A Proofs

Proof of Proposition 5.6

We require the following result.

Lemma A.1. *Consider the RNN defined by* (5.3) *and* (5.4) *with* $s_i(1) = 0$, $i \in [1:3]$*. Suppose that there exist* $\epsilon_1, \epsilon_2 \in [0, 1/2)$ *such that the following conditions hold:*

1. *If* $s_1(t-1) \geq 1 - \epsilon_1$*, then* $s_1(t) \geq 1 - \epsilon_1$*,*
2. *If* $s_4(t-1) = s_4(t-2) = s_4(t-3) = 1$*, then* $s_1(t) \geq 1 - \epsilon_1$*,*
3. *If Conditions 1 and 2 do not hold, and* $s_1(t-1) \leq \epsilon_2$*, then* $s_1(t) \leq \epsilon_2$*.*

Then, the RNN correctly classifies any given binary string according to the L_4 *language.*

Proof. Consider an arbitrary input string. Denote its length by l. We consider two cases.

Case 1: The string does *not* include '000' as a substring. In this case, the If-part in Condition 2 is never satisfied. Since $s_1(1) = 0$, Condition 3 implies that $s_1(t) \leq \epsilon_2$, for $t = 1, 2, 3 \ldots$, hence, $s_1(l+1) \leq \epsilon_2$. Recalling that the network output is $f_{out} = s_1(l+1)$, yields $f_{out} \leq \epsilon_2$.
Case 2: The string contains a '000' substring, say, $I(m-2)I(m-1)I(m) =$'000', for some $m \leq l$. Then, according to Condition 2, $s_1(m+1) \geq 1 - \epsilon_1$. Condition 1 implies that $s_1(t) \geq 1 - \epsilon_1$ for $t = m+1, m+2, \ldots$, so $f_{out} \geq 1 - \epsilon_1$.

Summarizing, if the input string includes a '000' substring, then $f_{out} \geq 1 - \epsilon_1 > 1/2$, otherwise, $f_{out} \leq \epsilon_2 < 1/2$, so the RNN accepts (rejects) all the strings that do (not) belong to the language. □

We now prove Proposition 5.6 by showing that the RNN defined by (5.3), (5.4), and (5.5) indeed satisfies the three conditions in Lemma A.1. Note that using (5.5) yields

$$s_1(t) = \sigma\left(15.2 s_1(t-1) + 8.4 s_2(t-1) + 0.2 s_3(t-1) + 3 s_4(t-1) - 7.6\right), \tag{A.1}$$

E. Kolman, M. Margaliot: Knowledge-Based Neurocomputing, STUDFUZZ 234, pp. 83–86.
springerlink.com

whereas substituting (5.4) in (5.6) and (5.7) yields

$$s_4(t) \in \{-1, 1\},\ s_3(t) \in \{0.015, 0.98\}, \quad \text{and} \quad s_2(t) \in [0, 0.8]. \tag{A.2}$$

Suppose that

$$s_1(t-1) \geq 1 - \epsilon_1. \tag{A.3}$$

Since $\sigma(\cdot)$ is a monotonically increasing function, we can lower bound $s_1(t)$ by substituting the minimal value for the expression inside the brackets in (A.1). In this case, (A.1), (A.2), and (A.3) yield

$$\begin{aligned} s_1(t) &\geq \sigma\big(15.2(1-\epsilon_1) + 8.4 \cdot 0 + 0.2 \cdot 0.015 - 3 - 7.6\big) \\ &= \sigma(-15.2\epsilon_1 + 4.6). \end{aligned}$$

Thus, Condition 1 in Lemma A.1 holds if $\sigma(-15.2\epsilon_1 + 4.6) \geq 1 - \epsilon_1$. It is easy to verify that this indeed holds for any $\epsilon_1 \in (0.01, 0.219)$.

To analyze the second condition in Lemma A.1, suppose that $s_4(t-1) = s_4(t-2) = s_4(t-3) = 1$. It follows from (5.3) and (5.5) that $s_3(t-1) = \sigma(3.8) = 0.98$, and

$$\begin{aligned} s_2(t-1) &\geq \sigma\left(-0.2 \cdot 0.8 + 4.5\sigma(3.8) + 1.5 - 4.7\right) \\ &= 0.73. \end{aligned}$$

Substituting these values in (A.1) yields

$$\begin{aligned} s_1(t) &= \sigma\big(15.2 s_1(t-1) + 8.4 s_1(t-1) + 0.2 \cdot 0.98 + 3 - 7.6\big) \\ &\geq \sigma\big(15.2 s_1(t-1) + 8.4 \cdot 0.73 + 0.2 \cdot 0.98 + 3 - 7.6\big) \\ &\geq \sigma(1.72), \end{aligned}$$

where the last inequality follows from the fact that $s_1(t-1)$, being the output of a Logistic function, is non-negative. Thus, Condition 2 in Lemma A.1 will hold if $\sigma(1.72) \geq 1 - \epsilon_1$, or $\epsilon_1 \geq 0.152$.

To analyze Condition 3 of the lemma, suppose that $s_1(t-1) \leq \epsilon_2$. Then (A.1) yields

$$s_1(t) \leq \sigma\big(15.2\epsilon_2 + 8.4 s_2(t-1) + 0.2 s_3(t-1) + 3 s_4(t-1) - 7.6\big).$$

We can upper bound this by substituting the maximal values for the expression inside the brackets. Note, however, that Condition 3 does not apply when $s_4(t-1) = s_4(t-2) = s_4(t-3) = 1$ (as this case is covered by Condition 2). Under this constraint, applying (A.2) yields

$$\begin{aligned} s_1(t) &\leq \sigma\big(15.2\epsilon_2 + 8.4 \cdot 0.8 + 0.2 \cdot 0.98 - 3 - 7.6\big) \\ &= \sigma(15.2\epsilon_2 - 3.684). \end{aligned}$$

Thus, Condition 3 of Lemma A.1 will hold if $\sigma(15.2\epsilon_2 - 3.684) \leq \epsilon_2$, and it is easy to verify that this indeed holds for any $\epsilon_2 \in (0.06, 0.09)$.

