UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
An Attention-Driven Computational Model of Human Causal Reasoning

Permalink
https://escholarship.org/uc/item/7bw8d284

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors
Bello, Paul
Lovett, Andrew
Briggs, Gordon
et al.

Publication Date
2018

eScholarship.org Powered by the California Digital Library
University of California

An Attention-Driven Computational Model of Human Causal Reasoning

Paul Bello, Andrew Lovett, Gordon Briggs, & Kevin O'Neill
{paul.bello, andrew.lovett.ctr, gordon.briggs.ctr, kevin.oneill} @nrl.navy.mil
U.S. Naval Research Laboratory, 4555 Overlook Ave. S.W.
Washington, DC 20375 USA

Abstract

Herein we describe CRAMM, a framework for Causal Reasoning via Attention and Mental Models. CRAMM develops and extends assumptions made by a previously developed counterfactual simulation model of human causal judgment. We implement CRAMM computationally and demonstrate how it robustly captures human causal judgments about simple two-object interactions at the level of underlying cognitive and perceptual processes, including data on eye-movements that serve as direct evidence for the role of counterfactuals in causal judgment.

Keywords: causal cognition; mental models; reasoning; attention; perception; cognitive architecture; computational model

Introduction

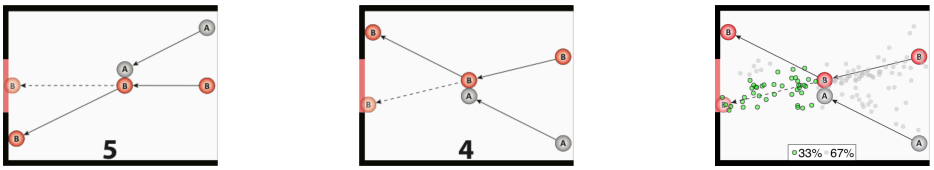
A series of recent papers have developed and defended a counterfactual simulation model (CSM) of human causal judgment (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2012, 2015; Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017). In broad strokes, the CSM assumes that human reasoners noisily sample from a generative model of Newtonian mechanics as the basis of mental simulations that track both actual and counterfactual interactions between physical objects. In turn, these simulations provide the ingredients for generating causal judgments. A recent paper extends this work with eye-movement data that corroborates central claims embodied in the CSM. Strikingly, participants make "counterfactual saccades" (see figure 1c) that seem to track what an object involved in a causal interaction would have done had the interaction failed to occur (Gerstenberg et al., 2017).

It is undeniably clear in the eye-tracking data that participants are extracting the ingredients for representing counterfactuals in an on-line fashion as they observe the stimuli in the experiment. The presence of counterfactual saccades in the eye-tracking data strongly suggests an interaction between processes supporting causal judgment and processes involved in visual perception. As it is currently expressed, the CSM makes no commitments about whether, where, or how perception makes contact with the notion of mental simulation it adopts. Consequently, the CSM offers no mechanistic explanation for why the eye-movement data looks the way that it looks other than assuming a correlation between counterfactual simulations and eye movements. In the absence of these details, we can only assert that the eye-movement data is consistent with, but not explained by the CSM. Getting closer to an explanation requires a more detailed account of how perceptual processes interact with task demands to produce judgments.

This paper reports first steps toward a detailed process model of causal reasoning called CRAMM that captures and explains both the behavioral and eye-tracking data reported in (Gerstenberg et al., 2017). CRAMM is implemented within the ARCADIA framework, which has been used to model the role of attention in a range of tasks, most pertinently in multiple object tracking (Bridewell & Bello, 2016; Lovett, Bridewell, & Bello, 2017). As input, CRAMM is presented with exactly the same stimuli (i.e., video clips) that Gerstenberg and colleagues presented to their human participants. Over the course of viewing, CRAMM attends to task-relevant objects, extracts relations between them, and builds up event structure in working memory.

