1562350

Московский государственный технический университет
имени Н.Э.Баумана.

На правах рукописи.

Пардаев Акбар Анварович

МЕТОДИКА ПРЕОБРАЗОВАНИЯ ОПИСАНИЯ ОПЕРАТИВНОЙ СИТУАЦИОННОЙ ОБСТАНОВКИ В ГРАФИЧЕСКИЕ ОБРАЗЫ ССОИ.

05.13.06 - Автоматизированные системы обработки информации и управления

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук

Москва 1998

Работа выполнена в Московском государственном техническом университете имени Н.Э.Баумана.

Научный руководитель: кандидат технических наук, доцент
Соломонов Л.А.

Официальные оппоненты:

1. Шемакин Юрий Иванович, д.т.н., профессор

2. Кинкулькин Аркадий Миронович, к.т.н., начальник отделения

Ведущее предприятие: НПО им. С.А.Лавочкина

Защита диссертации состоится "__ "_____199__ года на
[illegible]та Д 053.15.03 "Вычислительная и
[illegible]о государственного технического
[illegible] адресу: 107005, Москва, 2-я

[illegible] в библиотеке МГТУ имени

[illegible]_1998 г.

[illegible]га _____ Иванов С.Р.

1562350

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Концепция персональных информационных систем при своем воплощении требует решения двух проблем - формализации знаний и организации взаимодействия человека с техническими средствами АСУ.

В современных АСУ изменяется сам характер информации, с которой она имеет дело. В дополнение к информации фактографического и технико-экономического характера, которую уже умеют обрабатывать современные системы, должна все шире использоваться графическая, понятийная информация, организуемая в виде специализированных баз знаний пользователя о его предметной области - управляемой им системе, сфере его деятельности; информация, фиксирующая личный опыт принятия решений руководителем.

Актуальность проблемы. Прогресс в области построения сложных человеко-машинных систем привели к появлению в последние десятилетия особого класса человеко-машинных систем, называемых ситуационными системами отображения информации (ССОИ). В частности подобные системы применяются при решении задач:

- наблюдения за надводной, подводной, наземной, воздушной, космической обстановкой на достаточно большом ареале земной поверхности;
- управления (навигации) динамическими объектами в указанных средах;
- экологического мониторинга;
- управления сложными транспортными узлами и т.д.

В работе "Отображение информации" из серии "Организация взаимодействия человека с техническими средствами АСУ" показано, что взаимодействие человека-оператора (Ч-О) с реальным миром (РМ) осуществляется через информационные модели (ИМ), которые будем называть информационными моделями реального мира (ИМРМ). Схема данного взаимодействия изображена на рисунке 1.

1562350
НТБ МГТУ им. Н.Э. Баумана
1562350
Парфаев А.А. Методика прес
МГТУ им. Н.Э. Баумана
ОТДЕЛ 1

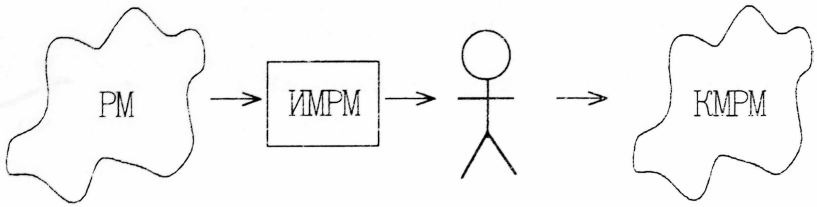


Рисунок 1. Взаимодействие Ч-О с реальным миром

ИМ, которые являются для Ч-О источником информации, на основе которой он формирует образ реального мира, как правило, подразделяются на:

- текстовые ИМ;
- графические ИМ.

Но наличие у Ч-О звукового канала восприятия говорит о том, что существует и звуковая информационная модель. Кроме этого применение в некоторых областях (архитектура, автомобилестроение и т.п.) предварительного макетирования позволяет сделать предположение и о пространственной ИМ.

Таким образом реальный мир воспринимается Ч-О через следующие типы ИМ:

- текстовый (ЕЯ);
- графический (ГО);
- звуковой (ЗР);
- пространственный (ПР).

