

**СОВРЕМЕННЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ  
ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ  
В РАЗВИТИИ ДИАГНОСТИЧЕСКИХ БАЗ ЗНАНИЙ  
ПО КОНКРЕТНОЙ ПРОБЛЕМЕ**

*Карп В.П.*, д.т.н., проф.

*Саруханов Б. А.*, аспирант

Московский Государственный Институт  
Радиотехники Электроники и Автоматики  
(технический университет)  
Москва, МИРЭА

# Технология формирования новых

знаний



Обсуждение конкретной проблемы

Эксперт

Когнитолог  
- аналитик

Декларативные знания

Интуитивные знания

Новые знания

Семантическая  
интерпретация ДБЗ  
(новые знания)

Задачи:

- ✓ Диагностика;
- ✓ Прогнозирование;
- ✓ Принятие управленческих решений;
- ✓ Выбор лечебного воздействия;
- ✓ ...

Алгоритм  
графического отображения  
связей в ДБЗ

База  
формализованных  
знаний

Диагностическая  
база знаний  
(ДБЗ)

База данных  
Data Mining  
«Consilium»

АРМ

ЭС

...

СППР

Пользователь



# Цель создания интеллектуальных технологий



**Обеспечить информационную поддержку  
пользователям для принятия решений  
в сложных ситуациях**

## **ЗАДАЧИ:**

**Обеспечить корректность анализа,  
учитывая особенности медицинских данных:**

- многообразие проявления заболевания,
- наличие большого количества пропущенных данных,
- их неоднозначность,
- зашумленность и т.д.

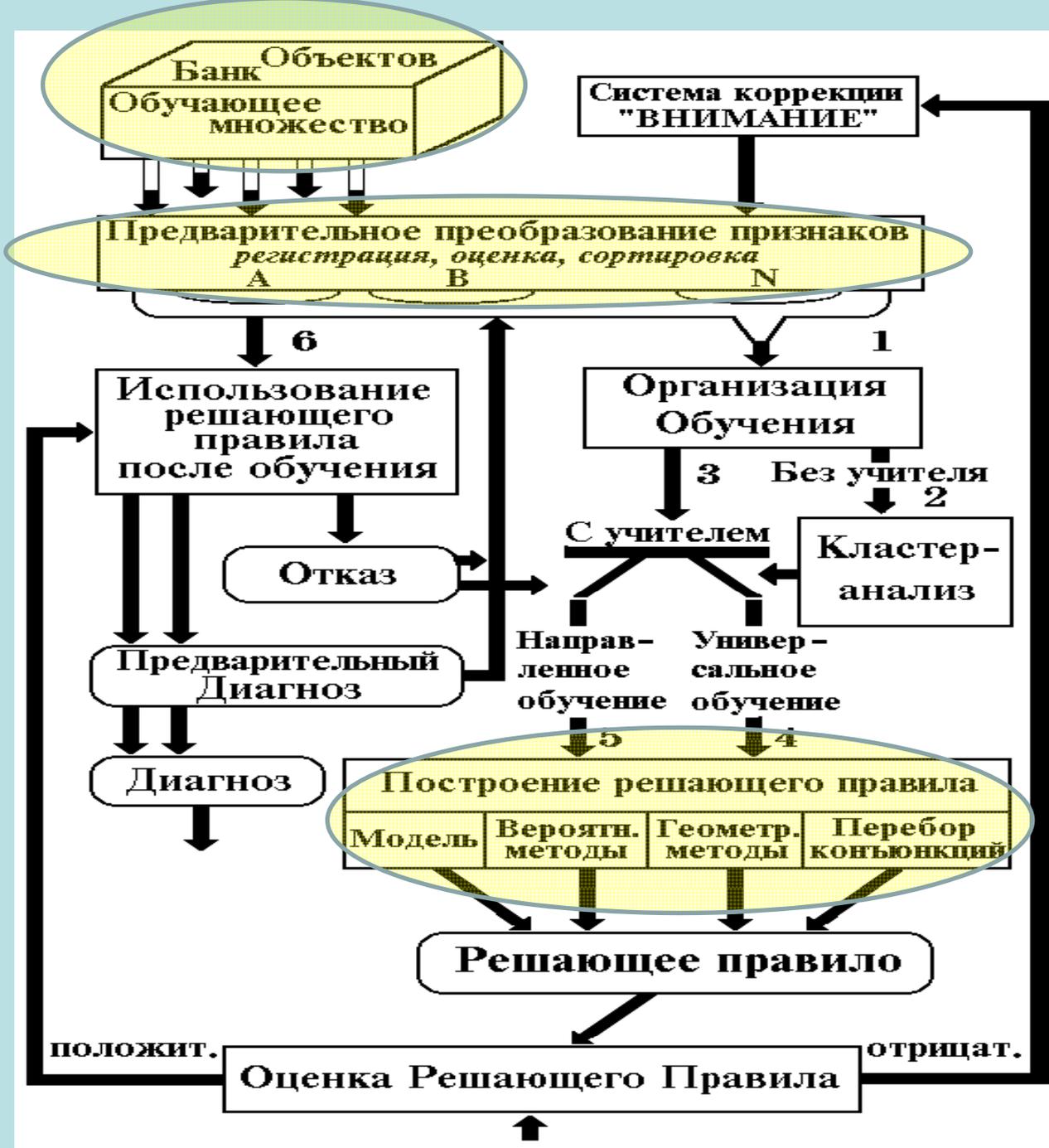
**НЕПРЕМЕННЫЕ ТРЕБОВАНИЯ** к данной группе аналитических разработок:

