Action and Procedure in Reasoning

van Benthem, J.F.A.K.

Publication date
2002

Document Version
Author accepted manuscript

Published in
The Dynamics of Judicial Proof

Link to publication

Citation for published version (APA):
van Benthem, J. F. A. K. (2002). Action and Procedure in Reasoning. In *The Dynamics of Judicial Proof* (pp. 243-259). Physica Verlag.

ACTION AND PROCEDURE IN REASONING

Johan van Benthem

Institute for Logic, Language and Computation, Universiteit van Amsterdam
Center for the Study of Language and Information, Stanford University
http://www.turing.wins.uva.nl/~johan/

Cardozo Law Review 22, 1575–1593.

1 The dynamic turn

Meaningful comparisons between ‘logic’ and ‘legal reasoning’ must evolve with their relata. In this paper, we explain some basics of 'logical dynamics', a current procedure-oriented view of reasoning and other cognitive tasks, using games as a model for many-agent interaction. Against this background, we speculate about possible new connections between logical dynamics and legal reasoning.

2 Law versus logic?

Many authors have claimed that there are systematic differences between logical reasoning and juridical reasoning. Here is a quote from the sixties, found in van Apeldoorn's well-known textbook for beginning Dutch lawyers:

> *"In law, as opposed to logic, if the premises*
> *of an argument are bad, so is the conclusion."*

In logic, hereditary sin works the other way around. Bad conclusions of valid arguments infect one or more premises, but bad premises may lead to perfectly true conclusions (witness many cases in the history of science). Anyway, this difference is only apparent. It makes little sense to use premises that are known to be false, even to logicians, while lawyers may appreciate standard logical refutation. Perhaps a more subtle difference was noted by Toulmin in “The Uses of Argument”:

> *“Logical proof turns on abstract mathematical form,*
> *legal reasoning turns on procedure (the 'formalities')."*

In other words, logic deals with static forms and Platonic relationships of implication between these, while the law must deal with ‘rationality in action’. This judgment was true, by and large, when Toulmin wrote his famous book in the 50s, but we shall se that matters are much more diverse these days. One indication of this is the lively modern literature at the interface of law, artificial intelligence and logic, documented e.g., in the proceedings of this conference (Nijboer, Tillers, Prakken). One can view AI as the 'strong arm' of logic, making its static forms 'work'. But in this paper, we want to discuss dynamic tendencies in logic itself.

Polemics aside, what specific features of legal reasoning might set it apart from logical reasoning in general? Generally speaking, these will reflect its task of delivering rationality under real-time constraints. Here are some positive prejudices to this effect from an outsider (often a good source of biased positive views):

- the crucial importance of good **procedure and timing**
- the role of **different parties**, not just one lonely Thinker
- inevitably **limited resources** for the reasoning process
- the aim for, not absolute, but **reasonable certainty**

Law has been so succesful in delivering 'real-time rationality' under these constraints that some philosophers see it as one main pillar of our western culture, which should not be 'reduced' to the other (being mathematics). But in recent years, logicians, too, are becoming increasingly interested in general cognitive mechanisms of **reasoning and information flow**, partly under the influence of computer science and AI. In this movement essential aspects are again: <u>resources</u> (notably, computational ones of time and effort), <u>preferences</u> (of both single and multiple agents), <u>procedure</u> and timing. This dynamic turn has noticeable influences on logical theory and practice, which makes traditional comparisons with legal reasoning less conclusive. (Note 1.)

3 Logical dynamics

Traditionally, logic is taken to be about implications between static propositions that are true or false about the world. The emphasis is then on 'truth conditions' for these propositions, and on sound and complete proof rules that manipulate these. To-day's 'dynamic turn' shifts the perspective towards general processes changing information states about the world – of which reasoning is an important example, but not the only one. When using language to communicate, people modify their own and other people's information, often in subtle ways. In a modern slogan:

Statements are actions – and therefore natural language is not a description language, but a kind of programming language for cognition.

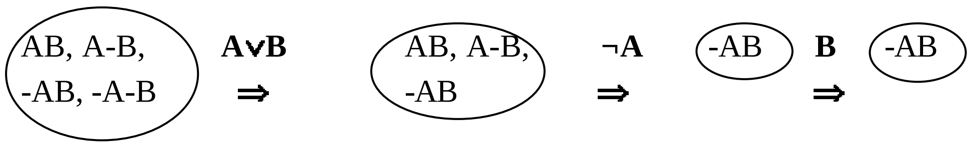
What we need to investigate then are the 'update conditions' for statements. Likewise, reasoning is a stepwise activity, whose success involves judicious use of timing and resources. Note that indeed the very terms "statement", and "reasoning" are ambiguous between 'activities' and the 'products' of those activities! Here is a warm-up example for the rest of this paper, demonstrating the new way of thinking. To a first approximation, one person's

information state is the set of all relevant possibilities that she entertains; *updates are actions that change this information set.*

<u>Propositional inference as an update process</u>

Consider two atomic propositions A, B. Their true/false combinations give 4 candidates for the real state of affairs. Information about the latter now comes in via utterances triggering updates, restricting this set. In the limit, only one option remains, and we know the real facts. Here is a video-strip illustrating successive information states and updates associated with the valid propositional inference

from two premises $A \vee B$ *and* $\neg A$ *to the conclusion* B *:*



The two premise updates add information. We can even measure their precise strength, via the numbers of possibilities removed: the second is more informative than the first. By contrast, the conclusion is a *fixed point*: updating with it adds no further information, the state remains the same. This kind of update mechanism explains the solution process in simple logic puzzles, or games like 'Master Mind'. Of course, there is much more to this than can be discussed here. (Note 2.)

