Presenting ECAI 2025: relevant articles and some pointers

Valentin Cuzin-Rambaud[1,2]

[1] Laboratoire d'InfoRmatique en Image et Systèmes d'information (LIRIS)
[2] Université Claude Bernard Lyon 1 (UCBL1)

2 december 2025

Table Of Contents

Provably Robust AI? A Formal Methods Perspective

- Author: **Martha Kwiatkowska**
- **PREMAP**[1] is a framework that verifies *the preimage* for the guarantee of the output of NN.
- **MIBP-Cert**[2] a certification method based *on mixed-integer bilinear programming (MIBP)* that provides provable robustness against perturbed data by computing bounds around parameters of the model during training.
- **Strategyproof Reinforcement Learning from Human Feedback (RLHF)**[3] is a paper that studies the robustness of RL algorithms against multiple labelers with different preferences. They propose strategyproof algorithms that maximize *the social welfare of labelers*.
- **Safe POMDP Online Planning among Dynamic Agents via Adaptive Conformal Prediction**[4] aboard the problems of Online Planning among Dynamic Agents, with *ACP*.

[1] https://arxiv.org/pdf/2408.09262
[2] https://arxiv.org/pdf/2412.10186v2
[3] https://arxiv.org/pdf/2503.09561
[4] https://arxiv.org/pdf/2404.15557

Provably Robust AI? A Formal Methods Perspective

- Author: **Martha Kwiatkowska**
- **PREMAP**[1] is a framework that verifies *the preimage* for the guarantee of the output of NN.
- **MIBP-Cert**[2] a certification method based *on mixed-integer bilinear programming (MIBP)* that provides provable robustness against perturbed data by computing bounds around parameters of the model during training.
- **Strategyproof Reinforcement Learning from Human Feedback (RLHF)**[3] is a paper that studies the robustness of RL algorithms against multiple labelers with different preferences. They propose strategyproof algorithms that maximize *the social welfare of labelers*.
- **Safe POMDP Online Planning among Dynamic Agents via Adaptive Conformal Prediction**[4] aboard the problems of Online Planning among Dynamic Agents, with *ACP*.

[1] https://arxiv.org/pdf/2408.09262
[2] https://arxiv.org/pdf/2412.10186v2
[3] https://arxiv.org/pdf/2503.09561
[4] https://arxiv.org/pdf/2404.15557

LIRIS

Provably Robust AI? A Formal Methods Perspective

- Author: **Martha Kwiatkowska**
- **PREMAP**[1] is a framework that verifies *the preimage* for the guarantee of the output of NN.
- **MIBP-Cert**[2] a certification method based *on mixed-integer bilinear programming (MIBP)* that provides provable robustness against perturbed data by computing bounds around parameters of the model during training.
- **Strategyproof Reinforcement Learning from Human Feedback (RLHF)**[3] is a paper that studies the robustness of RL algorithms against multiple labelers with different preferences. They propose strategyproof algorithms that maximize *the social welfare of labelers.*
- **Safe POMDP Online Planning among Dynamic Agents via Adaptive Conformal Prediction**[4] aboard the problems of Online Planning among Dynamic Agents, with *ACP*.

[1] https://arxiv.org/pdf/2408.09262
[2] https://arxiv.org/pdf/2412.10186v2
[3] https://arxiv.org/pdf/2503.09561
[4] https://arxiv.org/pdf/2404.15557

Provably Robust AI? A Formal Methods Perspective

- Author: **Martha Kwiatkowska**
- **PREMAP**[1] is a framework that verifies *the preimage* for the guarantee of the output of NN.
- **MIBP-Cert**[2] a certification method based *on mixed-integer bilinear programming (MIBP)* that provides provable robustness against perturbed data by computing bounds around parameters of the model during training.
- **Strategyproof Reinforcement Learning from Human Feedback (RLHF)**[3] is a paper that studies the robustness of RL algorithms against multiple labelers with different preferences. They propose strategyproof algorithms that maximize *the social welfare of labelers*.
- **Safe POMDP Online Planning among Dynamic Agents via Adaptive Conformal Prediction**[4] aboard the problems of Online Planning among Dynamic Agents, with *ACP*.

[1] https://arxiv.org/pdf/2408.09262
[2] https://arxiv.org/pdf/2412.10186v2
[3] https://arxiv.org/pdf/2503.09561
[4] https://arxiv.org/pdf/2404.15557

Provably Robust AI? A Formal Methods Perspective

- Author: **Martha Kwiatkowska**
- **PREMAP**[1] is a framework that verifies *the preimage* for the guarantee of the output of NN.
- **MIBP-Cert**[2] a certification method based *on mixed-integer bilinear programming (MIBP)* that provides provable robustness against perturbed data by computing bounds around parameters of the model during training.
- **Strategyproof Reinforcement Learning from Human Feedback (RLHF)**[3] is a paper that studies the robustness of RL algorithms against multiple labelers with different preferences. They propose strategyproof algorithms that maximize *the social welfare of labelers*.
- **Safe POMDP Online Planning among Dynamic Agents via Adaptive Conformal Prediction**[4] aboard the problems of Online Planning among Dynamic Agents, with *ACP*.

[1] https://arxiv.org/pdf/2408.09262
[2] https://arxiv.org/pdf/2412.10186v2
[3] https://arxiv.org/pdf/2503.09561
[4] https://arxiv.org/pdf/2404.15557

From Self-Organization to Control: Steering Robot Swarms

- Author: **Marco Dorigo**
- Swarmanoid[5] old project focuses on Heterogeneous Collaboration Robot: Hand bot, Foot bot, Eye bot.
- Self-organizing nervous systems for robot swarms[6] more recent project focus on Hierachical Organisation

[5] https://www.swarmanoid.org/index.php.html

[6] https://dipot.ulb.ac.be/dspace/bitstream/2013/384796/3/2401.13103v1.pdf

From Self-Organization to Control: Steering Robot Swarms

- Author: **Marco Dorigo**
- Swarmanoid[5] old project focuses on Heterogeneous Collaboration Robot: Hand bot, Foot bot, Eye bot.
- Self-organizing nervous systems for robot swarms[6] more recent project focus on Hierachical Organisation

[5] https://www.swarmanoid.org/index.php.html

[6] https://dipot.ulb.ac.be/dspace/bitstream/2013/384796/3/2401.13103v1.pdf

From Self-Organization to Control: Steering Robot Swarms

- Author: **Marco Dorigo**
- Swarmanoid[5] old project focuses on Heterogeneous Collaboration Robot: Hand bot, Foot bot, Eye bot.
- Self-organizing nervous systems for robot swarms[6] more recent project focus on Hierachical Organisation

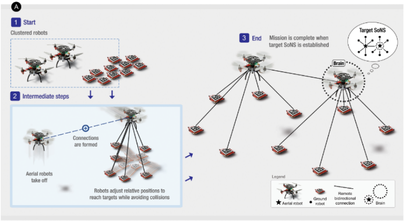

Figure: SonS: Establishing self-organized hierarchy

[5] https://www.swarmanoid.org/index.php.html
[6] https://dipot.ulb.ac.be/dspace/bitstream/2013/384796/3/2401.13103v1.pdf

Trustworthy Planning Agents for Collaborative Reasoning and Multimodal Generation