- **«прозрачность»**
- **«открытость»,**

которые позволяют эксперту–исследователю **контролировать процесс формирования знаний,**

**а именно:**

- иметь возможность получать промежуточные результаты,
- менять параметры анализа,
- наращивать аналитические модули,
- получать конечные результаты в виде, удобном для интерпретации, и т.п.



# Базовые определения признаков

**Признак** — свойство объектов, которое обуславливает различие или общность объекта с другими объектами. Признаки отражают качественные и/или количественные характеристики объектов. Формально признак – это переменная, область значений которой может быть как непрерывной, так и дискретной, в частности, нечисловой. Соответственно, признаки называются *непрерывными* или *дискретными*.

- **Непрерывные признаки** задаются диапазоном возможных значений.
- **Дискретные признаки** — задаются перечислением ограниченного числа возможных значений (в виде оценок, символов, слов).
- **Производные характеристики** — **признаки** объектов, сконструированные на множестве заданных исходных данных с использованием логико-арифметических функций.

# Конструирование производных характеристик

Пусть имеем два класса объектов - класс  $A$  и класс  $B$ . Каждый объект описан набором из  $k$  показателей; каждый показатель  $P_j$  задан векторами  $(x_1, \dots, x_k)$  длины  $k$  (по числу измерений), соответствующих моментам времени  $(t_1, \dots, t_k)$ .

Для каждого показателя  $P_j$  производится преобразование исходного  $k$ -мерного вектора измеренных данных в  $m$ -мерное пространство функций  $f_1, \dots, f_m$ , сконструированных из различных логических и арифметических функций. Преобразование имеет вид:

$$(x_1, x_2, \dots, x_k) \rightarrow (f_1(x_1, x_2, \dots, x_k), f_2(x_1, x_2, \dots, x_k), \dots, f_m(x_1, x_2, \dots, x_k)),$$

причем для каждого показателя может быть выбран свой набор преобразующих функций и свое значение  $m$ .

Таковыми функциями, например, могут являться:

$f_1 = \text{MAX}(x_1, x_2, \dots, x_k)$ ;  $f_2 = \text{MIN}(x_1, x_2, \dots, x_k)$ ; «размах»  $f_z = \text{MAX}(x_i) - \text{MIN}(x_i)$ ;  
«временные области монотонности» показателя  $P_j$ ;

параметры спектральных составляющих модели гармонических (квазигармонических) колебаний или характеристики периодических релаксационных колебаний и т.д.

# Схема перебора конъюнкций симптомов



## CONSILIUM V1.08

Протокол голосования по контрольным объектам

### РЕШАЕМАЯ ЗАДАЧА:

название задачи: ДВС-829 ДИАГНОСТИКА

НОВОРОЖДЕННЫХ

имя карты: ДВС-829

число признаков карты: 54

число симптомов: 383

число больных (всего): 829

### ПАРАМЕТРЫ АЛГОРИТМА:

малый параметр: 0.01

глубина перебора: 3

эталон достоверности в классе А: 5

эталон достоверности в классе В: 8

эталон "серости" (серые симптомы перебираются):  
0.700

ограничение на количество конъюнкций каждого  
ведущего симптома: 3

порог веса конъюнкций за класс А: 1.200

порог веса конъюнкций за класс В: 2.000

### КОЛИЧЕСТВО ОБЪЕКТОВ НА ОБУЧЕНИИ:

за класс А - 165 : Класс А (165)

