Understanding the impacts of uncertainty on optimal policies

Nicola Botta [2] [1]

[2]Potsdam Institute for Climate Impact Research

[1]Thanks to C. Ionescu, P. Jansson and to the Cartesian Seminar people.

Outline

- Results
- Their context
- Method
- A stylized emission problem
- Method (cont'd)

Results

ESD 2017-86, short summary:

> *We study the impact of uncertainty on optimal greenhouse gas (GHG) emission policies for a stylized emission problem. The results suggest that uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies. In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make early emission reductions less rewarding.*

Results

ESD 2017-86, short summary:

We study the impact of uncertainty on optimal greenhouse gas (GHG) emission policies for a stylized emission problem. The results suggest that uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies. In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make early emission reductions less rewarding.

Results

ESD 2017-86, short summary:

We study the impact of uncertainty on optimal greenhouse gas (GHG) emission policies for a stylized emission problem. The results suggest that uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies. In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make early emission reductions less rewarding.

Results

ESD 2017-86, short summary:

> *We study the impact of uncertainty on optimal greenhouse gas (GHG) emission policies for a stylized emission problem. The results suggest that uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies. In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make early emission reductions less rewarding.*

Results

ESD 2017-86, short summary:

We study the impact of uncertainty on optimal greenhouse gas (GHG) emission policies for a stylized emission problem. The results suggest that uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies. In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make early emission reductions less rewarding.

This is not very clear, is it?

The context in which these results have been derived

The context in which these results have been derived

- Global GHG emissions have to be reduced to avoid dangerous impacts of climate change.

The context in which these results have been derived

- Global GHG emissions have to be reduced to avoid dangerous impacts of climate change.
- Reducing emissions implies different costs and benefits for different countries.

The context in which these results have been derived

- Global GHG emissions have to be reduced to avoid dangerous impacts of climate change.
- Reducing emissions implies different costs and benefits for different countries.
- The highest global benefits are obtained if most countries reduce emissions by certain “optimal” amounts.

The context in which these results have been derived

- Global GHG emissions have to be reduced to avoid dangerous impacts of climate change.
- Reducing emissions implies different costs and benefits for different countries.
- The highest global benefits are obtained if most countries reduce emissions by certain “optimal” amounts.
- In this situation most countries face a free-ride option!

The context in which these results have been derived

- Global GHG emissions have to be reduced to avoid dangerous impacts of climate change.
- Reducing emissions implies different costs and benefits for different countries.
- The highest global benefits are obtained if most countries reduce emissions by certain "optimal" amounts.
- In this situation most countries face a free-ride option!

This is known since at least Hardin 1968.

Two sides of the GHG emission problem

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.
   - More than one decision maker.

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.
   - More than one decision maker.
   - Competition, free-riding options, uncertainty.

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.
   - More than one decision maker.
   - Competition, free-riding options, uncertainty.
   - $\Rightarrow$ game theory, mechanism design, cooperation dynamics.

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.
   - More than one decision maker.
   - Competition, free-riding options, uncertainty.
   - $\Rightarrow$ game theory, mechanism design, cooperation dynamics.

One would like to treat the problem as a whole but . . . how?

Two sides of the GHG emission problem

1. When and by how much global GHG emissions should be reduced to avoid potentially dangerous future states?
   - Temporal dimension, sequential decisions.
   - One decision maker.
   - The consequences of decisions are uncertain.
   - $\Rightarrow$ control theory, SDP, optimal policies under uncertainty.

2. How to make sure that (fair, agreed, optimal, etc.) emission reduction quotas are actually implemented?
   - Sequential and simultaneous decisions.
   - More than one decision maker.
   - Competition, free-riding options, uncertainty.
   - $\Rightarrow$ game theory, mechanism design, cooperation dynamics.

One would like to treat the problem as a whole but . . . how? Perhaps using open games?

Focus on 1: when and by how much to reduce emissions?

Decision processes

Decision processes

- Decisions on emission controls (reduce, increase, keep constant) are taken at subsequent points in time and . . .

Decision processes

- Decisions on emission controls (reduce, increase, keep constant) are taken at subsequent points in time and …
- …on the basis of observable current states (economic growth, cumulated emissions, current emission levels, etc.)

Decision processes

- Decisions on emission controls (reduce, increase, keep constant) are taken at subsequent points in time and ...

- ...on the basis of observable current states (economic growth, cumulated emissions, current emission levels, etc.)

- Taking a control in a given state yields a transition to a next state and associated reward (costs, benefits, damages) but ...

Transition uncertainty, decision problem

- …the outcome is not certain: one knows which next states can happen but not which one will happen!

Transition uncertainty, decision problem

- ... the outcome is not certain: one knows which next states can happen but not which one will happen!

- Under this uncertainty, the decision maker seeks emission controls that maximise a sum of rewards over a certain number of decision steps.

Transition uncertainty, decision problem

- . . . the outcome is not certain: one knows which next states can happen but not which one will happen!

- Under this uncertainty, the decision maker seeks emission controls that maximise a sum of rewards over a certain number of decision steps.

It's not really getting better, is it?

Sources of uncertainty

Sources of uncertainty

- Decisions to increase/reduce emissions might actually not be implemented.

Sources of uncertainty

- Decisions to increase/reduce emissions might actually not be implemented.
- Efficient technologies for lowering the costs of reducing emissions may or may not become available.

Sources of uncertainty

- Decisions to increase/reduce emissions might actually not be implemented.
- Efficient technologies for lowering the costs of reducing emissions may or may not become available.
- Exceeding a certain threshold of cumulated emissions may or may not yield potentially dangerous next states.

Sources of uncertainty

- Decisions to increase/reduce emissions might actually not be implemented.
- Efficient technologies for lowering the costs of reducing emissions may or may not become available.
- Exceeding a certain threshold of cumulated emissions may or may not yield potentially dangerous next states.
- What is the effect of these uncertainties on optimal emission policies?

Sources of uncertainty

- Decisions to increase/reduce emissions might actually not be implemented.
- Efficient technologies for lowering the costs of reducing emissions may or may not become available.
- Exceeding a certain threshold of cumulated emissions may or may not yield potentially dangerous next states.
- What is the effect of these uncertainties on optimal emission policies?

It seems that we need to compare optimal policies in two distinct cases. But what are policies?

Optimal policies

Optimal policies

- Informally, policies are functions that associates controls to states

Optimal policies

- Informally, policies are functions that associates controls to states
- A policy for a given decision step is a decision rule: it tells which control to take at that step for every state

Optimal policies

- Informally, policies are functions that associates controls to states
- A policy for a given decision step is a decision rule: it tells which control to take at that step for every state
- Two sequences of n policies can be compared for a given state x by computing their respective sums of rewards for n decision steps starting from x

Optimal policies

- Informally, policies are functions that associates controls to states
- A policy for a given decision step is a decision rule: it tells which control to take at that step for every state
- Two sequences of n policies can be compared for a given state x by computing their respective sums of rewards for n decision steps starting from x
- Optimal sequences of policies are sequences that are at least as good as any other sequence for any x

Optimal policies

- Informally, policies are functions that associates controls to states
- A policy for a given decision step is a decision rule: it tells which control to take at that step for every state
- Two sequences of n policies can be compared for a given state x by computing their respective sums of rewards for n decision steps starting from x
- Optimal sequences of policies are sequences that are at least as good as any other sequence for any x

Can we formalize these ideas?

Optimal policies

- Informally, policies are functions that associates controls to states
- A policy for a given decision step is a decision rule: it tells which control to take at that step for every state
- Two sequences of n policies can be compared for a given state x by computing their respective sums of rewards for n decision steps starting from x
- Optimal sequences of policies are sequences that are at least as good as any other sequence for any x

Can we formalize these ideas? Can we study the impacts of uncertainties on policies with rigorous methods?

Method

Sequential decision problems

Sequential decision problems

- A SDP can be specified in terms of four functions:

Sequential decision problems

- A SDP can be specified in terms of four functions:

$$State \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; Type$$

Sequential decision problems

- A SDP can be specified in terms of four functions:

$$State \ : \ (t \ : \ \mathbb{N}) \ \rightarrow \ Type$$

$$Ctrl \ : \ (t \ : \ \mathbb{N}) \ \rightarrow \ (x \ : \ State \ t) \ \rightarrow \ Type$$

Sequential decision problems

- A SDP can be specified in terms of four functions:

$$State \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; Type$$

$$Ctrl \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; (x \; : \; State \; t) \; \rightarrow \; Type$$

$$next \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; (x \; : \; State \; t) \; \rightarrow \; (y \; : \; Ctrl \; t \; x) \; \rightarrow \; {\color{red} M} \; (State \; (t+1))$$

Sequential decision problems

- A SDP can be specified in terms of four functions:

$$State \ : \ (t \ : \ \mathbb{N}) \ \rightarrow \ Type$$

$$Ctrl \ : \ (t \ : \ \mathbb{N}) \ \rightarrow \ (x \ : \ State \ t) \ \rightarrow \ Type$$

$$next \ : \ (t \ : \ \mathbb{N}) \ \rightarrow \ (x \ : \ State \ t) \ \rightarrow \ (y \ : \ Ctrl \ t \ x) \ \rightarrow \\ M \ (State \ (t+1))$$

Codomain of *State*, *Ctrl*!

