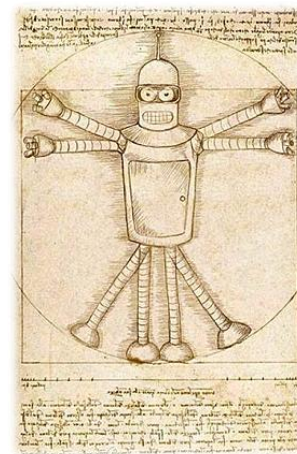


Нейроэволюция VS Случайность



Семинар по вычислительному интеллекту
Докладчик: Юрий Цой
г. Томск, 5 мая 2012 г.

<http://gai.narod.ru/TomskWorkshop/index.html>

Пролог

“Главная” цель:

Создание алгоритма эволюции нейронных сетей, который обучается на своем «опыте».

Подцель 0-го порядка:

Определить возможности простейшего НЭ алгоритма, работающего абсолютно случайно.

Введение

В большинстве (а то и во всех) НЭ алгоритмах используются эвристики

- Ограничения на структуру
- Выбор операторов
- Применение операторов (напр., удаление не случайных нейронов или связей, а по некоторому правилу)

Алгоритм

Нет эвристик и иных «подтяжек» для выбора операторов

Нет скрещивания

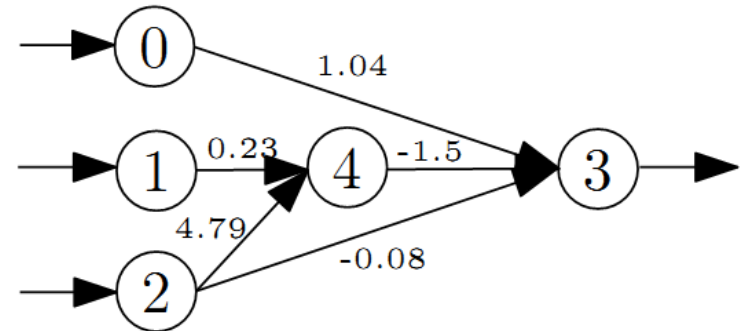
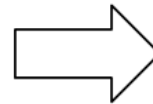
Есть мутации:

- +Node (“чистый” вариант)
- -Node
- +Connection
- -Connection
- Activation
- Weight change

Кодирование

gene1	gene2	gene3	gene4	gene5
1	4	4	3	2
3	2	3	2	4
0	3	2	4	0

Beginning node	Ending node
Connection weight	



+ Массив функций активаций для нейронов

Activation	Formula
Identity	$f(x) = x$
Linear	$f(\mathbf{x}) = S$
Sigmoid	$f(\mathbf{x}) = \frac{1}{1 + \exp(-aS)}$
Gaussian	$f(\mathbf{x}) = \exp\left(-\frac{\ \mathbf{x} - \mathbf{w}\ }{2a}\right)$

$$S = \mathbf{w}^T \mathbf{x}$$

Собираемая информация

Изменения особи:

- Название оператора («ДобавлениеСвязи», «МутацияВеса» и т.д.)
- Изменение приспособленности
- Поколение

В каждом поколении отслеживается среднее число узлов и связей.

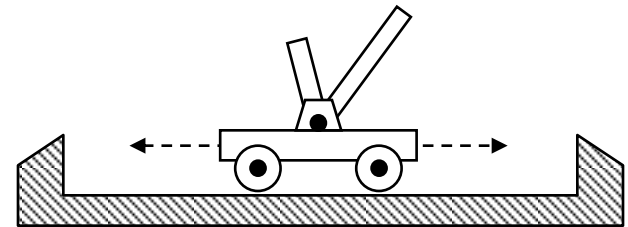
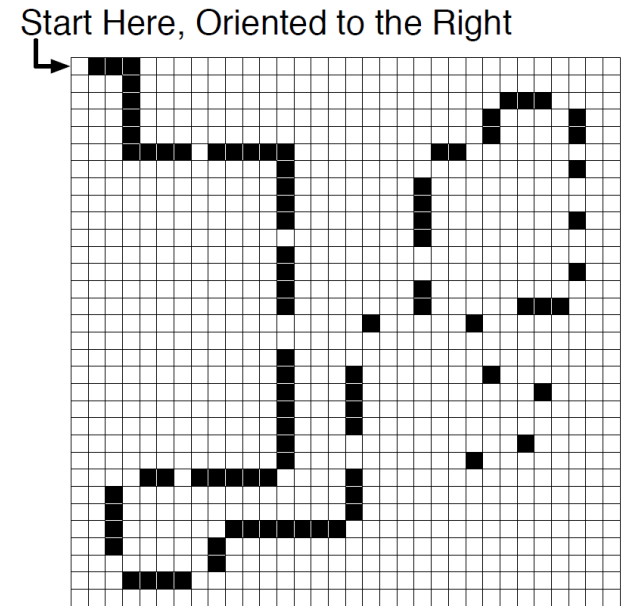
Добавление нейрона – нейтрально!

Анализ операторов

- **“Полезность”** операторов лучшей особи (“полезные” изменения приспособленности)
- **Распределение по поколениям** (как часто использовался оператор на различных этапах)
- **Переходы** (существуют ли шаблонные успешные комбинации операторов)

Тестовые задачи

- XOR
- Задачи из набора Proben1:
 - Cancer1
 - Card1
 - Diabetes1
 - Glass1
 - Heart1
 - Horse1
- Artificial Ant
- 2-Poles Balancing



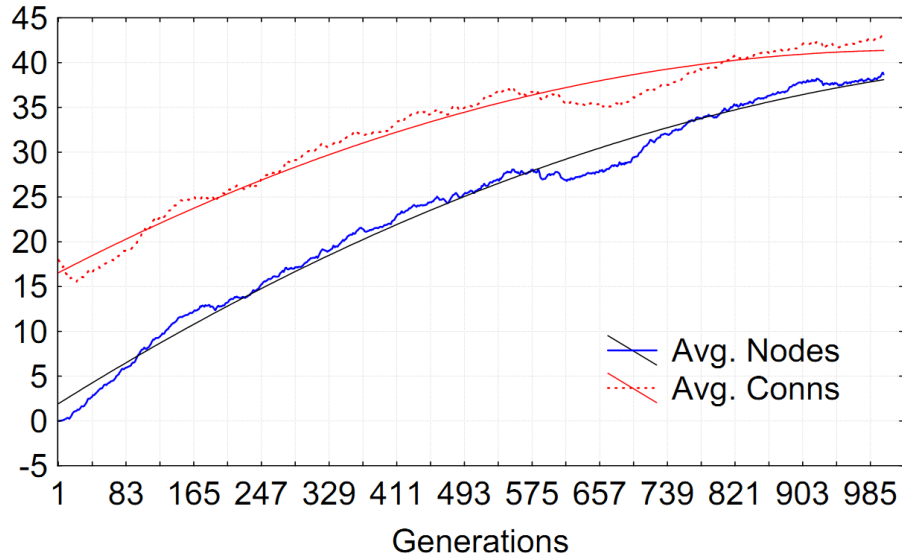
Тестовые задачи

Proben1

Problem's Name	Inputs	Classes	Training/Validation/Test Samples
cancer1	9	2	350/175/174
card1	51	2	345/173/172
diabetes1	8	2	384/192/192
glass1	9	6	107/54/53
heart1	35	2	460/230/230
horse1	58	3	182/91/91

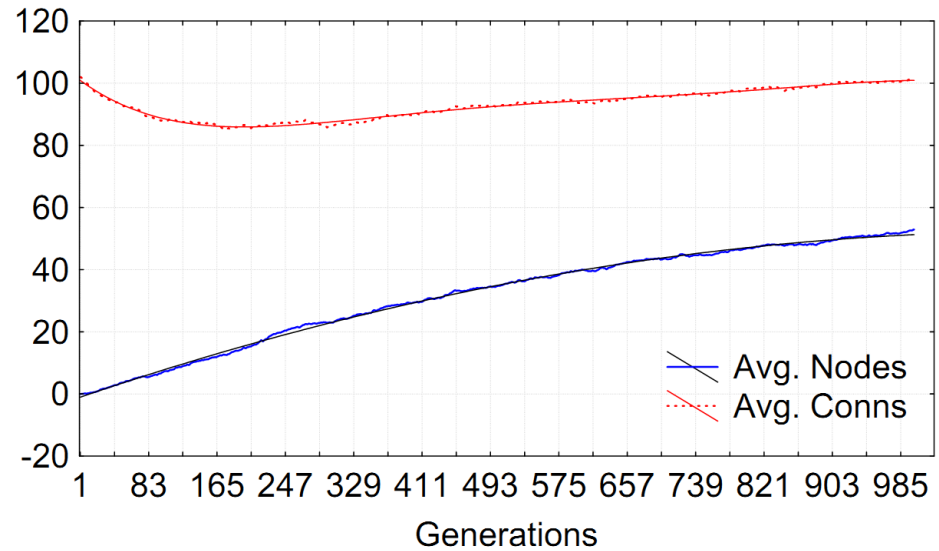
cancer1

Avg. Nodes = $1,878+0,0576*x-2,1353E-5*x^2$
 Avg. Conns = $16,5013+0,0477*x-2,2857E-5*x^2$



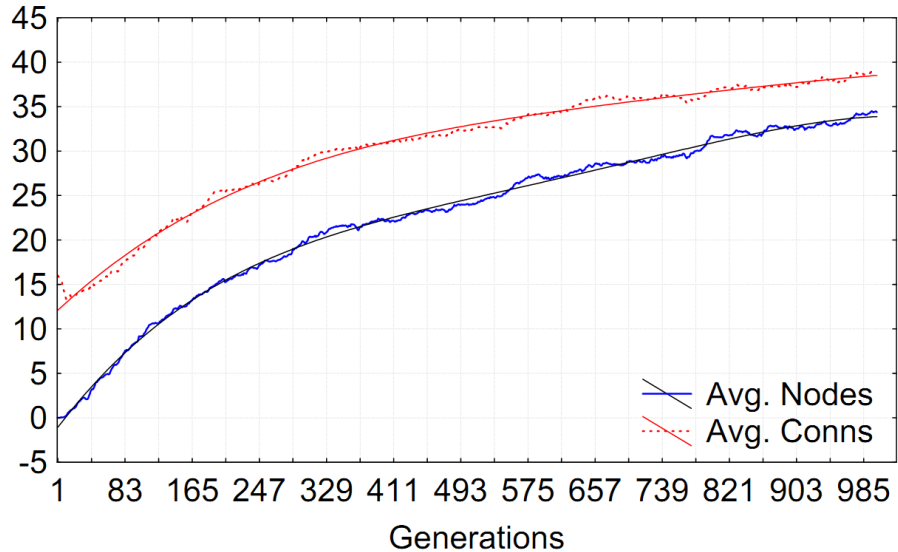
card1

Avg. Nodes = $-1,1234+0,0916*x-3,9248E-5*x^2$
 Avg. Conns = $101,0858-0,1971*x+0,0009*x^2-1,6092E-6*x^3+1,3704E-9*x^4-4,4167E-13*x^5$



