Guided Tour of Machine Learning in Finance

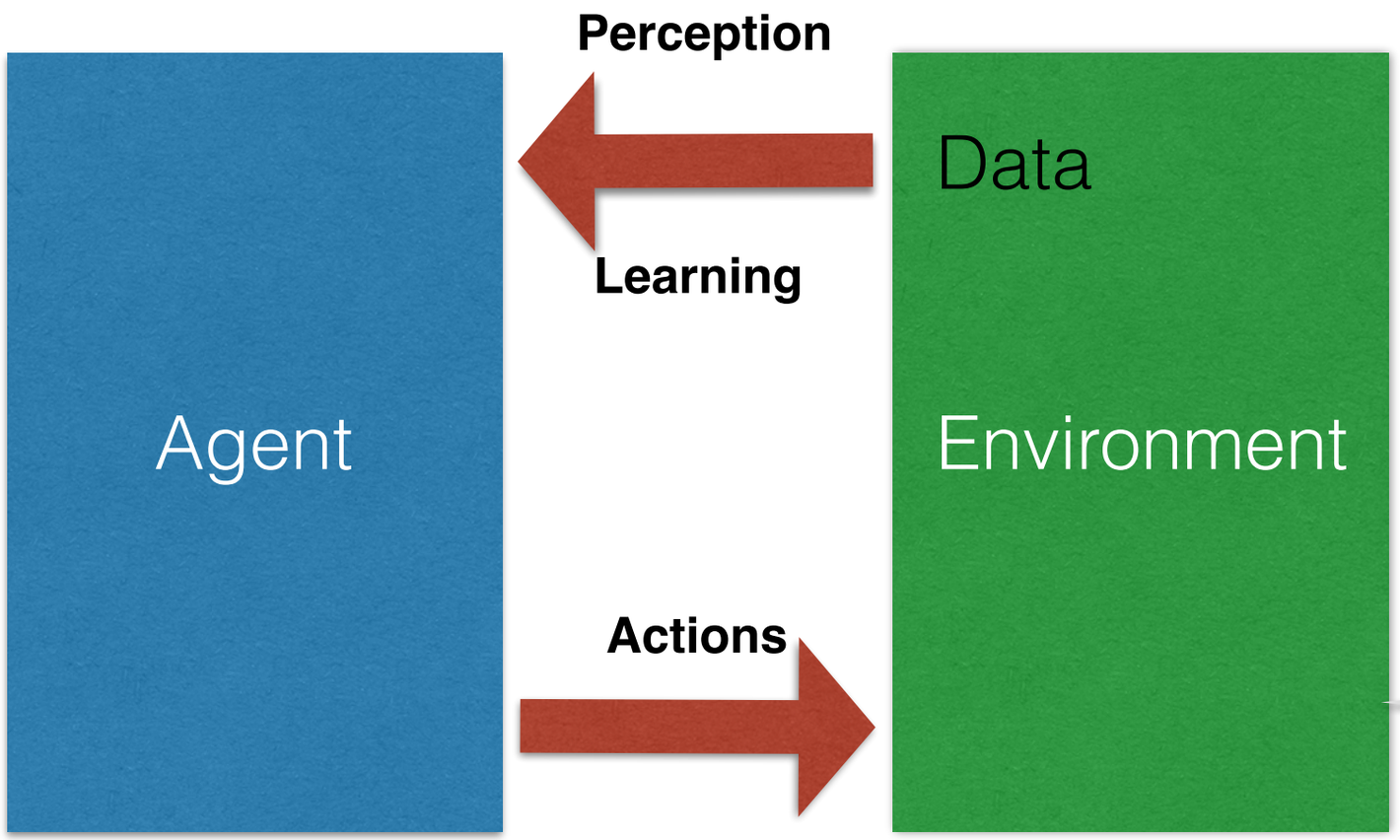
Week 4: Reinforcement Learning

4-2-2-Reinforcement Learning: core ideas

Igor Halperin

NYU Tandon School of Engineering, 2017

Types of ML tasks



The agent may not have access to streaming data from the environment (on-line learning) and learn instead in a batch mode (off-line) from data obtained from this environment.

