sciendo

STUDIES IN LOGIC, GRAMMAR AND RHETORIC 67 (80) 2022

DOI: 10.2478/slgr-2022-0011

This work is licensed under a Creative Commons Attribution BY 4.0 License
(http://creativecommons.org/licenses/by/4.0)

Peter Gärdenfors
Cognitive Science, Lund University
e-mail: peter.gardenfors@lucs.lu.se
ORCID: 0000-0001-7423-828X

REASONING WITH EXPECTATIONS ABOUT CAUSAL RELATIONS

Abstract. Reasoning is not just following logical rules, but a large part of human reasoning depends on our expectations about the world. To some extent, non-monotonic logic has been developed to account for the role of expectations. In this article, the focus is on expectations based on actions and their consequences. The analysis is based on a two-vector model of events where an event is represented in terms of two main components – the force of an action that drives the event, and the result of its application. Actions are modelled in terms of the force domain and the results are modelled with the aid of different domains for locations or properties of objects. As a consequence, the assumption that reasoning about causal relations should be made in terms of propositional structures becomes very unnatural. Instead, the reasoning will be based on the geometric and topological properties of causes and effects modelled in conceptual spaces.

Keywords: reasoning, expectations, non-monotonic logic, causal relations, event cognition, action modelling, conceptual spaces.

1. Introduction

We have expectations of different kinds. You expect there to be a floor when you enter a room; you expect that a bird can fly; you expect a door handle not to break when you press it; you expect your morning newspaper to arrive on time; and, more generally, you expect your actions to have more or less predictable consequences. The purpose of this article is to bring out some general principles of how our expectations about causal relations influence our reasoning.

Within philosophy, reasoning is the topic of logic. In classical logic, however, the role of expectations is evaded. There are two hidden assumptions that have led to this position: (1) It is assumed that logical inference is

a relation between sentences (or propositions), and (2) the logical validity of an argument is only dependent on the logical structure of the sentences in the premises and the conclusion, and independent of their meaning or the context in which the inference is drawn (Gärdenfors 1994). Deductive logic is informationally conservative: the information in the conclusion is implicit in the premises. Thus, in deductive reasoning, drawing a conclusion from a set of premises does not require the agent to exploit semantic knowledge about the premises.

In everyday reasoning, however, expectations figure ubiquitously. We make plans, understand the behavior and discourse of other people, and interpret stories in terms of what we expect. What makes something an expectation and not a standard logical premise is that it is *defeasible*. When facts override an expectation, it yields. Consequently, expectations do not fit with assumption (2).

In earlier work (Gärdenfors and Makinson 1994; Gärdenfors 1994), I have argued that much of *nonmonotonic logic* can be reduced to classical logic with the aid of an analysis of the expectations that function as hidden premises in the arguments. The guiding idea is that when we try to find out whether a conclusion C follows from a set of premises P, the background information that we use for the inference does not only contain the premises in P, but also information about what we expect in the given situation so that we end up in a larger set of assumptions. For instance, if we know that someone is a student at a university, we expect the person to be somewhere between 18 and 30 years old. Such expectations can be expressed in different ways, for example, by default assumptions, that is, statements about what is normal or typical. These expectations are not premises that have to be accepted, but they are defeasible in the sense that if the premises in P are in conflict with some of the expectations, we don't use them when determining whether C follows from P. Expectations are used basically in the same way as explicit premises in logical arguments; the difference is that the expectations are, in general, more defeasible than the premises. Consequently, the expectations used in nonmonotonic inferences need no special notation, but they can be expressed in the same language as regular beliefs.

Another problem is that expectations do not fit with assumption (1). When we plan, we reason about actions and their consequences. To take an everyday example, when I want to be in Malmö at 3 pm, I walk to the train station in Lund and expect that I can take a train there at 2:40 pm that will bring me in time. In planning situations, the actions and consequences are used in mental simulations of *events*. As I shall argue, it is very unnatural

to assume that such simulations are based on sentential structures or on probabilistic reasoning.

There are many kinds of expectations. Many relate to correlations between properties: You expect that the bark of a large dog to have a lower pitch than that of a small dog. You expect a red apple to be sweeter than a green one. Another field relating to properties is that of *category-based inferences* where psychologists, at least since Rips (1975), have studied how people judge inferences such as "Dogs have sesamoid bones: Thus, wolves have sesamoid bones" (for a model of such inferences, see Osta Velez and Gärdenfors 2020). And as I argue in section 5, understanding *narratives* also builds on reasoning by expectations.

