Multi-Agent Systems

Homework Assignment 3

MSc AI, VU

E.J. Pauwels

Version: November 24, 2023— Deadline: Wed November 29, 2023 (23h59)

3 Sequential Games with Perfect Information

3.1 Reduced centipede game

Consider a sequential 2-player game with the following game-tree: at each decision node the associated player needs to decide whether to continue (c) or stop (s). The tree (including utilities) and the players' rationality are common knowledge. Notice that the utility for **both** players increases along the game tree – and this is known to both players (perfect and complete information).

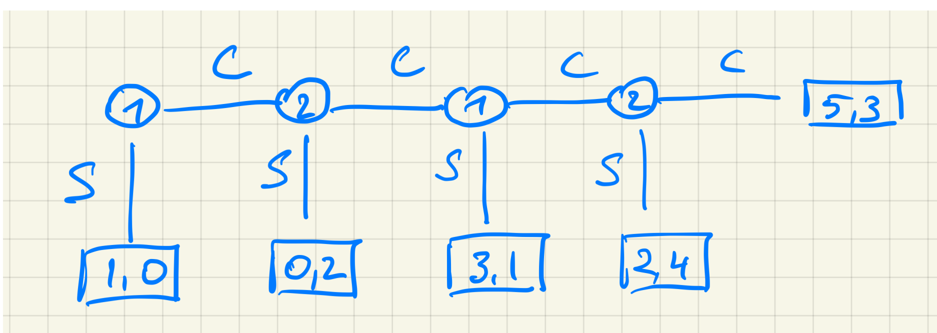


Figure 1: Shortened version of the centipede game.

Questions

1. Use backward induction to predict the (rational) outcome of this game. Does it make sense to you?

2. Write the normal form for this game and find all Nash equilibria in pure strategies (PNEs).

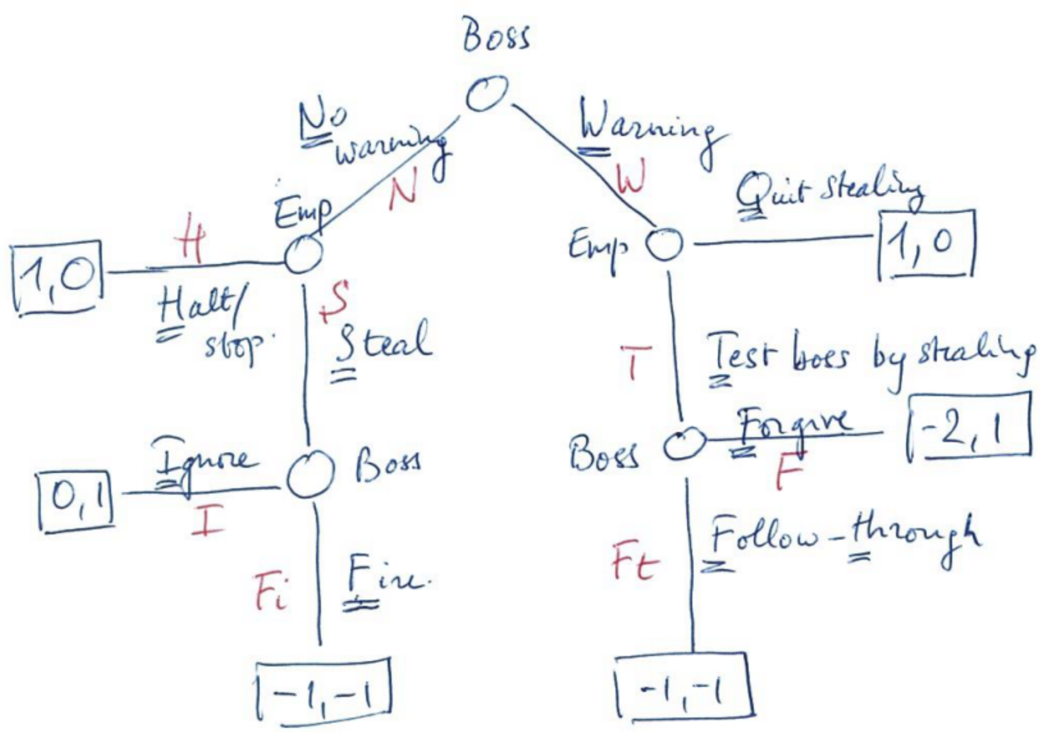
3. List all subgames and determine which of these PNEs are also subgame-perfect?