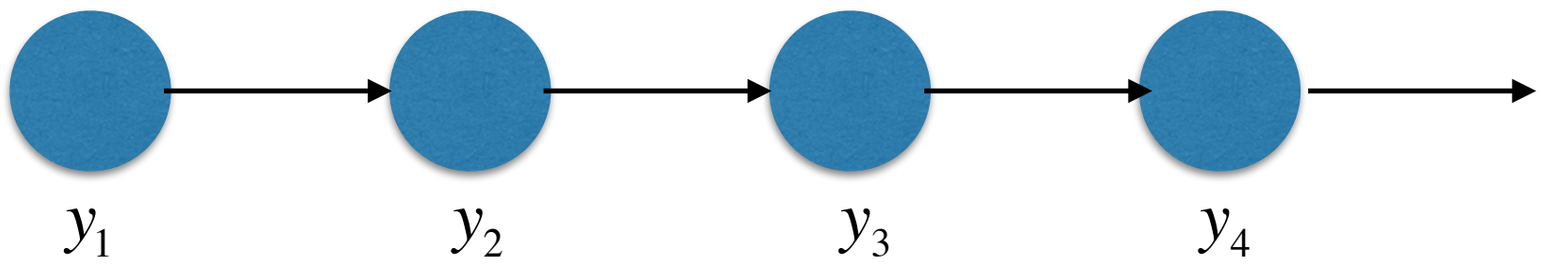
"**Perception tasks":** perception and learning from data. There is a fixed action, e.g. predict a loan default, classify an image, or translate a text. Regression and classifications are perception tasks. The output is a learned function of data $f(\mathbf{X})$

"**Action tasks**": the same as perception tasks, but there are multiple possible actions. For sequential (multi-step) problems, action tasks involve planning and forecasting the future.

State of environment as sequential data

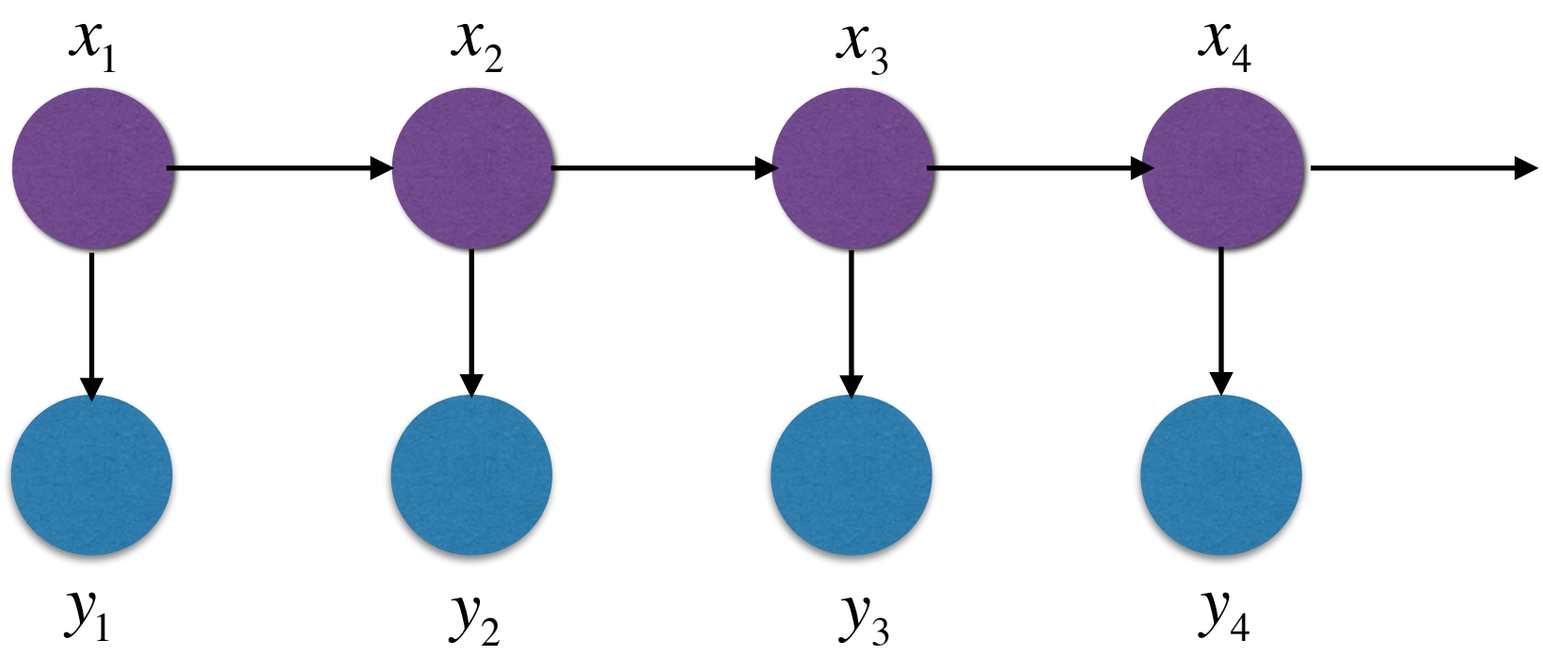
1. **Markov models:** $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$ - the observable environment

