

Commonsense Causal Explanation in a Legal Domain*

RINKE HOEKSTRA and JOOST BREUKER

*Leibniz Center for Law, University of Amsterdam, P.O. Box 1030 Amsterdam, 1000 BA,
The Netherlands*

E-mail: hoekstra@uva.nl; E-mail: breuker@science.uva.nl

Abstract. In this paper, we present an approach to commonsense causal explanation of stories that can be used for automatically determining the liable party in legal case descriptions. The approach is based on *LRICore*, a core ontology for law that takes a commonsense perspective. Aside from our thesis that in the legal domain many terms still have a strong commonsense flavour, the descriptions of events in legal cases, as e.g. presented at judicial trials, are cast in commonsense terms as well. We present design principles for representing commonsense causation, and describe a process-based approach to automatic identification of causal relations in stories, which are described in terms of the core ontology. The resulting causal explanation forms a necessary condition for determining the liability and responsibility of agents that play a role in the case. We describe the basic architecture and working of *DIRECT*, the demonstrator we are constructing to test the validity of our process oriented view on commonsense causation. This view holds that causal relations are in fact abstractions constructed on the basis of our commonsense understanding of physical and mental processes.

Key words: causation, commonsense, legal reasoning, liability, ontology

1. Introduction

Studies in cognitive science, such as in (psycho)linguistics (Clark 1996) or AI (Hobbs and Moore 1985), show that an enormous amount of commonsense knowledge is needed to understand everyday situations and events, whether directly perceived or described in natural language. Moreover, this understanding appears effortlessly and is highly accurate, even though scientific theories of physical, biological and psychological processes show that the ‘folk’ theories on which we base our day-to-day understanding are false and

*This is a significantly revised version of the paper presented at the LOAIT 2005 workshop in Bologna.

inaccurate under varying circumstances. We can easily predict the outcome of our own actions, explain the seemingly random occurrence of processes, and we are even capable of approximating an understanding of what other people think and believe. And all this, without even the slightest pretension of knowing exactly what is going on: we just have the *general idea*.

Our understanding of physical occurrences and agent behaviour requires not only the identification of objects and processes, but also some explanation of *why* certain events take place. To this end, we continuously construct models that identify the causal relations between the events that occur in situations.

In this paper we will describe an approach to capture this commonsense causation and will describe how it can relevantly contribute to the legal domain, a domain that is interspersed with commonsense reasoning (Section 2). Section 3 describes how commonsense stories and reports differ from the more traditional case-oriented approaches. Section 4 describes the generic framework for legal reasoning that underlies this approach. We will then describe how case descriptions (stories) can be causally enriched using *DIRECT* in Section 5 (see also (Breuker and Hoekstra 2004b; Hoekstra and Breuker 2005)). Furthermore, we will describe how it differs from existing work in Section 6.

2. Causation and liability

Causation plays a central role in virtually all scientific disciplines. In all cases its use is more precise than in commonsense – for instance, it is distinguished from covariance and correlation – but essentially it does not differ in its capacity to explain dependencies between events. However, reflection about its use in science has raised doubts about its status in theories. For instance, Russell believed that the notion of physical causality obscured rather than clarified thinking about physics. At the end of this section we will argue similarly when it comes down to commonsense, and describe how a too abstract notion of causation muddles discussions on liability theory in law.

2.1. CAUSATION OR CAUSALITY?

The definition of causality is a major issue in philosophical metaphysics, and has been for many centuries (Kim 1998; Davidson 2001; Schaffer 2003). We will not enter this debate (again) but refer to the work of Lehmann (2003) that reviews the main philosophical positions on causation, and is part of the work we report here. Philosophy primarily concerns itself with what causality *means*, i.e. it is focused on ontological questions regarding the existence and

properties of causality. Epistemological questions arise concerning how we can know about the occurrence of causality in the real world and, once we do know, what inferences we can draw from the causal relation between two events.

Lehmann (2003) adopts a distinction between the notions of causality and *causation*; he uses the term ‘causality’ to denote the ontological view, and reserves the term ‘causation’ for the *occurrence* of causality. At first sight, this may seem paradoxical as we usually assume the occurrence of a concept to be its instance. However, causal dependencies between events are not based on the recognition of an instance of causation. Causation is rather an abstract, reflective concept that summarises a more qualitatively distinct relation. In other words, after we perceive a collision between two billiard balls, we infer a transfer of power from the moving ball to the static ball. This allows us to say that the collision caused the second ball to move. In a nutshell, the basis for the approach presented here is that recognising the transfer of power at the collision is sufficient for our understanding; the assertion that the collision caused the ball to move is superfluous.

For this reason, the question whether causality is a real thing in the world, as is the major theme in philosophy, is of less importance to us. We rather give an account of causation, i.e. how we may understand relationships between events in terms of processes – e.g. transfer of energy or power. The intermediary role of processes can be conceived as a dependency: as causation between events. In this view, causation is no more and no less than the (abductive) inference that the occurrence of event *B* can be explained by a (sub-)process that is implied by an earlier or simultaneous event *A*. In summary, our focus on *causation* rather than *causality* allows us to respectfully avoid the unsettled philosophical debate on causality.

