Agent-Based Systems

Paolo Turrini

www.dcs.warwick.ac.uk/~pturrini p.turrini@warwick.ac.uk

Plan for Today

In this second lecture on **extensive games**, we are going to introduce the notion of **imperfect information**:

- **uncertainty** about the **current state** of the game being played
- **translation** to (and now also from) the **normal form**
- imperfect-information games with **perfect recall**
- subtle differences between **mixed** and **behavioural strategies**
- **Kuhn's Theorem**: condition under which they coincide
- **Computer Poker**: example for an application of these concepts

Much of this (and more) is also covered in Chapter 5 of the *Essentials*.

K. Leyton-Brown and Y. Shoham.
Essentials of Game Theory: A Concise, Multidisciplinary Introduction
Claypool Publishers, 2008. Chapter 5.

Terminology: Incomplete vs. Imperfect Information

Distinguish imperfect information from incomplete information:

- **Incomplete information**: uncertainty about the **game** itself, such as the utilities of other players (↪ Bayesian games)

  Example: Chess in practice (you don't know what your opponent can see)

- **Imperfect information**: uncertainty about current state of **play**, such as the actions taken by others (↪ today's games)

  Example: Poker (you don't know what cards your opponent drew)

Extensive Games with Imperfect Information

A strategic **imperfect-information game** (in extensive form) is a tuple $\langle N, A, H, Z, \underline{i}, \underline{A}, \sigma, \boldsymbol{u}, \sim \rangle$, where $\langle N, A, H, Z, \underline{i}, \underline{A}, \sigma, \boldsymbol{u} \rangle$ is a (finite) extensive form game (with perfect information), i.e.,

- $\mathbf{N} = \{1, \dots, n\}$ is a finite set of **players**;
- $\mathbf{A}$ is a (single) set of **actions**;
- $\mathbf{H}$ is a set of **choice nodes** (non-leaf nodes of the tree);

- $\mathbf{Z}$ is a set of **outcome nodes** (leaf nodes of the tree);
- $\underline{\mathbf{i}} : H \to N$ is the **turn function**, fixing whose turn it is when;
- $\underline{\mathbf{A}} : H \to 2^A$ is the **action function**, fixing the playable actions;

- $\sigma : H \times A \to H \cup Z$ is the (injective) **successor function**;
- $\boldsymbol{u} = (u_1, \dots, u_n)$ is a profile of **utility functions** $u_i : Z \to \mathbb{R}$;

and $\sim = (\sim_1, \dots, \sim_n)$ is a profile of **equivalence relations** $\sim_i$ on H for which $h \sim_i h'$ implies $\underline{i}(h) = \underline{i}(h')$ and $\underline{A}(h) = \underline{A}(h')$ for all $h, h' \in H$.

Modelling Imperfect Information

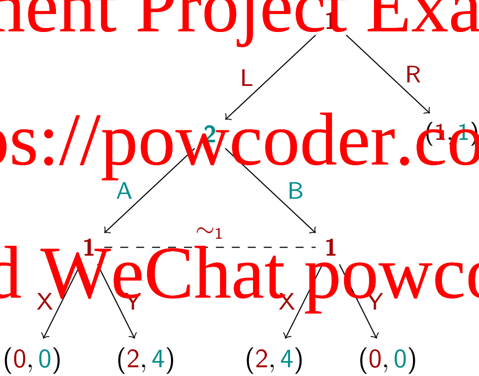
We use $\sim_i$ to relate states of the game that player i **cannot distinguish**.

Write $\mathbf{H_i} := \{h \in H \mid \underline{i}(h) = i\}$ for the set of choice nodes in which it is player i's turn.
We only really need to know $\sim_i$ on H_i (not $H \setminus H_i$).
Thus, for every player $i \in N$, $\sim_i$ is an equivalence relation on H_i such that $h \sim_i h'$ implies $\underline{i}(h) = \underline{i}(h')$ and $\underline{A}(h) = \underline{A}(h')$ for all $h, h' \in H_i$.

Remark: Constraints on $\underline{i}$ and $\underline{A}$ needed to ensure indistinguishability. Otherwise, you might infer extra information from $\underline{i}$ or $\underline{A}$.

Example

An imperfect-information game where Player 1 cannot tell apart the two lower choice nodes in which it is her turn:



Remark: In later examples we will omit the subscript in $\sim_1$, as it is always clear from context *who* is uncertain.

Equivalence Classes

The indistinguishability relation $\sim_i$ partitions the space H_i. Notation:

- The set of all choice nodes that are indistinguishable from node h as far as player i is concerned (**equivalence class**):

$$[\mathbf{h}]_{\sim_i} := \{h' \in H_i \mid h \sim_i h'\}$$

- The set of all such equivalence classes for player i (**quotient set**):

$$\mathbf{H}/\sim_i := \{[h]_\sim \mid h \in H_i\}$$

A **pure strategy** for player i maps any given **equivalence class** $[h]_{\sim_i}$ to an **action** that is playable in (all of!) the choice nodes in $[h]_{\sim_i}$.

