геометрической фигуры будем состоять из:

- множества плоских и неплоских частей, на которые можно разделить данную фигуру (2);
- перечня параметров данной фигуры (3);
- функций расчета объема и поверхности всей фигуры на заданном множестве параметров (4);
- подмножества параметров, по которым можно вычислить площадь каждой части фигуры (5).

**Описание комбинаций базовых фигур.** Описание комбинаций геометрических фигур основывается на описаниях простых геометрических тел и состоит из набора правил, определяющих, как следует поступить с параметрами базовых фигур, чтобы можно было вычислить объем и площадь поверхности их комбинации.

Рассмотрим совокупность геометрических тел, образующих сложную фигуру (рис.2, *a*), и опишем этапы вычисления ее объема и площади поверхности. Очевидно, объем фигуры равен сумме объемов геометрических тел, из которых она состоит. Площадь поверхности композитной фигуры равна сумме общих площадей поверхности исходных фигур минус сумма площадей их соприкасающихся частей, отмеченных на рис.2, *б* серым цветом. В случае, когда соприкасающиеся части плоские, т.е. соприкосновение полное, общую поверхность соприкосновения легко найти как сумму площадей поверхности соответствующих оснований базовых фигур. В случае неполного соприкосновения или пересечения двух фигур, возникающего, например, при касании плоского основания конуса с выпуклой поверхностью эллипсоида (рис.2, *в*), соприкасающаяся поверхность будет совпадать с поверхностью фигуры-пересечения. Для упрощения расчетов будем



Р и с . 2 . Сложные фигуры (а, в), состоящие из нескольких простых фигур, и способ их конструирования (б, г). Серым цветом отмечены соприкасающиеся поверхности. Общая площадь поверхности сцепления равна сумме площадей всех соприкасающихся поверхностей. В случае (б) все соприкасающиеся поверхности плоские. В случае (г) соприкасающаяся поверхность эллипса выпуклая, но при расчете площади поверхности сцепления, ее нужно считать сплюснутой и равной основанию конуса.

считать, что в месте соприкосновения выпуклая фигура деформируется (сплющивается) так, что поверхность ее сплющенной части совпадает с поверхностью основания первой фигуры (рис.2, г). Тогда площадь поверхности соприкосновения будет равна удвоенной площади плоской части первой фигуры. Предположим, что сплющивание не оказывает никакого влияния на объемы тел, следовательно, объем композитной фигуры, как и прежде, будет равен сумме объемов слагающих ее геометрических тел. Запишем все вышесказанное языком математики.

Определим оператор звездочка (\*), означающий, что две части  $s_1$  и  $s_2$  двух фигур  $F_1$  и  $F_2$  соприкасаются друг с другом. Если две соприкасающиеся поверхности плоские, то будем считать, что соприкосновение полное, то есть одна из частей  $s_1$  фигуры соприкасается с другой частью  $s_2$  фигуры всей своей поверхностью. Пусть поверхность  $s_2$  больше поверхность  $s_1$  и содержит ее. Отсюда следует, что площадь  $s_1$  будет не больше площади  $s_2$ , а площадь их соприкосновения будет равна удвоенной площади меньшей поверхности:

$$s_1 \subseteq s_2, \quad \text{area}(s_1) \leq \text{area}(s_2), \quad \text{area}(s_1 * s_2) = 2 \text{area}(s_1). \quad (6)$$

Если соприкосновение частичное, например, когда одна из поверхностей плоская, а вторая нет, то сплющив вторую поверхность, получим полное соприкосновение, площадь которого вычисляется аналогично (6), как удвоенная площадь плоской части одной из фигур.

Выведем правила вычисления объема и площади фигуры, образованной объединением двух базовых параметрических тел  $F_1$  и  $F_2$ . Очевидно, что объем композитной фигуры будет равен сумме объемов каждого тела минус объем их пересечения:

$$\text{volume}(F_1 \cup F_2) = \text{volume}(F_1) + \text{volume}(F_2) - \text{volume}(F_1 \cap F_2). \quad (7)$$

Площадь поверхности композитной фигуры будет равна сумме площадей поверхности каждого тела минус площадь поверхности их пересечения или, в случае, когда фигуры только соприкасаются, минус площадь поверхности соприкасающихся частей фигуры:

$$area(F_1 \cup F_2) = \begin{cases} area(F_1) + area(F_2) - area(F_1 \cap F_2), & F_1 \cap F_2 \neq 0; \\ area(F_1) + area(F_2) - area(s_1 * s_2), & F_1 \cap F_2 = 0. \end{cases} \quad (8)$$