In the stimuli used by Gerstenberg and colleagues, a red ball and a gray ball are on a collision course in a room with a gate on the left-hand side (see figure 1). They eventually collide, and the red ball either goes through the gate or doesn't. One of the tasks is to rate agreement with a statement such as "The gray ball [caused/prevented] the red ball [to/from] entering the gate" on a scale from 0 to 100, with 0 representing "not at all" and 100 representing "very much." At some point prior to the red and gray balls colliding, CRAMM is able to make a prediction about whether the red ball will eventually go through the gate. Subsequent collision with the gray ball may knock the red ball off course, ultimately invalidating the prediction. When this happens, the prediction becomes counterfactual, since it no longer describes the sequence of actual events as they unfolded. CRAMM maintains and updates information about the modal status of events, marking them as corresponding to predicted future events, actual events (observations), or counterfactual events that could have happened, but didn't.

With only the ability to rate agreement on statements such as "The gray ball [caused/prevented] the red ball [to/from] entering the gate" as either 0, 50, or 100, CRAMM fits the human data as well as the CSM does. CRAMM also gives a novel explanation for the eye-movement data, both within and across conditions in (Gerstenberg et al., 2017). As human participants do, prior to the collision between the red and gray balls in each stimulus clip, CRAMM tries to predict whether the red ball will go through the gate. It does so by performing spatial sweeps of covert attention down the trajectory being followed by the red ball toward the gate. Since overt attention typically follows covert attention, saccades are programmed and executed when narrow, slower sweeps of covert attention are required to resolve the question of whether the red ball will go through the gate. Following this, CRAMM is likely to generate counterfactual saccades (1) when there is a need



(a) Actual Miss / Counterfactual Hit

(b) Actual Miss / Counterfactual Close

(c) Eye-tracking Data

Figure 1: Two example stimuli from Gerstenberg et al. (2017), with eye-tracking data for the second stimulus.

to track the status of the relationship between the red ball's current trajectory and the gate, and (2), when narrow sweeps of covert attention are required to do so, as in cases where prior to the collision it is unclear whether the red ball will go through the gate (e.g., figure 1b).

The remainder of the paper begins with a brief review of the CSM, with additional review of the eye-tracking data collected by Gerstenberg and colleagues. After providing an analysis of what a computational model ought to explain, we describe CRAMM, focusing on its assumptions about perception and attention, including its commitment to the semantics for causal verbs described in (Goldvarg & Johnson-Laird, 2001). We show how CRAMM's assumptions are sufficient for capturing human ratings of causality along with qualitative trends in the eye-movement data via a computational implementation.

The Counterfactual Simulation Model

One way that the relationship between counterfactual dependence and causal judgment can be thought about is in terms of difference-making. The degree to which a candidate cause is judged as causal tracks the degree to which it made a difference to the outcome. In general, difference-making accounts rely on comparing actual outcomes to what would have (counterfactually) happened had things gone differently. The CSM frames the process we have been describing in terms of mental simulations. The idea here is that participants who view the videos sketched in figure 1 are using cognitive mechanisms that embody knowledge about physics to make predictions about different would-be events in the stimuli, including where the red ball will eventually end up. In general, the model of causal judgment implemented by the CSM works by generating noisy samples from the underlying physics engine used to produce the stimuli (seen in figures 1a-b), which can be thought of as implementing something like the ability to imagine or reconstruct happenings in the world. Each drawn sample removes the gray ball from the physics simulator, and adds a small degree of Gaussian noise to the trajectory of the red ball at each point directly after a collision with the gray ball would have occurred. The actual outcome is then recorded and stored. Each of these samples represents a counterfactual judgment, since what it captures as an outcome is the final status of the red ball given the absence of the gray ball. CSM calculates the probability that a candidate C causes a particular outcome E by computing:

$$P(C \rightarrow E) = P(E^* \neq E | S, remove(C))$$

In this case, C represents the presence of the gray ball in the sample, S represents what actually happened in the sample, E* represents whether the red ball went through the gate, and E represents the counterfactual outcome in which the gray ball wasn't present. The model predicts that participants' causal ratings will increase with their certainty that the counterfactual outcome would have differed from the actual outcome (compare figures 1a and 1b).

