

# Cognition, Control, and Learning for Everyday Manipulation Tasks in Human Environments

Michael Beetz



IAS  
Intelligent  
Autonomous Systems



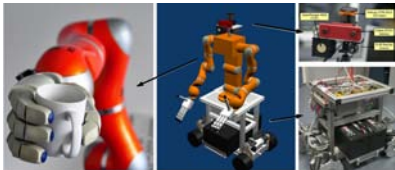
Japanese-German Symposium on Robotics





# Ultimate Goal

Realizing Personal Household Robots



with **compact** control programs performing an **open-ended set** of **everyday manipulation tasks** like

- ▶ laying the table and
- ▶ cleaning up

in **human environments** in a **natural, general, skillful, flexible** and **reliable** manner.

The robot is to enable people to stay **independent**, to keep them **active** and **participating** and thereby improve their **health state** and **quality of life**.





# Robots that Do It

## PR1, Willow Garage



## H.E.R.B., Intel Research



**note:** talks by Prof. Inaba and Prof. Dillmann

Japanese-German Symposium on Robotics

The Science of Everyday Activity

Michael Beetz





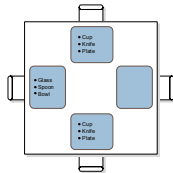
# Why It Is so Hard

The Household Robot Challenge — A Turing test for Action

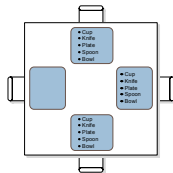
decide on

- ▶ **what to do?**
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ **for a single pick action:**
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?

**Typical breakfast:**



**Typical 3-person breakfast at a Thursday morning:**





# The Household Robot Challenge

A Turing test for Action

## decide on

- ▶ what to do?
- ▶ **which objects to take?**
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?



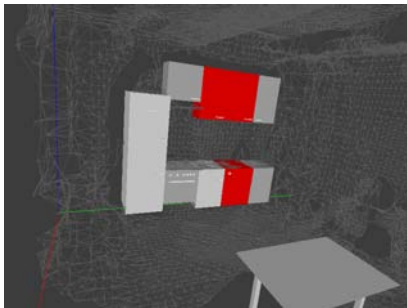


# The Household Robot Challenge

A Turing test for Action

## decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ **where to find the objects?**
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?





# The Household Robot Challenge

A Turing test for Action

## decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ **where to put them?**
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?





# The Household Robot Challenge

A Turing test for Action

## decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ **how to do it efficiently?**
- ▶ **for a single pick action:**
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?



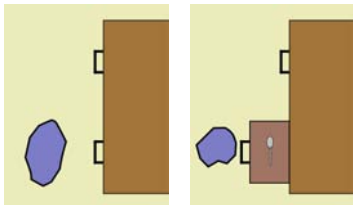
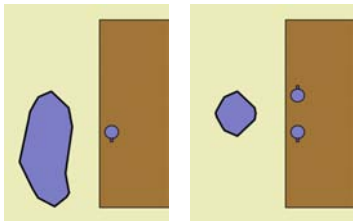


# The Household Robot Challenge

A Turing test for Action

## decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ **where to stand?**
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?





# The Household Robot Challenge

A Turing test for Action

## based on context

- ▶ object, object states
- ▶ environment
- ▶ scene
- ▶ task
- ▶ humans
- ▶ capabilities
- ▶ perceptual capabilities
- ▶ user preferences

## decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?





# The Household Robot Challenge

A Turing test for Action

with adopted contexts

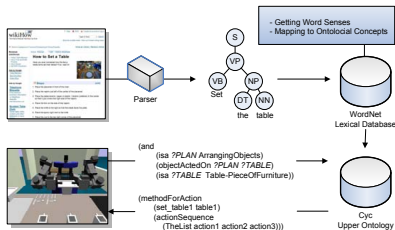
- ▶ novel Xs (tasks, objects, ...)
- ▶ context parameters change (preferences, capabilities, ...)

based on context

- ▶ object, object states
- ▶ environment
- ▶ scene
- ▶ task
- ▶ humans
- ▶ capabilities
- ▶ perceptual capabilities
- ▶ user preferences

decide on

- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - where to stand?
  - which hand(s) to use?
  - how to reach?
  - which grasp?
  - where to grasp?
  - how much force?
  - how to hold?





# Research Question

---

how can we specify robot controllers for such robots that achieve

- ▶ high performance,
- ▶ flexible and reliable, and
- ▶ natural

behavior and that can be implemented with

- ▶ reasonable programming effort?





## Research Question

---

how can we specify robot controllers for such robots that achieve

- ▶ high performance,
- ▶ flexible and reliable, and
- ▶ natural

behavior and that can be implemented with

- ▶ reasonable programming effort?

## Answer

- ▶ embed **cognitive capabilities**  
(learning, perception, reasoning, and planning)
- ▶ into the control programs
- ▶ in order to **achieve more flexibility, generality, reliability, and higher performance**

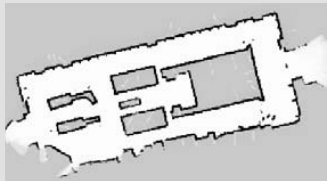




# Cognition-enabled Control — the Very Idea

Example: Map Acquisition and Map-based Navigation

## Model Acquisition



courtesy: Wolfram Burgard

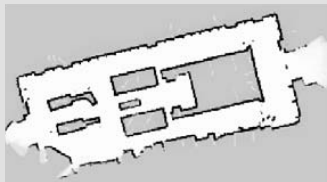




# Cognition-enabled Control — the Very Idea

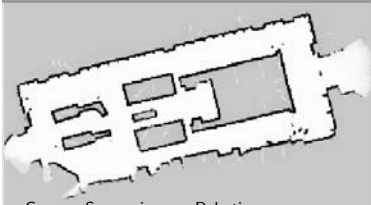
Example: Map Acquisition and Map-based Navigation

## Model Acquisition



courtesy: Wolfram Burgard

## Model Use

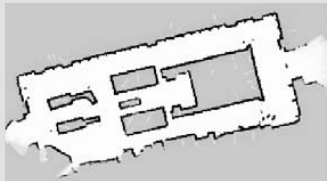




# Cognition-enabled Control — the Very Idea

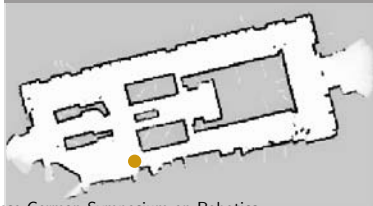
Example: Map Acquisition and Map-based Navigation

## Model Acquisition



courtesy: Wolfram Burgard

## Model Use



Where am I?

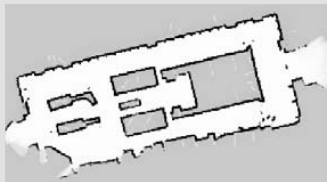




# Cognition-enabled Control — the Very Idea

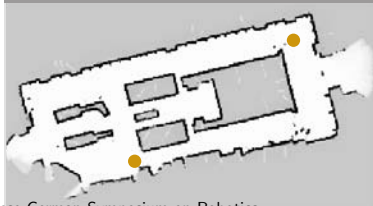
Example: Map Acquisition and Map-based Navigation

## Model Acquisition



courtesy: Wolfram Burgard

## Model Use



Where am I?

