Reinforcement learning for crop-management: a sequential decision-making under uncertainty approach

Philippe Preux
philippe.preux@inria.fr

SequeL

Talk prepared with Romain Gautron (CGIAR-CIAT CIRAD) and Odalric-Ambrym Maillard (Inria).

Disclaimer/General idea of this talk

- Create synergies between RL and Agronomy communities
- Simplistic models to be presented
- More sophisticated, realistic, settings are (obviously) within reach

Complementary session for Q&A

little time today

$\Rightarrow$ Q&A Thursday, 1pm

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment
- multiple dimensions of agricultural decision-making
  - economical
  - environmental
  - social
  - quality requirements

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment
- multiple dimensions of agricultural decision-making
  - economical
  - environmental
  - social
  - quality requirements
- constrainsts: money, labor power, climate, ...

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment
- multiple dimensions of agricultural decision-making
  - economical
  - environmental
  - social
  - quality requirements
- constrainsts: money, labor power, climate, ...
- sustainable-intensification: decision making at landscape level

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment
- multiple dimensions of agricultural decision-making
  - economical
  - environmental
  - social
  - quality requirements
- constrainsts: money, labor power, climate, ...
- sustainable-intensification: decision making at landscape level
- climate change: adaption and mitigation

Problem statement

"Crop-management is the logical and ordered combination of techniques applied to a plot to obtain a given production" [Sebillotte, 1974]

- ordered series of interdependent operations
- stochastic environment
- multiple dimensions of agricultural decision-making
  - economical
  - environmental
  - social
  - quality requirements
- constrainsts: money, labor power, climate, ...
- sustainable-intensification: decision making at landscape level
- climate change: adaption and mitigation

$\Rightarrow$ increasing complexity in objectives

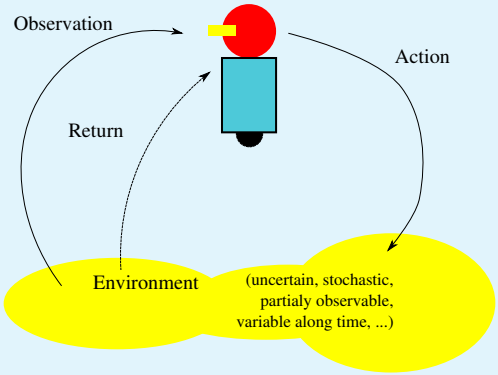
$\Rightarrow$ increasing complexity with uncertainty

$\Rightarrow$ a need for dynamic policies

Sequential decision-making under uncertainty

Reinforcement learning

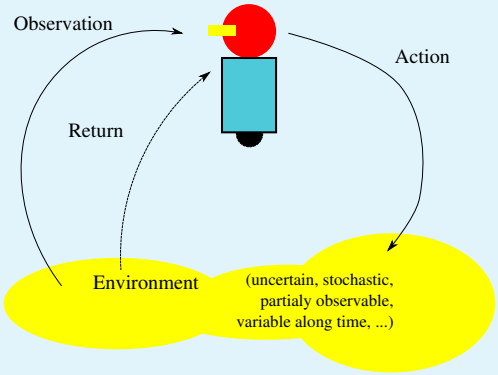
Learning to do.



Sequential decision-making under uncertainty

Reinforcement learning

Learning to do.

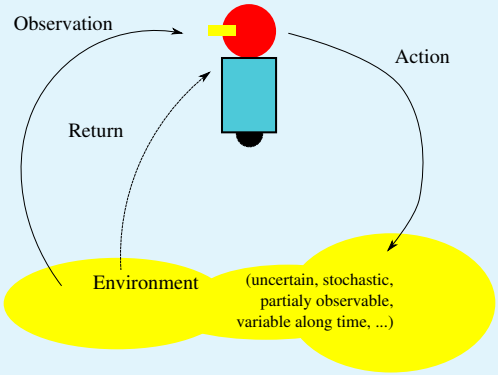


- Learning by trial-and-error.

