

Self-configuration from a Machine-Learning Perspective

Wolfgang Konen

Institute for Informatics, Cologne University of Applied Sciences
Steinmüllerallee 1, D-51643 Gummersbach, Germany

<http://www.gm.fh-koeln.de/~konen>

wolfgang.konen@fh-koeln.de

Contribution to **Dagstuhl Seminar 11181**

"Organic Computing - Design of Self-Organizing Systems"

organized by

Kirstie Bellmann, Andreas Herkersdorf and Michael G. Hinchey

Abstract

The goal of machine learning is to provide solutions which are trained by data or by experience coming from the environment. Many training algorithms exist and some brilliant successes were achieved. But even in structured environments for machine learning (e.g. data mining or board games), most applications beyond the level of toy problems need careful hand-tuning or human ingenuity (i.e. detection of interesting patterns) or both. We discuss several aspects how self-configuration can help to alleviate these problems. One aspect is the self-configuration by tuning of algorithms, where recent advances have been made in the area of SPO (Sequential Parameter Optimization). Another aspect is the self-configuration by pattern detection or feature construction. Forming multiple features (e.g. random boolean functions) and using algorithms (e.g. random forests) which easily digest many features can largely increase learning speed. However, a full-fledged theory of feature construction is not yet available and forms a current barrier in machine learning. We discuss several ideas for systematic inclusion of feature construction. This may lead to partly self-configuring machine learning solutions which show robustness, flexibility, and fast learning in potentially changing environments.

1 Introduction

How do we learn in new environments? It is a striking feature of human behaviour that humans can adapt rapidly to new situations or environments, can spot important patterns if only few examples are given and can perform meaningful generalizations. Computing machines show a much poorer performance in these disciplines. This is not only true for complex environments with noisy sensory information, but also for very clean and structured data.

As examples we will study in this work two cases with fairly structured information: board games and data mining.

Games are an ideal testbed for the study of learning and self-configuring systems. It is not so much that we are interested in construction world-leading AI players for complex games like chess (this problem is mainly solved today). The interesting point is nicely formulated by Simon Lucas [Luc08]:

The immediate goal of the research ... is to study the effectiveness of machine learning approaches to game playing: how well a machine can learn to play, rather than how well we can program it to play. In particular we are interested in how well the system can learn to play without any expert tuition, and without recourse to an expert opponent to practice against.

In games you need to discover a strategy. A strategy can be the detection of and reaction on the right pattern / feature. In principle, neural networks and reinforcement learning (RL) could solve the problem because they can learn arbitrary features. Indeed, some remarkable successes were achieved, e.g. Tesauro's TD-Gammon [Tes94]. Pollack and Blair showed an alternative approach based co-evolution [PB98], which was shortly followed by many neuroevolution contributions, starting with the work of Chellapilla and Fogel [CDF99].

But there are other cases in RL, where the learning progress is disappointingly slow [Tes92, KBB09], even when only simple games are considered. We have shown in [KBB09] examples from reinforcement learning where a simple toy problem is not solvable with the wrong features, but quickly solvable with the right features. It would be desirable to have a mechanism with which one can learn or construct such interesting features. We show below a general approach based on N-tuple systems, which was recently applied by Lucas [Luc08] to board game problems with remarkable success.

In data mining, our second example area, we have for small problems (e.g. UCI data) or for special problems robust learning solutions (e.g. random forests). But no solution is out there which works equally robust on a large variety of problems or on larger problems. For larger problems we usually still need careful hand-tuning or ingenuity (detect patterns) or both. It is the aim to replace some of this by self-configuration. Advances in this domain will have numerous applications, especially in the area of online data mining (stream data mining) with its often non-stationary environment conditions.

Recent data-mining developments try to use larger amount of features (many parallel features) and bring the optimal use of features into the optimization loop, as we try

in our current research project SOMA (Systematic Optimization of Models in IT and Automation).

Other developments, e.g. the IBM Watson project, where a supercomputer learns to answer questions from the Jeopardy quiz [FB⁺10], rely on massively parallel models with multi-hypothesis testing and multiple confidence estimation

In this paper we try to advocate the following (still unproven) conjecture:

Finding good features and using abundant features is key for learning in high-dimensional input spaces.

