

Controlling the spatio-visual context in situated natural language generation

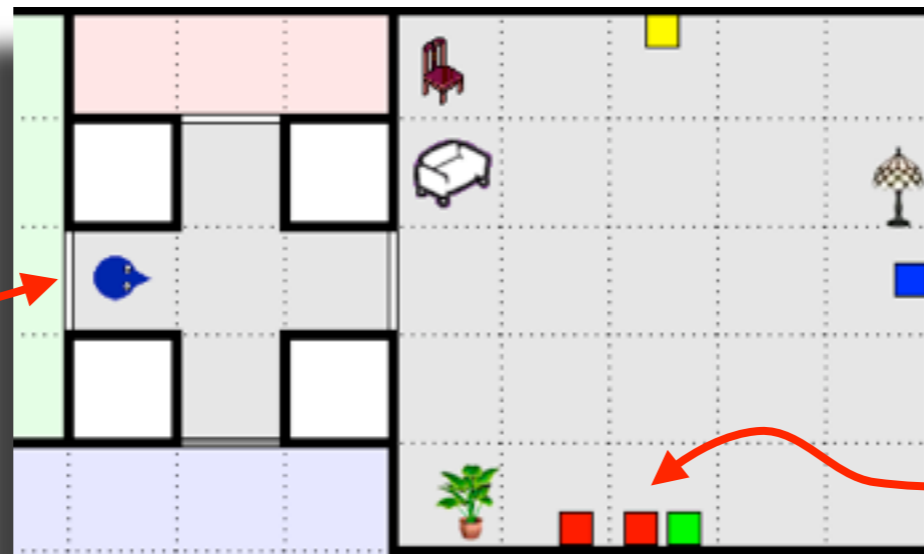
Konstantina Garoufi and Alexander Koller

Cluster of Excellence
on Multimodal Computing and Interaction
Saarland University, Germany

Conference “Space in Language”
Pisa, 9 October 2009

Motivation: space, language, and control

instruction follower's location



target

1. *walk forward and go through the door*
2. *(...) ok and then turn to your right*
3. *and then hit the button in the middle*



