



Personalised Self-Explanation by Robots: The Role of Goals versus Beliefs in Robot-Action Explanation for Children and Adults

Frank Kaptein, Joost Broekens, Koen Hindriks, and Mark Neerincx

Presenter: Fateme Zohairi



UNIVERSITÉ DU
LUXEMBOURG

December, 2020



Outline

- ▶ Introduction
- ▶ Motivation and Related Works
- ▶ Goal Hierarchy Trees
- ▶ User Study
- ▶ Result
- ▶ Discussion
- ▶ Conclusion



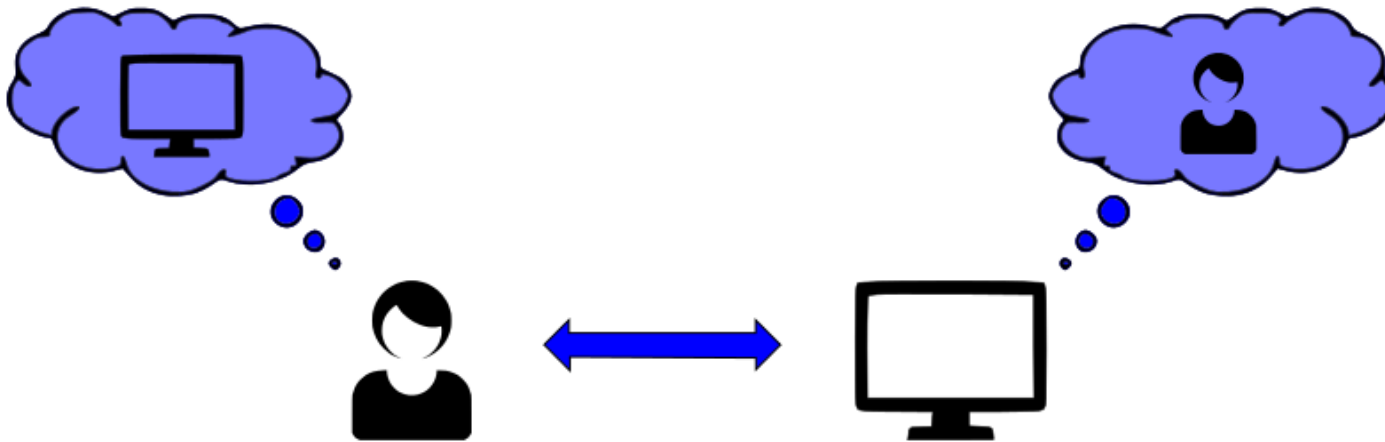
Introduction

- ▶ Explainable Artificial Intelligence (XAI)
- ▶ Folk Psychology
- ▶ BDI-Based (Belief, Desire, Intention) Agents
- ▶ PAL (a Personal Assistant for a Healthy Lifestyle) Project



Introduction

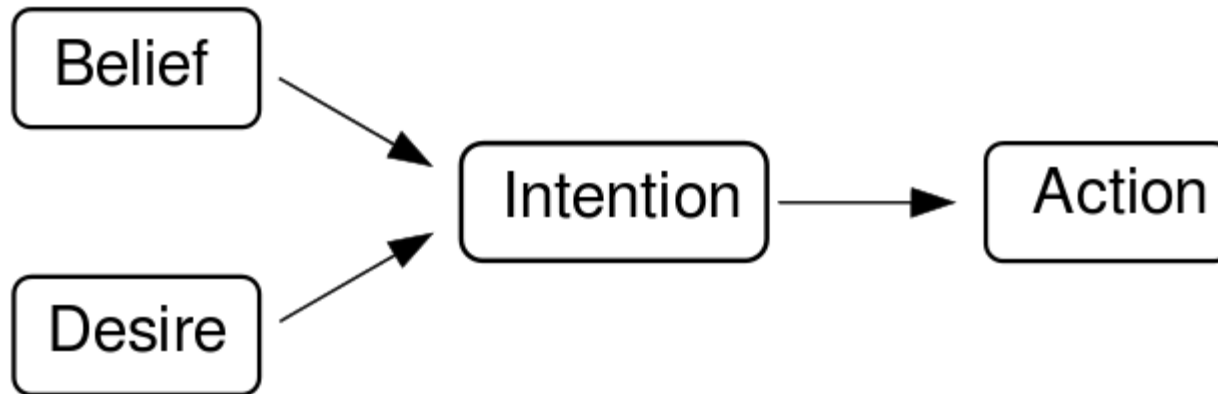
- Explainable Artificial Intelligence (XAI)





Introduction

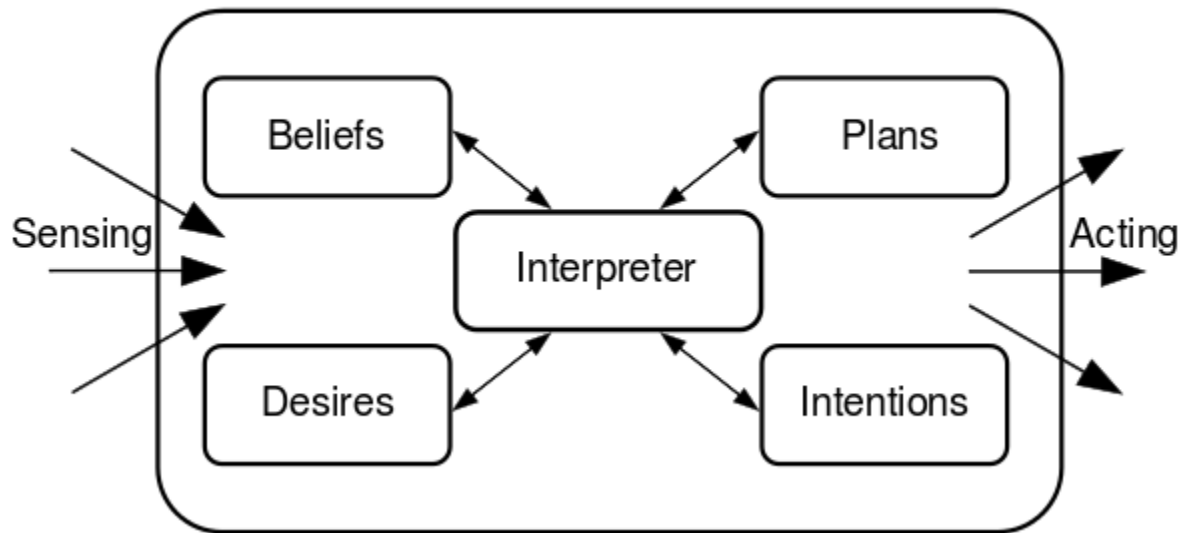
► Folk Psychology





Introduction

- BDI-based (Belief, Desire, Intention) Agents



Introduction

- ▶ PAL (a Personal Assistant for a Healthy Lifestyle) Project



Motivation and Related Works

- ▶ Goal-based and Belief-based Explanations
- ▶ Hypothesis





Motivation and Related Works

► Goal-based and Belief-based Explanations

For what purpose?

