Case Studies in Design Informatics

Lecture 5

The JAMES robot bartender: design and operation

Ron Petrick
School of Informatics
University of Edinburgh

james-project.eu

8 October 2015

Overview

1. Motivation for social interaction and robotics
2. Overview of the JAMES project
3. Design of the JAMES system
4. What does JAMES do (and why does it do it)?
5. Interacting with JAMES

⇒ Focus on design, implementation, and operation.

Two people walk into a bar...

Two people, A and B, each individually approach a bartender.

Bartender (to A):	How can I help you?
Person A:	A pint of cider, please.

Person C approaches the bartender and attracts his attention by gesturing.

Bartender (to C):	How can I help you?
Person C:	I'd like a pint of bitter.
Bartender:	(Serves C)
Bartender (to B):	What will you have?
Person B:	A glass of red wine.
Bartender:	(Serves B)
Bartender:	(Serves A)

Two people walk into *another* bar...

Two people, A and B, each individually approach a bartender.

Bartender (to A):	How can I help you?
Person A:	A pint of cider, please.

Person C approaches the bartender and attracts his attention by gesturing.

Bartender (to C):	Just a moment please.
Bartender:	(Serves A)
Bartender (to B):	What will you have?
Person B:	A glass of red wine.
Bartender:	(Serves B)
Bartender (to C):	Thanks for waiting. How can I help you?
Person C:	I'd like a pint of bitter.
Bartender:	(Serves C)

Two interactions

Two people, A and B, each individually approach a bartender

Bartender (to A):	How can I help you?	Bartender (to A):	How can I help you?
Person A:	A pint of cider, please.	Person A:	A pint of cider, please.

Person C approaches the bartender and attracts his attention by gesturing

Bartender (to C):	How can I help you?	Bartender (to C):	Wait a moment please
Person C:	I'd like a pint of bitter.	Bartender:	(Serves A)
Bartender:	(Serves C)	Bartender (to B):	What will you have?
Bartender (to B):	What will you have?	Person B:	A pint of Guinness.
Person B:	A pint of Guinness.	Bartender:	(Serves B)
Bartender:	(Serves B)	Bartender (to C):	Thanks for waiting.
Bartender:	(Serves A)		How can I help you?
		Person C:	I'd like a pint of bitter.
		Bartender:	(Serves C)

- Is one interaction "better" than the other?
- Does the purpose of the interaction matter?
- What if some (or all) of the participants are robots?

Two interactions (2)

Two people, A and B, each individually approach a bartender

Bartender (to A):	How can I help you?	Bartender (to A):	How can I help you?
Person A:	A pint of cider, please.	Person A:	A pint of cider, please.

Person C approaches the bartender and attracts his attention by gesturing

Bartender (to C):	How can I help you?	Bartender (to C):	Wait a moment please
Person C:	I'd like a pint of bitter.	Bartender:	(Serves A)
Bartender:	(Serves C)	Bartender (to B):	What will you have?
Bartender (to B):	What will you have?	Person B:	A pint of Guinness.
Person B:	A pint of Guinness.	Bartender:	(Serves B)
Bartender:	(Serves B)	Bartender (to C):	Thanks for waiting.
Bartender:	(Serves A)		How can I help you?
		Person C:	I'd like a pint of bitter.
		Bartender:	(Serves C)

- Both interactions result in the customers achieving their task goals.
- The first interaction is shorter.
- The second interaction is considered to be more **socially appropriate**.

Why social interaction?

- **Successful task interaction often relies on social interaction.**
  - May be several ways to achieve a task-based goal.
  - Appropriate social behaviour can lead to higher participant satisfaction.
- **Social interaction can be seen as an instance of joint action.**
  - Involves coordination of participant actions.
  - Inherently multimodal: speech, gesture, gaze, expression, etc.
- **Social interaction is often multi-party, dynamic, short-horizon.**
  - In contrast to one-on-one, companion-style relationships.
  - Interactions are often "one shot"; may not have an opportunity to recover from a poor interaction.

Some questions to consider...

