

III. СЕМАНТИЧЕСКИЕ МОДЕЛИ ДАННЫХ

3.1. Принципы семантического моделирования

Общий подход к проблеме семантического моделирования

1. Выявление **неформальных семантических концепций** (понятий) человеческого восприятия и осмысления объектов окружающего мира.
2. Синтез **формальных понятий модели**, которые могут использоваться для представления в системе реальных объектов определенных ранее семантических концепций в виде формальных объектов.
3. Определение набора **формальных правил для задания ограничений целостности**, предназначенных для переноса законов окружающего мира в систему формальных объектов.
4. Определение набора **формальных операторов** для манипулирования формальными объектами.

3.2. Модель данных «Сущность-Связь» (ER-модель)

3.2.1. Структуры

Структурные понятия ER-модели Чена

Множество сущностей (МСу) – E *СЛУЖАЩИЙ*

Множество связей (МСв) – R

$$R \subset \{ \langle e_1, e_2, \dots, e_n \rangle \mid e_1 \in E_1 \wedge e_2 \in E_2 \wedge \dots \wedge e_n \in E_n \}$$

$$\text{УПРАВЛЕНИЕ} = \{ \langle e_1, e_2 \rangle \mid e_1 \in \text{СЛУЖАЩИЙ} \wedge e_2 \in \text{СЛУЖАЩИЙ} \}$$

Роль - r

$$R \subset \{ \{ r_1 : e_1, r_2 : e_2, \dots, r_n : e_n \} \mid e_1 \in E_1 \wedge e_2 \in E_2 \wedge \dots \wedge e_n \in E_n \}$$

$$\begin{aligned} \text{УПРАВЛЕНИЕ} = \{ \{ \text{Начальник} : e_1, \text{Подчиненный} : e_2 \} \mid \\ \mid e_1 \in \text{СЛУЖАЩИЙ} \wedge e_2 \in \text{СЛУЖАЩИЙ} \} \end{aligned}$$

Множество значений (МЗн) – V *ЦПЧ или Строки*

Атрибут - f

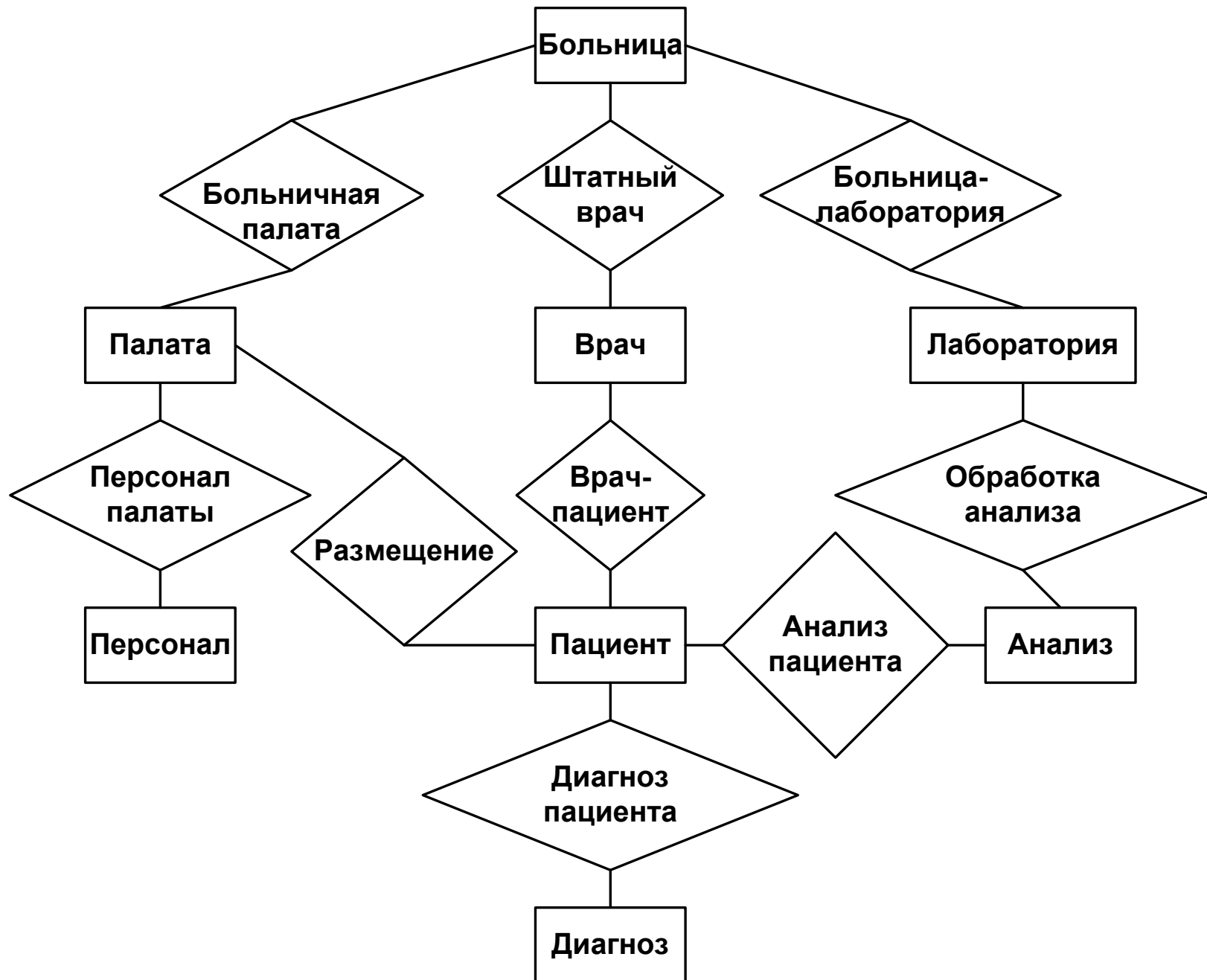
$$f : E \rightarrow V \quad \text{Фамилия: ПАЦИЕНТ} \rightarrow \text{Строки}$$

$$f : R \rightarrow V \quad \text{Номер койки: РАЗМЕЩЕНИЕ} \rightarrow \text{ЦПЧ}$$

$$f : E \rightarrow V_1 \times V_2 \times \dots \times V_n \quad \text{Адрес: ПАЦИЕНТ} \rightarrow \text{Строки} \times \text{ЦПЧ}$$

$$f : R \rightarrow V_1 \times V_2 \times \dots \times V_n \quad \begin{aligned} \text{Серия_Номер_Свидетельства: БРАК} \rightarrow \\ \text{ЦПЧ} \times \text{ЦПЧ} \end{aligned}$$

ER-диаграмма медицинской ПрО (структурный аспект)



Е
R

м
о
д
е
л
ь

С
Н
Е
Н
'
а

Проблема триализма в ER-модели

Выбирая форму данных, помните о возможном **триализме** явлений моделируемого мира. Так, брак между людьми можно представлять себе как атрибут, множество связей или множество сущностей. Проектируя схему БД, необходимо сделать выбор в пользу одного из этих представлений. При этом можно использовать следующие соображения.

Если вам достаточно информации о том, состоит человек в браке или нет, можно рассмотреть это явление как атрибут множества сущностей ЧЕЛОВЕК. Если вас к тому же интересует, с кем конкретно заключен брак, необходимо трактовать БРАК как множество связей между сущностями множества ЧЕЛОВЕК. И, наконец, если вам потребуется формулировать высказывания по поводу этого явления (о наличии у него определенных характеристик или связей), нужно использовать самостоятельное множество сущностей БРАК.

Проблема триализма в ER-модели

Атрибут

ЧЕЛОВЕК (...,Замужем/Женат?,...)