Sequential decision-making under uncertainty

Reinforcement learning

Learning to do.

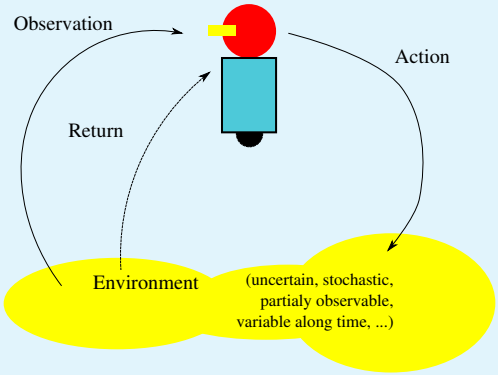


- Learning by trial-and-error.
- Trades-off short term and long term consequences of actions

Sequential decision-making under uncertainty

Reinforcement learning

Learning to do.

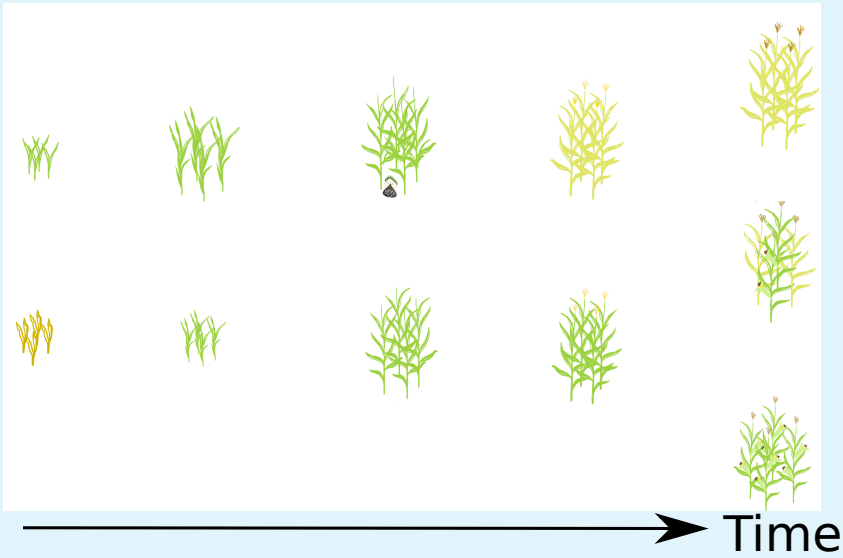


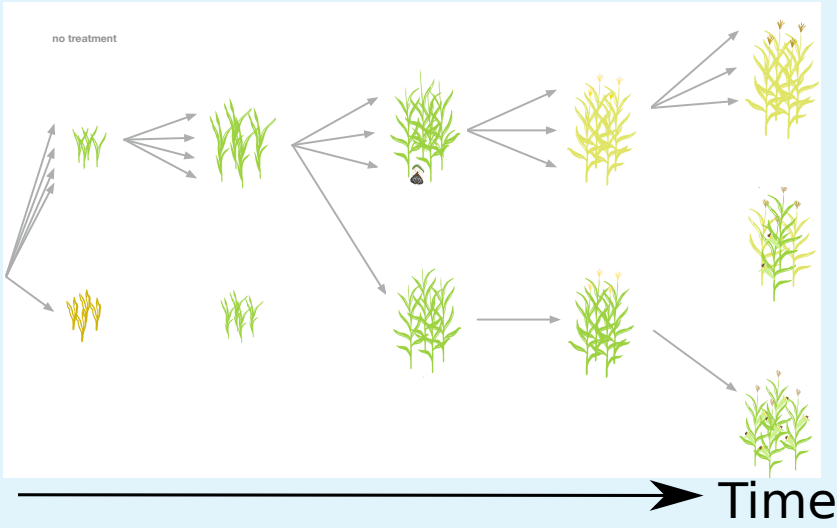
- Learning by trial-and-error.
- Trades-off short term and long term consequences of actions
- Reactive, and adaptive

