Requirements and A Framework for Learning Object Affordances (Target Kitty)

S. Hongeng

Development of Knowledge

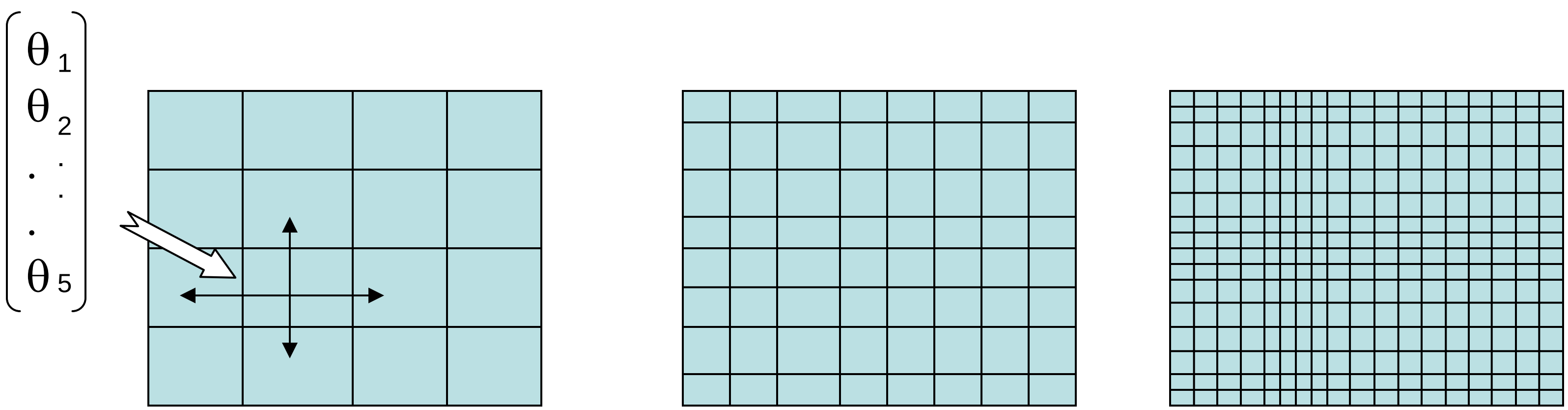
- Learn about body and motor control
  - Inverse kinematics, smooth/compliant control
  - Requires coordination between perception and motor control signal
  - Learn motor skills (skillfully approach a location from resting poses)
- Learn about intrinsic properties of objects
  - ***Actively*** probe objects to categorize properties
  - Segment boundary of an unknown object (to see the shape)
  - Probe for weight (heavy, light), size (graspable, non-graspable)
- Learn to discriminate means (non-goal directed)
- Learn about actions that can be done to object
  - Probe cause-and-effect relations among features of arm and object during the actions
  - develop means-end association (goal directed)
  - Shift from correlation to causation, enabling learning of object affordances
- Ontology genesis
  - Develop ontology (bootstraping actions and activities, develop theories)

Suggestions on Things to Focus

- Motion Control
  - Restrict the end effector to move on a 2D plane
  - Multi-level quantization of space (into grids) that allows the end effector to be moved in step-wise fashion
  - Use visual feedback to calibrate the motion (inverse kinematics is given)
- Assume innate drive (desire) to play with an object aiming to
  - gather (and categorize) information on object properties
  - discover a multimodal affordance representation (encoding object features, structures, arm poses, etc.)
- Focus on a framework for learning cause-and-effect relations among
  - Features related to arm
    - approaching trajectory, approaching poses, parts of the arms in contact with object, approaching directions, speeds of arm, force sense, etc.
  - Features related to object
    - Object pose, part of object in contact with the arm, movement direction after the contact, speed, rotating directions, etc.

Space Representation and Arm Control

- Multiple-scale space segmentation for better control of arm speed (e.g., to avoid collision)
- Motion control signal for moving the end effector from one grid to the neighboring grid is given or computed using inverse kinematics
- Skilled motion from a resting pose to reach a landmark is manually assigned (or learned)



Perception of Objects and Arm

- Initially, only objects of simple shapes will be used (brick, cube, ball, cylinder, cup)
  - Object localization will be based on figure/ground segmentation
  - Allows nudging action if there is ambiguity
  - How to detect shape/pose? Should we initially assume object models are given?
- Closed-loop arm control requires feedback from visual perception of arm poses
  - Assume that arm configuration is known (e.g., number of links and lengths)
  - Use proprioceptive data (joint angles) to constrain the search space for pose

Learning latent spatio-temporal structures of cause-effect relations

- Causal structures can be learned by constructing a model that best describes the observed data
- Search space is prohibitively large
  - There are many features
  - Feature combinations
  - There are many ways to represent features
    - A feature may be represented as quantitatively or qualitatively
- Bottom-up, statistical approach for discovering correlated data is slow, requiring lots of training sequences
- Higher level theories may help provide prior constraints on structures reducing the search space
  - E.g., relation between certain types of features may be more likely than others; one representation of feature may be more likely than others in certain contexts
  - Theories may also evolve
  - provide a framework for reasoning about alternative causal links, and hence allow for human-like learning process

End