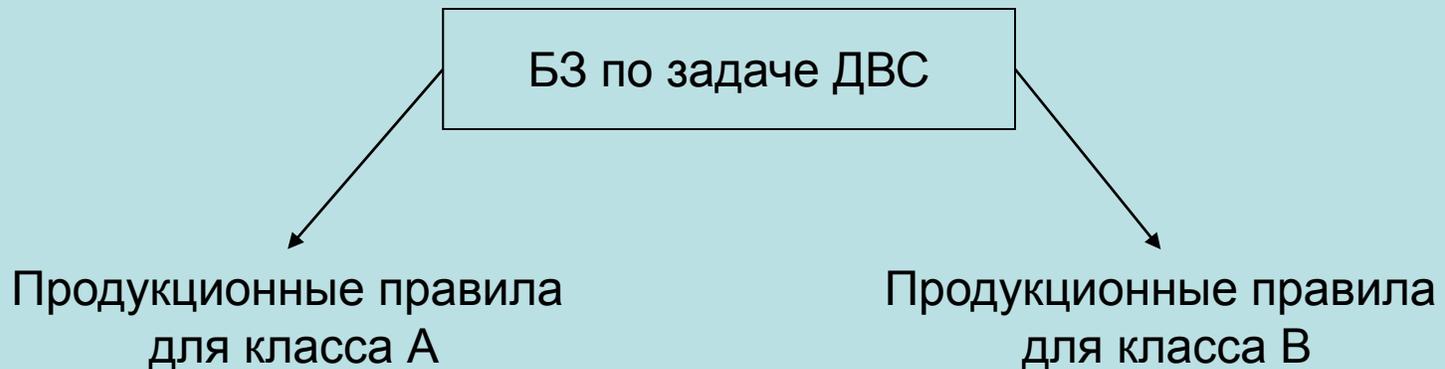
за класс В - 150 : Класс В (150)

### КОЛИЧЕСТВО ОБЪЕКТОВ НА КОНТРОЛЕ:

за класс А - 230

за класс В - 220

# Фрагменты компьютерной диагностической базы знаний (в виде продукционной модели)



№п/п	a	b	c	$W_i$
1	55	59	289	3,84
4	53	59	289	3,32
7	245	77	55	3,25
13	126	126	59	3,01
25	94	77	55	2,47
37	376	376	59	2,25

...

№п/п	a	b	c	$W_i$
1	290	290	125	-3,93
4	293	125	37	-3,78
12	76	125	27	-3,57
14	125	183	37	-3,45
30	86	206	27	-3,14

...

# Протокол голосования по контрольным объектам (%)

Перевес голосов	В процентах				
	объекты контроля			результат узнавания	
	из А	из В	всего	прав.	ошиб.
> 15	6,67	0	6,67	6,67	0
К > 10	5,78	0,89	6,67	5,78	0,89
Л > 5	15,3	0,67	16	15,3	0,67
А = 5	2,22	0,22	2,44	2,22	0,22
С = 4	3,33	0	3,33	3,33	0
С = 3	5,56	0	5,56	5,56	0
= 2	6	1,11	7,11	6	1,11
А = 1	3,11	0,22	3,33	3,11	0,22
= 0	1,33	0,44	1,78	0	0
К = -1	0,44	0,67	1,11	0,67	0,44
Л = -2	0,44	0,44	0,89	0,44	0,44
А = -3	0	0,89	0,89	0,89	0
С = -4	0,22	0,67	0,89	0,67	0,22
С = -5	0,22	0,67	0,89	0,67	0,22
< -5	0,44	3,33	3,78	3,33	0,44
В < -10	0	2,89	2,89	2,89	0
< -15	0	35,8	35,8	35,8	0
ИТОГО:	51,1	48,9	100	93,3	4,89

# Примеры диагностических продукционных правил в Базе Знаний за класс А

№п/п	a	b	c	$W_i$	Семантическое значение правила
1	55	59	289	3,84	Отсутствие <b>дородового</b> лечения у матери, что имеет связь с положительными нетрепонемными тестами у новорожденного.
4	53	59	289	3,32	Отсутствие <b>дородового</b> лечения у матери, что имеет связь с положительными нетрепонемными тестами у новорожденного.
7	245	77	55	3,25	Отсутствие <u>специфического</u> лечения <b>до беременности</b> и <u>профилактического</u> <b>во время беременности</b> , что имеет связь с индексом нетрепонемных тестов к родам $> 5$ .
13	126	126	59	3,01	Отсутствие специфического лечения <b>во время беременности</b> , что сопровождалось клиническими проявлениями заболевания у новорожденного.
25	94	77	55	2,47	Отсутствие <u>специфического</u> лечения <b>до беременности</b> и <u>профилактического</u> <b>во время беременности</b> , что имеет связь с родоразрешением на сроках 31 по 35 недель беременности.
37	376	376	59	2,25	Отсутствие <u>специфического</u> лечения <b>во время беременности</b> , что имеет связь с отношением показателей индекса нетрепонемных тестов у новорожденного к показателям у матери в пределах значений: 2-20/30-320; 30-320/0-1; 30-320/2-20; 30-320/30-320.