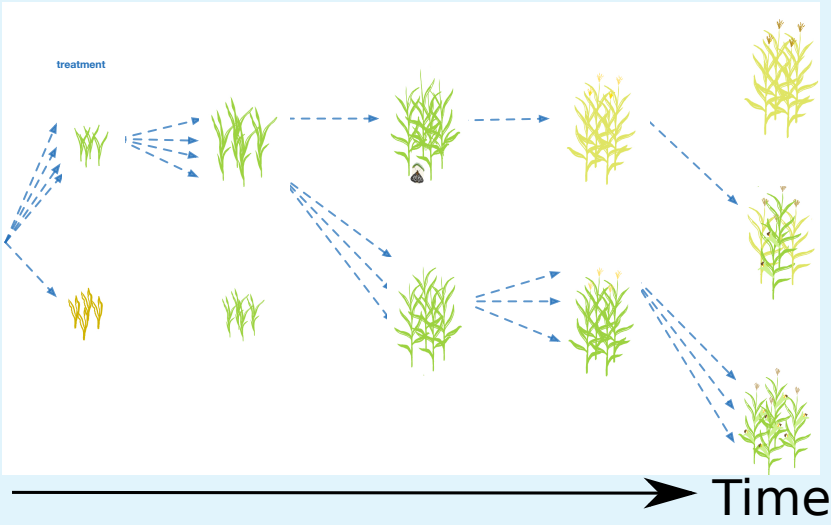
RL for agricultural (simplistic presentation)

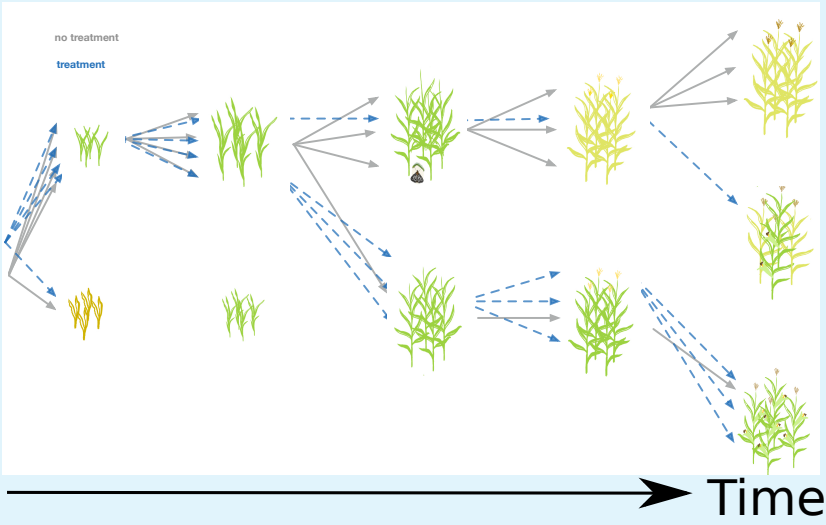
Series of actions to reach a given objective

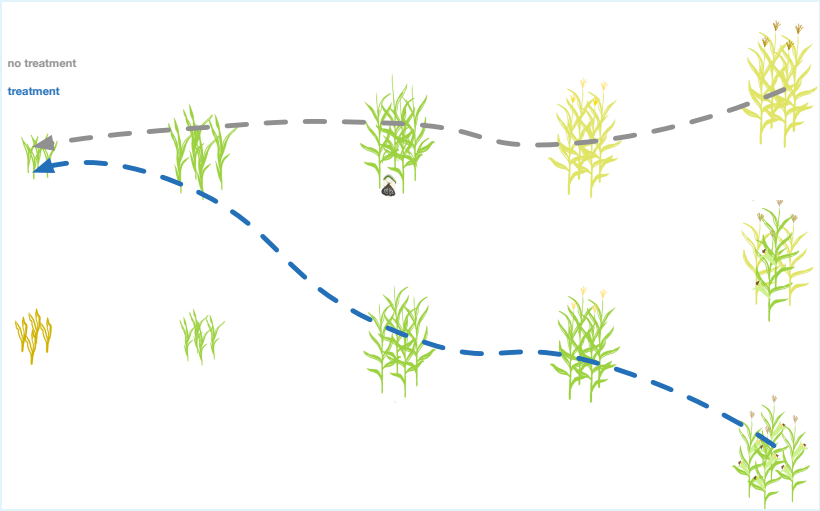
- actions
  - variety choice
  - tillage
  - sowing
  - weeding
  - ...
- observations (states)
  - photosynthetic activity
  - pests and diseases observations
  - growing stage
  - ...
- returns (rewards)
  - economical return
  - soil conservation
  - - plant fluorescence
  - ...


