Collaborative and Strategic Multiagent Learning
(Machine Learning in the Agents Group)

Michael Rovatsos

Centre for Intelligent Systems and their Applications

School of informatics

The University of Edinburgh

RPML meeting
12th November 2008

Agents & Multiagent Systems

- Common misconception (often reinforced by agents people): agents research is about making smart programs (that would be all of Informatics ...)

Agents & Multiagent Systems

- Common misconception (often reinforced by agents people): agents research is about making smart programs (that would be all of Informatics ...)
- Agents research is concerned with the coordination of intelligent capabilities in a single (single-agent) or several distributed (multiagent) systems

Agents & Multiagent Systems

- Common misconception (often reinforced by agents people): agents research is about making smart programs (that would be all of Informatics ...)
- Agents research is concerned with the coordination of intelligent capabilities in a single (single-agent) or several distributed (multiagent) systems
- When strictly collaborative, similar to very loosely coupled distributed systems; when not, distinct from other CS/AI technologies

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:
  - Collaborative learning (working together to improve local learning)

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:
  - Collaborative learning (working together to improve local learning)
  - Strategic learning (learning to act optimally when outcomes depend on others)

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:
  - Collaborative learning (working together to improve local learning)
  - Strategic learning (learning to act optimally when outcomes depend on others)
    - opponent modelling

Multiagent learning

- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:
  - Collaborative learning (working together to improve local learning)
  - Strategic learning (learning to act optimally when outcomes depend on others)
    - opponent modelling
    - communication in learning

Multiagent learning

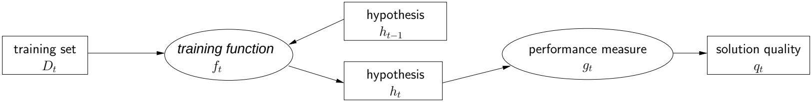
- Multiagent learning = learning in the presence of others (i.e. from, with, or about others)
- Parts of the area closely related to distributed learning, multi-strategy learning, transfer learning, mixed-initiative learning
- Work in my group in two areas:
  - Collaborative learning (working together to improve local learning)
  - Strategic learning (learning to act optimally when outcomes depend on others)
    - opponent modelling
    - communication in learning
- Some examples follow (joint work with Belesiotis, Collins, Figari, Fischer, Mhatre, Nickles, Pechoucek, Rafael, Sferopoulos, Tozicka, Weiss, Whitaker, Wolf)

Collaborative learning

- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning

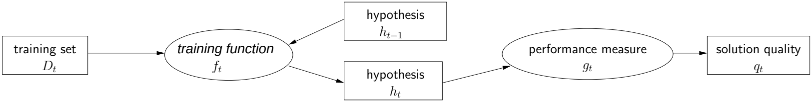
Collaborative learning

- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning
- Based on generic learning agent model:



Collaborative learning

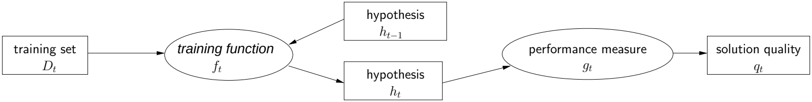
- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning
- Based on generic learning agent model:



- Assuming two interacting learners, we obtain a matrix of potential interactions

Collaborative learning

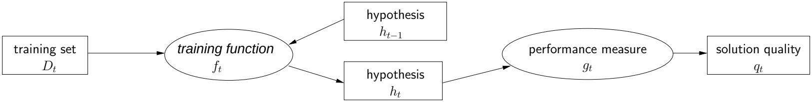
- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning
- Based on generic learning agent model:



- Assuming two interacting learners, we obtain a matrix of potential interactions
- Enables agents to make autonomous decisions about how to integrate information received from others

Collaborative learning

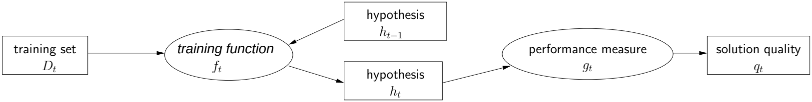
- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning
- Based on generic learning agent model:



- Assuming two interacting learners, we obtain a matrix of potential interactions
- Enables agents to make autonomous decisions about how to integrate information received from others
- Particularly interested in hypothesis merging in domains where data exchange is not possible/desirable

Collaborative learning

- MALEF: A generic multiagent learning framework that allows autonomous decision making in all stages of individual learning
- Based on generic learning agent model:



- Assuming two interacting learners, we obtain a matrix of potential interactions
- Enables agents to make autonomous decisions about how to integrate information received from others
- Particularly interested in hypothesis merging in domains where data exchange is not possible/desirable
- Applied to vessel surveillance and distributed brain tumor diagnosis

