UNIVERSITY OF BREMEN
INSTITUTE FOR
ARTIFICIAL INTELLIGENCE

Causal Observational Learning in Natural Settings for Robot Analysis of Human Interaction

Alexander Wich

Prüfer der Dissertation:

1. PD Dr. Holger Schultheis
   Universität Bremen
2. Prof. Michael Beetz, PhD
   Universität Bremen
3. Prof. Dr. Karinne Ramirez-Amaro,
   Chalmers University of Technology

Fachbereich: Informatik und Mathematik

Kolloquium: 14.10.2022

Abstract

Learning from observing others is a powerful human competence that robots lack. People can learn by analyzing others' interactions under almost any conditions because of reasoning capabilities, such as inferring causal relationships, predicting, adapting, and imagining. These capabilities allow people to attain causal understanding and harness observations for their benefit, such as anticipating others' behaviors, rehearsing them under different conditions, and imagining behavior not seen before. Possessing the four inference capabilities is essential for observational learning, but robots do not fully support them and require quality inputs to render inferences feasible.

To explore the viability of robots analyzing others' interactions in natural conditions, in this dissertation, we focus on formalizing human observational learning and then challenge and evaluate its potential, such as inferring hand behavior from everyday activities. The proof of principle comprises the identification of a formalism covering core capabilities of human observational learning, the instantiation of a framework serving as the object of study, the specification of a scenario that challenges its potential, the verification of proper functioning, and the utility determining inferences remain meaningful. The results show inferences can operate outside the formalism's functional design despite atypical conditions and breaking a foundational assumption. Moreover, under such conditions, inferences manage to find causal relationships which happen to be meaningful. By introducing this proof of principle and value, we know that robots equipped with the inference formalism operating outside the functional design do not necessarily fail and could provide valuable inferences.

Zusammenfassung

Das Lernen aus der Beobachtung anderer ist eine wichtige menschliche Fähigkeit, die Robotern fehlt. Menschen können lernen, indem sie die Interaktionen anderer unter fast allen Bedingungen analysieren, da sie über Denkfähigkeiten verfügen, mit denen sie z. B. kausale Beziehungen ableiten, vorhersagen, anpassen und sich vorstellen können. Diese Fähigkeiten ermöglichen es Menschen, kausale Zusammenhänge zu verstehen und Beobachtungen zu ihrem Vorteil zu nutzen, z. B. indem sie das Verhalten anderer vorhersehen, unter verschiedenen Bedingungen nachahmen und sich sogar Verhalten vorstellen können, das sie noch nie gesehen haben. Das Vorhandensein der vier Inferenzfähigkeiten ist für das Beobachtungslernen unerlässlich. Jedoch unterstützen Roboter sie nicht vollständig und benötigen qualitativ hochwertige Eingaben, um Inferenzen durchführen zu können.

Um die Machbarkeit der Roboteranalysen von Interaktionen unter natürlichen Bedingungen zu untersuchen, konzentrieren wir uns in dieser Dissertation auf die Formalisierung des menschlichen Beobachtungslernens, um dann dessen Potenzial zu testen und zu bewerten, z. B. das Ableiten von Handverhalten aus alltäglichen Aktivitäten. Der Prinzipienbeweis umfasst die Identifizierung eines Formalismus, der die Kernfähigkeiten des menschlichen Beobachtungslernens abdeckt, die Instanziierung eines Frameworks, das als Studienobjekt dient, die Spezifizierung eines Szenarios, das sein Potenzial herausfordert, die Verifizierung der Funktionstüchtigkeit und die Nützlichkeitsbestimmung von abgeleiteten Schlussfolgerungen. Die Ergebnisse zeigen, dass Schlussfolgerungen außerhalb des funktionalen Designs des Formalismus trotz atypischer Bedingungen und der Verletzung einer Grundannahme angewandt werden. Darüber hinaus gelingt es den Inferenzen unter solchen Bedingungen, kausale Beziehungen zu finden, die tatsächlich sinnvoll sind. Durch die Einführung dieses Prinzipien und Wertbeweises zeigen wir, dass Roboter, die mit dem Inferenzformalismus ausgestattet sind und außerhalb des funktionalen Designs arbeiten, nicht notwendigerweise scheitern sondern sogar wertvolle Inferenzen liefern können.

Acknowledgments

Words cannot express my gratitude to my supervisor, PD Dr. Holger Schultheis, who generously provided his invaluable guidance, encouragement, patience, and feedback throughout the stages of this dissertation. I also could not have undertaken this journey without my co-supervisor, Prof. Michael Beetz PhD, who granted me the opportunity of this journey. Joining the Institute for Artificial Intelligence (IAI) under his supervision was a profound experience and inspiring for me. Prof. Dr. Karinne Ramirez-Amaro, thanks for the valuable discussions and feedback. Additionally, this endeavor would not have been possible without the generous support from BECAL, who financed my research.

I am thankful to members of the IAI at the University of Bremen who contributed to the design and architecture of the robot platform this dissertation builds upon, Andrei Haidu, Asil Bozcuoğlu, Daniel Beßler, Jan-Hendrik Worch, Mihai Pomarlan, Gayane Kazhoyan, Sebastian Koralewski, Patrick Mania, Jörn Syrbe, Georg Bartels, Simon Stelter, Ference Benczédi, Mareike Picklum, Daniel Nyga, especially for the patience and assistance from Andrei Haidu, Daniel Beßler and Asil Bozcuoğlu.

I am grateful to my dear office friends, Lisset Salinas, Feroz Siddiky, and Michaela Kümpel, for what I learned with you and for inspiring me along the way. Words cannot express my gratitude to Lisset Salinas, my faithful companion throughout the entire journey.

I extend my sincere thanks to Heinz Riener, Christian Müller, Fereshta Yazdanil, Frank Ruge and Luis Carlos Díaz, who provided valuable comments and help editing my manuscript.

I express my deepest gratitude to my family, Patricia, Andreas, Margarita, Fritz, Christian, and Elisa, for keeping my spirits high and their unconditional support and encouragement throughout this research.

22 June, 2022,

Alexander Wich

Contents

List of Figures

List of Tables

Chapter one

Introduction

Learning indirectly from others' interactions is a powerful human competence. Learning via others empowers an observer to acquire valuable content quickly, such as know-how others' spent years perfecting and thus minimize risks of injuries when attempting to perform them. Learning from observations is unrestricted to physical proximity (works from anywhere), and allows acquiring abstract content, thus transcending the physical domain (is broad).

Observational learning involves analyzing others' interactions and eventually making the most out of observations in any present condition or hypothetical context not seen before. The social cognitive theory [Ban86, Ban77] establishes that observational learning allows humans to expand their knowledge at any distance in time and space by simply observing others.

However, there is a lack of approaches formalizing the core capabilities underlying human observational learning in robots. Also, the learning conditions where machines train demand authority over the inputs when observers have no control over observations. For example, data-driven approaches can find associations but not causal relationships and require large numbers of quality samples when such is rarely the case in observational settings.

This dissertation aims to identify, instantiate, and evaluate an approach that formalizes observational learning for its operation in natural scenarios where input observations are not controllable.

This chapter introduces the dissertation by first describing the background, followed by the research problem, the aim and objectives, and the significance and limitations.

1.1 Background

According to the Social Cognitive Theory (SCT), people support two acquisition modalities; one is direct experience (imitation) and indirect (observing others) [BGM66]. Both enable people to improve but are fundamentally different.

Learning by imitation is narrow in contrast to observation. Imitation comprises duplication as the end in itself, whereas observation transcends the act of mimicry [Dew16]. The theory in SCT defines indirect learning as

> "Observational learning (also known as vicarious learning, social learning, or modeling) is a type of learning that occurs as a function of observing, retaining, and replicating novel behavior executed by others. It is argued that reinforcement has the effect of influencing which responses one will partake in, more than it influences the actual acquisition of the new response" [Ban77]

Perhaps the most notorious distinction of observational learning is that acquisition can occur without exercising the act of imitation (enacting is optional). Another distinction is that observations – actions and consequences – are mentally modeled with the chance (option) of replicating them eventually (deferred rehearsal). Observation is unbounded to physical proximity (works from any distance). Allows learning the feedback directed to others, such as a child observing a parent's praise to another child (reinforcement is optional). Another uniqueness is that learning from observation involves new thinking (innovation), while imitation leverages replication. Table 2.1 contrast side by side both strategies.

	Direct	**Indirect**
Motor execution	required	optional
Execution	immediate	immediate or deferred
Bounded	physical proximity	anywhere
Reinforcement	required	optional
Involves	replication	innovation

Table 1.1: *A comparison of direct experience and indirect.*

Research in artificial intelligence directs a great deal of attention to learning from direct experience (or trial-and-error learning), vastly surpassing learning from indirect observation (or modeling). This is evident by the number of learning settings

developed for each.

The settings that learn from direct experience categorize as [RPCB20] learning with a teacher (type A), by doing (type B), and hybrids combining with a teacher and by doing (type C). Agents that learn with a teacher (type A) are supervised by an instructor (e.g., kinesthetic teaching, imitation learning, etc.). Those learning by doing (type B) are guided via the environment or the rewards encoded into agents (e.g., reinforcement learning, self-play, etc.). Last, those combining implicit and explicit assistance rely on both guidance modalities (e.g., reward-based reinforcement learning, apprenticeship learning, etc.). Beyond distinctions, settings (A, B, C) implicitly or explicitly assist learning.

In settings with a teacher (A), the student's movement is guided by the teacher (kinesthetic teaching) [BCG06], or the learner copies a teacher's behavior (imitation learning) [Sch99]. In learning, by doing (B), the machine gains experience by trial and error (reinforcement learning) [SB98], or the student plays against versions of itself (self-play) [Boz99]. In settings balancing explicit and implicit assistance (C), students imitate teachers and explore by trial and error (reward-based learning from demonstration) [KP10], or learn the teacher's intention and discover control policies on their own (apprenticeship learning) [AN04]. The literature review presented in Chapter 3 covers additional settings.

Unlike the number of different learning settings developed for trial-and-error learning, indirect learning is regarded informally as "learning without labels" [TWS19b] to distinguish the difference from the ones described earlier. This learning setting (type D) is pursued under various names, Learning by Observation [DB07], Observational Learning (OL) [TCH+13], Learning from Observation (LfO) [OMG13], and Imitation Learning from Observation (IfO) [TWS19a].

There are two remarks on the current body of research in the area of indirect learning. Observational learning is approached as if learning from direct experience. First, the same settings for learning from direct experience are also employed for observational learning. Second, the techniques formalizing learning from direct experience are also used to tackle observation. However, these research efforts conflict with what is already known about human observation as SCT advocates.

The settings employed by machines differ from those of humans. The environments where people observe others are natural. Natural because people do not control what they watch. Consequently, an observer cannot enforce the input quantity and quality. These are challenging learning conditions for an observer. Even harder, neither does nature arrange events in a particular manner to ease human learning. Hence, natural interactions are not necessarily intended to be learned. Nevertheless,

people manage to benefit from such natural sources of information. Regardless of learning conditions, people master daily chores (eating, dressing, cooking, etc.).

Besides the natural conditions in observational settings, another commonality is that interactions include the interplay between environment, others, and individual-self (social setting). Unlike other views that emphasize the importance of the environment and the internal states (e.g., behaviorism), the SCT proposes the individual-self as equally important (a.k.a. triadic reciprocal causation model [Ban86]). In this view, individuals are not simply reacting to the environment as they can act independently (without constant guidance). Reinforcement is therefore optional and implicit in the following list.

Natural settings, or vicarious setting (VS), features:

- s_0: inputs are natural, not necessarily intended to be learned
- s_1: inputs include interactions with the environment, others, and individual-self

The observational settings (VS) featuring s_0 and s_1 do not match machine learning settings A, B, C, and D (with a teacher, by doing, combining with a teacher, by doing, learning without labels). The differences render machine settings inadequate to capture the observational scenario. First, machine learning settings carry teaching intentions or require reinforcements (to make learning feasible). Second, these provide quality inputs to ensure quality inferences (correct examples, balanced classes of examples, etc.). Third, artificial settings offer nearly arbitrary learning opportunities (samples in thousands). Overall, VS features conditions closer to the real world than the artificial environment, thus rendering settings (A, B, C, D) inadequate to develop and evaluate observational learning.

Learning from direct experience centers on execution. Observational learning is not about execution but the process preceding it, namely acquisition. To acquire another's behavior means accepting someone's else conducts as one's own. According to SCT, people acquire others' behaviors via modeling. Figure 1.1 distinguishes the four processes undergone to own another's behavior: the observer perceives a model of interest, the retention process stores a model for an eventual mental or physical rehearsal, the production process guides rehearsals and the construction of responses, and the motivational process regulates the execution of modeled behaviors [Ban77].

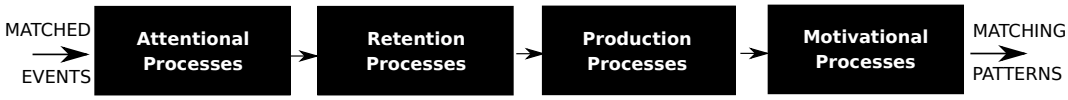


Figure 1.1: *Modeling in Observational Learning.*

Most research on learning from direct experience overlooks the motivational process, which is the center of advanced decisions made before execution. For instance, "What can I expect?", "What happens if?", "What would happen if things turned differently?", "What is the reason?". The motivational process requires a sort of language to support self-reflective dialogues with oneself. The self-reflective capacity of humans enables the auto-evaluation of thoughts and actions [Ban86].

SCT identifies the capabilities supporting modeling in humans to be the following:

> "People gain understanding of **causal relationships** and expand their knowledge by operating symbolically on the wealth of information derived from personal and vicarious experiences. They construct **possible solutions** to problems and evaluate their **likely outcomes**, without having to go through a laborious trial-and-error process. Through the medium of symbols, people can communicate with others at any distance in time and space. The other distinctive human capabilities (self-regulation, self-reflective, vicarious capacity) are founded on this advanced capacity for symbolization." [Ban01]

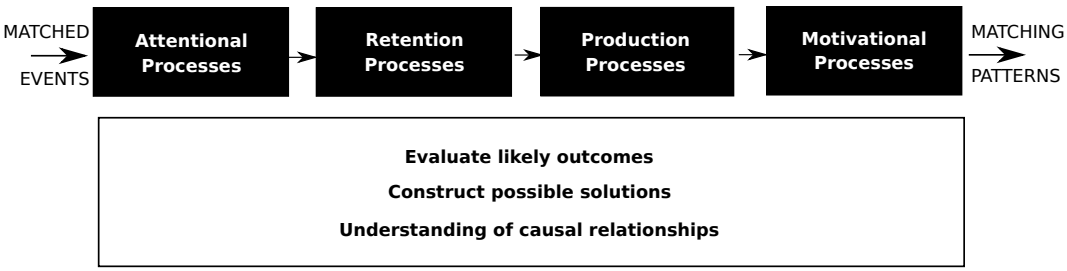


Figure 1.2: *Capabilities in Observational Learning.*

These capabilities allow people to detect faulty thinking (c_0), predict outcomes (c_1), adjust thoughts (c_2), and generate novel solutions (c_3), in short:

- c_0: use observations to infer (causal relationships),
- c_1: use observations to form outcome expectations (predict),
- c_2: use context-dependent observations in different conditions (adapt),
- c_3: use observations to conceive unseen conditions (imagine).

The first capability in the list calls for mechanistic reasoning "Does A cause B?." Crucially, this reasoning is exercised under natural conditions, not under the observer's

control. Consequently, capability c_0 requires robust causal inferences regardless of input conditions. This capability enables an observer to figure out what is relevant.

The second capability in the list underlines predictive analysis "Would A cause B?." Such analysis comprises rehearsing observations and forming outcome expectations. Capability c_1 requires predicting likely outcomes from past observations. This capability enables an observer to know what to expect.

The third capacity in the list involves reasoning over potentially different scenarios "What if A causes B?." Capability c_2 requires bending context-dependent observations to specific conditions. The capability to adapt context-dependent observations to different states enables indirect learning to rehearse models in situations that differ from the original.

The fourth capability in the list sparks innovation via imagination ("What would have been, had A cause B?"). Imagination involves a playful mind which imagines alternative (previously unseen) scenarios. c_3 requires the capability to revise observations and imagine scenarios not seen before. Observational learning requires anticipating the outcome of novel actions without ever attempting to exercise these.

Turning the attention to the construction of artificial learners. Two schools of thought drive the way artificial learners are built, those that construct on data, theory, and the hybrids in the mid-range. These are theory-driven (A), model-driven (B), data-driven (C), and instance-driven (D).

Each approach produces outcomes differently. Theory-driven relies on a priori theories to deduce outcomes (A). Model-driven uses data to update structures and infer outcomes (B). Data-driven uses data to build theories and induce outcomes (C). Instance-driven relates examples by similarity to derive outcomes (D).

Each approach involves a different procedure. A theory-driven agent starts with a priori principles (encoded once initially), serving as general rules to solve a class of problems (this agent begins with theories). A data-driven agent begins with data to induce theories for specific problems (this agent progresses from data to theory). A model-driven agent initiates with structures and then updates with data to infer outcomes (this agent updates theories with data). An instance-driven agent remains with data without establishing any theory (this agent lacks theories).

Each approach carries key assumptions upon which these are built. Theory-driven initiate with rules that, once deployed, should not require change (agents like these are data-blind because theories remain regardless of data). Data-driven find rules in data that, once learned, should not change (these are theory-blind because agents can create but not modify theories). Model-driven rests on structures assumed flexible

enough to capture diverse data (these are theory-modifiers because agents update theories with data). Instance-driven assumes similarity explains data (these agents are theory-free because no rules are involved).

The capabilities (c_0, c_1, c_2, c_3) introduced before are essential for observational learning. Still, conventional machine processing approaches (data-driven, theory-driven, model-driven, and instance-driven) do not fully support them. These four approaches support prediction (c_1) but no causal inference (c_0) because a truth-preserving mechanism is necessary but unavailable across approaches.

A truth-preserving mechanism is a strategy that counters potential bias when drawing interpretations from uncontrolled data. The strategy consists of identifying potential sources of bias and adjusting against them. Causal inferences require the preceding step to mitigate bias before finding relationships from data [SRMS97, Pea19].

Approaches relying on correlational inferences are prone to bias and thus unable to derive causal relationships. Not supporting any truth-preserving mechanism limits the four approaches to correlational inferences when causal is needed. This inability alone excludes approaches (A,B,C,D) as candidates to operationalize human-like observational learning. The difficulties in supporting the other capabilities are discussed and summarized in Chapter 2 (Table 2.2).

1.2 Problem Definition

Given the preceding description, this section emphasizes the current status and clarifies the two problems this dissertation addresses.

The approaches to formalize observational learning in machines span over data-driven [OMG13, PALvdP18, YMH+19], theory-driven [MD00, CB07, KNCC11], instance-driven [BCG06, GSNT19, LGA+17], and model-driven [HE16, TWS18]. Common across these approaches are inferences that rely on associations and support prediction.

The settings where machines aim to learn from observation include with a teacher (e.g., kinesthetic teaching, imitation learning, etc.), by doing (e.g., reinforcement learning, self-play, etc.), with a teacher and by doing (e.g., reward-based reinforcement learning, apprenticeship learning, etc.) – featuring the intention to assist learning and control the quality and quantity of the inputs (typically large collections of quality data).

However, conventional machine processing approaches and learning settings are inadequate to support human-like observational learning:

- The machine processing approaches (theory-driven, model-driven, data-driven, instance-driven) support only part of human observational capabilities (as identified by Social Cognitive Theory). The inferences provided by these approaches are inadequate to support observational learning because key processing capabilities are unsupported. The four approaches can draw correlational predictions, but none supports causal analysis, let alone the four capabilities (causal inference, predict, adapt, imagine).

- The settings where humans learn from observations differ from those in which machines attempt to learn from observation. Ensuring arbitrary quantity and quality of inputs is common across artificial settings (with a teacher, by doing, with a teacher, and by doing) but uncommon in natural scenarios where people observe (vicarious settings). Artificial settings carry teaching intentions that guide learning, whereas observations in natural environments are not intended to be learned. Such differences render machines settings (with a teacher, without, with a teacher, and by doing) insufficient to capture the necessary conditions of an observational scenario.

As a result, robots lack processing capabilities to analyze other's interactions and harness observations under natural conditions as humans do. Robots ill-equipped in terms of observational capabilities are inadequate to acquire others' know-how — for example, robots learning from observing human hand behavior in a table-setting scenario — where the quality and quantity of examples are not under the observer's control, and — where capabilities allow harnessing observations to explain hand manipulation (causal inference), anticipate hand behavior (predict), transfer other's behavior to different conditions (adapt), and innovate behavior not seen before (imagine). In such a scenario, causal inference is foundational to attaining a mechanistic understanding of the environment. The capability to predict is important for evaluating likely outcomes (basis to anticipate others' actions). The capability to adapt is crucial to benefit from past observations by bending these to present conditions (past conditions are unlikely to repeat identically). The capacity to imagine is critical to innovating in situations not seen before (problem-solving). Moreover, robots that rely on inferences sensitive to input conditions, such as large collections of quality input examples, are restricted from operating in those scenarios where the inputs meet known conditions.

To formalize human-like observational learning in robots and make them competent

in the analysis of human interaction, inferences require the capabilities to answer the following stereotypical questions:

1. Does X cause Y? (causal inference)
2. What would/will happen? (predict)
3. What would/will happen if? (adapt)
4. What would/will have happened if? (imagine)

Conventional machine processing approaches do not answer the four questions. This dissertation explores an inference approach that, in theory, covers the four processing capabilities and thus is positioned to answer these questions. However, there are no guarantees this proposition is feasible in practice, especially not under conditions for which it is not prepared to operate, such as with a small number of samples and models capturing only part of the relevant factors.

This thesis probes the potential of the approach with a scenario that challenges inferring causal relationships on hand-behavior from a person performing natural table-setting activities (unintended to be learned by a robot) considering only a few examples of uncontrolled quality (classes are not balanced, accidental samples are not discarded) using models capturing only part of the relevant factors (incomplete).

1.3 Aim and Objectives

Given existing approaches supporting only part of the core capabilities of human observation, this dissertation aims to "Identify and evaluate an approach formalizing the core capabilities of human observational learning and defy its potential with a natural scenario." Note that our aim is not making the approach work but rather to challenge its potential.

The research objectives to accomplish our aim are the following:

1. **To identify and instantiate an approach supporting the core observational capabilities of humans and a scenario challenging its potential.** This contribution establishes a framework (the object of our study) and sets the scene to challenge it.

2. **To evaluate the feasibility of the approach.** This evaluation verifies the proper functioning of the framework encompassing the validity and stability.

3. **To evaluate the utility of the approach.** This evaluation assesses the framework's capability to answer and explain causal questions.

1.4 Contributions

Concerning our aim and objectives presented in Section 1.3, the following describes our contributions:

1. **A framework and scenario for robot observational learning.** We identify an approach that covers core observational capabilities following the social cognitive theory. We cast the challenge of robot observational learning as a problem of structural causal models with non-parametric estimation and designed a model for hand manipulation. The do-calculus supports the core capabilities to infer causal relationships, predict, adapt, and imagine (using the dowhy [SK+19] library). The non-parametric estimators provide robust valuations despite uncontrolled data (with econml [Res19]). The design of a model for hand behavior encoded with a causal graph. Together, these components allow any time processing of causal relationships from uncontrolled inputs and support predict, adapt, and imagine to harness observations of hand behavior.

   A scenario that challenges the framework's potential. We set a scenario that probes the approach's potential. The scenario includes a machine (the observer), a human being (the observee), and hand interactions in a table-setting activity. To promote natural conditions, we ask participants to set the table without further instructions and use the raw data to prevent the quality and quantity of examples (keep accidental grasps in the data and unbalanced classes). To challenge the approach's potential feasibility and utility, we consider only hundreds of samples and employ an incomplete model. We employ a near photo-realistic virtual environment to record activities, enabling users to interact with a kitchen scenario using hand-held devices (unreal-engine, oculus rift). To allow robots access to virtual activities, we employ RobCoG and its kitchen scenario and use its data format to extract grasping actions and writing queries.

2. **A verification that probes the proper functioning of the approach.** We conducted an experimental study targeting human hand behavior and collected the validation dataset. Using the validation set, we evaluate the framework's ability to provide valid and stable results. We recover nontrivial causal relationships on hand behavior to verify the framework's validity (the case of showing an effect when there is none). We compromise the robustness perturbing the model, data, and both simultaneously to test the framework's stability.

3. **A comprehensive assessment of inferences on hand behavior.** We assess the framework's utility to harness observations by inspecting the capability to predict, adapt, and imagine. We collect a data set on natural table setting activities (free from experimental influence) and add an external data collection (thus hand behavior of two individuals). We postulate causal questions on human hand behavior and answer them with the framework. Robots employing the framework can realize the same analysis.

This study contributes to the body of knowledge on robot observational learning by proposing and probing an alternative approach prepared for any time reasoning under natural scenarios. This will help address the current shortage of research in this area and explore the boundaries of an unconventional approach in a social environment where learning by trial-and-error risks harming people and damaging robots, and exhaustive brute-force explorations are impracticable.

This investigation does not reach the real robot platform, and the results do not necessarily generalize to other settings. Moreover, the validation of the approach covers one type of error, namely, showing an effect when there is none, but does not verify the case – showing no effect where there is one. Furthermore, the assessment formulates three types of questions (predict, adapt, imagine) focusing on the first type and leaving the latter two for future explorations. Nonetheless, with this dissertation's proof of concept and value, we anticipate and test the boundaries of a promising inference approach to a scenario featuring real-world conditions.

1.5 Outline

The content of each chapter is as follows.

Chapter 2 distinguishes machines from humans. Machine and human learning settings and information processing are characterized and contrasted, and a set of

desirable properties are identified.

Chapter 3 provides an overview of the strategy and steps of this dissertation and facilitates foundational concepts that underlie the approach.

Chapter 4 establishes a framework enabling the analysis of causal relationships and core reasoning capabilities. The main components of the framework are introduced, including the model of hand behavior, the machinery for causal inference, and the estimators.

Chapter 5 describes the validity and stability of the framework. This chapter describes the ground-truth data, the verification for correct estimates, and the stability under varying conditions.

Chapter 6 inspects the ability to answer causal questions using the framework by targeting unconstrained hand manipulation of two individuals comprising the formulation of causal questions and their explanations.

Chapter 7 presents conclusions and proposes future venues for this research.

Parts of this thesis have been published in [PWYB18] and [WSB22].

Aim	Objectives	Contributions
Formalize and evaluate an approach supporting core capabilities in human observation and challenge its potential feasibility and utility with a natural scenario.	To identify and instantiate an approach supporting core observational capabilities of humans and a scenario challenging its potential. This establishes a framework (the object of our study) and sets the scenario to challenge it.	**Chapter 2, 4:** *Establish a framework and scenario.* • Established a framework for robot observational learning. • Set a scenario that challenges the framework's potential.
	To evaluate the feasibility of the approach. This verifies the proper functioning of the framework covering the correctness and robustness.	**Chapter 5:** *A verification probing the proper functioning of the approach.* • Designed an experimental study on human hand manipulation. • Verify the correctness. • Assess the robustness.
	To evaluate the utility of the approach. This assesses the capability of the framework to answer and explain causal questions.	**Chapter 6:** *A comprehensive assessment of inferences on hand behavior.* • Postulated causal questions targeting the hand behavior of two individuals. • Analyzed answers to causal questions and their explanations.

Figure 1.3: *A dissertation summary.*

Chapter two

Background

Robot analysis of human interaction has the potential to unlock social interaction. The benefits of robots observing human behaviors are twofold. First, having robots adopt people's know-how can potentially incorporate key knowledge for control systems in workspaces made by and for human beings. Further, robots adopting human behaviors act predictably to their human counterpart. Second, robots modeling human behavior are better positioned to anticipate a person's intention, thereby improving assistance in everyday tasks.

To make robots competent in the analysis of human interaction this dissertation investigates an approach to mechanize human-like observational learning and evaluates its feasibility and utility in an natural setting.

This chapter discusses the background work in the analysis of human interaction. It is structured in two parts: the approach of machine analysis of human interaction and how people learn from observing others. Each piece covers two dimensions, the learning setting, and the processing. First, the machine settings and processing are described in Section 2.1, followed by the human settings and capabilities in Section 2.2. Last, machine and human approaches are contrasted in Section 2.3, and a set of necessary characteristics and capabilities is identified.

2.1 Robot Analysis of Human Interaction (the machine way)

A learning setting comprises a student, a teacher, and an environment shown in Figure 2.1. These elements designate broad terms to cover a range of learning settings from the literature. The role of the student (a.k.a. learner, machine, robot) is to improve in some regard (a.k.a. learn, acquire). Improving requires some processing mechanism (a.k.a. inference, deduction). The role of the teacher is to share some knowledge (a.k.a. instructor, researcher), and the environment sets the context for the interaction, including but not limited to the interplay of student-teacher, student-environment, teacher-environment, to mention a few options.

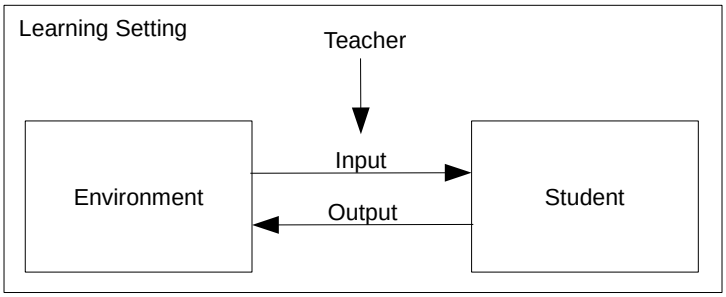


Figure 2.1: *A learning setting including a student, teacher, and environment.*

The following section characterizes the settings in which machines aim to improve (Section 2.1.1). Subsequently, the presentation describes the processing machines employ to improve (Section 2.1.2).

2.1.1 The settings where machines aim to understand others

The landscape in robot analysis of interaction presents several settings in which researchers teach machines. Some settings incorporate the explicit figure of a teacher, while others do not.

The literature categorizes three types of settings where machines learn [RPCB20]: A) learning from demonstration with a teacher, B) learning by doing without a teacher, and C) settings combining learning from demonstration and by doing.

Agents learning in settings of type A (with a teacher) are highly supervised, those in type B (without a teacher) are unsupervised, and those in type C are hybrid (combine

supervised and unsupervised). Figure 2.2 shows several settings in each category and how the three settings are related.

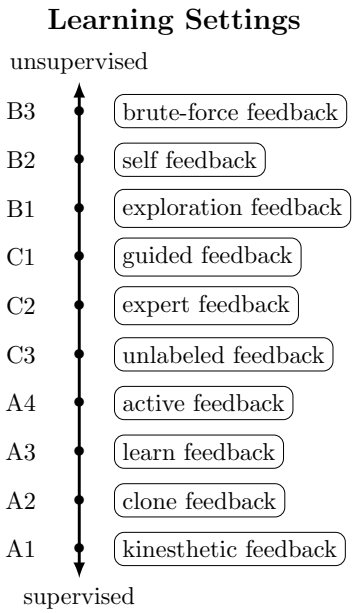


Figure 2.2: *The settings where machines aim to improve.*

The presentation covers the learning settings appearing in Figure 2.2. The first setting introduces those involving the figure of a teacher (type A), followed by those lacking one (type B), and last, those combining the presence and absence of supervision (type C). The settings within these categories, A1-A4, B1-3, and C1-C3, are described in subsequent sections.

A) Learning from demonstrations with a teacher

Robot Learning from Demonstration (LfD) or Robot Programming by Demonstration (PbD) is a paradigm for enabling robots to perform new tasks without the need for programming them every time [BCDS08]. This paradigm assumes robot controllers are derivable from a human's own performance [BG13]. The aim is to extend and adapt robot capabilities to novel situations, even by users without programming abilities [BG12].

Common across type A setting is guidance students receive from a teacher. The

presentation introduces settings A1-A4 ordered from the most supervised to the least.

A1) The student's movement is guided by the teacher (kinesthetic feedback)

The student learns movements as the teacher physically guides the student through the task [BCG06]. Because the teacher employs the student's own body to perform the demonstration, learners are not required to cope with the differences between the body of the teacher and their own. With kinesthetic control, students can replay examples provided by the teacher without first accommodating these to their bodies.

A2) The student imitates a teacher (clone feedback)

Dealing with issues that arise with demonstrations differing in the embodiment is often preferred than avoided because kinesthetic feedback is cumbersome to practice. To control a robot arm, teachers often require more degrees of freedom of their own. This issue is even worse in bimanual tasks where separate demonstrations per robot arm are necessary. Imitation learning [Sch99] aims to solve these issues at the cost of dealing with the difficulties that arise with demonstrations that differ in the embodiment.

In imitation learning (or behavioral cloning), the student copies a teacher's demonstration [CM95]. Unlike the kinesthetic approach, where the teacher guides the student's body, the student controls its body to mimic the teacher. This setting requires teachers to provide quality demonstrations because the student copies even unintentional aspects (e.g., mistakes, jerks). Another drawback with this setting is that students cannot outperform teachers.

A3) The student learns a cost function (learn feedback)

In the previous setting, teachers guide students on what needs to be learned—failing to do so results in unintentional learning. Manually avoiding over-imitation is only feasible for small state spaces (a.k.a. design the reward). To deal with over-imitation, students need the capability to discover the teaching intentions in demonstrations. Inverse Reinforcement Learning (IRL) aims to infer the teacher's

goal in a demonstration (i.e., cost function or reward) [Rus98] and ignores aspects irrelevant to the teacher's goal. This setting is frequently used in combination with others.

A4) The student learns interactively with a teacher (active feedback)

Active learning [Ang88] enables a bidirectional interaction between the student and teacher. This contrasts with the previous settings in that students can raise questions to teachers. For example, the students can ask feedback to label data points [HS14, CT12, BSC19] including corrective demonstrations [GSNT19, BLOD18], or indicate regions for which examples are lacking or limited. Aiming to increase interaction even further, Interactive Task Learning (ITL) proposes a setting where teachers provide an overview of the task using language, both the student and teacher can raise questions, after the essence of the task are understood by the student, a learner can then perform the task possibly through practice with the possibility to ask for further feedback from the instructors [LGA+17]. Works in this area include how to ask for help [KF13], and get human support with feature selection when learning tasks [PM98].

B) Learning by doing without a teacher

In the absence of a teacher (settings of type B), students learn by interacting with the environment. Note, however, that the role of the teacher is often implicit. For example, teaching efforts are encoded into the synthetic environment (modeling the scenario and dynamics) or into the agent (encoding rewards, distance metrics, loss functions, and so on).

B1) Reinforcement learning (exploration feedback)

In Reinforcement Learning (RL), students gain experience through trial and error, maximizing a reward [SB98]. This setting allows students to discover new control policies through free exploration of the state-action space, which contrasts with scenarios where students must follow the teacher. However, without any feedback

from teachers, the learning quest can take a long time to converge. Moreover, the external feedback required in RL settings involves a manual non-trivial engineering task.

B2) Learning by self-play (self-feedback)

Self-play is a setting that allows students to learn by doing without requiring reward engineering. Instead, policy improvement is achieved by engaging students into competitive environments playing against another version of themselves (or self-play or self-learning) [Boz99]. After some rounds, the best version of the student is returned. This approach, for instance, has achieved superhuman performance in the game of Go [SSS+17]. This setting relies on the availability of self-played trajectories which questions its applicability to real-world scenarios.

B3) Generative (brute-force feedback)

Generative settings produce synthetic distributions with competitive settings. The creation proceeds with a competitive environment where the student (i.e., an initial distribution) competes against an expert (i.e., a labeled dataset) to undergo unnoticed by a discriminator (playing the role of a teacher). The distribution mapping task completes when the sampling of the student proves indistinguishable from the expert distribution (the labeled dataset). Generative settings are also regarded as adversarial-settings being a well-known example of generative settings is Generative Adversarial Networks (GAN) [FCAL16].

C) Learning from demonstration and by doing

Combining supervised and unsupervised (type C) settings can overcome drawbacks or potentially solve multiple difficulties at once. Two major concerns were noted earlier when introducing the setting where students clone teacher's demonstrations. One of the issues is that a student is incapable of outperforming the teacher in settings of type A, leaving them no room for further improvements, while the other concerns with students' explorations taking time to converge due to lack of guidance

in settings of type B. Both concerns are no issue for students that learn by doing, as students learn from the interaction with the environment via free exploration without a teacher.

C1) Reinforcement learning with imitation learning (guided feedback)

The combination of reinforcement learning (B1) with imitation learning (A2) benefits from the strength of both settings by exploiting demonstrations to reduce convergence time. The role of demonstrations can either be to initiate [KP10] or guide [PVS05] exploration. This combination of setting is also known as reward-based LfD. In this setting, students require the teacher to guide on what needs to be learned (design the reward).

C2) Apprenticeship learning (expert feedback)

The combination of inverse reinforcement learning (A3) with reinforcement learning (B1) is known as apprenticeship learning via inverse reinforcement learning (AIRL) [AN04]. Unlike the previous setting, to avoid engineering manually the design of reward functions, AIRL enables students to determine a teacher's intention (without manual design) and to discover control policies on their own—opening the possibility for an apprentice to outperform an instructor's skill.

C3) Imitation from observation (unlabeled feedback)

In the settings described earlier, the feedback students receive from teachers is labeled. Labeled demonstrations comprise state and action information. The state information describes an instruction (e.g., a sequence of spatial locations), while the action information provides hints for the student to ease control (e.g., this aspect from the demonstration relates to your left arm). Such labels essentially facilitate mapping instructions to internal motor commands. Labeled examples are hardly transferred across robot platforms.

Unlike the labeled feedback provided by teachers in settings described earlier, Learning from Observation (LfO) [OMG13] (or Imitation Learning from Observation (IfO))

comprises state-only demonstrations without any action information (i.e., unlabeled feedback).

2.1.2 The processing with which machines aim to improve

The previous section characterized the settings where learners aim to improve. In the following the focus is on how learners process information.

There are mainly two opposing schools of thought in the way artificial learners are built, those that build upon data and those on theory. Both define the extreme of a spectrum that includes hybrids in the mid-range, as shown in Figure 2.3.

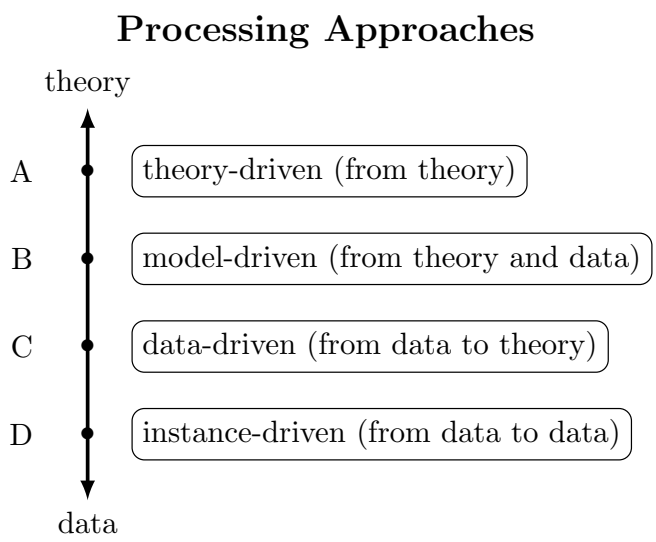


Figure 2.3: *The processing with which machines aim to improve.*

The four machine processing approaches shown in Figure 2.3 are those which begin with theories (theory-driven), those shaping initial theories (model-driven), learning theories (data-driven), and neglecting theory (instance-driven).

In the data end of the spectrum, the effort is on improving the quality of data to raise learners' performance on specific problems. The theory end of the spectrum focuses on building better theories to enhance an agent's performance for a class of problems. Efforts in the mid-range of the continuum dedicate to finding the minimal theory necessary to initialize intelligent behavior. Each is introduced in subsequent sections.

The data end of the continuum characterizes more reactive processing that rests on little initial structure (involve data). The theory end of the continuum features more deliberate reasoning starring considerable initial structures (involve theory), and the mid-range spans between reactive and deliberate (involve theory and data).

Data-driven approaches induce structure from input examples (learn a theory). Students that learn theory are built with statistical and machine learning methods (supervised, semi-supervised, and unsupervised). Data-driven approaches found great success in prediction tasks.

Theory-driven approaches embed initial structures that serve as general tools to confront a class of scenarios (begin with theory). Learners build with this approach typically begin with some language (such as logic reasoning, probabilistic logic, and other variations). Theory-driven approaches have found great success in tasks that demand complex reasoning.

Model-driven approaches initiate a minimal theory, which students can further expand and shape with data (develop initial theories). Model-driven success is in diagnostic tasks.

Instance-driven approaches build from and to examples without establishing any theory (skip the theory). Instance-driven approaches found success in tasks where a few examples are available (not enough to build theories).

The following sections emphasize machine processing within the established schools of thought, starting with theory-driven (A), then model-driven (B), data-driven (C), and instance-driven (D).

A) Theory-driven (from theory), or deductive reasoning

Theory-driven [BLS83] approaches equip agents with a set of general tools a priori. General tools are existing theories transformed into programs. In its pure form (as adopted here), empirical evidence is ignored.

An agent that incorporates physics principles, such as the law of motion, is theory-driven. Other theories built into agents are complex capabilities associated with cognition, such as language processing, planning, and relational reasoning, to mention a few. These agents initiate with theories regarded as invariant, thus exempting agents from modifying these. The law of motion remains unchanged, the language's syntax remains fixed, and so on.

Theory-driven approaches exhibit different kinds of deductive reasoning, some of

them described in the following. In deductive reasoning, conditionals apply on axiomatic premises ”if p then q.” The distinguishing feature of this sort of deduction is that inferences can be guaranteed valid (assuming premises are correct). This form of reasoning is strict as no exceptions to the rules are allowed, and costly because attaining axiomatic premises is not straightforward.

Deduction is supported by several variations of logic reasoning such as, propositional logic, first-order logic, and the set of standard logic which comply with the properties of i) law of excluded middle and double negation elimination, ii) law of noncontradiction, and the principle of explosion, iii) monotonicity of entailment and idempotency of entailment, iv) commutativity of conjunction, v) de Morgan duality where every logical operator is dual to another, being most standard logics two-valued systems.

Nonstandard logic systems reject, extend or restrict any of the five properties. For example, linear logic abandons the idempotency of entailment. Non-reflexive logic rejects or restricts the law of identity. Relevance logic, linear logic, and non-monotonic logic reject the monotonicity of entailment. Nondemonstrative logics are nonstandard because these provide no complete demonstration of a claim. Examples include probabilistic, statistical, abductive, and inductive reasoning.

Monotonicity [DLLW20] is a condition often assumed in practice. All standard logic, including most nonstandard logic systems, assume it. Monotonicity implicates that learning a new piece of knowledge cannot reduce the set of what is already known. On the other hand, non-monotonic logic enables reasoners to retract their conclusion based on further evidence, also known as defeasible reasoning [Nut03].

Defeasible reasoning is also known as reasoning to the default and regarded as an anytime algorithm. Unlike deductive reasoning, defeasible reasoning takes the conditional on authoritative (not necessarily axiomatic) premises “if p then (defeasibly) q.” Neither does default reasoning demonstrate claims (i.e., nondemonstrative) allowing to handle exceptions to rules. This sort of deduction has found applications in linguistics, knowledge representation, and in prima facie and ceteris paribus reasoning.

Another variation of nondemonstrative deductive reasoning is probabilistic logic [Hal90] where conditionals act on combinatorics “if p then (probably) q.” Probabilistic reasoning extends logic resulting in a more expressive formalism allowing to deal with uncertainty at the cost of computational complexity and the possibility of counter-intuitive results. Well-known examples of this formalism are fuzzy logic [JB96], Markov logic network [RD06], and evidential reasoning [Sha76], to mention a few.

Examples of deductive systems include languages (demonstrative logic [CM03,

WDF+09, SCV19, MQXE20], nondemonstrative logic [AB07], logic functional programming [MSC+20, Tar13], rule management systems (Drools [Pro12], jBPM [YZG+15]), propositional calculus (DPLL [NOT05], Chaff [VWM15], SMT solvers [BHM+13]), lambda calculus ([Chu40]), and proof system (calculus of structures [Gug07], deep inference [Sch77]).

B) Model-driven (from data and theory), or abductive reasoning

Model-driven approaches initialize minimal structures that learners further expand using data (develop theories). Model-driven approaches center on updating structures (e.g., belief networks) and, thereby theories.

Examples of model-driven agents are knowledge-based systems (or expert system, rule-based systems), graphical models, etc.

Model-driven approaches are understood as a form of abductive reasoning [KKT93] in which premises include data and theory "p and q are correlated, and q is sufficient for p; hence, if p then (abducibly) q as cause." Unlike deductive reasoning relying only on theories, authoritative premises in abductive reasoning comprise theory and data. Abduction begins with an observation and seeks the most plausible conclusion, also known as inference to the best explanation. Abduction belongs to nondemonstrative logic because claims are not verified. This sort of reasoning is suitable for diagnostics in automated planning (fault detection), analysis of competing hypotheses, belief revision, preference models, and historical linguistics (past reconstruction) [Jos91].

Most of graphical models comprise abductive reasoning. Bayesian networks represent knowledge with graphs and edges expressing parameters of the joint probabilities. A learner is initialized with prior distribution, which is then updated. Well-known specializations of the Bayesian network are hidden Markov models, neural networks, and variable-order Markov models. Other derivatives are the naive Bayes classifier, Markov random field (Markov network), conditional random field, restricted Boltzmann machine, to mention a few.

Also statistics is a sort of abductive reasoning with premises for the conditional encompasing data and presumptions "the frequency of qs among ps is high (or inference from a model fit to data); hence, (in the right context) if p then (probably) q." Statistics rely on data and a second authoritative premise, which are presumptions (analog to theories). For instance, inferential statistics (as a theory of belief) presumes observations originate from a population.

An abductive reasoning agent based on Bayesian inference (statistics) updates probabilities for hypotheses (theories) as more evidence (data) becomes available. Bayesian inference found applications in statistical classification and prediction of the next character upon a given series of symbols.

C) Data-driven (from data to theory), or inductive reasoning

Data-driven approaches equip agents with the capability to construct theories from examples. This approach is typically employed to find general principles from specific instances. Discovered principles (hypotheses or theories) are then employed to predict the occurrence of a next instance. Most machine learning relies on this inductive procedure. Unlike theory-based approaches, which incorporate theory a priori, data-driven learners discover the inexistent theories from examples.

Because no theory exists apriori, theories are built bottom-up from specific examples. Data-driven is a form of inductive reasoning where the premise is data "(inducibly) if p then q; hence if p then (deducibly-but-revisable) q." Induction differs from the deduction in that conclusions to be true cannot be guaranteed; rather, truth is probable according to the evidence given. Inductive reasoning offers probabilistic knowledge, which contrasts with deductive certainty hardly attained in practical systems.

Fitting an ML model is essentially induction (a.k.a. learn a function from training examples), where the resulting model is the generalization of specific instances seen in the training set. This approach enables to induce association rules from datasets. Various formats encode the learned association rules (i.e., induced theory) such as graphs or networks (e.g., neural networks), tree structures (e.g., decision trees), vectors, and matrices (support-vector machines, genetic algorithms, deep-learning), linear models (regression).

Some well-known examples of data-driven techniques are artificial neural networks (ANN) [GBC16], support vector machines (SVM) [CV95], random forest (RF) [LWZ12], relevance vector machines (RVM), classification and regression trees (CART) [KA17], linear discriminant analysis (LDA) [MK01], to mention a few.

