A Bayesian Approach to Context-based Recognition of Human Intention for Context-Adaptive Robot Assistance in Space Missions

Adrian Lubitz
alubitz@uni-bremen.de
University of Bremen
Bremen, Germany

Octavio Arriaga
arriagac@uni-bremen.de
University of Bremen
Bremen, Germany

Teena Hassan
thassan@uni-bremen.de
University of Bremen
Bremen, Germany

Nina Hoyer
nihoyer@uni-bremen.de
University of Bremen
Bremen, Germany

Elsa Andrea Kirchner
DFKI, Robotics Innovation Center
Bremen, Germany
elsa.kirchner@dfki.de
University of Duisburg-Essen
Duisburg, Germany
elsa.kirchner@uni-due.de

ABSTRACT

Long-duration space exploration (LDSE) missions with a high demand for intelligent and adaptable robotic assistance are expected to start in the next decades. Robot assistants deployed on these missions should be able to interpret the intentions of astronaut(s) and provide assistance accordingly, under continuously changing interaction context. This necessitates the development of robust and context-adaptive models for inferring human intentions. However, current state-of-the-art artificial intelligence (AI) algorithms rely on expensive data-driven machine learning (ML) models that provide inferences assuming a set of known contexts. In this paper, we present a context-adaptive hybrid architecture for intelligent and adaptable Human Computer Interaction (HCI). This architecture uses state-of-the-art machine learning models for the recognition of specific contexts within the robot's environment, and a two-layer *Bayesian network* to infer human intentions based on the current context. The Bayesian network enables the integration of human expert knowledge about the mission goals within the inference process and can also deal with missing or uncertain information about context. Thus, the proposed framework increases robustness and reduces the need for large and complex (high dimensional) human-robot interaction data, which is unavailable in space applications. Context-adaptive interpretation of human intention would promote an intuitive interaction with robots, which in turn could reduce the overall stress in astronauts while interacting with robot assistants in isolated, confined or extreme (ICE) conditions.

SpaceCHI 2.0, May 01, 2022, New Orleans, LA

CCS CONCEPTS

• **Computer systems organization** → *Data flow architectures*; • **Human-centered computing** → **HCI theory, concepts and models**; *Collaborative interaction*;

KEYWORDS

Bayesian networks, intention recognition, context, HCI, HRI, HMI

ACM Reference Format:
Adrian Lubitz, Octavio Arriaga, Teena Hassan, Nina Hoyer, and Elsa Andrea Kirchner. 2022. A Bayesian Approach to Context-based Recognition of Human Intention for Context-Adaptive Robot Assistance in Space Missions. In *SpaceCHI 2.0 workshop at ACM SIGCHI 2022*. ACM, New York, NY, USA, 5 pages.

1 INTRODUCTION

Exploring and inhabiting other planets is a topic which is largely present in science fiction novels and TV shows and fascinates people all over the globe. In recent years, this goal is shifting from science fiction to a scientifically attainable future vision through national space programs and private companies that are pushing the boundaries of what is possible [12] [5]. The advent of reusable rockets had a tremendous effect on the driving price factor for space missions [17]. Furthermore, the recent success in multi-robot cooperation for space missions [20] and the deployment of robotic systems like NASA's Ingenuity drone [24] showcase the enormous advancements in space robotics. All these are preparing the stage for manned missions for exploring the deep space in the coming decades. However, the human body has been shaped by evolution over thousands of years to adapt perfectly to the conditions on earth. New conditions in space and on other planets are challenging for humans and this creates a huge need for robotic assistants to support astronauts on long-duration space exploration (LDSE) missions or in the process of building space stations on planets or in orbits. Intelligent and adaptable robot assistants are crucial for success in such missions. Robots will be assisting humans in space not only with construction and maintenance, but also in health care related tasks. They will also be providing social interaction in isolated, confined, and extreme (ICE) conditions to help in dealing with cognitive health risks [13]. This necessitates the development of robust and context-adaptive models for inferring human intentions and providing appropriate assistance.

