

С. В. Булгаков

Агрегирование информационных моделей

Статья анализирует построение информационных моделей на основе агрегирования. Агрегирование рассматривается как информационная технология. Показана иерархичность агрегированных моделей. Описаны уровни агрегированных моделей. Описано понятие конфигуратора и оператора агрегатора

Ключевые слова: модели, информационные модели, информационные технологии, моделирование, агрегирование, структуры данных

S. V. Bulgakov

Aggregation information models

Article analyzes the construction of information models based on aggregation. Aggregation is considered as information technology. Shows the hierarchy of aggregated models. Describes the levels of aggregated models. Described the concept of the configurator and operator aggregator

Keywords: model, information model and information technology, simulation. aggregating data structure

Одним из основных способов построения моделей данных [1] является абстракция. Абстракция данных (data abstraction) — процедура организации данных без учета их внутреннего представления. Различают два вида абстракции: *обобщение и агрегация*. Они применяются для построения сложных моделей объектов и их анализа. Различают прямые и обратные процедуры [2, 3]. К прямым процедурам относят собственно обобщение, классификацию, агрегацию. К обратным относят — специализацию, экземпляцию, пошаговую детализацию.

Агрегация (агрегирование, называемое также композицией или включением) в области информационных технологий обозначает процессы создания составных объектов нового класса из уже существующих объектов подклассов путем их объединения. Агрегированные модели связаны с составляющими их частями «отношениями принадлежности» [4]. Составные объекты нового класса обычно включают свойства объектов их составляющих.

Переходя от агрегирования абстрактных объектов к информационным моделям можно выяснить ряд интересных свойств. Исходными информационными моделями могут быть простые информационные модели или информационные единицы (рис.1). Агрегирование информационных моделей может рассматриваться как информационная технология построения сложной информационной модели с заданными свойствами, которые задаются набором свойств исходных информационных моделей.

Таким образом, при построении информационной модели с необходимыми свойствами подбираются простые модели или информационные единицы с такими свойствами, а затем на основе агрегации строится требуемая информационная модель. Агрегативная информационная модель является многоуровневой. Поскольку агрегирование осуществляется снизу вверх, проведем анализ с нижних уровней.

Нижний уровень образуют информационные единицы. Следующий уровень — уровень простых информационных моделей. Следующий более высокий уровень — уровень составных моделей. Высший уровень — агрегативная информационная модель.

Следует остановиться на информационных единицах, которые задают систему описания. Это могут быть структурные информационные единицы [5], семантические информационные единицы [6], визуальные информационные единицы [7], телекоммуникационные информационные единицы [8], образовательные информационные единицы [9] и так далее. Сам по себе процесс агрегирования может быть рассмотрен как элемент стандартизации [10] построения информационных моделей.

Агрегативная информационная модель создается на основе схемы, элементами которой являются исходные модели. Такой подход позволяет изменять агрегативную информационную модель. Замена вложенных моделей позволяет динамически изменять поведение и свойства изменять агрегативной информационной модели. Таким образом, агрегирование можно рассма-

