

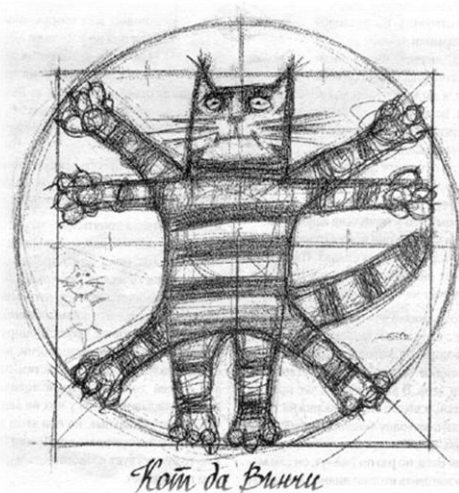
Эвристики в нейроэволюционных алгоритмах

Семинар по вычислительному интеллекту

Докладчик: Юрий Цой

г. Томск, 5 мая 2012 г.

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



Введение

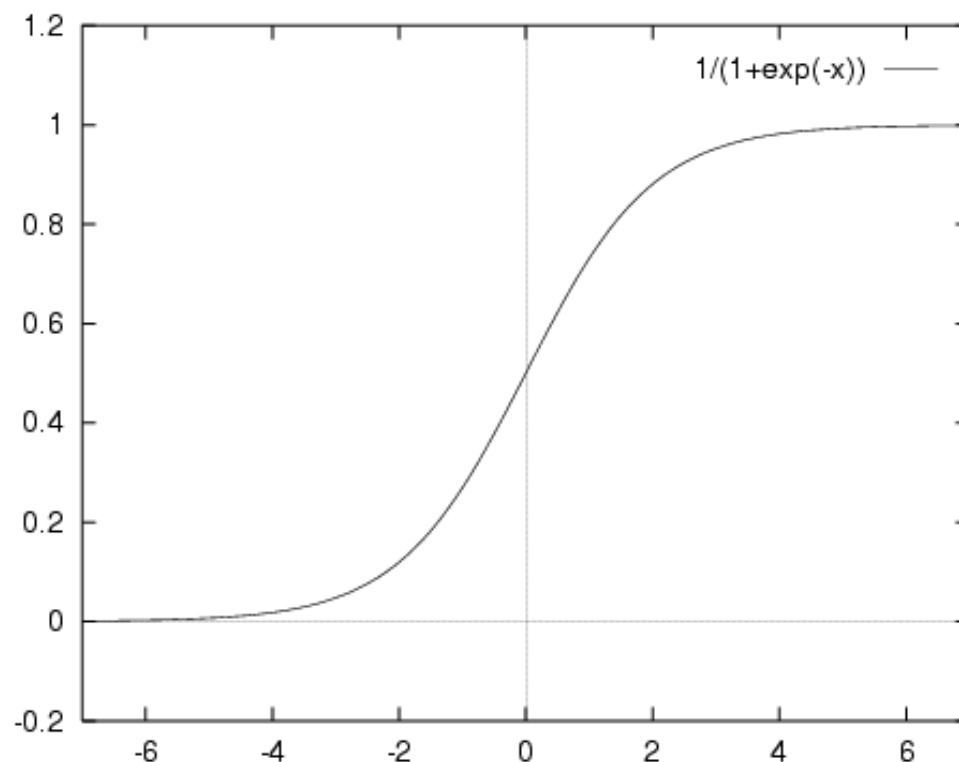
Причины:

- Очень большое пространство поиска со сложной структурой. У алгоритма без эвристик шансов показать хорошие результаты довольно невелики.
- Сложно (часто невозможно) определить вектор градиента.

