Fall 2010 Graduate Course on
Dynamic Learning

Chapter 1: Introduction

Byoung-Tak Zhang

Biointelligence Laboratory
School of Computer Science and Engineering &
Cognitive Science and Brain Science Programs
Seoul National University

http://bi.snu.ac.kr/

Outline

- Motivating examples
  - AI: Vision, Language, Robotics
  - BI: Brain and Cognitive Science
- Dynamical systems
  - Definition
  - Types
- Learning in dynamical systems
  - Sequential/online
  - Learning and action interleaved

Vision: Target Tracking

- Multiple people tracking (http://www.youtube.com/watch?v=sk9WY79w8oM&feature=related)

Speech Recognition

- The virtual personal assistant on your phone (Siri) (http://www.youtube.com/watch?v=MpjpVAB06O4)

Language Generation

Watch from 2:00 min

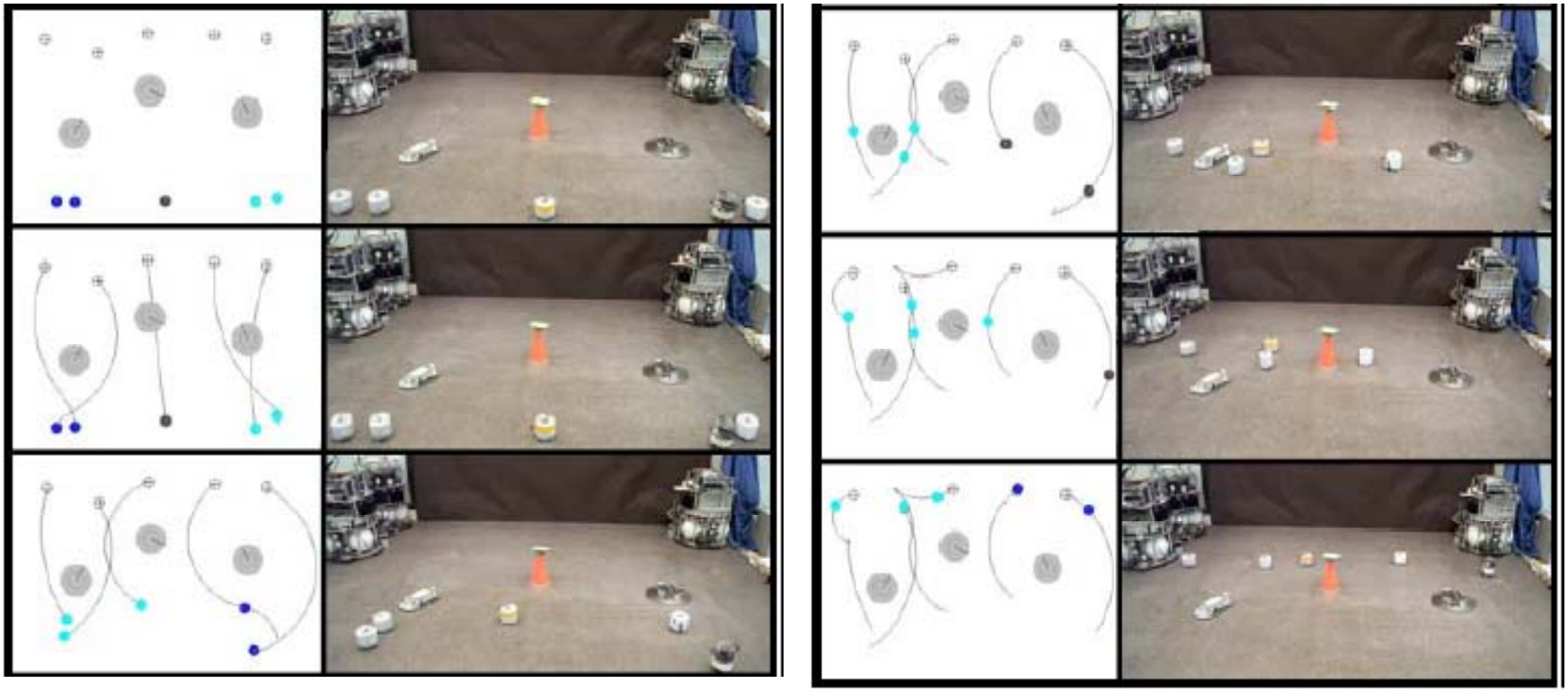
Multimodal conversational agents
(http://www.techfak.uni-bielefeld.de/ags/wbski/media/MAXMTV.MPG)