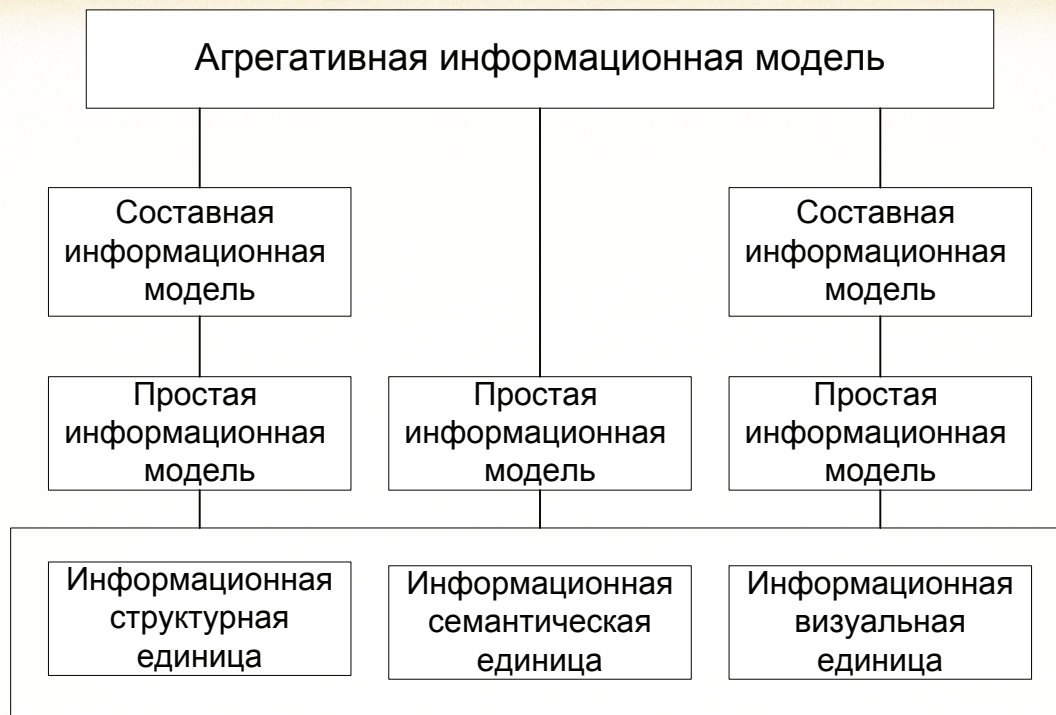


Рис.1. Структура агрегированной модели

тривать как информационный процесс идущий «снизу вверх» направленный на создание сложной модели из простых моделей [11].

Существует противоположный процесс идущий сверху — вниз [12]. Этот процесс заключается в создании простых специализированных моделей из сложных универсальных моделей. Этот процесс можно назвать дезагрегация или декомпозиция. Если дезагрегация осуществляется с сохранением связей между дезагрегированными моделями, то в итоге получается распределенная модель. Если на базе дезагрегирования реализуется методика делегирования функций в специализированные модели, то возникает модель сетцентрического управления [13]. При таком управлении поставленная перед сложным объектом задача перепоручается специализированному объекту более низкого уровня.

Необходимость агрегирования может вызываться различными целями и сопровождаться разными обстоятельствами, что приводит к различным способам агрегирования. Процесс агрегации может быть использован не только для построения информационной модели объекта, но и для построения информационной модели ситуации [14], в которой находится данный объект

Техника агрегирования основана на использовании определенных принципов и моделей исследуемого объекта. В зависимости от того что является доминантой агрегирования оно подразделяется на агрегирование по частям, агрегирование по свойствам, агрегирование по процессам, агрегирование по связям.

Выбранные модели определяют, какие части должны войти в состав системы (модель состава,

ва) и как они должны быть связаны между собой (модель структуры). Разные условия и цели агрегирования приводят к необходимости использовать разные модели, что в свою очередь определяет тип окончательного агрегата и технику его построения.

Результат агрегирования можно определить как установление отношений на заданном множестве элементов (процессов, свойств, структур). Благодаря значительной свободе выбора в том, что именно рассматривается в качестве элемента, как образуется множество элементов и какие отношения устанавливаются на этом множестве, получается разное количество и качественно разнообразное множество результатов агрегирования. Основные агрегаты, типичные для системного анализа: конфигуратор, агрегаты-операторы и агрегаты-структуры.

Агрегирование применяется как средство решения проблемы или средство построения необходимой информационной модели. В реальной жизни не бывает чистых проблем. Описание проблемы в терминах обозначает не саму проблему, а выбранную точку зрения на нее. Только совместное (агрегированное) описание в терминах нескольких качественно различающихся языков позволяет охарактеризовать явление с достаточной полнотой.