Where is L?

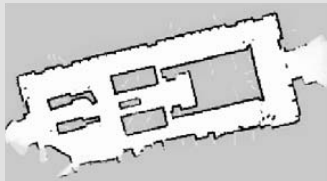




# Cognition-enabled Control — the Very Idea

Example: Map Acquisition and Map-based Navigation

## Model Acquisition



courtesy: Wolfram Burgard

## Model Use



Where am I?

Where is L?

How do I get there?





# Why Cognition-enabled Control?

## General Navigation Routine

```
routine navigate  $\langle tsk \rangle$   
  in parallel do continually estimate your position  
  whenever you are lost do relocalize  
  main process  
    if reachable(dest( $\langle tsk \rangle$ ))  
    then nav-plan  $\leftarrow$  compute-nav-plan(curr-pos, dest( $\langle tsk \rangle$ ))  
    execute nav-plan
```





# Why Cognition-enabled Control?

## General Navigation Routine

```
routine navigate  $\langle tsk \rangle$   
  in parallel do continually estimate your position  
  whenever you are lost do relocalize  
  main process  
    if reachable(dest( $\langle tsk \rangle$ ))  
    then nav-plan  $\leftarrow$  compute-nav-plan(curr-pos, dest( $\langle tsk \rangle$ ))  
    execute nav-plan
```

Cognitive mechanisms enable us to control the robot

- ▶ reliably
- ▶ flexibly
- ▶ efficiently

in concise control programs





# My Definition of Cognition

---

**Cognition** = information processing infrastructure that

- ▶ enables an **agent** *agt*
- ▶ to perform a set of **tasks** *tsk*
- ▶ better wrt **performance measure** *p*  
(typically generality, flexibility, reliability, performance, ...)
- ▶ based on
  - ▶ **experience** and **learning**
  - ▶ **knowledge/models** and **reasoning**
  - ▶ **forward models** and **planning/prediction**





# The Power of Cognitive Control

Motion Intelligence



- ▶ high speed
- ▶ high accuracy
- ▶ powerful
- ▶ fast feedback





# The Power of Cognitive Control

Motion Intelligence



- ▶ high speed
- ▶ high accuracy
- ▶ powerful
- ▶ fast feedback

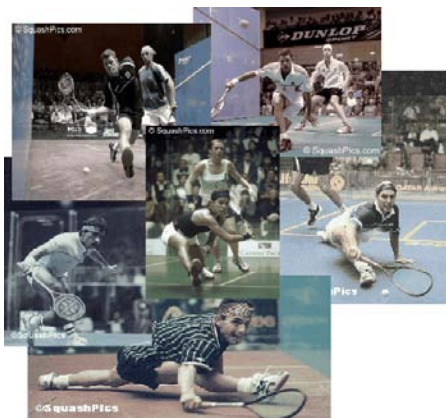
- ▶ slower
- ▶ less accurate
- ▶ less powerful
- ▶ slower feedback





# The Power of Cognitive Control

Motion Intelligence



## Characteristics

- ▶ acting in highly dynamic environments
- ▶ acting under uncertainty
- ▶ prediction-based motion control
- ▶ very fast & effective motion control
- ▶ acquisition of Squash playing skills





# How to Scale Cognitive Control?

## Model Acquisition



courtesy: Wolfram Burgard

## Model Use



Where am I?

Where is L?

How do I get there?



## with adopted contexts

- ▶ novel Xs (tasks, objects, ...)
- ▶ context parameters change (user preferences, capabilities)

### based on context

- ▶ object, object states
- ▶ environment
- ▶ scene
- ▶ task
- ▶ humans
- ▶ capabilities
- ▶ perceptual capabilities
- ▶ user preferences

### decide on




- ▶ what to do?
- ▶ which objects to take?
- ▶ where to find the objects?
- ▶ where to put them?
- ▶ how to do it efficiently?
- ▶ for a single pick action:
  - ▶ where to stand?
  - ▶ which hand(s) to use?
  - ▶ how to reach?
  - ▶ which grasp?
  - ▶ where to grasp?
  - ▶ how much force?
  - ▶ how to hold?





# Deployment of a Personal Household Robot

## Scenario

Mon	Tue	Wed	Thu	Fri	Sat	Sun
			acquire environment models & adapt skills			
Mon	Tue	Wed	Thu	Fri	Sat	Sun
			learn activities using www instructions			
Mon	Tue	Wed	Thu	Fri	Sat	Sun
			fully operational			
Mon	Tue	Wed	Thu	Fri	Sat	Sun
continue adaptation and learning						





# Research Questions and Approaches

---

- ▶ What is the “Science” of everyday manipulation activities?

Automated probabilistic models of everyday manipulation activities

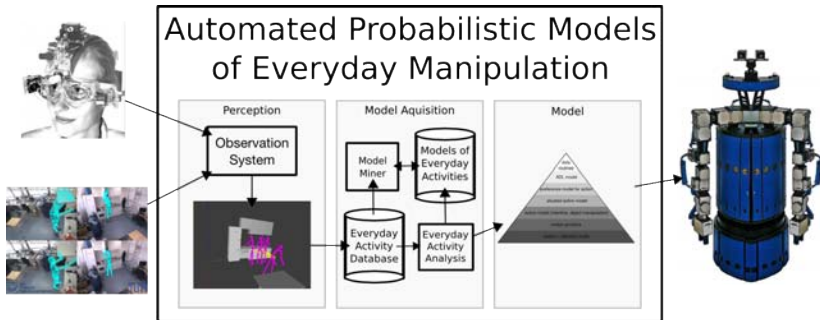
- ▶ How to scale autonomous robot control for household chores?

Web-enabled knowledge processing



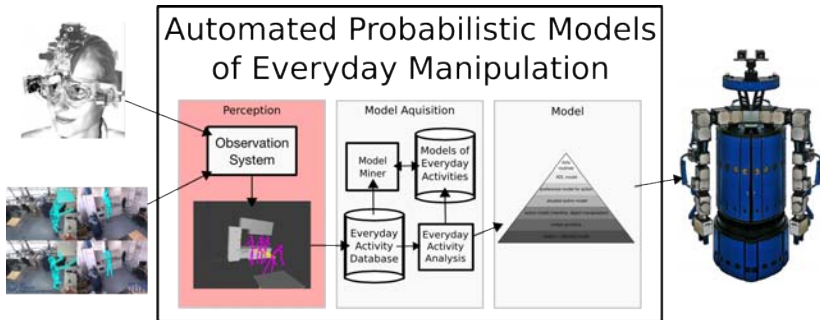


# Automated Probabilistic Activity Models



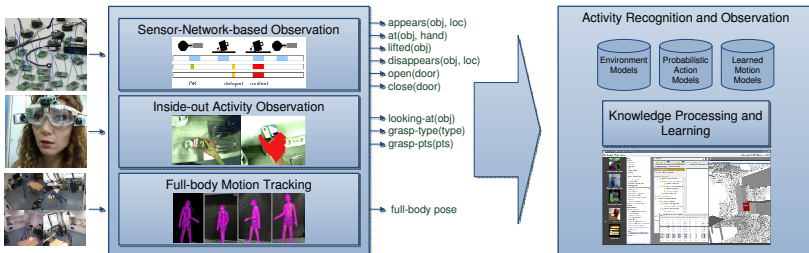
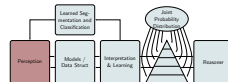