Graphics: Motion Planning

Watch from 3:00

- Real-time scalable motion planning for crowds (http://www.youtube.com/watch?v=ifimWFs5-hc)

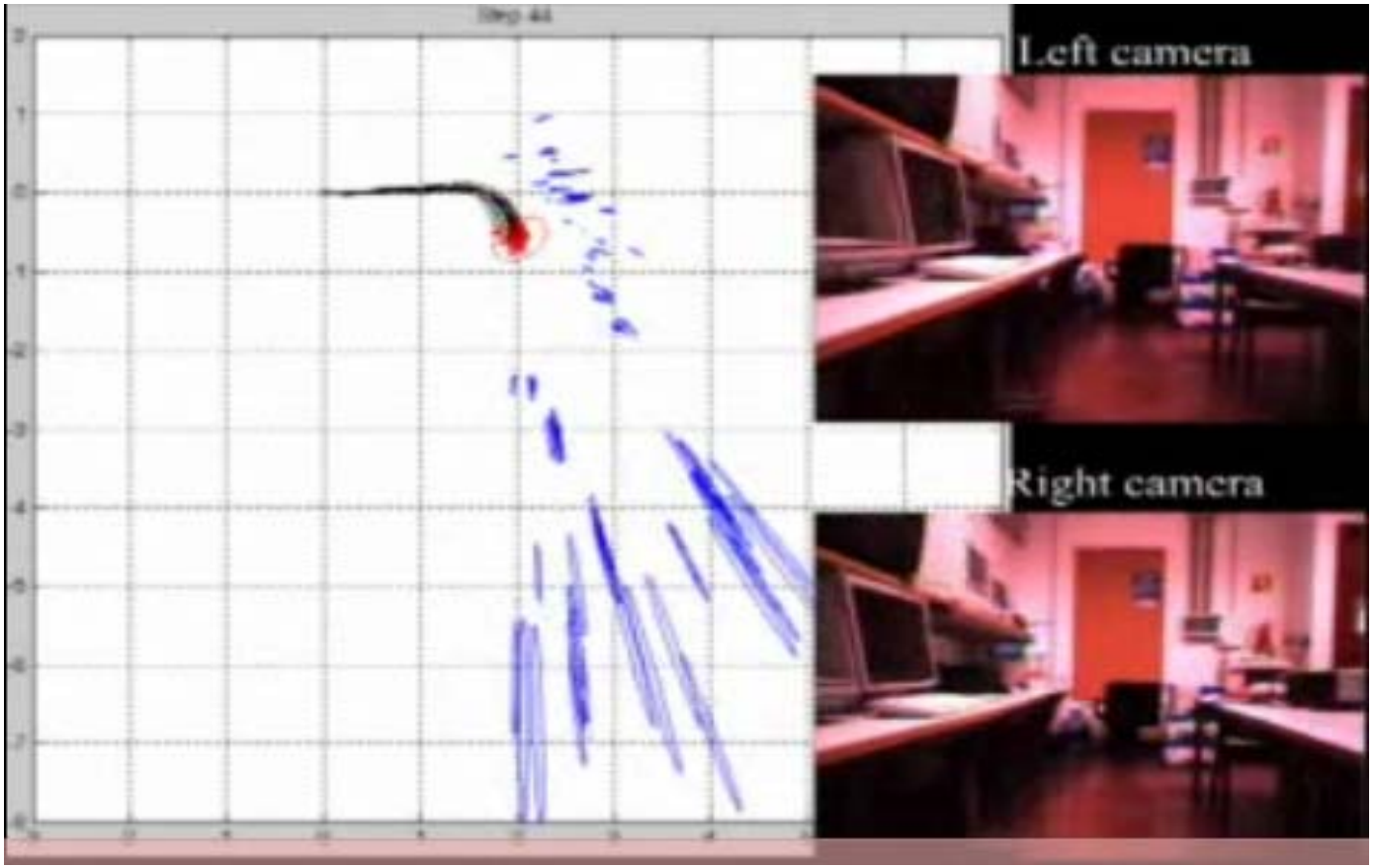
Robotics: Path/Motion Planning



Motion planning for multiple mobile robots

Example experiment on the Micro-Autonomous RoverS (MARS) test-platform involving 5 robots and 3 obstacles. (C.M. Clark *et al.*, *IEEE International Conference on Robotics and Automation*, 2003)