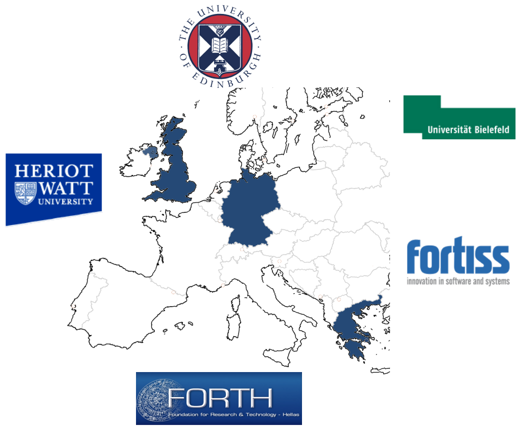
- Why should we care about appropriate social behaviour?
- Why should we care about designing a robot that is capable of such interactions?
- What impact does social interaction have on the design of a robot?
- Does the design of a robot impact the types of interactions (social or otherwise) that it can produce?
- Does appropriate social behaviour really have an impact on the interactions humans have with robots?
- ...?

Meet the bartender: JAMES

Image: fortiss GmbH

JAMES

Joint Action for Multimodal Embodied Social Systems · james-project.eu

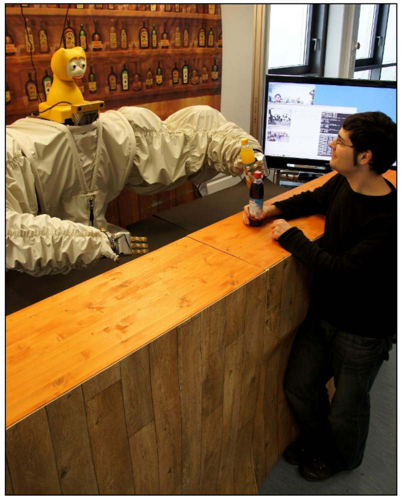


- 3.5 year project (2011–2014)
- Consortium of 5 European partners
- Focus on socially appropriate, multi-party, multimodal interactions in a robot bartending scenario.

Objectives

1. **Data collection and analysis:** Record and analyse the social and task-based behaviour of humans engaged in multimodal joint activities.
2. **Social modelling:** Design and train a model of social interaction, using annotated data from the human experiments.
3. **Representation and learning:** Endow the model with the ability to learn and adapt to human behaviours, and handle partial or uncertain information about the physical world and mental states of human users.
4. **Implementation:** Implement the model of social interaction on a physical robot platform, initially based on the JAST framework.
5. **Evaluation:** Evaluate the implemented robot system with multiple human users in a bartender scenario.

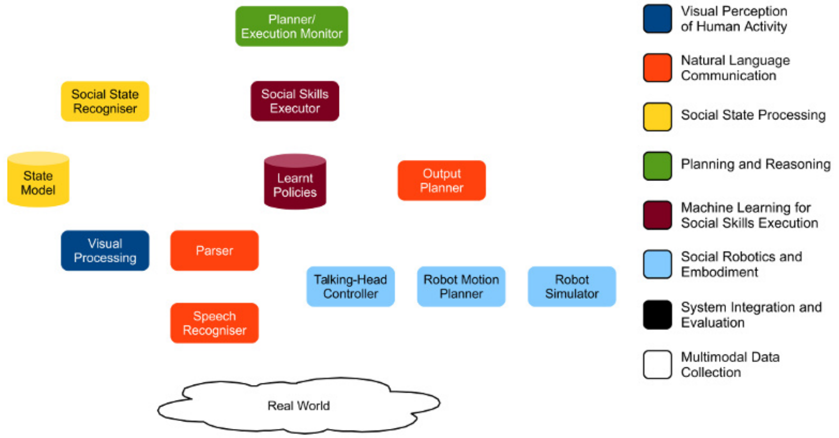
Target scenario: task-based social interaction



- Robot bartender must respond to user requests in a dynamic setting with multiple users and short interactions in German or English.
- Interactions incorporate both **task-based aspects** (e.g., ordering and serving drinks) and **social aspects** (e.g., managing multiple interactions).
- How important is **social interaction**?