Эвристики для весов связей

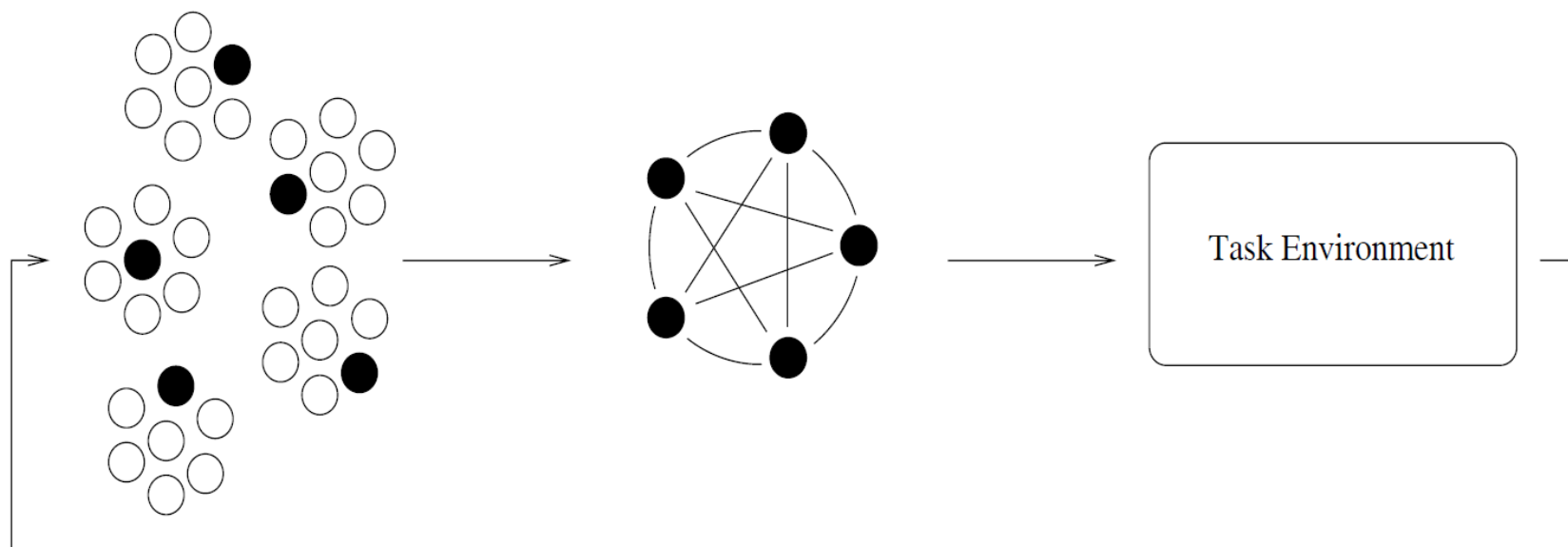
- Ограничение весов связей по модулю ($[-0,1; 0,1]$).

$$f(\mathbf{x}, \mathbf{w}) = \frac{1}{1 + \exp(-a\mathbf{x}^T \mathbf{w})}$$



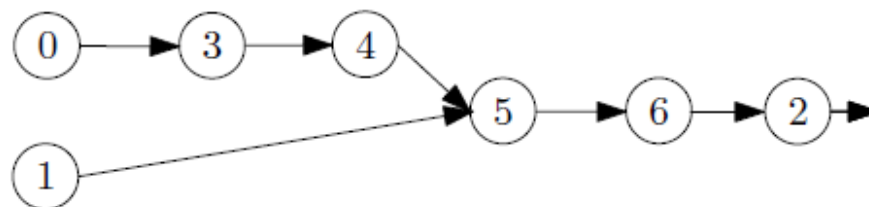
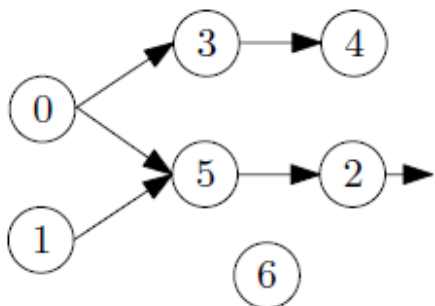
Эвристики для весов связей

- Декомпозиция задачи → сокращение размерности
- ESP (F. Gomez)