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards
  - stationary probability distribution governs transitions between these games (depending on all players' joint actions)

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards
  - stationary probability distribution governs transitions between these games (depending on all players' joint actions)
- Challenge: from a learner's perspective opponents could be learners, too (non-stationary environment)

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards
  - stationary probability distribution governs transitions between these games (depending on all players' joint actions)
- Challenge: from a learner's perspective opponents could be learners, too (non-stationary environment)
- Rational learners won't change behaviour in arbitrary ways

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards
  - stationary probability distribution governs transitions between these games (depending on all players' joint actions)
- Challenge: from a learner's perspective opponents could be learners, too (non-stationary environment)
- Rational learners won't change behaviour in arbitrary ways
- We can do better than in completely unstructured environments

Strategic multiagent learning

- Normally studied within framework of multiagent reinforcement learning (MARL):
  - every environment state is a multi-player game (in the game-theoretic sense) that determines agent rewards
  - stationary probability distribution governs transitions between these games (depending on all players' joint actions)
- Challenge: from a learner's perspective opponents could be learners, too (non-stationary environment)
- Rational learners won't change behaviour in arbitrary ways
- We can do better than in completely unstructured environments
- Hard to achieve balance between learning and teaching; also, learning always has to be done online

Opponent classification in repeated games

- Assume two-player interactions in repeated games in very large agent societies

Opponent classification in repeated games

- Assume two-player interactions in repeated games in very large agent societies
- Unlikely that each of them has a different strategy, and pointless to wait until we have played with everyone

Opponent classification in repeated games

- Assume two-player interactions in repeated games in very large agent societies
- Unlikely that each of them has a different strategy, and pointless to wait until we have played with everyone
- Group agents together by looking for similar behaviour (every opponent model is a finite-state machine)

Opponent classification in repeated games

- Assume two-player interactions in repeated games in very large agent societies
- Unlikely that each of them has a different strategy, and pointless to wait until we have played with everyone
- Group agents together by looking for similar behaviour (every opponent model is a finite-state machine)
- Main result: opponent classification boosts learning speed and improves average rewards when playing with unknown opponents

Opponent classification in repeated games

- Assume two-player interactions in repeated games in very large agent societies
- Unlikely that each of them has a different strategy, and pointless to wait until we have played with everyone
- Group agents together by looking for similar behaviour (every opponent model is a finite-state machine)
- Main result: opponent classification boosts learning speed and improves average rewards when playing with unknown opponents
- Later expanded this to a two-player soccer playing domain using Bayesian models rather than FSMs

Opponent modelling with temporarily stable strategies

- Similar situation as above, but rather than picking a strategy assume opponent occasionally switches between different strategies

Opponent modelling with temporarily stable strategies

- Similar situation as above, but rather than picking a strategy assume opponent occasionally switches between different strategies
- This can be used to model learning opponents to a certain extent (their strategies are stable iff they currently have a hypothesis about us)

Opponent modelling with temporarily stable strategies

- Similar situation as above, but rather than picking a strategy assume opponent occasionally switches between different strategies
- This can be used to model learning opponents to a certain extent (their strategies are stable iff they currently have a hypothesis about us)
- Easy to learn optimal best response strategies toward stable strategies, challenge: how to detect (and make best possible use of) transition phases

Opponent modelling with temporarily stable strategies

- Similar situation as above, but rather than picking a strategy assume opponent occasionally switches between different strategies
- This can be used to model learning opponents to a certain extent (their strategies are stable iff they currently have a hypothesis about us)
- Easy to learn optimal best response strategies toward stable strategies, challenge: how to detect (and make best possible use of) transition phases
- Ongoing work, will use hierarchical reinforcement learning methods and "machine games" studied in game theory

Nested honesty modelling in repeated games

- Imagine an opponent is being cooperative in situations where there is a temptation to deceive

Nested honesty modelling in repeated games

- Imagine an opponent is being cooperative in situations where there is a temptation to deceive
- This may be due to them assign some value to the "shadow of the future" (avoid appearing deceitful to escape future punishment)

Nested honesty modelling in repeated games

- Imagine an opponent is being cooperative in situations where there is a temptation to deceive
- This may be due to them assign some value to the "shadow of the future" (avoid appearing deceitful to escape future punishment)
- Appled nested Bayesian modelling ("I think that you think that I think ...") to estimate such honesty values

Nested honesty modelling in repeated games

- Imagine an opponent is being cooperative in situations where there is a temptation to deceive
- This may be due to them assign some value to the "shadow of the future" (avoid appearing deceitful to escape future punishment)
- Appled nested Bayesian modelling ("I think that you think that I think ...") to estimate such honesty values
- Estimation process biased by rationality considerations about the other player's behaviour in the game (in the honesty-modified game they should behave rationally)