In this article, my focus will be on expectations based on *actions* and their consequences and how they are involved in reasoning. My analysis will be based on a two-vector model of events that I have developed in collaboration with Massimo Warglien (Gärdenfors and Warglien 2012; Warglien et al. 2012; Gärdenfors et al. 2018) where an event is represented in terms of two main components – the force of an *action* that drives the event, and the *result* of its application. This model builds on the theory of conceptual spaces (Gärdenfors 2000, 2014) that are built up from different domains. Actions are modelled in terms of the force domain and the results are modelled with the aid of different domains for locations or properties of objects. This means that the assumption that reasoning about causal relation should be made in terms of propositional structures becomes very unnatural. Instead, the reasoning will be based on the geometric and topological properties of causes and effects.

Woodward (2011: 23–24) distinguishes between two main approaches to causal cognition: difference-making and the geometrical-mechanical approach: "Difference-making accounts focus on the idea that the distinctive feature of causes is that they make a difference for their effects. [...] By contrast, what I shall call geometrical-mechanical theories of causation focus on the idea that what is distinctive about causes is that they are 'connected' in some appropriate way (often thought to be specifiable in geometrical or mechanical terms) to their effects." It should be noted that difference-making accounts typically is expressed in terms of counterfactuals of the type "if the cause C had not been present, then the effect E, would not have occurred" (Hume 1748/2000). This means that causal reasoning that is based on difference-making, builds on sentential representations, that is, it fulfils assumption (1), while the geometrical-mechanical approach is typically represented in terms of different kinds of *forces* and, consequently, does not fulfil (1).

2. A model of events, actions and results

Real world events are complex phenomena, but the mental models we create of events focus on certain aspects. This section argues that mental representations of events exploited in language, physical thinking and planning can be described in geometric terms. Such a geometric perspective on events can be found in the work of several authors (e.g. Talmy 1988; Croft 2012; Wolff 2007, 2008, 2012; Gärdenfors and Warglien 2012; Gärdenfors et al. 2018). Building on previous work on conceptual spaces (Gärdenfors 2000, 2014; Gärdenfors and Warglien 2012; Warglien et al. 2012), I model events as complex structures that involve an action space based on forces and other spaces representing the results of actions.

2.1. The two-vector model

The two-vector model states that an event is represented in terms of two components – the force of an *action* that drives the event, and the *result* of its application. Both components are represented as vectors in spaces. The result of an event is modelled as a vector representing the change of properties of the patient before and after the event. (In the limiting case when nothing changes, that is, when the result vector is a point, the event is a *state*.) The force and result vectors are central, but more vectors and objects may be involved in representations of events as I show below.

As a simple example of the model, consider Victoria planning to move around the furniture in her living room. She wants to move the sofa away from the window and she expects that she can arrange this on her own. The force vector acting on the sofa (patient) is generated by the agent Victoria. The result vector is a change in the location of the sofa and thus a change in the properties of the sofa. The exact result vector depends on the properties of the sofa, for example its weight as well as other forces in the context, for example, friction. It may happen that the sofa is too heavy for Victoria, so that, against her expectation, it does not move when she pushes.

Following Gärdenfors and Warglien (2012), I put forward the following requirement on the cognitive representation of an event:

> *The two-vector condition*: A representation of an event must contain at least two vectors and one object; these vectors are a result vector representing a change in properties of the object and a force vector that causes the change.

The central object of an event will be called the *patient*. If there is an entity generating the force vector, it will be called the *agent*. Although typical event representations contain an agent, some need not involve any: for example, events of falling, drowning, dying, growing and raining. Agents and patients are objects (animate or inanimate) that have different properties. Following my theory of conceptual spaces (Gärdenfors 2000, 2014), they can be described as vectors of values from property dimensions. At least two spaces are needed to describe an event, an action space and a result space. The action space can be conceived as a space of forces (or, more generally, force patterns) acting upon some patient, the properties of which are described in the result space (Gärdenfors and Warglien 2012; Gärdenfors et al. 2018). The spaces represent different types of vectors: forces have a different nature than changes in properties.