D) Instance-driven (from data to data), or transductive reasoning

Instance-driven approaches known as transductive reasoning build the capability to relate specific examples to other instances [Vap06a]. Transductive reasoning employs specific instances to directly explain other observed examples without forming a theory (like done in data-driven approaches).

In this approach, conditionals apply to data "example p and q are similar if p then (analogous) q." Instance-driven contrasts with data-driven approaches in that the latter relies on large collections of data typically involved in theory formation. Hence, instance-driven reasoning is especially relevant when only a few examples are available. Transductive reasoning underlies case-based reasoning, analogical reasoning, transfer learning, and hypothesis testing.

Hypothesis testing in descriptive statistics is a sort of transductive reasoning which questions the similarities between two sets of examples. Another example where training data does not induce theory is transfer learning in ML (e.g., k-Nearest Neighbors algorithms [For65]). Instead, examples match other instances based on some distance metric.

Other instance-driven methods are Transductive Support Vector Machine (TSVM) [Vap06b], Bayesian Committee Machine (BCM) [Tre00], density-based clustering ([EKSX96, ZRL96]), network transduction [JY06], to mention a few. As no training is involved, instance-driven techniques naturally support adding new data points and thus allow incrementality.

2.2 Human Analysis of Human Interaction (the human way)

A capability often overlooked in humans is learning by observing others. Studies in neuroscience identified specialized neurons that activate when primates execute a task and also trigger when the primate observes someone else replicate that same task [RFG01, RC04]. These came to be known as the mirror neurons suggesting that action-execution and action-observation are linked. The experiments have also been replicated and confirmed in humans [SL13, RF14]. Moreover, neuroimaging studies found that brain waves largely overlap during action-execution and -observation [GARP03, HRD$^+$03, RES04] regarded as action-potentials in an observer's mind [RF14] with neurons displaying similar spiking modulation and directions [CK04, TRH07].

These findings led to direct and far-reaching implications. Some scientists have implicated the mirror neurons as the physiological mechanism for the perception-action coupling taken as the fundamental logic of the nervous system suggesting that the motor system is part of mental processing (a.k.a. motor cognition theory) [SD06]. Mirror neurons have also been linked to the mental simulation of actions in the observer's mind [RES07], and in accord, support the attribution of mental states of others (theory of mind) [KG06]. Mirror neurons are taken by some researchers as the neural basis for the human capacity for emotions such as empathy [Iac09]. Some attribute mirror neurons to form a neurological basis for understanding the actions of other people, social learning, imitation, and observational learning [IMSG+05, The06, UILK07, LRLAdO13, Rea14]. Others have also linked them to language [TPL02].

Long before discovering the mirror neurons, research in psychology proposed the Social Learning Theory (SLT) [Ban77]. In this theory, [Ban65] found that role models influence behavior and that imitation is not essential for learning to occur [BB71] (rather optional). Further research on the subject shows that people can learn through others' experiences [BR66] (by watching others), and that experience transmitted this way requires symbolic processing [BGM66]. Moreover, memory performance is governed more by information representation than associative strengthening processes [BJB74] (modeling is essential). Furthermore, behavior is learned symbolically by processing outcome information which precede action execution [BJ73] (outcome expectancies). According to this line of research, visual monitoring is to decrease discrepancies between conceptual behaviors and action patterns [CB82]. That conceptual behaviors translate into actions [CB87] guiding the production of outcomes [CB90]. Further advancements of the theory justified renaming the Social Learning Theory (SLT) to Social Cognitive Theory [Ban86].

The SCT theory attributes part of knowledge acquisition in people to the observation of others within a context of social interaction. This theory states that people observe role models' behaviors and their consequences, retain these observations and use them to guide subsequent behaviors [Ban86].

> "Observational learning (also known as vicarious learning, social learning, or modeling) is a type of learning that occurs as a function of observing, retaining, and replicating novel behavior executed by others. It is argued that reinforcement has the effect of influencing which responses one will partake in, more than it influences the actual acquisition of the new response" [Ban77]

The acquisition can be subtle in observational learning because visible changes in behavior are not necessarily apparent. This form of learning has been witnessed

to enhance interest. For example, individuals become curious about an object simply by watching others interact with it. Learning also manifests in behaviors simply becoming more frequently, such as when stimulating infants to learn new actions or practice those already in their skill set [Spe37]. Goal emulation is another manifestation of observational learning when, for example, individuals are captivated by the end result of observed behavior and aim for the same outcome, devising their own strategy. End state emulation leads to reflections about the performer's own abilities, and as a result, learning about the environment occurs [Hag09]. A related behavior enhancing observational learning is when individuals remain in close proximity to other individuals with more experience, thereby increasing learning opportunities by exposure to examples [Hey93].

2.2.1 The setting where people observe others

Contrasting imitation from observation is insightful to remark what characterizes observational settings. According to the SCT, people learn by direct experience (imitation) and indirectly (observation) [BGM66]. Both enable people to improve, but are fundamentally different.

The purpose of imitation learning is narrow in contrast to observation learning. Imitation comprises duplication as the end in itself, whereas observational learning transcends the act of mimicry [Dew16].

Perhaps the most notorious distinction of observational learning is that learning can occur without explicitly exercising the act of imitation. Imitation occurs in direct presence of a model, whereas learning from observation is not bounded to proximity. Instead, observed activities and effects are mentally modeled with the possibility to replicate them eventually. This distinction is the first listed in Table 2.1.

Feedback (reinforcements) directed to others can also be learned by simply witnessing them. For example, a child observes a parent's praise to another child. Learning reinforcements directed to others is also known as vicarious reinforcement (second distinction listed in the table).

Table 2.1 contrasts further differences which are described in the following. Third, the roles people adopt in either setting differs. Imitation evidences a clear teacher-student role while observation features the weaker observer-observee role instead (e.g., unaware of being watched). Fourth, teachers guide attention, thus alleviating students learning experience in imitation settings; in observational learning, on the other hand, attention is the responsibility of the observer (self-guided). Fifth,

Comparison	Imitation	Observation
1. Learning	direct experience	indirect
2. Reinforcement	required	optional
3. Roles	student and teacher	observer and observee
4. Attention	guided by teacher	observer's responsibility
5. Execution	immediate	immediate or deferred
6. Purpose	replication	replication or innovation

Table 2.1: *The distinction between imitation and observation.*

imitation occurs immediately in the presence of the model, while for observational settings, in the absence of the role model duplication can be deferred. Sixth, imitation leverages replication while observation fosters new thinking and behavior innovation.

The settings in which humans learn from observations are natural because people cannot control everything they observe. Because observations are mostly uncontrolled, input quantity and quality cannot be enforced (controlled). Moreover, storing observations for short-term or immediate use is valuable, but it is also imperative to exploit them in future (potentially different) situations. Furthermore, in future situations, exact conditions under which observations are acquired hardly repeat.

The setting in which people learn from observations characterizes by the following:

- inputs involve interactions of the environment, others, and oneself
- inputs are natural (of any distribution, quality, quantity, order)
- inputs are of immediate or delayed use
- inputs are unique (exact conditions rarely repeat)

Inputs of such characteristics challenge our understanding of how people learn from natural observations. Natural settings feature raw inputs that are far from ideal to learn with them. Nature does not arrange events in a particular manner to alleviate human learning. Nevertheless, people manage to benefit from natural sources of information. Despite learning conditions, people master daily chores (eat, dress, cook, etc.), travel, interact, communicate, and much more.

2.2.2 The cognitive capacity to understand others

A long-standing hypothesis supports that human logic reasoning is similar to propositional logic [O'B09]. Instead of general abstract and syntax reasoning, another hypothesis [CTFB05] proposes context-specific rules of inference. Alternative views advocate that people reason on mental representations (models) [JLB02], or that people inferences are based on probabilities [JL08].

Further studies [Eva12] identify that humans reason with two modalities, reactive and deliberate. The former is faster (intuitive, nondemonstrative), while the latter is slower (formal, demonstrative). A clear manifestation of both reasoning modalities, for instance, are mathematical proofs which are guided by intuition and proven by formal reasoning.

Experiments on deductive reasoning in humans hint that people do not comply fully with formal reasoning. For instance, [Eva05] shows that conditional inferences comply with modus ponens but not necessarily modus tollens. Likewise, research on human intuition does not identify a single but several mental strategies responsible for quick but error-prone reasoning [Rei16]. Findings like these suggest that human reasoning does not fully comply with the rational model of decision-making proposed in [vNM47] but rather incorporates biological biases [Kah11].

Several sorts of reasoning are identified in people. Deduction enables people to reason from known generalities (hypotheses, theories) to specifics (conclusions, outcomes). Moreover, deductive reasoning helps people discover when their knowledge conflicts with reality (test hypothesis).

Unlike a deduction, inductive reasoning allows people to start from specific observations and reach general principles (create understanding). On the other hand, abductive reasoning allows people to make educated guesses despite incomplete information (explain phenomena). Conversely, transductive reasoning permits people to relate aspects on the basis of similarity (find analogies).

Another kind of reasoning identified in humans is the understanding of mechanism. People understand causes and effects across a broad scope, for instance, in a physical realm (object A causes B), in a social context (behavior A causes B), in a personal context (my action A causes B). Causal reasoning is not limited to present conditions, as people's language reveals reasoning in hypothetical contexts ("what if we go left?") or retrospectively ("what would have been if turned right instead?").

Already at the age of two, toddlers display reasoning capabilities over cause and effect [Gop10] enabling them to understand their surroundings by simple observation

of other social agents [STG04], which include unintended behaviors for infants to learn [MWG12] (e.g., nose-picking behavior children acquire via observation).

The mental capacity to interpret and reason over behavior

According to the SCT:

> "People gain understanding of causal relationships and expand their knowledge by operating symbolically on the wealth of information derived from personal and vicarious experiences. They construct possible solutions to problems and evaluate their likely outcomes, without having to go through a laborious trial-and-error process. Through the medium of symbols, people can communicate with others at any distance in time and space. The other distinctive human capabilities (self-regulation, self-reflective, vicarious capacity) are founded on this advanced capacity for symbolization." [Ban01]

The essential capacities underlying observational learning in humans are [BO09]:

- **Symbolizing capacity**: Besides learning from trial-and-error experiences, people can learn indirectly via symbolization, model solutions to problems, and evaluate likely outcomes.
- **Self-regulation capacity**: Besides being shaped by the environment and innate states, people are ruled by their individual self (the third equally important component) which motivates and guides actions by setting goals and pursuing them.
- **Self-reflective capacity**: Human beings evaluate their own thoughts and actions. Individuals can distinguish accurate and faulty thinking by revising and adjusting their thoughts, generating novel ideas, acting on them or predicting occurrences from them.
- **Vicarious capacity**: People acquire knowledge and skills from various sources, including the indirect observation of others' actions and consequences. Learning on behalf of others gains insights into one's own activities.

A foundational aspect in SCT is that individuals possess advanced neural systems that enable the acquisition of skills and knowledge via direct experiences and indirectly

in terms of symbols [BO09]. This view contrasts with theories advocating for lack of symbolic reasoning in the human mind (symbolizing capacity).

Unlike other theories that emphasize the importance of the environment and the internal states (e.g., behaviorism), the SCT proposes the individual-self as equally important [Ban86]. In other words, individuals act on their own (self-regulation capacity) and are not only reacting to the environment.

According to the SCT, people verify how well thoughts match reality and change them when facing mismatches. The mental capabilities involved in thought verification are of self-reflective nature, including idea generation, occurrence prediction, result judgment, and change. The capacity to reflect upon oneself (cognitivism) stands in contrast to theories explaining human behavior with reinforcements and reflexes (e.g., behaviorism, connectivism). SCT describes four ways in which people verify thoughts against reality. These verifications are, executing actions (trial-and-error), another reality check is observing others and their effects (vicarious verification), asking what others believe (social verification), and detecting faulty thinking by deducing what follows from knowledge (logical verification). All of them support valid thinking, but cannot completely avoid invalid reasoning [BO09].

Another distinguishing aspect of the SCT is that humans can self-improve by simply observing others (vicarious capacity). According to [Ban86] vicariously observing people's actions and effects subsume learning from direct experience (behavioral, mental, and affective). This view contrast with theories prioritizing experiential learning as the critical acquisition mode when another modality exists (observational learning).

Because observational learning can occur in the absence of a model, observers retain symbolic representations of the observed behavior, which enables a later rehearsal (e.g., days after witnessing a behavior). Modeling is essential in the theory and further described in the subsequent section.

The modeling undergone to own another's behavior

Owning a behavior means accepting someone's else conducts as one's own. Owning others' behavior requires mental modeling. Owned behaviors can be adapted to situations very different from those in which these were initially acquired and can be innovated. Modeling is essential to own behaviors [Ban88].

To acquire behaviors, retaining models is essential in observational learning. Storing

mental models prepares for the eventual reproduction (prediction, adaptation, innovation) of behaviors which can occur without any external guidance (e.g., deferred in time).

The stages undergone to own another behavior are 1) exposure to the model and retention, 2) acquire the model's behavior, and 3) owning the behavior [Ban77]. In the first stage, observers perceive a model of interest concerned with what-to-model and encoding a representation for eventual (mental or physical) rehearsal. The second stage encompasses how-to-model involving the construction and reproduction of models. The third stage when-to-use relates to conditions that trigger the execution of modeled behaviors.

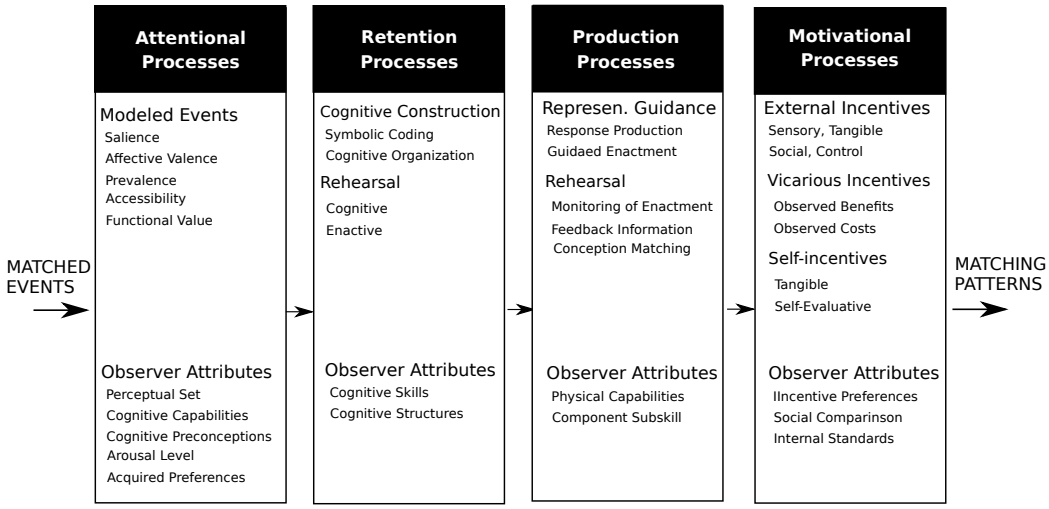


Figure 2.4: *Modeling in the Social Cognitive Theory.*

According to the Social Cognitive Theory, modeling is governed by the four processes shown in Figure 2.4 described as follows [Ban77]:

Attention (what to model): Observers select what to observe and what to extract. Influenced by several factors such as cognitive capability, preconceptions, value preference (accessibility, relevance, complexity).

Retention (model encoding): Observed actions and consequences are transformed to symbolic conceptions in preparation for eventual rehearsals. Recalling memories is not a plain retrieval of registered events since it involves the reconstruction of past experiences, including preconceptions and emotional states, which bias how information is stored and retrieved.

Production (how to model): Stored models are retrieved to guide construction and execution. When necessary, models are modified, matching conceptions to actions within an appropriate context. This process involves transformational and generative operations.

Motivational processes (when to use): Use of models is subject to incentives. For instance, behaviors resulting in rewarding outcomes are more likely executed (than those punished). Similarly, observed cost and benefit others experience can also inhibit (or discourage) action. Besides rewards, motivation is also self-produced via personal reflection.

Modeling draws a clear distinction between acquisition and performance because people do not imitate everything they observe. Neither do observers expect experiencing identical rewards (or punishments) models receive, however, anticipate similar outcomes [Way22]. A high commonality between observer and a role model leads to higher chances of learning behaviors [Ban88] (a.k.a. identification). Conversely, observers are unlikely to exercise behaviors when they do not understand the potential outcomes involved (e.g., pressing a red button in a nuclear facility). Self-efficacy as defined in [Ban89, Ban95] "belief in one's capabilities to organize and execute the courses of action required to manage prospective situations" also regulates execution.

Because several factors govern modeling, observational learning does not guarantee behaviors to be similar. Failures to match modeled behaviors may result from attention directed to irrelevant (or distracting) aspects, inadequate coding into memory, experiencing insufficient motivation, inability to retain observations or physically perform.

In children, deferred imitation is an essential developmental milestone (two-year-old) that manifests the emergence of representational capacity [Pia13]. Modeling is not limited to physical aspects of the world, as it includes the acquisition of abstract skills such as linguistic rules and hints on how to use knowledge.

To conclude, people learn from observing others by mentally modeling observations. First, observers guide their attention over a model of interest, encode a symbolic conception into memory, rehearse the model mentally, and guide learning.

2.3 Robots Analyzing Other's Interactions As Humans Do

The following section contrasts machines' analysis of human interactions against the way people learn from observing others and identifies what is missing to operationalize human-like observational learning in machines.

2.3.1 The Summary of Settings and Processing

Before the comparison is presented, the two preceding sections are summarized next.

Summary of machine analysis of other's interactions

In Section 2.1.1 three classes of settings where machines learn are distinguished. On one end of the spectrum (type A), learners rely on the role of a teacher (their demonstrations), while on the other end (type B), students are left unsupervised (by interaction with the environment). In the center between extremes are settings (type C) where students engage in supervised and unsupervised modalities (demonstrations and environment).

The more supervised a setting is, the more explicit guidance is (e.g., a teacher providing demonstrations). On the other hand, the less supervised a setting is, the more implicit guidance is (e.g., via rewards encoded into agents or a task embedded into the environment). The more semi-supervised a setting is, the more explicit and implicit supervision is balanced.

Just like the rationale behind learning settings in education, common across machine learning settings is the intention to make students succeed by carefully crafting learning conditions (aligned with learning goals, distraction-free, safe, predictable, full of learning opportunities). The intention to make students succeed is either explicit with teacher's guidance and/or implicit in the environment.

Under such conditions, learning can focus on how to execute instead of students' figuring out what to learn independently. In guided settings, teachers make learning for students easier by telling them exactly what to learn. Likewise, in unsupervised settings the focus remains on the how. Students are typically involved in learning all possible inputs, thus avoiding the decision of what needs to be mastered. The

rationale behind brute force learning is to make students succeed by preparing them for all possible input scenarios before deployment. The strategy primarily involves exhaustive learning of the input and state spaces via simulations in synthetic environments or extensive data collections. The learning focus in semi-supervised settings, learners explore variations of the how via exhaustive explorations that initiate on hints facilitated by teachers during supervised conditions.

Overall, regardless of the setting (supervised, semi-supervised, unsupervised), machine settings ensure students a sufficient quality and quantity of input examples. Moreover, the learning focus is set on how not what. Students are incapable to figure out what is relevant at their own.

Section 2.1.2 describes the processing with which machines aim to improve. Four approaches that differ in the authority attributed to data and theory were presented. Theory-driven progresses from theory, model-driven from theory and data, data-driven from data to theory, and instance-driven from data to data.

These approaches associate to different sorts of reasoning, namely, theory-driven to deductive reasoning, model-driven to abduction reasoning, data-driven to inductive reasoning, and instance-driven to transductive reasoning.

These approaches produce outcomes differently. Theory-driven relies on a priori theories to deduce outcomes. Model-driven updates theories with data to infer outcomes. Data-driven builds theories using data to induce outcomes. Instance-driven relates examples by similarity to derive outcomes.

Theory-driven stands on a priori theories that are supposed to not change (is data-blind). Model-driven rests on structures that are supposed to be flexibly updated with data. Data-driven centers on data that is supposed to represent a population (is theory-blind). Instance-driven assumes data similarity explains any input examples (is theory-independent).

Overall, learners build with processing approaches of type A (theory-driven) are theory evaluators, of type B (model-driven) are theory modifiers, of type C (data-driven) are theory builders, and of type D (instance-driven) are example mappers.

Summary of human analysis of other's interactions

Human reasoning hardly categorizes into a single class. Among various sorts of reasoning identified in human beings, some are deduction, abduction, induction and transduction.

An acquisition modality often overlooked in people is observation. People acquire knowledge or skills by simply observing others. Learning from observations encompasses virtually any context and inputs that are beyond observer's control (natural settings).

The mental ability to interpret and reason over behavior (modeling) requires advanced cognitive capabilities. These include the power to operate symbolic conceptions (symbolizing capacity), discover mismatches between thought and reality (understand causal relationships), construct possible solutions (adapt, innovate), and evaluate likely outcomes (predict).

2.3.2 The Comparison of Settings

The settings where humans learn from observations differ from those in which machines target learning. Such differences render machine settings inadequate for the evaluation of human-like observation.

The landscape in Figure 2.5 places the setting where humans learn from observations amongst the machines. The figure shows the human setting between C3 and A4, the first corresponding to active feedback (e.g., active learning) and the second to unlabeled feedback (e.g., imitation learning from observation). The human setting is shortened, hereafter, to vicarious-setting (VS).

The vicarious setting (VS) is between types of machine setting A and C for the following reasons. First, VS differs from type-A settings (where machines learn with a teacher) because demonstrations carry teaching intentions, whereas, in VS, feedback is not subject to such purposes. In other words, the input a learner receives in settings of type A (including IL, IRL, AL, ITL) conveys clear intentions to facilitate learning which is rarely the case in VS. This places VS outside the boundary of type A.

Second, VS diverges from settings where machines learn by doing (type B) because the social component (e.g., a human being) is missing. Consequently, type-B scenarios focus on the physical interactions with the environments, thereby neglecting the social aspect. The incentive to remove human interactions from the learning loop is leveraging uninterrupted and faster than real-time simulations in synthetic worlds. However, the SCT advocates for settings where the environment, others, and oneself are equally important. Settings omitting or prioritizing one component over the other conflicts with VS. This places VS outside the boundary of type B.

Third, the closest type of setting matching VS is when learning combines demonstra-

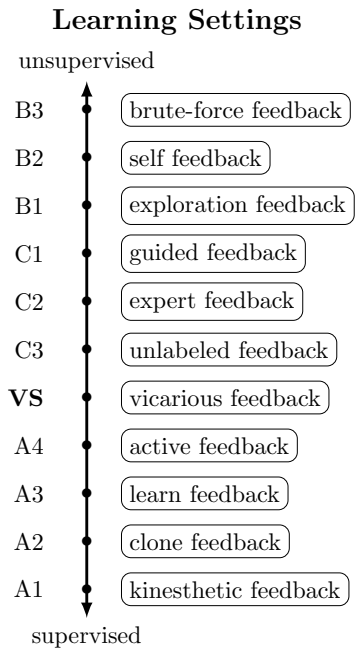


Figure 2.5: *The Human setting (VS) among the machine.*

tion and by doing (type C). The commonality between type C and VS is the flexibility to support both acquisition modalities, guidance with trial-and-error. However, the critical difference is that in settings of type C, reinforcement is a necessary condition for learning to occur, whereas, in VS, reinforcement is optional but can influence learning (e.g., via vicarious reinforcement). This dissociation between reinforcement and learning is rooted in the capacity of individuals to self-regulate. Settings requiring reinforcements to occur, such as type C, expose the lack of self-regulation, thus undervaluing the individual-self to the environment and intrinsic factors (breaking the triadic reciprocal causation model). This places VS outside the boundary of type C.

Two central issues underlie machine settings (type A, B, C): the arbitrary learning opportunities students are granted and the quality inputs they receive. Quality inputs are meant for students to succeed (e.g., curate inputs), just like crafting large numbers of examples is (e.g., ensure sufficient inputs). The number and quality of examples are such to make methods work. This is because settings are under the experimenter's control, which essentially allows controlling the quality and quantity of the input. For instance, in synthetic environments, trial-and-error attempts exceed the number a human could exercise in a lifetime.

Such scenarios contrast with human settings where learning opportunities are limited and potentially of any quality as observers have no control over them. For this reason, students observing in VS dedicate significant efforts to make the most out of every learning opportunity (e.g. exploiting past observations by adapting them to present conditions).

Assuming arbitrary quantity and quality of learning opportunities is mainstream in artificial settings but rare in real-world conditions. In VS, learning opportunities can be scarce and with no quality assurance, essentially because observations are not intended to be learned in natural settings. It is paramount for observational settings to reflect natural conditions and embrace the possibility of inputs with any quantity and quality.

Taken together, the settings where machines learn (type A, B, C) are inadequate to capture the observational scenario (VS). However, a setting matching the observational scheme is critical for developing and evaluating observational approaches.

The study of machine human-like observational learning requires an adequate setting that resembles the natural conditions where people learn from observations. The critical aspect of such a setting is the natural interplay of the environment, others, and oneself.

The following aspects are essential to resemble vicarious settings (VS):

- inputs include interactions of the environment, others, and individual-self
- inputs are natural, not necessarily intended to be learned

The first element of the list avoids breaking the triadic reciprocal causation model. Settings that comply with the triadic model allow for the interplay between environment-with-others, environment-with-oneself, environment-with-environment, and others-with-oneself. Beyond the physical aspects of environments, settings need to include other beings, and rewards are not mandatory (decoupled from learning) because of self-reflective capabilities.

The second element in the list leverages natural interactions. Natural interactions are not necessarily intended to be learned, opening the possibility to any input quantity and quality (learning opportunities). The condition abandons the strict teacher-student setup and favors a loose observer-observee interaction.

Vicarious settings (VS) leverage natural inputs that comprise no attempt to curate observations (any quality) nor attempt to facilitate learning experiences through guidance (embrace correct demonstrations and accidental, balanced, and unbalanced classes of examples, a few and many samples, etc.). The exact conditions are unlikely

to repeat.

2.3.3 The Comparison of Processing

Robots struggle learning in natural settings, for instance, a single unsuccessful trial-and-error is enough to break a robot. For this reason, self-improvement is mostly confined to safe (supervised, laboratory-like, synthetic) conditions. Human beings are also prone to injuries, and one of their strategies is to safely and quickly improve by observing others.

The capability to improve from observing others can potentially enable contact-free self-improvements under uncontrolled settings. None of the machine processing approaches in isolation (covered in Section 2.1) support observational learning (described in Section 2.2). The differences exposed between both settings highlight what is missing.

The capabilities underlying human observational learning are briefly summarized as follows. According to the SCT, people support two acquisition modalities; one is via direct experience (trial-and-error) and another indirect observation (modeling) [Ban77]. Moreover, the theory puts forward that humans possess the capacity to develop themselves [Ban86]. In particular, self-reflective capabilities allow people to detect faulty thinking (from here onwards denoted with the symbol c_0) [Ban01], predict occurrences from thought (symbolized by c_1) [BJ73], revise and adjust thoughts (c_2) [CB82], and generate novel ideas (c_3) [Ban95].

In short, the capabilities for human-like observational learning are:

- c_0: use natural-context-dependent observations to infer (causal relationships),
- c_1: use observations to form outcome expectations (predict),
- c_2: use context-dependent observations in different target conditions (adapt),
- c_3: use observations to imagine unseen conditions (retrospect).

Underlying the four capabilities to model others' behavior is the support for symbolic processing, a representative structure to encode models, and instantaneous reasoning across modeling stages.

The first capability in the list calls for a mechanical sort of reasoning "Does A cause B?." Crucially, this reasoning is exercised under natural conditions (i.e. not under the observer's control). In other words, capability c_0 requires robust inferences regardless

of input conditions. As observational learning primarily involves uncontrolled settings, observers need the capability to figure out how things work on their own (causal understanding).

The second capability in the list underlines predictive analysis "Would A cause B?." Such analysis comprises rehearsing observations and forming outcome expectations. Essentially, capability c_1 requires predicting likely outcomes from past observations. The execution of modeled behaviors is conditioned on anticipating the consequences of actions.

The third capacity in the list involves reasoning over potential scenarios ("What if A causes B?"). The evaluation of potential-outcomes leverage context-dependent observations in particular situations by adapting them. Essentially, capability c_2 requires bending context-dependent observations to specific conditions. Because observational learning deals with unique conditions (exact conditions hardly repeat), the capability to adapt accumulated observations to different states is paramount.

The fourth capability in the list sparks innovation via retrospection ("What would have been, had A cause B?"). Imagination involves a playful mind which innovates with alternative (previously unseen) scenarios inspired by observations. Essentially, c_3 requires the capability to revise observations and imagine scenarios not seen before. Observational learning requires anticipating the outcome of novel actions without ever attempting to exercise these.

Capabilities (c_0, c_1, c_2, c_3) are essential for observational learning, but none of the machine processing approaches (A, B, C, D) fully supports these, as shown in Table 2.2. A check-mark in the table indicates an approach exhibits the necessary features to support a particular capability, while a cross means the opposite.

Overall, Table 2.2 indicates all approaches support prediction (c_1) but none causal inference (c_0). This inability alone excludes candidates A, B, C, and D to operationalize human-like observational learning.

None of the machine processing supports causal inference (c_0) from natural observations because a mechanism to preserve validity is necessary but unavailable across approaches. Not supporting such a mechanism leaves the four approaches restricted to correlational inferences instead of causal. Techniques that rely on correlational inferences (i.e., ignore truth-preserving mechanisms) require supervision. Inputs are unnaturally supervised (i.e., preserves the teacher-student role), which conflicts with the observational setting.

Besides causal inference (c_0), observational learning calls for further reasoning (predict, adapt, imagine). Table 2.2 indicates that model-driven (B) support imagination

Observational Learning	**Core capabilities in SCT**			
Inference approach	c_0: **Causal?**	c_1: **Predict?**	c_2: **Adapt?**	c_3: **Imagine?**
Human-driven Social cognitive theory [Ban77, Ban86]	✓	✓	✓	✓
A) Theory-driven Demonstrative logic [CM03, WDF+09, SCV19] Nondemonstrative logic [AB07, GS14] Functional languages [MSC+20, Tar13]	×	✓	×	×
B) Model-driven RL with IL [KP10, PVS05] Apprenticeship learning [AN04, ZLZ18, ZLN14] IL from observation [OMG13, HE16, TWS18]	×	✓	×	✓
C) Data-driven Reinforcement learning [SB98, Boz99, SSS+17] Generative learning [MO14, FCAL16, KLA+20] Imitation learing [CM95, BCDS08, BG12, BG13] Inverse RL [Rus98, FLA16]	×	✓	×	×
D) Instance-driven Kinesthetic teaching [BCG06] Active learning [Set09, CT12, HS14, GSNT19] Interactive task learning [PM98, LGA+17]	×	✓	✓	×

Table 2.2: *The necessary capabilities to support human-like observational learning.*

(c_3) but not adapt (c_2). Conversely, instance-driven (D) covers adapt (c_2) but fails with imagination (c_3), whereas theory-driven (A) and data-driven (C) fails to support both (c_2, c_3).

As observational learning embraces two acquisition modalities (direct experiential and indirect symbolic). Hence, approaches relying on theory and data are natural candidates to support both modalities which discourage data-only approaches (data-driven, instance-driven) and theory-only (theory-driven) favoring hybrid approaches (model-driven) instead.

Prediction (c_1): the capability to form outcome expectations (requires data)

- Human-driven: Prediction regulates action execution, as poorly understood outcomes are less likely enacted.
- Theory-driven (A): Supports prediction by using built-in theories to evaluate inputs and produce forecasts. Because general theories are invariant, building and modifying theories are unsupported.
- Data-driven (C): Supports prediction by first learning a theory from data. During learning (training), theories can be extended and modified, but when deployed, these remain static.
- Model-driven (B): Supports prediction by updating structures (use theory and data). Theories can be modified and extended anytime.
- Instance-driven (D): Supports prediction by leveraging data (dis)similarity. No theory is involved.

Adapt (c_2): the capability to manage novel situations instantaneously (requires a structure and data)

- Human-driven: Adapting observations from one domain to another.
- Theory-driven (A): Does not adapt. Theories are not expected to change under any circumstance, even when experiential evidence conflicts with the theory (is data blind). The inability to update built-in theories with evidence renders this approach incapable of capturing novelty and adapting new conditions.
- Data-driven (C): Does not adapt. Adapting theories is only supported during theory formation (training) but not when deployed. Once deployed, adaptation means re-learning theories from scratch, which is a clear manifestation of non-instantaneous reasoning. The inability to adapt theories during deployment makes this approach inadequate for anytime reasoning.

- Model-driven (B): Does not adapt. Adapting structures anytime and (to some extent) incorporating novelty are supported. However, structures impose some form on data distributions that restrict reasoning to parametric inferences. The inability to support parametric-free reasoning makes this approach inappropriate for adapting processing to arbitrary context.

- Instance-driven (D): Does adapt. Anytime adaptation (no theory is involved) and incorporating novelty (no structure is involved) are supported. However, supporting structure is essential for modeling behaviors. The inability to reason over structures renders this approach unsuitable for modeling behaviors.

Imagine (c_3): the capability to anticipate outcomes of new events without the need for experience (requires theory and data)

- Human-driven: Wonder on alternative behaviors not witnessed before (but which could have been).

- Theory-driven (A): Does not support imagination. Context-free theories can anticipate outcomes beyond experience without regard to experiential evidence. However, data-blind imagination is unable to account for novel events. The inability to imagine with theory and data renders this approach unsuitable for anticipating novel events.

- Data-driven (C): Does not support imagination. This approach learns complex tasks via trial-and-error training relying only on data. However, anticipating outcomes that have never been experienced before requires a strategy that is not theory-blind. The inability to anticipate without experience renders this approach inappropriate to support imagination.

- Model-driven (B): Does support imagination. This approach features the necessary components (structure, theory, and data) to anticipate outcomes of new events. However, because the structures impose parametric assumptions on the data, imagination is limited to compliant distributions. The inability to imagine on arbitrary distributions is not supported with this approach.

- Instance-driven (D): does not support imagination. This approach anticipates outcomes of new events relying on evidential experiences without involving any theory. However, the ability to anticipate without direct experience (seeing before) is required in imagination. The inability to anticipate outcomes without seeing examples renders this approach unsuitable to support imagination.

2.4 Concluding Remarks

What is missing to enable the analysis of human interaction in robots is:

- a setting that features natural conditions to evaluate observational learning (any input quality and quantity)
- an approach enabling human-like observational learning (causal inference, predict, adapt, and imagine)

Section 3.2 introduces the strategy to address both issues.

Chapter three

Causal Observational Learning for Robots

This chapter outlines the dissertation and introduces key concepts and terminologies to ease the reading of subsequent chapters.

3.1 Dissertation Steps

To make robots competent in the analysis of human interaction, this dissertation identifies a promising approach. To investigate the approach, we compose a framework and scenario (objective A), verify the feasibility (objective B), and evaluate the utility (objective C). Objective A instantiates the object of our study, while B and C evaluate the approach.

To enable the analysis of human interaction in robots via human-like observational learning, this dissertation pursues the following aims:

A) Establish the Framework and Scenario:

1. Establish the Framework
   Design the Model
   Setup the Inference

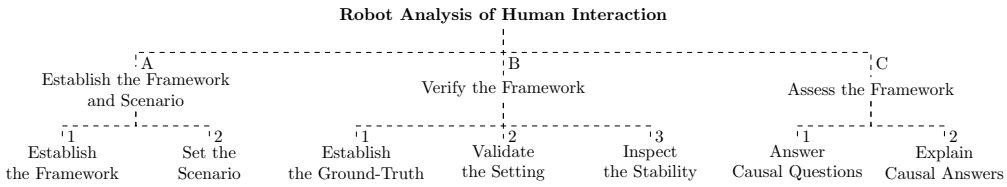


Figure 3.1: *A dissertation overview.*

Configure the Estimators

2. Set the Scenario

   Define the context

   Access the evidence

   Collect observations

B) Verify the Framework:

1. Establish the Ground-Truth
2. Validate the Setting
3. Inspect the Stability

C) Assess the Framework:

1. Answer Causal Questions
2. Explain Causal Answers

To achieve objective A, we cast robot analysis of human interaction as a problem of causal inference and set a challenging scenario for its evaluation. In particular, we formalize the approach with structural-causal models and non-parametric treatment-effect estimation. To evaluate the concept (objective B and C), we challenge the feasibility and utility of the approach on human behavior, employing an incomplete model and a limited number of samples.

There are two conditions the framework is not necessarily prepared to deal with, both being crucial in human-like observational learning. The first condition requires the framework to operate with an incomplete model given that causal inferences are only guaranteed with models fully capturing the set of relevant factors (complete models). The second condition limits the number of observations to only a few when treatment-effect estimators typically operate on large quantities of samples in the order of thousands.

The main motivation to impose both conditions is that human beings manage to acquire causal information despite incomplete models, especially children must deal with incompleteness. Moreover, people learn from even a single example, motivating to enforce few samples. These conditions resemble closely the challenge children confront, more than the challenge adults face.

When both conditions are enforced, the formalisms embedded into the framework do not necessarily operate as expected anymore. One of the conditions breaks an operational assumption of the formalism (a.k.a. unconfoundedness assumption). Consequently, bias-free inferences are not guaranteed anymore. The other condition lowers the number of samples to only hundreds thereby challenging estimation given that treatment effect typically target samples in the order of thousands.

Last, objectives B and C study the framework under both conditions. The feasibility of the approach (objective B) verifies the validity and stability of estimates, while the utility of the approach inspects the reasoning capabilities on a limited number of samples and an incomplete model.

In summary, to make robots competent in the analysis of human natural interaction, this dissertation mechanizes human-like observational learning (referring to the Social Cognitive Theory). To formalize human observational learning in machines, this dissertation selects structure causal models (SCM) with nonparametric treatment effects (A1). To challenge the feasibility and utility of the approach, this dissertation addresses a scenario involving natural activities which lack any teaching intention, involve a limited number of samples, and an incomplete model (A2). To verify the feasibility of the approach (objective B), an experimental study and two evaluations are performed (B1, B2, B3). To assess the utility of the approach (objective C), a set of questions are raised (C1) and their answers (C2) and explanations (C3) are studied.

3.2 Dissertation Strategy

This dissertation casts the challenge of observational learning as a problem of causal inference to mechanize the processing capabilities. The approach is confronted with a natural setting to evaluate the feasibility and assess the utility. The dissertation strategy shown in Figure 3.2 is described in the subsequent section.

The following provides further details on each of the strategies:

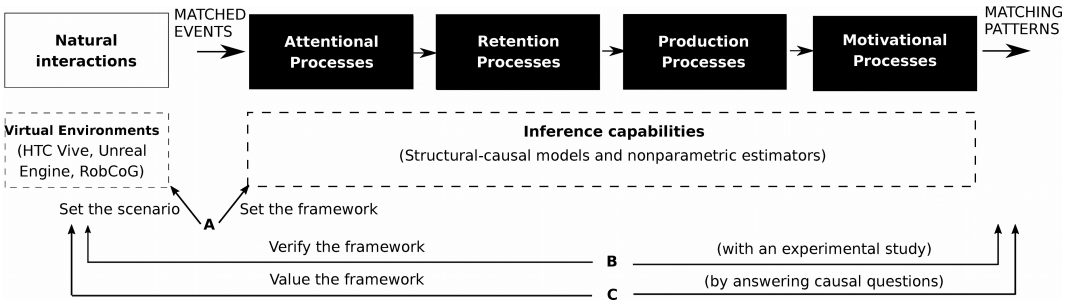


Figure 3.2: *The dissertation strategy.*

A framework for causal analysis and hypothetical reasoning (A). Having robots interpret and reason over human interactions is essential in household scenarios where service robots share workspaces with people.

Observational learning requires causal understanding of the things that are observed, and involves advanced reasoning capabilities to make the most out of observations (predict, adapt, imagine). However, conventional machine processing approaches are limited to correlational inferences (including data-driven, instance-driven, model-driven, and theory-driven approaches), they support only a subset of the necessary observational capabilities, impose parametric assumptions, and rely on supervision to ensure quality and quantity of inputs.

To make robots competent in the analysis of human interaction in natural settings, this dissertation proposes causal analysis and hypothetical reasoning via structural-causal models (SCM) and non-parametric treatment-effect estimation.

SCM solves the issue of causal inferences (instead of correlational) despite uncontrolled inputs (no supervision) because this formalism deals with bias instead of avoiding (or ignoring) as done by other approaches (solved with the identification and adjustment step). This formalism covers the necessary observational capabilities predict, adapt and imagine supporting instantaneous reasoning (solved with the do-calculus language). Non-parametric treatment-effect estimation solves the issue of robustness without imposing parametric assumptions on inferences (solved with robust estimation strategies).

The approach is composed of three components, one responsible for causal inferences and hypothetical reasoning (dowhy [SK+19]), the other for robust estimation (econml [Res19]), and a model encoding hand behavior (using a causal graph).

Chapter 4 establishes the framework describing the main components of the framework (the inferences, the estimators, the model) and scenario (kitchen-environment, table-setting activity, semantic rich data).

To challenge the approach, this dissertation investigates a social context. The scenario involves a machine (the observer), a real human being (the observee), and complex interactions (a table-setting activity). To study interactions, the focus is set on human behavior, in particular, hand behavior which could benefit robot control systems and anticipate human actions. To assess the capabilities to reason about models, a model of hand behavior is proposed and encoded with a causal graph.

A scenario is set with a near photo-realistic virtual-environment that enables people to interact via hand-held devices (Unreal-Engine, HTC Vive). The virtual activities are recorded in a kitchen scenario where users can collide against and grasp objects with the virtual-hands. To equip machines with access to virtual activities we use RobCoG, that stores virtual activities in a semantically-rich and robot-readable format.

The feasibility and utility of the approach (both further described in subsequent sections) are evaluated on a scenario involving natural activities. To avoid supervising (controlling) the inputs learners perceive, which is a common occurrence across machine-settings (supervised, semi-supervised, unsupervised), this dissertation addresses an activity lacking in teaching intentions (e.g. a person performing a table-setting unaware of being observed by a robot). Therefore, any quality and quantity of observations may result (e.g. examples can be good or bad, cover a great variety or not, and so forth). To make the scenario more natural and challenge the feasibility and utility of the approach even further, only a few samples and an incomplete model are considered. The number of observations are limited to only a few hundreds thereby challenging the estimators that operate in the number of thousands samples, and additionally, the inferences must rely on a model which does not fully capture the relevant factors in hand-behavior (breaking a crucial assumption in the formalism). Challenging the framework with the scenario under these conditions probes the potential of the approach.

A verification probing the proper functioning of the approach (B). Challenging the feasibility of the approach is crucial to probe on its practical viability. Verifying the approach on natural scenarios is especially valuable as these provide real support either against or in favor of observational abilities.

The feasibility of the approach is verified with ability to provide valid and stable

estimates despite the challenging task. The task involves recovering a non-trivial relationships of hand behavior in a table-setting activity. Two conditions make the task especially hard, the availability of samples is restricted to only a few (given that estimators operate with large samples), and the inferences must rely on a model which does not fully capture the relevant factors in hand-behavior (breaking a crucial assumption in the formalism). Challenging the framework with the scenario under these conditions probes the potential of the approach.

The framework's validity is verified with a randomized control trial, which includes the experimental design, data collection, and analysis. The framework's stability involves running refutation strategies [SK20] to destabilize the framework (analogous to sensitivity analysis in ML). The stability-tests compromise different aspects of the framework, such as the model, the data, or both.

Chapter 5 presents the verification for the validity and stability of the framework, including a randomized controlled experiment, a summary of the experimental data, and the analysis of validity and stability.

A comprehensive assessment of inferences on hand behavior (C). Investigating whether the approach is able to make the most out of observations provides insights on the suitability to automate human-like observational learning in machines. A comprehensive analysis provides insights on this matter. Harnessing observations requires the capabilities to predict, adapt, and imagine. To evaluate the utility of the framework, three types of inferences targeting human hand behavior are inspected. In particular, a set of causal questions on hand behavior are raised and answered with the framework. Informative answers hint at what robots equipped with the framework can achieve despite the challenging inference conditions.

Chapter 6 evaluates the framework's capability to answer and explain causal questions targeting two individuals free from any experimental design.

3.3 Dissertation Overview

The following sections provide an outlook on the main steps.

3.3.1 Set the Scenario

The scenario is portrayed in Figure 3.3. The figure features a robot observing a person perform a table setting. After observing several examples of a person grasping and placing objects around the kitchen, the robot is intrigued by the person's hand decision. Why does the person decide to use the right hand while sometimes the left?

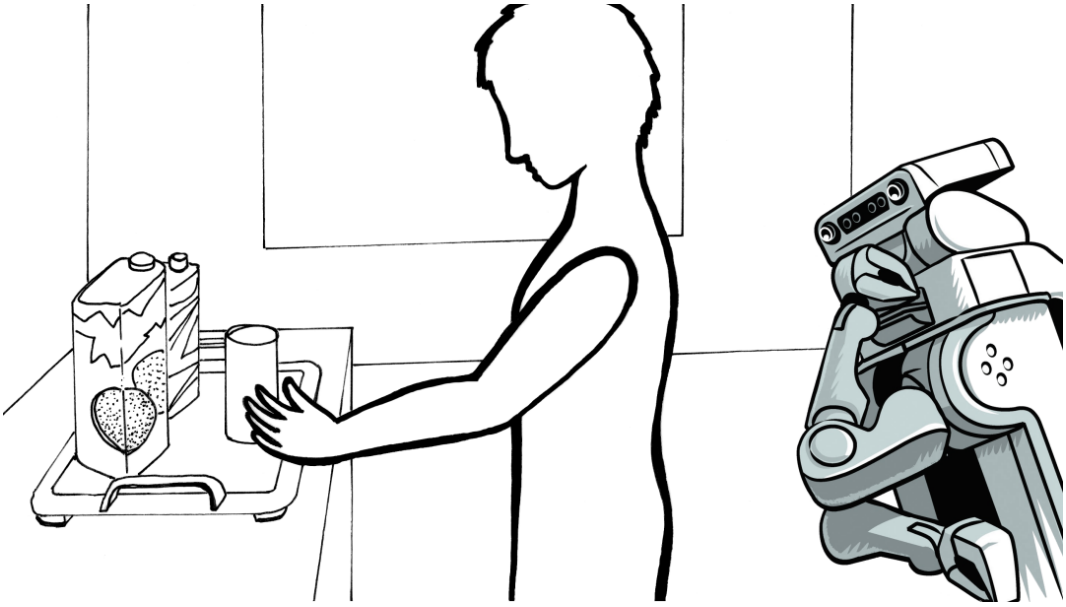


Figure 3.3: *The Scenario - A robot observes a person setting the table in a virtual kitchen: Why does this person sometimes choose the left hand over the right? Does the proximity of a target object matter in hand selection? Does the volume of an object influence the choice of hand? And if this is so, why?*

Developing robots that competently address this sort of question is a valuable pursuit. Robots capable to answer causal questions are in a position to improve themselves on behalf of others and thereby shortcut learning. For instance, robotic agents can improve their own control system by adopting strategies humans employ and thus provide human companions with better support in joint tasks.

Robots can hardly get help answering this sort of question by asking a person "What

is the policy driving your hand selection when setting a table?". People struggle to provide clear answers, and rightly so because human beings are not fully aware of decisions that underlie their own actions as part of them occur subconsciously. However, these questions are common when developing robot applications.