hit the middle button on the right wall in the next room



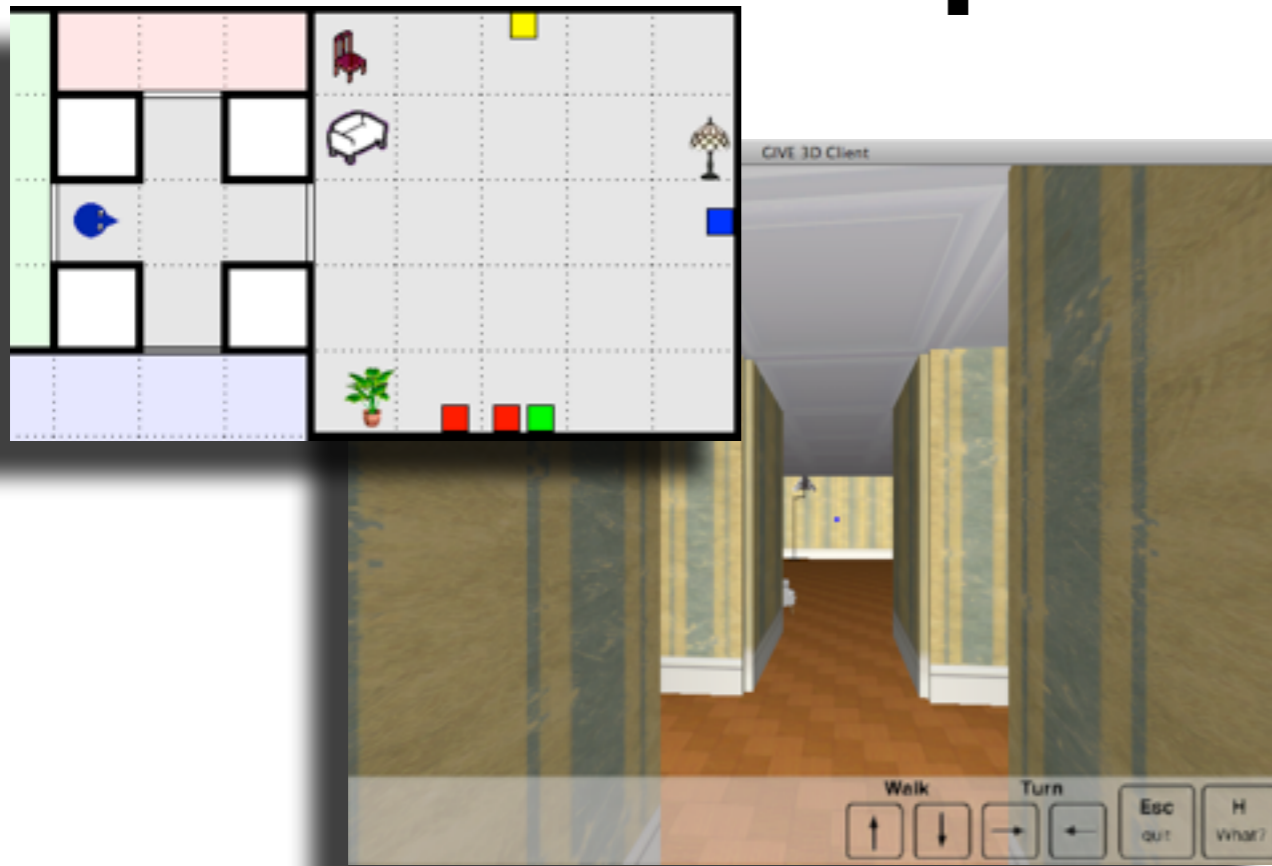
Goal: computational modeling

1. Implement a situated natural language generation system that can control the spatio-visual context of the scene in order to produce cognitively simpler distinguishing referring expressions
2. Evaluate the model with respect to quantitative and qualitative measures by participating in the GIVE Challenge



Byron et al. (2009)

How? With automated planning



Walk forward.

→
walk-forward(l1,l2)

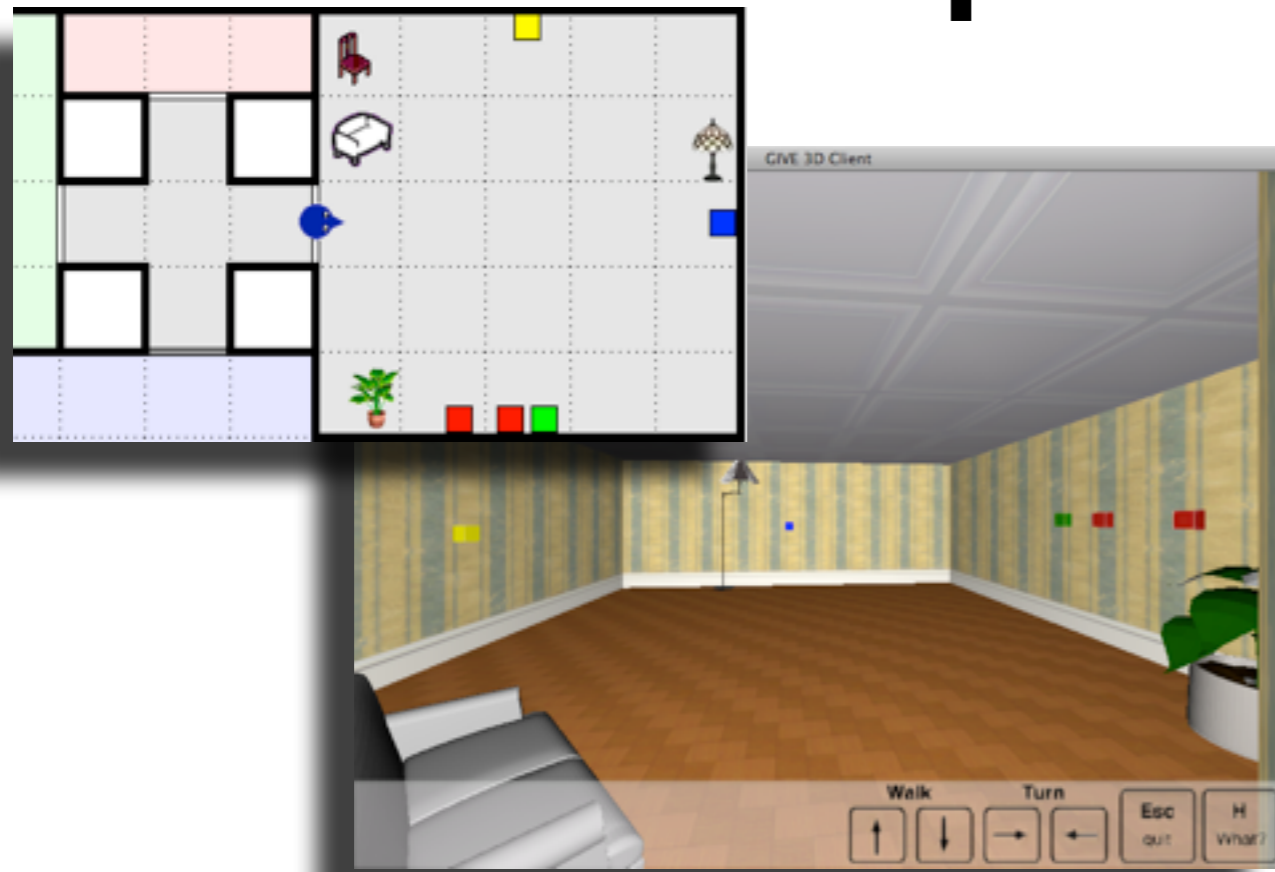
State
location(l1)
orientation(o1)
 \neg visible(l1,o1,b)

