LEGAL THEORY CONSTRUCTION
Jaap Hage and Giovanni Sartor

1 Introduction

The idea of legal theory construction has gained substantial popularity the last years, often under the heading of coherence theories of the law. In this paper we try to develop the beginnings of a method for theory construction.

In a series of papers, culminating in his book *Law's Empire*, Dworkin has developed an intuitively attractive picture of legal theory construction. This theory, called *The Model of Principles*, recognises three stages in constructing the law.[1] The first stage, the so-called pre-interpretative stage, consists of a preliminary identification of the rules, standards, and (generalised) decisions that make up the law. In this connection one may think of an inventory of the rules and standards that can be found in statutes, cases, and doctrinal literature. The second, interpretative, stage consists of an identification of the principles (in this connection including values and policies) that underlie (in the sense of explain), or are part of the legal phenomena identified in the first stage. The rules etc. identified in the first stage are to be seen as means to realise the principles identified in the second stage.

The purpose of the third, reforming, stage is to formulate (relevant parts of) the set of rules, including (generalised) decisions of cases, that best realises the principles identified in the second stage.

2 The elements of a legal theory

According to the Model of Principles, the test for a theory consists of a check whether the decisions and rules of the theory on the one hand, and the principles of the theory on the other hand mutually support each other. The rules and decisions should realize the principles, while the principles should explain the rules and cases.

The expression 'legal principle' is used in several senses. In one sense, legal principles can be seen as a kind of goals. Because there are more kinds of goals within the law than only legal principles, while their functions in a theory of the law is similar, we stop discussing the role of principles, and write about the role of goals instead. A goal in the sense we use it here is a generic state of affairs, a possible but not necessarily actual fact, that is to be realised as much as possible. Goals in this broad sense include values (e.g. justice, a healthy environment), human rights (e.g. protection of physical integrity and of privacy), and policies (e.g. full employment and a good health service).

Goals provide us with guidance when we must decide what to do. An action that promotes a goal should pro tanto be performed, while an action that detracts from a goal should pro tanto be avoided. Similarly, a rule that promotes a goal should pro tanto be adopted, while a rule that detracts from a goal should pro tanto be rejected.

Not all goals are equally important. Arguably, justice is more important for the law than the policy to increase the number of robins in Utopia Central Park. The promotion of more important goals should pro tanto be preferred to the promotion of less important goals. It is not uncommon that a solution both promotes one goal and detracts from another goal. If the one goal is more important than the other goal, pro tanto the more important goal determines whether the rule should be adopted or rejected. That is, if the rule promotes the more

[1] *Law's Empire*, p. 65f.

important goal, it should pro tanto be adopted, and if it detracts from the more important goal it should pro tanto be rejected.

The promotion of a goal by a rule is not an all or nothing matter. Sometimes a rule manages to achieve a goal completely, while at other times the contribution is only minimal. If a rule promotes a goal to a large extent, this is pro tanto a stronger reason to adopt this rule than if it only minimally promotes the goal. Analogously, if a rule detracts from a goal to a large extent, this is pro tanto a stronger reason to reject this rule than if it only minimally detracts from the goal.[2]

Most of what was written about the relation between rules and goals also holds for the relation between cases (including their outcomes) and goals.

3 Mutual adjustment

A good legal theory should be coherent in the sense that the cases and rules contained in it are the best means to realise the goals, while the goals provide the best possible explanation of the cases and rules. Such a theory can be achieved through mutual adjustment of cases and rules on the one hand, and goals and their relative importance on the other hand. However, there may be some cases and rules that are not so easy to explain by means of the given set of goals, for instance because they seem to detract from the goals, rather than to promote them. Moreover, it is possible that other rules and other solutions for some of the cases would even better promote the goals. One solution for this lack of coherence would be to disregard the offending cases and to replace the less than ideal rules by better ones. Another solution is to adopt different goals, that match the given rules and cases better. By ignoring some goals, or by adopting new ones, or by a combination of both approaches, it would be possible to obtain a better match between rules and cases on the one hand and the adopted goals on the other hand.

Instead of replacing some of the goals in the theory, it is also possible to tinker with the relative importance of the goals. Assume that some solution for a case promotes one goal and detracts from another goal. Then the theory is pro tanto coherent if the promoted goal is deemed the more important one. By changing the relative priorities of goals this can be achieved and thus it is possible to make the existing cases and rules better match the existing goals.