As the result component of the event represents changes in the properties of the patient, the result space can also be modeled as a vector space. The result vectors typically stand for changes of location or changes of object properties. For example, when Lucy opens the door, the agent Lucy exerts a force vector (action) on the door that leads to a change of the position of the door (result). Or in the event of the storm felling a tree, the force of the wind (action) leads to a change of the direction of the tree (result).

Events are represented not only as single instances, but more generally as event *categories*, for example, throwing a ball. The description of change vectors can be generalized to that of change *vector fields* by associating to each action force vector a result vector, taking into account the (counter-)force exerted by the patient and other contextual forces. An event category is a structure that represents how the agent space affects the patient vector field. Thus, we need to represent not only the mapping from one action to the patient vector field, but the way all actions in the action category of the agent space affect the patient. For example, an event category of pushing a table should be able to represent the effect of different, albeit similar, patterns of force on the different points in the table patient space.

Mathematically, such a mapping from actions to results can be seen as a function from a force vector that is the resulting combination of the action vector and other contextually given forces to a result vector (see Gärdenfors and Warglien 2012) for a more detailed description of the mapping). This mapping is part of the representation of an event category and it contains the central information about causal relations. The expectations about the effects of causes are generated by assumptions concerning the counterforces, typically that have some default values.

Denote by F_A the force vector exerted by the agent and F_P the counterforce vector exerted by the patient or by other forces such as gravitation. The resultant vector $F_R = F_A + F_P$ is what determines the result vector R.

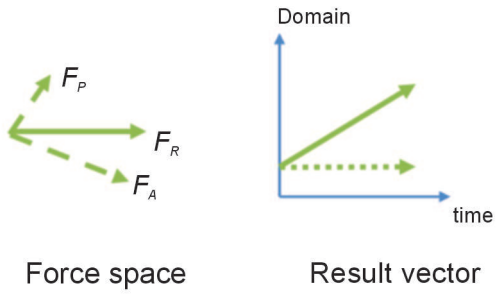


Figure 1. Causation expressed in the force and result spaces

The left part of figure 1 represents the force vectors and the right part shows how the location in the result domain changes over time as a consequence of the resultant force vector F_R. Following Hume's (1748/2000) definition of causation, a causal judgment has a counterfactual element. The hatched line in the right part of figure 1 represents the counterfactual situation that the change in the result domain would not have changed if F_R had been nil.

The events need not only involve physical forces, but also mental 'forces' can be causal variables (Talmy 1988; Leslie 1994). Humans interpret many mental factors (for example commands, threats, insults and persuasive arguments) as forces that can create a change in the physical, cognitive or emotional state of the addressee. For example, Wolff (2007: 19–22) presents two experiments where a woman intends to cross a street to meet a man and the directions of a police in the street crossing acts as an additional 'force' that enables or prevents the woman from reaching her goal. The results show that the subjects interpret the woman's intention as a force and they describe the various scenarios in the same terms as they would use for a situation where only physical forces are involved. Wolpert et al. (2003) present an analysis of how this kind of reasoning can be modeled in terms of control theory. Again, it should be noted that such models are analog and thereby non-propositional.[1]

The forces can also be medical, economic or social (Talmy 1988). For example, a person expects that if he takes an aspirin, this will cause his headache to go away, since the medicine acts a force causing a change in his physical state. And a woman may expect that if she is offered a high price

for her mother's wedding ring, this will make it possible for her to sell it. In this case, the price acts as a force *enabling* the woman to act. A social example is "The pressure from the villagers caused him to mow his lawn, even though he wanted to keep it as a meadow". To be sure, mental, medical, economic and social 'forces' are not of the same character as physical forces. A better term covering them all would perhaps be 'powers'. Nevertheless, extensive experimental evidence shows that human subjects treat all these types as causes in events (see e.g. Wolff 2007).

2.2. Representing actions

I next turn to a more detailed description of the action component of the model. Following Gärdenfors (2007) (see also Warglien et al. (2012) and Gärdenfors (2014)), I proposed that the human brain extracts the forces that lie behind different kinds of actions. This leads me to the following thesis:

> *Representation of actions*: An action is represented by a pattern of forces that generates it.

