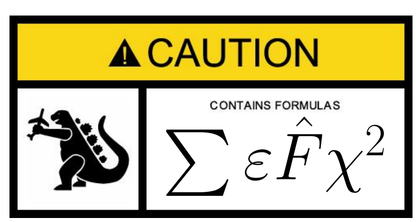
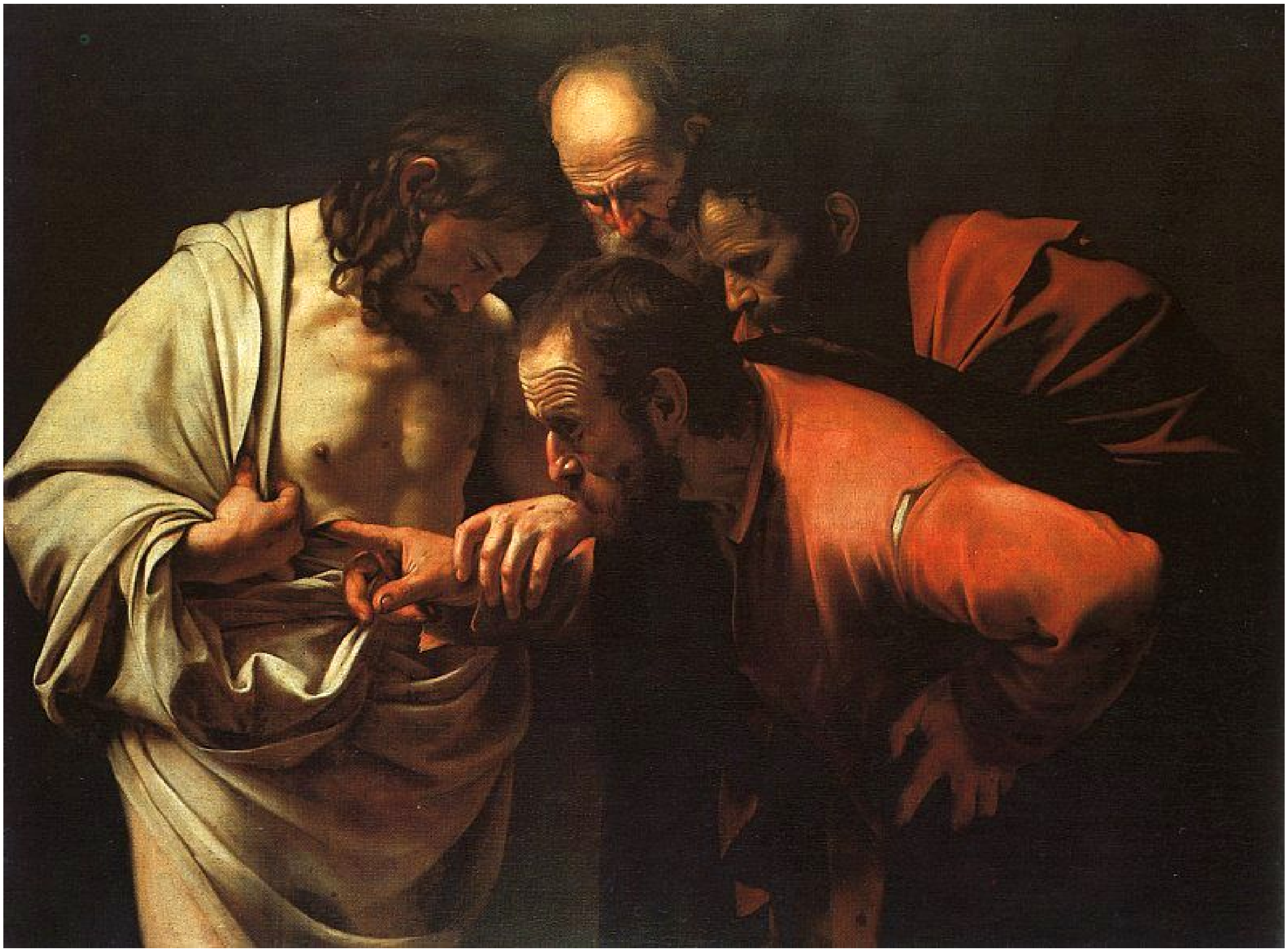
Do Robots Have to Reduce Uncertainty in Their Internal Representations?

Vladimir Sukhoy and Ryan Kirk
{sukhoy, rakirk}@iastate.edu

Image credit: `http://rsss.anu.edu.au/ maier/workshop.html`

This document is a project proposal for HCI 585X – the course on Developmental Robotics taught by Dr. Alexander Stoytchev in Spring 2011.





“The Incredulity of Saint Thomas,” by Caravaggio (AD 1601–1602)

24 ¶ But Thomas, one of the twelve, called Did’ymus, was not with them when Jesus came.

25 The other disciples therefore said unto him, We have seen the Lord. But he said unto them, Except I shall see in his hands the print of the nails, and put my finger into the print of the nails, and thrust my hand into his side, I will not believe.

26 ¶ And after eight days again his disciples were within, and Thomas with them: *then* came Jesus, the doors being shut, and stood in the midst, and said, Peace *be* unto you.

27 Then saith he to Thomas, Reach hither thy finger, and behold my hands; and reach hither thy hand, and thrust it into my side; and be not faithless, but believing.

28 And Thomas answered and said unto him, My Lord and my God.

29 Jesus saith unto him, Thomas, because thou hast seen me, thou hast believed: blessed *are* they that have not seen, and *yet have believed.*

(John 20, King James Version)

INSPIRATION

St. Thomas was among the Twelve Apostles of Jesus Christ. St. Thomas knew very well what will happen after the crucifixion. Yet, he still wanted to verify this knowledge. The Bible does not say if Thomas touched Jesus or not, but the consensus among the artists is that he did, as shown in Fig. 1. It is clear that visual and auditory verification occurred and that St. Thomas really wanted to perform tactile verification as well. Furthermore, St. Thomas sought to establish the temporal continuity and validate cause and effect relationships by confirming the presence of nail prints on the hands of Jesus.

St. Thomas was a verificationist at heart throughout the New Testament. After the death of Lazarus, Jesus wanted to return to Judea. Previously, judeans wanted to stone Jesus. Yet St. Thomas was full of doubts and was willing to pay the ultimate price to remove them: "Let us also go, that we may die with him" (John 11:16, King James Version). At the Last Supper, St. Thomas was uncertain of everyone's internal representations: "Lord, we know not whither thou goest; and how can we know the way?" (John 14:5, King James Version). In response to this, Jesus decreased uncertainty using the ultimate authority – he explained in detail how he is related to God the Father. It turned out that being a god was not enough to fully convince St. Thomas – his ultimate authority was his own somatosensory system.

These events occurred roughly 1965 years before Rich Sutton wrote his essay [1] about verification and AI. As Fig. 1 shows, the process of grounding St. Thomas' internal representation of resurrection in his own sensorimotor experience fascinated people through the centuries and inspired some really good art. The idea is not new.

Clearly, verification principle suggests that robots should be more like St. Thomas and less like St. Peter. Suppose that St. Thomas was a robot. What kind of robot was, then, St. Thomas? What kind of hardware did St. Thomas have? **What kind of software?**

On the software side, St. Thomas was often uncertain. Furthermore, St. Thomas was very certain that he was uncertain. St. Thomas was ready to die exploring if he could reduce uncertainty in his own representations (John 11:16). This suggests that: 1) exploration was St. Thomas' preferred means of knowledge acquisition (this also explains why he was not with the rest of the Disciples when Jesus came – St. Thomas would rather explore than just wait for Jesus), 2) the exploration strategy minimized the uncertainty between his internal representation and the world at all costs, and 3) his internal representation of the world was quite capable of remarkable transformations because the transition from unbelief to believing is quite a transformation indeed.

RESEARCH QUESTION

How can robot guide itself using reduction in the uncertainty of internal representations?

GOALS

Achieving the three goals of the project will shed the light on three different aspects of the research question.

Goal I : Develop Mathematical apparatus for self-detection based on Information Theory

The first goal of this project is **to develop mathematical apparatus that a robot can use to detect its own body** in the data it records from its senses. The key idea is to relate the uncertainty in various patterns present in the data to motor actions. Those patterns for which **the uncertainty is reduced** by the information stream used to control the motors represent the robotic body. The key idea is to combine Statistics with Information Theory to **estimate the mutual information** between different sources of information.

Goal II : Show that reduction in uncertainty can guide skill acquisition

The second goal of this project is **to demonstrate that a robot driven by a strategy that seeks to reduce uncertainty can acquire a real-world manipulation skill autonomously**. The specific skill to be acquired is manipulating – i.e., both detecting and pressing – doorbell buttons.

Goal III : Formulate a research question about phase transitions in robotic systems that learn and develop

The third goal of this project is **to pose a research question that clarifies possible roles of phase transitions in robotic learning and development**. Perhaps a phase transition can be demonstrated in a robotic system for self-detection or in a button-pressing robot. The research question must be grounded in research on phase transitions in Developmental Psychology and Neuroscience and **a survey of this research has to complement the question**.

