Agent-Based Systems

Paolo Turrini

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

The approach

We are going to look at exercises together, how to approach them and how to solve them.

I'm also going to point to resources with similar exercises.

The question

Consider now a variant of the game of chess where either the game terminates with a win for White, or a win for Black or it goes on forever. Does Zermelo's theorem work for this variant? Prove it, or give a counterexample.

This question is intentionally ambiguous. There is an easy answer: "No, because Zermelo's theorem is formulated for a different variant", and the one that requires work. When in doubt, it's always the one that requires work.

The question is asking you the following: formulate the new variant in terms of an extensive game, and prove whether it's determined, adapting the proof of Zermelo's Theorem.

Consider now a variant of the game of chess where either the game terminates with a win for White, or a win for Black or it goes on forever. Does Zermelo's theorem work for this variant? Prove it, or give a counterexample.

- Take Zermelo's proof
- See if/where it breaks under the new variant
- Try to adapt it
- Give the new result

Zermelo's Theorem

We want to show whether the following holds:

Theorem

In **(new variant of)** *Chess, one of the following must be true:*

- *White has a winning strategy*
- *Black has a winning strategy*
- *Both players have an at-least-drawing strategy* [a]

[a] Or, if you like, both players have a strategy to either win or make the game go on forever.

Proof (1/3)

Proof.

"Let us first recall that the game is finite, i.e., there is a natural number K such that every play of the game concludes after at most $2K$ rounds,
(K turns by White, K by Black)."

Proof (1/3)

Proof.

"Let us first recall that the game is finite, i.e., there is a natural number K such that every play of the game concludes after at most $2K$ rounds, (K turns by White, K by Black)."

false in the new variant!

Proof (1/3)

Proof.

"Let us first recall that the game is finite, i.e., there is a natural number K such that every play of the game concludes after at most $2K$ rounds, (K turns by White, K by Black)."

false in the new variant!

But we can instead say:

"We know that the set of board positions is finite. So, we know that if White can achieve a win, it can achieve it after at most $2K$ steps, no matter Black's strategy, for some natural number K. (the story is similar for reversed colours)."

Proof (1/3)

Proof.

"Let us first recall that the game is finite, i.e., there is a natural number K such that every play of the game concludes after at most $2K$ rounds,
(K turns by White, K by Black)."
false in the new variant!

But we can instead say:
"We know that the set of board positions is finite. So, we know that if White can achieve a win, it can achieve it after at most $2K$ steps, no matter Black's strategy, for some natural number K. (the story is similar for reversed colours)."

Assume there are exactly $2K$ such turns in every (finite!!) play of the game. Notice that if some plays are shorter, we can simply continue them by adding a "do nothing" move, preserving the result (so, for instance if we extend a play where White wins, we keep track of this).

Proof (2/3)

Proof.

For every k with $1 \leqslant k \leqslant K$ denote:

Proof (2/3)

Proof.

For every k with $1 \leqslant k \leqslant K$ denote:

- a_k the move implemented by White at their turn.
- b_k the move implemented by Black at their turn.

Proof (2/3)

Proof.

For every k with $1 \leqslant k \leqslant K$ denote:

- a_k the move implemented by White at their turn.
- b_k the move implemented by Black at their turn.

Denote W the fact that White wins (after $2K$ turns), $\neg W$ the fact that White does not.

Proof (2/3)

Proof.

For every k with $1 \leqslant k \leqslant K$ denote:

- a_k the move implemented by White at their turn.
- b_k the move implemented by Black at their turn.

Denote W the fact that White wins (after $2K$ turns), $\neg W$ the fact that White does not.

But then, the fact that White has a winning strategy can be written as:

$$\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

Proof (3/3)

Proof.

So, the fact that White has not a winning strategy can be written as:

$$\neg\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

Proof (3/3)

Proof.

So, the fact that White has not a winning strategy can be written as:

$$\neg \exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

This, using first-order logic, is equivalent to

$$\forall a_1 \exists b_1 \forall a_2 \exists b_2 \ldots \forall a_K \exists b_K (\neg W)$$

Proof (3/3)

Proof.

So, the fact that White has not a winning strategy can be written as:

$$\neg \exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

This, using first-order logic, is equivalent to

$$\forall a_1 \exists b_1 \forall a_2 \exists b_2 \ldots \forall a_K \exists b_K (\neg W)$$

But this says that Black is guaranteed at least a draw (as the game will go on forever)!

Proof (3/3)

Proof.

So, the fact that White has not a winning strategy can be written as:

$$\neg\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

This, using first-order logic, is equivalent to

$$\forall a_1 \exists b_1 \forall a_2 \exists b_2 \ldots \forall a_K \exists b_K (\neg W)$$

But this says that Black is guaranteed at least a draw (as the game will go on forever)!