Image: fortiss GmbH

Research themes in JAMES



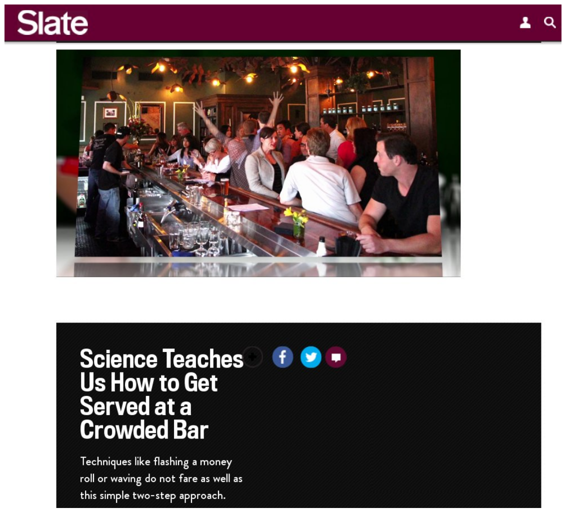
- How should we design our robot in order to achieve the project's objectives, given our requirements and constraints?

Design considerations: what do humans do?

Image: Universität Bielefeld

- Study how human customers interact with human bartenders to identify the (non-verbal) signals humans are using. How can these results be applied to the robot bartender?

Slate video



http://www.slate.com/articles/video/science/2013/09/how_to_get_served_at_a_bar_electronic_bartender_and_study_show_the_easiest.html

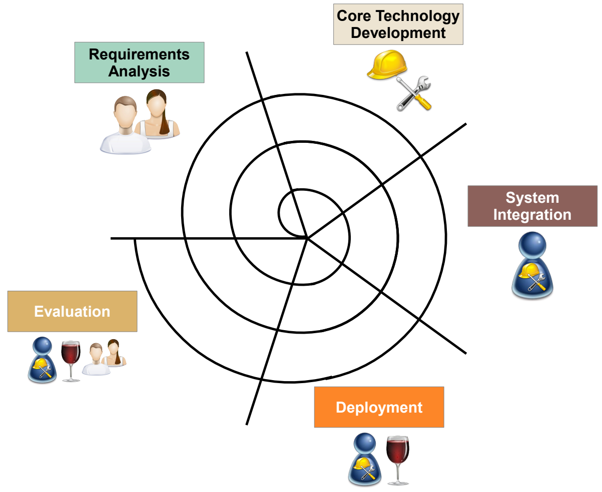
Design considerations: what should our bartender do?

- What activities are supported?
  - Asking customers for drinks
  - Clarifying drink requests
  - Handing over drinks
  - Keeping track of the order people arrive at the bar
  - Identifying groups
  - ...
- What activities are not supported?
  - Physically pouring drinks
  - Handling money
  - Small talk
  - ...

Design considerations: hardware and software

- Make use of existing and available robot hardware as much as possible
  - Cat head
  - Industrial arms
  - Lab infrastructure
- Supplement existing hardware
  - Kinect
  - 3D cameras
  - New robot hands
  - Tablets
- Use modern software development tools
  - Module-dependent choice of programming language: Java, Python, C++, C, …
  - Internet Communications Engine (ICE) for communication between components.
  - Modern build system (cmake)
  - Redistributable packages across a range of development and build environments (Linux, Mac, Windows).

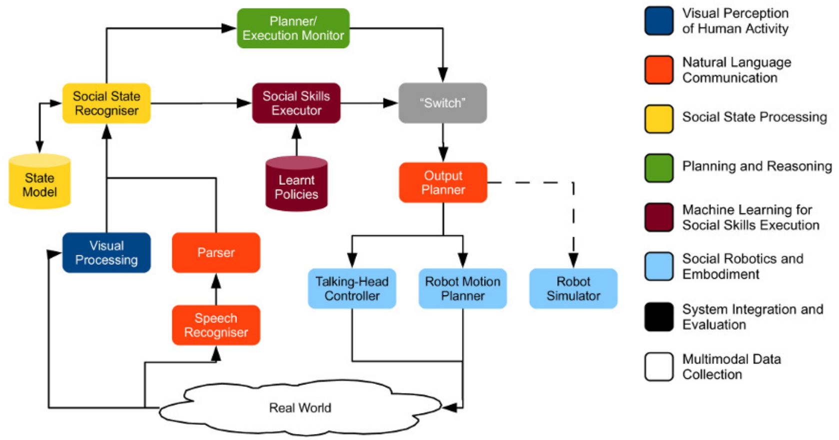
Design considerations: continual development