For questions, especially those relating to human behavior, robots must seek answers on their own. We believe this is possible because children do it. They direct their attention to role models and answer in this way countless questions on their own.

The setting where children manage to learn is the following: Children live a rather contemplative life as they still struggle to control their own limbs and thus exerting changes to the environment is challenging. Life unfolds in front of their eyes, they hardly regulate the inputs they get to witness (examples come in any distribution). Sometimes children are privileged to see examples repeatedly (large quantities of samples) while other times a single one (few samples) or even none (missing samples). Life does not favor a particular order in which examples are presented to them, nor are examples necessarily of a certain type (only successful), and neither the number of examples across classes is balanced (unbalanced). This characterizes (in a simplified manner) the setting children must deal with at early age. Regardless of the input distribution, children manage to process information and learn.

The settings robots must face are not much different than those children deal with. Robots are still clumsy when interacting with their environment but excel in contemplative tasks. Machines can store massive amounts of examples and recall them with the highest fidelity. In spite of this advantage, the robots struggle in the settings where children thrive. Unlike children, the inputs that robots process require a great deal of care. For instance, the number of examples needs to be massive, balanced out across classes, annotated with hints (labeled), of a certain type (correct), follow a certain distribution (parametric), distributions need to be stable, and more.

The scenario set in this dissertation aims to reflect (to a certain extent) some characteristics of the settings where children learn. The purpose is to challenge machine processing with more realistic inputs. The data originate from natural activities performed in a virtual kitchen. People who perform activities receive no other instruction than completing a table-setting. They are free to set the table for any number of guests or only for themselves, choose any meal and silverware, consider any order, place and arrange objects anywhere in the kitchen, use the left or right hand, stand closer or further away from targets, and so forth. Care has been taken to preserve raw data as much as possible. Considering only minimal transformations (binarize with median split) but no complex data transformations (feature engineering). Data is not filtered to examples that are correct and classes are not balanced out.

The virtual kitchen is fully furnished including a table, kitchen counters, drawers, and refrigerator. Moreover, it is equipped with several sizes of cups, glasses, bowls, and silverware. Furthermore, some drawers and the refrigerator are stuffed with objects such as cereal cartons, bread, salt, milk, juice, and more.

A person interacts with the virtual environment through VR devices. Users wear a headset that renders the near photo-realistic virtual kitchen, and virtual hands through which the interaction is exercised. The virtual hands can collide with objects and furniture. Pressing a button on the hand-controller triggers a grasp which allows objects to be held with either of the virtual hands.

Robots have access to the activities exercised in the virtual kitchen. The motions and interactions of users are recorded with a game engine and stored in a robot-compliant format (narrative episodic memories).

3.3.2 Establish the Framework

Chapter 4 describes the framework that confronts the scenario. Three components make the framework operational, the causal reasoner, the estimators, and the model. These elements as a unit compose the object studied in Chapter 5 (feasibility) and Chapter 6 (utility).

The framework incorporates desirable properties to confront the scenario. It features inferences that are free from bias (causal), end-to-end non-parametric processing (any distribution), and supports advanced reasoning (predict, adapt, transport). The question remains whether these properties realize in the scenario described earlier.

Even though the scenario is challenging enough two conditions made by the framework are strong. One relates to the completeness of models and another to the number of samples.

First, the framework guarantees bias-free estimates only when the model captures all the relevant factors (e.g. the set of common confounders is complete). Second, estimation typically operates on large data (with thousands of samples).

People, however, make decisions with incomplete information (models) and can learn from even a single observation. This motivates the study of the approach under these two conditions the framework is not designed for.

The first condition is enforced with the design of a model which does not fully capture all relevant confounders while the second condition is achieved by choosing estimators that operate with samples in the order of thousands but only hundreds

are provided to them.

Chapter 4 introduces the three main components that instantiate the approach. Section 4.1.2.1 defines the model that encodes the assumptions on hand behavior. Section 4.1.2.2 presents the causal reasoner which deals with confounding bias (identify and adjust) and supports three types of inferences (do-calculus). Section 4.1.2.3 describes the estimators responsible for the treatment-effect estimation on the challenging data.

3.3.3 Verify the Feasibility

Chapter 5 verifies the feasibility of the framework to derive correct and stable inferences despite the challenging scenario and the two conditions enforced (incomplete model and access to few samples).

Correct and stable are complementary and desirable for the inferences to exhibit. Inferences are correct when the framework derives valid estimates, and stable when estimates are consistent despite varying conditions.

Claiming both means that, the model, the causal reasoner, the estimators, once set, provided data and a target quantity, valid and stable inferences can be computed.

The framework provides valid inferences when the results reflect the expected outcome on ground-truth data. Failing to infer the ground-truth effects, raises a red flag for the validity of the framework.

The framework offers stable inferences when strategies are unable to introduce variations on the model and data to destabilize estimates. Variations to the model that give rise to instabilities manifest that assumptions encoded in the graph do not fully explain the evidence in data. Similarly, the undesirable instability that arises when varying the data indicate sensibilities to samples (e.g. to the order of samples).

The overall strategy to assess the feasibility of the framework consists of first establishing a ground-truth (Section 5.1), then verifying correctness against the validation-set (Section 5.2.1), and assessing the stability under varying conditions (Section 5.3.2).

Verifying for the correctness of the framework involves a randomized controlled trial (Section 5.1.1 and Section 5.1.2), collecting experimental evidence (Section 5.1.3), and the analysis of the results (Section 5.2.1 and Section 5.2.2).

Verifying for the stability of the estimates involves a set of strategies that intent to destabilize the framework (Section 5.3.1), and analyzing the impact these strategies have on the framework (Section 5.3.2 and Section 5.3.3).

3.3.4 Assess the Utility

Chapter 6 questions the ability of the framework to reason over hand behavior on the scenario set forth to investigate covering three types of inferences.

The evaluation consists in formulating a set of causal questions addressing hand behavior and answering these questions with estimates derived by the framework.

The hand behavior of two subjects is studied with the framework. Both perform natural hand manipulations in the context of virtual table settings (uncontrolled observations of human actions).

Three levels of inferences are covered (types of causal questions). The levels correspond to the ladder of causation (Figure 3.4). The inferences at the lowest level (type *what*) are suitable for prediction tasks. On the other hand, inferences at the first level (type *what-if*) adapt observations from one domain to another. The third level (type *what-would-if*) enables reasoning over unseen scenarios.

Exemplary questions targeting hand behavior are:

1. Does hand-distance affect hand-selection? (predict)

2. Would hand-selection be right-hand if we make sure that the hand-distance is far? (adapt)

3. Would hand-selection be left-hand had the hand-distance been close, given that the hand-selection is in fact right-hand and hand-distance is far? (imagine)

The interactions that are covered comprise person-environment (e.g. hand-object, furniture-hand) and person-person (e.g. distance-hand).

The chapter evaluating the utility of the framework (Chapter 6) studies the answers to causal questions on hand behavior. This chapter describes the access to activities in a virtual environment for robots (Section 4.2.2), summarize the data on uncontrolled hand-behavior of person A and B (Section 6.1.2), formulate causal questions on hand behavior (Section 6.1.3), report and analyze the answers framework derives (Section 6.2).

3.4 Dissertation Concepts

This section exists because the core concepts and methods this dissertation builds upon are developed outside the field of computer science, in particular epidemiology, economics and, experimental sciences.

3.4.1 Causal Inference

The landscape of causal inference can be divided into causal discovery and causal-effect estimation. Causal discovery aims to recover causal structure from data, i.e. building models from data. The estimation of causal effects on the other hand, aims to evaluate expressions using data, i.e. estimate causal quantities. This dissertation is framed within the latter also known as a treatment-effect estimation.

3.4.2 Causation not Correlation

A crucial distinction is the one between correlation and causation. Both make assertions over relationships, but one makes a weaker statement than the other. The weaker is correlations, which hints that A and B are related simply because evidence A and B co-variate. Correlations determine that A and B co-vary but do not distinguish whether A changes B or B changes A. Therefore, correlations are regarded as weak because these lack the power to convey direction. Correlations assert associations and not more.

The importance of direction is evident in decision-making. Suppose the direction is unknown, it is unclear which of the variables (A or B) should modulate to achieve the desired change. However, when the direction is known (A causes changes in B) it becomes clear that to influence B one modulates A.

Besides lacking the power to convey direction, another noteworthy issue is that correlations can be temporal. These can appear and vanish, often regarded as spurious.

Direction cannot be recovered with correlational approaches [Pea14]. This includes data-driven approaches such as statistics and machine learning. In statistics, practitioners are often reminded about this limitation with “correlation does not imply

causation (direction)". To express this limitation differently, causation cannot be inferred from data alone.

This exhibits a fundamental limitation in correlational approaches which provide inferences that are based on co-variation. These approaches operate at the lowest level in the ladder of causation (Figure 3.4).

Unlike correlation, causation conveys direction and is thus a stronger statement. Causal interpretation distinguishes whether A changes B, or vice versa. Causation can also implicate simultaneous co-occurrences such as A affecting both B and C, being A in this particular case a common cause for B and C.

Robot analysis of human interaction requires causal interpretation, not correlation.

3.4.3 Simpson's Paradox, or what is confounding bias

Simpson's paradox [Sim51] is introduced here to clarify the meaning of confounding bias. This phenomenon manifests with a contradiction when interpreting data. It occurs when identical data lead to opposing conclusions. Table 3.1 shows an example of this where data separated into groups contradict the whole population trend. The contradiction is highlighted in the table, On one hand, the combined trend suggests treatment (87%) and on the other, the segregated trend suggests no-treatment for group-2 (87%). Hence, depending on how data is summarized, reversed decisions emerge.

Gender	Treatment	Control
Female	81 out of 87 (93%)	192 out of 263 (73%)
Male	273 out of 350 (78%)	234 out of 270 **(87%)**
Combined	354 out of 437 **(81%)**	426 out of 533 (79%)

Table 3.1: *An example of Simpson's paradox. A clinical trial aims to determine whether a drug positively affects people who received the pill (treatment) against those who received a placebo pill (control). The numbers indicate the recovery rate.*

Any raw data is potentially affected. For instance, reversals affect a random uniformly distributed binary table of size $2 \times 2 \times 2$ with a probability of 1/60 [PP09]. The probability that reversals would occur at random in path models involving three variables is approximately ten percent [Koc15].

Data on hand manipulation is no exception to reversals either. Table 3.2 confirms the existence of such contradictions in hand manipulation. This table shows how

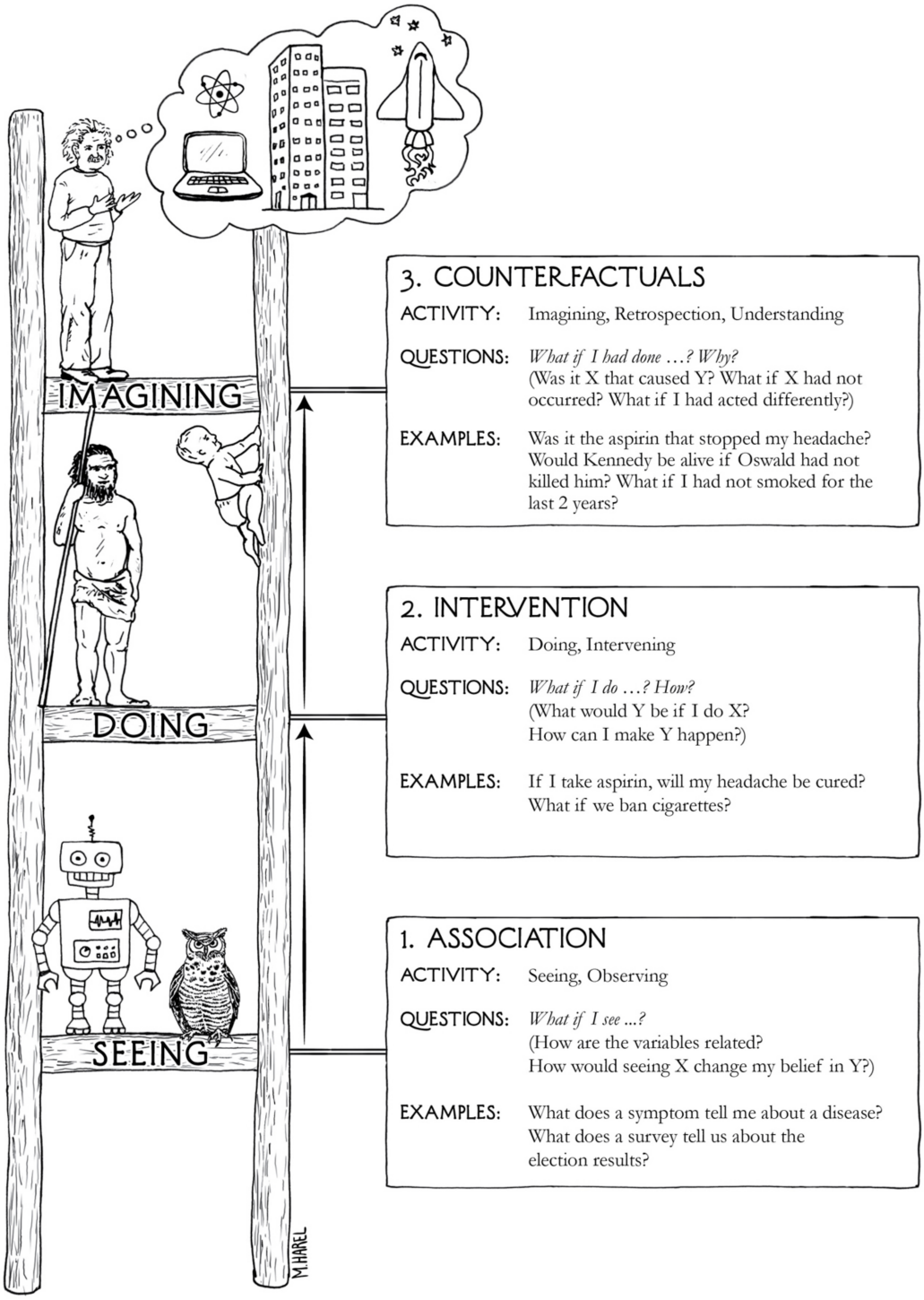


Figure 3.4: *An illustration showing the ladder of causation from the Book of Why [PM16] illustrated by Maayan Harel (www.maayanillustration.com)*

the numbers reverse when conditioning on the size of an object. Reversals have also been observed when conditioning on other covariates such as surface.

Object-volume	Hand-dist. close	Hand-dist. far
0.000082	0.79	**0.85**
0.000775	0.57	0.55
0.001379	0.16	0.14
0.001398	0.03	**0.41**
0.001911	0.70	0.31
0.001934	0.77	**0.95**
Combined	192 out of 270 (**0.71**)	71 out of 114 (0.61)

Table 3.2: *The confounding found on hand behavior. The numeric values indicate the rate a person chooses to grasp with the right-hand. Hand-distance indicates whether a grasp occurred close or far away w.r.t. the body. The trend in hand-selection reverses when conditioning on the volume of an object. The segregated trend disagrees with the combined trend, with three positive slopes contradicting the negative population trend.*

Like in the table showing the Simpson's paradox, the combined data conflicts with the segregated. Consulting the combined data, suggest hand-dist close (0.71 over 0.61), while the case of the segregated suggests differently depending on the volume of the object. The question is whether the robot should consult the combined or the segregated data. There is no way to figure out an answer to this question using data alone [Pea09].

Reversals are relevant because robots are expected to provide causal interpretations of contents stored in their memories, i.e. essentially frequency data such as the one shown in Table 3.2. Regardless of how memory is encoded in a robot platform the computation of expectancy or probabilities are affected. Ignoring reversals when processing data leads to potentially disparate interpretations where programmers or robots have to guess which of the trends in the table to follow.

These contradicting interpretations are not necessarily temporal in data. Reversals can also be systematic [Pea09], for instance, when introducing additional variables to the analysis trends can reverse back and forth. More data does not avoid this issue. Data alone cannot solve confounding [Pea14]. The solution to this paradox is appropriate modeling as proposed in [Pea00].

The Simpson's paradox is a special case of the co-variate selection problem known as confounding bias, which comprises higher dimensions and complex relationships. In the literature, the term bias relates to errors in numeric values such as noise in sensor measurements. Confounding, on the other hand, is the bias that occurs at the

level of relationships.

3.4.4 Experiments, Observations, and Quasi-Experiments

The distinction between experimental and observational is important.

- **Experimental**: The great value of experimental studies is the power to assert the causal direction between variables. Here, the evidence (data) needs to be controlled to ensure valid interpretations. Experimental studies achieve causal interpretations but require controlling the environment (conducting experiments).
- **Observational**: The value of observational studies is the power to find associations between variables. Here, the sources of evidence are uncontrolled. Observational studies attain correlational interpretations without the need to control the environment (by simple observation without running experiments).

Experimental studies offer a stronger interpretation than observational studies. In turn, observational studies afford arbitrary data while experimental studies are restricted to data originating from settings that can be controlled.

Experimental studies are the gold standard in the scientific method. Scientists dedicate important efforts to control the source of evidence because doing so grants causal interpretation. Controlling settings to gain experimental data is essentially a strategy which avoids confounding bias leaking into studies. The literature hints on the nature of data that originates from controlled environments (i.e. experiments) with controlled data, experimental data, validation data, or ground-truth data.

Because experimental studies require experimenters to exert control over some aspect of the environment, it is unclear whether robots could gain causal interpretations by running experimental studies themselves just like scientists do. This poses a hurdle in the development of robots that seek a causal understanding of their environment.

Robots might struggle to craft their own experimental evidence but excel when it comes to processing uncontrolled data. Machines can process any data but because the evidence is uncontrolled inferences remain at the correlational level. To put it differently, robots run observational studies. The literature hints at the nature of data originating from uncontrolled settings with terms such as non-experimental data, observational data, uncontrolled data, or raw data. The scientific community does

not accept causal interpretations of this kind of data, as it is regarded as confounded (or biased).

People on the other hand seem to cover the full experimental and observational spectrum. It is still not well understood how people gain causal information but the capability to draw interpretations on either uncontrolled (plain observations) or controlled evidence is unquestionable.

Machines that resemble people on this account can become more competent. Robots that acquire causal interpretations from observational evidence are unrestricted to laboratory conditions. Robots relying on causal interpretation are superior to machines that rely on a correlational understanding of the world.

Experimental studies are not always feasible. For instance, an experiment cannot compromise a subject's life. When scientists cannot run experiments, and still intend to gain causal interpretations, they attempt an alternative strategy (a.k.a. Quasi-experiments). The strategy consists in identifying and adjusting for variables that could potentially bias their studies. These steps are necessary because the evidence is uncontrolled. Quasi-experiments are observational studies that are treated as if these were experimental studies. This is granted only because confounding has been managed.

Intuitively, quasi-experiments can be understood as a thought experiment exercised on any evidence (either uncontrolled or controlled) to achieve causal interpretations. The approach investigated in this dissertation incorporates the capability to run quasi-experiments. The aim is to equip robots with the capability to draw causal interpretations on any evidence. Nonetheless, the two kinds of evidence are employed in this dissertation. Experimental data to validate the framework, and non-experimental to assess causal inferences.

3.4.5 Missing Data, or what is counterfactual

A core notion in causal inference is the one of counterfactual, or the missing data problem. This difficulty is described as follows. The fundamental problem of causal inference [Hol85] is that, one gets to ever observe one outcome and fail to observe what would have been if things had been different (a.k.a. the counterfactual world). In other words, for any observation, one can ever observe one outcome and not alternative outcomes (because only one of them occurs).

Dealing with this problem lies at the core of treatment-effect estimation. The chal-

lenge is to derive causal conclusions despite the missing data on the counterfactual. Estimators reconstruct the missing counterfactual considering factual outcomes and are further described in Section 4.1.2.3.

3.4.6 The Unconfoundedness Assumption, or completeness

A key assumption made in causal inference is the one of unconfoundedness which relates to completeness. To guarantee inferences that are free from bias, unconfoundedness must hold. Unconfoundedness requires all relevant factors to be known (i.e. modeled) such that when disentangling effects none of them is missing. Essentially, unconfoundedness requires that all common factors that potentially confound the effect of X on Y are encoded in the model.

Human beings, however, seem to acquire causal information despite incompleteness. People make decisions every day without taking all relevant variables into account. Children, who lack comprehensive knowledge, progressively build causal information.

To investigate causal inference under such conditions, this dissertation breaks the unconfoundedness assumption with a model not fully capturing the domain of interest, and consequently, not including all common confounders (Section 4.1.2.1). The question is whether the framework operating under incomplete models is still able to derive valid inferences, which is the main concern in Chapter 5.

3.4.7 Structural Causal Models and Potential Outcome Framework

There are two major alternatives to causal inference, the potential outcome framework [Rub74] and structural causal models [Pea00]. The potential-outcome framework is suitable in those cases where the causal structure of all covariates is unknown, or when approximations are desired (spare the need for modeling).

Structural causal models, on the other hand, are convenient for modeling the assumptions on the data distribution (encoded with graphs). Besides conveniently representing knowledge with graphs, another distinctive advantage is that the structural account can determine whether a causal quantity is identifiable or not.

This dissertation chooses structural causal models for the following reason. A causal graph serves as a debugging tool when developing applications given that robotics

already deals with multiple complex issues at once. Graphs allow to pinpoint meticulously adjustments sets and support the development of increasingly complex models. Another argument favoring graphs is the automatic generation of stability tests which serve as an automated diagnostic tool. Last, knowing when causal quantities are not identifiable is a great asset to have in the robot-human interaction or development phase.

This conclude the dissertation overview. The following Chapter 4 establishes the framework and scenario. Subsequently, Chapter 5 and Chapter 6 evaluate the framework.

Chapter four

Establish the Framework

This chapter describes the main components of the framework and defines the scenario. The main components of the framework are the causal graph, machinery for causal inference, and estimators. The scenario essentially defines the input for the framework. It includes the source of evidence, environment, activity, data collection, and pipeline that enables robots with access to virtual activities.

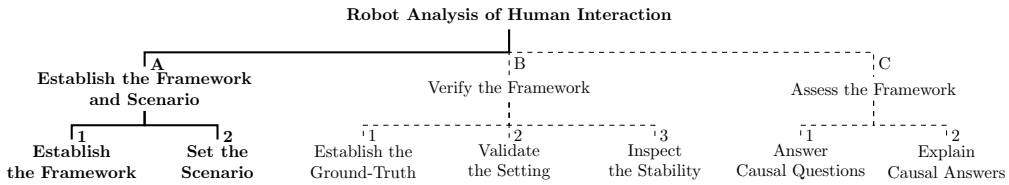


Figure 4.1: *The strategy to establish the framework and scenario.*

4.1 The Framework

The framework is presented in three sections. First an overview of the approach is provided in Section 4.1.1, then the main components of the framework are introduced in Section 4.1.2, and last the capabilities of the framework are highlighted Section 4.1.3.

The overview provided in Section 4.1.1 introduces the core steps underlying the approach of structural causal models (SCM) and facilitates the reading of subsequent sections.

Each of the introductory steps provided in the overview relate to one component of the framework, which are described more formally in Section 4.1.2.

The capabilities of the main components working as a unit are made explicit in Section 4.1.3, and two strong assumptions under which these operate are emphasized.

4.1.1 An Overview

The overview introduces three steps, the modeling of assumptions, identification and adjustment, followed by effect estimation. Each of these introductory steps relate to one component of the framework and each is further described in subsequent sections.

Modeling Assumptions

In comparison with the Bayesian view, causal inference adopts a different perspective. The Bayesian view ascertains a hypothesis with evidence. Often expressed with Bayes rule in the form Probability(hypothesis|evidence). Instead, causal inference explains evidence with hypothesis adopting the converse form of Bayes rules Probability(evidence|hypothesis).

In causal inference, hypotheses are assumed to be correct and thus regarded as assumptions rather than hypotheses. Assumptions are key building blocks in causal inference.

A compact representation to encode assumptions are graphs also referred in the literature as causal graphs or causal diagrams, in particular direct acyclic graph (DAG). The causal graph does not only encode assumptions but also participates in the inference process (potentially modifying such graphs).

A minimal causal graph is shown in Figure 4.2 comprising three nodes and three edges, where nodes correspond to concepts and edges to assumptions.

Concepts are abstractions which typically refer to features of the world (e.g. symbols X,Y,Z in the graph), and the edges assert a potential connection between two

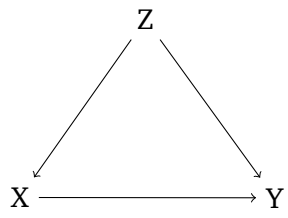


Figure 4.2: *A causal graph.*

concepts where the direction is given by the arrow (e.g. arrow from X to Y). An arrow pointing from X to Y asserts that the former may affect the latter (e.g. X affects Y). Nodes and arrows are the only elements required in modeling.

Modeling consists in identifying the data generating process that explains part of a domain. For example, Figure 4.2 can model a person grasping an object with the hand as follows. Let concept X represent the arm extension (near or far), Y refer to the choice of hand (left or right), and Z modeling the size of an object (small or large).

In this model, the concepts are the following:

$$\begin{aligned} &\text{X} \quad \text{hand-distance (D): person grasps near or far} \\ &\text{Y} \quad \text{hand-selection (H): person uses the left or right hand} \\ &\text{Z} \quad \text{object-volume (OV): person targets object of volume small or large} \end{aligned} \tag{4.1}$$

The edges in Figure 4.2 specify that X may affect Y and that Z might affect X and Y. The arrow from X to Y models a relationship between hand-distance and hand-selection. Namely that arm-extension (i.e. hand-distance) is one potential reason influencing hand-selection Y. This influence occurs, for example, when the person prefers to target distant objects with a particular hand (e.g. the dominant hand). The other arrow pointing to Y comes from Z stating a potential relationship between object-volume and hand-selection. This relationship occurs, for example, when the size of an object makes a person opt for the stronger limb (hand-selection). Moreover, the third arrow to explain in the graph points from Z (object-volume) to X (hand-distance). This arrow expresses that, for instance, large objects could restrict a person to shorter arm extensions (grasping closer to the body). Last, there is no arrow pointing from X to Z because hand-distance cannot alter the scale of an object, and neither can hand-selection affect the scale of an object (i.e. the arrow from Y to Z is also missing).

The relationships encoded in a causal graph make only weak assumptions. First, an arrow states that one concept affects another potentially but not for certain. Only the

possibility is hinted at by an arrow. Second, none of the arrows impose any functional form on the distribution. Only the topology of the graph is relevant. Hence, data is irrelevant at this point and enters the scene later in the estimation stage.

Causal graphs represent durable relationships. That is, the data generating process encoded in a graph is expected to endure across scenarios. For example, the potential influence of hand-distance (X) on hand-selection (Y) does not only hold for the activity of table-setting but also for activities such as tidying up the kitchen or meal preparation. Graphs are expected to endure across scenarios.

This concludes the introduction of causal graphs and modeling. The minimal model on hand behavior introduced in this section is extended and further described in Section Causal Graph. The following section provides the intuition on how causal graphs are used to confront confounding bias.

Identify and Adjust

Causal graphs are essential to deal with confounding bias. Building on the previous example consider again the graph of Figure 4.4, and in particular the arrow that points from X to Y which interprets as the question "Does X (hand-distance) affect Y (hand-selection) when targeting an object of Z (volume size)?"

Attempting to answer whether X affects Y by simply inspecting the values in Y can lead to a wrong (biased) answer. This is so because values observed in Y are not necessarily due to X alone, as Z can also be responsible of the values observed in Y. Even worse, Z can also indirectly affect Y through the path $Z \rightarrow X \rightarrow Y$ (i.e. mediated through X). Ignoring these complications, can lead to incorrect conclusions such as asserting that X affects Y when in fact Z (not X) is the major driver of change in Y. This is known as confounding bias. The bias that occurs at the level of relationships. To put it in terms of the domain, the wrong conclusion due to confounding bias would claim that hand-distance (X) affects hand-selection when in fact object-size (Z) truly drives hand-selection (Y).

The challenge to answer causal questions then is distinguishing the true change that X exerts on Y, given that the values observed in Y are potentially due to other reasons than X. It is easy to foresee where complex domains (graphs) lead to, and why automated path analysis is necessary.

To ensure correct conclusions, the strategy is to isolate the undesirable effects of confounding. The task consists in deriving paths from the graph (a.k.a. estimands)

which are free of bias. The process involves identifying confounders and adjusting when necessary.

The first step in the process identifies which variables need an adjustment (i.e. control). This step is selective because simply adjusting for all confounders can introduce instead of remove bias. Recall the classic example in the literature which warns against the control of colliders that risks confounding the expression by opening instead of closing a backdoor[1]. The identification step is responsible for providing a minimal adjustment set containing nodes (variables) that when controlled would not introduce bias anew. For instance, the minimal sufficient adjustment set to identify the effect of X on Y in the graph of Figure 4.2 without introducing bias anew is $\{Z\}$. In this case, object-volume (Z) is identified as an adjustment candidate because it affects hand-selection and hand-distance which hinders disentangling true change when analyzing the effect of X on Y. Not controlling this factor could lead to wrong conclusions.

Given the adjustment set, the next step in the strategy consists in controlling the variables that are identified. Graphically this step corresponds to deleting arrows from the graph. For instance, applying the adjustment step to $\{Z\}$ deletes the arrow pointing to X. This modification in the graph ensures that undesirable effects of path $Z \rightarrow X \rightarrow Y$ are blocked. Technically, the adjustment operation turns dependent variables into conditionally independent factors. The intuition underlying this operation is ensuring that when modulating the factor of interest X other variables refrain from co-variation (thereby allowing a ceteris paribus analysis).

The identification and adjustment steps have been mechanized by the backdoor-criterion [Pea00]. Given a causal diagram and a relationship of interest, the backdoor-criterion derives a bias-free expression applying the steps of identification and adjustment introduced earlier. For the graph of Figure 4.4, and the query whether X affects Y, the following estimand is derived:

$$\frac{\partial}{\partial X}(Expectation(Y|do(Z))) \tag{4.2}$$

The estimand (4.2) explicitly shows the adjustment operator applied to Z denoted by $do(\cdot)$. Because Z has been identified, the do-operator controls this variable to account for the undesirable effect of confounding bias.

Estimand (4.2) is simple because only a minimal graph is involved. However, more intricate assignments of the do-operator are possible. In general, the backdoor-criterion (and extensions of it) handle arbitrary graphs. Section The Components

[1] A collider is one of three types of junctions in path analysis [Pea09].

defines the backdoor criterion more formally and additionally introduces the do-calculus which supports the three levels of inferences.

In summary, given a causal graph and a relationship of interest, the identification and adjustment step derives an expression that is free of confounding bias. The next step evaluates this expression using data.

Estimate Effects

The estimand derived by the backdoor-criterion provides the appropriate condition (bias-free) to initiate the search for an answer in the data.

The estimation step computes a quantity known as the causal effect. This means that, continuing with the example described so far, given Z has been controlled satisfying the backdoor-criterion, the causal effect of X on Y is computed by a difference in expectations [2]

$$E(Y|do(X = x')) - E(Y|do(X = x'')) \tag{4.3}$$

where x' and x'' denote two binary levels (e.g. hand-distance close or far). The general case for non-binary feature levels is expressed differently and further described in Section Treatment Effect Estimation.

The challenging aspect of computing causal effects is that levels x' and x'' are both never known, only one of them is known while the other remains missing. This difficulty is known in the literature as the fundamental problem of causal inference [Hol85] and relates to a missing aspect of the world that cannot be observed. For example, when the person decides to target an object farther away from the body then $x' = far$ is a fact, and what is also true is that the potential world where grasping close to the body is not realized, i.e. the counter-factual $x'' = close$. Essentially, one of the terms in equation of causal effect (Definition 4.3) is never known. To evaluate causal effects from data estimators need to deal with missing data (the counterfactual world). One way estimators accomplish this is by reconstructing the counterfactual world using existing data.

Estimators compute the causal effect using statistical methods. The methods range from simple regression to more advanced techniques such as doubly robust learners introduced in Section Treatment Effect Estimation.

[2] As defined by [RR83] but written in terms of the do-operator [Pea00].

Suppose regression is employed to evaluate the causal effect (Definition 4.3) on a given set of data. The outcome of the estimation can either result in a positive, negative, or null effect. Positive when $E(Y|do(x')) - E(Y|do(x'')) > 0$, negative when $E(Y|do(x')) - E(Y|do(x'')) < 0$, or null when zero (i.e. confidence intervals contain the zero).

The sign of a causal effect indicates the direction of how X affects Y. Building on the example described so far, a positive effect of hand-distance (X) on hand-selection (Y) would indicate that the person employs the right hand when targeting objects farther away from the body. On contrary, a negative causal effect indicates that the right hand is rather engaged near the body. In the case of the absence of effects (a.k.a. null-effect), the evidence does not favor nor refute the existence of an effect.

The estimation step is further described in Section Treatment Effect Estimation, introducing the definition of causal effect adopted in the framework, as well as the selection of estimators and their parameters.

This concludes the overview of the main components involved in the inference process comprising the modeling of assumptions (causal graph), the steps to derive bias-free expressions (structural causal models), and the evaluation of estimands using data (treatment-effect estimation). The following sections provide further details on each of these components.

4.1.2 The Components

Each of the sections described earlier relate to one of the three components of the framework. Figure 4.3 illustrate the main components of the framework for which subsequent sections provide further details. In particular, Section Causal Graph introduces the modeling of the graph, Section Causal Inference describes the machinery enabling reasoning, and Section Treatment Effect Estimation presents the component responsible for data processing.

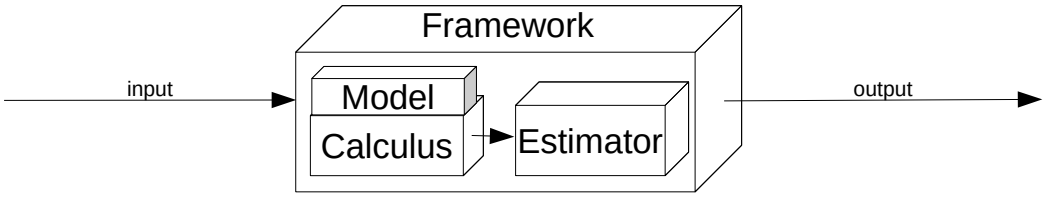


Figure 4.3: *The components of the framework.*

Causal Graph

Recall from Section An Overview that causal graphs encode assumptions about the world and are essential to account for bias. These assumptions express potential connections ("X could affect Y") rather than strong statements ("X affect Y"). Assumptions are potential relationships that are expected to endure across scenarios. For the construction of causal graphs, such relationships must be identified in the domain of interest. The identification of durable relationships in hand-manipulation is the main concern in this section.

Durable Relationships in Hand Behavior

The identification of durable relationships in hand behavior focuses primarily on two aspects which underlie nearly every hand manipulation. These are, the extent a person decides to stretch the arm, here referred to as hand-distance (e.g. near or far), and the choice of a particular hand referred to as hand-selection (e.g. left or right). Being the question whether the one factor influences the other.

Multiple factors can potentially influence this relationship, and thus confound observations. In the following a series of such factors are scrutinized.

Hand-dominance is perhaps the asymmetry most studied in human beings. This personal trait in humans can dictate a preference in hand-selection. Hand-dominance could also (affect hand-distance) be the reason why people accommodate themselves in a position such that targets placed further away are grasped with a particular hand thereby influencing arm extension (hand-distance). Hence, hand-dominance is a factor potentially affecting two variables (a.k.a. common confounder), in this case, hand-distance and hand-selection.

Another factor is the volume of an object (object-volume). This factor can potentially influence hand-distance and hand-selection as follows. The volume of a target object could hinder a person from stretching the arm too far out (avoiding collisions). Moreover, object-volume could favor the selection of one (stronger) hand over the other (the weaker arm of the person). Similar reasoning applies to other physical properties such as shape, weight, and specialized features like handles.

Another relevant factor potentially driving change in hand-distance and hand-selection is skill. People acquire skills through life. The skill required to manage a tool can condition hand-selection as follows. For example, using the non-dominant hand

to cut with a knife is difficult and thus people opt for the more skillful hand instead. A knife can also be the reason why people grasp this object farther away from their body to ensure safety (whereas a spoon would not). Hence, the identity of an object (object-category) is a common confounder to hand-selection and hand-distance.

Environmental elements can also affect hand behavior. Furniture imposes spatial constraints, for instance, a table imposes the height at which a person manipulates objects. Moreover, a piece of furniture can either facilitate (e.g. a table) or constrain access for a person (e.g. reaching deep inside a refrigerator). Beyond spatial constraints, the property of furniture, such as whether a surface slide (e.g. pulling a drawer) can also potentially influence hand manipulation, just like the surface-category (e.g. table, kitchen counter, refrigerator's shelf.) and surface-inside (e.g. difficult access or not) are also identified as common confounders.

Because so many factors can influence hand-selection, it is not straightforward to determine whether hand-distance (truly) influence hand-selection. To assess whether such a causal connection holds, researchers would certainly conduct experiments to exclude any bias getting into the study (Chapter 5 describes such an experiment). However, robots cannot run experiments, especially not on people. Unlike the experimental strategy researchers have available to them, robots confront the challenge to determine whether causal relationships hold from plain observations without regard to experimental evidence.

The following list identifies durable relationships in hand-manipulation:

1. **Hand-distance (D) on hand-selection (H)**: Hand selection occurs according to object proximity. For example, when targeting an object for which the left hand is closer than the right, the left hand is chosen. This behavior is supported by the kinesthetic hypothesis [GR00, GH04].

2. **Hand-dominance (DO) on hand-selection (H):** The preference of one hand over the other directly influences hand selection. [BRDB96, MRBR04] shows that the dominant hand is preferred for close to mid-line targets.

3. **Hand-dominance (DO) on hand-distance (D):** The preference on a particular hand could lead a person to extend the dominant arm further away than the non-dominant thereby affecting the proximity at which an object is grasped.

4. **Object-category (O) on hand-distance (D)**: The identity of an object influences the distance a person would stretch their hand relative to their body when grasping objects. For example, dangerous objects (eg. hot soup or a knife) are manipulated farther away from the body than fragile objects.

5. **Object-category (O) on hand-selection (H):** The identity of an object can lead to the selection of a particular hand. For example, tools such as knives or spoons are typically held with the dominant hand. [LF06, GFS15] show that the dominant hand crosses the midline more often when a task involves the manipulation of a tool.

6. **Surface (S) on hand-distance (D):** Surfaces can affect the distance at which a person manipulates objects. For example, handling a pot on a hot stove lead to safety distance to avoid burns.

7. **Surface (S) on hand-selection (H):** Spatial layouts can lead to hand selection. For example, a refrigerator's door opens comfortably for the right hand but less for the left.

8. **Object-volume (OV) on hand-selection (H):** The size of objects can affect hand selection. Large objects, but still manageable for a single hand, are likely handled with the more dexterous hand. [WS18, SW18] report that the non-dominant hand is for smaller objects while the dominant hand is for larger.

9. **Object-volume (OV) on hand-distance (D):** The size of objects could affect the extent to which a person stretches the arm. For instance, the handle of a frying pan could lead to a near body grasp.

10. **Surface-sliding (SS) on hand-distance (D):** Sliding surfaces can affect the distance at which a person takes an object. For example, when reaching into a sliding drawer, pulling the drawer open could affect hand distance.

11. **Surface-sliding (SS) on hand-selection (H):** Surfaces that slide can potentially drive hand selection. For example, preferences opening containers with a particular hand and reaching into the container with another or the same hand.

12. **Surface-inside (SC) on hand-distance (D):** Target objects located inside of containers are harder to reach than those on unconstrained surfaces. For example, reaching for a milk carton stored deep inside the refrigerator demands stretching the arm wide while on a table not.

13. **Surface-inside (SC) on hand-selection (H):** Targeting objects placed inside of containers can lead hand-selection. For example, when grabbing an object inside of a drawer, one hand opens the drawer while other reaches into.

These relationships are encoded in the causal graph of Figure 4.4. The nodes lacking ancestors in the graph are known as exogenous to the model because these are

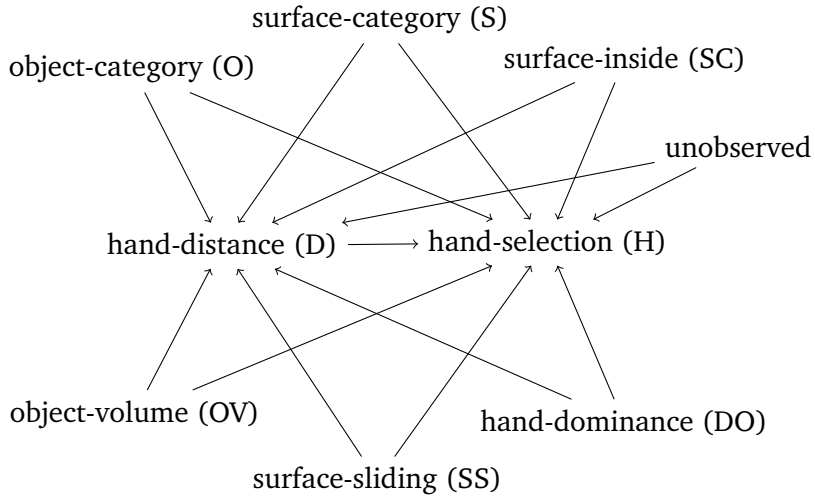


Figure 4.4: *A causal graph showing durable relationship in hand manipulation.*

factors for which ancestors are unmodeled. In the model, the majority of nodes are exogenous except for hand-distance and hand-selection.

The nodes having a parent relationship are endogenous to the model (descendants). This is the case for the nodes hand-distance (D) and hand-selection (H), each having multiple parent relationships.

Factors that do not change can be left out from the model. For example, this would be the case for a node representing the activity a person carries out (e.g. table-setting, preparing a meal, etc.). This node is not included in the graph because the context remains constant and thus is not included in the model.

Some factors are known to be relevant but cannot be measured. Such nodes can also be included in the graph despite the fact no data is available for them. For example, the weight of an object can be incorporated into the model but not measured in virtual environments. Such is the case for many other factors like pressure on fingers, sensation of texture, temperature, and so on. The advantage of incorporating nodes for which no data is available is that the formalism still manages to identify whether a causal relationship can be addressed free from bias via alternative paths.

This advantage is irrelevant in this dissertation because inferences are investigated under a model not fully capturing all relevant factors (however this feature is important for future robot applications). Because many relevant factors in hand behavior are not in the model, estimates that are bias-free cannot be guaranteed. For instance, despite the relevancy, all states in a person's mind are not modeled. A person in a hurry, for example, could exhibit different behavior than one which is not. A tired

person could choose the dominant hand more often, and so forth. Likewise, some environmental aspects relevant in hand behavior are also left unmodeled. For instance, the spatial arrangement of objects can influence hand-selection, a blocking object can make a person choose the less constrained hand for easier access [PWYB18]. The orientation of objects can also influence hand-selection by favoring an easier grasp, and so on. The node labeled "unobserved" in the causal graph of Figure 4.4 makes explicit that not all factors are modeled.

Exploring causal inference under conditions of an incomplete model is valuable. The list of factors influencing hand-behavior in people is arguably open-ended thus modeling all factors within this domain is rather ideal. However, beyond this argument, the main motivation to investigate causal inference under incomplete models is that humans evidence that extracting causal information despite incompleteness is possible.

Technically, when estimating causal effects, including the node "unobserved" to the model breaks the assumption that all relevant factors are known (a.k.a. unconfoundedness assumption). To put it differently, the backdoor-criterion does not guarantee bias-free estimands under such a condition. The main concern of Chapter 5 is to investigate whether causal estimates under such a condition can still be valid.

Before closing, a causal graph encodes implicit information. The missing arrows in a causal graph represent independence constraints. For example, in the causal graph of Figure 4.4, the arrow from object-volume to hand-dominance is missing because the scale of an object cannot the hand-dominance of a person. Likewise, the property of a surface (e.g. surface-inside) is not connected to the property of an object (e.g. object-volume) because these do not affect each other, and so on. The set of missing connections in a causal graph implies a set of independence constraints that can be verified with data (done in Section 5.3.2). For the causal graph in Figure 4.4 the set

of constraints are:

$$
\begin{aligned}
\text{hand-dominance} &\perp \text{object-volume} \\
\text{surface-category} &\perp \text{surface-inside} \\
\text{surface-category} &\perp \text{hand-dominance} \\
\text{surface-category} &\perp \text{surface-sliding} \\
\text{surface-category} &\perp \text{object-volume} \\
\text{surface-category} &\perp \text{object-category} \\
\text{surface-inside} &\perp \text{hand-dominance} \\
\text{surface-inside} &\perp \text{surface-sliding} \\
\text{surface-inside} &\perp \text{object-volume} \\
\text{surface-inside} &\perp \text{object-category} \\
\text{hand-dominance} &\perp \text{surface-sliding} \\
\text{hand-dominance} &\perp \text{object-volume} \\
\text{hand-dominance} &\perp \text{object-category} \\
\text{surface-sliding} &\perp \text{object-volume} \\
\text{surface-sliding} &\perp \text{object-category} \\
\text{object-volume} &\perp \text{object-category}
\end{aligned}
\tag{4.4}
$$

This concludes the design and definition of the model. This section identified durable relationships in hand behavior and encoded them into a causal graph. The model captures a variety of factors potentially influencing the decisions of hand-distance and hand-selection. The relationships of the model are supported with findings from the literature as well as examples. The model does not capture all factors that could influence hand behavior which is indicated explicitly by the node labeled “unobserved” signaling the unconfoundedness assumption is broken (in particular for the quantity hand-distance on hand-selection). Breaking unconfoundedness challenges causal inference to operate under a weaker assumption which resembles a more familiar condition children manage to overcome in practice (our main motivation to enforce it). The model defined in this section is plugged into the inference machinery which is the subject of concern in the following section.

Causal Inference

The previous section defined a causal graph that encodes assumptions made on hand behavior. The next section introduces the machinery that enables causal interpretation and advanced reasoning capabilities.

Causal Models

Much in the way boolean tables are models in propositional logic, and joint probabilities are in probabilistic logic, in causal inference, models are associated to graphs. Following the notation and definition in [Pea00], a causal model is defined as follows.

Definition 4.1.1 (Causal Model [Pea00, p. 203]) *A causal model is a triple $M =< U, V, F >$ where*

(i) U is a set of background variables that are determined by factors outside the model;

(ii) V is a set $V_1, V_2, \ldots, V_n$ of variables that are determined by variables in the model - variables in $U \cup V$; and

(iii) F is a set of functions $\{f_1, f_2, \ldots, f_n\}$ such that each f_i is a mapping from (the respective domains of) $U_i \cup PA_i$ to V_i, where $U_i \subseteq U$ and $PA_i \subseteq V/V_i$ and the entire set F forms a mapping from U to V. This is, each f_i in $v_i = f_i(pa_i, u_i)$, $i = \{1, \ldots, n\}$, assigns a value to V_i that depends on (the values of) a select set of variables in $V \cup U$, and the entire set F has a unique solution $V(u)$ [3].

Every causal model $M =< U, V, F >$ can associate to a directed graph $G(M)$ [Pea00, p. 203]. In the graph, the nodes correspond to variables (U and V) and edges (F) are directed from one member to another. Note that the model M, and neither the associated causal graph $G(M)$, impose any functional form on f_i. The graph defined in the previous section (shown in Figure 4.4) fully specifies a causal model $M =< U, V, F >$.