Sequential decision problems

- A SDP can be specified in terms of four functions:

$$State \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; Type$$

$$Ctrl \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; (x \; : \; State \; t) \; \rightarrow \; Type$$

$$next \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; (x \; : \; State \; t) \; \rightarrow \; (y \; : \; Ctrl \; t \; x) \; \rightarrow \; M \; (State \; (t+1))$$

Codomain of *State*, *Ctrl*! Dependently typed *Ctrl* and *next*!

Sequential decision problems

- M is a functor representing the problem's uncertainties:

Sequential decision problems

- M is a functor representing the problem's uncertainties:
  - $M = Id$ (deterministic uncertainty)

Sequential decision problems

- M is a functor representing the problem's uncertainties:
  - $M = Id$ (deterministic uncertainty)
  - $M = List$ (non-deterministic uncertainty)

Sequential decision problems

- M is a functor representing the problem's uncertainties:
  - $M = Id$ (deterministic uncertainty)
  - $M = List$ (non-deterministic uncertainty)
  - $M = Prob$ (stochastic uncertainty)

Sequential decision problems

- M is a functor representing the problem's uncertainties:
  - $M = Id$ (deterministic uncertainty)
  - $M = List$ (non-deterministic uncertainty)
  - $M = Prob$ (stochastic uncertainty)
- In many problems $M = Prob$ or $M = List$

Sequential decision problems

- M is a functor representing the problem's uncertainties:
  - $M = Id$ (deterministic uncertainty)
  - $M = List$ (non-deterministic uncertainty)
  - $M = Prob$ (stochastic uncertainty)
- In many problems $M = Prob$ or $M = List$

Monadic dynamical systems, Ionescu 2009!

Sequential decision problems

- The fourth function defines the problem's rewards

Sequential decision problems

- The fourth function defines the problem's rewards

$$\begin{aligned} \mathit{reward}\ :\ & (t\ :\ \mathbb{N})\ \rightarrow\ (x\ :\ \mathit{State}\ t)\ \rightarrow\ (y\ :\ \mathit{Ctrl}\ t\ x)\ \rightarrow \\ & (x'\ :\ \mathit{State}\ (t+1))\ \rightarrow\ \mathit{Val} \end{aligned}$$

Sequential decision problems

- The fourth function defines the problem's rewards

$$\begin{aligned} \mathit{reward} \; : \; (t \; : \; \mathbb{N}) \; \rightarrow \; (x \; : \; \mathit{State} \; t) \; \rightarrow \; (y \; : \; \mathit{Ctrl} \; t \; x) \; \rightarrow \\ (x' \; : \; \mathit{State} \; (t+1)) \; \rightarrow \; \mathit{Val} \end{aligned}$$

- In many problems $\mathit{Val} = \mathbb{R}$ and sums are discounted sums of real numbers!

Sequential decision problems

- The fourth function defines the problem's rewards

$$
\begin{aligned}
\mathit{reward} \; : \; (t \; : \; \mathbb{N}) \;\to\; (x \; : \; \mathit{State}\; t) \;\to\; (y \; : \; \mathit{Ctrl}\; t\; x) \;\to \\
(x' \; : \; \mathit{State}\,(t+1)) \;\to\; \mathit{Val}
\end{aligned}
$$

- In many problems $\mathit{Val} = \mathbb{R}$ and sums are discounted sums of real numbers!

Rewards depend on a current state, on a selected control but also on a next state!

Policies

- Policies are functions from states to controls

$$
\begin{aligned}
&\mathit{Policy} \quad : (t : \mathbb{N}) \rightarrow \mathit{Type} \\
&\mathit{Policy}\ t = (x : \mathit{State}\ t) \rightarrow \mathit{Ctrl}\ t\ x
\end{aligned}
$$

Policies

- Policies are functions from states to controls

$$\begin{aligned} &\mathit{Policy} \;\; : (t : \mathbb{N}) \to \mathit{Type} \\ &\mathit{Policy}\ t = (x : \mathit{State}\ t) \to \mathit{Ctrl}\ t\ x \end{aligned}$$

- Policy sequences are sequences of policies:

$$\begin{aligned} &\textbf{data}\ \mathit{PolicySeq} : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{Type}\ \textbf{where} \\ &\quad \mathit{Nil} \;\; : \mathit{PolicySeq}\ t\ Z \\ &\quad (::) : \mathit{Policy}\ t \to \mathit{PolicySeq}\ (t+1)\ n \to \mathit{PolicySeq}\ t\ (n+1) \end{aligned}$$

Policies

- Policies are functions from states to controls

$$
\begin{aligned}
&\mathit{Policy} \;\; : (t : \mathbb{N}) \rightarrow \mathit{Type} \\
&\mathit{Policy}\ t = (x : \mathit{State}\ t) \rightarrow \mathit{Ctrl}\ t\ x
\end{aligned}
$$

- Policy sequences are sequences of policies:

$$
\begin{aligned}
&\textbf{data}\ \mathit{PolicySeq} : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow \mathit{Type}\ \textbf{where} \\
&\quad \mathit{Nil} \;\; : \mathit{PolicySeq}\ t\ Z \\
&\quad (::) : \mathit{Policy}\ t \rightarrow \mathit{PolicySeq}\ (t+1)\ n \rightarrow \mathit{PolicySeq}\ t\ (n+1)
\end{aligned}
$$

A word-by-word translation from English into math.

Policies

- Policies are functions from states to controls

$$
\begin{aligned}
&\mathit{Policy} \quad : (t : \mathbb{N}) \to \mathit{Type} \\
&\mathit{Policy}\ t = (x : \mathit{State}\ t) \to \mathit{Ctrl}\ t\ x
\end{aligned}
$$

- Policy sequences are sequences of policies:

$$
\begin{aligned}
&\mathbf{data}\ \mathit{PolicySeq} : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{Type}\ \mathbf{where} \\
&\quad \mathit{Nil} \ : \mathit{PolicySeq}\ t\ Z \\
&\quad (::) : \mathit{Policy}\ t \to \mathit{PolicySeq}\ (t+1)\ n \to \mathit{PolicySeq}\ t\ (n+1)
\end{aligned}
$$

A word-by-word translation from English into math. Unfortunately, things are a bit more complicated: viability!

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; \mathit{Val} \; \rightarrow \; \mathit{Val} \; \rightarrow \; \mathit{Type}$$

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; \mathit{Val} \;\rightarrow\; \mathit{Val} \;\rightarrow\; \mathit{Type}$$

- ... on how she adds them

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; \mathit{Val} \; \rightarrow \; \mathit{Val} \; \rightarrow \; \mathit{Type}$$

- ... on how she adds them

$$\oplus \; : \; \mathit{Val} \; \rightarrow \; \mathit{Val} \; \rightarrow \; \mathit{Val}$$

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; Val \;\rightarrow\; Val \;\rightarrow\; Type$$

- ... on how she adds them

$$\oplus \; : \; Val \;\rightarrow\; Val \;\rightarrow\; Val$$

- and on how she measures uncertain outcomes

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; \mathit{Val} \; \rightarrow \; \mathit{Val} \; \rightarrow \; \mathit{Type}$$

- ... on how she adds them

$$\oplus \; : \; \mathit{Val} \; \rightarrow \; \mathit{Val} \; \rightarrow \; \mathit{Val}$$

- and on how she measures uncertain outcomes

$$\mathit{meas} \; : \; M \; \mathit{Val} \; \rightarrow \; \mathit{Val}$$

Optimal policies

- The notion of optimality for policy sequences depends on how the decision maker compares rewards

$$\sqsubseteq \; : \; \mathit{Val} \;\rightarrow\; \mathit{Val} \;\rightarrow\; \mathit{Type}$$

- ... on how she adds them

$$\oplus \; : \; \mathit{Val} \;\rightarrow\; \mathit{Val} \;\rightarrow\; \mathit{Val}$$

- and on how she measures uncertain outcomes

$$\mathit{meas} \; : \; M \; \mathit{Val} \;\rightarrow\; \mathit{Val}$$

Often $\mathit{Val} = \mathbb{R}$, $\sqsubseteq = \mathit{lift} \leqslant$, $\oplus = +$ and *meas* is the expected value function.