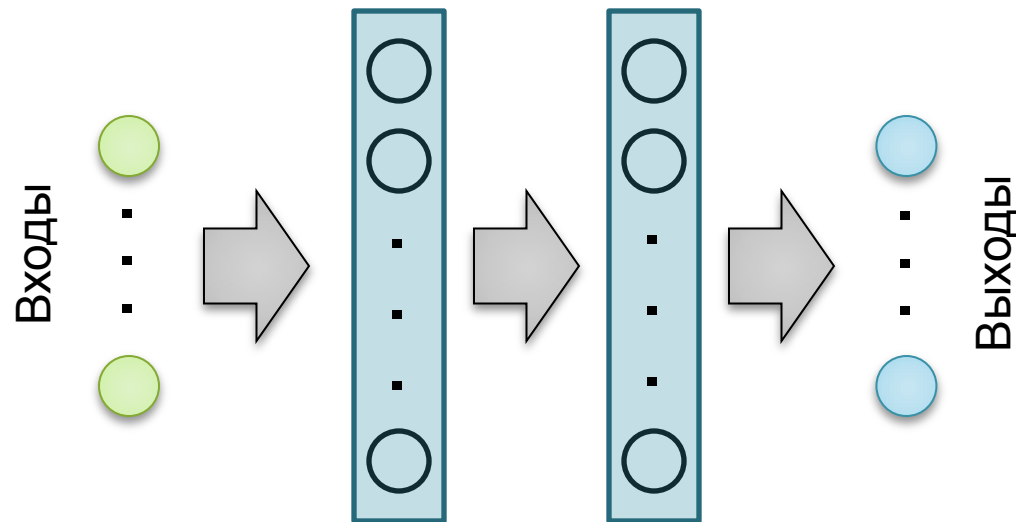
Эвристики для структуры

- Ограничение структуры.
- Ограничение направления поиска.
- Адаптация операторов:
 - Выбор.
 - Применение.



Ограничение структуры

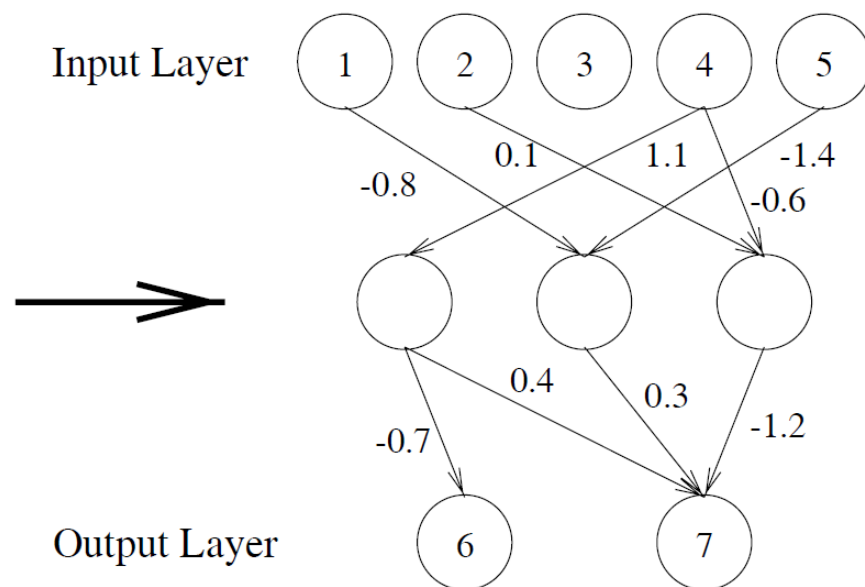
- Только многослойные ИНС (с ограниченным/фиксированным количеством слоев).



Ограничение структуры

- Только ИНС прямого распространения.
- SANE (D. Moriarty)

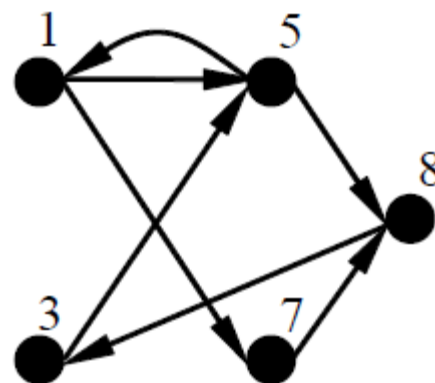
Label Weight					
4	1.1	6	-0.7	7	0.4
1	-0.8	5	-1.4	7	0.3
2	0.1	4	-0.6	7	-1.2