We can do exactly the same for Black.

Proof (3/3)

Proof.

So, the fact that White has not a winning strategy can be written as:

$$\neg\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (W)$$

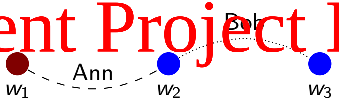
This, using first-order logic, is equivalent to

$$\forall a_1 \exists b_1 \forall a_2 \exists b_2 \ldots \forall a_K \exists b_K (\neg W)$$

But this says that Black is guaranteed at least a draw (as the game will go on forever)!

We can do exactly the same for Black.
Therefore, one of the three alternatives must hold. □

The question



Let $M, w \models \varphi$ be true if and only $w \in \varphi \subseteq W$ in model M made by worlds W equivalence relations $\{\sim_i\}_{i \in N}$, and function $V\{blue, red\} \rightarrow 2^W$.

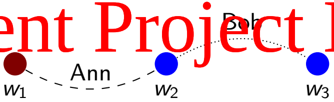
Let moreover $\neg$ and $\wedge$ be the logical correspondents of complement and intersection.

Is the following true or false in the model above?

- $M, w_3 \models K_{Ann} \neg K_{Bob} blue \wedge blue$

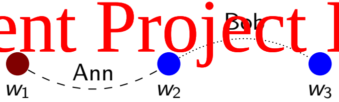
Approach: Simply a matter of applying the definitions and find out.

The answer



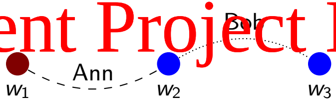
We wanna check whether $M, w_3 \models K_{Ann} \neg K_{Bob} blue \wedge blue$ is true.

.



We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$.

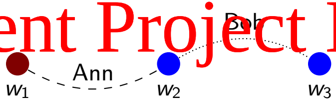
.



We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$.

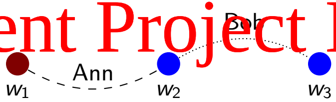
.

The answer



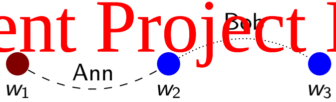
We wanna check whether $M, w_3 \models K_{Ann} \neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann} \neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$. Now we check the former.

.

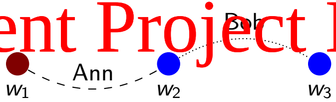


We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$. Now we check the former. $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true if and only if $\forall w : w_3 \sim_{Ann} w$ we have that $M, w \models \neg K_{Bob} blue$.

The answer

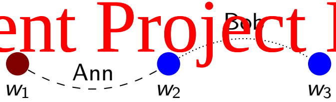


We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$. Now we check the former. $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true if and only if $\forall w : w_3 \sim_{Ann} w$ we have that $M, w \models \neg K_{Bob} blue$. Notice $\{w \mid w_3 \sim_{Ann} w\} = \{w_3\}$. So we need to check whether $M, w_3 \models \neg K_{Bob} blue$.



We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$. Now we check the former. $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true if and only if $\forall w : w_3 \sim_{Ann} w$ we have that $M, w \models \neg K_{Bob} blue$. Notice $\{w \mid w_3 \sim_{Ann} w\} = \{w_3\}$. So we need to check whether $M, w_3 \models \neg K_{Bob} blue$. This is true if there exists a w' with $w_3 \sim_{Bob} w'$ with $w' \in V(red)$.

The answer



We wanna check whether $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is true. By the inductive definition, it is true if and only if $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true AND $M, w_3 \models blue$. The latter holds, as $w_3 \in V(blue)$. Now we check the former. $M, w_3 \models K_{Ann}\neg K_{Bob} blue$ is true if and only if $\forall w : w_3 \sim_{Ann} w$ we have that $M, w \models \neg K_{Bob} blue$. Notice $\{w \mid w_3 \sim_{Ann} w\} = \{w_3\}$. So we need to check whether $M, w_3 \models \neg K_{Bob} blue$. This is true if there exists a w' with $w_3 \sim_{Bob} w'$ with $w' \in V(red)$. But this is false, so $M, w_3 \models K_{Ann}\neg K_{Bob} blue \wedge blue$ is false.

Question

"The game below is called the centipede game. We start in the choice node on the left. At each step, the player whose turn it is can choose between going right and going down:

$$
\begin{tikzcd}
\mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & \mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & (30, 30) \\
(5, 5) & (0, 15) & (35, 10) & (25, 35) &
\end{tikzcd}
$$

2.1) Intuitively, how would you play?

