Computational Intelligence
Lecture 21: Integrating Fuzzy Systems and Neural Networks

Farzaneh Abdollahi

Department of Electrical Engineering

Amirkabir University of Technology

Fall 2013

Reasons for Integrating Fuzzy and Neural Networks

- NN is able to learn and optimize.
- Fuzzy system is capable of reasoning, incorporating knowledge of experts easily.
- By using Fuzzy systems in NN:
  - High level human like rule of thinking and reasoning of fuzzy systems is brought into NN.
  - NN become more transparent.
  - Learning stage will be improved by defining smart initial values.
- By using NN in Fuzzy Systems:
  - Low level learning and computational power of NN will be used for Fuzzy systems
  - NN helps in automatic tuning of parameters and self adaptation.

Integrating Fuzzy Systems and NN

1. **Neural Fuzzy Systems**
   - It is using NN as a tool in fuzzy systems.
   - Fuzzy system is equipped with a kid of automatic tuning method without alerting tier functionality (fuzzification, defuzzification, inference engine)
   - NN is used for fuzzy set elicitation or/and realization of mapping fuzzy sets that is employed for fuzzy rules.
   - They are mostly used for control application

Integrating Fuzzy Systems and NN

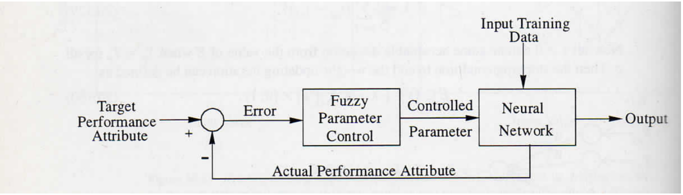
2. **Fuzzy Neural Networks**
   - It fuzzifies the conventional NN
   - A crisp neuron can become fuzzy.
   - Response of each neuron to its lower level activation signal can be a fuzzy relation rather than a sigmoid type
   - By considering knowledge of the system, the learning alg will be enhanced such as in activation fcn, weights, I/O data, NN model, and learning alg. can be fuzzified
   - These approaches are mostly used in pattern recognition

Integrating Fuzzy Systems and NN

3. **Fuzzy Neural Hybrid systems**
   - incorporating fuzzy and NN into hybrid system
   - They do their own job in serving different fcn.
   - They complement each other to achieve common goal
   - It is applicable for control as well as pattern recognition.

Fuzzy Control for Learning Parameter Adaptation [1]

- Objective: adaptively determine learning parameters in NN to improve performance of NN.
- A online fuzzy controller is used to adopt the learning parameter according to certain heuristic



Fuzzy Control of Back-Propagation Networks

- A general weight update rule in the back propagation alg. with momentum is $w(t+1) = -\alpha \nabla E(w(t)) + \eta w(t)$
- Objective: Automatically tuning learning parameters based on shape of the error surface in order to achieve faster convergence
- Consider
  - CE: change in error
  - CCE: change in CE, related to the acceleration of convergence
- IF-THEN rules
  - If CE is small with no sign changes in several consecutive time steps, THEN learning param. should be increased
  - IF CE changes the sign for several consecutive time steps, THEN learning param. should be reduced regardless of CCE
  - IF CE is very small and CCE is very small, with no sign change for several consecutive time steps, THEN learning param. and momentum gain should be increased.

Fuzzy Control of Back-Propagation Networks

- Sign changes param is $SC(t) = 1 - \|\frac{1}{2}[sgn(CE(t-1)) + sgn(E(t))]\|$
  - $sgn(.)$: sign fcn
  - $\frac{1}{2}$ is to ensure SC be 0 (no sign change) or 1 (one sign change)
- Let us consider the cumulative sum of SC in five steps: $CSC(t) = \sum_{m=t-4}^{t} SC(m)$
- We have two systems, one for adjusting η and one for α