Ограничение структуры

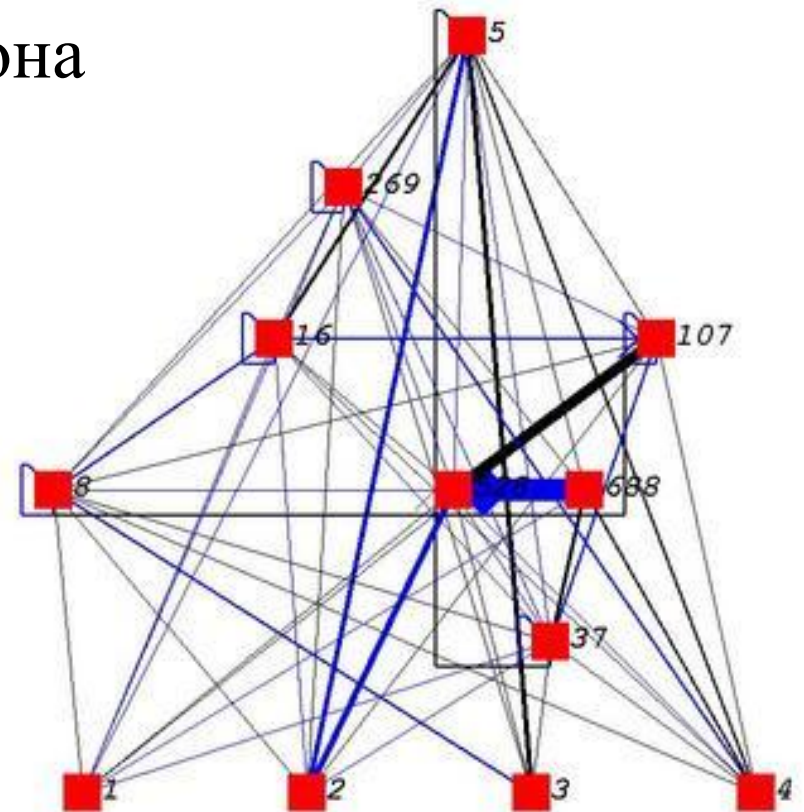
- Ограничение количества нейронов/связей
- Алгоритм Н. Kitano

0	0	0	0	1	0	1	0
0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1
0	0	1	0	0	0	0	0



Ограничение направления поиска

- Только «растущие» ИНС
- NEAT (K. Stanley)
 - Добавление нейрона
 - Добавление связи



Адаптация операторов

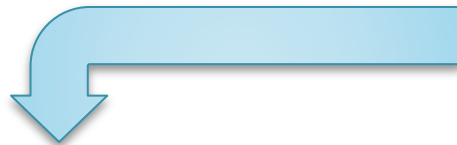
- Вероятность выбора
- Вручную (эвристики + по результатам экспериментов):
 - CE
 - SANE
 - ESP
 - NEAT
 - AGE
 - и многие другие

Часто вероятность мутации весов делают относительно большой

Часто вероятность добавления нейрона делают относительно небольшой

Адаптация операторов

- Вероятность выбора
- Подстраиваются на основе эвристик
- NEvA



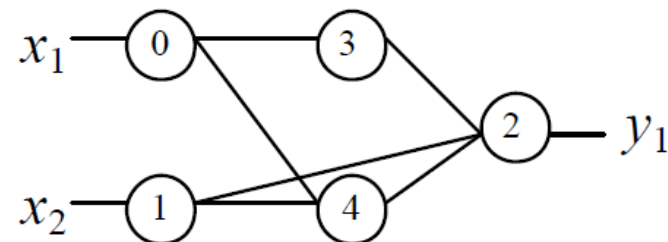
$$F_C = f_C^2 = 0,445,$$
$$F_N = F_C * f_N^2 = 0,160$$

$$p_{+conn} = (1 - F_C)(1 - F_C + \alpha F_C),$$

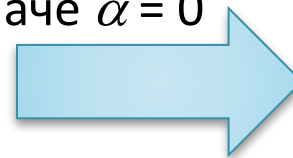
$$p_{-conn} = F_C(1 - F_N)(F_C + \alpha),$$

$$p_{+neur} = F_C F_N,$$

$$p_{-neur} = F_C(1 - F_C)(2 - F_N)(1 - \alpha).$$



$\alpha = 1$ если нет
скрытых узлов,
иначе $\alpha = 0$



$$p_{+conn} = 0,308,$$

$$p_{-conn} = 0,166,$$

$$p_{+neur} = 0,072,$$

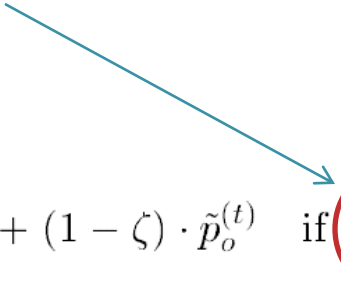
$$p_{-neur} = 0,454.$$

Адаптация операторов

- Вероятность выбора
- Подстраиваются на основе эвристик
- Igel, Kreutz, 2001 – обновление вероятностей операторов по «успешности»

