Introduction to Multiagent Systems Research

Michael Rovatsos
School of Informatics
The University of Edinburgh
mrovatso@inf.ed.ac.uk

EASSS 2014, Chania, Greece, 14th July 2014

What is agents research?

Agents research = research about agents

- What are agents?
  - Autonomous, situated, persistent entities that interact with the environment through “sensors” and “effectors”
- What are computational agents?
  - Those implemented on computer software or hardware
- What are intelligent agents?
  - Reactive, proactive, rational, socially able agents
- As the area grew out of AI, there is a strong bias toward associating “agency” with “intelligence”

Pitfalls of the "agents" concept

- The concept is excessively broad, under-specifies class of systems in question
  - What distinguishes them really from other systems?
  - Thermostat argument, vagueness of "autonomy"
- Anthropomorphism
  - That's OK when this is the research objective
- Scope of functionality
  - Does an agent have to be able to do "everything"?
- How important is the "multi" according to this view?
  - Does "agent" really say anything about "multiagent"?

Pitfalls of the "agents" concept

- Abuse as purely conceptual notion
  - Use of the metaphor for traditional systems ("agentification")
  - Use of agent-based methods (e.g. programming languages) for conventional systems
- Association with quality of solution
  - Concept suggests "comprehensive", not partial, solutions, and may sound better than "program"
- Debates sometimes not very fruitful
  - There seems to be a vague implicit concensus at least within the community - ambiguity can be a good thing

The "holistic position"

Agents research is about integration of system components in a single system . . .

- Integrating an system in its environment
  - Autonomous, situated and embodied agents
- Integrating a system with its user
  - Personal assistants
- Integrating various capabilities in one system
  - Agent architectures

The "holistic position"

. . . and about integration and interaction among separate system components

- Integrating systems with shared goals in the same environment
  - Cooperative and benevolent agents
- Integrating several systems with different goals in the same environment
  - Self-interested agents – really novel aspect of computation
- Agents research = the science of integration?
  - Maybe, in its infancy. Another reason for advocating breadth?

The history of agents

- The Turing Test (1950)
  - Implies embodiment and interaction with an “agent”
- A universal modular actor formalism for artificial intelligence (1973)
  - First “agent” model of an autonomous interacting entity
- The Contract-Net Protocol (1980)
  - A first concrete procedure for task allocation
- Deals Among Rational Agents (1984)
  - Self-interested agents appear

The history of agents

- The use of meta-level control for coordination in a distributed problem solving network (1983)
  - Distributed problem solving, blackboard systems
- A robust layered control system for a mobile robot (1986)
  - Subsumption architecture – “Elephants don’t play chess”
- Plans and resource-bounded rational reasoning (1988)
  - The Belief-Desire-Intention model of agency
- Intention is choice with commitment (1990)
  - Formal foundations for deliberative architectures

The history of agents

- Communicative actions for artificial agents (1995)
  - Formal foundations for agent communication
- On the acceptability of arguments (1995)
  - Abstract framework for agents that argue
- Provably optimal boundedly-rational agents (1995)
  - Decision-theoretic AI and meta-reasoning
- Collaborative plans for complex group action (1996)
  - Foundations of teamwork and joint planning

The history of agents

- Shopbot economics (1999)
  - Agents and electronic commerce
- Algorithm for optimal winner determination in combinatorial auctions (2002)
  - Generalising resource allocation problems
- When are elections with few candidates hard to manipulate? (2007)
  - Complexity analysis of social choice problems
- The complexity of computing a Nash equilibrium (2009)
  - How hard is equilibrium computation in general?

Technology context

This research has been pursued against the backdrop of . . .