To what end?

I want to ...

My goal is ...



Why a certain action over another?

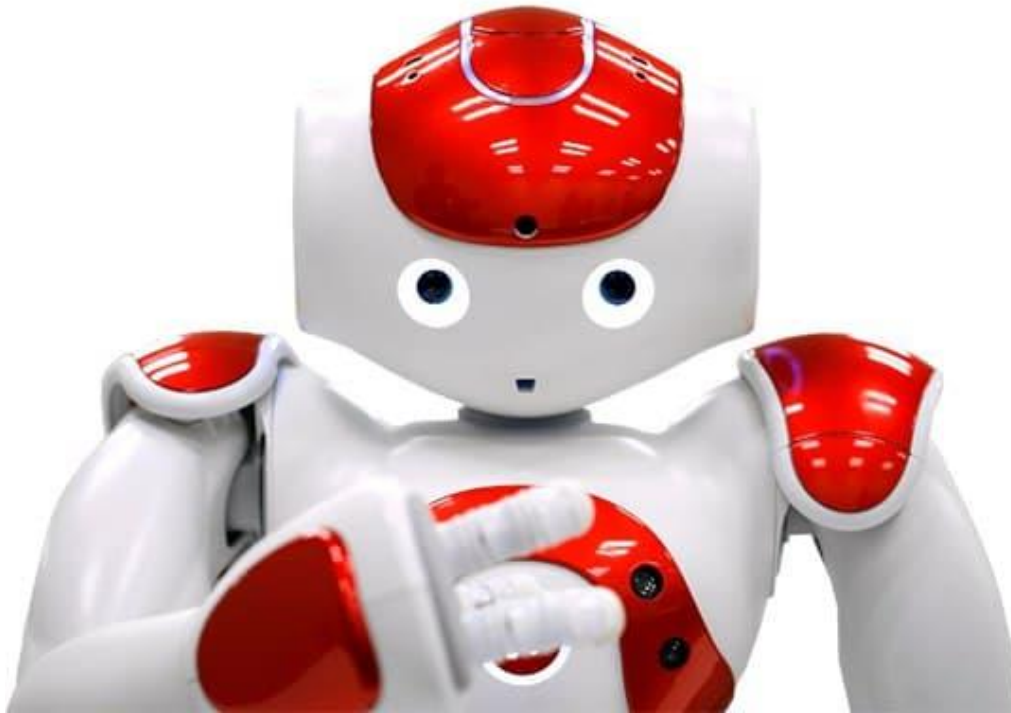
I think ...

I believe ...

Motivation and Related Works

► Hypothesis:

Adults have a stronger preference than children for goal-based over belief-based explanations.





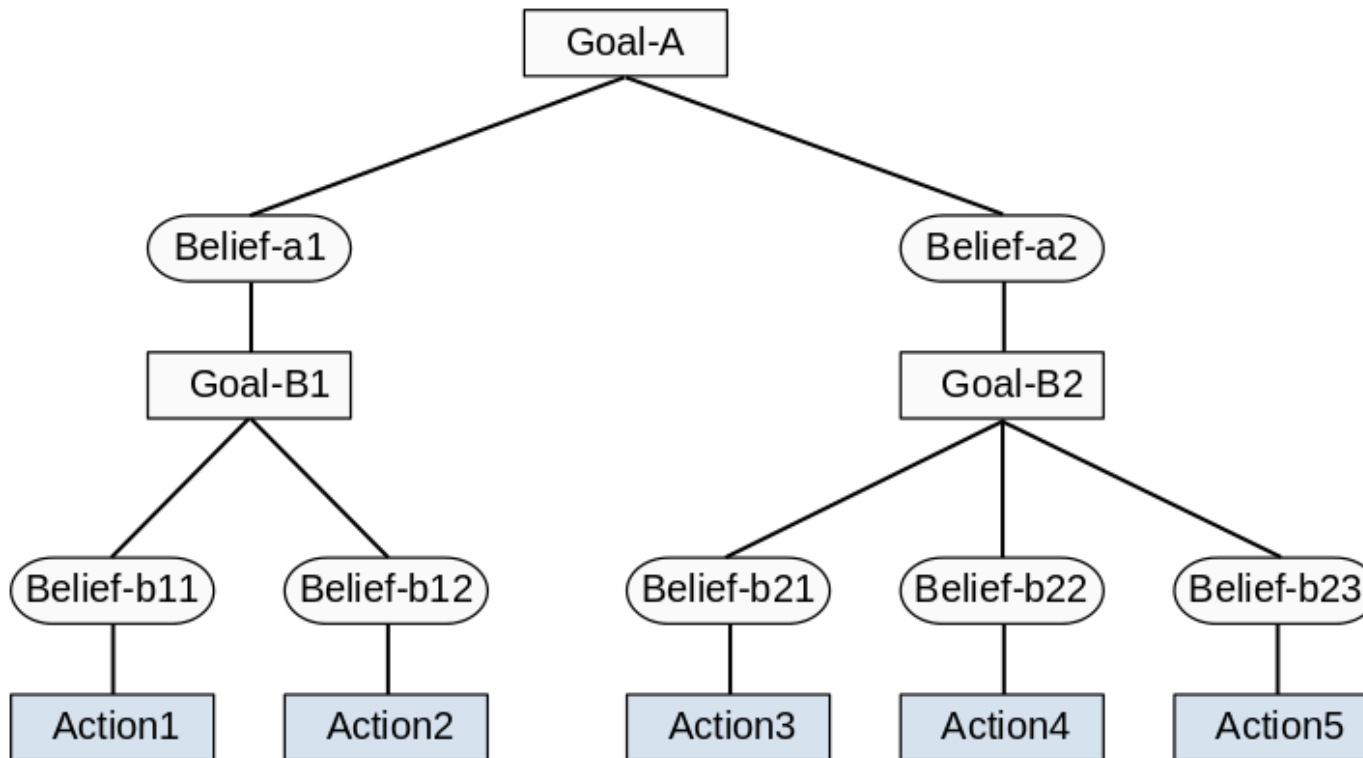
Goal Hierarchy Trees

- ▶ The Structure of a Goal Hierarchy Tree
- ▶ Goal-based and Belief-based Agent-action Explanations



