ATCI Final Exam
Date: 29/5/2014
Duration: 12h to 14h

Answer the exam in three different pages: questions from 1 to 5, 6 to 10 and 11 to 15. Do not forget to write your name in ALL the sheets.

1. Which are the main characteristics of the Fuzzy Inductive Reasoning methodology? Which are the main differences with respect Neural Network approaches? Justify all the answers.

2. Explain the cooperation versus the competition problem when designing Fuzzy Rule Base Systems (FRBS)?

3. Which learning approach do Slave and Mogul algorithms use? Describe the main components of its structure.

4. Which are the two main parameters in FIR modeling that affect directly to the curse of dimensionality problem? Explain briefly.

5. Why an angular coding is better suited to represent the TSK rule consequent parameters in MOGUL? Explain briefly.

6. How are obtained the hidden-layer weights in Extreme Learning Machines? And the output-layer weights?

7. Until 2006, deep architectures were thought to have worse results than shallow ones. Which models, introduced in 2006, changed this idea? Can you briefly describe them?

8. The derivative of the log-likelihood of the data for Restricted Boltzmann Machines can be expressed as the sum of two terms. How are these two terms named? Can both terms be computed analytically? Can both terms be computed efficiently? Why?

9. Which are the two approximations made by Contrastive Divergence (CD-k)?

10. What is the pre-training step? What is it used for? In which models (explained in the course) has been used?

11. Describe the computation that is performed by a heterogeneous neuron model. You may include some mathematical equation if you want, but a qualitative description in natural language is enough.

12. When extending the Breeder Genetic Algorithm (BGA) to deal with data types other than real-valued data, which are the three things (or choices) we must define for each new data type?

13. In NARX networks, what is the difference between training in open-loop and training in closed-loop? Is adequate to use the same training algorithm in both cases? Why?

14. In a comparison between BPTT and RTRL algorithms to compute the error gradient for training Recurrent Neural Nets (RNNs), which are the advantages of BPTT over RTRL? and which are the advantages of RTRL over BPTT?

15. Which are the components of a memory block in the recurrent LSTM architecture? What is the role of each one of them?