- Author: **Mohit Bansal**
- MAGDI: Structured Distillation of *Multi-Agent Interaction Graphs* Improves Reasoning in Smaller Language Models[7]
- MAGICORE: Multi-Agent, Iterative, Coarse-to-Fine Refinement for Reasoning[8]
- MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation[9]

[7] https://arxiv.org/pdf/2402.01620
[8] https://aclanthology.org/2025.emnlp-main.1660.pdf
[9] https://arxiv.org/pdf/2506.17113

Trustworthy Planning Agents for Collaborative Reasoning and Multimodal Generation

- Author: **Mohit Bansal**
- MAGDI: Structured Distillation of *Multi-Agent Interaction Graphs* Improves Reasoning in Smaller Language Models[7]
- MAGICORE: Multi-Agent, Iterative, Coarse-to-Fine Refinement for Reasoning[8]
- MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation[9]

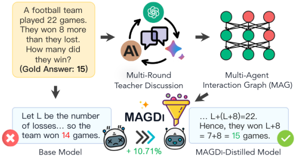

Figure: MAGDi: Overview

[7] https://arxiv.org/pdf/2402.01620
[8] https://aclanthology.org/2025.emnlp-main.1660.pdf
[9] https://arxiv.org/pdf/2506.17113

Trustworthy Planning Agents for Collaborative Reasoning and Multimodal Generation

- Author: **Mohit Bansal**
- MAGDI: Structured Distillation of *Multi-Agent Interaction Graphs* Improves Reasoning in Smaller Language Models[7]
- MAGICORE: Multi-Agent, Iterative, Coarse-to-Fine Refinement for Reasoning[8]
- MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation[9]

[7] https://arxiv.org/pdf/2402.01620
[8] https://aclanthology.org/2025.emnlp-main.1660.pdf
[9] https://arxiv.org/pdf/2506.17113

Trustworthy Planning Agents for Collaborative Reasoning and Multimodal Generation

- Author: **Mohit Bansal**
- MAGDI: Structured Distillation of *Multi-Agent Interaction Graphs* Improves Reasoning in Smaller Language Models[7]
- MAGICORE: Multi-Agent, Iterative, Coarse-to-Fine Refinement for Reasoning[8]
- MEXA: Towards General Multimodal Reasoning with Dynamic Multi-Expert Aggregation[9]

[7] https://arxiv.org/pdf/2402.01620
[8] https://aclanthology.org/2025.emnlp-main.1660.pdf
[9] https://arxiv.org/pdf/2506.17113

Many Facets of Proportionality in Multiwinner Voting

- Author: **Edith Elkind**
- *Justified Representation* in Approval-Based Committee Voting[10]
- Properties of *multiwinner voting rules*[11]

[10] https://ojs.aaai.org/index.php/AAAI/article/download/9324/9183

[11] https://link.springer.com/content/pdf/10.1007/s00355-017-1026-z.pdf

Many Facets of Proportionality in Multiwinner Voting

- Author: **Edith Elkind**
- *Justified Representation* in Approval-Based Committee Voting[10]
- Properties of *multiwinner voting rules*[11]

[10] https://ojs.aaai.org/index.php/AAAI/article/download/9324/9183
[11] https://link.springer.com/content/pdf/10.1007/s00355-017-1026-z.pdf

Many Facets of Proportionality in Multiwinner Voting

- Author: **Edith Elkind**
- *Justified Representation* in Approval-Based Committee Voting[10]
- Properties of *multiwinner voting rules*[11]

[10] https://ojs.aaai.org/index.php/AAAI/article/download/9324/9183

[11] https://link.springer.com/content/pdf/10.1007/s00355-017-1026-z.pdf

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

PAIS: Multi-Agent Pathfinding in Automatic Warehouse

- Author: **Roni Stearn**
- Deep Learning methods don't yet apply in industry to the MAPF problem for the following Challenges:
  - Real-Time Constraint => Fast Algorithms and Guarantee
  - Scale to Many Agents
  - LifeLong MAPF
- LACAM*[12] best theoretical solution, but not really applied (LifeLong and deployment industries Constraints).
- He has seen some warehouses where they use old deterministic solutions.
- Global critic of research orientation in the field, asking good questions about Learning, novel strategies.

[12] https://arxiv.org/pdf/2305.03632

Best Papers

- "FAIRGAME: a Framework for AI Agents Bias Recognition using Game Theory"[13]
- "Conditional Dominance Analysis for Classical Planning"[14]
- "Analysing Temporal Reasoning in Description Logics Using Formal Grammars"[15]
- PAIS best paper: "Optimizing Parcels Sorting Through Reinforcement Learning for Intralogistics"[16]

[13] https://arxiv.org/pdf/2504.14325

[14] https://openreview.net/pdf?id=UN8OUhaRzp

[15] https://arxiv.org/pdf/2508.00575

[16] https://boa.unimib.it/retrieve/8985846b-7938-4c90-9bb6-057eef713971/Roveda 20et 20al-2025-Frontiers 20in 20Artificial 20Intelligence 20and 20Applications-VoR.pdf

Workshops

- LLAIS: "VIRAL: Vision-grounded Integration for Reward design And Learning"[17]
- AIC: "Action matters for object representation learning from sequences"[18]
- SLLM: "Boosting Accuracy and Efficiency of Budget Forcing in LLMs via Reinforcement Learning for Mathematical Reasoning"[19]
- Multi-Objectives: "Multi-Objective Categorical Deep Q-Networks"[20]
- AIFin: Talk by Marcello Restelli "Reinforcement Learning in Algorithmic Trading" [21]

[17] https://llias2025.web.app/: https://arxiv.org/pdf/2505.22092
[18] https://codesign-lab.org/aic25/: https://hal.science/hal-05324845v1/file/Article_ECAI-AIC.pdf
[19] https://sites.google.com/view/sllm2025/home: https://arxiv.org/pdf/2510.21398
[20] https://modem2025.vub.ac.be/: https://modem2025.vub.ac.be/papers/MODeM_2025_paper_17.pdf
[21] https://sites.google.com/view/ECAI25-ai-fin

Workshops

- LLAIS: "VIRAL: Vision-grounded Integration for Reward design And Learning"[17]
- AIC: "Action matters for object representation learning from sequences"[18]
- SLLM: "Boosting Accuracy and Efficiency of Budget Forcing in LLMs via Reinforcement Learning for Mathematical Reasoning"[19]
- Multi-Objectives: "Multi-Objective Categorical Deep Q-Networks"[20]
- AIFin: Talk by Marcello Restelli "Reinforcement Learning in Algorithmic Trading" [21]

[17] https://llias2025.web.app/: https://arxiv.org/pdf/2505.22092
[18] https://codesign-lab.org/aic25/: https://hal.science/hal-05324845v1/file/Article_ECAI-AIC.pdf
[19] https://sites.google.com/view/sllm2025/home: https://arxiv.org/pdf/2510.21398
[20] https://modem2025.vub.ac.be/: https://modem2025.vub.ac.be/papers/MODeM_2025_paper_17.pdf
[21] https://sites.google.com/view/ECAI25-ai-fin

Workshops

- LLAIS: "VIRAL: Vision-grounded Integration for Reward design And Learning"[17]
- AIC: "Action matters for object representation learning from sequences"[18]
- SLLM: "Boosting Accuracy and Efficiency of Budget Forcing in LLMs via Reinforcement Learning for Mathematical Reasoning"[19]
- Multi-Objectives: "Multi-Objective Categorical Deep Q-Networks"[20]
- AIFin: Talk by Marcello Restelli "Reinforcement Learning in Algorithmic Trading" [21]