# Perception of Activities



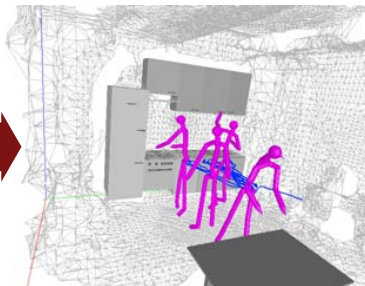
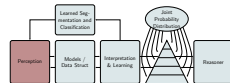


# Human action observation



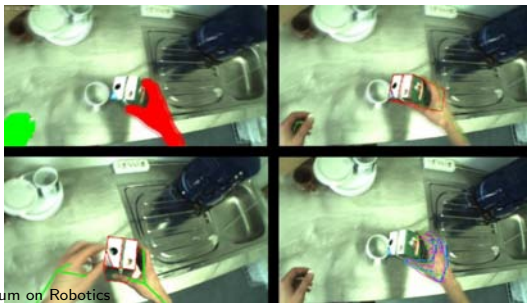
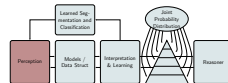


# Observing 'Setting the Table'



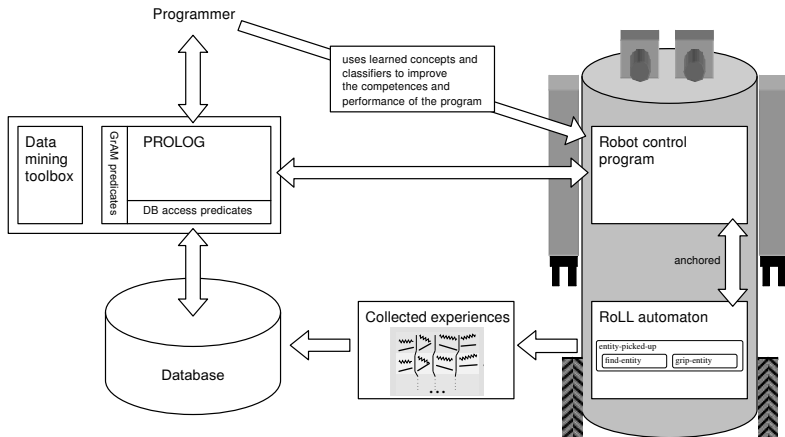


# Inside-out Activity Recognition





# Robot action observation



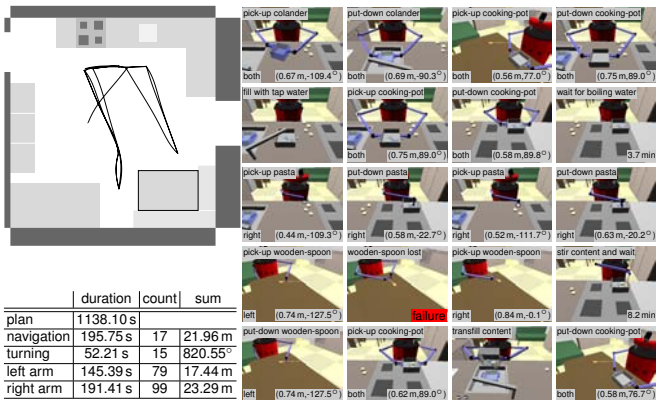
q





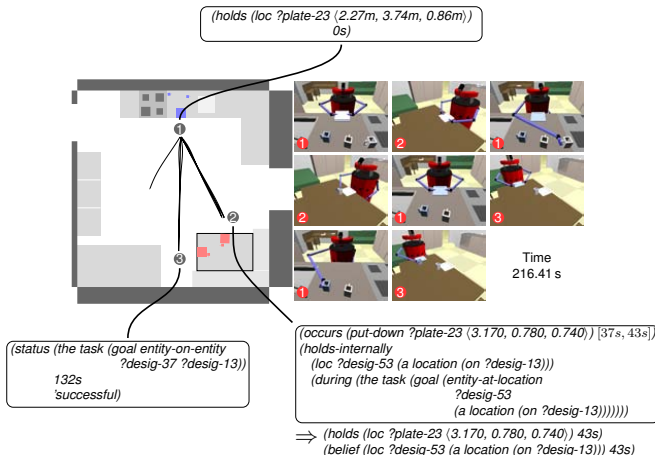
# Problem Solving Episodes

... as state trajectories in user-defined hybrid automata



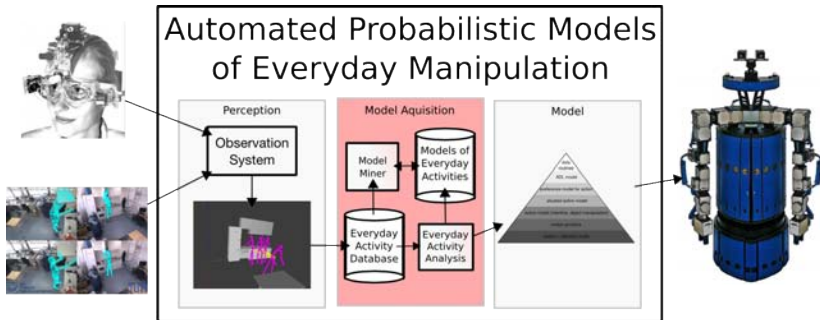


# Symbolic Representations of Episodes





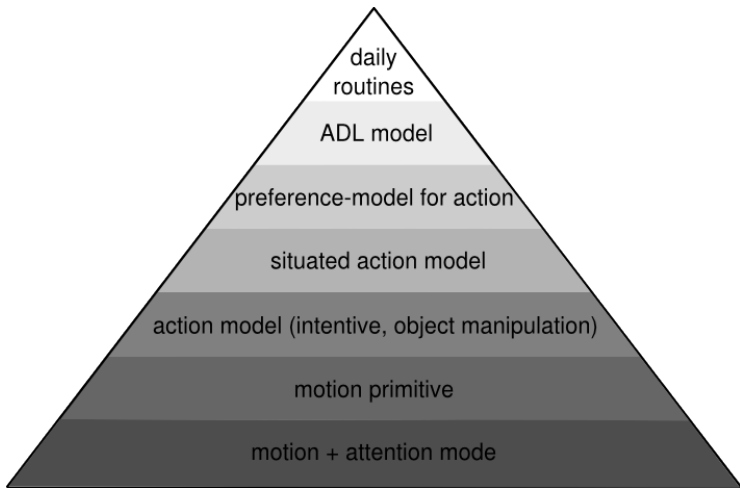
# Overview





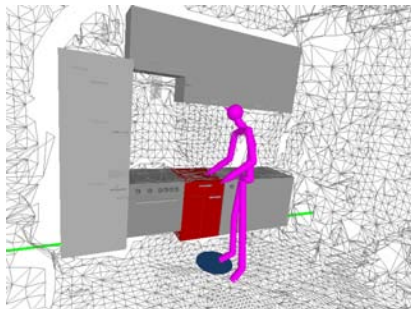
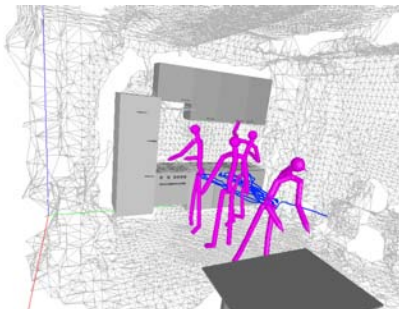
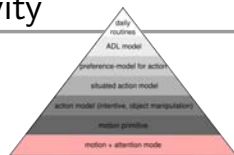
# Interpretation of Perceived Activity

---



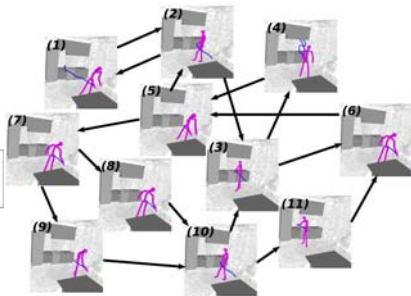
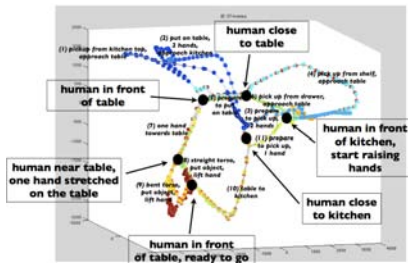
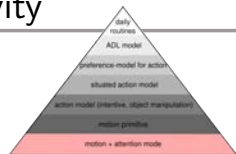


# Interpretation of Perceived Activity





# Interpretation of Perceived Activity

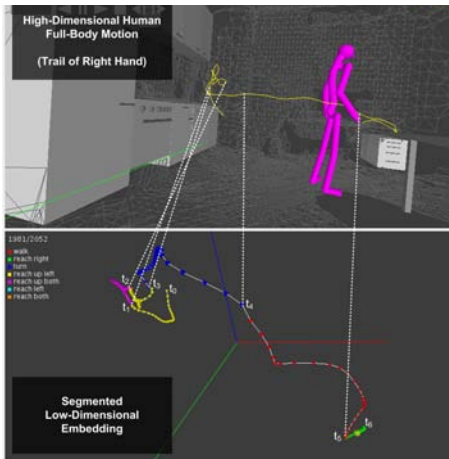
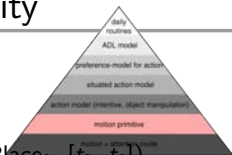


- Estimating Human Motion Capture Data of Table Setting Task
- Nonlinear Dimension Reduction Using ST-Isomap
- Paths in Lowdimensional Manifold Correspond to Actions
- Generation of motion FSM for Intention Recognition





# Interpretation of Perceived Activity



$\text{occurs}(\text{PickAndPlace}_8, [t_0, t_6])$   
 $\text{inActivity}(\text{PickAndPlace}_8, \text{SetTable}_1)$   
 $\text{handTrajectory}(\text{PickAndPlace}_8,$   
 $\quad \text{traj}(\text{FIR}, t_0, t_6): \blacksquare)$   
 $\text{actionT}(\text{PickAndPlace}_8, \text{PickAndPlace})$   
 $\text{performedBy}(\text{PickAndPlace}_8, P_1)$   