- the emergence of distributed computing
- the development of high-level programming languages
- enormous advances in the performance of computers
- radical shifts toward automation in all parts of society
- the evolution of the Internet and other global networks
- the increasingly ubiquitous presence of computers

Important to remember: most of agents research would be irrelevant without these

Agents research topics

- Agent architectures
  - How to structure the design of an agent
- Agent reasoning & learning mechanisms
  - What algorithms to run on a single agent
- Agent interaction and communication mechanisms
  - How to make agents talk to each other
- Agent-based simulation of humans
  - How to build lifelike/believable agents
- Agent programming languages
  - How to implement agents

Agents research topics

- Multiagent coordination methods
  - How to make agents consider others
- Multiagent collaboration techniques
  - How to make agents work together
- Multiagent reasoning mechanisms
  - How to make joint decisions and reach agreement
- Multiagent-based social simulation
  - How to reproduce the behaviour of human collectives
- Multiagent learning techniques
  - How to learn in the presence of other learners

Methods used

- Social science, psychology, philosophy, linguistics
- Knowledge representation and reasoning
- Search, planning, and constraint satisfaction
- Formal methods, logic and automated verification
- Software engineering, human-computer interaction
- Complexity and intractability analysis
- Game theory, economics, social choice theory
- Machine learning and pattern recognition
- Probabilistic reasoning, information & decision theory
- Efficient algorithms and combinatorial optimisation

Relationships to other fields of computing

- Decentralisation & interaction
  - Distributed systems, service-oriented computing, P2P etc
- Embodiment & autonomy
  - Robotics, sensor networks, mobile & pervasive computing
- Conflict of interest
  - Algorithmic game theory, eCommerce, game-playing AI
- Integration
  - Robotics, ”general AI”, collective intelligence, games
- Learning, emergence & self-organisation
  - Machine learning, evolutionary computation & swarms
- Use of methods
  - Theoretical computer science, numerical simulation, programming languages, human-computer interaction

Relationships to other disciplines

- Decentralisation & interaction
  - Sociology, social psychology, biology, linguistics
- Embodiment & autonomy
  - Electronics, mechanical & control engineering, design sciences, architecture
- Conflict of interest
  - Economics, politics, biology
- Integration
  - Design sciences, engineering, management science
- Learning, emergence & self-organisation
  - Biology, psychology, sociology, cybernetics
- Use of methods
  - Mathematics, philosophy, economics, biology

Single-agent reasoning: reactive approaches

- Dominating paradigm: Markov Decision Processes (MDPs)
  - Essentially stochastic finite-state machines with utility-based definition of design objective and preferences over states (and actions)
- Simple, domain-independent model, applied to many domains across AI
- Can be seen as a continuation of work on reactive architectures
  - subsumption architecture, situated automata, etc

Single-agent reasoning: deliberative approaches

- Dominant paradigm: Belief-Desire-Intention model of practical reasoning
- Based on Bratman's model of human practical reason:
  - Beliefs determine how desires (long-term preferences of agent) are transformed into intentions (concrete adoption of plans in continuous loop)
- Distinction between deliberation (commitment to ends) and means-ends reasoning (commitment to means)
- Intention theory enforces rationality constraints

Single-agent reasoning: deliberative approaches

- Planning and execution interleaved with deliberation and meta-level reasoning (intention reconsideration)
  - Essentially a method for reactive (continuous) planning, where existing intentions and plans together with external events trigger new intentions
- Paradigm emerged from knowledge-based AI/planning
  - Predecessors: production systems, SOAR, early agent programming languages (AGENT0, MetateM etc)

Multiagent reasoning: cooperative systems

- Setting: decentralisation of information and control when computational components are agents
- Cooperative and benevolent agents
  - goal is ultimately shared, no danger of malicious behaviour, no conflict of interest
- Dominating paradigm: “teamwork” among collaborating teammates
- Concerned with all stages of distributed problem solving
  - team formation, agreement on goals, agreement on means, joint execution and monitoring

Multiagent reasoning: cooperative systems

- Research issues:
  - How to ensure consistency of global state (joint intentions framework, partial global planning, blackboards)
  - How to allocate tasks and resources (contract nets, coalition structure generation)
  - How to deal with task execution and failure (communication mechanisms, joint plan repair and replanning etc)
  - How to combine partial results