Amirkabir University of Technology

Fuzzy Control of Back-Propagation Networks

- The Rules are

CCE \ CE	NB	NS	ZE	PS	PB
NB	NS	NS	NS	NS	NS
NS	NS	ZE	PS	ZE	NS
ZE	ZE	PS	PS	PS	ZE
PS	NS	ZE	PS	ZE	NS
PB	NS	NS	NS	NS	NS

(a)

CCE \ CE	NL	NL	ZE	PS	PL
NL	-0.01	-0.01	—	—	—
NS	-0.01	—	—	—	—
ZE	—	0.01	0.01	0.01	—
PS	—	—	—	—	-0.01
PL	—	—	—	-0.01	-0.01

(b)

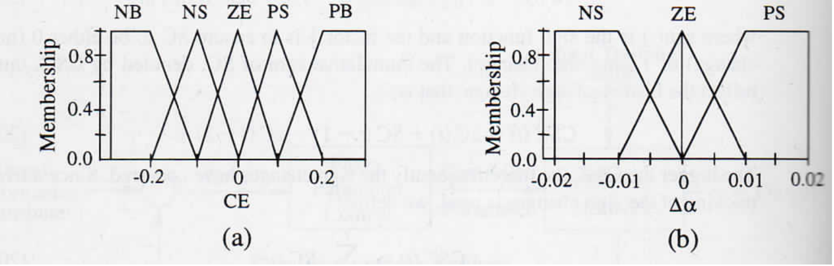
*(a) Decision table for the fuzzy control of learning rate α. Table contents represent the value of the fuzzy variable $\Delta\alpha$ for a given choice of values for CE and CCE, for $CSC(t) \leq 2$. (b) Decision table for the fuzzy control of momentum gain η. Table contents represent the value of the fuzzy variable $\Delta\eta$ for a given choice of values for CE and CCE, where — denotes no adaptation. The maximum value that η can take on is set to 1.

NB, Negative Big; NS, Negative Small; ZE, Zero; PS, Positive Small; PB, Positive Big.

- For instance: **IF CE** is NS and **CCE** is ZE, **THEN** $\Delta\alpha$ is PS
- Simulation results confirm dramatically faster convergence of this method comparing to regular BP alg

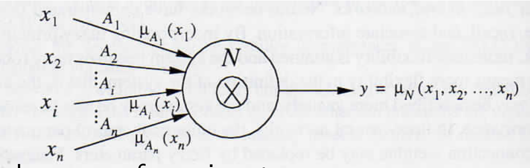
Fuzzy Control of Back-Propagation Networks

- Membership Function defined for fuzzy control



- (a): mem. fcn for CE and CCE, (b): mem. fcn for $\Delta\alpha$

Fuzzy Neuron



- This neuron has n nonfuzzy inputs $x_1, \ldots, x_n$
- The weights are fuzzy sets $A_i,\ 1 \leq i \leq n$
- i.e., the weighting operation is replaced by mem. fcn.
- Output is aggregation operation such as min, max, or any other t-norms of $\mu_{A_i}(x_i)$
- Output is a level of confidence.

$$\mu_N(x_1, \ldots, x_n) = \mu_{A_1}(x_1) \bigotimes \mu_{A_2}(x_2) \bigotimes \ldots \bigotimes \mu_{A_n}(x_n)$$

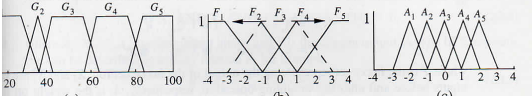
Example [2]

- Consider a fuzzy Controller with rules given in the Table, using the given mem. fcns

Δe = Change in error

e = error	G_1	G_2	G_3	G_4	G_5
F_1		1 A_1			
F_2		2 A_1		3 A_2	
F_3	4 A_2		5 A_3		6 A_4
F_4		7 A_4		8 A_5	
F_5				9 A_5	

*The rule number is given in the upper right-hand corner.