Optimal policies

- With $\sqsubseteq$, $\oplus$ and *meas*, one can compute the value of taking n decisions starting from some initial state and according to a policy sequence *ps*:

  $val \;:\; (x \;:\; State\ t) \;\rightarrow\; PolicySeq\ t\ n \;\rightarrow\; Val$

Optimal policies

- With $\sqsubseteq$, $\oplus$ and *meas*, one can compute the value of taking n decisions starting from some initial state and according to a policy sequence *ps*:

$$val \ : \ (x \ : \ State\ t) \ \rightarrow \ PolicySeq\ t\ n \ \rightarrow \ Val$$

$$val\ \{\, t\,\}\ \{\, n = Z\,\}\ x\ Nil = zero$$

Optimal policies

- With $\sqsubseteq$, $\oplus$ and *meas*, one can compute the value of taking n decisions starting from some initial state and according to a policy sequence *ps*:

$$val \; : \; (x \; : \; State \; t) \; \rightarrow \; PolicySeq \; t \; n \; \rightarrow \; Val$$

$$val \; \{t\} \; \{n = Z\} \; x \; Nil = zero$$

$$
\begin{aligned}
& val \; \{t\} \; \{n = m + 1\} \; x \; (p :: ps) = meas \; (fmap \; f \; mx') \; \textbf{where} \\
& \quad y \quad\;\; : \; Ctrl \; t \; x \\
& \quad y \quad\;\; = p \; x \\
& \quad mx' \; : \; M \; (State \; (t + 1)) \\
& \quad mx' = next \; t \; x \; y \\
& \quad f \quad\;\; : \; State \; (t + 1) \; \rightarrow \; Val \\
& \quad f \; x' = reward \; t \; x \; y \; x' \oplus val \; x' \; ps
\end{aligned}
$$

Optimal policies

- With $\sqsubseteq$, $\oplus$ and *meas*, one can compute the value of taking n decisions starting from some initial state and according to a policy sequence *ps*:

$$val \;:\; (x \;:\; State\ t) \;\rightarrow\; PolicySeq\ t\ n \;\rightarrow\; Val$$

$$val\ \{t\}\ \{n = Z\}\ x\ Nil = zero$$

$$\begin{aligned} & val\ \{t\}\ \{n = m+1\}\ x\ (p :: ps) = meas\ (fmap\ f\ mx')\ \textbf{where} \\ & \quad y \quad\;\; : \; Ctrl\ t\ x \\ & \quad y \quad\;\; = p\ x \\ & \quad mx' \; : \; M\ (State\ (t+1)) \\ & \quad mx' = next\ t\ x\ y \\ & \quad f \quad\;\; : \; State\ (t+1) \;\rightarrow\; Val \\ & \quad f\ x' = reward\ t\ x\ y\ x' \oplus val\ x'\ ps \end{aligned}$$

Complexity?

Optimal policies

- ... formalize the notion of optimality for policy sequences

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} \ : \ \mathit{PolicySeq}\ t\ n \ \rightarrow \ \mathit{Type} \\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ \mathit{ps} = (\mathit{ps}' \ : \ \mathit{PolicySeq}\ t\ n) \ \rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\quad (x \ : \ \mathit{State}\ t) \ \rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\quad \mathit{val}\ x\ \mathit{ps}' \sqsubseteq \mathit{val}\ x\ \mathit{ps}
\end{aligned}
$$

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} \ : \ \mathit{PolicySeq}\ t\ n \ \rightarrow \ \mathit{Type} \\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ ps = (ps' \ : \ \mathit{PolicySeq}\ t\ n) \ \rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\quad (x \ : \ \mathit{State}\ t) \ \rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\quad \mathit{val}\ x\ ps' \sqsubseteq \mathit{val}\ x\ ps
\end{aligned}
$$

- ... compute verified optimal policy sequences

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} : \mathit{PolicySeq}\ t\ n \to \mathit{Type} \\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ ps = (ps' : \mathit{PolicySeq}\ t\ n) \to \\
&\qquad (x : \mathit{State}\ t) \to \\
&\qquad \mathit{val}\ x\ ps' \sqsubseteq \mathit{val}\ x\ ps
\end{aligned}
$$

- ... compute verified optimal policy sequences

$$
\begin{aligned}
&\mathit{bi} : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{PolicySeq}\ t\ n \\
&\mathit{biLemma} : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{OptPolicySeq}\ (\mathit{bi}\ t\ n)
\end{aligned}
$$

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} \;:\; \mathit{PolicySeq}\ t\ n \;\to\; \mathit{Type}\\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ \mathit{ps} = (\mathit{ps}' \;:\; \mathit{PolicySeq}\ t\ n) \;\to\\
&\qquad\qquad\qquad\qquad\qquad\quad (x \;:\; \mathit{State}\ t) \;\to\\
&\qquad\qquad\qquad\qquad\qquad\quad \mathit{val}\ x\ \mathit{ps}' \sqsubseteq \mathit{val}\ x\ \mathit{ps}
\end{aligned}
$$

- ... compute verified optimal policy sequences

$$
\begin{aligned}
&\mathit{bi} \;:(t \;:\; \mathbb{N}) \;\to\; (n \;:\; \mathbb{N}) \;\to\; \mathit{PolicySeq}\ t\ n\\
&\mathit{biLemma} \;:(t \;:\; \mathbb{N}) \;\to\; (n \;:\; \mathbb{N}) \;\to\; \mathit{OptPolicySeq}\ (\mathit{bi}\ t\ n)
\end{aligned}
$$

- and, finally, best decisions!

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} \;:\; \mathit{PolicySeq}\ t\ n \;\rightarrow\; \mathit{Type} \\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ ps = (ps' \;:\; \mathit{PolicySeq}\ t\ n) \;\rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\;\; (x \;:\; \mathit{State}\ t) \;\rightarrow \\
&\qquad\qquad\qquad\qquad\qquad\;\; \mathit{val}\ x\ ps' \sqsubseteq \mathit{val}\ x\ ps
\end{aligned}
$$

- ... compute verified optimal policy sequences

$$
\begin{aligned}
&\mathit{bi} \;:\; (t \;:\; \mathbb{N}) \;\rightarrow\; (n \;:\; \mathbb{N}) \;\rightarrow\; \mathit{PolicySeq}\ t\ n \\
&\mathit{biLemma} \;:\; (t \;:\; \mathbb{N}) \;\rightarrow\; (n \;:\; \mathbb{N}) \;\rightarrow\; \mathit{OptPolicySeq}\ (\mathit{bi}\ t\ n)
\end{aligned}
$$

- and, finally, best decisions!

Propositional types, proofs as computations!

Optimal policies

- ... formalize the notion of optimality for policy sequences

$$
\begin{aligned}
&\mathit{OptPolicySeq} \;:\; \mathit{PolicySeq}\ t\ n \;\to\; \mathit{Type}\\
&\mathit{OptPolicySeq}\ \{t\}\ \{n\}\ \mathit{ps} = (\mathit{ps}' \;:\; \mathit{PolicySeq}\ t\ n) \;\to\\
&\qquad\qquad\qquad\qquad\qquad (x \;:\; \mathit{State}\ t) \;\to\\
&\qquad\qquad\qquad\qquad\qquad \mathit{val}\ x\ \mathit{ps}' \sqsubseteq \mathit{val}\ x\ \mathit{ps}
\end{aligned}
$$

- ... compute verified optimal policy sequences

$$
\begin{aligned}
&\mathit{bi} \;:\; (t \;:\; \mathbb{N}) \;\to\; (n \;:\; \mathbb{N}) \;\to\; \mathit{PolicySeq}\ t\ n\\
&\mathit{biLemma} \;:\; (t \;:\; \mathbb{N}) \;\to\; (n \;:\; \mathbb{N}) \;\to\; \mathit{OptPolicySeq}\ (\mathit{bi}\ t\ n)
\end{aligned}
$$

- and, finally, best decisions!

Propositional types, proofs as computations! How to implement *bi* and *biLemma*?

Wrap-up

Given a problem (M, $State$, $Ctrl$, $next$, $reward$ and $meas$) we can:

Wrap-up

Given a problem (*M*, *State*, *Ctrl*, *next*, *reward* and *meas*) we can:

- Compute verified optimal policy sequences.

Wrap-up

Given a problem (*M*, *State*, *Ctrl*, *next*, *reward* and *meas*) we can:

- Compute verified optimal policy sequences.
- Compute all possible trajectories from an initial state under a given policy sequence.

Wrap-up

Given a problem (M, *State*, *Ctrl*, *next*, *reward* and *meas*) we can:

- Compute verified optimal policy sequences.
- Compute all possible trajectories from an initial state under a given policy sequence.

For "small" problems!

Let's take a decision!

- Jump to "A stylized emission problem"
- Jump to "Method (cont'd)"
- Stop here!

A stylized emission problem

Controls

Controls

- At each decision step the decision maker has just two options: low or high emissions:

Controls

- At each decision step the decision maker has just two options: low or high emissions:

$$Ctrl\ t\ x = \{Low, High\}$$

Controls

- At each decision step the decision maker has just two options: low or high emissions:

$$Ctrl\ t\ x = \{Low, High\}$$

- The idea is that low emissions, if implemented, increase the cumulated emissions less than high emissions.

Controls

- At each decision step the decision maker has just two options: low or high emissions:

$$Ctrl\ t\ x = \{Low, High\}$$

- The idea is that low emissions, if implemented, increase the cumulated emissions less than high emissions.

- Without lost of generality, we can take these increases to be zero and one.

States

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:
  - The amount of cumulated emissions.

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:
  - The amount of cumulated emissions.
  - The current emission level $E = \{Low, High\}$.