The thesis speaks of a pattern of forces since, for bodily motions, several body parts are involved; and thus, several force vectors are interacting (by analogy with Marr and Vaina's (1982) differential equations). One can represent these patterns of forces in principally the same way as the patterns of shapes described in Gärdenfors (2014, Section 6.3).

In order to identify the structure of the action space, similarities between actions should be investigated. This can be accomplished by the same methods used for investigating similarities between objects. Just as there, the dynamic properties of actions can be judged with respect to similarities: for example, walking is more similar to running than to waving. A large set of such similarity ratings can serve as input for one of several related statistical techniques, such as multidimensional scaling and principal component analysis, that turn similarities into spatial structures. I assume that the notion of betweenness is meaningful in the action space, allowing the following thesis (in analogy with the thesis about properties in Gärdenfors (2000, 2014)):

> *Thesis about action concepts*: An action concept is represented as a convex region in the action space.

It is natural to interpret convexity as the assumption that, given two actions in the region of an action concept, any linear morph between those

actions will fall under the same concept. For some empirical support for the thesis, see Giese and Lappe (2002) and Gärdenfors (2014).

3. The causal mapping

The main reason for introducing the event model is that it is a natural way of capturing how we reason about causation: the action *causes* the result. Davidson (1969) writes that "events have a unique position in the framework of causal relations between events". Most accounts of causation analyze the relation between the action and the effect as a relation between two events (see e.g. Zacks and Tversky 2001, Casati and Varzi 2008). In contrast, the model presented here describes causation as a relation *within* an event by introducing a distinction between forces and changes of states. There are many similarities between the event model presented here and Wolff's (2007, 2008, 2012) dynamics model. His affector vector corresponds to the force vector, his patient vector to the counterforces, and he also includes a result vector. The two models have been developed for slightly different purposes: the two-vector model is intended as a general model of events while Wolff focuses on causal reasoning. Another difference is that his result vector is presented as being of the same kind as the force vectors. In contrast, in the model presented here, causes and effects are modeled as entities of different types: they are modelled in different spaces – causes to the force space and results to change in location space (in the case of movements) or in some property space (color, size, shape, weight, temperature, etc).

When determining the expectations generated from an event model, the mapping between the force space and the result space is the most central part of the event model. A problem is that externalities, such as friction and other counterforces, make it difficult to determine the result vector, given only the force vector of the agent. For example, pushing a coffin sometimes results in the coffin moving, sometimes not; taking a medicine sometimes cures a patient, sometimes not. As above, the expectations are determined by assumptions about the relevant counterforces.

Given our ignorance of the counterforces in a situation and the limited knowledge about the relevant causal relations, it is often very hard to precisely predict the outcome of an action. Still, the effect of actions can be understood in qualitative terms.

The two-vector model of events still has testable consequences. For example, experimental studies suggest that individuals can perform intu-

itive addition of force vectors when observing how simultaneous sources of force affect a patient trajectory (Wolff 2007). Classical experiments based on Michotte's (1963) launching paradigm show that subjects' attributing of causality in an event depends on the angle of the trajectory of a hit object B with a hitting object A. This shows that individuals judge whether an animation captures a single event or two disjoint ones on the ground of the mapping of forces to movements. More recent research has shown the perception of such a mapping to be remarkably accurate, and a good predictor of 'causal impression' (White 2012). In these cases, the two-vector model of events predicts well how individuals perceive causal events.

As an illustrative example, consider a malevolent bus driver who wants to make a passenger fall down. In order to do so, he abruptly brakes while driving at high speed.[2] In this case one can distinguish a series of three events: (i) The bus driver's intention acts as a force, that results in his pressing the brake pedal. In this event the result is an action that itself belongs to the force domain. (ii) The force exerted on the brake pedal, has as a result that the bus quickly diminishes its speed. The brake pedal exerts a counterforce to the driver's foot, but the force of the foot overcomes it. (iii) The rapid slowing down of the bus results in the passenger falling down. In this event, there is no agent. The force involved comes from the kinetic energy of the passenger. The friction from the feet on the floor generates a counterforce that make the legs of the passenger slow down, while the rest of his body continues at the same speed, which leads to the fall. The experience of the bus driver makes him foresee that events (ii) and (iii) will be consequences of event (i).