2.2. LIABILITY

The attribution of responsibility is one of the most prominent tasks in legal reasoning. And although highly specialised, commonsense reasoning plays an important role in establishing the liable party in court. Hart and Honoré (1985) define legal responsibility as the “liability of a person to be punished, forced to compensate, or otherwise subjected to a sanction by the law”. In other words, if a judge or other authorised party acknowledges someone’s liability to be punished, the person is considered to be held legally responsible. Interpreting and explaining written accounts of events (e.g. official reports) is certainly one of the primary tasks of the court. The causal relation between events is not just an incidental aspect of legal reasoning, but can be considered the second most important after the application of norms; at least for determining liability, see Valente (1995). As described in Lehmann (2003),

(legal) causality, responsibility, liability and guilt are closely intertwined concepts in legal reasoning aimed at compensating for wrong doing. The work described here is based on Lehmann (2003), but deviates from his more recent work of Lehmann et al. (2004) and Lehmann and Gangemi (2006) in that it has a stronger AI flavour.

Lehmann did a thorough investigation of the main aspects of legal causation in Artificial Intelligence, and proposed a high-level ontology for representing causation in law, cf. Lehmann (2003). Following Hart and Honoré (1985), he identifies the following grounds for the attribution of legal responsibility to a person for a given harm:

1. the *conduct* of a person,
2. the *causal connection* between the conduct of the person and the given harm,
3. the *fault* legally implied by the conduct of the person.

These grounds can be combined to form the following 7 categories: (1) conduct, causation, and fault; (2) conduct and causation (strict liability); (3) conduct occasioning harm and fault; (4) conduct occasioning harm; (5) conduct and fault (common in criminal law); (6) conduct; (7) purely legal ground (insurance or guarantee). *Occasioning* is a weak counterfactual causal connection: it can be tested by the *sine qua non* rule.

The discourse on liability and legal causation in legal theory suffers from causal overloading. This overloading takes place on three levels: *terms*, *conclusions* and *policy*. First, terms such as conduct, fault and harm can only be used as labels for actual events after careful consideration and (non-legal) causal reasoning. Restricting a theory on liability attribution to the (causal) relation between those terms is therefore biased with respect to some particular interpretation of events, and does no real justice to the complexity of the legal and causal reasoning underlying that interpretation. Second, the 7th category of Hart & Honoré, i.e. a purely legal ground, is a good example of where complex causal and legal reasoning is obfuscated by its conclusion, i.e. which party is deemed liable. Although it is certainly true that parents are liable for damage caused by their children (and as such are not liable on the bases of conduct, causal connection or fault by *themselves*), there still needs to be ‘conduct’ or ‘fault’ (e.g. breaking someone’s window) by some *other person* (the child). This secondary chain of reasoning is perhaps not legally relevant, but is certainly necessary for establishing the liability relation between the broken window and the liable party. As a rule, legal causation is thus always established *after the fact*, i.e. after the reasoning process is completed. Third, ‘causal’ relations between conduct and harm are not always necessarily causal as far as commonsense is concerned, but can be based on policy or legal tradition. For instance, in Dutch law a fault may be

almost retroactively deemed to have legally caused some harm, based on criteria such as the vulnerability of the victim.

An example of the overloaded causal relation is *negative causation*, or negligence. Negative causation concerns the case where some event would not have occurred if an agent had acted appropriately. Negative causation occurs in the case where an actor *should* have performed some action, but didn't. It is therefore a *normative* statement, comparing perceived behaviour to some idealised standard behaviour.

Liability is not only determined on the basis of successful actions, *attempts* to act can be a ground as well. For instance, negative causation often does not hold when an agent unsuccessfully attempted to prevent the harm. Attempts to act do not result in the desired action, but rather some rough approximation of it. For these reasons, the distinction between commonsense causation and *legal* causation needs to be scrupulously guarded. The latter has an overlap with the concept of liability, where the former has not.

In the same way, *physical* causation and causation in *fact* (or factual causation) are separate concepts: causation in fact represents the non-legal part of (established) legal causation and covers both physical and agent causation. The construction of a (partial) commonsense causal chain is always necessary for determining liability, even for cases where the liability is not based on some form of legal causation. Thus, in order to reach the higher goal of automatic liability attribution, we need to have an explicit representation of both commonsense causation, and of commonsense *legal* causation.

3. Cases as stories

Traditional approaches to reasoning on legal cases in AI&Law (e.g. Ashley 1990, 1997)) generally concern themselves with a comparison of cases as described in the verdict of a judge. Such case descriptions contain only those elements which were deemed relevant by the judge to support the verdict: they are *compiled facts* used as *arguments*. As argued in the previous section, we believe that for the causal explanation needed in liability attribution, at least all information available to a judge should enter under consideration. This information typically takes the form of (official) reports by police officers, eyewitness reports, written statements etc., in short: stories told by people.

These stories generally relate of effects in the physical world caused by agents, and are interspersed with views on the motives and intentions behind actions. In a more extreme form, communication actions do not have any significant physical influence; they rather influence the mental state of other

agents, i.e. they are purely *interpersonal*. Communication transfers intentions (motives, plans) and beliefs to other people; it is the most direct manifestation of what is often called interpersonal causation (Hart and Honoré 1985; Lehmann 2003): some agent can provide reasons or draw attention to reasons which influence the conduct of another agent, who causes some event. Interpersonal causation is thus tightly connected to the notions of coercion and authority.

As long as the views expressed in stories are in accord with commonsense, they appear to be acceptable to a reader. Obviously, ambiguities about the causes of actions are well known in legal cases. Nonetheless, actions can only be described in relation to their physical effect: every action minimally involves some movement of the body. Similarly, interpersonal processes are dependent of physical conduits (be it air or phone wires) and are thus subject to the same spatial and temporal bounds. Different perceptions of the same events often correspond on what happened ‘physically’, but diverge on the underlying intentions and motives (as e.g. in a quarrel).