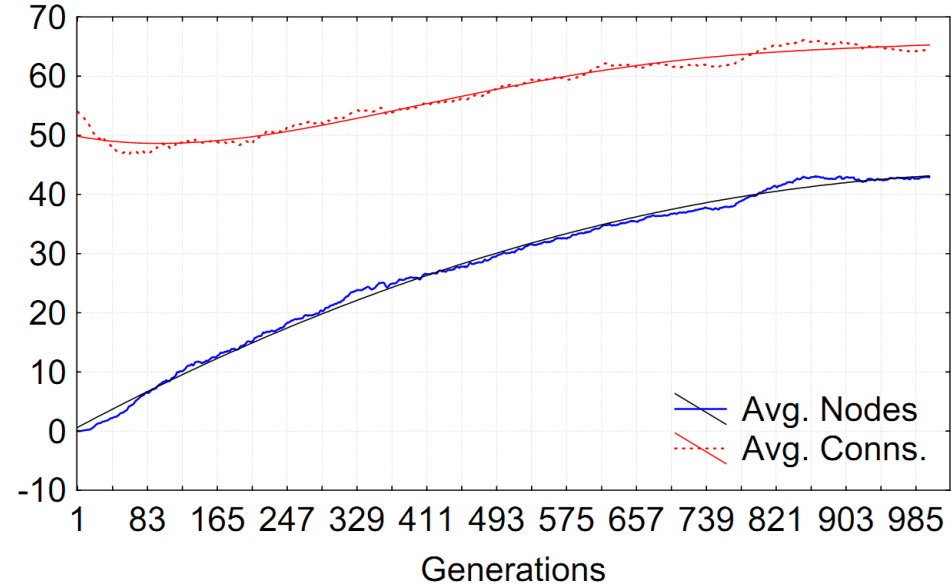
diabetes1

Avg. Nodes = $-1,1732+0,1208*x-0,0002*x^2+2,5794E-7*x^3-1,0139E-10*x^4$
 Avg. Conns = $12,0319+0,0859*x-0,0001*x^2+1,1117E-7*x^3-3,5453E-11*x^4$



glass1

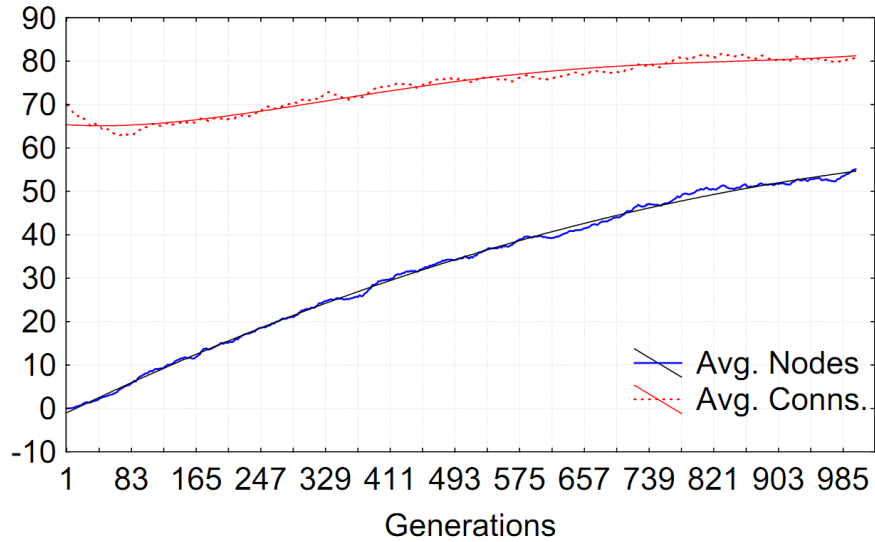
Avg. Nodes = $0,5601+0,0768*x-3,4169E-5*x^2$
 Avg. Conns. = $49,8757-0,0274*x+0,0002*x^2-2,0097E-7*x^3+7,463E-11*x^4$



heart1

$$\text{Avg. Nodes} = -1,0677 + 0,0873 * x - 3,1598E-5 * x^2$$

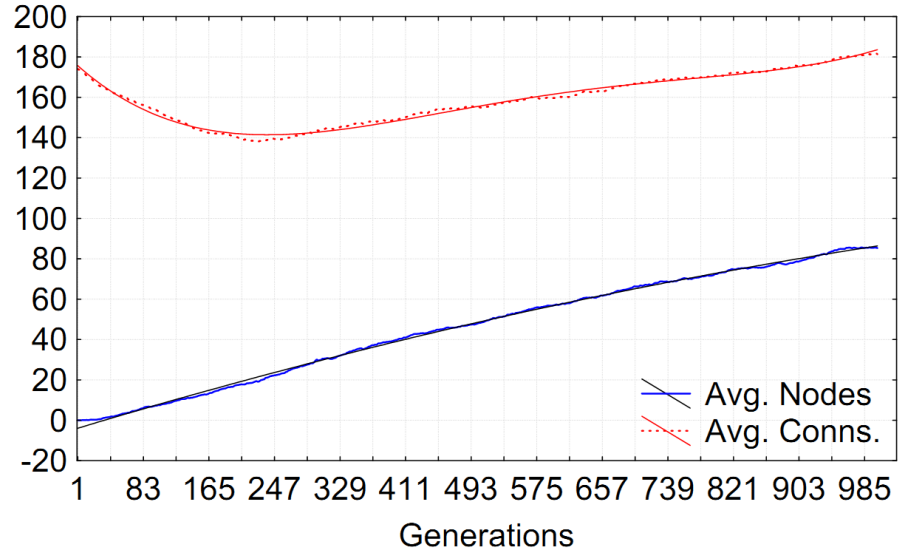
$$\text{Avg. Conns.} = 65,378 - 0,0112 * x + 0,0001 * x^2 - 1,9487E-7 * x^3 + 8,2407E-11 * x^4$$



horse1

$$\text{Avg. Nodes} = -4,0727 + 0,1195 * x - 2,909E-5 * x^2$$

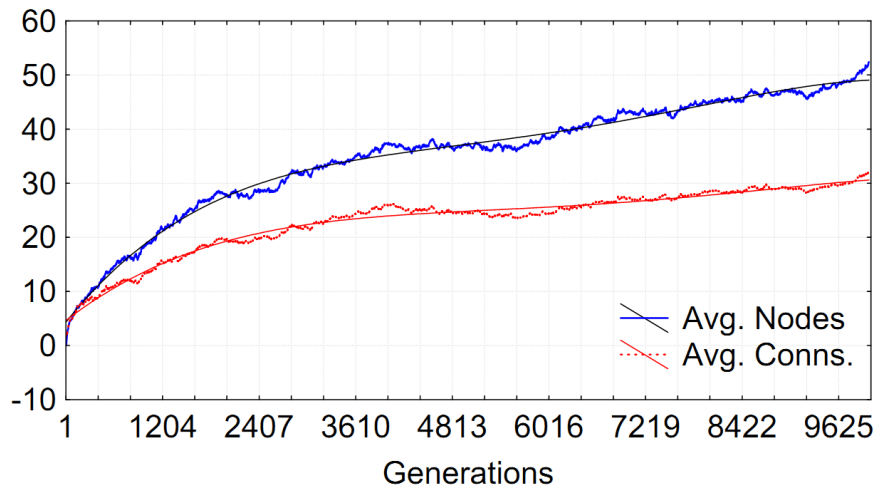
$$\text{Avg. Conns.} = 175,9998 - 0,3522 * x + 0,0012 * x^2 - 1,3519E-6 * x^3 + 5,5142E-10 * x^4$$



XOR

$$\text{Avg. Nodes} = 4,3279 + 0,0184 * x - 4,2758E-6 * x^2 + 4,7365E-10 * x^3 - 1,8569E-14 * x^4$$

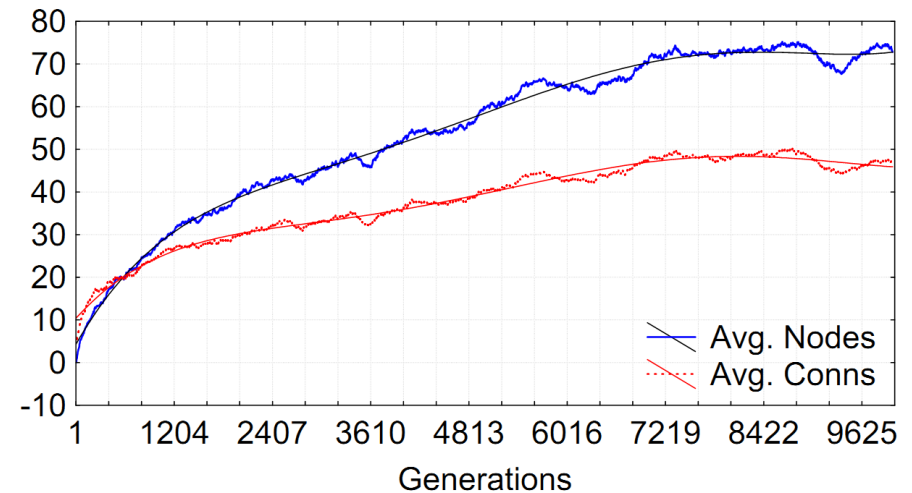
$$\text{Avg. Conns.} = 4,6177 + 0,0114 * x - 2,5266E-6 * x^2 + 2,5456E-10 * x^3 - 9,0336E-15 * x^4$$



2-Poles

$$\text{Avg. Nodes} = 4,2851 + 0,0336 * x - 1,2521E-5 * x^2 + 2,662E-9 * x^3 - 2,5904E-13 * x^4 + 9,1343E-18 * x^5$$