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:
  - The amount of cumulated emissions.
  - The current emission level $E = \{Low, High\}$.
  - The availability of effective technologies for reducing emissions $T = \{Available, Unavailable\}$.

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:
  - The amount of cumulated emissions.
  - The current emission level $E = \{\mathit{Low}, \mathit{High}\}$.
  - The availability of effective technologies for reducing emissions $T = \{\mathit{Available}, \mathit{Unavailable}\}$.
  - A "state of the world" $W = \{\mathit{Good}, \mathit{Bad}\}$.

States

- At each step, the decision maker has to choose between low and high emissions on the basis of four data:
  - The amount of cumulated emissions.
  - The current emission level $E = \{\mathit{Low}, \mathit{High}\}$.
  - The availability of effective technologies for reducing emissions $T = \{\mathit{Available}, \mathit{Unavailable}\}$.
  - A "state of the world" $W = \{\mathit{Good}, \mathit{Bad}\}$.
- Thus, states are just tuples

$$\mathit{State}\ t = (\{0\,..\,t\}, E, T, W)$$

Decision process

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world
- In this state, the probability that the world turns bad is low.

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world
- In this state, the probability that the world turns bad is low.
- But if the cumulated emissions increase beyond a critical threshold $crE\ :\ \mathbb{R}$, the probability that the world becomes bad increases.

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world
- In this state, the probability that the world turns bad is low.
- But if the cumulated emissions increase beyond a critical threshold $crE : \mathbb{R}$, the probability that the world becomes bad increases.
- Once the world has reached a bad state, there is no chance to turn back to a good state.

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world
- In this state, the probability that the world turns bad is low.
- But if the cumulated emissions increase beyond a critical threshold $crE \; : \; \mathbb{R}$, the probability that the world becomes bad increases.
- Once the world has reached a bad state, there is no chance to turn back to a good state.
- Similarly, the probability that effective technologies become available is low in the beginning and increases after a critical number of decision steps $crN \; : \; \mathbb{N}$.

Decision process

- The process starts with zero cumulated emissions, high emissions, unavailable technologies and in a good world
- In this state, the probability that the world turns bad is low.
- But if the cumulated emissions increase beyond a critical threshold $crE\ :\ \mathbb{R}$, the probability that the world becomes bad increases.
- Once the world has reached a bad state, there is no chance to turn back to a good state.
- Similarly, the probability that effective technologies become available is low in the beginning and increases after a critical number of decision steps $crN\ :\ \mathbb{N}$.
- Once available, effective technologies stay available for ever.

Uncertainties

Uncertainties

- Low/high emission decisions might not be implemented.

Uncertainties

- Low/high emission decisions might not be implemented.

- After crN decision steps, efficient technologies may or may not become available.

Uncertainties

- Low/high emission decisions might not be implemented.
- After *crN* decision steps, efficient technologies may or may not become available.
- Exceeding *crE* may or may not turn the world into a bad state.

Uncertainties

Uncertainties

- Low/high emission decisions might not be implemented:

Uncertainties

- Low/high emission decisions might not be implemented:
  - pLL – the probability of implementing low emissions when the current emissions measures are low and low emissions are chosen.

Uncertainties

- Low/high emission decisions might not be implemented:
  - pLL – the probability of implementing low emissions when the current emissions measures are low and low emissions are chosen.
  - pLH – the probability of implementing low emissions when the current emissions measures are high and low emissions are chosen.

Uncertainties

- Low/high emission decisions might not be implemented:
  - pLL – the probability of implementing low emissions when the current emissions measures are low and low emissions are chosen.
  - pLH – the probability of implementing low emissions when the current emissions measures are high and low emissions are chosen.
  - pHL – the probability of implementing high emissions when the current emissions measures are low and high emissions are chosen.

Uncertainties

- Low/high emission decisions might not be implemented:
  - *pLL* – the probability of implementing low emissions when the current emissions measures are low and low emissions are chosen.
  - *pLH* – the probability of implementing low emissions when the current emissions measures are high and low emissions are chosen.
  - *pHL* – the probability of implementing high emissions when the current emissions measures are low and high emissions are chosen.
  - *pHH* – the probability of implementing high emissions when the current emissions measures are high and high emissions are chosen.

Uncertainties

- Low/high emission decisions might not be implemented:
    - pLL – the probability of implementing low emissions when the current emissions measures are low and low emissions are chosen.
    - pLH – the probability of implementing low emissions when the current emissions measures are high and low emissions are chosen.
    - pHL – the probability of implementing high emissions when the current emissions measures are low and high emissions are chosen.
    - pHH – the probability of implementing high emissions when the current emissions measures are high and high emissions are chosen.
- Constraints: $pLH \leqslant pLL$ and $pHL \leqslant pHH$.

Uncertainties

- After crN decision steps, efficient technologies may or may not become available:

Uncertainties

- After crN decision steps, efficient technologies may or may not become available:
  - $pA1$ – the probability that effective technologies become available when the number of decision steps is $\leqslant crN$,

Uncertainties

- After crN decision steps, efficient technologies may or may not become available:
  - $pA1$ – the probability that effective technologies become available when the number of decision steps is $\leqslant crN$,
  - $pA2$ – the probability that effective technologies become available when the number of decision steps is $> crN$.

Uncertainties

- After *crN* decision steps, efficient technologies may or may not become available:
  - *pA1* – the probability that effective technologies become available when the number of decision steps is $\leqslant$ *crN*,
  - *pA2* – the probability that effective technologies become available when the number of decision steps is $>$ *crN*.
- Constraint: *pA1* $\leqslant$ *pA2*.

Uncertainties

- Exceeding *crE* may or may not turn the world into a bad state:

Uncertainties

- Exceeding *crE* may or may not turn the world into a bad state:
  - *pS1* – the probability of staying in a good world when the cumulated emissions are $\leqslant$ *crE*,

Uncertainties

- Exceeding *crE* may or may not turn the world into a bad state:
  - *pS1* – the probability of staying in a good world when the cumulated emissions are $\leqslant crE$,
  - *pS2* – the probability of staying in a good world when the cumulated emissions are $> crE$.

Uncertainties

- Exceeding crE may or may not turn the world into a bad state:
  - $pS1$ – the probability of staying in a good world when the cumulated emissions are $\leqslant crE$,
  - $pS2$ – the probability of staying in a good world when the cumulated emissions are $> crE$.
- Constraint: $pS2 \leqslant pS1$.

Transition function

Transition function

- Given crN, crE and the probabilities $pLL \ldots pS2$, the transition function can be defined by cases.

Transition function

- Given crN, crE and the probabilities $pLL \ldots pS2$, the transition function can be defined by cases.

- For instance, if

  - the current state is $x = (e, H, U, G)$,

Transition function

- Given crN, crE and the probabilities $pLL \ldots pS2$, the transition function can be defined by cases.
- For instance, if
  - the current state is $x = (e, H, U, G)$,
  - the decision maker has opted for low emissions,

Transition function

- Given crN, crE and the probabilities $pLL \dots pS2$, the transition function can be defined by cases.
- For instance, if
  - the current state is $x = (e, H, U, G)$,
  - the decision maker has opted for low emissions,
  - e is smaller or equal to crE and

Transition function

- Given crN, crE and the probabilities $pLL \ldots pS2$, the transition function can be defined by cases.
- For instance, if
  - the current state is $x = (e, H, U, G)$,
  - the decision maker has opted for low emissions,
  - e is smaller or equal to crE and
  - t is smaller or equal to crN ...

Transition function

- the result of $\mathit{next}\ t\ x\ L$ is a probability distribution with the assignments:

Transition function

- the result of *next t x L* is a probability distribution with the assignments:

$$
\begin{array}{lcl}
prob\,(e, \quad L, U, G) & = & pLH \ast (1 - pA1) \ast pS1 \\
prob\,(e+1, H, U, G) & = & (1 - pLH) \ast (1 - pA1) \ast pS1 \\
prob\,(e, \quad L, A, G) & = & pLH \ast pA1 \ast pS1 \\
prob\,(e+1, H, A, G) & = & (1 - pLH) \ast pA1 \ast pS1 \\
prob\,(e, \quad L, U, B) & = & pLH \ast (1 - pA1) \ast (1 - pS1) \\
prob\,(e+1, H, U, B) & = & (1 - pLH) \ast (1 - pA1) \ast (1 - pS1) \\
prob\,(e, \quad L, A, B) & = & pLH \ast pA1 \ast (1 - pS1) \\
prob\,(e+1, H, A, B) & = & (1 - pLH) \ast pA1 \ast (1 - pS1)
\end{array}
$$

Rewards

Rewards

The idea is that:

- Being in a bad world yields less benefits (more damages) than being in a good world.

Rewards

The idea is that:

- Being in a bad world yields less benefits (more damages) than being in a good world.
- Low emissions yield less benefits (more costs, less growth) than high emissions.

Rewards

The idea is that:

- Being in a bad world yields less benefits (more damages) than being in a good world.
- Low emissions yield less benefits (more costs, less growth) than high emissions.
- Implementing low emissions when effective technologies are unavailable costs more than implementing emissions when these technologies are available.