Goal Hierarchy Trees

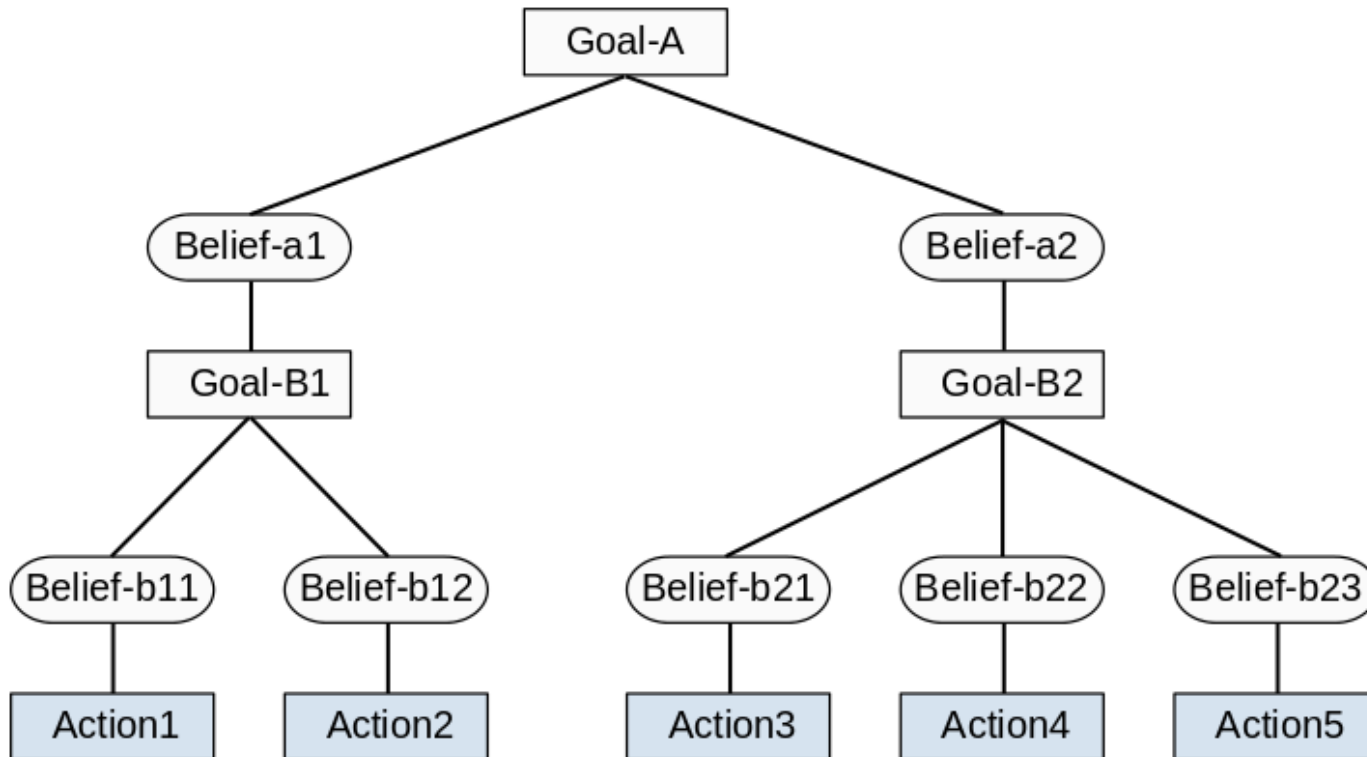
► The Structure of a Goal Hierarchy Tree





Goal Hierarchy Trees

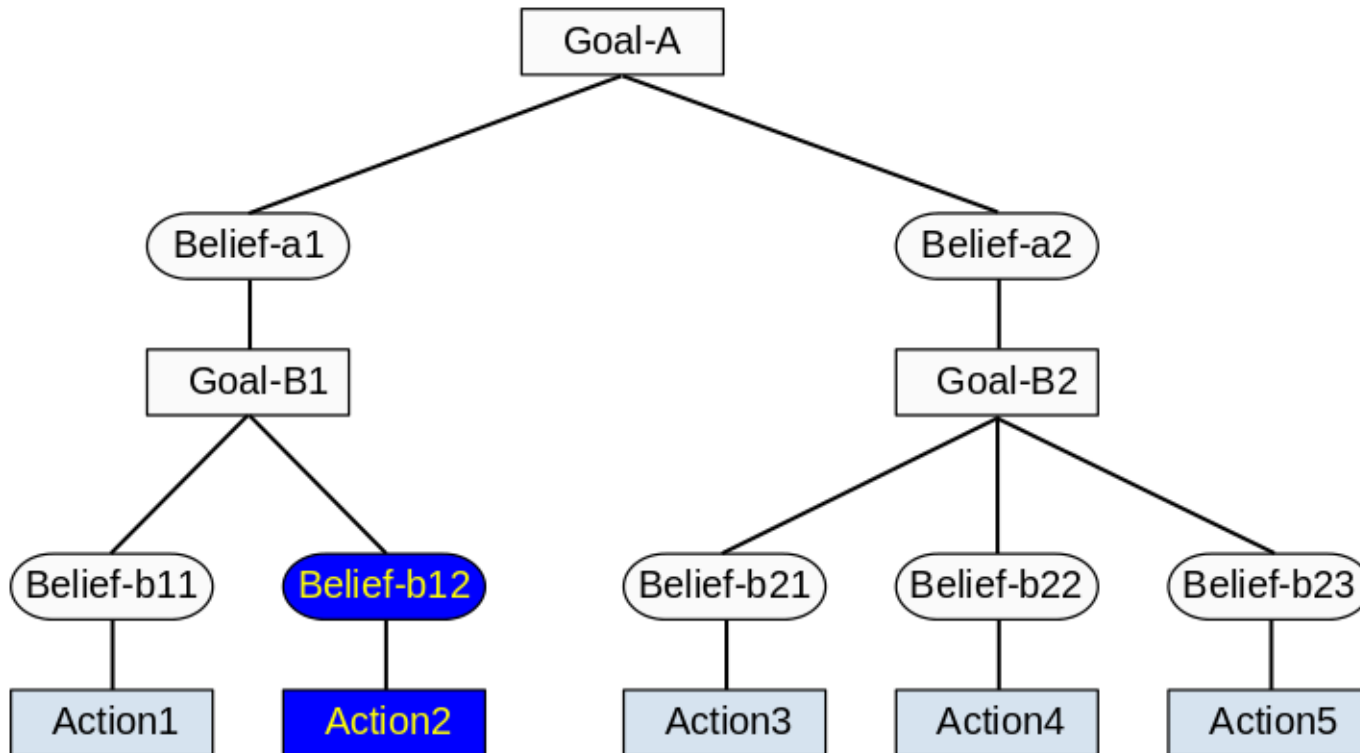
► Goal-based and Belief-based Agent-action Explanations