Current state-of-the-art artificial intelligence (AI) algorithms rely on expensive data-driven machine learning (ML) models that provide inferences assuming a set of known contexts. In space applications, these data-driven approaches face the challenge that sufficient data are not available for training such models. Data recording is a tedious, costly and time consuming task. Existing work on Bayesian context-based intention recognition [9, 10, 14–16, 21] introduce complex network structures which make quick adaptation and re-configuration hard and impose a need for the human expert to have knowledge about the underlying network structure. The proposed cognitive architectures for space missions in [2, 7] rely on very complex algebras and therefore introduce a very large barrier to start using the system. Data-driven intention recognition systems [11, 23] have the drawback of relying on large and complex (high dimensional) human-robot interaction data. In sum, all approaches mostly lack the flexibility needed to quickly create or adapt a mission scenario. Furthermore, non-Bayesian approaches miss the capability of uncertainty estimation.

In this paper, we present a context-adaptive hybrid architecture for intelligent and adaptable Human-Computer Interaction (HCI). This architecture uses state-of-the-art machine learning models for the recognition of specific contexts within the robot's environment, and a two-layer *Bayesian network* to infer human intentions based on the predicted context. This approach allows to use data-driven models where available and enables human experts (we will continue to use the term human experts for experts with knowledge about the mission and the environment but no expert knowledge about Bayes theorem and the underlying two-layer Bayesian network) to integrate knowledge about the mission goals. The context-based two-layer Bayesian network structure allows us to apply some useful assumptions to dramatically reduce the effort for creating and adapting these networks by human experts. Furthermore, the Bayesian approach to context-based intention recognition is able to efficiently handle multimodal data as well as missing or uncertain evidence in the inference process. Essentially, the provided framework reduces the need for large and complex (high dimensional) human-robot interaction data, which is unavailable in space applications. This in turn reduces the complexity in creating and adjusting models for context-based Bayesian intention recognition while providing robust results. We hypothesize that the approach presented in this paper will reduce the stress in astronauts in ICE conditions by enabling robust, intuitive and naturalistic HCI.

2 CHALLENGES IN SPACE

Making HCI effective, efficient, and natural is crucial to the success of future space exploration missions[4]. In space missions, we face the problem that, for most parts, it is impossible to record data beforehand. Even if it is possible, recording such data is very costly due to the limited time and resources in space missions. This lack of data makes it very hard to apply data-driven approaches like modern deep learning models. Furthermore, in space, the computational

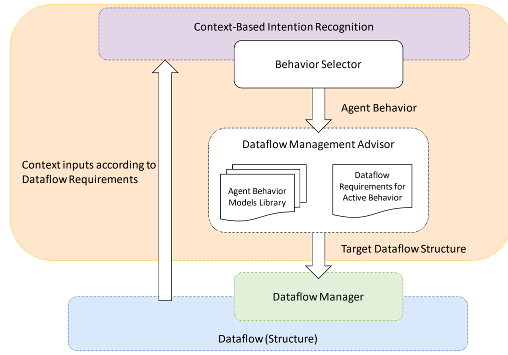

Figure 1: Overview of the proposed software framework for context-adaptive HCI.

hardware faces challenges like massive shock and vibration during launch, as well as extreme temperatures and radiation [18]. In addition, space missions have a very long planing phase. These two facts lead to the problem that hardware in space missions is mostly outdated by years. Retraining large models on this hardware is barely feasible. Balancing the computational load on these systems therefore becomes very important. A context-sensitive adaptation to specific scenarios can preserve the generality to tackle multiple scenarios as well as balance the computational load accordingly. In sum, HCI in space missions is different from HCI on earth in a multitude of aspects, and therefore it is important to preserve flexibility and adaptability but also to enable the system to use perception modules that work in both scenarios.