# Примеры диагностических продукционных правил в Базе Знаний за класс В

№п/п	a	b	c	$W_i$	Семантическое значение правила
23	125	60	27	3,27	<u>Адекватное</u> лечение <b>во время беременности</b> (диагноз подтвержден положительными трепонемными тестами), что имеет связь с отсутствием клинических проявлений заболевания у новорожденного
31	51	125	96	3,13	<u>Специфическое</u> лечение матери, проведенное <b>до беременности</b> , при котором родоразрешение наступило на сроках > 37 недели беременности, что имеет связь с отсутствием клинических проявлений заболевания у новорожденного.
36	13	51	125	3,11	Отсутствие клинических проявлений заболевания у новорожденного, что имеет связь с ранней постановкой на дородовой учет (на сроках до 12 нед. беременности) и <u>специфическим</u> лечением проведенным <b>до беременности</b> .
40	51	125	183	3,05	<u>Специфическое</u> лечение матери, проведенное <b>до беременности</b> , при котором во втором триместре беременности наблюдались положительные нетрепонемные тесты, что сопровождалось отсутствием клинических проявлений заболевания у новорожденного.
47	51	96	183	2,98	<u>Специфическое</u> лечение матери, проведенное <b>до беременности</b> , при котором во втором триместре беременности наблюдались положительные нетрепонемные тесты, что сопровождалось родоразрешением на сроках > 37 нед. беременности.

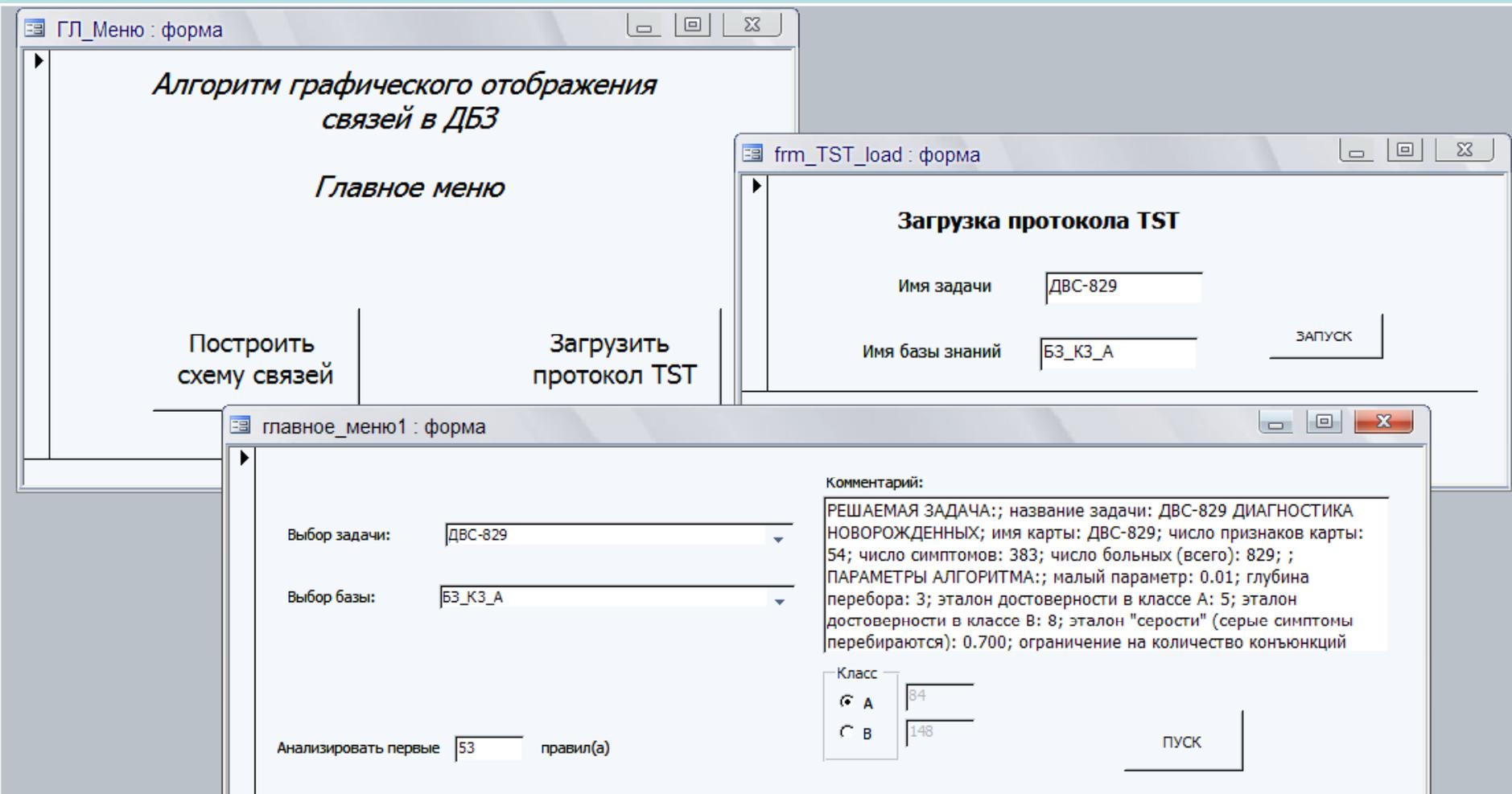
# Распределение объектов по результатам компьютерной и врачебной диагностики