3.2 Boss and stealing employee

A boss notices that one of her employees has been stealing company material lately. The material was not all that valuable, so she is inclined to let it pass, preferring to keep the employee around rather than firing him and having to hire and retrain a replacement. Nevertheless she wants the stealing to stop.

She is therefore thinking to issue a warning at the next company meeting: the next person caught stealing company property will be fired immediately. She envisages the following game tree with pay-offs (see fig below).

1. Analyse this game using backward induction.
2. What are the pure actions for the two players (boss and employee)? Construct the normal form matrix.
3. Use this matrix to identify all the pure Nash equilibria of the normal form game.
4. Determine the subgame-perfect equilibrium (equilibria?) by eliminating all the Nash equilibria that fail to induce a NE in subgames.
5. Compare to the solution based on backward induction.



SOLUTIONS

3.1 Reduced centipede game

Backward induction (BI) predicts that the first player will immediately stop the game, yielding a utility of 1 for first player, 0 for 2nd. This does not feel like a good prediction for what

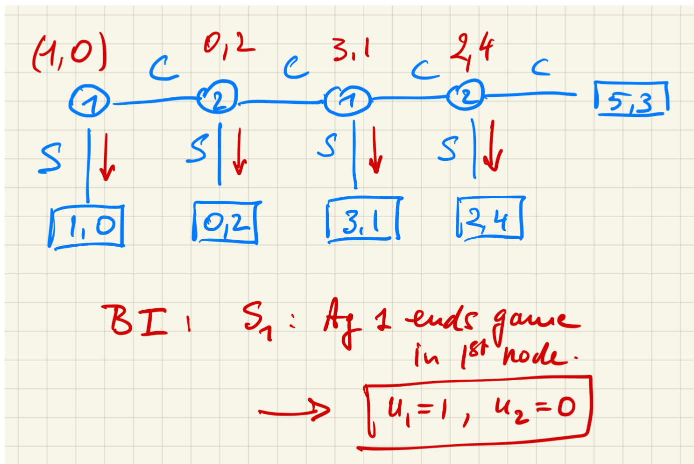


Figure 2: Centipede game with non-trivial subgames for future reference.

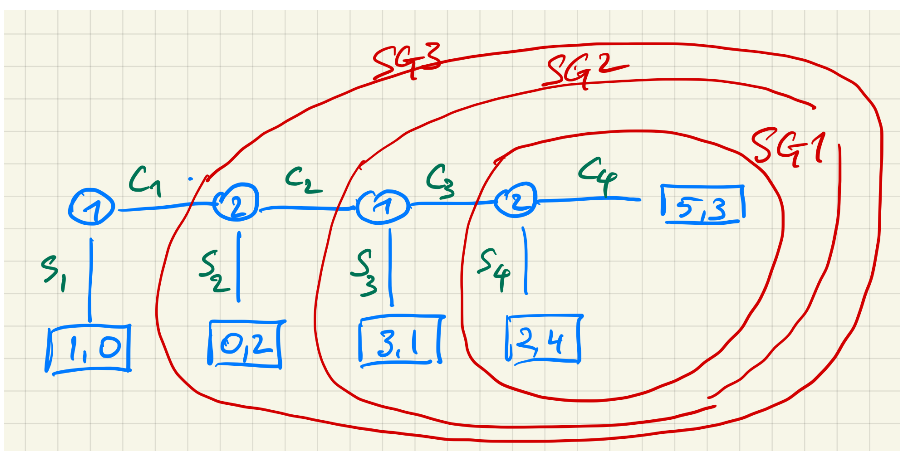


Figure 3: Centipede game with non-trivial subgames for future reference.

Each player has two decision nodes yielding the following pure strategies:

- player 1 : $\{s_1, c_1\} \times \{s_3, c_3\} = \{s_1 s_3, s_1 c_3, c_1 s_3, c_1 c_3\}$
- player 2 : $\{s_2, c_2\} \times \{s_4, c_4\} = \{s_2 s_4, s_2 c_4, c_2 s_4, c_2 c_4\}$

Conversion to normal form yields the following matrix, with four Nash equilibria as indicated.