The preceding example is also misleading, however, in that most realistic communicative settings games involve updating of *many-agent information states*. For instance, when a question is asked and an answer is given, two agents learn much more than a simple factual update about the content of the question. Under normal circumstances, the questioner conveys that he does not know the answer, and that he expects the answerer to know. The answerer achieves 'common knowledge' of the answer, which means that both parties now know that they both know the answer, and that the other knows this, etcetera. Such 'logical overtones' matter. Information about other agents' knowledge or ignorance can be crucial to further action. Also, 'who knew what when' is crucial to establishing innocence or guilt. Humans are remarkably good at keeping track of such subtleties. Compare the case where everyone knows that your partner is unfaithful. This is a nuisance, but one you can live down. But if the bad situation is common knowledge, it may be time to draw your gun, restore your injured honour (and get a good lawyer…). Updating information states for many agents is a subtle process, that we will not pursue here. (Note 3.) (Note 4.) Instead, we now turn to a many-agent model with useful concrete intuitions that will be the focus for the rest of this paper.

4 Games as a model for many-agent logical dynamics

A desire to improve one's skills in winning arguments and debates is a motivation for many students of logic (whether well-founded or not…), and it has indeed been a source of inspiration for logic since Antiquity. There even seems to be a plausible intuition of a 'valid inference' as a *guaranteed winning strategy in debate*. Even so, games have never been a major recognized paradigm for inference in logic, such as 'semantic validity' or 'provability'. But these days, this minor current is turning into a full-fledged research program using concrete games as a procedural model for a variety of logical tasks: semantic evaluation, model construction, proof, comparison, communication. (Pioneers in this movement have been Lorenzen, Ehrenfeucht, and Hintikka. By now, there are many strands and directions in this field.) At the same time, contacts are increasing between logicians and other communities interested in games: philosophers, linguists, computer scientists, and of course, game-theorists in economists, starting from the tradition of von Neumann– Morgenstern and Nash. Shared concerns across these communities are rational deliberation, communication, decision and action for groups of agents. These involve typical game features like

- interaction between players, their moves and turns,
- players' rights and duties, intentions, resources,
- explaining behaviour via strategies, and stable behaviour in terms of 'strategic equilibria', where the players have no interest in deviating.

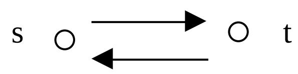
(Note 5.) Games are not just a metaphor here (as they were in the old days of Huizinga's 'homo ludens', Wittgenstein's 'language games' or NN's 'games people play'): but a concrete method of modelling. We will show this for logical tasks in the coming three sections, while also pointing out some broader implications for cognition and legal reasoning in particular.

5 Disputes about facts

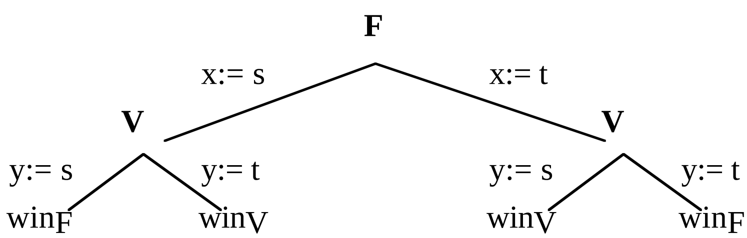
5.1 Evaluation games Here is the simplest logic game. Suppose two parties disagree about a the truth of a statement in some situation under discussion: *Verifier* **V** claims it is true, *Falsifier* **F** that it is false. Evaluation games describe their moves of defense and attack – with a schedule driven by the assertion at issue:

atoms	*test* to determine who wins
disjunction	**V**'s *choice*
conjunction	**F**'s *choice*
negation	*role switch* between the two players
existential quantifiers	let **V** *pick an object*
universal quantifiers	do the same for **F**

E.g., consider two objects s, t , with R = {<s, t>, <t, s>}:



Here is the game of perfect information for the logical formula $\forall x\ \exists y\ Rxy$ (say, the assertion that "everyone has an enemy") in this situation – pictured as a tree of possible moves, with the scheduling read from top to bottom:



Falsifier starts, Verifier must respond. (Another way of thinking about this particular game is as 'matching pennies'.) Clearly, there are 4 possible plays here, with 2 wins for each player. Yet, the players are not evenly matched. Evidently, **V** is the lucky one: after all, she is defending the truth of the matter…

5.2 Truth, winning strategies, and determinacy The point is that **V** can always win, by choosing the appropriate response to whatever **F** plays: she has a *winning strategy*, a pattern of behaviour that will guarantee a certain outcome, in this case: 'winning". By contrast, **F** has no winning strategy: **V** may always thwart him. But neither does he have a 'losing strategy': he cannot force V to finish him off… Thus, players' powers of determining the outcomes of a game may be quite different. Here is the fundamental connection between the relevant key notions of logic and game theory underlying our simple observation:

Proposition A statement A is *true* if and only if Verifier has a *winning strategy* for its evaluation game.

