



EUI Working Papers

LAW 2009/08

DEPARTMENT OF LAW

WHY LAWYERS ARE NICE (OR NASTY)
A GAME - THEORETICAL ARGUMENTATION EXERCISE

Giovanni Sartor, Michel Rudnianski, Antonino Rotolo, Régis Riveret and Eunáte Mayor

**EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF LAW**

Why Lawyers Are Nice (or Nasty)
A Game-Theoretical Argumentation Exercise

**Giovanni Sartor, EUI - Florence, giovanni.sartor@eui.eu
Michel Rudnianski, CNAM/ORT - Paris, michel.rudnianski@cnam.fr
Antonino Rotolo, CIRSFID - Bologna, antonino.rotolo@unibo.it
Régis Riveret, University of Aberdeen, r.riveret@abdn.ac.uk
Eunate Mayor, EUI - Florence, eunate.mayor@eui.eu**

This text may be downloaded for personal research purposes only. Any additional reproduction for such purposes, whether in hard copy or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper or other series, the year, and the publisher.

The author(s)/editor(s) should inform the Law Department of the EUI if the paper is to be published elsewhere, and should also assume responsibility for any consequent obligation(s).

ISSN 1725-6739

©2009 G. Sartor, M. Rudnianski, A. Rotolo, R. Riveret, E. Mayor

Printed in Italy

European University Institute

Badia Fiesolana

I - 50016 San Domenico di Fiesole (FI)

Italy

eui.eu

cadmus.eui.eu

Astract:

This contribution introduces a novel approach to study legal interactions, legal professions, and legal institutions, by combining argumentation, game theory and evolution. We consider a population of lawyers, having different postures, who engage in adversarial argumentation with other lawyers, obtaining outcomes according the existing context and their chosen strategies. We examine the resulting games and analyse the evolution of the population.

Keywords:

law, game theory, evolution, argumentation, litigation

TABLE OF CONTENTS

1. Introduction	1
2. The example	1
3. The Argumentation Framework	1
4. The history supporting the game	2
5. The costs of the parties	3
6. States of the world	4
7. Judges' assessment capacity	5
8. Probability to win a case	6
9. Payoffs	7
10. Lawyer's types	8
11. Encounters and games	9
12. Changing payoffs and games	12
13. Evolutionary analysis	13
14. Conclusion	17
15. References	18

1. Introduction

We shall focus on lawyers interacting in a civil proceedings, and assume their interaction can be modelled as a game, where they may adopt different argumentation strategies. The lawyers' outcomes depend non only on their strategies, but also on the strategies their adversaries are playing. Moreover such outcomes also depend on the context in which the argumentation is taking place. In particular, we shall consider the following contextual variables: the accuracy of the judges (namely their ability to recognise true factual propositions, given certain evidential arguments), the costs of participating in the proceedings and providing evidence, and the value of the case.

We shall also assume that the population of lawyers is divided into different types, each one endowed with a certain posture, defined by two dimensions: honesty and aggressivity. We shall first analyse the strategy adopted by each of type of lawyer when encountering different types kind of counterparts, and the payoffs obtained in each one of these encounters. By assuming that the more successful types of lawyers tend to persist and expand in the population, we shall move to dynamic analysis. This offers an approach for understanding what kind of strategies (and what postures) will prevail and persist in a certain population, under certain conditions, though we will here only provide some preliminary considerations on this issue.

2. The example

In this paper we shall focus on medical liability. More precisely, we shall consider a conflict between a patient and a doctor, concerning the issue of whether the doctor negligently damaged the patient, and should therefore compensate him.

Patient, if he sues *Doctor*, can argue (argument A_1) that he is entitled to a compensation from *Doctor* since (a) he came to harm as a consequence of the operation performed by *Doctor* (*Harm*), and (b) there was negligence in *Doctor's* behaviour (*Negligence*).

Doctor, if she enters the proceeding to defend herself, in order to rebut the *Patient's* claim can reply in one of these ways : (1) deny *Harm* (argument A_2), (2) deny *Negligence* (argument A_3), deny both *Harm* and *Negligence* (arguments A_2 and A_3).

In the following we shall speak of *Patient* and *Doctor* meaning also the patient's lawyers and the doctor's lawyer. The lawyers advice patients and doctors on their processual strategy, making the choices for them, and the clients' gains and losses are gains and losses also for their lawyers (for simplicity's sake we shall not consider ways of mapping clients' gains and losses into lawyers' outcomes).

3. The Argumentation Framework

In the sequel, we assume that these cases can be associated with the set of arguments $\{A_1, A_2, A_3\}$ above. As we observed, arguments A_2 and A_3 attack and are attacked by A_1 : according to A_2 there was no harm while according to A_1 there was; similarly, according to A_3 there was no negligence, while according to A_1 there was. We assume a Dung's argumentation framework ($Args, Attacks$) where $Args$ is a set of arguments and $Attacks$ is a set of attack relations between arguments [1].

Definition 3.1 *An argumentation framework is a pair $AF = (Args, Attacks)$ where $Args$ is a set of arguments, and $Attacks$ is a binary relation on $Args$, i.e., $Attacks \subseteq Args \times Args$.*

We assume that each case c of confrontation between the parties is an encounter in which the latter exchange certain arguments $Args(c)$, having attacks $Attacks(c)$, where $Args(c) \subseteq Args$ and $Attacks(c) \subseteq Attacks$. Thus each case c generates an argumentation framework $\langle Args(c), Attacks(c) \rangle$, which we denote as $AF(c)$. Note that different cases can generate the same argumentation framework (different lawyers or the same ones, may exchange the same arguments in different cases). Considering a case c , we distinguish its argumentation framework $AF(c)$ and its maximal argumentation framework $MaxAF(c)$. While $AF(c)$ only refers to the arguments really exchanged between the parties in case c , $MaxAF(c)$ takes into account all arguments that might possibly have been exchanged: $MaxAF(c) = \langle MaxArgs(c), MaxAttacks(c) \rangle$, where $MaxArgs(c)$ is the set of all argument which the parties could have advanced in c , and $MaxAttacks(c)$ is the set of all attacks between such arguments. In every case c_i we shall consider in this paper, we assume that $MaxArgs(c_i)$ is the same, namely:

$$MaxAF(c_i) = (\{A_1, A_2, A_3\}, \{Attacks(A_2, A_1), Attacks(A_1, A_2), Attacks(A_3, A_1), Attacks(A_1, A_3)\})$$

On the contrary the argumentation framework will depend on the particular case under consideration. For instance, for a particular case c_a , the argumentation framework $AF(c_a)$ may be:

$$AF(c_a) = (\{A_1, A_3\}, \{Attacks(A_3, A_1), Attacks(A_1, A_3)\})$$

In other words, in the argumentation framework $AF(c_a)$ *Patient* claims compensation for having been damaged by *Doctor's* negligence, and *Doctor* replies by arguing that there was no negligence (she does not contest that there was a damage).