 $\text{objectActedOn}(\text{PickAndPlace}_8, \text{Cup}_3)$   
 $\text{target}(\text{PickAndPlace}_8, \text{Table}_1)$

$\text{occurs}(\text{Open}_{23}, [t_0, t_1])$   
 $\text{subEventOf}(\text{Open}_{23}, \text{PickAndPlace}_8)$   
 $\text{subEventT}(\text{Open}_{23}, \text{OpenCupboard})$   
 $\text{objectActedOn}(\text{Open}_{23},$   
 $\quad \text{OverheadCupboard}_2)$

$\text{occurs}(\text{PickUp}_8, [t_1, t_2])$

⋮





# Interpretation of Perceived Activity

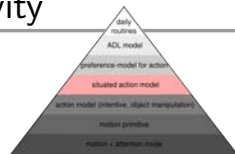


- ▶ Turing test for action:
  - ▶ where should the robot stand to perform an action?
- ▶ **ARPLACE**: the set of positions that an action can be performed successfully from
  - ▶ can be learned from experience
  - ▶ can be reasoned about

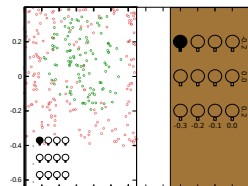
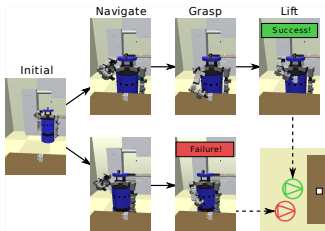




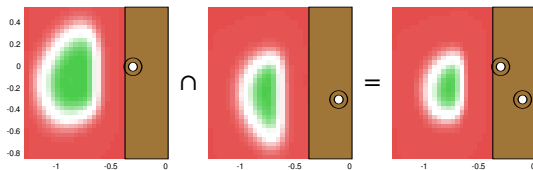
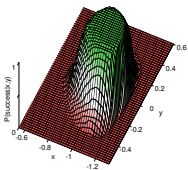
# Interpretation of Perceived Activity



Learn ARPLACES:



Use ARPLACES:





# Mining Env. Knowledge from Activity Data

Example

## Step 1: Extracting Positions

positions where pick and place actions occur while setting the table



## Step 2: Cluster Positions into Locations

cluster positions into locations



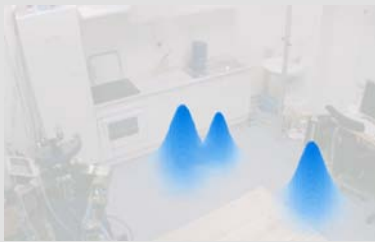


# Mining Env. Knowledge from Activity Data

Example

## Step 3: Detailed Location Representations

approximate positions of a location as a probability distribution



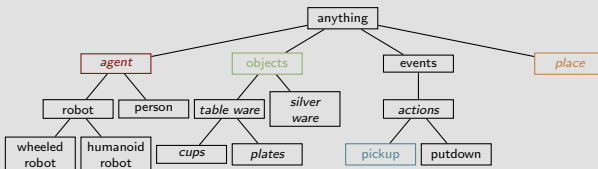


# Mining Env. Knowledge from Activity Data

## Example

### Step 4: Inferring the roles of locations

do(pickup(agent, object), s)  $\rightsquigarrow$  place





# Mining Env. Knowledge from Activity Data

## Example

### Step 4: Inferring the roles of locations

- ▶ causal structure between places and actions?