– множество сущностей

Связь

ЧЕЛОВЕК – множество сущностей

БРАК (Муж, Жена) – множество связей

Сущность

ЧЕЛОВЕК – множество сущностей

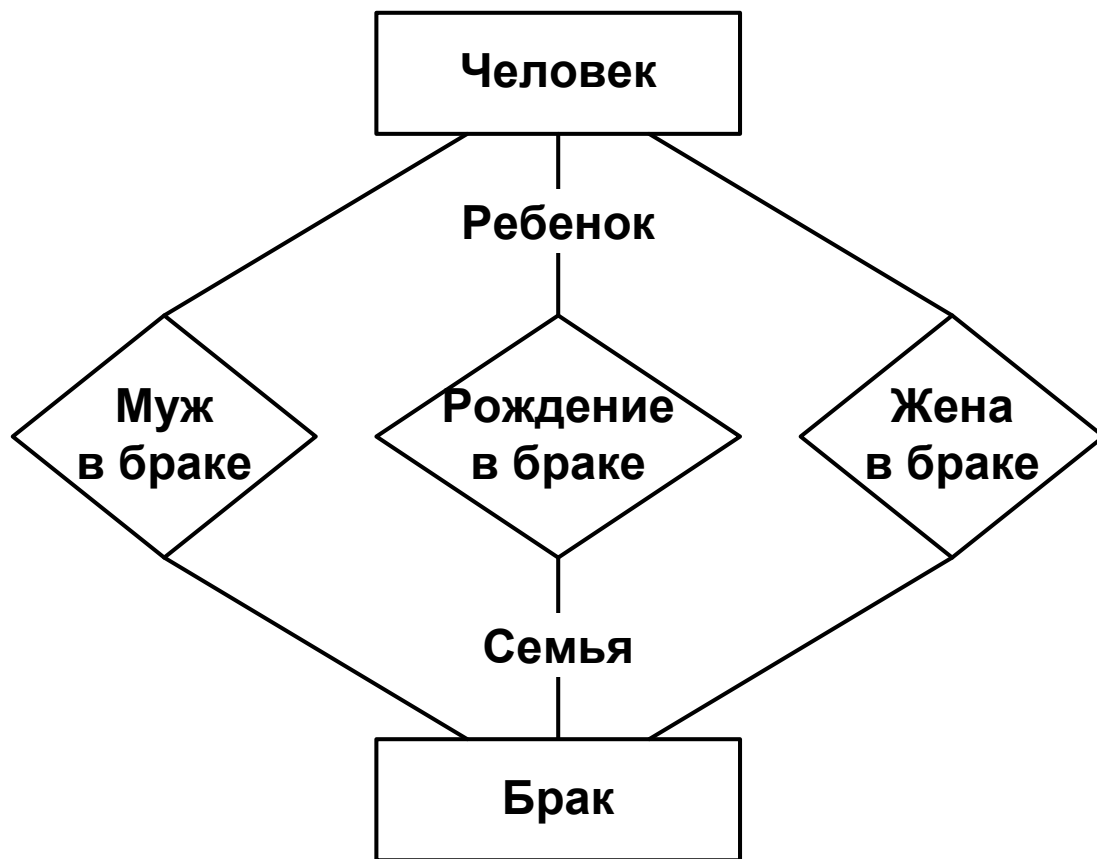
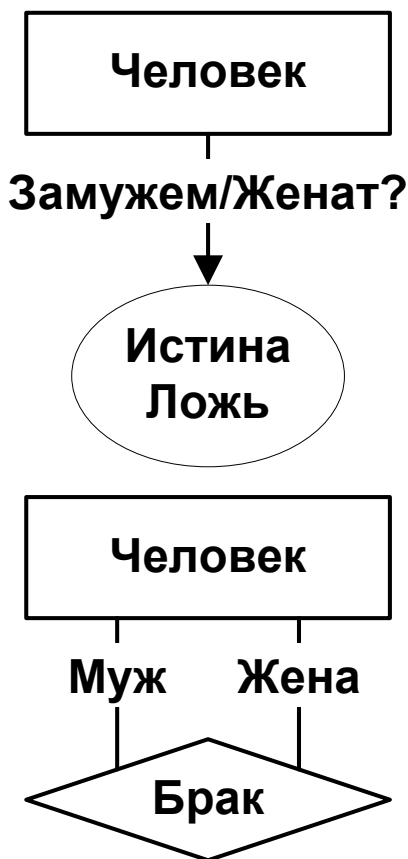
БРАК – множество сущностей

РОЖДЕНИЕ_В_БРАКЕ –

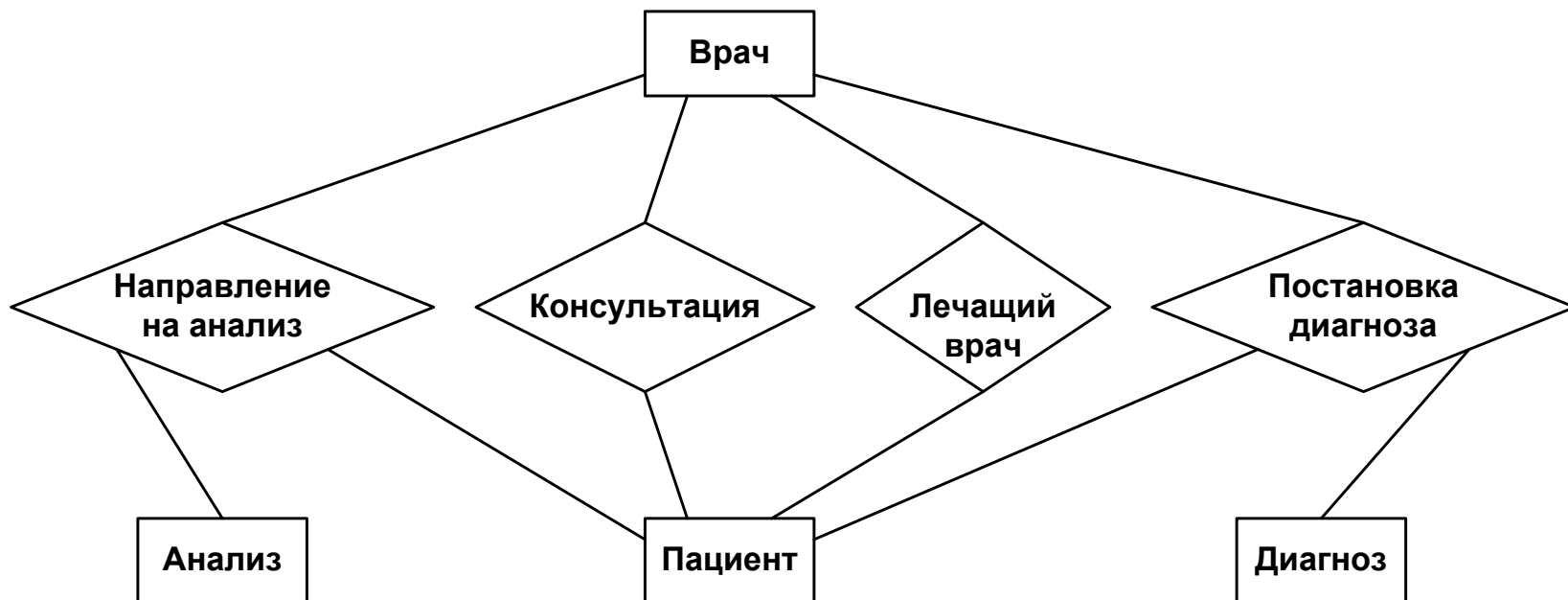
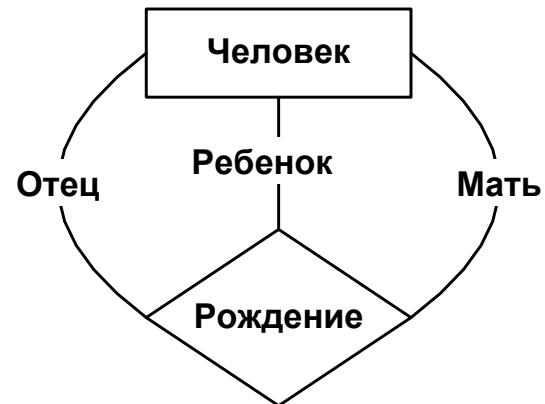
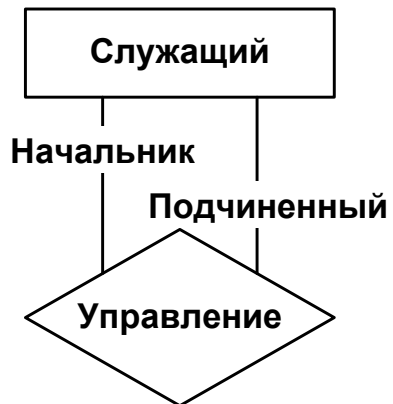
множество связей