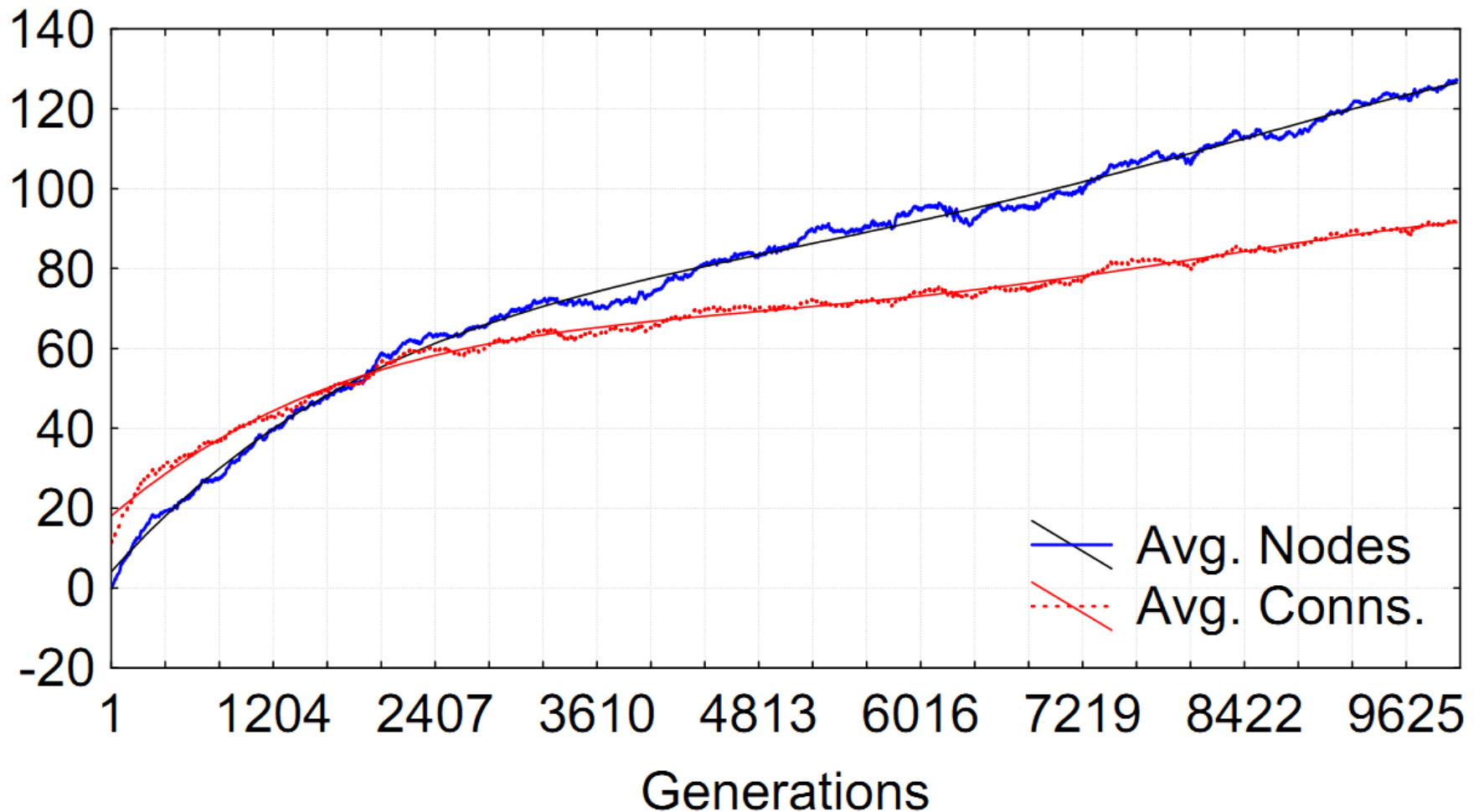
$$\text{Avg. Conns.} = 10,3818 + 0,0212 * x - 8,6806E-6 * x^2 + 1,8542E-9 * x^3 - 1,7623E-13 * x^4 + 5,9992E-18 * x^5$$



Artificial Ant

$$\text{Avg. Nodes} = 4,0982 + 0,0377 * x - 7,5334 \text{E-}6 * x^2 + 7,8887 \text{E-}10 * x^3 - 2,9027 \text{E-}14 * x^4$$

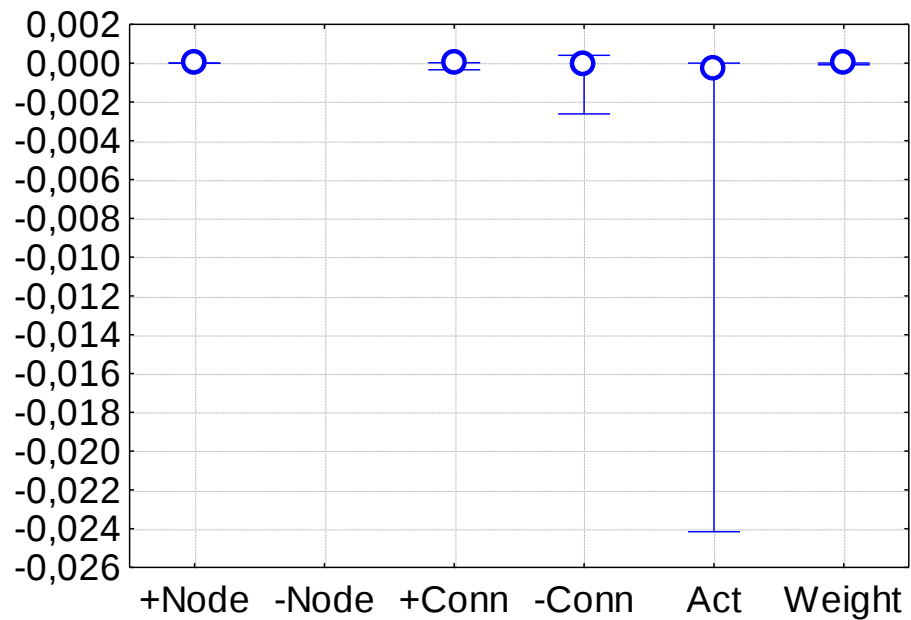
$$\text{Avg. Conns.} = 18,0709 + 0,0288 * x - 6,5764 \text{E-}6 * x^2 + 7,1354 \text{E-}10 * x^3 - 2,702 \text{E-}14 * x^4$$



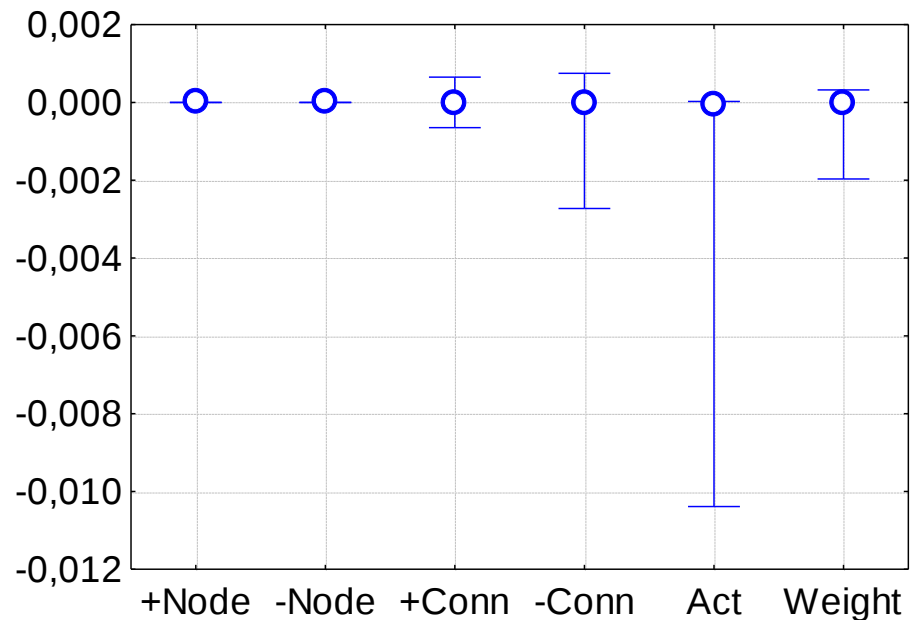
Результаты анализа размеров ИНС

- В большинстве случаев рост числа нейронов почти **квадратичный**.
- Для задач с большим начальным количеством связей, в первых поколениях происходит удаление связей, т.е. алгоритм «пытается» уменьшить размерность задачи, отсекая ненужное.
- **Эвристика:**
Если количество связей в начале велико, то сначала необходимо удалить часть связей. В противном случае нужно добавлять нейроны и связи.