При агрегировании возникает проблема о допустимой минимизации описания явления. Риск неполноты описания недопустим, риск переопределения связан с излишними затратами на обработку и анализ.

Приведенные соображения приводят к понятию конфигуратора как средства, включающего качественно различные языки или средства опи-

сания модели и обладающего тем свойством, что число этих средств описания минимально необходимо для заданной цели. Продемонстрируем смысл этого понятия на примерах.

Конфигуратором для задания любой точки n -мерного пространства является совокупность ее координат. Обратим внимание на эквивалентность разных систем координат (разных конфигураторов) и на предпочтительность ортогональных систем, дающих независимое описание на каждом «языке» конфигулятора.

Конфигуратором для описания поверхности любого трехмерного тела на «плоскостных» языках является совокупность трех ортогональных проекций, принятая в техническом черчении и начертательной геометрии. Обратим внимание на невозможность уменьшения числа проекций и на избыточность большего числа «точек зрения».

Опыт проектирования сложных организационно-технических систем [15, 16] показывает, что для синтеза организационной системы конфигуратор состоит из описания распределения власти (структуры подчиненности), распределения ответственности (структура функционирования) и распределения информации (организация связи и памяти систем, накопления опыта, обучения, истории). Все три структуры не обязаны совпадать топологически, хотя связывают одни и те же части системы.

Необходимо подчеркнуть, что конфигуратор является содержательной моделью высшего концептуального уровня. Определив средства описания пользователь тем самым определяет тип модели. Как всякая модель конфигуратор имеет целевое назначение, поэтому при смене цели может утратить свойства конфигулятора. Как каждая модель, конфигуратор в простых случаях очевидно адекватен, в то время как адекватность в других случаях подтверждает практика.

Подобно тому, как можно говорить о различных уровнях декомпозиции, т.е. степени детализации модели рассматриваемого объекта, явления, проблемы, можно говорить и о различных уровнях агрегированности (рис.1) модели как о балансе между конкретностью и абстрактностью информационной модели и информационного процесса [17].

При агрегировании модели учитывается тот факт, что каждый структурный уровень описывается специфическими для него системообразующими законами [15]. Естественно, что высший уровень агрегированности соответствует исходному уровню декомпозиции — модели-основанию. По мере повышения уровня агрегированности системы происходит следующее:

- изменяются законы, определяющие поведение системы, что влечет за собой изменение структуры и, возможно, типа модели;

- упрощается информационное обеспечение модели за счет уменьшения степени детализации (положительный аспект);
- из-за уменьшения степени детализации модели вне рассмотрения могут оказаться некоторые важные эффекты, представляющие интерес для пользователя;
- состав потенциальных пользователей модели изменяется;
- проще решается проблема целостности описания объекта;
- закономерности, описывающие поведение системы на высоких уровнях агрегирования все более удаляются от фундаментальных законов природы.

Последняя особенность может приводить к отрыву модели от реального объекта. Кроме того, переменные, используемые при модельном описании объекта, становятся более абстрактными, что вновь может увеличить трудности информационного обеспечения модели.

При построении моделей сложных систем рекомендуется на каждом этапе получать ряд завершенных моделей различного уровня агрегированности, представляющих в каждом случае целостный объект. При этом детальность проработки модели каждого уровня может быть различной и зависит от конкретной постановки практической задачи, которую предполагается решить с помощью модели. Модели высоких уровней агрегированности могут играть роль логических схем, помогать обосновывать и документировать модели нижележащих уровней, проверять непротиворечивость общей логики моделирования.

Одна из наиболее частых ситуаций, требующих агрегирования, состоит в том, что совокупность данных, с которыми приходится иметь дело, плохо обозрима [18], с этими данными трудно «работать». Необходимость упрощения работы с сложной совокупностью данных приводят к необходимости агрегирования. В данном случае на первый план выступает такая особенность агрегирования, как уменьшение размерности: агрегат объединяет части в нечто единое с исключением несущественного.