Summarizing, for $\epsilon_1 \in [0.152, 0.219)$ and $\epsilon_2 \in (0.06, 0.09)$, the trained RNN satisfies all the conditions of Lemma A.1. This completes the proof of Proposition 5.6. □

Proof of Proposition 6.1

The proof is similar to the proof of Proposition 5.6, namely, we show that the RNN defined by (5.4), (6.4), and (6.5) indeed satisfies the three conditions in Lemma A.1. Note that using (6.5) yields

$$\begin{aligned} s_1(t) = \sigma\big(&5.04\alpha s_1(t-1) + 2\alpha s_2(t-1) + 2\alpha s_3(t-1) \\ &+ 0.76\alpha s_4(t-1) - 3.52\alpha\big), \end{aligned} \tag{A.4}$$

whereas substituting (5.4) in (6.1) and (6.2) yields

$$s_4(t) \in \{-1, 1\}, \quad \text{and} \quad s_2(t), s_3(t) \in \{0.12, 0.88\}. \tag{A.5}$$

Suppose that

$$s_1(t-1) \geq 1 - \epsilon_1. \tag{A.6}$$

Since $\sigma(\cdot)$ is a monotonically increasing function, we can lower bound $s_1(t)$ by substituting the minimal value for the expression inside the brackets in (A.4). In this case, (A.4), (A.5), and (A.6) yield

$$\begin{aligned} s_1(t) &\geq \sigma\big(5.04\alpha(1-\epsilon_1) + 0.24\alpha + 0.24\alpha - 0.76\alpha - 3.52\alpha\big) \\ &= \sigma(-5.04\alpha\epsilon_1 + 1.24\alpha). \end{aligned}$$

Thus, Condition 1 in Lemma A.1 holds if $\sigma(-5.04\alpha\epsilon_1 + 1.24\alpha) \geq 1 - \epsilon_1$. It is easy to verify that this indeed holds for any $\alpha \geq 5.7$, with $\epsilon_1 = 0.036$.

To analyze the second condition in Lemma A.1, suppose that $s_4(t-1) = s_4(t-2) = s_4(t-3) = 1$. It follows from (6.4) and (6.5) that $s_3(t-1) = \sigma(2) = 0.88$, and $s_2(t-1) = \sigma(5.26\sigma(2) - 2.62) = 0.88$. Substituting these values in (A.4) yields

$$\begin{aligned} s_1(t) &= \sigma\big(5.04\alpha s_1(t-1) + 1.76\alpha + 1.76\alpha + 0.76\alpha - 3.52\alpha\big) \\ &\geq \sigma(0.76\alpha), \end{aligned}$$

where the inequality follows from the fact that $s_1(t-1)$ is non-negative. Thus, Condition 2 in Lemma A.1 will hold if $\sigma(0.76\alpha) \geq 1 - \epsilon_1$, and it is easy to verify that this indeed holds for any $\alpha \geq 5.7$, with $\epsilon_1 = 0.036$.

To analyze Condition 3 in Lemma A.1, suppose that $s_1(t-1) \leq \epsilon_2$. Then (A.4) yields

$$s_1(t) \leq \sigma\big(5.04\alpha\epsilon_2 + 2\alpha s_2(t-1) + 2\alpha s_3(t-1) + 0.76\alpha s_4(t-1) - 3.52\alpha\big).$$

We can upper bound this by substituting the maximal values for the expression inside the brackets. Note, however, that Condition 3 does not apply when $s_4(t-1) = s_4(t-2) = s_4(t-3) = 1$ (as this case is covered by Condition 2). Under this constraint, applying (A.5) yields

$$\begin{aligned} s_1(t) &\leq \sigma\big(5.04\alpha\epsilon_2 + 1.76\alpha + 1.76\alpha - 0.76\alpha - 3.52\alpha\big) \\ &= \sigma(5.04\alpha\epsilon_2 - 0.76\alpha). \end{aligned}$$

Thus, Condition 3 of Lemma A.1 will hold if $\sigma(5.04\alpha\epsilon_2 - 0.76\alpha) \leq \epsilon_2$ and it is easy to verify that this indeed holds for any $\alpha \geq 5.7$, with $\epsilon_2 = 0.036$.

Summarizing, for $\alpha \geq 5.7$ the designed RNN satisfies all the conditions of Lemma A.1 for the specific values $\epsilon_1 = \epsilon_2 = 0.036$. This completes the proof of Proposition 6.1. $\square$

B Details of the LED Recognition Network

The 24-6-10 ANN was trained using MATLAB's Neural Network Toolbox. Parameter values were initialized using the "init" command with "net.layers{i}. initFcn" set to "initnw" (implementing the Nguyen-Widrow algorithm [121]). Training was performed using the "trainlm" command (Levenberg-Marquardt backprop), with "net.performFcn" set to "msereg" (that is, using the regularization factor $\sum w_{ij}^2$ [69]).

The parameters of the trained ANN are as follows.[1] The weights from the inputs to the hidden neurons are:

$$W = \begin{pmatrix} 0.23 & -0.04 & 0.45 & 0.30 & -0.17 & -0.52 & -0.14 & \ldots \\ -1.31 & -0.25 & -0.06 & 0.77 & 0.70 & 0.73 & 1.07 & \ldots \\ -1.09 & -2.05 & -1.86 & 1.58 & 0.60 & -0.15 & -0.63 & \ldots \\ 2.99 & 0.59 & -0.17 & 0.40 & -0.79 & 1.08 & -2.50 & \ldots \\ -0.57 & -2.02 & -0.25 & -0.65 & -0.09 & 2.08 & 2.90 & \ldots \\ -0.49 & 0.89 & 0.02 & -0.44 & -0.62 & -1.65 & 0.55 & \ldots \end{pmatrix}.$$

