Universität Bielefeld

Applied Informatics, Bielefeld University, Germany

CRC 673 Alignment in Communication

A Computational Model for the Alignment of Hierarchical Scene Representations in Human-Robot Interaction

Agnes Swadzba
in collaboration with Constanze Vorwerg, Sven Wachsmuth, Gert Rickheit

Introduction

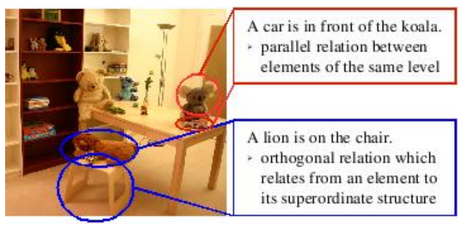
- scenario: talking to a robot about
  a pen lying beside a book on a table that should be placed back into a drawer under the desk
- scene representation on different granularity levels
- implicit learning of room layouts from verbal user descriptions
- alignment of spatial structures between a robot and the user

Empirical Foundation

- people's descriptions reflect their mental representation of the perceived scenes
- psycholinguistic study:
  - 10 German native speakers were shown a photo of a real room and have to describe it freely
  - objects are localized relative to their supporting room structure
  - orthogonal relation (**rel**$_{\perp}$): *a lion is on the chair*
  - parallel relation (**rel**$_{\parallel}$): *a car lies in front of the koala*

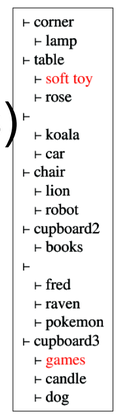
Computational Model

- verbal descriptions: sequence of *orthogonal* and *parallel* relations between objects

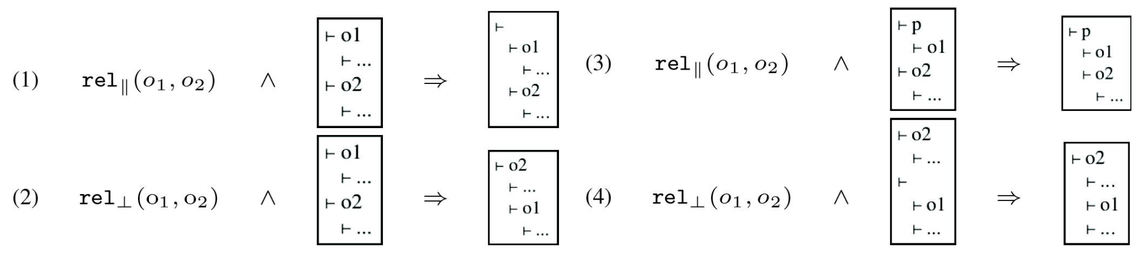


In the corner is a lamp . Soft toys are on the table , a rose is on the table , and a car is in front of the koala . A lion is on the chair . A small robot lies in front of the lion . In the left cupboard (cupboard2) are books . Also there are games in the cupboard2 . Next to fred is a raven . Below the raven are the pokemon . In the right cupboard (cupboard3) are games . Also a candle is in cupboard3 . Above the candle is a dog .

- hierarchical organization of related objects in set of dependency trees T (parent node holds superordinate structure of subordered elements)
  - manipulate-operations:
    $\texttt{obj}(\text{“o”})$, $\texttt{child}(n_o, n_p)$, $\texttt{delete}(n_p)$
  - query-operation:
    $\texttt{bool} = \texttt{ischild}(n_o, n_p)$

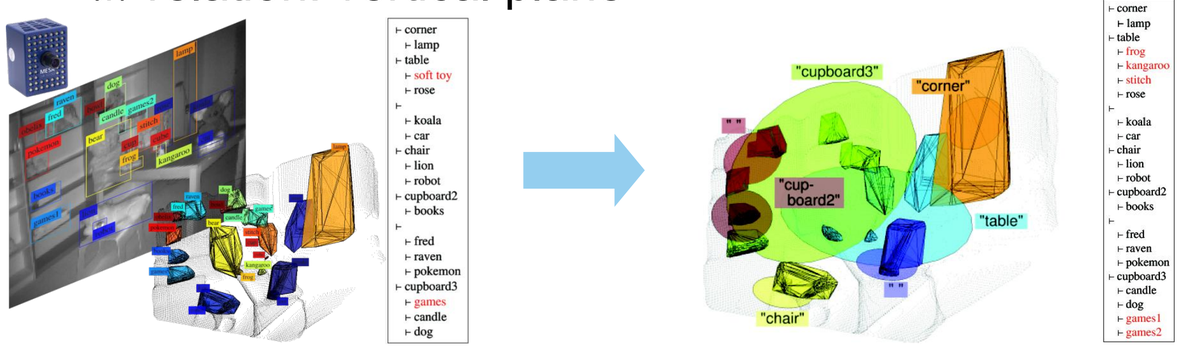


- rules:



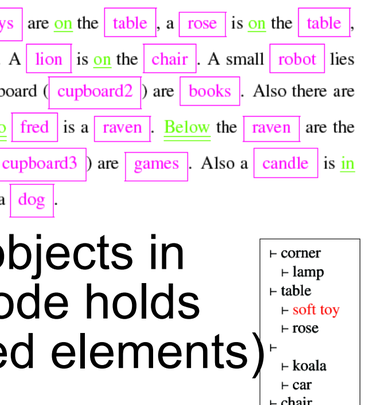
Computing Potential Planar Patches

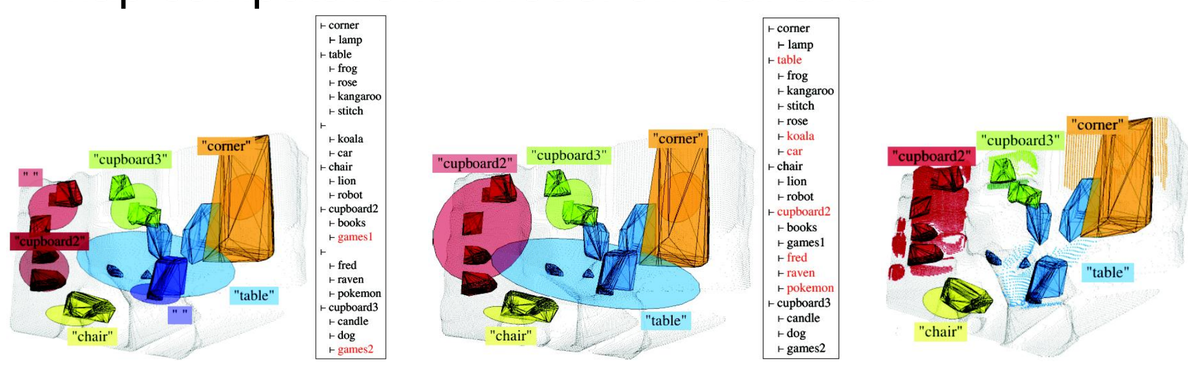
- planar surfaces:
  supporting structures of movable objects
- confirmed objects $O_{\text{con}} \in T$ and potential objects $O_{\text{pot}} \notin T$
- compute for each parent node from O_{con}
  plane parameters (P: $\boldsymbol{n} \cdot \boldsymbol{x} - d = 0$)
  - *on*-relation: horizontal plane
  - *in*-relation: vertical plane



Adaptation to Real Data

- extract bottom-up real planar patches
- correct misleadingly assigned objects
- introduce relations
  (not mentioned explicitly, e.g., “table”)
- map computational model on real data





Results (for 10 subjects of the experiment)