Eye-Tracking Causality

The CSM suggests that if participants are using runnable mental simulations of counterfactual situations, there should be signatures of the process revealed in their eye movements. To test the prediction, Gerstenberg and colleagues (2017) captured eye-movements while participants viewed 18 videos that varied over two dimensions. The first dimension was whether the red ball missed entering the gate (e.g., figure 1a), entered the gate, or was a close call regardless of whether it entered. The second dimension was arranged similarly, but instead ranged over counterfactual hits, misses, and close calls (see figure 1b for a close call). Participants were randomly assigned to one of three conditions. In the outcome condition, participants rated whether the red ball completely missed the gate (when it missed) or whether it went through the center of the red gate (when it didn't miss) on a scale from 0 ("not at all") to 100 "very much"). In the counterfactual condition, participants rated whether the red ball would have gone through the gate had the gray ball not been present. In the causal condition, participants were asked to rate either (1) whether the gray ball prevented the red ball from going through the gate when the red ball missed, or (2) whether the gray ball caused the red ball to go through the gate when it went through. The CSM captured judgments well, producing mean agreement ratings of r = .87, r = .90, and r = .92 in the outcome, counterfactual, and causal conditions, respectively.

Figure 1c shows a sample of participants' combined saccades for one clip. Gerstenberg and colleagues predicted that participants ought to run mental simulations that characterize where the red ball would have gone if the gray ball hadn't been present in both the causal and counterfactual conditions, while predicting that no such simulation is necessary for the outcome condition. Moreover, Gerstenberg and colleagues predicted a relationship between counterfactual saccades and certainty in causal judgments. In "counterfactual close" cases, where it is unclear whether the ball would have gone through the gate, they predicted that participants would

engage in a greater degree of mental simulation to determine whether the red ball would have gone through the gate had the gray ball not been present, and therefore they would make more counterfactual saccades. Both predictions about counterfactual saccades were born out in the data.

Causal Reasoning via Attention and Mental Models (CRAMM)

Each of these predictions about eye movements make good sense with respect to the counterfactual simulation model, but it remains entirely unclear why, *qua* predictions, they should necessarily follow from the model as it is stated. Perhaps the easiest and most direct answer this question is that the CSM is ultimately a model of judgment and not of the perceptual and executive control processes that support judgment. It is sometimes true, for example, that participants needn't construct or maintain any sort of predictions or make any kind of counterfactual inferences about where the red ball could have gone, had the gray ball not been present. Success on the task in the outcome condition merely requires one to follow where the red ball actually goes, and no more. This is in fact what the eye movement data reveals. However, once asked to make causal or counterfactual judgments, just tracking actual outcomes is no longer sufficient, and task-related demand for prediction and inference returns. Moreover, the fact that the eye-movement data shows more counterfactual saccades in cases where the fate of the red ball isn't clear prior to the collision suggests that these eye-movements are sensitive to processes that track task-relevant uncertainty. Perception is thus selective with respect to the task at hand. What is missing from the CSM is a story about the mechanisms from which such perceptual selectivity arises, and how it might make contact with high-level cognition.

CRAMM provides an initial set of hypotheses about the interaction between attention, perception, and the simulation process thought to be involved in online causal reasoning. CRAMM commits to a view of representation and capacity-limited reasoning based on mental models (Goldvarg & Johnson-Laird, 2001). In addition, rather than sampling from runs of an actual physics engine to generate counterfactuals, CRAMM encodes, maintains, and updates a small number of discrete possibilities over the course of stimulus viewing. These possibilities correspond to short-term episodic traces that represent objects, events, and relations that have been parsed out of the video stimuli during the viewing process. We first describe CRAMM in an implementation-neutral way, and follow with a highly abbreviated discussion of CRAMM's implementation in the ARCADIA cognitive system.

CRAMM: A High-Level Picture

Given what was gleaned from the human eye-movement data across the various conditions, we determined that task instructions impose top-down constraints on perception via selective attention. Against this backdrop, CRAMM is built so that it can be "configured" by top-down constraints imposed in the *task set*. The task set provides an interface layer between task instructions in natural language, recruited background knowledge, plans, and perception. Instructions are parsed into a task set, which provides top-down constraints on attentional deployment. Information in the task set is used to "configure" the cognitive system for the task it describes by specifying what the relevant objects, features, relations, and events are, given a specific task, along with providing a set of priorities for attentional selection. Currently, the mapping from instructions to task set to configuring the cognitive system is performed by hand (but see Future Directions).

