CHIST-ERA Projects Seminar 2019

Topic: Life-long Learning Intelligent Systems (2017 call)

ALLIES, DELTA, LIHLITH

Eneko Agirre

(LIHLITH, University of the Basque Country)

Bucharest, April 4th, 2019

Programme co-funded by the EUROPEAN UNION

Introduction: Life-long Learning Intelligent Systems

- **Machine learning systems rely on human experts to:**
  - ✓ Select proper data
  - ✓ Tune the meta-parameters
  - ✓ Choose the training/development/evaluation sets
  - ✓ Choose the evaluation protocol

Our aim:

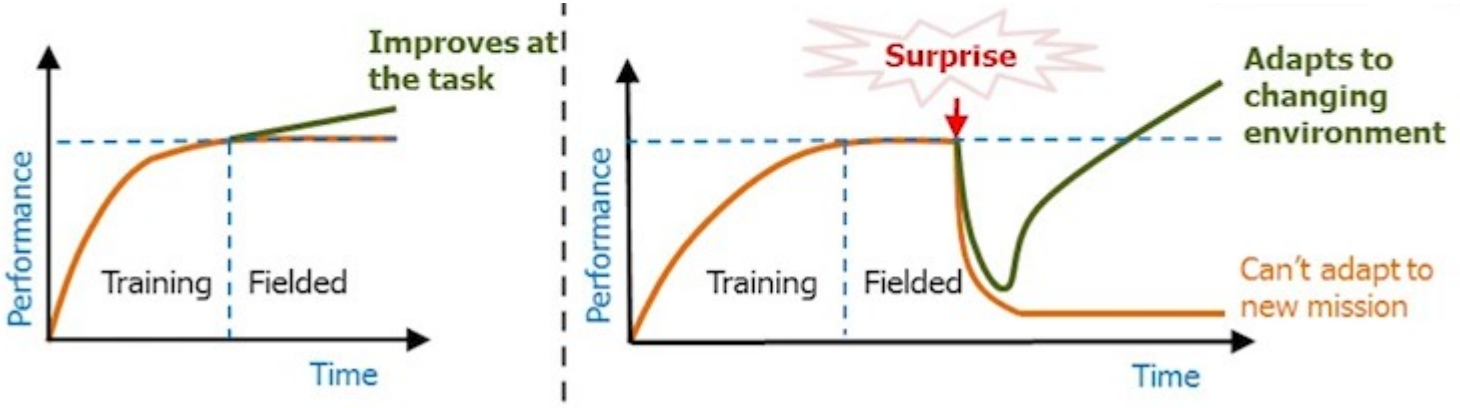
build autonomous intelligent systems sustaining their performance across time without machine learning experts

Objectives

- Enable the development of autonomous systems:

  « tuned once live forever »

- Initial system training requires developers (machine learning experts) and domain experts
- After deployment, autonomous systems may require the help of domain experts (but not developers / machine learning experts)
  ➥ **active / reinforcement learning**
- System should be able to adapt to unknown environments



Source: L2M DARPA program https://www.darpa.mil/staff/dr-hava-siegelmann

Different aspects of Life-long learning covered

- ✓ Classic machine learning (supervised / unsupervised)
- ✓ Deep learning
- ✓ Reinforcement learning
- ✓ Active learning
- ✓ Interactive learning

Key Challenges

Auto-evolution

Autonomous systems have to:

- ✓ Identify « new » information that has to be modeled
- ✓ Look for relevant adaptation data
- ✓ Balance the importance of « old » and « new » data in the adaptation process
- ✓ Collect unsupervised data (technical and legal issues)

Key Challenges

Auto-evaluation

- ✓ Enable automatic selection of evaluation data
- ✓ Enable automatic labeling of the data
- ✓ Find a metric that is measurable and related to the objective function
- ✓ Balance the ratio of « old » and « new » data within the auto-evaluation set

Key Challenges

Auto-evaluation

- ✓ How to reduce the cost of external supervision?
- ✓ Can the system ask questions on what it thinks is important?
- ✓ Should the user specify what is important?

