Neural Program Synthesis from Diverse Demonstration Videos

Shao-Hua Sun¹, Hyeonwoo Noh², Sriram Somasundaram¹, Joseph Lim¹

shaohuas@usc.edu  shgusdngogo@postech.ac.kr  sriramso@usc.edu  limjj@usc.edu

Contribution

- Program synthesis for explicit modeling of underlying programs
- Architecture for summarizing diverse demonstrations
- Auxiliary objectives enhancing program synthesis

Our approach: Neural Program Synthesis

Q: Could we infer underlying program of a behavior from diverse demonstration videos?

Underlying program

Don’t let them get close.

Problem Statement

Q: Could we infer underlying program of a behavior from diverse demonstration videos?

Experimental Results

- Advantage of explicit modeling of program
- Effect of summarizer / multi-task loss
  - Karel
    - Methods: Execution / Program Sequence
      - Induction baseline: (62.8% / 69.1%)
      - Synthesis baseline: (64.1% / 42.4%)
      - + summarizer (ours): (68.6% / 45.3%)
      - + multi-task loss (ours-full): 72.1% / 48.9%
  - ViZDoom
    - Methods: Execution / Program Sequence
      - Induction baseline: (35.1% / 60.6%)
      - Synthesis baseline: (48.2% / 39.9%)
      - Ours-full: 78.4% / 62.5%

Experimental Setting

- Karel:
  - Movement
  - Put / pick marker
- ViZDoom:
  - Movement
  - Kill monster
  - Select weapon

- Induction baseline (Few-shot imitation)
  - Implicit modeling of programs by learning to imitate
- Synthesis baseline
  - Program synthesis without summarizer / multi-task loss

Qualitative Results

- Karel
  - Underlying program
  - Ours-full
  - Baseline

- ViZDoom
  - Underlying program
  - Synthesized program

- Effect of summarizer with more demonstrations
  - Advantage of summarizer is bigger with more demonstrations
    - Methods: k=3, k=8, k=10
      - Synthesis baseline: 58.5%, 60.1%, 64.1%
      - Improvement: 2.1%, 3.0%, 4.5%

- If-else experiment - inferring condition
  - Single IF-ELSE statement with 1 condition, 2 different actions
  - Inferring correct condition is essential
    - Methods: Execution / Program Sequence
      - Induction baseline: 26.5% / 28.3%
      - Synthesis baseline: 59.9% / 44.4%
      - Ours-full: 89.4% / 69.1%