Простейший способ агрегирования состоит в установлении отношения эквивалентности между агрегируемыми элементами, т.е. образования классов [19]. Классификация является важным и многофункциональным явлением в человеческой практике вообще и в моделировании в частности. С практической точки зрения одной из важнейших является проблема определения, к какому классу относится данный конкретный элемент. Все свидетельствует о том, что агрегирование в классы является эффективной, но далеко не тривиальной процедурой. Если представить класс как результат действия агрегата-оператора, то такой оператор имеет вид «ЕСЛИ <условия на агрегируемые признаки>, ТО <имя класса>».

Другой тип агрегата-оператора возникает, если агрегируемые признаки фиксируются в числовых шкалах. Тогда появляется возможность задать отношение на множестве признаков в виде числовой функции многих переменных, которая и является агрегатом. Свобода выбора в задании функции, агрегирующей переменные, является кажущейся, если этой функции придается какой-то реальный смысл. Редкий пример однозначности агрегата-функции дает стоимостной анализ экономических систем. Если все участвующие факторы удастся выразить в терминах денежных расходов и доходов, то агрегат оказывается их алгебраической суммой. Вопрос состоит лишь в том, в каких случаях можно использовать этот агрегат, не обращаясь к другим системам ценностей, а когда следует вернуться к конфигуратору, включающему не только финансовые критерии, но и политические, моральные, экологические и т.д.

Важный пример агрегирования данных дает статистический анализ. Среди различных агрегатов, называемых в этом случае статистиками, т.е. функциями выборочных значений, особое место занимают достаточные статистики. Достаточные статистики — это такие агрегаты, которые извлекают всю полезную информацию об интересующем нас параметре из совокупности наблюдений. Однако при агрегировании обычно потери информации неизбежны, и достаточные статистики являются в этом отношении исключением. В таких случаях становятся важными оптимальные статистики, т.е. позволяющие свести неизбежные в этих условиях потери к минимуму в некотором заданном смысле.

Примером статистического агрегирования является факторный анализ, в котором несколько переменных сводятся в один фактор. Именно потому, что при рассмотрении реальных данных самым важным является построение модели агрегата при отсутствии информации, необходимой для теоретического анализа статистики, некто Тьюки предложил назвать эту область «анализом данных», оставляя за математической

статистикой задачи алгоритмического синтеза и теоретического анализа статистик.

Важной формой агрегирования является образование структур. К тому, что о моделях структур уже было сказано, необходимо добавить следующее. Как и любой вид агрегата, структура является моделью системы и, следовательно, определяется тройственной совокупностью: объекта, цели и средств моделирования. Это и объясняет многообразие типов структур (сети, матрицы, деревья и т.п.) возникающих при выявлении, описании структур (познавательные модели).

При синтезе мы создаем, определяем, называем структуру будущей, проектируемой системы. Если это не абстрактная, а реальная система [19], то в ней возникнут и начнут «работать» не только те связи, которые спроектированы, но и множество других, не предусмотренных схемой моделирования. Поэтому при проектировании агрегативной информационной модели важно задать ее структуру во всех существенных отношениях, так как в остальных отношениях структуры сложатся сами, стихийным образом. Совокупность всех существенных отношений определяется конфигуратором, и отсюда вытекает, что «проект любой модели должен содержать разработку стольких структур, сколько языков включено в ее конфигуратор». Например, проект организационной системы должен содержать структуры распределения власти, распределения ответственности и распределения информации. Подчеркнем, что хотя эти структуры могут сильно отличаться топологически, они лишь с разных сторон описывают одну и ту же систему и, следовательно, не могут быть не связаны между собой.

Выводы. Агрегирование является видом информационной технологии, позволяющей получать новые информационные модели. Агрегирование является инструментом построения интегрированных систем, процессов, технологий, моделей и данных.