	Исход беременности										
	Диагноз ВС		Мертворождение		Прерывание на поздних сроках		Перевед. из гр. В	Здоровые	Всего		
	Активный	Скрытый	Сиф.	Не сиф.	Самопроизв	Мед. аборт			из А	из В	
>15	24	3			1		2		30	0	30
>5	8	17	18		13	14	6		78	0	78
5	2			1		3			6	0	6
4		2			3	1			6	0	6
3	2	3		1	4	3			13	0	13
2	3	2		2	2	4	1		14	0	14
1				4	2	1			7	0	7
0 = 0	1	1		4				2	6	2	8
-1				2					2	0	2
-2		1						1	1	1	2
-3									0	0	0
-4								3	0	3	3
-5								1	0	1	1
<-5		2						26	2	26	28
<-15								117	0	117	117
Всего	40	31	18	14	27	26	9	150	165	150	315

# Идея построения кластеров связей в базе знаний



# Интерфейс пользователя программы



# Кластеры связей симптомов класса А

(10-0)	197	61				
(14-0)	191	99				
(16-0)	32	268		376	376	(14-0)
(17-0)	197	77		126	77	(30-0)
(18-0)	77	94		126	126	(32-0)
		<b>55</b>		<b>59</b>		
		53 (44-0)		53 (44-0)		
		268 (16-0)	<b>289 (47-0)</b>	48 (41-0)		
				243 (10-0)		
		77 (41-0)	<b>245 (41-0)</b>			
			<b>297 (30-0)</b>	376 (14-0)		
			<b>237 (59-2)</b>	4 (13-0)		
			<b>5 (18-0)</b>			
			<b>27 (77-5)</b>	305 (11-0)		
			<b>4 (16-0)</b>			
			<b>305 (13-0)</b>	37 (12-0)		
			<b>303 (13-0)</b>			
			<b>371 (12-0)</b>			

<b>55</b>		<b>59</b>
53 (44-0)	<b>289 (47-0)</b>	53 (44-0)
	245 (41-0)	
	297 (30-0)	
	237 (59-2)	
	5 (18-0)	
	27 (77-5)	
	4 (16-0)	
	305 (13-0)	
	303 (13-0)	
	371 (12-0)	

### При полном отсутствии лечения у матерей до родов (53):

✓ у них к родам сохраняются высокие титры нетрепонемных тестов (н/т 4+), что объясняет высокие значения индексов н/т > 5 у новорожденных (237, 245, 289, 297);  
 ✓ особенно ярко это проявляется у молодых матерей (до 21 года) (4,5), причем для новорожденных характерно – высокие индексы н/т > 5 с уровнем их значений равными или даже выше материнских (303,305,371).

### **До беременности (55):**

Отсутствие специфического лечения у матерей, у которых индекс значений нетрепонемных тестов в околородовый период >5 (268), влечет за собой положительные нетрепонемные тесты у новорожденных (289) .

(10-0)	197	61				
(14-0)	191	99				
(16-0)	32	268	376	376	(14-0)	
(17-0)	197	77	126	77	(30-0)	
(18-0)	77	94	126	126	(32-0)	
		<b>55</b>	<b>59</b>			
		268 (16-0)	289	48 (41-0)		
				243 (10-0)		
		77 (41-0)	245			
			297	376 (14-0)		
			237	4 (13-0)		
			5			
			27	305 (11-0)		
			4			
			305	37 (12-0)		
			303			
			371			

### **Во время беременности (59):**

Отсутствие у матери специфического и профилактического лечения (77) может привести к рождению детей с результатами индекса значений н/т выше, чем у матери (376), при явно выраженных клинических проявлениях у новорожденного (126).