4. The history supporting the game

The above framework induces a game consisting basically of two stages: first *Patient* presents his argument and then *Doctor* replies, attacking *Patient's* argument. Thus *Doctor's* choice will depend on *Patient's* choice. Table 1 gives the possible sequences of arguments, called histories, and the factual propositions concerned by the arguments of the parties. We assume that each party will provide evidence for every factual propositions he or she claims to be true, regardless of whether that proposition is true or false. For instance, history H_3 corresponds to the processual

History	<i>Patient's</i> arguments	<i>Doctor's</i> arguments	Propositions to be evidenced
H_1	\emptyset	\emptyset	\emptyset
H_2	A_1	\emptyset	H, N
H_3	A_1	A_2	$H, N, \neg H$
H_4	A_1	A_3	$H, N, \neg N$
H_5	A_1	$[A_2, A_3]$	$H, N, \neg H, \neg N$

Table 1: Histories and corresponding factual propositions

interaction where *Patient* argues that *Doctor* should compensate him, since she has negligently harmed him, and *Doctor* counterargues that she has not harmed *Patient*.

5. The costs of the parties

We assume that each party *Patient* involved in legal proceedings has to support the following processual costs:

- Participation cost (*PartCost*), which is a fixed cost (a tax) to be paid for participating in the proceedings (i.e., for suing or resisting).
- Contested participation surcharge (*ContPartSurch*). This surcharge has to be paid in addition to the *PartCost*, if there is an adversarial interaction (it covers lawyer's and other costs involved in replying to the arguments of the adversarial party).
- Evidence cost (*EvCost*). This cost has to be paid by a party *Patient* for each factual proposition appearing within *Patient's* arguments (*EvCost* for a factual proposition covers the expert or lawyer's work required for building or presenting the evidence related to that proposition). Thus if there are n factual propositions in *Patient's* arguments, *Patient* has to pay $n * EvCost$.
- Contested evidence surcharge *ContEvSurch*. This surcharge has to be paid when the evidence provided on a factual proposition is contested through counterevidence, namely evidence for the negation of that factual proposition (*ContEvSurch* covers contestation of the the counterevidence, providing additional evidence etc.). Thus if there are n factual propositions in *Patient's* arguments of which m are contested, in addition to $n * EvCost$, *Patient* has also to pay $m * ContEvSurch$.

We suppose that all such unit costs are the same for both parties. Thus the processual cost (*ProcCost_{Pat}*) to be sustained by a party *Patient* in a particular processual history is given by the formula

$$ProcCost_{Pat} = X * PartCost + Y * ContPartSurch + W * EvCost + Z * ContEvSurch$$

where

- X equals 1 or 0 depending on whether the party has taken part in the proceedings,
- Y is 1 or 0 depending on whether the adversary participates in the proceedings,
- W is the number of factual proposition upon which the party provides evidence,
- Z is the number of such factual proposition on which counterevidence is provided.

Let us then consider the costs for each history:

- In history H_1 (when there is no litigation), for both parties there is no *PartCost*, *ContPartSurch*, *EvCost* or *ContEvSurch* (X , Y , W and Z are 0), thus $ProcCost_{Pat} = ProcCost_{Doc} = 0$.
- In history H_2 (when *Patient* sues and *Doctor* does not reply, i.e, does not participate in the process) *Patient* pays *PartCost* plus *EvCost* on 2 items (for the 2 factual propositions he has to prove, namely *Harm* and *Negligence*), while *Doctor* pays nothing: $ProcCost_{Pat} = PartCost + 2 * EvCost$; $ProcCost_{Doc} = 0$.
- In history H_3 (when *Patient* sues, and *Doctor* denies harm) *Patient* pays *PartCost*, plus *ContPartSurch*, plus 2 *EvCost*, plus 1 *ContEvSurch*, *Doctor* pays *PartCost*, plus *ConPartSurch* (evidence surcharge for the only factual proposition which is contested, namely *Harm*), plus 1 *EvCost* and 1 *ContEvSurch* (also \neg *Harm* is contested, by *Patient's* evidence for *Harm*): $ProcCost_{Pat} = PartCost + ContPartSurch + 2 * EvCost + ContEvSurch$; $ProcCost_{Doc} = PartCost + ContPartSurch + EvCost + ContEvSurch$.

- In history H_4 (when *Patient* sues, and *Doctor* denies negligence) the costs are the same as in H_3 .
- In history H_5 (when *Patient* sues, and *Doctor* denies both harm and negligence) both parties have to pay $PartCost$, $ContPartSurch$, 2 $EvCost$, and 2 $ContEvSurch$: $ProcCost_{Pat} = ProcCost_{Doc} = PartCost + ContPartSurch + 2 * EvCost + 2 * ContEvSurch$.

Let us assume that in the cases considered the processual costs have the values indicated in Table 2.

	$PartCost$	$ConPartSurch$	$EvCost$	$ConEvSurch$
Value	1	4	1	2

Table 2: Values for processual costs

Then Table 3 shows the corresponding total costs sustained by the parties in each processual history.

History	Patient Costs	Doctor Costs
$H_1(\emptyset, \emptyset)$	0	0
$H_2(A_1, \emptyset)$	3	0
$H_3(A_1, A_2)$	9	8
$H_4(A_1, A_3)$	9	8
$H_5(A_1, [A_2, A_3])$	11	11

Table 3: Costs of the parties

Besides the costs mentioned above, the economical parameters of a case, also include the effects of winning or losing on the substance of the case. Thus we shall consider that in case of victory *Patient* is awarded a compensation $Comp$. In this case the doctor loses and has to pay the compensation. If *Doctor* wins no compensation has to be taken into account. We assume that $Comp$ always equals 10 (Table 4).

	Patient wins	Doctor wins
Patient	10	0
Doctor	-10	0

Table 4: Gains and losses

6. States of the world

Two parameters should be taken into consideration to determine the state of the world: (a) whether *Patient* has suffered harm as a consequence of the operation (H), (b) whether *Doctor* has been negligent in performing the operation (N). For simplicity's sake, we treat negligence too as a fact (though establishing negligence requires matching the agent's behaviour with standards of reasonable case). Thus there are for possible states of the world: $\langle H, N \rangle$; $\langle H, \neg N \rangle$;

$\langle \neg H, N \rangle$; $\langle \neg H, \neg N \rangle$. As we shall see in the next session, the states of the world influence the judicial assessment of the facts of the case and consequently, the probability that the judge accepts the arguments based upon such facts.

