Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)
Demonstrations Track

# PROBANCH: a Modular Probabilistic Anchoring Framework

**Andreas Persson**[1*], **Pedro Zuidberg Dos Martires**[2], **Luc De Raedt**[1,2] and **Amy Loutfi**[1]
[1]Center for Applied Autonomous Sensor Systems (AASS), Örebro University, Sweden
[2]Department of Computer Science and Leuven.AI, KU Leuven, Belgium
{andreas.persson, amy.loutfi}@oru.se, {pedro.zudo, luc.deraedt}@kuleuven.be

## Abstract

Modeling object representations derived from perceptual observations, in a way that is also semantically meaningful for humans as well as autonomous agents, is a prerequisite for joint human-agent understanding of the world. A practical approach that aims to model such representations is *perceptual anchoring*, which handles the problem of mapping sub-symbolic sensor data to symbols and maintains these mappings over time. In this paper, we present PROBANCH, a modular data-driven anchoring framework, whose implementation requires a variety of well-orchestrated components, including a probabilistic reasoning system.

## 1 Probabilistic Perceptual Anchoring

Perceptual anchoring has been defined as the *"[...] process of creating and maintaining the correspondence between symbols and sensor data that refer to the same physical objects"* [Coradeschi and Saffiotti, 2000]. Traditionally, the anchoring process depends on perceptual observations to create and maintain *anchors* (i.e., representations of objects). In the absence of perceptual inputs, e.g., in case of *object occlusions*, this approach is not viable. To remedy this, we extended perceptual anchoring with probabilistic reasoning [Persson *et al.*, 2020; Zuidberg Dos Martires *et al.*, 2020] – resulting in the PROBANCH framework[1].

**Example 1:** *Consider the situation in the top most panel on the right of Figure 1.* PROBANCH *observes three objects and associates each of them with an anchor:* `block-1`*,* `mug-1`*, and* `ball-1`*. In the following two panels, the* `ball` *is occluded by the* `mug`*, which in turn becomes occluded by the* `block`*. Note that the* `block` *is later re-classified as a* `box`*. We illustrate the occlusion by plotting samples from the probability distributions of the possible positions of the* `mug` *and* `ball` *in black and yellow, respectively. Occluded objects are tracked via their relationship with observed objects using logical rules, i.e., the position of an occluded object is logically inferred through the position of the occluding object. For instance, the* `mug` *is tracked via its relationship with the* `box`*, which is observed and tracked by the perceptual component of* PROBANCH*. As the logic rules, which describe the occlusion relationships between objects, can be recursive, observing the* `box` *does also allow* PROBANCH *to track the* `ball`*. Note that the occlusion relation between the* `mug` *and the occluded* `ball` *continues to persist after the* `mug` *has been re-observed, and the corresponding sensor data has correctly been matched with the* `mug-1` *anchor. In the last panel, also the* `ball` *is revealed and re-identified as the same initial* `ball-1` *anchor.*

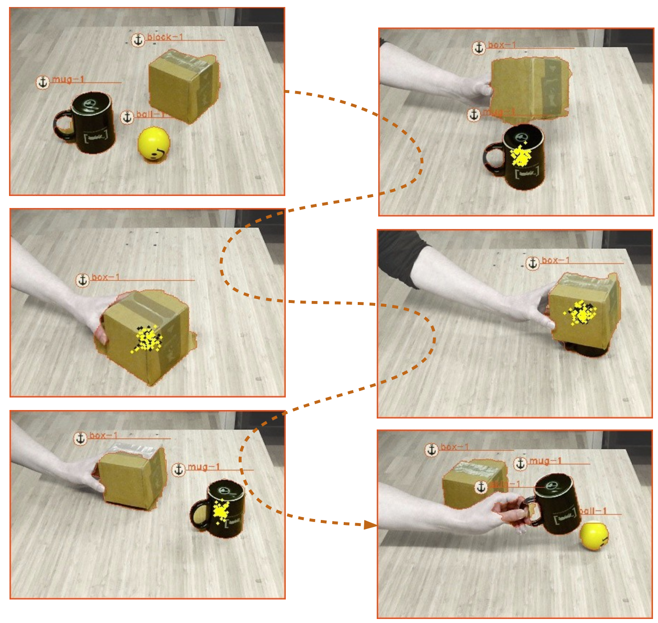


Figure 1: Relational tracking in the presence of transitive occlusions using PROBANCH. Link to video: https://vimeo.com/388874421

A key contribution of PROBANCH is that it allows to reason about occluded objects using a combination of logical, probabilistic and neural-symbolic methods. Other key features of PROBANCH, are modularity and adaptability – meaning that any system of PROBANCH can be interchanged or modified with minimum requirements for re-training the framework.