ЛИТЕРАТУРА

1. Цветков В.Я. Информатизация: Создание современных информационных технологий. Часть 1. Структуры данных и технические средства. М., ГКНТ, ВНТЦентр, 1990. 118 с.
2. Иванников А.Д., Кулагин В.П., Тихонов А.Н. . Цветков В.Я. Прикладная геоинформатика . М.: МаксПресс 2005. 360 с.
3. Цветков В.Я. Модели и моделирование. М.: Госинформобр, 2006. 94 с.
4. Цветков В.Я. Логика в науке и методы доказательств. LAP LAMBERT Academic Publishing GmbH & Co. KG, Saarbrücken, Germany 2012. 84 с.
5. Tsvetkov V.Ya. Information objects and information Units // European Journal of Natural History, 2009. №2. p.99
6. V. Ya. Tsvetkov. Semantic Information Units as L. Florodi's Ideas Development // European Researcher, 2012. Vol.(25). №7. pp.1036-1041.
7. Лютий А.А. Язык карты: сущность, система, функции. М.: ГЕОС, 2002, Изд. 2-е. 2002. 327 с.
8. Цветков В.Я. Информационные единицы сообщений // Фундаментальные исследования, 2007. №12. С.123-124
9. Кудж С.А., Цветков В.Я. Информационные образовательные единицы // Дистанционное и виртуальное обучение, 2014. №1. С.24-31.
10. Цветков В.Я. Особенности развития информационных стандартов в области новых информационных технологий // Информационные технологии, 1998. №8. С.2-7.
11. Поляков А.А., Цветков В.Я. Прикладная информатика М.: Янус-К, 2002. 392 с.
12. Цветков В.Я. Семиотический подход к организации моделей данных // Известия высших учебных заведений. Геодезия и аэрофотосъемка, 2000. №5. С.142-146.

13. Тихонов А.Н., Иванников А.Д., Соловьёв И.В., Цветков В.Я., Кудж С.А. Концепция сетецентрического управления сложной организационно-технической системой. М.: МаксПресс, 2010. 136 с.
14. V. Ya. Tsvetkov. Information Situation and Information Position as a Management Tool // *European Researcher*, 2012. Vol.(36). № 12-1, pp.2166-2170.
15. Соловьёв И.В. Сложная организационно-техническая система как инструмент исследования искусственных антропогенных систем // *Дистанционное и виртуальное обучение*, 2014. №1. С.5-23.
16. Кудж С.А., Соловьёв И.В., Цветков В.Я. Когнитивные модели и методы. Краткий словарь-справочник. М.: МГТУ МИРЭА. 95 с.
17. Цветков В.Я. Информационное управление. LAP LAMBERT Academic Publishing GmbH & Co. KG, Saarbrücken, Germany 2012. 201 с.
18. Цветков В.Я. Когнитивные аспекты построения виртуальных образовательных моделей // *Перспективы науки и образования*, 2013. №3. С.38-46.
19. Цветков В.Я. Формальная и содержательная классификация // *Современные наукоёмкие технологии*, 2008. №6. С. 85-86.