To facilitate the following discussion, we will introduce two abbreviations for parts of a full theory of the law. We will use the term ‘case part’ for the cases with their solutions and the rules included in the theory, and we will use the term ‘goal part’ to denote the part of a legal theory that consists of both goals in the broad sense and their relative importance. In this way we can write about the mutual adjustment of the case part and the goal part of a theory if we mean to talk about the mutual adaptation of on the one hand rules and cases with their solutions and on the other hand goals and their relative importance.

Given a set of goals, the solution for a particular case will promote some (zero or more) of these goals, detract from some other goals, and will be neutral with regard to the rest. Every goal to which the solution for this case contributes provides a reason for (the rightness) of the solution for this case, while every goal from which the solution detracts provides a reason against this solution. Whether the solution for a case is right all

[2] The line of reasoning exposed here is elaborated formally in Hage 2000 and 2001.

things considered depends on the balance of these reasons. If the reasons why the solution is right (the pro-reasons) outweigh the reasons why the solution is wrong (the con-reasons), the solution is *right*. If the balance of reasons goes the other direction, the solution is *wrong*. If the reasons pro and con a conclusion are more or less in balance, the solution of the case is *indifferent*.

To make a decision about the rightness of the solution for a case, we must balance reasons and most often this will just be a matter of decision making. Such decisions about the relative weight of (sets of) reasons are expressed by what we will call *weighing knowledge.* Weighing knowledge is also part of a theory of the law, and we will include it in the goal part (as opposed to the case part) of the theory.

The solution for a case may be right in the sense that the reasons pleading for it outweigh the reasons pleading against it, but that does not guarantee that there is no better solution for that case. One might think of a set of possible solutions for a case of a certain kind. Some of these solutions are right, some of them indifferent, and some of them wrong. But not all the right solutions are equally good. Some of them are better than other ones. Now we will call the solution for a case *outstanding* if there is no better possible solution.[3]

The distinction between right solutions for a case and outstanding ones has a counterpart on the negative side. A solution for a case may be not only wrong but even so wrong that there is no comparable worse solution. Such solutions that have no comparable worse alternative will be called 'terrible' Terrible solutions compare to merely wrong ones just like outstanding solutions compare to right ones.

Given a set of goals, a legal theory can be improved by replacing terrible solutions by merely wrong solutions, wrong solutions by indifferent ones, indifferent solutions by right ones, and by replacing merely right solutions by outstanding solutions. A theory can also be improved by improving the goals. Given a fixed cases part, improvement is accomplished by making the actual solutions for cases that were originally classified as terrible to be merely wrong, wrong solutions into at least indifferent, indifferent solutions to be at least right and by making merely right solutions to be outstanding. Improvements in the goal part of a theory can take two forms. One is by making modifications in the set of goals by adding new goals or removing old ones. The other is by making changes in the relative importance of the goals, or in the weighing knowledge.

4 Improving the goal part of a theory

To illustrate how the goal part of a theory can be improved to accommodate a set of pre-existing cases, we use three cases from Dutch Tort Law concerning dangerous situations.[4]

4.1 The cases

THE CELLAR-FLAP CASE[5]

Defendant left the cellar-flap of his café open to facilitate the transport of drinks. Plaintiff dropped into the cellar and suffered damages. The court applied some version of the Learned Hand formula. If the costs of avoiding or taking away the danger are less than the damages when the danger is realised multiplied by the chance that the danger is realised, it is negligent

[3] Probably some solutions are not well comparable, so it is not necessarily the case that for every set of competing solutions there is one outstanding solution, or one set of outstanding solutions that are equally good. So a set of competing solutions may have more than one outstanding solutions that are not comparable.

[4] The style of analysis used for this purpose is based upon the model of Sartor (2002) and Bench-Capon and Sartor (forthcoming).

[5] HR 5-11-1965, NJ 1966, 136.

to let the danger continue to exist. This led to the decision that defendant's behaviour was negligent and that plaintiff could recover the damages.

THE STUMBLING BAKER[6]

While delivering bread, a baker stumbled over a piece of rope that was stretched across the path to the front door of a house. Two children who lived there, four and five years old, saw the bakes approach the rope, but did not remove the rope, or warn the baker. (The children had not stretched the rope.) The baker sued the (parents of the) children for negligence. The court decided that a duty of care only exists for those who have realised the presence of the danger, unless there is a special duty of care, based upon, for instance, a special relationship toward the victim or with the place of the danger. Defendant's behaviour was not deemed negligent, and plaintiff could not recover the damages.