Robotics: Localization

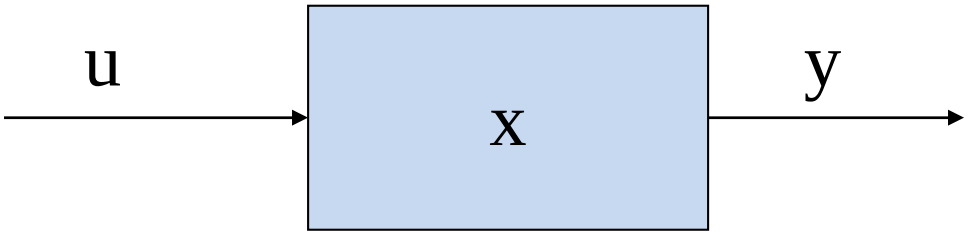


- Real-time localization in indoor with stereo vision (http://www.youtube.com/watch?v=m3L8OfbTXH0&NR=1)

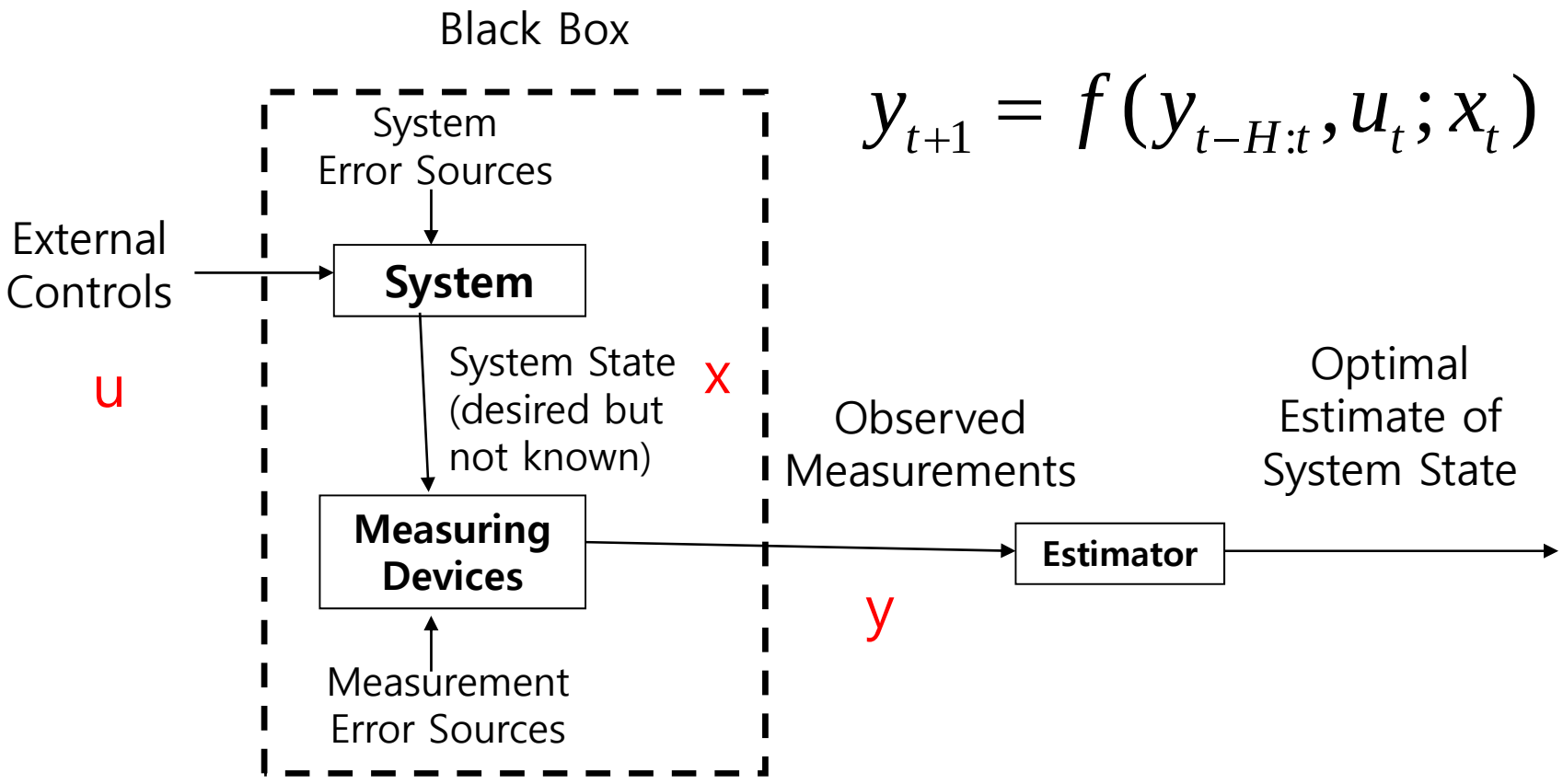
Dynamical Systems

Dynamical systems

- System: a (physical) entity which can be manipul ated with actions, called inputs (u) and that, as a consequence of the actions, gives a measurable reaction, called output (y)
- Dynamic: the system changes over time; in gene ral, the output does not only depend on the inp ut, but also on the current "state" of the system (x), i.e. on the system history