If $t \bmod \tau = 0$ then

for each $o \in \Omega$ do

$$\tilde{p}_o^{(t+1)} := \begin{cases} \zeta / q_{\text{all}}^{(t,\tau)} \cdot \sum_{i=0}^{\tau-1} q_o^{(t)} + (1 - \zeta) \cdot \tilde{p}_o^{(t)} & \text{if } q_{\text{all}}^{(t,\tau)} > 0 \\ \zeta / |\Omega| + (1 - \zeta) \cdot \tilde{p}_o^{(t)} & \text{else} \end{cases}$$


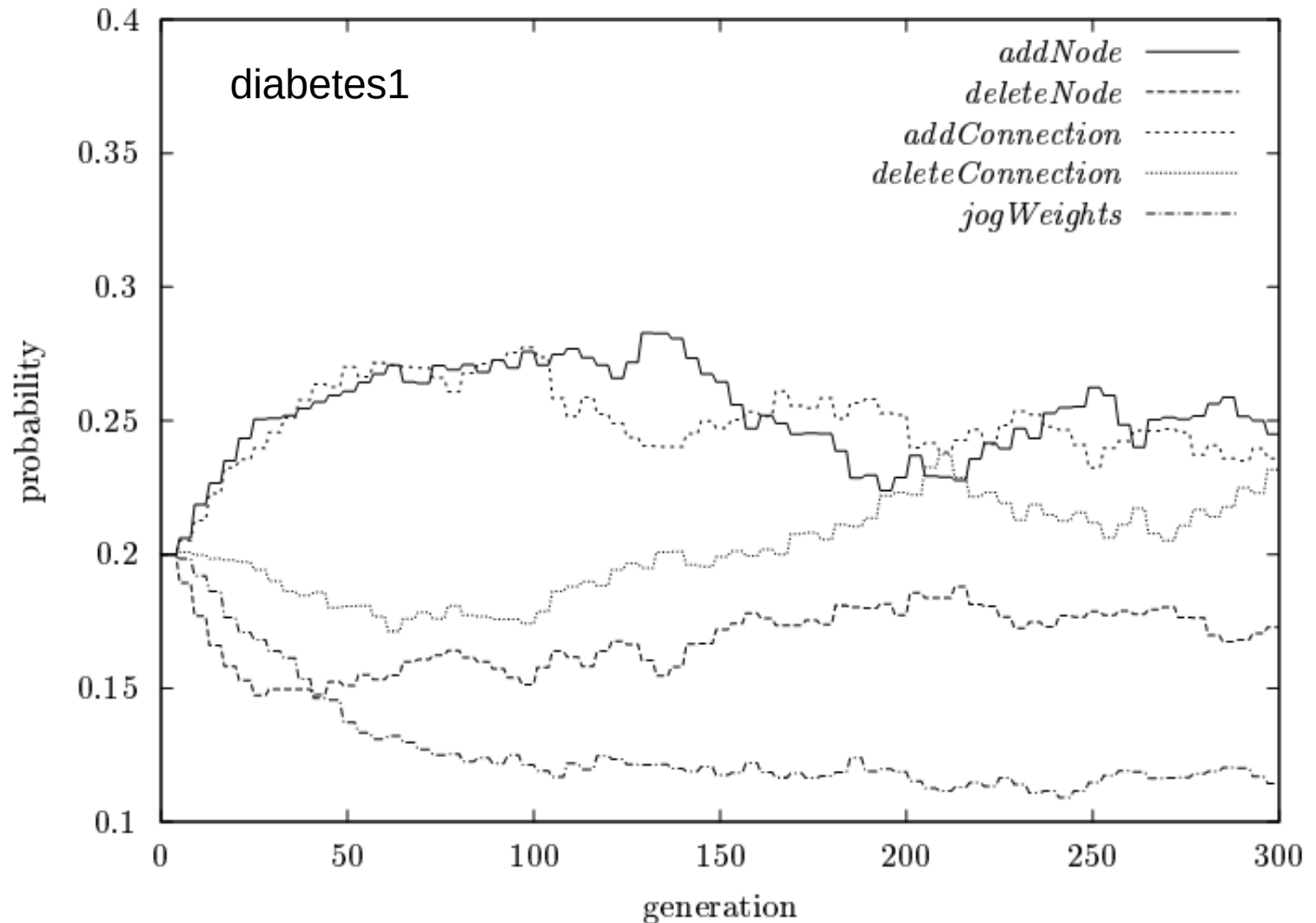
for each $o \in \Omega$ do

$$p_o^{(t+1)} := p_{\min} + (1 - |\Omega| \cdot p_{\min}) \frac{\tilde{p}_o^{(t+1)}}{\sum_{o' \in \Omega} \tilde{p}_{o'}^{(t+1)}}$$

