

A Goal-Oriented Interface to Consumer Electronics Using Planning and Commonsense Reasoning

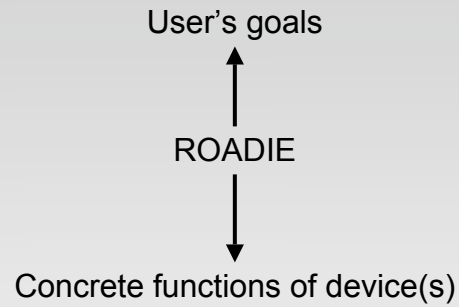
Henry Lieberman and Jose Espinosa,
MIT Media Lab, IUI 2005

Presented by: Charles Rich
January 30, 2008

Motivation

- People overwhelmed by complexity of current home electronics
 - e.g., Canon S500 digital camera has 15 buttons, 2 dials, 4 x 2 mode switches, 3 menus with 5 choices in each mode (each with 3 values), 7 on-screen mode icons, etc.
- *Authors' diagnosis:* Manufacturers are attempting to maintain 1-1 correspondence between “functions” and “controls” (which worked for simpler devices)

Proposed Solution



Roadie's Capabilities

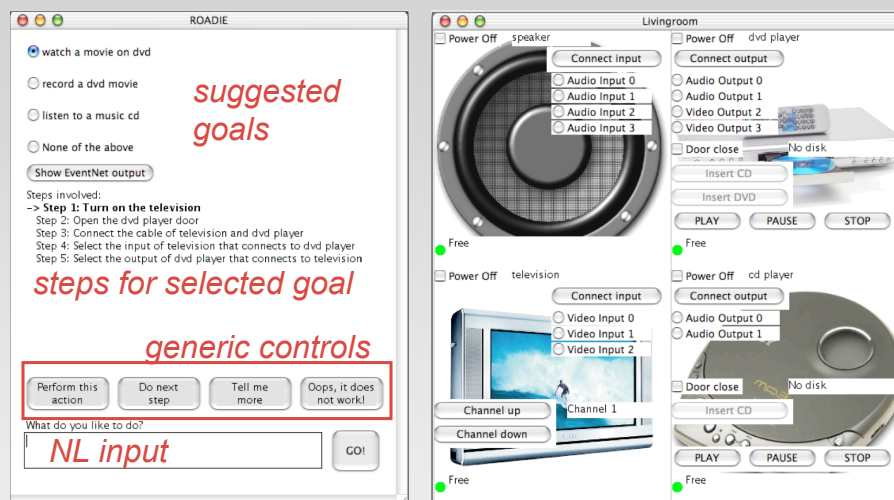
- Guide user through normal operation
- Help troubleshoot when things go wrong
- Future capabilities:
 - adapt to user preferences of use
 - learn new goals (e.g., by composing devices)



Device Requirements

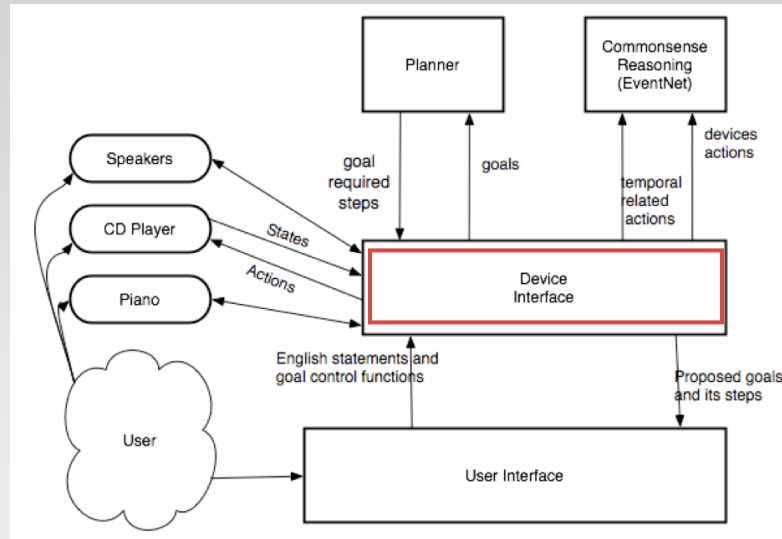
- Devices must provide *software interface* to:
 - control functions
 - query state
- Current state of consumer electronics industry: Standards exist (e.g., UPnP), but not fully supported by devices yet