The rest of the paper is organized as follows: Sec. 2 looks at self-configuration and self-learning (w/o teacher) in games and looks specifically at the N-tuple approach. Sec. 3 is about self-configuration in data mining (DM), it discusses the desirable characteristics of self-configuring, robust, and flexible DM solutions. It looks briefly at tuning and at feature construction in DM. Finally in Sec. 4 we discuss our findings and present a set of open research questions.

2 Self-configuration in games

2.1 Strategy example

Reinforcement learning is a remarkable example of self-organization in game play: Just by playing the game against itself, an agent can learn emergent behaviour, e.g. to play on a master level for some games, as was shown with Tesauro's TD-gammon [Tes94]. But reinforcement learning alone is not the complete solution. It is necessary to detect / construct the right features, otherwise the success may be slowed down dramatically or may be completely blocked [KBB09].

A simple example might illustrate the point: In the game Nim-3 either player can take 1, 2 or 3 pieces from a collection of initially n pieces and winner is the one who picks the last piece. The right feature to detect is of course: "Leave your opponent m stones, where m is a number divisible by 4." As long as you do not spot this pattern or feature, only complicated calculations find the right move for sufficiently large n . With this feature, the formulation of a winning strategy is trivial for arbitrary n .

We investigated in [KBB09] a similar but non-trivial case for a small game (TicTacToe) and found similar results (Fig. 1): Some feature sets (T2, T4) make TD-learning very fast, while others (T0, T1) completely block successful TD-learning, among them the raw board position (T0). The feature sets T2 and T4 contained human-designed features.

2.2 N-tuple systems

The question is how to find such or other features given a new task or a new game **without** the need for human design. Lucas [Luc08] has recently proposed to use the old idea of N-tuple systems for the first time also for the detection / construction of useful features

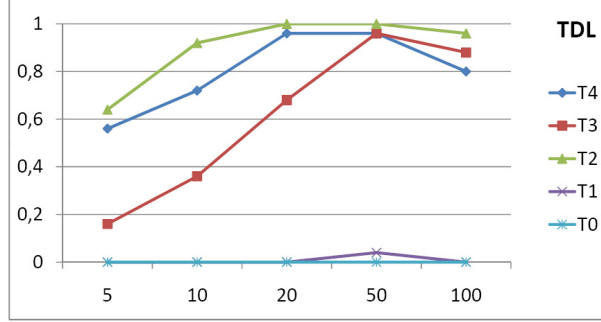


Figure 1: The influence of different feature sets T0,...,T4 on the speed of TD-learning. The x-axis shows the number of training games (in thousands), the y-axis shows the success rate of the trained agent [KBB09].

in games. As an example he considered the game Othello: An N-tuple is here a chain of length N formed by a random walk on the board. Fig. 2 shows a 3-tuple example. The piece situation along the chain is taken as adress into a LUT (look-up table). If the game possesses symmetries (here: 8-fold, reflection and rotation) then all symmetric N-tuple positions do also activate a LUT entry. Therefore we see 8 red arrows in Fig. 2. Each LUT entry d has a weight $l(d)$ connected with it. Given a board position b which activates the LUT entries in set $D(b)$, the output of the N-tuple is simply the sum

$$v(b) = \sum_{d \in D(b)} l(d) \quad (1)$$

Given K such randomly formed N-tuples, the output of the network given a board position b is the sum over all N-tuples

$$V(b) = \sum_{k=1}^K v_k(b) \quad (2)$$

The goal is that $V(b)$ approximates the game's value function, i.e. that it assigns to each board position b the correct probability of Black to win from this b . The weights $l(d)$ are trained by temporal difference learning (TD-learning).

The important message from [Luc08]: Lucas formed randomly 30 N-tuples and he succeeded in generating a strong-playing Othello agent within only 1250 games (!) of training. This is remarkable, since other reinforcement learning approaches to Othello usually need millions of games: van Eck and van Wezel [vEvW05] use Q-learning and need 15,000,000 games and Szubert et al. [SJK09] use coevolution together with TD-learning and need 4,500,000 games.