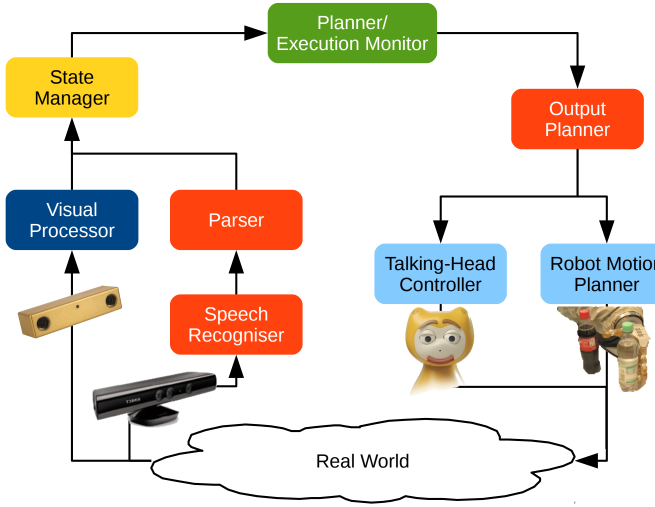
Design considerations: other factors

- Distributed design
  - Global agreement about the aims of the system.
  - No global control over the internal design of individual components/modules.
  - Individual components in the system are the responsibility of particular partners.
  - Interfaces are agreed upon in a pairwise fashion between individual components.
- Practical considerations
  - Budgets are tight
  - Time is tight
  - Research agendas are (somewhat) fixed
- Geographical considerations
  - The development team was spread over Europe
  - The robot was in Munich and isn't very portable…
  - Testing? Debugging? Evaluation?
- …

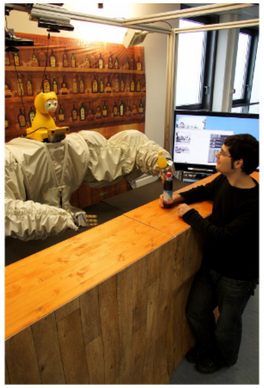
JAMES system architecture



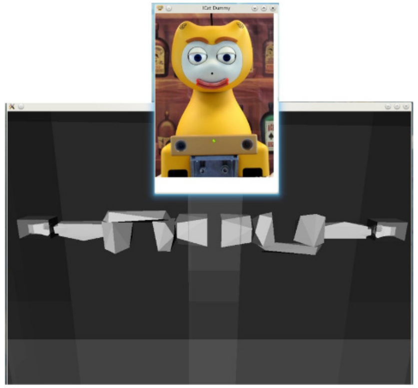
Simplified architecture



Hardware and simulation



Physical hardware (shown in video)

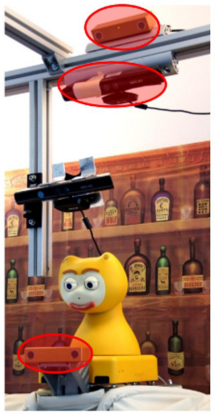


Simulated hardware (used for demo)

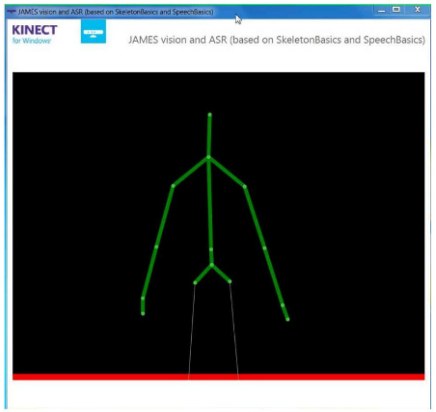
Credit: A. Gaschler and M. Giuliani, fortiss GmbH

- Provides the robot's embodiment and supplies the primary means for the robot to interact with the real world.

