SUGILITE: Supporting Co-adaptive Human-Agent Relationship through Programming by Demonstration using Existing GUls.

Toby Jia-Jun Li, Igor Labutov, Nancy Li, Tom M. Mitchell, Brad A. Myers
Carnegie Mellon University
tobyli@cs.cmu.edu

http://natprog.org/sugilite.html
We want to empower **end users** to program **intelligent agents** using a combination of **demonstration** and **verbal instruction** on **existing third-party GUIs**
Back-end Services  \(\rightarrow\) API call  \(\rightarrow\) GUI  \(\leftarrow\) Display  \(\leftarrow\) Direct Manipulation  \(\leftarrow\) User
Back-end Services

API call

Agent

Conversation / Automation Rules

Needs to be customizable and programmable

GUI

User
Back-end Services

API call

Agent

Conversation / Automation Rules

End user development (EUD)

GUI

User
Back-end Services

API call

Agent

Conversation / Automation Rules

End user development (EUD)

GUI

User

No open API available

Lack of technical expertise
Back-end Services

Direct Manipulation at Runtime

Semantic Understanding

Agent

Verbal Instruction in Natural Language

GUI

Demonstration

User
We enable users to use the most natural modality, or a mix of both

**Demonstration:** *What* the user would do in the current context

**Verbal Instruction:** *Why* the user would do this

* Users are able to describe their intentions by referring to contents shown on GUIs
We enable users to use the most natural modality, or a mix of both

**Demonstration:** *What* the user would do in the current context

![Diagram showing Disambiguate and Ground]

**Verbal Instruction:** *Why* the user would do this

* Users are able to describe their intentions by referring to contents shown on GUIs
Video
(CLICK (ARG_MIN numeric_time (AND (hasClass Button) (below (hasText "Joe's Steakhouse"))))))
Things I would love to discuss at this workshop

• How can we better design a multi-modal and mixed-initiative “meta-interface” to support a more natural end user development experience?

• What information about GUIs and user interaction traces can be better leveraged for inferring user intentions from demonstrations?

• What strategies can we use to allow users to be more expressive when describing their intentions, logic and goals using GUIs as the grounding?

• What design principles can we create for the new role of GUIs, where they are no longer the sole interaction point with users, but used in conjunction with other forms of interfaces (e.g., conversational), used as the grounding for programming, or controlled indirectly by intelligent agents?