CAUSTIC SODA[7]

Employees of a community centre placed a bag with household refuse along the street, in order to be taken away by the cleansing department. The bag held a container with a liquid, which was, unknown to the employees, caustic soda. A cleaner put the bag into the dustcart, and due to the cart's mechanism, part of the caustic soda was swept into his face, as a consequence of which he became blind. The cleaner sued the operator of the community centre for the damages.

Even though the employees of the community centre were unaware that the liquid in the bad was caustic soda, their behaviour was held to be negligent, because the court assumed a duty of care not to place a container with an unknown liquid in it, only protected by a cardboard box and a plastic bag, along the street to be taken away by the cleansing department, unless one has good reasons to assume that the liquid is not dangerous, or one keeps the bag under control and warns those who want to handle the bag for its possibly dangerous contents. Plaintiff could therefore recover his damages.

4.2 Basic notions

The basic idea in improving the goal part of a theory is that of constructing a theory which can in the best (in the most coherent) way integrate (explain, combine, connect) the various input elements which are provided to the lawyer. These input elements include previously decided cases and existing written law.

OUTCOMES

We may first characterize our theory construction exercise by specifying its outcome. The outcome would be the decision that is at issue in the examples we have just described, that is whether a person who caused damages is required to compensate the person who suffered them. Therefore we consider only two possible outcomes: C (compensation is granted), and ~C, (no compensation).

FACTORS

Decided cases are essential building blocks of legal theories. Initially, a case can be seen as a set of facts, together with a decision (an outcome)

[6] HR 22-11-1974, NJ 1975, 149.
[7] HR 8-1-1982, NJ 1982, 614.

made on the basis of those facts. But this has not typically been found to be the most useful way of representing cases for case based reasoning purposes. Case facts are in themselves neutral and not necessarily relevant to the outcome. Explanation of outcomes has usually therefore been in terms of *factors*.[8] Factors are an abstraction from the facts, in that a given factor may be held to be present in a case on the basis of several different fact situations, and moreover, factors are taken to strengthen the case for one or other of the parties to the dispute. In our analysis, we will have factors pleading for compensation and factors pleading against compensation.

For example, the fact that the author of the damage intended to cause it, or behaved negligently, or could cheaply have prevented the damage, are all factors favouring the conclusion that there should be compensation. On the other hand, the fact that the damage was not foreseen, or that all due care was taken, or that the expenses required to prevent damage were higher than the expected damage itself, or that the damaged person could avoid the damage at less costs than the author could, are factors pleading against compensation.

From a factor-based perspective, a case can be seen as constituted by a set of reasons (factors) pleading for and against its conclusion. For example in the Cellar-Flap case, the fact that the defendant (denoted as δ) caused a danger, when acting in his own interests may be seen as a factor for compensation. Another factor pleading for compensation might be that the defendant could have prevented the danger at little costs. In our examples, we assume to have the following factors:

_ δCreatedDanger: defendant created a danger
_ δCouldCheaplyPrevent: defendant could cheaply prevent the damage
_ δDidNotKnowDanger: defendant did not know the danger
_ δCouldNotKnowDanger: defendant could not know the danger

Clearly all such factors are reasons favouring certain conclusions. Let us represent that by a set of factor links, where each factor link is expressed by an arrow. An upward going arrow ↑ denotes that the factor favours the following outcome (in casu compensation), while a downward going arrow ↓ denotes that the factor disfavours the outcome. Note that in our representation favouring one of the two possible conclusions (C) amounts to disfavouring the alternative conclusion (~C). In our example we assume the following factor-links:

_ δCreatedDanger ↑ C
_ δCouldCheaplyPrevent ↑ C
_ δDidNotKnowDanger ↑ ~C
_ δCouldNotKnowDanger ↑ ~C

GOALS

Goals explain why certain factors promote a certain outcome, i.e. why they are factors rather than meaningless facts. We take it that a factor favours an outcome, because deciding for that outcome in a case where that factor is present promotes some goal, which the legal system promotes. Note that one factor may be appealed to on the basis of different goals. For example, the creation of a dangerous situation (δCreatedDanger) is a factor for compensation, since imposing liability on persons who create dangers may lead people to be more careful, and in this way to reduction of damage. Another goal is fairness, i.e., one's gains should not be to the loss of other people, or those who profit from a certain activity should bear its costs.

[8] Ashley 1991.

If serious damages can be prevented against modest costs, the value of fairness makes the fact that the damages could be prevented cheaply into a factor pleading for compensation.