The proof of this is not hard (a simple induction on formulas), but it goes one step beyond the intentions of this paper. But even so, we see some important aspects of the dynamics of two interacting logical agents, which do not arise on standard semantic accounts of truth at all. There is 'necessary effort' and 'mutual power':

(1) *Fairness* guarantees a win with optimal patterns of behaviour. It is not a panacea for lazy people. Verifier might play stupidly, and lose the game. Logic is about what you can win, not what you cannot help winning.

(2) *Joint powers*. Even if one player has a winning strategy, outcomes may still depend on the other, who might decide e.g., where to make his final stand! (Think of the losing party at least chosing the final battlefield with honor.) In general, then, both parties are involved in determining the final result.

Both points are important to applications of the model. In particular, it is a typical game-theoretic thought that all players are stakeholders in the final outcome.

In addition to these general points, there are also interesting technical aspects to the above connection. In particular, logical laws now acquire game-theoretic import! E.g., consider the classical logical law of *Excluded Middle* $A \vee \neg A$, often taken to be a triviality. The fact that Verifier has a winning strategy for this statement in every model means she can choose to play either A as Verifier, or $\neg A$ as Verifier (i.e.: A as Falsifier), and then still have a winning strategy for the remainder. But this is another way of stating a well-known notion from game theory:

Excluded Middle $A \vee \neg A$ expresses *determinacy* of evaluation games:
i.e., one of the two players must always have a winning strategy.

This is just the tip of an iceberg. The preceding observation holds because of a few very general features of our evaluation games. In the first decade of this century, Zermelo (yes, the famous set-theorist) already proved that

All *zero-sum* two-player games that have *finite depth* are determined.

Zermelo was mainly concerned with finite-depth games like chess, where his result implies that one of the two players has a non-losing strategy. (It is not yet known whether this is Black or White. Abstract existence results for strategies may still defy the best current computers.) By contrast, current evaluation games for more complex languages than first-order predicate logic may go on *forever*, and determinacy becomes a more subtle issue. This technical issue again has a clear practical repercussion. Short-term disputes and debates may have finite depth – if only by stipulation of some kind of limit to repetitions (as happens in chess). But when we move to a higher level, and consider macro-games like 'language use' (the 'operating system' of cognition), or indeed 'law' (the operating system of justice), infinite patterns and strategies become essential.

5.4 Points of relevance to legal reasoning The preceding analysis shows that one meaningful type of legal episode, two-party disputes about facts, can be modelled as games with precise rules and winning conventions. Different roles are necessary here, players are both stakeholders (antagonistic in our particular example, but cooperative in others), and the truth of the matter is reflected in their available strategies. These strategies correspond to their rational behaviours, which they might display in court when interacting with the other party.

Next, we turn from games about the question *"Did* things actually happen this way?" to logical games about "*Could* things have happened this way?". This might be the object of a lawyer's strategy, pointing out scenarios that are consistent with the innocence of her client. The logical issue here is no longer *Model Checking*, but *Satisfiability*. Given one or more assertions A, but no specific situation, does there exist a model making all these assertions true? Answering the Satisfiability question is equivalent to answering a question, not of truth, but *validity*:

"Does A have a model?" is equivalent to "Is the negation ¬A valid?".

6 Disputes about consistency

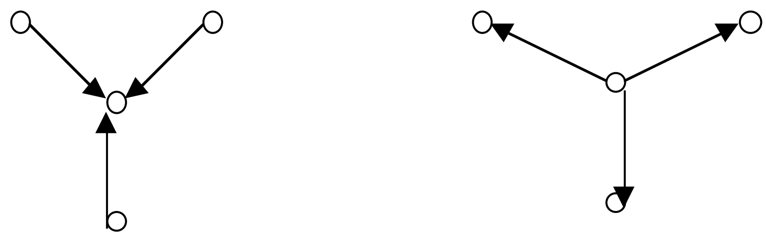
6.1 Model construction games Consider two parties disagreeing about the consistency of a story, without an external source for checking. Logical *construction games* model this situation. We only describe their outline, because our further points can be understood without details. (Note 6.) One player is *Builder*, who claims that a model exists. Her moves introduce objects and postulate facts concerning the model in statu nascendi. The other player is *Critic*, who searches for inconsistencies in this process, issues challenges, and makes sure every requirement is met. The *schedule* of moves in construction games is much freer than in evaluation games. At any stage, a number of assertions may have to be made true and/or false, and players can pick any one of these. Another difference with the preceding evaluation games are the *asymmetries* between players. E.g., if an existential formula $\exists x\, A(x)$ has to be made true, Builder must introduce a new object that can serve as a witness, while, if it is false, Critic only has to issue challenges to the effect that no existing object satisfies A . Also, Builder has no special interest in meeting every requirement (we are not talking about construction of physical buildings with possible legal action afterwards), but Critic does.