2.2) Calculate the backwards induction outcome.

2.3) Transform the game into normal form.

2.4) Reflect on the normal form game you obtained. What is the solution concept that corresponds to backwards induction?"

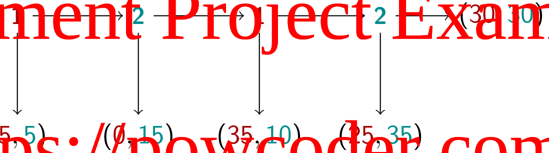
2.1) Intuitively, how would you play?

$$
\begin{tikzcd}
\mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & \mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & (30, 30) \\
(5, 5) & (0, 15) & (35, 10) & (25, 35) &
\end{tikzcd}
$$

Approach: the question is asking to try out the game and think what can happen.

Answer: if I think that my opponent is rational, and I'm rational and she knows I am and I know she knows... then I go down immediately.

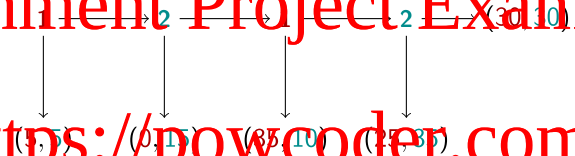
2.2) Calculate the backwards induction outcome.



Approach: list the strategies and the payoff and apply the definition

Answer: Each player has four strategies (DR), (DD), (RR), (RD). Each backwards induction outcome will have player 2 playing D before the terminal node, as $35 > 30$. But then player 1 will have to choose between 25 by choosing R and 35 by choosing D. So 1 will choose D before player 2. But then player 2 is choosing between 10 (R) and 15 (D), so will go D. Finally 1 will choose between 0 and 5, so 1 is going D in the beginning of the game.

2.3) Trasform the game into normal form.



Approach: Write the elements of the normal form game and translate each of them applying the definition.

Answer: The extensive game above corresponds to the normal form game (N, A, u) where:

- $N = \{1, 2\}$ (the players)
- $A = \{((x, y), (x', y')) \mid x, y, x', y' \in \{R, D\}\}$ (the action profiles)
- and u (the utility) is defined as follows...

$$\begin{tikzcd} 1 \arrow[r] \arrow[d] & 2 \arrow[r] \arrow[d] & 1 \arrow[r] \arrow[d] & 2 \arrow[r] \arrow[d] & (30, 30) \\ (5, 5) & (0, 15) & (35, 10) & (25, 35) & \end{tikzcd}$$

for $x, y, x', y' \in \{R, D\}$

$u_1((D, y), (x', y')) = u_2((D, y), (x', y')) = 5$

$u_1((R, y), (D, y')) = 0, \; u_2((R, y), (D, y')) = 15$

$u_1((R, D), (R, y')) = 35, \; u_2((R, D), (R, y')) = 10$

$u_1((R, R), (R, D)) = 25, \; u_2((R, R), (R, D)) = 35$

$u_1((R, R), (R, R)) = u_2((R, R), (R, R)) = 30$

Notice: Giving the matrix representation is a perfectly fine answer.

2.4) Reflect on the normal form game you obtained. What is the solution concept that corresponds to backwards induction?

$$
\begin{tikzcd}
\mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & \mathbf{1} \arrow[r] \arrow[d] & \mathbf{2} \arrow[r] \arrow[d] & (30, 30) \\
(5, 5) & (0, 15) & (35, 10) & (25, 35) &
\end{tikzcd}
$$

Approach: Think about how the backwards induction reasoning can be replicated in the normal form representation.

Answer: Iterated elimination of weakly dominated strategies. Take for instance (R, R) by 2. It is not strictly dominated. But if 1 plays (R, R) then it is! (namely by (R, D))

The question

1.1) Consider the following game:

	L	R
T	1, 5	10, 0
B	1, -10	-2, -20

1.1.a) Calculate all the pure strategy Nash equilibria.

1.1.b) Calculate all the mixed strategy Nash equilibria.

1.1.c) Calculate the outcomes surviving IESDS.

1.1.a) Calculate all the pure strategy Nash equilibria.

	L	R
T	1, 5	10, 0
B	1, -10	-2, -20

Approach: for each outcome, calculate whether that outcome is a pure NE.

Answer: TL is (no profitable deviation for either player), BL is (same), TR isn't (column is better off deviating to TL), BR isn't (column is better off deviating to BL)

1.1.b) Calculate all the mixed strategy Nash equilibria.

	L	R
T	1, 5	10, 0
B	1, −19	−2, −20

Approach: Look at the game first! and ask yourself if some actions are dominated: no mixed NE profile assigns non-zero probability to dominated actions!