[17] https://llias2025.web.app/: https://arxiv.org/pdf/2505.22092
[18] https://codesign-lab.org/aic25/: https://hal.science/hal-05324845v1/file/Article_ECAI-AIC.pdf
[19] https://sites.google.com/view/sllm2025/home: https://arxiv.org/pdf/2510.21398
[20] https://modem2025.vub.ac.be/: https://modem2025.vub.ac.be/papers/MODeM_2025_paper_17.pdf
[21] https://sites.google.com/view/ECAI25-ai-fin

Workshops

- LLAIS: "VIRAL: Vision-grounded Integration for Reward design And Learning"[17]
- AIC: "Action matters for object representation learning from sequences"[18]
- SLLM: "Boosting Accuracy and Efficiency of Budget Forcing in LLMs via Reinforcement Learning for Mathematical Reasoning"[19]
- Multi-Objectives: "Multi-Objective Categorical Deep Q-Networks"[20]
- AIFin: Talk by Marcello Restelli "Reinforcement Learning in Algorithmic Trading" [21]

[17] https://llias2025.web.app/: https://arxiv.org/pdf/2505.22092
[18] https://codesign-lab.org/aic25/: https://hal.science/hal-05324845v1/file/Article_ECAI-AIC.pdf
[19] https://sites.google.com/view/sllm2025/home: https://arxiv.org/pdf/2510.21398
[20] https://modem2025.vub.ac.be/: https://modem2025.vub.ac.be/papers/MODeM_2025_paper_17.pdf
[21] https://sites.google.com/view/ECAI25-ai-fin

Workshops

- LLAIS: "VIRAL: Vision-grounded Integration for Reward design And Learning"[17]
- AIC: "Action matters for object representation learning from sequences"[18]
- SLLM: "Boosting Accuracy and Efficiency of Budget Forcing in LLMs via Reinforcement Learning for Mathematical Reasoning"[19]
- Multi-Objectives: "Multi-Objective Categorical Deep Q-Networks"[20]
- AIFin: Talk by Marcello Restelli "Reinforcement Learning in Algorithmic Trading" [21]

[17] https://llias2025.web.app/: https://arxiv.org/pdf/2505.22092
[18] https://codesign-lab.org/aic25/: https://hal.science/hal-05324845v1/file/Article_ECAI-AIC.pdf
[19] https://sites.google.com/view/sllm2025/home: https://arxiv.org/pdf/2510.21398
[20] https://modem2025.vub.ac.be/: https://modem2025.vub.ac.be/papers/MODeM_2025_paper_17.pdf
[21] https://sites.google.com/view/ECAI25-ai-fin

Model Science, getting serious about verification, explanation and control of AI systems[22]

- Author: Przemyslaw Biecek
- Growing Interest in the Foundation model (for fine-tuning use) and more generally Big model, motivates the shift from Data Science to Model Science
- Proposing a New Perspective focused on the Analysis of the model.

The proposed new domain is standing on four main pillars.

[22] https://arxiv.org/pdf/2508.20040

Model Science, getting serious about verification, explanation and control of AI systems[22]

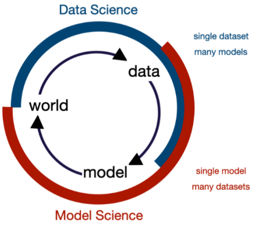

Figure: Complementarity and scope of Model Science

- Author: Przemyslaw Biecek
- Growing Interest in the Foundation model (for fine-tuning use) and more generally Big model, motivates the shift from Data Science to Model Science
- Proposing a New Perspective focused on the Analysis of the model.

The proposed new domain is standing on four main pillars.

[22] https://arxiv.org/pdf/2508.20040

Model Science, getting serious about verification, explanation and control of AI systems[22]

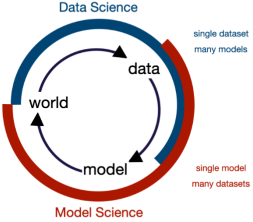

Figure: Complementarity and scope of Model Science

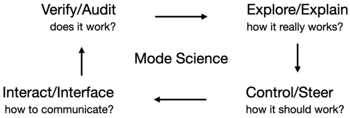

Figure 2. Four main pillars of model analysis

- Author: Przemyslaw Biecek
- Growing Interest in the Foundation model (for fine-tuning use) and more generally Big model, motivates the shift from Data Science to Model Science
- Proposing a New Perspective focused on the Analysis of the model.

The proposed new domain is standing on four main pillars.

[22] https://arxiv.org/pdf/2508.20040

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Proposed **5 levels of model evaluation:**

1. training data
2. optimistic test data
3. realistic test data
4. adversarial samples with limited access
5. adversarial samples with full access

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Proposed **5 levels of model evaluation:**

1. training data
2. optimistic test data
3. realistic test data
4. adversarial samples with limited access
5. adversarial samples with full access

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Proposed **5 levels of model evaluation:**

1. training data
2. optimistic test data
3. realistic test data
4. adversarial samples with limited access
5. adversarial samples with full access

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Proposed **5 levels of model evaluation:**

1. training data
2. optimistic test data
3. realistic test data
4. adversarial samples with limited access
5. adversarial samples with full access

Verification: does it work ?

- Evaluation of the model is not rigorous enough.
- Misalignment between the score announced and the true behaviors on a specific task in the real world.
- Whisper and Copilot examples.

We need to have a better audit of Models

The majority of AI models are less good at specific context validation. Need rigorous verification and especially in critical domains (medical, legal, education...).

Proposed **5 levels of model evaluation:**

1. training data
2. optimistic test data
3. realistic test data
4. adversarial samples with limited access
5. adversarial samples with full access

Explanation: how it really works ?

Observation of the growing interest in explainable AI (XAI), because it's really important to better interpret how the Model really works (for users and developers).

[23] Explainable AI Methods - A Brief Overview (2022)

Explanation: how it really works ?

Observation of the growing interest in explainable AI (XAI), because it's really important to better interpret how the Model really works (for users and developers).

We need to be aware of progress in XAI[23]

- Representation et Attention analysis
- Feature attribution
- Case-Based explanations
- Exploring with Synthetic data
- Model Auditing (red teaming)

[23] Explainable AI Methods - A Brief Overview (2022)

Explanation: how it really works ?

Observation of the growing interest in explainable AI (XAI), because it's really important to better interpret how the Model really works (for users and developers).

We need to be aware of progress in XAI[23]

- Representation et Attention analysis
- Feature attribution
- Case-Based explanations
- Exploring with Synthetic data
- Model Auditing (red teaming)

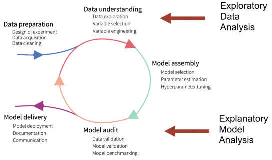

Figure 4. Explanatory model analysis can play a similar role to exploratory data analysis. The difference is that it takes place at a different part of the model training cycle.

[23] Explainable AI Methods - A Brief Overview (2022)

Control: how it should work ?

This Pillar is in continuity of verification and explanation $\rightarrow$ we want to make it work in the right manner.

Control: how it should work ?

This Pillar is in continuity of verification and explanation $\rightarrow$ we want to make it work in the right manner.

It's preferable to fine-tune instead of retraining because of the cost reasons.

Interface: how to interact with it ?