The events and situations we experience are filtered by perceptual mechanisms; i.e. selective, attentional mechanisms, and in particular the application of perceptual and conceptual categories. Verbal reports of these experiences often leave out explicit accounts of perceived causation, because commonsense allows us to reconstruct this relation almost effortlessly. These stories, thus, usually do not contain ‘self evident’ causal relationships, i.e. similar to the original observation of actual events, causation needs to be inferred.

Philosophy maintains two conflicting views on physical causation. The first being Russell’s original view that causation is a secondary, and largely redundant concept (see Section 2.1). Second, philosophy often sees causation as a primary concept, right next to space, time etc. which can be experienced a priori. For our commonsense approach, we understand causation to be secondary, and causal relations are attributed to events we can ‘explain’ by referring to some known process. We do not ‘know’ theoretical causal dependencies, but rather recognise operationalised processes.

Commonsense has a huge number of physical processes in stock, which – despite the fact that humans may easily agree and have consistent interpretations – we will not be able to represent (the famous knowledge-acquisition bottleneck). However, our research is in essence not intended to enable machines to actually *read* reports in legal case documentation, and to identify (legal) responsibility and liability, but rather to render theoretical views on causation and liability in computational terms for the purpose of verifying their consistency and demonstrating their consequences.

In the following sections we will not deal with the question of sufficiency of evidence for establishing legal causation and liability, we will focus

instead on laying the groundwork for a system that will enable us to determine liability. Input to this system are the commonsense stories we have described in this section. These stories are incomplete accounts of some chain of events, resulting in an undesirable situation. Explaining these stories in terms of physical causation is then the first step in the automatic detection of liability and causal relations (i.e. by *DIRECT*).

4. A generic framework for legal reasoning

The causal explanation of stories is embedded in a generic framework for legal reasoning. As any knowledge-based application, it is defined by both an epistemological and an ontological view of the task domain. The former could be seen as the interpretation of the concepts in a domain for a particular purpose. The latter defines those knowledge-structures the system is to reason about. In our case, the epistemological framework is the Functional Ontology of Law, described in (Valente 1995). Though called ‘ontology’, *FOLaw* does not contain traditional ontological categories (Breuker and Hoekstra 2004c). These are provided by the *LRICore* ontology for law: an ontology founded on commonsense principles, see Breuker and Hoekstra (2004a) and Breuker et al. (2004). This section sketches these two components that make up our generic framework for legal reasoning.

4.1. THE FUNCTIONAL ONTOLOGY OF LAW

A general breakdown of the types of knowledge and their dependencies, involved in the legal domain is described in (Valente 1995). His Functional Ontology of Law is an epistemological framework that describes the legal system as an instrument to influence the behaviour of (part(s) of) society, in order to reach certain social goals, i.e. it exists to fulfil a certain *function*. In this sense, the legal system can be viewed as a “social device which reacts to social behaviour, operating in- and on society”.

Essentially, the *FOLaw* describes the path along which a formal account travels as it is under consideration of the legal system. First, the account is described using commonsense terms. As it ‘enters’ the legal system, it is enriched with legal vocabulary, i.e. it is *interpreted*, or legally *explained*. The causal relations between events described in the account are identified. This interpreted causal account is then considered for norm violations, using some body of normative knowledge (rules, legislation, precedents), and liability is determined for agents (in)directly responsible for the violation or some harm. A reactive measure is then taken on the basis of the outcome of this process.

The separation of the detection and establishing of a causal account and the attribution of responsibility is considered rather controversial in the field of AI and Law. In Section 2 we have described why we think this distinction still is very useful for building ontology based legal reasoning systems, and in Section 5 we will discuss a more detailed procedure for causal explanation within this framework.

4.2. THE LRICORE ONTOLOGY

Legal accounts are always based on some abstraction of real-world occurrences, uninterpreted accounts are expressed in the commonsense vocabulary people use in everyday communication. Legal norms, and definitions of legal concepts we commonly find in legislation are expressed using a mixture of these commonsense terms and (domain) specific legal jargon; a heterogeneous abstraction of commonsense descriptions. For this reason, it is our conviction that legal knowledge representation should be strongly based on those concepts which are part of both the legal professionals' and the citizens' understanding of the legal field. Not only does this facilitate mapping between an ontology and legal texts, it enables the direct communication on internal reasoning processes of a legal knowledge-based system to its users.

Reasoning on official reports and accounts therefore requires some common unifying vocabulary: a *core ontology of law*. Generally speaking, a core ontology covers the common vocabulary of a field, e.g. medicine, law, physics and so on (Valente and Breuker 1996; van Heijst et al. 1997). These fields in turn, usually consist of multiple domains such as criminal law, environmental law and private law. A core ontology specifies the common conceptual denominators of a field, i.e. those abstract concepts that are part of all (or an important majority of) domains. In law concepts like norm, role, document, liability, etc. are present in all legal domains, both implicitly and explicitly.

Except for CYC, existing foundational ontologies (e.g. SUMO, Sowa's upper ontology (Sowa 2000) and DOLCE (Gangemi et al. 2002)) do not generally take an explicit commonsense stance. Even in DOLCE, which is based upon human perception, cultural imprints and social conventions, the commonsense perspective is not explicitly developed. Although it does model the world of commonsense, it does not model using the vocabulary and structure of commonsense: it is rather a framework for describing commonsense things, than a commonsense framework. An example of this is its use of meta-properties such as the distinction between endurants and perdurants.

*LRI*Core is represented in OWL DL using the Protégé ontology editor and its OWL Plugin, Figure 1 shows the main categories. More on *LRI*Core is available at http://wiki.leibnizcenter.org/open/index.php/LRI_Core.