*Contact Author

[1]https://github.com/probabilistic-anchoring/probanch

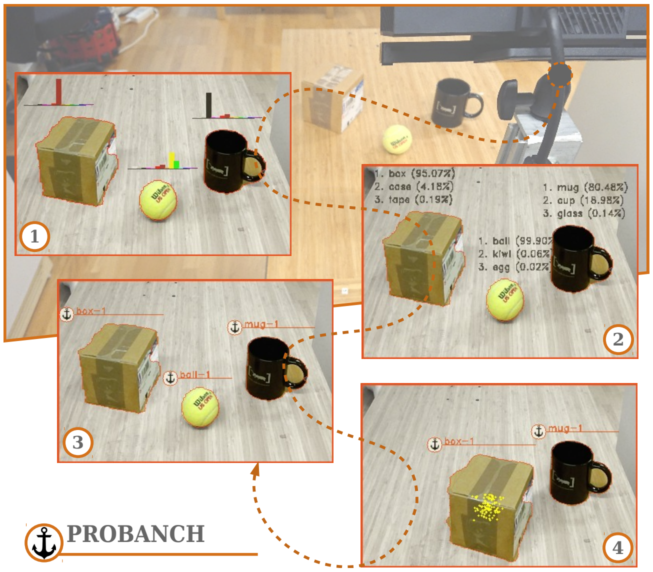


Figure 2: A conceptual overview of PROBANCH.

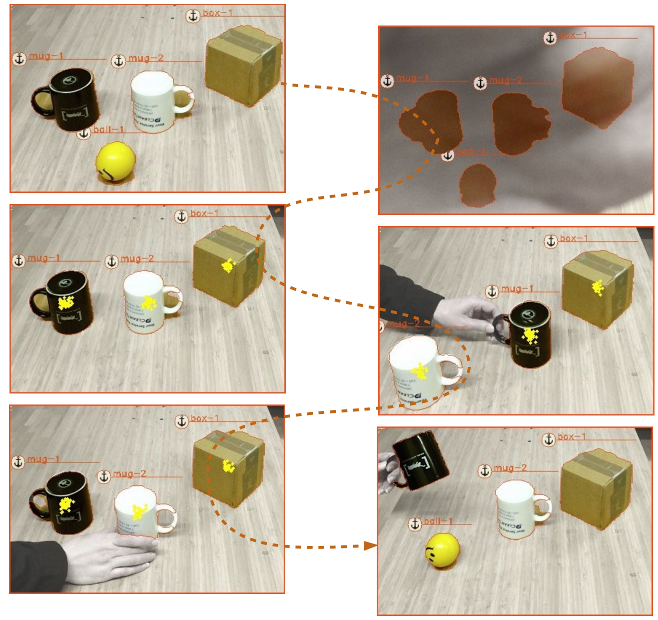


Figure 3: Relational tracking in the presence of mulit-modal probability distributions. Link to video: https://vimeo.com/388872843

## 2 PROBANCH: Framework Description

The PROBANCH framework architecture is a modular architecture implemented with the libraries and communication protocols available in the Robot Operating System (ROS)[2]. PROBANCH is, at its core, a sensor-driven perceptual *bottom-up anchoring system* [Loutfi *et al.*, 2005]. However, the overall framework consists of several systems, which are conceptually represented in Figure 2.