The values w_{ij}, $j \in [8:24]$, $i \in [1:6]$, are omitted since they satisfy $|w_{ij}| \leq 8.3E-5$. The hidden neurons biases are

$$\mathbf{b} = (0.33, -0.59, 1.63, -2.20, -1.90, 1.59)^T.$$

The weights from the hidden to the output neurons are:

$$C = \begin{pmatrix} -0.43 & -0.32 & 0.62 & 0.95 & 0.02 & -0.38 & -0.89 & 0.1 & 0.07 & 0.46 \\ -0.22 & -0.69 & -0.07 & 0.32 & 0.05 & 0.05 & 0.43 & -0.59 & 0.59 & 0.13 \\ -0.43 & 0.24 & 0.25 & 0.13 & 0.12 & 0.21 & 0.21 & -0.04 & -0.42 & -0.26 \\ -0.38 & -0.43 & -0.31 & 0.22 & 0.01 & 0.57 & 0.07 & 0.12 & -0.23 & 0.35 \\ 0.34 & -0.14 & -0.22 & 0.85 & -0.49 & 0.28 & -0.24 & -0.49 & -0.17 & 0.28 \\ 0.28 & -0.18 & 0.27 & -0.28 & 0.2 & 0.47 & -0.26 & -0.66 & -0.27 & 0.44 \end{pmatrix},$$

and the output neurons biases are:

$$\mathbf{\Phi} = -\left(0.74, 0.78, 1.13, 0.92, 0.89, 0.71, 0.45, 0.68, 0.74, 0.96\right)^T.$$

[1] All the numerical values were rounded to two decimal digits, without affecting the classification accuracy.

E. Kolman, M. Margaliot: Knowledge-Based Neurocomputing, STUDFUZZ 234, p. 87.
springerlink.com

References

[1] D. W. Aha, D. Kibler, and M. K. Albert, “Instance-based learning algorithms,” *Machine Learning*, vol. 6, pp. 37–66, 1991.

[2] R. Alquezar and A. Sanfeliu, “An algebraic framework to represent finite state machines in single-layer recurrent neural networks,” *Neural Computation*, vol. 7, pp. 931–949, 1995.

[3] H. C. Andersen, A. Lotfi, and L. C. Westphal, “Comments on “Functional equivalence between radial basis function networks and fuzzy inference systems”,” *IEEE Trans. Neural Networks*, vol. 9, pp. 1529–1531, 1998.

[4] R. Andrews, J. Diederich, and A. B. Tickle, “Survey and critique of techniques for extracting rules from trained artificial neural networks,” *Knowledge-Based Systems*, vol. 8, pp. 373–389, 1995.

[5] P. J. Angeline, G. M. Saunders, and J. B. Pollack, “An evolutionary algorithm that constructs recurrent neural networks,” *Neural Computation*, vol. 5, pp. 54–65, 1994.

[6] A. Asuncion and D. Newman, “UCI machine learning repository,” 2007. [Online]. Available: http://www.ics.uci.edu/~mlearn/MLRepository.html

[7] M. Ayoubi and R. Isermann, “Neuro-fuzzy systems for diagnosis,” *Fuzzy Sets and Systems*, vol. 89, pp. 289–307, 1997.

[8] M. Banerjee, S. Mitra, and S. K. Pal, “Rough fuzzy MLP: Knowledge encoding and classification,” *IEEE Trans. Neural Networks*, vol. 9, pp. 1203–1216, 1998.

[9] N. Barakat and J. Diedrich, “Eclectic rule-extraction from support vector machines,” *Int. J. Computational Intelligence*, vol. 2, pp. 59–62, 2005.

[10] P. Baranyi and Y. Yam, “Fuzzy rule base reduction,” in *Fuzzy IF-THEN Rules in Computational Intelligence: Theory and Applications*, D. Ruan and E. E. Kerre, Eds. Kluwer, 2000, pp. 135–160.

[11] R. Bellman, *Dynamic Programming.* Princeton University Press, 1957.

[12] J. M. Benitez, J. L. Castro, and I. Requena, “Are artificial neural networks black boxes?” *IEEE Trans. Neural Networks*, vol. 8, pp. 1156–1164, 1997.

[13] E. Bernado, X. Llora, and J. M. Garrel, “XCS and GALE: A comparative study of two learning classifier systems on data mining,” in *Advances in Learning Classifier Systems: Proc.* 4^{th} *Int. Workshop on Learning Classifier Systems (IWLCS’01)*, 2001, pp. 115–132.

[14] C. M. Bishop, *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

[15] A. D. Blair and J. B. Pollack, "Analysis of dynamical recognizers," *Neural Computation*, vol. 9, pp. 1127–1142, 1997.

[16] A. Blanco, M. Delgado, and M. C. Pegalajar, "Fuzzy automaton induction using neural network," *Int. J. Approximate Reasoning*, vol. 27, pp. 1–26, 2001.

[17] M. Boden, J. Wiles, B. Tonkes, and A. Blair, "Learning to predict a context-free language: analysis of dynamics in recurrent hidden units," in *Proc. 9^{th} Int. Conf. Artificial Neural Networks (ICANN'99)*, 1999, pp. 359–364.

[18] Z. Boger and H. Guterman, "Knowledge extraction from artificial neural network models," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics (SMC'97)*, 1997, pp. 3030–3035.

[19] B. K. Bose, "Expert system, fuzzy logic, and neural network applications in power electronics and motion control," *Proc. IEEE*, vol. 82, pp. 1303–1323, 1994.

[20] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Chapman & Hall, 1984, pp. 18–55.

[21] B. G. Buchanan, G. L. Sutherland, and E. A. Feigenbaum, "Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry," in *Machine Intelligence*, B. Meltzer and D. Michie, Eds. Edinburgh University Press, 1969, vol. 4, pp. 209–254.

[22] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 2000.

[23] J. L. Castro, L. D. Flores-Hidalgo, C. J. Mantas, and J. M. Puche, "Extraction of fuzzy rules from support vector machines," *Fuzzy Sets and Systems*, vol. 158, pp. 2057–2077, 2007.

[24] J. L. Castro, C. J. Mantas, and J. M. Benitez, "Interpretation of artificial neural networks by means of fuzzy rules," *IEEE Trans. Neural Networks*, vol. 13, pp. 101–116, 2002.