# Кластеры связей симптомов класса В

(0-32)	189	189				
(0-32)	189	194				
(2-74)	96	51				
(1-59)	27	56				
(0-38)	27	60				
(0-39)	194	194				
(1-84)	27	76				
(0-77)	290	290				
(0-77)	293	290		76	56	(0-45)
		<b>125</b>		<b>183</b>		
		56 (0-51)	23 (0-63)	96 (0-60)		
		293 (0-64)	<b>37 (1-74)</b>			
		368 (0-62)				
		191 (0-44)				
		160 (0-41)				
		56 (4-32)	<b>13 (0-42)</b>	96 (0-42)		
		51 (0-32)		206 (0-32)		
		27 (2-77)	<b>42 (1-63)</b>			
			<b>217 (0-33)</b>	206 (0-33)		

Адекватная пенициллиновая терапия у матерей до беременности (**51,56,60**) и профилактическое лечение во время беременности (**76**) влекут за собой низкую степень позитивности н/т (2+,3+) во втором триместре беременности (**189, 194**). При данных условиях высоко вероятен благополучный исход беременности на сроках от 37 недель (**96**) с отрицательными н/т у новорожденных (**290,293**).

<b>125</b>		<b>183</b>
	<b>23 (0-63)</b>	
	<b>37 (1-74)</b>	
	<b>13 (0-42)</b>	
	<b>42 (1-63)</b>	
	<b>217 (0-33)</b>	

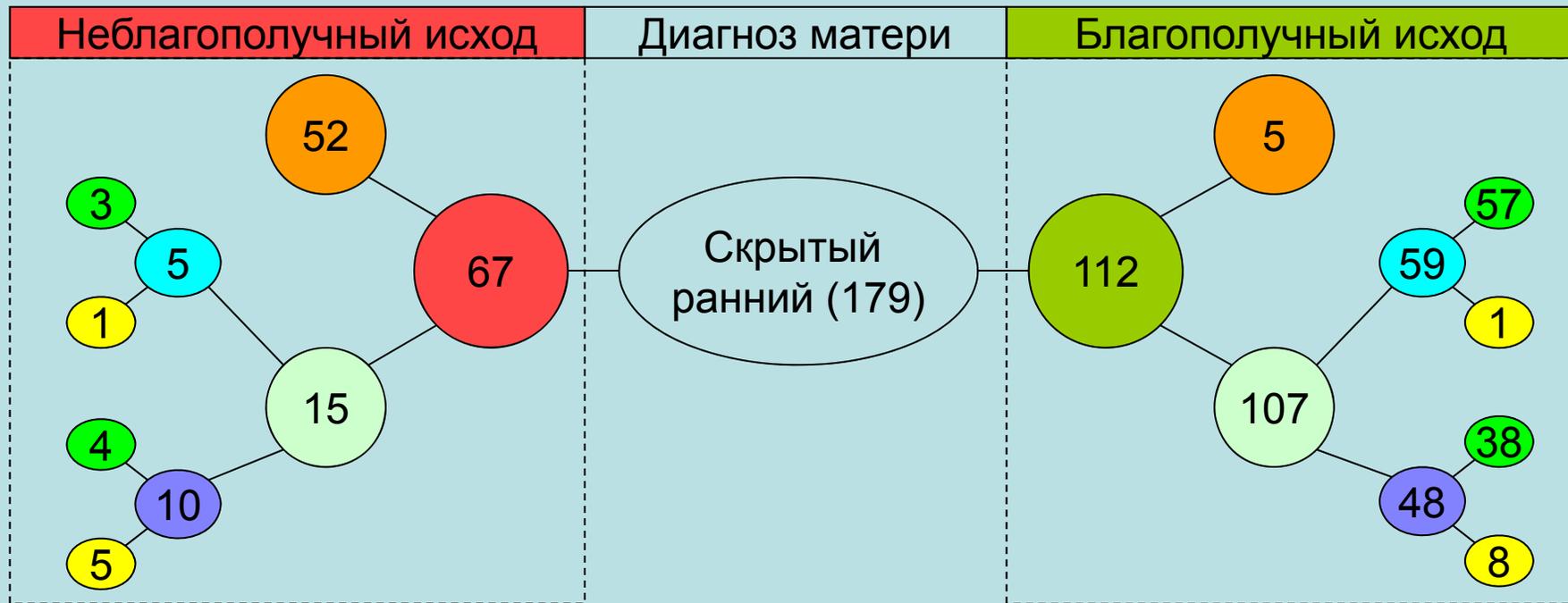
*Адекватное лечение матери до беременности (56) и частое посещение женской консультации (>4 посещений)(23) обеспечивает рождение ребенка без клинических проявлений заболевания (125).*

(0-32)	189	189					
(0-32)	189	194					
(2-74)	96	51					
(1-59)	27	56					
(0-38)	27	60					
(0-39)	194	194					
(1-84)	27	76					
(0-77)	290	290					
(0-77)	293	290		76	56	(0-45)	
		<b>125</b>		<b>183</b>			
		56 (0-51)	<b>23</b>	96 (0-60)			
		293 (0-64)	<b>37</b>				
		368 (0-62)					
		191 (0-44)					
		160 (0-41)					
		56 (4-32)	<b>13</b>	96 (0-42)			
		51 (0-32)			206 (0-32)		
		27 (2-77)	<b>42</b>				
			<b>217</b>	206 (0-33)			