3 A NOVEL FRAMEWORK FOR HCI IN SPACE

To cope with the outlined challenges in space missions and yet provide an effective HCI, a software framework is being developed in the project `KiMMI SF` [6]. The aim of the project is the development of a software framework for flexible and context-adaptive HCI (see Figure 1). The framework shall offer online reconfigurable data handling. For this, methods for context-dependent evaluation, selection and weighting of data sources will be used. With the help of the proposed framework, a robotic system can independently recognize which context is active and select known or learned models to be applied accordingly. With the described adaptability, a context-based HCI can be realized.

In this framework, the *Context-Based Intention Recognition* module plays an important role as it enables the assistance system to effectively, efficiently and naturally infer the intention of an astronaut. We propose a hybrid Bayesian approach for context-based intention recognition. The context is provided by the pretrained, well-known perception models included in the Perception for Autonomous Systems (PAZ) library [1] and other state-of-the-art libraries. As depicted in Figure 1, the Intention Recognition module uses context to infer an intention. To do so it relies on an underlying network of context variables, where one context variable describes a specific aspect of the whole context. The whole context is defined by all hidden and observable states within the HCI scenario. Thus it can include context variables representing Human 6D Pose, Robot 6D Pose, Object Poses, Gestures, Human Emotion and many more.

Context variables can depend on other context variables or directly on sensor values. The inferred intention is in turn used by the *Behavior Selector* module, which selects the most appropriate behavior for the robotic system corresponding to the inferred human intention. Depending on the selected behavior, the data flow can be adjusted in order to provide context-adaptive HCI and balance the overall load of the system.

4 ARCHITECTURE OF THE TWO-LAYER BAYESIAN NETWORK FOR CONTEXT-BASED INTENTION RECOGNITION

To address the challenges outlined in Section 2, we propose a two-layer Bayesian network, depicted in Figure 2 for intention recognition. Other approaches like [16] use three-layer Bayesian networks modeling context, intentions and actions, and [21] even introduce a fourth layer. However, we believe that the simple two-layer structure provides several advantages. Firstly, it is extremely flexible in a sense that every context variable can be treated in the same way. Systems that use a third layer for actions, give the action recognition component a special meaning. However, this is not always useful. Consider situations where no action is being performed or no action is observable by the system. For example, during an environmental crisis, environmental context may lead to the conclusion that the astronaut needs help, whereas the actual action performed by the astronaut has only little influence on that. Secondly, by modeling everything as context, it is easy to set the importance of specific context variables for a specific intention. This reduces the complexity of the Bayesian network and requires the human experts to only possess knowledge about the mission goals and scenario and eliminates the need for them to know about Bayes theorem and the underlying two-layer Bayesian network. Finally, the proposed structure enables us to make some assumptions to reduce the number of values that should be set by a human expert. The first assumption is that all context variables are independent and all intention variables are independent. This enables us to generically model context and intentions as depicted in Figure 2. Context dependencies can still be modeled in the underlying dataflow structure which allows to use one context variable to produce another. The second assumption is that intentions are binary, which allows us to only consider the positive case and treat the negative case as its complement. This already reduces the number of values which need to be set by half. In contrast, each context variable can have as many instantiations as needed to model the scenario sufficiently. The third assumption is that all context variables have an individual and independent influence on a specific intention (single condition assumption). That is, we assume that we can approximate the conditional probability as the mean over all single condition probabilities $P(I_m|C_{k,l})$, where I_m is the *m-th* intention and $C_{k,l}$ is the *l-th* instantiation of the *k-th* context variable. Under this assumption, to calculate the conditional probability of the *m-th* intention given all contexts with their first instantiation, we would need to calculate:

$$P(I_m|C_{1,1}, C_{2,1}, ..., C_{k,1}) = \frac{\sum_{\hat{k}=1}^{k} P(I_m|C_{\hat{k},1})}{k} \quad (1)$$