It seems that the use of abundant features greatly influences the speed of learning, the speed at which self-organization of emergent behaviour appears. Why is this the case? At present we can only formulate hypotheses: As Lucas [Luc08] notes, the N-tuple system

... is somewhat similar to the kernel trick used in support vector machines (SVMs) and is also related to Kanervas sparse distributed memory model [Kan88]. The low

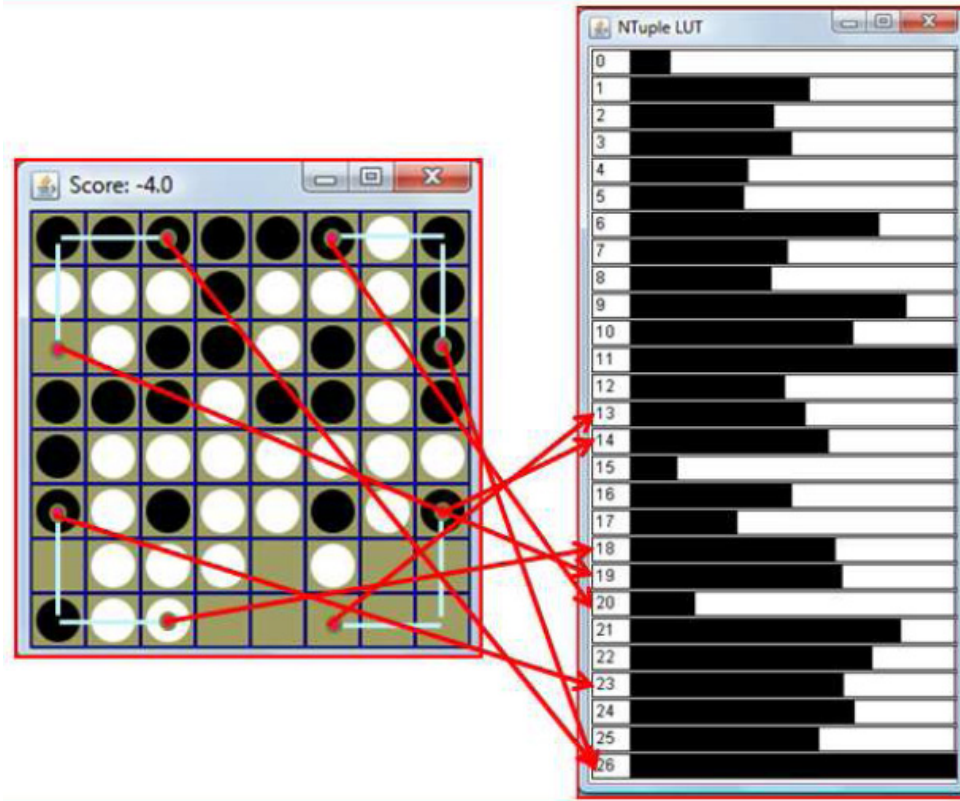


Figure 2: 3-tuple example from [Luc08]. LUT cells are indexed (addressed) with the ternary code {white=0, empty=1, black=2}. Each LUT cell d has a weight $l(d)$ assigned (black bar; the higher, the better for Black).

dimensional board is projected into a high dimensional sample space by the N-tuple indexing process.

The high-dimensional sample space can help to make the game strategy easier to learn. It reduces the probability that weights receive conflicting signals from different input situations. The indexing operation into the LUT performs a non-linear mapping to high-dimensional feature space. The weights of the LUT can be seen as the weights of a single-layer perceptron.

In fact, the N-tuple system has an architecture close to Rosenblatt's perceptron [Ros62]. The original perceptron had hidden units which formed fixed random boolean functions on the subset of the input. N-tuples are also random boolean functions on a subset of the input. This means that N-tuple systems share the strengths (fast training) and the weaknesses (theoretical limitations, "parity problem") with the perceptron. But note that the N-tuple system has learned a rather complex Othello value function. The main differences between the original perceptron and Lucas' approach are that the target signal is delivered by the TD-learning concept and that the number of hidden units (number of LUT entries in all LUTs) is potentially very large. Is it that the vast amount

of LUT entries has an influence on the self-organizational capabilities of the system? We want to investigate this question in the near future.