For a concrete example, the reader may want to think about Builder's moves and Critic's challenges in a police station where people ponder the case of

<u>The Gang of Four</u>
Somewhat unorthodox (but let us hope, admissible) detective work has yielded the following facts about the power structure in a certain gang with four members:

- *some member either commands or is commanded by everyone else*
- *if you are commanded by someone, you do not command anyone*

Can there be gangs satisfying this description? Builder must construct a situation that do (pictures are a good means of displaying these), while Critic is a colleague trying to take this construction apart. Indeed, introducing objects that satisfy these two assertions, and that can stand up to the obvious challenges corresponding to the quantifiers “everyone else” and “you”, “someone”, can be done in two ways:



These two patterns correspond to the two different winning strategies which Builder in fact possesses. Notice that we are interested in *both* structures here, but a lawyer trying to show that the assertions are consistent would only need one. Conversely, if a lawyer argued that this description is inconsistent – so her client could not have belonged to any such organisation – just one picture would do.

But now, a third report comes in (this is really a kind of logic police force):

- *everyone takes orders from exactly one person.*

This time Critic has a winning strategy, being in essence a formal proof that the three assertions together form a contradiction. There are indeed various strategies for this purpose, depending how one ‘pinpoints’ the contradiction.

6.2 Strategies, models and proofs

The main theoretical result about construction games is this

<u>Proposition</u> Builder has a *winning strategy* in the construction game
if and only if the initial set of assertions is *consistent*.

A more detailed analysis of the proof reveals that, as with evaluation games,

Each construction game is *determined*:
either Builder or Critic has a winning strategy.

The reason is not as simple as for evaluation games, since construction games can have *infinite* branches, corresponding to the construction of infinite models that satisfy the given assertions. This is needed to take care of predicate-logical validity in its entirety, although many practical applications will lead to finite game trees. (Infinitely large criminal organisations seem rare, unless one thinks of Crime as one party in a never-ending game against Justice.) More concretely, there is again a lot of interesting fine-structure to the interactions between players in a construction game. In particular, different strategies carry important information:

> Builder's winning srategies (if any) are the different *models*,
> Critic's winning strategies (if any) are inconsistency *proofs*.

Also, despite this antagonistic description, the very asymmetry of players' moves shows that players do not necessarily have conflicting interests. We can view Critic just as well as someone who helps Builder schedule his tasks in a way which ensures solidity of the construction – and heartily rejoices in her winning strategy. It is often quite convenient to delegate one's coordinating tasks to other parties.

A final logical point of interest is this. Generally speaking, the model evaluation task and the model construction task are of *different complexit*y (in terms of the order of magnitude of the time steps required):

> Construction games have *higher computational complexity* than evaluation games. Given a finite model, the evaluation games take a polynomial number p(n) of time steps (where n measures the length of the input assertions). Construction games may take forever – and even if they do terminate, they may take any (exponential) amount of time f(n) .

Thus, checking for consistency is more complex than checking for truth. A related logical point is this. *Lying* is a more sophisticated skill than telling the truth – and typically, it comes many years later in a child's cognitive development.

In actual argumentation, and legal settings, the two tasks may occur intertwined. As philosophers point out, we check the facts (by perception, or other means) when we can without too much effort, and 'fantasize' in all other cases. This can be modelled by *mixed ga*mes, which would presumably be closer to courtroom reality.

7 Dialogues, debates, and procedural bias

Historically the oldest logic games are 'Lorenzen dialogues'. These are related to the preceding construction games, but they present some interesting further points, that seem closer to practical argumentation. Again, we refer to the literature for the precise definition of these games, stating merely that we now have

> a formal debate between *Proponent* and *Opponent*
> arguing for c.q. against some proposition at issue.

Dialogue games have *logical rules* breaking down formulas much like those in our earlier games. E.g., to defend a disjunction, Proponent has to chose a disjunct and defend that – while for a conjunction, Opponent chooses the conjunct. To defend an existential quantifier, Proponent must choose a 'witness' object, while Opponent chooses a 'challenge' object if a universal quantifier is defended by Proponent. Winning or losing occurs when one player has run out of legitimate things to say (according to the dialogue conventions): the first to encounter this impasse loses. (This is a well-known convention in many games – though not, e.g., in checkers, where it generates a draw. But in debates, having no further moves is indeed often interpreted as a player's defeat.) Given all these stipulations, the structure of the eventual 'adequacy assertion' for dialogue games remains as we have seen before:

> *validity of a formula A amounts to Proponent's having*
> *a winning strategy – viewed in more detail, Proponent's*
> *winning strategies correspond precisely to logical proofs.*

A bit more generally, one can start the dialogue game with a situation where Proponent makes a claim (the conclusion), while Opponent has already granted some assertions (the premises) which proponent can exploit in further play.