$$p\left(y^{(1)}, y^{(2)}, y^{(3)}, \ldots, y^{(T)}\right) = \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right) \rightarrow \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(n-1)}\right)$$

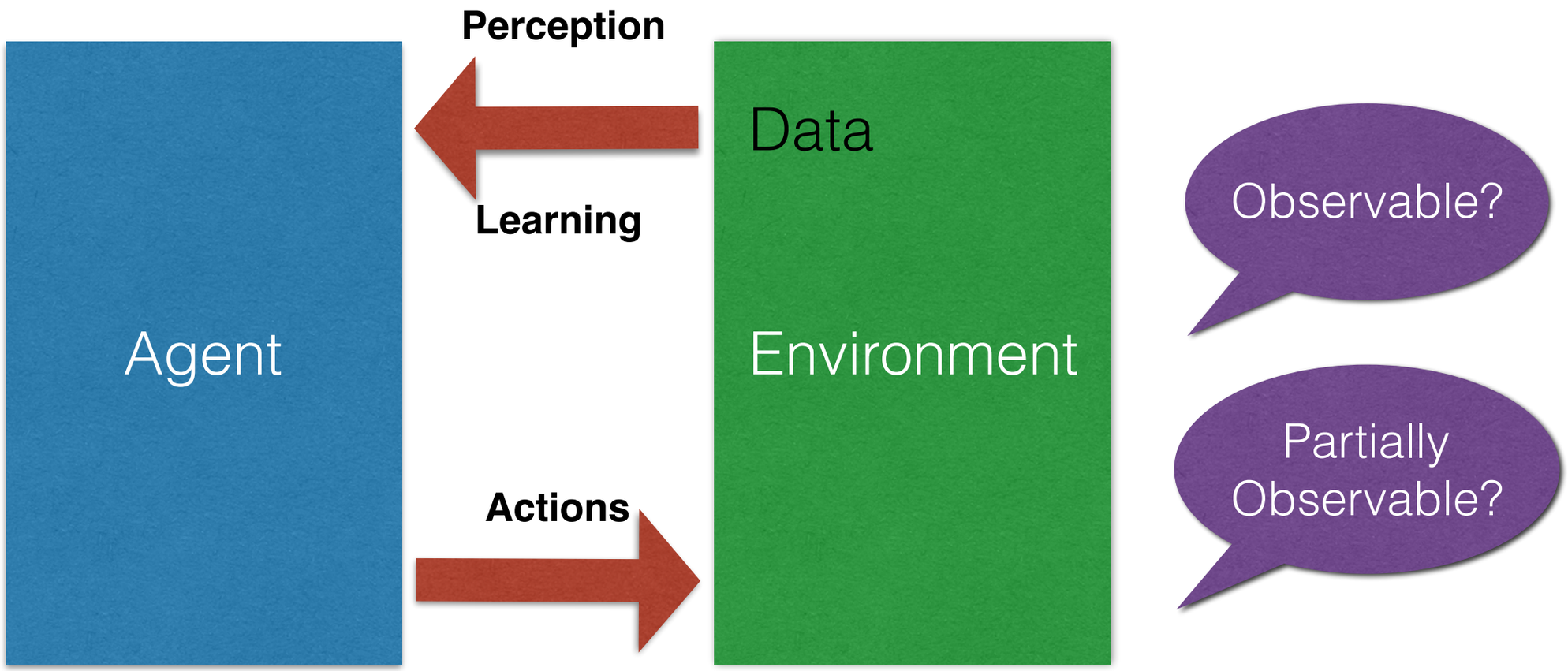