Goal Hierarchy Trees

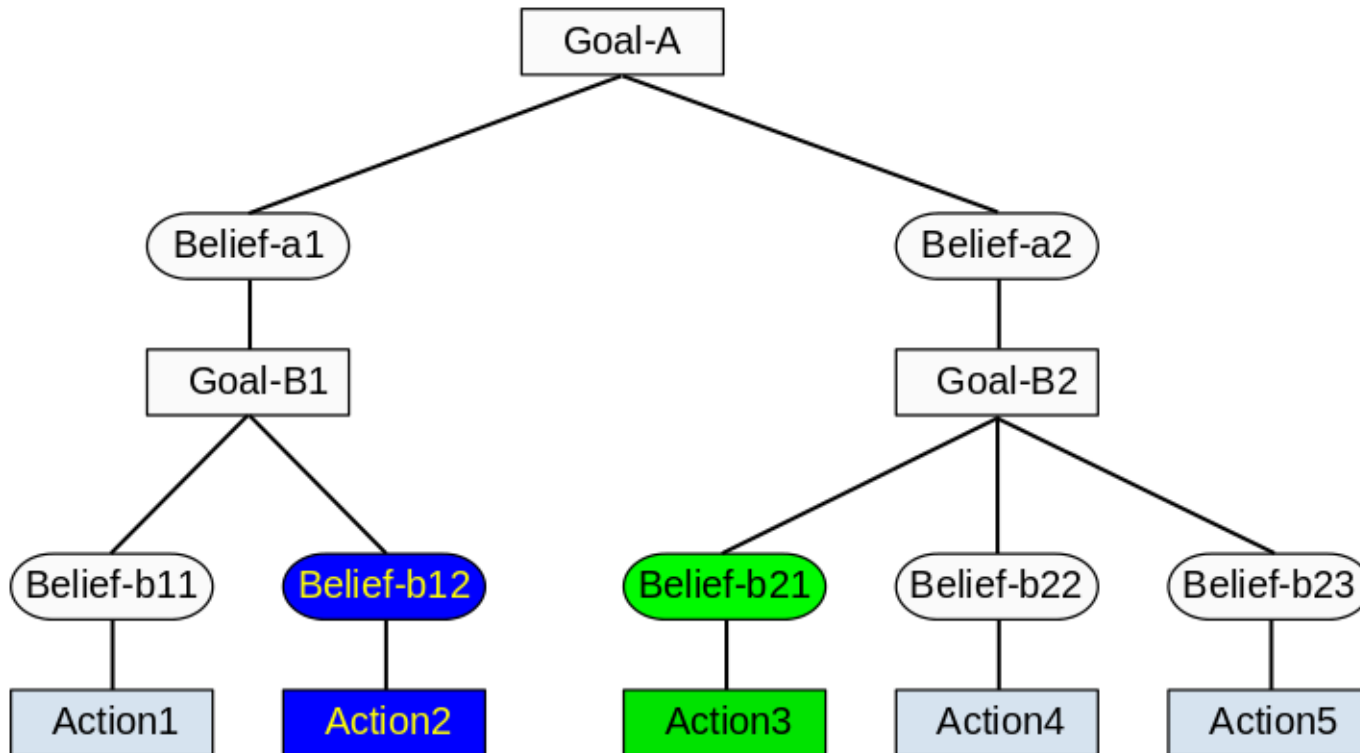
► Belief-based Explanation Algorithm





Goal Hierarchy Trees

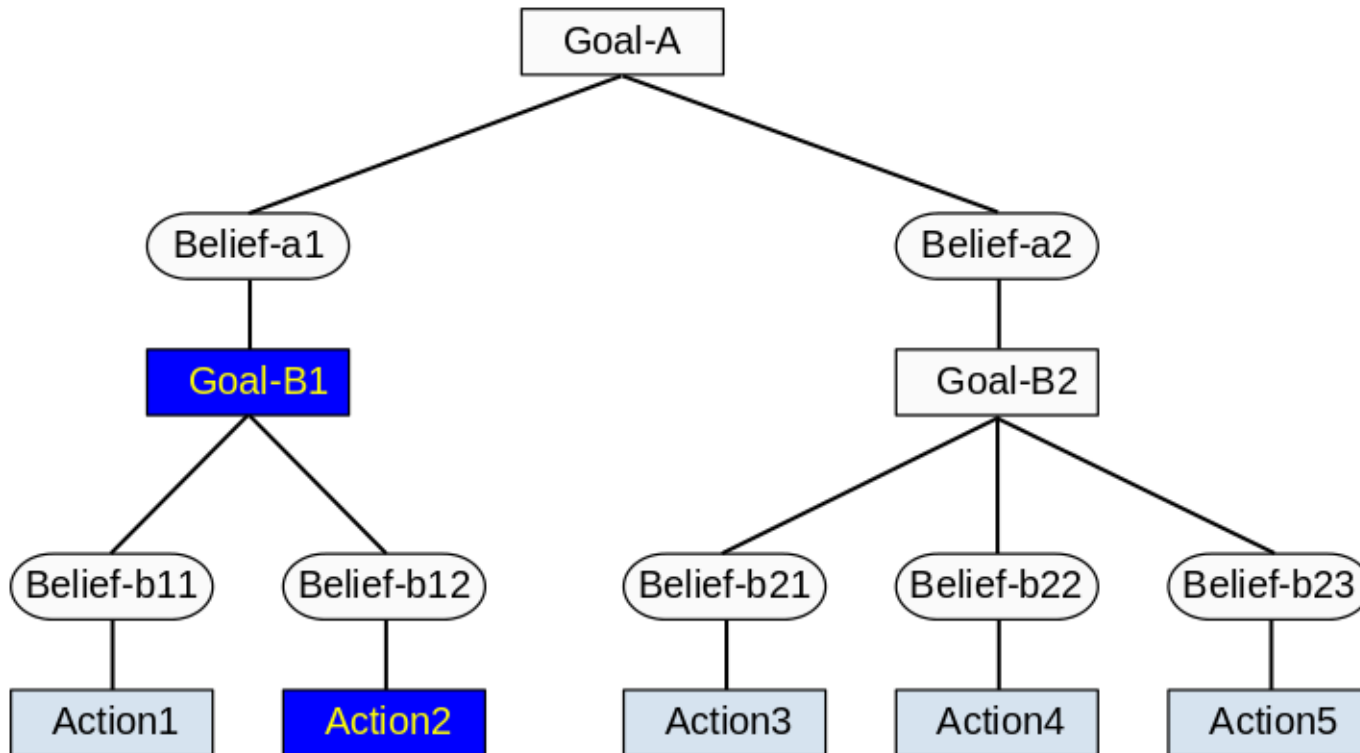
► Belief-based Explanation Algorithm





Goal Hierarchy Trees

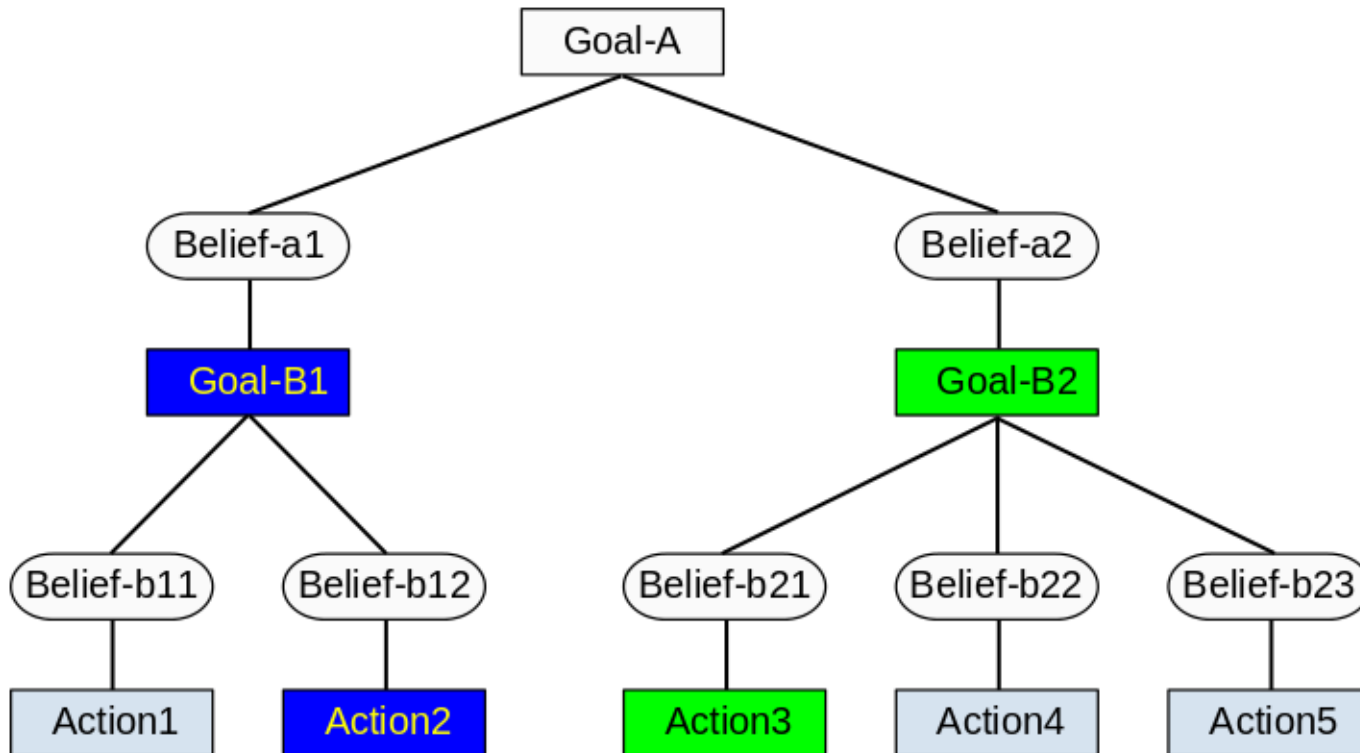
► Goal-based Explanation Algorithm





Goal Hierarchy Trees

► Goal-based Explanation Algorithm





User Study

- ▶ Participants
- ▶ Designing a Goal Hierarchy Tree
- ▶ Set-up and Materials
- ▶ Variable and Design
- ▶ Procedure