But the main point we wish to stress in this section is one which comes out only in the details of such argumentation games. The notion of 'validity' in the preceding assertion is not as stable as one might expect! For, in addition to the non-controversial logical rules regulating a dialogue, one must agree on what may be called *procedural conventions*. These involve crucial features of *scheduling*:

> *who wins what depends essentially on procedure*:
> in particular, timing and rights of defense or attack.

Example: Classical versus intuitionistic logic

Consider Excluded Middle $A \vee \neg A$, taken for granted so far. If Proponent is to win a dialogue game for this, she must be allowed to switch defenses. Without that, she has to make one choice (left or right), but there is no guarantee of a further win. If she chooses A, and Opponent attacks that, she has no response. If she chooses $\neg A$, and Opponent attacks by asserting A, she has no winning response either. (We forego some technical details of Lorenzen dialogues at this point.) But, if she can switch defenses, she can first defend $\neg A$, forcing Opponent to assert A in order to attack that, and then switch her defense of the initial $A \vee \neg A$ to A, exploiting the fact that Opponent himself has already granted this.

Is being able to answer the same attack twice a justified procedural right? Lorenzen himself did not think so, and opted for intuitionistic, instead of classical logic. And indeed, Proponent's shifty defense of Excluded Middle does not really suggest great moral fibre… In any case, the above general point will have become clear: procedure determines logical powers. Is this a legally relevant point? I am not sure, but consider this situation. The evidence on the table consists of three assertions

$$\neg(A\&B) \qquad \neg A \rightarrow C \qquad \neg B \rightarrow C$$

Say, your client cannot have been at locations A and B at the same time. If she was not at point A she is guilty of misdemeanor C. If she was not at B, she is guilty of the same misdemeanor C. Does it follow that she is guilty of C? Yes, if your opponent is allowed to argue according to classical logic. No, if the stricter proof rules of intuitionistic logic are applied. For then, the first premise does not imply the disjunction of $\neg A$ and $\neg B$, as you may lack the evidence to determine which one obtains, and hence you may be at a loss when challenged to produce a choice at some stage of the argumentation game. Think of castigating your opponents in court for being unable to come up with a precise location where she was not…

Different legal procedures can have systematic effects on outcomes of trials – at least if one believes the Proceedings of this conference (see [ref.]), where to my great interest and edification, guilty persons are advised to try their luck in US courts, and innocent persons in European ones... Startling outcome relativities are well-known for formal procedures such as voting schemes in social choice theory, even though its effects have not been fully taken in by practising politicians and decision makers. It is at least interesting that similar dependencies – especially that on the almost inevitable practical curtailing of repetition rights for attacks and defenses in debate – can have systematic logical consequences, too.

This concludes our discussion of logic games and their import to legal reasoning. But this picture would be seriously incomplete without at least mention of some further themes that arise in game theory. These add more realism to the game perspective on cognitive activities: of which we will consider three main aspects.

8 Preferences, equilibrium, and fair outcomes

One crucial aspect of games has remained implicit so far. Like deontic reasoning, games refer essentially to players' *preferences between different outcomes*. With just winning and losing, we have a degenerate case of this: presumably, players prefer situations where they win over those where they lose. But in general, there may be much more finer-grained preferences between outcomes, either in the form of numerical valuations, or in the form of comparative preferences between them. Players will seek maximal gain and minimal loss in these terms, and choose their strategies accordingly. This preference pattern is the basis for the fundamental game-theoretic notion of *Nash Equilibrium*, an acceptable pattern of behaviour for all participants, which can be considered ‘stable’ and justified. More technically,

> A Nash equilibrium profile is a choice of strategy for each player such that the resulting outcome cannot be improved by anyone switching his strategy (while the others keep theirs fixed).

In our determined logic games, Nash outcomes consist of a winning strategy for one player, and any strategy for the other. Of course, other players may still have reasons for choosing their strategies. If they prefer some sites for defeat to others, the equilibrium will be more subtle. Equilibria are the game-theorist’s favourite tool for explaining stable rationally justified behaviour. In particular, if there is just one Nash equilibrium, a game will have a *value*, which is what players would get if this equilibrium is played. This is a good candidate for a fair outcome. This notion may be generalized by means of probabilistic considerations in case there is more than one equilibrium. (This happens, e.g., in the notorious Prisoner’s Dilemma – perhaps not surprisingly, a case with a juridical setting of sorts.) (Note 7.)

A legal counterpart to the notion of fair value of a game might be an assessment of the parties' chances, and their preferences between various outcomes – while equilibrium pay-offs would correspond to some kind of imposed setllement. Conversely, negotiated settlement is a juridical concept that might find further applications in game theory, or even logic when time constraints become important.