3 Self-configuration in machine learning and data mining

3.1 What is desirable?

Machine learning models have made considerable progress during the last decades: Random forests (RF) are more robust and flexible than single decision trees; support vector machines (SVM) can better handle large numbers of input variables and build implicitly high-dimensional feature spaces. But still: there is no free lunch, expert knowledge is required to get high quality results for complicated, perhaps noisy tasks. Usually the expert is needed to (a) build the right features (preprocessing), (b) select the right features, and (c) select the right model and tune model and preprocessing parameters.

It would be desirable to have a complete solution which is able to configure itself on the task just by looking at the data and by self-configuring the options in (a) - (c). This becomes even more important if we want to build systems for *online data mining* or *stream data mining*, where new data are continuously streaming in and the underlying process might also be non-stationary, thus requiring self-configuration or re-calibration in daily routine.

3.2 Tuning

Tuning, i.e. parameter adjustment, can be seen as a form of self-configuration. Although the problem is simply stated: "There are N adjustable parameters - find the best values for them given the current task", it is seldom done systematically in its general form. Why? - It quickly gets complex, if N becomes larger (curse of dimensionality) and/or if it contains mixed boolean, integer and real-valued parameters. Furthermore, for complex tasks the modeling step is often time-consuming. Only few parameter adjusting runs may be feasible within the given budget.

Recently, some advances in tuning have been made: Bartz-Beielstein's sequential parameter optimization (SPO) [BBLP10, BB10] builds meta models (surrogate models) with the help of Kriging or other algorithms and makes it possible to get good optimization results within only a few runs of the real model. Other interesting tuning methods are CMA-ES, REVAC, BFGS, and other. A recent comparison of different tuners for DM can be found in [KKF⁺11]. We present exemplary results from [KKF⁺11] in Fig. 3.

3.3 Feature construction and feature selection

According to Fayyad's definition [FPSS96]:

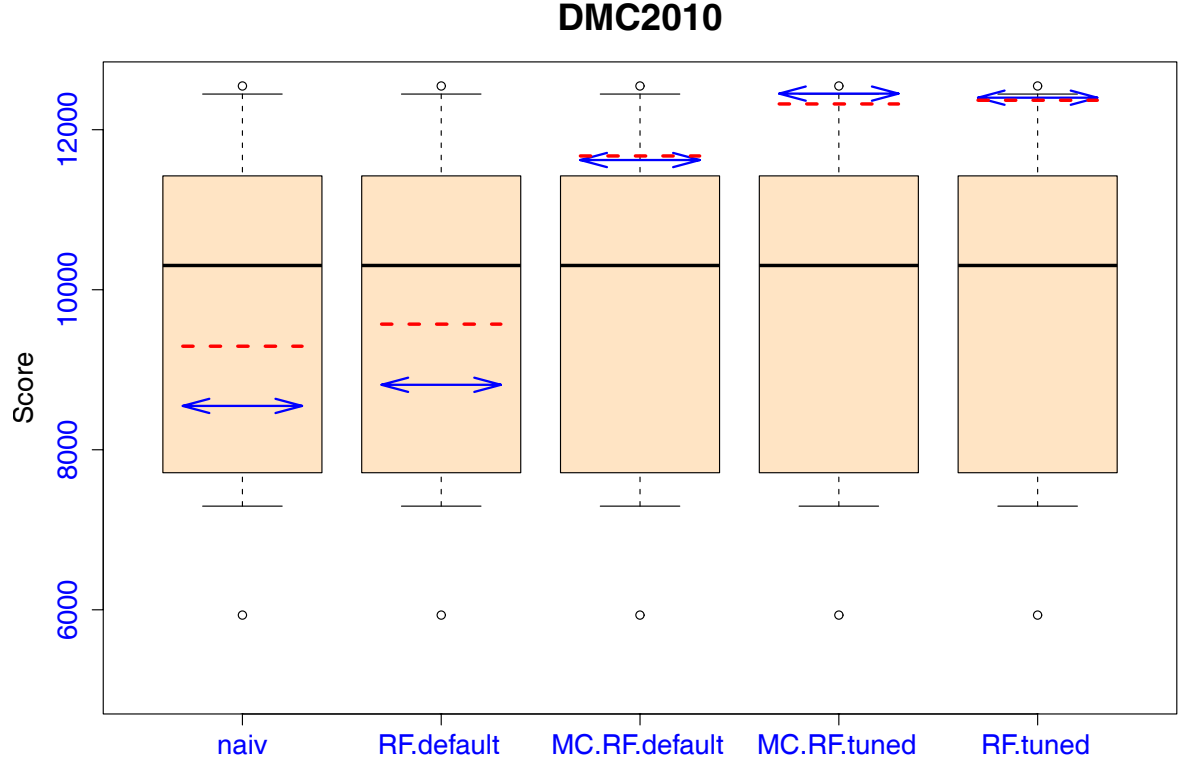


Figure 3: Our tuning results for the Data Mining Cup 2010 benchmark [KKF⁺11]. The red dashed lines show the score of our models on the training data (10-fold cross validation), the blue arrows show the score on independent test data. The boxplot shows for comparison the spread of score among the competition participants. It is clearly visible, that the models with default parameter settings (columns 2 and 3) do not produce results of high quality, while the tuned models (columns 4 and 5) have high-quality results close to the winner of the Data Mining Cup 2010 benchmark (MC = MetaCost, RF = Random Forest).

Data mining is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.

we know that data mining is tightly connected with pattern finding.

There are numerous (infinitely many) ways to construct features from input data. Besides the transformation for (spatial or temporal) continuous signals (Fourier transform, wavelet transforms, other), which we do not consider here, some examples are

1. N-tupel systems (see above)
2. GP (genetic programming) [Koz92]
3. PCA, GHA (variance, assumes linearity) [Oja82, San89]
4. SFA (slowness, assumes linearity in a higher dim space) [Wis98, WS02]
5. Kernel PCA [SSM98], KHA (nonlinear) [KFS03]

Example 1 and 2 are of the class "form many features 'at random' and find somehow the ones which are good for the task", while the remaining examples belong to the class "form features guided by some general principles and select the important ones".

We had some good experience with SFA on a gesture classification task which could not be successfully solved by a solely PCA-based feature construction, see Fig. 4 and Fig. 5 [KKH10].

4 Discussion

Some evidence has been presented that feature construction can be very relevant for success and fast convergence (both in time and in number of training examples) of machine learning problems. We have no proof that this is always the case, but it is my working hypothesis that:

"Self-configuring machine learning systems will require one or several modules for feature construction if they are to work robustly and flexible on a large variety of problems."

In order to strengthen this hypothesis, one should build machine learning systems containing explicit feature construction modules and compare their performance with systems where the construction of useful representations (features) is implicit and is delegated to the model-building step.

However, there is currently no generally accepted theory or framework of feature construction for arbitrary data mining or machine learning problems. Even worse, the problem is ill-defined right at the beginning, because there is no way to define exhaustively what a feature might look like given a set of N input variables. But the problem seems important and although I do not have a full solution right now, I would like to discuss