- When an AI model is designed for a user, the Human-computer interaction matters so that the user can answer "why?" and "how to change?" the output.
- Allowing exploration of the Model
- Sourcing
- Visualization

The UX design influences the perception of the user about the Model (does it work? / can I trust it ?)

Interface: how to interact with it ?

- When an AI model is designed for a user, the Human-computer interaction matters so that the user can answer "why?" and "how to change?" the output.
- Allowing exploration of the Model
- Sourcing
- Visualization

The UX design influences the perception of the user about the Model (does it work? / can I trust it ?)

Interface: how to interact with it ?

- When an AI model is designed for a user, the Human-computer interaction matters so that the user can answer "why?" and "how to change?" the output.
- Allowing exploration of the Model
- Sourcing
- Visualization

The UX design influences the perception of the user about the Model (does it work? / can I trust it ?)

Interface: how to interact with it ?

- When an AI model is designed for a user, the Human-computer interaction matters so that the user can answer "why?" and "how to change?" the output.
- Allowing exploration of the Model
- Sourcing
- Visualization

The UX design influences the perception of the user about the Model (does it work? / can I trust it ?)

Interface: how to interact with it ?

- When an AI model is designed for a user, the Human-computer interaction matters so that the user can answer "why?" and "how to change?" the output.
- Allowing exploration of the Model
- Sourcing
- Visualization

The UX design influences the perception of the user about the Model (does it work? / can I trust it ?)

Conclusion

Resume

The paper introduces a new manner to approach the evaluation of AI Models by some recommendations and observations classified in 4 classes: **(1) Verify** → **(2) Explanation** → **(3) Control** → **(4) Interface**. This new "field" name Model Science is complementary to the Data Science approach, but focuses on the Model instead of the Data.

Own opinion

- The Validation, Explanation, and Control pillars are relevant...
- ...while the Interface pillar has not convinced me.
- The paper is mainly oriented towards LLM.

Conclusion

Resume

The paper introduces a new manner to approach the evaluation of AI Models by some recommendations and observations classified in 4 classes: **(1) Verify** → **(2) Explanation** → **(3) Control** → **(4) Interface**. This new "field" name Model Science is complementary to the Data Science approach, but focuses on the Model instead of the Data.

Own opinion

- The Validation, Explanation, and Control pillars are relevant...
- ...while the Interface pillar has not convinced me.
- The paper is mainly oriented towards LLM.

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

Full-History Graphs with Edge-Type Decoupled Networks for Temporal Reasoning[24]

- Author: Osama Mohammed
- In Many cases, the real-world data can be modelised by a Graph (relational). We can study the evolution over time of the Data, and thus, the Dynamic Graph is a modelization often used for temporal and structural data.
- The goal is to capture **instantaneous relations** (the links between objects, and how it's evolve in time) in two cases: Waymo and Elliptic++ Datasets.
- The Contribution:
  - a new modeling of the Dynamic Graph via **Full-History Graph (FHG) Representation**
  - **Edge-Type Decoupled Network(ETDNet):** a new architecture of Temporal Graph Neural Network that exploits the FHG.

[24] https://arxiv.org/pdf/2508.03251?

How to make Temporal Graph Neural Networks (TGNN) ?

2 strategies of TGNNs:

- Snapshot Methods
- Integrated methods
  - Memory-Bank temporal GNNs
  - Continuous-time attention kernels
  - Spatial-temporal separation

Snapshot methods tend to lose the order of observations and fail with long-range temporal data.

Memory-bank temporal GNNs and Continuous-time attention kernels methods mix structural and temporal information => Blur problem

EDTNet is a Spatial-temporal separation method using 2 channels for temporal and structural information, respectively.

How to make Temporal Graph Neural Networks (TGNN) ?

2 strategies of TGNNs:

- Snapshot Methods
- Integrated methods
  - Memory-Bank temporal GNNs
  - Continuous-time attention kernels
  - Spatial-temporal separation

Snapshot methods tend to lose the order of observations and fail with long-range temporal data.

Memory-bank temporal GNNs and Continuous-time attention kernels methods mix structural and temporal information => Blur problem

EDTNet is a Spatial-temporal separation method using 2 channels for temporal and structural information, respectively.

How to make Temporal Graph Neural Networks (TGNN) ?

2 strategies of TGNNs:

- Snapshot Methods
- Integrated methods
  - Memory-Bank temporal GNNs
  - Continuous-time attention kernels
  - Spatial-temporal separation

Snapshot methods tend to lose the order of observations and fail with long-range temporal data.

Memory-bank temporal GNNs and Continuous-time attention kernels methods mix structural and temporal information => Blur problem

EDTNet is a Spatial-temporal separation method using 2 channels for temporal and structural information, respectively.

How to make Temporal Graph Neural Networks (TGNN) ?

2 strategies of TGNNs:

- Snapshot Methods
- Integrated methods
  - Memory-Bank temporal GNNs
  - Continuous-time attention kernels
  - Spatial-temporal separation

Snapshot methods tend to lose the order of observations and fail with long-range temporal data.

Memory-bank temporal GNNs and Continuous-time attention kernels methods mix structural and temporal information => Blur problem

EDTNet is a Spatial-temporal separation method using 2 channels for temporal and structural information, respectively.

How to make Temporal Graph Neural Networks (TGNN) ?

2 strategies of TGNNs:

- Snapshot Methods
- Integrated methods
  - Memory-Bank temporal GNNs
  - Continuous-time attention kernels
  - Spatial-temporal separation

Snapshot methods tend to lose the order of observations and fail with long-range temporal data.

Memory-bank temporal GNNs and Continuous-time attention kernels methods mix structural and temporal information => Blur problem

EDTNet is a Spatial-temporal separation method using 2 channels for temporal and structural information, respectively.

Full-History Graph Representation

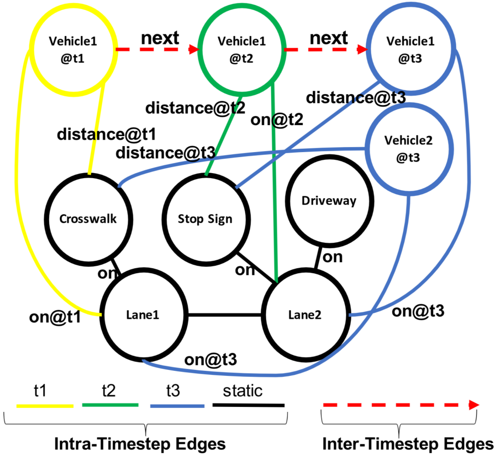

Figure 1: Example full-history graph with three timesteps ($t=1, 2, 3$). Dynamic entity *Vehicle 1* appears as three nodes linked by dashed red temporal edges; static road objects interact with it in each timestep.

- 2 types of nodes: statics and dynamics
- 2 types of edges: Intra-Timestep and Inter-Timestep

Edge-Type Decoupled Network (ETDNet)

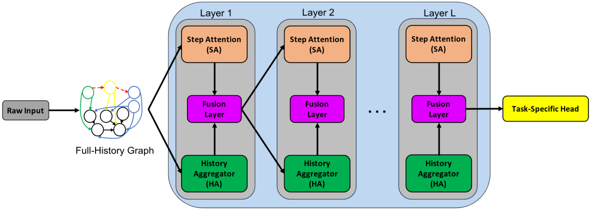

Figure 3: End-to-end **ETDNet** encoder. Raw data are converted to a *full-history graph*; each layer then applies *Step Attention* (**SA**) on domain edges and *History Attention* (**HA**) on temporal edges. A residual *Fusion Layer* merges the output of the two messages and feeds the next layer. After L layers, a lightweight *task-specific head* produces the final predictions.