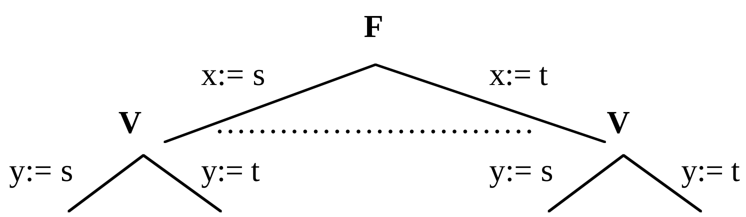
9 Games with imperfect information

The dynamics of playing games is rooted in what may be called 'future ignorance'. We do not know precisely which move our opponent is going to make, and therefore, our strategy must be prepared for eventualities. Logic games have this uncertainty, even though they have perfect information in all other respects. At each stage, players know exactly where they are in the game tree, including all moves that have been played by them and their opponents. But in real games, there may also be 'past' or 'present' ignorance. In a card game, I do not know exactly which cards you were dealt in the first move. A game like "Cluedo" does not permit me to observe which card you are showing to your neighbour. Or, in Prisoner's Dilemma, I cannot see what you are doing right now in the other prison cell… In such *games of imperfect information*, the notion of a strategy (suitably adapted to cope with my current uncertainties) still makes sense – but the situation does get more complicated. In particular these games may be *non-determined*: neither player has a winning strategy. Outcomes will then depend on *cleverness plus chance.*

Here is a concrete example of such a game, derived from a recently proposed variation on evaluation games. Consider the earlier game of Section 5 : $\forall x\ \exists y\ Rxy$. Now change this to

$$\forall x\ \exists y/x\ \ Rxy$$

where Verifier, in her response to Falsifier's initial move, does not know (failed to notice, or forgot) which object was chosen. In the game tree, this is indicated by a dotted line, indicating the two positions which Verifier cannot distinguish:



win_F win_V win_V win_F

Mow, neither **V** nor **F** has a winning strategy. **F** does not have one for the same reason as before: whatever he does, **V** might make the right response, and win. But also, **V** does not have a winning strategy now, because there is no uniform recommendation which she can use in both cases across the dotted line. Choosing s is winning in one case but losing in the other, and so is choosing t .

Is this second type of ignorance relevant to legal procedure? It depends on how it arises. If a player cannot win because of imperfect information about moves made by others (say, what did the opposing party tell the judge?), this may seem unfair, and in conflict with full disclosure. But certain moves by other parties in court may indeed be private to some extent, and in that case, imperfect information is an inevitable fact of life. If the present or past ignorance comes from defective memory, however or other personal handicaps, this seems irrelevant. (That is, unless we wish to compensate some legal actors for their stupidity – the way we sometimes take pity on unworthy opponents in a game.) If legal games have precise boundaries on the 'imperfect information' which they tolerate, this might by itself be an interesting defining feature setting them apart from games in general.

In any case, as in the preceding section, even non-determined games of imperfect information can have a *value*, representing what players can reasonably hope to achieve if the game is played many times. So, even in this game-theoretic scenario, there is a radical option of

> *computing the merits of the case beforehand,*
> *and then pay/punish both parties accordingly....*

10 Repeated games

One noteworthy feature of our logic games which we have not emphasized so far, is their 'one-shot character'. We cannot play all possible runs, which gives the actual sequence of events its drama. Say, Verifier won one run of evaluation game. Does this prove the assertion was true? No, as we have not tested whether she had a winning strategy, reacting succesfully to *every* play by Falsifier. So, did she win by accident or for deeper reasons? Here is where Falsifier has an interesting role to play. It is his interest to offer the best possible counter-play, in order to maximize his chances of winning – but in doing so, he also offers the best possible guarantee that Verifier really has a winning strategy! This one-shot character seems characteristic of law-suits, or soccer matches, and other types of ‘decisive event’. Nevertheless, one must realize that it does not give complete certainty. (Note 8.)

In game theory, *repeated games* are essential for achieving fairness and strategic equilibrium... Repeated games reveal more of players' strategies, and they give players a fair chance for changing some of the actions they took (perhaps in some randomized fashion). Thus, players will get their fair value in the long run. Of course, this would take too much time for realistic decision making – even when the separate games themselves are short. In legal procedure, some sort of repetitions are allowed (appeals), but they do not seem to have the same spirit. Again, the better way of dealing with this need for fairness need would be by proper adjudication. Nevertheless, repeated litigation can be surprising in its own right. Instead of proper analysis, here is an old anecdote from the logic textbooks.

The sophist Corax had taught a pupil Euathlos: the fee would be the first money his student made in winning a law-suit. But Euathlos never entered court. So, Corax brought a case against him, arguing:

> *“Either I win, and you have to pay me (by the verdict),*
> *or I lose, and you also have to pay, by our contract.”*

But the obviously well-taught Euathlos produced a ‘counter-dilemma’:

> *“Either I lose (and I need not pay, by our contract),*
> *or I win, and need not pay because of the verdict.”*

The logical point of the anecdote is just the clever use of dilemma and counter-dilemma in arguing for one’s position. As for the legal point, a lawyer once told me this case should be easy to solve. The student wins the first case, since he had not promised his teacher to engage in lawsuits (there is no breach of faith). But then, the teacher must start a *second* lawsuit to collect his fee, and he will win that. (Note 9.) Having two lawsuits never came up in logical discussion of this puzzle!