МУЖ - множество связей

ЖЕНА - множество связей



Разновидности множеств связей

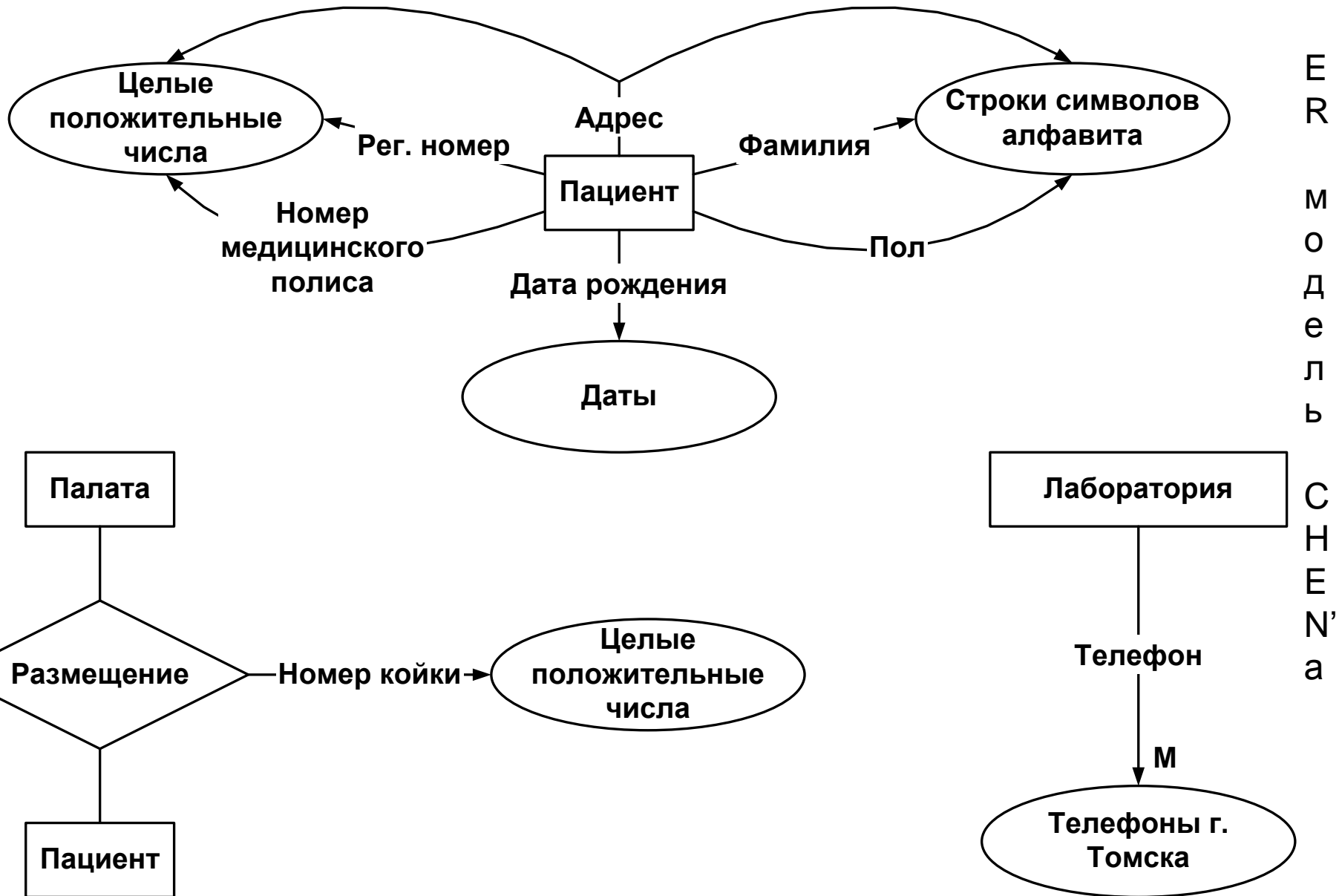


Е
R

М
О
Д
Е
Л

С
Н
Е
Н
'
а

Разновидности атрибутов



Экстенционал БД в форме графа

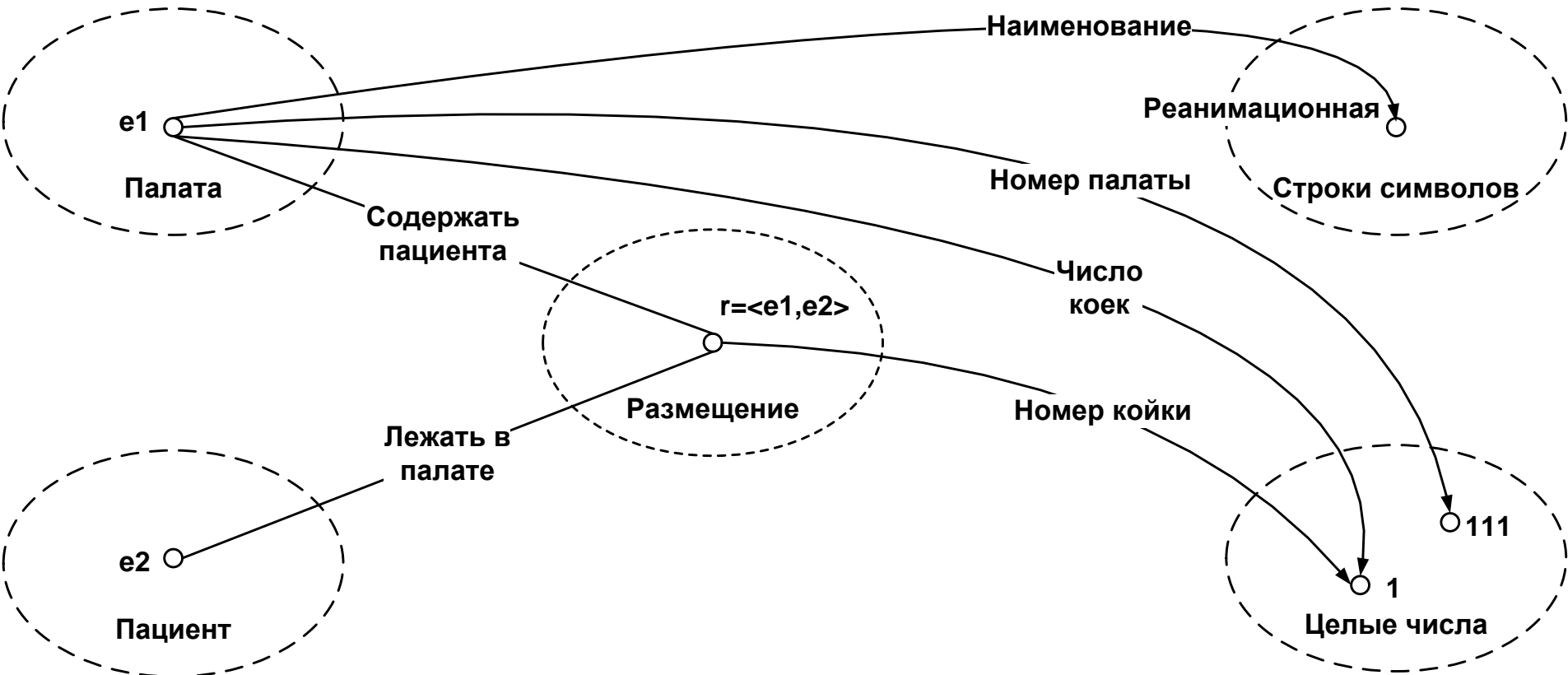
Множества сущностей

Роли

Множества связей

Атрибуты

Множества значений



Экстенсионал БД в форме таблиц

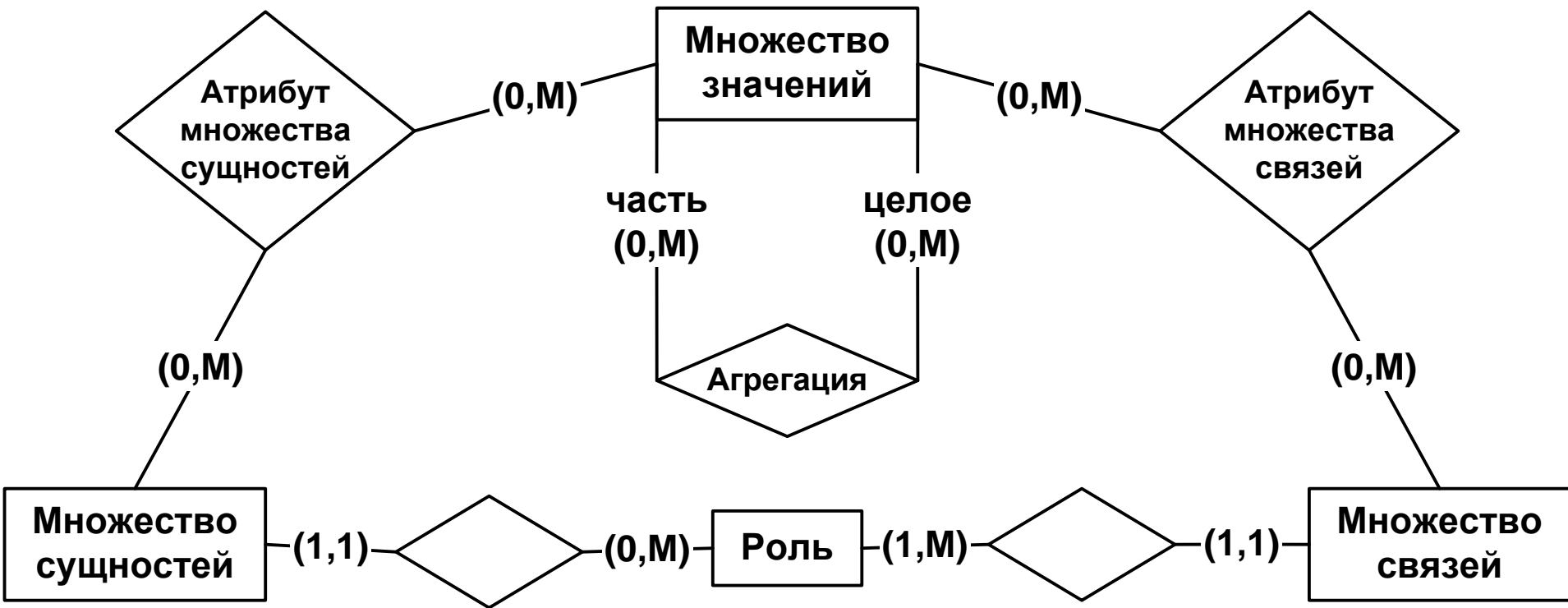
Множественность

Сущности	Атрибуты	Номер палаты	Наименование	Число коек
	ПАЛАТА	Целые числа	Строки символов	Целые числа
	e1	111	Реанимационная	1
	e5	555	После-операционная	2

Множественность

Содержать пациента		Лежать в палате
РАЗМЕЩЕНИЕ		Номер койки
ПАЛАТА	ПАЦИЕНТ	Целые числа
e1	e2	1
e5	e10	1
...