7. Judges' assessment capacity

We assume that the parties have full knowledge of the existing states of the world. Namely, they know whether there was harm and negligence (we need to make this assumption to keep the example simple enough, though it must obviously be relaxed to model cases where the parties are uncertain, or have false beliefs, about the existence of the state of affairs they affirm or contest).

On the contrary, judges do not have such knowledge, and have to decide on the basis of the evidence and arguments provided by the parties. However, we assume that judges have cognitive capacities with regard to factual circumstances, so that they are more likely to accept a factual proposition when it is true rather than when it is false, and when evidence for it is provided rather than when it is not. By considering that each proposition may be true or false, and non-contested (only evidence for it is provided) or contested (besides evidence for it, also evidence against is provided), we shall specify the assessment capacity with regard to the two parameters of a factual proposition: whether it is true or false and whether it is non-contested or contested. Thus altogether we need to consider four possible status a proposition may be, when facing its judicial assessment: $\langle True, \neg Contested \rangle$, $\langle True, Contested \rangle$, $\langle False, \neg Contested \rangle$, $\langle False, Contested \rangle$.

To specify the probability of judicial acceptance in each of those status, we introduce a function $PrAcc$, defining the probability that the judge accepts a proposition ϕ according to the status of ϕ , denoted as $Status(\phi)$. Though the (average) values for $PrAcc$ need to be established through empirical inquiry, we can make the following general considerations.

Firstly, the probability of judicial acceptance of a true and non contested proposition ($PrAcc(\phi)$ when $Status(\phi) = \langle True, \neg Contested \rangle$) must be very high.

Secondly, the probability of judicial acceptance of a true and contested proposition ($PrAcc(\phi)$ when $Status(\phi) = \langle True, Contested \rangle$), must be lower than the probability of acceptance of a true and not contested one. However, it must still be higher than .5 (given that judges have some cognitive capacity).

Thirdly, the probability of judicial acceptance of a false and non-contested proposition ($PrAcc(\phi)$ when $State(\phi) = \langle False, \neg Contested \rangle$) depends on the possibility for the judge to get evidence not provided by the parties. When, as usually in private law, the judge does not have this possibility, the judge would tend to align with what is falsely indicated by the uncontested evidence provided by the lying party. Thus this probability too must be higher than .5.

Fourthly, the probability of judicial acceptance of a false and contested proposition ($PrAcc(\phi)$ when $State(\phi) = \langle False, Contested \rangle$) must be lower than .5 (assuming that judges have some cognitive capacity).

In particular, we assume that in the cases we are considering $PrAcc$ is defined as in Table 5.

Thus, according to these values, when it is true that *Patient* has suffered harm and provided evidence for it, two situations must be distinguished: if the evidence for harm is uncontested (*Doctor* provides no evidence against harm), there is a probability 1 that the judge will be persuaded that there is harm, while if the evidence is contested (*Doctor* provides evidence against harm), the chance that the judge will be convinced goes down to .8.

<i>Patient</i>	Evidence on ϕ	$PrAcc(\phi)$
<i>True</i>	$\neg Contested$	1
<i>True</i>	<i>Contested</i>	.8
<i>False</i>	$\neg Contested$.9
<i>False</i>	<i>Contested</i>	.2

Table 5: Judges' accuracy

8. Probability to win a case

To determine what chance *Patient* has of convincing the judge to accept his argument A_1 and thus to grant him a compensation, we have to consider his chances of convincing the judge that operative facts H (harm) and N (negligence) have both occurred. *Doctor*, as we know, can object by providing evidence that there is no damage (A_2), that there is no negligence (A_3), or that there is no damage and that there is no negligence ($[A_2, A_3]$).

Let us now consider what are the chances for *Patient* to convince the judge, given a state of the world and a processual history. For simplicity's sake we assume that the probability that the judge accept two propositions ϕ and χ is the product of the probabilities that he accepts each one of them, namely $PrAcc(\phi \wedge \chi) = PrAcc(\phi) * PrAcc(\chi)$ (the acceptances of the two propositions are independent events). Thus, we obtain Tables 6 and 7 below, where rows are the denoted by argument histories, and columns by the states of the world. The tables indicate the chances of winning of *Patient* and *Doctor*, given the corresponding history and state of the world.

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	0	0	0	0
$H_2(A_1, \emptyset)$	1	.9	.9	.81
$H_3(A_1, A_2)$.8	.72	.2	.18
$H_4(A_1, A_3)$.8	.2	.72	.18
$H_5(A_1, [A_2, A_3])$.64	.16	.16	.04

Table 6: Patient's chances of winning

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	1	1	1	1
$H_2(A_1, \emptyset)$	0	.1	.1	.19
$H_3(A_1, A_2)$.2	.28	.8	.82
$H_4(A_1, A_3)$.2	.8	.28	.82
$H_5(A_1, [A_2, A_3])$.36	.84	.84	.96

Table 7: Doctor's chances of winning

Let us consider for instance the fourth row in Table 6. It corresponds to the argument history $\langle A_1, A_3 \rangle$ where *Patient* asks to be compensated since there was harm and negligence, and *Doctor* denies negligence. In case both harm and negligence are true, *Patient* has .8 chances of convincing the judge. This is because, according to the values indicated in Table 6, he has 1.0 chances of convincing the judge that there was harm (a true factual proposition on which no counterevidence is provided) and .8 chances of convincing the judge that there was no negligence (a true factual proposition on which counterevidence is provided), and $1.0 * .8 = .8$.

In case there was harm and no negligence, and *Doctor* defends herself by claiming that there was no negligence, *Patient*'s chances of success would go down to .2 (i.e., $1 * .2$). This is because *Patient* still has 1.0 chances of convincing the judge that there was harm, but his chances of convincing the judge that there was negligence go down to .2, given that there was no negligence and that *Doctor* is providing evidence to that extent. Similarly, when there was no harm but there was negligence, and *Doctor* denies only negligence, *Patient* has .72 (i.e., $.9 * .8$) chances of success, given that his chances of convincing the judge are .9 for harm (an uncontested false factual proposition), and .8 for negligence (a contested true factual proposition). Finally, when there was no harm and no negligence, and *Doctor* denies only negligence, then *Patient* has .18 (i.e., $.9 * .2$) chances, his chances of convincing the judge are .9 for harm (an uncontested false factual proposition), and .2 for negligence (a contested false factual proposition). Analogous considerations apply to Table 7.