Rewards

Without loss of generality, we can take the benefits of being in a good world for a step to be one and define

$$
\begin{array}{lcl}
\mathit{reward}\ t\ x\ y\ (e, H, U, G) & = & 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, U, B) & = & b + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, G) & = & 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, B) & = & b + h \\
\mathit{reward}\ t\ x\ y\ (e, L, U, G) & = & 1 + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, U, B) & = & b + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, G) & = & 1 + \mathit{la} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, B) & = & b + \mathit{la}
\end{array}
$$

where $h, b, \mathit{lu}, \mathit{la} : \mathbb{R}$ fulfil $b \leqslant 1$, $0 \leqslant \mathit{lu} \leqslant \mathit{la} \leqslant h$.

Rewards

Without loss of generality, we can take the benefits of being in a good world for a step to be one and define

$$
\begin{aligned}
\mathit{reward}\ t\ x\ y\ (e, H, U, G) &= 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, U, B) &= b + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, G) &= 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, B) &= b + h \\
\mathit{reward}\ t\ x\ y\ (e, L, U, G) &= 1 + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, U, B) &= b + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, G) &= 1 + \mathit{la} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, B) &= b + \mathit{la}
\end{aligned}
$$

where $h, b, \mathit{lu}, \mathit{la} : \mathbb{R}$ fulfil $b \leqslant 1$, $0 \leqslant \mathit{lu} \leqslant \mathit{la} \leqslant h$. Notice that

- The minimal cost of implementing low emissions is $h - \mathit{la}$

Rewards

Without loss of generality, we can take the benefits of being in a good world for a step to be one and define

$$
\begin{aligned}
reward\ t\ x\ y\ (e, H, U, G) &= 1 + h \\
reward\ t\ x\ y\ (e, H, U, B) &= b + h \\
reward\ t\ x\ y\ (e, H, A, G) &= 1 + h \\
reward\ t\ x\ y\ (e, H, A, B) &= b + h \\
reward\ t\ x\ y\ (e, L, U, G) &= 1 + lu \\
reward\ t\ x\ y\ (e, L, U, B) &= b + lu \\
reward\ t\ x\ y\ (e, L, A, G) &= 1 + la \\
reward\ t\ x\ y\ (e, L, A, B) &= b + la
\end{aligned}
$$

where $h, b, lu, la : \mathbb{R}$ fulfil $b \leqslant 1$, $0 \leqslant lu \leqslant la \leqslant h$. Notice that

- The minimal cost of implementing low emissions is $h - la$
- The step costs of being in a bad world are $1 - b$

Rewards

Without loss of generality, we can take the benefits of being in a good world for a step to be one and define

$$
\begin{array}{lcl}
\mathit{reward}\ t\ x\ y\ (e, H, U, G) & = & 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, U, B) & = & b + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, G) & = & 1 + h \\
\mathit{reward}\ t\ x\ y\ (e, H, A, B) & = & b + h \\
\mathit{reward}\ t\ x\ y\ (e, L, U, G) & = & 1 + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, U, B) & = & b + \mathit{lu} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, G) & = & 1 + \mathit{la} \\
\mathit{reward}\ t\ x\ y\ (e, L, A, B) & = & b + \mathit{la}
\end{array}
$$

where $h, b, \mathit{lu}, \mathit{la} : \mathbb{R}$ fulfil $b \leqslant 1$, $0 \leqslant \mathit{lu} \leqslant \mathit{la} \leqslant h$. Notice that

- The minimal cost of implementing low emissions is $h - \mathit{la}$
- The step costs of being in a bad world are $1 - b$
- $1 - b < h - \mathit{la} \Rightarrow$ reducing emissions is never a best choice!

Completing the specification

Completing the specification

- *State*, *Ctrl*, *next* and *reward* as discussed.

Completing the specification

- *State*, *Ctrl*, *next* and *reward* as discussed.
- $M = \mathit{Prob}$, $\mathit{Val} = \mathbb{R}$, $\sqsubseteq = \mathit{lift} \leqslant$, $\oplus = +$, *meas* is the expected value.

Completing the specification

- *State*, *Ctrl*, *next* and *reward* as discussed.
- $\mathit{M} = \mathit{Prob}$, $\mathit{Val} = \mathbb{R}$, $\sqsubseteq = \mathit{lift} \leqslant$, $\oplus = +$, *meas* is the expected value.
- 9 decision steps, starting in $(0, \mathit{H}, \mathit{U}, \mathit{G})$: zero cumulated emissions, high emissions, unavailable efficient technologies and world in a good state.

Completing the specification

- *State*, *Ctrl*, *next* and *reward* as discussed.
- $M = \mathit{Prob}$, $\mathit{Val} = \mathbb{R}$, $\sqsubseteq = \mathit{lift} \leqslant$, $\oplus = +$, *meas* is the expected value.
- 9 decision steps, starting in $(0, H, U, G)$: zero cumulated emissions, high emissions, unavailable efficient technologies and world in a good state.
- $\mathit{crE} = 4$ and $\mathit{crN} = 2$.

Completing the specification

- *State*, *Ctrl*, *next* and *reward* as discussed.
- $M = Prob$, $Val = \mathbb{R}$, $\sqsubseteq = lift \leqslant$, $\oplus = +$, *meas* is the expected value.
- 9 decision steps, starting in $(0, H, U, G)$: zero cumulated emissions, high emissions, unavailable efficient technologies and world in a good state.
- $crE = 4$ and $crN = 2$.
- $b = 0.5$, $lu = 0.1$, $la = 0.2$, $h = 0.3$.

Results

Logical results

Logical results

- If the state of the world is bad, reducing emissions can never be optimal: Après moi le déluge.

Logical results

- If the state of the world is bad, reducing emissions can never be optimal: Après moi le déluge.
- Conversely, reducing emissions can only pay off if it allows avoiding transitions to a bad world.

Logical results

- If the state of the world is bad, reducing emissions can never be optimal: Après moi le déluge.
- Conversely, reducing emissions can only pay off if it allows avoiding transitions to a bad world.
- At the last decision step it is always optimal to select high emissions.

Logical results

- If the state of the world is bad, reducing emissions can never be optimal: Après moi le déluge.
- Conversely, reducing emissions can only pay off if it allows avoiding transitions to a bad world.
- At the last decision step it is always optimal to select high emissions.
- *crE* $= 4 \Rightarrow$ it takes at least 5 steps to achieve states in which the sum of the cumulated emissions exceeds *crE* and, the probability of a transition to a bad world increases from *pS1* to *pS2*.

Logical results

- If the state of the world is bad, reducing emissions can never be optimal: Après moi le déluge.
- Conversely, reducing emissions can only pay off if it allows avoiding transitions to a bad world.
- At the last decision step it is always optimal to select high emissions.
- *crE* $= 4 \Rightarrow$ it takes at least 5 steps to achieve states in which the sum of the cumulated emissions exceeds *crE* and, the probability of a transition to a bad world increases from *pS1* to *pS2*.
- *crN* $= 2 \Rightarrow$ it takes 3 steps to achieve states in which the probability that effective technologies for reducing GHG emissions become available increases from *pA1* to *pA2*.

Deterministic case, basic facts

Deterministic case, basic facts

- $pS2 = pA1 = 0$, $pS1 = pA2 = pLL = pLH = pHL = pHH = 1$.

Deterministic case, basic facts

- *pS2* $=$ *pA1* $= 0$, *pS1* $=$ *pA2* $=$ *pLL* $=$ *pLH* $=$ *pHL* $=$ *pHH* $= 1$.
- Effective technologies become available (with certainty) after 4 steps.

Deterministic case, basic facts

- $pS2 = pA1 = 0$, $pS1 = pA2 = pLL = pLH = pHL = pHH = 1$.
- Effective technologies become available (with certainty) after 4 steps.
- The state of the world turns bad (with certainty) after 6 steps at high emissions.

Deterministic case, basic facts

- *pS2* $=$ *pA1* $= 0$, *pS1* $=$ *pA2* $=$ *pLL* $=$ *pLH* $=$ *pHL* $=$ *pHH* $= 1$.
- Effective technologies become available (with certainty) after 4 steps.
- The state of the world turns bad (with certainty) after 6 steps at high emissions.
- For any given policy sequence there is exactly one possible state-control trajectory.