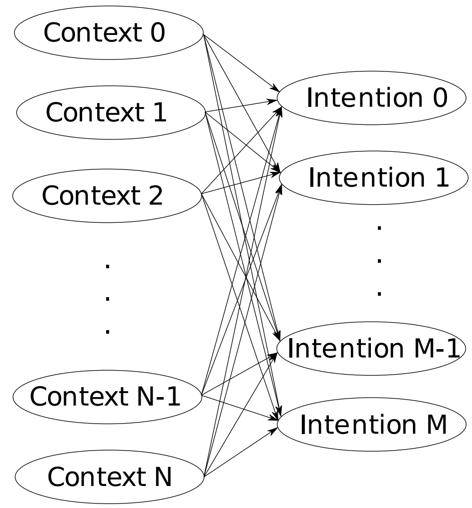

Figure 2: The architecture of the two-layer Bayesian network allows for a high degree of flexibility

The single condition assumption changes the number of values to be set from

$$V(i, j, c, n) = \sum_{\hat{j}=1}^{j} c_{\hat{j}} + i \times \prod_{\hat{i}=1}^{i} n_{\hat{i}} \times \prod_{\hat{j}=1}^{j} c_{\hat{j}} \quad (2)$$

to

$$V(i, j, c) = (i + 1) \times \sum_{\hat{j}=1}^{j} c_{\hat{j}} \quad (3)$$

in the optimal case, where V is the number of values to be set, i is the number of intentions, j is the number of contexts, c_j is the number of context instantiations for the *j-th* context, n_i is the number of intention instantiations for the *i-th* intention. In Equations 2 and 3, $\sum^j c_j$ describes the values needed to set the *apriori* probabilities of all context instantiations, while the rest describes the values needed to fill the CPTs. The single condition assumption does definitely not cover all cases. For example, the fact that the astronaut and a tool are in a specific position should jointly provide a very high probability for the intention *pick up tool*, but the positions of each of them individually should give a rather low probability for *pick up tool*. To handle such cases, we can create conditional probabilities for conditions containing multiple context instantiations that have a combined meaning for a specific intention.

Optionally, the previously inferred intention can be reused as a context variable describing the current intention. This recursive structure implicitly introduces a temporal dependency under the Markov assumption.

5 ADVANTAGES OF CONTEXT-BASED BAYESIAN INTENTION RECOGNITION

The context-based Bayesian intention recognition in the proposed architecture comes with a handful of advantages which we want to present in this section. The most important advantage of the proposed architecture is the flexibility of the hybrid approach. Context variables can be produced by any algorithm, might it be data-driven, heuristic, probabilistic or any other approach. The Bayesian Network will fuse all these context information and jointly infer a probability distribution over all intentions. The Bayesian network itself can be used as an expert system in which human experts define the CPTs in the described minimized way or as a data-driven system which is trained with recorded data from a given scenario if applicable. Specific knowledge about the mission can easily be incorporated, using the Bayesian network as an expert system. The compact architecture of the two-layer Bayesian network as well as the simplifications made on basis of the single condition assumption allows for quick adaptation to new HCI scenarios. The proposed way of creating the CPTs is intuitive and helps to avoid the unintuitive and tedious way of setting CPTs. This makes the system more easily configurable by experts of the space mission without the need for a deeper knowledge about Bayesian networks and CPTs. The quick and easy creation and adaptation of the scenario is especially important in unknown environments which occur regularly in space missions. Missing input data is still a very challenging problem for supervised learning algorithms [8]. The two-layer Bayesian network offers a very intuitive solution to that problem by assigning a priori probabilities to every context variable. These can easily be estimated by a human expert or taken from observations. Additionally, the Bayesian network can be extended easily to include a context variable which is observable but was not yet defined in the network. This helps to iteratively define the Bayesian network towards the final mission. Furthermore, this approach comes with an inherent uncertainty estimation because of its probabilistic nature. Finally, the approach is interpretable, i.e. it is possible to identify the context variables that predominantly influenced the inference of the intention and its uncertainty. This information can be used to generate explanations to understand the decisions made or behaviors shown by the robot assistant. Explainability and uncertainty estimates for predictions would increase the trust in the system dramatically [19, 22].

