AGENT-BASED MODELING OF AGRICULTURAL LAND USE ON TERAGRID

Anand Padmanabhan¹, Wenwu Tang², Shaowen Wang¹

CyberInfrastructure and Geospatial Information Laboratory (CIGI)
National Center for Supercomputing Applications (NCSA)
¹University of Illinois at Urbana-Champaign

²University of North Carolina at Charlotte
Outline

- Science Problem
- Background
- Agent-based Land Use Model
- Experiments
- Conclusions and Future work
Science Problem

- Investigation of complex land use systems
  - Midwest is the world’s bread basket
- Policies and technologies affect the production and use of crops
  - E.g. price supports, biofuel mandates
  - Careful assessment of environmental impacts, economic costs and benefits of policies is essential

Photo sources: http://www.ars.usda.gov/is/graphics/photos/k5051-5.jpg
http://www.isgs.illinois.edu/nsdihome/webdocs/landcover/Revised2007IllinoisCDLMap_screen.jpg
Study Area

Map of study area (the Midwest) for land use simulation

Map showing agricultural land parcel data for Champaign county in Illinois
Agent-based models (ABM) are widely applied to gain better understanding of dynamics of complex adaptive systems, characterized by self-organization, nonlinearity, path-dependence, emergence, adaptation (Levin 1998, Parker et al. 2003; Bennett and Tang 2006; 2008)

Macro level studies have been conducted using mathematical models and empirical observations

By simulating the effect of policy at micro level ABM are used to understand macro-level observations
Agents simulate farmers with decision making capabilities

- Farmers are guided and constrained by their individual objectives
- Machine learning algorithms are used to represent adaptive decision making of agents
- Farmers get influenced by their social network

Overall landuse impact of their decisions on broader ecosystems can be evaluated
Outline

- Science Problem
- Background
- Agent-based Land Use Model
- Experiments
- Conclusions and Future work
Agent-based Models (ABM) Introduction

- Computational approach for the modeling of complex adaptive spatial systems (Bennett and Tang 2008)
  - Agents and their spatially-explicit environments
  - Agent-agent and agent-environment interactions

Adapted from Russell and Norvig (1995)
ABM Characteristics

- Agents tend to be independent, while they could be influenced by other agents through interaction
  - Decentralized and concurrent

- Agents are adaptive
  - Individuals adjust their behavior to improve performance in response to stimuli (Holland 1975)
  - Bounded rationality (Simon 1956)
  - Spatially explicit cognition and learning behavior

- ABM are compute- and data-intensive
Intelligent Agents

- Agents in land use model are infused with learning capabilities
- Multiple techniques are available for integrating machine learning and ABM (Russel and Norvig 1995)
  - Predefined behavior
    - Inductive
  - Learned behavior
    - Reinforced mechanisms
      - Individual-level: Reinforcement learning (used in this study)
      - Population-level: Evolutionary learning
The objective of a reinforcement-learning agent is to maximize its rewards:

\[
\text{Max} \quad : Q(s, a) \\
\forall s \in S, \forall a \in A
\]

Learning action-value functions (Q learning)

\[
Q(s_t, a_t) = Q(s_t, a_t) + \alpha \delta_t \\
\delta_t = r(s_t) + \gamma \max_{a_{t+1} \in A(s_{t+1})} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)
\]

Where
- \( S \): States
- \( A \): Actions
- \( Q \): Action-value functions
- \( \alpha \): Learning rate
- \( \gamma \): Discount factor
- \( \delta_t \): Temporal-difference errors at time \( t \)
ABM - Computational Challenges

- The agent-based land use model is compute- and data-intensive due to:
  - Frequent agent-agent and agent-environment interactions;
  - Multiple model repetitions necessary to obtain statistically reasonable results for a number of stochastic parameters involved; and
  - Assimilation of a number of large spatio-temporal datasets by every agent.