[25] N. Chomsky, "Three models for the description of language," *IRE Trans. Information Theory*, vol. IT-2, pp. 113–124, 1956.

[26] N. Chomsky and M. P. Schützenberger, "The algebraic theory of context-free languages," in *Computer Programming and Formal Systems*, P. Bradford and D. Hirschberg, Eds. North-Holland, 1963, pp. 118–161.

[27] I. Cloete, "Knowledge-based neurocomputing: Past, present, and future," in *Knowledge-Based Neurocomputing*, I. Cloete and J. M. Zurada, Eds. MIT Press, 2000, pp. 1–26.

[28] I. Cloete, "VL_1ANN: Transformation of rules to artificial neural networks," in *Knowledge-Based Neurocomputing*, I. Cloete and J. M. Zurada, Eds. MIT Press, 2000, pp. 207–216.

[29] I. Cloete and J. M. Zurada, Eds., *Knowledge-Based Neurocomputing*. MIT Press, 2000.

[30] M. W. Craven, "Extracting comprehensible models from trained neural networks," Ph.D. dissertation, Dept. of Computer Science, University of Wisconsin–Madison, 1996.

[31] Da Ruan, "Online experiments of controlling nuclear reactor power with fuzzy logic," in *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE'99)*, 1999, pp. 1712–1717.

[32] A. Darbari, "Rule extraction from trained ANN: A survey," Institute of Artificial intelligence, Dept. of Computer Science, Dresden Technology University, Tech. Rep., 2000.

[33] D. Dubois, H. T. Nguyen, H. Prade, and M. Sugeno, "Introduction: The real contribution of fuzzy systems," in *Fuzzy Systems: Modeling and Control*, H. T. Nguyen and M. Sugeno, Eds. Kluwer, 1998, pp. 1–17.

[34] W. Duch, "Therapeutic implications of computer models of brain activity for Alzheimer's disease," *J. Medical Informatics Technologies*, vol. 5, pp. 27–34, 2000.

[35] W. Duch, R. Adamczak, K. Crabczewski, M. Ishikawa, and H. Ueda, "Extraction of crisp logical rules using constrained backpropagation networks-comparison of two new approaches," in *Proc. European Symp. Artificial Neural Networks (ESANN'97)*, 1997, pp. 109–114.

[36] W. Duch, R. Adamczak, and K. Grabczewski, "Extraction of logical rules from neural networks," *Neural Processing Letters*, vol. 7, pp. 211–219, 1998.

[37] W. Duch, R. Adamczak, K. Grabczewski, and G. Zal, "Hybrid neural-global minimization method of logical rule extraction," *J. Advanced Computational Intelligence*, vol. 3, pp. 348–356, 1999.

[38] T. Etchells and P. J. G. Lisboa, "Orthogonal search-based rule extraction (OSRE) for trained neural networks: A practical and efficient approach," *IEEE Trans. Neural Networks*, vol. 17, pp. 374–384, 2006.

[39] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annual Eugenics*, vol. 7, pp. 179–188, 1936.

[40] J. A. Fodor and Z. W. Pylyshyn, "Connectionism and cognitive architecture: A critical analysis," in *Connections and Symbols*, S. Pinker and J. Mehler, Eds. MIT Press, 1988, pp. 3–72.

[41] P. Frasconi, M. Gori, M. Maggini, and G. Soda, "Representation of finite-state automata in recurrent radial basis function networks," *Machine Learning*, vol. 23, pp. 5–32, 1996.

[42] L.-M. Fu, "Knowledge-based connectionism for revising domain theories," *IEEE Trans. Systems, Man and Cybernetics*, vol. 23, pp. 173–182, 1993.

[43] L.-M. Fu, *Neural Networks in Computer Science*. McGraw-Hill, 1994.

[44] L.-M. Fu, "Learning capacity and sample complexity on expert networks," *IEEE Trans. Neural Networks*, vol. 7, pp. 1517 –1520, 1996.

[45] L.-M. Fu and L.-C. Fu, "Mapping rule-based systems into neural architecture," *Knowledge-Based Systems*, vol. 3, pp. 48–56, 1990.

[46] M. Fukumi, Y. Mitsukura, and N. Akamatsu, "Knowledge incorporation and rule extraction in neural networks," in *Proc. Int. Conf. Artificial Neural Networks (ICANN'01)*, ser. Lecture Notes in Computer Science, G. Dorffner, H. Bischof, and K. Hornik, Eds., vol. 2130. Springer, 2001, pp. 1248–1254.

[47] S. I. Gallant, "Connectionist expert systems," *Communications of the ACM*, vol. 31, pp. 152–169, 1988.

[48] T. Ganchev, D. K. Tasoulis, V. M. N., and N. Fakotakis, "Generalized locally recurrent probabilistic neural networks for text-independent speaker verification," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'04)*, vol. 1, 2004, pp. 41–44.

[49] C. L. Giles, S. Lawrence, and A. C. Tsoi, "Noisy time series prediction using recurrent neural networks and grammatical inference," *Machine Learning*, vol. 44, pp. 161–183, 2001.

[50] C. L. Giles, C. B. Miller, D. Chen, H. H. Chen, G. Z. Sun, and Y. C. Lee, "Learning and extracting finite state automata with second-order recurrent neural networks," *Neural Computation*, vol. 4, pp. 393–405, 1992.

[51] C. L. Giles, C. W. Omlin, and K. K. Thornber, "Equivalence in knowledge representation: automata, recurrent neural networks, and dynamical fuzzy systems,"

in *Recurrent Neural Networks: Design and Applications*, L. R. Medsker and L. C. Jain, Eds. CRC Press, 1999.

[52] M. Gori, M. Maggini, E. Martinelli, and G. Soda, "Inductive inference from noisy examples using the hybrid finite state filter," *IEEE Trans. Neural Networks*, vol. 9, pp. 571–575, 1998.