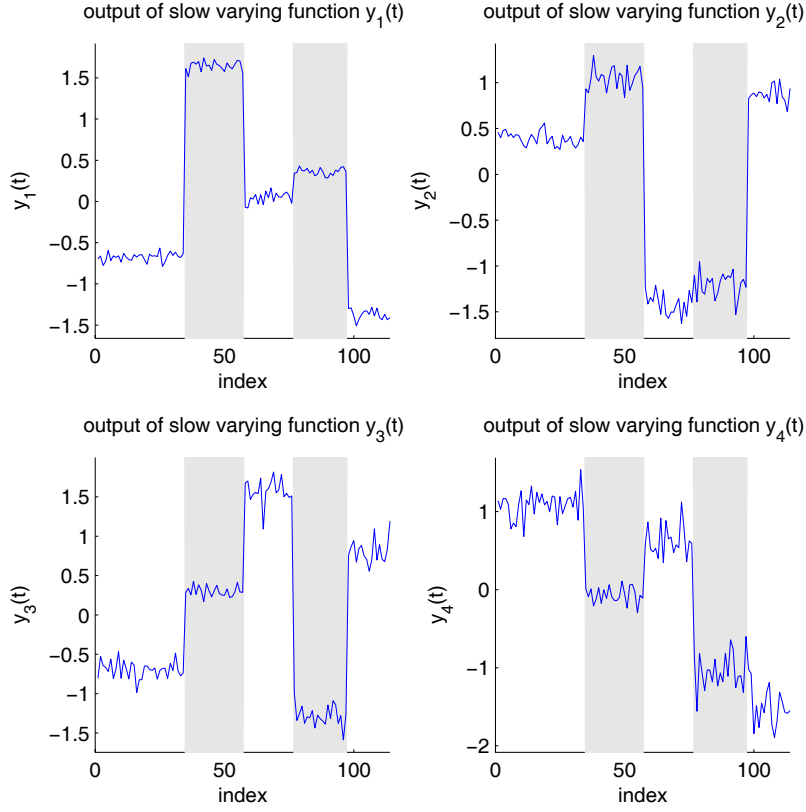


Figure 4: Gesture classification with SFA: Output of the first four SFA feature detectors y_1, y_2, y_3, y_4 for the different gesture classes (ordered along the x-axis) [KKH10].

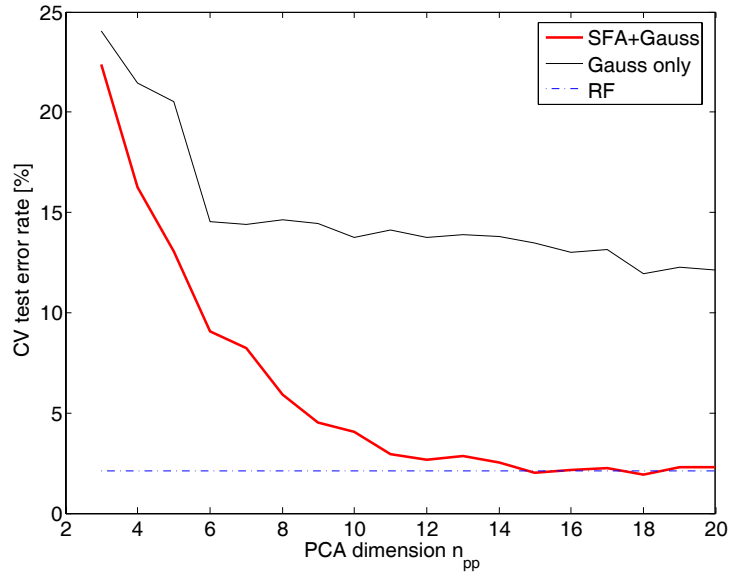


Figure 5: "SFA+Gauss": Classification error in gesture recognition with SFA working on n_{pp} preprocessed PCA features. "Gauss only": Gauss classifier on the PCA features alone. "RF": random forest classifier.

it more deeply here in the workshop and perhaps we can bring together a framework of interesting ideas for self-configuration in the context of feature construction. To start the discussion I have put together some questions:

- Which road to follow: random feature formation (try many and throw away many) or more careful feature construction driven by (complex) guiding principles?
- Is GP (Genetic Programming) a solution?
- Which general guiding principles can be used for feature formation / construction?
 - variance (PCA)
 - slowness (SFA)
 - information gain or other, guided by supervised information
 - ...
- Self-configuration: Given a large number of constructed features (from different approaches), are there *fast* procedures to decide which features are most promising for the given task?
- Are there procedures which might switch features on-line if the task is non-stationary?
- Which is the role of hierarchy? Could we initially use random feature formation, gain some experience from the environment and then combine the most promising features to form more complex features which are better adapted to the task?
- Can one translate the random N-tuple approach to real-valued variables?

5 Conclusion

Self-configuration in machine learning is still difficult, if problem scale gets bigger. Tuning should be part of the solution, but it is not the solution. Feature construction seems not solvable in the general sense, but it should not be neglected. Conjecture (to be proven):

Rich-enough feature sets lead to robustness, flexibility and fast learning in potentially changing environments.

Advances in this area will have numerous applications, especially in the area of online data mining (stream data mining) with often non-stationary environment conditions.

6 Acknowledgements

This work has been supported by the Bundesministerium für Bildung und Forschung (BMBF) under the grant SOMA (AiF FKZ 17N1009, "Ingenieurnachwuchs") and by the Cologne University of Applied Sciences under the research focus grant COSA.