User Study

► Participants

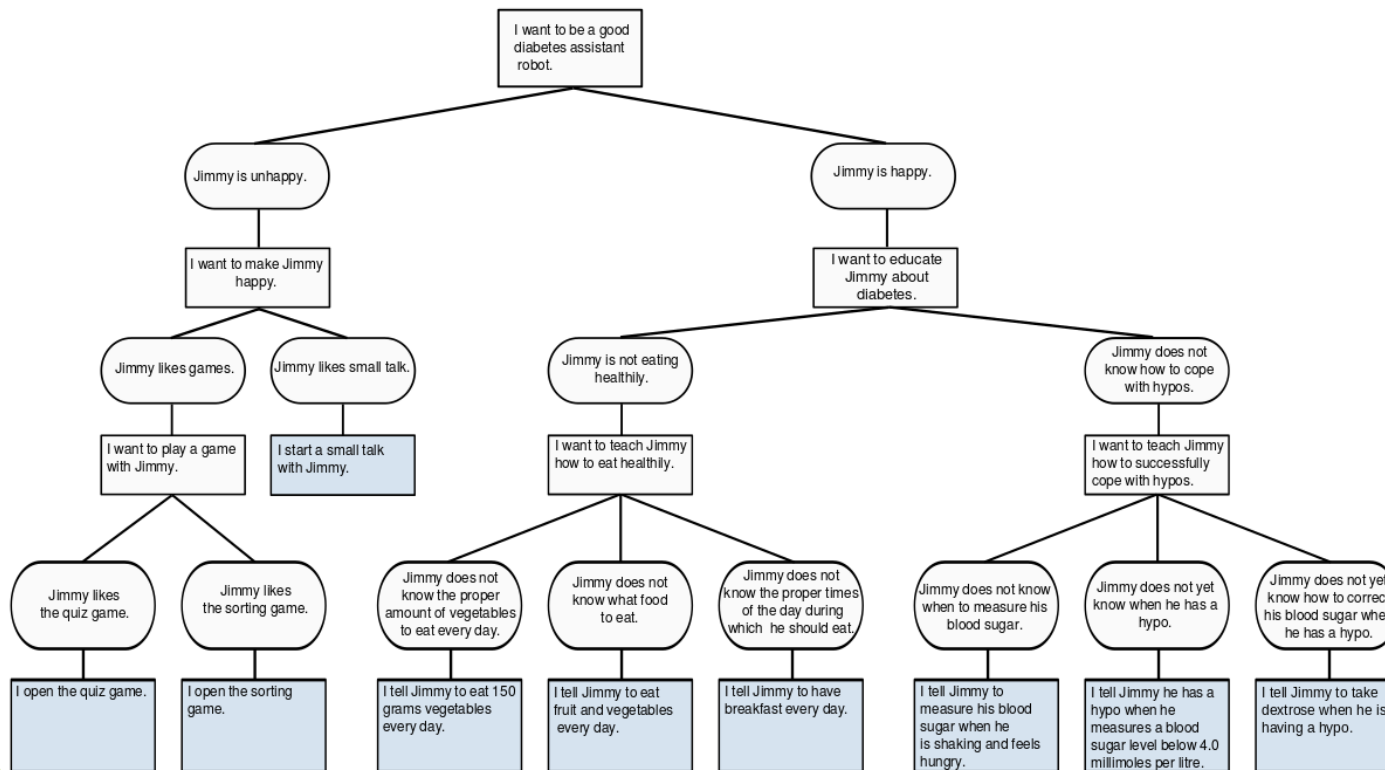
$$7 \times \text{[Female Icon]} + 12 \times \text{[Male Icon]} = 19 \text{ Children (aged 8-11)}$$

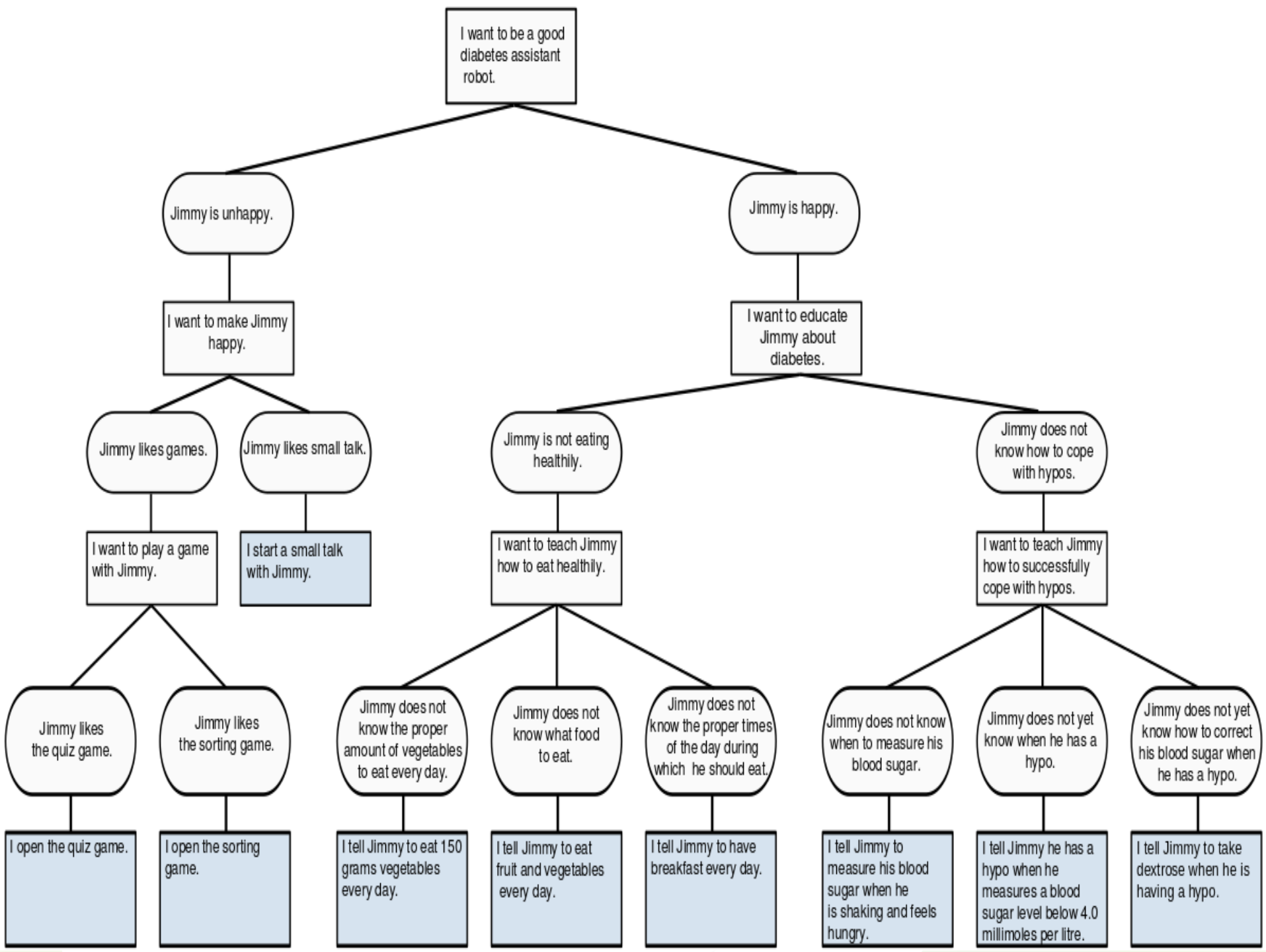
$$11 \times \text{[Female Icon]} + 8 \times \text{[Male Icon]} = 19 \text{ Adults (aged 35-48)}$$