Метасхема ER-модели



3.2.2. Ограничения целостности

1) Ограничения целостности на значения атрибутов

Множество допустимых значений атрибутов можно задавать:

1) указанием соответствующего множества значений

Фамилия: ПАЦИЕНТ -> Строки

Номер койки: РАЗМЕЩЕНИЕ -> ЦПЧ

**2) сравнением с константой или значением атрибута
(возможны более сложные выражения)**

Назначенное время (a) < 24, где a ∈ АНАЛИЗ

3) перечислением значений

Пол (a) ∈ {'м', 'ж'}, где a ∈ ПАЦИЕНТ

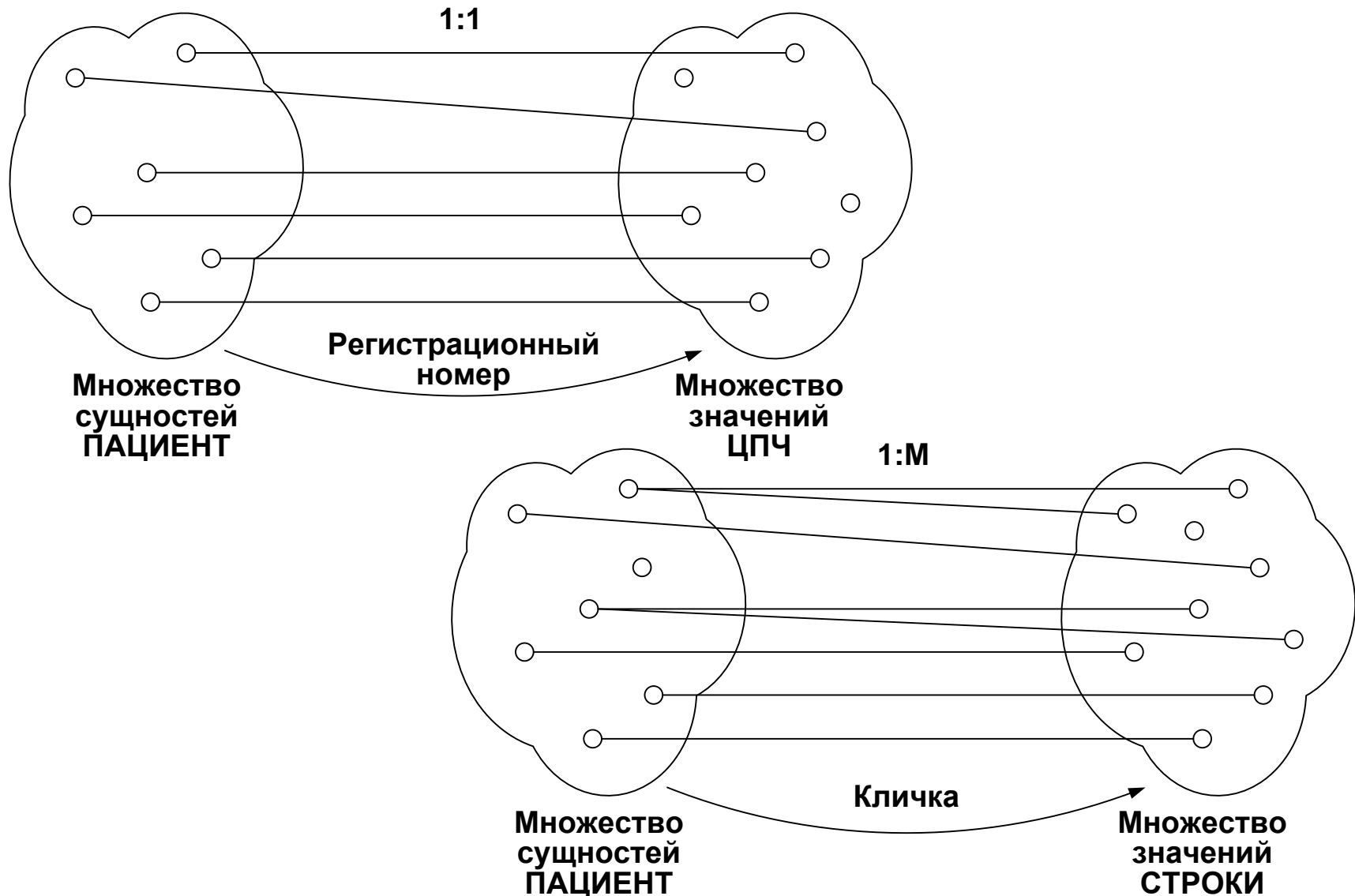
**4) более сложным логическим выражением, включающим
в виде атомов конструкции 2 – 3**

Назначенное время (a) >= 0 AND

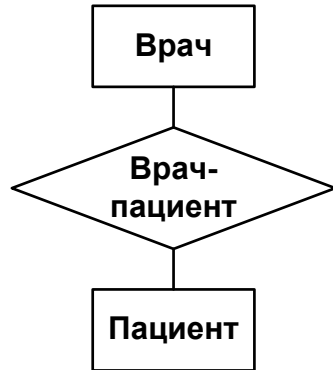
Назначенное время (a) < 24, где a ∈ АНАЛИЗ

2А) Ограничения целостности на отображения между атрибутами одного множества сущностей или связей

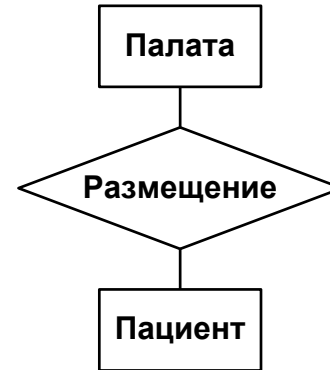
Ключ сущности



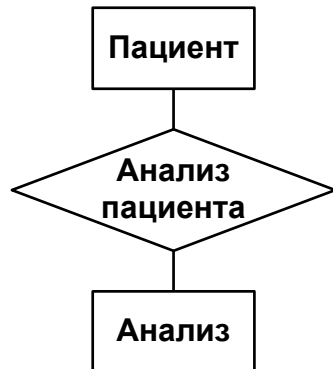
Уникальная идентификация связи



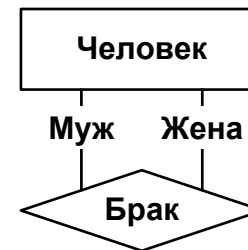
по всем сущностям,
в ней участвующим



по одной сущности типа ПАЦИЕНТ или
по сущности типа ПАЛАТА и значению атрибута связей
Номер койки

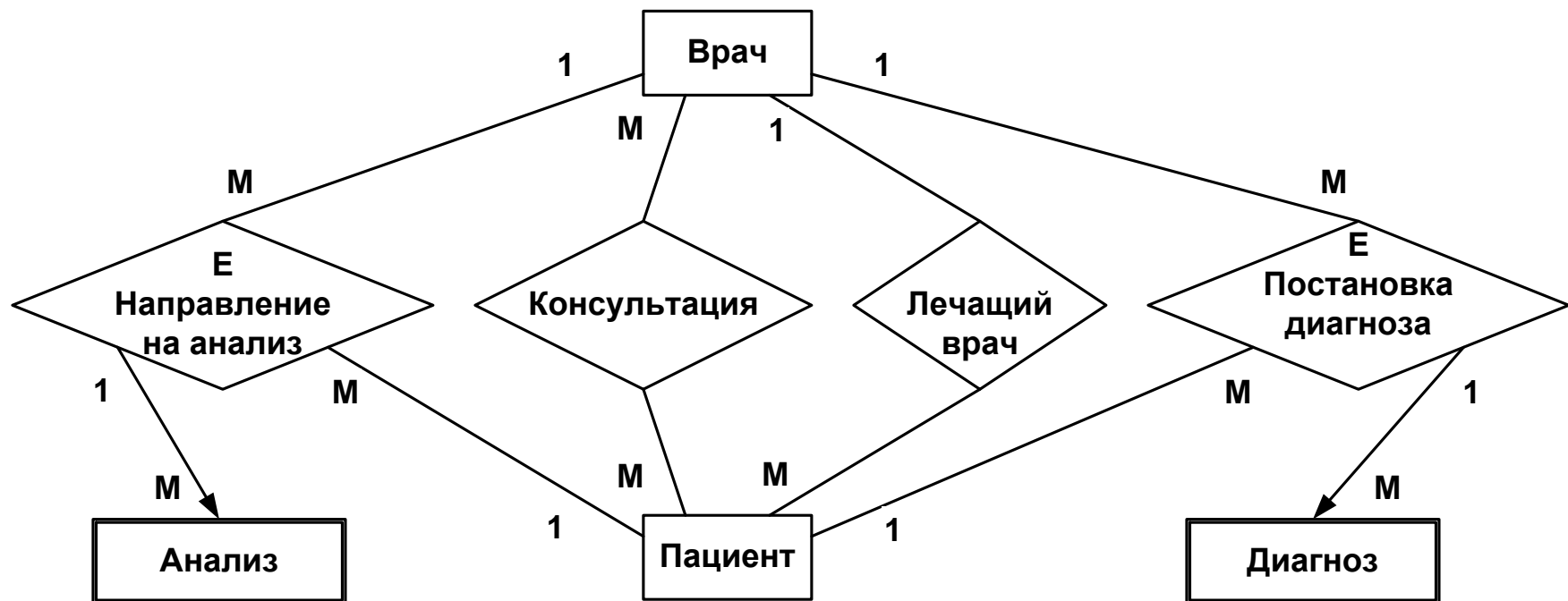
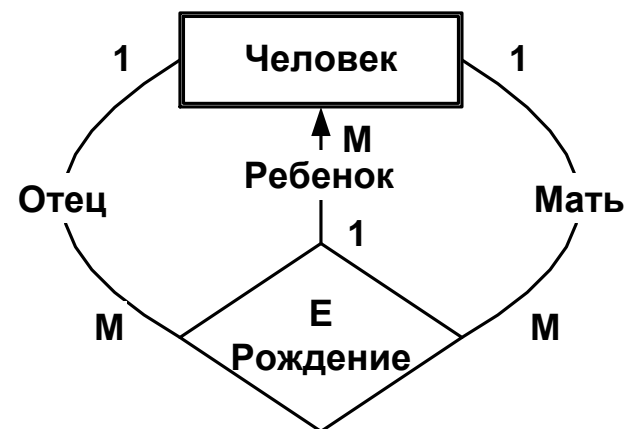
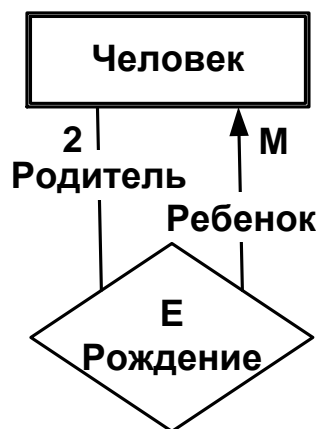
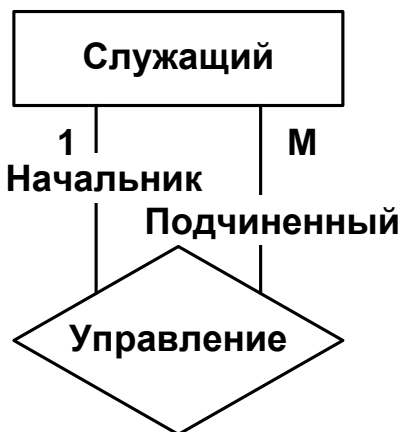