6 CONCLUSION AND FUTURE WORK

In this paper, we proposed a general architecture for robust, flexible and context-adaptive HCI. We also presented a context-based Bayesian approach to recognize intentions of humans using a two-layer Bayesian network. The individual variables describing the HCI context is in turn recognized using data-driven machine learning models. Thus, this approach helps to balance computational load in robotic systems designed for space missions and reduces the need for complex interaction data, which is unavailable in space applications. In addition to context-adaptability during runtime, we proposed a simplified way to configure the two-layer Bayesian network with less values to set in contrast to filling the complete CPTs. This simplified approach would make it easier and faster for human mission experts to extend and adapt the network to changing requirements of the missions. Furthermore, the inspectability and uncertainty estimation ability of Bayesian networks could promote increased trust in robot assistants and enhance the acceptance to deploy them in LDSE missions. Finally, the context-adaptive human intention recognition can facilitate intuitive interaction with the robots, and thus potentially reduce system-induced stress during interaction in ICE conditions. So far, we have implemented and tested the proposed HCI framework for very small use cases in simulation, as a proof of concept. Ongoing work focuses on extending the framework to include more context variables and human intentions, in order to realize more elaborate HCI scenarios. Future work would focus on deploying the framework on a physical robot built for assisting astronauts in space, e.g. the RH5 humanoid robot [3]. Subsequently, experiments would be conducted to investigate how our context-adaptive hybrid architecture for intelligent and adaptable HCI influences the human factors and the conclusions that can be derived for future space missions.

ACKNOWLEDGMENTS

This work was supported through a grant of the German Federal Ministry of Economic Affairs and Climate Action (BMWK, FKZ 50 RA 2022).

REFERENCES

[1] Octavio Arriaga, Matias Valdenegro-Toro, Mohandass Muthuraja, Sushma Devaramani, and Frank Kirchner. 2020. Perception for Autonomous Systems (PAZ). *arXiv:2010.14541 [cs]* (Oct. 2020). http://arxiv.org/abs/2010.14541 arXiv: 2010.14541.

[2] Pete Bonasso, Stephen Hart, Ana Huaman, and James Kramer. 2018. CASE: A COGNITIVE ARCHITECTURE FOR SPACE EXPLORATION. (2018), 8. https://robotics.estec.esa.int/i-SAIRAS/isairas2018/Papers/Session%209c/2_Bonasso-69-92-Bonasso-Russell.pdf

[3] Julian Esser, Shivesh Kumar, Heiner Peters, Vinzenz Bargsten, Jose de Gea Fernandez, Carlos Mastalli, Olivier Stasse, and Frank Kirchner. 2021. Design, analysis and control of the series-parallel hybrid RH5 humanoid robot. (Jan. 2021). https://doi.org/10.48550/arXiv.2101.10591

[4] Terrence Fong and Illah Nourbakhsh. 2005. Interaction challenges in human-robot space exploration. *Interactions* 12, 2 (March 2005), 42–45. https://doi.org/10.1145/1052438.1052462

[5] Louis Friedman. 2021. Humans and Robots in Space Exploration 63. *Handbook of Bioastronautics* (2021), 831. Publisher: Springer.

[6] Nina Hoyer. 2020. KiMMI SF - Adaptive software framework for context-sensitive, intuitive man-machine-interaction. https://robotik.dfki-bremen.de/en/research/projects/kimmi-sf/

[7] Terry Huntsberger. 2011. *Cognitive architecture for mixed human-machine team interactions for space exploration.* https://doi.org/10.1109/AERO.2011.5747552 Journal Abbreviation: IEEE Aerospace Conference Proceedings Pages: 11 Publication Title: IEEE Aerospace Conference Proceedings.

[8] Niels Ipsen, Pierre-Alexandre Mattei, and Jes Frellsen. 2020. How to deal with missing data in supervised deep learning? (2020), 6.