[53] M. W. Goudreau, C. L. Giles, S. T. Chakradhar, and D. Chen, "First-order versus second-order single-layer recurrent neural networks," *IEEE Trans. Neural Networks*, vol. 5, pp. 511–513, 1994.

[54] S. Guillaume, "Designing fuzzy inference systems from data: An interpretability-oriented review," *IEEE Trans. Fuzzy Systems*, vol. 9, pp. 426–443, 2001.

[55] A. Gupta and S. M. Lam, "Generalized analytic rule extraction for feedforward neural networks," *IEEE Trans. Knowledge and Data Engineering*, vol. 11, pp. 985–992, 1999.

[56] M. K. Hani, S. M. Nor, S. Hussein, and N. Elfadil, "Machine learning: The automation of knowledge acquisition using Kohonen self-organising map neural network," *Malysian J. Computer Science*, vol. 14, pp. 68–82, 2001.

[57] M. E. Hasselmo, "Runaway synaptic modification in models of cortex: Implications for Alzheimer's disease," *Neural Networks*, vol. 7, pp. 13–40, 1991.

[58] M. E. Hasselmo, "A computational model of the progression of Alzheimer's disease," *MD Computing*, vol. 14, pp. 181–191, 1997.

[59] M. H. Hassoun, *Fundamentals of Artificial Neural Networks.* MIT Press, 1995.

[60] D. O. Hebb, *The Organization of Behavior.* Wiley, 1949.

[61] M. Holena, "Piecewise-linear neural networks and their relationship to rule extraction from data," *Neural Computation*, vol. 18, pp. 2813–2853, 2006.

[62] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, 2nd ed. Addison-Wesley, 2001.

[63] D. Horn, N. Levy, and E. Ruppin, "Neuronal-based synaptic compensation: A computational study in Alzheimer's disease," *Neural Computation*, vol. 8, pp. 1227–1243, 1996.

[64] D. Horn, E. Ruppin, M. Usher, and M. Herrmann, "Neural network modeling of memory deterioration in Alzheimer's disease," *Neural Computation*, vol. 5, pp. 736–749, 1993.

[65] S. H. Huang, "Dimensionality reduction in automatic knowledge acquisition: A simple greedy search approach," *IEEE Trans. Knowledge and Data Engineering*, vol. 15, pp. 1364–1373, 2003.

[66] S. H. Huang and H. Xing, "Extract intelligible and concise fuzzy rules from neural networks," *Fuzzy Sets and Systems*, vol. 132, pp. 233–243, 2002.

[67] L. Hunter and T. Klein, "Finding relevant biomolecular features," in *Proc. 1^{st} Int. Conf. Intelligent Systems for Molecular Biology*, 1993, pp. 190–197.

[68] O. Intrator and N. Intrator, "Interpreting neural-network results: A simulation study," *Computational Statistics and Data Analysis*, vol. 37, pp. 373–393, 2001.

[69] M. Ishikawa, "Structural learning and rule discovery," in *Knowledge-Based Neurocomputing*, I. Cloete and J. M. Zurada, Eds. MIT Press, 2000, pp. 153–206.

[70] P. Jackson, *Introduction to Expert Systems.* Addison-Wesley, 1986.

[71] H. Jacobsson, "Rule extraction from recurrent neural networks: A taxonomy and review," *Neural Computation*, vol. 17, pp. 1223–1263, 2005.

[72] M. Jamshidi, A. Titly, L. A. Zadeh, and S. Boverie, Eds., *Applications of Fuzzy Logic: Towards High Machine Intelligence Quotient Systems.* Prentice Hall, 1997.

[73] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Systems, Man and Cybernetics*, vol. 23, pp. 665–685, 1993.

[74] J.-S. R. Jang and C.-T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems," *IEEE Trans. Neural Networks*, vol. 4, pp. 156–159, 1993.

[75] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.* Prentice-Hall, 1997.

[76] W. Jassem and W. Grygiel, "Off-line classification of Polish vowel spectra using artificial neural networks," *J. Int. Phonetic Association*, vol. 34, pp. 37–52, 2004.

[77] Y. Jin, "Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement," *IEEE Trans. Fuzzy Systems*, vol. 8, pp. 212–221, 2000.

[78] Y. Jin, W. Von Seelen, and B. Sendhoff, "On generating FC3 fuzzy rule systems from data using evolution strategies," *IEEE Trans. Systems, Man and Cybernetics*, vol. 29, pp. 829–845, 1999.

[79] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, pp. 263–292, 1979.

[80] A. Kandel and G. Langholz, Eds., *Hybrid Architectures for Intelligent Systems.* CRC Press, 1992.

[81] N. K. Kasabov, J. Kim, M. J. Watts, and A. R. Gray, "FuNN/2–a fuzzy neural network architecture for adaptive learning and knowledge acquisition," *Information Sciences*, vol. 101, pp. 155–175, 1997.

[82] T. Kavzoglu and P. M. Mather, "The use of backpropagating artificial neural networks in land cover classification," *Int. J. Remote Sensing*, vol. 24, pp. 4907–4938, 2003.

[83] E. Khan and F. Unal, "Recurrent fuzzy logic using neural networks," in *Advances in Fuzzy Logic, Neural Networks, and Genetic Algorithms*, T. Furuhashi, Ed. Springer-Verlag, 1995, pp. 48–55.

[84] S. C. Kleene, "Representation of events in nerve nets and finite automata," in *Automata Studies*, ser. Annals of Mathematics Studies, C. E. Shannon and J. McCarthy, Eds. Princeton University Press, 1956, vol. 34, pp. 3–41.

[85] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14^{th} Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, 1995, pp. 1137–1145.

[86] J. F. Kolen, "Fool's gold: Extracting finite state machines from recurrent network dynamics," in *Advances in Neural Information Processing Systems 6*, J. D. Cowan, G. Tesauro, and J. Alspector, Eds. Morgan Kaufmann, 1994, pp. 501–508.

[87] J. F. Kolen and J. B. Pollack, "The observer's paradox: Apparent computational complexity in physical systems," *J. Experimental and Theoretical Artificial Intelligence*, vol. 7, pp. 253–277, 1995.