Time


no treatment
Time


treatment
Time

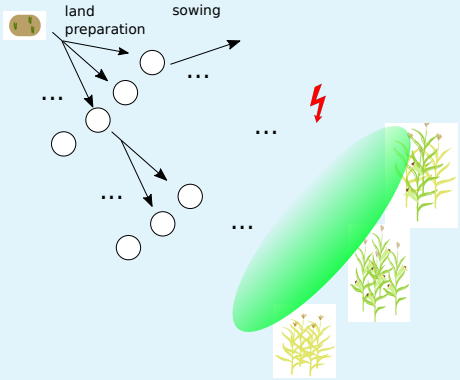

no treatment
treatment
Time



Back-propagate the consequences of the action.

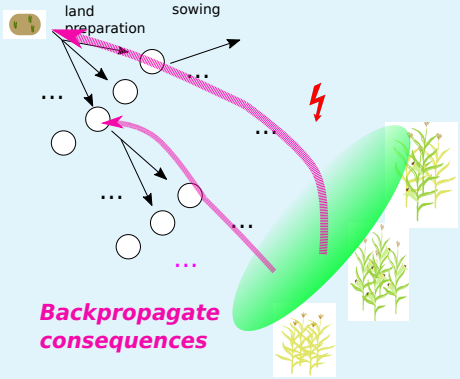
Reinforcement learning

Sequence of actions



Reinforcement learning

Sequence of actions

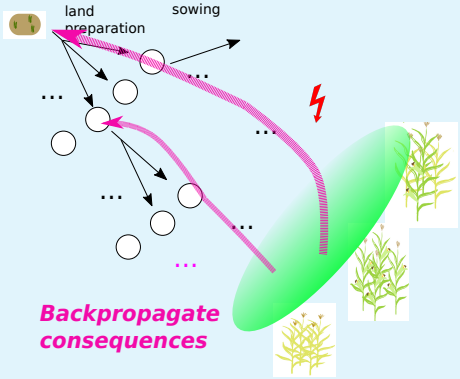


Back-propagate the consequences of actions

Learn the value of actions.

Reinforcement learning

Sequence of actions



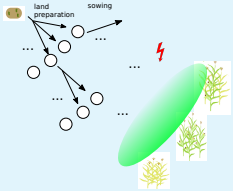
Back-propagate the consequences of actions

Learn the value of actions.

Explore new actions.

Reinforcement learning

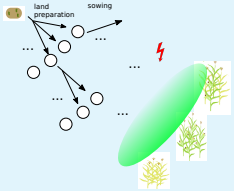
Autonomous dynamical system



A system evolving along time, driven by farmers' actions and environmental hazards (weather, pests, ...).

Reinforcement learning

Autonomous dynamical system



A system evolving along time, driven by farmers' actions and environmental hazards (weather, pests, ...).

Goal:

- increase the yield
- maintain/improve the quality of the soil
- meeting budget / HR / material constraints
- ...

Reinforcement learning

- Many successive decisions to make along time.

Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.

Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our **goal** is to design, and implement a system able to determine which decision to take, and when: a policy.

Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.

Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.
  - Set of states.