Чтобы не усложнять расчетные формулы, будем считать, что объем пересечения двух тел всегда равен нулю:

$$volume(F_1 \cap F_2) \equiv 0. \quad (9)$$

Следовательно, выражение (7) примет следующий вид:

$$volume(F_1 \cup F_2) = volume(F_1) + volume(F_2). \quad (10)$$

Аналогично поступим и с формулой вычисления площади поверхности композитной фигуры (8). Будем считать, что два тела не пересекаются, а только полностью соприкасаются своими плоскими частями. Тогда из (6) следует, что выражение (8) можно записать следующим образом:

$$area(F_1 \cup F_2) = area(F_1) + area(F_2) - 2 \cdot area(s_1). \quad (11)$$

Объединение двух композитных фигур также будет являться композитной фигурой, объем и площадь поверхности которой рассчитываются по формулам (10) и (11).

**Дополнительные условия.** Как уже говорилось, каждой части геометрического тела соответствует определенный набор параметров (5), используя которые можно вычислить ее площадь. При соединении двух частей предполагается, что одна из них не больше второй. Следовательно, такие же условия следует наложить на параметры. Например, рассмотрим соединенные окружности с диаметром  $d$ , и прямоугольника со сторонами  $a$  и  $b$ . Если окружность должна быть не больше прямоугольника, то ее диаметр должен быть не больше его сторон:

$$(d \leq a) \wedge (d \leq b). \quad (12)$$

Наоборот, если прямоугольник должен быть не больше окружности, то его диагональ должна быть не больше диаметра:

$$\sqrt{a^2 + b^2} \leq d. \quad (13)$$

Таким образом, при соединении двух поверхностей, каждой из которых соответствует определенный набор параметров, на параметры должны быть наложены дополнительные условия.

Описание сложной геометрической фигуры включает:

- описание каждой базовой фигуры;
- перечень соприкасающихся частей фигур;
- дополнительные условия, наложенные на параметры соприкасающихся поверхностей.

По данному описанию можно восстановить алгоритм расчета объема и площади поверхности композитной геометрической фигуры, используя формулы (10), (11) и дополнительные условия, подобные (12) и (13).

**Реализация на языке XML.** Вся информация, хранящаяся в XML-документе, при помощи элементов разметки (тегов) разделяется на структурные блоки. Способ интерпретации каждого тега и перечень его атрибутов описывается в дополнительном файле – схеме XML-документа.

*XML*-документ, хранящий описания базовых и сложных фигур и алгоритмы расчета их объемов и площадей, будет состоять из трех основных блоков:

- 1) блока с описанием базовых геометрических фигур;
- 2) блока с описанием сложных геометрических фигур;
- 3) блока с таксономическими названиями видов фитопланктона и соответствующих им сложных фигур, аппроксимирующих их форму.

Структура первого блока соответствует структуре описания базового геометрического тела. Блок содержит параметры фигуры, название частей фигуры, формулы расчета объема фигуры и площади каждой ее части. Все математические выражения реализуются на языке *MathML* [9]. Каждая базовая фигура имеет уникальное название, по которому осуществляется ссылка на нее. Т.к. базовые фигуры изменяются редко, блок описания геометрических фигур помещается в отдельный файл и однократно передается участникам проекта.

Второй блок *XML*-документа хранит описание сложных фигур. В начале блока хранятся идентификаторы (названия или числовые коды) всех базовых фигур, из которых состоит сложная фигура. Далее указываются пары соприкасающихся фигур, т.е. перечисляются пары соприкасающиеся частей этих фигур. Здесь же определяются дополнительные условия, накладываемые на параметры частей фигур.

Третий блок *XML*-документа содержит список видов одноклеточных водорослей и соответствующих им геометрических фигур, аппроксимирующих их форму. Т.к. форма одноклеточных водорослей очень изменчива, то клетки одного вида могут описываться разными геометрическими телами. Поэтому одному виду могут соответствовать несколько геометрических тел. В данном блоке возможно использование спецификации *Darvin Core* [7] применяемой при описании таксономической информации.