(a) Roslin, T. (1268) (b) Buoninsegna, D. (1308-11) (c) Verrocchio, A. (1476-83) (d) St. Bartholomew Altar, Master of (1501) (e) Conegliano, C. (1505) (f) Dürer, A. (1511)

(g) Mazzolino, L. (1520) (h) Salviati, C. (1572) (i) Vasari, G. (1572) (j) Caravaggio (1601-02) (k) Terbrugghen, H. (1604)

(l) Rubens, P. (1613-15) (m) Honthorst, G. (1620) (n) Rembrandt H. (1634) (o) Guercino (not dated, between 1591 and 1666)

Fig. 1: Some famous art depicting the process of reducing the uncertainty regarding the resurrection of Jesus Christ. In most of these works, St. Thomas is using the tactile receptors in his fingertip to establish temporal continuity and validate cause and effect relationships. Individual works shown: (a) “The Incredulity of Thomas” by Toros Roslin (miniature), Yerevan, Matendaran, No. 10675; (b) “Doubting Thomas” by Duccio di Buonisegna (tempera on wood, 55.5×50.5 cm) Museo dell’Opera del Duomo, Siena; (c) “Christ and Doubting Thomas” by Adrea del Verrocchio (bronze, height: 230 cm), Orsanmichele, Florence; (d) St. Thomas Altarpiece (oak, 143×106 cm – central panel, 143×47 cm – each wing), Wallraf-Richartz Museum, Cologne; (e) “Incredulity of St Thomas with Bishop Magno” by Cima da Conegliano (tempera and oil on panel, 215×151 cm), Gallerie dell’Accademia, Venice; subreffig:durer “Small Passion: 33. The Incredulity of St Thomas” by Albrecht Dürer (woodcut), British Museum, London; (g) “The Incredulity of St Thomas” by Ludovico Mazzolino (oil on panel), Galleria Borghese, Rome; (h) “The Incredulity of St Thomas” by Cecchino del Salviati (oil on wood transferred to canvas, 275×234 cm), Musée du Louvre, Paris; (i) “Incredulity of St Thomas” by Giorgio Vasari (oil on panel), Santa Croce, Florence; (j) “The Incredulity of Saint Thomas” by Caravaggio (oil on canvas, 107×146 cm), Sanssouci, Potsdam; (k) “The Incredulity of Saint Thomas” by Hendrick Terbrugghen (oil on canvas, 108.8×136.5 cm), Rijksmuseum, Amsterdam; (l) “The Incredulity of St Thomas” by Pieter Pauwel Rubens (oil on wood, 143×123 cm – central panel, 146×55 cm – side panels), Koninklijk Museum voor Schone Kunsten, Antwerp; (m) “The Incredulity of St Thomas” by Gerrit van Honthorst (oil on canvas, 125×99 cm), Museo del Prado, Madrid; (n) “The Incredulity of St Thomas” by Rembrandt Harmenszoon van Rijn (oil on wood, 53×51 cm), Pushkin Museum, Moscow. (o) “Doubting Thomas” by Guercino (oil on canvas), Residenzgalerie, Salzburg.

WHY?

Artificial Intelligence has enjoyed tremendous success over the last twenty five years. Its tools and techniques are now main stream within computer science, and at the core of so many of the systems we use every day. Search algorithms, the backbone of traditional AI, are used throughout operating systems, compilers and networks. More modern machine learning techniques are used to adapt these same systems in real-time. Satisfiability of logic formulas has become a central notion in understanding computability questions and once esoteric notions like semantic ontologies are being used to power the search engines that have become organizers of the world's knowledge, replacing libraries, encyclopedias, and automating business interfaces. And who would have guessed that AI powered robots in people's homes would now be counted in the millions. So much accomplishment to bring pride to us all.

But at the same time Artificial Intelligence has not yet succeeded in its most fundamental ambitions. Our systems are still fragile when outside their carefully circumscribed domains. The best poker playing program can't even understand the notion of a chess move, let alone the conceptual idea of animate versus inanimate. A six year old child can discuss all three domains, but may not be very good at any of them compared to our specialized systems. The challenge for AI, still, is to capture the fundamental nature of generalized perception, intelligence, and action. Worthy challenges for AI that would have tremendous practical impact, are, in my opinion:

- **the generic visual object recognition capabilities of a two year old child**
- **the manual dexterity of a six year old child**
- **the social interaction and language capablities of a ten year old child**

So much work for all of us to be challenged by.

Rodney Brooks[1]

The field of Artificial Intelligence, much like AIs in John Gibson's science fiction novels, is currently fractured [3]. In the early days, AI and Robotics were about machines with human-like intelligence. The goal of AI was to provide the software, while Robotics sought to provide the hardware.

The situation is now quite different: Computer Vision does not have much to say to Genetic Algorithms, while Decision Theory is quite disjoint from Natural Language Processing. It can be assumed that only the advances in integrated systems and successful applications of first principles can bridge the gaps between many disjoint sub-fields of AI and Robotics. Only these advances can lead to the integrated intelligence with human-like capabilities.

Human development and the human brain are the key areas where the inspiration for a unified approach towards artificial intelligence must come from. If there is agreement on anything in AI, then the fact that humans are intelligent must be it. Therefore, the key integrating principles, if they can be at all isolated, must be present in human development and the human brain.

Human development can really be thought of as series of linear periods of growth interrupted by brief periods of dynamic change. Here, a case will be made for the idea that these dynamic periods of growth can be thought of as phase shifts within the brain. Conceptualizing these periods of growth as phase shifts allows for the inclusion of methods from physics that deal with phase shifts in matter. These methods from Physics center on the conservation of energy, the notion of entropy and the quantification of uncertainty. The field of information theory commonly employs methods to quantify uncertainty within a data set. The goal of this project will be to explore what role, if any, uncertainty has in dynamic periods of growth within humans and within robots. The experiment will focus upon replicating dynamic periods of growth within humans using a robot that learns through by uncertainty driven exploration.

INFANT FEASIBILITY TEST

The goal of this section is to show that all three goals are accomplished by two year old human infants.

1) **Self-detection** is a basic capability of animals and human infants. For the purposes of this test, it is sufficient to note that even two month old infants are capable of self-detection in a TV screen [4].
2) **Button-pressing** is a trivial task for a two year old. According to Piaget [5], the tertiary circular relations emerge between 12 and 18 months of age. For instance, if pressing one button on the telephone does not activate it, infants press other buttons until the desired effect is achieved [6, p. 112]. There is also evidence that even 9 months old infants can predict that a bright light will flash or there will be an interesting sound when an experimenter presses a colored button [7].
3) **Phase transitions.** There is significant evidence that phase transitions routinely occur in the human brain [8]. During rhythmic movements, human motor commands exhibits traits that can be explained by a model based on a dynamical systems undergoing a phase transition [8]. Another example is that hypnic jerk – an involuntary movement that often occurs at the onset of sleep – can be explained by a behavior of the dynamical systems near its instability [9]. A

[1] This essay was prepared for the workshop on the Future of AI held at the 25$^{\text{th}}$ anniversary of AAAI [2].

Fig. 2: The target audience for the project, top to bottom: 1) IROS 2011 – IEEE/RSJ International Conference on Intelligent Robots and Systems, paper submission deadline: Mar 14th, 2011; 2) ICDL-EPIROB 2011 – First Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, paper submission deadline: March 28th, 2011; 3) Humanoids 2011 – 11th IEEE-RAS International Conference on Humanoid Robots, paper submission deadline: May 29th, 2011.

popular theory [10] of cognition and action uses an approach based on dynamical systems, which can exhibit phase transitions.

Finally, there is evidence that visual attention can be modelled using the approach based on the change in uncertainty of the internal model [11].

Target Audience

The target audience for this project is **not human**. The project is intended primarily for scientific conferences and journals. Due to the tight timeline, it is not likely that a journal paper can be fully completed and submitted before the final report is due. Therefore, a short list of scientific conferences in Robotics and AI with paper submission deadlines that are close to the project due date was composed. These conferences, shown in Fig. 2 are the target audience of this project. The workshops held at various conferences also belong to the target audience if they match the scope of the project.

Related Work

Self-detection

The comprehensive review of the related work for self-detection is a prerequisite for publishing. Luckily, a number of reviews are available in literature, e.g. [12]. During the project, the related work should be further surveyed and the relationship with the proposed entropy-based methodology must be established.

Buttons

The analysis of the related work on pressing buttons in Robotics produced several categories for the proposed approaches. These categories highlight different aspects of the manipulation problem and the proposed methodology for solving it.

1) Detecting buttons is hard, pressing them is straightforward. Work in this category is focused on a single aspect of the problem: detecting buttons using vision. Once a button is detected, it is assumed to be easily pressable. Due to the narrow focus of this line of work, the feedback that a button might generate is often completely ignored. The evaluation is often performed for elevator buttons [13] [14] [15]. The fact that buttons in a typical elevator are arranged in a grid pattern and have numeric labels is often used to boost the performance of the learning algorithm [13] [14].

2) Both detecting and pressing buttons is very hard. Another category of the related work assumes that detecting and manipulating buttons, switches, levers, knobs, and similar widgets designed for humans is intractable for robots. To help robots solve these problems, different types of environmental augmentations are proposed. For example, reflective markers [16] or RFID tags [17] [18] can be attached to the widgets. A tag may inform the robot where the widget is, how to activate it and what happens when it is activated. The main focus of this line of work is on robotic applications enabled by different types of environmental augmentations [18].

3) Understanding social context is crucial for both pressing and detecting buttons. The third category of the related work focuses on social aspects of manipulation. These approaches seek to interpret human-provided social cues associated with robotic actions. For instance, a robot can learn how to detect buttons when humans point at it [19] [20] and learn to press it from human demonstrations [21].