References

- [BB10] Thomas Bartz-Beielstein. [SPOT: An R Package For Automatic and Interactive Tuning of Optimization Algorithms by Sequential Parameter Optimization](#). Technical Report arXiv:1006.4645. CIOP Technical Report 05-10, Cologne University of Applied Sciences, June 2010. [6](#)
- [BBLP10] Thomas Bartz-Beielstein, Christian Lasarczyk, and Mike Preuss. The sequential parameter optimization toolbox. In Bartz-Beielstein et al., editors, *Experimental Methods for the Analysis of Optimization Algorithms*, pages 337–360. Springer, Berlin, Heidelberg, New York, 2010. [6](#)
- [CDF99] K. Chellapilla and D. D. Fogel. Evolving neural networks to play checkers without expert knowledge. *IEEE Transactions on Neural Networks*, pages 1382–1391, 1999. [2](#)
- [FB⁺10] D. Ferrucci, E. Brown, et al. [Building Watson: An Overview of the DeepQA Project](#). *AI Magazine*, page 60, 2010. [3](#)
- [FPSS96] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54, 1996. [6](#)
- [Kan88] P. Kanerva. *Sparse Distributed Memory*. MIT Press, Cambridge, Mass., 1988. [4](#)
- [KBB09] Wolfgang Konen and Thomas Bartz-Beielstein. [Reinforcement learning for games: failures and successes](#). In *Proc. GECCO’09, Montreal*, pages 2641–2648, New York, NY, USA, 2009. ACM. [2](#), [3](#), [4](#)
- [KFS03] Kwang In Kim, Matthias O. Franz, and Bernhard Schölkopf. Kernel Hebbian algorithm for iterative kernel principal component analysis. Technical report, Max Planck Institute for Biological Cybernetics, 2003. [8](#)
- [KKF⁺11] Wolfgang Konen, Patrick Koch, Oliver Flasch, Thomas Bartz-Beielstein, Martina Frieze, and Boris Naujoks. [Tuned Data Mining: A Benchmark Study on Different Tuners](#). CIOP Technical Report 02-11, Cologne University of Applied Sciences, Feb 2011. [6](#), [7](#)
- [KKH10] P. Koch, W. Konen, and K. Hein. [Gesture Recognition on Few Training Data using Slow Feature Analysis and Parametric Bootstrap](#). In *2010 International Joint Conference on Neural Networks*, 2010. [8](#), [9](#)
- [Koz92] J.R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge MA, 1992. [8](#)
- [Luc08] Simon M. Lucas. Learning to play othello with n-tuple systems. *Australian Journal of Intelligent Information Processing*, pages 1–20, 2008. [2](#), [3](#), [4](#), [5](#)

- [Oja82] E. Oja. A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15:267–273, 1982. 8
- [PB98] J. Pollack and A. Blair. Co-evolution in the successful learning of backgammon strategy. *Machine Learning*, 32:225–244, 1998. 2
- [Ros62] F. Rosenblatt. *Principles of Neurodynamics*. Spartan Books, New York, 1962. 5
- [San89] T. D. Sanger. Optimal unsupervised learning in a single-layered linear feed-forward network. *Neural Networks*, 2:459–473, 1989. 8
- [SJK09] Marcin Szubert, Wojciech Jaskowski, and Krzysztof Krawiec. Coevolutionary temporal difference learning for othello. In *Proceedings of the 5th international conference on Computational Intelligence and Games, CIG’09*, pages 104–111, Piscataway, NJ, USA, 2009. IEEE Press. 4
- [SSM98] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, 10:1299–1319, July 1998. 8
- [Tes92] Gerald Tesauro. Practical issues in temporal difference learning. *Mach. Learning*, 8:257–277, 1992. 2
- [Tes94] Gerald Tesauro. TD-gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation*, 6:215–219, 1994. 2, 3
- [vEvW05] N.J. van Eck and M. van Wezel. Reinforcement learning and its application to othello. Technical Report EI 2005-47, Econometric Institute Report, Erasmus University Rotterdam, 2005. 4
- [Wis98] L. Wiskott. **Learning Invariance Manifolds**. In *Proc. of the 5th Joint Symp. on Neural Computation, May 16, San Diego, CA*, volume 8, pages 196–203, San Diego, CA, 1998. Univ. of California. 8
- [WS02] L. Wiskott and T. Sejnowski. **Slow Feature Analysis: Unsupervised Learning of Invariances**. *Neural Computation*, 14(4):715–770, 2002. 8