agent	plate	cup	table ware	knife	fork	silver ware	action	place
R	T	F	T	F	F	F	PICK	P1
R	F	T	T	F	F	F	PICK	P1
R	F	F	F	T	F	T	PICK	P3
R	T	F	T	F	F	F	PUT	P2
R	F	F	F	F	T	T	PUT	P2

- ▶ example decision rule:  
 action = pick up  
 $\wedge$  object = tableware  
 $\rightarrow P0$

- ▶ example class assertion  
 $\exists \text{actionPosition}^- . (\text{Pickup} \sqcap$   
 $\forall \text{object} . \text{Tableware}) \sqsubseteq P0$

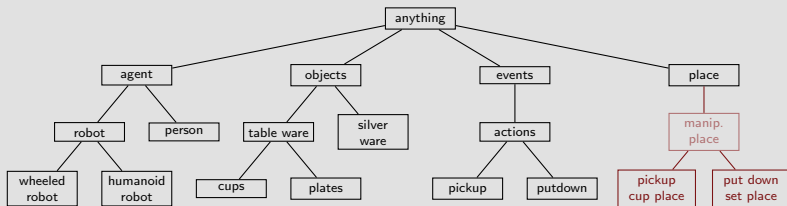




# Mining Env. Knowledge from Activity Data

## Example

### Step 4: Inferring the roles of locations



*pickup cup place* knowledge representation is linked to

- ▶ a map object
- ▶ to robot actions (control programs)
- ▶ recognition procedure



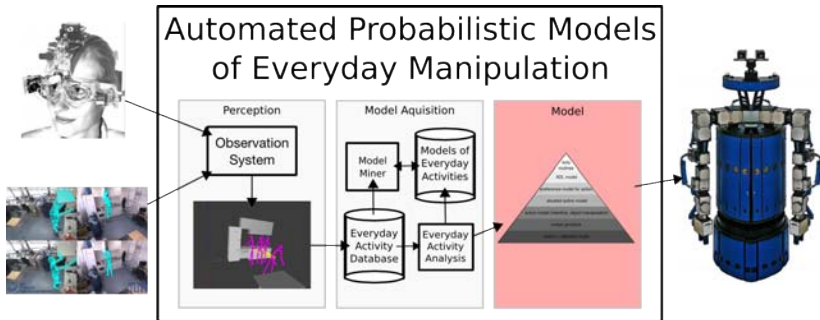


# Using Action-related Knowledge





# Overview





# Statistical Relational Models

---

## first-order representation

- ▶ quantification
- ▶ object classes
- ▶ relations
- ▶ expressiveness





# Statistical Relational Models

---

## first-order representation

- ▶ quantification
- ▶ object classes
- ▶ relations
- ▶ expressiveness

## Bayesian approach

- ▶ joint probability distribution
- ▶ diagnostic inferences
- ▶ causal inferences
- ▶ temporal inferences





# Statistical Relational Models

---

## first-order representation

- ▶ quantification
- ▶ object classes
- ▶ relations
- ▶ expressiveness

## Bayesian approach

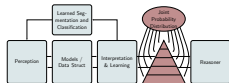
- ▶ joint probability distribution
- ▶ diagnostic inferences
- ▶ causal inferences
- ▶ temporal inferences

# Markov Logic





# Probability Distributions over Everyday Activities



given:

very large factbase on observed activity episodes

learn joint probability distribution over factbase:

$P(\text{what, who, where, when, how})$

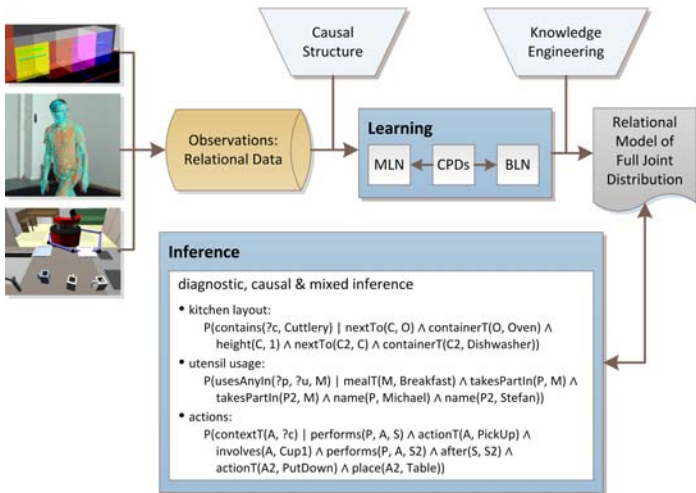
then you can infer:

- ▶  $P(\text{what, who, how} \mid \text{where, when})$
- ▶  $P(\text{where} \mid \text{who, when})$
- ▶ ...



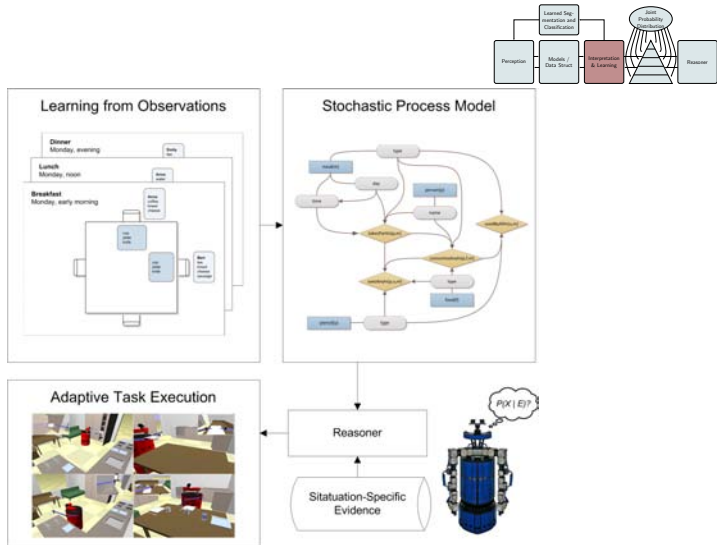


# Probabilistic Models from Observational Data





# Learning Table Setting with Markov Logic





## Example – Inferring Action Contexts

---

Given that a person has taken objects from the cupboard and has placed them on the table, we can conclude that the table is probably being set and assist in performing this task.