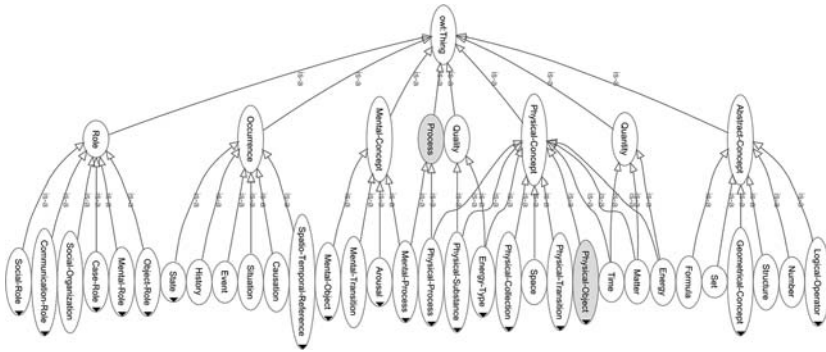


Figure 1. Top two layers of LRICore.

5. Commonsense causation in DIRECT

The notions of causation and responsibility described in the previous section, are operationalised in a system called *DIRECT* (see Figure 2). The input for *DIRECT* is a simple story, typical for legal case descriptions, which consists of unrelated events. It is *DIRECT*'s task to identify causal and intentional relations between these events. *DIRECT* is powered by (extensions of) *LRICore*, i.e. *LRICore* is the knowledgebase we use for interpreting story fragments.

The machine-interpretable representation of a story is an explicit account of what occurred during a certain period of time, i.e. it is expressed in terms of instances (i.e. states) of individuals, and the events that change these individuals. Figure 3 illustrates the relation between story and ontology. We define a *Story* to be a temporally ordered collection of situations. It can be conceived of as a filmstrip or storyboard like sequence of situation models. A *Situation* is a state of affairs: a fixed (spatial) configuration in a collection of states of individuals, during some bounded period of time. Part of these situations, a *State* is the instance (occurrence) of an individual at a certain time and place. In other words, an individual is the identity holder for a set of

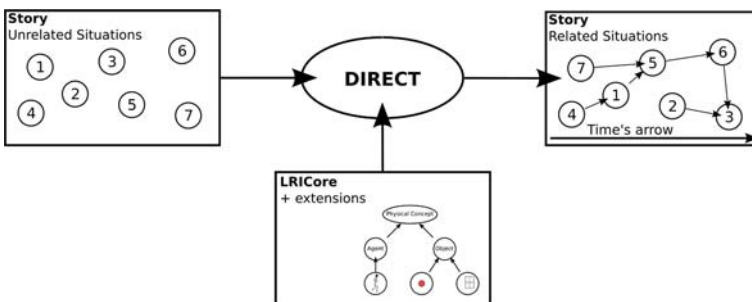


Figure 2. General architecture of DIRECT.

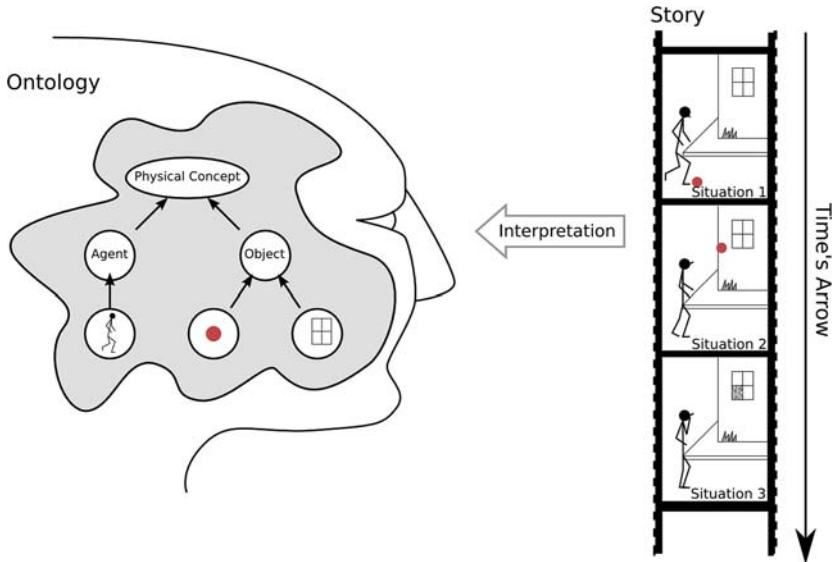


Figure 3. Instances in a story interpreted in terms of an ontology.

states. Since the ontology defines concepts such as e.g. *Person*, *Ball*, *Window*, we are capable of recognising the states in our story as instances of these concepts. An *Event* is then the occurrence of a transition between two situations: at least one individual changes state, and consequently the event brings about a new situation.

To construct a causal explanation within this framework, *DIRECT* interprets these single facts (events and states) in a story by finding the underlying processes which may account for the changes that occur between the situations it describes. The core ontology represents our knowledge of the world; the mapping between instances in a story and terms of the ontology is an interpretation of those instances.

We hold the hypothesis that an account of events and states in terms of processes is both necessary and sufficient for the identification of causal relationships. As explained in Section 3 an operational representation in terms of processes to explain events in stories is much closer to the everyday conception of occurrences than abstract causal dependencies (see also Section 6). In everyday communication we explain events in terms of processes: “Tom kicked the ball, and the window broke” we automatically interpret as “The window broke *because* Tom kicked the ball”. The order of statements in a sentence is interpreted as an ordering in time. The story in Figure 3 is a conjunction of situations; what *DIRECT* does is provide a translation from *and* to *because* in terms of processes.