Depending on which condition an instance of CRAMM is assigned to, instructions will be compiled down into the corresponding task set. Once so configured, CRAMM traces a path up from perception through judgment by populating working memory with event sequences. Each sequence is represented as a possibility: a collection of events that are modally tagged as either predicted, actual, or counterfactual. Once encoded, a procedure for deploying attention internally to working memory representations matches possibilities against causal concept definitions to produce judgment. The procedure below, less item 11, describes CRAMM's behavior given the task of responding to either a counterfactual or causal rating question. When configured by task instructions to merely report the final location of the red ball, as in the outcome condition, steps 4 through 6, step 8, and step 10 are never executed by CRAMM, but step 11 is. We explore the notion of "scanning" mentioned in items 4, 6, and 11 in the next section.

1. Focus and encode gate, ball A, and ball B.
2. Focus on red ball B.
3. Build trajectory information for B and use it to smooth pursue.
4. Rough scan to check whether B will enter gate.
5. Store result in WM.
6. If WM reflects uncertainty, perform tighter scan.
7. Detect and encode collision between B and other objects.
8. Mark all no-collision representations in WM as counterfactual.
9. Observe actual outcome and mark as actual.
10. Match WM contents against causal concepts and make judgment.
11. (Outcome Condition only) Perform scan from the terminal position of the red ball to the gate.

Attention: Overt and Covert CRAMM assumes basic capacities for selectively focusing on objects based on feature information encoded in the task set, such as color descriptors. Along with objects, CRAMM also assumes a capacity for covertly attending to regions of space, regardless of whether an object is present at the region. CRAMM distinguishes overt from covert deployments of attention, meaning that it is possible in principle for a CRAMM implementation to be attending to one portion of the visual field while its

"eyes" or fovea is located elsewhere. This can be seen in the top panel of figure 3. The neon green dot represents where CRAMM's "eyes" are, while the blue box represents where CRAMM's covert attention is currently directed. CRAMM assumes the capacity for planning and executing simple ballistic saccades, along with tracking an object via smooth pursuit. Crucially, CRAMM assumes that overt attention in the form of eye movements follows the deployment of covert attention, if the latter remains relatively stable for a sufficient period of time to program and execute an eye movement.

Scanning: Relational Verification CRAMM assumes that at least some relations can be verified (i.e., determined to be true or false) through operations normally thought to be perceptual in nature. For example, in the causal judgment task under consideration, whether the red ball will or actually does go through the gate depends on its trajectory, and the spatial relationship between the ball's outer contour and the area marked out by the gate. To verify whether this relation holds, CRAMM performs an initially wide sweep of covert spatial attention from the current position of the red ball along its current trajectory toward the gate. This can be seen in figure 2b, where the blue box represents the relatively wide spatial region being attended and shifted leftward down the red ball's trajectory until a determination can be made about whether or not the blue box intersects the area marked out by the red gate. If there is partial overlap, the relation may or may not be true. These values are stored in CRAMM's working memory, and can drive further processing.

Monitoring and Control Finally, CRAMM assumes that perception is active in task-relevant ways by virtue of what is in the task set. For example, if one of the goals of a human participant in this task is to determine whether the red ball will go through the gate, but a wide sweep of spatial attention leaves two possibilities in working memory (figure 2b top), a control signal will be generated, and a narrower sweep of spatial attention will be performed (by shrinking the size of the blue box) until a determination about the relation is made (figure 2b bottom). The process of relational verification proceeds on a coarse-to-fine basis in an online manner depending on need. It should be noted that it takes far less time to perform a wide sweep of spatial attention than it does to perform a narrow sweep, since the latter consists in moving a smaller blue box the same amount of distance. In many of these instances, enough time passes while covert attention is being swept such that an eye movement can be programmed and follows covert attention (see figure 2b, bottom panel).