How Roadie Looks to User



simulated device controls

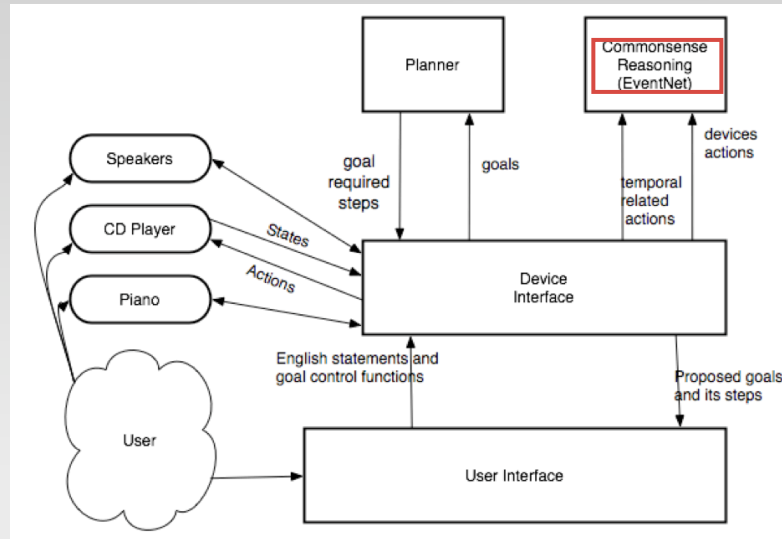
Roadie Architecture



Device-Specific Knowledge

- All task types supported by device (primitive actions and goals), represented in two forms:
 - English string (for matching in EventNet), e.g.,
“play the music CD”
 - Logical specification (for planner), e.g.,
(play-music-CD [cd-player-device]
[speaker-device])
Also preconditions/postconditions, etc.

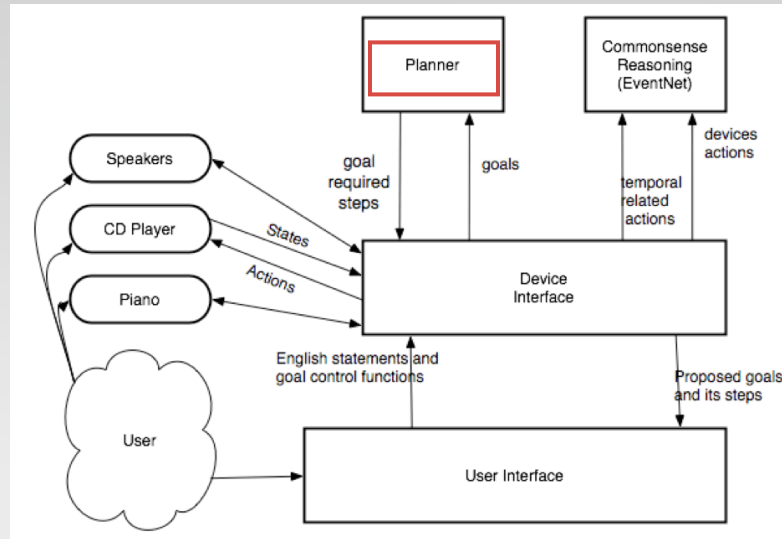
Roadie Architecture



Commonsense Knowledge - EventNet

- EventNet built on top of OpenMind
 - 770,000 English sentences describing everyday life
 - contributed by volunteers on the Web (little or no quality control)
- Used to:
 - infer user's goals from actions (plan recognition)
 - infer user's goal from natural language input
 - suggest device functions to achieve goals