[88] E. Kolman and M. Margaliot, "Extracting symbolic knowledge from recurrent neural networks–a fuzzy logic approach," *Fuzzy Sets and Systems*, to appear. [Online]. Available: www.eng.tau.ac.il/~michaelm

[89] E. Kolman and M. Margaliot, "A new approach to knowledge-based design of recurrent neural networks," *IEEE Trans. Neural Networks*, to appear. [Online]. Available: www.eng.tau.ac.il/~michaelm

[90] E. Kolman and M. Margaliot, "Are artificial neural networks white boxes?" *IEEE Trans. Neural Networks*, vol. 16, pp. 844–852, 2005.

[91] E. Kolman and M. Margaliot, "Knowledge extraction from neural networks using the all-permutations fuzzy rule base: The LED display recognition problem," *IEEE Trans. Neural Networks*, vol. 18, pp. 925–931, 2007.

[92] V. Kreinovich, C. Langrand, and H. T. Nguyen, "A statistical analysis for rule base reduction," in *Proc.* 2^{nd} *Int. Conf. Intelligent Technologies (InTech'2001)*, 2001, pp. 47–52.

[93] C.-M. Kuan and T. Liu, "Forecasting exchange rates using feedforward and recurrent neural networks," *J. Applied Econometrics*, vol. 10, pp. 347–364, 1995.

[94] K. J. Lang, "Random DFA's can be approximately learned from sparse uniform examples," in *Proc.* 5^{th} *ACM Workshop Computational Learning Theory*, 1992, pp. 45–52.

[95] Y. LeCun, "Learning processes in an asymmetric threshold network," in *Disordered Systems and Biological Organization*, E. Bienenstock, F. Fogelman-Soulie, and G. Weisbuch, Eds. Springer-Verlag, 1986, pp. 233–340.

[96] S.-W. Lee and H.-H. Song, "A new recurrent neural network architecture for visual pattern recognition," *IEEE Trans. Neural Networks*, vol. 8, pp. 331–340, 1997.

[97] G. Leng, T. M. McGinnity, and G. Prasad, "An approach for on-line extraction of fuzzy rules using a self-organizing fuzzy neural network," *Fuzzy Sets and Systems*, vol. 150, pp. 211–243, 2005.

[98] C.-T. Lin and C. S. G. Lin, "Neural-network-based fuzzy logic control and decision system," *IEEE Trans. Computers*, vol. 40, pp. 1320–1336, 1991.

[99] R. Linggard, D. J. Myers, and C. Nightingale, Eds., *Neural Networks for Vision, Speech, and Natural Language*. Chapman & Hall, 1992.

[100] X. Llora, "Genetic based machine learning using fine-grained parallelism for data mining," Ph.D. dissertation, Enginyeria i Arquitectura La Salle, Ramon Llull University, 2002.

[101] X. Llora and E. Goldberg, "Bounding the effect of noise in multiobjective learning classifier systems," *Evolutionary Computation*, vol. 11, pp. 279–298, 2003.

[102] F. Maire, "Rule-extraction by backpropagation of polyhedra," *Neural Networks*, vol. 12, pp. 717–725, 1999.

[103] E. H. Mamdani, "Applications of fuzzy algorithms for simple dynamic plant," *Proc. IEE*, vol. 121, pp. 1585–1588, 1974.

[104] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man-Machine Studies*, vol. 7, pp. 1–13, 1975.

[105] P. Manolios and R. Fanelly, "First order recurrent neural networks and deterministic finite state automata," *Neural Computation*, vol. 6, pp. 1155–1173, 1994.

[106] M. Margaliot and G. Langholz, "Hyperbolic optimal control and fuzzy control," *IEEE Trans. Systems, Man and Cybernetics*, vol. 29, pp. 1–10, 1999.

[107] M. Margaliot and G. Langholz, *New Approaches to Fuzzy Modeling and Control - Design and Analysis*. World Scientific, 2000.

[108] J. McCarthy, "What is artificial intelligence?" 2007. [Online]. Available: http://www-formal.stanford.edu/jmc/whatisai.pdf

[109] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence," 1956. [Online]. Available: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

[110] W. S. McCulloch and W. H. Pitts, "A logical calculus of the ideas imminent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–137, 1943.

[111] K. McGarry, S. Wermter, and J. Maclntyre, "Hybrid neural systems: From simple coupling to fully integrated neural networks," *Neural Computing Surveys*, vol. 2, pp. 62–93, 1999.

[112] L. R. Medsker and L. C. Jain, Eds., *Recurrent Neural Networks: Design and Applications.* CRC Press, 1999.

[113] C. B. Miller and C. L. Giles, "Experimental comparison of the effect of order in recurrent neural networks," *Int. J. Pattern Recognition and Artificial Intelligence*, vol. 7, pp. 849–872, 1993.

[114] D. A. Miller and J. M. Zurada, "A dynamical system perspective of structural learning with forgetting," *IEEE Trans. Neural Networks*, vol. 9, pp. 508–515, 1998.

[115] W. T. Miller, R. S. Sutton, and P. J. Werbos, Eds., *Neural Networks for Control.* MIT Press, 1990.

[116] M. L. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry.* MIT Press, 1969.

[117] S. Mitra, R. K. De, and S. K. Pal, "Knowledge-based fuzzy MLP for classification and rule generation," *IEEE Trans. Neural Networks*, vol. 8, pp. 1338–1350, 1997.

[118] S. Mitra and Y. Hayashi, "Neuro-fuzzy rule generation: Survey in soft computing framework," *IEEE Trans. Neural Networks*, vol. 11, pp. 748–768, 2000.

[119] S. Mitra and S. K. Pal, "Fuzzy multi-layer perceptron, inferencing and rule generation," *IEEE Trans. Neural Networks*, vol. 6, pp. 51–63, 1995.

[120] A. Murray, Ed., *Applications of Neural Networks.* Kluwer, 1995.

[121] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Proc.* 3^{rd} *Int. Joint Conf. Neural Networks*, 1990, pp. 21–26.