The event model, can handle *what-if* questions, that is, counterfactual reasoning concerning what would have happened if an action would have been different. For example: "If I had dropped the glass on the ceramic floor instead of on the mat, then it would have broken". Such reasoning can be computationally modeled by simulations of various changes in the force and counterforce vectors and using the mapping function and assumed counterforces to predict a result. Simulations use similarity measures and operations for projecting forward and backwards to understand the causes and consequences. In contrast, probabilistic models of causation have deep-going problems in handling what-if reasoning (Pearl 2018; Gärdenfors to appear).

Similar to counterfactual reasoning, humans often reason in terms of *omissive* causation, that concerns events that do not occur. For example, the fact that a person did not take a vaccination, caused that the person got

sick with Covid19. Omissive causation is a problem for many other models of causation, but the two-vector model can also explain this form (for related solutions see Talmy 1988, Wolff et al. 2010, and Wolff and Thorstad 2017). Again, probabilistic models have problems to account for omissive causation.

4. Three qualitative principles governing causal relations

The formal nature of event mappings has been little investigated. Although other theories of events (Talmy 1988; Croft 2012; Wolff 2007, 2008) also recognize such a mapping, they do not analyze it. Gärdenfors et al. (2018), however, present an analysis of three general principles for event mappings, that constrains the relation between the force vector and the result vector. All three principles are expressed in qualitative terms, which reflects the qualitative nature of event cognition and reasoning about causation.

Three mathematical properties of the event mapping function correspond to the three principles: (1) The function being *monotonic* can be expressed as that larger forces lead to larger results (this relates to qualitative causal thinking); (2) the function being *continuous* can be expressed as that small changes in forces lead to small changes of the result; and (3) the function being *convex* can be expressed as that intermediary results are caused by intermediary forces. A strong motivation for investigating them is that human causal thinking typically satisfies these properties. I next turn to a brief description of the three principles.

4.1. Monotonicity

A general constraint for expectations based on causal relations is that whenever counterforces and other external factors are kept constant in a given situation, then increasing the force involved in the action will also lead to a larger result (or at least not decrease it). For example, if I press the gas pedal harder in my car, I expect it to run faster.

This constraint captures an important part of our reasoning about how a change of an action will change the outcome. The constraint implies at least a qualitative predictability of the effects of actions and it is a mechanism that generates a variety of expectations. It is a central component in understanding causality (Hume 1748/2000; Wolff 2007, 2008) and in making causal inferences.

The constraint enables qualitative causal inferences. First of all, it makes it possible to draw ordinal inferences about the relationship between

causes and effects. For example, since different individuals may react with different intensity to a medical treatment, it is difficult to make quantitative predictions. One may, however, still expect that larger doses of the treatment will have larger effects, independently of the person who is treated. Mill (1843) called this form of causal inference 'the method of concomitant variations'.

Mathematically, this constraint corresponds to the *monotonicity* of the mapping function. A function is said to be monotonous when $f(x) \leq f(y)$, whenever $x \leq y$. This property thus depends on an ordering relation on the forces. As long as all forces act in the same direction such an ordering exists. However, in higher dimensional spaces such an ordering function may not exist.

To return to the case of a medical treatment, a physician may not know the exact mechanism of why a pain reliever works, but she may from experience have learnt enough about the causal mapping from the doses of the medicine to its effects to be able to determine the amount of the medicine that will have the desired effect. Thus, following Mill's considerations, knowledge about the causal mapping is to a large extent a matter of experience.

4.2. Continuity

When the aim is to change the effect of an action only by a small amount, it can be achieved by applying a correspondingly small change of force. For example, when turning the control for a heater on a stove a little more to the right, one expects the heat also to increase just a little, and not lead to a drastic change that would destroy the food. And when a tennis ball is hit a little harder, it will fly a little faster and further, but not move wide out of the court.

Mathematically, this constraint corresponds to the *continuity* of the mapping function. This can be defined in terms of a nearness relation on the space, which is easily defined for the force space.[3]

While the constraint captures a very general principle of causal thinking, it is not always true that small changes in the force lead to small change in the result. Sometimes small changes lead to phase transitions. For example, if you are gradually increasing your arm force when bending a wooden stick, there is a point where the stick breaks. At the transition point, a very small change of effort produces a large effect. In more general terms, a discontinuous phase transition occurs when an obstructing counterforce is suddenly overcome, and a drastically different result is achieved.