$$P(\text{contextT}(C, ?cT) \mid \text{next}(S1, S2) \wedge \text{next}(S2, S3) \wedge \text{context}(A1, C) \wedge \\ \text{actionT}(A1, \text{Pickup}) \wedge \text{performs}(P, A1, S1) \wedge \text{involves}(A1, \text{Cup}) \wedge \\ \text{place}(A1, \text{Cupboard}) \wedge \text{context}(A2, C) \wedge \text{actionT}(A2, \text{Putdown}) \wedge \\ \text{performs}(P, A2, S2) \wedge \text{involves}(A2, \text{Cup}) \wedge \text{place}(A2, \text{Table}))$$

$$= \langle \text{SettingTable: } 0.76, \text{CleaningUp: } 0.08, \text{Cooking: } 0.16 \rangle$$




## Example – Inferring People's Intentions

- ▶ (cont'd) What is the intended action of a person approaching a cupboard when setting the table for breakfast, and what type of object does it involve?

$$\begin{aligned}
 & P(\text{actionT}(A, ?a), \text{involves}(A, ?o) \mid \\
 & \quad \text{intention}(P, A, S1) \wedge \text{context}(A, C) \wedge \text{contextT}(C, \text{SettingTable}) \wedge \text{next}(S1, S2) \wedge \\
 & \quad \text{performs}(P, A1, S1) \wedge \text{context}(A1, C) \wedge \text{actionT}(A1, \text{Move}) \wedge \text{place}(A1, \text{Cupboard})) \\
 & = \langle \langle \text{Pickup: } 0.61, \text{Move: } 0.20, \text{Putdown: } 0.19 \rangle, \\
 & \quad \langle \text{Cup: } 0.35, \text{Cuttlery: } 0.18, \text{Plate: } 0.16, \dots \rangle \rangle
 \end{aligned}$$

- ▶ What is the intention of a person picking up a cup from the cupboard in the same context?

$$\begin{aligned}
 & P(\text{actionT}(A2, ?a2), \text{place}(A2, ?p2), \text{actionT}(A3, ?a3) \mid \\
 & \quad \text{next}(S1, S2) \wedge \text{next}(S2, S3) \wedge \text{contextT}(C, \text{SettingTable}) \wedge \text{context}(A1, C) \wedge \\
 & \quad \text{actionT}(A1, \text{Pickup}) \wedge \text{performs}(P, A1, S1) \wedge \text{involves}(A1, \text{Cup}) \wedge \text{place}(A1, \text{Cupboard}) \wedge \\
 & \quad \text{context}(A2, C) \wedge \text{intention}(P, A2, S1) \wedge \text{context}(A3, C) \wedge \text{intention}(P, A3, S2)) \\
 & = \langle \langle \text{Move: } 0.73, \text{Putdown: } 0.20, \text{Pickup: } 0.07 \rangle, \\
 & \quad \langle \text{Table: } 0.74, \text{Cupboard: } 0.09, \dots \rangle, \\
 & \quad \langle \text{Putdown: } 0.63, \text{Move: } 0.24, \text{Pickup: } 0.13 \rangle \rangle
 \end{aligned}$$





## Example – Inferring Meal Types. . .

Given the things people use during their meals, what are they most likely to be consuming and what is the most likely type of meal?

- ▶ One person is taking part in the meal and using a bowl, a cup, no glass and no plate.

$$\begin{aligned}
 &P(\text{mealT}(M, ?mt), \text{consumesAnyIn}(P1, ?ft, M) \mid \\
 &\quad \neg \text{usesAnyIn}(P1, \text{Glass}, M) \wedge \text{usesAnyIn}(P1, \text{Bowl}, M) \wedge \\
 &\quad \neg \text{usesAnyIn}(P1, \text{Plate}, M) \wedge \text{usesAnyIn}(P1, \text{Cup}, M)) \\
 &= \langle \langle \text{Breakfast: } 0.87, \text{ Lunch: } 0.09, \text{ Dinner: } 0.03 \rangle , \\
 &\quad \langle \text{Cereals: } 0.87, \text{ Tea: } 0.83, \text{ Coffee: } 0.22, \text{ Soup: } 0.12, \text{ Bread: } 0.09, \dots \rangle \rangle
 \end{aligned}$$

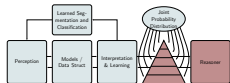
- ▶ What if another person joined and used only a fork, a glass and a plate?

$$\begin{aligned}
 &P(\text{mealT}(M, mt), \text{consumesAnyIn}(P1, ft, M) \mid \dots \wedge \text{usesAnyIn}(P2, \text{Plate}, M) \wedge \\
 &\quad \text{usesAnyEIn}(P2, \text{Knife}, M) \wedge \text{usesAnyIn}(P2, \text{Glass}, M) \wedge \\
 &\quad \neg \text{usesAnyIn}(P2, \text{Bowl}, M) \wedge \neg \text{usesAnyEIn}(P2, \text{Teaspoon}, M) \wedge \\
 &\quad \text{usesAnyEIn}(P2, \text{Fork}, M) \wedge \neg \text{usesAnyEIn}(P2, \text{Spoon}, M)) \\
 &= \langle \langle \text{Lunch: } 1.00 \rangle , \\
 &\quad \langle \text{Soup: } 1.00, \text{ Coffee: } 1.00, \text{ Bread: } 0.03, \dots \rangle \rangle
 \end{aligned}$$





# Reasoning patterns:



$$P(\text{state}(\text{object}, t_i) \mid \text{at}(\text{objects}, \text{locations}, t_i))$$

- ▶ What is the state of a cup that is on the table?

$$P(\text{stateOfUtensilIn}(U, C, ?s) \mid \text{utensilT}(U, \text{Cup}), \\ \text{placeOfUtensilIn}(U, C, \text{Table})) \\ = \langle \text{Dirty: } 0.3577, \text{ Clean: } 0.6423 \rangle$$

$$P(\text{location}(\text{object}, t_i) \mid \text{state}(\text{object}, \text{dirty}, t_i))$$

- ▶ What is the location of a dirty plate?

$$P(\text{placeOfUtensilIn}(U, C, ?s) \mid \text{utensilT}(U, \text{Plate}), \\ \text{stateOfUtensilIn}(U, C, \text{Dirty})) \\ = \langle \text{Dishwasher: } 0.9289, \text{ Table: } 0.0711 \rangle$$



# Web-enabled Knowledge Processing for Personal Assistive Robots





# Egg Cracking

---

A cook is cracking a raw egg against a glass bowl. Properly performed, the impact of the egg against the edge of the bowl will crack the eggshell in half. Holding the egg over the bowl, the cook will then separate the two halves of the shell with his fingers, enlarging the crack, and the contents of the egg will fall gently into the bowl. The end result is that the entire contents of the egg will be in the bowl, with the yolk unbroken, and that the two halves of the shell are held in the cook's fingers.

What happens if: The cook brings the egg to impact very quickly? Very slowly? The cook lays the egg in the bowl and exerts steady pressure with his hand? The cook, having cracked the egg, attempts to peel it off its contents like a hard-boiled egg? The bowl is made of looseleaf paper? of soft clay? The bowl is smaller than the egg? The bowl is upside down? The cook tries this procedure with a hard-boiled egg? With a coconut? With an M & M?