2. **Dynamic hidden (latent) variables models**, with an unobservable state $x^{(t)}$

- The hidden state is used for predictions $p\left(y^{(t)} \mid y^{(t-1)}\right) \rightarrow p\left(y^{(t)} \mid x^{(t)}\right)$.



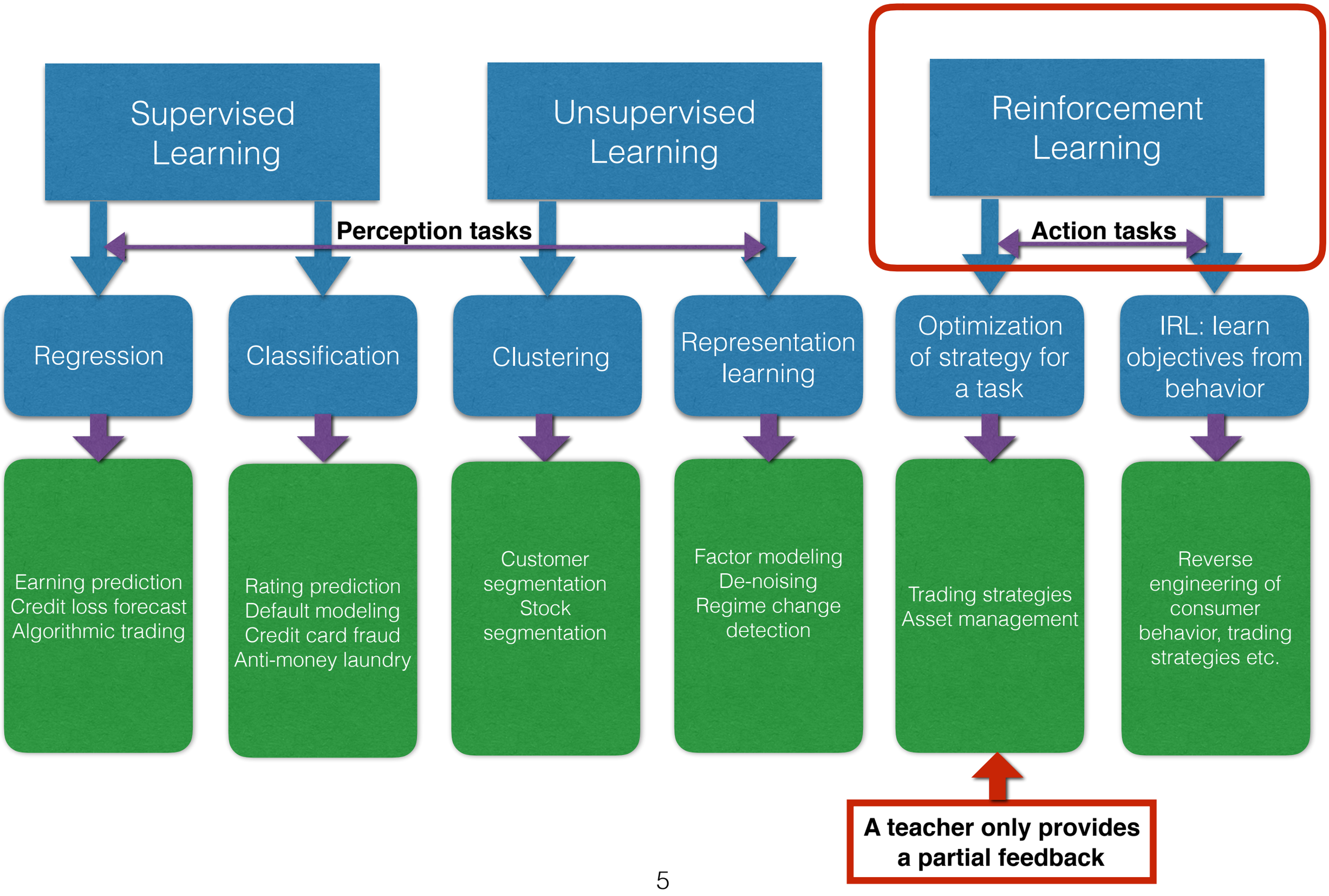
Observable vs Partially Observable Environment