Так как различные типы ИМ характеризуют один и тот же реальный мир, то можно поставить задачу преобразования одного типа ИМ в другой (данное утверждение справедливо лишь в фиксированной предметной области). Наибольший интерес на современном этапе представляет переход от текстового представления ИМ в графическое.

Цель работы. Разработка методики преобразования естественно-языкового (ЕЯ) описания оперативной ситуационной обстановки в графические образы (ГО) ССОИ.

Общая методика исследований. В работе использованы методы теории математической лингвистики, теории алгоритмов, теории ситуационного управления, теории исчисления предикатов первого порядка, теории нечетких множеств. Для исследования состоятельности и эффективности предлагаемых методов и средств выполняется практическая реализация макетной версии системы интерактивного преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ.

Научная новизна. В диссертации разработана методика преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ.

Разработаны и выносятся на защиту:

- методика преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ;
- функциональная грамматика;
- методика формирования лексики заданной ПО;
- методика структуризации характеристик объектов заданной ПО;
- метод переходных коэффициентов;
- алгоритмы синтаксического и семантического анализов, размещения объектов в кадре (планирование кадра) и генерации сценария на ЯФОС.

Практическая ценность. Разработанные модели, алгоритмы, методики и программные средства использовались при создании комплекса программного и математического обеспечения системы отображения коллективного пользования объекта 14П04 - КПМО СОКП в РНИИКП.

Апробация работы. Содержание отдельных частей и диссертации в целом было доложено:

- на заседаниях аттестационной комиссии при ежегодной аттестации аспирантов кафедры "Автоматизированные системы отображения информации и управления" МГТУ имени Н.Э. Баумана;

– на заседании кафедры "Автоматизированные системы системы отображения информации и управления" МГТУ имени Н.Э. Баумана;

– на комиссиях по приемке этапов работ в РНИИКП.

Программное обеспечение, созданное в рамках настоящей диссертационной работы демонстрировалось:

– на международной выставке Softool 91 на ВВЦ;

– на выставке с международным участием ImproGraph 91 на ВВЦ;

– на выставке с международным участием Softool 92 на ВВЦ.

Публикации. Основное содержание диссертации опубликовано в 3 печатных работах. Результаты разработок и исследований, проведенных по теме диссертационной работы включены в 3 отчета по НИР.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения и приложений, изложенных на 201 страницах машинописного текста, содержит список литературы, включающий 53 наименования.

СОДЕРЖАНИЕ РАБОТЫ

В первой главе диссертации проведен обзор методов формализации ЕЯ, дан анализ видов компьютерной графики, показаны системы нового направления - когнитивной графики. Ставится формальная постановка задачи преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ.

К свойствам ЕЯ, которые затрудняют моделирование относятся:

1. Сложность ЕЯ.
2. Недетерминированность соответствий и переходов.
3. Нечеткость многих множеств и отношений.
4. Неравноправие различных феноменов и конструкций.
5. Изменчивость, которой подвержен и запас слов, и смыслы выражений, и допустимость тех или иных сочетаний.

6. Невозможность использования в современных языковых процессорах столь широкого круга сведений, каким пользуется при анализе и синтезе текстов человек.

При построении модели некоторого выбранного ЕЯ, то, как всегда при моделировании, с целью уменьшения трудностей приходится ограничивать круг учитываемых явлений.

Сокращение сложности и устранение изменчивости обычно делаются за счет ограничения на множество допустимых на вход языкового процессора текстов по двум показателям: по тематике и по строению.

Моделирование процессов порождения графического образа ситуаций, описываемых на естественном языке, требует понимания когнитивных процессов, связанных как с вербальной (языковой), так и невербальной информацией. В процессе понимания некоторой пространственной сцены используется не только языковое описание сцены, но и ее образное представление.

В настоящее время существует большое число прикладных программ, которые работают с текстово-графической информацией. Все они входят в широкий класс систем машинной графики, что обусловило появление графических систем, которые можно разделить на несколько классов.

1. Деловая графика.
2. Научная графика.
3. Автоматизация чертежных и конструкторских работ.
4. Машинная живопись.
5. Компьютерная анимация.
6. Когнитивная графика.