Visual processing



Full system (shown in video)

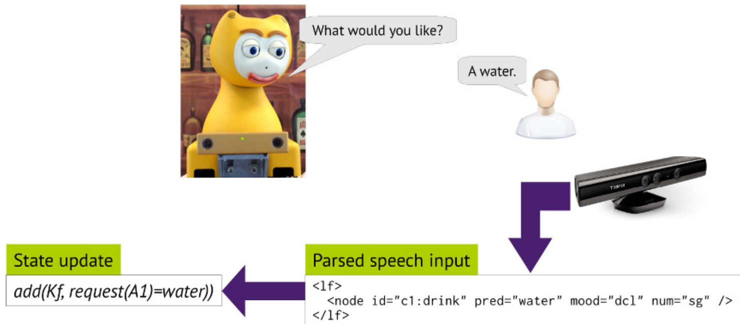


Kinect-only (used for demo)

Credit: M. Pateraki and M. Sigalas, FORTH

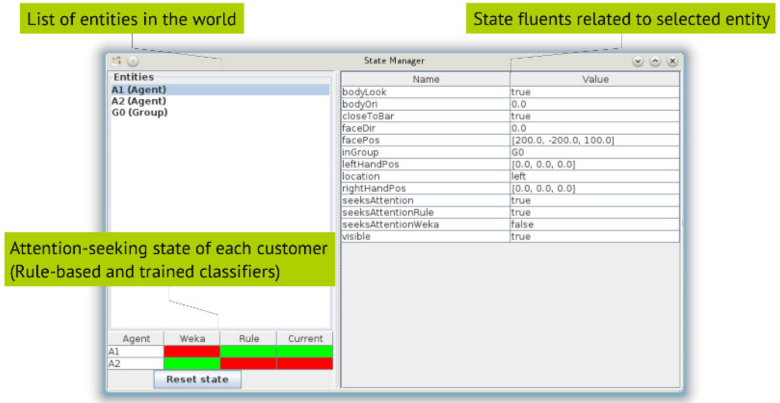
- Provides visual information about the customers in a scene, including their location and body posture.

Speech recognition and parsing



- Attempts to identify and understand what a user has said to the robot using natural language.

State management



- Provides an abstract description of what the robot currently believes about the world and the interaction.

Planning and reasoning

action *ask-drink(?a : agent)*
preconds: $K(inTrans = ?a) \wedge \neg K(ordered(?a)) \wedge \neg K(otherAttnReq) \wedge \neg K(badASR(?a))$
effects: $add(K_f, ordered(?a))$, $add(K_v, request(?a))$

action *serve (?a : agent, ?d : drink)*
preconds: $K(inTrans = ?a) \wedge K(ordered(?a)) \wedge K_v(request(?a)) \wedge K(request(?a) = ?d) \wedge \neg K(otherAttnReq) \wedge \neg K(badASR(?a)) \wedge K(ackOrder(?a))$
effects: $add(K_f, served(?a))$

Plan: *greet(A), ask-drink(A), ack-order(A), serve-drink(A, request(A)), bye(A)*

A customer appears

ROBOT: Hello there!

ROBOT: What would you like?

CUSTOMER: A water.

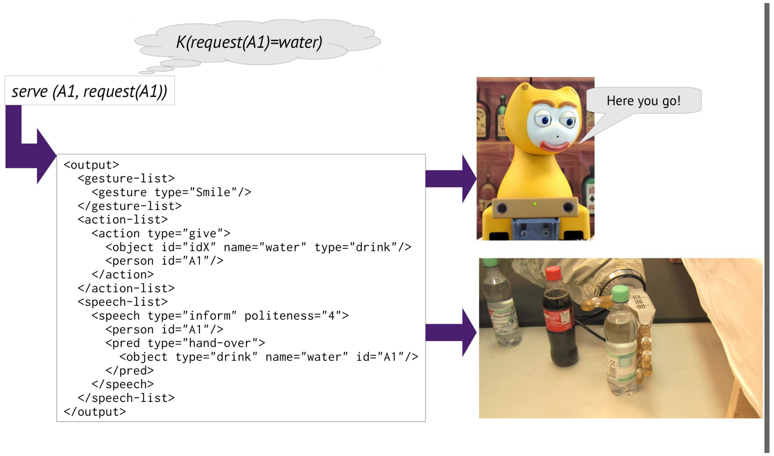
ROBOT: Okay!

ROBOT: *[Serves drink]* Here you go.

ROBOT: Good-bye!