(a) (b) (c)

- (a): mem, fcn. for error, (b): mem. fcn. for change in error, (c): mem. fcn. for output

Example Cont'd

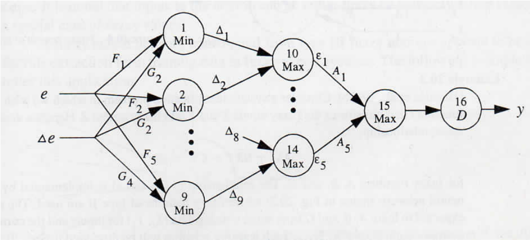
- The fuzzy controller accepts singleton inputs:
  - $e =$ error
  - $\Delta e =$ Change in error
- The fuzzy rules are:

$$
\begin{aligned}
\Delta_1 &= min(\mu_{F_1}(e), \mu_{G_2}(\Delta e)) \\
\Delta_2 &= min(\mu_{F_2}(e), \mu_{G_2}(\Delta e)) \\
\vdots &= \vdots \\
\Delta_9 &= min(\mu_{F_5}(e), \mu_{G_4}(\Delta e))
\end{aligned}
$$

- The outcome of the rules are combined as follows: $\epsilon_1 = \max(\Delta_1, \Delta_2)$, $\epsilon_2 = \max(\Delta_3, \Delta_4)$, $\epsilon_3 = \Delta_5$, $\epsilon_4 = \max(\Delta_6, \Delta_7)$, $\epsilon_5 = \max(\Delta_8, \Delta_9)$

Example Cont'd

- Each ϵ_k is assigned into its A_k, $1 \leq k \leq 5$..
- Then $A = \bigcup(e_k A_k)$, where union is max here.
  - y is defuzzified output from controller using center of gravity of A



- F_i, G_j, and A_p can be adjusted using the BP alg. as we discussed in the previous slides

Hybrid Neuro-Fuzzy [3]

- **A hybrid intelligent system**: is combination of at least two intelligent technologies.
- "Lotfi Zadeh is reputed to have said that
  - a good hybrid would be British Police, German Mechanics, French Cuisine, Swiss Banking, and Italian Love"
  - BUT British Cuisine, German Police, French Mechanics, Italian Banking and Swiss Love would be a bad one.
- ∴Our goal: selecting the right components for building a good hybrid system.
- A heterogeneous neuro-fuzzy system is hybrid system consisting of a neural network and a fuzzy system working as independent components.

Example: Diagnosing Myocardial Perfusion from Cardiac Images.

- A set of cardiac images, the clinical notes, and physicians interpretation are available.
- Diagnosis in modern cardiac medicine is based on the analysis of SPECT (Single Proton Emission Computed Tomography) images.
- By injecting a patient with radioactive tracer, two sets of SPECT images are obtained:
  - one is taken 10-15 minutes after the injection when the stress is greatest (stress images)
  - The other is taken 2-5 hours after the injection (rest images). Distribution of the radioactive tracer in the cardiac muscle is proportional to the muscles perfusion.

Example Cont'd

- cardiologist detects abnormalities in the heart function by comparing stress and rest images.
- Unfortunately a visual inspection of the SPECT images is highly subjective; physicians interpretations are therefore often inconsistent and susceptible to errors.
- For this study, 267 cardiac diagnostic cases is used, including two SPECT images (the stress image and the rest image), and each image is divided into 22 regions.
- The regions brightness, which in turn reflects perfusion inside this region, is expressed by an integer number between 0 and 100.
- Thus, each cardiac diagnostic case is represented by 44 continuous features and one binary feature that assigns an overall diagnosis normal or abnormal.