4) Pressing and detecting buttons must be learned together. The last category of the related work differs from other categories in two ways: 1) both pressing and detecting tasks are regarded as challenging, but solvable; 2) the visual model for detecting buttons is trained from multimodal events produced by pressing them. Previously, it was shown that a robot can bootstrap the visual model for detecting buttons by exploring them with pushing behaviors [22]. Our work builds on these ideas and shows how both skills can be developed together in real time.

GPU Programming: GPU programming [23] was used in this work to achieve real time performance for the visual pipeline. Image convolutions are extensively used in the pipeline and implementing them on GPU [24] was a key enabling factor for the speedup. Previously, SIFT [25] and SURF [26] visual pipelines were implemented using GPU to improve their performance.

Developmental Psychology: E.J. Gibson [27]. Experience obtained while exploring objects stimulates further interest [28]. 7-11 months old infants are not interested in seeing object or people who manipulate objects until the infants have had the chance to play with them. Infants as young as 9 m.o. can predict interesting events when an experimenter presses a colored button [7]. Infants perform repetitive movements when they learn to manipulate [5].

EXPERIMENTAL SETUP

A research question will be developed as a process of conducting secondary research in the fields of human sciences and engineering. This question will relate to combining the academy's knowledge about phase shifts that occur within the human experience with established methodology for modeling learning and development in robotics. This design will be reliant upon an experimental, robotic apparatus that is a part of Dr. Alexander Stoytchev's lab. Using a robot is important because it gives the experimenters unparalleled access to the inner working of the intelligent system's mind. The goal will be to show that phase changes occur both within categorization and within self-detection through the use of the same algorithm. As mentioned, an algorithm is going to be developed based upon reducing uncertainty pursuant to the laws of entropy. The results, as measured by the robot, will also be analyzable by the experimenters. At this time, research is still being done into the best means of determining whether phase shifts have occurred within learning; however, evaluation of results will be based upon the robots recorded sensory-modalities, the robots internal model and the performance of the robot.

EVALUATION

Definition of Success

This project succeeds if it produces peer-reviewed publications. Fig. 3(a) shows the high-level approach that motivated the proposed evaluation criterion.

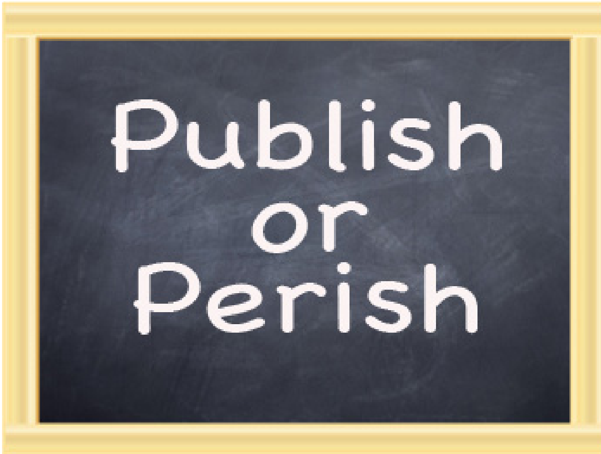
Due to the short timeline, it is **not realistic** to expect acceptance decisions for submitted papers before the final report is due. Therefore, the project success will be computed from the number of papers n submitted to peer-reviewed conferences and journals. The function $\mathrm{Success} : \mathbb{N} \rightarrow \mathbb{R}$ that computes project success (measured in %) is

$$\mathrm{Success}(n) = \log_2(1 + n) \times 100\%. \tag{1}$$

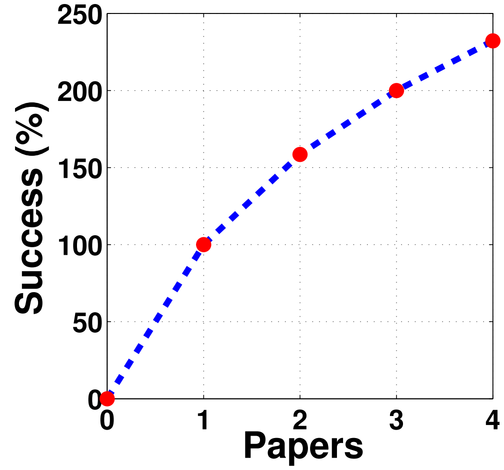
For example, if the project produces no paper submissions, $\mathrm{Success} = 0\%$. For a single submission, $\mathrm{Success} = 100\%$. For two submitted papers, $\mathrm{Success} \approx 158.5\%$, etc. $\mathrm{Success}$ for possible project outcomes is plotted in Fig. 3(b).

Phase Transitions

Literature review is included in Goal III.



(a) Fundamental Rule



(b) Success Function

Fig. 3: Definition of the project success: (a) the fundamental rule that determines project success (image credit: `http://converge.baderrutter.com/`); (b) project success as a function of the number of paper submissions to peer-reviewed journals and conferences according to the success function (1). As more papers are written, the project becomes more successful.

Discussion

Goal I : Mathematical Apparatus for Self-detection

The self-detection problem:

Given a feature, the robot needs to decide if it is self or not self.

Information theory is used to solve this problem. The key idea is to determine whether the stream of information from the feature is influenced by the stream of information that controls movement of the robot. If this is true, the feature is self. If this is not true then the feature is not self.

More specifically, the relationship between movement of the feature, quantified by a variable M, and a the temporal delay since the last motor command, quantified by a variable D, is used. The mutual information $I(M;D) \geq 0$ quantifies the amount of information shared between the two variables. The criterion for self-detection:

$$\text{Feature} = \begin{cases} \text{Self} & \text{if } I(M;D) > 0, \\ \text{Not Self} & \text{if } I(M;D) = 0. \end{cases}$$

The mutual information $(M;D)$ can be written using the Shannon entropy [29] function H as follows:

$$I(M;D) = H(M) + H(D) - H(M,D), \tag{2}$$

where the entropy $H(X)$ quantifies the uncertainty in X:

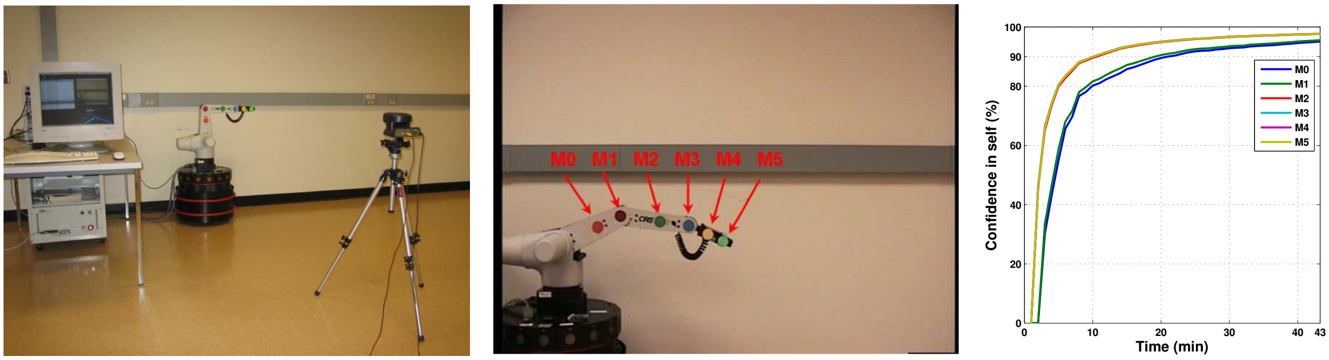
$$H(X) = H(p(x_1), \ldots, p(x_m)) = -\sum_{i=1}^{m} p_i \log p_i, \tag{3}$$

where $p_i = p(x_i)$.

The robot does not know true distributions for M, D, or the true joint distribution. It can only estimate these distributions from the data. For instance, the temporal delay can be discretized using a histogram, while movement can be reduced to a binary variable. For more details on statistical estimation, see the Appendix.

Preliminary evaluation was performed on the dataset from [30]. The dataset consists of the timestamped color marker tracking coordinates, recorded while the robot was performing motor babbling. Fig. 4(a) and 4(b) give a visual summary of the self-detection experiment used to collect this dataset. For a detailed description of the dataset and the experiment, the reader is referred to [30, Chapter 5].

Preliminary results for entropy-based self-detection show that the method can be applied to a variety of different datasets. For one of the CRS+ A251 datasets, the results are shown in Fig. 4(c).



(a) Robot (b) Color Markers (c) Preliminary Results

Fig. 4: Preliminary results for the dataset collected using the CRS+ A251 robot: (a) the robot used to collect the dataset; (b) six color markers used for self-detection; (c) preliminary results for the entropy-based approach. Preliminary results indicate high confidence in the self hypothesis. The experimental setup is described in more detail in [30, Chapter 5], from which (a) and (b) are reprinted with permission.

Goal II : Detecting and Pressing Buttons

This section summarizes the current state of the work on Goal II. The methodology is capable of learning to press and detect buttons in real time. More work is still needed for integration of the various components of the system and a new experiment needs to be conducted to produce publishable results. Some of these experiments were already conducted, those are mentioned in the past tense.

Learning Framework: To illustrate how the robot applies the learning framework illustrated in Fig. 5, we describe the steps in chronological order performed for one pushing behavior.

1) The robot generates a number of random candidates. The robot chooses the next behavior to be performed from these candidates.
2) The model for predicting press locations predicts where each candidate might press the environment.
3) The predicted press locations are assigned the 10×10 pixel patches in the image. The result gives the spread of patches that may be explored.
4) The visual model uses the visual features associated with each of the patches to predict the distribution of acoustic outcomes expected for pressing each of these patches.
5) The current exploration strategy uses the predicted distributions to select the patch to explore.
6) The robot performs the candidate pushing behavior targeted at the selected patch.
7) The outcome of the performed behavior is recorded: this includes both the observed press location and the acoustic outcome of the behavior.
8) The observed press location and the behavior parameters are used to update the model for predicting the press locations.
9) The acoustic outcome of the behavior and the visual features for the patch that the robot pressed update the visual model.
10) The robot returns to step 1 and performs the next behavior.