Before introducing the rules that manipulate causal graphs, further concepts and notations are required. Actions are graph operations that can modify models and are denoted by operator $do(\cdot)$[4]. For instance, applying an action $do(X = x)$ to the model

[3] Uniqueness is ensured in acyclic topologies [Pea00, p. 203].
[4] Already mentioned before as the adjustment operator.

results in the modified submodel $M_x =< U, V, F_x >$ where X is set to a particular level (graphically this operation removes arrows from the graph). The outcome Y of action $do(X = x)$ is denoted by $Y_x(u)$ (the potential response). This denotes the potential response of the factual and the counterfactual world (introduced conceptually in Section 3.4.5). A counterfactual simulates a hypothetical modification via an external (spontaneous) action denoted by the equality $Y_x(u) = y$ where the potential response of Y, to $X = x$, is equated to the counterfactual y. A deterministic causal model M generalizes to the probabilistic formulation by specifying a probability function $P(\cdot)$ defined over the domain of U, i.e. the probabilistic causal model is a tuple $< M, P(u) >$.

Inference Rules

This section introduces the component in the framework which is responsible to derive bias-free estimands and support advanced reasoning capabilities.

Structural causal models (SCM) provides bias-free equations and support three levels of inferences. The three levels of inferences are made explicit in the ladder of causation pictured in Figure 3.4, these are:

1. Association: (seeing / predicting)
   How would seeing X change my belief in Y?

2. Intervention: (doing / adapting)
   What would Y be if I do X?

3. Counterfactuals: (imagining / retrospecting)
   What if X had not occurred?

where the first level is plain prediction and the last two invoke hypothetical reasoning. The first level draws inferences on a distribution that is unmodified (raw, original), whereas the second and third draw inferences on a modified distribution (truncated). The capability to operate on unmodified and especially on modified distributions distinguishes this formalism from data-driven approaches. This capability might hold the key to adapting and transporting interpretations from one domain to another and is the main concern in Chapter 6.

The capability to account for confounding bias and support the three levels of inferences rest on sound algebraic machinery known as the do-calculus. The do-

calculus enables a language for causal reasoning and rests on three rules which are introduced next.

Theorem 4.1.1 (Rules of do-Calculus [Pea00, p. 85]) *Let X, Y, and Z be arbitrary disjoint sets of nodes in a causal DAG G. Let $G_{\overline{X}}$ denote the graph obtained by deleting from G all arrows pointing to nodes in X (also denoted as $\hat{x}$). Likewise, let $G_{\underline{X}}$ denoted the graph obtained by deleting from G all arrows emerging from nodes in X. To represent the deletion of both incoming and outgoing arrows, we use the notation $G_{\overline{X}\underline{Z}}$. Last, the expression $P(y|\hat{x}, z) \triangleq P(y, z|\hat{x})/P(z|\hat{x})$ stands for the probability of $Y = y$ given that X is held constant at x and that (under this condition) $Z = z$ is observed. These elements are part of the rules of do-calculus as introduced next.*

Let G be the directed acyclic graph associated with a causal model as defined in 4.1.1, and let $P(\cdot)$ stand for the probability distribution induced by that model. For any disjoint subsets of variables X, Y, Z, and W, the following rules are used.

Rule 1 (Insertion/deletion of observations):

$$P(y|\hat{x}, z, w) = P(y|\hat{x}, w) \quad if (Y \perp\!\!\!\perp Z)|X, W)_{G_{\overline{X}}}. \tag{4.5}$$

Rule 2 (Action/observation exchange):

$$P(y|\hat{x}, \hat{z}, w) = P(y|\hat{x}, z, w) \quad if (Y \perp\!\!\!\perp Z)|X, W)_{G_{\overline{X}\underline{Z}}}. \tag{4.6}$$

Rule 3 (Insertion/deletion of actions):

$$P(y|\hat{x}, \hat{y}, w) = P(y|\hat{x}, w) \quad if (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \overline{Z(W)}}}, \tag{4.7}$$

where Z(W) is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

The three rules include the do-operator on $\hat{x}$ which isolates X by removing arrows in the graph that point inwards to it resulting in the subgraph $G_{\overline{X}}$. Rule 1 ensures that $do(X = x)$ truly isolates X in the graph. Rule 2 preserves the condition such that an external intervention $do(X = x)$ has the same effect on Y as seeing $Z = z$, achieved by blocking all backdoor paths from Z to Y $G_{\overline{X}}$. Rule 3 licenses the introduction or removal of external interventions $do(X = x)$ while preserving the probability $Y = y$, achieved by deleting all outwards arrows from Z resulting in subgraph $G_{\overline{XZ}}$ except for non-ancestors of W-nodes. Further details on the do-calculus are available here [Pea00, p. 86].

These rules are not only responsible for enabling the three levels of inferences but

also identifying and adjusting against confounding. Unlike the graphical intuition provided earlier in Section Identify and Adjust, the following corollary expresses the strategy to identify and adjust against confounding in terms of algebraic expressions (instead of graphically).

Corollary 4.1.1.1 (Identifiable [Pea00, p. 86]) *A causal effect* $q = P(y_1, \ldots, y_k|\hat{x}_1, \ldots, \hat{x}_m)$ *is identifiable in a model characterized by a graph* G *if there exists a finite sequence of transformations, each conforming to one of the inference rules in Theorem 4.1.1, that reduces* q *into a standard (i.e., "hat-free") probability expression involving observed quantities.*

The three rules and the corollary derived from these rules conform the language of causality and provide the core operations to account for bias and reason hypothetically.

This means that a bias-free equation for causal effect q exists when, after reducing the equation using the three rules, the probability expression has no term with a hat ($\hat{x}$ or $do(\cdot)$). In case the reduction is possible, the resulting adjusted expression holds only observed quantities.

Rules 1 to 3 derive all identifiable causal effects, shown to be complete in [HV06]. However, [Pea00, p. 86] notes that a procedure to determine whether a sequence of rules reducing arbitrary causal effect expressions has not been found. For this reason, identification based on a graph criterion is desirable over the algebraic provided in Corollary 4.1.1.1. Moreover, a graphical criterion is especially beneficial for robot applications where substantial domain knowledge is necessary. The graphical criterion for the identification of causal effects is as follows.

Theorem 4.1.2 (Criterion [TP02]) *A sufficient condition for identifying the causal effect* $P(y|do(x))$ *is that there exists no bi-directed path (i.e., a path composed entirely of bi-directed arcs) between* X *and any of its children*[5].

This theorem states that the expression of causal effect $P(y|do(x))$ is identifiable when every child of X on the paths towards Y is not reachable from X via a bi-directed path [6]. This criterion generalizes the back-door and front-door criterion (being the last two special cases of Definition 4.1.2)[7].

[5] To gain in performance, before applying this criterion one may delete from the causal graph all nodes that are not ancestors of Y.

[6] [SP08] shows examples of graphs where $P(y|do(x))$ is not identifiable.

[7] Othe well-known identification methods that rely on adjustment sets are the mediation formula. Some alternative identification methods which are not based on graphs are regression discontinuity, instrumental variables, and difference-in-difference.

The complete graphical criterion has been established in [SP08] which covers the identifiability of conditional intervention $P(y|do(x), z)$ for arbitrary X, Y, and Z. This is, an arbitrary quantity invoking the do-operator is reducible in polynomial time on semi-Markovian models, and in the case an expression is reducible, the estimand is provided [SP08].

To summarize, the inference capabilities of structural causal models (SCM) described in this section are the following. First, SCM provides the mechanism to control confounding via a criterion that identifies a set of variables that, when adjusted, results in an unbiased estimand. Second, SCM enables the evaluation of policies on non-experimental data via the do-calculus and graphical criterion supporting three levels of inferences. Last, the completeness of the do-calculus provides a means to detect whether any given quantity can have a bias-free estimand.

Treatment Effect Estimation

The previous section describes how structural-causal-models provides a bias-free equation known as the estimand. The next concern is solving the equation. Solving the equation means deriving a value for the equation using data.

Causal Effect

The causal effect is computed by modulating the levels of a variable known as the treatment. Modulating the treatment variable is a safe strategy only because the estimand provided by SCM (described in previous section) already counters undesirable effects. Treatment-effect estimation aims to derive the quantity known as the causal effect and is the main concern described in this section.

The literature provides several ways to define the causal effect (a.k.a. target quantity or treatment-effect). One of them defines the causal effect over a population Average-Treatment-Effect (ATE) [KSV+18], the other defines the Individual-Treatment-Effects (ITE) [Imb15], and another targets subgroups with the Conditional-Average-Treatment-Effect (CATE) [KSV+18].

The ATE presumes populations respond equally to a given treatment a presumption known as treatment-effect homogeneity. However, subgroups in a population could respond differently and thus lead to imprecise results [AS18]. On the other hand,

treatment heterogeneity considers subgroups in the population that can react differently to the same treatment [Ang04] and are thus better suited than ATE for such populations. This is the case of ITE and CATE.

This dissertation favors treatment heterogeneity over homogeneity (i.e. ITE and CATE, over ATE) because human hand-behavior is context and purpose-specific thus expecting sub-groups to exist. However, this poses another challenge because targeting causal effects at the individual (ITE) or subgroup (CATE) level implicates the need for more data than studies powered at the population level (ATE). Hence, CATE is favored over ITE. The causal effect is defined as follows.

Definition 4.1.2 (Causal Effect) *"Given two disjoint sets of variables, X and Y, the causal effect of X on Y, denoted either as $P(y|\hat{x})$ or as $P(y|do(x))$, is a function from X to the space of probability distributions on Y. For each realization x of X, $P(y|\hat{x})$ gives the probability of $Y = y$ induced by deleting from the model all equations corresponding to variables in X and substituting $X = x$ in the remaining equations" [Pea00, p. 70].*

Following the more conventional definition in the literature [RR83] which adopts the specialized version of Definition 4.1.2 for the case of two distinct realizations x' and x'' of X (e.g. the case of binary variables) which expresses causal effect as a difference $E(Y|do(x')) - E(Y|do(x''))$. Based on this more common definition, the Conditional Treatment Effect Estimation reported in subsequent sections is defined as follows.

Definition 4.1.3 (Conditional Treatment Effect Estimation)

$$E\left[Y^{e}_{i,x=1} - Y^{e}_{i,x=0}|Z\right] \tag{4.8}$$

where outcome Y (e.g. hand-selection) is computed by estimator e (described below) for data point i (e.g. instant of time the person triggers a grasp) conditioned on the treatment x (e.g. hand-distance) provided the set of observed features Z (e.g. variables in Figure 4.4) [Res19].

Estimators and Parameters

The previous section defines the causal effect (CATE), which is the target that estimators aim to evaluate with data. This section introduces the estimators that are responsible to evaluate the causal effects using data.

Estimators are algorithms that read data to evaluate the effect of a treatment variable X on an outcome variable Y while controlling for a set of features $\{Z, W\}$, including those effects that vary as a function of Z (i.e. heterogeneous treatment) [Res19].

Besides the flexibility to provide heterogeneous treatment effects, all estimators evaluated in subsequent sections assume that estimands are bias-free (i.e. these operated under unconfoundedness assumption). This assumption is true when the causal graph captures the relevant factors for the target quantity under study. This is so when the set of Z includes all factors that could affect treatment X and outcome Y.

Estimators are composite ML algorithms that distinguish two stages where each stage solves a different task. One stage involves a prediction task, whereas the other builds upon the former stage (i.e. predictors) to compute the causal effect (a.k.a. solve the moment equation) as shown in Figure 4.5.

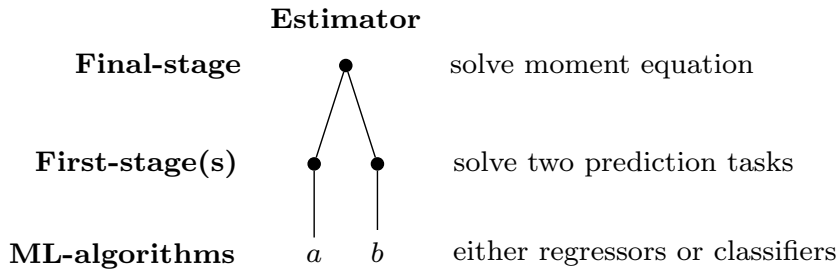


Figure 4.5: *Estimators are multi-stage ML algorithms. At the lowest level, the first-stages deal with prediction tasks with parameters set algorithmically. The final-stage guarantees desirable properties with the parameters set analytically.*

The prediction stage is the first-stage and the subsequent is the final-stage. It is the final-stage that confers properties to an estimator and thus the selection of estimators is based on them. To grant the properties at the final-stage, the first-stage(s) require cross-fitting. Hence, only ML-algorithms with cross-fitting are considered at the first-stages.

The selection strategy of the estimators starts with an initial set of candidate estimators from which unsuitable options are discarded leading to the final selection. The selection of the final-stage is analytic while the first-stage(s) is algorithmic.

Out of the fifteenth estimators available in the library [Res19] four are chosen. Subsequent sections justify the selection of these estimators. The selection depends

primarily on the design, desirable statistical properties, and other technical aspects. Figure 4.6 outlines the selection of estimators.

The presentation follows Figure 4.6 from top to bottom. The selection begins by considering four designs under which estimators operate. The choice of an estimator within each operational design is driven by the final-stage properties. The properties include the capability of an estimator to provide valid confidence-intervals and the strategy for robustness of an estimator. Subsequently, the selection of ML-algorithms for the first-stages are presented. Figure 4.6 outlines the selection of estimators described in the following sections.

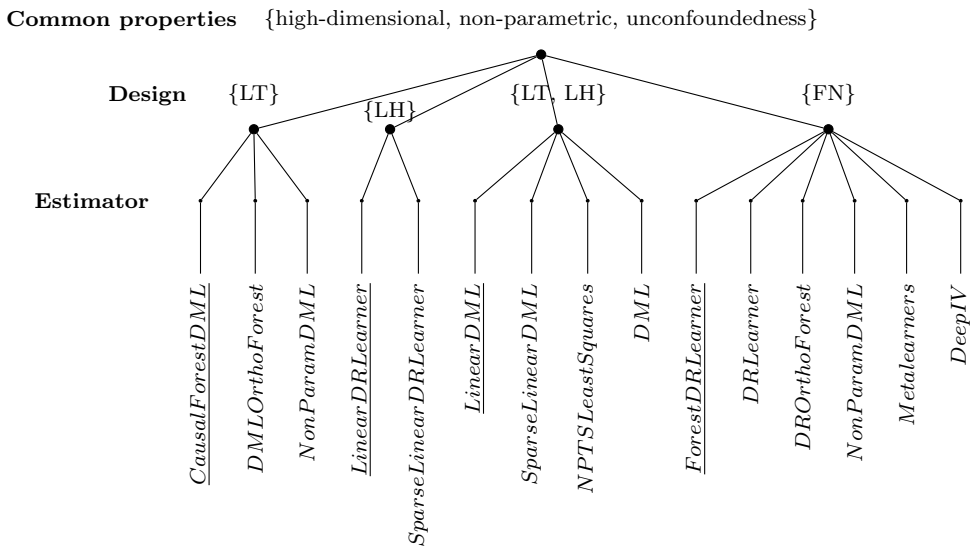


Figure 4.6: *The selection of estimators relies on the design and final-stage properties. Four estimators are selected (highlighted). The resulting selection covers one estimator within each design. The hyperparameters at the first-stage(s) are set with a cross-validated grid search. The design Linear Treatment (LT), Linear Heterogeneity (LH), and Fully non-parametric (FN) indicate operational assumptions introduced in the following section.)*

Estimator Selection

Recall that the chosen target quantity is CATE (Definition 4.1.3), and thus estimators supporting heterogeneous treatment-effect estimation are considered. Within the scope of heterogeneous treatment-effect estimators, some estimators are flexible

while others impose some form on heterogeneity. Each design dictates a different assumptions under which estimators operate [Res19]:

Linear in treatment (LT): The outcome Y is a linear function of the treatment X. The estimators that impose linearity on the structure are CausalForestDML, DMLOrthoForest, and DML.

Linear in heterogeneity (LH): The size of the outcome Y is a linear function of observable characteristics (Z,W). The estimators that operate with this design are LinearDRLearner and SparseLinearDRLearner.

Linear in the treatment and heterogeneity (LT, LH): This design makes both of the previous assumptions. Outcome Y is a linear function of treatment X, and the size of the outcome Y is a linear function of observable characteristics (Z,W). The estimators that incorporate both assumptions are LinearDML, SparseLinearDML and NonparametricTwoStageLeastSquares.

Fully non-parametric (FN): Unlike previous designs, estimators imposing no parametric form are flexible but at the expense of data and computational cost. The estimators following this design are ForestDRLearner, DRLearner, DROrthoForest, NonParamDML, including several meta-learners.

Less flexible designs are suitable when knowledge about heterogeneity is known and can be exploited. Otherwise, when heterogeneity is unknown, the last option applies. When treatment is known to be linear the first operational design is adequate, when the linearity in heterogeneity is known then the second operational design applies, and when both are known the third design is a suitable option. Leaving the last option for cases when knowledge is not available.

The linearity of the first three designs is not limited to linear functions, as non-linear relationships are also covered including additively separable linear functions (e.g. treatment and treatment-squared). The estimators operating under the first three designs cannot adequately capture a fully non-parametric relationship in heterogeneity whereas the last option does, however, at the expense of more data and higher computational costs.

This dissertation covers the four operational designs for the following reasons. Including a fully non-parametric design (i.e. the last design in the list) is important because heterogeneity in human hand behavior is unknown. Even when the underlying heterogeneity is known, it is reasonable to investigate this design because service robots observing people in household scenarios would lack this information. For

these reasons, an estimator operating under the most flexible design is included in the selection.

The reason that favors estimators operating under the first three designs is the following. Estimators dealing with observational evidence typically expect a quantity of data in the order of thousands but only hundreds of samples are considered here. Data scarcity is precisely one of the challenging aspects of the scenario set forth to investigate. Moreover, the impact of samples in observational-data estimation is not well understood. Designs that impose some assumptions are better suited to handle fewer samples and provide faster estimation rates than the fully non-parametric design. For this reason, estimators within each of the first three designs are included in the selection.

Including the four designs is important because estimators are challenged with operating under the condition of broken unconfoundedness. Therefore, at least one estimator from within each design is chosen thus covering the full operational design space of estimators.

The next choice concerns the selection of estimators within each of the operational designs. This choice rests on the final-stage properties.

The first property considers is whether estimators afford the construction of valid confidence intervals. Confidence intervals are central to hypothesis testing and consequently for this work. Therefore, only the estimators that can guarantee valid confidence intervals are considered. This excludes meta-learners such as DRLearner, NonParamDML, NonparametricTwoStageLeastSquares, to mention a few.

Estimators that construct valid confidence intervals are DROrthoForest, ForestDR-Learner, DMLOrthoForest, and CausalForestDML. CausalForestDML is preferred over DMLOrthoForest because the former supports the computation of multiple outcomes while the latter does not. ForestDRLearner is preferred over DROrthoForest because the former is less demanding in computational resources than the latter. Table 4.1 summarizes the selection of estimators.

The capability to cope with high-dimensional spaces is another important property to consider. Some estimators deal with high dimensional (sparse) data while others do not. Sparse estimators are preferred over the non-sparse versions when the number of features is comparable to the number of samples, which is not the case in the data considered here. Therefore, the non-sparse versions LinearDML and LinearDRLearner are favored over SparseLinearDML and LinearDRLearner correspondingly.

To summarize, because the underlying heterogeneity in the data is unknown four estimators are considered, one operating under each design thus covering the full

operational space. Moreover, only the estimators that construct valid confidence intervals are included. Furthermore, because the number of features is not comparable to the samples the non-sparse versions are chosen. Taken together, the selected estimator are:

Operational Design	**Selected Estimator**
Linear in treatment	CausalForestDML (CFDML) [NSS18]
Linear in heterogeneity	LinearDRLearner (LDRL) [BR05]
Linear in treatment and in heterogeneity	LinearDML (LDML) [CCD+17]
Fully non-parametric	ForestDRLearner (FDRL) [FS19],

Table 4.1: *The list of selected estimators and their corresponding operational design.*

Parameter Selection

Recall estimators are composite ML-algorithms that involve two stages, each solving a different task. The final-stage models solve the moment equation whereas first-stage models deal with a prediction task. It is the final stage that confers desirable properties to the estimator.

Two major classes of final-stage models coexist, namely Double Machine Learning (DML) and Doubly Robust Learners (DRL). The name and abbreviation of the estimators listed in Table 4.1 conveniently embed the sub-strings DML and DRL indicating the class an estimator belongs to. The four abbreviations indicate that two estimators belong to DML while the other two correspond to DRL.

Unlike conventional methods in statistics, both classes DML and DRL operate without presuming any parametric form on distributions and apply to large dimensional spaces [CCD+17].

DML and DRL define the final-stage of an estimator and guarantees favorable statistical properties such as small mean-squared-error, asymptotic normality, and the construction of confidence intervals. Only a few constraints need to be satisfied for these approaches to guarantee favorable properties.

DML and DRL differ in the strategy to achieve robustness. DML predicts the outcome and treatment from the controls separately while DRL performs a joint prediction. Subsequently, these approaches combine the first-stage predictive models into the final-stage model to compose the heterogeneous treatment effect.

The final-stage models are:

- Linear Double Machine Learning (LDML) - a.k.a. LinearDML:

  Relies on an unregularized final linear model and supports confidence intervals via asymptotic normality arguments [CCD+17].

- Linear Doubly Robust Learner (LDRL) - a.k.a. LinearDRLearner:

  In this case, the final-stage model is a Linear Regression which fits a standard ordinary linear regression (OLS) [BR05]. This class is valid even if the CATE model is not linear in heterogeneity [CCD+17].

- Forest Double Machine Learning (FDML) - a.k.a. CausalForestDML: This class of estimators employs a causal forest for the final-stage model [WA18, ATW19], delivering confidence intervals via Bootstrap-of-Little-Bags (BLB) [ATW19].

- Forest Doubly Robust Learner (FDRL) - a.k.a. ForestDRLearner: Employs a Subsampled Honest Forest regressor as the final model [WA18, ATW19], and offers confidence intervals via the BLB as described in [OSW19].

Unlike the final-stage, the first-stage models are cross-validated with grid-search to optimize the hyperparameters. Grid-search makes an exhaustive exploration over the parameter space and accounts overfitting with cross-validation thereby sacrificing fewer samples in the model validation. The ML-algorithms are configured according to the parameter space and fitted over multiple partitions of data ($splits = 3$) with cross-validation splitting strategy k-fold. The performance is measured by the average score over these training partitions based on the negative mean-squared-error (i.e. neg. MSE). Last, the performance of cross-validated ML-algorithms is assessed on test data to derive the best ML-algorithm and corresponding parameters.

The parameters space for grid-search depends on the final-stage model in as follows. Final-stage models allow for nearly arbitrary machine-learning algorithms to be selected at the first-stages while maintaining statistical properties with the final model (e.g. small mean squared error, asymptotic normality, and the construction of valid confidence intervals).

The capability to support arbitrary algorithms is made explicit in Table 4.2 with the entry *any* for the two DML estimators. DRL on the other hand, requires a classifier for one of the first stages and a regressors for the other.

The number of grids required to run the experiments is the following. Each estimator listed in Table 4.1 requires the selection of two algorithms for the first-stage. Hence, a total of eight algorithms need to be selected. This would typically involve the

Estimator	Final-stage	First-stage-1	First-stage-2	Inference
LDML	Regressor	any	any	Bootstrap
LDRL	Regressor	Classifier	Regressor	Bootstrap
CFDML	Causal Forest	any	any	BLB
FDRL	Honest Forest	Classifier	Regressor	BLB

Table 4.2: *The stages of the selected estimators, final and first.*

specification of one hyper-parameter space per algorithm thus reaching 8 exploration grids in total.

As Double Machine Learning (DML) supports arbitrary algorithms [Res19], specifying separate exploration grids are not necessary because a single hyper-parameter grid can be shared across stages. As a result, the two selected DML estimators require a single exploration grid instead of four.

DRL on the other hand enforces a type of algorithm for each of the stages. In this case, two hyper-parameter grids are necessary in this case, one per stage. One parameter-grid explores regressors the other classifiers. Therefore, specifying two hyper-parameter grids cover both DRL estimators (LinearDRL and ForestDRL) shared across the first-stages 1 and 2.

Last, considering that DML admits any of the exploration grids defined for Doubly Robust Learners (DRL), one can spare the need for one hyper-parameter grid. This is, the parameter space can be reused across estimator classes. Therefore, a total of two grids cover the four estimators (eight first-stage algorithms). Both grids are shared across DML and DRL estimators.

The hyper-parameter grid exploring regressors is:

Estimator Space	Parameter Space
Random Forest Regressor	max_depth: [3, None] min_samples_leaf: [10, 50]
Gradient Boosting Regressor	n_estimators: [50, 100] max_depth:[3] min_samples_leaf: [10, 30]

Table 4.3: *The hyper-parameter grid explores regressors optimizing the negative mean-squared error across* 3 *folds.*

The hyper-parameter grid exploring classifiers is the following:

The best ML algorithms and parameters for the estimators are presented in Figure 4.7.

Estimator Space	Parameter Space
Random Forest Classifier	max_depth: [3, 5] min_samples_leaf: [10, 50]
Gradient Boosting Classifier	n_estimators: [50, 100] max_depth: [3] min_samples_leaf: [10, 30]

Table 4.4: *The hyper-parameter grid explores classifiers optimizing the negative mean-squared error across* 3 *folds.*

In the case of LinearDML, the best regressor corresponds to Gradient-Boosting and classifier to Random-Forest with the corresponding parameters as indicated in the figure. For LinearDRL the same algorithms are selected but differ in the parameter selection. CausalForestDML on the other hand, selects a RandomForest-Regressor and -Classifier with the corresponding parameters as shown in the figure. Last, ForestDRL selects two RandomForest-Regressor with the corresponding parameters shown in the figure. These hyper-parameters are optimized on ds-v, the parameters found on the remaining collection of data are listed in the appendix.

Point Estimates and Confidence Intervals

The results presented in subsequent chapters report point estimates and confidence intervals (CI). A point estimate is a single value that provides a representative measure, whereas CI is a range of values that measures the uncertainty of estimates. Both are commonly reported together to show the reliability of estimates. The level of confidence that true effects are within a range of values is set to three sigmas ($alpha = 0.95$).

Estimators deliver analytic confidence intervals via different inference methods. This is technically possible because these estimators correctly adjust for the reuse of data across the multiple stages that are involved in the computation of estimates (data reuse is cross-validated over stages).

Two inference methods are employed for the computation of confidence intervals: Bootstrap and Bootstrap-of-Little-Bags. As summarized in Table 4.2, the confidence intervals for LDML and LDRL are computed with bootstrap, whereas for FDRL and CFDML with BLB.

Bootstrap computes confidence intervals training multiple versions of the original

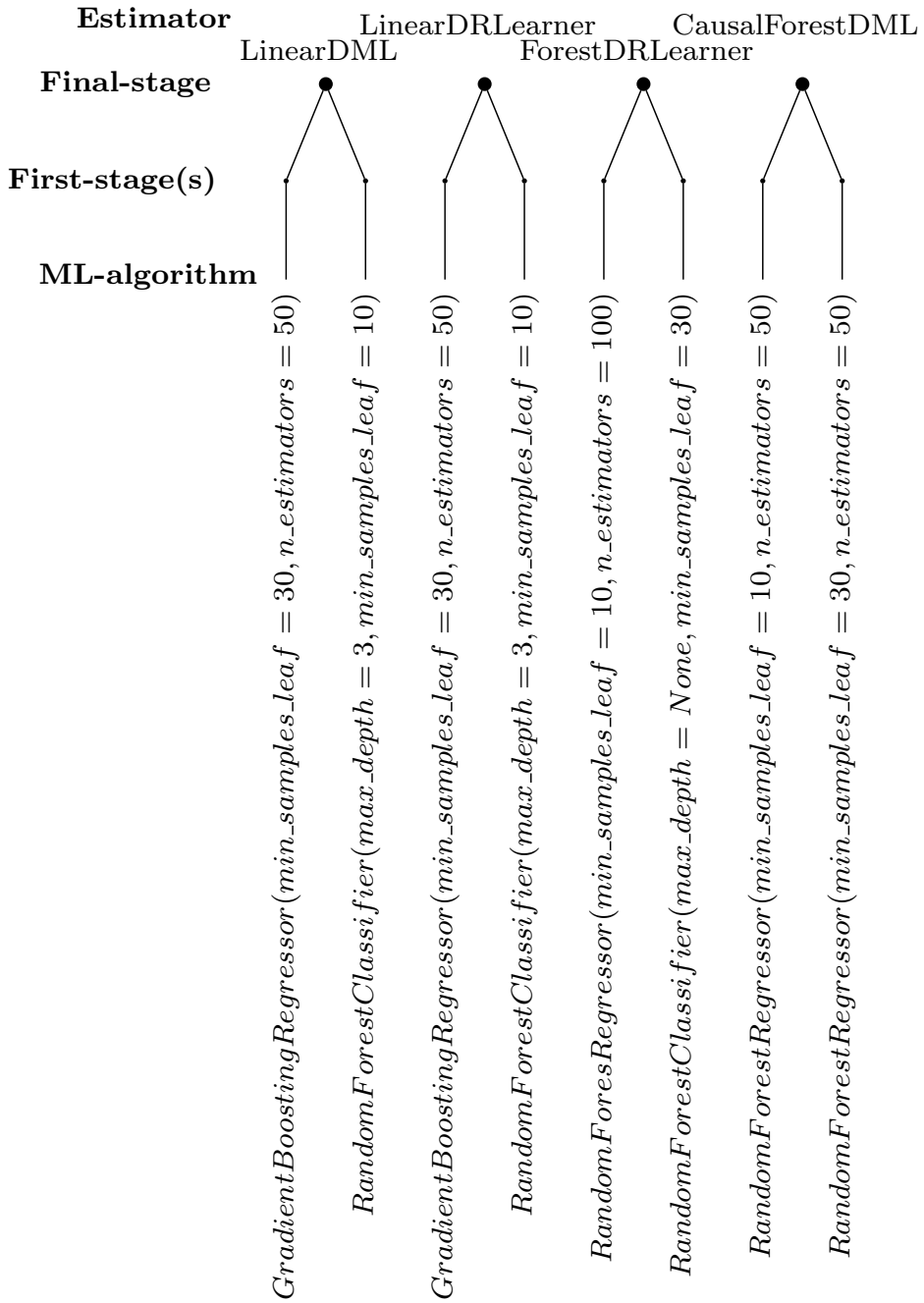


Figure 4.7: *The selection of ML algorithms and corresponding parameters for the first-stages of the estimators optimized with grid search (score neg. MSE) and cross-validated on k-fold (splits 3).*

estimator using random sub-samples with replacement (n_bootstrap_samples=100). Subsequently, the intervals are calculated based on the quantiles of the estimated distribution across the multiple samples. The resulting interval relies on the estimate over the entire dataset (bootstrap_type=pivot).

BLB on the other hand constructs confidence intervals and quantifies the uncertainty of estimates delivered by non-parametric models based on Honest Random Forest [ATW19]. BLB calculates the covariance of the parameter vector with an objective Bayesian debiasing correction to ensure that variance quantities are positive. This inference method applies to final-stage forest-based models such as CausalForestDML and ForestDRLearner.

In summary, four robust estimators that operate under the unconfoundedness assumption are selected. All of them offer the construction of valid confidence intervals and aim for a small mean squared error. Two estimators correspond to the class of DML while the other two belong to DRL. The algorithms and parameters for the first-stages are optimized with grid search. The methods to compute the confidence intervals are based on bootstrap inference for (LDML and LDRL) and Subsampled Honest Forest for (CFDML and FDRL) [SK+19].

4.1.3 The Capabilities

This section summarizes the key capabilities of the framework and highlights two conditions under which the framework is expected to operate properly.

The three components of the framework described in previous sections are, the model presented in Section Causal Graph, the graph operations introduced in Section Causal Inference, and the estimators described in Section Treatment Effect Estimation. These components - working as a unit - offer a set of capabilities that are necessary to enable robot analysis of human interaction.

Some remarks on how the framework resembles the core capabilities of human observation follows below:

- The framework enables causal reasoning as opposed to correlational interpretation. Causal reasoning has been shown to exist in children at an early age and is a foundational capability identified in Social Cognitive Theory (SCT), supporting observational learning in humans.
- The framework enables end-to-end non-parametric processing because neither the inference machinery nor the estimators impose any form on the input

distribution. This resembles the capability of people to deal with any input distribution. Decoupling inferences from distributions avoid well-known issues with nonstationary distributions (concept and population drifts). This leverages hypothetical reasoning (adapt and generalize), enabling interpretation of past observations in another domain. The SCT identifies hypothetical reasoning as an essential capability supporting human observational learning.

- The framework provides anytime reasoning instead of learning every time from scratch, as done with data-driven approaches. This resembles the people's capability to adapt or interpret quickly in novel scenarios which is crucial for observational learning.

The capabilities featured by a framework support the acquisition process known as a modeling in the theory of SCT. The retention process encodes models using causal graphs; the production process manipulates models using the do-calculus, the attentional and motivational process queries with the inference language. The framework establishes the foundation to support robot observational learning and realizes the strategy shown in Figure 3.2.

However, two conditions enforced by the framework misalign with human capabilities. One relates to the completeness of models and the other to the number of samples.

- First, SCM guarantees bias-free estimands under the condition of models fully capturing the relevant factors for a given quantity of interest, a condition known as the unconfoundedness assumption. People, however, make decisions with incomplete information. Especially children cope with incompleteness more than adults do. This motivates studying the framework under incomplete models.

- Second, estimators are expected to operate with non-experimental data and thus typically large quantities of samples are assumed (over thousands of samples). Methods suitable for large quantities of samples are not necessarily successful processing few data. People, however, show the capability to learn from few examples, even a single one. Moreover, observers do not control everything they watch rending the quantity and quality of observations uncontrollable. This motivates studying the framework on few samples.

The first condition is enforced with a model not fully capturing the relevant factors in hand behavior as described in Section Causal Graph. The second condition is realized by drawing inferences on samples in the order of hundreds. This is well below the

thousands common in treatment-effect estimation targeting non-experimental data. The following section introduces the scenario.

4.2 The Scenario

The previous sections establish the framework that processes evidence. The focus on the following sections is on the input for the framework (Figure 4.8) which include the source of evidence, scenario, activity, data collection, and how robots can access virtual environments.

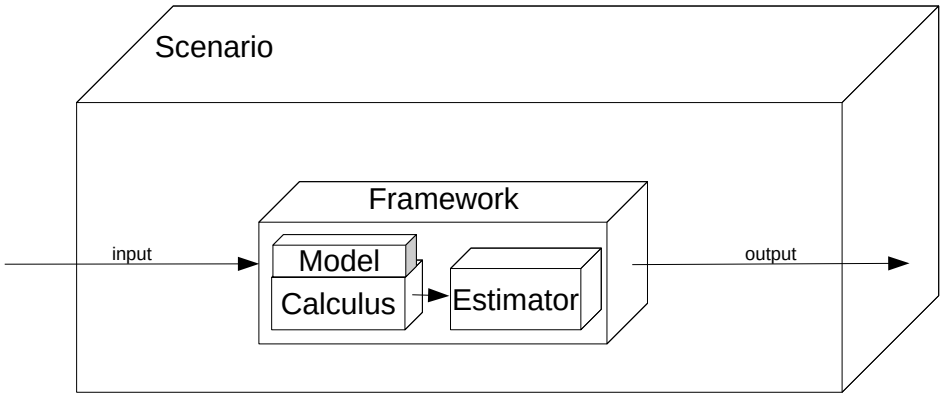


Figure 4.8: *The scenario and the framework.*

4.2.1 The Collection of Evidence

The collection of data on which the framework estimates causal effects originate from a virtual environment. Virtual environments provide nearly unrestricted access to contextual cues without bothering participants with intrusive markers potentially compromising natural behaviors.

A headset and two hand controllers (VR devices) track user motions in three-dimensional space. A game engine is responsible for logging streams of data and renders near photo-realistic scenarios as shown in Figure 4.9.

The activity people perform in the virtual environments are table settings. A table setting essentially involves a person moving objects from one place to another. This

Figure 4.9: *A set of images showcasing a table setting in a virtual kitchen as experienced by the participants. The last picture of the sequence features a completed table setting.*

includes searching, transporting, and placing objects in the kitchen - opening and closing containers, and so forth.

The three collections of table settings considered in this dissertation employ the same recording equipment, virtual kitchen layout, initial position of objects, and instruction to set the table. Each collection of data corresponds to a single person. Note that the unit of analysis are not participants (only three) but rather grasping actions (order of hundreds). However, each collection of evidence differs considerably as follows.

The data collection [PWYB18] is recorded with the purpose to study how people arrange objects on the table (denoted here with Ds-1). The dataset [HB19] focus on the time people spent setting a table under various modalities (denoted with Ds-2). The collection of data Ds-v has been specifically recorded to validate the framework. The key distinction across datasets is that Ds-1 and -2 originate from uncontrolled sources of evidence (further described in Chapter 6) while Ds-v stems from a controlled experiment (discussed in Chapter 5).

Because people are not restricted nor instructed to perform table settings in a particular way, the behavior is natural. Consequently, each set of data is different in Spatio-temporal aspects such as the preference of objects selected for the breakfast, the surfaces on which objects are placed, etc. The dashes in the summary Table 4.5 are evidence of such asymmetries across data collections.

The features of the data are either binary, categorical, or numeric. Table 4.6 shows the correspondence between the causal graph of Figure 4.4 and the features of the data in Table 4.5:

The feature transformations are as follows. The distance of a grasp (denoted by D) is measured as soon as the user triggers a grasp with the motion controllers. Hand-distance is the only feature that requires a transformation because not all selected estimators (introduced in Section Treatment Effect Estimation) support continuous treatment values. Discretizing with a median split overcomes this technical restriction.

User interaction			Ds-1	Ds-2	Ds-v
Object (O):	**OV (m^3)**		**384**	**174**	**137**
Silverware	8.2×10^{-5}		83	70	29
Glass	1.22×10^{-3}		83	-	23
Milk	1.40×10^{-3}		18	8	24
Juice	1.93×10^{-3}		88	-	21
Bowl	2.40×10^{-3}		19	36	16
Cereal	6.30×10^{-3}		93	28	24
Tray	9.95×10^{-3}		-	32	-
Surface (S):	**SC**	**SS**	**384**	**174**	**137**
DiningTable	F	F	12	18	-
FrdgArea	T	F	-	1	1
FrdgDrBtmShlf	T	F	5	2	22
FrdgGlassShlf	T	T	41	-	21
IslndArea	F	F	-	8	21
IslndDrwBtmLft	T	T	-	16	-
LabFloor	F	F	5	-	2
OvenArea	F	F	-	2	-
OvenDrwRight	T	T	41	16	22
SinkArea	F	F	11	24	1
SnkDrwLftBtm	T	T	-	-	1
SnkDrwLftMid	T	T	41	-	20
SnkDrwLftTop	T	T	41	42	25
Tray	F	F	184	42	-
Hand-selection (H):	**DO**		**384**	**174**	**137**
Left (H_0)	F		121	96	64
Right (H_1)	T		263	78	73
Hand-distance (D):			**384**	**174**	**137**
Close (HD_F)	0		192	87	69
Far (HD_T)	1		192	86	68

Table 4.5: *The three collections of evidence (ds-v, ds-1, ds-2) and corresponding frequencies of occurrences for different levels of features. Abbreviation* OV *refers to the volume of an object.* SC *corresponds to surfaces inside containers such as the refrigerator's shelf.* SS *refers to surfaces that slide, such as drawers.* DO *refers to hand dominance.* SnkDrwLftMid *refers to the middle drawer out of three stacked one over another spatially located left of the sink.* FrdgDrBtmShlf *refers to the bottom shelf on the fridge door. Last, levels* T *and* F *abbreviate true and false.*

Concept	Feature	Symbol	Level
hand-distance	either far or close w.r.t. head	D	binary
hand-selection	either left or right	H	integer
hand-dominance	either left or right	HD	integer
object-category	category name	O	integer
object-volume	size of an object	OV	float
surface-category	category name	S	integer
surface-sliding	either sliding or not	SS	integer
surface-container	either inside of a container or not	SC	integer

Table 4.6: *The correspondence between concepts in the model and features in the data.*

A median split over the straight-line distance between the tracked headset and handheld VR devices determines the distance. The resulting splits for the validation set are $(0.56, 0.77)$ meters for a close grasp and $(0.77, 0.85)$ meters for a distant grab. Likewise, for Ds-1 the bins corresponding to close/far are $(0.59, 0.73)$ and $(0.73, 0.92)$. Similarly, the splits for Ds-2 are $(0.47, 0.63)$ and $(0.63, 0.77)$, correspondingly. The frequencies for hand-distance are based on these splits as shown in Table 8.1.

4.2.2 The Access to Virtual Environments

This section describes how robots have access to virtual environments, in particular, the infrastructure enabling access to evidence produced in the virtual environment.

Robots have access to natural human behaviors in virtual environments where people immersed in near photo-realistic scenarios perform activities. Virtual environments enable access to nearly any interaction of users within the virtual scenario. A headset and hand-controllers (VR devices) track user motions in three-dimensional space. A game engine is responsible for logging streams of data and rendering a near photo-realistic virtual kitchen scenario. The gear users employ to interact with the virtual environment is the HTC Vive Pro shown in Figure 4.10.

To access and extract information from virtual environments, this dissertation employs the pipeline shown in Figure 4.11. With this pipeline, robots have access to activities recorded in virtual environments. The main components of this pipeline are described next.

A game engine (Unreal Engine) renders near photo-realistic scenarios and logs streams of data. The tool described in [BBH+18] annotates raw streams of data

Figure 4.10: *The setup includes the HTC Vive gear (a headset, two motion controllers, two mounted base stations), the Unreal Game Engine, and RobCoG*[8]*.*

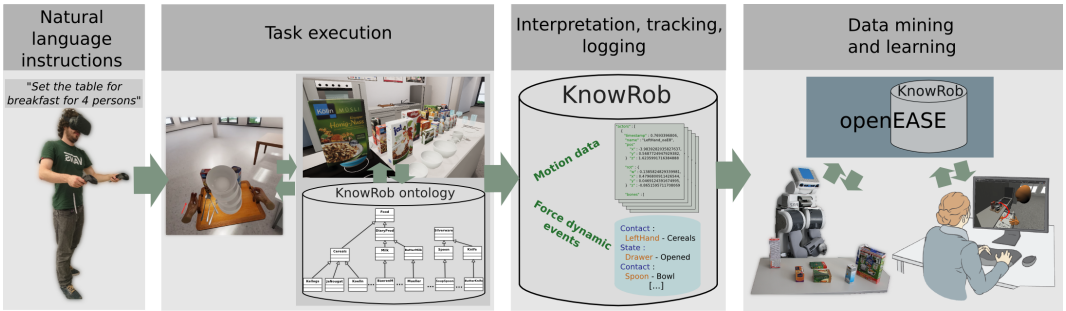


Figure 4.11: *The RobCoG pipeline providing robots with access to virtual activities.*

with semantics such as collisions, changes in object states, and so on. Annotations comply with an ontology especially designed for robots and human household domain [BBH+18].

To make recordings in virtual environments accessible and consumable for robots, data complies to a machine-understandable format. The collections of semantically enriched data are stored in a robot understandable format known as narrative-enabled episodic memories (NEEM). Robots and scientists have access to such data either on a cloud platform (openEASE[9]) [BKF+18] or directly with [Gay20]. The focus of this dissertation is on the processing of such data.

The pipeline outputs data enrich with semantics. As a result, robots can query for collisions that occur in the virtual kitchen, for instance, when a virtual hand contacts an object, extract motion trajectories, and so forth.

Of particular relevance are grasping events. Grasping events are instances at which a person holds a target object with a virtual hand. Grasps are triggered by pressing a button on the hand controller. Grasps are not necessarily successful unless the person performs a stable grasp of the object. To achieve a successful grasp, the hand must be sufficiently close to the target such that the fingers surround the object. Targets do not simply clamp to the virtual hand, instead, participants regulate the pressing and releasing of the virtual hand on a continuous scale.

Robots can extract grasping actions that are performed during virtual table settings issuing a query such as the one shown in Figure 4.12.

The pipeline shown in Figure 4.11 produces the collection of evidence studied in the subsequent chapters. The focus of this dissertation is the processing of this information. The collection of evidence produced with the pipeline includes grasping actions of movable targets such as objects interacted with during table-setting activities among other contextual information.

4.3 Concluding Remarks

To formalize robot analysis of human interaction this chapter introduces a framework and scenario. The framework aims to resemble some capacities in human reasoning while the scenario aims to incorporate challenging conditions in which children manage to learn.

[9] www.open-ease.org

```
% the query
entity(A,[an, action,
[type, knowrob:'
  GraspingSomething'],
[ start_time , Start ],
[ end_time , End ] ] ).
% the result:
Event = knowrob:'
  GraspingSomething_CLYMGOHV',
Start = knowrob:'timepoint_107
  .052307'
End = knowrob:'timepoint_110
  .65230699999991'
```

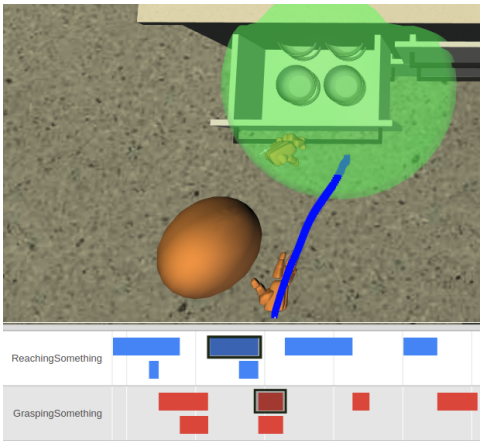


Figure 4.12: *A query featuring how events are extracted from virtual activities. The keyword* event *identifies a grasping action while* start *and* end *the time when the user issues a grasp. The image displays OpenEASE visualizing a person about to grasp a bowl. The bars visualized below the image are segments of reaching and grasping actions.*

The framework incorporates the core inference capabilities to interpret as children do. These are, the capability to reason causally, predict, adapt, and imagine (supported with the do-calculus). Other desirable properties are the (non-parametric) processing of any inputs and anytime reasoning.

The components that integrate the framework are three, a model, the inference machinery, and the estimators. The model encodes minimal assumptions in a causal graph (Section Causal Graph). The inference machinery provides the mechanism to confront confounding and articulates three levels of reasoning (Section Causal Inference). The estimators evaluate expressions using data (Section Treatment Effect Estimation).

These components relate as follows. The inference machinery relies on the model to confront confounding bias and articulate three levels of reasoning. The calculus inflicts selective changes to the model using the do-operator to leverage bias-free estimands. The estimators evaluate the resulting estimands using data despite the missing counterfactuals.

The scenario challenges the framework by inferring human hand behavior from natural activities. The examples on which the framework draws inferences are natural (uncontrolled), and the number of these is limited to only a few, just like scenarios where children observe role models without being explicitly taught by them.

Two conditions enforced by the framework are misaligned with the scenario, thereby

challenging its potential feasibility and utility. One of them requires that models capture the relevant aspects, while the other rests on the availability of large numbers of samples. Both conditions are not met in the scenario. The first condition is broken by employing a model capturing only part of the relevant aspects, and the second by limiting the number of samples to only hundreds.

The aim is not investigating a framework that works but rather to challenge its potential. The question is, whether the framework operating under these conditions still manages to derive valid and stable estimates. To this end, chapters 5 and 6 investigate the feasibility and utility of the approach.

Chapter five

Feasibility of the Framework

This chapter aims at verifying the feasibility of the framework introduced in the previous chapter. Verifying the proper functioning is necessary because two operational assumptions of the framework are unmet, thus compromising the feasibility to operate as expected.