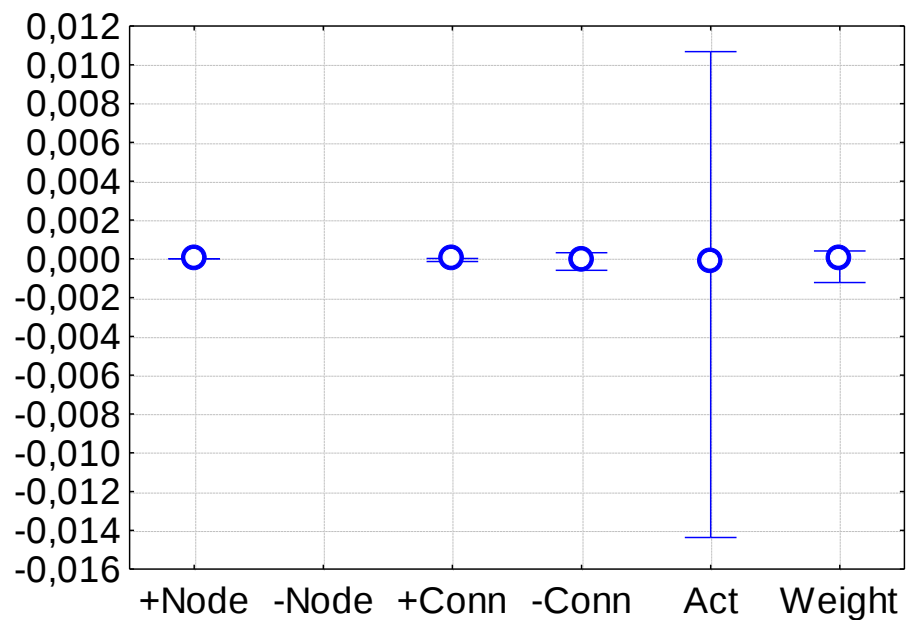
cancer1



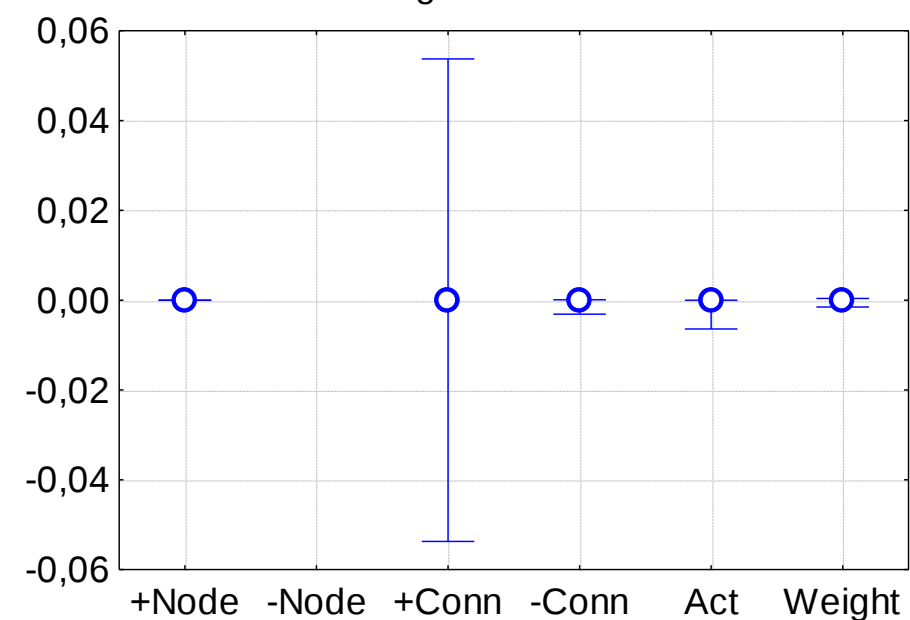
card1



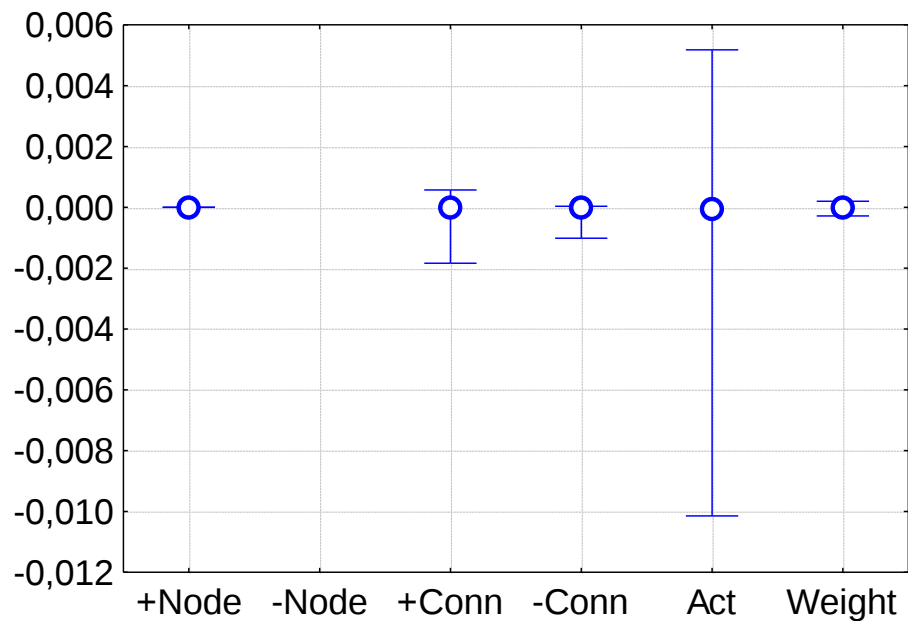
diabetes1



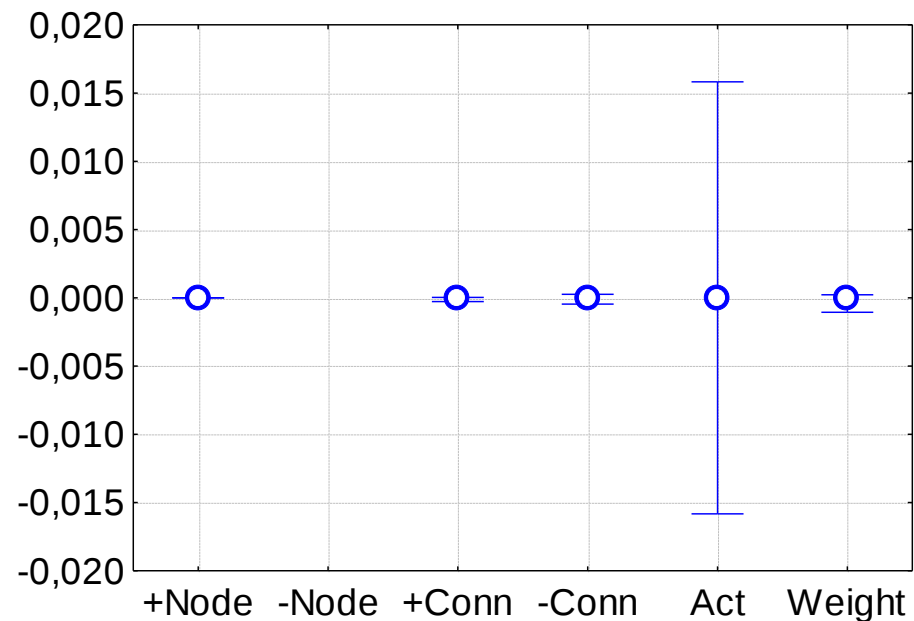
glass1



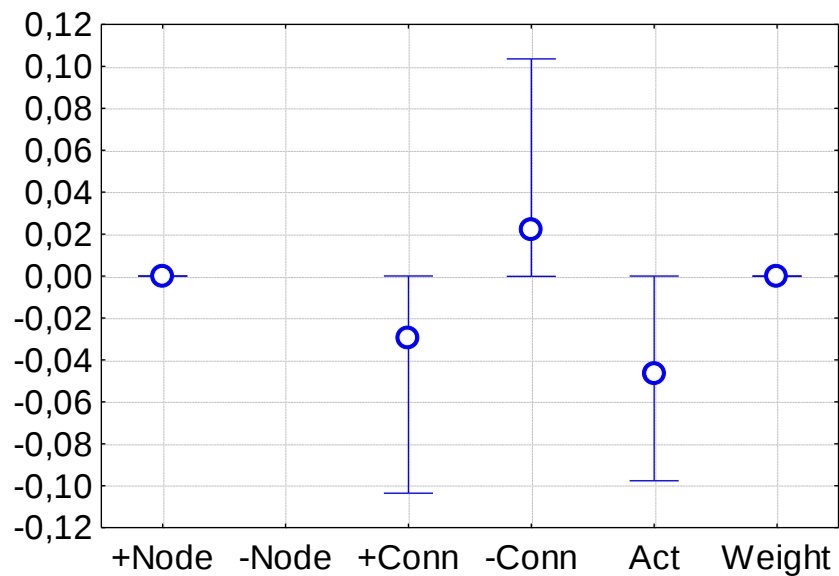
heart1



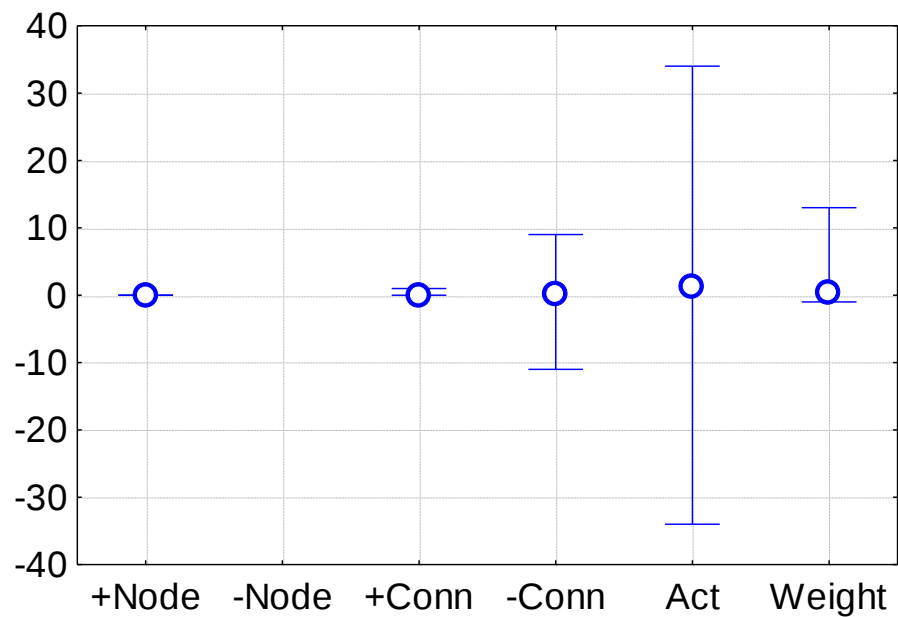
horse1



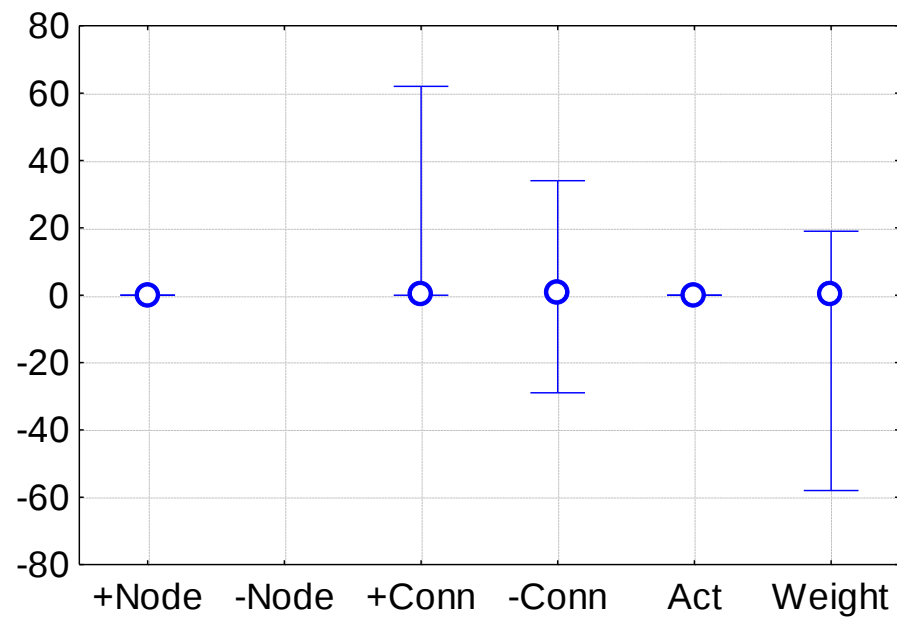
XOR



2-Poles Balancing



Artificial Ant



Результаты анализа полезности операторов

- Информация о полезности операторов может **существенно отличаться** для различных проблем.