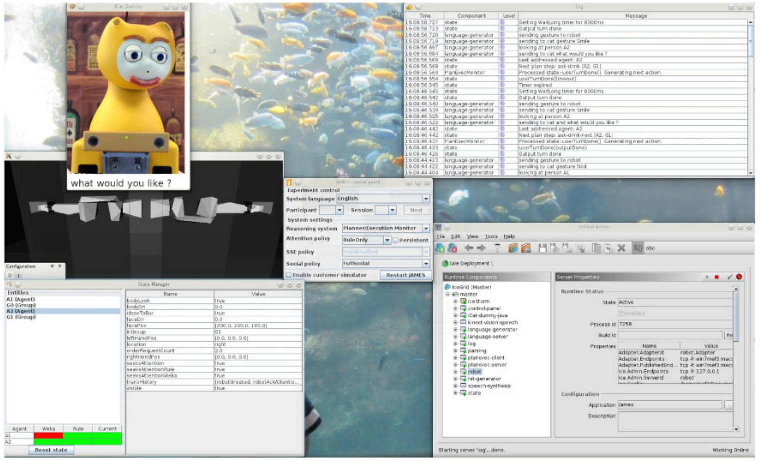
- Decides what action the robot should perform in the current context.

Multimodal output planning



- Processes a high-level action into lower-level actions that control the cat head (speech, head posture) and robot arms.

Software interface



- Monitors and controls all components of the system.

Target scenario: revisited

Two people, A and B, each individually approach a bartender.	
Bartender (to A):	**How can I help you?**
Person A:	A pint of cider, please.
Person C approaches the bartender and attracts his attention.	
Bartender (**nods at A**, then to C):	**Just a moment please**.
Bartender:	(**Serves A**)
Bartender (to B):	**What will you have?**
Person B:	A glass of red wine.
Bartender (**nods at B**):	(**Serves B**)
Bartender (to C):	**Thanks for waiting.**
	How can I help you?
Person C:	I'd like a pint of bitter.
Bartender (**nods at C**):	(**Serves C**)

- **Challenge:** customers should be able to interact with JAMES in a manner similar to how they would interact with a human bartender.

Robot bartender actions

- Available actions

`greet(?a)`	greet an agent ?a
`ask-drink(?a)`	ask agent ?a for a drink order
`ack-order(?a)`	acknowledge agent ?a's drink order
`serve(?a,?d)`	serve drink ?d to agent ?a
`wait(?a)`	tell agent ?a to wait
`ack-wait(?a)`	thank agent ?a for waiting
`not-understand(?a)`	alert agent ?a that it was not understood
`bye(?a)`	end an interaction with agent ?a
...	

- Inspired by data collected from customers in the human studies.
- The choice of which action the robot should perform is determined by using artificial intelligence automated planning techniques.

Automated planning actions

```
action ask-drink(?a : agent)
     preconds: K(inTrans = ?a) & !K(ordered(?a))
               !K(otherAttnReq) & !K(badASR(?a))
     effects:  add(Kf,ordered(?a)),
               add(Kv,request(?a))

action serve(?a : agent, ?d : drink)
     preconds: K(inTrans = ?a) & K(ordered(?a)) &
               Kv(request(?a)) & K(request(?a) = ?d)
               !K(otherAttnReq) & !K(badASR(?a)) &
               K(ackOrder(?a))
     effects:  add(Kf,served(?a))
```


A simple interaction

`greet(A1),`	Greet customer A1
`ask-drink(A1),`	Ask A1 for drink order
`ack-order(A1),`	Acknowledge A1's order
`serve(A1,request(A1)),`	Give the drink to A1
`bye(A1).`	End the transaction

- Simplest possible standard interaction in the single customer case.

Interacting with two customers

`wait(A2),`	Tell customer A2 to wait
`greet(A1),`	Greet customer A1
`ask-drink(A1),`	Ask A1 for drink order
`ack-order(A1),`	Acknowledge A1's order
`serve(a1,request(A1)),`	Give the drink to A1
`bye(A1),`	End A1's transaction
`ack-wait(A2),`	Thank A2 for waiting
`ask-drink(A2),`	Ask A2 for drink order
`ack-order(A2),`	Acknowledge A2's order
`serve(A2,request(A2)),`	Give the drink to A2
`bye(A2).`	End A2's transaction

- Each customer's interaction is similar to the single customer case with additional actions to manage the interaction order.