Deterministic case, policies

- *Const High* policies:

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
```

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```
  - Expected sum of rewards = 9.7.

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```
  - Expected sum of rewards = 9.7.
- *Const Low* policies:

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    `[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),`
    `((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7`
  - Expected sum of rewards $= 9.7$.
- *Const Low* policies:
  - trajectories, probabilities, rewards
    `[((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L),`
    `((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5`

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:

    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```
  - Expected sum of rewards = 9.7.
- *Const Low* policies:
  - trajectories, probabilities, rewards

    ```
    [((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L),
     ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5
    ```
  - Expected sum of rewards = 10.5

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    `[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7`
  - Expected sum of rewards = 9.7.
- *Const Low* policies:
  - trajectories, probabilities, rewards
    `[((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5`
  - Expected sum of rewards = 10.5
- Optimal policies:

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```
  - Expected sum of rewards = 9.7.
- *Const Low* policies:
  - trajectories, probabilities, rewards
    ```
    [((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L),
     ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5
    ```
  - Expected sum of rewards = 10.5
- Optimal policies:
  - trajectories, probabilities, rewards
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),L),
     ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),H), ((5,H,A,G)  )], 100%, 11.3
    ```

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```
  - Expected sum of rewards = 9.7.
- *Const Low* policies:
  - trajectories, probabilities, rewards
    ```
    [((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L),
     ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5
    ```
  - Expected sum of rewards = 10.5
- Optimal policies:
  - trajectories, probabilities, rewards
    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),L),
     ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),H), ((5,H,A,G)  )], 100%, 11.3
    ```
  - Expected sum of rewards = 11.3

Deterministic case, policies

- *Const High* policies:
  - trajectories, probabilities, rewards:

    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
     ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 100%, 9.7
    ```

  - Expected sum of rewards = 9.7.
- *Const Low* policies:
  - trajectories, probabilities, rewards

    ```
    [((0,H,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,U,G),L), ((0,L,A,G),L),
     ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G),L), ((0,L,A,G), )], 100%, 10.5
    ```

  - Expected sum of rewards = 10.5
- Optimal policies:
  - trajectories, probabilities, rewards

    ```
    [((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),L),
     ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),L), ((4,L,A,G),H), ((5,H,A,G)  )], 100%, 11.3
    ```

  - Expected sum of rewards = 11.3
- Optimal policies dictate postponing emission reductions until effective technologies for reducing emissions become available!

Uncertainty on implementability, basic facts

Uncertainty on implementability, basic facts

- *pS2* $=$ *pA1* $= 0$, *pS1* $=$ *pA2* $= 1$ but ...

Uncertainty on implementability, basic facts

- *pS2* $=$ *pA1* $= 0$, *pS1* $=$ *pA2* $= 1$ but ...
- ... *pLL* $=$ *pHH* $= 0.9$ and *pLH* $=$ *pHL* $= 0.7$
- Effective technologies still become available after 4 steps and the state of the world turns bad after 6 steps at high emissions but ...

Uncertainty on implementability, basic facts

- *pS2* $=$ *pA1* $= 0$, *pS1* $=$ *pA2* $= 1$ but ...
- ... *pLL* $=$ *pHH* $= 0.9$ and *pLH* $=$ *pHL* $= 0.7$
- Effective technologies still become available after 4 steps and the state of the world turns bad after 6 steps at high emissions but ...
- ... a policy (optimal or not) now entails $2^9 = 512$ possible trajectories.

Uncertainty on implementability, policies

- *Const High* policies:

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

  - Expected sum of rewards $= 9.904$.

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

  - Expected sum of rewards = 9.904.
- Optimal policies:

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

  - Expected sum of rewards = 9.904.

- Optimal policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )], 23.4%, 11.2
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )],  7.8%, 11.3
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G), )],  7.8%, 11.1
...
```

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

  - Expected sum of rewards = 9.904.
- Optimal policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )], 23.4%, 11.2
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )],  7.8%, 11.3
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G), )],  7.8%, 11.1
...
```

  - Expected sum of rewards = 11.085.

Uncertainty on implementability, policies

- *Const High* policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )], 38.7%,  9.7
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((8,L,A,B)  )],  4.3%,  9.6
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),H), ((4,H,A,G),H),
 ((4,L,A,G),H), ((5,H,A,G),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B), )],  3.3%, 10.1
...
```

  - Expected sum of rewards $= 9.904$.
- Optimal policies:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )], 23.4%, 11.2
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G), )],  7.8%, 11.3
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),L), ((2,L,U,G),L), ((2,L,A,G),L),
 ((2,L,A,G),L), ((2,L,A,G),H), ((2,L,A,G),H), ((3,H,A,G),H), ((4,H,A,G), )],  7.8%, 11.1
...
```

  - Expected sum of rewards $= 11.085$.
- Under uncertainty on implementability, optimal policies dictate earlier emission reductions!

More uncertainties

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?
- We want to estimate the impacts of:

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?
- We want to estimate the impacts of:
  - Uncertainty on the availability of efficient technologies:

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?
- We want to estimate the impacts of:
  - Uncertainty on the availability of efficient technologies:
    - There is a small probability that technologies become available before 4 steps and a small probability that technologies do not become available available even after 4 steps!

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?
- We want to estimate the impacts of:
  - Uncertainty on the availability of efficient technologies:
    - There is a small probability that technologies become available before 4 steps and a small probability that technologies do not become available available even after 4 steps!
  - Uncertainty on the consequences of exceeding the critical cumulated emission threshold crE:

More uncertainties

- What happens to optimal policies if we account for more uncertainties in the decision problem?
- We want to estimate the impacts of:
  - Uncertainty on the availability of efficient technologies:
    - There is a small probability that technologies become available before 4 steps and a small probability that technologies do not become available available even after 4 steps!
  - Uncertainty on the consequences of exceeding the critical cumulated emission threshold crE:
    - There is a small probability that the world turns bad before 6 high emission steps and a small probability that the world doesn't turns bad even after crE has been exceeded!

Uncertainty on the availability of technologies

Uncertainty on the availability of technologies

- *pLL*, *pHH*, *pLH*, *pHL*, *pS1* and *pS2* as before but ...

Uncertainty on the availability of technologies

- *pLL*, *pHH*, *pLH*, *pHL*, *pS1* and *pS2* as before but ...
- ... *pA1* $= 0.1$ and *pA2* $= 0.9$ instead of 0 and 1.

Uncertainty on the availability of technologies

- *pLL*, *pHH*, *pLH*, *pHL*, *pS1* and *pS2* as before but ...
- ... *pA1* $= 0.1$ and *pA2* $= 0.9$ instead of 0 and 1.
- $2^n * (n + 1) = 5120$ possible trajectories for a policy sequence for $n = 9$ steps!

Uncertainty on the availability of technologies

- *pLL*, *pHH*, *pLH*, *pHL*, *pS1* and *pS2* as before but ...
- ... *pA1* $= 0.1$ and *pA2* $= 0.9$ instead of 0 and 1.
- $2^n * (n+1) = 5120$ possible trajectories for a policy sequence for $n = 9$ steps!
- Optimal policies entail the same most likely trajectories. The expected sum of rewards is almost the same!

Uncertainty on the consequences of exceeding *crE*

Uncertainty on the consequences of exceeding *crE*

- *pLL*, *pHH*, *pLH*, *pHL*, *pA1* and *pA2* as before but ...

Uncertainty on the consequences of exceeding *crE*

- *pLL*, *pHH*, *pLH*, *pHL*, *pA1* and *pA2* as before but ...
- ... *pS1* $= 0.9$ and *pS2* $= 0.1$ instead of 1 and 0.

Uncertainty on the consequences of exceeding *crE*

- *pLL*, *pHH*, *pLH*, *pHL*, *pA1* and *pA2* as before but ...
- ... *pS1* $= 0.9$ and *pS2* $= 0.1$ instead of 1 and 0.
- 51200 possible trajectories for a 9-steps policy sequence!

Uncertainty on the consequences of exceeding *crE*

- *pLL*, *pHH*, *pLH*, *pHL*, *pA1* and *pA2* as before but ...
- ... *pS1* $= 0.9$ and *pS2* $= 0.1$ instead of 1 and 0.
- 51200 possible trajectories for a 9-steps policy sequence!
- For *Const High* policies the most likely trajectory is unchanged but ...

Uncertainty on the consequences of exceeding *crE*

- Optimal policies look now quite different:

Uncertainty on the consequences of exceeding *crE*

- Optimal policies look now quite different:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G),)],  5.9%, 11.3

[((0,H,U,G),H), ((1,H,U,B),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.5%,  7.2

[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.3%,  7.7
...
```

Uncertainty on the consequences of exceeding *crE*

- Optimal policies look now quite different:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G),)],  5.9%, 11.3

[((0,H,U,G),H), ((1,H,U,B),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.5%,  7.2

[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.3%,  7.7
...
```

  - Expected sum of rewards $= 9.543$

Uncertainty on the consequences of exceeding *crE*

- Optimal policies look now quite different:
  - trajectories, probabilities, rewards

```
[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,G),H), ((3,H,U,G),L), ((3,L,A,G),L),
 ((3,L,A,G),L), ((3,L,A,G),L), ((3,L,A,G),H), ((4,H,A,G),H), ((5,H,A,G),)],  5.9%, 11.3

[((0,H,U,G),H), ((1,H,U,B),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.5%,  7.2