9. Payoffs

The payoffs of the parties depend on their chances of winning or losing the case, and on their processual costs. *Patient*'s payoff is defined by the difference between the expected value of his victory (the value of the compensation times his chances of winning) and his processual costs, while *Doctor*'s payoffs only includes negative entries, i.e., the expected cost for compensating *Patient* and her processual costs. If the parties do not litigate, nothing happens (no compensation is paid and no cost is suffered). Tables 8 and 9 below indicates the payoffs for *Patient* and the doctor.

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	0	0	0	0
$H_2(A_1, \emptyset)$	7	6	6	5.1
$H_3(A_1, A_2)$	-1	-1.8	-7	-7.2
$H_4(A_1, A_3)$	-1	-7	-1.8	-7.2
$H_5(A_1, [A_2, A_3])$	-4.6	-9.4	-9.4	-11

Table 8: Patient's expected payoffs

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	0	0	0	0
$H_2(A_1, \emptyset)$	-10	-9	-9	-8.1
$H_3(A_1, A_2)$	-17	-15.2	-10	-9.8
$H_4(A_1, A_3)$	-17	-10	-15.2	-9.8
$H_5(A_1, (A_2, A_3))$	-17.4	-12.6	-12.6	-11.4

Table 9: Doctor's expected payoffs

To explain the meaning of Tables 8 and 9 consider the fourth line of Table 8, where *Patient* argues for harm and negligence and *Doctor* denies negligence (history H_4). In column $\langle H, N \rangle$ both factual propositions claimed by *Patient* are true. Consequently, *Patient* has an expected payoff of -1 : indeed, *Patient*'s expected compensation is 8 (his chance of getting the 10 compensation is .8, since he must convince of H which is true but contested) from which we must deduct processual costs of 9. In the following entry on the same row, the state of the world is $\langle H, \neg N \rangle$. Thus *Patient* gets -7 , since in such a state history H_4 gives him an expected victory

of 2 (because he has only .2 chances of persuading the judge that there was negligence, this being a false and contested factual proposition) against costs of 9. In column $\langle \neg H, N \rangle$ of the same row, by claiming that there no negligence, *Doctor* argues for the wrong fact (she should counterargue by denying H , which is false, rather than N , which is true). Thus *Patient* has a .72 chance of winning 10, which, once the costs have been taken into account, gives an expected outcome of -1.8 . Finally in the last entry of this row (there was no harm and no negligence), the patient's expected payoffs is -7.2 : indeed, *Patient* has only 18 chances of winning the compensation, since his argument contains two false factual propositions, one of which is contested by *Doctor*.

The payoffs for *Doctor* can be computed in the same way, the compensation won by *Patient* being a loss for *Doctor* (who has also to sustain her expenses). Note that *Doctor* always suffers some losses as a consequence of the process, but her losses are limited when the facts are on her side. It is interesting to notice that some choices make both parties worse off. For instance, in the second column (when there is harm but no negligence), by choosing to counterargue on both harm and negligence (H_5), *Doctor*, while worsening the situation of *Patient* (who has to pay the surcharge for contested evidence also on the issue of harm, besides having a slight decrease in his chances of winning), also worsens her own position. In fact, by contesting also harm *Doctor* diminishes her expected loss for having to have to pay a compensation (since the chance that the judge accepts that there is *Harm* diminish when *Harm* is contested, even when *Harm* is true), but this advantage does not match her additional costs for bringing evidence on a new fact).

10. Lawyer's types

We assume that the lawyers are characterised by two features, aggressiveness and honesty. An honest lawyer only attacks or defends himself by providing arguments based on true factual propositions (we assume that both parties know what really happened). Thus he will refuse to engage in any action requiring him to lie. A non-honest lawyer may use arguments based upon false factual propositions.

As far as aggressiveness is concerned, we distinguish between two levels. An aggressive lawyer always attacks or defends, using the most comprehensive attack strategy available to him or her. A non-aggressive lawyer only attacks or resists when the expected gains exceed the expected costs.

By combining the two features, we obtain the following four *postures*, each one defining a lawyer's type:

- Honest and non aggressive ($\langle Ho, \neg Ag \rangle$). The lawyer advances an argument whenever (1) the argument is right (it is based upon true factual propositions) and (2) it is cost-effective (its expected gains outweigh its expected costs).
- Honest and aggressive ($\langle Ho, Ag \rangle$). The lawyer advances an argument whenever it is right, regardless of its cost.
- Non honest and non aggressive ($\langle \neg Ho, \neg Ag \rangle$). The lawyer advances an argument whenever it is cost-effective, regardless of its rightness.
- Non honest and aggressive ($\langle \neg Ho, Ag \rangle$). The lawyer advances every argument available, regardless of its rightness and its cost.

Thus, an honest patient (whatever his aggressiveness level) would not start a case when he knows that there was no fault of the doctor while an honest doctor would not provide arguments in a case where she knows that she negligently caused damage. Similarly, an honest and non-aggressive patient's lawyer —given the processual costs indicated above—would not sue a aggressive and

<i>Pat</i> \ <i>Doc</i>	<i>Ho</i> , \neg <i>Ag</i>	<i>Ho</i> , <i>Ag</i>	\neg <i>Ho</i> , \neg <i>Ag</i>	\neg <i>Ho</i> , <i>Ag</i>	Total <i>Pat</i>
<i>Ho</i> , \neg <i>Ag</i>	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_1 : 0, 0$	21
<i>Ho</i> , <i>Ag</i>	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : -4.6, -17.4$	16.4
\neg <i>Ho</i> , \neg <i>Ag</i>	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_1 : 0, 0$	21
\neg <i>Ho</i> , <i>Ag</i>	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : -4.6, -17.4$	16.4
Total <i>Doc</i>	-40	-40	-40	-34.8	

Table 10: Game 1, state of the world $\langle H, N \rangle$

dishonest doctor, since his outcome in such a case would be negative (given that the aggressive and dishonest doctor would always counterargue, on all points).

11. Encounters and games

In the sequel, we shall first consider matrix games, in which the players are opposing lawyers. We shall distinguish different games according to the state of the world in which the encounter of the lawyers is taking place. The player's strategic sets (the moves available to them) will be the players' postures $\{\langle Ho, \neg Ag \rangle, \langle Ho, Ag \rangle, \langle \neg Ho, \neg Ag \rangle, \langle \neg Ho, Ag \rangle\}$. In other words, we have $4 * 4$ possible encounters between lawyers postures in each one of the 4 possible states of the world ($\langle H, N \rangle$; $\langle H, \neg N \rangle$; $\langle \neg H, N \rangle$; $\langle \neg H, \neg N \rangle$).