As mentioned in Section 1, the process of perceptual anchoring is to create and maintain the link between sensor data and symbols that refer to the same physical object. In PROBANCH, this process happens in multiple stages; each stage is constituting a system (or module). Triggered by *RGB-D* sensory data, a *perceptual system* segments the sensory data into *percepts* (i.e., the sensor data from an individual object), and measures for each percept *attributes*, e.g., the $\mathbb{R}^3$ position of an object, or the HSV-color histogram over an object (as exemplified in Figure 2.①). A *symbolic system* establishes, subsequently, the percept-symbol linkage, e.g., a certain peek in a color histogram is mapped to the corresponding symbol `yellow`. In addition, a *convolutional neural network*, such as GoogLeNet [Szegedy *et al.*, 2015], is used to *semantically categorize* objects (as shown in Figure 2.②).

Given the percept-symbol linkage of an unknown object, the challenge of any *anchoring system* is to apply a *matching function* to determine whether this object matches any existing anchor (or not). In PROBANCH, a Support Vector Machine (SVM) classifier is used to determine whether an unknown object matches an existing anchor (or not), and whether the object should be *maintained* as an existing matching anchor, or a new anchor should be *created* for the object. An example of anchored objects, identified by their unique symbols (e.g., `mug-1`), is depicted in Figure 2.③.

[2] https://www.ros.org/

PROBANCH is also equipped with a *probabilistic reasoning system* based on *Dynamic Distributional Clauses* [Nitti *et al.*, 2016], which allows PROBANCH to probabilistically reason about anchors and their relationships over time. For instance: `occluded_by(ball-1,box-1):t` (in Figure 2.④), means that `ball-1` is occluded by `box-1` at time `t`. These temporally labeled predicates are defined in distributional clauses, which allows PROBANCH to *track* objects even when some objects are occluded. The distributional clauses define the conditions under which certain relationships hold. These clauses extend the clauses used in programming languages, such as Prolog, with discrete and continuous distributions.

**Example 2:** *As an additional example of probabilistic perceptual anchoring, consider Figure 3. In this example, we initially prepare the state in a probabilistic multi-modal state by covering the RGB-D sensor while hiding the yellow* `ball` *underneath one of the larger objects. The objects are then shuffled around, and the occluded object is tracked. Once the occluded object (the* `ball`*) is re-observed, the probability distribution collapses.*

## 3 The Modularity of PROBANCH

The modular architecture of PROBANCH allows one to extend or replace single systems easily. For instance, in [Persson *et al.*, 2020], we used *connected component segmentation* on organized point cloud data for segmenting arbitrary object instances in tabletop scenarios [Trevor *et al.*, 2013]. In [Zuidberg Dos Martires *et al.*, 2020], we replaced this initial approach in favor of the combined *depth seeding* and *region refinement networks* [Xie *et al.*, 2019]. Virtually any system within PROBANCH can, likewise, be interchanged, modified, or extended. A possible extension of PROBANCH could, for

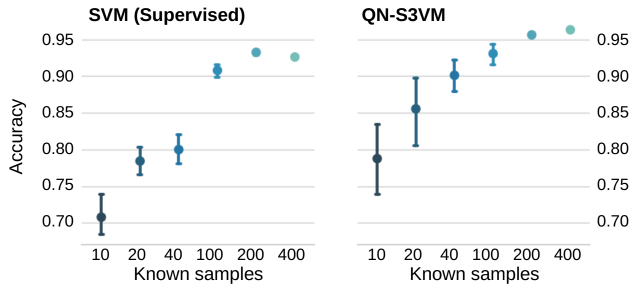


Figure 4: Average classification accuracy of learning the anchoring of objects through supervised classification (left), and semi-supervised classification (right).

example, be to take into account spatial attributes, such as the distance between two objects – similarly to the spatial object features explored in *context-based 3D anchoring* [Ruiz-Sarmiento *et al.*, 2017; Günther *et al.*, 2018].

To demonstrate the adaptability, we completely replaced the *matching function* used in [Persson *et al.*, 2020; Zuidberg Dos Martires *et al.*, 2020] (as part of the *anchoring system* of PROBANCH for determining whether to create or maintain an anchor, seen in Figure 2.③), which was learned using a supervised SVM, by a new matching function learned using a semi-supervised SVM in combination with sparse quasi-Newton optimization (QN-S3VM) [Gieseke *et al.*, 2012]. In such a semi-supervised training regime, the classifier can be learned with a reduced number of labeled known samples combined with a greater number of unlabeled samples. As an illustration, we include a brief experimental comparison, seen in Figure 4. By the result, we observe that an accuracy of about 90% (with ±2% deviation), was attained for the anchor matching function when using the QN-S3VM classifier with as few as 40 known labeled samples in combination with 562 unlabeled samples. By increasing the number of labeled known samples to 400 (combined with 202 unlabeled samples), we can further report an overall best average classification accuracy of 96.1% achieved by the QN-S3VM classifier (which is comparable to the accuracy of the previously used classifier trained with a considerably larger dataset).