The framework operates as expected when the model captures the relevant factors in the domain (entirely), and the number of samples is thousands. However, our setting challenges the framework to operate on few examples (only hundreds) and with an incomplete model.

Our purpose for enforcing such a mismatch is to resemble human beings' natural condition and capability. The rationale is that people manage to draw robust and valid inferences despite incomplete mental models and few examples (sometimes even from a single one). The framework must handle such conditions to equip robots with the necessary capability to observe as humans do.

Two complimentary evaluations verify the feasibility of the framework - one involves checking for correctness (validity) and the other for robustness (stability). Validity ensures that the framework derives correct estimates, while stability checks that estimates remain robust despite perturbing the setting. Each of them is further described in the following.

The verification for validity checks that confidence intervals contain an expected value while stability perturbs components of the framework to assess whether estimates remain within certain bounds. Both evaluations are independent and presented

separately.

The two verifications are carried out on the same collection of data. An experimental study is performed to acquire the ground-truth data. The experiment's purpose is only to influence hand behavior minimally such that an expected outcome is implanted in the data. The framework operates properly when providing ground-truth data, and the desired outcome is recovered across evaluations.

Overall, the verification strategy shown in Figure 5.1 establishes ground-truth data to verify the framework's feasibility. Then, using the validation data, the estimates provided by the framework are verified for correctness and robustness.

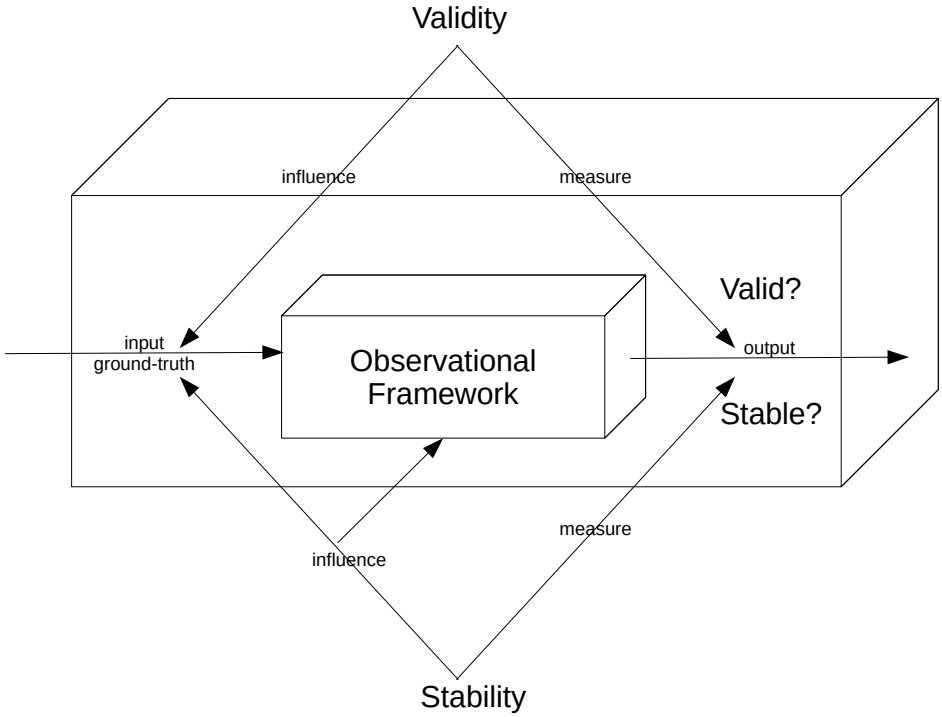


Figure 5.1: *The framework's feasibility comprises the evaluation of the validity and stability of ground-truth data.*

The presentation of this chapter encompasses three sections, as shown in Figure 5.1. Section 5.1 describes the experimental study and the resulting ground truth data. Section 5.2 verifies the capability of the framework to derive inferences that are correct, while Section 5.3 evaluates the framework's inferences to remain stable. Last, Section 5.4 concludes on the validity and stability of the framework.

In particular, the experimental study described in Section 5.1 includes the experimental procedure (Section 5.1.1), defines the experimental design (Section 5.1.2), summarizes the evidence collection (Section 5.1.3).

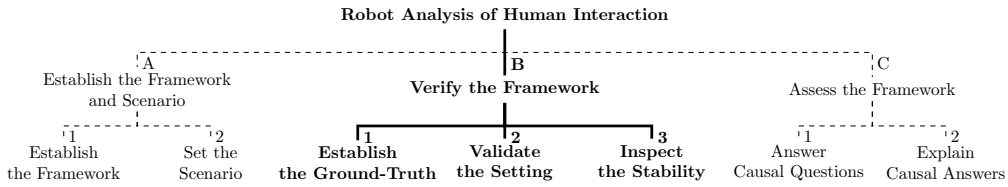


Figure 5.2: *The evaluation strategy consists of an experiment establishing the ground-truth data while the validity and stability verify the proper functioning and value of the framework.*

The verification for correctness presented in Section 5.2 provides a detailed analysis of a single hypothesis (Section 5.2.1), extends the analysis to further hypotheses (Section 5.2.2), and last, Section 5.2.2 summarizes the results.

The evaluation of stability presented in Section 5.3 first introduces the perturbation strategies (Section 5.3.1), describes the stability for a single hypothesis (Section 5.3.2), and evaluates further hypotheses (Section 5.3.3).

5.1 Establish the Ground Truth

Certifying the framework recovers valid and stable estimates requires ground-truth data. The rationale for running an experiment is influencing some aspect of the evidence such that one can expect a given outcome on the resulting data. Then, the framework can be tested by recovering the expected outcome when using the ground-truth data.

The main concern of this section is describing the acquisition of ground-truth data. Approving or disproving the framework's ability to recover the expected outcome is the subject of subsequent sections.

An experimental study establishes the ground truth data. The empirical research presented in this section follows the standard methodology. Section 5.1.1 defines the experimental procedure (describes the context, formulates the theoretical model, states the hypothesis to be tested), Section 5.1.2 describes the experimental setup (sketches the experimental setting, states the equipment, describes the experimental protocol, and method for the analysis), and Section 5.1.3 summarizes the collection of evidence.

5.1.1 Experimental Procedure

The experiment aims to produce a collection of evidence against which the framework can be verified. The investigation involves human behavior in the context of a table-setting activity.

Essentially, a table setting involves displacing objects from one spot in the kitchen to another. Transporting an object requires a person to hold a given target with the hand. Grasping an object has the purpose of retaining a target in hand.

How people select their hands in everyday chores is not fully understood. Every time a person grasps an object, a particular hand is chosen. Understanding hand selection is problematic because many factors can have an influence.

The graph in Figure 5.3 specifies a set of factors that could affect hand selection. The model expresses potential connections ("X could affect Y"). In the graph, an arrow pointing from X to Y asserts that the former could affect the latter. For instance, the relationship connecting object-volume with hand-selection means that the size of an object could make a person favor one hand over another.

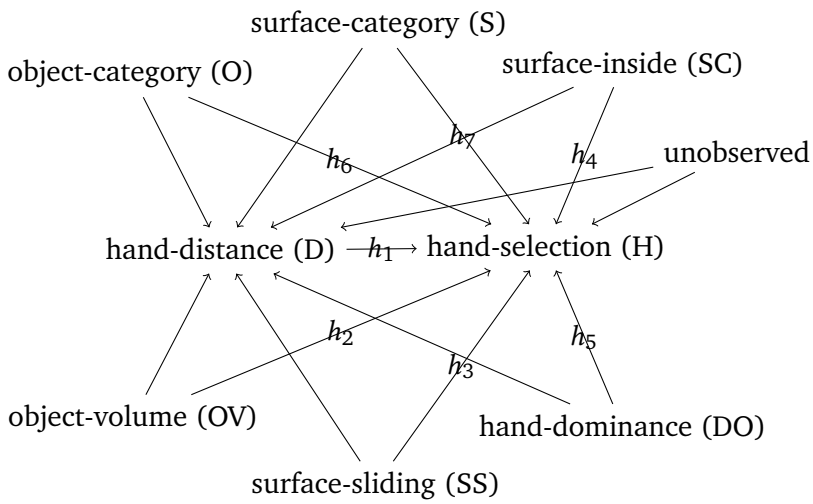


Figure 5.3: *A theoretical model for the experimental study.*

Moreover, a person might decide on

- using the stronger hand for large objects (object-volume),
- favor the most proximate hand to a given target (hand-distance),
- choose the more skillful hand to accomplish a delicate task (object-category),

- opt for the spatially less-constraint hand for an easier grasp (surface-category),
- assign a particular hand with an assistive role (surface-sliding, surface-inside),
- prefer a hand over the other (hand-dominance).

The model of Figure 5.3 defines the variables involved in the experimental study and the hypotheses space. The theoretical model of hand behavior hypothesizes why a person decides to grasp an object with one hand and not the other. The model also hypothesizes on why one might grab objects near the body instead of far. Section 4.1.2 describes the construction of this model.

The graph clarifies with the node labeled "unobserved" that the model does not include all relevant elements. For example, the fact a certain configuration of objects arranged on a surface can drive hand-selection, i.e., favor or hinder one hand over the other. But the model does not capture this factor. Similarly, the person's mental state when grabbing an object is neither included in the model. Unobserved factors are relevant, and not including these in the model can bias inferences.

Every relationship in the model of Figure 5.3 is, in principle, a candidate hypothesis for the experimental study. Ideally, one would like to validate all hypotheses included in the graph, but running experiments is costly and not always feasible.

Experiments typically manipulate one variable to determine whether change causes variation in another, given other factors remain unchanged. If variations are significant, it hints at the causal factor being the experimental influence thereby supporting a given hypothesis.

An experiment can either enhance or diminish variation. Enhancing allows testing for the presence of effects by inhibiting variation, whereas diminishing allows testing for the absence by reducing variation indirectly.

Any of these strategies is sufficient to validate the proper functioning of the framework. To attest the framework's feasibility to operate out of the design (hundred samples and unmodeled factors), it is enough verifying the framework can succeed at least in one of the two strategies, namely, correctly recover the variation, or, successfully deriving no unwanted variation.

In either of the cases, failing to infer valid variation suggests that atypical conditions are too challenging. Meaning, robots equipped with the approach would need near-perfect models and more data to draw inferences properly, which is a red flag for the approach.

This thesis favors the experimental strategy that diminishes variation (test for the absence) over the design that enhances (test for the presence). The reason is

that testing for the presence involves thirteen experiments (one experiment per hypothesis) to fully validate the graph, whereas testing for the absence requires two. The experimental design that enhances variation is thus not pursued.

The limitations for testing only for the absence of variation (or vice versa), but not both, account only for one type of mistake, namely, showing variation when there is none, but does not verify the case – showing no variation where there is one.

Nonetheless, as stated earlier, the objective is not "verifying inferences make no mistake" but whether the approach can operate out of operational design and address near natural behavior. Ensuring that inferences overcome one of the mistakes is sufficient to reveal that deriving behavior under atypical conditions is feasible.

With the chosen experimental strategy one could validate all hypotheses included in the graph with two experiments, one targeting hand-distance and another hand-selection. However, only one experiment is feasible in practice.

The experiment targeting hand-distance is abandoned for the following difficulties. Instructing participants with "Set a table, grasping objects near your body (max. 20 cm away)." is hard to follow. Restricting hand-distance forces a participant to think about aspects they would typically not consider making forgetting the restriction easy while performing the activity, or inhibiting participants specializing on interactions where restrictions are clearly met (max. 20 cm). Another difficulty arises when attempting to comply fully with the instruction. For instance, grasping an object deep inside a refrigerator without stretching the arm beyond 20 cm is not possible. One cannot expect compliance in such a situation. One way to sidestep this issue is to avoid such interactions. For example, changing the virtual environment by removing the refrigerator, or the objects placed deep inside it. Figuring out all situations where the participants clash with impossible instances is itself a non-trivial task. For example, one would need to restrict interactions with bottom-level drawers but not necessarily with top-level drawers. The depth at which objects are placed inside containers depends on the participant's arm length, and so on. Overall, these issues lead to unnatural hand behavior.

Featuring natural behavior is crucial for investigating the viability of the approach. An experiment featuring near natural conditions is paramount to challenge the approach's practical viability. Because the experiment addressing hand-selection features near natural conditions, and the other targeting hand-distance compromise natural behavior with artificial laboratory conditions, only the former is embarked.

For these reasons, to leverage natural behaviors in participants and validate multiple hypotheses with few experiments, hypotheses are tested for the absence of effects, thereby involving an experimental design that diminishes effects instead of enhanc-

ing them. Testing for the absence of effects calls for an experimental design that diminishes effects.

The experimental procedure selected to test for the absence of effects is the Null Hypothesis Significance Testing (NHST) [Per16]. NHST defines the null hypothesis as the absence of an effect and the alternative hypothesis set to the presence of an effect. Consequently, near-zero values are expected for the null hypothesis instead of different-than-zero.

To evaluate the NHST, effects are neutralized by conducting a randomized control trial (RCT). RCT asymptotically neutralizes any effect of measured, unmeasured, and unknown factors. Therefore, RCT diminishes any effects on the selected target to randomize, and as a result, evidence collected with this experimental design can test for no-effects on multiple relationships.

The target for the randomized control trial (RCT) is hand-selection. Consequently, all effects of factors influencing hand-selection are neutralized. In terms of the graph in Figure 5.3, this means RCT neutralizes the arrows pointing to hand-selection. The arrows are labeled to distinguish that labeled hypotheses in the graph are testable while those unlabeled are not.

In terms of the graph in Figure 5.3, RCT neutralizes the arrows pointing to hand-selection. These arrows are labeled to distinguish that the set of labeled hypotheses in the graph are testable while those unlabeled are not. For instance, one can expect a null effect for object-volume on hand-selection (labeled with h_2) but not necessarily on hand-distance (thus unlabeled in the graph).

Technically, the framework allows targeting any of the hypotheses modeled in the graph of Figure 5.3 (which is done in Chapter 6). However, the experimental evidence can only serve the purpose of validating the subset of labeled hypotheses in the graph.

The set of testable hypotheses with the experimental data is:

h_1) Does hand-distance affect hand-selection?

h_2) Does object-volume affect hand-selection?

h_3) Does surface-sliding affect hand-selection?

h_4) Does surface-container affect hand-selection?

h_5) Does hand-dominance affect hand-selection?

h_6) Does object-category affect hand-selection?

h_7) Does surface-category affect hand-selection?

5.1.2 Experimental Setup

The theoretical model shown in Figure 5.3 hints at the variables to measure during the experiment. The set of variables includes the three body parts (head and hands) along with contextual cues such as the target object grasped with the hand, the surface on which the target is placed, and properties such as the size of objects or whether a surface slides.

$$\begin{aligned}
&\text{hand-distance (D): distance of the hand w.r.t. head}\\
&\text{hand-selection (H): the person triggers a left or right grasp}\\
&\text{hand-dominance (HD): the person's hand dominance}\\
&\text{object-volume (OV): the volume box enclosing a 3D object}\\
&\text{object-category (O): the object's categorical name}\\
&\text{surface-category (S): the surface's categorical name}\\
&\text{surface-container (SC): the surface is inside a container or not}\\
&\text{surface-sliding (SS): surface slides or not}
\end{aligned} \tag{5.1}$$

Table settings are recorded within a virtual environment. The participant starts at the center of the kitchen layout with all movable objects initially stored inside furniture. The participant directs visual attention to any spot in the kitchen simply by moving the head around. Similarly, participants can freely walk around the scenario and control their virtual hands. Additionally, users can either grasp or release objects with the handheld device by gradually pressing or loosening a trigger. This functionality enables users to hold an object with a virtual hand and thus complete a table setting. Figure 5.4 shows the end state of a table setting.

The physical setup in the robot laboratory of the Institute for Artificial Intelligence Bremen University comprises two roof-mounted base stations opposing diagonally and curtains avoiding bright light. This setup complies with the official guideline [1].

The gear users employ to interact with the virtual environment is the HTC Vive Pro with the original handheld controllers. These VR devices are responsible for tracking user motions in three-dimensional space. The equipment tracking the user's body parts and interactions within the virtual scenario is shown in Figure 5.5.

A game engine (Unreal Game Engine) is responsible for recording user interactions (e.g., collisions and state changes) and rendering a near photo-realistic virtual-kitchen scenario. Figure 5.6 features the interaction in the virtual kitchen from a

[1] https://www.vive.com/eu/support/vive/category_howto/tips-for-setting-up-the-base-stations.html

Figure 5.4: *An example of a table setting featuring the milk, cereal, and juice carton; the glass and spoon; the head, hand-left, and hand-right.*

Figure 5.5: *A virtual reality gear comprising a head-set and two motion controllers.*

first perspective view as experienced by a user. A tour around the virtual kitchen is featured in the video [Wic22a].

Figure 5.6: *A first-person-view perspective of the interaction in the virtual kitchen.*

The following tools are employed to extract grasping instances from virtual demonstrations. The game engine logs streams of data with the tool described in [HB21] raw data annotated with semantics such as collisions and changes in object states. Semantics formats data as an episodic narrative that links to a specialized ontology for kitchen scenarios, including a taxonomy of categorical concepts [BBH+18]. The semantically annotated data enables extracting instances of grasping actions. Robots can also access such data directly with [Gay20]. Together these tools allow robots to access virtual activities.

Figure 5.7: *A sequence of actions leading to a table-setting. From left to right: targeting a bowl, a cereal carton, a bowl, a cereal carton, and the last image of the table set.*

The experimental protocol is as follows. The participant first grants permission to make his data public anonymously. After ensuring the person is wearing the VR gear properly, the participant is invited to explore the virtual-kitchen scenario for five minutes before starting the experiment. To begin, the experimenter asks the participant to stand on a particular marker on the laboratory floor. This physical

location in the lab coincides with the center of the virtual-kitchen layout, thereby avoiding participants spawning inside furniture. After the participant stands upon the patch on the physical ground floor, the kitchen scenario is initialized anew. The instruction for the participant is to - set the table for breakfast. To realize the experimental design, the experimenter instructs the participant to use a single (either left, right) or both hands before engaging in the activity. The sequence of instructions is chosen randomly [2] by the experimenter. The case of both hands is a distracting element to avoid suspicion by the participant on the matter being investigated. After the participants conclude the activity, the experimental protocol is followed anew:

1. the participant is asked again to stand on the patch,
2. the kitchen scenario is reinitialized, and
3. the instructions are repeated, except that due to randomization, a different constraint is communicated to the participant.

The unit of analysis is grasping actions a person performs during a table-setting activity in the virtual environment, not the number of participants. Even corrective grasps, such as picking up objects that tumbled or fell to the floor by mistake, are included in the population of the study. By definition, touching (or colliding against an object) is not a grasping action and, therefore, not part of the population.

The method for the study is a diagnostic analysis. The framework introduced in Chapter 4 is the measurement device under examination. The aim is to determine whether the framework recovers the expected effects from the experimental data. To quantify the analysis, the framework computes point estimates and confidence intervals at an alpha level of two sigmas. The point estimates and confidence intervals are computed with the four estimators introduced in Chapter 4 (listed in Table 4.1). These estimators are non-parametric and differ in the assumption made on heterogeneity. The confidence interval is computed with bootstrap for two estimators (LDRL and LDML) while Bootstrap-of-Little-Bags (BLB) for the others (CFDML and FDRL). Bootstrap provides confidence intervals from training multiple versions of the original estimator using random sub-samples with replacement ($n\_bootstrap\_samples = 100$). The quantiles are then calculated on the distribution of estimates across multiple samples over the entire dataset ($bootstrap\_type = pivot$). On the other hand, BLB constructs confidence intervals and quantifies the uncertainty of estimates based on Honest Random Forest [ATW19] by calculating the covariance of the parameter vector with an objective Bayesian debiasing correction to ensure that variance quantities

[2] `https://www.random.org/`

are positive. Further information regarding the underlying design of each estimator is available in the previous chapter (Table 4.2).

5.1.3 Collect the Evidence

The experimental data denoted with ds-v contains 26 table settings involving 126 grasping actions. A video presenting a table setting performed by the participant is featured in [Wic22b]. The frequencies are shown in Table 5.1 summarize the interactions of the participant during the experiment. Out of 126 grasps performed by the person, the target selected most of the time is a spoon (on 23 occasions). The person primarily targeted objects inside the sink-drawer-left-top (24 times) which corresponds to the drawer storing silverware such as the spoon. The hyphens in the table indicate no interaction. For example, the person never grasped the tray. For half of the grasps, the person stretched the arm near its body and the other far. Due to a median split, these frequencies are balanced out. The frequencies of hand selection are experimentally controlled, resulting in 61 and 65 for the left hand and right, respectively.

Despite randomization, the frequencies on hand selection are not completely balanced out. The explanation of this difference in frequencies is as follows. One cannot expect the total grasps performed with the left-hand to match the right-hand because the participant's actions are unrestricted. Once a particular hand is randomized and fixed at the beginning of the activity, the participant is free to choose any object to have on the table. Also, the order in which these are brought to the table varies, and so on. For this reason, a table-setting carried out with the left-hand does not necessarily coincide with a table-setting performed with the right hand. Henceforth, both numbers will likely not match due to the activity's uncontrolled nature but are expected to be similar.

5.2 Verify the Validity of the Framework

Recall that because hand-selection is randomized in the experiment, only weak (near to null) supporting evidence is expected on the validation data (ds-v). Particularly those hypotheses relating to hand-selection identified earlier in the graph of Figure 5.3 as testable.

User interaction			*Ds-v*
Object (O):	**Vol. (OV)**		**126**
Spoon	4.53×10^{-5}		23
Fork	2.48×10^{-5}		4
Glass	9.97×10^{-4}		21
Milk	1.22×10^{-3}		22
Juice	1.93×10^{-3}		19
Bowl	2.40×10^{-3}		15
Cereal	5.00×10^{-3}		22
Tray	9.95×10^{-3}		-
Surface (S):	**SC**	**SS**	**126**
DiningTable	F	F	-
FrdgArea	T	F	1
FrdgDrBtmShlf	T	F	20
FrdgGlassShlf	T	T	19
IslndArea	F	F	20
IslndDrwBtmLft	T	T	-
LabFloor	F	F	2
OvenArea	F	F	-
OvenDrwRight	T	T	20
SinkArea	F	F	1
SnkDrwLftBtm	T	T	1
SnkDrwLftMid	T	T	18
SnkDrwLftTop	T	T	24
Tray	F	F	-
Hand (H):	**Domin. (DO)**		**126**
Left (H_0)	F		61
Right (H_1)	T		65
Distance (D):			**126**
Close (HD_0)	F		63
Far (HD_1)	T		63

Table 5.1: *The collection of experimental evidence (ds-v) and corresponding frequencies of occurrences for different levels of features. The abbreviation* OV *refers to the volume of an object.* SC *corresponds to surfaces inside containers, such as the refrigerator's shelf.* SS *refers to surfaces that slide, such as drawers.* DO *refers to hand dominance.* SnkDrwLftMid *refers to the middle drawer out of three drawers stacked one over another spatially located left of the sink.* FrdgDrBtmShlf *refers to the bottom shelf on the fridge door. Last, levels* T *and* F *abbreviate true and false.*

Meaning the framework detecting the presence of effects for any of the (testable) hypotheses on the validation data indicates a red flag on the framework's validity. On the contrary, ideal results consist of estimates indicating no support for such data.

The following presentation dedicates a detailed analysis of hypothesis h_1 in Section 5.2.1 and extends over the remaining hypotheses in Section 5.2.2.

5.2.1 Analysis of a single Hypothesis

The hypothesis studied in subsequent sections is whether hand-distance drives hand-selection or not (h_1). This hypothesis involves two decisions people make (unconsciously) when grasping objects with their hands. These are the distance at which a person decides to grasp an object (i.e., the arm extension) and the choice of hand.

More formally, the hypothesis h_1 interprets as follows: "Does *hand-distance* (X) affect *hand-selection* (Y) given observable (Z) and unobservable confounders (U)?".

X	hand-distance (D): a person grasps near or far
Y	hand-selection (H): a person uses the left or right hand
Z_1	hand-dominance (DO): a person's hand dominance
Z_2	object-volume (OV): object's size in terms of volume
Z_3	object-category (O): object's categorical name
Z_4	surface-category (S): surface's categorical name
Z_5	surface-inside (SC): surfice is inside a container or not
Z_6	surface-sliding (SS): surface slides or not

Where X is the treatment (a.k.a. action or exposure), Y the outcome, and the rest of the variables are the relevant covariates (a.k.a. confounders). The backdoor-criterion (Definition 4.2) derives the following estimand for hypothesis H_1:

$$\frac{\partial}{\partial D}(E(H|O, OV, S, SS, SC)) \tag{5.2}$$

The estimand for H_1 is then evaluated with data using four methods. The four estimators that report the results are non-parametric, suitable for high-dimensional data, provide valid confidence intervals and operate under the unconfoundedness assumption.

After computing the effects on the validation set, four estimators recovered no

Hypothesis	Samples	Estimator	Effects	Conf. Int.
1	126	LDML	0.12	(-0.38,0.63)
		LDRL	-0.10	(-0.72,0.52)
		CFDML	0.04	(-0.28,0.36)
		FDRL	-0.00	(-0.29,0.29)

Table 5.2: *The effects for hypothesis h_1 on experimental evidence ds-v at an alpha level of two sigmas.*

evidence in support of the hypothesis. This is indicated in Table 5.2 by the confidence intervals, including the zero value for all estimators. The result is clear - the framework recovers the expected effect.

Despite targeting a few hundred samples with an incomplete model (breaking the unconfoundedness assumption), the framework correctly detects the expected null-effect repeatedly with four estimators operating under a different design.

To conclude, the approach recovers the expected effect on the validation set. Moreover, four estimators operating under different assumptions agree on the outcome.

This completes the validation of the framework on hypothesis h_1. The analysis provided here is the first of a series of hypotheses covered next.

5.2.2 Analysis of multiple Hypotheses

The analysis in the previous section confirmed the absence of the effect for hypothesis h_1. In the following, the focus is on the remaining hypotheses h_2 to h_7. These are tested employing an identical experimental setting and collection of data as before—only the hypothesis changes.

The effects shown in Table 5.3 cover the set of testable hypotheses that potentially influence hand-selection (i.e., those arrows labeled in Figure 5.3). Each row of the table reports the effects for a particular hypothesis. The first row repeats hypothesis h_1 described in the previous section (reported in Table 5.2), while results for h_2 to h_7 are presented next.

After computing the effects on the experimental evidence, the estimators recover no evidence in support for hypotheses h_2 to h_7. This is indicated in Table 5.3 by the confidence intervals, including the zero value for all estimators. These results show that the framework recovers the expected absence of effects for the remaining hypotheses.

Hypothesis	Samples	Estimator	Effects	Conf. Int.
1	126	LDML	0.12	(-0.38,0.63)
		LDRL	-0.10	(-0.72,0.52)
		CFDML	0.04	(-0.28,0.36)
		FDRL	-0.00	(-0.29,0.29)
2	110	LDML	0.00	(-0.64,0.64)
		LDRL	0.09	(-0.32,0.49)
		CFDML	0.11	(-0.29,0.50)
		FDRL	0.06	(-0.26,0.38)
3	126	LDML	-0.08	(-0.52,0.36)
		LDRL	-0.04	(-0.37,0.29)
		CFDML	-0.11	(-0.48,0.26)
		FDRL	-0.07	(-0.35,0.20)
4	126	LDML	0.12	(-1.74,1.98)
		LDRL	-0.09	(-0.36,0.19)
		CFDML	0.26	(-0.57,1.09)
		FDRL	-0.07	(-0.27,0.12)
6	122	LDML	-0.66	(-3.15,1.84)
		LDRL	0.05	(-0.16,0.26)
		CFDML	-0.67	(-3.97,2.63)
		FDRL	0.09	(-0.07,0.25)
7	121	LDML	3.18	(-2.52,5.88)
		LDRL	0.04	(-0.23,0.31)
		CFDML	0.97	(-1.46,3.39)
		FDRL	0.02	(-0.26,0.30)

Table 5.3: *The effects for hypothesis h_1 to h_7 at alpha-level of two sigmas on the experimental evidence ds-v.*

The confidence intervals reported in Table 5.3 show DRL estimators (DRL,FDRL) provide narrower confidence intervals than their DML counterparts (LDML, CFDML). In turn, estimators operating under linear-heterogeneity (LDRL) exhibit narrower confidence intervals than the variant linear-treatment (LDML), being the only exception h_1 covered in previous section. A possible explanation is an underlying heterogeneity being indeed linear (case of LDRL) but not the treatment (case of LDML). Here again, beyond distinctions, all estimators recovered the correct quantity.

The reason some of the hypothesis are computed on less samples than others is due to a technical reason. Variables must contain at least 10 examples per class when set as the treatment variable. Because of extreme imbalances in data, several cases do not comply with this requirement. To solve this issue, samples are dropped, and for this reason some of the effects are computed on a smaller number of samples. Moreover, hand dominance (hypothesis h_5) is constant (not expected to change in a single person) and thus safely removed from the analysis.

Results on the Validity of the Framework

The results reported in Table 5.3 support the validity of the framework. The framework correctly detects the expected null-effect despite the freedom in the activity (free choice of objects, order, placement, to some extent reflected by the asymmetries in Table 5.1). Moreover, the framework operated with a model capturing only part of the relevant factors, thus compromising the capability to provide bias-free estimates. Furthermore, the estimators typically operate on data at the level of thousands of samples managed by dealing with only hundreds. Despite small quantities of data, the framework is sensible enough to detect the expected null-effect.

The results indicate that the framework recovers non-trivial effects from human behavior. Consequently, robots equipped with this framework could in principle draw inferences from behavioral examples in natural conditions.

It is worth mentioning that Table 5.3 demonstrates the framework's capability to detect the absence of effects. A related question remaining unanswered is whether the framework is sensible enough to detect the presence of effects. It could be the case that the framework is insensible to detect any effect at all. This possibility does not hold as shown by the effects in Table 8.3 reporting non-null values for those relationships unaffected by the experiment. This section concludes the validation of the framework. Next, the stability of the framework is verified.

5.3 Verify the Stability of the Framework

In the following, the focus is on the framework's ability to remain stable under varying conditions. Verifying for stability consists of perturbing the framework to discover instabilities. Instabilities signal difficulties for the framework to provide stable estimates. Perturbing the inference process is paramount for generalization and thus robot applications. To some extent perturbing the framework resembles the simulation of potential scenarios where key assumptions are stressed. This evaluation is similar to sensitivity analysis in machine learning, only stronger as more than data is involved.

Essentially stability verifies outputs after perturbing some components of the framework. The overall strategy is (shown in Figure 5.8) consists of destabilizing the framework and then measuring the output estimates again. The evaluation for stability involves two measurements, the estimates provided by the framework before perturbing the original setting and after. Ideally, in a robust framework, the second measurement (the re-computed values) differs only slightly from the original estimates (the baseline values). Consequently, significant differences between the baseline and the re-computed estimates indicate instabilities.

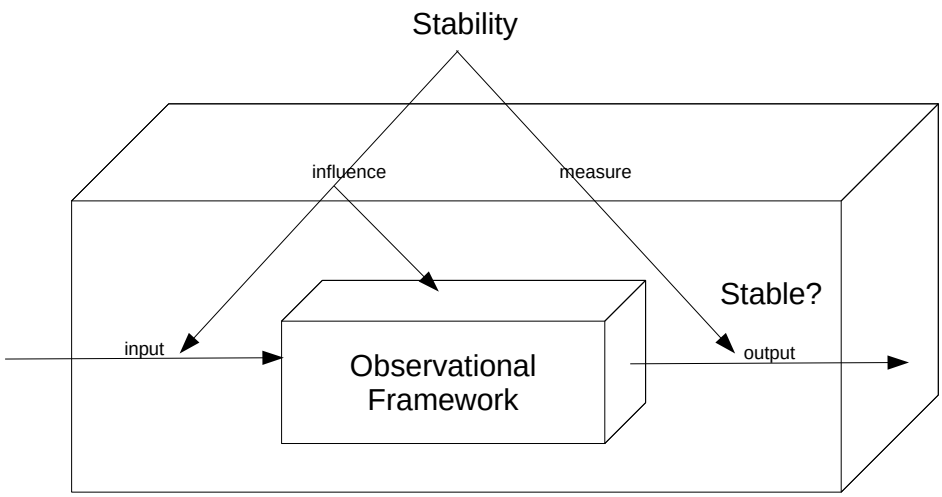


Figure 5.8: *The strategy to assess the stability of the framework.*

The model and data are the two aspects of the framework targeted to perturb the inference processes. Models because these are expected to be applicable beyond scenarios for which they were initially designed. Data because the inference process is supposed to be robust against varying input conditions.

Instabilities that arise due to perturbation to the model signal difficulties to generalize over scenarios other than those initially conceived (e.g., relevant factors not fully captured by the model). Therefore, it is crucial to verify a model under different conditions than originally planned. For example, does an additional node in the model result in a more stable estimate?

Instabilities that arise when perturbing samples indicate sensibilities to assumptions on the data (e.g., order in the data points or class imbalances). Therefore, it is crucial to verify whether the estimators of the framework can handle any data. For instance, does altering the order of data points drastically change estimates?

To summarize, perturbation strategies target the model, the data, or both, by introducing specific modifications to the structure of the model or the sample distribution. Perturbation strategies are more formally described in the following section.

5.3.1 Perturbation Strategies

The following methods are employed to perturb the framework [SK20]:

Check-1 Add Random Common Cause (RCC) graphically this method adds several nodes with an arrow pointing to the treatment and another to the outcome (i.e., a common cause) in an attempt to discover factors that should be part of the model but are missing. The estimates should not change after adding an independent variable with associated random generated data to the graph.

Check-2 Dummy Outcome Refuter (DOR) replaces the original values of the outcome variable with a custom distribution which typically is a uniformly random distribution. This strategy virtually acts as a placebo thus (near) zero values are expected in this verification.

Check-3 Data Subsets Validation (DSV) modifies the data by replacing original values with random subsets. This modification should not change estimates significantly.

The point estimates reported after perturbation are based on 100 simulations. Only some of the strategies require parameters to configure. Check-2 is such a case, where the parameter effect-strength on the outcome is set to 0.02. In the case of Check-3, the parameter subset-fraction is set to 0.9.

Once refutation strategies are set, these are automated tests that can serve the purpose of diagnostic evaluations in robot applications. For instance, such verification can

alert on a mismatch of the model and novel input data (i.e., a novel scenario the robot encounters).

5.3.2 Stability of a single Hypothesis

Table 5.4 summarizes the results of three refutation strategies. The first column (effect) corresponds precisely to the original values presented in Table 5.2 (before perturbation). In contrast, the remaining columns correspond to the re-computed estimates after inflicting changes with a given refutation strategy (columns Check-1 to Check-3). This table reports the stability for hypothesis h_1 in particular.

Hypothesis	Samples	Estimator	Effect	Conf. Int.	Check-1	Check-2	Check-3
1	126	LDML	0.12	(-0.38,0.63)	-0.06	0.01	0.11
		LDRL	-0.10	(-0.72,0.52)	0.03	-0.03	-0.02
		CFDML	0.04	(-0.28,0.36)	0.01	0.00	0.03
		FDRL	-0.00	(-0.29,0.29)	0.02	0.01	0.04

Table 5.4: *The stability of estimates at alpha-level two sigmas for the hypothesis h_1 on 126 samples from the experimental evidence ds-v. Re-computed values within the confidence intervals are an indicator of reliability, except for Check-2 (DOR), where near-zero values - regardless of prior estimates - indicate stability.*

The re-computed effects (Check-1) are calculated after adding a common cause with random data. These should not differ from the original values. This is, the re-computed effect remains within the confidence interval. Indeed such is the case, as Table 5.4 reports re-computed estimates well within the original confidence interval for all estimators, presenting LDML as the largest value but well within the interval.

Unlike Check-1 (and Check-3) where prior baseline values matter, in the case of check-2, near zero values - regardless of prior estimates - indicate stability. When randomizing the outcome, close to zero values are reported across estimators indicating robustness.

Last, removing random subsets of data (Check-3) exhibits estimates well within the confidence intervals.

Overall, the estimates after perturbing the system are within the confidence intervals. These results support the stability of the framework against three perturbation strategies for hypothesis h_1. The next section verifies further hypotheses.

5.3.3 Stability of multiple Hypotheses

The previous section shows the stability of the framework for a single hypothesis. A greater portion of the causal graph is tested by targeting more hypotheses.

Table 5.5 reports the stability for hypotheses h_1 to h_7, where each row corresponds to a hypothesis. Note that the first row of the table repeats the results for h_1 presented earlier in Table 5.4.

H	Samples	Estimator	Effect	Conf. Int.	Check-1	Check-2	Check-3
1	126	LDML	0.12	(-0.38,0.63)	-0.06	0.01	0.11
		LDRL	-0.10	(-0.72,0.52)	0.03	-0.03	-0.02
		CFDML	0.04	(-0.28,0.36)	0.01	0.00	0.03
		FDRL	-0.00	(-0.29,0.29)	0.02	0.01	0.04
2	110	LDML	0.00	(-0.64,0.64)	-0.57	-0.07	-0.44
		LDRL	0.09	(-0.32,0.49)	-0.26	-0.01	-0.26
		CFDML	0.11	(-0.29,0.50)	-0.22	0.03	-0.23
		FDRL	0.06	(-0.26,0.38)	-0.15	0.01	-0.12
3	126	LDML	-0.08	(-0.52,0.36)	-0.03	0.00	-0.05
		LDRL	-0.04	(-0.37,0.29)	-0.21	0.00	-0.18
		CFDML	-0.11	(-0.48,0.26)	0.01	-0.03	-0.00
		FDRL	-0.07	(-0.35,0.20)	-0.02	0.00	-0.03
4	126	LDML	0.12	(-1.74,1.98)	-0.12	0.00	-0.11
		LDRL	-0.09	(-0.36,0.19)	-0.14	-0.06	-0.21
		CFDML	0.26	(-0.57,1.09)	-0.12	-0.02	-0.15
		FDRL	-0.07	(-0.27,0.12)	-0.11	0.01	-0.11
6	122	LDML	-0.66	(-3.15,1.84)	-1.22	0.30	-0.86
		LDRL	0.05	(-0.16,0.26)	0.06	-0.01	0.08
		CFDML	-0.67	(-3.97,2.63)	0.91	0.04	0.74
		FDRL	0.09	(-0.07,0.25)	0.06	-0.02	0.10
7	121	LDML	3.18	(-2.52,5.88)	-1.15	0.30	-2.51
		LDRL	0.04	(-0.23,0.31)	-0.08	-0.01	0.08
		CFDML	0.97	(-1.46,3.39)	0.96	-0.15	0.61
		FDRL	0.02	(-0.26,0.30)	-0.08	0.03	-0.04

Table 5.5: *The stability of estimates at alpha-level two sigmas for hypotheses (H) affecting hand-selection (h_1 to h_7) using* 4 *estimators targeting experimental evidence* $ds - v$ *on* 126 *samples.*

Like the previous section, the column *effect* corresponds to estimates before perturbing the framework, whereas the remaining columns are re-computed values after destabilization (the last three columns in the table). Generally, the smaller the check values are, the more stable estimates are (except for Check-2 where a near zero value is a good indicator).

Overall, Table 5.5 reports that the framework is robust across hypotheses against refutation methods Check-1, Check-2, and Check-3 for hypotheses h_1 to h_7. This is indicated by re-computed estimates under column *check*, which are within the confidence interval, with no exceptions. However, estimator LDML and CFDML are close in providing estimates that fall outside the range, specially in the case of h_2 and h_7, followed by LDRL then FDRL also in the case of hypotheses h_2 and h_7. These hypotheses relate to variables object-volume and surface respectively.

The framework deems robust as values indicate stable estimates within the range of the confidence intervals. These results mean that the hand manipulation model is robust against domain variations and thus has good chances to transfer across scenarios.

5.4 Concluding Remarks

This chapter evaluates the feasibility of the approach by first establishing a ground truth by conducting an experimental study (the experimental procedure, design, and data collection given in Section 5.1). Provided the validation data, the framework's capability to derive valid inferences is verified (Section 5.2 first describes a single hypothesis and extends the analysis to others). Last, the framework's stability is perturbed (Section 5.3 introduces the perturbation strategies, describes the robustness of a single hypothesis then extends to others).

The tables present the main results of the chapter. Table 5.1 summarizes the validation data, Table 5.3 shows the effects for hypotheses h_1 to h_7, and Table 5.5 reports their stability.

The numbers reported in these tables indicate that the framework recovers the null effect on the validation data. Four estimators operating under different operational designs agree on this result. Moreover, the estimates remain stable against the three refutation strategies.

It should be noted that Table 5.3 indicates the framework detects the expected

absence of effects (i.e., a null effect). However, it could be the case the framework is insensible to detect any effect at all. This possibility does not hold given the effects shown in Table 8.3 reporting non-null values where the framework detects the presence of effects hypotheses h_8 to h_{13}).

Note that hypotheses h_8 to h_{13} are not validated because the experimental study targets hand-selection but not hand-distance. An experiment controlling for the latter is challenging (if not infeasible), as described by the experimental design in Section 5.1. Because these hypotheses were not influenced experimentally, any effect can be expected for them, not only the null value. Table 8.3 in the appendix report that, in indeed, hypothesis h_9 derives a negative effect, hypotheses h_{10} a positive (except for one estimator), and null values for the rest. This result validates the experiment and confirms the framework's sensitivity detecting non-null effects (the presence of effects).

Before concluding, it is worth contextualizing the conditions under which the framework provides the reported numbers. First, the framework operates with a model capturing part of the relevant variables (not all); thus, potentially biased, instead of bias-free, estimates are expected because the latter is not guaranteed anymore (i.e., the unconfoundedness assumption is broken). Second, the non-parametric estimators typically target samples in the order of thousands, but only hundreds are considered – even worse, the number of samples per class is unbalanced. These conditions are enforced because people can learn from a few examples and with incomplete mental models (especially children). It is under these conditions the framework draw inferences challenging its feasibility.

Last, the validation of the approach covers one type of error, namely, showing an effect when there is none, but does not verify the case – showing no effect where there is one.

This section concludes the evaluation of feasibility given by the evaluation of validity and stability. Next, Chapter 6 investigates the utility of the framework, where the setting engages with examples from individuals who freely choose their hands in table-setting activities.

Chapter six

Utility of the Framework

The previous chapter verified the framework's ability to derive valid and reliable inferences. The present chapter evaluates the reasoning capabilities of the framework. The framework faces the scenario and conditions confronted previously in Chapter 5 (few samples and a model not fully capturing the relevant factors), except that the data source is not experimentally influenced. Like in the previous chapter, these conditions make the evaluation unique and are thus further emphasized below.

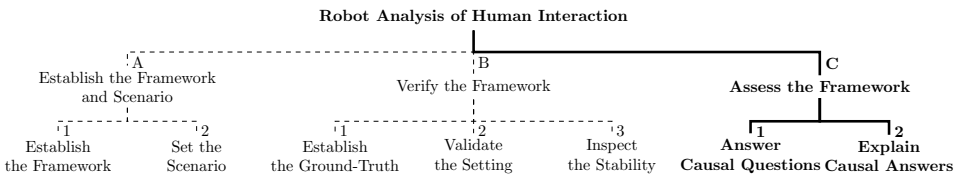


Figure 6.1: *The strategy to evaluate the framework's utility.*

This chapter evaluates the utility of the framework. The evaluation consists in formulating causal questions, answering them with the framework, and inspecting the answers. The framework is useful when answers to causal questions are meaningful and informative despite the challenging conditions of the scenario. The causal questions aim to uncover hand behavior in two subjects. Both of them perform table settings, and are not instructed how to perform the activity (this promotes natural behaviors).

In principle, robots that infer and explain why X matters on Y have the potential to

uncover human behavior themselves. Assessing this potential of the framework is the goal. All answers reported in this chapter are technically attainable by robots equipped with the framework.

The inferences described in this chapter draw on data not influenced by any experimental design, also regarded as observational or uncontrolled evidence. The two collections of uncontrolled evidence employed to evaluate causal questions are introduced in Section 6.1.2.

Drawing causal relationships from uncontrolled evidence is only viable because the framework provides the mechanisms to account for confounding bias, in particular with the identification and adjustment steps described in Sections 4.1.1.2 and 4.1.2.2. The framework essentially automates quasi-experiment for robots.

However, two conditions challenging the approach's viability are crucial to emphasize here. First, because the framework incorporates an incomplete model, the underlying formalism does not guarantee bias-free estimates. Thus, the results should be interpreted as bias-reduced estimates instead of bias-free. Nonetheless, the previous Chapter 5 confirmed the framework manages to recover valid estimates despite an incomplete model (Table 5.3). Second, the small number of samples in the order of hundreds challenges the capability of the framework to derive meaningful estimates (ranging from 100 - 400 samples).

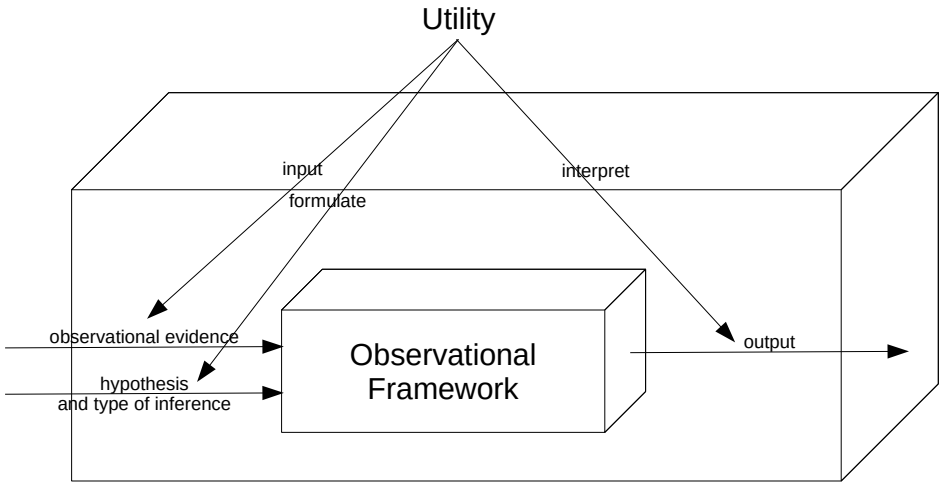


Figure 6.2: *The strategy to assess the framework's utility.*

An overview of the evaluation is shown in Figure 6.2. The framework supports three levels of inferences: what, what-if, and what-would-if. Each of these levels can be assessed by formulating a set of hypotheses, e.g., targeting hand behavior in table setting activities. Behavior is studied by observation; no evidence is experimentally

perpetrated. Given a set of causal queries and a collection of observational evidence, the task for the framework is to answer causal questions on hand behavior.

The outline of the chapter is as follows. First, the three levels of inferences are introduced with a single hypothesis (Section 6.1.1). The collection of uncontrolled evidence is described in Section 6.1.2. The hypotheses space is introduced in Section 6.1.3. The analysis initiated on a single hypothesis (Section 6.1.1) is extended over multiple hypotheses and two collections of evidence (Section 6.2), where the analysis of hand behavior for Person-A (Section 6.2.1) and Person-B (Section 6.2.2) are covered. The chapter concludes by providing insights on hand manipulation recovered with the framework.

6.1 Causal Interpretations as-if Experimental Observations

This section evaluates three types of causal questions, which are answered with the framework by supplying examples of hand behavior. A robot equipped with the framework is technically in a position to realize the same analysis presented in subsequent sections.