On the basis of the payoffs indicated in Tables 8 and 9 and the judges' level of accuracy indicated in Table 5, we obtain the games indicated in Tables 10, 11, 12 and 13. Let us explain one of these games, for instance the first one in Table 10, taking place in state of the world $\langle H, N \rangle$, namely when it is true that *Patient* was harmed and *Doctor* was negligent. The row player in the table is *Patient's* lawyer and the column player is *Doctor's* lawyer. Each entry of a column represents one encounter (one judicial case) and indicates the history taking place, followed by the corresponding payoffs for *Patient* and *Doctor*.

For instance, the first row of Table 10 lists the encounters which an honest and non-aggressive lawyer, working for a patient, can make. In the first column, such lawyer, *Patient*, encounters, a similarly honest and non aggressive doctor's lawyer, *Doctor*. In such an encounter, *Patient* will sue and *Doctor* will not counterargue (history H_2 , namely, argumentation (A_1, \emptyset)). In fact *Doctor*, being honest, cannot counterargue (this would require her to assert false facts, namely that there was no harm or no negligence from *Doctor*). On the other hand *Patient* knows *Doctor's* posture, and therefore will sue her without being afraid of her counterattack. *Patient's* payoff will then be 10 minus his cost of 3 (which is low since there is no adversarial litigation: it will include 1 for the fixed processual cost, and 2 for providing evidence on the two factual propositions *Patient* has to prove).

The second encounter in the first row of Table 10 confront *Patient* (always honest and non-aggressive) with *Doctor* which is now honest and aggressive. The outcome is the same as in the first column, for the same reason (an honest lawyer cannot counterargue on the basis of false facts).

The third encounter in the first row gives the same outcome, but for a different reason. Here *Patient* knows that if he attacks then he will create a situation where for *Doctor* the best choice is not to counterargue, since being non-aggressive she prefers the most cost-effective choice. In fact, of all histories which start with A_1 , A_1, \emptyset is the story where *Doctor* gets the best payoff (-10 rather than -17 or -17.4), as can be seen in Table 10. Therefore, *Patient* can sue *Doctor*

$Pat \setminus Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_2 : 6, -9$	$H_1 : 0, 0$	$H_2 : 6, -9$	$H_1 : 0, 0$	12
$\neg Ho, Ag$	$H_2 : 6, -9$	$H_5 : -9.4, -10$	$H_2 : 6, -9$	$H_5 : -9.4, -12.6$	-6.8
Total Doc	-18	-10	-18	-12.6	

Table 11: Game 1, state of the world $\langle H, \neg N \rangle$

$pat \setminus Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_2 : 6, -9$	$H_1 : 0, 0$	$H_2 : 6, -9$	$H_1 : 0, 0$	12
$\neg Ho, Ag$	$H_2 : 6, -9$	$H_5 : -9.4, -10$	$H_2 : 6, -9$	$H_5 : -9.4, -12.6$	-6.8
Total Doc	-18	-10	-18	-12.6	

Table 12: Game 1, state of the world $\langle \neg H, N \rangle$

$Pat \setminus Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_2 : 5.1, -8.1$	$H_1 : 0, 0$	$H_2 : 5.1, -8.1$	$H_1 : 0, 0$	1.2
$\neg Ho, Ag$	$H_2 : 5.1, -8.1$	$H_5 : -11, -11.4$	$H_2 : 5.1, -8.1$	$H_5 : -11, -11.4$	-11.8
Total Doc	-16.2	-11.4	-16.2	-11.4	

Table 13: Game 1, state of the world $\langle \neg H, \neg N \rangle$

without being afraid that *Doctor* counterargues, namely, without considering that in histories $H_3(A_1, A_2)$, $H_4(A_1, A_3)$ and $H_5(A_1, (A_2, A_3))$, *Patient* would get a payoff lower than the payoff he would obtain if he would not attack, namely in history H_1 . For a given state of the world and a given pair of postures, this situation corresponds to the kind of games which are called Stackelberg games. i.e. a non simultaneous game in which one of the players who is called the leader plays first, and therefore has to anticipate that the other who plays after him, and is for that reason called the follower will know when she will her decision, the decision taken by the leader. In the case under consideration *Patient* is the leader, and *Doctor* the follower. In the third encounter considered here, *Patient* case he knows that *Doctor*'s best response to A_1 will be \emptyset (which gives her the best possible payoff after A_1 , while *Doctor*'s response to \emptyset will be \emptyset). Since (A_1, \emptyset) gives *Patient* the payoff 7 while (\emptyset, \emptyset) gives him 0, he will play A_1 .

The fourth encounter in the first row gives a different outcome. In this case as *Doctor* is non-honest and aggressive, *Patient* anticipates that *Doctor* would reply to A_1 with (A_2, A_3) , producing history H_5 , with a payoff of -4.6 for *Patient*. Thus *Patient* will not play A_1 , but rather \emptyset , producing history H_1 , where he has payoff 0. This shows how by being apparently irrational (or, if you prefer, by being credibly able to implement a threat also when it would be against her interest to do so), an aggressive lawyer can obtain a better outcome than if she was expected to always make the most cost-effective choice. In other terms, *Doctor*'s determination to counterargue regardless of costs successfully deters a non-aggressive *Patient*, who only reasons on the basis of his gains and costs to take her to court. Note, however, that the aggressive and dishonest *Doctor* would suffer a sever loss when encountering an aggressive *Patient*, who would sue *Doctor* anyway, determining, together with her reaction, history H_5 , with *Doctor*'s outcome of -17.4 .

Let us now examine Table 13, which considers a different factual situation, namely, where there is no harm nor negligence $\langle \neg H, \neg N \rangle$. Let us focus on the third row, which concerns possible encounters of a non-honest and non-aggressive *Patient*. When *Patient* meets an honest and non aggressive *Doctor*, the outcome will be history $H_2(A_1, \emptyset)$, where *Patient* has an expected outcome of 5.1, obtained by subtracting from the expected compensation 8.1 (obtained by multiplying the compensation of 10 by the probability .81 to get it) the processual and evidence costs (3). This is because *Patient* knows that in such a case, if he attacks (with A_1) the best answer for the *Doctor* would be not to respond, even if she is right. By not responding she gets a payoff of -8.1 , while by responding she would get then -9.8 or -11.4 . Therefore *Patient* can sue *Doctor* without being deterred by the possibility of a counterargument (which would bring him a loss of -7.2 or even of -11). The situation is different in the second column, where *Patient*, meets an aggressive *Doctor*. In this case, it is better for *Patient* not to sue, since the aggressive *Doctor*, (both honest or not honest) in this state of the word would counterargue in any case. In order to avoid a loss, a non-aggressive and non-honest *Patient* will not sue, getting a 0 outcome.