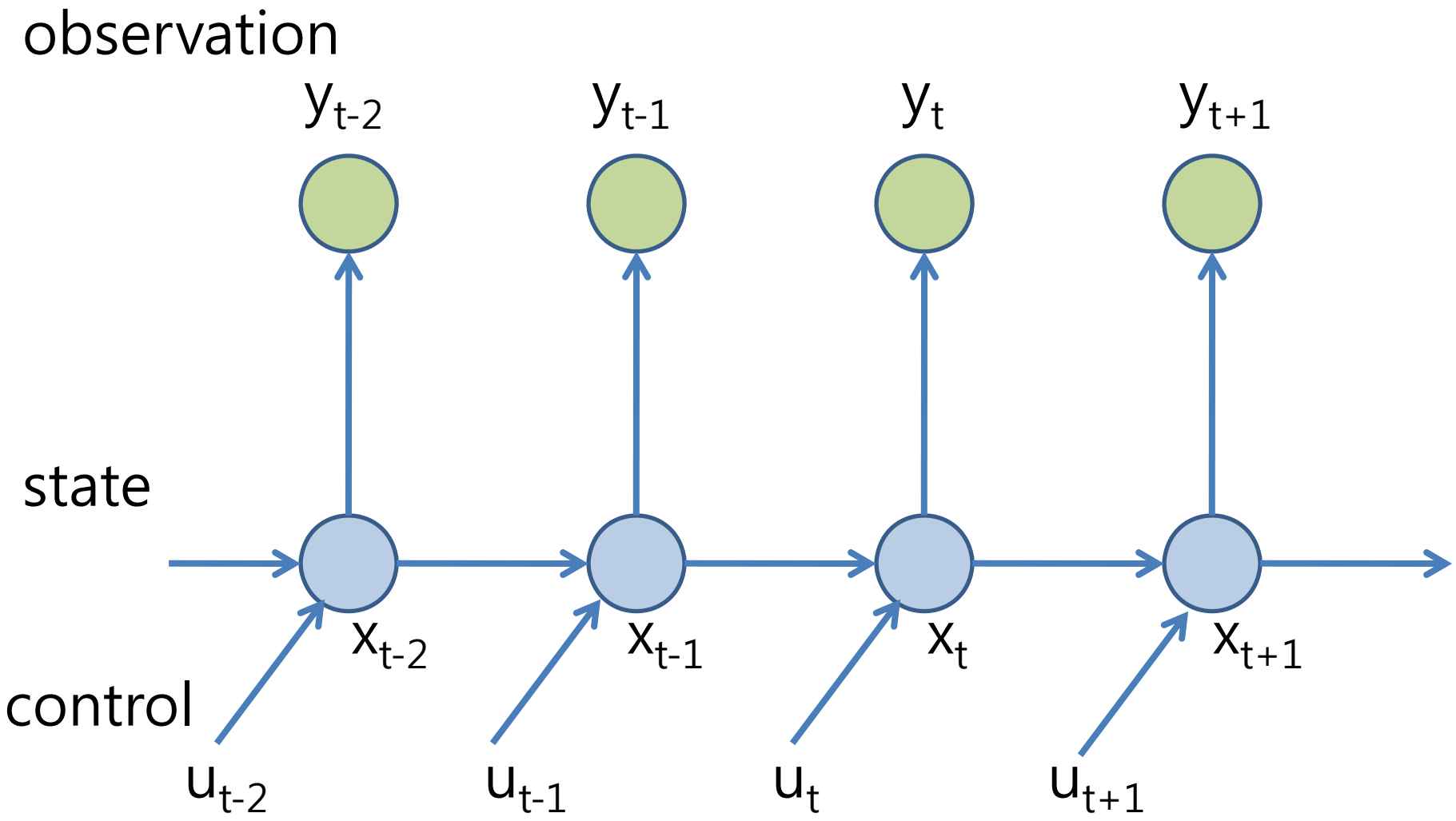


Dynamical system



$$y_{t+1} = f(y_{t-H:t}, u_t; x_t)$$

Graphical model



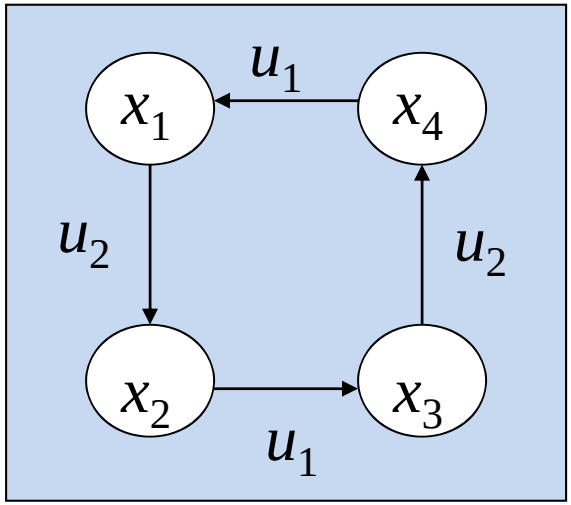
Modeling a dynamical system

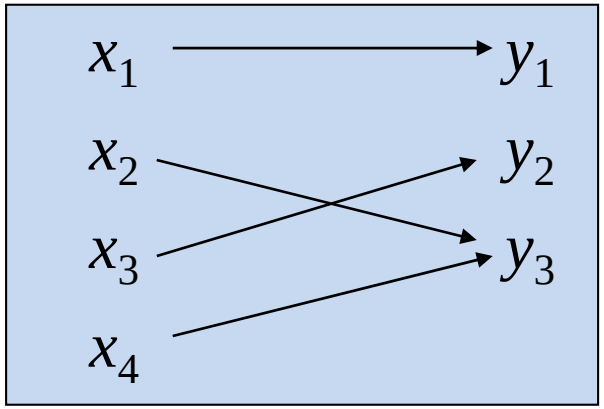
- Two ingredients:
  - State transition function

$$x_{t+1} = f(x_t, u_t)$$

  - Observation function

$$y_t = g(x_t)$$





Main classes of dynamical systems

- Continuous / discrete
- Linear / nonlinear
- Time invariant / variant systems
- Single / multiple input / outputs
- Deterministic / stochastic

Discrete and continuous systems

- Discrete: the time set is the set of integer numbers (t=1,2,…,k,…). The system is typically modeled with difference equations

$$x_{k+1} = f(x_k, u_k),$$

$$y_k = h(x_k)$$

- Continuous: the time set is the set of non-negative real numbers. The system is typically modeled with differential equations

$$\frac{dx}{dt} = f(x_t, u), \qquad x_{t_o} = x_o, \qquad y_t = h(x_t)$$

Learning in Dynamical Systems

Dynamic Learning

- Target system
  - Changing environment
  - Moving target (can be nonstationary)
  - Sequential decision making
  - Action (control) selection
  - Prediction/forecasting
- Learning system
  - Sequential observation of data
  - Online learning (vs. batch learning)
  - Interactive learning (action interleaved)
  - Sequential/recursive estimation
  - Short-term vs. long-term objectives