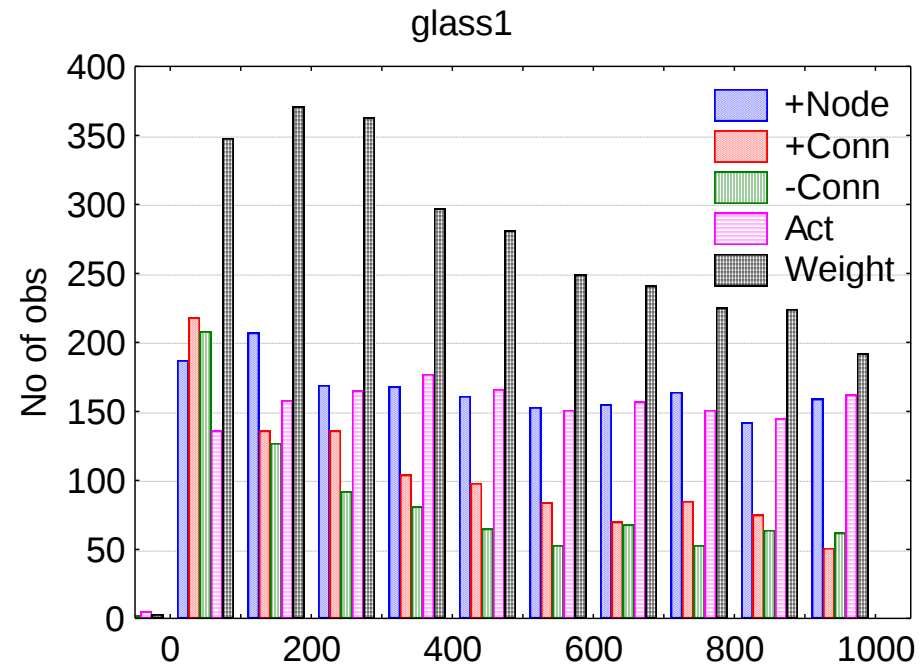
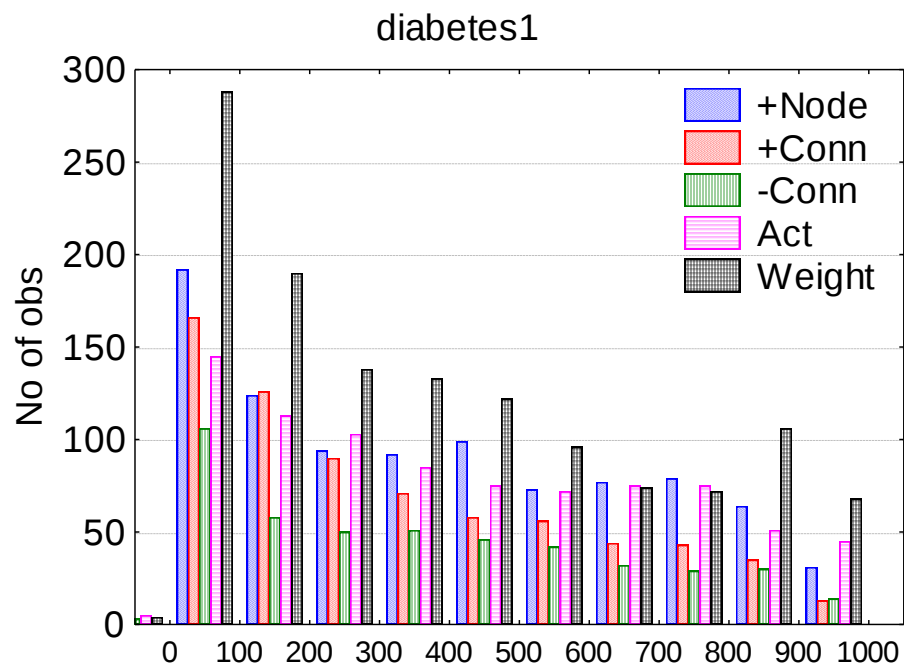
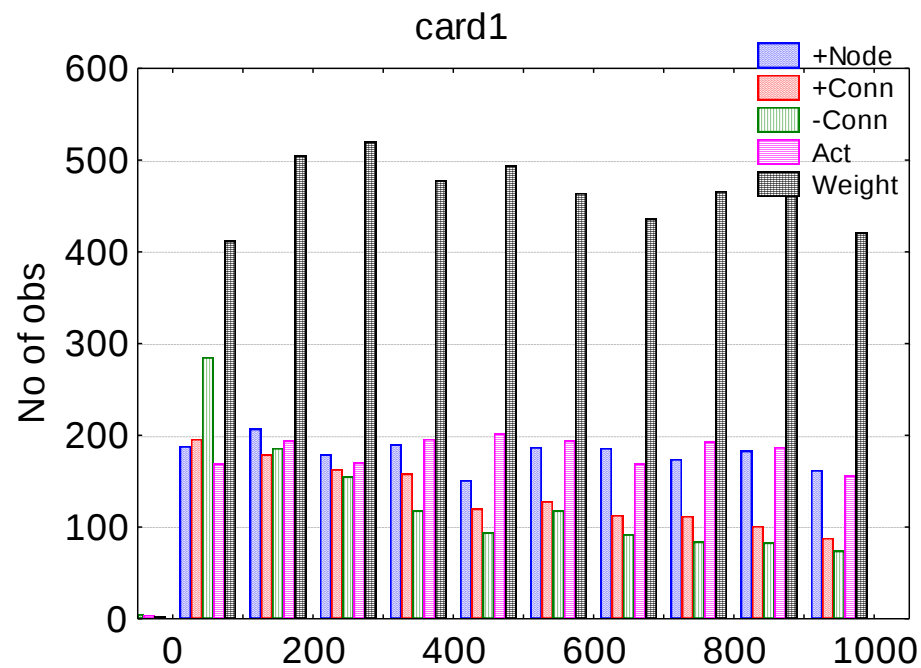
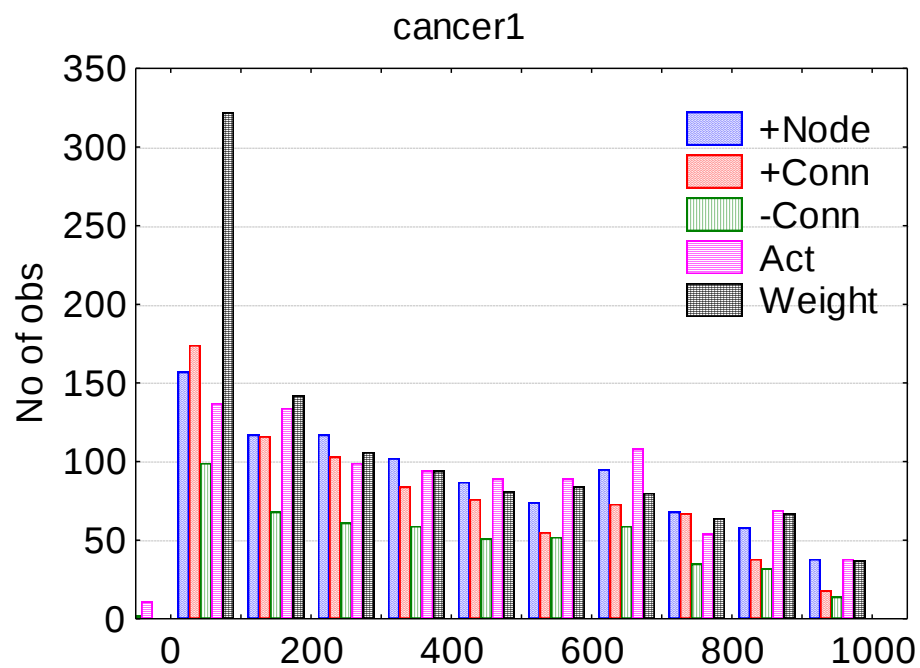
- **Эвристика:**