11 The law as a game

11.1 General considerations The claim in this paper is not that legal events *are* games in some deep objective sense – even though there are some game-like features to what one sees happening in court. The main claim is rather that it makes sense to *analyze* legal events as games, as a means of bringing out some salient characteristics – for reasons analogous to those making game-theoretic analysis of logic illuminating. There is no unique way of doing this. One can think of single trials as games, but also of types of legal activity (say, the functioning of a court) as games in a broader sense. For an analogy, think of the distinction between the specific programs you are running on your computer versus the operating system: a macro-program enabling the former events to take place at all. Moreover, the contact can go both ways. One can take notions and results from logic games or general games – and then suggest legal counterparts. This has been the main thrust of our paper. We have merely made some suggestions here and there, and more case studies would be required to prove the utility of this perspective. (Note 10.)

But things become more interesting if one also moves in the opposite direction, looking at existing legal procedures, and then tries to extract their game-theoretic and logical import. Legal practice contains several ideas that seem of general interest in games, while legal games might also have special distinguishing features that set them apart as a natural subclass of games in general. It is not the purpose of this paper to explore this in detail, but here are some points that come to mind. First consider the *effects of having more than two players*, not just the contestants, but also lawyers, judges, expert witnesses, etc. A much-studied issue in game theory is powers of *coalitions* of players to achieve outcomes that are beyond each member's individual powers. This fits well with current logical trends in studying collective information, group obligation, etc. More specifically, what happens to logic games, or general games when we add 'juridical players', such as *judges* or *referees*?

Perhaps the most decisive peculiarity of legal games is their *reasoning with bounded resources*, obeying constraints on time and effort. Disputes cannot take forever, judgments have to be arrived at in effectively bounded time. This requires not just *decidability*, but indeed much *lower complexity* than in the usual systems that logicians can afford (in theory). One can study what things keep legal debate functioning within these constraints. This is precisely where the connection with Artificial Intelligence has already provided valuable experience. But it would be of interest to see if this is just computational 'implementation', or whether there are features of general fundamental interest that can inform new logical theorizing. (Note 11.) In particular, resource bounds and methods for enforcing them, e.g. in terms of 'feasible strategies', have been studied in game theory, but these may not be quite the ones that drive legal practice. (Note 12.)

12 Conclusions

Reasoning and many other logical tasks (evaluation, model construction, and others) can be cast as games. This leads to new dynamic views on logic as a many-agent theory of cognitive activity, with a shift in basic notions and results. The resulting paradigm invites fresh comparisons with legal procedure. Through this comparison, one also gets into broader connections between game theory and legal reasoning. Our emphasis here has not been on specific systems that effect this junction, but rather on general issues that seem worth pondering. If these contacts work well, they should work in several directions: first, a better logical understanding of 'legal games', but also, ideas from legal practice may enrich logic games, and games in general. This idea of mutual influence is not new by itself, but we may have added new twists. (Note 13.) Finally a warning seems in order. In my experience succesful interdisciplinary work seldom consists in solving problems of one field by using methods from another. It is rather that members of two intellectual communities meet, fall in love, and produce *common offspring*, in the form of *new questions*, and new types of result. In the long run, the joint venture of having children produces the stronger bond than temporary adaptation of lovers' individual habits…

13 Notes

(1)

There are many further analogies between basic issues in computation/cognition and in law. We mention just a few for perspective: (a) rule subsumption vs pattern recognition, (b) worst vs average case performance, (c) avoiding errors of two types (false positives, false negatives), (d) protocols for achieving secrecy, etc.

(2)

For many consequences of this dynamic view on inference, see van Benthem 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford. Note that reasoning still makes sense in the update setting. It provides a 'red thread' of significant assertions through successive updates, as may be seen in realistic problem solving. This is one instance of the general issue how abstract 'information' is turned into concrete 'knowledge'. Moreover, the strict order dependence of premise updates may be unrealistic. Consider the following information: $A \rightarrow \neg B$, $B \vee C$, A . Most people would first combine the third premise with the first, and then use the second to arrive at the facts $A, \neg B, C$. On the other hand, real-time argumentation makes such choices irrevokable, which is another form of order-dependence.

(3)

J. Gerbrandy, 1998, *Bisimulations on Planet Kripke*, ILLC, A'dam.

A. Baltag, 1999, 'A Logic of Communication', CWI-ILLC, A'dam.

(4)

Many other notions are naturally subject to updating. Consider agents' preferences. These, too, may change because of incoming information – and logical calculi performing 'upgrades' have been proposed by a.o. Spohn and Veltman.

(5)

A recent survey of the logic/games interface: J. van Benthem, *"Logic and Games"*, electronic course notes, Stanford and Amsterdam, 1999. See also the home page 'Logic and Games in Amsterdam', http://www.cwi.nl/~pauly/games.html.