4.3. Convexity

Imagine that you are throwing a ball at a basket. You can control the forces of your arms in the throw. If you have tried force x and observed that the ball was short of the basket and tried force y and observed that the ball went too far, then you expect that a force of a strength between x and y will lead to an intermediary result.

The third constraint can be formulated as that the causal mapping f is convexity preserving: if the force vector z is between force vectors x and y, then the result $f(z)$ is between the results $f(x)$ and $f(y)$. In other words, intermediate forces lead to intermediate results. Therefore, this constraint depends on the fact that betweenness is defined for the force and result spaces.[4]

Convexity is fundamentally involved in the process of *learning* new concepts (see Gärdenfors 2000, Ch. 3). From a cognitive point of view, convexity is closely related to *generalization* since, by interpolation, it allows inferences over whole regions from a limited number of observations. In general, feedback control mechanisms presume that the mapping from actions to result is convexity preserving (e.g. Shiriaev et al. 2007).

To sum up this section, the three qualitative constraints do not uniquely determine the mapping from causes to results, but they add rich structure to it. The constraints make it possible to generate robust expectations about causal consequences of actions even if counter-forces and other contextual factors are unknown. In this way, the constraints considerably strengthen human causal reasoning.

5. Reasoning about causes and effects

In analytic philosophy, causes and effects are expressed as sentences and they are seen as the same kind of entities. In the event model presented here they are instead modelled as force and result vectors. Thus, causes and effects are of different kinds. The relation between the vectors – the causal connection – obeys certain mathematical principles that have been described above.

The event model can also explain some variations of causation that are typically expressed by the verbs "enable" and "prevent" and the preposition "despite" which are central in Talmy's (1988) force dynamics. Wolff (2007) presents an analysis of these notions in terms of a force vector model and mine is an extension of his. In particular, Wolff does not distinguish the result space from the force space. As I shall show, distinguishing between these vector enables a more general and systematic account of different causal situations.

(i) Enable

In this case, the two force vectors work together towards a resulting goal. The single patient vector is not sufficient to make the result vector reach the goal (marked by the counterfactual result vector R_P in the right part of figure 2). But adding the agent force vector F_A, the goal is reached. The hatched line marks the counterfactual result vector that would have been caused by the F_P alone.

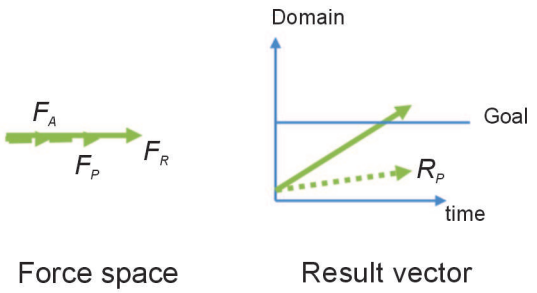


Figure 2. Enabling by adding two force vectors that work in the same direction

(ii) Prevent

In cases of this type the agent vector taken alone is sufficient to make the result vector reach the goal (marked by the counterfactual result vector R_A in the right part of figure 3). But adding a patient force vector (or a vector from another source) F_P, the goal is no longer reached. The hatched line marks the counterfactual result vector that would have been caused by the F_A alone.

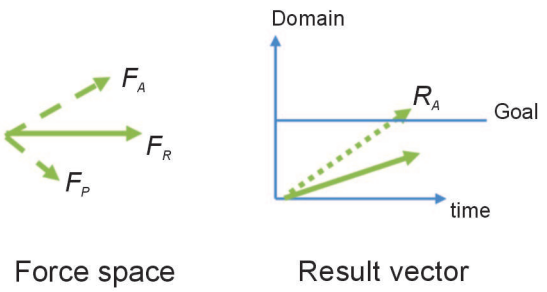


Figure 3. A result is prevented since the patient force vector counteracts the agent vector

(iii) Despite

'Despite' event situations can be seen as a special form of causation where the counterforce vector F_P goes in the opposite direction of the agent vector F_A, but the goal is nevertheless reached.

