# Introduction to Advanced AI Topics

Vision

Natural Language Processing

Robotics

# Advanced Topics Overview

- Agents have sensors and actuators
- Sensors:
  - Seeing (visual input) ⇒ Image Processing and Computer Vision
  - Hearing (audio input) ⇒ Natural Language Processing
- Actuators:
  - Moving and manipulating ⇒ Robotics

# Computer Vision

Perception

# Definition: versus graphics

- Graphics
  - Have world model $W$
  - Generate the sensory stimulus from the model $S = f(W)$
- Vision
  - Generate the model from the sensors: $W = f^{-1}(S)$
  - To think about: f() doesn't have a proper inverse. Why?

# Ambiguity in sensory input

- Girls playing with dollhouses
- Or giants playing with people?

• Many possible world models to choose from.
• Vision works on choosing the best model given the input.

# Definition: versus image processing

- Image Processing
  - A transformation of data to other data
  - e.g., smoothing

- Computer Vision
  - Reduction in data to a (more useful) abstraction
  - e.g., digit / face recognition

# Applications

- Surveillance – can we detect objects or people as they move around our field of vision?
- Handwriting recognition – from handwritten addresses to barcodes
- Content based Image Retrieval – query for images using without any text features. "Show me similar pictures"
- Automated Driving – speaks for itself

# Natural Language Processing

## Communication

N.B. We will go over this area in more detail towards the end of the course

# Definition of NLP

- Examines communication in human languages.
  - Theoretical and practical aspects.
  - Similar to vision, has production and understanding affects
    - Understanding: speech / text to meaning
    - Generation: meaning to speech / text
  - Both processes have inherent ambiguity

# Not so great newspaper headlines

- Squad helps dog bite victim.
- Helicopter powered by human flies
- Portable toilet bombed; police have nothing to go on.
- British left waffles on Falkland Islands.
- Teacher strikes idle kids.

# Sample Applications

- Restaurant Query converts English queries into SQL.
- MS Dictation converts speech into text
- Babelfish translates Web pages to different languages
- Summarizing multiple news articles from the web

# Robotics

Planning in the real world environment

# Getting around

- Effectors
  - Sensors on effectors? Is the output noisy?
  - Low-level: need to build higher-level abstractions

# Problems

- Localization – where am I?
  - Mobile robots but also robotic arms
- Mapping – what does my environment look like?
- Moving – how do I get from here to my goal? What type of plan do I have execute?

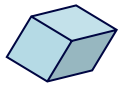
# Applications of robotics

- Robotic Flight – robotic helicopter, unmanned piloting
- Path planning for exploration
- Rock climbing, perhaps difficult even for some of us

# Summary

- All three areas deal with search:
  - Vision: search for most likely world *w* given input sensor *s*
  - Natural Language Processing: given an input utterance / text *i*, find most likely meaning *m*
  - Robotics:
    - Localization: given unknown input configuration / location, determine configuration.
    - Planning: given goal *g* and state *s* output plan *p* to reach *g* from *s*

# Summary

- All three areas use heuristics :
  - Vision: trihedral structure 
  - Natural Language Processing: grammars of language, most frequent meanings
  - Robotics: decomposition of problems into cells, maximizing distance between obstacles
- Many of these heuristics involve probability, which we will get to soon