Driving exploration using the visual model: Our previous work [22] laid the foundation of the learning framework for pressing and detecting buttons. However, the framework described in [22] was not yet complete. The last link – i.e. the mechanism for driving exploration through the visual model – was missing. The design of a complete framework is be described below.

To use the visual model for exploration, the robot has to press image locations that appear interesting. This problem is solved by generating a number of candidate pushing behaviors at random, predicting where each of these candidates may press the experimental fixture, using the visual model to select the most interesting press location, and executing the corresponding behavior. More formally, given the behavior parameters b, the model predicted the tracking coordinates $x \in \mathbb{R}^2$ of the fingertip at the moment of press should the robot execute the pushing behavior parametrized by b.

To satisfy the requirements of the proposed learning framework, the model should support real time predictions and updates. The robot used a model based on Locally-Weighted Projected Regression (LWPR) [31] was used in this experiment. This algorithm is based on nonparametric regression with locally-linear models. It supports incremental training and its complexity is linear on the number of inputs. It was previously used for learning dynamic movement primitives – an approach for motor control and skill learning. This algorithm satisfies the requirements of the learning framework: it supports incremental updates and its predictions are computed in real time. The model is trained using the outcomes of pushing behaviors performed by the robot. The skill of pressing visually interesting locations was developed together with other abilities during the experiment

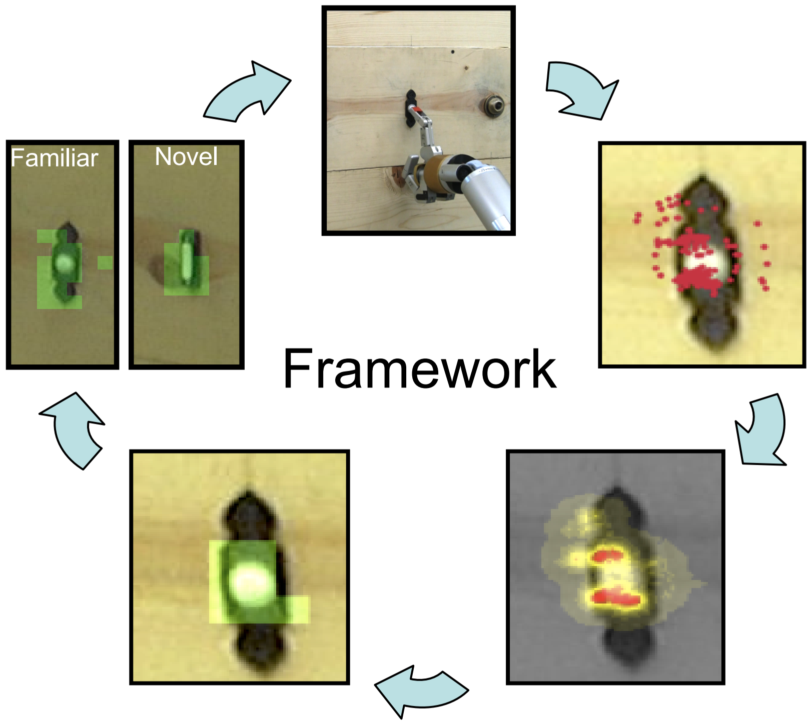


Fig. 5: Summary of the learning framework. The robot starts by pressing buttons. Auditory feedback from buttons combined with the fingertip tracking gives point clouds. These point clouds are used to extract the probability distribution of the auditory events. This distribution is thresholded to train the visual model for identifying the part of a button that has to be pressed to produce sound. This visual model is applied to the frames from the robot's camera and used to select further pushing behaviors.

itself.

Exploration Strategy:

> The need to get information from the environment is as strong as to get food from it, and obviously useful for survival. The search is terminated not by externally provided rewards and punishments, but by internal reduction of uncertainty. The products of the search have the property of reducing the information to be processed. [32, pg 144]

Exploration strategies based on uncertainty reduction were proposed before [33] [11]. A computational model based on the principle of uncertainty reduction agreed with human attention and gaze patterns [11].

To learn about the buttons, the robot explored them with pushing behaviors. The goal of the exploration strategy was to select the next pushing behavior. The choice was made in two stages. First, the visual servoing model for predicting the press locations predicted what patches are pressable. Next, from these patches the exploration strategy used the visual model for detecting buttons and the visual memory of to select the one patch to press.

In previous work [34], the robot also explored doorbell buttons. However, the exploration was performed using prioprioception alone. The visual model was built offline, in batch mode [22]. In contrast, in this work the exploration is guided by vision in real time.

The exploration strategy used the visual model for detecting buttons and the visuospatial memory for the outcomes of pressing them. The visual model predicts where the buttons may be. The memory records where they are. The goal of exploration is to reduce the uncertainty of the memory with respect to the visual model. In other words, the robot's goal is to ensure that what seems to be a button is indeed a button and vice versa.

More formally, let S be a set of patches that the robot can press. When the robot pressed a patch $p \in S$, it recorded the outcome o in the set of outcomes $O = \{buzzer, no\ buzzer\}$. The exploration strategy uses two probability distributions of these outcomes: the predictions $\Pr(V = o|p)$ according to the visual model and the memorized outcome frequencies $\Pr(M = o|p)$ recorded when the robot pressed p. The goal of exploration is to minimize the uncertainty between these two distributions. Conditional entropy $H(M|V)$ is used to quantify this uncertainty:

$$H(M \mid V) = \sum_{o_1, o_2 \in O} \Pr(V = o_1, M = o_2) \log \frac{\Pr(V = o_1)}{\Pr(V = o_1, M = o_2)}.$$

The marginal probabilities are estimated over S assuming uniform prior (all patches are alike):

$$\Pr(V = o) = \sum_{p \in S} \frac{1}{|S|} \Pr(V = o|p).$$

In addition, the joint probabilities are assumed to be conditionally independent given a particular patch:

$$\Pr(V{=}o_1, M{=}o_2) = \sum_{p \in S} \frac{1}{|S|} \Pr(V{=}o_1|p) \times \Pr(M{=}o_2|p).$$

Additive smoothing was used to avoid zero probabilities [35].

The visual model gives the estimates of $\Pr(V = o|p)$ for each individual patch using the classifier trained as described in Section [ref]. The visuospatial memory estimates the probabilities using the frequencies of the outcomes recorded when the robot pressed the patch, i.e.,

$$\Pr(M = o|p) = \frac{N(o, p)}{N(buzzer, p) + N(no\ \ buzzer, p)},$$

where $N(buzzer, p)$ is the number of times pressing the patch p triggered the buzzer and $N(no\ \ buzzer, p)$ is the number of times p was pressed in silence.

The **exploration** strategy sought to select a patch p to press to minimize the expected uncertainty of the memory given the predictions of the visual model given that p will be pressed. More formally,

$$p = \operatorname*{argmin}_{p \in S} E[H(M \mid V, p)],$$

where the expectation is calculated over the possible outcomes of pressing p

$$E[H(M \mid V, p)] = \sum_{o \in O} \Pr(M = o|p) H(M \mid V, p, o).$$

Here, $H(M \mid V, p, o)$ is obtained by incrementing the outcome counter for the patch p and computing the conditional entropy value with the new counters.

For **exploitation**, the robot selected p to maximize both the memorized probability $\Pr(M = buzzer \mid p)$ and the visual model prediction $\Pr(V = buzzer \mid p)$ while minimizing the entropy of the memorized distribution $H(M \mid p)$, where

$$H(M \mid p) = -\sum_{o \in O} \Pr(M = o \mid p) \log \Pr(M = o \mid p).$$

This was achieved by

$$p = \operatorname*{argmax}_{p \in S} \Pr(M = buzzer \mid p) + \Pr(V = buzzer \mid p) - H(M \mid p).$$

With this criterion type of criterion for exploitation the robot would be able to press a button in the dark using its memory. The exploitation strategy instantly can adapt to the loss of the visual model. Preliminary benchmarks of the exploitation strategy, shown in Fig. 6, indicate that the robot can press a button reliably.

Spatial Representation of Multimodal Events: The tracking coordinates of the fingertip were recorded for each visual frame recorded by the robot's camera. These coordinates were associated with the events detected by the robots across different modalities to map them to space. For auditory events, the example point cloud is shown in Fig. 8(a). The robot can press buttons with parts of its finger distinct from the fingertip. Several examples of pressing buttons with various finger parts are shown in Fig. 7 (these examples were extracted from the dataset used in our earlier work [22]). The tracking position of the fingertip does not have to be exactly over the button in visual space when it is pressed.

The point clouds by themselves are not sufficient for detecting novel buttons. To train a visual model, the image needs to be partitioned and the data from the point clouds needs to be converted to a class label. In prior work [22], this was done by computing the density of auditory events and thresholding it over image patches. For the proposed learning framework the old mechanism had to be modified for two reasons: 1) whenever a new auditory event is recorded, the density changes for a significant region in the image which comes at extra computational cost, and 2) the densities did not map clearly to the probability of of ringing a button when pressing it in certain locations – rather, it selected the regions that were more likely to ring when pressed than other regions.