Answer:
We know that no mixed strategy Nash equilibrium has column play R with non-zero probability. So, all of them have L played with probability 1. Given this, Row is indifferent between T and B. So there are infinitely many Nash equilibria, where Column plays L and Row plays any mixture of T and B.

1.1.c) Calculate the outcomes surviving IESDS.

	L	R
T	1, 5	10, 0
B	1, −19	−2, −20

Approach: Apply the algorithm

Answer:
Let the algorithm start by picking Column (order does not matter, by Gilboa's theorem). L strictly dominates R. So, we eliminate R. So the reduced game is $\{(TL),(BL)\}$, as no other strategy can be eliminated.

Question

Ann has a left shoe, Bob and Charles have a right shoe. A coalition is winning if it has a member with a left shoe and one with a right shoe.

- Model this scenario as a coalitional game.
- Calculate the core of the game.

Approach: list the elements of a coalitional game and apply the definitions.

Answer $N = \{A, B, C\}$
$v = \{A, B\} = \{A, C\} = \{A, B, C\} = 1$, $v(C') = 0$ for any other coalition C'.

Take now an allocation (x, y, z) where Ann gets less than 1, say $1 - p$. But then one between Bob and Charles is getting strictly less than p (remember the allocation has to sum up to $v(N) = 1$). Suppose it's Bob. Then Bob and Ann can deviate to a two person coalition and get a bit more both! (for instance splitting the profit of Charles). You can easily check that the allocation where Ann is getting everything is a stable imputation. So the core is nonempty and Ann is getting everything.

Question

3) Consider the following agents' preferences in a stable marriage problem (from top to bottom):

$w_1 : (m_2, m_3, m_1)$
$w_2 : (m_3, m_2, m_1)$
$w_3 : (m_1, m_3, m_2)$

$m_1 : (w_2, w_3, w_1)$
$m_2 : (w_2, w_3, w_1)$
$m_3 : (w_1, w_3, w_2)$

3a) Calculate the outcome of the deferred acceptance algorithm.

3b) Which agents would be better off by misrepresenting their preferences?

Answer

$w_1 : (m_2, m_3, m_1)$
$w_2 : (m_3, m_2, m_1)$
$w_3 : (m_1, m_3, m_2)$

$m_1 : (w_2, w_3, w_1)$
$m_2 : (w_2, w_3, w_1)$
$m_3 : (w_1, w_3, w_2)$

Answer

At round one, both m_1 and m_2 propose to w_2 and m_3 proposes to w_1. w_2 rejects m_1. At round two, only m_1 is rejected and proposes to w_3, who happily accepts.
The stable matching is $(m_1, w_3), (m_2, w_2), (m_3, w_1)$. No player is better off misrepresenting their preferences (by permuting their profiles).

Question

1a) Consider a seal-bid second price auction with one object. You value the object 10. There are two opponents, who bid natural numbers uniformly at random, within their budget. Both you and your opponents have a budget of 10.
Is truth-telling (bidding your value) a dominant strategy for you?

2a) At what price do you expect the object to be sold?

Question

1a) Consider a second price auction with one object. You value the object 10. There are two opponents, who bid natural numbers uniformly at random, within their budget. Both you and your opponents have a budget of 10.
Is truth telling a dominant strategy for you?

Answer Yes. If you win, you pay the price of the second-highest bidder. Therefore bidding more or bidding less won't be an improvement. If you lose, bidding more or bidding less won't be an improvement either. Bidding less won't make you win, and winning bidding more will make you pay more than your value.

2b) At what price do you expect the object to be sold?
Answer
6.818181818. Reasoning by cases. [1]

[1] There's 21 cases where there's a 10 involved so that would be the max [(Any number from 0 to 9, 10)=10 cases, (10 , 0 to 9)= another 10, (10,10)= last one for 21], then 19 cases where the max is 9 [(0 to 8, 9) = 9 cases , + (9, 0 to 8) = 9 cases + (9,9) = 1 case for 19] and you can keep going down, with the number of cases for each integer being max being 2 less, i.e, 17 cases where max = 8, 15 where max = 7 and so on for 3 cases where the max is 1 and only 1 case where the max is 0 (0,0). So then you can just average it by doing (10*21)+(9*19)+(8*17)+.....+(1*3)+(0*1) then dividing by total number of cases (11*11) = 6.818181818 recurring.

Suggested resources with exercises

Answer all the questions posed in the slides. Many of them are exercises.

Many other resources you can use.
Among our textbook, the following have very good exercises:

- For AI: Russell and Norvig's textbook.
- For Game Theory: Maschler, Solan, Zamir's textbook.