Studying Deliberated Judgements

Olivier Cailloux
LAMSADE, Université Paris-Dauphine

Context and goal of this poster

Context

- Internal deliberation facing a decision problem
- Considering an individual i

Goal

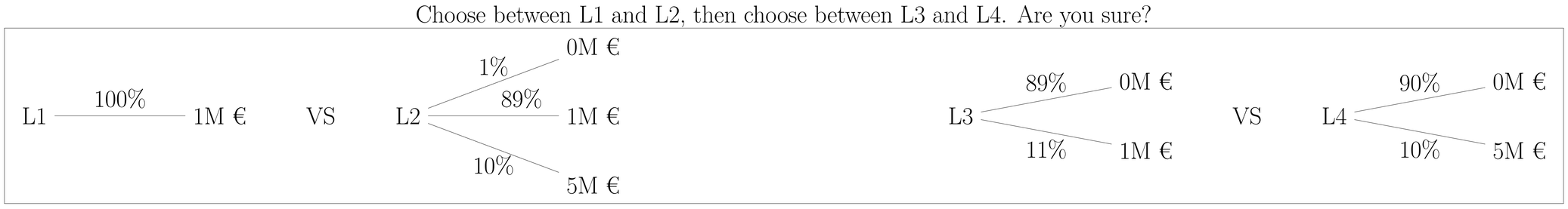
- Introduce the notion of Deliberated Judgement
- Motivate studying it
- Sketch how

Deliberated judgement: a missing conception of "preference"

- Descriptive approach
  - Observe people's epistemic position / choice without interference
- Normative approach
  - How you ought to reason / choose
  - Can't be validated through observation of individuals
- *Deliberated* judgement (or preference)
  - i's position after having considered all arguments

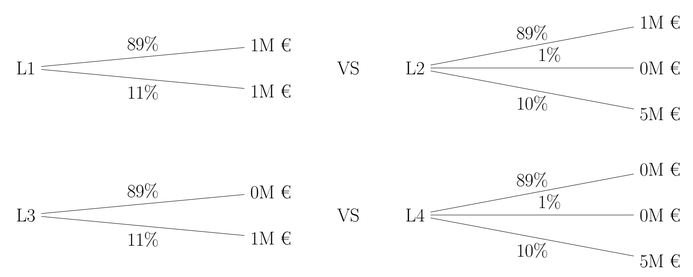
Deliberation can change your mind

Choose between L1 and L2, then choose between L3 and L4. Are you sure?



- First observation (Bernouilli): don't be content with maximizing (untransformed) expected revenue!
- Second observation: i could be intuitively attracted by L1 $\succ$ L2 and L3 $\succ$ L4 (Allais's problem)
- Including Savage
- And might change her mind when given a reasoning pro expected utility
- "There is, of course, an important sense in which preferences, being entirely subjective, cannot be in error"
- ... "but in a different, more subtle sense they can be." (Savage, *The Foundations of Statistics*)

$\Rightarrow$ Systematic decision principles might help deliberate



Study deliberated judgement

The proposed research program aims at the following.

1. Define Deliberated Judgment (DJ) formally
   - Given a set of arguments
   - Of an individual i

   $\Rightarrow$ The position that is stable facing counter-arguments
2. Define the concept of a model of someone's DJ

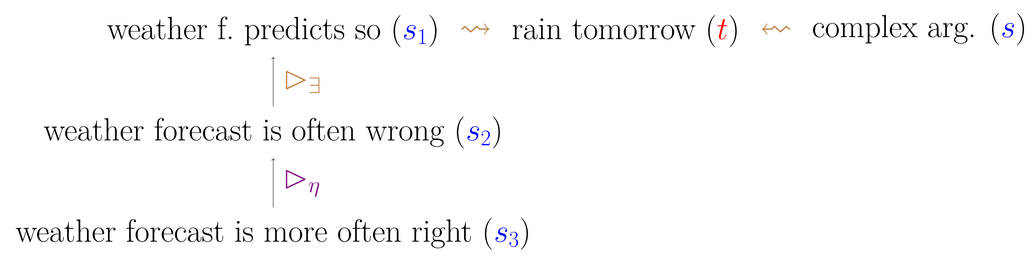
   $\Rightarrow$ A model articulates claims concerning i's DJ and argues for its claim
3. Define validity of a model

   $\Rightarrow$ Correctly captures i's DJ
4. Study conditions for falsifying models using observable data only

   $\Rightarrow$ Let models debate, use i as a judge

Example of a situation and a model of it

Notation	Here	Description
T	$\{t\}$	The topic, containing propositions about which i deliberates
S	$\{s, s_1, s_2, s_3\}$	The arguments
$\rightsquigarrow \subseteq S \times T$	$\{(s,t),(s_1,t)\}$	Support as considered by i
$\rhd_\exists \subseteq S \times S$	$\{(s_2,s_1)\}$	Attacks as considered by i in some perspective
$\rhd_\eta \subseteq S \times S$	$\{(s_3,s_2)\}$	Attacks as considered by the model η



Application: test axioms of decision theory

- Axioms considered appropriate normatively?
  - But some (Allais, Ellsberg) disagree
- Proposal: build models resting on those axioms
- Test models: their convincing power will give us indications about the reasonableness of the axioms for "normal" people (meaning, not scientists studying decision theory)

Application: test conceptions of justice

- Philosophers have proposed sophisticated conceptions of justice (Rawls, Nozick, ...)
- Individual's shallow intuitions about justice are observed and used to confront Rawls or others (Experimental Social Choice)
- Proposal: study reactions of individuals to arguments of philosophers rather than just shallow intuitions
- Move towards Reflective equilibrium (Goodman, Rawls)

Olivier Cailloux and Yves Meinard. *A formal framework for deliberated judgment.* Under revision at Theory and Decision. arXiv:1801.05644 [cs.AI].