Fully observable environment: video games, walking robots, self-driving cars, Go game, …

Partially observable environment: video games, finance, …

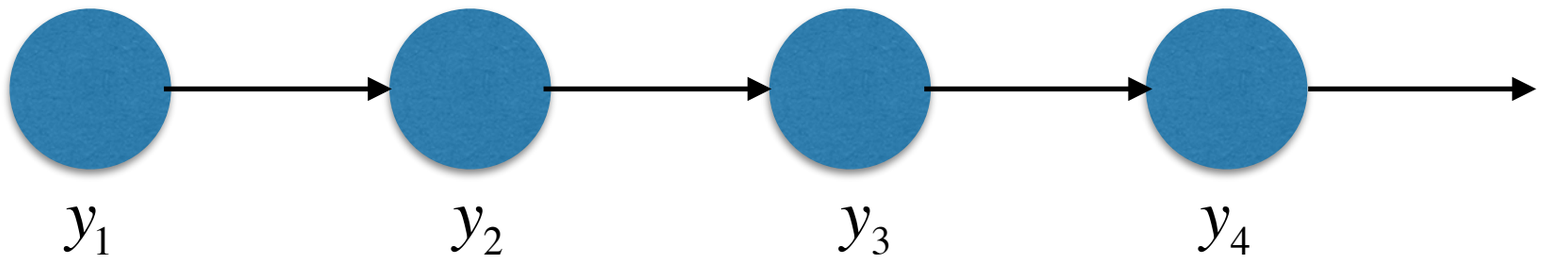
Machine Learning in Finance



Markov Decision Processes

1. **Markov models:** $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$ - the observable environment

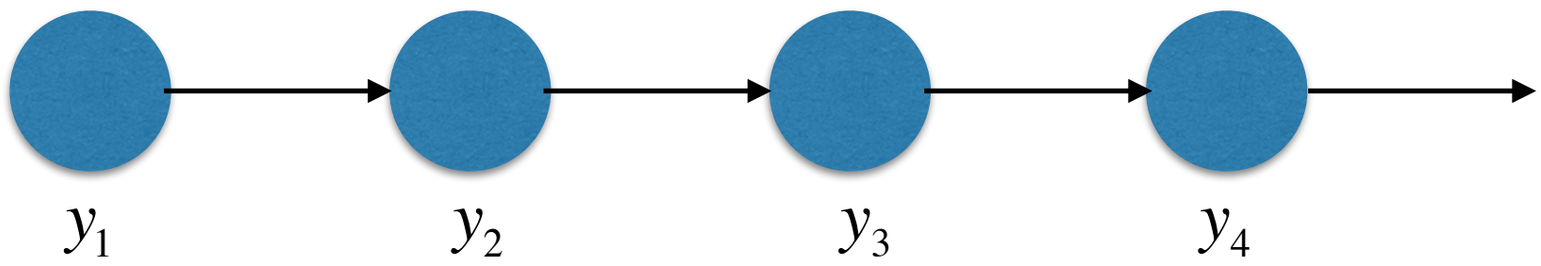
$$p\left(y^{(1)}, y^{(2)}, y^{(3)}, \ldots, y^{(T)}\right) = \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right) \to \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(n-1)}\right)$$