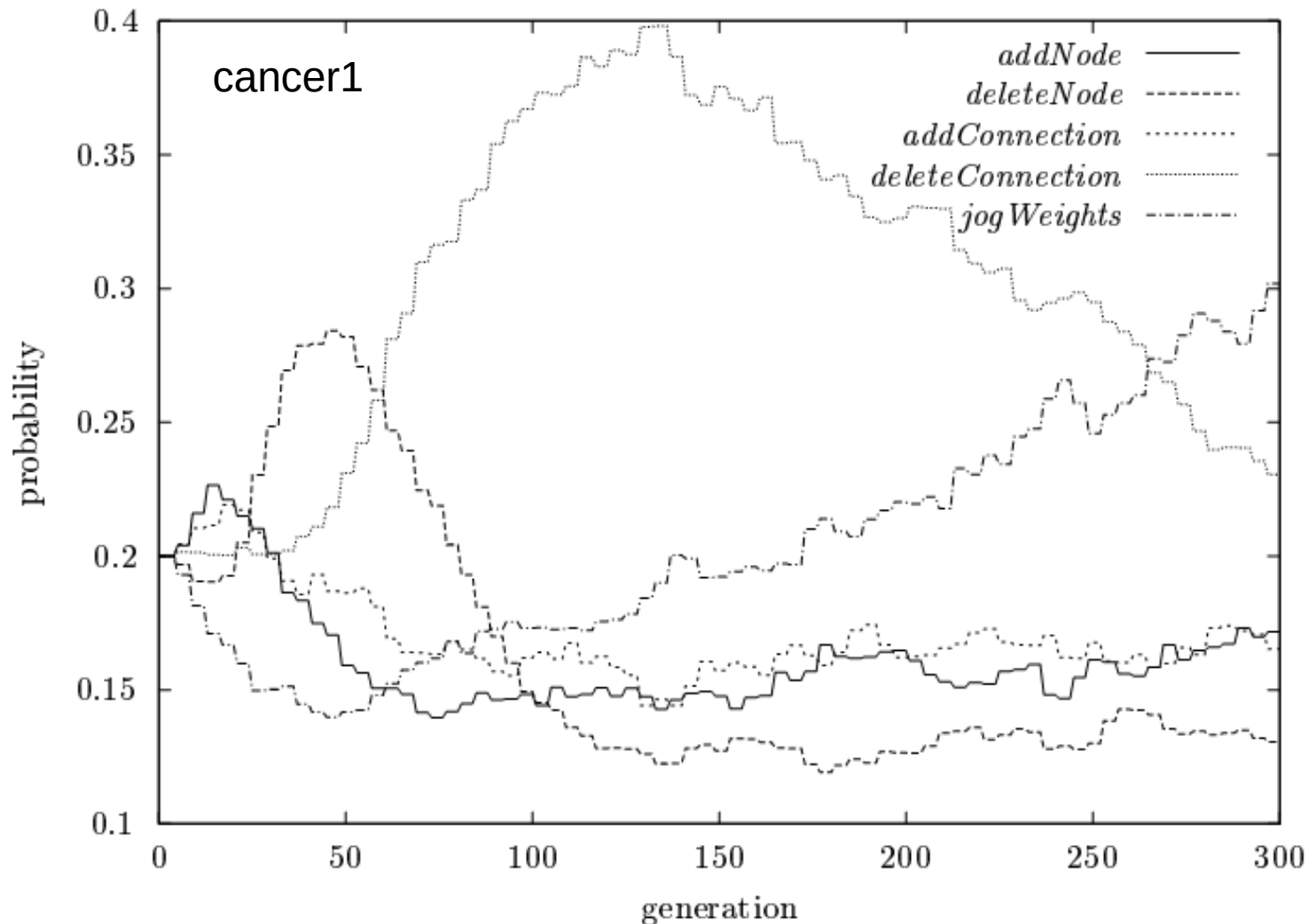
else

$$p_o^{(t+1)} := p_o^{(t)}$$

Igel, C., and Kreutz, M. (2003)



Igel, C., and Kreutz, M. (2003)



Адаптация операторов

- Применение операторов (какие связи, нейроны удалять, куда добавить новую связь и т.д.)
- Yao, Liu, 1996

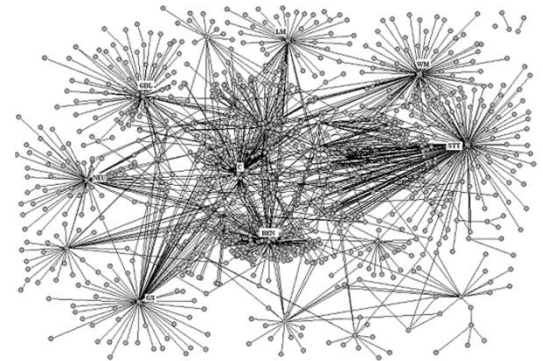
«Важность» связи:

$$test(w_{ij}) = \frac{\sum_{t=1}^T \xi_{ij}^t}{\sqrt{\sum_{t=1}^T (\xi_{ij}^t - \bar{\xi}_{ij})^2}}$$

$$\xi_{ij}^t = w_{ij} + \Delta w_{ij}^t(w)$$

Адаптация операторов

- Применение операторов (какие связи, нейроны удалять, куда добавить новую связь и т.д.)
- SARN
(безмасштабные сети)



Вероятность $P(i)$ выбора i -го узла сети зависит от степени d_i его связанности

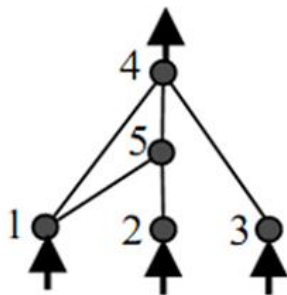