Recovering when things go wrong

- Low-confidence speech recognition / timeouts

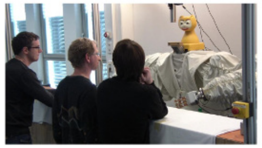
`ask-drink(A1)`	Ask A1 for drink order
???	A1 was not understood
[Replan]	[Replan]
`not-understand(A1)`	Alert A1 not understood
`ask-drink(A1)`	Ask A1 again for drink order
…	[Continue with old plan]

- Overanswering

`greet(A1)`	Greet A1
???	A1 says "I'd like a beer"
[Replan]	[Replan]
`serve(A1,request(A1))`	Serve A1 their drink
`bye(A1)`	End the transaction with A1
…	

- Interactions often lead to unexpected outcomes. A mechanism is needed to recover from such situations. Replanning is used in JAMES.

A more complex interaction



Three customers:
A1 and A2 in group G1
A3 is alone (singleton group G2)
Bartender serves members of G1 in sequence, then deals with G2.

Other social behaviour:
- First-come/first-served ordering
- All orders are acknowledged immediately
- If a new customer arrives while the bartender is occupied, it nods at them and serves them later

Social behaviour is based on the observation of bartenders in real bars (Huth et al., 2012); see Foster et al. (2013) for details on the planning domain.

wait(A3, G1)	Tell G2 to wait (with a nod)
greet(A1, G1)	Greet group G1
ask-drink(A1, G1)	Ask A1 for drink order
ack-order(A1, G1)	Acknowledge A1's order
ask-drink(A2, G1)	Ask A2 for drink order
ack-order(A2, G1)	Acknowledge A2's order
serve(A1, request(A1), G1)	Give the drink to A1
serve(A2, request(A2), G2)	Give the drink to A2
bye(A2, G1)	End G1's transaction
ack-wait(A3, G2)	Acknowledge G2's wait
ask-drink(A3, G2)	Ask A3 for drink order
ack-order(A3, G2)	Acknowledge A3's order
serve(A3, request(A3), G3)	Give the drink to A3
bye(A3, G2)	End G2's transaction

JAMES interaction video

Image/video: fortiss GmbH

`http://youtu.be/8k7Pd-CbbhE`
`http://james-project.eu/`

Experimental results

- Action selection time is typically quite short (<0.1s), which doesn't impact the system's reaction time.
  - Less than 2s is usually okay.
  - Robot motions are slow.
  - Frequent replanning to recover from problems in the interaction.
- **Study 1**: system tested with 2 customers at a time in a drink ordering scenario (31 participants $\times$ 3 interactions each), 95% success rate on delivering correct drinks.
- **Study 2**: more complex scenario (3 customers at a time, 40 participants), group detector, task only vs. social domain, 87% success rate.

Measure	Mean	Median	Min	Max
Drinks served	2.63	3	1	3
Low ASR turns	6.25	6	1	14
Timeouts	1.05	0	0	12
Order requests	7.10	6	1	22
Response time	14.7	12.9	1.8	52.8
Time to first drink	51.0	46.1	33.0	110.1
Total time (s)	111.6	109.0	41.7	214.5
Total system turns	13.2	13	6	35
Group orders	0.75	1	0	2
Acknowledgements	2.55	2	1	7

Measure	Task-only (sd)	Full social (sd)
Time to first drink	55.6 (17.0)	46.3 (15.5)
Total system turns	11.0 (4.8)	15.5 (5.8)

Category	Pre (sd)	Post (sd)	Change (sd)
Anthropomorphism	**2.74 (0.95)**	**2.02 (0.64)**	**-0.73 (0.90)**
German	2.34 (0.78)	1.90 (0.60)	-0.43 (0.86)
English	3.51 (0.76)	2.24 (0.65)	-1.27 (0.71)
Animacy	**3.04 (0.90)**	**2.30 (0.70)**	**-0.73 (0.96)**
German	2.83 (0.97)	2.25 (0.66)	-0.59 (1.01)
English	3.43 (0.59)	2.42 (0.78)	-1.01 (0.82)
Likeability	**4.43 (1.01)**	**3.70 (1.17)**	**-0.74 (1.34)**
German	4.27 (1.03)	4.05 (0.96)	-0.22 (1.01)
English	4.74 (0.92)	3.04 (1.27)	-1.70 (1.38)
Perc. Intelligence	**3.77 (0.80)**	**2.87 (0.83)**	**-0.90 (0.86)**
German	3.52 (0.78)	2.92 (0.90)	-0.60 (0.81)
English	4.22 (0.65)	2.76 (0.70)	-1.46 (0.64)
Perc. Safety	**4.15 (1.10)**	**3.97 (1.12)**	**-0.18 (1.20)**
German	4.12 (1.11)	4.17 (1.08)	+0.05 (0.82)
English	4.22 (1.12)	3.62 (1.14)	-0.60 (1.65)