Edge-Type Decoupled Network (ETDNet)

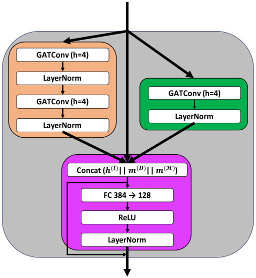

Figure 4: Inside one ETD layer. **(left)** SA stack: K_s GATConv sub-layers, each with H_s heads, followed by LayerNorm and a residual arrow. **(right)** HA block: H_t-head multi-head attention over the B-length window, followed by LayerNorm. **(bottom)** Fusion Layer concatenates $h^{(\ell)}$, m^D, and $m^{\mathcal{H}}$, applies a $3d \to d$ linear projection, ReLU, and LayerNorm, then outputs $h^{(\ell+1)}$. Arrows show residual connections.

- **Step Attention (SA)**: focus only on intra-timesteps edges, for the local information
- **History Attention (HA)** focuses only on inter-timesteps edges, to treat temporal cues.
- **Fusion Layer (FL)**: FCN which merge embeddings of SA and HA.

Within ETDNet, we always separate spatial interactions with temporal continuity, avoiding the blur problem.

Edge-Type Decoupled Network (ETDNet)

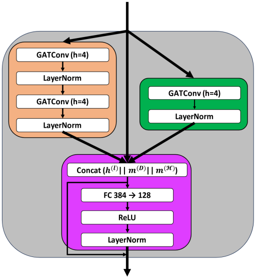

Figure 4: Inside one ETD layer. **(left)** SA stack: K_s GATConv sub-layers, each with H_s heads, followed by LayerNorm and a residual arrow. **(right)** HA block: H_t-head multi-head attention over the B-length window, followed by LayerNorm. **(bottom)** Fusion Layer concatenates $h^{(\ell)}$, m^D, and $m^{\mathcal{H}}$, applies a $3d \to d$ linear projection, ReLU, and LayerNorm, then outputs $h^{(\ell+1)}$. Arrows show residual connections.

- **Step Attention (SA)**: focus only on intra-timesteps edges, for the local information
- **History Attention (HA)** focuses only on inter-timesteps edges, to treat temporal cues.
- **Fusion Layer (FL)**: FCN which merge embeddings of SA and HA.

Within ETDNet, we always separate spatial interactions with temporal continuity, avoiding the blur problem.

Edge-Type Decoupled Network (ETDNet)

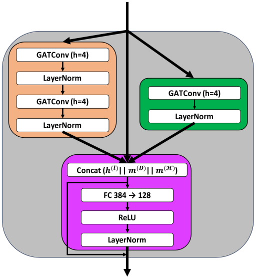

Figure 4: Inside one ETD layer. **(left)** SA stack: K_s GATConv sublayers, each with H_s heads, followed by LayerNorm and a residual arrow. **(right)** HA block: H_t-head multi-head attention over the B-length window, followed by LayerNorm. **(bottom)** Fusion Layer concatenates $h^{(\ell)}$, m^D, and $m^{\mathcal{H}}$, applies a $3d \to d$ linear projection, ReLU, and LayerNorm, then outputs $h^{(\ell+1)}$. Arrows show residual connections.

- **Step Attention (SA)**: focus only on intra-timesteps edges, for the local information
- **History Attention (HA)** focuses only on inter-timesteps edges, to treat temporal cues.
- **Fusion Layer (FL)**: FCN which merge embeddings of SA and HA.

Within ETDNet, we always separate spatial interactions with temporal continuity, avoiding the blur problem.

Edge-Type Decoupled Network (ETDNet)

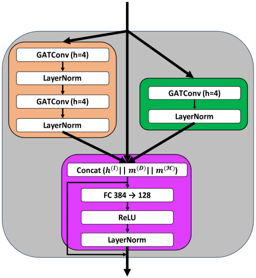

Figure 4: Inside one ETD layer. **(left)** SA stack: K_s GATConv sublayers, each with H_s heads, followed by LayerNorm and a residual arrow. **(right)** HA block: H_t-head multi-head attention over the B-length window, followed by LayerNorm. **(bottom)** Fusion Layer concatenates $h^{(\ell)}$, m^D, and $m^{\mathcal{H}}$, applies a $3d \to d$ linear projection, ReLU, and LayerNorm, then outputs $h^{(\ell+1)}$. Arrows show residual connections.

- **Step Attention (SA)**: focus only on intra-timesteps edges, for the local information
- **History Attention (HA)** focuses only on inter-timesteps edges, to treat temporal cues.
- **Fusion Layer (FL)**: FCN which merge embeddings of SA and HA.

Within ETDNet, we always separate spatial interactions with temporal continuity, avoiding the blur problem.

Experiments

Table 2: Driver-intention classification on Waymo (macro–F_1 %, mean $\pm$ std over 3 runs).

Method	**F1 (Speed)**	**F1 (Direction)**	**Joint Acc.**
LSTM (per-agent)	82.5 ± 0.3	77.6 ± 0.4	70.2 ± 0.5
ST-GCN [33]	83.1 ± 0.4	78.2 ± 0.5	71.0 ± 0.4
SemanticFormer [23]	83.9 ± 0.4	79.5 ± 0.4	72.4 ± 0.5
GMT [1]	84.5 ± 0.3	80.9 ± 0.4	73.6 ± 0.4
STGNN [13]	83.5 ± 0.5	78.9 ± 0.6	71.5 ± 0.5
TGN [19]	84.9 ± 0.3	81.6 ± 0.3	74.1 ± 0.4
ETDNet (ours) [16]	$\mathbf{85.7 \pm 0.3}$	$\mathbf{82.9 \pm 0.3}$	$\mathbf{75.6 \pm 0.4}$

Table 3: Elliptic++ fraud detection (timesteps 41–49). Mean $\pm$ std. over 3 seeds; higher = better.

Model	**ROC-AUC**	**AUPRC**	**F1 (illicit)**
GCN [11]	0.833 ± 0.004	0.595 ± 0.006	56.5 ± 0.9
GraphSAGE [8]	0.828 ± 0.005	0.631 ± 0.008	58.4 ± 0.8
GAT [25]	0.810 ± 0.006	0.589 ± 0.009	55.2 ± 1.1
TGN [19]	0.878 ± 0.003	0.659 ± 0.007	60.4 ± 0.6
TGAT [32]	0.868 ± 0.004	0.648 ± 0.006	59.0 ± 0.5
EvolveGCN [17]	0.860 ± 0.005	0.641 ± 0.007	57.3 ± 0.7
DyGFormer [10]	0.882 ± 0.003	0.675 ± 0.005	61.0 ± 0.4
ETDNet (ours) [16]	$\mathbf{0.884\pm0.002}$	$\mathbf{0.863\pm0.004}$	$\mathbf{88.1\pm0.5}$

Table 6: Training efficiency on **Waymo** and **Elliptic++**.

Model	**Waymo**		**Elliptic++**	
	Params (M)	Epoch (s)	Params (M)	Epoch (s)
ST-GCN	0.15	0.62	0.15	0.18
TGN	0.90	1.74	0.90	0.52
TGAT	1.20	1.95	1.20	0.60
ETDNet	0.30	**1.50**	0.30	**0.40**

Experiments

Table 2: Driver-intention classification on Waymo (macro–F_1 %, mean ± std over 3 runs).