We leave to the reader the task of analysing the other rows of Tables 10, 11, 12, 13 and move to Table 14, which gives us the total outcomes for lawyers going through all encounters indicated in Tables from 10 to 13. The first column indicates the outcomes for a patient's lawyer *Patient*. The most profitable posture for *Patient* is to be non-honest and non-aggressive: indeed such a posture enables him to sue whenever convenient, benefiting from the behaviour of non-aggressive *Doctors*, who will refrain from replying when such restraint provides them with better payoff. At the same time, *Patient* refrain from suing aggressive *Doctors*, to avoid losses. The worst outcome would be obtained by a dishonest and aggressive *Patient*, who would attack his aggressive counterparts also when he is wrong (when the facts are against him), and suffer severe losses.

Let us now consider doctors (second column of Table 14). The most convenient posture

	<i>Patient</i>	<i>Doctor</i>	<i>Pat/Doc</i>
$Ho, \neg Ag$	21.0	-92.2	-71.2
Ho, Ag	16.4	-71.4	-55.0
$\neg Ho, \neg Ag$	55.2	-92.2	-37.0
$\neg Ho, Ag$	-9.0	-71.4	-80.4

Table 14: Totals

for *Doctor* consists in being aggressive. An aggressive *Doctor* would do better than a non-aggressive one, since her aggressiveness (the fact that she would counterargue) would deter non-aggressive *Patients* from attacking her. Honesty for an aggressive *Doctor* brings advantages and disadvantages that balance out exactly. In state $\langle H, D \rangle$, an aggressive and dishonest *Doctor* would take advantage of non-aggressive *Patients* (who would be deterred from attacking her, even if she is wrong), but would suffer losses against aggressive *Patients*.

Finally let us consider the third column of Table 14, which provides the total payoffs, indicating what posture would be most convenient when adopted by a lawyer both when representing a patient and when representing a doctor, namely, when playing the *Patient* and the *Doctor* role with the same frequency. It appears that a non-honest and non-aggressive lawyer obtains the best outcome, followed by an honest and aggressive one, then an honest and non-aggressive one, and finally, a non-honest and aggressive one. However, an even better outcome would be achieved by a lawyer who could credibly change his/her posture according to the role he/she is playing, namely, a lawyer able to be non-honest and non-aggressive when representing a patient and being instead aggressive when representing a doctor. Such a lawyer would avoid suing (even when he/she is right) when he/she knows that attack would be followed by a reprisal (a counterargument), but would be able to deter non-aggressive patients from suing her.

12. Changing payoffs and games

Let us assume that, from a societal perspective, the best judicial framework is the one in which virtue is rewarded, namely, in which honest parties obtain a better deal than non-honest ones. This is not the case with previous framework, where dishonest lawyers can obtain the best outcomes. However, things do not need to be this way, there does not need to be a conflict between justice and self-interest. By changing the costs related to processual activity, we can create a new framework where dishonesty will no more be advantageous. This is what happens for instance if, instead of the costs represented in Table 2, we have the costs represented in Table 15, where the surcharge for having a contested case goes down from 4 to 1, and similarly the surcharge for contested evidence goes down from 3 to 1. This change produces the payoffs represented in Tables 16 and 17 and consequently the games represented in Tables 18, 19, 20 and 21.

	<i>PartCost</i>	<i>ConPartSurch</i>	<i>EvCost</i>	<i>ConEvSurch</i>
Amount	1	1	1	1

Table 15: New Costs

This change determines a whole new set of payoffs, as can be seen in Tables 16 and 17. The payoffs *Patient* obtains by suing when he is right (in state $\langle H, N \rangle$) are higher, also when he faces counterarguments, than those he obtains by not suing. Hence, he is not deterred by the reaction of an aggressive and dishonest *Doctor*. Similarly, an honest and non aggressive *Doctor* will find it convenient to counterargue whenever unjustly sued.

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	0	0	0	0
$H_2(A_1, \emptyset)$	7	6	6	5.1
$H_3(A_1, A_2)$	3	2.2	-3	-3.2
$H_4(A_1, A_3)$	3	-3	2.2	-3.2
$H_5(A_1, (A_2, A_3))$.4	-4.4	-4.4	-6

Table 16: Patient's new expected payoffs

History	$\langle H, N \rangle$	$\langle H, \neg N \rangle$	$\langle \neg H, N \rangle$	$\langle \neg H, \neg N \rangle$
$H_1(\emptyset, \emptyset)$	0	0	0	0
$H_2(A_1, \emptyset)$	-10	-9	-9	-8.1
$H_3(A_1, A_2)$	-13	-11.2	-6	-5.8
$H_4(A_1, A_3)$	-13	-6	-11.2	-5.8
$H_5(A_1, (A_2, A_3))$	-12.4	-7.6	-7.6	-6.4

Table 17: Doctor's new expected payoffs

$Pat \setminus Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : .4, -12.4$	21.4
Ho, Ag	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : .4, -12.4$	21.4
$\neg Ho, \neg Ag$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : .4, -12.4$	21.4
$\neg Ho, Ag$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_2 : 7, -10$	$H_5 : .4, -12.4$	21.4
Total Doc	-40	-40	-40	-49.6	

Table 18: Game 2, state of the world $\langle H, N \rangle$

This leads the figures displayed in Tables 18, 19, 20, 21. In these encounters honest and non-aggressive lawyers act exactly in the same way as dishonest and non-aggressive ones (since there are no incentives for dishonesty, justice and self-interest are on the same side), when playing both roles. Only slightly worse outcomes are obtained by honest and aggressive lawyers, since, in state $\langle \neg H, \neg N \rangle$ they would challenge both factual propositions falsely claimed by *Patient*, while it would be slightly more convenient to focus just on one of such propositions (since the increased costs for providing evidence on the other would outweigh the additional amount of expected compensation). Under such a cost structure, aggressive and dishonest lawyers would fare very badly: they would have lost any deterrence capacity, and all of their unjust attacks would be duly countered, so that they would suffer severe losses (on deterrence, see [4]).