Example Cont'd

- The entire SPECT data set consists of 55 cases classified as: normal (positive examples) and 212 cases classified as abnormal (negative examples).
- The SPECT data set
- The training set has 40 positive and 40 negative examples.
- The test set is represented by 15 positive and 172 negative examples.
- Let us employ a BP NN to classify the SPECT images into normal and abnormal:
  - A NN with one hidden neuron
  - In 44 neurons in the input layer: each image is divided into 22 regions, input neurons.
  - Two output neurons: SPECT images are to be classified as either normal or abnormal.
  - A good generalisation is obtained with as little as 5 to 7 neurons in the hidden layer.

Example Cont'd

- After test
  - About 25 percent of normal cardiac diagnostic cases are misclassified as abnormal
  - over 35 percent of abnormal cases are misclassified as normal
  - the overall diagnostic error exceeds 33 percent.
- A neural network is only as good as the examples used to train it $\leadsto$ the training set may lack some important examples.
- In real clinical trials, the ratio between normal and abnormal SPECT images is very different, the misclassification of an abnormal cardiac case could lead to infinitely more serious consequences than the misclassification of a normal case.
- In order to achieve a small classification error for abnormal SPECT images, we might agree to have a relatively large error for normal images.

Example Cont'd

- The neural network produces two outputs.
  - The first output corresponds to the possibility that the SPECT image belongs to the class normal
  - The second to the possibility that the image belongs to the class abnormal.
- For example
  -

Example Cont'd

- The neural network produces two outputs.
  - The first output corresponds to the possibility that the SPECT image belongs to the class normal
  - The second to the possibility that the image belongs to the class abnormal.
- For example
  - the first (normal) output is 0.92 and the second (abnormal) is 0.16 $\rightsquigarrow$ the SPECT image is classified as normal, the risk of a heart attack for this case is low.

Example Cont'd

- The neural network produces two outputs.
  - The first output corresponds to the possibility that the SPECT image belongs to the class normal
  - The second to the possibility that the image belongs to the class abnormal.
- For example
  - the first (normal) output is 0.92 and the second (abnormal) is 0.16 $\rightsquigarrow$ the SPECT image is classified as normal, the risk of a heart attack for this case is low.
  - The normal output is 0.17, and the abnormal output is much higher, say 0.51$\rightsquigarrow$ the SPECT image is classified as abnormal, the risk of a heart attack in this case is rather high.

Example Cont'd

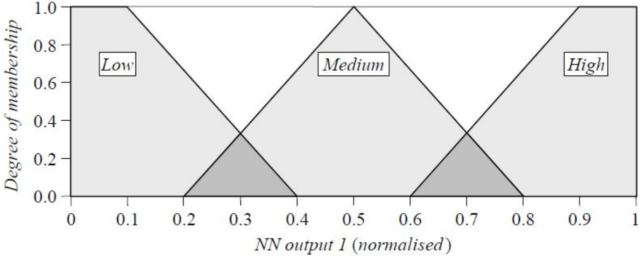
- The neural network produces two outputs.
  - The first output corresponds to the possibility that the SPECT image belongs to the class normal
  - The second to the possibility that the image belongs to the class abnormal.
- For example
  - the first (normal) output is 0.92 and the second (abnormal) is 0.16 $\rightsquigarrow$ the SPECT image is classified as normal, the risk of a heart attack for this case is low.
  - The normal output is 0.17, and the abnormal output is much higher, say 0.51$\rightsquigarrow$ the SPECT image is classified as abnormal, the risk of a heart attack in this case is rather high.
  - The two outputs are close: the normal output is 0.51 and the abnormal 0.49 we cannot confidently classify the image.