Simulation, Mental Models, and Causal Concepts The notion of mental simulation used by the CSM and similar models predicated on a "physics engine in the head" has been criticized for appearing to be under-constrained (Davis & Marcus, 2015). One promising alternative account of mental simulation is the mental model theory of human reasoning (Goldvarg & Johnson-Laird, 2001). The theory supposes that untutored reasoners rarely (spontaneously) represent what is not immediately given as a premise or in perception, and have trouble considering any more than three possibilities simultaneously. The mental model theory is thus in broad accord with basic facts about capacity limitations in working memory. The theory makes contact with causal language by supposing that the core meanings of various words, causal verbs in this case, denote unique sets of possibilities, such as those shown in table 1. A and B are descriptions of event occurrences, with logical negation applying to generate descriptions of non-occurrences. Each row/pair represents a possibility consistent with Cause or Prevent. In CRAMM, the two different schematized sets of possibilities corresponding to "Cause" and "Prevent" are compiled down into task set information where they contribute to top-down guidance of attention. CRAMM ultimately matches the contents of working memory against these schemata to produce judgments.

Table 1: Causal concepts and the possibilities they pick out.

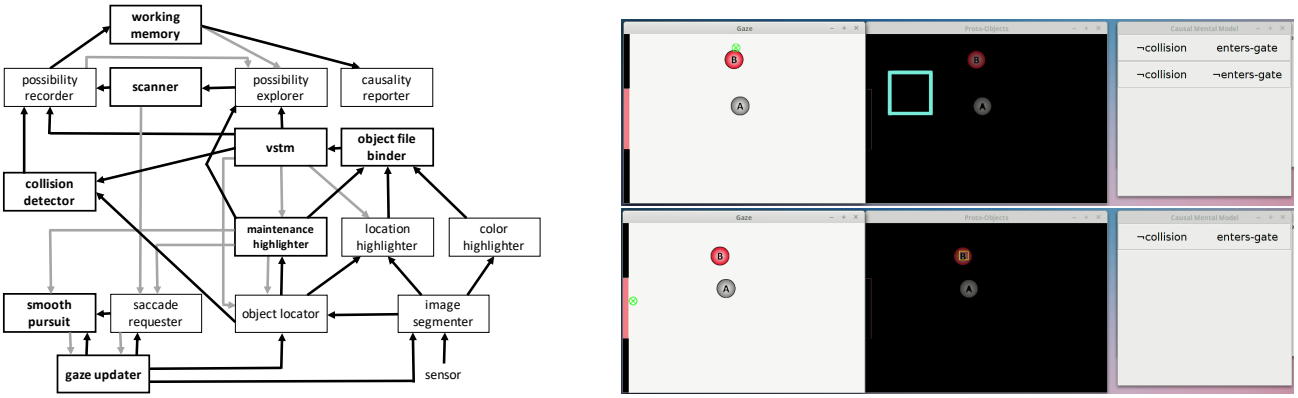
Cause		Prevent	
A	B	A	$\neg B$
$\neg A$	B	$\neg A$	B
$\neg A$	$\neg B$	$\neg A$	$\neg B$

Implementation in ARCADIA

We implemented CRAMM in the ARCADIA framework. ARCADIA provides a unique set of capabilities for building attention-centric models of cognitive phenomena. On each cycle, ARCADIA receives sensory input from the environment, and chooses a focus of attention, which is broadcast system-wide in a way that enables integrated processing (see Bridewell & Bello, 2016 for details).

Figure 2a specifies the flow of information between components in our ARCADIA implementation of CRAMM. Boxes that are boldfaced represent components that are influenced by the current focus of attention. As frames from the video clips used as stimuli in Gerstenberg et al. (2017) come into the system, pre-attentive segmentation processes identify possible objects to focus on in the scene. This information is passed through a series of "highlighters" that capture top-down effects from task goals on attentional capture. Highlighters generate requests for attention.