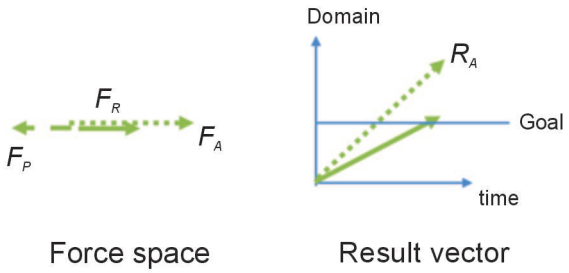


Figure 4. A goal is reached despite a counteracting patient force

Similar to what-if reasoning and omissive causation, probabilistic models have severe problems explaining causation involving "enable", "prevent" and "despite". This is a further argument in favor of the two-vector model. Rather than logical inferences involving sentences, reasoning with causes and effects involves a mental simulation where a force vector representing the cause is combined with expected counterforces and other contextual factors within the event frame to predict a result vector. What makes the simulations cognitively difficult is that the knowledge about the relevant contextual factors may be scant. This is where the expectations are added to fill the information gaps.

A common example of such mental simulations is how our minds work when we understand *narratives*. The process is similar independently of whether we are listening to a story, reading a novel, or watching a movie. If you, for example, in a movie see a woman enter a car alone at her home and then, in the next scene, see her come out of the car in front of an office building, your expectations lead you infer that she has been driving between the two places. Our minds simulate the intervening process and thereby fill in the gaps in the story. Good narratives typically do not express what is expected, only that which is not expected. And good jokes often involve violations of expectations.

6. Conclusion

One aim of this paper has been to demonstrate that reasoning with expectations is much broader than what traditional propositional logic ac-

counts for. I have argued that expectations function as extra premises, although defeasible, in an argument. Here I have focused on expectations about causation and events, but there are several other types of expectations that also widen the logical domain, for example the category-based inductive inferences that have been studied in psychology (Osta Velez and Gärdenfors 2020). Also induction as it has been analyzed in philosophy has been presumed to involve a relation between sentences. However, it is more natural to see inductive inferences as a relation between concepts (Gärdenfors and Stephens 2017).

Following Woodward (2011), one can distinguish between difference-making and the geometrical-mechanical approaches to causal cognition. This article follows the geometrical-mechanical approach since the event model that I have outlined build on vectors in spaces and forces underlying actions.

In the dominating accounts of reasoning within philosophy, it is assumed that reasoning involves sentences (or propositions). However, given the geometrical-mechanical account of how expectations are involved in reasoning with causes and effects, mental simulations play a central role. These simulations involve different kinds of variables – physical, mental, social, etc. – rather than sentential structures. Since many animal species can plan their actions (Gulz 1991), without having any linguistic abilities, I conclude that mental simulations based on expectations are cognitively more fundamental than pure logical reasoning.

NOTES

[1] Further connections to reasoning about causation in psychology and linguistics can be found in Gärdenfors (2021).

[2] This example was suggested by a reviewer.

[3] The precise definition is: A mapping $f\colon X \to Y$ between topological spaces is called *continuous* if the pre-image under f of any open subset of Y (denoted $f^1(Y)$) is an open subset of X. It should be noted that any metric induces a nearness relation.

[4] Again, a metric induces a betweenness relation. If S is a space with a metric d, then z in S lies *between* x, y in S if $d(x, y) = d(x, z) + d(z, y)$. However, betweenness can be defined for many other structures, for example, graphs.

REFERENCES

Casati, R. and Varzi, A. (2008). "Event concepts". In: Shipley, T.F. and Zacks, J. (eds.), *Understanding Events: From Perception to Action.* New York: Oxford University Press, 31–54.

Croft, W. (2012). *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.

Davidson, D. (1969). The individuation of events. In: Rescher, N. (ed.), *Essays in the Honor of Carl Hempel*. Dordrecht: Springer, 216–234.

Gärdenfors, P. (1994). "The role of expectations in reasoning". In: Masuch, M. and Polos, L. (eds.), *Knowledge Representation and Reasoning Under Uncertainty*, Berlin: Springer-Verlag, 1–16.

Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought* Cambridge Mass.: The MIT Press.

Gärdenfors P. (2007). "Evolutionary and developmental aspects of intersubjectivity". In: Liljenström, H. and Århem, P. (eds.), *Consciousness Transitions: Phylogenetic, Ontogenetic and Physiological Aspects.* Amsterdam: Elsevier, 281–305.

Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*, Cambridge Mass.: The MIT Press.

Gärdenfors, P. (2021). "Causal reasoning and event cognition as evolutionary determinants of language structure". *Entropy* 23(7), 843.