[122] V. Novak, "Are fuzzy sets a reasonable tool for modeling vague phenomena?" *Fuzzy Sets and Systems*, vol. 156, pp. 341–348, 2005.

[123] H. Nunez, C. Angulo, and A. Catala, "Rule extraction from support vector machines," in *Proc.* 10^{th} *European Symp. Artificial Neural Networks*, 2002, pp. 107–112.

[124] C. W. Omlin and C. L. Giles, "Extraction and insertion of symbolic information in recurrent neural networks," in *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration*, V. Honavar and L. Uhr, Eds. Academic Press, 1994, pp. 271–299.

[125] C. W. Omlin and C. L. Giles, "Constructing deterministic finite-state automata in recurrent neural networks," *J. ACM*, vol. 43, pp. 937–972, 1996.

[126] C. W. Omlin and C. L. Giles, "Extraction of rules from discrete-time recurrent neural networks," *Neural Networks*, vol. 9, pp. 41–52, 1996.

[127] C. W. Omlin and C. L. Giles, "Symbolic knowledge representation in recurrent neural networks: Insights from theoretical models of computation," in *Knowledge-Based Neurocomputing*, I. Cloete and J. M. Zurada, Eds. MIT Press, 2000, pp. 63–115.

[128] C. W. Omlin, K. K. Thornber, and C. L. Giles, "Fuzzy finite-state automata can be deterministically encoded into recurrent neural networks," *IEEE Trans. Fuzzy Systems*, vol. 6, pp. 76–89, 1998.

[129] R. P. Paiva and A. Dourado, "Interpretability and learning in neuro-fuzzy systems," *Fuzzy Sets and Systems*, vol. 147, pp. 17–38, 2004.

[130] D. B. Parker, "Learning logic," Center for Computational Research in Economics and Management Science, MIT, Tech. Rep., 1985.

[131] W. Pedrycz, *Fuzzy Control and Fuzzy Systems*, 2nd ed. Wiley, 1993.

[132] W. Pedrycz, A. Kandel, and Y.-Q. Zhang, "Neurofuzzy systems," in *Fuzzy Systems: Modeling and Control*, H. T. Nguyen and M. Sugeno, Eds. Kluwer, 1998, pp. 311–380.

[133] W. Pedrycz and A. F. Rocha, "Fuzzy-set based models of neurons and knowledge-based networks," *IEEE Trans. Fuzzy Systems*, vol. 1, pp. 254–266, 1993.

[134] J. B. Pollack, "Cascaded back-propagation on dynamic connectionist networks," in *Proc. 9^{th} Annual Conf. Cognitive Science Society*, 1987, pp. 391–404.

[135] J. B. Pollack, "The induction of dynamical recognizers," *Machine Learning*, vol. 7, pp. 227–252, 1991.

[136] I. Rashkovsky and M. Margaliot, "Nicholson's blowflies revisited: A fuzzy modeling approach," *Fuzzy Sets and Systems*, vol. 158, pp. 1083–1096, 2007.

[137] J. A. Reggia, E. Ruppin, and D. L. Glanzman, Eds., *Neural Modeling of Brain and Cognitive Disorders*. World Scientific, 1996.

[138] T. Robinson, M. Hochberg, and S. Renals, "The use of recurrent networks in continuous speech recognition," in *Automatic Speech and Speaker Recognition: Advanced Topics*, C. H. Lee, F. K. Soong, and K. K. Paliwal, Eds. Kluwer, 1996, pp. 233–258.

[139] P. Rodriguez, "Simple recurrent networks learn context-free and context-sensitive languages by counting," *Neural Computation*, vol. 13, pp. 2093–2118, 2001.

[140] P. Rodriguez, J. Wiles, and J. L. Elman, "A recurrent neural network that learns to count," *Connection Science*, vol. 11, pp. 5–40, 1999.

[141] F. Rosenblatt, "The perceptron: A perceiving and recognizing automaton," Project PARA, Cornell Aeronautical Laboratory, Tech. Rep., 1957.

[142] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, pp. 23–38, 1998.

[143] V. Rozin and M. Margaliot, "The fuzzy ant," *IEEE Computational Intelligence Magazine*, vol. 2, pp. 18–28, 2007.

[144] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representation by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D. E. Rumelhart and J. L. McClelland, Eds. MIT Press, 1986, pp. 318–362.

[145] D. E. Rumelhart and J. L. McClelland, Eds., *Parallel Distributed Processing*. MIT Press, 1986, vol. 2.

[146] E. Ruppin and J. Reggia, "A neural model of memory impairment in diffuse cerebral atrophy," *British J. Psychiatry*, vol. 166, pp. 19–28, 1995.

[147] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. Prentice Hall, 2003.

[148] K. Saito and R. Nakano, "Medical diagnostic expert system based on PDP model," in *Proc. 1^{st} IEEE Int. Conf. Neural Networks*, 1988, pp. 255–262.

[149] K. Saito and R. Nakano, "Rule extraction from facts and neural networks," in *Proc. 2^{nd} Int. Conf. Neural Networks (INNC90)*, 1990, pp. 379–382.

[150] K. Saito and R. Nakano, "Second-order learning algorithm with squared penalty term," *Neural Computation*, vol. 12, pp. 709–729, 2000.

[151] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM J. Research and Development*, vol. 3, pp. 211–229, 1959.

[152] E. Saund, "Dimensionality-reduction using connectionist networks," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, pp. 304–314, 1989.

[153] G. Schmitz, C. Aldrich, and F. Gouws, "Extraction of decision trees from artificial neural networks," in *Knowledge-Based Neurocomputing*, I. Cloete and J. M. Zurada, Eds. MIT Press, 2000, pp. 369–401.

[154] T. J. Sejnowski and C. R. Rosenberg, "Parallel networks that learn to pronounce English text," *Complex Systems*, vol. 1, pp. 145–168, 1987.

[155] S. Sestito and T. Dillon, "Knowledge acquisition of conjunctive rules using multilayered neural networks," *Int. J. Intelligent Systems*, vol. 8, pp. 779–805, 1993.