Key Challenges

Evaluation of life-long-learning

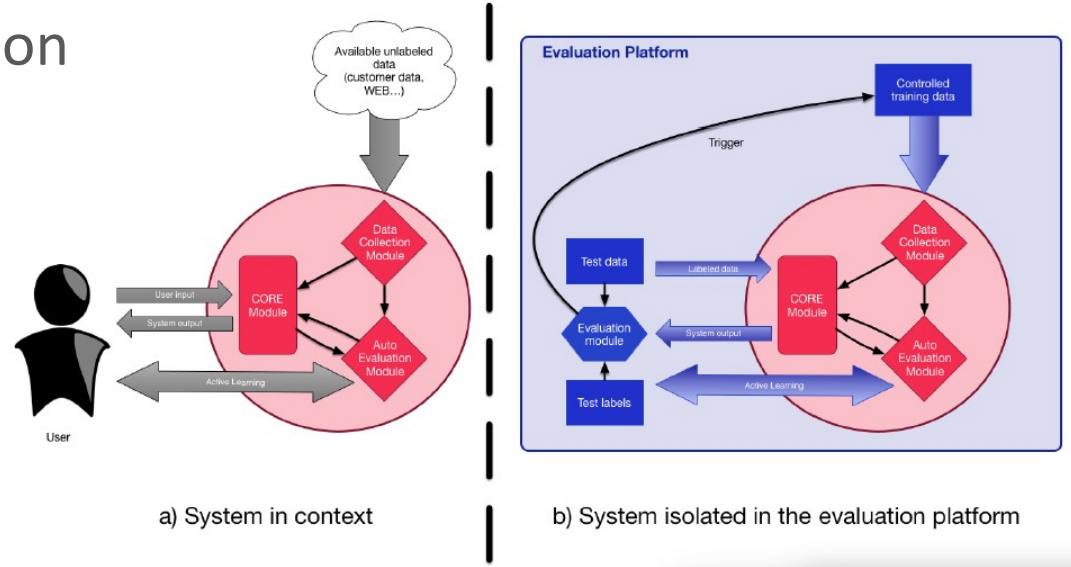
- ✓ No established protocols
- ✓ How to evaluate the task itself
- ✓ How to evaluate the task across time
- ✓ Difficulty of the evaluation depends on the task & on the definition of lifelong learning
- ✓ Generalization of the life-long learning evaluation across tasks

ALLIES
Autonomous Life-long Learning Intelligent Systems

- How to implement autonomous systems?
- How to evaluate autonomous systems?

Enable systems that don't require machine learning experts to sustain their performance across time

- Lifecycle of an autonomous system includes:
  - ✓ Initial training
  - ✓ unsupervised adaptation
  - ✓ active learning
  - ✓ interactive learning



ALLIES Achievements (1st year)

- Evaluation platform for lifelong learning V1
- Lifelong learning systems for
  - ✓ Machine Translation – V1
  - ✓ Speaker Diarization – V1
- Metrics, protocols and scenarios (to be released next week)
- Publications
  - ✓ 2 international journals
  - ✓ 9 internation conferences

DELTA
Dynamically Evolving Long-Term Autonomy

- Goal: Adapt reinforcement learning to the lifelong setting
  - Relax common stationarity assumptions
  - Develop state-of-the-art algorithms for **planning**, **exploration** and **task decomposition**
- Key ideas:
  - Reinforcement learning is a powerful technique for adaptive sequential decision making
  - Current algorithms are poor at handling changing environments and tasks
- Application domain: **microgrid management**

DELTA achievements (1st year)

❖**Y1 achievements**:

- Developed a planning algorithm for actions with decreasing reward
- Developed an exploration algorithm that adapts to changes in the environment
- Developed a task decomposition algorithm based on a partition of the state space
- Published a first version of the evaluation platform

LIHLITH: Learning to Interact with Humans by Lifelong Interaction with Humans

- Goal : Lifelong Learning methods that use dialogues to improve dialogue systems over time
  - Improve the quality on deployment across time
  - Lower the cost of deployment to new domains
- Key ideas
  - Systems designed to get feedback from user
  - In particular in LIHLITH, improve dialogue management, question answering, knowledge induction
- Development of evaluation protocols and benchmarks for reproducibility
- Open source and industrial valorisation

LIHLITH achievements (1st year)

- Evaluation of LL and dialogue systems:
  - Survey paper on dialogue evaluation, position papers
  - Framework for LL evaluation: task-based dialogue
  - Dataset: Accessing FAQ access using dialogue
- Technology ready to incorporate LL:
  - Knowledge induction system
  - Question answering system
  - Dialogue sytem
- Special session on LL and dialogue at IWSDS 2019
- Dissemination:
  - 3 research papers, 4 position papers (2 collaborative)

Sustainability / Valorization

- Evaluation protocols for life-long learning
  - ✓ ALLIES: ready to be released
  - ✓ DELTA: incorporated in the simulator (based on reward)
  - ✓ LIHLITH: proposal ready, under review at Semeval
- Scientific reproducible evaluation platform
  - ✓ ALLIES/DELTA: already released
  - ✓ LIHLITH: in preparation during 2019
- Benchmarking: open evaluation will happen during 2019-2020
  - ✓ ALLIES: diarization (Dec. 2019, IBERSPEECH/ALBAYZIN community)
  - ✓ ALLIES MT 2020 (WMT community, to be confirmed)
  - ✓ LIHLITH dialogue 2019 (Semeval, to be confirmed)
- Open-source software
- Industrial demonstrator (LIHLITH)

Topic Challenges and roadmap

- End of 2019:
  Our evaluation campaigns for LL will be the first of their kind

- 2019-2020: Involvement of different communities
  - Open campaigns on
    MT, diarization, dialogue, micro-grids
  - Attract the interest of LL community at large
- 2020: Feedback from participants and communities
  - More complex simulators
  - More changes introduced across time
  - Larger datasets with more domains

Role of the CHIST-ERA Support

Features of CHIST-ERA which are most helpful

- Having meetings with related projects
- Middle point between national projects and large European-wide projects
- Appropriate size (# partners, # projects) for the topic
- Light administrative overhead

Questions

Questions ?