	$s_2 s_4$	$s_2 c_4$	$c_2 s_4$	$c_2 c_4$
$s_1 s_3$	1) 1 0	2) 1 0	1 0	1 0
$s_1 c_3$	3) 1 0	4) 1 0	1 0	1 0
$c_1 s_3$	0 2	0 2	3 1	3 1
$c_1 c_3$	0 2	0 2	2 4	5 3

Figure 4: Centipede, full game in normal form, with 5 pure NE (numbered for future reference).

Next, we determine the pure NE's for the three non-trivial subgames:

SG1: $\boxed{s_4 \text{ is NE.}}$

SG2:

	c_4	s_4
c_3	5 3	2 4
s_3	3 1	3 1

$\boxed{NE(SG_2) = (s_3, s_4)}$

SG3: Player 1: s_3, c_3

Player 2: $\{s_2, c_2\} \times \{s_4, c_4\}$

$= \{s_2 s_4, s_2 c_4, c_2 s_4, c_2 c_4\}$

	$s_2 s_4$	$s_2 c_4$	$c_2 s_4$	$c_2 c_4$
s_3	0 2	0 2	3 1	3 1
c_3	0 2	0 2	2 4	5 3

$\boxed{NE(SG_3) = \{(s_3, s_2 s_4), (s_3, s_2 c_4)\}}$

Figure 5: PNEs for subgames

Checking for sub-game perfect Nash eq. Do the Nash equilibria in the game induce Nash eq. in all of the subgames? This is only the case for $NE = (s_1 s_3, s_2 s_4)$.