[156] R. Setiono, "Extracting rules from neural networks by pruning and hidden-unit splitting," *Neural Computation*, vol. 9, pp. 205–225, 1997.

[157] R. Setiono and M. Tanaka, "Neural network rule extraction and the LED display recognition problem," *IEEE Trans. Neural Networks*, to appear.

[158] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis.* Cambridge University Press, 2004.

[159] E. H. Shortliffe, *Computer-Based Medical Consultations: MYCIN.* Elsevier, 1976.

[160] H. T. Siegelmann, "Foundations of recurrent neural networks," Ph.D. dissertation, Dept. of Computer Science, Rutgers, 1993.

[161] P. Smolensky, "On the proper treatment of connectionism," *Behavioral and Brain Science*, vol. 11, pp. 1–23, 1988.

[162] S. Snyders and C. W. Omlin, "What inductive bias gives good neural network training performance?" in *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks*, 2000, pp. 445–450.

[163] S. Snyders and C. W. Omlin, "Inductive bias in recurrent neural networks," in *Proc.* 6^{th} *Int. Work-Conf. Artificial and Natural Neural Networks*, 2001, pp. 339–346.

[164] J. W. C. Sousa and U. Kaymak, *Fuzzy Decision Making in Modeling and Control.* World Scientific, 2002.

[165] T. Terano, K. Asai, and M. Sugeno, *Applied Fuzzy Systems.* AP Professional, 1994.

[166] A. B. Tickle, R. Andrews, M. Golea, and J. Diederich, "The truth will come to light: Directions and challenges in extracting the knowledge embedded within trained artificial neural networks," *IEEE Trans. Neural Networks*, vol. 9, pp. 1057–1068, 1998.

[167] A. B. Tickle, M. Golea, R. Hayward, and J. Diederich, "The truth is in there: Current issues in extracting rules from trained feedforward artificial neural networks," in *Proc.* 4^{th} *IEEE Int. Conf. Neural Networks*, 1997, pp. 2530–2534.

[168] A. B. Tickle, M. Orlowski, and J. Diederich, "DEDEC: A methodology for extracting rules from trained artificial neural networks," in *Rules and Networks*, R. Andrews and J. Diederich, Eds. QUT Publication, 1996, pp. 90–102.

[169] P. Tino and M. Koteles, "Extracting finite-state representation from recurrent neural networks trained on chaotic symbolic sequences," *Neural Computation*, vol. 10, pp. 284–302, 1999.

[170] P. Tino and V. Vojtec, "Extracting stochastic machines from recurrent neural networks trained on complex symbolic sequences," *Neural Network World*, vol. 8, pp. 517–530, 1998.

[171] M. Tomita, "Dynamic construction of finite-state automata from examples using hill-climbing," in *Proc.* 4^{th} *Ann. Conf. Cognitive Science*, 1982, pp. 105–108.

[172] G. G. Towell and J. W. Shavlik, "Extracting refined rules from knowledge-based neural networks," *Machine Learning*, vol. 13, pp. 71–101, 1993.

[173] G. G. Towell, J. W. Shavlik, and M. Noordenier, "Refinement of approximate domain theories by knowledge based neural network," in *Proc.* 8^{th} *Natl. Conf. Artificial Intelligence*, 1990, pp. 861–866.

[174] E. Tron and M. Margaliot, "Mathematical modeling of observed natural behavior: A fuzzy logic approach," *Fuzzy Sets and Systems*, vol. 146, pp. 437–450, 2004.
[175] E. Tron and M. Margaliot, "How does the Dendrocoleum lacteum orient to light? a fuzzy modeling approach," *Fuzzy Sets and Systems*, vol. 155, pp. 236–251, 2005.
[176] H. Tsukimoto, "Extracting rules from trained neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 377–389, 2000.
[177] J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," *J. Clinical Epidemiology*, vol. 49, pp. 1225–1231, 1996.
[178] R. L. Watrous and G. M. Kuhn, "Induction of finite-state languages using second-order recurrent networks," *Neural Computation*, vol. 4, pp. 406–414, 1992.
[179] Z. Wei, C. Wu, and L. Chen, "Symbolism and connectionism of artificial intelligence," in *Proc. IEEE Asia-Pacific Conf. on Circuits and Systems (IEEE APCCAS)*, 2000, pp. 364–366.
[180] P. J. Werbos, *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting.* Wiley, 1994.
[181] J. Wiles and J. L. Elman, "Learning to count without a counter: A case study of dynamics and activation landscapes in recurrent networks," in *Proc.* 17^{th} *Annual Conf. Cognitive Science Society*, 1995, pp. 482–487.
[182] R. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, vol. 8, pp. 229–256, 1992.
[183] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Computation*, vol. 1, pp. 270–280, 1989.
[184] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.
[185] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Sytems, Man and Cybernatics*, vol. 3, pp. 28–44, 1973.
[186] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Information Sciences*, vol. 30, pp. 199–249, 1975.
[187] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. Fuzzy Systems*, vol. 4, pp. 103–111, 1996.
[188] Z. Zeng, R. M. Goodman, and P. Smyth, "Learning finite state machines with self-clustering recurrent networks," *Neural Computation*, vol. 5, pp. 976–990, 1993.
[189] D. Zhang, X.-L. Bai, and K.-Y. Cai, "Extended neuro-fuzzy models of multilayer perceptrons," *Fuzzy Sets and Systems*, vol. 142, pp. 221–242, 2004.
[190] Z.-H. Zhou, "Rule extraction: Using neural networks or for neural networks," *J. Computer Science and Technology*, vol. 19, pp. 249–253, 2004.
[191] Z.-H. Zhou, S.-F. Chen, and Z.-Q. Chen, "A statistics based approach for extracting priority rules from trained neural networks," in *Proc.* 3^{rd} *Int. Joint Conf. Neural Networks (IJCNN'00)*, 2000, pp. 401–406.
[192] J. M. Zurada, *Introduction to Artificial Neural Systems.* West Publishing Company, 1992.

Index