User Study

► Designing a Goal Hierarchy Tree







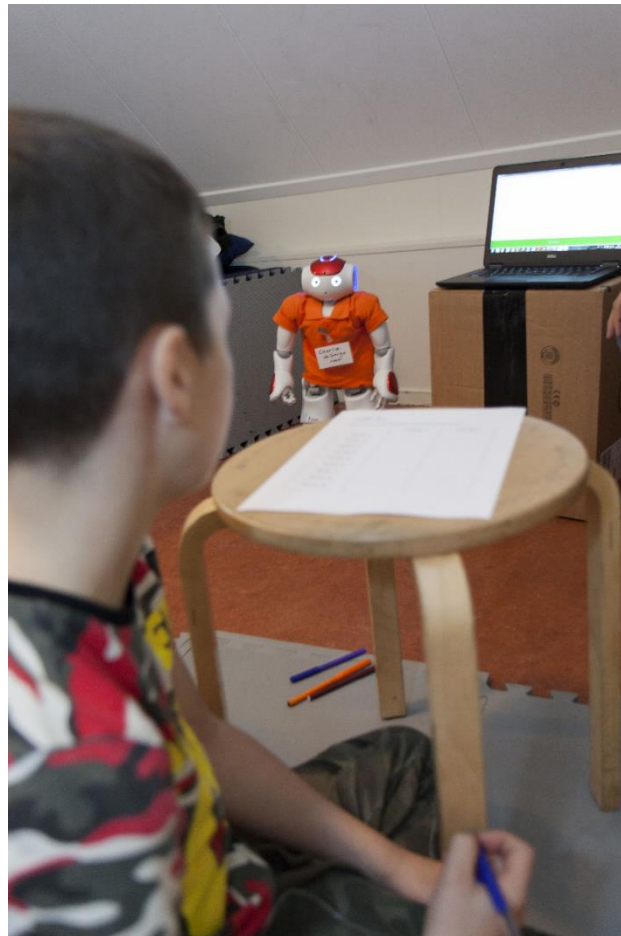
User Study

► Set-up and Materials

Action: 'I tell Jimmy to take dextrose when he is having a hypo.'

Explanation 1: 'Jimmy does not yet know how to correct his blood-sugar when he has a hypo.'

Explanation 2: 'I want to teach Jimmy how to successfully cope with hypos.'





User Study

► Variable and Design

Table I: Distribution of children and adults over the 4 conditions

	Random Seed of Scenarios			
	Normal Order of Explanations		Reversed Order of Explanations	
	Normal Order of Scenarios	Reversed Order of Scenarios	Normal Order of Scenarios	Reversed Order of Scenarios
Children	6	5	4	4
Adults	5	4	5	5



User Study

► How to counterbalance the conditions

Normal Order of Explanations	Senario 8	Senario 2	Senario 4	Senario 7	Senario 3	Senario 9	Senario 6	Senario 1	Senario 5
	Action 8	Action 2	Action 4	Action 7	Action 3	Action 9	Action 6	Action 1	Action 5
	Goal 8	Belief 2	Belief 4	Goal 7	Belief 3	Goal 9	Belief 6	Goal 1	Goal 5
	Belief 8	Goal 2	Goal 4	Belief 7	Goal 3	Belief 9	Goal 6	Belief 1	Belief 5
	Normal Order of Scenarios								
Normal Order of Explanations	Senario 5	Senario 1	Senario 6	Senario 9	Senario 3	Senario 7	Senario 4	Senario 2	Senario 8
	Action 5	Action 1	Action 6	Action 9	Action 3	Action 7	Action 4	Action 2	Action 8
	Goal 5	Goal 1	Belief 6	Goal 9	Belief 3	Goal 7	Belief 4	Belief 2	Goal 8
	Belief 5	Belief 1	Goal 6	Belief 9	Goal 3	Belief 7	Goal 4	Goal 2	Belief 8
	Reversed Order of Scenarios								
Reversed Order of Explanations	Senario 8	Senario 2	Senario 4	Senario 7	Senario 3	Senario 9	Senario 6	Senario 1	Senario 5
	Action 8	Action 2	Action 4	Action 7	Action 3	Action 9	Action 6	Action 1	Action 5
	Belief 8	Goal 2	Goal 4	Belief 7	Goal 3	Belief 9	Goal 6	Belief 1	Belief 5
	Goal 8	Belief 2	Belief 4	Goal 7	Belief 3	Goal 9	Belief 6	Goal 1	Goal 5
	Normal Order of Scenarios								
Reversed Order of Explanations	Senario 5	Senario 1	Senario 6	Senario 9	Senario 3	Senario 7	Senario 4	Senario 2	Senario 8
	Action 5	Action 1	Action 6	Action 9	Action 3	Action 7	Action 4	Action 2	Action 8
	Belief 5	Belief 1	Goal 6	Belief 9	Goal 3	Belief 7	Goal 4	Goal 2	Belief 8
	Goal 5	Goal 1	Belief 6	Goal 9	Belief 3	Goal 7	Belief 4	Belief 2	Goal 8
	Reversed Order of Scenarios								



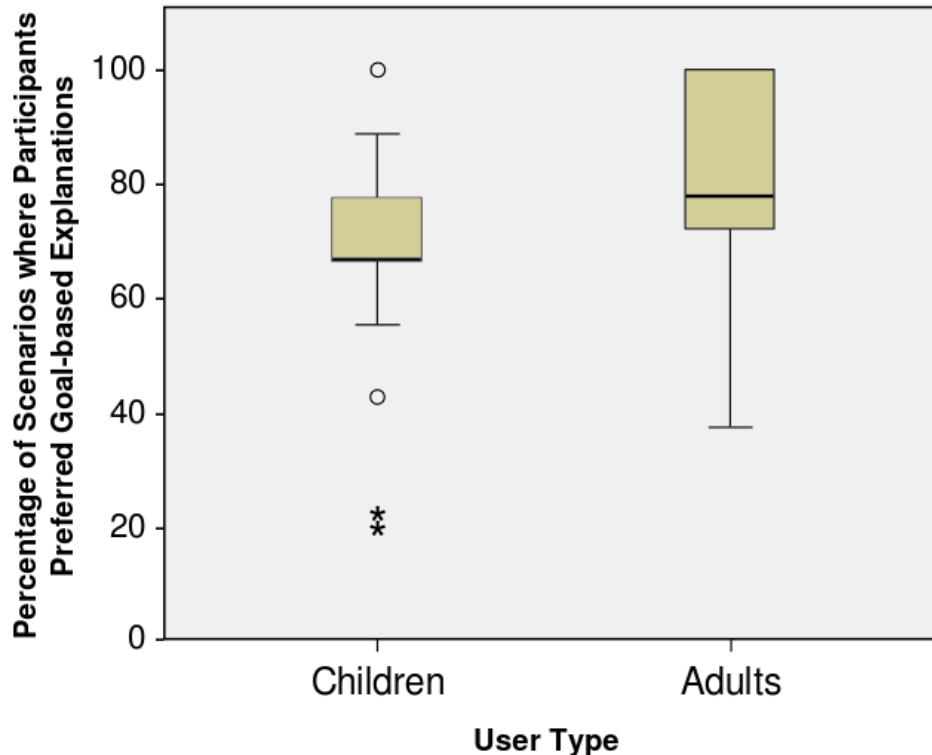
User Study

► Procedure





Results



In one- sample Wilcoxon signed rank test for children:

med = 0.667,
95% CI = [0.667,0.778],
p = .007

and for adults:

med = 0.778,
95% CI = [0.667,1.0],
p < .000).