На рисунке 2 показана схема взаимодействия человека-оператора (Ч-О) в ситуационной системе отображения информации (ССОИ) с реальным миром, которая описывается моделью опосредствованного управления предметами реального мира. На рисунке введены следующие сокращения:

ДМР - динамическая модель результата;

КМПРМ - концептуальную модель предметов реального мира;

Sp - решающая система;

СМ - сенсорно-моторная система;

ПРМ - предметы реального мира;

ИМПРМ - Информационная модель предметов реального мира;

М - Манипуляторная система.

Отличительной чертой этой модели является опосредствованное восприятие человеком ситуации реального мира через ИМПРМ и опосредствованное же воздействие на эту ситуацию через манипуляторную систему. Среда Ч-О и среда реального мира не соприкасаются. Этим определяется исключительная роль ИМПРМ, ведь всю информацию, которую получает человек-оператор о развитии ситуации в реальном мире, он получает только из ИМПРМ (непосредственно предметы реального мира «своими глазами» он не видит).

Информационная модель реального мира хранит информацию, необходимую для решения всех штатных задач Z ССОИ для всех ситуаций S, возникающих в проблемной среде, которая называется информационной моделью максимальной насыщенности I^{max}.

Восприятие человеком-оператором реального мира производится через ИМПРМ, что позволяет Ч-О сформировать свою концептуальную модель. На основании этого можно сделать вывод, что концептуальная модель (К) у Ч-О формируется при восприятии информационной модели (ИМ) максимальной насыщенности. В свою очередь концептуальная модель состоит из концептуальной модели о ситуации и объектах реального мира, что можно записать в виде:

$$K = [Ki^{Obj}] \bigcup [Ki^{St}]$$

$$I^{max} \Rightarrow^{C} [Ki^{Obj}] \bigcup [Ki^{St}] \qquad (1)$$

где $\Rightarrow^{C}$ - операция соответствия.

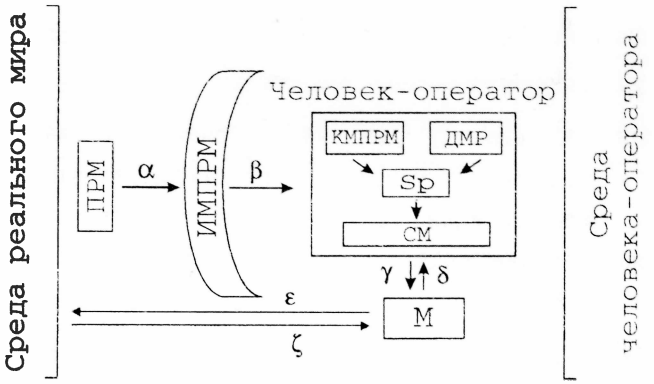


Рисунок 2. Модель опосредствованного управления предметами реального мира

ИМ максимальной насыщенности можно разбить на совокупность документальных ($I i^{d}$), картографических ($I i^{k}$), пространственных ($I i^{p}$) и реальных ($I i^{r}$) ИМ:

$$I^{max} = [I i^{d}] \bigcup [I i^{k}] \bigcup [I i^{p}] \bigcup [I i^{r}] \qquad (2)$$

При проектировании ССОИ необходимо учитывать возможности восприятия человека (I_o). Из этого утверждения вытекает условие реализации ССОИ:

$$I^{max} \leq I_o \qquad (3)$$

Необходимо учесть условие адекватного отражения информационной моделью ситуации (принцип адекватности), которое записывается в виде:

$$I^{max} \Leftrightarrow^{Sm} [K i^{Obj}] \bigcup [K i^{St}] \qquad (4)$$

где $\Leftrightarrow^{sm}$ - операция семантического соответствия.