# Why Web-enabled Robots?

---

3 reasons:





# Why Web-enabled Robots?

---

3 reasons:

- ▶ scaling





# Why Web-enabled Robots?

---

3 reasons:

- ▶ scaling
- ▶ scaling





# Why Web-enabled Robots?

---

3 reasons:

- ▶ scaling
- ▶ scaling
- ▶ **scaling**





# Why Web-enabled Robots?

---

3 reasons:

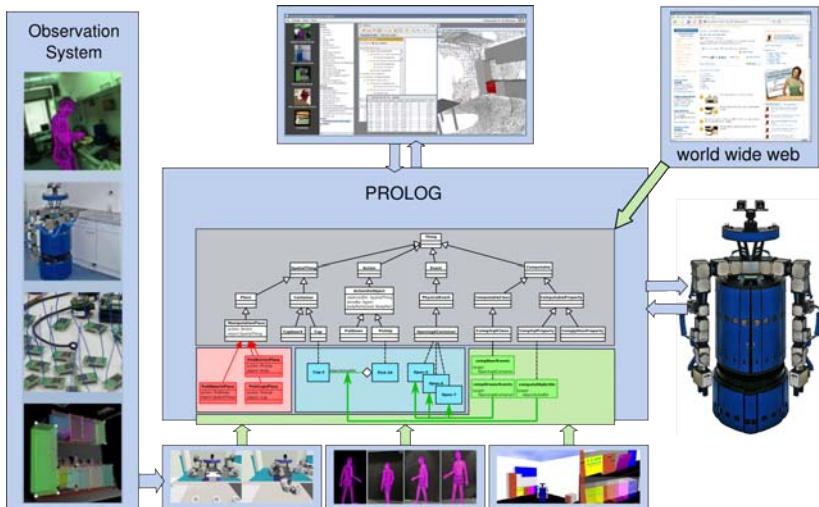
- ▶ scaling
- ▶ scaling
- ▶ **scaling**

Tom Mitchell: *In fact, I hereby offer to bet anyone a lobster dinner that by 2015 we will have a computer program capable of automatically reading at least 80% of the factual content across the entire English-speaking web, and placing those facts in a structured knowledge base.*





# COTESYS Knowledge Processing System





# About eggs

---

- ▶ Important knowledge when handling eggs
  - ▶ What is the meaning of the word 'egg'?
  - ▶ How does an egg look like?
  - ▶ What is the form of an egg?
  - ▶ Common knowledge about eggs
  - ▶ How to handle eggs?
  - ▶ Generally: How do technical devices work?
  
- ▶ Knowledge can be looked up automatically using the internet





# The meaning of the word 'egg'

- ▶ Meaning
- ▶ Look
- ▶ Form
- ▶ Common
- ▶ How-to
- ▶ Function

What is the meaning of the word 'egg'?

→ WordNet explains noun and verb usages of a word

WordNet Search - 3.0 - [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

## Noun

- [S:](#) (n) **egg** (animal reproductive body consisting of an ovum or embryo together with nutritive and protective envelopes; especially the thin-shelled reproductive body laid by e.g. female birds)
- [S:](#) (n) **egg**, **eggs** (oval reproductive body of a fowl (especially a hen) used as food)
- [S:](#) (n) **testis**, **testicle**, **orchis**, **ball**, **ballock**, **bollock**, **nut**, **egg** (one of the two male reproductive glands that produce spermatozoa and secrete androgens) "*she kicked him in the balls and got away*"

## Verb

- [S:](#) (v) **egg** (throw eggs at)
- [S:](#) (v) **egg** (coat with beaten egg) "*egg a schnitzel*"

[WordNet home page](#)





# The look of an egg

- ▶ Meaning
- ▶ Look
- ▶ Form
- ▶ Common
- ▶ How-to
- ▶ Function

How does an egg look like?

→ Google Images can be searched for images showing eggs

Google

Moderater SafeSearch ist aktiviert

Bilder Anzeigt:   Ergebnisse 1 - 18 von etwa 15.600.000 (0,13 Sekunden)

 <p>Chicken or Egg? 641 x 609 - 7k - jpg <a href="http://idology.wordpress.com">idology.wordpress.com</a></p>	 <p>Eggs 310 x 414 - 13k - jpg <a href="http://www.bonappetit.com">www.bonappetit.com</a></p>	 <p>Inhalation of egg particles ... 600 x 400 - 13k - jpg <a href="http://www.allergyadvisor.com">www.allergyadvisor.com</a></p>
 <p>Sunny-side up egg for stress relief 400 x 325 - 10k - jpg <a href="http://www.poppadget.net">www.poppadget.net</a></p>	 <p>Cracking Open the Egg 500 x 333 - 40k - jpg <a href="http://healthybirds.umd.edu">healthybirds.umd.edu</a></p>	 <p>... not a painted egg. 619 x 800 - 80k - jpg <a href="http://witchdoctor.wordpress.com">witchdoctor.wordpress.com</a></p>



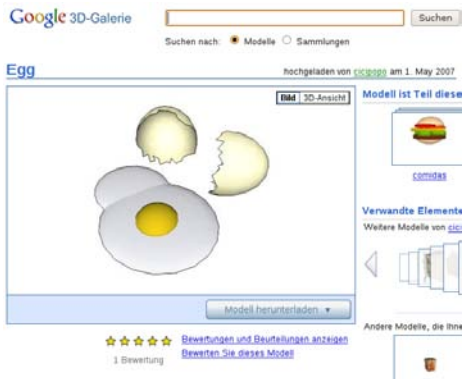


# The form of an egg

- ▶ Meaning
- ▶ Look
- ▶ **Form**
- ▶ Common
- ▶ How-to
- ▶ Function

What is the form of an egg?

→ Look up 3D CAD models at Google 3D Warehouse





# Common knowledge about eggs

- ▶ Meaning
- ▶ Look
- ▶ Form
- ▶ Common
- ▶ How-to
- ▶ Function

Common knowledge about eggs  
→ Browse OpenMind or OpenCyc


**Teaching robots the stuff we all know**  
 Think of Open Mind as a young child, learning from everyone on the Web.

Welcome David! 0 accepted / 2 under review

Search:   [Information](#) [Logout](#)

**Search Results for egg**

Author	Knowledge
mykel	You might cause a kitchen floor to become dirty by making a egg become dropped
enna	The picture h0099.gif is of a carton of eggs
enna	An object that you might find in a house is a carton of eggs
killroy6	Do not include the whole egg when making scramble eggs, use only the yolk and egg white
u108306	You frequently want a egg to be fresh
u108306	You generally find a egg in the same room as a fridge
u108306	You almost always want a egg to be whole



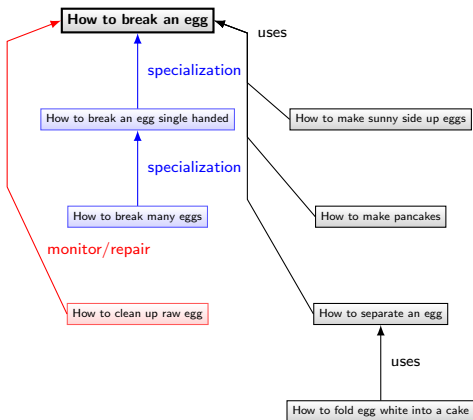


# How to handle eggs

- ▶ Meaning
- ▶ Look
- ▶ Form
- ▶ Common
- ▶ **How-to**
- ▶ Function

How to handle eggs?