In Mann-Whitney test

the preference towards goal-based explanations, rather than belief- based explanations, was greater for adults (med = 0.778) than for children (med = 0.667),

U = 112.5, p = .042, r = .33.

Adults prefer goal-based explanations significantly more than children.

Adults have a significantly higher preference for goal-based action explanations (Median = 0.778) than children (Median = 0.667).



Discussion

- ▶ Significant preference for goal-based explanations in both user groups
- ▶ Significant preference for goal-based explanations in adults more than their children
- ▶ Limitation of the generalisability of the results
 - ▶ Future work includes replicating of this study with more divers scenarios, contexts, and users.



Conclusion

- ▶ Adults have a significantly higher preference for goal-based explanations than children.
- ▶ Children and adults perceived the self-explanations of intelligent agents differently.
- ▶ This work is a necessary step towards providing personalised explanations in human-robot and human-agent interaction.



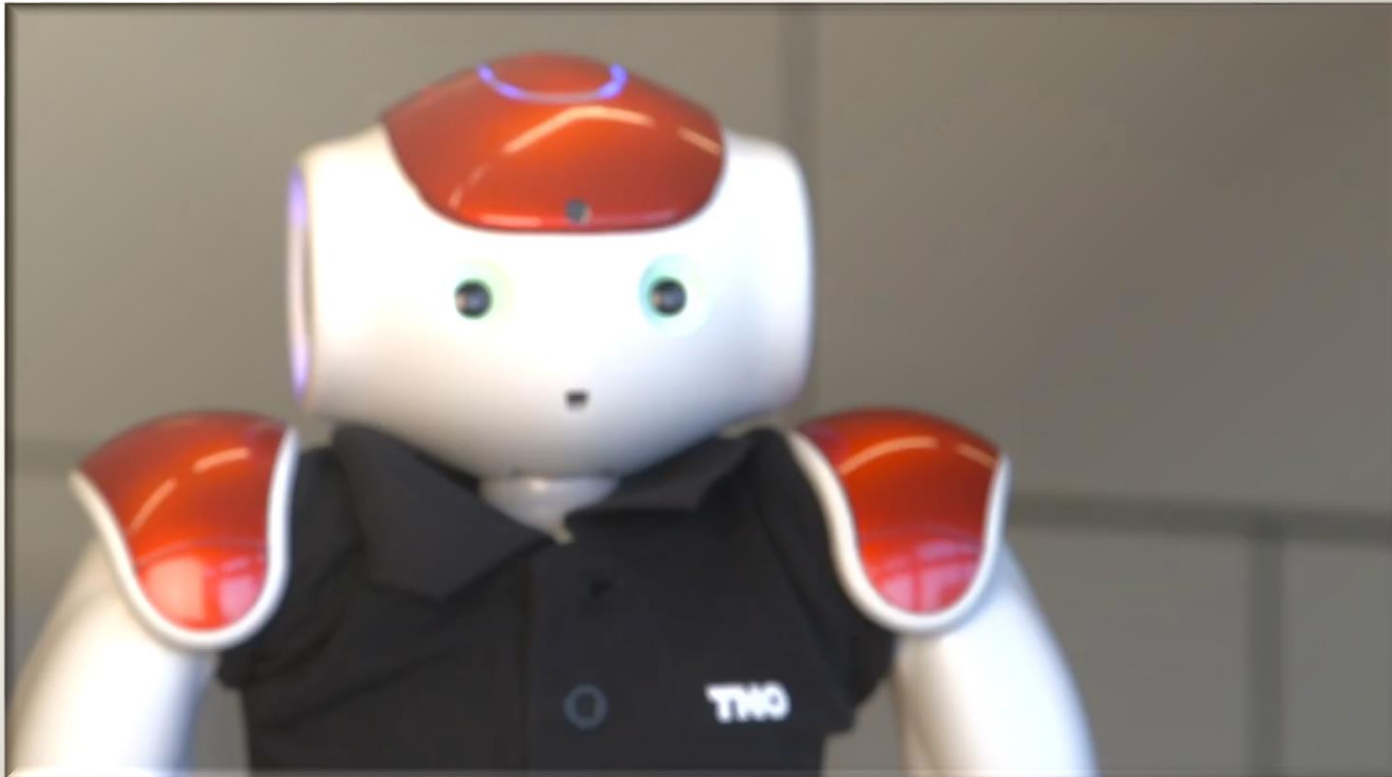
Citations

► Cited by 27 (10 in 2020)

- Anjomshoe, S., Najjar, A., Calvaresi, D., & Främling, K. (2019). Explainable agents and robots: Results from a systematic literature review. In 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019 (pp. 1078-1088). International Foundation for Autonomous Agents and Multiagent Systems (Cited by 88).
- De Visser, E. J., Peeters, M. M., Jung, M. F., Kohn, S., Shaw, T. H., Pak, R., & Neerincx, M. A. (2020). Towards a theory of longitudinal trust calibration in human–robot teams. *International journal of social robotics*, 12(2), 459-478 (Cited by 19).
- Neerincx, M. A., van der Waa, J., Kaptein, F., & van Diggelen, J. (2018, July). Using perceptual and cognitive explanations for enhanced human-agent team performance. In *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 204-214). Springer, Cham (Cited by 25).
- Clabaugh, C., & Matarić, M. (2019). Escaping oz: Autonomy in socially assistive robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 2, 33-61 (Cited by 14).
- ...



TNO's Robot, Charlie



Netherlands Organisation for Applied Scientific Research (TNO)

<https://www.tno.nl/en/tno-insights/articles/charlie-the-ultimate-buddy-for-diabetic-children/>

THANK YOU FOR YOUR ATTENTION