$$P(i) = \begin{cases} \frac{d_i + \gamma}{\sum_k (d_k + \gamma)}, & m > 0 \\ \frac{d_{\max} - d_i + \gamma}{\sum_k (d_{\max} - d_k + \gamma)}, & m \leq 0 \end{cases}$$

Адаптация операторов

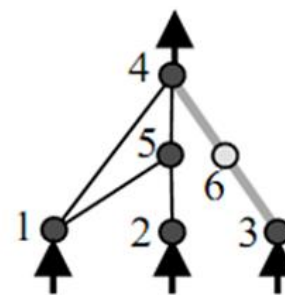
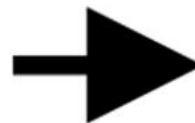
- Добавление нейрона
 - Добавление нейрона со связями
 - Вставка нейрона

1	2	3	4	5	6
1-→4	2-→4	3-→4	2-→5	5-→4	1-→5
	DIS				

1	2	3	4	5	6	8	9
1-→4	2-→4	3-→4	2-→5	5-→4	1-→5	3-→6	6-→4
	DIS	DIS					



Mutate Add Node



Эвристики для функции приспособленности

- Введение регуляризации
 - По весам связей

$$\|\mathbf{w}\| \quad \text{или} \quad \sum_i |\mathbf{w}_i|$$

- По размеру сети

$$N_{\text{нейронов}} \quad \text{или} \quad N_{\text{связей}}$$

$$\text{или} \quad f(N_{\text{нейронов}}, N_{\text{связей}})$$

- По весам связей и размеру сети

«Обобщения»