по одной сущности
типа АНАЛИЗ

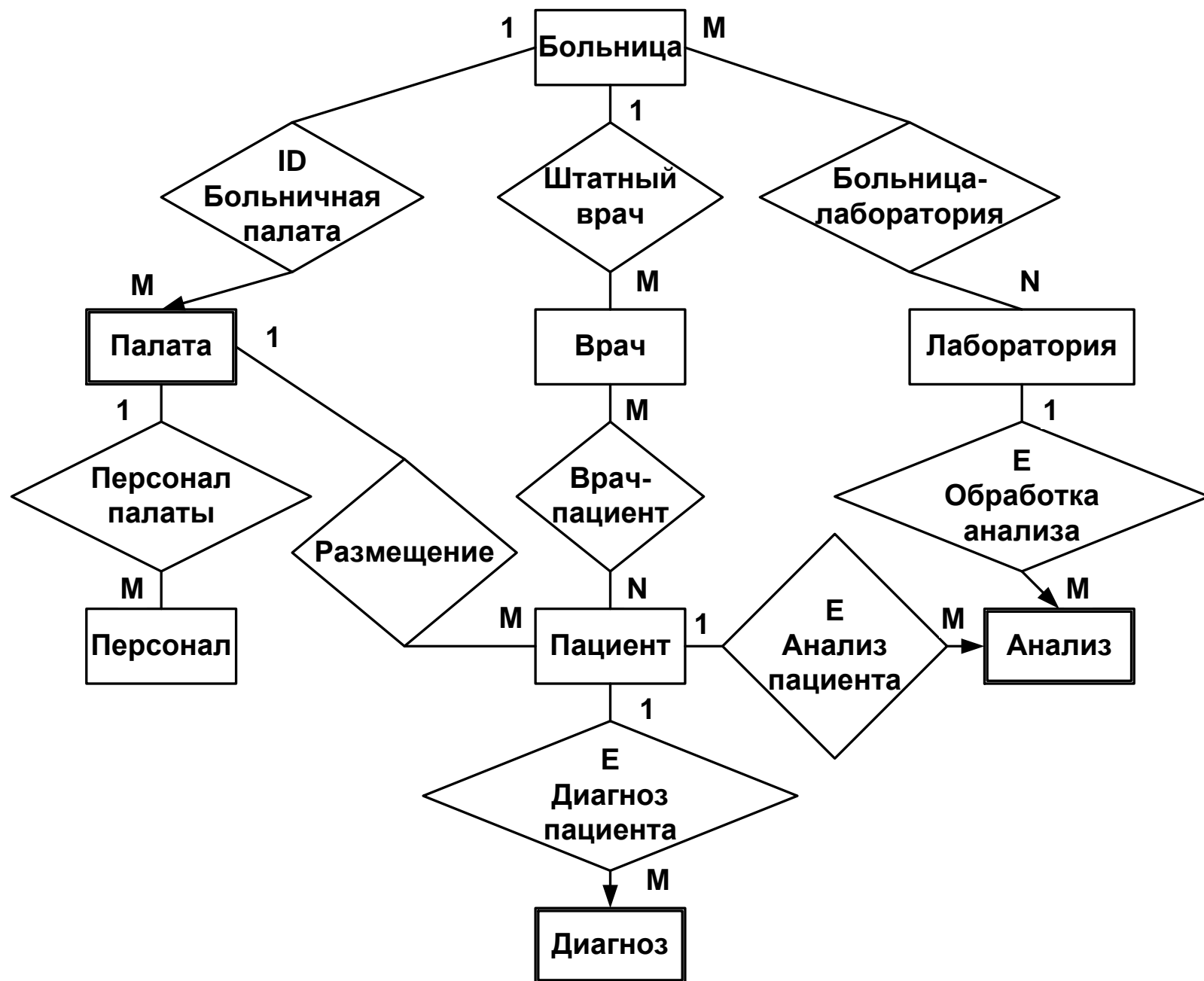


по одной сущности типа ЧЕЛОВЕК в роли Муж или Жена и
значению атрибута связей Дата регистрации брака или
по значению атрибута связей Серия и Номер Свидетельства

2Б) Ограничения целостности на отображения между множествами сущностей в контексте множества связей



Полная ER-диаграмма медицинской предметной области



ER
м
о
д
е
л
ь
С
Н
Е
Н
'
а

3.2.3. Операции

Операции навигационного языка

- 1) Создание сущностей (с возможностью задания экземпляров атрибутивных отображений)**
- 2) Изменение сущностей (изменение экземпляров атрибутивных отображений)**
- 3) Удаление сущностей**
- 4) Создание связей (с указанием сущностей для всех ролей и экземпляров атрибутивных отображений)**
- 5) Изменение связей (изменение экземпляров атрибутивных отображений)**
- 6) Удаление связей**
- 7) Селекция сущности по указателю текущей**
- 8) Селекция связи по указателю текущей**
- 9) Селекция сущности по указателю текущей и по данным**
- 10) Селекция связи по указателю текущей и по данным**
- 11) Селекция сущности по указателю текущей и по связи**
- 12) Селекция связи по указателю текущей и по связи**
- 13) Селекция сущности по указателю текущей, по данным и по связи**
- 14) Селекция связи по указателю текущей, по данным и по связи**
- 15) Выборка значений атрибутов для текущей сущности или связи**

Спецификационный язык CABLE (ChAin Based Language)

$$\left[\begin{array}{l} \{ \langle \text{Множество сущностей} \rangle \\ \langle \text{Множество связей} \rangle \} \end{array} \right] [.] [\langle \text{Условие} \rangle]$$

1. Получить фамилии врачей-окулистов

OUTPUT ВРАЧ.Фамилия

SELECT ВРАЧ.Специальность = 'окулист'

2. Получить фамилии пациентов, лежащих в больнице Святой Елены.

OUTPUT ПАЦИЕНТ.Фамилия

SELECT БОЛЬНИЦА.Название = 'Святая Елена' / ПАЛАТА

3. Получить фамилии и специальности врачей, у которых есть пациенты больницы Святой Елены, выполнявшие анализы в лаборатории «Друг желудка».