So another component for a theory of the law would be a list of the goals that are included in the legal system. Let us assume that for our purposes the relevant values are:

_ Reduction of damage To avoid damages where reasonably possible
_ Responsibility To have people answer for their acts and their consequences
_ Fairness To have people pay for costs which are related to their gains

Obviously, one may extend or refine this list, and this is also a very important feature of legal decision-making, at least at the highest level (for example, in debates in front of a constitutional court). However, we may assume that in most cases the list of values will be available to the lawyer as something given.

TELEOLOGICAL LINKS

Let us now move to the next element of our theories, the connection between goals and factor-links. The basic idea is that a factor-link is supported by the fact that it promotes some goal. For example, we may say that creation of danger leads to compensation, because by linking creation of danger to compensation goals are promoted, such as Reduction of damage and Fairness.

In general, we may say that a teleological link $(\alpha \uparrow \gamma) \Uparrow G$ expresses the idea that the general adoption of factor link $\alpha \uparrow \gamma$ (its being used by legal agents as a standard for their reasoning and practice) would advance the achievement of goal G.[9] So for example, to say that linking the creation of danger to compensation promotes reduction of damage, we write

$(\delta\text{CreatedDanger} \uparrow C) \Uparrow$ Reduction of damage

Similarly, we can say that this link promotes fairness, and therefore we can write:

$(\delta\text{CreatedDanger} \uparrow C) \Uparrow$ Fairness

Information of this type provides the teleological background of the lawyer's work.

4.3 Heuristics for theory construction

Now we will indicate how theory construction can take place by providing a set of theory constructors, i.e., of operators that use the available legal materials for constructing a legal theory. The constructors are only mentioned. There is no room to illustrate the operation of all of them.

The first constructor is simply the *input constructor*, which takes and element (a case, a factor, a factor link, a goal, a teleological link, a rule) and adds it to the theory

The second constructor is what we may call *rule-constructor*. This constructor takes a set of factors pointing in the same direction and links them into a rule having the same outcome. The sense of this operation is that the rule's antecedent is assumed to provide a normally sufficient condition for its conclusion (unless an exception or a conflict with another rule exists). In contrast, a factor is only a contributory condition: it favours a certain outcome, but is in itself not sufficient to bring the outcome about.

[9] A goal-background may include many other components, such as a specification of the relative importance of the goals, of their relations (achieving some goals may impact positively or negatively on others), etc. Here, however, the goal-background will be limited to elementary teleological links. Moreover, for simplicity's sake, we assume that single goals do not interfere with each other, and that all factor-links promoting the same goal do that to the same degree.

For example, we could transform the factor-link δDidNotKnowDanger ↑ ~C (the fact that one did not know that a danger was present favour's the conclusion that one should not be liable) into the rule

δDidNotKnowDanger ⇒ ~C

Rules constructed out of factor-links inherit all teleological links with supported those factor-links. This corresponds to a third constructor, the *teleological constructor*, which assigns to a rule a list of all values which were promoted by the factors out of which the rule was build. So, if we have the factors

(α ↑ γ) ⇑ V1
(β ↑ γ) ⇑ V2

the rule we construct out of these factors will promote both values:

(α and β ⇒ γ) ⇑ [V1, V2]

The fourth theory constructor, which we call *rule-preference-from-value-preference,* consists in introducing preferences between rules on the basis of preferences between values. The assumption is that rules promoting more important values are stronger than those promoting less important values.

The fifth theory constructor, which we call *rule-broadening,* consists in introducing a more general rule on the basis of a more specific one, already contained in the theory. So, if the theory contains α & β ⇒ γ one may heuristically expand the theory with α ⇒ γ or β ⇒ γ.

The sixth theory constructor, which we call *rule-preference-from-case,* consists in introducing preferences between rules when these preferences help explaining the precedents.

4.4 Constructing a theory for tort liability

Let us now see how we can build a theory by using the input knowledge and some of the theory constructors we described in previous pages. First of all, as lawyers working in a civil law jurisdiction, we assume that our background includes the statutory rule which establishes that in general if one culpably performs an illegal action and thereby causes damage to another, then one needs to compensate it. Let us express this rule as

IllegalBehaviour & Culpability & CausedDamage ⇒ C

So, using the input constructor, we add the rule to our theory. Then we may enter some factors. For example, in relation to the Cellar-flap case, one may enter factors δCreatedDanger and δCouldCheaplyPrevent, the corresponding factor links, and teleological links. Then one may want to build directly a rule combining those factors

δCreatedDanger & δCouldCheaplyPrevent ⇒ C

Such a rule will allow us to explain the decision in the Cellar-flap case and also in the Caustic Soda case. However, it will fail to provide an explanation for the Stumbling Baker case. Note, however, that, at least in civil law systems, one needs to connect factors to statutory rules. This means that the two factors above should be seen as factors contributing to the realization of the pre-conditions of the above-mentioned rule, rather then as independently favoring a certain conclusion. So one could add the following factors to the theory:

δCreatedDanger ↑ IllegalBehaviour
δCouldCheaplyPrevent ↑ Culpability

They would be associated to the teleological links

(δCreatedDanger ↑ IllegalBehaviour) ⇑ [Responsibility, Fairness]

(δCouldCheaplyPrevent ↑ Culpability) ⇑ Efficiency

Therefore, the fact that one created a danger for another can be seen as leading to the conclusion that one has acted illegally (since one has endangered other people's rights). Similarly, the fact that damage could be cheaply avoided can be seen as factor leading to the conclusion that one has acted culpably.

The resulting theory succeeds in explaining cases Cellar-flap and Caustic Soda (which we have therefore added to the theory). However, it still fails to explain the Stumbling Baker case. The intuitive reason why in this last case no responsibility was assumed is that the children were so young that they did not realize that there was a danger. This suggests the introduction of the factor-link

δDidNotKnowDanger ↑ ~C,

and the corresponding teleological link

(δDidNotKnowDanger ↑ ~C) ⇑ Fairness.

There are two ways to make it explain the decision in the Stumbling Baker case (no compensation). One is to add the rule δDidNotKnowDanger ⇒ ~C to the theory and add the weighing knowledge that this rule outweighs the main rule:

(δDidNotKnowDanger ⇒ ~C) > (Tort & Culpability & CausedDamage ⇒ C)

Attractive as this theory may seem, it has one obvious disadvantage, namely that it does not explain the Caustic Soda case. The employees of the community center were not aware of the danger they created, and yet compensation was granted. Somehow, if both the Caustic Soda and the Stumbling Baker case are to be maintained in our theory, we need to distinguish them. The crucial difference would be that in the Caustic Soda case, the employees should have thought about the danger they created (although in fact they did not), while the children in the case of the Stumbling Baker could not be expected to think of the present danger.

This suggests the introduction of the factor

δCouldNotKnowDanger ↑ ~Culpability,

and the corresponding factor link

(δCouldNotKnowDanger ↑ ~Culpability) ⇑ Responsibility.

To explain both the Caustic Soda case and the Stumbling Baker case, the following weighing information is needed[10]:

[δDidNotKnowDanger, δCouldNotKnowDanger] >
  [δCouldCheaplyPrevent, δCreatedDanger]
[δCouldCheaplyPrevent, δCreatedDanger] > [δDidNotKnowDanger]

With these factors and this weighing information added to the theory, we can derive Culpability in the Caustic Soda case, because there the employees were not aware of the danger, but could have been aware of it, and we cannot derive Culpability in the Stumbling Baker Case, because the children did not know nor could know about the danger. Moreover, these factors and the weighing information allow the introduction of the rule:

δDidNotKnowDanger & δCouldNotKnowDanger ⇒ ~Culpability

[10] Actually other constructions are possible too, but to keep our exposition relatively simple we ignore them.

5 Conclusion

We started from the assumption that legal theory construction can be seen as the process of mutually adapting two parts of a theory of the law. We have given a theoretical reconstruction of how tinkering with goals, their relative importance, and with cases, may lead to a coherent theory. Finally we went into more detail by showing how it is possible to use heuristics for theory construction to develop a theory that explains a number of cases from Dutch tort law.

Because of limitations of length, our exposition had to be incomplete. The interested reader is referred to the already mentioned papers by (Bench-Capon and) Sartor and by Hage to obtain more details.

References

Ashley, K.D. (1991). Reasoning with cases and hypotheticals in HYPO. *International Journal of Man-Machine Studies*, vol. 34, p. 753-796.

Bench-Capon, T.J.M. and G. Sartor (forthcoming), *A Model of Legal Reasoning with Cases Incorporating Theories and Values*, accepted for publication in the Artificial Intelligence Journal.

Dworkin, R. (1986). *Laws Empire*, Fontana, London.

Hage, J.C. (2000). Goal-based Theory Evaluation. J. Breuker e.a. (eds.), *Proceedings of the Thirteenth Jurix Conference*, IOS Press, Amsterdam, p. 59-72.

Hage, J.C. (2001). Formalizing legal coherence. *Proceedings of the 8th International Conference on Artificial Intelligence and Law*, ACM, New York, p. 22-31.

Sartor, G. (2002). Teleological Arguments and Theory Based Dialectics. *Artificial Intelligence and Law* 10 (1-3), p. 95-112.