6.1.1 The Three Levels of Inferences

The stereotypical questions — what, what-if, what-would-if — correspond to levels in the ladder of causation (shown in Figure 3.4) and are presented as follows:

- what: *"Would the hand-selection be right-hand if the hand-distance is far?"*
- what-if: *"Would the hand-selection be right-hand if we make sure that the hand-distance is far?"*
- what-would-if: *"Would the hand-selection be right-hand had the hand-distance been far, given that the hand-selection, is in fact, left-hand and the hand-distance close?"*

What and Why

Suppose a robot witnesses several examples of table-setting activities and aims to set the table for breakfast itself. Neither robot nor the programmer has the recipe on how to manipulate objects with their hands. The task for the robot is then to understand a person's hand behavior by examining examples of hand manipulation to find answers to questions like the following:

> "Would the hand-selection be right-hand if the hand-distance is far?"

Alternative formulations of the same question are "Does hand-distance-far cause hand-selection-right?" or "Does hand-selection-right occur because of hand-distance-far?". The general formulation of these questions derives from "Would the Y be y if the X is x?".

Answering this question involves a prediction task. The collection of table settings witnessed by the robot conforms to the general distribution of table settings. With a select operation or equivalently conditioning the distribution, restricting cases to those where hand-distance is far would obtain the desired quantity. Except that the resulting quantity is potentially confounded. Select operations on table structures or conditioning in the case of probabilities, risk selection bias.

Robots operating on conventional approaches risk providing the wrong answer to the question. The framework presented in Chapter 5 enables robots to adjust automatically against confounding bias and overcome misleading answers. Unlike conventional methods, the quantities predicted with the framework make corrective adjustments to counter bias.

What-if and Why

Suppose a robot has already witnessed several table-setting activities and sees a person standing far away from a given target object. The next reasoning task for the robot is to answer what-if questions such as:

> "Would the hand-selection be right-hand if we make sure that the hand-distance is far?"

Alternatively, the previous question formulated in terms of symbols is "Would the Y be y if we make sure that the X is x?"

Note that answering this question does not involve a plain prediction task. In predictions tasks inferences draw on the general distribution, which does not necessarily reflect the present condition observed by the robot ("hand-distance is far"). This is because distributions are general, corresponding to table settings under varying conditions (i.e., not only those where hand-distance is far).

It is tempting to accommodate the general distribution to present conditions by issuing a select operation (e.g., conditioning examples to those where hand-distance is far). But select operations, or conditioning distributions, risk confounding bias. Robots relying on conditioning operations over general distributions are bounded to confounding bias. Answering the question requires inferences of the second order in the ladder of causation (Figure 3.4).

The framework presented in Chapter 5 allows adapting the general distribution to the present conditions without incurring selection bias. To answer the question raised before, the framework adapts the general distribution by issuing an intervention with the do-operator, and then draws inferences on the interventional (or truncated) distribution instead.

Different than the inferences (type what) described in the previous section, what-if infers on the interventional distribution. The intervention is hinted at the question by "make sure". What-if inferences collapse the potential world to a more specific condition than inferences of type what, particularly the distribution, is truncated with hand-distance set to far.

What-would-if and Why

Like in the previous case, suppose a robot already witnessed several table settings, watches a person standing far away from a given target object, and in addition, witnesses the person use the right hand to grasp an object. This time the reasoning task for the robot comprises answering the following question:

> "Would hand-selection be left-hand had hand-distance been close, given that hand-selection is, in fact, right-hand and hand-distance is far?"

The formulation in terms of symbols of the previous question is, "Would the Y be y_1 had the X been x_1, given that the Y is in fact y_0 and X is x_0?".

This question ponders on retrospective reasoning. The robot observes the person grasping an object with the hand but wonders about an alternative scenario. A

scenario where the person acted differently, i.e., a counterfactual scenario. Answering the counterfactual question requires a robot to imagine scenarios.

Attempting to answer this question with conventional methods is infeasible. A select operation (or conditioning) on the general distribution leads to an unsolvable conflict. The question requires to condition the distribution to cases where hand-selection is right-hand and those cases where the hand is left. Conventional methods fall short of proving answers to counterfactual questions.

The framework presented in Chapter 5 prepares robots to answer the counterfactual question. Like the inferences described in the previous section (what-if), what-would-if draws estimates on a modified distribution. The key difference being what-would-if infers on the counterfactual distribution instead of the interventional distribution.

Technically the distribution is modified with the do-operator assigning variables to specific feature levels (and overcoming the conflict mentioned before). The distribution is first adapted to the present condition (intervened) as described in the previous section, but then further do-operations are carried out on the distribution resulting in the counterfactual distribution.

Inferences concerning What and Why are covered after introducing the data collection and query space. Unfortunately, we cannot cover what-if and what-would-if promptly and adequately in time, even though the framework supports them. We leave our work documented so that others can continue.

6.1.2 The Collection of Uncontrolled Evidence

Two collections of evidence are presented next. Robots have access to these sets of data through the platform described earlier (Section 4.2.2). These data collections correspond to natural table-setting activities, which are in contrast to the experimentally controlled evidence (of Section 5.1.3) employed to validate the framework.

One of the data collections [PWYB18] is recorded to study how people arrange objects on the table (denoted here as ds-1). The other dataset [HB19] studies the time spent setting the table under various modalities (denoted by ds-2). Ds-1 and -2 originate from uncontrolled activities. Note that the unit of analysis is not the number of participants (only two) but grasping actions in the order of hundreds.

A table-setting activity involves a person moving objects from one place to another. Setting the table includes searching, transporting, placing objects in the kitchen - opening and closing containers, etc.

Hand behavior in each collection of data is different in spatio-temporal aspects such as the preference of objects for breakfast, the surfaces on which objects are placed, the hand employed to hold them, the distance at which these are grasped, etc. The dashes in the summary Table 6.1 evidence such asymmetries across data collections.

The frequencies are shown in Table 6.1, summarizing interactions during the table-setting activity. Out of 384 grasps in ds-1, the juice-carton is grasped most of the time (on 88 occasions). Moreover, the person did not interact with the tray, as hinted by the hyphen indicating no interaction. Furthermore, many grasps target objects located inside the refrigerator-glass-shelf (41 times), which is the initial position for the juice-carton. On nearly two-thirds of the occasions, the person grasps objects with the right-hand (263) instead of the left (121). The number of times this person stretches the arm near and far away from the body is balanced out due to a median split over the distance (described in Section 4.2.1).

In the case of $ds-2$, the frequencies in Table 6.1 report that out of 174 grasps, the silverware is targeted mostly (70 occasions). Unlike the previous person lacking in interactions with the tray, this person opted for the tray 32 times, and most of the grasps target objects that are placed on the tray (42 times). For more than half of the occasions (0.55) the person employs the left-hand (96 times) instead of the right (78). Like before, the number of times this person stretches the arm near and far away from the body is balanced out due to a median split over the distance.

The number of samples in either of these collections is small. Estimators targeting observational data typically operate with thousands of examples. Here only a few hundred are considered (384 and 174 samples). Moreover, both collections present disparate frequencies in classes, including none. Drawing causal interpretation on such data with conventional methods risks selection bias.

To summarize, the pipeline (introduced in Section 4.2.2) enables robots access to virtual activities. The activities are stored in machine-understandable format and processed with the framework. The evidence collections $ds-1$ and $ds-2$ are regarded as observational because the recorded activities are free from experimental design. Each collection belongs to the hand behavior of a single person and is denoted with Person-A and Person-B for $ds-1$ and $ds-2$, respectively.

The estimates derived on the validation data ds-v (presented in Chapter 4) are expected to derive specific quantities (i.e., the null-effect) because of the experimental design. This is not the case for observational data ds-1 and ds-2, on which inferences can derive any effect, including the null.

This concludes the description of the observational evidence. The next step is to process the collection of observational evidence using the reasoning capabilities

User interaction			*Ds-1*	*Ds-2*
Object (O):	**OV (m^3)**		**384**	**174**
Silverware	8.2×10^{-5}		83	70
Glass	1.22×10^{-3}		83	-
Milk	1.40×10^{-3}		18	8
Juice	1.93×10^{-3}		88	-
Bowl	2.40×10^{-3}		19	36
Cereal	6.30×10^{-3}		93	28
Tray	9.95×10^{-3}		-	32
Surface (S):	**SC**	**SS**	**384**	**174**
DiningTable	F	F	12	18
FrdgArea	T	F	-	1
FrdgDrBtmShlf	T	F	5	2
FrdgGlassShlf	T	T	41	-
IslndArea	F	F	-	8
IslndDrwBtmLft	T	T	-	16
LabFloor	F	F	5	-
OvenArea	F	F	-	2
OvenDrwRight	T	T	41	16
SinkArea	F	F	11	24
SnkDrwLftBtm	T	T	-	-
SnkDrwLftMid	T	T	41	-
SnkDrwLftTop	T	T	41	42
Tray	F	F	184	42
Hand (H):	**Domin. (DO)**		**384**	**174**
Left (H_0)	F		121	96
Right (H_1)	T		263	78
Distance (D):			**384**	**174**
Close (HD_0)	F		192	87
Far (HD_1)	T		192	86

Table 6.1: *The collections of uncontrolled evidence (ds-1, ds-2) and corresponding frequencies of occurrences for different levels of features. The abbreviation* OV *refers to the volume of an object.* SC *corresponds to surfaces inside containers, such as the refrigerator's shelf.* SS *refers to surfaces that slide, such as drawers.* DO *refers to hand dominance.* SnkDrwLftMid *refers to the middle drawer out of three drawers stacked one over another spatially located left of the sink.* FrdgDrBtmShlf *refers to the bottom shelf on the fridge door. Last, levels* T *and* F *abbreviate true and false.*

provided by the framework.

6.1.3 The Query Space

This section introduces the notation that identifies the formulations of causal questions. The query space comprises the expressiveness of the do-calculus, the graph, and feature levels. The formulations of causal questions in subsequent sections follow the notation

$$Type^{Hypothesis}_{Formulation} = \{A, B, C\}^{\{h_1, h_2, \ldots, h_n\}}_{\{1,2,\ldots,m\}}$$

where type refers to three types of inferences supported by the do-calculus (denoted letters A, B, C), the superscript to relationships in the causal graph (set of hypotheses h_i), and the subscript is a simple enumeration that depends on particular data feature levels chosen for a type of query.

Causal questions that belong to inference type what, what-if, and what-would-if, are identified by letters A^j_i, B^j_i, C^j_i, respectively.

The coding of questions includes the numbering of hypotheses as a superscript j to preserve the association with the graph. To this end, Figure 6.3 identifies the hypotheses space with a number assigned to each arrow of the graph. For instance, the superscript in $A^{j=h_1}$ associates a causal question to hypothesis h_1 in the graph.

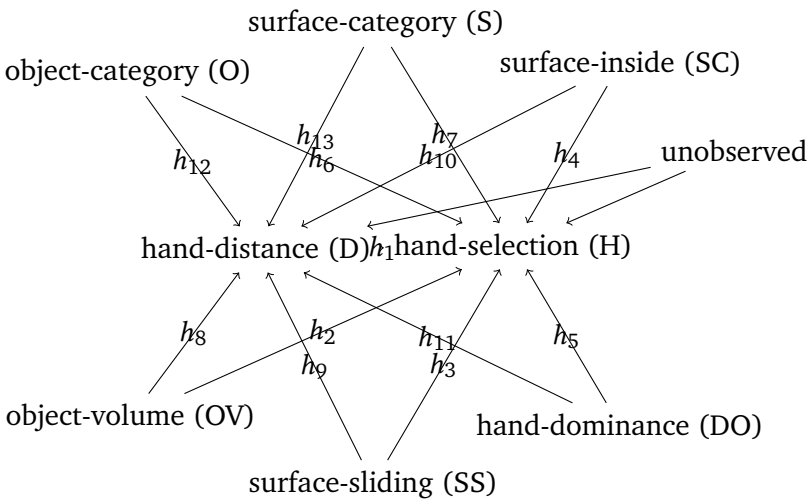


Figure 6.3: *An enumeration identifying relationships in the causal graph.*

Last, because hypotheses are associated with multiple formulations of causal questions, further notation is necessary. The subscript i in A_i identifies a particular

variation of a causal question. Selecting alternative feature levels formulates a different question. For example, inferences of type what-if such as "Would the Y be y if we make sure that X is x." specifies feature levels x and y. The specifications of these values result in different formulations distinguished with subscripts such as $A_{i=2}$ and $A_{i=1}$. When only a single formulation exists, the subscripts are removed (A^{h_1} instead of $A_1^{h_1}$).

The causal question "Would the Y=hand-selection be y=right-hand if the X=hand-distance is x=far?", identified by $A_{i=1}^{j=h_1}$ or simply $A_1^{h_1}$, associated with the stereotypical inference of type what (letter A), corresponds to hypothesis h_1 in the graph (Figure 6.3), where variables (X, Y) correspond to nodes hand-distance (D) and hand-selection (H) along with their feature levels ("right-hand," "far").

Out of the three types of inferences, what, what-if, and what-would-if, the following section starts with reporting the what while subsequent sections cover the remaining types.

6.2 Utility - Analysis of What and Why

The main concern is to report insights the framework can provide robots with. Subsequent sections report on multiple causal questions targeting the hand behavior of two individuals. A robot equipped with the framework is technically in a position to articulate the causal questions formulated here and, unattended infer the corresponding estimates presented in the following sections.

The presentation begins by introducing the formulation of causal questions in Section 6.2. To deal with a large number of questions, the presentation reports estimates separately per collection of evidence, those corresponding to the hand behavior of the Person-A (Section 6.2.1) followed by those of Person-B (Section 6.2.2). To further simplify the presentation, each report subdivides into two groups. One group reports the effects on hand-selection for Person-A and Person-B) followed by those on hand-distance correspondingly for Person-A and Person-B).

The Set of What and Why Questions

This section addresses the first of the three types of inferences supported by the framework. The first type of inference articulates causal questions of the type "What is the effect of X on Y?." Answering causal questions of type-what involve a prediction task with inferences drawing on the general distribution. This type of inference does not specify the do-operator to ensure external conditions (barring those potentially introduced by the identification process). The subscript is not necessary because a single variation of feature level is considered for this type of question (hence A^j instead of A_i^j).

The list of type what-questions targeting hand behavior is the following:

A^{h_1}) Does hand-distance (D) affect hand-selection (H)? Why?

A^{h_2}) Does object-volume (OV) affect hand-selection (H)?

A^{h_3}) Does surface-sliding (SS) affect hand-selection (H)?

A^{h_4}) Does surface-inside (SC) affect hand-selection (H)?

A^{h_5}) Does hand-dominance (DO) affect hand-selection (H)?

A^{h_6}) Does object-category (O) affect hand-selection (H)?

A^{h_7}) Does surface-category (S) affect hand-selection (H)?

A^{h_8}) Does object-volume (OV) affect hand-distance (D)?

A^{h_9}) Does surface-sliding (SS) affect hand-distance (D)?

$A^{h_{10}}$) Does surface-container (SC) affect hand-distance (D)?

$A^{h_{11}}$) Does hand-dominance (DO) affect hand-distance (D)?

$A^{h_{12}}$) Does object-category (O) affect hand-distance (D)?

$A^{h_{13}}$) Does supporting-surface (S) affect hand-distance (D)?

Related to each of the questions in the list is the query "Why does X affect Y?." Suppose the answer to a given causal question is positive, then the question why provides a causal explanation providing insights on "Which are the conditions to which hand-distance and hand-selection are sensible?." Explanations of either positive or negative answers hint on the primary factor driving the effect.

A robot that can answer the list of questions introduced earlier is in a position to reason over people's hand behavior. The list showcases only a subset of the addressable questions. The lists corresponding to what-if and what-would-if are made explicit subsequently.

6.2.1 What and Why on Hand-Behavior (Person A)

To answer the list of what-and-why questions, the framework infers on examples of hand manipulation performed by Person A (ds-1). The effects are reported in two groups. The first group of questions includes those factors that potentially influence hand-selection. In the graph, this group of questions corresponds to those nodes pointing to hand-selection. Instead, the second group is associated with those factors potentially affecting hand-distance (i.e., nodes in the graph pointing to hand-distance).

The Effects on Hand-Selection (Group 1)

Table 6.2 reports the first group of answers to the causal questions A^{h1} to A^{h7}. The first column of the table identifies the causal question with the notation introduced earlier. The second column indicates the number of samples. The third specifies the estimator. The fourth report is the estimated effect. The fifth shows the confidence interval. The sixth is a reference to a figure presenting the sensibility of the effect. Each row in the table answers a question on the list, i.e., the first row of the table answers the first question A^{h_1} on the list, the second answers A^{h_2}, and so forth.

In particular, row A^{h4} in Table 6.2 reports a point estimate 0.19 with values ranging between $(0.01, 0.38)$ using estimator LDML on 384 samples from dataset ds-1. Correspondingly, the other estimators derive effects 0.52, 0.24, 0.52 for LDRL, CFDML, and FDRL.

Each estimator operates under different operational assumptions and is thus not expected to coincide in the estimation. The estimators are listed in order from the most strict operational design to the least (LDML, LDRL, FDML, FDML) as described in Chapter 4.

A point estimate and confidence interval of 0.19 $(0.01, 0.38)$ means that modulating the treatment variable surface-inside (SC) from false to true causes an increase of 0.19

Type: What	Samples	Estimator	Effect	Conf. Int.	Sensitivity
$A^{h1} : (D \rightarrow H)$	384	LDML	0.02	(-0.29,0.33)	Figure6.4a
		LDRL	0.07	(-0.15,0.29)	
		CFDML	0.09	(-0.05,0.23)	
		FDRL	0.08	(-0.15,0.30)	
$A^{h2} : (OV \rightarrow H)$	384	LDML	-0.08	(-0.30,0.13)	Figure6.4b
		LDRL	-0.17	(-0.40,0.05)	
		CFDML	-0.20	(-0.41,0.01)	
		FDRL	-0.15	(-0.46,0.15)	
$A^{h3} : (SS \rightarrow H)$	384	LDML	-0.09	(-0.57,0.39)	Figure6.4c
		LDRL	-0.11	(-0.32,0.09)	
		CFDML	-0.03	(-0.38,0.32)	
		FDRL	-0.07	(-0.20,0.06)	
$A^{h4} : (SC \rightarrow H)$	384	LDML	0.19	(0.01,0.38)	Figure6.4d
		LDRL	0.52	(0.42,0.62)	
		CFDML	0.24	(0.04,0.44)	
		FDRL	0.52	(0.40,0.64)	
$A^{h6} : (O \rightarrow H)$	384	LDML	1.64	(-1.07,3.80)	Figure6.4e
		LDRL	0.07	(-0.19,0.33)	
		CFDML	0.59	(-0.16,1.34)	
		FDRL	0.02	(-0.25,0.28)	
$A^{h7} : (S \rightarrow H)$	371	LDML	-4.50	(-6.51,-2.49)	Figure6.4f
		LDRL	-0.12	(-0.75,0.51)	
		CFDML	-0.07	(-1.18,1.04)	
		FDRL	0.06	(-0.24,0.36)	

Table 6.2: *The type-what inferences for factors potentially affecting hand-selection in Person A. The framework estimates these effects with inferences type-what for hypotheses h1 to h7 on observational evidence ds-1 (384 samples).*

in the expected value of the outcome hand-selection (H) over the data distribution represented by dataset ds-1 with the increase ranging over the interval $(0.01, 0.38)$.

When a confidence interval includes only positive values, the causal relationship is positive, indicating that modulating treatment levels result in positive changes in the expected value of the outcome. An example of such a case shown in Table 6.2 is A^{h4}. Conversely, intervals ranging over negative numbers indicate a negative causal relationship where modulating treatment levels decreases the expected value of the outcome (an example of such a case is A^{h9} shown later). Last, the relationship is neutral when the confidence when interval range between positive and negative numbers, in which case, the interval includes the zero value. Neutral indicates that modulating treatment levels result in a positive and negative change in the outcome variable. An example of a neutral relationship is A^{h_1}.

The estimates on dataset ds-1 for group-1 are shown in Table 6.2 report:

- A^{h_1}: Increasing the variable hand-distance (D) from close to far causes an increase of 0.08 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.15, 0.30)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of 0.02 with values ranging over $(-0.29, 0.33)$, LDRL 0.07 $(-0.15, 0.29)$, and FDML 0.09 $(-0.05, 0.23)$. The causal direction of hand-distance (D) on hand-selection (H) is neutral (4 estimators) with the tendency positive.

- A^{h_2}: Increasing the variable object-volume (OV) from 0 to 0.001 (e.g. from spoon to glass) causes an increase of -0.15 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.46, 0.15)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of -0.08 with values ranging over $(-0.30, 0.13)$, LDRL 0.17 $(-0.40, 0.05)$, and FDML -0.20 $(-0.41, 0.01)$. The causal direction of object-volume (OV) on hand-selection (S) is neutral (4 estimators) with the tendency negative.

- A^{h_3}: Modulating the variable surface-sliding (SS) from false to true (e.g. table to drawer) causes a change of -0.07 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.20, 0.06)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of -0.09 with values ranging over $(-0.57, 0.39)$, LDRL -0.11 $(-0.32, 0.09)$, and FDML -0.03 $(-0.38, 0.32)$. The causal direction of surface-sliding (SS) on hand-selection (H) is neutral (4 estimators) with the tendency negative.

- A^{h_4}: Modulating the variable surface-inside (SC) from false to true (e.g. from a sink-area to a fridge) causes a change of 0.52 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(0.40, 0.64)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of 0.19 with values ranging over $(0.01, 0.38)$, LDRL 0.52 $(0.42, 0.62)$, and FDML 0.24 $(0.04, 0.44)$. The causal direction of surface-inside (SC) on hand-selection (H) is positive (4 estimators).

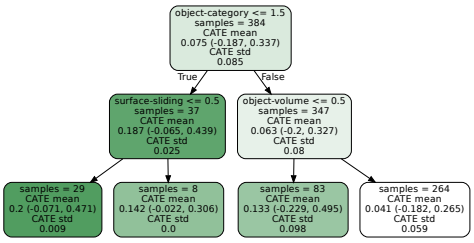
- A^{h_6}: Modulating the variable object-category (O) from milk to bowl causes a change of 0.02 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.25, 0.28)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of 1.64 with values ranging over $(-1.07, 3.80)$, LDRL 0.07 $(-0.19, 0.33)$, and FDML 0.59 $(-0.16, 1.34)$. The causal direction of object-category (O) on hand-selection (H) is neutral (4 estimators) with the tendency positive.

- A^{h_7}: Modulating the variable surface-category (S) from table to fridge-shelf causes a change of 0.06 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.24, 0.36)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of -4.50 with values ranging over $(-6.51, -2.49)$, LDRL -0.12 $(-0.75, 0.51)$, and FDML -0.07 $(-1.18, 1.04)$. The causal direction of surface-category (S) on hand-selection (H) is neutral (4 estimators) with the tendency negative.

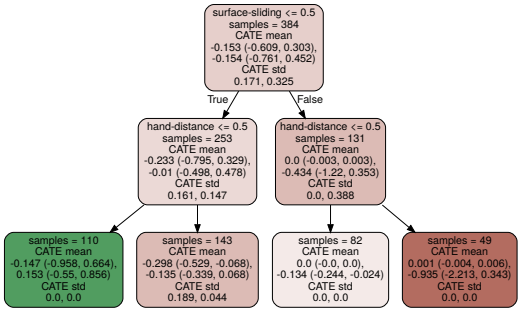
The previous paragraphs concerned the question "Does X cause Y?". In what follows, the concern is explaining change by inspecting the sensitivities of the effects "How does X influence Y?." This is addressed by interpreting a given effect with a tree structure.

Each effect reported in Table 6.2 has a corresponding tree structure interpreting the sensibility of the effect in Figure 6.4. The last column of the table references the figure presenting the sensitivity of the given effect.

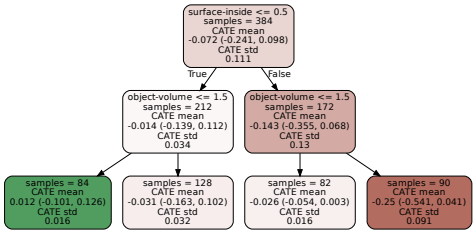
The tree structure ranks the factors according how sensible a given effect is. The more sensible an effect is to a given factor, the higher up a factor appears in the tree. Hence, the root node of a tree indicates the primary factor driving an effect. For instance, the tree of Figure 6.4a shows that the effect of hand-distance on hand-selection (h_1) is sensible to object-category (root node in the tree), more than it is w.r.t. object-volume and surface-sliding as these appear ranked lower in the tree. Nonetheless, because the tree is limited in-depth, the three factors are relevant in explaining the sensibility of the effect.



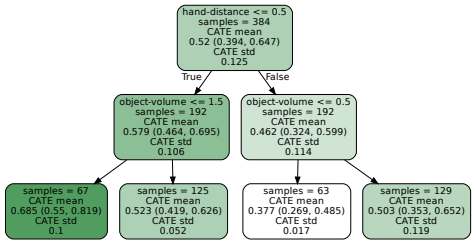
(a) *The effect of hand-distance on hand-selection ($h1$) is sensible to object-category in the positive direction.*



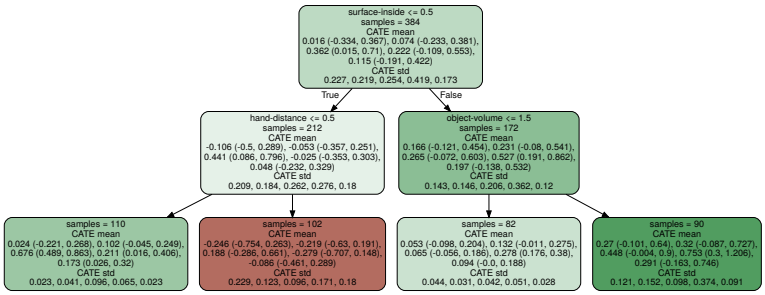
(b) *The effect of object-volume on hand-selection ($h2$) is sensible to surface-sliding in the negative direction.*



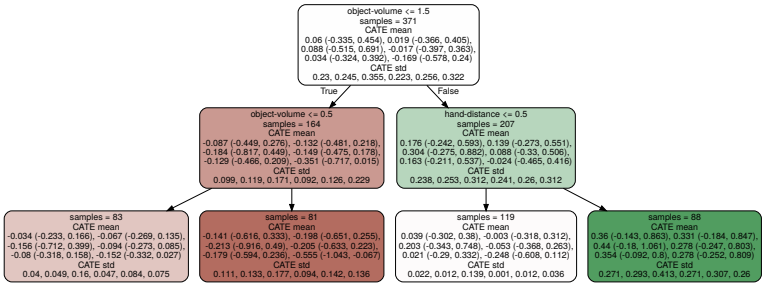
(c) *The effect of surface-sliding on hand-selection ($h3$) is sensible to surface-inside in the negative direction.*



(d) *The effect of surface-inside on hand-selection ($h4$) is sensible to hand-distance in the positive direction.*



(e) *The effect of object-category on hand-selection ($h6$) is sensible to surface-inside in the positive direction.*



(f) *The effect of surface-category on hand-selection ($h7$) is sensible to object-volume in the positive and negative directions.*

Figure 6.4: *The sensitivities for hand-selection in Person A. Each tree interprets the sensitivity of an effect reported in Table 6.2. Common across these trees is the evidence ds-1 and the estimator FDRL with a depth limit of 2.*

The color of a node in the tree indicates the direction in which a factor impacts a given effect. The color green indicates a positive change in the outcome, red a negative, and white a neutral. The root node in the tree of Figure 6.4a shows object-category in green, indicating the factor has a positive impact on the relationship between hand-distance and hand-selection (drives positive change in the expected outcome).

The tree structure is limited to depth 2 and constructed with the estimator imposing the least assumptions in the estimation process (i.e. FDRL).

The sensibilities for the effects reported in Table 6.2 presented in Figure 6.4 indicate the following[1]:

- A^{h_1}: The effect of hand-distance on hand-selection ($h1$) is sensible to object-category in the positive direction shown in Figure 6.4a.
- A^{h_2}: The effect of object-volume on hand-selection ($h2$) is sensible to surface-sliding in the negative direction shown in Figure 6.4a.
- A^{h_3}: The effect of surface-sliding on hand-selection ($h3$) is sensible to surface-inside in the negative direction shown in Figure 6.4c.
- A^{h_4} The effect of surface-inside on hand-selection ($h4$) is sensible to hand-distance in the positive direction shown in Figure 6.4d.
- A^{h_6} The effect of object-category on hand-selection ($h6$) is sensible to surface-inside in the positive direction shown in Figure 6.4e.
- A^{h_7} The effect of surface-category on hand-selection ($h7$) is sensible to object-volume in the positive and negative direction shown in Figure 6.4f.

This concludes reporting the change on hand-selection using dataset ds-1 (Person A). The next section addresses hand-distance.

The Effects on Hand-Distance (Group 2)

The next group of estimates describe the effects on hand-distance instead of hand-selection covered in the previous section. These are hypotheses $h8$ to $h13$ in the causal graph in Figure 6.3.

The estimates on dataset ds-1 for group-2 shown in Table 6.3 report:

[1]Note, only the highest-ranked factor in the tree is considered in summary (i.e., the root node).

Type: What	Samples	Estimator	Est. Effect	Conf. Inter.	Sensitivity
A^{h_8}	384	LDML	-0.22	(-0.45,0.01)	Figure 6.5a
		LDRL	-0.25	(-0.52,0.02)	
		CFDML	-0.28	(-0.53,-0.02)	
		FDRL	-0.36	(-0.67,-0.04)	
A^{h_9}	384	LDML	-0.41	(-1.56,0.73)	Figure 6.5b
		LDRL	-0.40	(-0.51,-0.30)	
		CFDML	-0.83	(-1.25,-0.41)	
		FDRL	-0.40	(-0.53,-0.27)	
$A^{h_{10}}$	384	LDML	0.33	(0.01,0.66)	Figure 6.5c
		LDRL	0.31	(0.20,0.42)	
		CFDML	0.46	(0.24,0.68)	
		FDRL	0.30	(0.16,0.44)	
$A^{h_{12}}$	384	LDML	1.08	(-2.28,4.43)	Figure 6.5d
		LDRL	-0.19	(-0.50,0.11)	
		CFDML	-0.61	(-3.35,2.12)	
		FDRL	-0.22	(-0.53,0.09)	
$A^{h_{13}}$	371	LDML	-1.34	(-3.77,1.10)	Figure 6.5e
		LDRL	0.55	(0.32,0.78)	
		CFDML	-0.22	(-1.14,0.70)	
		FDRL	0.53	(0.26,0.80)	

Table 6.3: *The type-what inferences for factors potentially affecting hand-distance in Person A. The framework estimates these effects with inferences of type* what *targeting hypotheses* $h8$ *to* $h13$ *on observational evidence ds-1 (384 samples).*

- A^{h_8}: Increasing the variable object-volume (OV) from 0.002 to 0.001 (size of a spoon to of a glass) causes an increase of −0.36 in the expected value of the outcome hand-distance (D) over the data distribution ds-1 with changes ranging over interval (−0.67, −0.04) using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.22 with values ranging over (−0.45, 0.01), LDRL −0.25 (−0.52, 0.02), and FDML −0.28 (−0.53, −0.02). The causal direction of object-volume (OV) on hand-distance (D) is neutral (4 estimators) with tendency negative.

- A^{h_9}: Modulating the variable surface-sliding (SS) from false to true (e.g. sink-surface to drawer) causes a change of −0.40 in the expected value of the outcome hand-distance (D) over the data distribution ds-1 with changes ranging over interval (−0.53, −0.27) using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.41 with values ranging over (−1.56, 0.73), LDRL −0.40 (−0.51, −0.30), and FDML −0.83 (−1.25, −0.41). The causal direction of surface-sliding (D) on hand-distance (D) is negative (3 estimators) and neutral (1 estimator).

- $A^{h_{10}}$: Modulating the variable surface-inside (SC) from false to true (e.g. table to fridge) causes a change of 0.30 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval (0.16, 0.44) using estimator FDRL. Correspondingly, estimator LDML reports a change of 0.33 with values ranging over (0.01, 0.66), LDRL 0.31 (0.20, 0.42), and FDML 0.46 (0.24, 0.68). The causal direction of surface-inside (SC) on hand-distance (D) is positive (4 estimators).

- $A^{h_{12}}$: Modulating the variable object-category (O) from milk to bowl causes a change of −0.22 in the expected value of the outcome hand-distance (D) over the data distribution ds-1 with changes ranging over interval (−0.53, 0.09) using estimator FDRL. Correspondingly, estimator LDML reports a change of 1.08 with values ranging over (−2.28, 4.43), LDRL −0.19 (−0.50, 0.11), and FDML −0.61 (−3.35, 2.12). The causal direction of object-category (O)) on hand-distance (D) is neutral (4 estimators) with tendency negative.

- $A^{h_{13}}$: Modulating the variable surface-category (S) from table to fridge causes a change of 0.53 in the expected value of the outcome hand-distance (D) over the data distribution ds-1 with changes ranging over interval (0.26, 0.80) using estimator FDRL. Correspondingly, estimator LDML reports a change of −1.34 with values ranging over (−3.77, 1.10), LDRL 0.55 (0.32, 0.78), and FDML −0.22 (−1.14, 0.70). The causal direction of surface-category (S) on hand-distance (D) is positive (2 estimator), neutral (2 estimators) with tendency negative.

The previous paragraphs report the effects a set of factors has on hand-distance. In what follows, the concern is explaining effects by clarifying the underlying sensibilities using the tree structures shown in Figure 6.5.

The sensibilities for the effects reported in Table 6.3 presented in Figure 6.5 indicate the following:

- A^{h_8}: The effect of hand-distance on hand-selection (h_8) is sensible to surface-sliding in the negative direction shown in Figure 6.5a.
- A^{h_9}: The effect of object-size on hand-selection (h_9) is sensible to surface-inside in the negative direction shown in Figure 6.5b.
- $A^{h_{10}}$: The effect of surface-sliding on hand-selection (h_{10}) is sensible to surface-sliding in the positive direction shown in Figure 6.5c.
- $A^{h_{12}}$: The effect of hand-dominance on hand-selection (h_{12}) is sensible to hand-selection in the negative direction shown in Figure 6.5d.
- $A^{h_{13}}$: The effect of object-category on hand-selection (h_{13}) is sensible to hand-selection in the positive and negative directions shown in Figure 6.5e.

This concludes the report for the second group of estimates that relate to hand-distance in person A.

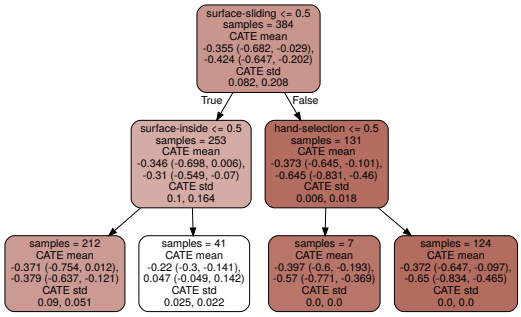
6.2.2 What and Why on Hand-Behavior (Person B)

Following the previous section, the effects on hand-selection are reported first (group-1), followed by those on hand-distance (group-2). The key difference with the previous section is that inferences are drawn on the behavior of Person B, i.e., dataset ds-2. This collection of data contains fewer examples than dataset ds-1.

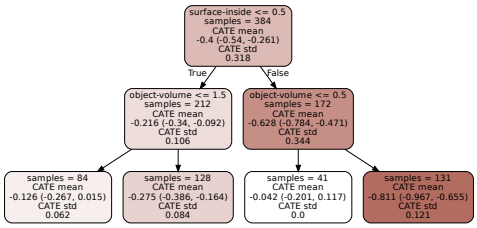
The Effects on Hand-Selection (Group 1)

Table 6.4 reports the estimates on hand-selection for person B, and Figure 6.4 presents the corresponding sensibilities of the effects.

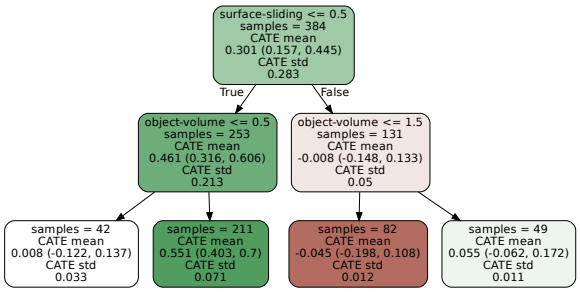
The estimates on dataset ds-2 for group-1 shown in Table 6.4 report:



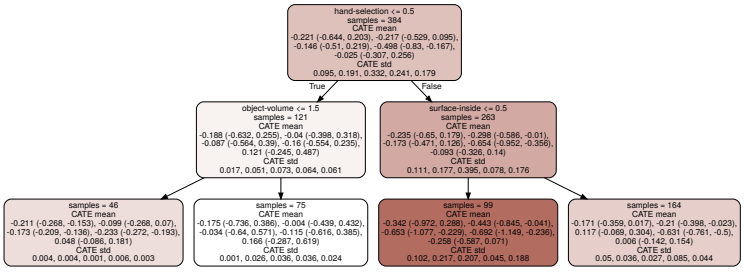
(a) *The effect of object-volume on hand-distance ($h8$) is sensible to surface-sliding in the negative direction.*



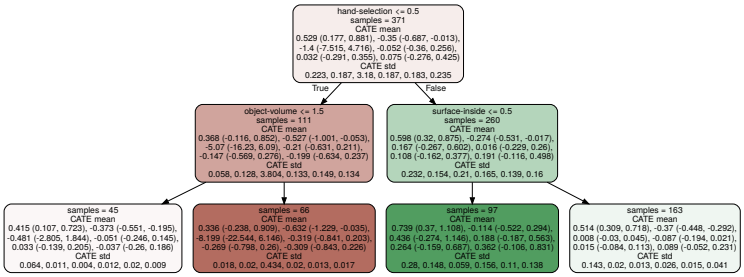
(b) *The effect of surface-sliding on hand-distance ($h9$) is sensible to surface-inside in the negative direction.*



(c) *The effect of surface-inside on hand-distance ($h10$) is sensible to surface-sliding in the positive direction.*



(d) *The effect of object-category on hand-distance ($h12$) is sensible to hand-selection in the negative direction.*



(e) *The effect of surface-category on hand-distance ($h13$) is sensible to hand-selection in the positive and negative direction.*

Figure 6.5: *The sensitivities for hand-distance in Person A. Each tree interprets the sensitivity of an effect reported in Table 6.3. Common across these trees is the evidence ds-1 and the estimator FDRL with a depth limit of 2.*

Type: What	Samples	Estimator	Effect	Conf. Int.	Sensitivity
$h_1 : (A \rightarrow H)$	174	LDML	-0.17	(-0.54,0.21)	Figure 6.6a
		LDRL	-0.16	(-0.51,0.20)	
		CFDML	-0.14	(-0.38,0.09)	
		FDRL	-0.17	(-0.48,0.14)	
$h_2 : (A \rightarrow H)$	166	LDML	-0.09	(-0.43,0.24)	Figure 6.6b
		LDRL	-0.16	(-0.49,0.17)	
		CFDML	-0.12	(-0.43,0.18)	
		FDRL	-0.10	(-0.39,0.20)	
$h_3 : (A \rightarrow H)$	174	LDML	0.22	(-0.58,1.01)	Figure 6.6c
		LDRL	0.06	(-0.10,0.22)	
		CFDML	-0.11	(-0.87,0.66)	
		FDRL	0.04	(-0.13,0.22)	
$h_4 : (A \rightarrow H)$	174	LDML	0.33	(-0.45,1.12)	Figure 6.6d
		LDRL	0.06	(-0.11,0.22)	
		CFDML	0.01	(-0.74,0.76)	
		FDRL	0.08	(-0.09,0.26)	
$h_6 : (A \rightarrow H)$	166	LDML	-2.22	(-6.81,2.38)	Figure 6.6e
		LDRL	-0.07	(-0.23,0.08)	
		CFDML	0.23	(-3.83,4.30)	
		FDRL	-0.12	(-0.28,0.05)	
$h_7 : (A \rightarrow H)$	158	LDML	-1.14	(-6.03,3.76)	Figure 6.6f
		LDRL	-0.14	(-0.41,0.14)	
		CFDML	-0.70	(-3.53,2.13)	
		FDRL	-0.00	(-0.18,0.17)	

Table 6.4: *The type-what inferences for factors potentially affecting hand-selection in Person B. The framework provide these effects for hypotheses $h1$ to $h7$ on observational evidence ds-2 (174 samples).*

- A^{h_1}: Increasing the variable hand-distance (D) from close to far causes an increase of −0.17 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval (−0.48, 0.14) using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.17 with values ranging over $(-0.54, 0.21)$, LDRL −0.16 $(-0.51, 0.20)$, and FDML −0.14 $(-0.38, 0.09)$. The causal direction of hand-distance (D) on hand-selection (H) is neutral (4 estimators) with tendency negative.

- A^{h_2}: Increasing the variable object-volume (OV) from 0.003 to 0.007 (size of a spoon to of a glass) causes an increase of −0.10 in the expected value of the outcome hand-selection (H) over the data distribution ds-1 with changes ranging over interval $(-0.39, 0.20)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.09 with values ranging over (−0.43, 0.24), LDRL −0.16 $(-0.49, 0.17)$, and FDML −0.12 $(-0.43, 0.18)$. The causal direction of object-volume (OV) on hand-selection (S) is neutral (4 estimators) with tendency negative.

- A^{h_3}: Modulating the variable surface-sliding (SS) from false to true (e.g. sink-surface to drawer) causes a change of 0.04 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval $(-0.13, 0.22)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of 0.22 with values ranging over $(-0.58, 1.01)$, LDRL 0.06 $(-0.10, 0.22)$, and FDML −0.11 $(-0.87, 0.66)$. The causal direction of surface-sliding (SS) on hand-selection (H) is neutral (4 estimators) with tendency positive.

- A^{h_4}: Modulating the variable surface-inside (SC) from false to true (e.g. table to drawer) causes a change of 0.08 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval $(-0.09, 0.26)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of 0.33 with values ranging over $(-0.45, 1.12)$, LDRL 0.06 $(-0.11, 0.22)$, and FDML 0.01 $(-0.74, 0.76)$. The causal direction of surface-inside (SC) on hand-selection (H) is neutral (4 estimators) with tendency positive.

- A^{h_6}: Modulating the variable object-category (O) from bowl to cereal-carton causes a change of −0.12 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval (−0.28, 0.05) using estimator FDRL. Correspondingly, estimator LDML reports a change of −2.22 with values ranging over $(-6.81, 2.38)$, LDRL −0.07 $(0.23, 0.08)$, and FDML 0.23 $(-3.83, 4.30)$. The causal direction of object-category (O) on hand-selection (H) is neutral (4 estimators) with tendency negative.

- A^{h_7}: Modulating the variable surface-category (S) from table to drawer causes a change of -0.00 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval $(-0.18, 0.17)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of -1.14 with values ranging over $(-6.03, 3.76)$, LDRL -0.14 $(-0.41, 0.14)$, and FDML -0.70 $(-3.53, 2.13)$. The causal direction of surface-category (S) on hand-selection (H) is neutral (4 estimators) with a tendency negative.

Figure 6.6 shows the sensitivities for the effects reported in Table 6.4. Each effect has a corresponding tree interpreting its sensibility concerning factors of the study. The root node of the tree hints at the primary driver of a given effect. The color of the nodes indicates the direction in which the factor drives change in the outcome, green for positive changes in the outcome, red for negative, and white for neutral.

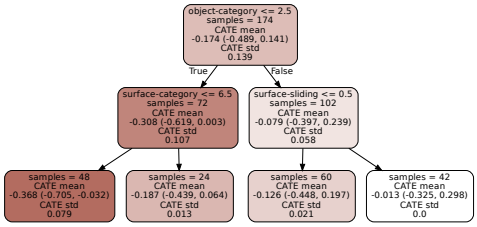
The sensibilities for the effects reported in Table 6.5 presented in Figure 6.6 indicate:

- A^{h_1}: The effect of hand-distance on hand-selection (h_1) is sensible to object-category in the negative direction shown in Figure 6.6a.
- A^{h_2}: The effect of object-size on hand-selection (h_2) is sensible to hand-distance in the positive direction shown in Figure 6.6b.
- A^{h_3}: The effect of surface-sliding on hand-selection (h_3) is sensible to object-volume in the positive direction shown in Figure 6.6c.
- A^{h_4}: The effect of surface-inside on hand-selection (h_4) is sensible to object-volume in the positive direction shown in Figure 6.6d.
- A^{h_6}: The effect of object-category on hand-selection (h_6) is sensible to hand-distance in the negative direction shown in Figure 6.6e.
- A^{h_7}: The effect of surface-category on hand-selection (h_7) is sensible to object-volume in the negative direction shown in Figure 6.6f.

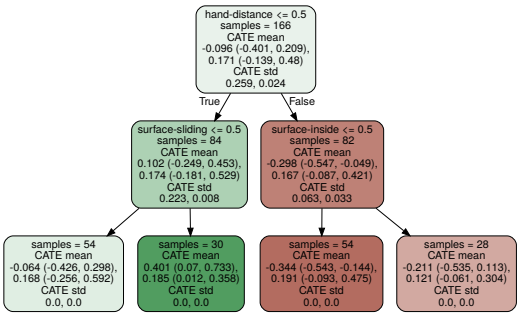
This concludes by reporting the effects on hand-selection for person B involving dataset ds-2. The next section concerns change in hand-distance.

The Effects on Hand-Distance (Group 2)

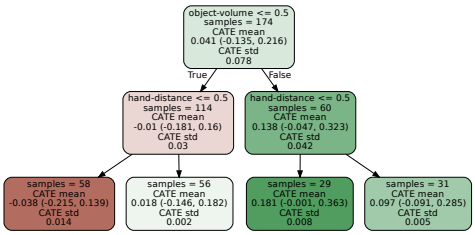
The effects covered in this section involve those potentially influencing hand-distance in Person B. Hypotheses $h8$ to $h13$ are covered.