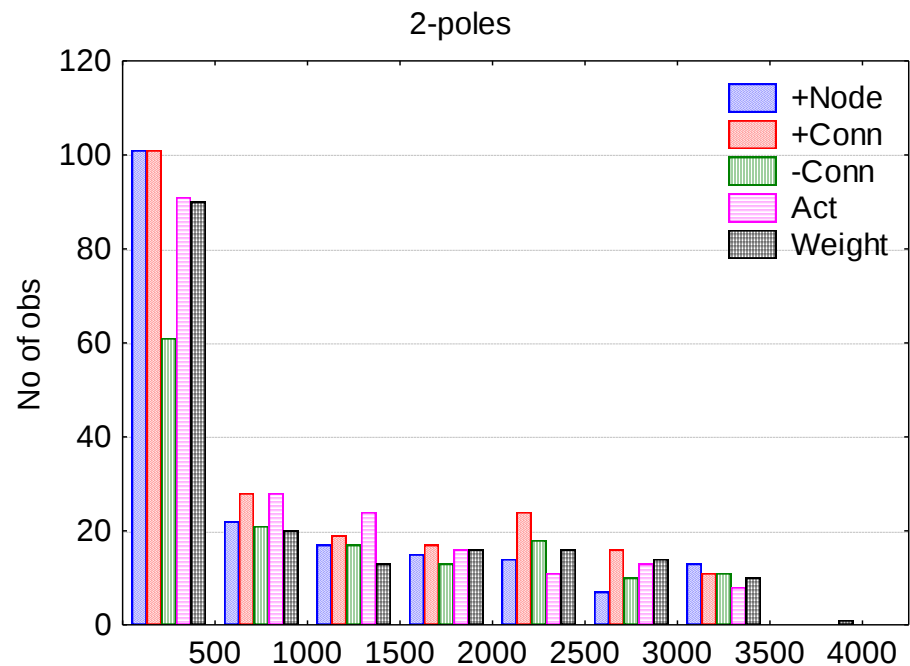
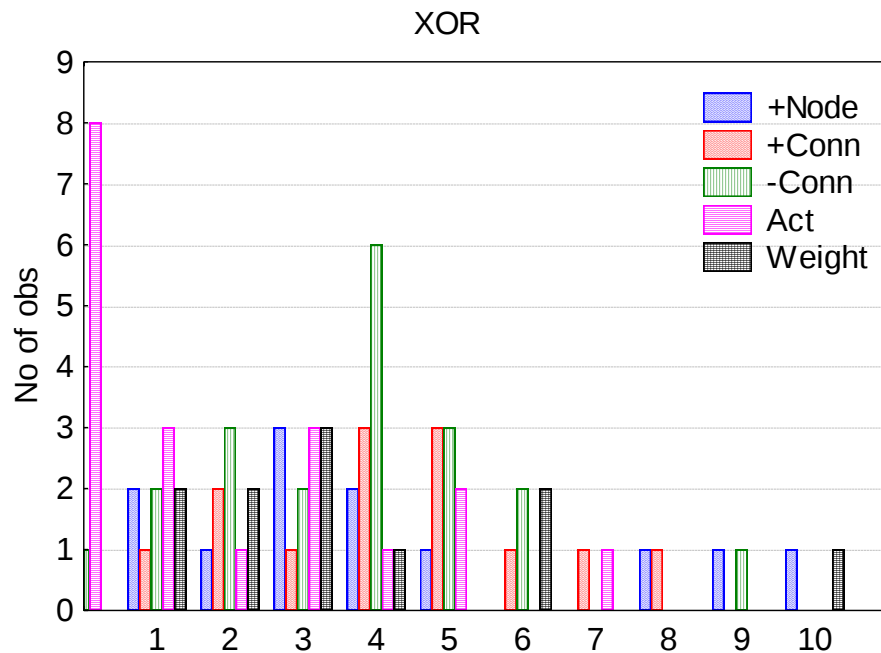
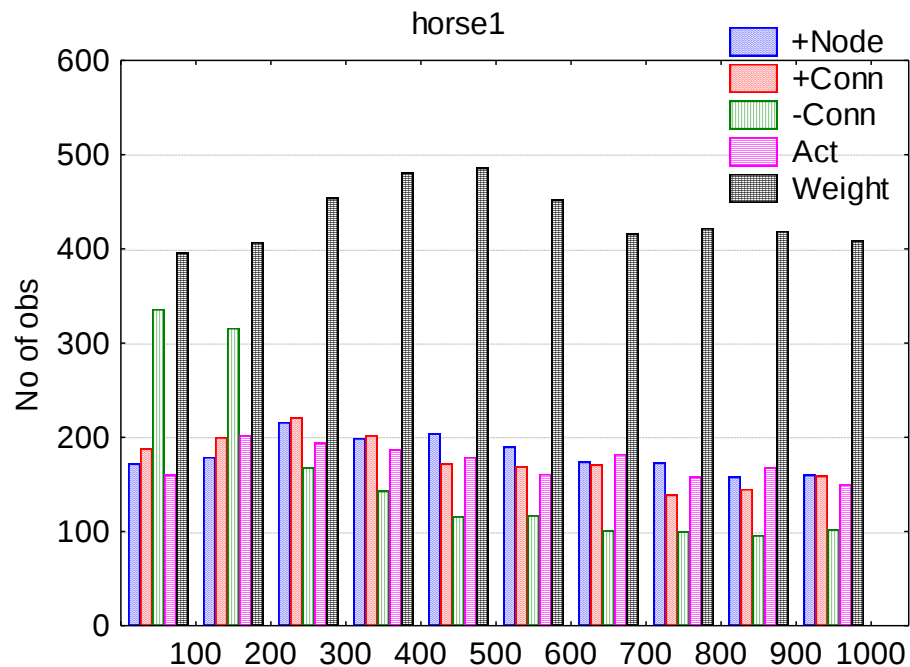
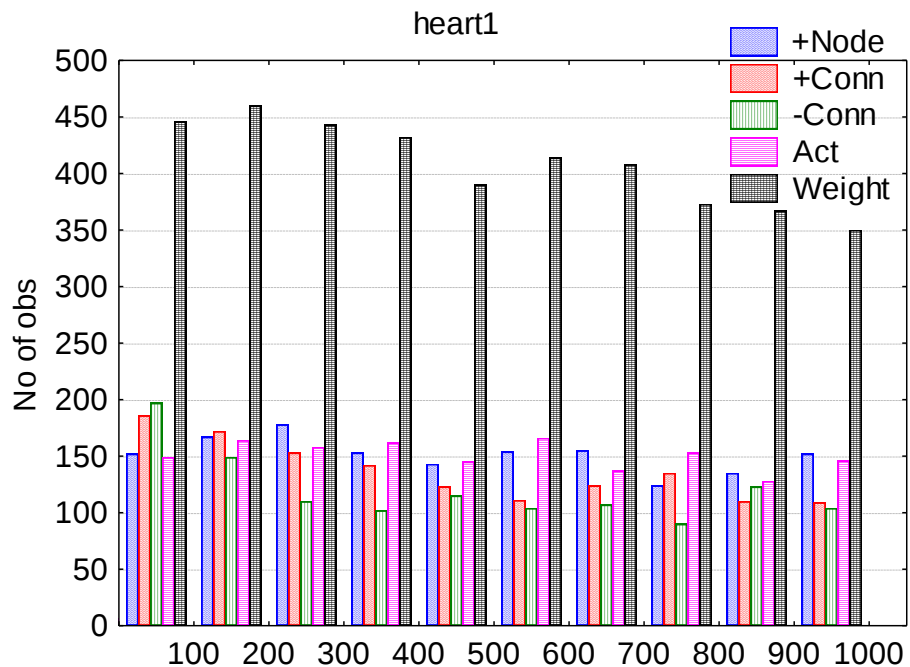
*Изменение функции **активации** (недооцененное во многих исследованиях) очень полезно.*



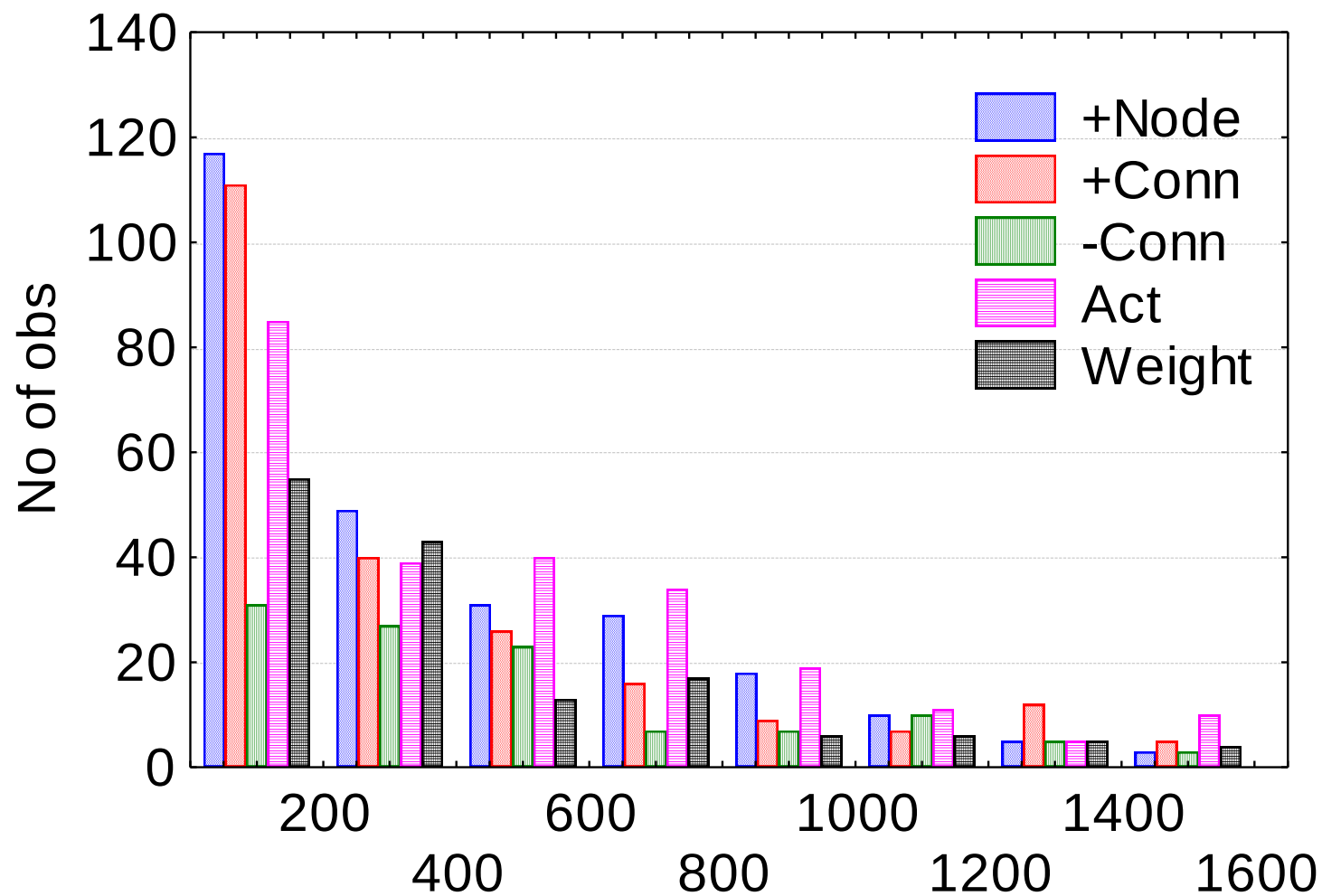
Результаты анализа полезности операторов

- Ни в одной задаче у лучших особей **не был** использован **оператор удаления нейрона**, несмотря на то, что в целом количество использований этого оператора было такое же, как и у других операторов!





Art. ant



Результаты анализа частоты использования операторов

- И опять **оператор удаления нейрона не был использован** лучшими особями.
- **Успешное применение** операторов изменяется со временем.
- **Эвристика:**
Уменьшение вероятности мутации с ростом числа поколений.
- **Эвристика:**
Оператор мутации веса связи должен иметь большую вероятность.