- Ограничение роста размера сети (размерности задачи)
 - большая вероятность мутации весов связей
 - малая вероятность добавления нового нейрона
 - добавление штрафных коэффициентов в функцию приспособленности
- Повышение качества работы операторов модификации ИНС
 - Адаптация к структуре ИНС
 - Зависимость от приспособленности
 - Аналогии

СПИСОК ИСТОЧНИКОВ

1. Angeline, P., Saunders, G., and Pollack, J. (1993). An evolutionary algorithm that constructs recurrent neural networks. *IEEE Trans. on Neural Networks*, (5):54-65.
2. Barabasi, A.-L. (2003). *Linked: How Everything Is Connected to Everything Else and What It Means*. Plume.
3. Barrera, F. (2007). Busqueda de la estructura optima de redes neurales con algoritmos geneticos y simulated annealing. verificacion con el benchmark proben1. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 11(34):41–61.
4. Fogel, D. B. (1993). Using evolutionary programming to create neural networks that are capable of playing tic-tac-toe. In *International Conference on Neural Networks*, pages 875–880, San Francisco, CA. IEEE Press.
5. Gomez, F. (2003). *Robust Non-Linear Control through Neuroevolution*. PhD thesis, Department of Computer Sciences.
6. Gruau, F. (1994). *Neural network synthesis using cellular encoding and the genetic algorithm: Unpublished PhD thesis*. PhD thesis, lUniversite Claude Bernard, Lyon.
7. Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation (2nd Edition)*. Prentice Hall.
8. Hesser, J. and Manner, R. (1990). Towards an optimal mutation probability for genetic algorithms.
9. Igel, C., and Kreutz, M. (2003) Operator adaptation in evolutionary computation and its application to structure optimization of neural networks. *Neurocomputing*. Vol. 55 (1-2). P. 347-361.

СПИСОК ИСТОЧНИКОВ

10. In Schwefel, H.-P., M. R., editor, Proceedings of the 1st Conference on Parallel Problem Solving from Nature. LNCS No. 496, pages 23–32, Berlin: Springer-Verlag.
11. Kitano, H. (1990). Designing neural network using genetic algorithm with graph generation system. Complex Systems, (4):461–476.
12. Koza, J. R. (1992). Genetic Programming: On the Programming of Computers by Natural Selection. MIT Press, Cambridge, MA, USA.
13. Luke, S. (2009). Essentials of Metaheuristics. Lulu.
14. Mattiussi, C., Durr, P., Marbach, D., and Floreano, D. (2011). Beyond graphs: A new synthesis. Journal of Computational Science, 2(2):165–177.
15. Moriarty, D. and Miikkulainen, R. (1996). Efficient reinforcement learning through symbiotic evolution. Machine Learning, (22):11–32.
16. Prados, D. (1992). New learning algorithm for training multilayered neural networks that uses genetic-algorithm techniques. Electronics Letters, 28(16):1560–1561.
17. Prechelt, L. (1994). Proben1 - a set of neural network benchmark problems and benchmarking rules. Technical Report 21/94, Fakultät für Informatik, Universität Karlsruhe, Karlsruhe, Germany.
18. Stanley, K. (2006). Patterns without development, 2006. Technical report CS-TR-06-01, University of Central Florida.

СПИСОК ИСТОЧНИКОВ

19. Stanley, K. and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10(2):99–127.
20. Tsoy, Y. and Spitsyn, V. (2004). Using genetic algorithm with adaptive mutation mechanism for neural networks design and training. *Optical memory and neural networks*, 13(4):225–232.
21. Tsoy, Y. R. (2008). Computational regulatory networks and self-adaptive neuroevolutionary algorithm. In 11th Int. Conf. with Int. Participation on Artificial Intelligence (CAI-2008), volume 3, pages 50–57. Moscow: LENAND. In Russian.
22. Yao, X. and Liu, Y. (1996). Evolutionary artificial neural networks that learn and generalise well. In Proc. of the 1996 IEEE International Conference on Neural Networks (ICNN'96), volume on Plenary, Panel and Special Sessions, pages 159–164.

Цой Ю. Эвристики в нейроэволюционных алгоритмах

СПАСИБО ЗА ВНИМАНИЕ!