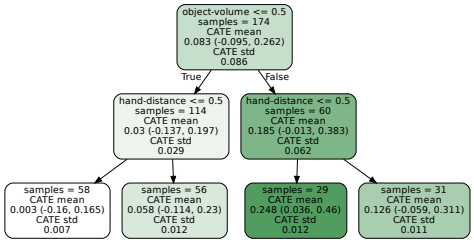
(a) *The effect of hand-distance on hand-selection (h1) is sensible to object-category in the negative direction.*



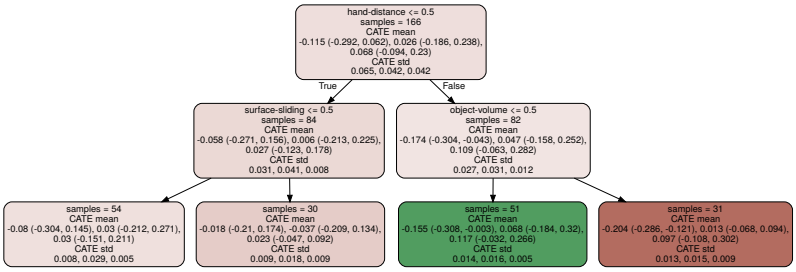
(b) *The effect of object-volume on hand-selection (h2) is sensible to hand-distance in the positive direction.*



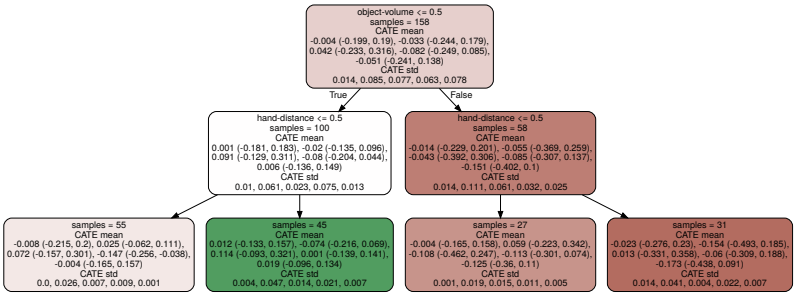
(c) *The effect of surface-sliding on hand-selection (h3) is sensible to object-volume in the positive direction.*



(d) *The effect of surface-inside on hand-selection (h4) is sensible to object-volume in the positive direction.*



(e) *The effect of object-category on hand-selection (h6) is sensible to hand-distance in the negative direction.*



(f) *The effect of surface-category on hand-selection (h7) is sensible to object-volume in the negative direction.*

Figure 6.6: *The sensitivities for hand-selection in Person B. Each tree interprets the sensitivity of an effect reported in Table 6.2). Common across these trees is the evidence ds-2 and the estimator FDRL with a depth limit of 2.*

Type: What	Samples	Estimator	Est. Effect	Conf. Int.	Sensitivity
A^{h_8}	166	LDML	-0.23	(-0.56,0.11)	Figure 6.7a
		LDRL	1.38	(-1.75,4.50)	
		CFDML	-0.13	(-0.44,0.18)	
		FDRL	-0.48	(-1.30,0.34)	
A^{h_9}	174	LDML	-0.48	(-1.25,0.29)	Figure 6.7b
		LDRL	0.06	(-0.10,0.22)	
		CFDML	0.06	(-0.74,0.85)	
		FDRL	0.06	(-0.10,0.22)	
$A^{h_{10}}$	174	LDML	-0.49	(-1.27,0.29)	Figure 6.7c
		LDRL	0.07	(-0.09,0.23)	
		CFDML	0.04	(-0.82,0.90)	
		FDRL	0.07	(-0.08,0.22)	
$A^{h_{12}}$	166	LDML	-0.98	(-4.77,2.81)	Figure 6.7d
		LDRL	-0.11	(-0.27,0.05)	
		CFDML	0.06	(-3.41,3.53)	
		FDRL	-0.14	(-0.30,0.03)	
$A^{h_{13}}$	158	LDML	-1.99	(-7.66,3.68)	Figure 6.7e
		LDRL	-0.02	(-0.27,0.24)	
		CFDML	0.16	(-1.78,2.11)	
		FDRL	-0.01	(-0.23,0.21)	

Table 6.5: *The type-what inferences on factors potentially affecting hand-distance in Person B. The framework estimates these effects for hypotheses* $h8$ *to* $h13$ *on observational evidence ds-2 (174 samples).*

The estimates derived with dataset ds-2 for group-2 shown in Table 6.5 indicate:

- A^{h_8}: Increasing the variable object-volume (OV) from 0.003 to 0.010 (e.g. size of a cereal-carton to tray) causes an increase of −0.48 in the expected value of the outcome hand-distance (D) over the data distribution ds-2 with changes ranging over interval $(-1.30, 0.34)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.23 with values ranging over (−0.56, 0.11), LDRL 1.38 $(-1.75, 4.50)$, and FDML −0.13 $(-0.44, 0.18)$. The causal direction of object-volume (OV) on hand-distance (D) is neutral (4 estimators) with tendency negative.

- A^{h_9}: Modulating the variable surface-sliding (SS) from false to true (e.g. table to drawer) causes a change of 0.06 in the expected value of the outcome hand-distance (D) over the data distribution ds-2 with changes ranging over interval $(-0.10, 0.22)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.48 with values ranging over $(-1.25, 0.29)$, LDRL 0.06 (−0.10, 0.22), and FDML 0.06 $(-0.74, 0.85)$. The causal direction of surface-sliding (D) on hand-distance (D) is neutral (4 estimators) with tendency positive.

- $A^{h_{10}}$: Modulating the variable surface-inside (SC) from false to true (e.g. table to drawer) causes a change of 0.07 in the expected value of the outcome hand-selection (H) over the data distribution ds-2 with changes ranging over interval $(-0.08, 0.22)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.49 with values ranging over $(-1.27, 0.29)$, LDRL 0.07 (−0.09, 0.23), and FDML 0.04 $(-0.82, 0.90)$. The causal direction of surface-inside (SC) on hand-distance (D) is neutral (4 estimators) with tendency positive.

- $A^{h_{12}}$: Modulating the variable object-category (O) from bowl to cereal-carton causes a change of −0.14 in the expected value of the outcome hand-distance (D) over the data distribution ds-2 with changes ranging over interval (−0.30, 0.03) using estimator FDRL. Correspondingly, estimator LDML reports a change of −0.98 with values ranging over $(-4.77, 2.81)$, LDRL −0.11 $(-0.27, 0.05)$, and FDML 0.06 $(-3.41, 3.53)$. The causal direction of object-category (O)) on hand-distance (D) is neutral (4 estimators) with tendency negative.

- $A^{h_{13}}$: Modulating the variable surface-category (S) from table to drawer causes a change of −0.01 in the expected value of the outcome hand-distance (D) over the data distribution ds-2 with changes ranging over interval $(-0.23, 0.21)$ using estimator FDRL. Correspondingly, estimator LDML reports a change of −1.99 with values ranging over $(-7.66, 3.68)$, LDRL −0.02 $(-0.27, 0.24)$, and

FDML 0.16 $(-1.78, 2.11)$. The causal direction of surface-category (S) on hand-distance (D) is neutral (4 estimators) with a tendency negative.

The previous paragraphs describe the effects potentially driving hand-distance. Next, the sensibilities of these estimates are interpreted using a tree structure shown in Figure 6.7.

The sensibilities for the effects reported in Table 6.5 are presented in Figure 6.7 and indicate:

- A^{h_8}: The effect of hand-distance on hand-selection (h_8) is sensible to hand-selection in the negative direction shown in Figure 6.7a.
- A^{h_9}: The effect of object-size on hand-selection (h_9) is sensible to object-volume in the positive direction shown in Figure 6.7b.
- $A^{h_{10}}$: The effect of surface-sliding on hand-selection (h_{10}) is sensible to object-volume in the positive direction shown in Figure 6.7c.
- $A^{h_{12}}$: The effect of hand-dominance on hand-selection (h_{12}) is sensible to hand-selection in the negative direction shown in Figure 6.7d.
- $A^{h_{13}}$: The effect of object-category on hand-selection (h_{13}) is sensible to hand-selection in the negative direction shown in Figure 6.7e.

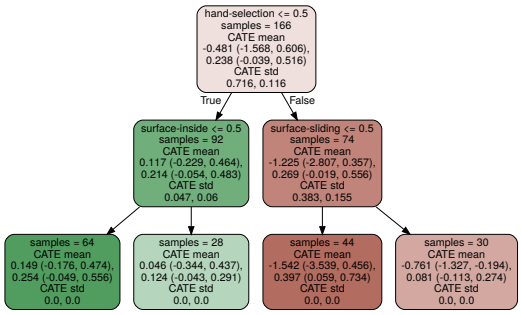
This concludes the report for the second group of estimates describing hand-distance in person B.

6.3 Discussion

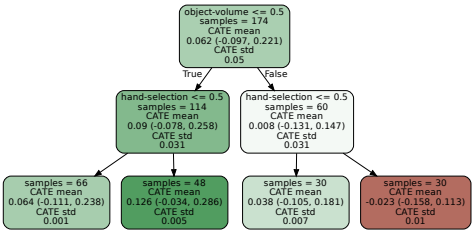
Overall, the framework discovered three causal relationships where confidence intervals do not include the null. Two of them are positive (A^{h_4}, $A^{h_{10}}$) and one negative (A^{h_9}). The interpretation of these causal discoveries is the following:

- A^{h_4}: The causal relationship of surface-inside (SC) on hand-selection (H) is positive for person's A hand-behavior. In particular, modulating the surface-inside (SC) from false to true (from a sink-area to a fridge) causes a positive change in hand-selection (H) over the data distribution ds-1 (person A).

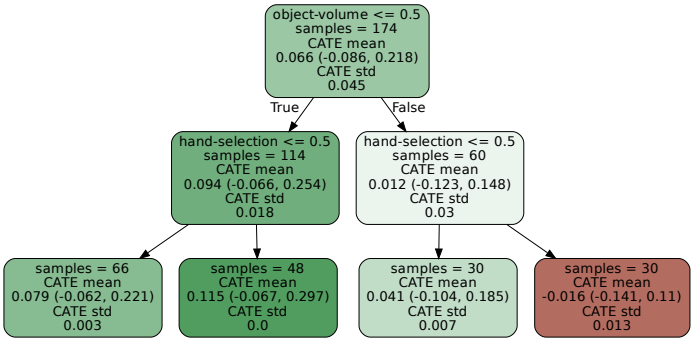
  Modulating from a sink-area (surface-inside is false) to a fridge (surface-inside is true) means that grasping actions in unconstrained spaces are contrasted to



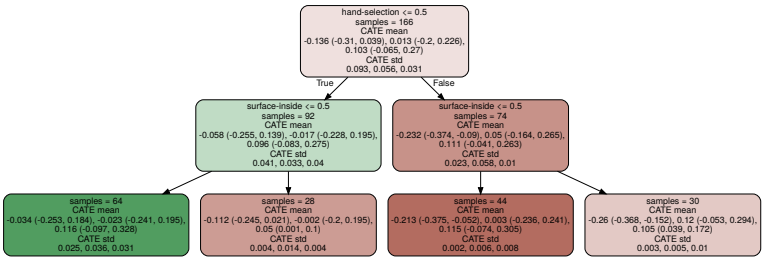
(a) *The effect of object-volume on hand-distance ($h8$) is sensible to hand-selection in the negative direction.*



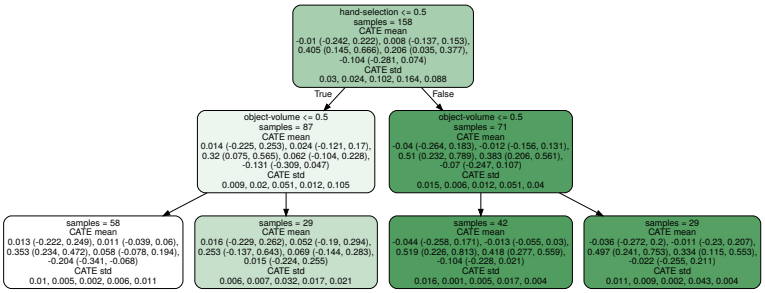
(b) *The effect of surface-sliding on hand-distance ($h9$) is sensible to object-volume in the positive direction.*



(c) *The effect of surface-inside on hand-distance ($h10$) is sensible to object-volume in the positive direction.*



(d) *The effect of object-category on hand-distance ($h12$) is sensible to hand-selection in the negative direction.*



(e) *The effect of surface-category on hand-distance ($h13$) is sensible to hand-selection in the positive direction.*

Figure 6.7: *The sensitivities for hand-distance in Person B. Each tree interprets the sensitivity of an effect reported in Table 6.3). The evidence ds-2 and the estimator FDRL with depth limit 2 underlie the construction of these trees.*

those in constrained spaces. For instance, the sink-area is an unconstrained space whereas the fridge is constrained. A positive effect on hand-selection means that grasping objects in constrained places causes person A to favor the right hand over the left. In other words, person's A hand behavior involves the dominant hand to reach into difficult places such as, for instance, reaching into the refrigerator. This causal effect agrees with intuition.

- A^{h_9}: The causal relationship of surface-sliding (SS) on hand-distance (D) is negative. In particular, modulating the variable surface-sliding (SS) from false to true (e.g. from a table to a drawer) causes a negative change in hand-distance (D) over the data distribution ds-1.

  Modulating from a table to a drawer means that grasping objects placed on non-sliding surfaces are contrasted to those that slide. For instance, the table is an non-sliding surface (surface-sliding is false) whereas the drawer is one that slides (surface-sliding is true). A negative effect on hand-distance means that grasping objects on surfaces that slide causes person A to grasp targets closer to the body than farther away. In other words, person's A hand behavior grasps objects close to the body when objects are located on surfaces that slide such as, for instance, a drawer containing silverware. At first glance this effect seemed to defy intuition, however, the effect makes sense given that pulling a drawer (or other furniture involving sliding surfaces) brings the objects closer. This causal effect is meaningful.

- $A^{h_{10}}$: The causal direction of surface-inside (SC) on hand-distance (D) is positive. In particular, modulating surface-inside (SC) from false to true (e.g. from a table to a fridge) causes a positive change in hand-selection (H) over the data distribution ds-1.

  Modulating the context from unconstrained spaces (such as a table) to constrained spaces (such as a refrigerator) resulted in a positive effect. A positive effect on hand-distance means that grasping objects in constrained places causes person A to target objects farther away from the body than closer. In other words, person's A hand behavior indicates greater arm extension when interacting with objects that are inside of containers such as, for instance, reaching deep into the refrigerator. This causal effect also agrees with intuition.

6.4 Concluding Remarks

This chapter assesses the framework's utility by uncovering hand behavior under the same conditions investigated in the previous chapter. Two collections of evidence recording uncontrolled hand behavior are introduced in Section 6.1.2, the hypotheses space supported by the framework is described in Section 6.1.3, and the formulation of causal questions targeting hand behavior are given in Section 6.2. Last, the analysis of hand behavior for persons A and B are presented in Section 6.2.1 and Section 6.2.2, respectively.

Unlike the null effects derived from the validation set in the previous chapter, the results of this chapter show that when the framework targets observational data (i.e., not experimentally influenced), the effects can differ from the null. The results show positive, neutral, and negative effects across two subjects' hand behavior. This result confirms the framework's capability to derive effects in all possible directions.

However, more importantly, the framework provided some confidence intervals which do not include the null value. Meaning the framework can estimate consistently, such that the presence of a causal effect is detectable. Consequently, the possibility of the framework being overly insensible, hinted at in the previous chapter, is herewith discarded.

The framework uncovered three causal relationships in hand behavior (discussed in Section 6.3). First, grasping objects in constrained spaces favor the dominant hand (h4). Second, targeting objects on sliding surfaces causes near-to-body grasps (h9). Third, grasping in constrained locations causes grabs farther away from the body (h10).

The framework's capability to detect causal effects (confidence intervals not including the zero) indicates that inferences from only hundreds of samples deem practical; otherwise, the confidence intervals would all include the null value, and a low-powered study would have been announced. This result confirms the approach can detect non-null effects from only a few samples.

Point estimates often agree across estimators; when not, a single estimator mostly disagrees with the rest. The estimator most frequently in disagreement is FDML, followed by LDML, LDRL, and FDRL.

The confidence intervals show the estimator imposing linearity on the treatment provides the widest intervals (LDML), especially when treatment is categorical confidence intervals are wide. Assuming the strongest operational assumption, linearity in treatment and heterogeneity seems inadequate because estimator FDML also

struggles to handle these cases. Therefore, imposing either assumption (linearity in treatment-and-heterogeneity and linearity in treatment) seems inadequate for capturing hand behavior in table setting activities (i.e., answer the thirteen causal questions). The narrowest confidence intervals belong to estimators assuming only linearity on heterogeneity (case of LDRL) and those making no assumption (case of FDRL).

Last, the sensibility analysis of the reported effects showed that the composition of change can be interpreted. Explaining effects using tree-structures enables the derivation of explicit rules. The tree structures employed to interpret effects are limited in depth to keep interpretation manageable, but robots could benefit from deeper analysis to explain intricate effects even further.

Chapter seven

Conclusion

The following concludes the dissertation by outlining the research findings concerning the aim and objectives and discussing the value and contribution. Then, the limitations of the study and opportunities for future research follow.

This dissertation challenges the potential of a formalism that fits existing theory on human indirect learning to enable robot analysis of human interaction in natural settings. We set three objectives to realize our aim, namely, formalize observational learning and a setting to challenge it (A), then verify the feasibility (B), and evaluate the utility (C).

The contributions aiming for these objectives are the following. We formalized observational learning in accord with the Social Cognitive Theory and established a framework (A1). We challenged this formalism to draw inferences under conditions people handle naturally, but machines do not (A2). We verified this formalism addresses the challenge with an experimental study (B1) evaluating for validity (B2) and stability (B3). We uncovered behavior by answering (C2) and explaining (C3) causal questions (C1).

To formalize observational learning (A), we first identify a formalism. To identify the formalism that fits the SCT we contrast this theory to the machines on two dimensions, processing and setting, considering four machine processing approaches and ten learning settings (in Chapter 2 Figure 2.5 shows the mismatch of human observational setting among the machine and Figure 2.2 exhibits the missing capabilities in processing approaches). To formalize observational learning following the Social Cognitive Theory we propose structural-causal models with non-parametric estimators. First, because this formalism enables the core inference capabilities

identified in humans as given by SCT (causal inference, predict, adapt, imagine), and unlike inference that rely on correlations (data-driven, theory-driven, model-driven, instance-driven), this formalism supports a truth-preserving mechanism (identification and adjustment) to counter confounding bias and thus enable causal inferences. Second, it consists of a representation to encode models, a calculus to process observations, and a language to articulate inferences, which taken together, support the acquisition process known as modeling in the SCT, where the retention process encodes models with causal graphs, the production process manipulates models via the calculus, the attentional and motivational process queries with the inference language. The two arguments given above justify the formalism is in line with SCT.

Further, to study the formalism, we instantiate a framework (A1) encompassing three components, structural-causal models, estimators, and a model. In particular, we build the framework upon the implementation of structural-causal models from the [SK+19] library, the estimators from the [Res19] library, and design a model where we identify relationships in hand behavior using findings from the literature and encode them into a causal graph (Figure 4.4). We justify the selection of four nonparametric estimators (Figure 4.7) that offer valid confidence intervals but differ in the operational assumption of heterogeneity. The three components that specify the framework (defined in Sections 4.1.2.1, 4.1.2.2, 4.1.2.3), and the identification described before, realize contribution A1.

Furthermore, to investigate the potential of the formalism (A2), we challenge the framework with a scenario (defined in Section 4.2). The scenario follows our characterization derived in Section 2.3 (vicarious setting). We set the scenario to address natural activities (a table setting), leave the data as unmodified as possible (any balance of classes, keep accidental actions in the set), consider only a few examples (hundreds instead of thousands), and leverage partial confounding using a model capturing only part of relevant aspects. The acquisition pipeline comprises recording activities in virtual environments (Unreal Engine, HTC Vive), providing robots with access to virtual activites (Figure 4.11), and extracting events via queries (Figure 4.12). The dataset recording natural interactions in virtual table settings (ds-1) have been used to study human-object arrangement in kitchen scenarios published in [PWYB18]. Taken together, the framework (A1) and the scenario that challenges the proper functioning of the framework (A2) realize objective A.

To evaluate the feasibility of the approach (B), we establish a validation set (B1), we verify the validity (B2) and stability of the inferences (B3). First, we verify validity with an experimental study (B1) targeting a subject's hand behavior (experimental procedure 5.1.1, setup 5.1.2, data 5.1.3). Second, we verify multiple hypotheses (Sections 5.2.2) with results indicating the framework derives valid effects for non-

trivial relationships in hand-behavior limited to a few samples and a model not fully specifying the relevant factors (B2). Third, we verify the framework's robustness by perturbing its stability using three strategies and assess the sensibility of inferences to remain stable. The results indicate the framework shows resilience for the three strategies refuting stability (B3). The experimental study (Section 5.1), the verification for validity (Table 5.3), and robustness (Table 5.5) accomplish objective B.

To evaluate the approach's utility (C), we formulate questions (C1) and use the framework to uncover behavior (C2, C3). First, we formulate thirteen causal questions (Section 6.2), target two individuals' hand interactions performing table settings (Person A and B), and employ the framework to derive causal relationships (Section 6.2.1 and Section 6.2.2, respectively). Despite the low number of examples and partial confounding, the results indicate the framework detects three relationships which are informative (C2). We interpret the composition of effects using tree structures to find their explanation (C3). The thirteen questions addressing hand behavior, their corresponding effects (Table 6.2 and Table 6.3 on Person A, Table 6.4 and Table 6.5 on Person B), and the explanation trees (Figure 6.4 and Figure 6.5 on Person A, Figure 6.6 and Figure 6.6 on Person B) contribute with objective C. Our study reporting the effects discovered on uncontrolled hand manipulation in table setting activities is published in [WSB22].

We believe that examining the potential of human-like observational learning for robots, comprising the identification of a formalism in accord to human indirect learning, establishing a framework, challenge its potential with a scenario, evaluating the feasibility, and assessing the utility, described in this work, show signs that the formalism can operate out of its operational design when we have broken a foundational assumption (a model capturing only part of the relevant aspects) and confront atypical conditions (only hundreds of samples, unbalanced classes) on a demanding task (infer a person's hand behavior) in the service robotics domain. To the best of our knowledge, there is no work available yet formalizing the four core capabilities given in SCT with structural-causal models and probing the formalism's potential with a natural scenario including a human being, uncontrolled behavior, a small number of examples, a model capturing only part of relevant aspects, and prioritizing the focus on inferences. The selected formalism covers the core inference capabilities serving as a platform for the acquisition process (modeling in SCT), a gap in the body of theory mentioned in Chapter 1. In principle, robots equipped with the formalism could potentially analyze human interaction in natural conditions (infer causal relationships), anticipate others' behavior (predict), accommodate past observations to different target conditions (adapt), and envision scenarios not seen before (imagine). In practice, however, one would wonder about the

boundaries of its applicability for two reasons. First, the formalism guarantees bias-free inference when models capture the relevant aspects (but what if not?). Second, the estimators typically operate with samples in thousands (but what if not?). These two circumstances resemble conditions children manage to overcome and therefore enforced. With our contributed proof of principle, beyond the theory, practitioners can know that inferences do not necessarily break under such conditions. Our evaluations indicate that even when models do not capture the relevant aspects, and data availability is in the order of hundreds, the formalism has a chance to operate as usual.

The proof of principle encompassing feasibility and utility renders viable. However, viability does not necessarily transfer to other settings where varied outcomes are expected. Therefore, replicating our experimental study is necessary. An avenue worth exploring (but left untested) is probing the potential of counterfactual reasoning (imagine). Nonetheless, with our contributed proof of concept and value, we anticipate research on robot applications by testing the boundaries of a promising inference formalism. We found that the formalism complies with indirect human learning, can be instantiated into a framework, and the proof of principle indicates the feasibility and utility remain uncompromised despite the challenge, which encourages further research.

Chapter eight

Appendix

8.1 Data Collections

Table 8.1 summarizes three data collections, one of them is experimentally influenced while the other two are not. The first is regarded as experimental evidence or validation data (Ds-v), whereas the other two are referred as observational evidence (Ds-1, Ds-2).

User interaction			Ds-1	Ds-2	Ds-v
Object (O):	**OV (m^3)**		**384**	**174**	**137**
Silverware	8.2×10^{-5}		83	70	29
Glass	1.22×10^{-3}		83	-	23
Milk	1.40×10^{-3}		18	8	24
Juice	1.93×10^{-3}		88	-	21
Bowl	2.40×10^{-3}		19	36	16
Cereal	6.30×10^{-3}		93	28	24
Tray	9.95×10^{-3}		-	32	-
Surface (S):	**SC**	**SS**	**384**	**174**	**137**
DiningTable	F	F	12	18	-
FrdgArea	T	F	-	1	1
FrdgDrBtmShlf	T	F	5	2	22
FrdgGlassShlf	T	T	41	-	21
IslndArea	F	F	-	8	21
IslndDrwBtmLft	T	T	-	16	-
LabFloor	F	F	5	-	2
OvenArea	F	F	-	2	-
OvenDrwRight	T	T	41	16	22
SinkArea	F	F	11	24	1
SnkDrwLftBtm	T	T	-	-	1
SnkDrwLftMid	T	T	41	-	20
SnkDrwLftTop	T	T	41	42	25
Tray	F	F	184	42	-
Hand-selection (H):	**DO**		**384**	**174**	**137**
Left (H_0)	F		121	96	64
Right (H_1)	T		263	78	73
Hand-distance (D):			**384**	**174**	**137**
Close (HD_F)	0		192	87	69
Far (HD_T)	1		192	86	68

Table 8.1: *The three collections of evidence (ds-v, ds-1, ds-2) and corresponding frequencies of occurrences for different levels of features. Abbreviation* OV *refers to the volume of an object.* SC *corresponds to surfaces inside containers such as the refrigerator's shelf.* SS *refers to surfaces that slide, such as drawers.* DO *refers to hand dominance.* SnkDrwLftMid *refers to the middle drawer out of three stacked one over another spatially located left of the sink.* FrdgDrBtmShlf *refers to the bottom shelf on the fridge door. Last, levels T and F abbreviate true and false.*

8.2 Validity and Stability

This section summarizes estimates obtained with the framework when providing experimentally controlled data (ds-v). The first Table 8.2 summarizes the validation and stability described in Chapter 5. The second Table 8.3 shows the effects corresponding to hypotheses h_7 to h_13 for which no particular effect can be expected.

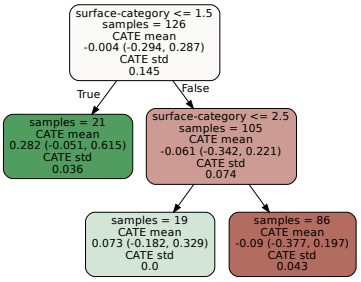
The following tree structures present the sensibilities of effects corresponding to each of tables presented before. The trees are constructed using the estimator LDRL with max depth 2.

Table 8.2: *The Validity and Stability for the group of hypotheses (H) h_1 to h_7 on experimental evidence ds-v at the alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

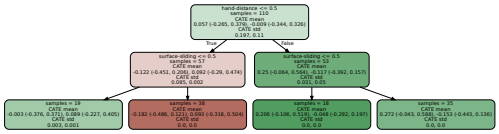
H	Samples	Estimator	Effects	Conf. Int.	RCC	DOR	DSR
1	126	LDML	0.12	(-0.38,0.63)	-0.06	0.01	0.11
		LDRL	-0.10	(-0.72,0.52)	0.03	-0.03	-0.02
		CFDML	0.04	(-0.28,0.36)	0.01	0.00	0.03
		FDRL	-0.00	(-0.29,0.29)	0.02	0.01	0.04
2	110	LDML	0.00	(-0.64,0.64)	-0.57	-0.07	-0.44
		LDRL	0.09	(-0.32,0.49)	-0.26	-0.01	-0.26
		CFDML	0.11	(-0.29,0.50)	-0.22	0.03	-0.23
		FDRL	0.06	(-0.26,0.38)	-0.15	0.01	-0.12
3	126	LDML	-0.08	(-0.52,0.36)	-0.03	0.00	-0.05
		LDRL	-0.04	(-0.37,0.29)	-0.21	0.00	-0.18
		CFDML	-0.11	(-0.48,0.26)	0.01	-0.03	-0.00
		FDRL	-0.07	(-0.35,0.20)	-0.02	0.00	-0.03
4	126	LDML	0.12	(-1.74,1.98)	-0.12	0.00	-0.11
		LDRL	-0.09	(-0.36,0.19)	-0.14	-0.06	-0.21
		CFDML	0.26	(-0.57,1.09)	-0.12	-0.02	-0.15
		FDRL	-0.07	(-0.27,0.12)	-0.11	0.01	-0.11
6	122	LDML	-0.66	(-3.15,1.84)	-1.22	0.30	-0.86
		LDRL	0.05	(-0.16,0.26)	0.06	-0.01	0.08
		CFDML	-0.67	(-3.97,2.63)	0.91	0.04	0.74
		FDRL	0.09	(-0.07,0.25)	0.06	-0.02	0.10
7	121	LDML	3.18	(-2.52,5.88)	-1.15	0.30	-2.51
		LDRL	0.04	(-0.23,0.31)	-0.08	-0.01	0.08
		CFDML	0.97	(-1.46,3.39)	0.96	-0.15	0.61
		FDRL	0.02	(-0.26,0.30)	-0.08	0.03	-0.04

Table 8.3: *The Validity and Stability for the group of hypotheses (H) h_7 to h_{13} on experimental evidence ds-v at an alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

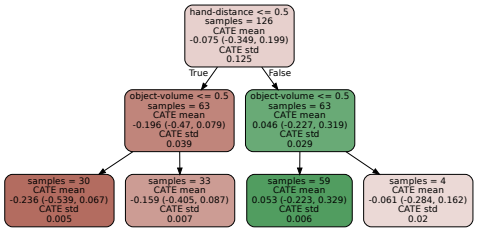
H	Samples	Estimator	Effect	Conf. Int.	RCC	DOR	DSR
8	110	LDML	-0.03	(-0.38,0.32)	-0.48	0.01	-0.42
		LDRL	-0.15	(-0.42,0.12)	-0.50	-0.00	-0.48
		CFDML	-0.15	(-0.45,0.15)	-0.44	0.03	-0.44
		FDRL	-0.09	(-0.42,0.24)	-0.52	0.00	-0.53
9	126	LDML	-0.37	(-0.72,-0.03)	-0.52	0.01	-0.43
		LDRL	-0.25	(-0.49,-0.00)	-0.24	0.01	-0.24
		CFDML	-0.36	(-0.68,-0.04)	-0.35	0.01	-0.32
		FDRL	-0.27	(-0.48,-0.06)	-0.21	-0.00	-0.22
10	126	LDML	-0.96	(-2.55,0.63)	0.04	-0.01	0.05
		LDRL	0.41	(0.19,0.63)	0.04	-0.01	0.05
		CFDML	0.69	(0.30,1.09)	0.05	0.12	-0.10
		FDRL	0.41	(0.19,0.63)	-0.03	-0.01	-0.03
12	122	LDML	-2.41	(-5.27,0.45)	-3.43	0.36	-1.64
		LDRL	0.03	(-0.16,0.22)	-0.02	-0.02	0.02
		CFDML	-0.83	(-2.46,0.79)	-0.96	0.17	-0.63
		FDRL	0.02	(-0.10,0.13)	-0.03	0.02	0.00
13	121	LDML	-1.01	(-5.79,3.16)	0.10	-0.40	-0.15
		LDRL	0.29	(0.03,0.56)	0.28	-0.02	0.26
		CFDML	0.01	(-0.53,0.55)	0.16	-0.09	0.09
		FDRL	0.14	(-0.13,0.42)	0.23	-0.02	0.18



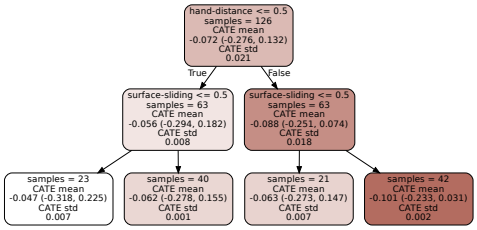
(a) *The effect of hand-distance on hand-selection ($h1$) is sensible to surface-category in the positive and negative direction.*



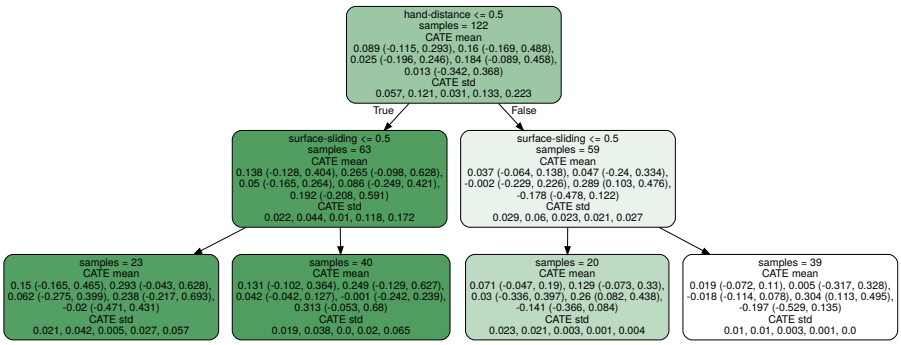
(b) *The effect of object-volume on hand-selection ($h2$) is sensible to hand-distance in the positive direction.*



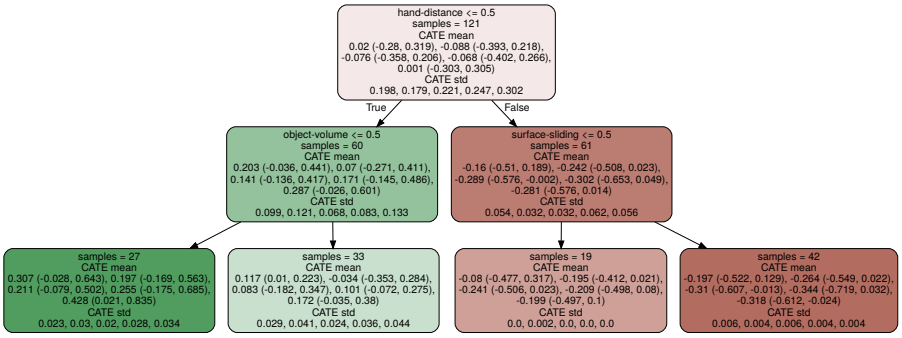
(c) *The effect of surface-sliding on hand-selection ($h3$) is sensible to hand-distance in the negative direction.*



(d) *The effect of surface-inside on hand-selection ($h4$) is sensible to hand-distance in the negative direction.*

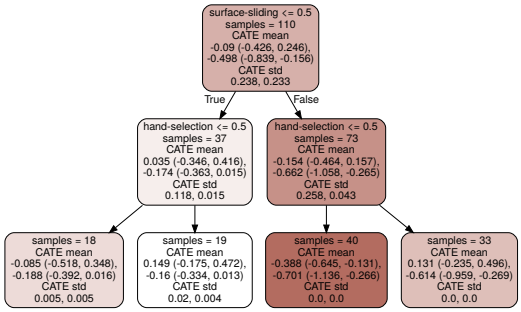


(e) *The effect of object-category on hand-selection ($h6$) is sensible to hand-distance in the positive direction.*

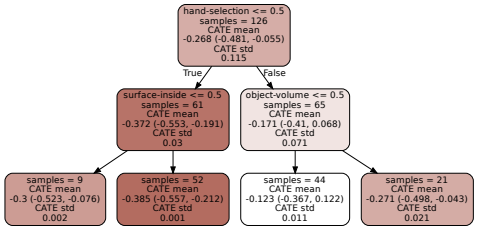


(f) *The effect of surface-category on hand-selection ($h7$) is sensible to hand-distance in the negative direction.*

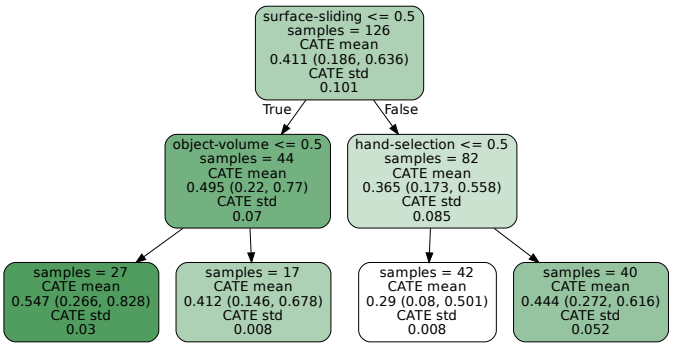
Figure 8.1: *The sensitivities for hypotheses group 1 ($h1 - h7$) on dataset ds-v.*



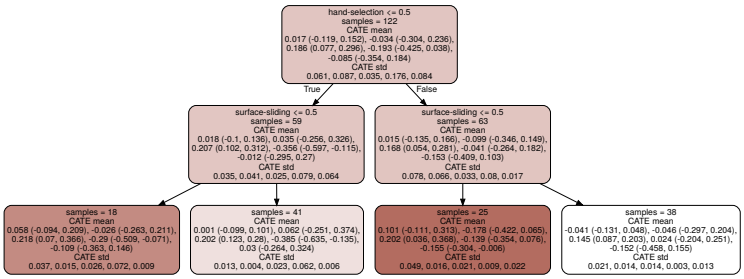
(a) *The effect of object-volume on hand-distance ($h8$) is sensible to surface-sliding in the negative direction.*



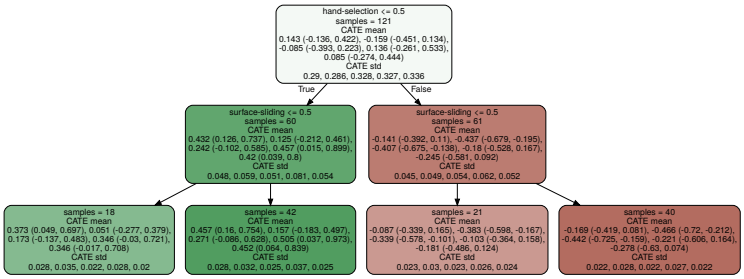
(b) *The effect of surface-sliding on hand-distance ($h9$) is sensible to hand-selection in the negative direction.*



(c) *The effect of surface-inside on hand-distance ($h10$) is sensible to surface-sliding in the positive direction.*



(d) *The effect of object-category on hand-distance ($h12$) is sensible to hand-selection in the negative direction.*



(e) *The effect of surface-category on hand-distance ($h13$) is sensible to hand-selection in the positive and negative direction.*

Figure 8.2: *The sensitivities for hypotheses group 2 ($h8$ – $h13$) on dataset ds-v.*

8.3 Utility for Person A

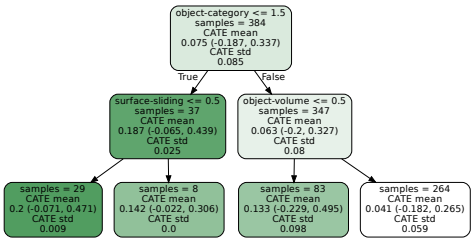
The effects reported in this section correspond to the hand behavior of the person A, h_1 to h_7 for hand-selection in Table 8.4 while h_8 to h_13 to hand-distance in Table 8.5.

Table 8.4: *The Validity and Stability for a group of hypotheses (H) h_1 to h_7 on experimental evidence ds-1 at the alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

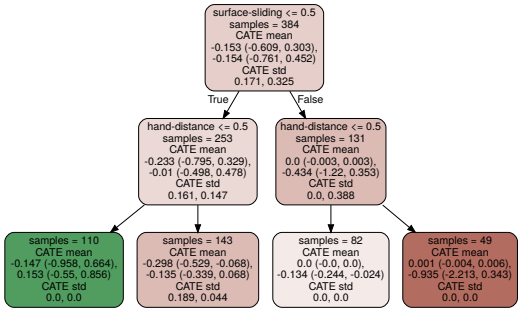
H	Samples	Estimator	Effects	Conf. Int.	RCC	DOR	DSR
1	384	LDML	0.02	(-0.29,0.33)	0.14	-0.01	0.10
	384	LDRL	0.07	(-0.15,0.29)	0.15	-0.01	0.15
	384	CFDML	0.09	(-0.05,0.23)	0.17	-0.00	0.17
	384	FDRL	0.08	(-0.15,0.30)	0.16	0.01	0.17
2	384	LDML	-0.08	(-0.30,0.13)	-0.20	-0.00	-0.19
	384	LDRL	-0.17	(-0.40,0.05)	-0.36	-0.02	-0.33
	384	CFDML	-0.20	(-0.41,0.01)	-0.41	-0.01	-0.41
	384	FDRL	-0.15	(-0.46,0.15)	-0.37	-0.00	-0.25
3	384	LDML	-0.09	(-0.57,0.39)	0.01	-0.01	0.01
	384	LDRL	-0.11	(-0.32,0.09)	0.28	0.06	0.14
	384	CFDML	-0.03	(-0.38,0.32)	0.09	-0.01	0.12
	384	FDRL	-0.07	(-0.20,0.06)	0.01	0.01	0.01
4	384	LDML	0.19	(0.01,0.38)	0.45	-0.00	0.46
	384	LDRL	0.52	(0.42,0.62)	0.35	-0.01	0.32
	384	CFDML	0.24	(0.04,0.44)	0.26	0.00	0.23
	384	FDRL	0.52	(0.40,0.64)	0.45	-0.00	0.45
6	384	LDML	1.64	(-1.07,3.80)	-0.33	0.26	-0.76
	384	LDRL	0.07	(-0.19,0.33)	0.12	-0.41	0.72
	384	CFDML	0.59	(-0.16,1.34)	0.70	0.06	0.67
	384	FDRL	0.02	(-0.25,0.28)	0.16	0.62	0.31
7	371	LDML	-4.50	(-6.51,-2.49)	0.07	0.01	-2.26
	371	LDRL	-0.12	(-0.75,0.51)	0.01	-0.71	0.55
	371	CFDML	-0.07	(-1.18,1.04)	-0.20	0.13	-0.26
	371	FDRL	0.06	(-0.24,0.36)	-0.03	0.30	0.14

Table 8.5: *The Validity and stability for a group of hypotheses h_8 to h_{13} on experimental evidence ds-1 at the alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

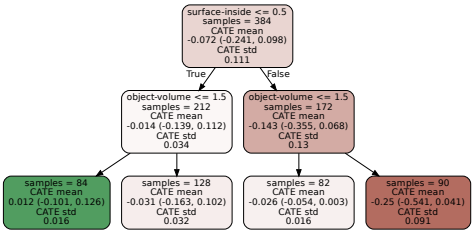
H	Samples	Estimator	Effects	Conf. Int.	RCC	DOR	DSR
8	384	LDML	-0.22	(-0.45,0.01)	-0.16	0.03	-0.15
	384	LDRL	-0.25	(-0.52,0.02)	-0.27	0.00	-0.26
	384	CFDML	-0.28	(-0.53,-0.02)	-0.27	-0.00	-0.27
	384	FDRL	-0.36	(-0.67,-0.04)	-0.27	-0.03	-0.26
9	384	LDML	-0.41	(-1.56,0.73)	-0.23	-0.08	-0.31
	384	LDRL	-0.40	(-0.51,-0.30)	-0.40	-0.01	-0.41
	384	CFDML	-0.83	(-1.25,-0.41)	-0.85	-0.04	-0.78
	384	FDRL	-0.40	(-0.53,-0.27)	-0.41	0.01	-0.41
10	384	LDML	0.33	(0.01,0.66)	0.18	-0.01	0.16
	384	LDRL	0.31	(0.20,0.42)	0.65	0.02	0.50
	384	CFDML	0.46	(0.24,0.68)	0.40	-0.01	0.34
	384	FDRL	0.30	(0.16,0.44)	0.19	-0.00	0.17
12	384	LDML	1.08	(-2.28,4.43)	2.58	-2.40	3.99
	384	LDRL	-0.19	(-0.50,0.11)	-0.26	-0.60	-0.69
	384	CFDML	-0.61	(-3.35,2.12)	-0.10	0.16	-0.12
	384	FDRL	-0.22	(-0.53,0.09)	-0.24	-0.64	-0.40
13	371	LDML	-1.34	(-3.77,1.10)	-0.01	0.30	-0.54
	371	LDRL	0.55	(0.32,0.78)	0.37	0.02	0.56
	371	CFDML	-0.22	(-1.14,0.70)	-0.04	-0.00	0.04
	371	FDRL	0.53	(0.26,0.80)	0.45	-0.02	0.43



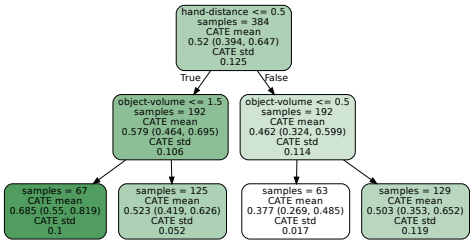
(a) *The effect of hand-distance on hand-selection ($h1$) is sensible to object-category in the positive direction.*



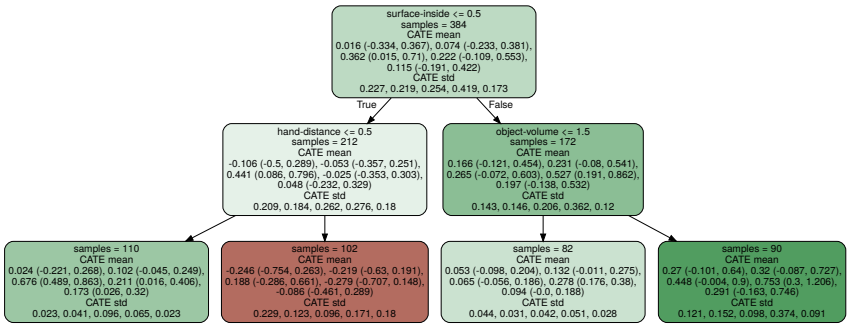
(b) *The effect of object-volume on hand-selection ($h2$) is sensible to surface-sliding in the negative direction.*



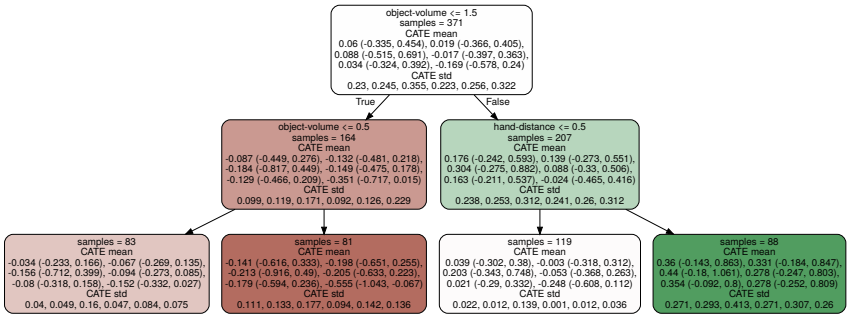
(c) *The effect of surface-sliding on hand-selection ($h3$) is sensible to surface-inside in the negative direction.*



(d) *The effect of surface-inside on hand-selection ($h4$) is sensible to hand-distance in the positive direction.*

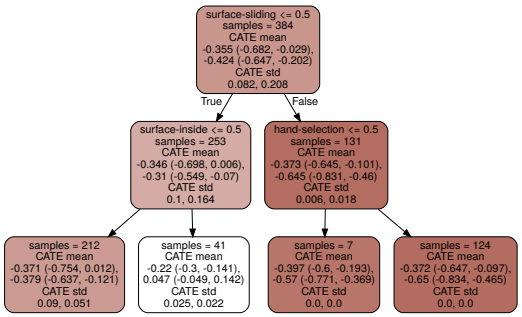


(e) *The effect of object-category on hand-selection ($h6$) is sensible to surface-inside in the positive direction.*

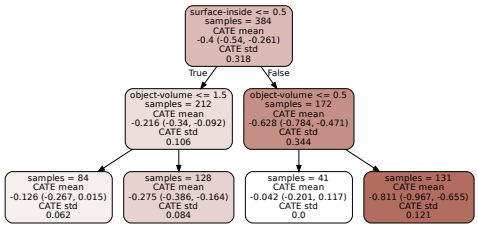


(f) *The effect of surface-category on hand-selection ($h7$) is sensible to object-volume in the positive and negative direction.*

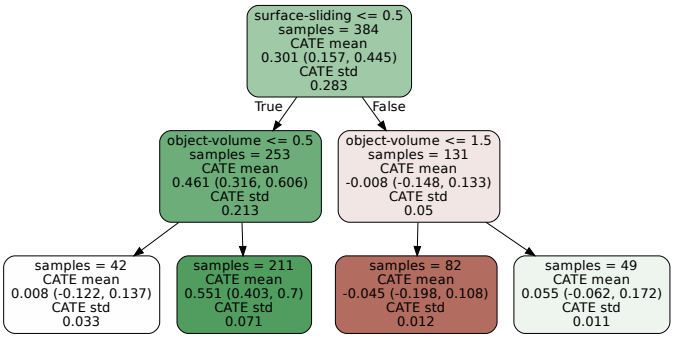
Figure 8.3: *The sensitivities for hypotheses group 1 ($h1 - h7$) on dataset ds-1.*



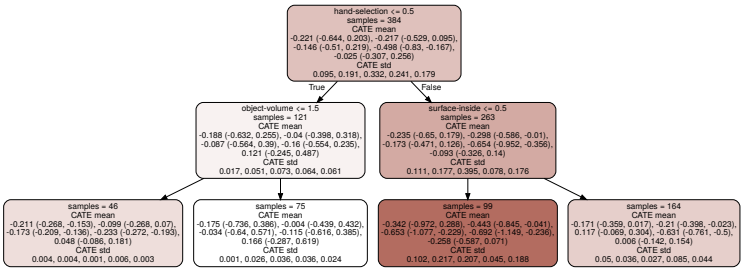
(a) *The effect of object-volume on hand-distance ($h8$) is sensible to surface-sliding in the negative direction.*



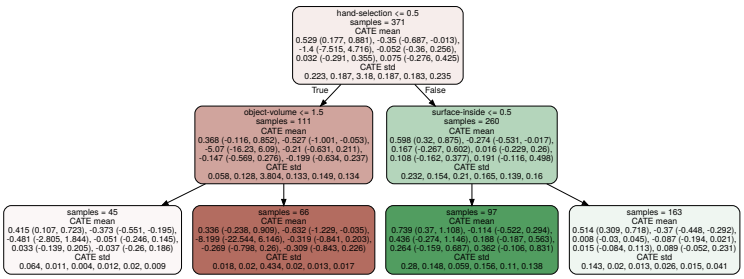
(b) *The effect of surface-sliding on hand-distance ($h9$) is sensible to surface-inside in the negative direction.*



(c) *The effect of surface-inside on hand-distance ($h10$) is sensible to surface-sliding in the positive direction.*



(d) *The effect of object-category on hand-distance ($h12$) is sensible to hnd-selection in the negative direction.*



(e) *The effect of surface-category on hand-distance ($h13$) is sensible to hand-selection in the positive and negative direction.*

Figure 8.4: *The sensitivities for hypotheses group 2 ($h8$ – $h13$) on dataset ds-1.*

8.4 Utility for Person B

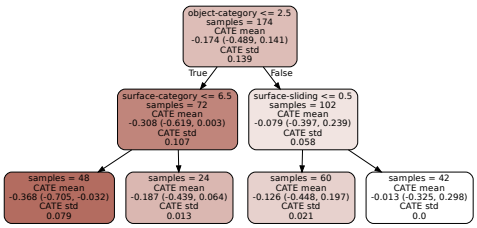
The effects reported in this section correspond to the hand behavior of the person A, h_1 to h_7 for hand-selection in Table 8.6 while h_8 to h_13 to hand-distance in Table 8.7.

