NeuroSymbolic AI Assignment 2025/2026: Ordered List Reasoning using **Semantic Loss**

Professor: Francesco Ricca
Teaching Assistant: Salvatore Fiorentino

Problem Description

In this assignment you will work with a simple but fundamental reasoning task: given a list of five integers, determine the *sorted position* of each element.

Formally, for an input list:

$$[x_1, x_2, x_3, x_4, x_5]$$

the system must output:

$$[p_1, p_2, p_3, p_4, p_5]$$

where each $p_i \in \{0, 1, 2, 3, 4\}$ denotes the position of x_i in the ascending ordering of the list ($0 =$ smallest, $4 =$ largest).

Objective

Students must design one system capable of predicting the positional ranking of integers: A NeuroSymbolic model using **Semantic Loss** to enforce ordering constraints during learning.

Assignment Tasks

1. **Symbolic Component** Define the logic constraints expressing the ordering properties of the five integers.

2. **Neural Component** Build a neural network that takes 5 integer as input and outputs a distribution over the 5 possible positions for each integer.

3. **NeuroSymbolic Component: Semantic Loss Integration** Combine the neural model with your logical constraints using Semantic Loss. Your implementation should:

   - Convert the propositional constraints to a differentiable Semantic Loss term.
   - Integrate this loss with the neural network's training objective.
   - Ensure that the final loss function is of the form:

     $$\mathcal{L} = (1 - \lambda) \cdot \mathcal{L}_{\text{supervised}} + \lambda \cdot \mathcal{L}_{\text{semantic}}$$

     where $\lambda \in [0, 1]$ controls the influence of symbolic knowledge.

4. **Experiments Part** Conduct a full experimental evaluation including:

   - Training the Semantic Loss model using your logical constraints.
   - Evaluating the system on:
     - Accuracy (exact match)
     - Per-position accuracy
     - Precision / Recall / F1
     - **Training time** (mandatory)
   - Plotting learning curves.
   - Providing a discussion on:
     - Differences in accuracy
     - Differences in convergence speed
     - Examples where constraints correct neural mistakes
     - Typical misprediction patterns in both models
   - Ablation study with different values of λ (e.g. $0 - 0.5 - 1$)

   Additional suggestions:

   - Evaluate performance under multiple training set sizes (e.g. 100, 1k, 10k samples).

- Discuss whether Semantic Loss improves robustness with reduced data.

5. **Theory Part** Prepare a short slide deck (8–12 slides) covering:
   - The theoretical foundations of Semantic Loss.
   - The connection between propositional logic and differentiable loss terms.
   - Your architecture and training methodology.
   - Experimental results and comparisons.
   - Critical reflections on advantages and limitations of Semantic Loss.

Submission

You have two submission slots:

- Last lecture of the course: **19 December**
- First exam session (appello): **16 January**