Method	F1 (Speed)	F1 (Direction)	Joint Acc.
LSTM (per-agent)	82.5 ± 0.3	77.6 ± 0.4	70.2 ± 0.5
ST-GCN [33]	83.1 ± 0.4	78.2 ± 0.5	71.0 ± 0.4
SemanticFormer [23]	83.9 ± 0.4	79.5 ± 0.4	72.4 ± 0.5
GMT [1]	84.5 ± 0.3	80.9 ± 0.4	73.6 ± 0.4
STGNN [13]	83.5 ± 0.5	78.9 ± 0.6	71.5 ± 0.5
TGN [19]	84.9 ± 0.3	81.6 ± 0.3	74.1 ± 0.4
ETDNet (ours) [16]	**85.7 ± 0.3**	**82.9 ± 0.3**	**75.6 ± 0.4**

Table 3: Elliptic++ fraud detection (timesteps 41–49). Mean ± std. over 3 seeds; higher = better.

Model	ROC-AUC	AUPRC	F1 (illicit)
GCN [11]	0.833±0.004	0.595±0.006	56.5±0.9
GraphSAGE [8]	0.828±0.005	0.631±0.008	58.4±0.8
GAT [25]	0.810±0.006	0.589±0.009	55.2±1.1
TGN [19]	0.878±0.003	0.659±0.007	60.4±0.6
TGAT [32]	0.868±0.004	0.648±0.006	59.0±0.5
EvolveGCN [17]	0.860±0.005	0.641±0.007	57.3±0.7
DyGFormer [10]	0.882±0.003	0.675±0.005	61.0±0.4
ETDNet (ours) [16]	**0.884±0.002**	**0.863±0.004**	**88.1±0.5**

Table 6: Training efficiency on **Waymo** and **Elliptic++**.

Model	**Waymo**		**Elliptic++**	
	Params (M)	Epoch (s)	Params (M)	Epoch (s)
ST-GCN	0.15	0.62	0.15	0.18
TGN	0.90	1.74	0.90	0.52
TGAT	1.20	1.95	1.20	0.60
ETDNet	0.30	**1.50**	0.30	**0.40**

Experiments

Table 2: Driver-intention classification on Waymo (macro–F_1 %, mean ± std over 3 runs).

Method	F1 (Speed)	F1 (Direction)	Joint Acc.
LSTM (per-agent)	82.5 ± 0.3	77.6 ± 0.4	70.2 ± 0.5
ST-GCN [33]	83.1 ± 0.4	78.2 ± 0.5	71.0 ± 0.4
SemanticFormer [23]	83.9 ± 0.4	79.5 ± 0.4	72.4 ± 0.5
GMT [1]	84.5 ± 0.3	80.9 ± 0.4	73.6 ± 0.4
STGNN [13]	83.5 ± 0.5	78.9 ± 0.6	71.5 ± 0.5
TGN [19]	84.9 ± 0.3	81.6 ± 0.3	74.1 ± 0.4
ETDNet (ours) [16]	**85.7 ± 0.3**	**82.9 ± 0.3**	**75.6 ± 0.4**

Table 3: Elliptic++ fraud detection (timesteps 41–49). Mean ± std. over 3 seeds; higher = better.

Model	ROC-AUC	AUPRC	F1 (illicit)
GCN [11]	0.833±0.004	0.595±0.006	56.5±0.9
GraphSAGE [8]	0.828±0.005	0.631±0.008	58.4±0.8
GAT [25]	0.810±0.006	0.589±0.009	55.2±1.1
TGN [19]	0.878±0.003	0.659±0.007	60.4±0.6
TGAT [32]	0.868±0.004	0.648±0.006	59.0±0.5
EvolveGCN [17]	0.860±0.005	0.641±0.007	57.3±0.7
DyGFormer [10]	0.882±0.003	0.675±0.005	61.0±0.4
ETDNet (ours) [16]	**0.884±0.002**	**0.863±0.004**	**88.1±0.5**

Table 6: Training efficiency on **Waymo** and **Elliptic++**.

Model	Waymo		Elliptic++	
	Params (M)	Epoch (s)	Params (M)	Epoch (s)
ST-GCN	0.15	0.62	0.15	0.18
TGN	0.90	1.74	0.90	0.52
TGAT	1.20	1.95	1.20	0.60
ETDNet	0.30	**1.50**	0.30	**0.40**

Conclusion

Resume

The paper introduces Full-History Graphs and ETDNet (dual-branch temporal-GNN), which permit to learn fast, and have a better expressivity of temporal and spatial information of the dynamic graph. The Model is evaluated on 2 datasets with different scales of time, showing its capacity to deal with dense local information and long-range temporal data.

Own opinion

This is my first read about the research in Temporal-GNN...

Conclusion

Resume

The paper introduces Full-History Graphs and ETDNet (dual-branch temporal-GNN), which permit to learn fast, and have a better expressivity of temporal and spatial information of the dynamic graph. The Model is evaluated on 2 datasets with different scales of time, showing its capacity to deal with dense local information and long-range temporal data.

Own opinion

This is my first read about the research in Temporal-GNN...

Towards Language-Augmented Multi-Agent Deep Reinforcement Learning (LAMARL)[25]

- Author: Maxime Toquebiau
- MARL and communication: emergent language
- They argue the need for learning communication through a natural language to enable complex reasoning in embodied and social settings.
- Language-Augmented MARL (LAMARL) is a learning method where agent learn both communication through pre-defined language and learn to solve the task (optimize their behavior policy).

[25] https://arxiv.org/pdf/2506.05236

Towards Language-Augmented Multi-Agent Deep Reinforcement Learning (LAMARL)[25]

- Author: Maxime Toquebiau
- MARL and communication: emergent language
- They argue the need for learning communication through a natural language to enable complex reasoning in embodied and social settings.
- Language-Augmented MARL (LAMARL) is a learning method where agent learn both communication through pre-defined language and learn to solve the task (optimize their behavior policy).

[25] https://arxiv.org/pdf/2506.05236

Towards Language-Augmented Multi-Agent Deep Reinforcement Learning (LAMARL)[25]

- Author: Maxime Toquebiau
- MARL and communication: emergent language
- They argue the need for learning communication through a natural language to enable complex reasoning in embodied and social settings.
- Language-Augmented MARL (LAMARL) is a learning method where agent learn both communication through pre-defined language and learn to solve the task (optimize their behavior policy).

[25] https://arxiv.org/pdf/2506.05236

Towards Language-Augmented Multi-Agent Deep Reinforcement Learning (LAMARL)[25]

- Author: Maxime Toquebiau
- MARL and communication: emergent language
- They argue the need for learning communication through a natural language to enable complex reasoning in embodied and social settings.
- Language-Augmented MARL (LAMARL) is a learning method where agent learn both communication through pre-defined language and learn to solve the task (optimize their behavior policy).

[25] https://arxiv.org/pdf/2506.05236

how to communicate in MADRL ?

- ~~emergent communication~~
- grounding communication
- ~~LLM-based multi-agents~~
- Language Augmented Reinforcement Learning

how to communicate in MADRL ?

- ~~emergent communication~~
- grounding communication
- ~~LLM-based multi-agents~~
- Language Augmented Reinforcement Learning

how to communicate in MADRL ?