Результаты анализа частоты использования операторов

- Количество операций **добавления нейрона** оказалось на удивление большим (в силу нейтральности)
- Отсутствие обновлений лучшей особи свидетельствует о **стагнации** поиска.

Анализ переходов

Operator Name	Fitness change	Generation
+Node	0	2
Weight	$-6.73E - 06$	3
Weight	$7.24E - 06$	5
-Conn	-0.01	6
+Node	0	7
+Node	0	7
Act	0	7
Weights	$-1.28E - 06$	7
+Conn	0	9
Weight	$-4.02E - 05$	9
-Conn	$3.79E - 05$	11

Задача

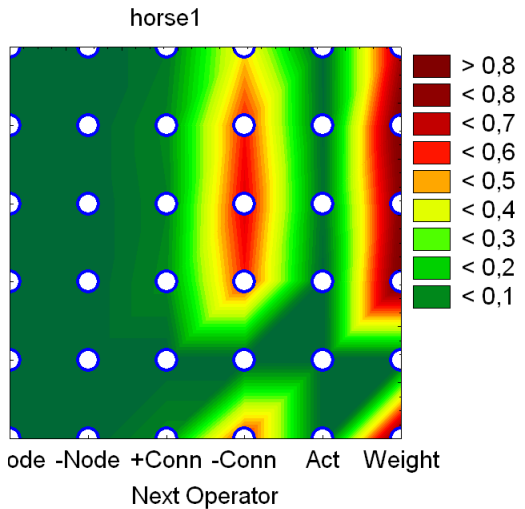
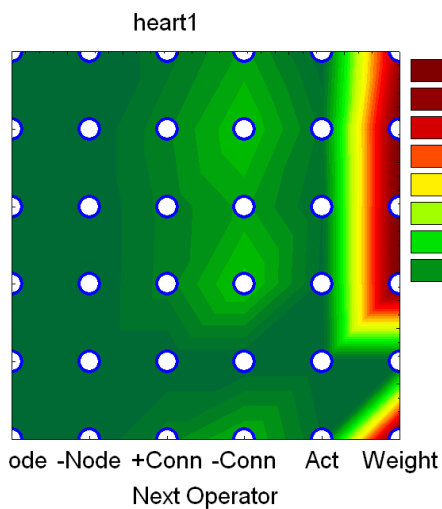
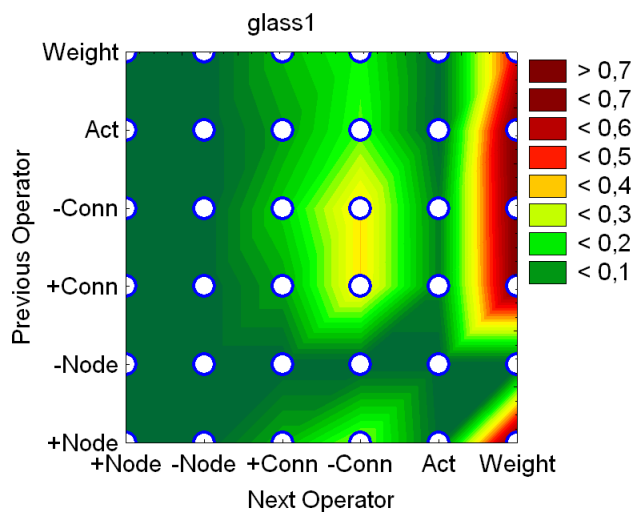
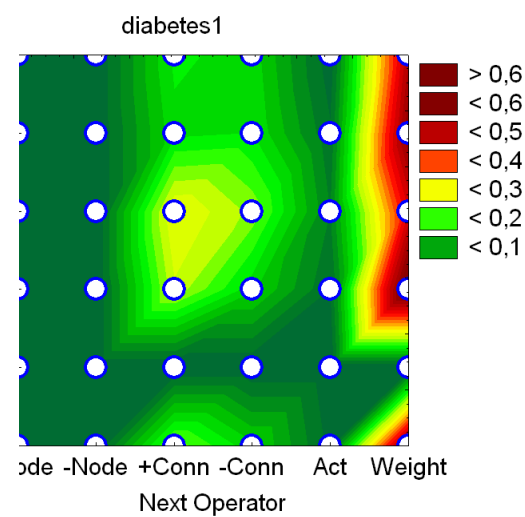
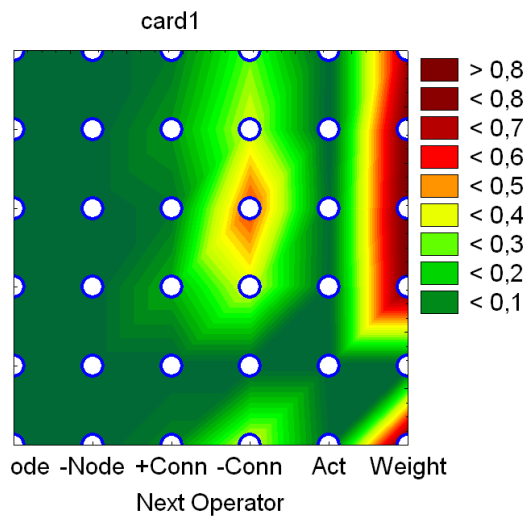
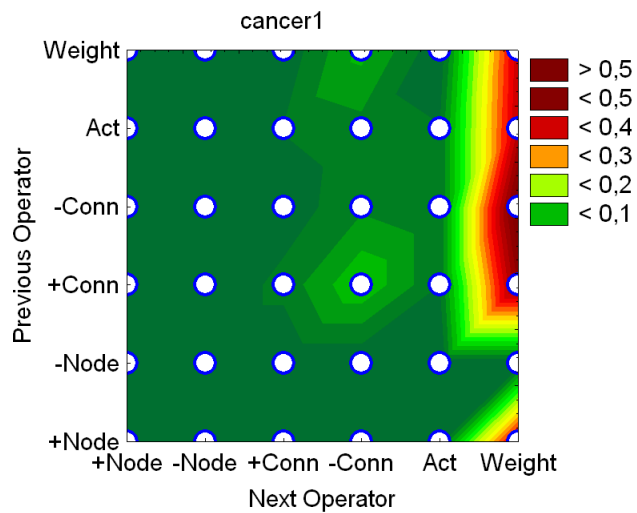
МИНИМИЗАЦИИ