# Распределение объектов по характеристике «лечение»

	Диагноз ВС		Мертворождение		Прерывание на поздних сроках				
	активный	скрытый	сифил.	несифил.	самопроизв	медаборт	перевед	здоровые	всего
<b>Лечение до беременности</b>									
Адекватное + профлечение		1		1				60	62
Адекватное без профлечения	2	2		3	3	4		21	35
Неадекватное без профлечения								1	1
<b>Лечение во время беременности</b>									
Адекватное + профлечение	3	3		4		1		41	52
Адекватное без профлечения		2			4	3		2	11
Неадекватное + профлечение	1							3	4
Неадекватное без профлечения	5	2			8	7		6	28
<b>Отсутствие дородового леч.</b>	29	19	18	5	2	10	9	9	101
<b>неизвестно</b>	0	2	0	1	10	1	0	7	21
<b>всего</b>	40	29	18	13	17	25	9	150	315

# Исход беременности в зависимости от вариантов лечения матери



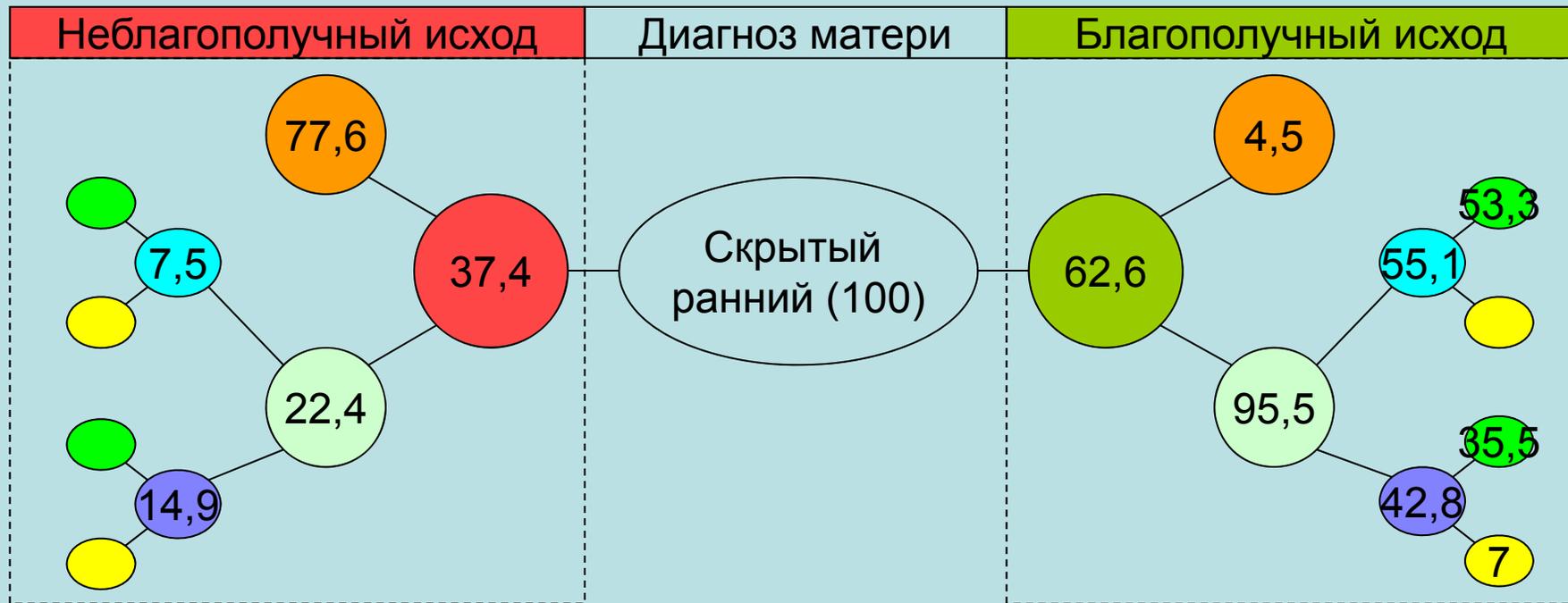
*Лечение матери*

- – не проводилось
- – проводилось

*Период проведения лечения*

- – до беременности
- – во время беременности
- – неадекватное лечение
- – адекватное лечение