<u>So:</u> a function $\alpha_i : H_i/\sim_i \to A$ with $\alpha_i([h]_{\sim_i}) \in \underline{A}(h')$ for all $h' \in [h]_{\sim_i}$.

Alternative Definition of Pure Strategies

Conceptually, pure strategies map **equivalence classes** to **actions**.
But mathematically, we could just as well choose to define pure strategies as mapping **choice nodes** to **actions**:

> *A pure strategy is a function $\alpha_i : H_i \to A$ with $\alpha_i(h) \in A(h)$ for all $h \in H_i$, such that $\alpha_i(h) = \alpha_i(h')$ whenever $h \sim_i h'$.*

Thus, we can think of an imperfect-information game as a standard extensive game where certain strategies are not permitted.

Remark: This does *not* mean that imperfect-information games are special cases of extensive games (the opposite is true!).

Translation to Normal Form

We can translate imperfect-information games into normal-form games, just as we did for (perfect-information) extensive games.

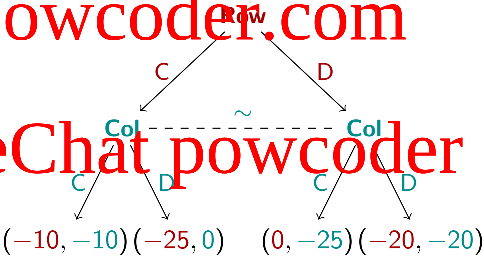
Clear once you understand that incomplete-information games are just extensive games with restricted pure strategies ($\hookrightarrow$ previous slide).

Thus: full machinery available (mixed Nash equilibria, ...).

We have seen that not every normal-form game can be translated into a corresponding (perfect-information) extensive game . . .

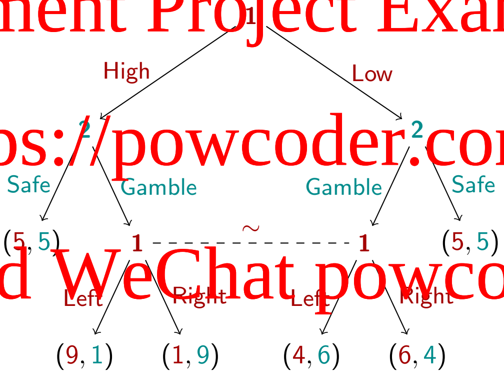
In contrast, *every* NF-game can be translated into an II-game. Use indistinguishability relation to "hide" sequential nature. Example:

	C	D
C	-10 / -10	0 / -25
D	-25 / 0	-20 / -20



Exercise: High-Risk/Low-Risk Gamble

Consider the following imperfect-information game:



This meets our definitions. Still: what's "wrong" with this game?

Perfect Recall

In the game on the previous slide, in the final round Player 1 is assumed to forget the action she played in the first round . . .

An imperfect-information game has **perfect recall**, if $h \sim_i h_0$ implies $h = h_0$ for the root node h_0 and all players $i \in N$, and if the following hold for all $i \in N$, all choice nodes $h, h' \in H$, and all actions $a, a' \in A$:

(*i*) if $\sigma(h, a) \sim_i \sigma(h', a')$, then $h \sim_i h'$

(*ii*) if $\sigma(h, a) \sim_i \sigma(h', a')$ and $\underline{i}(h) = i$, then $a = a'$

Thus, in a perfect-recall game no player i can resolve indistinguishabilityby inspecting the history of (*i*) nodes visited or (*ii*) actions she played.

Note: Every **perfect-information** extensive game has **perfect recall**.

Remark: Games **w/o perfect recall** can make sense in certain contexts. Think of having different agents play on your behalf in different rounds and suppose communication between them is limited.

Mixed vs. Behavioural Strategies

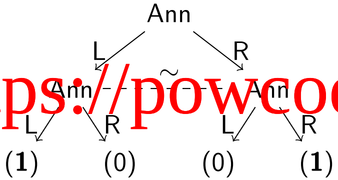
Let $H_i/\!\sim_i = \{[h_i^1]_{\sim_i}, \ldots, [h_i^m]_{\sim_i}\}$ be the set of equivalence classes of choice nodes where it is player i's turn.
Recall: A **pure strategy** for i is some $\alpha_i \in \underline{A}(h_i^1) \times \cdots \times \underline{A}(h_i^m)$.
Thus: A **mixed strategy** for i is some $s_i \in \mathbf{\Pi}(\underline{A}(h_i^1) \times \cdots \times \underline{A}(h_i^m))$.

Clean definition. Nice. But are mixed strategies the right concept? More natural to assume players mix **locally** in each choice node . . .