OUTPUT ВРАЧ.Фамилия, Специальность

(SELECT БОЛЬНИЦА.Название = 'Святая Елена' / ПАЛАТА /
ПАЦИЕНТ)

\cap (SELECT ЛАБОРАТОРИЯ.Название = 'Друг желудка' /
АНАЛИЗ / ПАЦИЕНТ)

3.2.4. Назначение ER-модели

Критерии качества семантической модели

- 1. Модель должна обладать достаточной общностью и ясностью для того, чтобы в ней можно было легко представить любые явления и законы моделируемого мира.**
- 2. Разрыв между этой моделью и моделями, реализуемыми в СУБД, не должен быть большим, желательно, чтобы можно было воспользоваться максимально формальными правилами преобразования структур и ограничений целостности из одной модели в другую.**

3.2.5. Модификации ER-модели Цена

Общая схема модификаций ER-модели Чена

ER-модель Чена

+

Специализации и категоризации

=

Enhanced ER-модель (EER-модель)

-

Множества связей степени больше 2

Атрибуты множеств связей

Многозначные атрибуты

Атрибутные отображения в декартово произведение множеств значений

=

ER-модель Баркера (Oracle Designer)

-

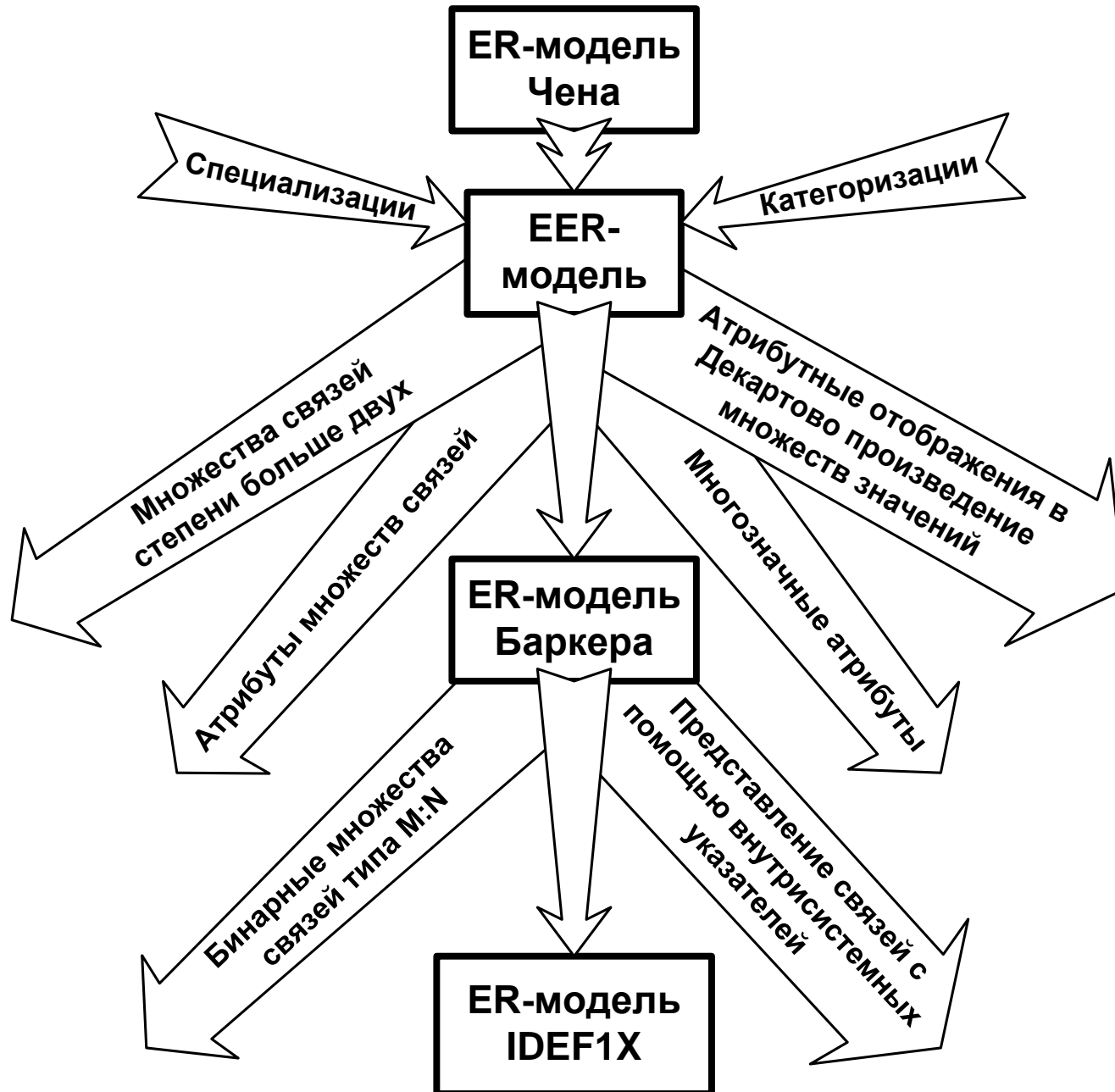
Бинарные множества связей типа M:N

Представление связей с помощью внутрисистемных указателей
(необходимо дублировать атрибуты множеств сущностей)

=

ER-модель IDEF1X (ERWin)

Модификации ER-модели



Специализации и категоризации

1. Специализации и категоризации являются традиционной естественной формой представления информации для человека.
2. Специализация предполагает перенос определений некоторых атрибутов и множеств связей на уровень суперкласса, благодаря этому фиксируется важная информация о том, что данные атрибуты и типы связей имеют смысл не только для экземпляров конкретного подкласса, но и для суперкласса в целом, а, значит, и для всех остальных его подклассов.
3. Специализации и категоризации позволяют исключить дублирующие определения атрибутов и множеств связей для «родственных» множеств сущностей, в результате чего удастся избежать некоторых ошибок и сократить время, затрачиваемое на проектирование.
4. Процесс семантического моделирования зачастую ведется в терминах различных абстракций одного и того же объекта (например, конкретный человек может быть представлен и сущностью типа ВРАЧ, и сущностью типа ПАЦИЕНТ одновременно), а специализации и категоризации позволяют в таких случаях восстановить целостность объекта реального мира.
5. Выделение подкласса в процессе специализации является своего рода гарантией применимости связей и значений атрибутов только к тем сущностям, для которых они актуальны.
6. Специализация призвана упростить процесс моделирования, позволяя постепенно увеличивать сложность схемы, в то время как генерализация предоставляет проектировщику возможность работать с большими схемами, выбирая различные уровни абстракции.

Ограничения целостности на специализации и категоризации

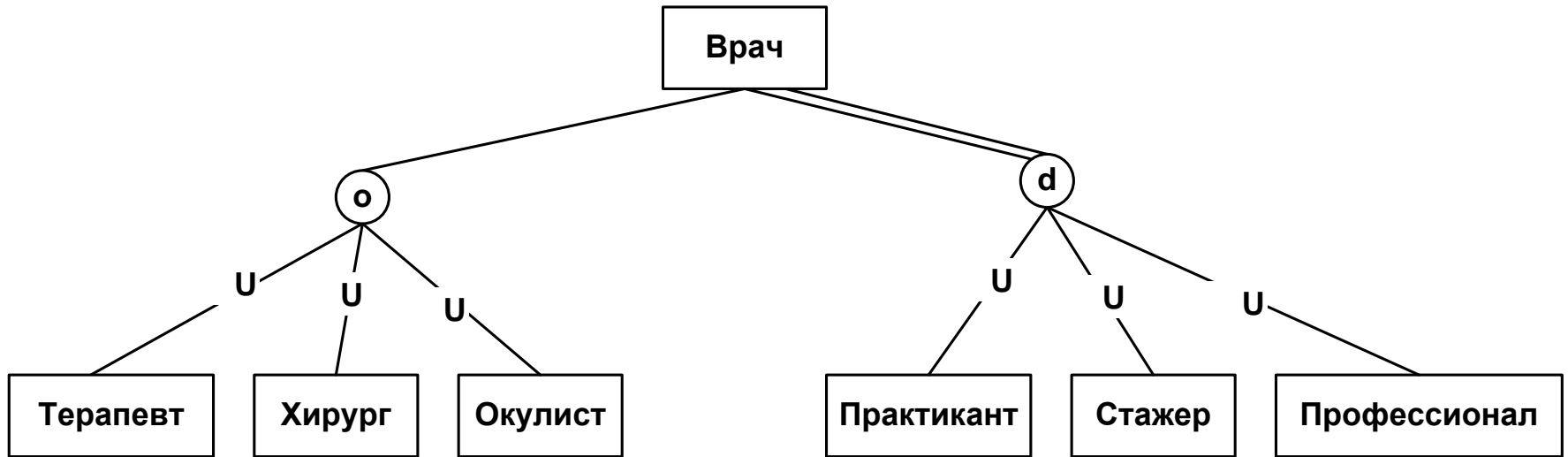
На специализации могут быть наложены ограничения.

Первое ограничение называется **ограничением непересечения**. Оно гласит, что если подклассы некоторой специализации не пересекаются, то каждый объект может быть представлен сущностью только одного из подклассов данной специализации. Если подклассы специализации пересекаются, один объект может быть представлен сущностями сразу нескольких подклассов специализации.