Reinforcement learning

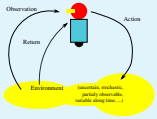
- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.
  - Set of states.
  - Set of actions possible in each state.

Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.
  - Set of states.
  - Set of actions possible in each state.
  - Probability of transition between states.

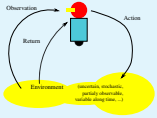
Reinforcement learning

- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.
  - Set of states.
  - Set of actions possible in each state.
  - Probability of transition between states.
- Control of this chain to perform a certain task = optimize a certain criteria: Markov decision problem



Reinforcement learning

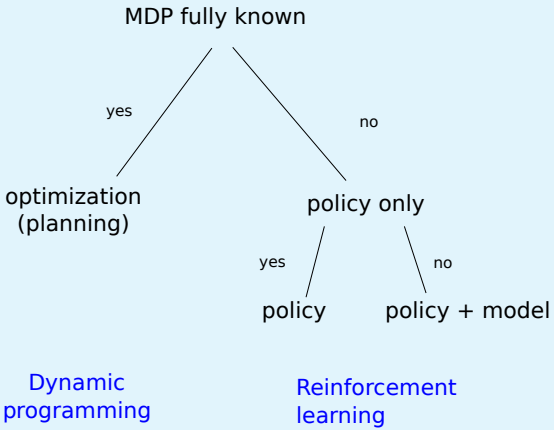
- Many successive decisions to make along time.
- Each decision has short and longer terms consequences.
- Our goal is to design, and implement a system able to determine which decision to take, and when: a policy.
- The system to control is modeled by a Markov chain.
  - Set of states.
  - Set of actions possible in each state.
  - Probability of transition between states.
- Control of this chain to perform a certain task = optimize a certain criteria: Markov decision problem



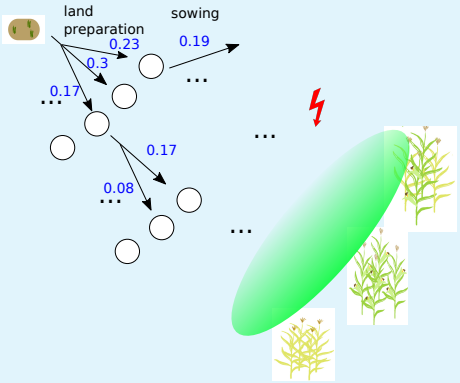
- Optimization criterion: crop yield, net revenue, soil conservation, ...

Markov decision problem

Different approaches



Learning a model



RL: renewed opportunities

- Almost nothing since the 90s ...:
  - ✓ Dynamic programming approaches, Bellman, 1950s
  - ✓ Modern RL: [Garcia, 1999] model-free approach off-line setting
- Renewed opportunities:
  - ✓ A new data context: sensors, remote sensing, drones, social networks...
  - ✓ New tools and computational resources.
  - ✓ Safe RL methods: cVaR, demonstrations from experts...
  - ✓ Many advances in machine learning more generally.

$\Rightarrow$ New possibilities for RL

$\Rightarrow$ We can envision RL for ag in an online, model-based, safe setup

Potential

- RL methods can fit farmers' reasoning
- RL methods can enhance decision-making: reactive policy, risk sensitive (various criterions)
- RL as learning tool
- Meta agent: dealing with policy at landscape matrix
- RL and environmental challenges

Challenges

- An unusual frame for RL
- **Need for RL oriented data**
- **Problem of model training and validation**
- Making explicit policies
- Variable objective and constraints
- **Interconnecting RL and Agronomy communities**

Take home message

- RL is a promising approach to improve crop management
- RL may be applied to other problems for agriculture
- requires RL oriented quality data
- some RL challenges to be tackled
- current efforts in Malawi by CGIAR/CIRAD/Inria

- **key is to create synergies**: collaborations are welcome
- Please check Romain's paper "Will Reinforcement Learning provide a fresh start to crop-management Decision Support Systems?" (under review)