Course Objectives

- Review existing models for dynamic learning
- Develop a sequential Bayesian framework for dynamic learning
- Emphasis on Monte Carlo filtering methods, such as particle filters
- Establish the connection between particle filtering and evolutionary optimization
- Construct graphical models for dynamic Bayesian learning from video (language + vision) stream data

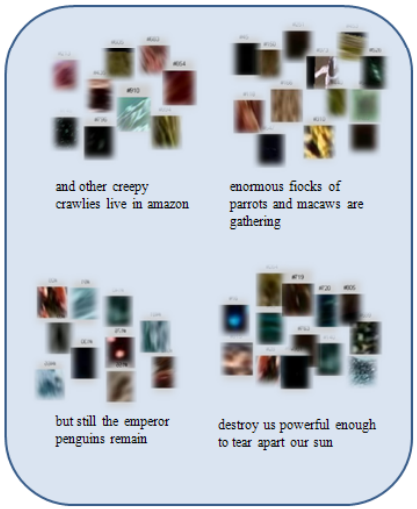
Project

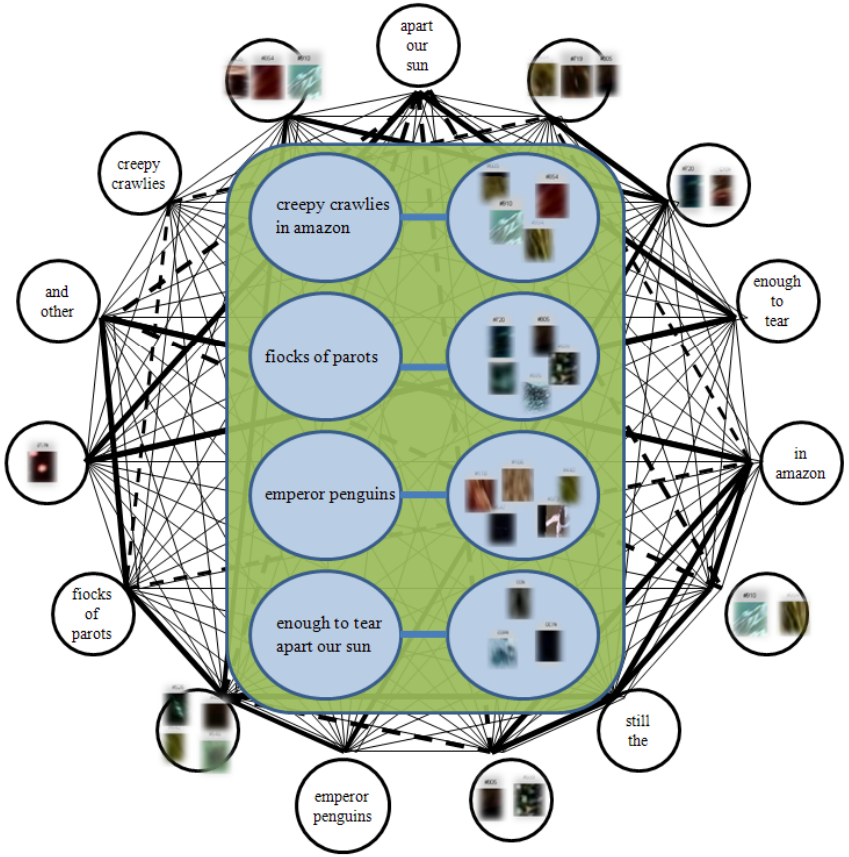
- Given
  - A sequence of *T* image-text pairs of Y(t) = (V(t), L(t)), t=1,…, *T*
  - V(t): a vector of visual words, L(t): a vector of linguistic words
  - E.g., from a 20-minute episode of *Friends*
- Construct
  - A dynamic system that learns to estimate the (mental memory) states to generate the future image-text sequences from a historical context of size *H*:

$$y_{t+1} = f(y_{t-H:t}, u_t; x_t)$$

- To demonstrate
  - Visual storytelling: Given a sequence of pairs of (V(h), L(h)), h = t, t-1, t-2, …, t-H, generate (V(r),L(r)) for r = t+1, t+2, t+3, …, t+R
  - L2V translation (mental imagery): Given a series of texts, generate a series of images.
  - V2L translation (scene description): Given a series of images, generate a series of texts.

Dynamic Learning Memory





Project: Visual Storytelling