Commonsense causation is defined as the relation that holds between two events iff there exists some intermediary process(es) that binds the consequent

state of the first event to the antecedent of the second. In general, two types of causation are minimally necessary for establishing legal causation in fact: *physical causation* which describes physical processes and *agent causation* which describes the actions of rational agents. As described in Sections 2 and 3, legal theory, e.g. (Lehmann 2003; Hart and Honoré 1985), introduces two additional forms of causation: *interpersonal causation*, which describes the effect one agent might have on another (e.g. as a consequence of communication), and *negative causation*, representing the connection between an effect and some agent *not* acting (e.g. negligence). As the latter are combinations of the former, we will focus our discussion on the detection of physical causation and our ideas about agent causation.

5.1. PROCESSES AND ACTIONS

Events and states occur at some time and place, i.e. they happen against a four-dimensional canvas of space and time. These coordinates provide us with clues for the discovery of causal relationships: location limits causal propagation. We distinguish between *mental* events and *physical* events, the former constitutes a change in the mental state of a person, and the latter is a change in the physical world. Physical causation occurs when some physical event causes another physical event: i.e. one or more some physical processes can be recognised, which accounts for the events. Physical processes consist of other physical processes or of transitions which are defined by pre- and post-conditions on physical entities, e.g. *Standstill – to – Moving* (see Figure 4) is a physical transition since the restrictions on the *from* and *to* properties include membership of the *Physical – Object* class.

Commonsense covers a wide variety of differing notions of process: from mechanics (movement), radiation of light, heat exchange, etc. to some primitive chemistry and biology (e.g. digestion). Some events are inherently hidden, they do inhibit change, but do not create differences which can be easily detected. The chemical equilibrium, for instance, is important in explaining certain occurrences but does not in itself constitute any discrete change in the state of an object. This is a matter of granularity. By increasing the detail of the description of such processes, the different components (i.e. transitions) of the process can be separated.

Agent causation, or mental causation, is defined as the causal relation between a mental event and a physical event. The recognition of agent causation is an important prerequisite for detecting liability as conduct is either some action, or a prolonged state caused by an action (e.g. the possession of illegal firearms). To detect this kind of causation we need to recognise the physical events which are brought about by the actions of intentional agents.

An *Action* is a special type of physical process which consist of both transitions which operate on only mental objects and of transitions on physical objects: the *decision* to act, and the act *itself*. Transitions with conditions that cross the gap between the physical and the mental do not exist as this would imply that an object can be both mental and physical at the same time. Physical actions are operations by agents on the physical world, i.e. they are changes caused by (intentional) muscle movement. Because of these physical characteristics of actions it is sufficient to know a process to be an action if it was initiated by an agent. Furthermore, effects of actions are local to agents, and do not have a physical predecessor.

Actions differ from physical processes in that these operations are guided by some instigating motive. The action needs to be accompanied by some intention, the will to achieve some goal, the agent's *reason* for acting. This intention should not be confused with legal intent, as the latter is rather a plan-like structure with goals that lie beyond those of a single action. Without intention an act would be 'just' a process. Although knowledge of intention is certainly necessary to be able to distinguish between legally relevant types e.g. voluntary and involuntary actions, negligence, coercion etc. it is of no particular relevance to the detection of the causal chain between the (physical) event brought about by the action and its physical effect. This detection of physical causal relations is the basis for more complex causal discovery, and (in the end) a prerequisite for the attribution of liability.

A *Transition* describes an allowed (possible, meaningful) change in the value of a single property defined on some individual. The restriction on the pre or post condition of a transition can be any complex OWL class definition. An example of a simple transition can be seen in Figure 4.

We define *Process* in a similar fashion; it is a functionally coherent aggregate of one or more transitions and sub-processes. When executed, it consumes or produces resources (e.g. energy, time) and changes individuals, i.e. it produces new states. Processes manifest themselves in stories as the occurrence of one or more events, transitions can only occur as a single event. Figure 5 illustrates how events and states in a story are mapped onto transitions and processes.

Although a standard OWL description classifier is able to infer the type (i.e. transition) of events between simple states, and can infer whether conjunctions of identified transitions belong to a particular process, more complex (structural or spatial) constraints cannot be expressed in OWL. For

```
class Standstill-to-Moving Transition and
                           from some (Physical-Object and (acceleration has zero)) and
                           to some (Physical-Object and (acceleration has plus))
```

Figure 4. The definition of the Standstill-to-Moving transition.

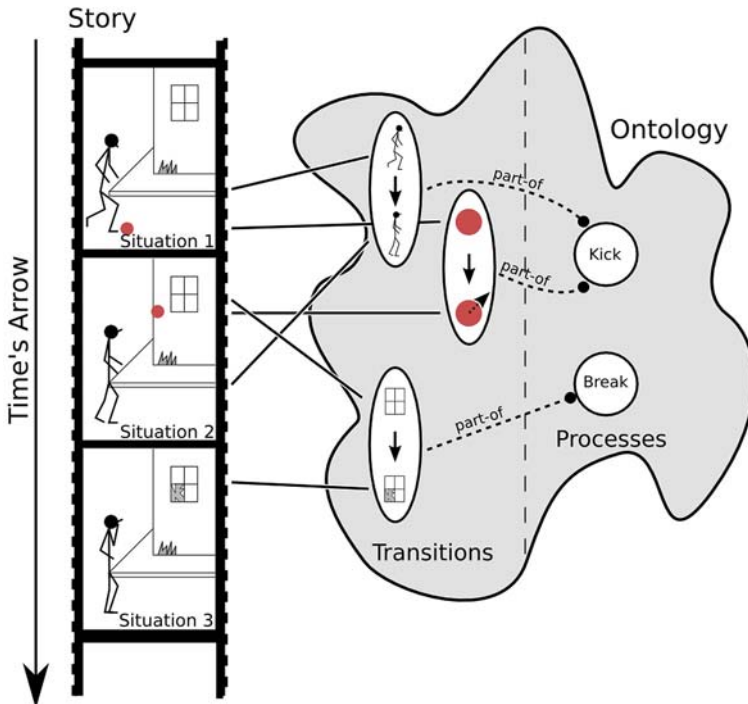


Figure 5. Matching states and events to predefined transitions and processes.