[9] Siddarth Jain and Brenna Argall. 2018. Recursive Bayesian Human Intent Recognition in Shared-Control Robotics. *Proceedings of the ... IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE/RSJ International Conference on Intelligent Robots and Systems* 2018 (Oct. 2018), 3905–3912. https://doi.org/10.1109/IROS.2018.8593766

[10] R. Kelley, A. Tavakkoli, C. King, A. Ambardekar, M. Nicolescu, and M. Nicolescu. 2012. Context-Based Bayesian Intent Recognition. *IEEE Transactions on Autonomous Mental Development* 4, 3 (Sept. 2012), 215–225. https://doi.org/10.1109/TAMD.2012.2211871

[11] Zhiguang Liu and Jianhong Hao. 2019. Intention Recognition in Physical Human-Robot Interaction Based on Radial Basis Function Neural Network. *Journal of Robotics* 2019 (April 2019), e4141269. https://doi.org/10.1155/2019/4141269 Publisher: Hindawi.

[12] NASA. 2017. NASA's Journey to Mars. https://www.nasa.gov/content/nasas-journey-to-mars

[13] NASA. 2021. Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders. https://humanresearchroadmap.nasa.gov/Risks/risk.aspx?i=99

[14] Dimitris Panagopoulos, Giannis Petousakis, Rustam Stolkin, Grigoris Nikolaou, and Manolis Chiou. 2021. A Bayesian-Based Approach to Human Operator Intent Recognition in Remote Mobile Robot Navigation. (Sept. 2021). https://arxiv.org/abs/2109.12045v1

[15] LuÃŋs Pereira and The Anh Han. 2009. *Intention Recognition via Causal Bayes Networks Plus Plan Generation.* Vol. 5816. https://doi.org/10.1007/978-3-642-04686-5_12 Pages: 149.

[16] LuÃŋs Pereira and The Anh Han. 2010. Intention Recognition with Evolution Prospection and Causal Bayes Networks. 1–33. https://doi.org/10.1007/978-94-007-0093-2_1

[17] Vidya Sagar Reddy. 2018. The SpaceX Effect. *New Space* 6, 2 (June 2018), 125–134. https://doi.org/10.1089/space.2017.0032 Publisher: Mary Ann Liebert, Inc., publishers.

[18] Liz Ronchetti. 2011. Challenges in Space Qualified Hardware. https://www.markimicrowave.com/blog/challenges-in-space-qualified-hardware/

[19] Lindsay Sanneman and Julie A. Shah. 2020. Trust Considerations for Explainable Robots: A Human Factors Perspective. *arXiv:2005.05940 [cs]* (May 2020). http://arxiv.org/abs/2005.05940 arXiv: 2005.05940.

[20] Roland U Sonsalla, Florian Cordes, Leif Christensen, Thomas M Roehr, Tobias Stark, Steffen Planthaber, Michael Maurus, Martin Mallwitz, and Elsa A Kirchner. 2017. FIELD TESTING OF A COOPERATIVE MULTI-ROBOT SAMPLE RETURN MISSION IN MARS ANALOGUE ENVIRONMENT. (2017), 8.

[21] Karim A. Tahboub. 2006. Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition. *Journal of Intelligent and Robotic Systems* 45, 1 (Jan. 2006), 31–52. https://doi.org/10.1007/s10846-005-9018-0

[22] Matias Valdenegro-Toro. 2021. I Find Your Lack of Uncertainty in Computer Vision Disturbing. *arXiv:2104.08188 [cs]* (April 2021). http://arxiv.org/abs/2104.08188 arXiv: 2104.08188.

[23] Zhichao Wang, Bin Wang, Hong Liu, and Zhaodan Kong. 2017. Recurrent convolutional networks based intention recognition for human-robot collaboration tasks. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1675–1680. https://doi.org/10.1109/SMC.2017.8122856

[24] Alexandra Witze. 2021. Lift off! First flight on Mars launches new way to explore worlds. *Nature* 592, 7856 (April 2021), 668–669. https://doi.org/10.1038/d41586-021-00909-z Bandiera_abtest: a Cg_type: News Number: 7856 Publisher: Nature Publishing Group Subject_term: Planetary science, Atmospheric science.