In this experiment, a different approach is used for working with the point clouds. Instead of the auditory events, the press locations corresponding to the proprioceptive events are used. Each of these events can be assigned to two distinct classes depending on whether the push triggered the buzzer or not. This distinction provides both positive and negative examples in the point cloud, as shown in Fig. 8(b). The presence of both types of examples allows to estimate the probability of ringing given that the robot presses a certain location in the image. Finally, the ability to verify its own knowledge by pressing interesting image patches eliminated the need for synthetic negative examples – the robot is now able to generate negative examples by itself.

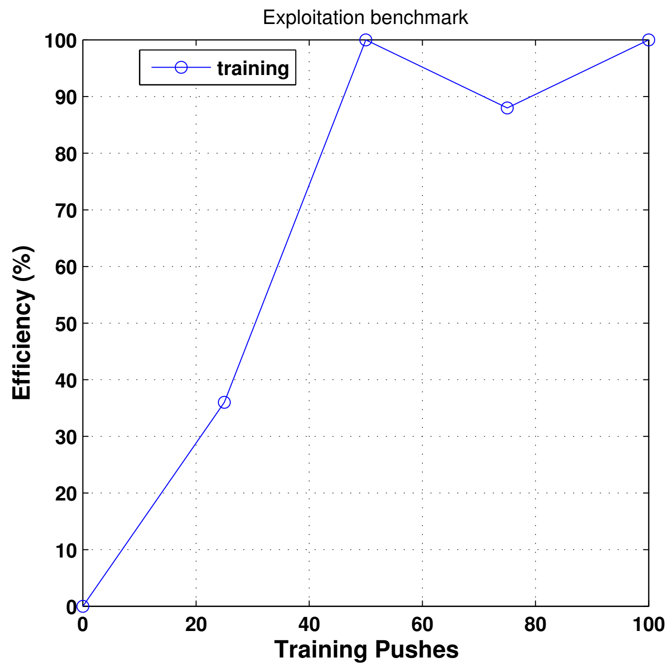


Fig. 6: Exploitation efficiency for a single button. The x-axis is the number of training pushes. The y-axis is the percentage of exploitation pushes that successfully pressed the button and triggered the buzzer. For each data point the robot performed the evaluation on 25 pushes. The sensory data recorded during exploitation was not used for any training.

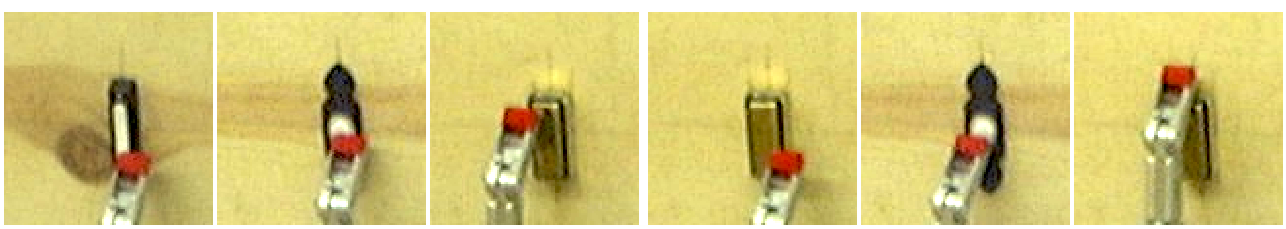
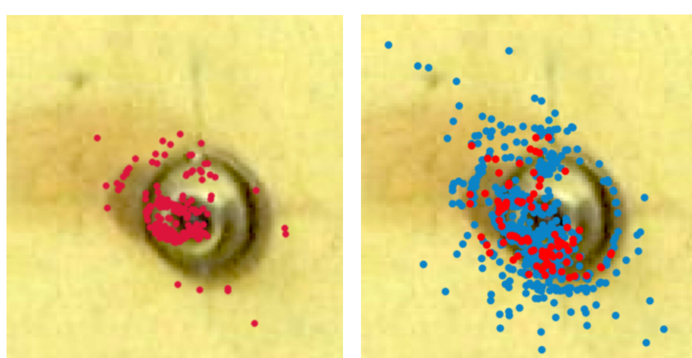


Fig. 7: Still frames from the dataset used in prior work [22] showing how the robot could press buttons. The color marker attached to the fingertip did not have to occlude the functional component to press it.



(a) Auditory events (b) Button press locations

Fig. 8: Spatial localizations used for training the visual model. (a) auditory events in visual space used for the visual model in [22]. (b) button press locations that ring (red points) and did not ring (teal points) used for training the visual model in this work.

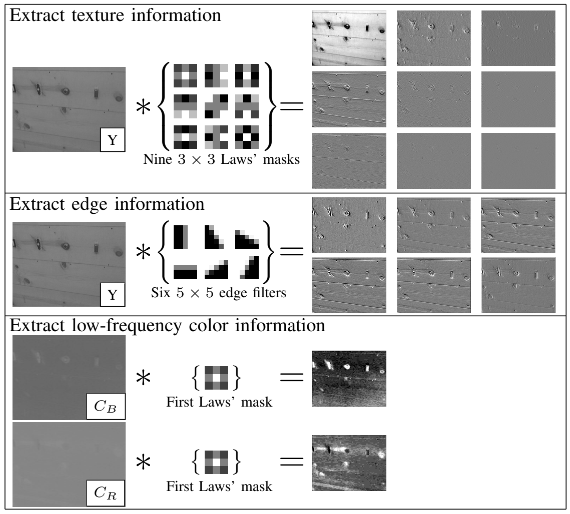


Fig. 9: Illustration of the convolutions used to extract the visual features. First, the raw RGB image is converted into YC_bC_r. Next, the resulting channels are convolved with convolution masks to extract texture, edge, and low-frequency color information. The resulting 17 images are used to compute the features. The convolved images were clamped at the 1st and the 99th percentile for better visualization. See text for more details.

Visual features: As the robot pressed buttons, it built the visual model for detecting them. These visual features were first described in [36] for detecting grasp points in novel objects. Subsequently they were used in our previous work [22] for detecting buttons. In this work the features remain the same, however, the process of computing them was refined to meet the demands of the enhanced framework described in this paper. In particular, features are now computed in real-time so that the robot can use them to learn during exploration rather than offline.

The visual features extract texture, edge, and low-frequency color information from vision. First, the image is converted to YC_bC_r. Next, 17 convolutions are performed: 15 with the Y channel, and 1 with C_b and C_r channels. Texture information is extracted from the Y channel by convolving it with nine 3×3 Laws' masks [37]. Similarly, edge information is extracted from the Y channel by applying six 5×5 oriented edge filters. Low-frequency color information is extracted from C_b and C_r channels using the first Laws' mask. Fig. 9 illustrates this procedure, which results in 17 convolved images. These images are squared before the visual features are extracted from them.

To compute the visual features, the 640×480 image was partitioned 10×10 pixel patches. For a patch p, the 459-dimensional feature vector was computed. The first 17 elements of this feature vector were obtained by summing up the 100 pixel values corresponding to p in the 17 convolved images. In addition to this, the features for neighbouring 24 patches in a 5×5 patch window centered at p is included in the feature vector. This gives $24 \times 17 = 408$ more elements in the feature vector.

Information from two extra spatial scales is included in the 459-dimensional feature vector to capture information about more global properties of the image. For the first scale, the raw 640×480 RGB image is downscaled by 3x to 214×160 pixels. First, the operations illustrated in Fig. 9 are applied to the scaled image to produce the 17 convolved image for this spatial scale. Next, an image patch of size roughly 10×10 pixels, centered at the original patch p, is identified and the its pixels are summed up to produce 17 more elements in the feature vector. The procedure is repeated for another spatial scale, with a zoom factor of 9x.

The contents of the 459-dimensional feature vector computed for a 10×10 pixel patch in the original RGB image can be summarized with a formula:

$$\left(\underset{\text{patch}}{1} + \underset{\text{neighbors}}{24} + \underset{\text{scales}}{2}\right) \times \underset{\text{convolutions}}{17} = \underset{\text{features}}{459}\,.$$

Real-time visual pipeline: The learning framework introduced in this paper requires simultaneous training of a variety of models, including the visual model. This requirement sets this work apart from prior work [22], where the visual model was trained offline, in batch mode. The visual features had to be computed in real time during the experiment because the visual model learned from these features drove the robot – i.e., the robot used it to select where to push next.

To fulfill these requirements, the old pipeline [36] [22] was rewritten in C to run natively. Next, OpenMP was used used to distribute computations across multiple CPU cores. At this point, the pipeline ran at 3 Hz – i.e., it computed 459-dimensional feature vectors for each 10×10 pixel patch in the 640×480 image 3 times per second. For further performance enhancements, the most computationally expensive stage of the pipeline – image convolutions – were reimplemented on GPU [24]. The massively parallel architecture of the GPU with its large number of cores significantly improved the performance. On the system that controlled the robot, the visual pipeline runs at 15 Hz. This performance was sufficient to computed the visual features in real-time from the video recorded by the camera in the robot's eye.

Visual model: The visual model was trained from 459-dimensional feature vector computed for a 10×10 pixel patch in the image and a binary class label for this patch. The class label was positive when the number of presses in the patch that triggered the buzzer exceeded the number of presses that didn't. If the patch was never pressed, it was not used for training.

The goal of the visual model was to predict the probability of belonging to a button's functional component given only the feature vector for an image patch. In prior work [22] [36], a Logistic Regression-based classifier was used. The proposed learning framework illustrated in Fig. 5 requires incremental real-time updates to the visual model. To achieve this goal, several options were evaluated. Currently, a Logistic Regression-based classifier is planned to be used to maintain continuity with previous work.