A **behavioural strategy** for i is some $s_i \in \mathbf{\Pi}(\underline{A}(h_i^1)) \times \cdots \times \mathbf{\Pi}(\underline{A}(h_i^m))$.

Issue: *Can we work with behavioural instead of mixed strategies?*
Definition: Two strategies for player i are called **outcome-equivalent** if, for every partial profile of pure strategies of the other players, the induced probability distributions over outcomes are the same.

In this one-player game, Ann is asked to play two actions in sequence and assumed to forget what she did after the first action got played. She wins (utility 1) if she chooses the same action twice.



Pure strategies: LL, LR, RL, RR

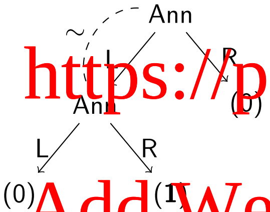
Suppose we want a non-pure strategy that maximises expected utility:

- **Mixed strategy** $(\frac{1}{2}, 0, 0, \frac{1}{2})$ does the job: $\text{EU} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 = 1$
- Any **behavioural strategy** $((\mathbf{p}, 1-p), (\mathbf{q}, 1-q))$ must specify probabilities p to choose L first and q to choose L second:
  $\text{EU} = p \cdot q + (1-p) \cdot (1-q) = 1 + 2pq - p - q < 1$ unless $p = q = 0, 1$

Another Example: Mixed $\neq$ Behavioural Strategy

Recall: Just now we saw that there exist mixed strategies that do not admit any outcome-equivalent behavioural strategy.

It gets worse. Here is another one-player game:



So Ann wins if she manages to play two distinct actions. But she forgets what (and whether!) she played.

Her pure strategies are L and R (more precisely, $H/\sim \to \{L, R\}$).

Playing the **behavioural strategy** $(\frac{1}{2}, \frac{1}{2})$, she wins 25% of the time.
Playing whatever **mixed strategy** (picking L or R), she never wins.
Thus: There exist behavioural strategies that do not admit any outcome-equivalent mixed strategy. *End of story?*

Kuhn's Theorem

Good news:

Theorem (Kuhn, 1953)

In a (finite) imperfect-information game with **perfect recall**, *for any given* **mixed strategy** *of a given player, there exists an* **outcome-equivalent behavioural strategy** *for the same player.*

Proof on next slide.

Remark: The converse holds as well (and sometimes is considered part of Kuhn's Theorem). Proof omitted.

Thus: We can freely move between our two types of randomised strategies. Nice.

Harold W. Kuhn
(1925–2014)

H.W. Kuhn.
Extensive Game and the Problem of Information.
In: H.W. Kuhn and A.W. Tucker (eds.), *Contr. to the Th. of Games*, 1953.

Proof

Claim: For any **perfect-recall game**, for any given mixed strategy s_i of player i we can find an outcome-equivalent behavioural strategy $s_i^\star$.

Proof: We need to fix $s_i^\star(h)(a)$ for any $h \in H_i$ and $a \in \underline{A}(h)$.

Let $p(s_i, h)$ denote the probability of passing through h when player i plays s_i and the others play pure strategies consistent with reaching h.

Let $p(s_i, \sigma(h, a))$ be definied analogously. Fix:

$$s_i^\star(h)(a) \; := \frac{p(s_i, \sigma(\mathbf{h}, \mathbf{a}))}{p(s_i, \mathbf{h})} \quad \left(\text{and } s_i^\star(h)(a) := \frac{1}{|\underline{A}(h)|} \text{ if } p(s_i, h) = 0 \right)$$

Clear: probabilities add up to 1 in each node: $\sum_{a \in \underline{A}(h)} s_i^\star(h)(a) = 1$ ✓

But to be a well-defined behavioural strategy, $s^\star$ also must respect $\sim_i$:

$$h \sim_i h' \quad \Rightarrow \quad s_i^\star(h)(a) = s_i^\star(h')(a)$$

Due to **perfect recall**, the actions played by i on the path to h are the same as those on the path to h'. Nothing else affects probabilities. ✓

The well-known recent work on building an intelligent agent to play Poker (for two players and a simplified set of rules) models the game as an extensive imperfect-information game and attempts to compute Nash equilibria in behavioural strategies.

The main challenge is in dealing with the sheer size of such games. This is tackled using a mix of game theory, abstraction techniques, combinatorial optimisation, and machine learning.

T.W. Sandholm.
Solving Imperfect-Information Games.
Science, 347(6218):122–123, 2015.

Summary

This conclues our review of extensive games in general and of imperfect-information games in particular:

- model: basic **extensive games** + **indistinguishability** relation
- same **expressive power** as normal form: translation possible
- **behavioural strategies** more natural than usual mixed strategies
- **Kuhn's Theorem**: for imperfect-information games of perfect recall, we can rewrite any mixed strategy as a behavioural strategy

What next? Cracking poker with Counterfactual Regret Minimisation.