# Исход беременности (%) в зависимости от вариантов лечения матери



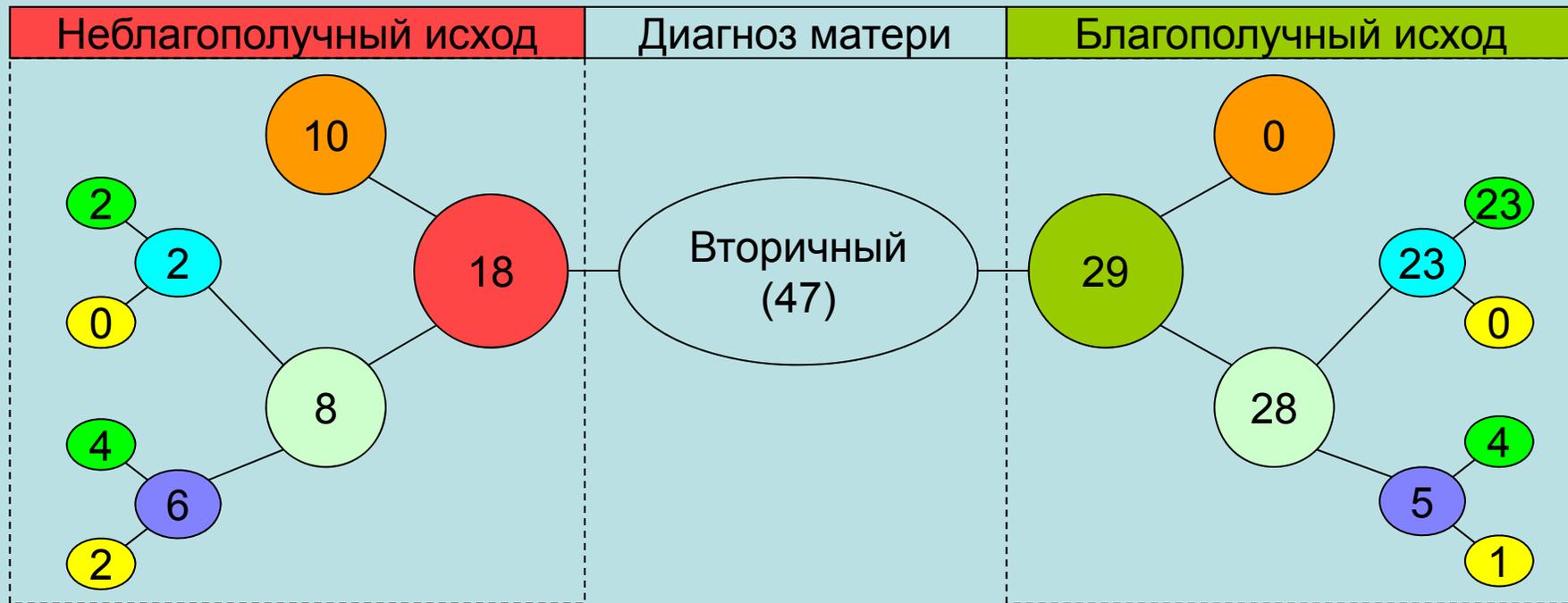
*Лечение матери*

- – не проводилось
- – проводилось

*Период проведения лечения*

- – до беременности
- – во время беременности
- – неадекватное лечение
- – адекватное лечение

# Исход беременности в зависимости от вариантов лечения матери



## Лечение матери

- – не проводилось
- – проводилось

## Период проведения лечения

- – до беременности
- – во время беременности
- – неадекватное лечение
- – адекватное лечение

# Рекомендации по эффективному ведению беременных

Постановка на дородовый учет на ранних сроках беременности  
и частое посещение женской консультации

Выявление диагноза у матери на сроках < 12 нед. беременности  
и своевременное начало лечения

Выбор адекватного лечения беременной.  
Положительные нетрепонемные тесты во 2 триместре беременности  
при их низких значениях в третьем

***Благополучный исход  
беременности***

- Разработанные алгоритмы позволяют по-новому взглянуть на знания предметной области, представленные в диагностических правилах.
- Практическое применение разработанных алгоритмов для реальных данных конкретной задачи показало его эффективность в решении задачи оптимизации БЗ.
- Это позволяет рассматривать данный алгоритм как один из способов повышения качества Базы Знаний по конкретной проблеме.

# Технология формирования новых

знаний



Обсуждение конкретной проблемы

Эксперт

Когнитолог  
- аналитик

Декларативные знания  
Интуитивные знания  
Новые знания

Семантическая интерпретация ДБЗ (новые знания)

- Задачи:
- ✓ Диагностика;
  - ✓ Прогнозирование;
  - ✓ Принятие управленческих решений;
  - ✓ Выбор лечебного воздействия;
  - ✓ ...

База формализованных знаний

Алгоритм графического отображения связей в ДБЗ

Диагностическая база знаний (ДБЗ)

База данных  
Data Mining  
«Consilium»

АРМ    ЭС    ...    СППР



Пользователь

СПАСИБО  
за внимание