- ~~emergent communication~~
- grounding communication
- ~~LLM-based multi-agents~~
- Language Augmented Reinforcement Learning

how to communicate in MADRL ?

- ~~emergent communication~~
- grounding communication
- ~~LLM-based multi-agents~~
- Language Augmented Reinforcement Learning

how to communicate in MADRL ?

- ~~emergent communication~~
- grounding communication
- ~~LLM-based multi-agents~~
- Language Augmented Reinforcement Learning

LAMARL settings

- Cooperative (common-reward games)
- Partially Observable
- Unconstrained Communication
- Full Communication
- Centralised Training Decentralised Execution (CTDE)

Oracle function to describe observations with vocabulary

An Oracle function defines the vocabulary:

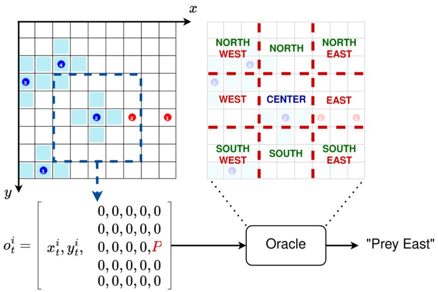

Figure 2. Illustration of the oracle's process for describing observations. In the top-left part is a screenshot of the MA-gym Predator-Prey task (agents in blue and preys in red). Agents observe their position in the grid and the objects in their 5×5 observation range (shown in dark blue). The oracle generates a language description describing the location of observed preys.

The function gives a simple but Structured Combinatorial Language which gives relevant information.

Oracle function to describe observations with vocabulary

An Oracle function defines the vocabulary:

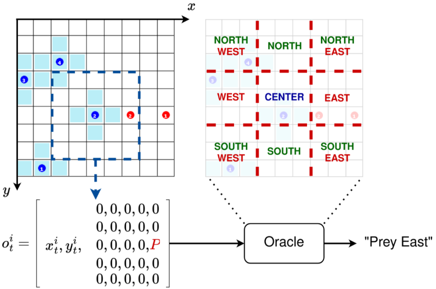

Figure 2. Illustration of the oracle's process for describing observations. In the top-left part is a screenshot of the MA-gym Predator-Prey task (agents in blue and preys in red). Agents observe their position in the grid and the objects in their 5×5 observation range (shown in dark blue). The oracle generates a language description describing the location of observed preys.

The function gives a simple but Structured Combinatorial Language which gives relevant information.

Language-Augmented Agents

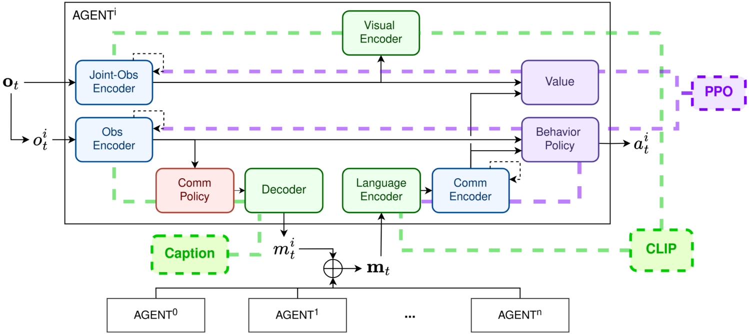

Figure 1. Language-augmented agent architecture. Each module represents a neural network with a specific purpose. **Encoder modules** receive incoming information (observations or communication) and embed it in latent representations used in further modules, with dashed self-pointing arrows indicating the use of recurrent neural networks to allow some memorization over multiple time steps. The **communication policy** selects what information in the observation should be communicated. **Language modules** provide language capabilities: the decoder generates messages, the language encoder transforms incoming messages in a compact latent representation, and the visual encoder is used for training the language encoder (as described in Section 4.3). Finally, **reinforcement learning modules** take information from both observations and communication to select actions and generate values. Colored dashed lines indicate the three training objectives and which modules they impact.

3 Objectives Learning

Multi-Agent PPO with centralised critic (joint observations) and Decentralised policy (local observations), but both take incoming joint messages.

Generating Language with Captioning given observation-description tuple (o, l) with $l = (t_0, ..., t_K)$ a token sequences: the decoder learn to generate l based on selected observations from the communication policy:

$$\mathcal{L}^{capt}(\theta^i) = -\sum_{k=1}^{K} p(t_k) \log \hat{p}(t_k | o, t_0, ..., t_{k-1}, \theta^i)$$

Encoding language with contrastive learning to interpret incoming joint messages, use of CLIP loss like:

$$\mathcal{L}^{CLIP}(\theta^i) = \sum_{j,k} CLIP(j, k)$$

$$CLIP(j, k) = \begin{cases} -\mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j = k, \\ \mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j \neq k. \end{cases}$$

Complete Objective :
$\mathcal{L}(\theta^i) = \beta^\pi \mathcal{L}^\pi(\theta^i) + \beta^V \mathcal{L}^V(\theta^i) + \beta^{\mathrm{capt}} \mathcal{L}^{\mathrm{capt}}(\theta^i) + \beta^{\mathrm{CLIP}} \mathcal{L}^{\mathrm{CLIP}}(\theta^i)$

3 Objectives Learning

Multi-Agent PPO with centralised critic (joint observations) and Decentralised policy (local observations), but both take incoming joint messages.
Generating Language with Captioning given observation-description tuple (o, l) with $l = (t_0, ..., t_K)$ a token sequences: the decoder learn to generate l based on selected observations from the communication policy:

$$\mathcal{L}^{capt}(\theta^i) = -\sum_{k=1}^{K} p(t_k) \log \hat{p}(t_k | o, t_0, ..., t_{k-1}, \theta^i)$$

Encoding language with contrastive learning to interpret incoming joint messages, use of CLIP loss like:

$$\mathcal{L}^{CLIP}(\theta^i) = \sum_{j,k} CLIP(j,k)$$

$$CLIP(j,k) = \begin{cases} -\mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j = k, \\ \mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j \neq k. \end{cases}$$

Complete Objective :
$\mathcal{L}(\theta^i) = \beta^{\pi}\mathcal{L}^{\pi}(\theta^i) + \beta^{V}\mathcal{L}^{V}(\theta^i) + \beta^{\mathrm{capt}}\mathcal{L}^{\mathrm{capt}}(\theta^i) + \beta^{\mathrm{CLIP}}\mathcal{L}^{\mathrm{CLIP}}(\theta^i)$

3 Objectives Learning

Multi-Agent PPO with centralised critic (joint observations) and Decentralised policy (local observations), but both take incoming joint messages.
Generating Language with Captioning given observation-description tuple (o, l) with $l = (t_0, ..., t_K)$ a token sequences: the decoder learn to generate l based on selected observations from the communication policy:

$$\mathcal{L}^{capt}(\theta^i) = -\sum_{k=1}^{K} p(t_k) \log \hat{p}(t_k | o, t_0, ..., t_{k-1}, \theta^i)$$

Encoding language with contrastive learning to interpret incoming joint messages, use of CLIP loss like:

$$\mathcal{L}^{CLIP}(\theta^i) = \sum_{j,k} CLIP(j, k)$$

$$CLIP(j, k) = \begin{cases} -\mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j = k, \\ \mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j \neq k. \end{cases}$$