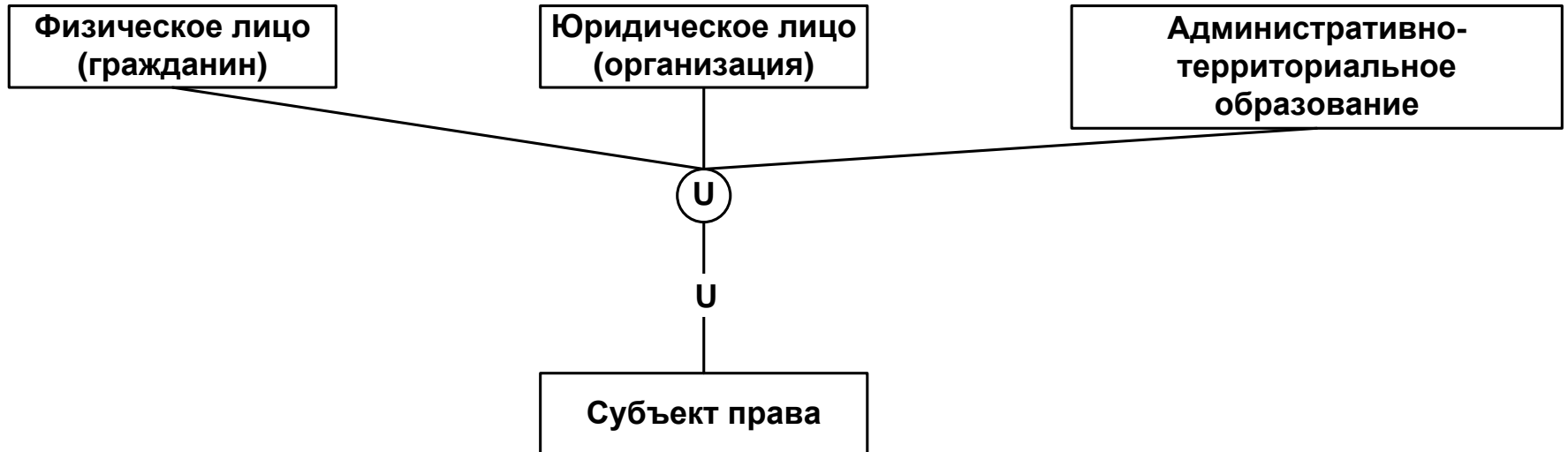
Второе ограничение специализации называется **ограничением участия**, оно может быть полным или частичным. Специализация с полным участием означает, что каждый объект, представленный сущностью суперкласса, должен быть представлен сущностью хотя бы одного подкласса этой специализации. Специализация с частичным участием означает, что объект, представленный сущностью суперкласса, не обязательно должен быть представлен сущностью какого-либо подкласса этой специализации.

Категоризация может быть дополнительно детализирована с учетом **полного или частичного участия** объектов суперклассов в категории. При полном участии каждый объект всех суперклассов должен быть представлен сущностью данной категории. При частичном участии это ограничение устраняется, и всем объектам всех суперклассов иметь сущности данной категории не обязательно.