Roadie Architecture

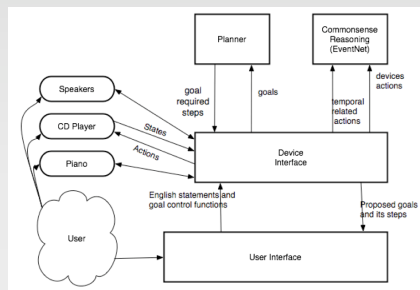


Planner

- GraphPlan
 - general-purpose STRIPS-style planner
 - very flexible
 - good for handling unexpected situations (troubleshooting)
 - depends on adequate formalization of actions
 - produces partial-order plan

Typical Scenario

- User turns on DVD player (using front panel switch)
- English description “turn the DVD player on” is reported by device interface to Roadie
- Roadie sends description to EventNet

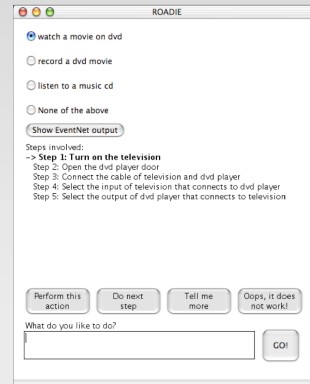


Typical Scenario (cont'd)

- EventNet returns matches:
 - “watch hours of world’s best nature programs”
 - “hit play”
 - “insert your recorded cd”
 - “listen to music”
 - “insert disk”
 - “insert dvd”
 - “leave the room”
 - “push television”
 - “turn on home theater projector”
 - etc.

Typical Scenario (cont'd)

- Roadie filters EventNet results (not exactly clear how) to present suggested goals:
 - “watch a movie on dvd”
 - “record a dvd movie”
 - “listen to a music cd”

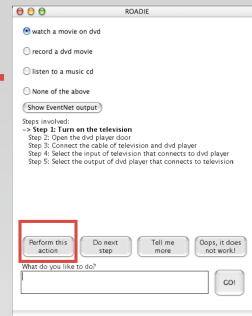


Typical Scenario (cont'd)

- User chooses goal “record a dvd movie”
- Roadie uses planner to compute plan for goal and displays it:
 1. Turn on recorder
 2. Connect the cable of the recorder and the DVD player
 3. Open the DVD player door
 4. Select the DVD player output that connects to the speaker
 5. Select the speaker input that connects to the DVD player
 6. Insert the movie DVD
 7. Close the DVD player door

Typical Scenario (cont'd)

- User clicks “Perform this action”
- Roadie starts executing plan
 - Roadie turns on recorder
 - Roadie stops at cable connection step because cannot do itself
- Roadie asks user to connect cable of the recorder and the DVD player
- User clicks “Tell me more”
- Roadie displays photo

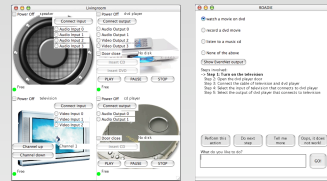


Resolving Conflicting Goals

- A common source of difficulty, especially in networked environments
- E.g., user types “I want to watch a movie” while DVD player in use player music CD
- Roadie suggest moving music CD to CD player to free up DVD player

Evaluation

- Six participant user study (not enough for statistical validity)
- Compared using simulated device interfaces alone vs. devices plus Roadie
- Compared # clicks and time
- Best result:
 - without Roadie: 60 clicks, 440 sec
 - with Roadie: 23 clicks, 202 sec



Related Work

- Consumer Electronics
 - smart remotes: context-aware, universal
 - no paradigm shift to user goals
 - smart homes: mostly sensor-focused
 - ditto
 - natural language control of devices
 - Roadie (EventNet) approach is “rougher”
- Other applications of OpenMind
- Other debugging interfaces (using reflection)

Related Work - DiamondHelp

- Overall share same philosophy:

User's goals
↑
Roadie/DiamondHelp
↓
Concrete functions of device(s)
- Roadie uses *first-principles planner* vs. DiamondHelp *hierarchical task network*
- DiamondHelp does not have:
 - natural language processing
 - debugging support
- DiamondHelp has better user interface design

Remaining Problems/Future Directions

- Biggest problem is where the formal planning knowledge (device-specific STRIPS model) comes from
 - programming by example?
 - learned from experience?
- Similar techniques can be applied to customize and extend Roadie's informal knowledge