Goal III : Phase Transitions

Phase Transitions: Phase transitions occur in order to adjust micro-instability in exchange for reduction of macro-instability [38]. This pattern of exchanging localized entropy for a reduction in overall system entropy is useful because it can be extended to a host of synthetic and biological systems. Furthermore, phase transitions can typically be modeled algorithmically because of the predictable exchange of localized instability for maximum system stability. Robinson, et al. [38] give the example that piles of sand will naturally self-organize into the optimal angle of declination and go on to discuss how phase transitions can be thought of as natural phenomenon within all self-organizing systems.

Development: Various models of human development exist and range from the Metaphysical to the Psychofreudian to the archetypal to the biological to the social; many forms of developmental models exist and each model type varies in epistemological foundation. One sort of grounded approach is one that can be examined through the use of logical positivism. Here, the focus is on the examination of models that were created based upon examining quantifiable changes in human development over time. Three forms of easily verifiable changes that usually occur within human development occur within memory, within learning and within behavior. Human development across time, therefore, will be characterized by the associated differences in memory, learning and behavior before and after a developmental period has been achieved. Explaining why humans develop so quickly at some points compared to others is difficult and usually results in suggesting that the brain has grown in sophistication enough to allow for the new growth to occur, or that certain other critical environmental events happened to the individual in order to cause the drastic growth. Structural-functional arguments for human development have long been sought after, but have never been more than casually tied to actual human development. Critical environmental conditions can make a big difference in development, but environmental conditions alone cannot explain what happens within the internal milieu during a period of dynamic development.

Structural-functional approach: Phrenology used to be considered a verifiable and useful science that attempted to pair the etiology of unique qualities of being human to various regions of the brain that were distinct from regions of animal brains. Unfortunately for Phrenologists, it was later shown that various regions of the brain could account for a multitude of behaviors. More recently, attempts have been made to pair changes in human development with changes in brain development. This makes sense because, after all, without a brain an individual is not very intelligent and probably will not develop to their full potential. However, this approach is also antiquated and does not take into account the complexities of human intelligence. The cognitive revolution moved the field towards an exploration of human memory. It was discovered that there were distinct 'types' of memory that may or may not have structural origins. Eventually, connectionist theories of memory and theories of higher order thought (HOT) were developed that called into question the structural organs of memory itself. Modern neuroscience has suggested that, in a sense, all memory may originate from the same neuronal function. There may be different areas of the brain that store information differently, resulting in the experience of phenomenologically distinct memory types, but the biological function of the neurons in these regions is not substantially different. Finally, it has recently been demonstrated that phase transitions occur within the neuronal connections inside the human brain; mean field theories strive to model phase transitions by looking at structure across neurons and across the brain [38]. If structural differences alone cannot account for the differences in memory, learning and behavioral outcomes, then what can be used to explain how these distinct capacities develop? Furthermore, how can these capacities be explained while maintaining focus upon a logical positivist approach. Physicists have taught that entropy naturally increases within a system. Neuroscientists have taught that all brain regions exhibit similar neuronal functions but that some regions may be responsible for the boosting' of intellectual development. If the brain is a system, bounded by entropy as quantified by uncertainty, that follows one neuronal function, then perhaps this function is that of reducing uncertainty.

Prior art in Psychology: If one were to adopt the idea that uncertainty drove human development and the development of intelligence, then how would that idea relate to existing research on memory, learning and behavioral outcomes? Concepts from Social Psychology, Cognitive Science, Behaviorism, Neuroscience and even Therapeutic Psychology can be explained through the lens of uncertainty. Across these disciplines, the examination of human intelligence is common, but two concepts in particular appear important to human development across all of these disciplines: the "sense of self" and "concept inference". Inference of concepts comes from the ability to classify perceptions. Here perceptions are distinguished from stimuli since perceptions are stimuli that have already been non-consciously pre-processed. Concept formation requires distinguishing various attributes of a one perception from the attributes of other perceptions. What results is the classification of perceptions based upon their perceptual attributes. Classification is an important characteristic of intelligence that is examined, in depth, by many branches of psychology. Behaviorism is interested in the association of one concept with another and the connection that a stimuli can have on the participants' behavior/response. Behaviorism is concerned with learning curves and extinction rates which measure the persistence of learning over time. In both cases, it could be said that a phase transition occurs within the participant when they learn to associate or disassociate. In cognitive psychology, categorization is often used as dependent variable to measure changes within the internal milieu. Exemplar theory and prototype theory both agree that the human brain creates 'models' of the world that new perceptions are compared against for categorical purposes. Piaget mentioned that humans tend to either absorb or 'assimilate' information into fresh categories, or that humans tend to fit or 'accommodate' into existing constructs. If this is the case, then distinctive information becomes important since it can be used to simply and efficiently separate prior distributions from conditional distributions. The principle of saliency is well established and falls in line with the tendency to use unique information to categorize. Using salient features would reduce the amount of information (as measured by information compression) to categorize new perceptions into existing constructs. This makes sense and can help explain humans' tendency to create stereotypical classifications: humans are simply classifying concepts/events based upon the one attribute thought to give the most mutual information about the new perception based upon the individual's previous knowledge of the concept/event. Information gain can also be used to explain the inability of humans to perceive differences that fall below a critical threshold known as the Just Noticeable Difference. The familiarity effect relates to humans' positive affect bias towards information that is familiar and this may relate to information gain because it takes less effort to process information that is familiar. A reduction in cognitive effort may be an intrinsic motivator and may act as a natural reward. Signal detection theory relates to the effort needed to detect the difference between two concepts. If two concepts are too similar they will commonly be classified as the same concept or, at the least, will not be distinguished from each other. Cognitive psychology is concerned with the likelihood of committing a Type I or a Type II error and certain social systems (such as the justice system) can be defined by their predisposition towards a particular form of error. The mutual information between two categories would, using information gain theory, determine the tendency to commit Type I and Type II error. Self-detection and a sense of self are vital to the development of a self-identity. The ability to detect one's self has been shown to arise in humans as young as four months, but this sense of self is not immediately complete [39]. Self-detection and self-identity are not distinct to humans, but are concepts exhibited in other animals considered to be intelligent, such as great apes, dolphins and elephants. Understanding the self will reduce the amount of information required to process ones' environment. Tool use requires some self-concept because the tool user will need to understand what their body is capable of doing in order to know how to use a tool. The ability to transfer skills learned during the performance of one task to the performance of another task requires understanding that the same person, the self, was the actor committing both activities. Thus, generalizing tool use from one situation to another also requires a sense of self. Finally, the development of a sense of self-identity is important for the development of the ability to perceive others as individuals. Perceiving others as individuals is necessary for the development of society since social interactions often require understanding the needs of others. In both categorization and self-detection, the principle of information gain could be applied to explain the development of these processes. An information gain approach would utilize the minimal information necessary to separate a signal from a prior distribution. As mentioned earlier, a characteristic of reduction of entropy within a system is a phase shift/ transition. These transitions can be measured physically through changes in state or through changes in behavior and/or interaction patterns or through measuring differences in memory and learning outcomes. Measuring transitions here requires inter-observer reliability to be high and it also requires that the instruments used be standardized and appropriate for measuring changes within a system. Phase transitions can also sometimes be quantified through measuring artifacts and changes in temporal signals. Examples of temporal signals include wave patterns within a physical system, input patterns into a computational system and electroencephalograph (EEG) patterns within human systems. There are clearly a number of ways of measuring phase transitions. Here, a method will be used that can reliably compare phase transitions within the human system to transitions within a synthetic, robotic system. This method, discussed in depth in the methods section, will utilize a traditional approach to information processing that strives to compress, classify and ultimately categorize information. Such an approach has been used both in robotics and in analyzing EEG data. Furthermore, this is thought to provide the most consistent results for analysis.

Putting together the pieces: The minimum information required to classify an event would be that information that distinguishes it from an existing distribution (read: existing concept, category or idea). Thus, stimuli come in and are all processed serially through some innate discretization, dimensional- reduction, pre-processing, invariant pattern/algorithm. For simplicity, this function will be referred to as the neuronal function. Processing stimuli in such a way allows for the comparison of information from different modalities to be tied into a contiguous conception of reality. As intelligence arises, it develops through particular stages. Not all capacities will be immediately present. These capacities are probably absent at first and arise as a result of discontinuous development within an intelligent system. Here, two distinct periods of dynamic development will be examined: the development of the "sense of self" and that development of "concept inference". The development of both of these distinctive qualities of intelligence can be tied to differences in memory, learning and behavioral outcomes. The human infant is a ball of potential that, at birth, comes pre-wired with the minimal biological structures and the minimal information necessary in the brain's 'innate invariant neuronal patterns' for the developmental lifecycle for the normal human to occur. Of course, a host of mitigating social and developmental factors can vastly change the realized/actualized potential of an individual; however, what is clear, is that we all start in very similar forms. At birth, an infant is a bundle of reflexes that lead it to certain 'survival instincts', imprinting potentials and, importantly, exploratory behaviors. These exploratory behaviors are all that are necessary for the neuronal function to begin to lead the human down the path of development towards the path of its innate potential.

Hypotheses: It is expected that phase transitions will occur within a synthetic system where that system is utilizes uncertainty-driven exploration. Furthermore, it is thought that robotic development that attempts to reduce entropy within a system will demonstrate phase shifts at similar points to where humans exhibit phase transitions. Two periods of development will be of interest here: "concept inference" and "self detection". If uncertainty driven exploration and information gain can explain the discontinuous periods of human development, then a synthetic system built to utilize these techniques should exhibit similar characteristic phase shifts in developmental patterns.

