Grounding Natural Language Advice in a Machine Learning Video Game

Igor Karpov
Advice Demo
Example Advice

- “if there is a wall ahead then turn right slightly and walk forward“

- (if_rule
  <Antecedent>condition(exist <Actor>wall <Modifier>ahead )
  <Consequent>command(walk <Modifier>forward))

- what are the conditions in terms of sensors?
  - wall_sensor_1 > 0.5

- what are the actions in terms of outputs?
  - turn_actuator += 0.1 and walk_actuator += 0.1

- can we learn lexical alignment? can we learn the grammar?
Current System

Human advice: if there is a wall ahead, turn right slightly and walk forward

OpenCCG-based semantic parser

Advice as a formalized command: \{ \text{if } \text{wall}_1 > 0.5 \text{ then turn } 0.7 \}\}

lex/yacc neural network construction

Advice as a neural network controller

Advice injection

Combined advice/behavior genotype

Real-time neuroevolution in-game

Desired behavior
Semantic parsing

- Translates NL sentences into a logic form (using OpenCCG)
- A probabilistic CCG can be estimated from labeled examples using maximum likelihood methods (Zettlemoyer & Collins, Mooney and others)
- Can a system learn the correct grammar and grounding from user interaction and reinforcement in service of a task?
Advice Injection

Existing Network

Advice Network

Advice Injection
Neuroevolution

- Advice network structures are on equal footing with the rest of the genotype
  - Good advice will survive
  - Bad advice will die out
  - Reusable modules can get reused for other purposes
Experiments

- Learning words without advice
  - Learn to say three words (enemies!, friends!, walls)
  - Allow use of these words as input and compare with non-verbal populations

- Learning simple advice grounding
  - Given an advice structure, evolve its grounding assuming that it is beneficial
  - Each word is represented by a number of weights in the sensor or action space
Vision for the future

OpenCCG-based semantic parser

Grounded advice representation

Advice injection

Combined advice/behavior genotype

Real-time neuroevolution

Questions, confirmation, grounded grammar

Human advice: “if there is a wall ahead, turn right slightly and walk forward”

Human reinforcement: “good job!”

Desired behavior