REFERENCES

1. Tsvetkov V.Ia. *Informatizatsiia: Sozdanie sovremennykh informatsionnykh tekhnologii. Chast' 1. Struktury dannykh i tekhnicheskie sredstva* [Informatization: Creation of modern information technologies. Part 1. Data structures and technical facilities]. Moscow, GKNT, VNTTsentr, 1990. 118 p.
2. Ivannikov A.D., Kulagin V.P., Tikhonov A.N. . Tsvetkov V.Ia. *Prikladnaia geoinformatika* [Applied Geoinformatics]. Moscow, MaksPress, 2005. 360 p.
3. Tsvetkov V.Ia. *Modeli i modelirovanie* [Models and modeling]. Moscow, Gosinformobr, 2006. 94 p.
4. Tsvetkov V.Ia. *Logika v nauke i metody dokazatel'stv* [Logic in science and techniques of evidence]. LAP LAMBERT Academic Publishing GmbH & Co. KG, Saarbrücken, Germany, 2012. 84 p.
5. Tsvetkov V.Ya. Information objects and information Units. *European Journal of Natural History*, 2009, no.2, p.99
6. V. Ya. Tsvetkov. Semantic Information Units as L. Florodi's Ideas Development. *European Researcher*, 2012. Vol.(25), no.7, pp.1036-1041.
7. Liutyi A.A. *Iazyk karty: sushchnost', sistema, funktsii* [Language of the map: the nature, the system functions]. Moscow, GEOS, 2002. 327 p.
8. Tsvetkov V. Ia. Information unit messages. *Fundamental'nye issledovaniia - Basic research*, 2007, no.12. pp.123-124 (in Russian)
9. Kudzh S.A., Tsvetkov V.Ia. Information educational units. *Distsionnoe i virtual'noe obuchenie - Distance and virtual training*, 2014, no.1, pp.24-31 (in Russian).
10. Tsvetkov V.Ia. Features of development of information standards in the field of new information technologies. *Informatsionnye tekhnologii - Information technologies*, 1998, no.8, pp.2-7 (in Russian).
11. Poliakov A.A., Tsvetkov V.Ia. *Prikladnaia informatika* [Applied Informatics]. Moscow, Ianus-K, 2002. 392 p.
12. Tsvetkov V.Ia. Semiotic approach to the organization of data models. *Izvestiia vysshikh uchebnykh zavedenii. Geodeziia i aerofoto"emka - Proceedings of higher educational institutions. Geodesy and air photography*, 2000, no.5, pp.142-146 (in Russian).
13. Tikhonov A.N., Ivannikov A.D., Solov'ev I.V., Tsvetkov V.Ia., Kudzh S.A. *Kontseptsiiia setetsentricheskogo upravleniia slozhnoi organizatsionno-tekhnicheskoi sistemoi* [Concept of network-centric management of complex technical-organizational system]. Moscow, MaksPress, 2010. 136 p.
14. V. Ya. Tsvetkov. Information Situation and Information Position as a Management Tool. *European Researcher*, 2012. Vol.(36), no.12-1, pp.2166-2170.
15. Solov'ev I.V. A complex of organizational-technical system as a research tool artificial anthropogenic systems. *Distsionnoe i virtual'noe obuchenie - Distance and virtual training*, 2014, no.1, pp.5-23 (in Russian).
16. Kudzh S.A., Solov'ev I.V., Tsvetkov V.Ia. *Kognitivnye modeli i metody. Kratkii slovar'-spravochnik* [Cognitive models and methods. Quick reference dictionary]. Moscow, MGTU MIREA. 95 p.
17. Tsvetkov V.Ia. *Informatsionnoe upravlenie* [Information management]. LAP LAMBERT Academic Publishing GmbH & Co. KG, Saarbrücken, Germany, 2012. 201 p.
18. Tsvetkov V.Ia. Cognitive aspects of building a virtual educational models. *Perspektivy nauki i obrazovaniia - Perspectives of science and education*, 2013, no.3, pp.38-46 (in Russian).
19. Tsvetkov V.Ia. Formal and substantive classification. *Sovremennye naukoemkie tekhnologii - Modern high technologies*, 2008, no.6, pp. 85-86 (in Russian).

Информация об авторе Булгаков Сергей Владимирович (Россия, Москва)

Кандидат технических наук
Заместитель начальника НИЧ
Московский государственный институт
радиотехники, электроники и автоматики
E-mail: cvj2@mail.ru

Information about the author Bulgakov Sergei Vladimirovich (Russia, Moscow)

PhD in Technical Sciences
Deputy Head of Research Department
Moscow State Institute of Radio Engineering, Electronics
and Automation
E-mail: cvj2@mail.ru