Table 8.6: *The Validity and Stability for the group of hypotheses (H) h_1 to h_7 on experimental evidence ds-2 at the alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

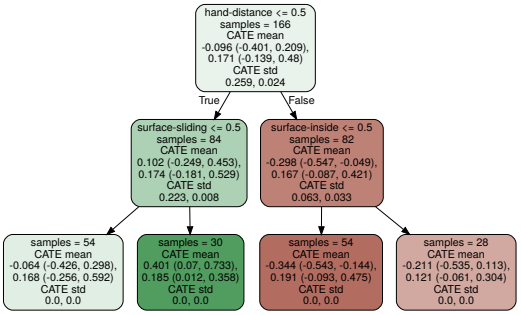
H	Samples	Estimator	Effect	Conf. Int.	RCC	DOR	DSR
1	174	LDML	-0.17	(-0.54,0.21)	-0.31	0.00	-0.30
	174	LDRL	-0.16	(-0.51,0.20)	-0.32	-0.02	-0.30
	174	CFDML	-0.14	(-0.38,0.09)	-0.29	0.03	-0.30
	174	FDRL	-0.17	(-0.48,0.14)	-0.30	0.01	-0.28
2	166	LDML	-0.09	(-0.43,0.24)	-0.18	-0.02	-0.19
	166	LDRL	-0.16	(-0.49,0.17)	-0.38	-0.04	-0.83
	166	CFDML	-0.12	(-0.43,0.18)	-0.26	0.00	-0.24
	166	FDRL	-0.10	(-0.39,0.20)	-0.22	0.10	-0.41
3	174	LDML	0.22	(-0.58,1.01)	0.06	0.00	0.06
	174	LDRL	0.06	(-0.10,0.22)	0.18	0.04	0.17
	174	CFDML	-0.11	(-0.87,0.66)	0.12	-0.00	0.09
	174	FDRL	0.04	(-0.13,0.22)	0.06	0.02	0.07
4	174	LDML	0.33	(-0.45,1.12)	0.06	0.01	0.06
	174	LDRL	0.06	(-0.11,0.22)	0.19	0.03	0.17
	174	CFDML	0.01	(-0.74,0.76)	0.13	0.00	0.10
	174	FDRL	0.08	(-0.09,0.26)	0.06	0.00	0.07
6	166	LDML	-2.22	(-6.81,2.38)	-1.09	0.22	-1.82
	166	LDRL	-0.07	(-0.23,0.08)	-0.04	-0.02	-0.04
	166	CFDML	0.23	(-3.83,4.30)	-0.05	-0.01	-0.02
	166	FDRL	-0.12	(-0.28,0.05)	-0.04	0.01	-0.05
7	158	LDML	-1.14	(-6.03,3.76)	17.99	-3.70	18.19
	158	LDRL	-0.14	(-0.41,0.14)	-0.18	-0.01	-0.16
	158	CFDML	-0.70	(-3.53,2.13)	-1.86	0.14	-1.41
	158	FDRL	-0.00	(-0.18,0.17)	-0.10	0.00	-0.08

Table 8.7: *The Validity and Stability for the group of hypotheses (H) h_8 to h_{13} on experimental evidence ds-2 at the alpha level of two sigmas. The last three columns correspond to Random Common Cause (RCC), Dummy Outcome Refuter (DOR), Subsets Validation (DSV).*

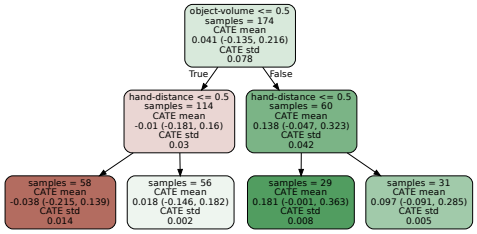
H	Samples	Estimator	Effect	Conf. Int.	RCC	DOR	DSR
8	166	LDML	-0.23	(-0.56,0.11)	-0.11	0.02	-0.12
	166	LDRL	1.38	(-1.75,4.50)	-0.11	0.62	-4.05
	166	CFDML	-0.13	(-0.44,0.18)	-0.13	0.01	-0.13
	166	FDRL	-0.48	(-1.30,0.34)	-0.13	-0.06	-0.15
9	174	LDML	-0.48	(-1.25,0.29)	-0.48	-0.00	-0.48
	174	LDRL	0.06	(-0.10,0.22)	0.05	-0.00	0.06
	174	CFDML	0.06	(-0.74,0.85)	0.08	-0.01	0.07
	174	FDRL	0.06	(-0.10,0.22)	0.05	-0.01	0.06
10	174	LDML	-0.49	(-1.27,0.29)	0.04	-0.02	0.07
	174	LDRL	0.07	(-0.09,0.23)	0.02	-0.07	-0.08
	174	CFDML	0.04	(-0.82,0.90)	0.00	0.04	-0.03
	174	FDRL	0.07	(-0.08,0.22)	0.04	0.01	0.05
12	166	SLDML	-0.98	(-4.77,2.81)	-0.20	0.62	-0.40
	166	LDRL	-0.11	(-0.27,0.05)	-0.10	-0.01	-0.11
	166	CFDML	0.06	(-3.41,3.53)	0.42	-0.54	-0.24
	166	FDRL	-0.14	(-0.30,0.03)	-0.12	-0.00	-0.12
13	158	SLDML	-1.99	(-7.66,3.68)	-7.30	-0.68	-10.64
	158	LDRL	-0.02	(-0.27,0.24)	0.01	0.03	0.01
	158	CFDML	0.16	(-1.78,2.11)	0.43	0.05	0.30
	158	FDRL	-0.01	(-0.23,0.21)	0.01	0.01	0.00



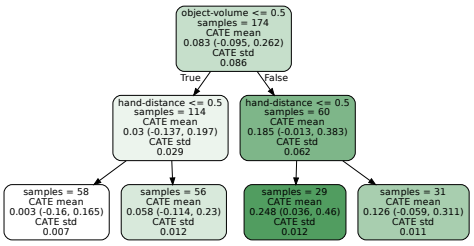
(a) *The effect of hand-distance on hand-selection ($h1$) is sensible to object-category in the negative direction.*



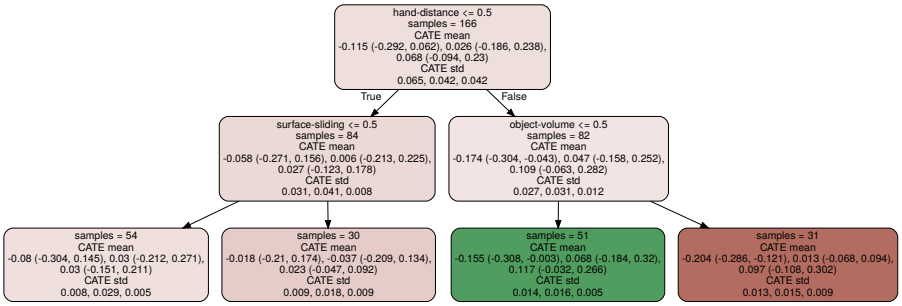
(b) *The effect of object-volume on hand-selection ($h2$) is sensible to hand-distance in the positive direction.*



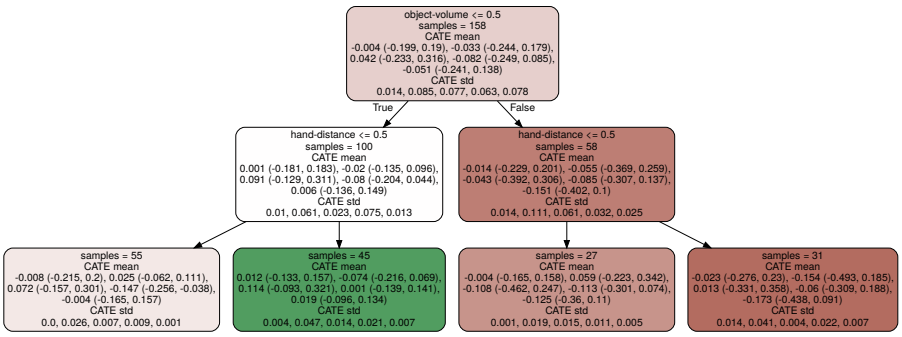
(c) *The effect of surface-sliding on hand-selection ($h3$) is sensible to object-volume in the positive direction.*



(d) *The effect of surface-inside on hand-selection ($h4$) is sensible to object-volume in the positive direction.*

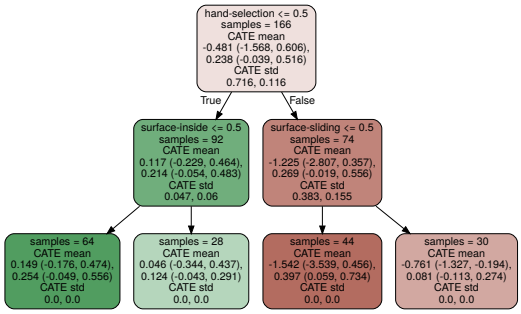


(e) *The effect of object-category on hand-selection ($h6$) is sensible to hand-distance in the negative direction.*

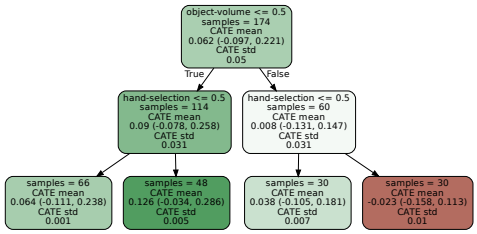


(f) *The effect of surface-category on hand-selection ($h7$) is sensible to object-volume in the negative direction.*

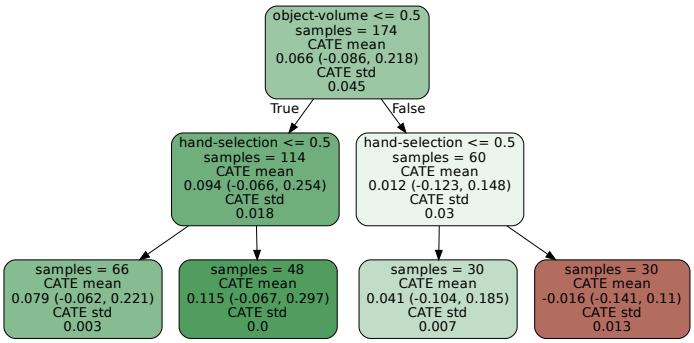
Figure 8.5: *The sensitivities for hypotheses group 1 ($h1$ – $h7$) on dataset ds-2.*



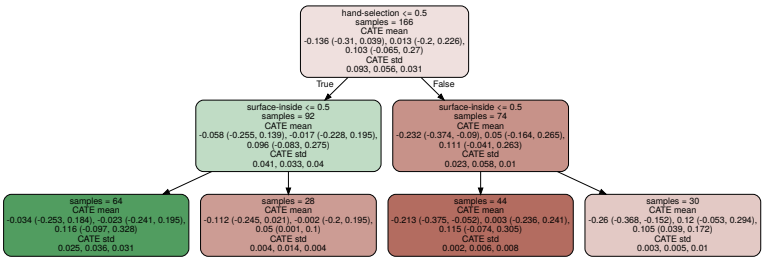
(a) *The effect of object-volume on hand-distance ($h8$) is sensible to hand-selection in the negative direction.*



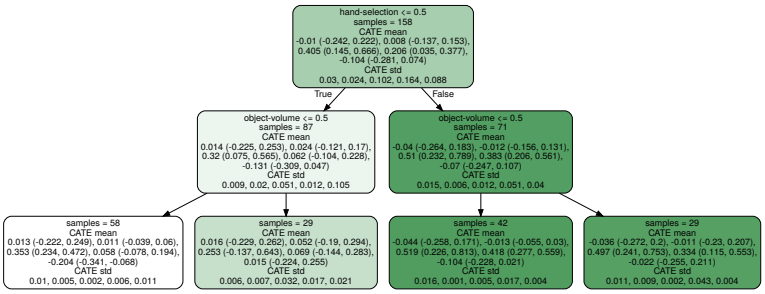
(b) *The effect of surface-sliding on hand-distance ($h9$) is sensible to object-volume in the positive direction.*



(c) *The effect of surface-inside on hand-distance ($h10$) is sensible to object-volume in the positive direction.*



(d) *The effect of object-category on hand-distance ($h12$) is sensible to hand-selection in the negative direction.*



(e) *The effect of surface-category on hand-distance ($h13$) is sensible to hand-selection in the positive direction.*

Figure 8.6: *The sensitivities for hypotheses group 2 ($h8$ – $h13$) on dataset ds-2.*

Acronyms

BLB Bootstrap-of-Little-Bags 89–91, 93, 112

DML Double Machine Learning 88–90, 93

DRL Doubly Robust Learners 88–90, 93

FDML Forest Double Machine Learning 89, 136, 138, 139, 143, 147, 148, 151, 152, 155

FDRL Forest Doubly Robust Learner 89, 138, 139, 141, 143, 147, 148, 151, 155, 156

LDML Linear Double Machine Learning 89, 136, 138, 139, 143, 147, 148, 151, 155

LDRL Linear Doubly Robust Learner 89, 136, 138, 139, 143, 147, 148, 151, 155, 156

SCT Social Cognitive Theory 2–5, 27, 30, 31, 36, 39, 41, 93, 94, 157–159

Bibliography

[AB07] Grigoris Antoniou and Antonis Bikakis. Dr-prolog: A system for defeasible reasoning with rules and ontologies on the semantic web. *IEEE Transactions on Knowledge and Data Engineering*, 19(2):233–245, 2007.

[AN04] Pieter Abbeel and Andrew Y. Ng. Apprenticeship learning via inverse reinforcement learning. In Carla E. Brodley, editor, *Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004), Banff, Alberta, Canada, July 4-8, 2004*, volume 69 of *ACM International Conference Proceeding Series*. ACM, 2004.

[Ang88] Dana Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.

[Ang04] Joshua D. Angrist. Treatment Effect Heterogeneity in Theory and Practice. *The Economic Journal*, 114(494):C52–C83, march 2004.

[AS18] Sarah Abraham and Liyang Sun. Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *SSRN Electronic Journal*, 2018.

[ATW19] Susan Athey, Julie Tibshirani, and Stefan Wager. Generalized Random Forests. *The Annals of Statistics*, 47(2):1148 – 1178, 2019.

[Ban65] Albert Bandura. Influence of models' reinforcement contingencies on the acquisition of imitative responses. *American Psychological Association*, 1(6):589–595, 6 1965.

[Ban77] A. Bandura. *Social learning theory*. Prentice-Hall, 1977.

[Ban86] A. Bandura. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall series in social learning theory. Prentice-Hall, 1986.

[Ban88] Albert Bandura. Organisational Applications of Social Cognitive Theory. *Australian Journal of Management*, 13(2):275–302, 12 1988.

[Ban89] Albert Bandura. Human agency in social cognitive theory. *American Psychologist*, 44(9):1175–1184, 1989.

[Ban95] Albert Bandura. *Self-Efficacy in Changing Societies*. Cambridge University Press, 4 1995.

[Ban01] Albert Bandura. Social cognitive theory of mass communication. *Media Psychology*, 3(3):265–299, 8 2001.

[BB71] Albert Bandura and Peter G. Barab. Conditions governing nonreinforced imitation. *Developmental Psychology*, 5(2):244–255, 1971.

[BBH+18] Michael Beetz, Daniel Beßler, Andrei Haidu, Mihai Pomarlan, Asil Kaan Bozcuoglu, and Georg Bartels. Knowrob 2.0 – a 2nd generation knowledge processing framework for cognition-enabled robotic agents. In *International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018.

[BCDS08] A. Billard, S. Calinon, R. Dillmann, and S. Schaal. Robot programming by demonstration. In B. Siciliano and O. Khatib, editors, *Handbook of Robotics*. Springer, 2008. In press.

[BCG06] Aude G. Billard, Sylvain Calinon, and Florent Guenter. Discriminative and adaptive imitation in uni-manual and bi-manual tasks. *Robotics and Autonomous Systems*, 54(5):370–384, 2006. The Social Mechanisms of Robot Programming from Demonstration.

[BG12] Aude Billard and Daniel Grollman. *Imitation Learning in Robots*, pages 1494–1496. Springer US, Boston, MA, 2012.

[BG13] A. Billard and D. Grollman. Robot learning by demonstration. *Scholarpedia*, 8(12):3824, 2013. revision 138061.

[BGM66] Albert Bandura, Joan E. Grusec, and Frances L. Menlove. Observational learning as a function of symbolization and incentive set. *Child Development*, 37(3):499–506, 1966.

[BHM+13] Armin Biere, Marijn Heule, Hans Van Maaren, Toby Walsch, C Clark Barrett, Roberto Sebastiani, Sanjit A. Seshia, and Cesare Tinelli. Chapter 12 Satisfiability Modulo Theories, 10 2013.

[BJ73] Albert Bandura and Robert W. Jeffrey. Role of symbolic coding and rehearsal processes in observational learning. *Journal of Personality and Social Psychology*, 26(1):122–130, 1973.

[BJB74] Albert Bandura, Robert Jeffery, and Daniel L. Bachicha. Analysis of memory codes and cumulative rehearsal in observational learning. *Journal of Research in Personality*, 7(4):295–305, 1974.

[BKF+18] Asil Kaan Bozcuoglu, Gayane Kazhoyan, Yuki Furuta, Simon Stelter, Michael Beetz, Kei Okada, and Masayuki Inaba. The exchange of knowledge using cloud robotics. *Robotics and Automation Letters*, 3(2):1072–1079, April 2018.

[BLOD18] Andrea Bajcsy, Dylan P. Losey, Marcia K. O'Malley, and Anca D. Dragan. Learning from physical human corrections, one feature at a time. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2018, Chicago, IL, USA, March 05-08, 2018*, pages 141–149, 2018.

[BLS83] Gary F. Bradshaw, Patrick W. Langley, and Herbert A. Simon. Studying Scientific Discovery by Computer Simulation. *Science*, 222(4627):971–975, 12 1983.

[BO09] Jennings Bryant and Mary Beth Oliver, editors. *Media Effects*. Routledge, 1 2009.

[Boz99] S. Bozinovski. Crossbar adaptive array: The first connectionist network that solved the delayed reinforcement learning problem. In *Artificial Neural Nets and Genetic Algorithms*, pages 320–325, Vienna, 1999. Springer Vienna.

[BR66] Albert Bandura and Theodore L. Rosenthal. Vicarious classical conditioning as a function of arousal level. *Journal of Personality and Social Psychology*, 3(1):54–62, 1966.

[BR05] Heejung Bang and James M. Robins. Doubly Robust Estimation in Missing Data and Causal Inference Models. *Biometrics*, 61(4):962–973, 6 2005.

[BRDB96] D. V. M. Bishop, V. A. Ross, M. S. Daniels, and P. Bright. The measurement of hand preference: A validation study comparing three groups of right-handers. *British Journal of Psychology*, 87(2):269–285, 5 1996.

[BSC19] Kalesha Bullard, Yannick Schroecker, and Sonia Chernova. Active learning within constrained environments through imitation of an expert questioner. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 2045–2052, 2019.

[CB82] Wayne R. Carroll and Albert Bandura. The role of visual monitoring in observational learning of action patterns. *Journal of Motor Behavior*, 14(2):153–167, 1982. PMID: 15155177.

[CB87] Wayne R. Carroll and Albert Bandura. Translating cognition into action. *Journal of Motor Behavior*, 19(3):385–398, 1987. PMID: 14988054.

[CB90] Wayne R. Carroll and Albert Bandura. Representational guidance of action production in observational learning: A causal analysis. *Journal of Motor Behavior*, 22(1):85–97, 1990. PMID: 15111282.

[CB07] S. Calinon and A. Billard. What is the teacher's role in robot programming by demonstration? - Toward benchmarks for improved learning. *Interaction Studies. Special Issue on Psychological Benchmarks in Human-Robot Interaction*, 8(3):441–464, 2007.

[CCD+17] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, and Whitney Newey. Double/debiased/neyman machine learning of treatment effects. *American Economic Review*, 107(5):261–65, May 2017.

[Chu40] Alonzo Church. A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5(2):56–68, 6 1940.

[CK04] Paul Cisek and John F. Kalaska. Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature*, 431(7011):993–996, #oct# 2004.

[CM95] Rui Camacho and Donald Michie. Behavioral cloning A correction. *AI Mag.*, 16(2):92, 1995.

[CM03] William F. Clocksin and Christopher S. Mellish. *Programming in Prolog*. Springer, Berlin, 5 edition, 2003.

[CT12] Maya Cakmak and Andrea Lockerd Thomaz. Designing robot learners that ask good questions. In Holly A. Yanco, Aaron Steinfeld, Vanessa Evers, and Odest Chadwicke Jenkins, editors, *International Conference on Human-Robot Interaction, HRI'12, Boston, MA, USA - March 05 - 08, 2012*, pages 17–24. ACM, 2012.

[CTFB05] Leda Cosmides, John Tooby, Laurence Fiddick, and Gregory A. Bryant. Detecting cheaters. *Trends in Cognitive Sciences*, 9(11):505–506, 11 2005.

[CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 9 1995.

[DB07] Yiannis Demiris and Aude Billard. Special issue on robot learning by observation, demonstration, and imitation. *IEEE Trans. Syst. Man Cybern. Part B*, 37(2):254–255, 2007.

[Dew16] John Dewey. *Democracy and Education*. Macmillan, London, 1916. Published online at `http://www.ilt.columbia.edu/academic/texts/dewey/d_e/contents.html`.

[DLLW20] Dun Deng, Fenrong Liu, Mingming Liu, and Dag Westerståhl, editors. *Monotonicity in Logic and Language - Second Tsinghua Interdisciplinary Workshop on Logic, Language and Meaning, TLLM 2020, Beijing, China, December 17-20, 2020, Proceedings*, volume 12564 of *Lecture Notes in Computer Science*. Springer, 2020.

[EKSX96] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[Eva05] Jonathan St. B. T. Evans. Deductive reasoning. In Keith J. Holyoak and Robert G. Morrison, editors, *The Cambridge Handbook of Thinking and Reasoning*, pages 169–184. Cambridge University Press, Cambridge, England, 2005.

[Eva12] Jonathan St. B. T. Evans. *Dual-Process Theories of Deductive Reasoning: Facts and Fallacies*. Oxford University Press, 3 2012.

[FCAL16] Chelsea Finn, Paul F. Christiano, Pieter Abbeel, and Sergey Levine. A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models. *CoRR*, abs/1611.03852, 2016.

[FLA16] Chelsea Finn, Sergey Levine, and Pieter Abbeel. Guided cost learning: Deep inverse optimal control via policy optimization. *CoRR*, abs/1603.00448, 2016.

[For65] E. W. Forgy. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. *Biometrics*, 21:768–769, 1965.

[FS19] Dylan J. Foster and Vasilis Syrgkanis. Orthogonal statistical learning. *CoRR*, abs/1901.09036, 2019.

[GARP03] J Grèzes, J.L Armony, J Rowe, and R.E Passingham. Activations related to "mirror" and "canonical" neurones in the human brain: an fmri study. *NeuroImage*, 18(4):928–937, 2003.

[Gay20] Gayane Kazhoyan and Alina Hawkin and Sebastian Koralewski and Andrei Haidu and Michael Beetz. Learning motion parameterizations of mobile pick and place actions from observing humans in virtual environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[GBC16] Ian J. Goodfellow, Yoshua Bengio, and Aaron C. Courville. *Deep Learning*. Adaptive computation and machine learning. MIT Press, 2016.

[GFS15] Claudia L. R. Gonzalez, Jason W. Flindall, and Kayla D. Stone. Hand preference across the lifespan: effects of end-goal, task nature, and object location. *Frontiers in Psychology*, 5, 1 2015.

[GH04] Carl Gabbard and Casi Rabb Helbig. What drives children's limb selection for reaching in hemispace? *Experimental Brain Research*, 156(3):325–332, 6 2004.

[Gop10] Alison Gopnik. How babies think. *Scientific American*, 303(1):76–81, 7 2010.

[GR00] Carl Gabbard and Casi Rabb. What determines choice of limb for unimanual reaching movements? *The Journal of General Psychology*, 127(2):178–184, 4 2000.

[GS14] Alejandro Javier García and Guillermo Ricardo Simari. Defeasible logic programming: Delp-servers, contextual queries, and explanations for answers. *Argument Comput.*, 5(1):63–88, 2014.

[GSNT19] Reymundo A. Gutierrez, Elaine Schaertl Short, Scott Niekum, and Andrea Lockerd Thomaz. Learning from corrective demonstrations. In *14th ACM/IEEE International Conference on Human-Robot Interaction, HRI 2019, Daegu, South Korea, March 11-14, 2019*, pages 712–714, 2019.

[Gug07] Guglielmi. A system of interaction and structure. *ACMTCL: ACM Transactions on Computational Logic*, 8, 2007.

[Hag09] M. E. Haggerty. Imitation in monkeys. *Journal of Comparative Neurology and Psychology*, 19(4):337–455, 7 1909.

[Hal90] Joseph Y. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46(3):311–350, 1990.

[HB19] Andrei Haidu and Michael Beetz. Automated models of human everyday activity based on game and virtual reality technology. In *International Conference on Robotics and Automation (ICRA)*, Montreal, Canada, 2019.

[HB21] Andrei Haidu and Michael Beetz. Automated acquisition of structured, semantic models of manipulation activities from human vr demonstration. In *IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, China, 2021. Finalist for Best Paper Award on Human-Robot Interaction.

[HE16] Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. In Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pages 4565–4573, 2016.

[Hey93] C.M. Heyes. Imitation, culture and cognition. *Animal Behaviour*, 46(5):999–1010, 11 1993.

[Hol85] Paul W. Holland. Statistics and causal inference. Technical Report 85-40, Educational Testing Service, Princeton, New Jersey, 1985. With discussions by Clark Glymour and CLive Granger.

[HRD+03] Farsin Hamzei, Michel Rijntjes, Christian Dettmers, Volkmar Glauche, Cornelius Weiller, and Christian Büchel. The human action recognition system and its relationship to broca's area: an fmri study. *NeuroImage*, 19(3):637–644, 2003.

[HS14] Bradley Hayes and Brian Scassellati. Discovering task constraints through observation and active learning. In *IROS*, pages 4442–4449. IEEE, 2014.

[HV06] Yimin Huang and Marco Valtorta. Identifiability in Causal Bayesian Networks: A Sound and Complete Algorithm. In *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA*, pages 1149–1154. AAAI Press, 2006.

[Iac09] Marco Iacoboni. Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, 60(1):653–670, 2009. PMID: 18793090.

[Imb15] Guido Imbens. *Causal inference for statistics, social, and biomedical sciences : an introduction*. Cambridge University Press, New York, NY, 2015.

[IMSG+05] Marco Iacoboni, Istvan Molnar-Szakacs, Vittorio Gallese, Giovanni Buccino, John C Mazziotta, and Giacomo Rizzolatti. Grasping the intentions of others with one's own mirror neuron system. *PLOS Biology*, 3(3):null, 02 2005.

[JB96] R. I. John and S. C. Bennett. Fuzzy sets and community transport. In J. F. Baldwin, editor, *Fuzzy Logic*, chapter 13. John Wiley and Sons Ltd., 1996.

[JL08] Philip Johnson-Laird. *How We Reason*. Oxford University Press, 10 2008.

[JLB02] P. N. Johnson-Laird and Ruth M. J. Byrne. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4):646–678, 2002.

[Jos91] J. Josephson. Abduction: Conceptual analysis of a fundamental pattern of inference. Technical Research Report 91-JJ-DRAFT, Laboratory for Artificial Intelligence Research, The Ohio State University, 1991.

[JY06] Kevin A. Janes and Michael B. Yaffe. Data-driven modelling of signal-transduction networks. *Nature Reviews Molecular Cell Biology*, 7(11):820–828, 11 2006.

[KA17] Martin Krzywinski and Naomi Altman. Classification and regression trees. *Nature Methods*, 14(8):757–758, 8 2017.

[Kah11] Daniel Kahneman. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, 2011.

[KF13] Matthew Klenk and Kenneth D. Forbus. Exploiting persistent mappings in cross-domain analogical learning of physical domains. *Artif. Intell.*, 195:398–417, 2013.

[KG06] Christian Keysers and Valeria Gazzola. Towards a unifying neural theory of social cognition. In S. Anders, G. Ende, M. Junghofer, J. Kissler, and D. Wildgruber, editors, *Understanding Emotions*, volume 156 of *Progress in Brain Research*, pages 379–401. Elsevier, 2006.

[KKT93] Antonis Kakas, Robert Kowalski, and Francesca Toni. Abductive logic programming. *Journal of Logic and Computation*, 6?(2?):719–770, 1993.

[KLA+20] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In *CVPR*, pages 8107–8116. Computer Vision Foundation / IEEE, 2020.

[KNCC11] Petar Kormushev, Dragomir N. Nenchev, Sylvain Calinon, and Darwin G. Caldwell. Upper-body kinesthetic teaching of a free-standing humanoid robot. In *ICRA*, pages 3970–3975. IEEE, 2011.

[Koc15] Ned Kock. How Likely is Simpson's Paradox in Path Models? *IJeC*, 11(1):1–7, 2015.

[KP10] Jens Kober and Jan Peters. Imitation and reinforcement learning. *IEEE Robotics Autom. Mag.*, 17(2):55–62, 2010.

[KSV+18] Sören R. Künzel, Bradly C. Stadie, Nikita Vemuri, Varsha Ramakrishnan, Jasjeet S. Sekhon, and Pieter Abbeel. Transfer learning for estimating causal effects using neural networks. *CoRR*, abs/1808.07804, 2018.

[LF06] Pascale Leconte and Jacqueline Fagard. Which factors affect hand selection in children's grasping in hemispace? combined effects of task demand and motor dominance. *Brain and Cognition*, 60(1):88–93, 2 2006.

[LGA+17] John E. Laird, Kevin A. Gluck, John R. Anderson, Kenneth D. Forbus, Odest Chadwicke Jenkins, Christian Lebiere, Dario D. Salvucci, Matthias Scheutz, Andrea Thomaz, J. Gregory Trafton, Robert E. Wray, Shiwali Mohan, and James R. Kirk. Interactive task learning. *IEEE Intell. Syst.*, 32(4):6–21, 2017.

[LRLAdO13] Angel Lago-Rodriguez, Virginia Lopez-Alonso, and Miguel Fernández del Olmo. Mirror neuron system and observational learning: Behavioral and neurophysiological evidence. *Behavioural Brain Research*, 248:104–113, 7 2013.

[LWZ12] Yanli Liu, Yourong Wang, and Jian Zhang. New machine learning algorithm: Random forest. In *Information Computing and Applications*, pages 246–252. Springer Berlin Heidelberg, 2012.

[MD00] Jun Morimoto and Kenji Doya. Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML, pages 623–630, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[MK01] A.M. Martinez and A.C. Kak. PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):228–233, 2001.

[MO14] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *CoRR*, abs/1411.1784, 2014.

[MQXE20] Hongyuan Mei, Guanghui Qin, Minjie Xu, and Jason Eisner. Neural datalog through time: Informed temporal modeling via logical specification. *CoRR*, abs/2006.16723, 2020.

[MRBR04] Carla M. Mamolo, Eric A. Roy, Pamela J. Bryden, and Linda E. Rohr. The effects of skill demands and object position on the distribution of preferred hand reaches. *Brain and Cognition*, 55(2):349–351, 7 2004.

[MSC+20] Clare F. Macrae, Ioana Sovago, Simon J. Cottrell, Peter T. A. Galek, Patrick McCabe, Elna Pidcock, Michael Platings, Greg P. Shields, Joanna S. Stevens, Matthew Towler, and Peter A. Wood. Mercury 4.0: from visualization to analysis, design and prediction. *Journal of Applied Crystallography*, 53(1):226–235, 2 2020.

[MWG12] Andrew N. Meltzoff, Anna Waismeyer, and Alison Gopnik. Learning about causes from people: Observational causal learning in 24-month-old infants. *Developmental Psychology*, 48(5):1215–1228, 2012.

[NOT05] Robert Nieuwenhuis, Albert Oliveras, and Cesare Tinelli. Abstract DPLL and abstract DPLL modulo theories. In F. Baader and A. Voronkov, editors, *Proceedings of the 11th International Conference on Logic for*

Programming, Artificial Intelligence and Reasoning (LPAR'04), Montevideo, Uruguay, volume 3452 of *Lecture Notes in Artificial Intelligence*, pages 36–50. Springer, 2005.

[NSS18] Denis Nekipelov, Vira Semenova, and Vasilis Syrgkanis. Regularized orthogonal machine learning for nonlinear semiparametric models, 2018.

[Nut03] Donald Nute. Defeasible logic. In *Lecture Notes in Computer Science*, pages 151–169. Springer Berlin Heidelberg, 2003.

[O'B09] David P. O'Brien. Human reasoning includes a mental logic. *Behavioral and Brain Sciences*, 32(1):96–97, 2 2009.

[OMG13] Santiago Ontañón, José L. Montaña, and Avelino J. Gonzalez. Towards a Unified Framework for Learning from Observation, 9 2013.

[OSW19] Miruna Oprescu, Vasilis Syrgkanis, and Zhiwei Steven Wu. Orthogonal random forest for causal inference. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 4932–4941. PMLR, 09–15 Jun 2019.

[PALvdP18] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *CoRR*, abs/1804.02717, 2018.

[Pea00] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, England, 2000.

[Pea09] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[Pea14] Judea Pearl. Comment: Understanding Simpson's Paradox. *The American Statistician*, 68(1):8–13, 1 2014.

[Pea19] Judea Pearl. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, 62(3):54–60, 2019.

[Per16] Cyril Pernet. Null hypothesis significance testing: a short tutorial. *F1000Research*, 4:621, 10 2016.

[Pia13] Jean Piaget. *Play, Dreams And Imitation In Childhood*. Routledge, 7 2013.

[PM98] Richard W. Pew and Anne S. Mavor. *Modeling Human and Organizational Behavior: Application to MilitarySimulations*. The National Academies Press, Washington, DC, 1998.

[PM16] Judea Pearl and Dana MacKenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, 2016.

[PP09] Marios G. Pavlides and Michael D. Perlman. How likely is Simpson's Paradox? *The American Statistician*, 63(3):226–233, 8 2009.

[Pro12] Mark Proctor. Drools: A rule engine for complex event processing. In *Applications of Graph Transformations with Industrial Relevance*, pages 2–2. Springer Berlin Heidelberg, 2012.

[PVS05] Jan Peters, Sethu Vijayakumar, and Stefan Schaal. Natural actor-critic. In João Gama, Rui Camacho, Pavel Brazdil, Alípio Jorge, and Luís Torgo, editors, *Machine Learning: ECML 2005, 16th European Conference on Machine Learning, Porto, Portugal, October 3-7, 2005, Proceedings*, volume 3720 of *Lecture Notes in Computer Science*, pages 280–291. Springer, 2005.

[PWYB18] Lisset Salinas Pinacho, Alexander Wich, Fereshta Yazdani, and Michael Beetz. Acquiring knowledge of object arrangements from human examples for household robots. In *KI 2018: Advances in Artificial Intelligence: 41st German Conference on AI, Berlin, Germany, September 24–28, 2018, Proceedings*, volume 11117, page 131. Springer, 2018.

[RC04] Giacomo Rizzolatti and Laila Craighero. The mirror-neuron system. *Annual Review of Neuroscience*, 27(1):169–192, 2004. PMID: 15217330.

[RD06] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine Learning*, 62(1-2):107–136, 1 2006.

[Rea14] Sara Reardon. Monkey brains wired to share. *Nature*, 506(7489):416–417, 2 2014.

[Rei16] Daniel Reisberg. *Cognition : exploring the science of the mind*. W.W. Norton & Company, New York, 2016.

[RES04] Vassilis Raos, Mina N. Evangeliou, and Helen E. Savaki. Observation of action: grasping with the mind's hand. *NeuroImage*, 23(1):193–201, 2004.

[RES07] Vassilis Raos, Mina N. Evangeliou, and Helen E. Savaki. Mental simulation of action in the service of action perception. *Journal of Neuroscience*, 27(46):12675–12683, 2007.

[Res19] Microsoft Research. EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation. https://github.com/microsoft/EconML, 2019. Version 0.11.

[RF14] Giacomo Rizzolatti and Leonardo Fogassi. The mirror mechanism: recent findings and perspectives. *Phil. Trans. R. Soc.*, 369(1644):20130420, 6 2014.

[RFG01] Giacomo Rizzolatti, Leonardo Fogassi, and Vittorio Gallese. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2(9):661–670, 9 2001.

[RPCB20] Harish Ravichandar, Athanasios S. Polydoros, Sonia Chernova, and Aude Billard. Recent advances in robot learning from demonstration. *Annual Review of Control, Robotics, and Autonomous Systems*, 3(1):297–330, 2020.

[RR83] PAUL R. ROSENBAUM and DONALD B. RUBIN. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[Rub74] D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. In *Journal of Educational Psychology*, volume 66, pages 688–701, 1974.

[Rus98] Stuart Russell. Learning agents for uncertain environments (extended abstract). In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, COLT' 98, page 101–103, New York, NY, USA, 1998. Association for Computing Machinery.

[SB98] R.S. Sutton and A.G. Barto. Reinforcement learning: An introduction. *IEEE Transactions on Neural Networks*, 9(5):1054–1054, 1998.

[Sch77] K. Schütte. *Proof Theory*. Springer-Verlag, Berlin, 1977.

[Sch99] Stefan Schaal. Is imitation learning the route to humanoid robots? *Trends in cognitive sciences*, 3(6):233–242, 1999.

[SCV19] Stephan Schulz, Simon Cruanes, and Petar Vukmirović. Faster, higher, stronger: E 2.3. In Pascal Fontaine, editor, *Proc. of the 27th CADE, Natal, Brasil*, number 11716 in LNAI, pages 495–507. Springer, 2019.

[SD06] Jessica A. Sommerville and Jean Decety. Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychonomic Bulletin & Review*, 13(2):179–200, 4 2006.

[Set09] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.

[Sha76] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

[Sim51] E. H. Simpson. The interpretation of interaction in contingency tables. *J. of the Royal Statistical Society, Series B*, 13(2):238–241, 1951.

[SK+19] Amit Sharma, Emre Kiciman, et al. DoWhy: A Python package for causal inference. https://github.com/microsoft/dowhy, 2019.

[SK20] Amit Sharma and Emre Kiciman. Dowhy: An end-to-end library for causal inference. *arXiv preprint arXiv:2011.04216*, 2020.

[SL13] Robert P. Spunt and Matthew D. Lieberman. The busy social brain: Evidence for automaticity and control in the neural systems supporting social cognition and action understanding. *Psychological Science*, 24(1):80–86, 2013. PMID: 23221019.

[SP08] Ilya Shpitser and Judea Pearl. Complete identification methods for the causal hierarchy. *J. Mach. Learn. Res.*, 9:1941–1979, jun 2008.

[Spe37] Kenneth W. Spence. Experimental studies of learning and the higher mental processes in infra-human primates. *Psychological Bulletin*, 34(10):806–850, 12 1937.

[SRMS97] Peter Spirtes, Thomas Richardson, Chris Meek, and Richard Scheines. Using path diagrams as a structural equation modelling tool, 1 1997.

[SSS+17] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 10 2017.

[STG04] David M. Sobel, Joshua B. Tenenbaum, and Alison Gopnik. Children's causal inferences from indirect evidence: Backwards blocking and

Bayesian reasoning in preschoolers. *Cognitive Science*, 28(3):303–333, 2004.

[SW18] Christian Seegelke and Peter Wühr. Compatibility between object size and response side in grasping: the left hand prefers smaller objects, the right hand prefers larger objects. *PeerJ*, 6:e6026, 12 2018.

[Tar13] M. Tarver. *The Book of Shen*. Upfront Publishing Limited, 2013.

[TCH+13] John G. Taylor, Vassilis Cutsuridis, Matthew Hartley, Kaspar Althoefer, and Thrishantha Nanayakkara. Observational Learning: Basis, Experimental Results and Models, and Implications for Robotics. *Cognitive Computation*, 5(3):340–354, 2 2013.

[The06] The European Commission. *Observational Learning in Cognitive Agents*. The European Commission, 2006.

[TP02] Jin Tian and Judea Pearl. A general identification condition for causal effects. In *Eighteenth National Conference on Artificial Intelligence*, pages 567–573, USA, 2002. American Association for Artificial Intelligence.

[TPL02] Hugo Théoret and Alvaro Pascual-Leone. Language acquisition: Do as you hear. *Current Biology*, 12(21):R736–R737, 10 2002.

[Tre00] Volker Tresp. A Bayesian Committee Machine. *Neural Comput*, 12(11):2719–2741, 2000.

[TRH07] Dennis Tkach, Jacob Reimer, and Nicholas G. Hatsopoulos. Congruent activity during action and action observation in motor cortex. *Journal of Neuroscience*, 27(48):13241–13250, 2007.

[TWS18] Faraz Torabi, Garrett Warnell, and Peter Stone. Behavioral cloning from observation. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 4950–4957. International Joint Conferences on Artificial Intelligence Organization, 7 2018.

[TWS19a] Faraz Torabi, Garrett Warnell, and Peter Stone. Imitation learning from video by leveraging proprioception. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 3585–3591. International Joint Conferences on Artificial Intelligence Organization, 7 2019.

[TWS19b] Faraz Torabi, Garrett Warnell, and Peter Stone. Recent advances in imitation learning from observation. In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 6325–6331. ijcai.org, 2019.

[UILK07] Lucina Q. Uddin, Marco Iacoboni, Claudia Lange, and Julian Paul Keenan. The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, 11(4):153–157, 4 2007.

[Vap06a] Vladimir Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer New York, 2006.

[Vap06b] Vladimir Vapnik. Transductive inference and semi-supervised learning. In Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien, editors, *Semi-Supervised Learning*, pages 452–472. The MIT Press, 2006.

[vNM47] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1947.

[VWM15] Yakir Vizel, Georg Weissenbacher, and Sharad Malik. Boolean satisfiability solvers and their applications in model checking. *Proceedings of the IEEE*, 103(11):2021–2035, 11 2015.

[WA18] Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 6 2018.

[Way22] Wayne W. LaMorte, MD, PhD, MPH. The social cognitive theory, 2022. `https://sphweb.bumc.bu.edu/otlt/MPH-Modules/SB/BehavioralChangeTheories/BehavioralChangeTheories5.html`, Last accessed on 2017-11-30.

[WDF+09] Christoph Weidenbach, Dilyana Dimova, Arnaud Fietzke, Rohit Kumar, Martin Suda, and Patrick Wischnewski. SPASS version 3.5. In *Automated Deduction – CADE-22*, pages 140–145. Springer Berlin Heidelberg, 2009.

[Wic22a] Alexander Wich. The Virtual Kitchen Layout. `https://youtu.be/eCSCdqxNMUg`, 4 2022.

[Wic22b] Alexander Wich. A Virtual Table Setting. `https://youtu.be/ektCjpsFdQE`, 4 2022.

[WS18] Peter Wühr and Christian Seegelke. Compatibility between physical stimulus size and left-right responses: Small is left and large is right. *Journal of Cognition*, 1(1):17, 2 2018.

[WSB22] Alexander Wich, Holger Schultheis, and Michael Beetz. Empirical Estimates on Hand Manipulation are Recoverable: A Step Towards Individualized and Explainable Robotic Support in Everyday Activities. In *AAMAS*, pages 1718–1720. ACM, 2022.

[YMH+19] Chao Yang, Xiaojian Ma, Wenbing Huang, Fuchun Sun, Huaping Liu, Junzhou Huang, and Chuang Gan. Imitation learning from observations by minimizing inverse dynamics disagreement. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[YZG+15] Dongjin Yu, Qi Zhu, Dalong Guo, Binbin Huang, and Jianwen Su. jbpm4s: A multi-tenant extension of jbpm to support bpaas. In Joonsoo Bae, Suriadi Suriadi, and Lijie Wen, editors, *AP-BPM*, volume 219 of *Lecture Notes in Business Information Processing*, pages 43–56. Springer, 2015.

[ZLN14] Jiangchuan Zheng, Siyuan Liu, and Lionel M. Ni. Robust Bayesian Inverse Reinforcement Learning with Sparse Behavior Noise. In *AAAI*, 2014.

[ZLZ18] QiJie Zou, Haoyu Li, and Rubo Zhang. Inverse Reinforcement Learning via Neural Network in Driver BehaviorModeling. In *2018 IEEE Intelligent Vehicles Symposium, IV 2018, Changshu, Suzhou, China, June 26-30, 2018*, pages 1245–1250. IEEE, 2018.

[ZRL96] T. Zhang, R. Ramakrishnan, and M. Livny. BIRCH: An efficient data clustering method for very large databases. In *Proceedings of ACM-SIGMOD International Conference of Management of Data*, pages 103–114, 6 1996.