Gärdenfors, P. (to appear). "Conceptual engineering for engineers (and for philosophers): The role of conceptual spaces". In: Isaac, M.G., Koch, S. and Sharp, K. (eds.), *New Perspectives on Conceptual Engineering*. Vol. 1: *Foundational Issues in Conceptual Engineering*. Berlin: Springer.

Gärdenfors, P. and Makinson, D. (1994). "Nonmonotonic inference based on expectations". *Artificial Intelligence* 65, 197–245.

Gärdenfors, P. and Stephens, A. (2017). "Induction and knowledge-what". *European Journal for Philosophy of Science* 8, 471–491, DOI 10.1007/s13194-017-0196-y.

Gärdenfors, P. and Warglien, M. (2012). "Using conceptual spaces to model actions and events". *Journal of Semantics* 29, 487–519.

Gärdenfors, P., Jost, J. and Warglien, M. (2018). "From actions to events: Three constraints on event mappings". *Frontiers in Psychology*, 14 August 2018, DOI: 10.3389/fpsyg.2018.01391.

Giese, M. A. and Lappe, M. (2002). "Measurement of generalization fields for the recognition of biological motion". *Vision Research* 42, 1847–1858.

Gulz, A. (1991). *The Planning of Action as a Cognitive and Biological Phenomenon*, Lund University Cognitive Studies, 2. Lund: Lund University.

Hume, D. (1748/2000). *An Enquiry Concerning Human Understanding*. Oxford: Clarendon Press.

Leslie, A. M. (1994). "ToMM, ToBy, and agency: Core architecture and domain specificity". In: Hirschfeld, L.A. and Gelman, S.A. (eds), *Mapping the Mind: Domain Specificity in Cognition and Culture.* New York: Cambridge University Press, 139–148.

Marr, D. and Vaina, L. (1982). "Representation and recognition of the movements of shapes", *Proceedings of the Royal Society in London*, B214, 501–524.

Michotte, A. (1963). *The Perception of Causality*. New York: Methuen.

Mill, J. S. (1843). *A System of Logic*. London: John W. Parker.

Osta Velez, M. and Gärdenfors, P. (2020). "Category-based induction in conceptual spaces". *Journal of Mathematical Psychology* 96, DOI 10.1016/j.jmp.2020.1s 02357.

Pearl, J. (2018). "Theoretical impediments to machine learning: With seven sparks from the causal revolution". arXiv:1801.04016vi

Rips, L. J. (1975). "Inductive judgments about natural categories". *Journal of Verbal Learning and Verbal Behavior* 14, 665–681.

Shiriaev, A. S., Freidovich, L. B., Robertsson, A., Johansson, R., and Sandberg, A. (2007). "Virtual-holonomic-constraints-based design of stable oscillations of Furuta pendulum: Theory and experiments". *IEEE Transactions on Robotics* 23(4), 827–832.

Talmy, L. (1988). "Force dynamics in language and cognition". *Cognitive Science* 12, 49–100.

Warglien, M., Gärdenfors, P. and Westera, M. (2012). "Event structure, conceptual spaces and the semantics of verbs". *Theoretical Linguistics* 38, 159–193.

White, P. A. (2012). "Visual impressions of causality: Effects of manipulating the direction of the target object's motion in a collision event". *Visual Cognition* 20(2), 121–142.

Wolff P. (2007). "Representing causation". *Journal of Experimental Psychology: General* 13, 82–111.

Wolff, P. (2008). "Dynamics and the perception of causal events". In: Thomas, S. and Zacks, J. (eds.), *Understanding Events: How Humans See, Represent, and Act on Eve*nts. Oxford: Oxford University Press, 555–587.

Wolff, P. (2012). "Representing verbs with force vectors". *Theoretical Linguistics* 38, 237–248.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). "A unifying computational framework for motor control and social interaction". *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358(1431), 593–602.

Woodward J. (2011). "A philosopher looks at tool use and causal understanding". In: McCormack, T., Hoerl, C. and Butterfill, S. (eds.), *Tool Use and Causal Cognition*. Oxford: Oxford University Press, 18–50.

Zacks, J. M. and Tversky, B. (2001). "Event structures in perception and conception". *Psychological Bulletin* 127(1), 3–21.