Advice-taking in MARL

- Assume a third party that provides agents with some (reward-irrelevant) information about how they should play with each other

Advice-taking in MARL

- Assume a third party that provides agents with some (reward-irrelevant) information about how they should play with each other
- This information source might be unreliable (deliberately or due to lack of knowledge of the interaction problem)

Advice-taking in MARL

- Assume a third party that provides agents with some (reward-irrelevant) information about how they should play with each other
- This information source might be unreliable (deliberately or due to lack of knowledge of the interaction problem)
- Developed algorithm that can take such advice into account in a robust way

Advice-taking in MARL

- Assume a third party that provides agents with some (reward-irrelevant) information about how they should play with each other
- This information source might be unreliable (deliberately or due to lack of knowledge of the interaction problem)
- Developed algorithm that can take such advice into account in a robust way
- Based on heuristic that constantly monitors whether advice is useful and whether other agents are following advice, too

Communication in MARL

- Assume communication actions are available in MARL domains (very low cost, no immediate impact on reward)

Communication in MARL

- Assume communication actions are available in MARL domains (very low cost, no immediate impact on reward)
- Can agents learn to use these signals to improve their coordination abilities?

Communication in MARL

- Assume communication actions are available in MARL domains (very low cost, no immediate impact on reward)
- Can agents learn to use these signals to improve their coordination abilities?
- Ongoing work: currently able to show that listening or speaking can be learnt when the respective other functionality is in place, but hard to learn both at the same time

Communication in MARL

- Assume communication actions are available in MARL domains (very low cost, no immediate impact on reward)
- Can agents learn to use these signals to improve their coordination abilities?
- Ongoing work: currently able to show that listening or speaking can be learnt when the respective other functionality is in place, but hard to learn both at the same time
- In the future we want to extend this to non-cooperative domains to investigate deception in communication

Learning of dialogue strategies

- Assume a set of (temporally extended) dialogue strategies is given, how to learn when to use which one of them?

Learning of dialogue strategies

- Assume a set of (temporally extended) dialogue strategies is given, how to learn when to use which one of them?
- Orthogonal to previous cases: learning about communication rather than communication to improve learning performance

Learning of dialogue strategies

- Assume a set of (temporally extended) dialogue strategies is given, how to learn when to use which one of them?
- Orthogonal to previous cases: learning about communication rather than communication to improve learning performance
- Developed mechanism for induction of dialogue patterns (hybrid symbolic-numerical methods)

Learning of dialogue strategies

- Assume a set of (temporally extended) dialogue strategies is given, how to learn when to use which one of them?
- Orthogonal to previous cases: learning about communication rather than communication to improve learning performance
- Developed mechanism for induction of dialogue patterns (hybrid symbolic-numerical methods)
- Applied hierarchical reinforcement learning for dialogue strategy selection

Learning of dialogue strategies

- Assume a set of (temporally extended) dialogue strategies is given, how to learn when to use which one of them?
- Orthogonal to previous cases: learning about communication rather than communication to improve learning performance
- Developed mechanism for induction of dialogue patterns (hybrid symbolic-numerical methods)
- Applied hierarchical reinforcement learning for dialogue strategy selection
- Evaluated in domain of strategic linkage between web sites

Summary

- Collaborative multiagent learning: mainly interested in

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems
  - methods for cautious/strategic information sharing among learners

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems
  - methods for cautious/strategic information sharing among learners
- Strategic multiagent learning: mainly interested in

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems
  - methods for cautious/strategic information sharing among learners
- Strategic multiagent learning: mainly interested in
  - trying to tackle the fundamental problem of non-stationarity by developing appropriate heuristics (e.g. stereotyping, switching, appropriate state abstractions)

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems
  - methods for cautious/strategic information sharing among learners
- Strategic multiagent learning: mainly interested in
  - trying to tackle the fundamental problem of non-stationarity by developing appropriate heuristics (e.g. stereotyping, switching, appropriate state abstractions)
  - learning how to communicate better, and communicating to learn better (closely related to my other interest which is agent communication)

Summary

- Collaborative multiagent learning: mainly interested in
  - increasing learner autonomy in distributed learning systems
  - methods for cautious/strategic information sharing among learners
- Strategic multiagent learning: mainly interested in
  - trying to tackle the fundamental problem of non-stationarity by developing appropriate heuristics (e.g. stereotyping, switching, appropriate state abstractions)
  - learning how to communicate better, and communicating to learn better (closely related to my other interest which is agent communication)
- Any questions?