13. Evolutionary analysis

In this section we are going to develop an evolutionary analysis of the lawyers' behaviour leaning on evolutionary game theory, an approach originating from biological research and extended to human behaviour. The basic idea is that the superior (inferior) fitness, with regard to their environment, of a certain type of agents determines the evolution of the type's proportion in the considered population, but also that such fitness depends on the composition of the population. For instance, given the payoffs of Table 8, if the population was only made of aggressive and dishonest lawyers, a non-aggressive lawyer would have a higher fitness than an aggressive

$Pat \backslash Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_2 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, Ag$	$H_4 : -3, -6$	$H_5 : -3, -6$	$H_2 : -3, -6$	$H_1 : -4.4, -7.6$	-13.4
Total Doc	-6	-6	-6	-7.6	

Table 19: Game 2, state of the world $\langle H, \neg N \rangle$

$Pat \backslash Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_2 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, Ag$	$H_4 : -3, -6$	$H_5 : -3, -6$	$H_2 : -3, -6$	$H_1 : -4.4, -7.6$	-13.4
Total Doc	-6	-6	-6	-7.6	

Table 20: Game 2, state of the world $\langle \neg H, N \rangle$

$Pat \backslash Doc$	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$	Total Pat
$Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
Ho, Ag	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, \neg Ag$	$H_1 : 0, 0$	$H_1 : 0, 0$	$H_2 : 0, 0$	$H_1 : 0, 0$	0
$\neg Ho, Ag$	$H_4 : -3.2, -5.8$	$H_5 : -6, -6.4$	$H_2 : -3.2, -5.8$	$H_5 : -6, -6.4$	-18.4
Total Doc	-5.8	-6.4	-5.8	-6.4	

Table 21: Game 2, state of the world $\langle \neg H, \neg N \rangle$

lawyer, being able to avoid destructive clashes. On the contrary, if the population was mostly composed of non-aggressive lawyers, an aggressive lawyer would have a higher fitness being able to deter his adversaries from attacking or resisting him. Given a certain initial profile of the population (namely, the proportions of the different types of agents), it is possible to study how the population evolves, and in particular whether it will reach certain stable states or equilibria.

Let us first provide the general framework for an evolutionary analysis, and then apply it to our problem. So consider a population comprised of N interacting species $i = 1, 2, \dots, n$. Let $u_{i,q}$ be the *individual outcome* of an individual of species i after interacting with an individual of species q , and let $(\theta_1; \dots; \theta_n)$ be the *population's profile*, specifying the proportions $\theta_1, \dots, \theta_n$ of the various species in the total population. We need to consider two aspects, each species average outcome and the population's average outcome.

- The species i 's *average outcome* u_i , usually called species i 's fitness, obtained by summing the utility i obtains against each other species q multiplied by q 's frequency, i.e., $u_i = \sum_{q=1}^n \theta_q u_{i,q}$;
- the *population's average outcome* u_T , obtained by summing the average outcome of each species, times the frequency of that species, i.e., $u_T = \sum_{i=1}^n \theta_i u_i$

Note that u_i , θ_i , and u_T , are functions of time, but for the sake of simplicity, we shall not mention time explicitly, unless necessary. A species growth depends on its own outcome, while the evolution of its proportion in the population depends on its relative outcome $u_i - u_T$.

Thus the fundamental equation of the Replicator Dynamics considers that for every species i , its differential change in frequency θ'_i (its tendency to increase or decrease its frequency) is given by $\theta'_i = \theta_i(u_i - u_T)$. For simplicity's sake we shall consider only the case when all lawyers play the same role, representing patients and doctors with the same frequency, thus leading to a symmetric game. We also assume that these lawyers may adopt each of the 4 postures introduced above— $\{Post_{Ho, \neg Ag}, Post_{Ho, Ag}, Post_{\neg Ho, \neg Ag}, Post_{\neg Ho, Ag}\}$ —, and therefore that these 4 postures represent the species into which the lawyers' population is broken down.

To apply this model to our lawyer's example, first of all we have to identify the outcome that each lawyer's posture obtains when facing another lawyer's posture. This can be done by computing first the outcome each posture obtains, in each one of the roles, when meeting each other posture in the different states of the world, which we assume to be equiprobable. By summing outcomes obtained in each state of the world and role, we obtain Table 22 which indicates the outcome a species of lawyers obtains when meeting any other species. The outcome

Posture	$Ho, \neg Ag$	Ho, Ag	$\neg Ho, \neg Ag$	$\neg Ho, Ag$
$Ho, \neg Ag$	-3.0	-3.0	-29.1	-36.1
Ho, Ag	-3.0	-3.0	-3.0	-46.0
$\neg Ho, \neg Ag$	14.1	-3.0	-12.0	-36.1
$\neg Ho, Ag$	24.1	-40.2	24.1	-88.4

Table 22: Total outcome per encounter

one species obtains depends on the species of the counterpart, so that species 1 may a better performance that species 2 when encountering species 3, while being inferior to species 2 when encountering species 4. For instance a non-honest and aggressive lawyer has a better performance

than an honest and aggressive one when encountering an honest and non-aggressive partner, but has an inferior performance when encountering another non honest and aggressive lawyer.

On the basis of Table 22, and our assumption that at the initial time the lawyers are equally distributed among the 4 species (i.e. proportion of each species within the population is .25), we can compute the average outcome for the whole population, according to the formula $u_T = \sum_{i=1}^n \theta_i u_i$, which gives -15.225 . The evolution in the proportion of a species i in a population is determined by $u_i - u_T$, namely, by the difference between i 's average outcome and the average outcome u_T of the whole populations: the species proportion will increase when this difference is positive and decrease when it is negative. This also applies to social attitudes, like the lawyers' postures: the more successful lawyers will remain in the profession and that their postures will imitated by their less successful colleagues. The species whose average outcome is higher than

<i>Posture</i>	u_i	$u_i - u_T$
$Ho, \neg Ag$	-17.800	-2.575
Ho, Ag	-13.750	1.475
$\neg Ho, \neg Ag$	-9.250	5.975
$\neg Ho, Ag$	-20.100	-4.875

Table 23: Outcomes for each species of lawyers

the population's average are, as we you can see in the table above, $\langle \neg Ho, \neg Ag \rangle$, which gets and $\langle Ho, Ag \rangle$. It follwos that that at the initial time (when each species covers 1/4 of the population) the proportion of these two species will tend to increase, while the proportion of the other species will decrease. However, this may not lead to the extinction of all less successful species, since such species may increase their average outcome under different population profiles (for instance, when they only represent an inferior proportion of the population). The precise details of such a dynamic will have to be worked out in further studies. Here we can however tackle the issue of whether the considered species are evolutionarily stable, namely, whether a particular posture would persist in case it were adopted at a given moment by the whole population. In other words, supposing that all lawyers adopt the same posture at a given moment, will they be able to successfully oppose any invader (i.e. any other posture). If such is the case we shall say that the posture selected is an *evolutionary stable strategy* (ESS). More exactly, $u_{i,j}$ being the average outcome a species i obtains when meeting species j , we can say that a species k is a strictly evolutionary stable strategy (ESS) if $\forall i \neq k$:

- either $u_{k,k} > u_{i,k}$
- or $u_{k,k} = u_{i,k}$ and then $u_{k,i} > u_{i,i}$

One can similarly define a weak ESS by simply replacing the second strict inequality here above by $u_{k,i} \geq u_{i,i}$.