(6)

More precisely, construction games may be represented by the well-known logical technique of *semantic tableaus*, viewed this time as 'dynamic objects'.

(7)

This section only scratches the surface of a complex interaction between game theory and logic. There are many further topics of investigation here. Preferences in games also suggest *deontic dynamics*. One might make preferences themselves an issue for gaming, providing mechanisms for changing them.

(8)

In computer science, one has intermediate cases, where a game would be played a sufficient number of times (i.e., a sufficient number of branches of the full game tree is traversed) to make it *highly plausible* that Verifier has a winning strategy. The most famous algorithm for achieving optimal performance in one-shot situations comes from a judicial setting, however, viz. Cake Cutting. This seems to derive from old Germanic procedures in dividing the loot of a raid. One party divides, the other gets the first choice – as in still visible in the Dutch expression "kiezen of delen" ("do you wish to choose, or divide up?").

(9)

In Amsterdam, Ron Allen gave a more substantial legal analysis, reproduced here from a private communication: "The contract entered into explicitly calls for the fee to "be the first money his student made in winning a law-suit." The suit does not call for a fee if his student wins a suit; it calls for a fee if he makes money winning a suit. When the teacher sues the student, whether the student "wins" or "loses", the student will not make money; therefore, no payment would be due under the contractual provision. Since no payment would be due under the contractual provision no matter how the lawsuit against the student comes out, obviously the lawsuit has no basis and will be dismissed. This result is unfair only if the student somehow misled the teacher. For example, perhaps the student was only interested in a learning about the law, but never intended to practice it. In order to get a free legal education, however, perhaps he feigned an interest in practice, inculcating the belief in the teacher that the student intended to practice, thus inducing the teacher to enter into this contract. Well, the law handles this as well. It is called "fraud in the inducement." If the teacher can prove that there was such fraud, he can recover his damages. And in American courts, maybe punitive damages as well."

(10)

One case was Josephson’s lecture on legal abduction procedures at the Amsterdam conference. H is an abductive conclusion from D if (a) H implies D (together with some background theory), (b) D are correct data, (c) there is no 'better' hypothesis H’ which also derives D . Unpacking this in a logic game between ‘Proposer’ and ‘Critic’, one gets moves corresponding to the three clauses: (a) attack the inference from H to D: i.e., be Builder in a construction game for {H, ¬D}, (b) attack at least one of D, (c) using he quantifier form of "being best": attack (c) by presenting some H’ for which you claim that it also derives D, and that it is better. Proposer can then attack either conjunct of this, again entirely via the rules for a standard logic game. The latter may be cast as a game for checking if some proposed hypothesis H is really a best explanation for a given data D . Some further interesting points in Josephson’s account. (a) One of the options for ‘Critic’ vs ‘Proposer’ is to derive a false consequence from H. This seems to correspond to a further requirement that "H has to be true", or at least "consistent with what is known". (b) Proposer can dispose of a whole bunch of alternative hypotheses at once. This is not needed in our game: it would tell him to reveal more of his strategy than is warranted. (But mentioning a lot of potential points for your opponent, even if you do *not* have a strong refutation is a well-known rhetorical trick. At least, your opponent incurs the odium of saying something 'predictable'.) (c) The requirement for a defense lawyer is that she should produce an *alternative explanation*, after the prosecution has come up with explanation H (normally, the guilt of the accused). This seems stronger than what would be minimally required: viz. showing the evidence to be consistent with the negation of H . (d) One commentator described the judge’s task as choosing between alternative 'stories' in which the data D ‘fit’. Are these like models, hypotheses, or a mixture of both ways of thinking?

(11)

What logicians already know is that restrictions to various fine-structure formats of assertion, and accompanying ‘lightweight calculi’ can improve performance in consistency checking and proof search dramatically. It might be of interest to see whether these correspond to anything in legal reasoning.

(12)

Winning strategies may be too costly to execute, so we must sometimes settle for less. This can be modelled by assigning costs to actions in a game tree, and then computing optimal strategies through the tree given initial resources of players.

(13)

There have been suggestions for ‘science courts’, where legal-style debate would be used to get best current opinions on issues that have been under scientific debate for a very long time – and where some temporary resolution would be useful, e.g. when preparing funding decisions.

14 References

A. Baltag, 1999, ‘A Logic of Communication’, CWI, Amsterdam.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

J. van Benthem, 1999, Logic and Games, lecture notes, Stanford and Amsterdam, http://www.turing.wins.uva.nl/~johan/teaching/.

J. Gerbrandy, 1998, *Bisimulations on Planet Kripke*, Ph.D. thesis, Institute for Logic, Language and Computation, University of Amsterdam.

H. Prakken & G. Sartor, 1999, 'Argument-based extended logic programming with defeasible priorities', Faculty of Law, Free University Amsterdam & The Queen's University of Belfast.

Lambèr Royakkers, *Representing Legal Rules in Deontic Lo*gic, Ph. D. thesis, KU Brabant, Tilburg, 1996.