Complete Objective :
$\mathcal{L}(\theta^i) = \beta^\pi \mathcal{L}^\pi(\theta^i) + \beta^V \mathcal{L}^V(\theta^i) + \beta^{\mathrm{capt}} \mathcal{L}^{\mathrm{capt}}(\theta^i) + \beta^{\mathrm{CLIP}} \mathcal{L}^{\mathrm{CLIP}}(\theta^i)$

3 Objectives Learning

Multi-Agent PPO with centralised critic (joint observations) and Decentralised policy (local observations), but both take incoming joint messages.
Generating Language with Captioning given observation-description tuple (o, l) with $l = (t_0, ..., t_K)$ a token sequences: the decoder learn to generate l based on selected observations from the communication policy:

$$\mathcal{L}^{capt}(\theta^i) = -\sum_{k=1}^{K} p(t_k) \log \hat{p}(t_k | o, t_0, ..., t_{k-1}, \theta^i)$$

Encoding language with contrastive learning to interpret incoming joint messages, use of CLIP loss like:

$$\mathcal{L}^{CLIP}(\theta^i) = \sum_{j,k} CLIP(j, k)$$

$$CLIP(j, k) = \begin{cases} -\mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j = k, \\ \mathrm{cosim}(E^V(\mathbf{o}), E^L(\mathbf{l})) & \text{if } j \neq k. \end{cases}$$

Complete Objective :
$\mathcal{L}(\theta^i) = \beta^{\pi}\mathcal{L}^{\pi}(\theta^i) + \beta^{V}\mathcal{L}^{V}(\theta^i) + \beta^{\mathrm{capt}}\mathcal{L}^{\mathrm{capt}}(\theta^i) + \beta^{\mathrm{CLIP}}\mathcal{L}^{\mathrm{CLIP}}(\theta^i)$

Experiments

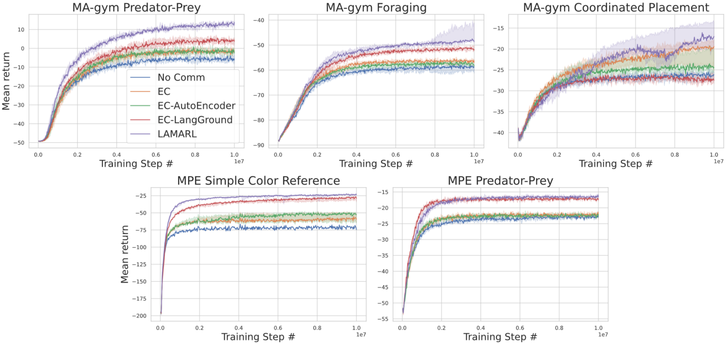

Figure 3. Training performance of LAMARL agents against baselines (7 runs each, with median and 95% confidence interval). LAMARL agents consistently beat the emergent communication baselines.

Experiments

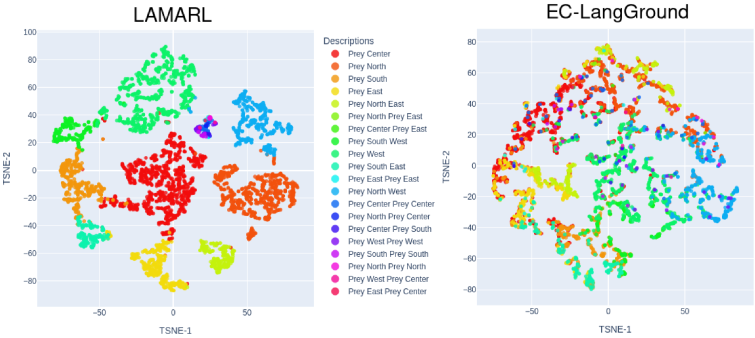

Figure 4. Visualizations of embeddings produced by the communication policy in `LAMARL` (left) and `EC-LangGround` (right). By explicitly learning to generate language utterances, `LAMARL` learns better structured representation, as shown by the more distinct clusters and the silhouette scores: 0.59 for `LAMARL` and 0.06 for `EC-LangGround`.

Experiments

Agent version	Language learning	Communication strategy
LAMARL	Yes	Learned language
Lang+Oracle	Yes	Oracle
Lang+No Comm	Yes	No communication
No Lang+Oracle	No	Oracle
Observations	No	Raw observations
No Comm	No	No communication

Table 1. Properties of variants compared in the ablation study.

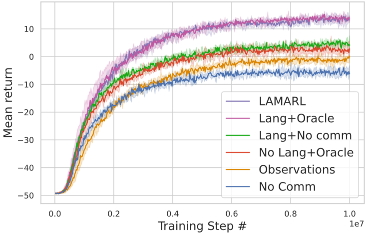

Figure 5. Training performance of the ablated version on Predator-Prey 18×18 (7 runs each, with median and 95% confidence interval). `LAMARL` agents achieve similar performance as `Oracle` agents, showing that they succeed in using the language properly. All other ablations are inferior.

LIRIS

Experiments

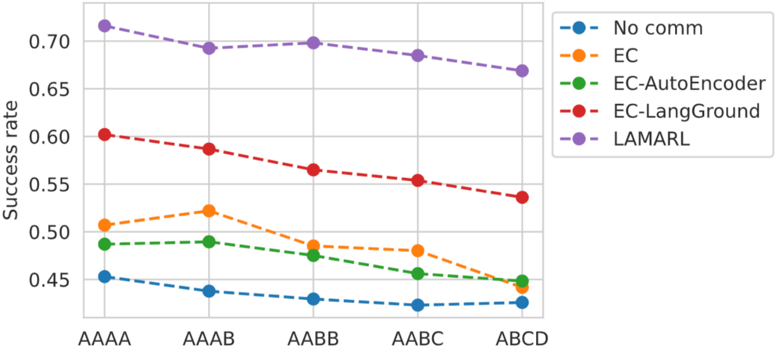

Figure 6. Evaluation performance in the zero-shot teaming experiment. Letters indicate the team composition. Results show the success rate over the evaluation run (1.2×10^6 episodes).

Experiments

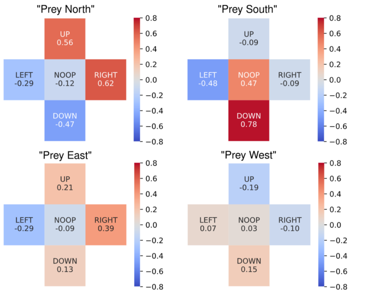

Figure 7. Change in action probabilities depending on the input message. Positive values (resp., negative values) indicate that the action will be taken more often (resp., less often) if the corresponding message is received, compared to if no message is received.

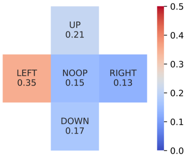

Figure 9. Action probabilities of agents evaluated in the interaction experiment, when no message is received.

Conclusion

Resume

LAMARL is a framework of Communication-MADRL for learning both to communicate between agents and to behave cooperatively. Using a combined Objective function to learn (1) to complete a given embodied task, while (2) generating language messages, and (3) understanding language messages

Own opinion

I found the approach interesting, and the evaluation is quite complete.

Conclusion

Resume

LAMARL is a framework of Communication-MADRL for learning both to communicate between agents and to behave cooperatively. Using a combined Objective function to learn (1) to complete a given embodied task, while (2) generating language messages, and (3) understanding language messages

Own opinion

I found the approach interesting, and the evaluation is quite complete.

Thanks for your attention

Let's talk ! Any questions?