→ Detailed How-Tos can be found at [wikiHow.org](http://wikiHow.org), [eHow.com](http://eHow.com) or [kuechengoetter.de](http://kuechengoetter.de) (videos)





# How to break an egg

"How to break an egg" - a How-To from wikihow.com:

Generally a How-to can include:

✓ Things you'll need

🎯 Steps needed

✳ Tips

⚠ Warnings

📺 Instruction video

💬 Discussion

The screenshot shows the wikiHow page for "How to Break an Egg". The page includes a search bar, navigation links, and a list of steps. The main content area features a photograph of a white egg and a list of steps: 1. Grasp the egg in your dominant hand. 2. Hold it between the thumb and first two fingers. 3. Now tap the egg firmly onto a hard surface to crack the shell. Below the steps is an "edit" link.





# How to separate an egg

”How to spearate an egg” - a How-To video from küchengötter.de:

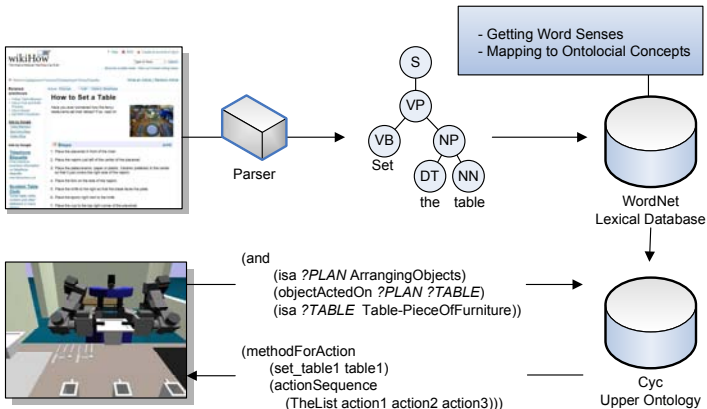
The screenshot shows the website küchengötter.de. The main content area features a video player titled "Eier richtig trennen" (Separating eggs correctly). The video title is "Eier richtig trennen" and it was created on 21.07.08 by küchengötter. The description asks how to separate eggs cleanly and correctly, mentioning a chef named Bodo Hasenberg. The video player shows hands cracking an egg into a glass bowl. The sidebar on the right contains "Küchenpraxis Tipps" (Kitchen Practice Tips) with three items: "Orangescreme selber machen" (Homemade orange cream), "Rosenkohl richtig putzen und verarbeiten" (Cleaning and processing cauliflower), and "Sauce Hollandaise richtig zubereiten" (Preparing Hollandaise sauce). There is also a "Jetzt kostenlos registrieren" (Register now for free) section.





# Import of Task Descriptions from the WWW

**Goal:** Convert natural language instructions into logical formulas and store them in the KB.





# Problems with Web Instructions

Web instructions require commonsense knowledge to be executable

- ▶ vague instructions  
*stir occasionally*
- ▶ qualitative spatial specifications  
*in front of the chair*
- ▶ missing aspects  
*put the plate in front of the chair*
- ▶ missing instructions  
*switch off the oven*



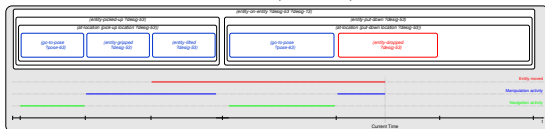


# Plan Debugging Overview

Logged  
Problem-Solving  
Episode



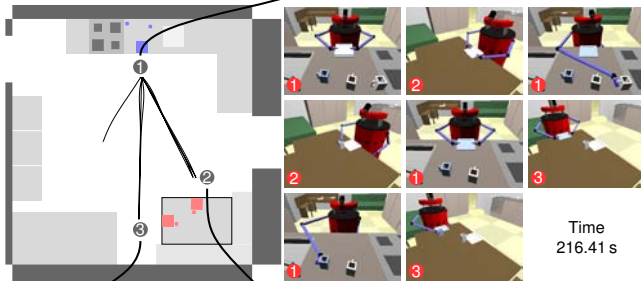
Generation of Virtual Episode Representation





# Plan Debugging Overview

*(holds (loc ?plate-23 (2.27m, 3.74m, 0.86m))  
0s)*



Time  
216.41 s

*(status (the task (goal entity-on-entity  
?desig-37 ?desig-13))  
132s  
'successful)*

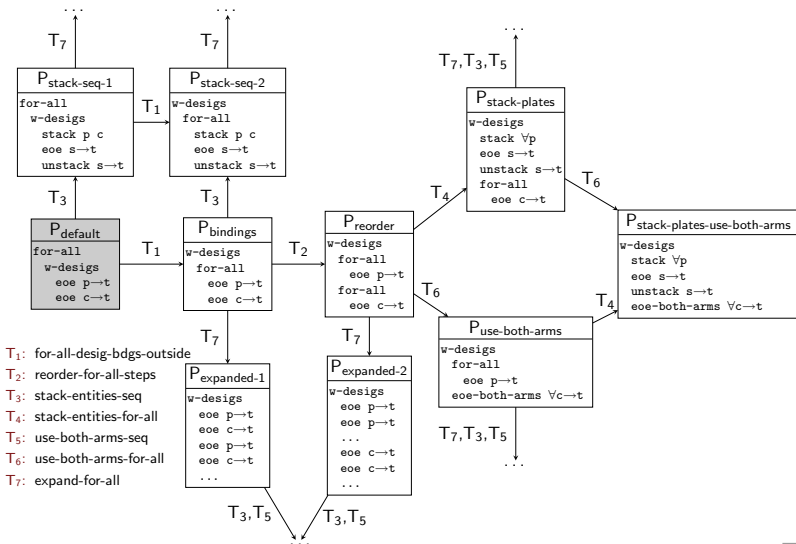
*(occurs (put-down ?plate-23 (3.170, 0.780, 0.740)) [37s, 43s])  
(holds-internally  
(loc ?desig-53 (a location (on ?desig-13)))  
(during (the task (goal (entity-at-location  
?desig-53  
(a location (on ?desig-13)))))))*

⇒ *(holds (loc ?plate-23 (3.170, 0.780, 0.740)) 43s)  
(belief (loc ?desig-53 (a location (on ?desig-13))) 43s)*





# Plan Optimization





# Conclusions

---

- ▶ personal household robot
- ▶ action-centered knowledge representation and processing
  - ▶ inferring symbolic models from observed action
  - ▶ action-related concepts
- ▶ probabilistic hybrid (continuous/discrete) first-order representations
- ▶ web-enabled knowledge processing