Разработка стандартного набора сложных геометрических фигур будет состоять из нескольких этапов. На первом шаге выбираются базовые геометрические примитивы, описание которых отсылается всем участникам проекта. Далее создается *XML*-документ, содержащий описание сложных геометрических фигур и списки видов фитопланктона, который считается текущей, но не окончательной реализацией стандарта. Копия текущей реализации также отсылается каждому участнику проекта. Участник имеет право внести в текущую реализацию *XML*-документа определенные изменения, создавая локальную реализацию стандарта. Все локальные реализации отсылаются рабочей комиссии, которая на их основе разрабатывает следующую версию стандарта, которая снова отсылается всем участникам проекта, и т.д. Процедура обмена продолжается до тех пор, пока все или большинство участников проекта не будут согласны с внесенными изменениями и не будут предлагать новых изменений. В этом случае текущая реализация будет считаться окончательной реализацией стандартных наборов геометрических фигур, служащих для аппроксимации формы микроводорослей и расчета их объемов и площадей поверхности.

**Выводы.** 1. Используя предложенную схему описания геометрических тел, исследователь может с легкостью переводить естественное описание композитных фигур на язык, понятный компьютеру. Например, естественное описание «основание конуса соединено с верхним основанием цилиндра»

ра» может быть преобразовано в XML-описание соприкосновения двух фигур их плоскими частями.

2. Используя предложенную схему, ученый получает возможность конструировать произвольные геометрические тела из набора базовых фигур – своеобразных кубиков конструктора, не заботясь о выводе формул расчета их объемов и площадей поверхности.

3. По XML-схеме описания формы одноклеточных водорослей машина может сопоставить определенным видам одноклеточных водорослей геометрические фигуры, аппроксимирующие их форму, и восстановить алгоритмы расчета объемов и площадей поверхности клеток.

4. Обмен XML-документами, описывающих форму микроводорослей, не только значительно ускорит процесс создания стандартной методики расчета биомассы и площади поверхности одноклеточных водорослей, но и позволит апробировать полученные результаты на редких видах микроводорослей или видах нетипичной формы, ограниченно распространенных в Мировом океане.

Представленная структура XML-документа содержит описание геометрических фигур и алгоритмы расчета их объемов и площадей поверхности. В то же время пользователю важно видеть, как выглядит сложная фигура. Поэтому дальнейшая работа будет заключаться в разработке структуры XML-файла, позволяющей восстанавливать форму геометрического тела с указанием необходимых промеров. Помимо этого, автор разрабатывает технологию аппроксимации формы микроводорослей трехмерными моделями [11, 12], которые позволят наиболее точно рассчитать объем и площадь поверхности микроводорослей.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Киселев И.А.* Методы исследования планктона // Жизнь пресных вод СССР.– М.-Л.: АН СССР, 1950.– 280 с.
2. *Брянцева Ю.В., Лях А.М., Сергеева А.В.* Расчет объемов и площадей поверхности одноклеточных водорослей Черного моря. Препринт.– Севастополь: ИнБИОМ, 2005.– 25 с.
3. *Hillebrand H., Durselen C.-D., Kirschlel D., Pollingher U., Zohary T.* Biovolume calculation for pelagic and benthic microalgae // J. of Phycology.– 1999.– 35.– P.403-424.
4. *Sun J., Liu D.* Geometric models for calculating cell biovolume and surface area for phytoplankton // J. of Plankton Research.– 2003.– 25, №11.– P.1331-1346.
5. *Extensible Markup Language (XML).*– <http://www.w3c.org/XML/>
6. *MarineXML Portal.*– <http://www.marinexml.net/>
7. *Darwin Core.*– <http://darwincore.calacademy.org/>
8. *OBIS: Ocean Biogeographic Information System.*– <http://www.iobis.org/>
9. *W3C Math Home.*– <http://www.w3c.org/Math/>
10. *Virtual Reality Modeling Language.*– <http://tecfa.unige.ch/guides/vrml/vrml97/>
11. *Lyakh A.M.* The free form deformation of phytoplankton models // Lecture Notes in Computer Science.– 2002.– 2331.– P.194-201.
12. *Лях А.М.* Деформации пространства и их применение в моделировании формы клеток фитопланктона // Уч. Зап. Таврич. Нац. Ун-та.– 2005. (в печати).

Материал поступил в редакцию 27.07.2004 г.  
После доработки 17.03.2005 г.