It appears that the honest and aggressive posture is the only posture which is evolutionarily stable. If all lawyers would adopt this posture, no other posture could do better then them, and invade the population. In fact (see Table 22), when $\langle Ho, Ag \rangle$ encounters another $\langle Ho, Ag \rangle$ he gets a payoff (-3) which is not inferior to the payoffs any other lawyer would obtain when meeting $\langle Ho, Ag \rangle$. More exactly, $\langle Ho, \neg Ag \rangle$ and $\langle \neg Ho, \neg Ag \rangle$ also obtain -3 against $\langle Ho, Ag \rangle$, while $\langle \neg Ho, Ag \rangle$ obtains an inferior payoff (-46). Moreover $\langle Ho, Ag \rangle$ when meeting $\langle \neg Ho, \neg Ag \rangle$ obtains a payoff (3) equal to the payoff $\langle \neg Ho, \neg Ag \rangle$ obtains meeting another $\langle \neg Ho, \neg Ag \rangle$, and

the same happens with regard to $\langle Ho, \neg Ag \rangle$. Thus we can say that $\langle Ho, Ag \rangle$ is weakly evolutionary stable.

The other postures are not evolutionarily stable: a population of non-aggressive lawyers ($\langle Ho, \neg Ag \rangle$ or $\langle \neg Ho, \neg Ag \rangle$) could be successfully infiltrated by $\langle \neg Ho, \neg Ag \rangle$. In fact, $\langle Ho, \neg Ag \rangle$ or $\langle \neg Ho, \neg Ag \rangle$ when meeting their own type get a payoff inferior to the payoff (24.1) $\langle \neg Ho, Ag \rangle$ gets against a non-aggressive counterpart. A population of $\langle \neg Ho, Ag \rangle$ can be infiltrated by their non-aggressive counterparts, who get, against $\langle \neg Ho, Ag \rangle$, a payoff higher than the payoff $\langle \neg Ho, Ag \rangle$ obtains against his own type (-88.4).

14. Conclusion

In this paper we have analysed legal argumentation through a game-theoretical framework. To that aim we have built upon previous work on argumentation, in particular upon Dung's abstract argumentation framework ([1]) and upon the game-theoretic analyses of argumentation developed by some of the authors of this paper ([3]; [2]). We have provided a model where chances of success of an argument are determined by three elements, the real state of the world, the assessment capacity of the judges, and the argumentative and probatory activities of the parties. We have also represented the costs of legal argumentation, where such costs depend on the decision to sue or resist, but also upon the number of facts on which evidence is to be provided, and on whether there is adversarial confrontation, in particular with regard to particular empirical facts. This model has provided an analysis of argumentation strategies, where the chosen strategies do not depend just on costs and expectations of success, but also on the postures of the involved partners. Finally, we have considered evolutionary dynamics, showing how the population of lawyers may evolve over time, with different postures becoming prevalent.

This new framework has enabled us to show that given certain hypotheses concerning the costs of proceeding, the most successful posture for a lawyer is to be non-honest and non-aggressive, followed by being honest and aggressive, then by being honest and non-aggressive, and finally by being non-honest and aggressive. In other words, given that framework, being non-honest pays only when one is non-aggressive, while aggressiveness only pays when coupled with honesty. We have also shown that by changing the external variable (in particular by reducing the costs of adversarial contest), dishonesty may lose its edge. Our dynamical analysis, while being still very preliminary, also leads to interesting results, such as the emergence (given a certain cost structure) of an environment where non-honest lawyers tend to prevail, followed by non-honest and aggressive ones, while the frequency of the honest and non-aggressive tends to decrease. More important, we the ideas presented in the present paper pave the way for future developments where AI and law can be combined not only with argumentation theory but also with legal sociology and (behavioural) economics.

Firstly, an in-depth analysis could be provided of how contextual variables may impact on argumentation strategies and more generally on the parties' choices. Here we have considered only one possible change in payoffs, namely, a decrease in the costs for adversarial argumentation. Other changes may be considered, namely, increasing or decreasing costs in a different way, putting all costs upon the losing party, adding a penalty in addition to the costs upon the losing party, increasing or decreasing the accuracy of the judges, etc. Moreover, different attitudes to risk and loss could also impact on agent's choices. The option of making an agreement, and the way of managing the corresponding negotiation space should also be considered.

Secondly, the pattern here proposed could be developed by relaxing our knowledge assumptions about the lawyers, namely, the hypothesis that they always know the relevant state of the

world (whether H and N have taken place) and the posture taken by their counterparty (whether the lawyer they encounter is honest and aggressive). Abandoning this assumptions would lead to different games, more general, realistic and complex. A further development would consist in relaxing the assumption that the dispute only concerns facts, i.e., in considering that different views of the applicable law (norms and interpretations) may be advanced by the parties, who have then to take into account the chance that the judge will accept one view rather than another.

Thirdly, the present framework can be extended in a more abstract way, enabling to assess the impact of certain external variables on argumentation strategies, and to determine what strategies will be developed and become prevalent, in different contexts and with different cases. This would require to deal to a larger extent with the mathematical technicalities associated with evolutionary game theory (see [5]).

Finally, we need to develop appropriate simulation tools. While the calculations necessary for the present papers required only some spreadsheets, subsequent developments will require an advanced simulation platform.

The developments just mentioned may be significant both for empirical research on the behaviour of lawyers, parties, judges, and other agents involved in the legal domain. But they will also matter at the policy level, for anticipating the impact of new policies on the attitudes of legal operators and on the working of the legal system, as well as for identifying the ways in which legal interactions can be improved by acting upon the variables which impact on the success (and thus on the reproduction) of certain processual strategies.

15. References

- [1] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n -person games. *Artificial Intelligence*, 77:321–57, 1995.
- [2] Regis Riveret, Henry Prakken, Antonino Rotolo, and Giovanni Sartor. Heuristics in argumentation: A game-theoretical investigation. In *Computational Models of Argument. Proceedings of COMMA-08*, Amsterdam, 2008. IOS.
- [3] Regis Riveret, Nino Rotolo, Giovanni Sartor, Henry Prakken, and Bram Roth. Success chances in argument games: A probabilistic approach to legal disputes. In Arno R. Lodder, editor, *Proceeding of Legal Knowledge and Information Systems - JURIX 2007*, pages 99–108. IOS, Amsterdam, 2007.
- [4] Michel Rudnianski. Deterrence typology and nuclear stability: A game theoretic approach. In *Defence Decision Making*, pages 137–68. Springer, Heidelberg, 1991.
- [5] Jörgen Weibull. *Evolutionary Game Theory*. MIT, Cambridge, Mass. 1995.