Multiagent reasoning: non-cooperative systems

- Generalisation of previous scenario allowing self-interest
  - This small change makes a world of difference!
- Main research themes:
  - How to behave in the presence of other self-interested agents to further one's goals
  - How to reach agreement in the presence of conflict of interest
- Area stands in tradition of game theory
  - The two questions above roughly correspond to non-cooperative and cooperative game theory
  - Relies on mathematical models of agents, actions, outcomes and preferences

Multiagent reasoning: non-cooperative systems

- Can be broken down into more concrete sub-topics:
  - Strategic behaviour: how to act given different preferences of others
  - Resource/task allocation: how to distribute goods given agents' preferences
  - Coalition formation: how to form groups and split the gain obtained from collaboration
  - Social choice: how to make group decisions given everyone's preferences
  - Bargaining: how to negotiate effectively in the presence of conflict of interest
- Computational focus (unlike standard game theory)
  - algorithms for mechanism design, equilibrium computation, coalition structure formation, winner determination in auctions, representation languages

Multiagent approaches to standard AI problems

There is also much work on applying an "agents" approach to more traditional AI problems:

- Distributed search
- Distributed constraint satisfaction
- Distributed knowledge representation
- Distributed theorem proving
- Distributed planning
- Distributed game-playing
- Distributed machine learning
- Distributed sensing and acting

Multiagent interaction and coordination methods

- Agent communication languages, protocols, semantics, ontologies
  - A lot of standardisation work in 90s and agent plaftorms
  - Inherent problem: addresses only part of a whole MAS
  - Considered “done” by many researchers, a standard part of applied MAS design
- Norms, institutions, and organisations
  - System-level models of behavioural expectations as a control mechanism
  - Mostly modelling, logical specification, few algorithms
  - Biased toward “regimentation” of open systems (tension!)

Multiagent interaction and coordination methods

- Trust and reputation
  - Dominant paradigm for mutual modelling in open MAS
  - For some just a multiagent learning problem
  - For others lots of “human” aspects
- Argumentation
  - Integration of conflicting knowledge, non-monotonic reasoning tradition
  - Dominant approach for logic-based dealing with uncertainty
  - Little account of strategic behaviour and practical algorithms

Agent systems engineering

- Agent-oriented software engineering (beyond AI agenda)
  - Methodologies, programming languages, tools, verification
- Mostly based on deliberative models and declarative programming
  - Much current work part of research on BDI
- Main contribution: agent-style programming paradigm
  - Not established as mainstream programming paradigm
- Usual problems of evaluation
  - Needs large community to assess practical value
  - Multiagent programming contest is useful

Human-oriented agent systems

- Also outside standard AI agenda
- Virtual agents
  - User assistants & tutors, storytelling, believable behaviour
  - Related to animation, user modelling, intelligent tutoring
  - Usually rather applied and interdisciplinary
- Social simulation
  - Use agent-based methods for social science
  - Usually large-scale systems, simple agents
  - Problem of validity of artificially generated data
- Hard to identify domain-independent performance measures, require human-based validation (often qualitative)

A Critical View

- Very high diversity within the area, many sub-areas lack clear problem and solution formulations
  - A lot of work overlaps with other fields
- Fundamental barriers to cross-fertilisation
- Few “reference systems” within which to test new components
  - Multiagent mechanisms assume “given” agents
- Several attempts to introduce benchmarks through competitions
  - RoboCup (+Rescue), ART, TAC/CAT, progr. contest
- Novel application areas raise new challenges
  - electronic markets, social networks, human-based computation, mobile computing

Conclusions

- Agents and multiagent systems: a very diverse, inclusive, dynamically changing field
- Characterised by mix of formal and informal, theoretical and applied
- A variety of diverse approaches, “big achievements” of area not always clear
- Hard to get a big picture of the whole field, increasing fragmentation
- Methodological problems that arise from these issues

Conclusions

Some personal thoughts

- An exciting field to be in, with very open-minded people
- Sometimes dangerous temptation to cover everything
- Tension between “small” and “big” approaches
- We have the right tools for future computer applications
- It is time we deliver the technology that shows that!

The End

Thank you for your attention and participation!