Каждая ИМ ($I i^{d}$, $I i^{k}$, $I i^{p}$, $I i^{r}$) должна соответствовать концептуальной модели:

$$I j \Leftrightarrow K j \qquad (5)$$

$I_j = I_i^d, I_i^k, I_i^p, I_i^r;$

$K_j = K_i^{Obj}, K_i^{St}.$

Обычно условие (3) для ИМ максимальной насыщенности не выполняется, а поэтому, учитывая принцип симультантности, применяют правило декомпозиции. Декомпозиция бывает: пространственная, временная, семантическая и комбинированная [9]. Тогда ИМ максимальной насыщенности разворачивается в некоторое множество (цепочку) ситуаций I_i^{St}. Эту цепочку ситуаций называют *сценарием*, а информационные модели ситуаций I_i^{St} *кадрами*. Вообще говоря, для некоторых I_i^{St} условие (3) также может не выполняться. В этом случае вновь применяют правило декомпозиции и так до тех пор пока условие (3) не будет выполняться для всех I_i^{St}. Все вышесказанное можно записать как:

$$I^{max} \rightarrow [Sci] \rightarrow [Kdi] \Rightarrow I^{max} \Leftrightarrow [Kdi]; \qquad (6)$$

где:

Sci - сценарий;

Kdi - кадр.

Если ИМ максимальной насыщенности написана на ЕЯ и на основании принципа симультантности имеем:

$$[I^{max}(ЕЯ)] \rightarrow [I^{max}(ОЕЯ)] \rightarrow [I^{max}(ЯОС)] \rightarrow [I^{max}(ЯФОС)] \rightarrow [ГО] \qquad (7)$$

где:

ОЕЯ - ограниченный ЕЯ;

ЯОС - язык описания ситуаций;

ЯФОС - язык формализованного описания ситуаций;

$[I^{max}(j)]$ - описание I^{max} на соответствующем языке,

j = [ЕЯ, ОЕЯ, ЯОС, ЯФОС]

[ГО] - графическое изображение.

На каждое описание I^{max} оказывают влияние некоторые факторы, которые будут называться *граничными условиями* На описание на ЕЯ оказывают влияние общеязыковые требования, правила грамматики, лексики, стилистики языка, на котором написано это описание (в данном случае русского языка). На описание на ОЕЯ оказывают влияние предметная область, в некоторой степени грамматика ЯФОС.

Таким образом задача преобразования ЕЯ описания в ГО сводится к решению системы вербально-логических уравнений:

1. [ЕЯ] → [ОЕЯ];
2. [ОЕЯ] → [ЯФОС];

При выполнении условий принципа адекватности:

$$I^{max} \Leftrightarrow^{Sm} [Ki^{Obj}] \bigcup [Ki^{St}]$$

$$Ij \Leftrightarrow Kj$$

При выполнении ограничения:

$$I^{max} \leq Io$$

Во второй главе описывается методика формирования лексики заданной ПО и разработанная функциональная грамматика. Приводится разработанная концепция словарей трех контекстов предметной области (ПО). Сформирована методика преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ.

Грамматика G = (Vт, Vн, P, S, F) позволяет провести переход представления ситуации из концептуального контекста в вербальный. Где:

Vт - терминальный словарь;

Vн - нетерминальный словарь;

P - правила грамматики G;

S = ПРЕДЛОЖЕНИЕ;

F = [f1, f2, f3, f4, f5];

fi - семантические функции имеют вид:

fi(A, L, T, b1, b2, ... bm); где:

A - переменная, определяющая предметную область;

L - переменная, определяющая место действия ситуации;

T - переменная, определяющая время действия ситуации;

bi - переменные, определяющие и характеризующие объекты;

m - количество переменных, для каждой функции различно;

Для преобразования вербального представления ситуации (описание ситуации в вербальном контексте) в графическое необходимо проанализировать текст описания, т.е. потребуется провести семантический анализ. Семантические значения слов тесно связаны с их синтаксическими значениями в предложении, т.е. потребуется провести синтаксический анализ, а для этого необходимо знать морфологические характеристики слов, которые можно почерпнуть из вербального словаря где они хранятся. Операцию отбора слов из текста определим как *лексический анализ*

Таким образом лексический, синтаксический и семантический анализы определяют последовательность обработки текста, которая необходима для решения задачи преобразования вербального представления ситуации в графическое.

Из задачи преобразования ЕЯ описания в ГО вытекает, что объекты заданной предметной области, и их характеристики представляются в двух формах: вербальном и графическом. Говоря о вербальной форме представления, необходимо помнить, что речь идет о текстовом виде. При формализации поставленной задачи необходимо ввести дополнительную форму представления объектов - формальную. Итак получаем необходимое количество форм представления объектов и их характеристик - вербальная, формальная и графическая. Введем понятие *контекста*, которое означает ту или иную форму представления объектов и их характеристик.