Design, human-robot interaction, and JAMES

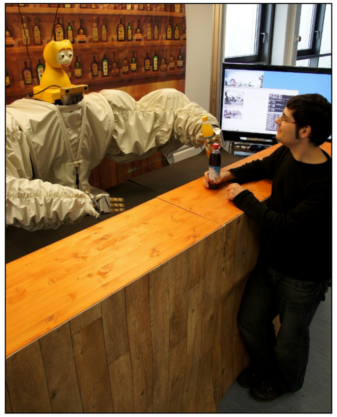


Image: fortiss GmbH

- Design decisions at every stage of the project and at all levels of software development helped shape the resulting JAMES system.
- Design choices were often motivated by decisions about (practical) tradeoffs.
- Social interaction places additional requirements on the design of a robot system: achieving a task goal isn't always enough.
- Expect surprises when working in the real world.

The other JAMES robot

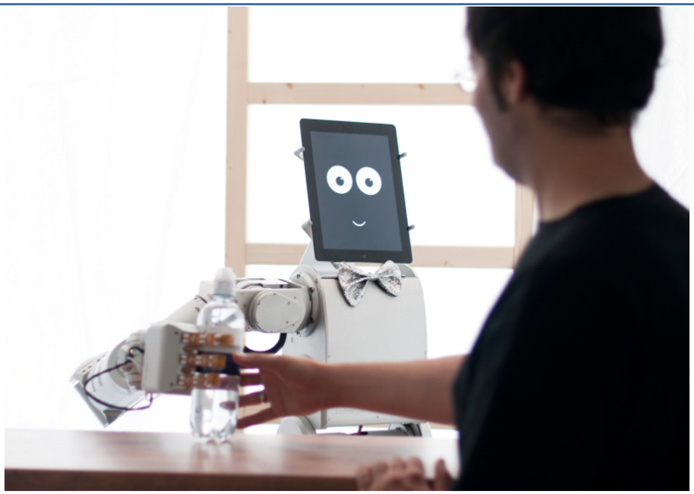


Image: fortiss GmbH

- Does embodiment make a difference?

References

Baltzakis, H., Pateraki, M., and Trahanias, P. (2012). Visual tracking of hands, faces and facial features of multiple persons. *Machine Vision and Applications*, pages 1–17.

Foster, M. E., Gaschler, A., Giuliani, M., Isard, A., Pateraki, M., and Petrick, R. P. A. (2012). Two people walk into a bar: Dynamic multi-party social interaction with a robot agent. In *Proceedings of the International Conference on Multimodal Interaction (ICMI 2012)*, pages 3–10.

Giuliani, M., Petrick, R. P. A., Foster, M. E., Gaschler, A., Isard, A., Pateraki, M., and Sigalas, M. (2013). Comparing task-based and socially intelligent behaviour in a robot bartender. In *Proceedings of the International Conference on Multimodal Interaction (ICMI 2013)*.

Loth, S., Huth, K., and De Ruiter, J. P. (2013). Automatic detection of service initiation signals used in bars. *Frontiers in Psychology*, 4(557).

Nau, D., Ghallab, M., and Traverso, P. (2004). *Automated Planning: Theory and Practice*. Morgan Kaufmann Publishers.

Petrick, R. P. A. and Foster, M. E. (2013). Planning for social interaction in a robot bartender domain. In *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS 2013), Special Track on Novel Applications*, pages 389–397.

Petrick, R. P. A., Foster, M. E., and Isard, A. (2012). Social state recognition and knowledge-level planning for human-robot interaction in a bartender domain. In *AAAI 2012 Workshop on Grounding Language for Physical Systems*.

For more information on the **JAMES Project** visit `http://james-project.eu/`.