action: walk-forward(x,y)

precondition: location(x)

effect: \neg location(x), location(y)

How? With automated planning



Turn right.

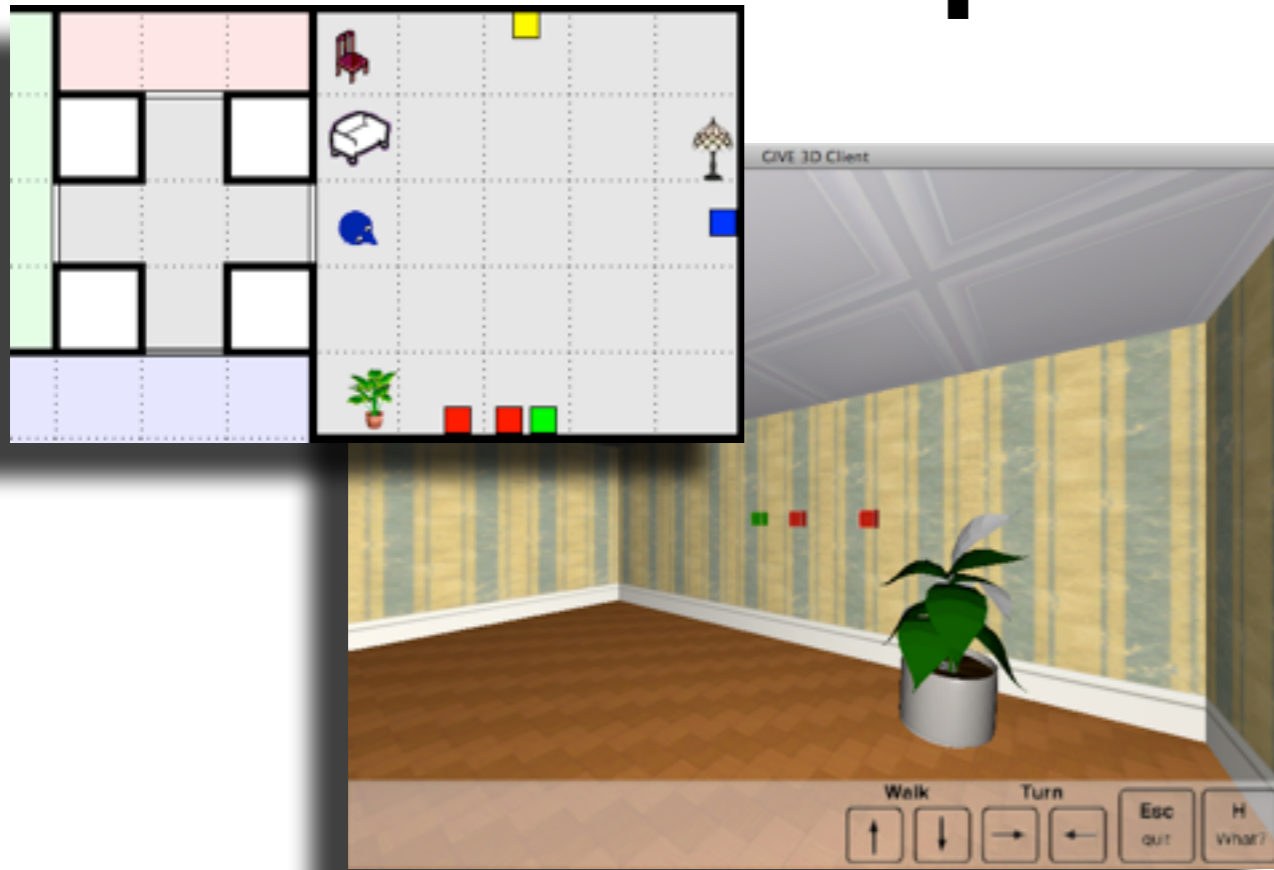


turn-right(o1,o2)

State
location(l2)
orientation(o1)
 \neg visible(l2,o1,b)

action: turn-right(x,y)
precondition: orientation(x)
effect: \neg orientation(x), orientation(y)

How? With automated planning



State
 location(l2)
 orientation(o2)
 visible(l2,o2,b)

action: in-the-middle(x)
precondition: middle(x)
effect: $\forall y. \neg \text{middle}(y) \rightarrow \neg \text{distractor}(y)$

Hit the button in the middle.