Множество значений *значащих элементов* (объекты и их характеристики заданной предметной области) Ψ конкретного контекста K_j назовем *словарем* этого контекста и обозначим D^j.

$$D^j = \bigcup_i \Psi_i^{K_j}$$

где:

$\Psi_i^{K_j}$ - множество значений значащих элементов $j^{го}$ контекста K_j.

Слово в словаре D^j будем обозначать d^j и называть *записью*, которая состоит из *полей*.

Для решения задачи преобразования «текст-рисунок» используются три контекста и соответстсвенно три словаря:

1. Вербальный D^B;
2. Формальный D^M;
3. Графический $D^Г$.

Методика преобразования состоит из следующих этапов:

I. Извлечение вербальных знаний. Данный этап заключается в исследовании и выделении подмножества ЕЯ, описывающего оперативную ситуационную обстановку.

II. Означивание вербальных знаний. Это означает заполнение значениями словарей каждого конкретного контекста.

III. Порождение описаний ситуций на ограниченном естественном языке. На этом этапе формируются описания на основе правил грамматики $G = (V_т, V_н, P, S, F)$.

Первый этап подразделяется на следующие стадии:

1) Первичная обработка текстов:
   a) выбор текстов по заданной предметной области;
   b) составление словника исходных текстов.
2) Выделение группы языковых единиц, идентифицирующих объекты:
   a) отбор из словника слов, принадлежащих классу существительных;
   b) отбор из языковой группы существительных слов в именительном падеже;

c) отбор из языковой группы существительных в именительном падеже слов, идентифицирующих объекты;

3) Выделение группы языковых единиц, характеризующих объекты:

a) отбор из словника слов, принадлежащих классу прилагательных;

b) отбор слов, характеризующих объекты;

c) провести классификацию прилагательных по признакам объектов.

В третьей главе приводятся правила формирования лексики заданной ПО. Описываются разработанные методика структуризации характеристик объектов заданной ПО, метод переходных коэффициентов и алгоритмы поддержки преобразования ЕЯ описания ситуации в ГО.

В системах управления ЕЯ выступает прежде всего как средство описания тех объектов, которыми управляют, ситуаций, складывающихся с этими объектами. Что побуждает поиск того множества языковых единиц, с помощью которых можно выражать понятия и отношения важные для управления. В третьей главе были сформулированы правила, с помощью которых формируется такое множество.

1. Лексические исследования. Выделить подмножество существительных, глаголов, прилагательных и т.п.
2. Грамматическая формализация. Сформировать синтаксическую структуру, определить тип предложений, установить порядок членов предложения.
3. Означивание вербальных знаний. Статистические исследования числовых значений ЛП «СЛЕВА», «В ЦЕНТРЕ», «В ПРАВОМ НИЖНЕМ УГЛУ» и т.п.

Для решения этих задач были поставлены два эксперимента. Тексты заданий приведены в приложении 1. В первом эксперименте от экспериментируемых требовалось:

1. Сформировать композицию (ситуацию) из геометрических фигур.
2. Описать воображаемую ситуацию на ЕЯ (русском).
3. Изобразить эту ситуацию на листе формата А4.

4. Составить картинный словарь.
5. Составить тезаурус.

Во втором эксперименте полученные описания на ЕЯ были розданы экспериментируемым, по которым последние должны были:

1. Восстановить изображения ситуаций по описаниям.
2. Провести грамматические преобразования (привести к заданному типу) предложения описаний.

По результатам экпериментов были сформированы:

- лексический состав терминального словаря V_T грамматики G:
- грамматические правила P грамматики G построения предложений, описывающих ситуации:
- метод переходных коэффицентов, основанный на правилах теории нечетких множеств.

В четвертой главе описывается макетная реализация системы автоматизированного проектирования сценариев, которая доказывает эффективность предложенных подходов в решении задачи преобразования ЕЯ описания оперативной ситуационной обстановки в ГО ССОИ.

В заключении формулируются основные выводы, полученные в работе.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ.