An attentional strategy specifies the relative priority of different requests for attention, so that an element can be selected as the focus of attention. In the case of modeling the Gerstenberg et al. results, we developed two different attentional strategies corresponding to attentional priorities for the outcome condition, and the causal and counterfactual conditions, respectively. In the former condition, attentional priority is not given to the relation between the red ball and the gate until immediately before the ball reaches the gate, at which point a quick scan determines how centrally the ball will enter the gate. In the latter conditions, tracking the status of this relation is crucial for constructing and maintaining actual and counterfactual possibilities used to drive



(a) CRAMM Model

(b) The ARCADIA implementation in action

Figure 2: A: An information-flow diagram detailing the collection of ARCADIA components used in implementing CRAMM. B: An example of CRAMM, mid-processing, tasked with answering a causal question. Top: The neon green dot represents CRAMM's fovea, while the blue box represents CRAMM's current focus of (covert) attention. On the far right, there are two possibilities in working memory after a coarse sweep. Bottom: After slower, tighter scanning, uncertainty about the "enters-gate" relation is resolved, but scanning has resulted in a counterfactual saccade.

causal judgment. Once attended, objects get placed into visual short-term memory and assigned a visual index in ARCADIA's *object locator* component (Lovett et al., 2017) so that their spatial positions can be tracked. The implementation also has components that attend and encode collisions between objects.

The *maintenance highlighter* component generates requests to maintain attention on the current focus. In this case, the component keeps ARCADIA's attention fixated on the red ball long enough to (1) execute a saccade to the red ball, (2) extract trajectory information and bind it to the representation of the red ball in visual short-term memory, and (3) begin smooth-pursuit tracking. Further task set information is included in the *possibility explorer* and *recorder* components, which implement top-down biasing from the task set. For these specific stimuli in the causal and counterfactual conditions, the *possibility explorer* is configured to verify whether or not "will enter" is true of the red ball and the gate.

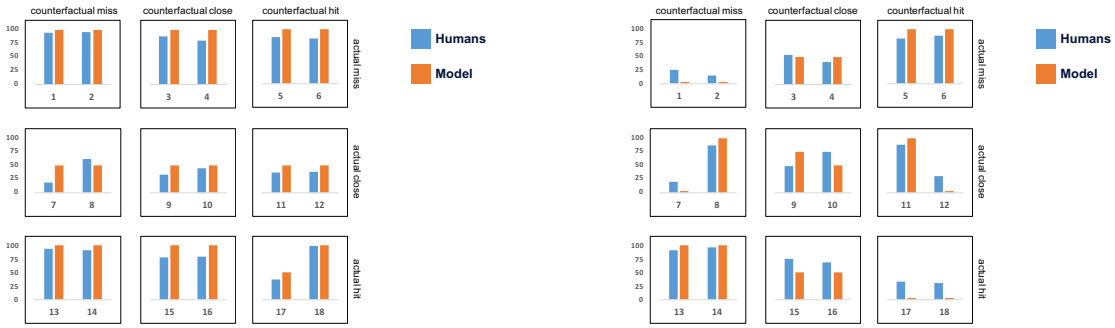
A parameterized request to perform an initial wide scan is issued to the *scanner* component, which projects covert attention along the trajectory associated with the red ball, with the scan result being attended and memorized in working memory. *Possibility explorer* continues to issues requests for tighter scans as long as no resolution to the question of "enters-gate" is encoded in working memory. *Saccade requester* requests overt attention in the form of an eye-movement whenever there is a large discrepancy (over time) between the current location of covert attention and ARCADIA's fovea. Such requests are most commonplace in "counterfactual close" trials, when a tight scanning window is required to determine whether or not the red ball will go through the gate. Because scanning requires more time on these trials, there are more opportunities to program saccades along the red ball's trajectory.

As shown in figure 2b, the entire process results in ordered possibilities that correspond to mental models of what could have and actually did happen. Finally, a *causality reporter* component matches the possibilities in working memory to the causal concept schemata (shown in table 1) via cosine similarity to generate numeric scores. In a handful of counterfactual close cases, the scanning procedure fails to resolve uncertainty in working memory via scanning before the collision takes place, leading to partial matches for both prevention and causation (i.e. a "50" rating on the numeric scale).