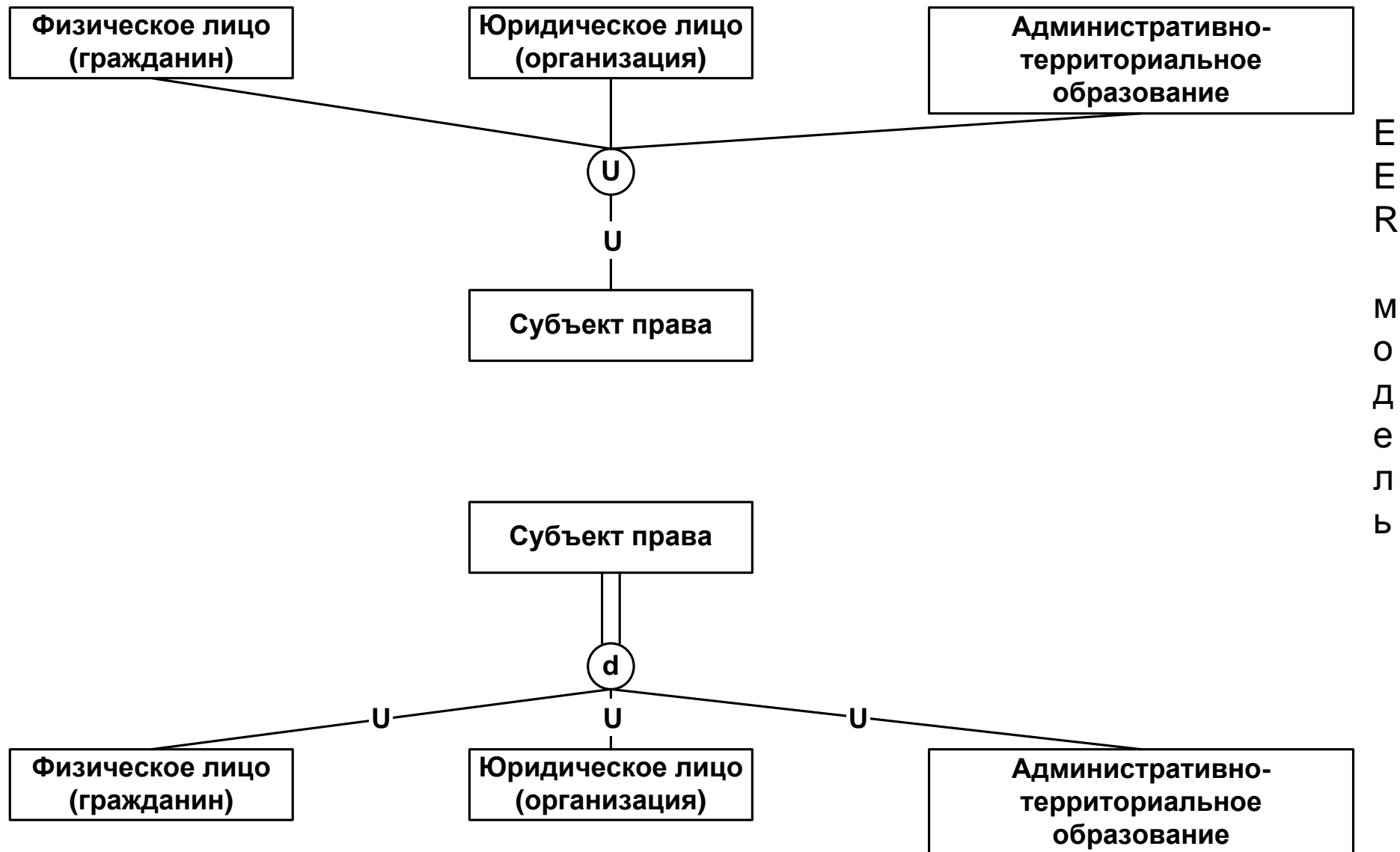
Специализации на ER-диаграмме



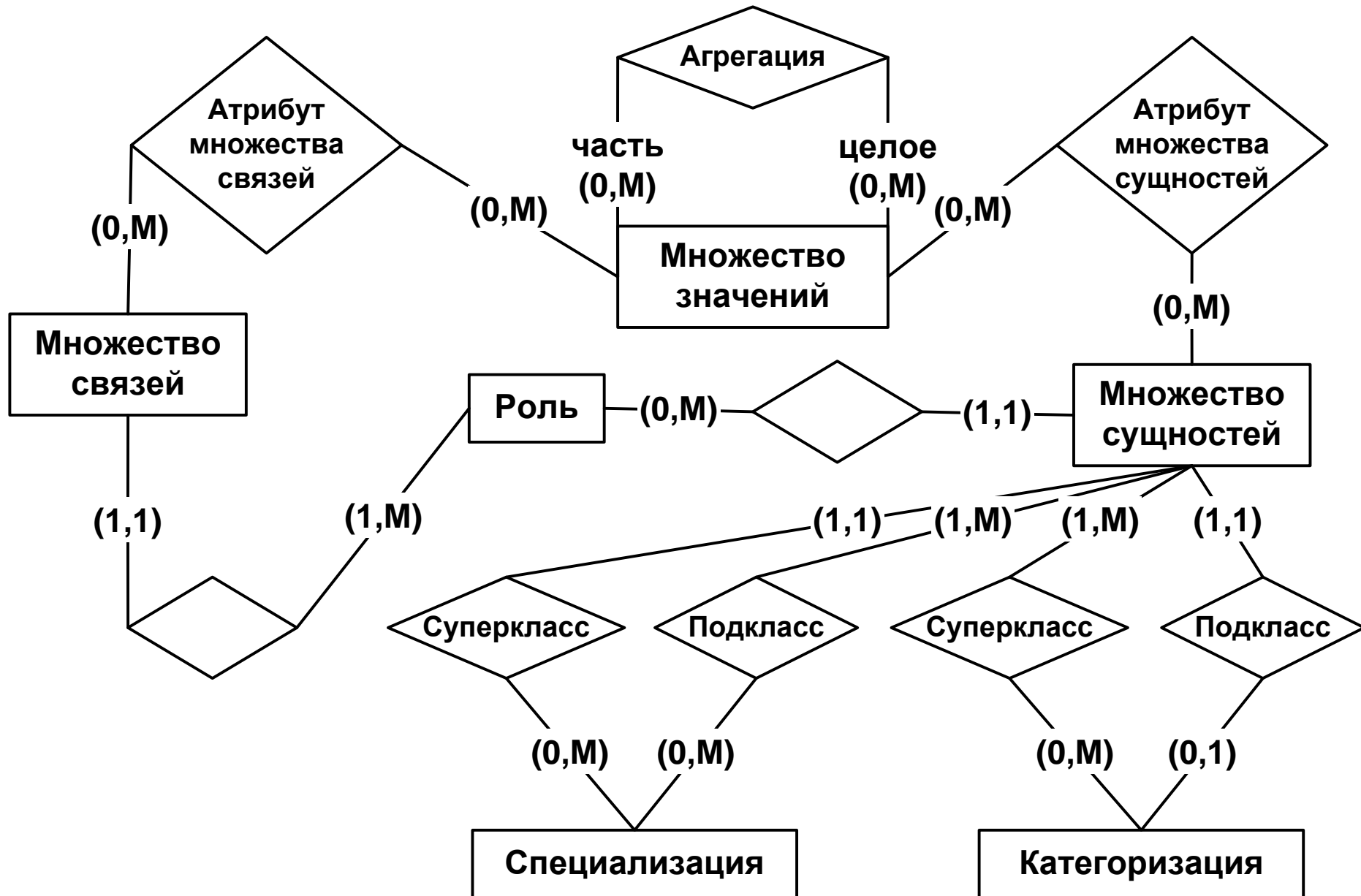
Категоризация на ER-диаграмме



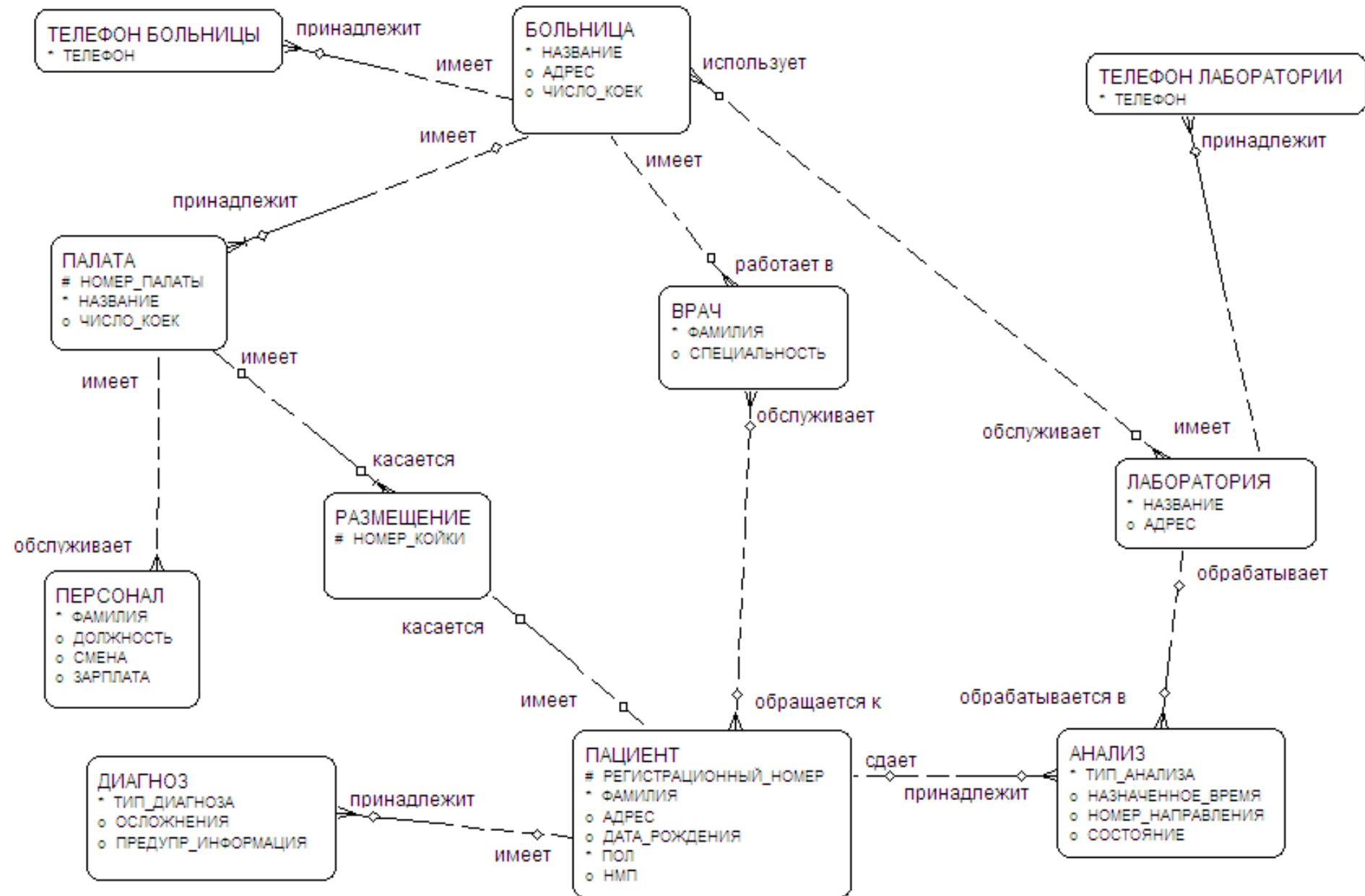
Сравнение частичной категоризации и полной непересекающейся специализации



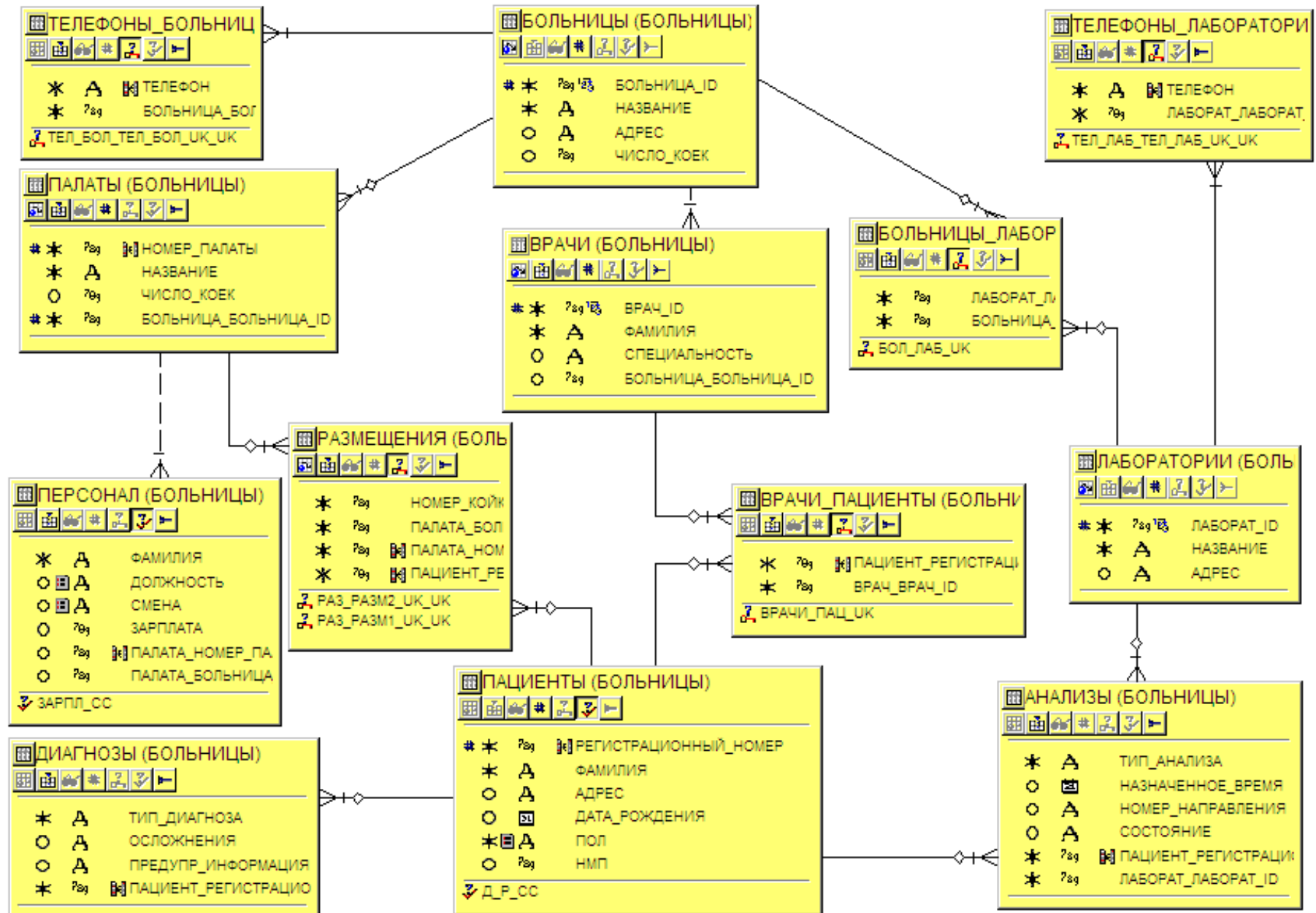
Метасхема EER-модели



ER-диаграмма медицинской предметной области в нотации Баркера (Oracle Designer)



Сгенерированная реляционная схема медицинской предметной области (Oracle Designer)



ER-диаграмма медицинской предметной области в нотации IDEF1X (Design/IDEF)