Конкретно научные и практические результаты состоят в следующем:

1. Разработана методика преобразования естественно-языкового описания оперативной ситуационной обстановки в графические образы ССОИ.
2. При разработке методики преобразования естественно-языкового описания в графические образы были выделены основные задачи, которые состоят в следующем:
   - анализ текста;
   - выработка концепции лингво-графической базы знаний.

Решение данных задач позволило провести требуемое преобразование.

3. Было установлено, что для проведения анализа текста необходимо предварительно:
   - выделить подмножество естественного языка;
   - сформировать правила построения предложений, поддающихся формальной обработке.
4. Был предложен процесс вербализации заданной предметной области. На его основе была разработана методика выделения подмножества естественного языка, которая формирует лексику этой предметной области.
5. Были выделены и сформулированы правила построения предложений, на основе которых была построена формальная грамматика. Применение данной грамматики позволяет проводить синтаксический и семантический анализы предложений.
6. Выработка концепции лингво-графической базы знаний потребовала решения следующих проблем:
   - способ представления объектов и их характеристик;
   - разработка структур словарей, входящих в состав лингво-графической базы знаний;
   - заполнение соответствующими значениями словарей, иначе говоря, провести означивание заданной предметной области.
7. Были выделены три формы представления объектов и их характеристик: вербальный, графический и формальный, которые описываются в соответсвующих базах данных (словарях).
8. На основе выделенных форм представления и характера задачи преобразования были разработаны структуры словарей.
9. Для проведения означивания была разработана методика структуризации объектов и их характеристик.
10. Опираясь на выделенную лексику и на проведенную структуризацию объектов и их характеристик, была разработана

методика означивания вербальных знаний о объектах, их характеристиках и пространственных отношениях.

11. Для решения задачи неоднозначности, возникающей при отображении некоторых характеристик объектов и пространственных отношений, был предложен метод переходных коэффицентов.
12. Разработан пакет алгоритмов поддержки преобразования естественно-языкового описания в графические образы, в который входят алгоритмы:
    - морфологического, синтаксического и семантического анализа;
    - размещения графических образов;
    - генерации сценариев.

    Данные алгоритмы были разработаны на основе методики формирования лексики, построенной грамматики, методике структуризации объектов и их характеристик, разработанных структурах словарей и методе переходных коэффицентов.
13. В практической части диссертации, на основе созданных программных комплексов, показана состоятельность разработанной методики преобразования естественно-языкового описания оперативной ситуационной обстановки в графические образы ССОИ.
14. Разработанные методики, алгоритмы и программные средства использовались при создании комплекса программного и математического обеспечения системы отображения коллективного пользования объекта 14П04 - КПМО СОКП в РНИИКП.

Основное содержание диссертации отражено в следующих работах:

1. Архив графических и картографических инфомационных моделей: Отчет о НИР / НПП "Экотек", Рук. Ю.Н. Филиппович - Инв. № 3/06-3. - М., 1990. - Книга 6. - 56 с.

2. Комплекс программных средств создания и ведения базы графических данных объекта 14П04 - КПМО СОКП: Отчет о НИР / РНИИ КП, Рук. Ю.Н. Филиппович - № ГР 07256; Инв. № к8532.- М., 1991. - Книга 9. - 287 с.
3. Макет комплекса программных средств создания и ведения архива графических и картографических информационных моделей: Отчет о НИР / НПП "Экотек", Рук. Ю.Н. Филиппович - Инв. № 3/03-01. -М., 1990. - Книга 3. - 63 с.
4. Пардаев А.А.., Филиппович Ю.Н., Современные системы компьютерной анимации //Компьютерная хроника. - 1992.. - №5-6. - С. 5.
5. Филиппович Ю.Н., Пардаев А.А. SAGO - система создания и ведения архива графических образов: Руководство пользователя. - М.: НПП "Фрегат", 1991. - 15 с.
6. Черкасова Г.А., Пардаев А.А. Комплекс программ автоматизированного построения контекстно-семантических тезаурусов//Тезисы докладов третьей всесоюзной конференции по созданию машинного фонда русского языка. - М., 1989. - С. 93.

Подписано к печати 19.06.98г. Зак.113 Объем 1 п.л. Тир 100

Типография МГТУ им. Н.Э.Баумана