[((0,H,U,G),H), ((1,H,U,G),H), ((2,H,U,B),H), ((3,H,U,B),H), ((4,H,A,B),H),
 ((5,H,A,B),H), ((6,H,A,B),H), ((7,H,A,B),H), ((8,H,A,B),H), ((9,H,A,B), )],  2.3%,  7.7
...
```

  - Expected sum of rewards = 9.543

- Under uncertainty on the consequences of exceeding *crE*, precautionary policies become sub-optimal: optimal policies dictate later emission reductions!

Wrap-up

Wrap-up

- Uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies.

Wrap-up

- Uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies.

- In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make precautionary policies sub-optimal.

Wrap-up

- Uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies.

- In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make precautionary policies sub-optimal.

- The results are rigorous: optimality of "optimal" policies is machine-checked.

Wrap-up

- Uncertainties about the implementability of decisions on emission reductions (or increases) call for more precautionary policies.
- In contrast, uncertainties about the implications of exceeding critical cumulated emission thresholds tend to make precautionary policies sub-optimal.
- The results are rigorous: optimality of "optimal" policies is machine-checked.

I hope that my first slide is a little bit clearer now!

Method (cont'd)

Generic, verified backwards induction

Generic, verified backwards induction

- The challenge is implementing total functions

$$bi : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow PolicySeq\ t\ n$$

Generic, verified backwards induction

- The challenge is implementing total functions

$$bi \;:\; (t \;:\; \mathbb{N}) \;\rightarrow\; (n \;:\; \mathbb{N}) \;\rightarrow\; PolicySeq\; t\; n$$

and

$$biLemma \;:\; (t \;:\; \mathbb{N}) \;\rightarrow\; (n \;:\; \mathbb{N}) \;\rightarrow\; OptPolicySeq\; (bi\; t\; n)$$

Generic, verified backwards induction

- The challenge is implementing total functions

$$bi \; : (t \; : \; \mathbb{N}) \; \rightarrow \; (n \; : \; \mathbb{N}) \; \rightarrow \; PolicySeq \; t \; n$$

and

$$biLemma \; : (t \; : \; \mathbb{N}) \; \rightarrow \; (n \; : \; \mathbb{N}) \; \rightarrow \; OptPolicySeq \, (bi \; t \; n)$$

for arbitrary M, $\oplus$, $\sqsubseteq$, $State$, $Ctrl$, $next$, $reward$ and $meas$.

Generic, verified backwards induction

- The challenge is implementing total functions

$$bi : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow PolicySeq\ t\ n$$

and

$$biLemma : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow OptPolicySeq\ (bi\ t\ n)$$

for arbitrary M, $\oplus$, $\sqsubseteq$, *State*, *Ctrl*, *next*, *reward* and *meas*.

- As it turns out, if $\oplus$, $\sqsubseteq$ and *meas* fulfill minimal requirements, the implementation directly follows from Bellman's principle of optimality.

Generic, verified backwards induction

- The challenge is implementing total functions

  $$bi : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{PolicySeq}\ t\ n$$

  and

  $$\mathit{biLemma} : (t : \mathbb{N}) \to (n : \mathbb{N}) \to \mathit{OptPolicySeq}\ (bi\ t\ n)$$

  for arbitrary M, $\oplus$, $\sqsubseteq$, *State*, *Ctrl*, *next*, *reward* and *meas*.

- As it turns out, if $\oplus$, $\sqsubseteq$ and *meas* fulfill minimal requirements, the implementation directly follows from Bellman's principle of optimality.

Now we get to the meat of the theory!

Bellman's principle of optimality

- The key idea for understanding Bellman's principle is the notion of optimal extension of a policy sequence:

Bellman's principle of optimality

- The key idea for understanding Bellman's principle is the notion of optimal extension of a policy sequence:

$$
\begin{aligned}
&\mathit{OptExt} &&: \mathit{PolicySeq}\ (t+1)\ m \ \rightarrow\ \mathit{Policy}\ t \ \rightarrow\ \mathit{Type} \\
&\mathit{OptExt}\ ps\ p &&= (x\ :\ \mathit{State}\ t)\ \rightarrow\ (p'\ :\ \mathit{Policy}\ t)\ \rightarrow \\
& && \quad \mathit{val}\ x\ (p' :: ps) \sqsubseteq \mathit{val}\ x\ (p :: ps)
\end{aligned}
$$

Bellman's principle of optimality

- The key idea for understanding Bellman's principle is the notion of optimal extension of a policy sequence:

$$
\begin{aligned}
&\mathit{OptExt} &&: \mathit{PolicySeq}\ (t+1)\ m \ \to\ \mathit{Policy}\ t \ \to\ \mathit{Type} \\
&\mathit{OptExt}\ \mathit{ps}\ p &&= (x\ :\ \mathit{State}\ t)\ \to\ (p'\ :\ \mathit{Policy}\ t)\ \to \\
& && \quad \mathit{val}\ x\ (p' :: \mathit{ps}) \sqsubseteq \mathit{val}\ x\ (p :: \mathit{ps})
\end{aligned}
$$

- With this notion, Bellman's principle can be expressed as

$$
\begin{aligned}
\mathit{Bellman}\ :\ & (\mathit{ps}\ :\ \mathit{PolicySeq}\ (t+1)\ m)\ \to\ \mathit{OptPolicySeq}\ \mathit{ps}\ \to \\
& (p\ \ :\ \mathit{Policy}\ t)\ \to\ \mathit{OptExt}\ \mathit{ps}\ p\ \to \\
& \mathit{OptPolicySeq}\ (p :: \mathit{ps})
\end{aligned}
$$

Bellman's principle of optimality

- The key idea for understanding Bellman's principle is the notion of optimal extension of a policy sequence:

$$
\begin{aligned}
&\mathit{OptExt} &&: \mathit{PolicySeq}\ (t+1)\ m \ \rightarrow\ \mathit{Policy}\ t\ \rightarrow\ \mathit{Type}\\
&\mathit{OptExt}\ ps\ p &&= (x\ :\ \mathit{State}\ t)\ \rightarrow\ (p'\ :\ \mathit{Policy}\ t)\ \rightarrow\\
& &&\quad \mathit{val}\ x\ (p' :: ps) \sqsubseteq \mathit{val}\ x\ (p :: ps)
\end{aligned}
$$

- With this notion, Bellman's principle can be expressed as

$$
\begin{aligned}
\mathit{Bellman}\ :\ &(ps\ :\ \mathit{PolicySeq}\ (t+1)\ m)\ \rightarrow\ \mathit{OptPolicySeq}\ ps\ \rightarrow\\
&(p\ \ :\ \mathit{Policy}\ t)\ \rightarrow\ \mathit{OptExt}\ ps\ p\ \rightarrow\\
&\mathit{OptPolicySeq}\ (p :: ps)
\end{aligned}
$$

- We can prove the principle (implement *Bellman*) if ...

Bellman's principle of optimality (cont'd)

- … $\sqsubseteq$ is reflexive and transitive.

Bellman's principle of optimality (cont'd)

- ... $\sqsubseteq$ is reflexive and transitive.
- $\oplus$ is monotone w.r.t. $\sqsubseteq$:

$$monotonePlusLTE \; : \; a \sqsubseteq b \; \rightarrow \; c \sqsubseteq d \; \rightarrow \; (a \oplus c) \sqsubseteq (b \oplus d)$$

Bellman's principle of optimality (cont'd)

- ... $\sqsubseteq$ is reflexive and transitive.
- $\oplus$ is monotone w.r.t. $\sqsubseteq$:

$$monotonePlusLTE \ : \ a \sqsubseteq b \ \rightarrow \ c \sqsubseteq d \ \rightarrow \ (a \oplus c) \sqsubseteq (b \oplus d)$$

- *meas* fulfills a monotonicity condition (Ionescu 2009):

$$\begin{aligned} measMon \ : \ & \{A \ : \ Type\} \ \rightarrow \\ & (f \ : \ A \ \rightarrow \ Val) \ \rightarrow \ (g \ : \ A \ \rightarrow \ Val) \ \rightarrow \\ & ((a \ : \ A) \ \rightarrow \ (f \ a) \sqsubseteq (g \ a)) \ \rightarrow \\ & (ma \ : \ M \ A) \ \rightarrow \\ & meas \ (fmap \ f \ ma) \sqsubseteq meas \ (fmap \ g \ ma) \end{aligned}$$

Bellman's principle of optimality (cont'd)

- ... $\sqsubseteq$ is reflexive and transitive.
- $\oplus$ is monotone w.r.t. $\sqsubseteq$:

$$monotonePlusLTE \;:\; a \sqsubseteq b \;\rightarrow\; c \sqsubseteq d \;\rightarrow\; (a \oplus c) \sqsubseteq (b \oplus d)$$