## 4 Conclusion

Perceptual anchoring allows for the modeling of object representations that are semantically meaningful to humans and autonomous agents alike. Such representations are useful for tasks that require a mutual human-agent understanding, e.g., human-robot interaction tasks, or collaborative object manipulation tasks. In this paper, we introduced the probabilistic perceptual anchoring framework PROBANCH, which combines ideas from perceptual anchoring, machine learning, and probabilistic programming. As such, it can be regarded as an implementation of a *neural-symbolic vision system*.

## Acknowledgements

The work is supported by Vetenskapsrådet (2016-05321), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. PZD is supported by the Research Foundation-Flanders and Special Research Fund of the KU Leuven.

## References

[Coradeschi and Saffiotti, 2000] Silvia Coradeschi and Alessandro Saffiotti. Anchoring symbols to sensor data: preliminary report. In *Proc. of the 17th AAAI Conf.*, pages 129–135, Menlo Park, CA, 2000. AAAI Press.

[Gieseke *et al.*, 2012] Fabian Gieseke, Antti Airola, Tapio Pahikkala, and Oliver Kramer. Sparse quasi-newton optimization for semi-supervised support vector machines. In *ICPRAM (1)*, pages 45–54, 2012.

[Günther *et al.*, 2018] Martin Günther, J.R. Ruiz-Sarmiento, Cipriano Galindo, Javier González-Jiménez, and Joachim Hertzberg. Context-aware 3d object anchoring for mobile robots. *Robotics and Autonomous Systems*, 110:12–32, 2018.

[Loutfi *et al.*, 2005] Amy Loutfi, Silvia Coradeschi, and Alessandro Saffiotti. Maintaining coherent perceptual information using anchoring. In *Proc. of the 19th IJCAI Conf.*, pages 1477–1482, Edinburgh, UK, 2005.

[Nitti *et al.*, 2016] Davide Nitti, Tinne De Laet, and Luc De Raedt. Probabilistic logic programming for hybrid relational domains. *Machine Learning*, 103(3):407–449, 2016.

[Persson *et al.*, 2020] Andreas Persson, Pedro Zuidberg Dos Martires, Luc De Raedt, and Amy Loutfi. Semantic relational object tracking. *IEEE Transactions on Cognitive and Developmental Systems*, 12(1):84–97, 2020.

[Ruiz-Sarmiento *et al.*, 2017] J.R. Ruiz-Sarmiento, Martin Günther, Cipriano Galindo, Javier González-Jiménez, and Joachim Hertzberg. Online context-based object recognition for mobile robots. In *2017 IEEE Int. Conf. on Autonomous Robot Systems and Competitions (ICARSC)*, pages 247–252, 2017.

[Szegedy *et al.*, 2015] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1–9, 2015.

[Trevor *et al.*, 2013] Alexander JB Trevor, Suat Gedikli, Radu B Rusu, and Henrik I Christensen. Efficient organized point cloud segmentation with connected components. In *3rd Workshop on Semantic Perception Mapping and Exploration (SPME), Karlsruhe, Germany*, 2013.

[Xie *et al.*, 2019] Christopher Xie, Yu Xiang, Arsalan Mousavian, and Dieter Fox. The best of both modes: Separately leveraging rgb and depth for unseen object instance segmentation. In *Conf. on Robot Learning (CoRL)*, 2019.

[Zuidberg Dos Martires *et al.*, 2020] Pedro Zuidberg Dos Martires, Nitesh Kumar, Andreas Persson, Amy Loutfi, and Luc De Raedt. Symbolic learning and reasoning with noisy data for probabilistic anchoring. *arXiv:2001.08603*, 2020.