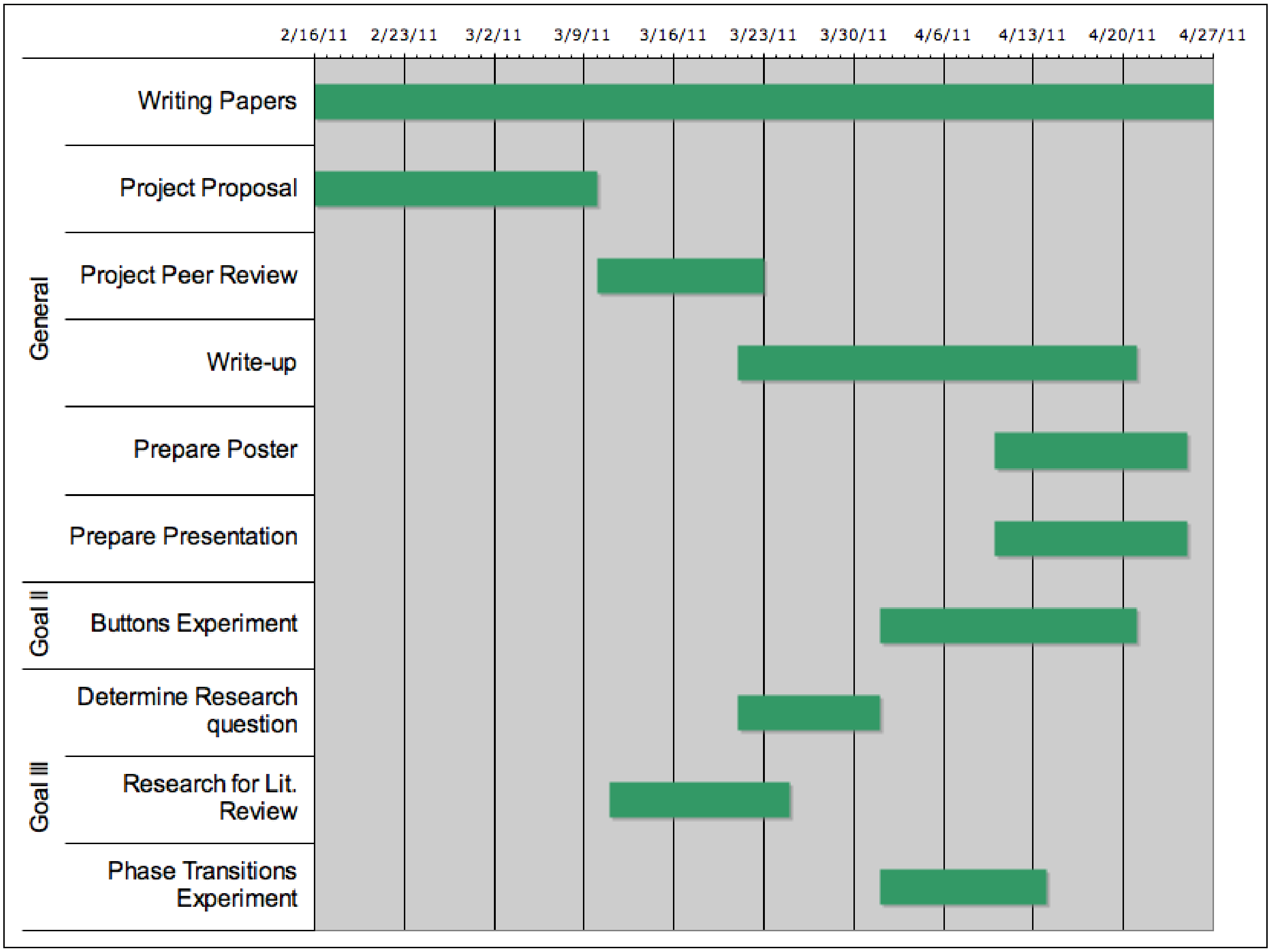


Fig. 10: Project schedule.

SCHEDULE

Preliminary project schedule is shown in Fig. 10. The key activity for the success for the project is writing papers. Different goals of the project are at different levels of readiness for publication, therefore the amount of time designated for goal-specific milestones varies depending on a goal. Research Goal III is currently the least developed, therefore the bulk of planning effort was devoted to that goal.

STATEMENT OF CONFIDENCE

Our team has a unique combination of experience that makes us uniquely qualified to engage in investigating this topic. In other words, the the estimated lower bound on the level of confidence in the team is very high and the null hypothesis can be rejected outright.

Ryan's undergraduate career was at Drake University where he was always in the idea of Artificial Intelligence. It was during Ryan's first Psychology class that Ryan's interest in intelligence really blossomed and he began to realize that artificial intelligence can be modeled through many different methods. Unlike what he had originally pictured, it did not solely involve using computers. he also grew interested in social thinking, group processes and business structures. Ryan became fascinated with the idea of consciousness. Ryan's current interests are a marriage of his prior intellectual interests; he is currently interested in studying the nature of complex systems through creating experimental models of such systems. Ryan believes it is possible to model natural processes using a variety of techniques that model behaviors ranging in complexity from simplest cause-effect, correlative relationship models to more complex nonlinear models of novel phenomenon. Ryan has recently begun conducting research in creating models of intelligent systems through using concepts from A.I. and Machine Learning to organize and structure knowledge systems.

Vladimir received his Bachelors degree in Applied Mathematics from Donetsk National University (Donetsk, Ukraine) in 2004. He is currently a PhD student in Computer Engineering and works at the Developmental Robotics Laboratory at Iowa State University, Ames. His research interests are in the areas of developmental robotics, human-computer interaction, computational perception, and machine learning.

Vladimir is really interested in this project it is *great fun*.

REFERENCES

[1] R. Sutton, *Verification, The Key to AI*, http://webdocs.cs.ualberta.ca/~sutton/IncIdeas/KeytoAI.html.
[2] R. Brooks, *The AI Journey: Future Challenges*, 2004, http://www.ai.rutgers.edu/aaai25/brooks.htm.
[3] P. McCorduck, *Machines Who Think*, 2nd ed. Peters, Ltd., 2004.
[4] J. Watson, "Detection of self: The perfect algorithm," in *Self-awareness in Animals and Humans: Developmental Perspectives*. Cambridge University Press, 1994, pp. 131–148.
[5] J. Piaget, *The origins of intelligence in children*. Intl. Univ. P., 1952.
[6] E. Hutchison, *Dimensions of human behavior. The changing life course*. Sage Publications, 2003.
[7] P. Hauf and G. Aschersleben, "Action-effect anticipation in infant action control," *Psych. Research*, vol. 72, no. 2, pp. 203–210, 2008.
[8] J. Kelso, *Dynamic patterns: the self-organization of brain and behavior*. MIT Press, 1995.
[9] A. Phillips and P. Robinson, "A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system," *J Biol Rhythms*, vol. 22, pp. 167–179, April 2007.
[10] E. Thelen and L. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action (Cognitive Psychology)*. MIT Press, 1996.
[11] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Vision Research*, vol. 49, no. 10, pp. 1295–1306, 2009.
[12] A. Stoytchev, "Self-detection in robots: a method based on detecting temporal contigencies," *Robotica*, vol. 29, pp. 1–21, 2011.
[13] E. Klingbeil, B. Carpenter, O. Russakovsky, and A. Ng, "Autonomous operation of novel elevators for robot navigation," in *Proc. of ICRA*, 2010, pp. 751–758.
[14] J. Miura, K. Iwase, and Y. Shirai, "Interactive teaching of a mobile robot," in *Proc. of ICRA*, 2005, pp. 3378–3383.
[15] K.-T. Song and T.-Z. Wu, "Visual servo control of a mobile manipulator using one-dimensional windows," in *Proc. of Industrial Electronics Society*, vol. 2, 1999, pp. 686–691.
[16] R. Katsuki, J. Ota, T. Yamura, T. Mizuta, T. Arai, T. Ueyama, and T. Nishiyama, "Handling of objects with marks by a robot," in *Proceedings of IROS 2003*, October 2003, pp. 130–135.
[17] T. Deyle, H. Nguyen, M. Reynolds, and C. Kemp, "RFID-guided robots for pervasive automation," *Pervasive Computing*, vol. 9, no. 2, pp. 37–45, 2010.
[18] H. Nguyen, T. Deyle, M. Reynolds, and C. Kemp, "PPS-tags: Physical, Perceptual and Semantic tags for autonomous mobile manipulation," in *Proc. of the IROS Workshop on Semantic Perception for Mobile Manipulation*, 2009.
[19] A. Thomaz, "Socially guided machine learning," Ph.D. dissertation, Massachusetts Institute of Technology, 2006.
[20] C. Breazeal and A. Thomaz, "Learning from human teachers with socially guided exploration," in *Proceedings of ICRA*, 2008, pp. 3539–3544.
[21] J. Lieberman, "Teaching a robot manipulation skills through demonstration," Master's thesis, Massachusetts Institute of Technology, 2004.
[22] V. Sukhoy and A. Stoytchev, "Learning to detect the functional components of doorbell buttons using active exploration and multimodal correlation," in *In Proceedings of the 2010 IEEE-RSJ Conference on Humanoid Robots*, Nashville, TN, 2010, pp. 572–579.
[23] D. Luebke, M. Harris, N. Govindaraju, A. Lefohn, M. Houston, J. Owens, M. Segal, M. Papakipos, and I. Buck, "GPGPU: general-purpose computation on graphics hardware," in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, 2006.
[24] V. Podlozhnyuk, "Image convolution with CUDA," *NVIDIA Corporation white paper*, Jan 2007.
[25] S. Heymann, K. Müller, B. Smolic, A. Fröhlich, and T. Wiegand, "SIFT implementation and optimization for general-purpose GPU," in *WSCG*, 2007, pp. 317–322.
[26] T. Terriberry, L. French, and J. Helmsen, "GPU accelerating speeded-up robust features," 2008.
[27] E. Gibson, "Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge," *Annual review of psychology*, vol. 39, no. 1, pp. 1–42, 1988.
[28] P. Hauf, G. Aschersleben, and W. Prinz, "Baby do-baby see!: How action production influences action perception in infants," *Cognitive Development*, vol. 22, no. 1, pp. 16 – 32, 2007.