example, we could express that transition *Standstill – to – Moving* can change a property (e.g. *acceleration*) of some physical object on the condition that it is connected to *some* other object, and remains connected to *some* object. But we cannot express that these objects should be *the same*. Even worse, OWL does not even allow us to express that the precondition holds for the same object as the post-condition. For this reason, SWRL rules are specified that constrain the identities of the OWL individuals involved in the execution of processes and transitions.

We currently employ a simple iterative deepening generate-and-test strategy to causal discovery. For sure this process can be optimised using more advanced heuristics, which we are currently investigating. *DIRECT* will allow for generating and testing hypothetical causal explanations using different techniques, and evaluating those techniques. Hypothetical explanations (i.e. chains of instantiated transitions and/or processes) are generated and checked for validity with respect to the ontology using a freely available description classifier (currently Pellet) and SWRL/Jena rules.

The causal analysis of a story of unconnected events, results in an explicit account of what happened, telling us which events were causally related and *why*. We can then use this model to determine whether this enrichment of a case with commonsense causal explanations is sufficient for liability

attribution. In the following section we will discuss how the emphasis on process oriented representation of commonsense causation differs from other approaches AI.

6. Existing approaches

In Artificial Intelligence, causation has been the direct or indirect object of study in a number of specialised areas. Worth mentioning in this respect is the work of (Pearl 2000), which finds its basis in a probabilistic representation of physical causation, more knowledge-intensive approaches to physical causation from the fields of Qualitative Reasoning (QR) and Model Based Reasoning (MBR) (e.g. Horn 1990), ontology-based approaches, and a multitude of formalisms for agent causation (communication) within multi-agent systems.

6.1. FORMAL CAUSAL REASONING

The work by Pearl (2000) stands out in AI as the most comprehensive and explicit modelling of what we may call “computational causality”. The major substance of his work is in probabilistic (Bayesian) modelling of quantifiable dependencies between occurrences, e.g. the use of drugs and their effects on patients. In this sense it fits the tradition in science to enable the identification of causes or causal factors in a well founded, formal way. This work is also relevant for legal reasoning about cases that have a probabilistic basis (data about (co-)occurrences), as for instance in claims about compensation for damage of health suffered due to the consumption of certain industrial products. However, his approach is not particularly well suited for the kind of causation that we have in mind and which (Pearl 2000, Ch. 10) coined “actual cause”: i.e. “an event recognised as responsible for the production of a certain outcome ... Human intuition is extremely keen in detecting and ascertaining this type of causation and hence is considered the key to construct explanations ... and the ultimate criterion (known as “cause in fact”) for determining legal responsibility” (p. 309). Although his view on the conceptual basis for (actual) causation is similar to ours: “Clearly, actual causation requires information beyond that of necessity and sufficiency: the actual process mediating between the cause and the effect must enter into consideration. (p. 309)”, it differs in that for us this consideration is not a mere additional element but primary and sufficient for deciding on causation in fact. His somewhat exploratory work on factual causation is again aimed at a formal, ‘correct’ modelling of causal dependency rather than the commonsense perspective that we want to represent and formalise.

6.2. QUALITATIVE AND MODEL BASED REASONING (QR AND MBR)

Another potential source of inspiration for our work can be found in the representation and reasoning involved in modelling the structure and behaviour of systems (see for a current overview Bredeweg and Struss (2004)). QR started as an approach to commonsense reasoning and was for a while coined as “naive physics” (Hayes 1985) in his influential “second naive physics manifesto” argues that causality is not a “useful, self-contained theory”, but “that it is an umbrella term for a large variety of particular relationships” (p. 19). We agree that causality (or causation) is not a primary term in an ontology of common sense (e.g. it is not part of *LRICore*) and see the variety of relationships as the variety of kinds of processes we know of.

In QR, a model of the structure of a system – e.g. an electronic circuitry – is used to simulate the propagation of behavioural changes through the system (including structural changes of the system itself). This propagation yields behavioural traces of chains of events, which represent the predictions about the behaviour of the system, given some initial state. Events in QR are connected by two kinds of relations: influences and proportionalities. Proportionalities represent definitional or inherent dependencies. For instance, an increase in volume of a substance corresponds linearly with the weight of a substance. Influences are actual causes. Processes are viewed as the origin of causation (Forbus 1984). The ‘scientifically’ correct distinction between dependencies does not completely match commonsense intuitions. Students have great difficulty in applying this distinction and in everyday reasoning it hardly appears to matter.

QR is particularly useful for modelling well known and stable physical structures, i.e. devices, and it is here that model based reasoning – an applied offshoot of QR – finds its applications. Although QR has been used in other domains than physics, e.g. in modelling ecological and social systems, its views are limited to purely *physical causation*, i.e. causation as implied by physical processes. It does not provide a means to model agents, let alone the processes initiated by them.