+Node → Weight

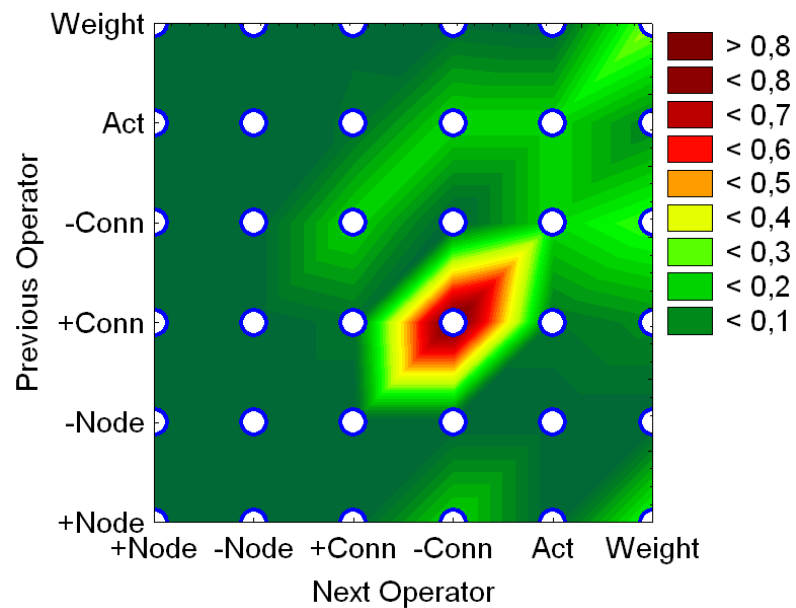
1.0

Weight → -Conn

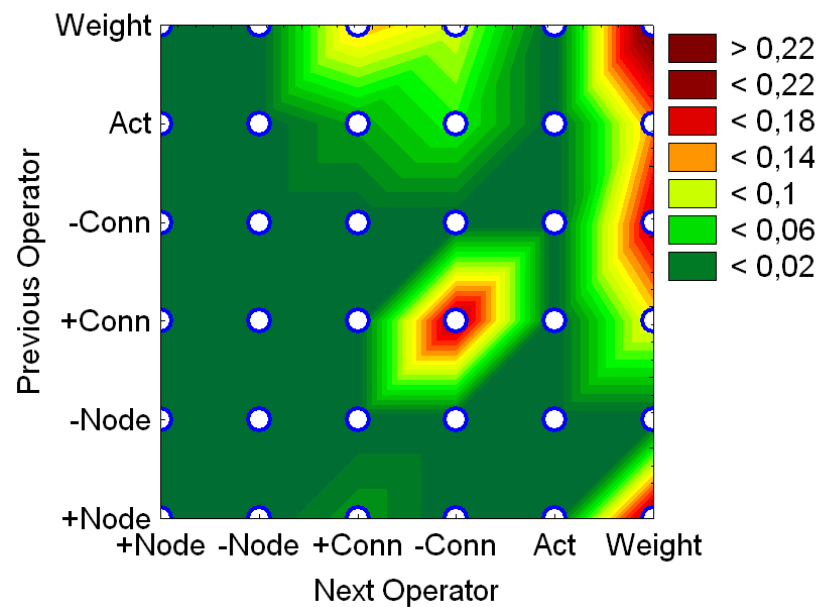
0.5



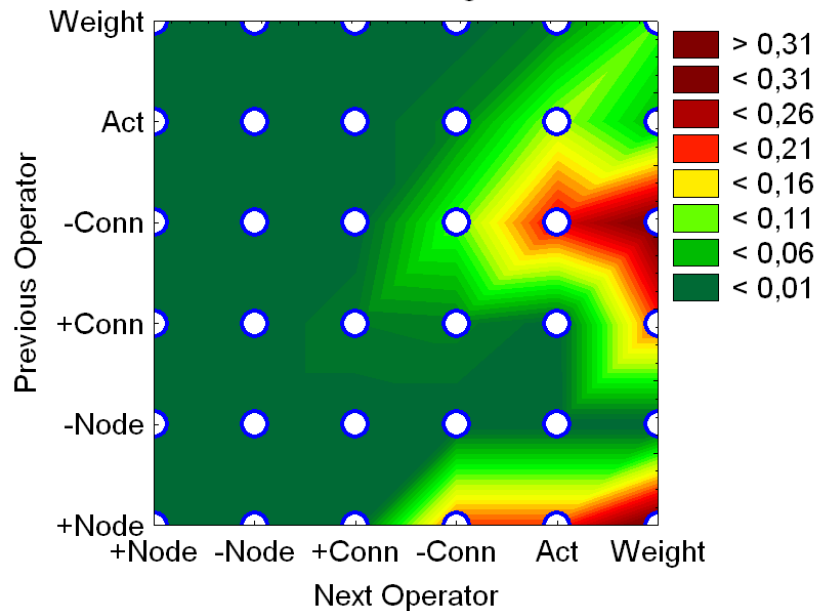
XOR



Artificial Ant



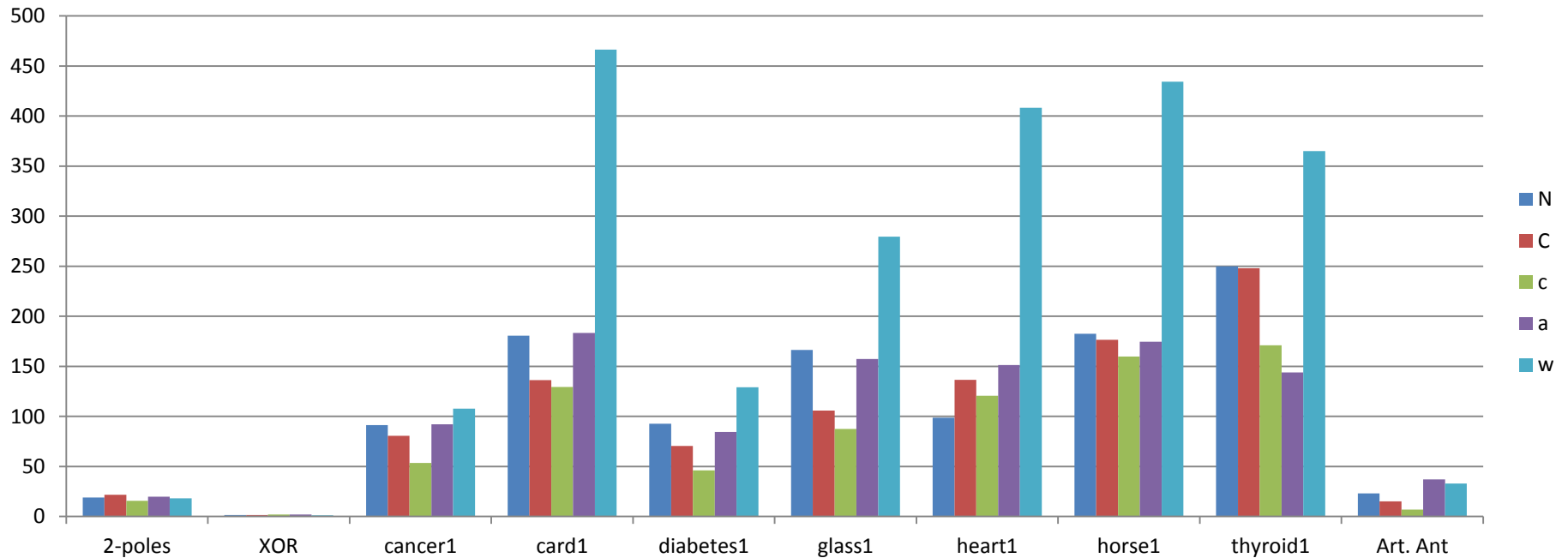
2-Poles Balancing



Результаты анализа переходов

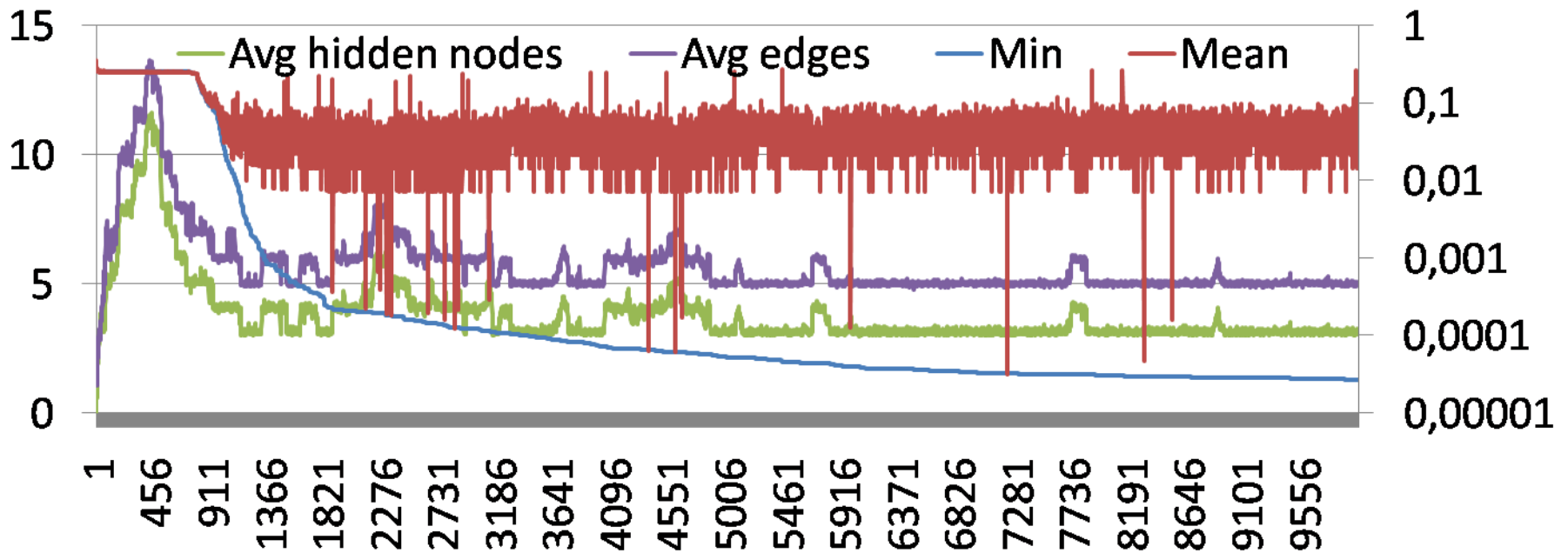
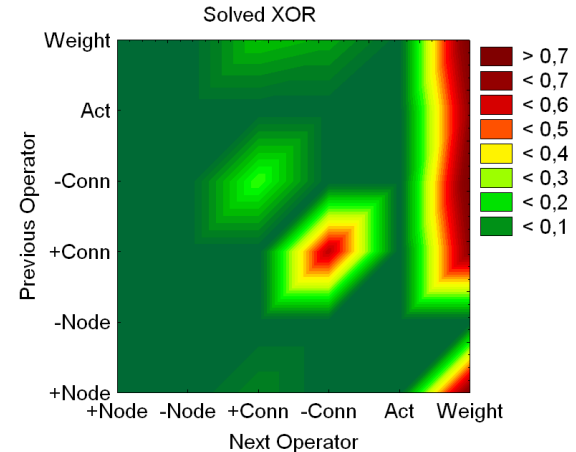
- Матрицы успешных переходов **похожи**, даже для разных по типу задач.
- **Эвристика:**
Решение различных задач может иметь много общего.
- Если решение задачи неуспешно, то матрица переходов, возможно, будет «**аномальной**».
- **Удаление связи** в целом полезно (сдерживающий фактор)

Количество применений операторов



Перенос опыта?

XOR, увеличенная
вероятность мутации



Модификация

Выбор оператора осуществляется в соответствии с полученной матрицей переходов.

	N	n	C	c	a	w
N	0	0	0,071429	0	0,004651	0,748768
n	0	0	0	0	0	0
C	0	0	0	0,625	0	0,705882
c	0	0	0,25	0	0	0,785714
a	0	0	0	0	0,004329	0,695652
w	0	0	0,153846	0,117647	0,011111	0,782772

Алгоритм

Algorithm 2 Selection of operators using successful probability matrix

- 1: $\mathbf{T} \leftarrow$ matrix of successful transitions probabilities.
 - 2: $Op_{-1} \leftarrow$ operator acted last
 - 3: $\mathbf{v} \leftarrow$ vector of transitions probabilities extracted from $\mathbf{T}$ corresponding to Op_{-1}
 - 4: $\mathbf{v} \leftarrow \mathbf{v} + 0.1$ {Increase each element in $\mathbf{v}$ by 0.1 to enable selection of operators with 0 probability.}
 - 5: $\mathbf{v} \leftarrow (\sum_i \mathbf{v}_i)^{-1} \mathbf{v}$ {Normalize probabilities.}
 - 6: $Op \leftarrow$ operator selected according to the distribution $\mathbf{v}$.
-

- Успеха не было...

Анализ результата

1. Различная **инициализация** весов.
2. Матрица переходов показывает лишь **усредненную** вероятность и не учитывает наличие различных этапов эволюционного поиска.
3. Необходимо динамическое (например, по Байесу) **отслеживание применимости операторов**, а не только учет условных вероятностей успеха.



Тестирование

Problem	Proben1	GA	Prunning	$\tau = 15$	
cancer1 (9)	1.38	1.24	1.1	1.78 (4.6)	4.02
card1 (51)	14.05	14.27	13.7	15.64 (51)	14.77
diabetes1 (8)	24.10	23.70	20.8	25.00 (8)	30.63
glass1 (9)	32.7	47.62	30.2	32.07 (7.9)	53.23
heart1 (35)	19.72	21.87	18.5	20.00 (34.2)	21.26
horse1 (58)	29.19	26.44	26.9	30.66 (58)	27.36

Задача «Artificial Ant» решена в 20% запусков (считается сложной).

Не удалось решить задачу балансирования 2 шестов и XOR.

Заключение

Экспериментально получены эвристики как известные, так и новые:

- *Адаптация к размерности задачи.*
- *Полезность изменения активации.*
- Малая полезность оператора удаления нейрона.
- Большая вероятность мутации весов.
- Снижение вероятности мутации с течением времени.
- *Наличие общих закономерностей в различных задачах.*

Заключение

- Задача «переноса опыта» в лоб не решается.
- Использование кроссинговера и другого оператора добавления нейрона может изменить результат.



Нейроэволюция VS Случайность

Спасибо за внимание!