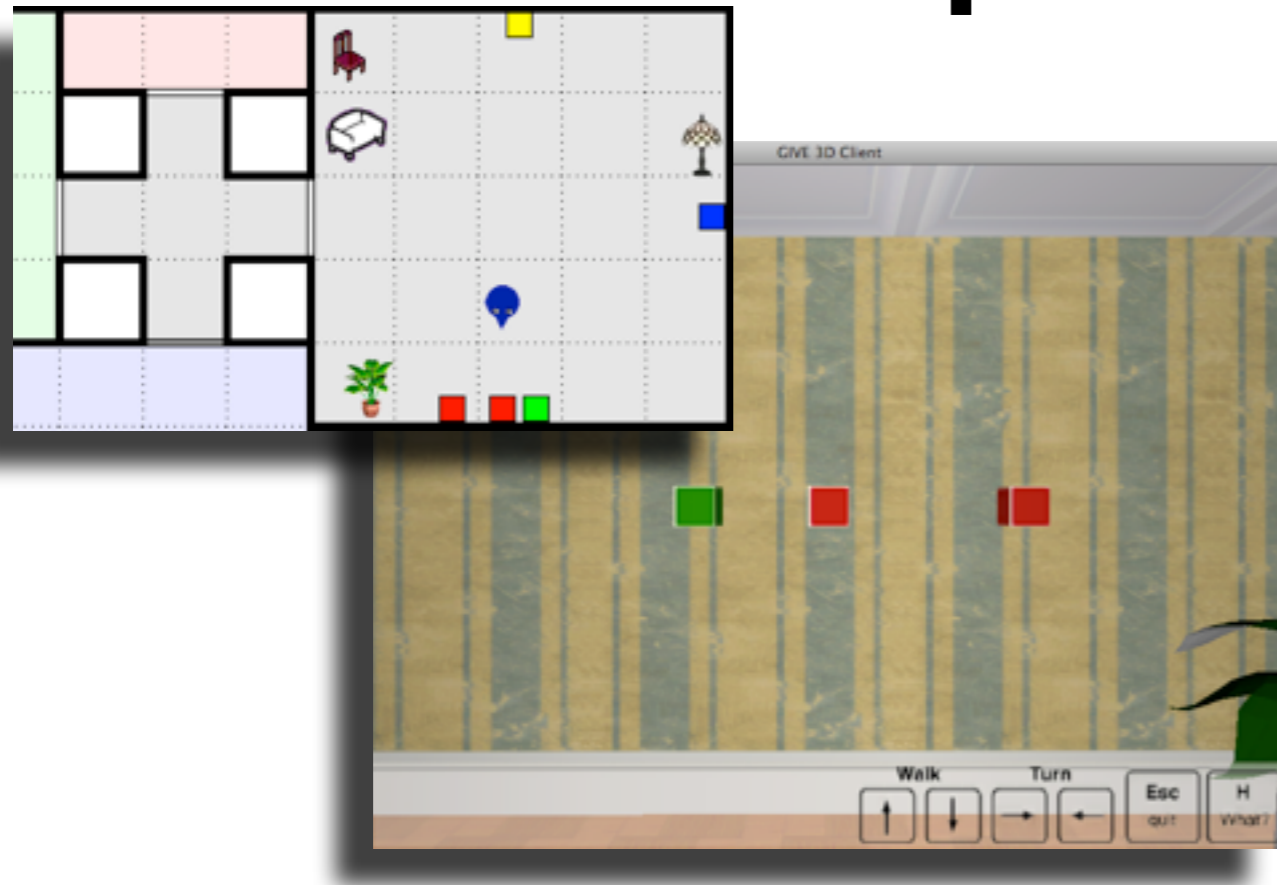


hit(b)
 button(b)
 middle(b)

action: hit(x)
effect: to-do(hit(x))

action: button(x)
precondition: visible(x)
effect: $\forall y. \neg \text{button}(y) \vee \neg \text{visible}(y) \rightarrow \neg \text{distractor}(y)$

How? With automated planning



State

to-do(hit(b))



Challenges

- Interleaving linguistic and non-linguistic effects of communicative actions
- Plan generation, execution monitoring, and error recovery
- The representation of paths and motion in spatial cognition and language