Example Cont'd

- Let us use fuzzy logic for decision-making in medical diagnosis.
- To build a fuzzy system, we first need to determine input and output variables, define fuzzy sets and construct fuzzy rules.
  - Two inputs: NN output 1 and NN output 2
  - One output: the risk of a heart attack.
  - The inputs are normalised to be within the range of [0, 1]
  - The output can vary between 0 and 100 percent

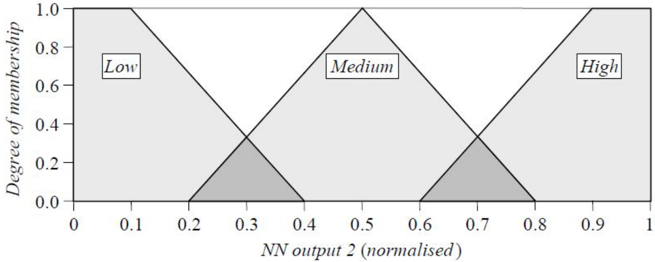
Example Cont'd

- Fuzzy sets of the neural network output normal



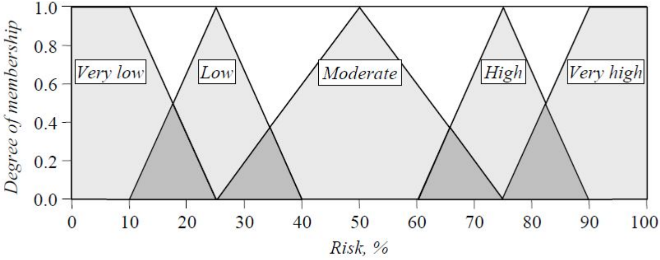
Example Cont'd

- Fuzzy sets of the neural network output abnormal



Example Cont'd

- Fuzzy sets of the linguistic variable Risk



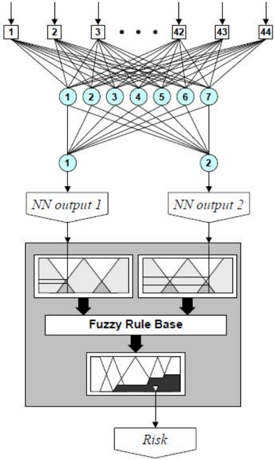
Example Cont'd

- Fuzzy rules for assessing the risk of a heart decease

1. If (NN_output1 is Low) and (NN_output2 is Low) then (Risk is Moderate)
2. If (NN_output1 is Low) and (NN_output2 is Medium) then (Risk is High)
3. If (NN_output1 is Low) and (NN_output2 is High) then (Risk is Very_high)
4. If (NN_output1 is Medium) and (NN_output2 is Low) then (Risk is Low)
5. If (NN_output1 is Medium) and (NN_output2 is Medium) then (Risk is Moderate)
6. If (NN_output1 is Medium) and (NN_output2 is High) then (Risk is High)
7. If (NN_output1 is High) and (NN_output2 is Low) then (Risk is Very_low)
8. If (NN_output1 is High) and (NN_output2 is Medium) then (Risk is Low)
9. If (NN_output1 is High) and (NN_output2 is High) then (Risk is Moderate)

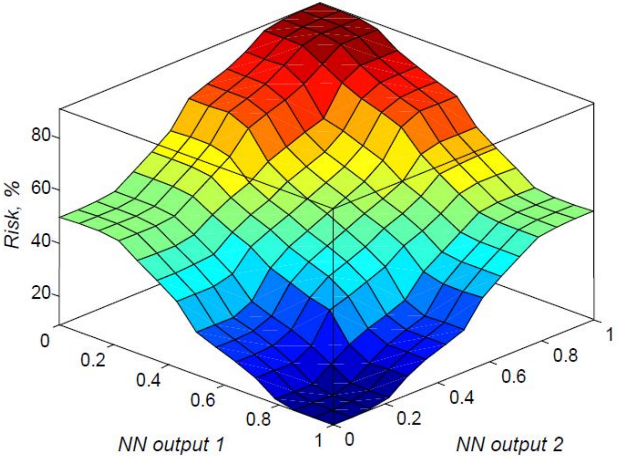
Example Cont'd

- Structure of the neuro-fuzzy system



Example Cont'd

- Three-dimensional plot for the fuzzy rule base



Example Cont'd

- The systems output is a crisp number that represents a patients risk of a heart attack.
- A cardiologist can now classify cardiac cases with greater certainty when the risk is quantified
- For instance, if the risk is low, say, smaller than 30 percent, the cardiac case can be classified as normal, but if the risk is high, say, greater than 50 percent, the case is classified as abnormal
- BUT cardiac cases with the risk between 30 and 50 percent cannot be classified as neither normal nor abormal rather, such cases are uncertain