- *meas* fulfills a monotonicity condition (Ionescu 2009):

$$\begin{aligned} measMon \;:\; & \{A \;:\; Type\} \;\rightarrow \\ & (f \;:\; A \;\rightarrow\; Val) \;\rightarrow\; (g \;:\; A \;\rightarrow\; Val) \;\rightarrow \\ & ((a \;:\; A) \;\rightarrow\; (f\ a) \sqsubseteq (g\ a)) \;\rightarrow \\ & (ma \;:\; M\ A) \;\rightarrow \\ & meas\ (fmap\ f\ ma) \sqsubseteq meas\ (fmap\ g\ ma) \end{aligned}$$

Remember that M is a functor. Thus, it has a $fmap : (A \;\rightarrow\; B) \;\rightarrow\; (M\ A \;\rightarrow\; M\ B)$!

Bellman's principle of optimality (cont'd)

- ... $\sqsubseteq$ is reflexive and transitive.
- $\oplus$ is monotone w.r.t. $\sqsubseteq$:

$$monotonePlusLTE \ : \ a \sqsubseteq b \ \rightarrow \ c \sqsubseteq d \ \rightarrow \ (a \oplus c) \sqsubseteq (b \oplus d)$$

- *meas* fulfills a monotonicity condition (Ionescu 2009):

$$\begin{aligned} measMon \ : \ & \{A \ : \ Type\} \ \rightarrow \\ & (f \ : \ A \ \rightarrow \ Val) \ \rightarrow \ (g \ : \ A \ \rightarrow \ Val) \ \rightarrow \\ & ((a \ : \ A) \ \rightarrow \ (f \ a) \sqsubseteq (g \ a)) \ \rightarrow \\ & (ma \ : \ M \ A) \ \rightarrow \\ & meas \ (fmap \ f \ ma) \sqsubseteq meas \ (fmap \ g \ ma) \end{aligned}$$

Remember that M is a functor. Thus, it has a $fmap \ : (A \ \rightarrow \ B) \ \rightarrow \ (M \ A \ \rightarrow \ M \ B)$!

Proof idea: $val \ x \ (p' :: ps') \sqsubseteq val \ x \ (p' :: ps) \sqsubseteq val \ x \ (p :: ps)$ and transitivity of $\sqsubseteq$.

Generic, verified backwards induction (cont'd)

Generic, verified backwards induction (cont'd)

- How can we take advantage of Bellman's principle?

Generic, verified backwards induction (cont'd)

- How can we take advantage of Bellman's principle?
- Assume that we can compute optimal extensions of arbitrary policy sequences:

$$\begin{aligned} \mathit{optExt} \quad &: \mathit{PolicySeq}\,(t+1)\,n \;\rightarrow\; \mathit{Policy}\;t \\ \mathit{optExtLemma} \quad &: (\mathit{ps} \;:\; \mathit{PolicySeq}\,(t+1)\,n) \;\rightarrow \\ &\quad \mathit{OptExt}\;\mathit{ps}\;(\mathit{optExt}\;\mathit{ps}) \end{aligned}$$

Generic, verified backwards induction (cont'd)

- How can we take advantage of Bellman's principle?
- Assume that we can compute optimal extensions of arbitrary policy sequences:

$$\begin{aligned}
&\mathit{optExt} &&: \mathit{PolicySeq}\;(t+1)\;n \;\rightarrow\; \mathit{Policy}\;t\\
&\mathit{optExtLemma} &&: (\mathit{ps}\;:\;\mathit{PolicySeq}\;(t+1)\;n)\;\rightarrow\\
& &&\;\;\mathit{OptExt}\;\mathit{ps}\;(\mathit{optExt}\;\mathit{ps})
\end{aligned}$$

- Then the implementation

$$\begin{aligned}
&\mathit{bi}\;t\;Z &&= \mathit{Nil}\\
&\mathit{bi}\;t\;(n+1) &&= \mathit{optExt}\;\mathit{ps} :: \mathit{ps}\;\textbf{where}\\
&\quad \mathit{ps}\;:\;\mathit{PolicySeq}\;(t+1)\;n\\
&\quad \mathit{ps} = \mathit{bi}\;(t+1)\;n
\end{aligned}$$

Generic, verified backwards induction (cont'd)

- How can we take advantage of Bellman's principle?
- Assume that we can compute optimal extensions of arbitrary policy sequences:

$$
\begin{aligned}
&\mathit{optExt} &&: \ \mathit{PolicySeq}\ (t+1)\ n \ \rightarrow \ \mathit{Policy}\ t \\
&\mathit{optExtLemma} &&: \ (\mathit{ps}\ :\ \mathit{PolicySeq}\ (t+1)\ n) \ \rightarrow \\
& && \quad \mathit{OptExt}\ \mathit{ps}\ (\mathit{optExt}\ \mathit{ps})
\end{aligned}
$$

- Then the implementation

$$
\begin{aligned}
&\mathit{bi}\ t\ Z &&= \mathit{Nil} \\
&\mathit{bi}\ t\ (n+1) &&= \mathit{optExt}\ \mathit{ps} :: \mathit{ps}\ \textbf{where} \\
&\quad \mathit{ps}\ :\ \mathit{PolicySeq}\ (t+1)\ n \\
&\quad \mathit{ps} = \mathit{bi}\ (t+1)\ n
\end{aligned}
$$

can be verified:

Generic, verified backwards induction (cont'd)

- How can we take advantage of Bellman's principle?
- Assume that we can compute optimal extensions of arbitrary policy sequences:

$$\begin{aligned}
&\mathit{optExt} &&: \mathit{PolicySeq}\ (t+1)\ n \ \rightarrow\ \mathit{Policy}\ t \\
&\mathit{optExtLemma} &&: (\mathit{ps}\ :\ \mathit{PolicySeq}\ (t+1)\ n)\ \rightarrow \\
& && \quad \mathit{OptExt}\ \mathit{ps}\ (\mathit{optExt}\ \mathit{ps})
\end{aligned}$$

- Then the implementation

$$\begin{aligned}
&\mathit{bi}\ t\ Z &&= \mathit{Nil} \\
&\mathit{bi}\ t\ (n+1) &&= \mathit{optExt}\ \mathit{ps} :: \mathit{ps}\ \textbf{where} \\
&\quad \mathit{ps}\ :\ \mathit{PolicySeq}\ (t+1)\ n \\
&\quad \mathit{ps} = \mathit{bi}\ (t+1)\ n
\end{aligned}$$

can be verified:

But it is very inefficient: the curse and blessing or recursion!

Generic, verified backwards induction (cont'd)

- The task is to implement

$$\mathit{biLemma} : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow \mathit{OptPolicySeq}\,(\mathit{bi}\ t\ n)$$

Generic, verified backwards induction (cont'd)

- The task is to implement

$$biLemma : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow OptPolicySeq\,(bi\ t\ n)$$

- $n = 0$: reflexivity of $\sqsubseteq$.

Generic, verified backwards induction (cont'd)

- The task is to implement

$$biLemma : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow OptPolicySeq\ (bi\ t\ n)$$

- $n = 0$: reflexivity of $\sqsubseteq$.
- $n = m + 1$: induction on n

$$
\begin{array}{l}
biLemma\ t\ (m+1) = Bellman\ ps\ ops\ p\ oep\ \textbf{where} \\
\quad \begin{array}{lcl}
ps & : & PolicySeq\ (t+1)\ m \\
ps & = & bi\ (t+1)\ m \\
ops & : & OptPolicySeq\ ps \\
ops & = & biLemma\ (t+1)\ m \\
p & : & Policy\ t \\
p & = & optExt\ ps \\
oep & : & OptExt\ ps\ p \\
oep & = & optExtLemma\ ps
\end{array}
\end{array}
$$

Generic, verified backwards induction (cont'd)

- The task is to implement

$$biLemma : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow OptPolicySeq\ (bi\ t\ n)$$

- $n = 0$: reflexivity of $\sqsubseteq$.
- $n = m + 1$: induction on n

$$\begin{aligned}
&biLemma\ t\ (m+1) = Bellman\ ps\ ops\ p\ oep\ \textbf{where} \\
&\quad ps \quad : PolicySeq\ (t+1)\ m \\
&\quad ps \quad = bi\ (t+1)\ m \\
&\quad ops\ : OptPolicySeq\ ps \\
&\quad ops = biLemma\ (t+1)\ m \\
&\quad p \quad\ : Policy\ t \\
&\quad p \quad\ = optExt\ ps \\
&\quad oep\ : OptExt\ ps\ p \\
&\quad oep = optExtLemma\ ps
\end{aligned}$$

The question is if and under which conditions we can compute optimal extensions of arbitrary policy sequences.

Generic, verified backwards induction (cont'd)

- The task is to implement

$$biLemma : (t : \mathbb{N}) \rightarrow (n : \mathbb{N}) \rightarrow OptPolicySeq\ (bi\ t\ n)$$

- $n = 0$: reflexivity of $\sqsubseteq$.
- $n = m + 1$: induction on n

$$\begin{array}{l}
biLemma\ t\ (m+1) = Bellman\ ps\ ops\ p\ oep\ \textbf{where} \\
\quad \begin{array}{lcl}
ps & : & PolicySeq\ (t+1)\ m \\
ps & = & bi\ (t+1)\ m \\
ops & : & OptPolicySeq\ ps \\
ops & = & biLemma\ (t+1)\ m \\
p & : & Policy\ t \\
p & = & optExt\ ps \\
oep & : & OptExt\ ps\ p \\
oep & = & optExtLemma\ ps
\end{array}
\end{array}$$

The question is if and under which conditions we can compute optimal extensions of arbitrary policy sequences.

But this is for another talk!

Thanks for your attention!