Experiment and Results

The data to be explained in Gerstenberg's causal judgment task is as follows: first, causal ratings should be driven primarily by differences between counterfactual and actual outcomes. This predicts middling ratings for the "counterfactual close" cases (e.g., figure 1b), and similarly predicts low ratings in cases where the gray ball has no appreciable causal effect on the outcome either way. In terms of eye-movements, more counterfactual saccades should be generated in the causal and counterfactual conditions than in the outcome condition, since there is no need to gather or maintain information about counterfactuals in the case of the latter. Additionally, the number of counterfactual saccades should vary as a function of certainty about what would have happened had the gray ball not been present.

We ran the model over all 18 stimuli in each of the three conditions in Gerstenberg et al. (2017). The model fit the data competitively with $r(16) = .93$, $r(16) = .88$, and $r(16) = .92$ in the outcome, counterfactual, and causal conditions, respectively, all $ps < .001$. The attentional strategy for the outcome condition differed from those used in the causal and counterfactual condition in that there wasn't a need to determine mid-viewing whether or not the red ball was going to enter the gate. Because this was the case, there was only one, quick scan requested immediately before the red ball reached the gate, and thus no counterfactual saccades were produced.



(a) Results for the simulation of the outcome condition.

(b) Results for the simulation of the causal condition

Figure 3: Simulation results for the outcome condition (r = .93), and the causal condition (r = .92)

For the other two conditions, counterfactual saccades were produced. As predicted, the greatest proportion of counterfactual saccades was produced when fine-grained scans were required to verify the truth of "enters-gate." This happened when, prior to collision with the gray ball, there was uncertainty about whether or not the red ball would have gone in, mirroring the pattern found by Gerstenberg and colleagues in the human data. To confirm this effect, we converted the model's causal judgments to certainty scores, with a causal score of 50 becoming a certainty score of 0 (unsure whether the red ball would have entered the gate), and a causal score of 0 or 100 becoming a certainty score of 100. There was a significant negative correlation between the model's certainty and the number of counterfactual saccades it produced, $r(16) = -.69$, $p = .002$.

Future Directions

While the CRAMM model provides a first few tenuous steps toward a process model of embodied causal reasoning, much work remains. Near-term research on the structure of task sets, and the relationship between instructions (in natural language) and elements of task sets is currently in progress. We are currently extending connections between CRAMM and the causal mental model theory to handle other causal terms such as "helping" and "allowing." Finally, the stimuli treated in this paper are very simple, with all of the relevant events and relations able to be processed on-line while the clips are being watched by the system. Simply adding one more ball to the equation adds remarkable complexity (Gerstenberg et al., 2015). Participants in those experiments often need to view the clips three or more times in order to make judgments with any degree of confidence. We postulate that an initial episodic trace is laid down during the first viewing and is elaborated upon additional viewings. Adding a richer set of facilities to parse, store, and reconstruct event sequences is yet another necessary, and near-term priority to bring CRAMM closer to a general purpose computational theory of embodied causal reasoning.

Acknowledgments

The authors would especially like to thank Tobias Gerstenberg for generously lending his stimuli and data, but most importantly his time and interest in our project. We would also like to thank Sangeet "Sunny" Khemlani, and audiences at Rensselaer Polytechnic Institute, MIT, and the Cambridge Institute for the Future of Intelligence for their valuable questions and comments. This work was supported by National Research Council Research Associateships awarded to GB and AL, a grant from the Naval Research Laboratory awarded to PB, and by grant N0001416WX00762 awarded to PB by the Office of Naval Research. The views expressed in this paper are solely those of the authors and should not be taken to reflect any official policy or position of the United States Government or the Department of Defense.

References

Bridewell, W., & Bello, P. (2016). A theory of attention for cognitive systems. In *Fourth Annual Conference on Advances in Cognitive Systems.*

Davis, E., & Marcus, G. (2015). The scope and limits of simulation in cognitive models. *CoRR*, *abs/1506.04956*. Retrieved from `http://arxiv.org/abs/1506.04956`

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2012). Noisy Newtons: Unifying process and dependency accounts of causal attribution. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society.*

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2015). How, whether, why: Causal judgments as counterfactual contrasts. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society.*

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *28*(12), 1731–1744.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive Science*, *25*(4), 565–610.

Lovett, A., Bridewell, W., & Bello, P. (2017). Goal-directed deployment of attention in a computational model: A study in multiple-object tracking. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society.* London.