6.3. AGENT TECHNOLOGY

Agent causation is not entirely explained by causes but by reasons, and in particular the *intention* to perform an action. The notion of intention, although more recent, has given rise to as much philosophical controversy as causation. In AI the notion has been operationalised to model actions, and in particular communication between artificial agents. The languages by which these agents communicate typically provide constructs for representing

belief-states, actions, plans, data etc. The most prominent example being the belief, desire, intention (BDI) model, cf. Georgeff et al. (1999). The BDI model is a representation of the inner workings of agents (including people), and can contribute to a better understanding of how intention influences action. Nonetheless, there are a number of reasons why agent technology cannot be used (initially) for detecting commonsense causation. First of all, agent technology is concerned with simulation, real-time behaviour of agents, where liability attribution requires postdiction. Where artificial agents do planning, we need to do plan recognition. Furthermore, agent technology does not adhere to a particular theory on physical causation. As we have argued in the previous section, the bulk of actions undertaken by people directly correspond to an intention and can therefore be primarily explained using physical constructs.

6.4. ONTOLOGY

The ontological approach to causation in fact described in Lehmann et al. (2004) and Lehmann and Gangemi (2006), defines causal dependencies within the framework of the DOLCE ontology, see e.g. Gangemi et al. (2002). Causal relations relate very simple (types of) events that change a single aspect of a single object. Between these (sets of) simple events, three existential dependencies are identified: *structural*, *causality* and *circumstantial* dependencies. Physical causation is defined as the relation that holds between two individual events that satisfy both the causality and the circumstantial constraints. The framework allows for the classification of some description of a number of events as an instance of *physical causation in fact*.

The ontology contains no theory of the (physical) world; i.e. descriptions are not expressed using domain knowledge. For instance, the semantics of a state is not represented intensionally, but only through its name (e.g. *being-wounded*), it contains no description of *what it means* to be in that state.

The lack of such descriptions makes this approach less useful for explaining the existence of a causal relation. We argue that by defining causation solely at the abstract level of dependencies, important information which can explain causal relations in stories and descriptions is lost. Furthermore a system built on such a representation is less flexible with respect to practical use in which accounts are generally under-specified.

7. Discussion and future work

We presented a process based approach for ontology-based discovery of causal relations in stories, and shown how it can form the groundwork for

the automatic determination of liability. We have argued that existing approaches from philosophy and AI do not represent actual causation in a way which is in accordance with commonsense, whereas the stories that lie at the heart of legal cases are entirely of a commonsense nature. Furthermore, we have discussed the importance of commonsense causal reasoning for establishing legal causation and determining the liable party on the basis of official reports. Legal theoretical discussions on the role of causation in liability attribution obfuscate the difference between these two notions. This impedes a proper understanding of what knowledge needs to be added to stories to make attribution possible.

The commonsense stance of *LRICore* distinguishes it from other established foundational ontologies such as *DOLCE* and *SUMO*. Based on psychological evidence, it is closer to our everyday interpretation of reality than other more science, or philosophy-based approaches. It is our conviction that legal knowledge representation requires such psychologically plausible commonsense conceptualisation of the world. Together with *FOLaw*, it provides a generic framework for legal reasoning which has a strong epistemological, ontological and legal theoretical grounding.

We adopt a definition of causation which is very close to the commonsense interpretation of cause and effect. It is based on a non-probabilistic ontological conceptual representation of agent- and physical causation expressed using commonly understood processes. This representation is used by *DIRECT* to construct causal explanations from partial accounts – such as stories – of temporally ordered series of events, using predefined processes.

We are currently experimenting with stories that contain simple physical causal relations and agents without elaborate intentions. As we have discussed in Sections 5.1 and 6, we believe these relations are often sufficient for basic connections between action and effect. In law, an example are relatively simple torts, such as causing damage, where intention is used to determine the *degree* of liability (see also Section 2.2), and the act itself is enough to commit the tort.

We will gradually extend *DIRECT* to be able to handle a set of more complex stories which include more elaborate notions such as intentions, plans, coercion, persuasion and other legally relevant elements. This will involve the addition of more legal and domain-dependent concept definitions, such as processes, actions and objects to *LRICore*.

For sure the work presented here only scratches the surface of the complexity involved in not only legal reasoning but also commonsense physical reasoning. However, it forms the premise for more elaborate commonsense legal reasoning, and enables us to formulate and test our hypothesis that a process-based causal explanation of events and states is both necessary and sufficient for liability attribution.

Notes

- ¹ DIRECT:Discovery of REsponsibility and CausalTY.
- ² Later he admitted that this was maybe too strong an opinion.
- ³ Although (Lehmann et al. 2005) lists only 5 combinations, we rather adhere to the full list of Hart and Honoré (1985).
- ⁴ More specifically, in law anything causal is in the end caused by people: legal causation always originates from some conduct.
- ⁵ See <http://www.cyc.com>
- ⁶ Suggested Upper Merged Ontology; <http://ontology.teknowledge.com>
- ⁷ Descriptive Ontology for Linguistic and Cognitive Engineering; <http://www.loa-cnr.it/DOLCE.html>
- ⁸ See <http://protege.stanford.edu>
- ⁹ The notion of ‘individual’ in OWL is rather confusing. Because of open-world semantics, an OWL individual is not an identity holder (it is rather an instance). In the following, we will use the word individual in its more traditional meaning.
- ¹⁰ The term mental causation is rather confusing; we have argued in amongst others (Breuker and Hoekstra 2004a) that thinking is purely *intentional* and therefore does not follow the laws of cause and effect.
- ¹¹ The example uses Manchester OWL Syntax, see: http://www.co-ode.org/resources/reference/manchester_syntax/
- ¹² See <http://www.mindswap.org/2003/pellet>.
- ¹³ See <http://jena.sourceforge.net> and <http://www.w3.org/Submission/SWRL/>
- ¹⁴ If the substance is not a contained gas.