Markov Decision Processes

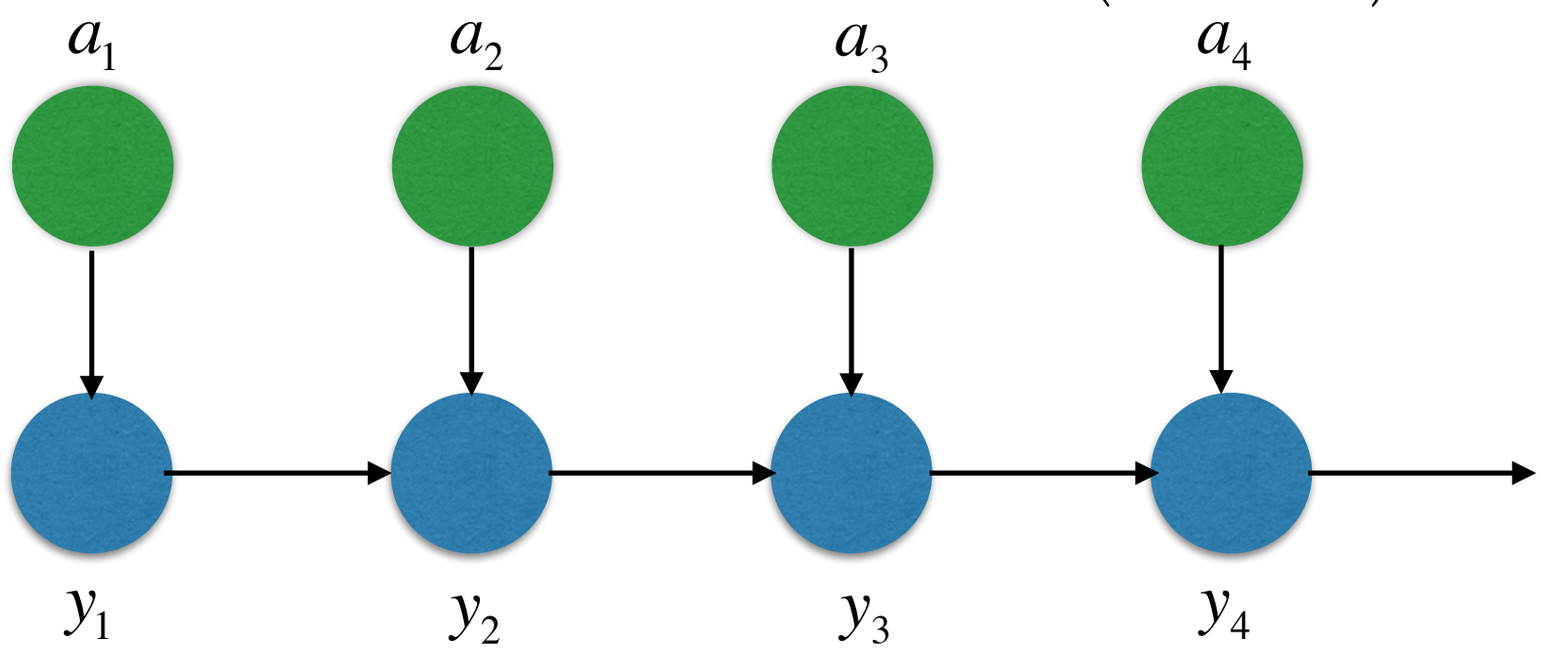
1. **Markov models:** $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$ - the observable environment

$$p\left(y^{(1)}, y^{(2)}, y^{(3)}, \ldots, y^{(T)}\right) = \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right) \rightarrow \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(n-1)}\right)$$



2. **Markov Decision Processes (MDP):** a Markov process controlled by agent's action $a^{(t)}$

- Agent's actions impact transition probabilities: $p\left(y^{(t+1)} \mid y^{(t)}\right) \rightarrow p\left(y^{(t+1)} \mid y^{(t)}, a^{(t)}\right)$



Control question

Select all correct answers

1. Markov Decision Processes model decisions as a Markov Process.
2. A fully observable environment can be described by a Markov process.
3. A partly observable environment can be described by a Dynamic Latent Variable model.
4. Decisions for a partially observable environment should be modeled using Latent Variables, to further decrease the number of observables in the system. This eases on data requirements, and if all components of the environment and actions are made unobservable, such model may not even need any data at all.

Correct answers: 2, 3