[29] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, July,October 1948.

[30] A. Stoytchev, "Robot tool behavior: A developmental approach to autonomous tool use," Ph.D. dissertation, College of Computing, Georgia Institute of Technology, August 2007.

[31] S. Vijayakumar, A. D'Souza, and S. Schaal, "Incremental online learning in high dimensions," *Neural Computation*, vol. 17, no. 12, pp. 2602–2634, Jan 2005.

[32] E. Gibson, *Principles of Perceptual Learning and Development*. Appleton-Century-Crofts, 1969.

[33] J. Schmidhuber, "Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts," *Connection Science*, vol. 18, no. 2, pp. 173–187, June 2006.

[34] V. Sukhoy, J. Sinapov, L. Wu, and A. Stoytchev, "Learning to press doorbell buttons," in *In Proceedings of the 9th IEEE Intl. Conf. on Development and Learning (ICDL)*, Ann Arbor, MI, August 2010.

[35] S. Chen and J. Goodman, "An empirical study of smoothing techniques for language modeling," in *Proceedings of the 34th annual meeting on Association for Computational Linguistics (ACL)*, 1996, pp. 310–318.

[36] A. Saxena, J. Driemeyer, and A. Y. Ng, "Robotic grasping of novel objects using vision," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.

[37] K. Laws, "Textured image segmentation," Ph.D. dissertation, University of Southern California, 1980.

[38] P. Robinson, C. Rennie, A. Phillips, J. Kim, and J. Roberts, "Phase transitions in physiologically-based multiscale mean-field brain models," in *Modeling Phase Transitions in the Brain*, ser. Springer Series in Computational Neuroscience, D. A. Steyn-Ross and M. Steyn-Ross, Eds. Springer New York, 2010, vol. 4, pp. 179–201.

[39] P. Rochat, "Five levels of self-awareness as they unfold early in life," *Consciousness and Cognition*, vol. 12, no. 4, pp. 717 – 731, 2003.

[40] G. Miller, "Note on the bias of information estimates," in *Information Theory in Psychology: Problems and Methods*, 1955, pp. 95–100.

[41] L. Paninski, "Estimation of entropy and mutual information," *Neural Comput.*, vol. 15, pp. 1191–1253, June 2003.

[42] A. Antos and I. Kontoyiannis, "Convergence properties of functional estimates for discrete distributions," *Random Structures & Algorithms*, vol. 19, no. 3-4, pp. 163–193, 2001.

[43] G. Oehlert, "A note on the delta method," *American Statistician*, vol. 46, no. 1, pp. 27–29, 1992.

[44] C. McDiarmid, "On the method of bounded differences," in *Surveys in combinatorics*, Siemons, Ed. Cambridge University Press, 1989, pp. 148–188.

Appendix

Mathematical Apparatus for Self-detection

A number of theorems that allow to establish statistical properties of the information sources associated with different features are formulated. The proofs are omitted.

The process of self-detection is a statistical test for $I > 0$ using the estimated distributions for M and D. To perform the test, I is estimated by an estimator $\hat{I}$ from the data. $\hat{I}$ is computed using the entropy estimators $\hat{H}$ according to the formula (2):

$$\hat{I}(M; D) = \hat{H}(M) + \hat{H}(D) - \hat{H}(M, D),$$

For a discrete random variable $X \in \{x_1, \ldots, x_m\}$, the entropy can be estimated from N measurements using the histogram with the bin counters $c_1, \ldots, c_m$ with $c_i \in \mathbb{Z}$, $c_i \geq 0$, and $c_1 + \ldots + c_m = N$. One possible estimator uses (3) with the probability estimates from the histogram:

$$\hat{H}_{\text{MLE}} = -\sum_{i=1}^{m} \frac{c_i}{N} \log \frac{c_i}{N}.$$

$\hat{H}_{\text{MLE}}$ is the *maximum likelihood* estimator – i.e., it maximizes the probability of observing a histogram given that $H = \hat{H}_{\text{MLE}}$. This probability is the *likelihood function*. $\hat{H}_{\text{MLE}}$ maximizes it, which is why it is the maximum likelihood estimator.

It may seem counter-intuitive, but $\hat{H}_{\text{MLE}}$ is not the best possible estimator. In particular, it can be expected to underestimate the true value of H that is not known. In other words, it has a negative negative bias. A better estimator, proposed by Miller [40], applies a correction to $\hat{H}_{\text{MLE}}$ to improve it.

$$\hat{H}_{\text{MM}} = \hat{H}_{\text{MLE}} + \frac{m-1}{2N} = -\sum_{i=1}^{m} \frac{c_i}{N} \log \frac{c_i}{N} + \frac{m-1}{2N}.$$

The key idea is to bind the statistical properties of $\hat{H}_{\text{MM}}$ with the condition $I > 0$ using the bounds on the bias of $\hat{H}_{\text{MM}}$. The following propositions sketch the overall argument. The proofs are skipped. The overall argument is structured following [41] and [42]. The argument is the framework of the delta method [43] in Statistics.

Proposition 1: Let $p = (p_1, \ldots, p_m)$ and $\hat{p} = (\hat{p}_1, \ldots, \hat{p}_m)$ such that $p_1 + \ldots + p_m = \hat{p}_1 + \ldots + \hat{p}_m = 1$. Also suppose that if $p_i = 0$ then $\hat{p}_i = 0$. Then, $H(\hat{p})$ can be written as follows:

$$H(\hat{p}) = H(p) - \sum_{i=1}^{m} (\hat{p}_i - p_i) \log_2 p_i + \mathbf{R}, \tag{4}$$

where the remainder term $\mathbf{R}$ is

$$\mathbf{R} = H(\hat{p}) - H(p) + \sum_{i=1}^{m} (\hat{p}_i - p_i) \log_2 p_i. \tag{5}$$

Proposition 2:

$$\mathbf{R} = -D_{\text{KL}}(\hat{p}||p).$$

Property 1: If $\hat{p}_i = \frac{c_i}{N}$, then

$$E\left[\chi^2(\hat{p}, p)\right] = \frac{m-1}{N}.$$

Proposition 3:

$$-\log_2\left(1 + \frac{m-1}{N}\right) \leq \text{Bias}(\hat{H}_{\text{MLE}}) \leq 0.$$

Proposition 4:

$$-\log_2\left(1 + \frac{m-1}{N}\right) + \frac{m-1}{N} \leq \text{Bias}(\hat{H}_{\text{MM}}) \leq \frac{m-1}{N}.$$

The following proposition is a famous probabilistic inequality proven by McDiarmid [44].

Proposition 5: **McDiarmid Inequality.** Suppose $X_1, \ldots, X_N$ are independent random variables on $\mathbb{R}$, and $\hat{F} : \mathbb{R} \rightarrow R$ satisfies

$$\sup_{x \in \mathbb{R}^N, x'_j \in \mathbb{R}} \left|\hat{F}(x) - \hat{F}(x')\right| \leq c_j, \quad 1 \leq j \leq N, \tag{6}$$

where

$$\begin{aligned} x &= (x_1, \ldots, x_N), \\ x' &= (x_1, \ldots, x_{j-1}, x'_j, x_{j+1}, \ldots, x_N). \end{aligned}$$

Then, for any $\varepsilon > 0$,

$$\Pr\left(\hat{F}(X_1, \ldots, X_N) - E\left[\hat{F}(X_1, \ldots, X_N)\right]\right) \leq 2e^{-2\varepsilon^2 / \sum_{j=1}^{N} c_j^2}$$

In other words, if there is a bound on the change of the estimator when a single value in the sample changes, then the probability of the estimator deviating from its estimate by more than ε decreases exponentially as a function of ε.

The following proposition that gives one way to get the p-value for the null hypothesis that a feature is not self. It can be proven using the diversity representation of the histogram and McDiarmid inequality [41] [42].

Proposition 6: For $\hat{H}_{\text{MLE}}$ estimated from N measurements,

$$\Pr(|\hat{H}_{\text{MLE}} - E\hat{H}_{\text{MLE}}| > \varepsilon) \leq 2\exp\left(\frac{-N\varepsilon^2}{2\log_2^2 N}\right).$$

$$\Pr(|\hat{H}_{\text{MM}} - E\hat{H}_{\text{MM}}| > \varepsilon) \leq 2\exp\left(\frac{-N\varepsilon^2}{2\log_2^2 N}\right). \tag{7}$$

Proposition 7:

$$p \leq 6\exp\left(\frac{-N(a/3)^2}{2\log_2^2 N}\right) = p^*, \tag{8}$$

where $p = \Pr(\hat{I} = \hat{I}_{\text{DATA}}|I = 0)$ is the p-value for the null hypothesis $I = 0$.