- High-performance and high-throughput approaches are needed.
Outline

- Science Problem
- Background
- Agent-based Land Use Model
- Experiments
- Conclusions and Future work
Land Use Modeling Framework

- **Agent**
  - Sensors
  - Learning Algorithm
  - Knowledgebase (Learned and Predefined)
  - State (Environment and Events)
  - Effectors

- **LandAgent Framework**
  - Stimulus
  - Iterate
  - Response
  - t
  - t+1
  - Stimulus
  - Response

- **Environment**
  - Spatial Neighborhood Graph
  - Parcel
  - Ethanol Plants
  - Examples
  - Line
  - Polygon
  - Point
  - Spatial Datatypes

- **Events**
  - Datasets
  - Geographic Datasets
  - e.g. Change in Policy, Price
LandAgent Simulation Model
Characteristics of Agents

- **Context-aware**
  - Agents can perceive contextual information in their environments

- **Adaptable**
  - Agents have ability to adapt their behavior to achieve their goals in response to external events or changes in their environment

- **Cooperation**
  - Agents can communicate state and knowledgebase
Model Execution

- **Step 1: Read in input datasets**
  - Includes farmers, parcels and events
  - Number of iteration in a run is determined by number of events

- **Step 2: During every training iteration**
  - Send state information (knowledgebase) to neighbors
    - Neighbors are determined by spatial proximity
    - MPI used
  - Conduct Q-learning on result from the previous iteration
    - Rank-based selection is applied
    - Enhance learning using knowledge shared by neighbors
    - Update knowledgebase
Model Execution (cont)

- **Step 3:** During every validation cycle
  - Land use decision made on deterministic and probabilistic selection
  - Calculate the objective function

- **Step 4:** For every event iterate through decision making steps (2-3) for each farmer
  - The first $\frac{3}{4}$ events are used for training the agent
  - The last $\frac{1}{4}$ events are used for validation

- **Step 5:** Repeat steps 2-4 a number of times to ensure convergence of the learning algorithm
Outline

- Science Problem
- Background – Agent-based Model
- Agent-based Land Use Model
- Experiments
- Conclusions and Future work
Experiment Design

- Experiments are designed with a goal to evaluate the following aspects of the framework:
  - Agent’s learning performance
  - Effects of social network on agent’s learning
  - Ability of the framework to handle multiple spatial scales
  - Scalability to large number of processors

- Spatial datasets consist of land parcel data.

- Price and crop yield data for the past fifty years
  - Data source: http://www.farmdoc.illinois.edu/

- Initial experiments consider two crop types - corn and soybean

- Experiments were conducted on NCSA Abe
Result Maps

- Results show land use pattern from a validation run
  - Champaign county
Learning Performance

- **Scenarios**
  - Without communication among neighbors (i.e. learning from agent’s own behavior)
  - With communication (i.e. neighbors also influence learning)
    - Four nearest neighbors are identified for every parcel
    - A farmer’s neighborhood graph is defined by the union of the neighbors of all the parcels s/he owns
Learning Performance

Graphs averaged with over 2750 farmers and 10000 parcels of farm land for every run.
Performance Testing at Multiple Spatial Scales

- Evaluate runtime performance
  - mass spatial scales
    - County
      - Campaign dataset has over 2750 farmers and 10000 parcels
    - State
      - Over 275,000 farmers and 1M land parcels
      - Data extrapolated using Champaign’s datasets
  - Compare performance by including
    - Reinforcement learning (farmer only)
    - Reinforcement learning (farmer + neighbors)
      - Performance scaling with and without communication
- Scalability testing (2048 processors)
Scaling Performance
Reinforcement Learning (self)

Execution time (in secs)

Number of CPUs

8 16 32 64 128 256 512 1024 2048

Spatial Scale: State
Spatial Scale: County

Scaling Performance
Reinforcement Learning (self + neighbor)

Execution time (in secs)

Number of CPUs

8 16 32 64 128 256 512 1024 2048

Spatial Scale: State
Spatial Scale: County
Speedup
(only self reinforcement)
Outline

- Science Problem
- Background
- Agent-based Land Use Model
- Experiments
- Conclusions and Future work
Summary

- Understanding complex land use systems is important.
- Agent-based simulation models are well suited to gain such understanding.
  - By simulating the effect of events on individual farmers, overall macro-level impact may be understood.
  - Machine learning algorithms are applied to agents.
    - Reinforcement learning.
    - Social networks (neighbors) influence agents.
- Since simulated agent behavior is stochastic, a large number of repetitions are needed to get valid results.
- Both high-throughput and high-performance computing approaches can be used.
Future work

- Enhance the underlying model
  - Incorporate policy events, extreme environmental events
  - Conduct larger scale performance evaluation experiments
- Develop strategies to decrease communication overhead
  - Use spatial knowledge of agent locations to decompose agents among processors
Thank You!