Example Cont'd

- Let us employ cardiologist knowledge:
- Use the fact that in normal heart muscle, perfusion at maximum stress is usually higher than perfusion at rest:

Example Cont'd

- Let us employ cardiologist knowledge:
- Use the fact that in normal heart muscle, perfusion at maximum stress is usually higher than perfusion at rest:
  - If perfusion inside region i at stress is higher than perfusion inside the same region at rest, then the risk of a heart attack should be decreased.
  - If perfusion inside region i at stress is not higher than perfusion inside the same region at rest, then the risk of a heart attack should be increased.

Example Cont'd

- These heuristics can be implemented in the diagnostic system as follows:
  - **Step 1:** Present the neuro-fuzzy system with a cardiac case.

Example Cont'd

- These heuristics can be implemented in the diagnostic system as follows:
  - **Step 1:** Present the neuro-fuzzy system with a cardiac case.
  - **Step 2:** If the systems output is less than 30, classify the presented case as **normal** and then stop. If the output is greater than 50, classify the case as **abnormal** and stop. Otherwise, go to Step 3.

Example Cont'd

- These heuristics can be implemented in the diagnostic system as follows:
  - **Step 1:** Present the neuro-fuzzy system with a cardiac case.
  - **Step 2:** If the systems output is less than 30, classify the presented case as **normal** and then stop. If the output is greater than 50, classify the case as **abnormal** and stop. Otherwise, go to Step 3.
  - **Step 3:** For region 1, subtract perfusion at rest from perfusion at stress. If the result is positive, decrease the current risk by multiplying its value by 0.99. Otherwise, increase the risk by multiplying its value by 1.01. Repeat this procedure for all 22 regions and then run the neuro-fuzzy system.

Example Cont'd

- These heuristics can be implemented in the diagnostic system as follows:
  - **Step 1:** Present the neuro-fuzzy system with a cardiac case.
  - **Step 2:** If the systems output is less than 30, classify the presented case as **normal** and then stop. If the output is greater than 50, classify the case as **abnormal** and stop. Otherwise, go to Step 3.
  - **Step 3:** For region 1, subtract perfusion at rest from perfusion at stress. If the result is positive, decrease the current risk by multiplying its value by 0.99. Otherwise, increase the risk by multiplying its value by 1.01. Repeat this procedure for all 22 regions and then run the neuro-fuzzy system.
  - **Step 4:** If the new risk value is less than 30, classify the case as normal; if the risk is greater than 50, classify the case as **abnormal**; otherwise classify the case as **uncertain**.

Example Cont'd: Results

- Only 3 percent of abnormal cardiac cases are misclassified as normal.
- The system performance on normal cases has not been improved (over 30 percent of normal cases are still misclassified as abnormal),
- Up to 20 percent of the total number of cases are classified as uncertain
- $\therefore$ The neuro-fuzzy system can actually achieve even better results in classifying SPECT images than a cardiologist can.

J. J. Choi, P. Apabshahi, R. J. Marks, and T. P. Caudell, "Fuzzy parameter adaptation n neural systems," in *Int. Joint Conf. Neural Networks*, pp. 232–238, 1992.

Y. Hayashi, E. Czogala, and J. J. Buckley, "Fuzzy neural network controller," in *IEEE Int. Conferemce of Fuzzy Systems*, pp. 197–202, 1992.

M. Negnevitsky, "Key note speak of INTELLI 2013 conference." `http://www.iaria.org/conferences2013/filesINTELLI13/INTELLI%20Keynote_Negnevitsky.pdf`. Date of Acess: Dec, 2013.