References

- Ashley, K. (1990). Modeling Legal Argument. MIT Press.
- Aleven, V. (1997). Teaching Case Based Argumentation Through an Example and Models. Ph.D. thesis, University of Pittsburgh, Pittsburgh, PA, USA.
- Bredeweg, B. and Struss, P. (2004). Current Topics in Qualitative Reasoning, *AI Magazine*, 24, 13–16, special issue on Qualitative Reasoning.
- Breuker, J. and Hoekstra, R. (2004a). Core Concepts of Law: Taking Common-sense Seriously. In: *Proceedings of Formal Ontologies in Information Systems*, FOIS-2004. IOS-Press.
- Breuker, J. and Hoekstra, R. (2004b). DIRECT: Ontology based Discovery of Responsibility and Causality in Legal Case Descriptions. In: Gordon, T. (ed.), *Legal Knowledge and Information Systems. Jurix 2004: The Seventeenth Annual Conference*. Amsterdam, 59–68, IOS Press.
- Breuker, J. and Hoekstra, R. (2004c). Epistemology and Ontology in Core Ontologies: FOLaw and LRI-Core, Two Core Ontologies for Law. In: *Proceedings of EKAW Workshop on Core ontologies*. <http://sunsite.informatik.rwthachen.de/Publications/CEUR-WS/>, CEUR.
- Breuker, J., Valente, A., and Winkels, R. (2006). Use and Reuse of Legal Ontologies in Knowledge Engineering and Information Management. *Artificial Intelligence and Law* 12:241–277.
- Clark H. (1996) Using Language. Cambridge University Press.
- Davidson, D. (2001). *Essays on Actions and Events*, 2nd Edition. Oxford University Press: Oxford.

- Forbus, K. D. (1984). Qualitative Process Theory, *Artificial Intelligence*, 24, 85–168.
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening Ontologies with DOLCE. In Gomez-Perez, A. and Benjamins, V. (eds.), *Proceedings of the EKAW-2002*, 166–181. Springer.
- Georgeff, M., Pell, B., Pollack, M., Tambe, M. and Wooldridge, M. (1999). The Belief-Desire-Intention Model of Agency: In Müller, J., Singh, M. P. and Rao, A. S. (eds.), *Proceedings of the 5th International Workshop on Intelligent Agents V: Agent Theories, Architectures, and Languages (ATAL-98)*, Vol. 1555. Springer-Verlag: Heidelberg, 1–10.
- Hart, H. and Honoré, T. (1985). *Causation in the Law*, 2nd edition. Oxford University Press: New York.
- Hayes, P. J. (1985). The Second Naive Physics Manifesto. In Hobbs J.R., Moore R.C. (eds) *Formal Theories of the Common Sense World*, 1–36. Norwood, Ablex Publishing Corporation.
- Hobbs, J. R. and Moore, R. C. (eds.) (1985). *Formal Theories of the Common Sense World*. Ablex Publishing Company: Norwood.
- Hoekstra, R. and Breuker, J. (2005). Processes as Causal Glue in a Framework for Ontology-Based Responsibility Attribution: In: *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques*, in conjunction with ICAIL 2005. Bologna, Italy.
- Horn, W. (ed.) (1990) *Causal AI Models: Steps Toward Applications*. Hemisphere Publishing Corporation.
- Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation, Representation and Mind*. MIT Press: Cambridge, Massachusetts.
- Lehmann, J. (2003). *Causation in Artificial Intelligence and Law: A Modelling Approach*. Ph.D. thesis, University of Amsterdam.
- Lehmann, J., Borgo, S., Masolo, C. and Gangemi, A. (2004). Causality and Causation in DOLCE: In Varzi, A. C. and Vieu, L. (eds.), *Formal Ontology in Information Systems Proceedings of the International Conference FOIS 2004*. Torino, 273–284, IOS Press.
- Lehmann, J., Breuker, J. and Brouwer, B. (2005). CausatiO^{NT}: Modeling Causation in AI& Law: In Benjamins, V., Casanovas, P., Breuker, J. and Gangemi, A. (eds.), *Law and the Semantic Web, Lecture Notes in Artificial Intelligence (3369)*. Springer-Verlag: Berlin, Heidelberg, 77–96.
- Lehmann, J. and Gangemi, A. (2006). An Ontology of Physical Causation as a Basis for Assessing Causation in Fact and Attributing Legal Responsibility. *Artificial Intelligence and Law. Special Issue on Legal Ontologies and Artificial Intelligence Techniques*.
- Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press: Cambridge.
- Schaffer, J. (2003). The Metaphysics of Causation/*Stanford Encyclopedia of Philosophy*.
- Sowa, J. F. (2000). *Knowledge Representation: Logical Philosophical, and Computational Foundations*. Brooks Cole Publishing Co: Pacific Grove, CA.
- Valente, A. (1995). *Legal Knowledge Engineering: A Modelling Approach*. IOS Press: The Netherlands, Amsterdam.
- Valente, A. and Breuker, J. (1996) Towards Principled Core Ontologies. In: *Proceedings of the Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop*. Banff, Alberta, Canada.
- van Heijst, G., Schreiber, A. T. and Wielinga, B. (1997). Using Explicit Ontologies for KBS Development, *International Journal of Human-Computer Studies*, 46(2/3): 183–292.