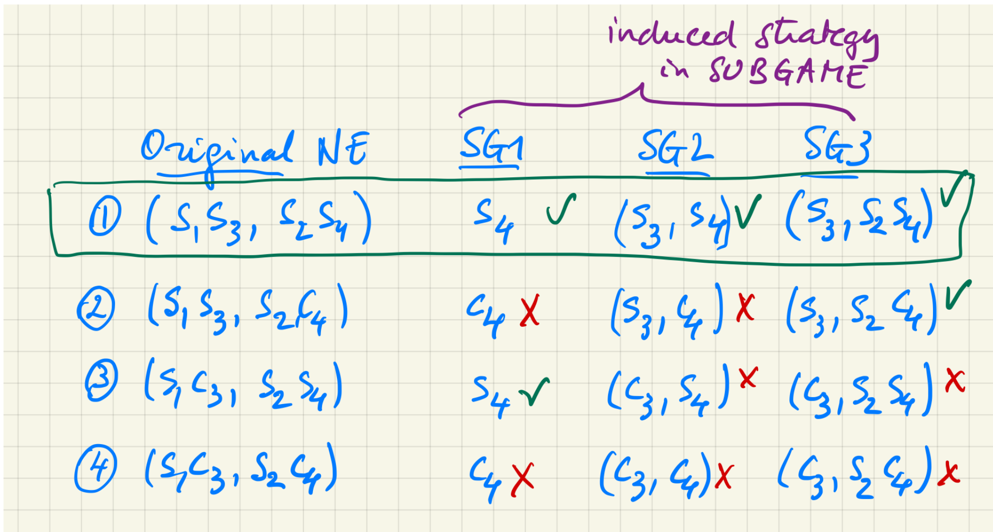


Figure 6: Checking for subgame perfection

3.2 Boss and stealing employee

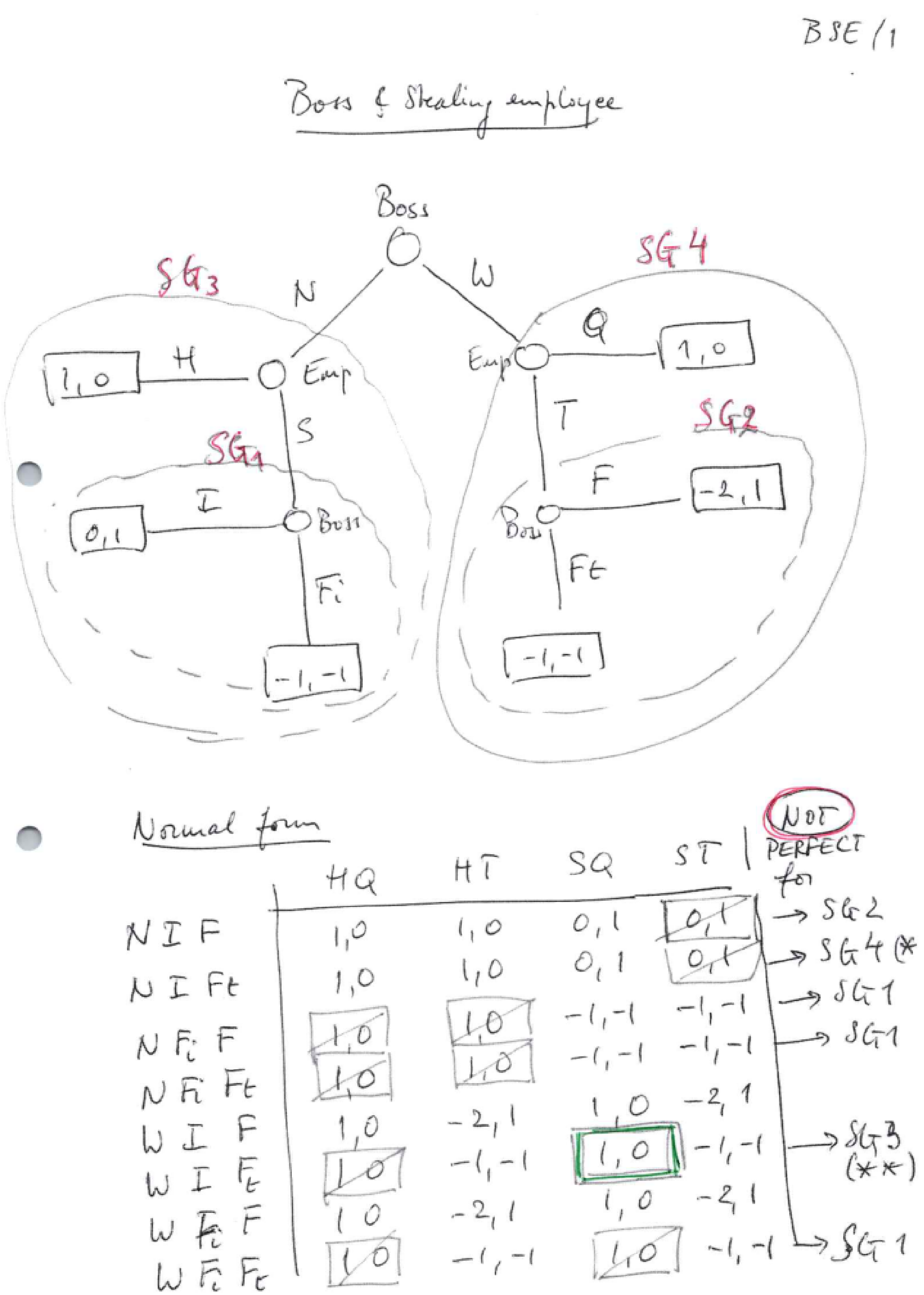


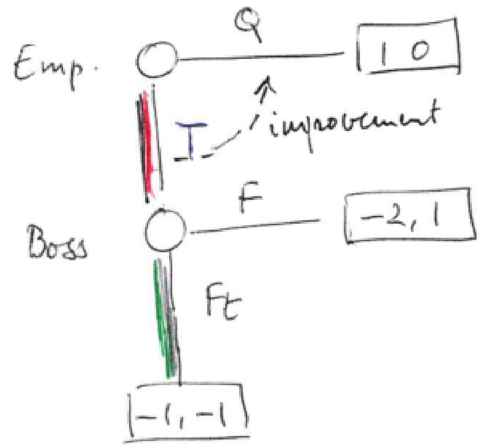
Figure 7: TOP: Game tree with subgames. Bottom: The normal form game with all Nash equilibria (black boxes) and the SPNE (boxed in green). Last column indicates in which subgame the NE fails to induce a subgame NE.

BSE/2

(*)

Consider (NI F_t, ST) for SG4 : induces the strategy below.

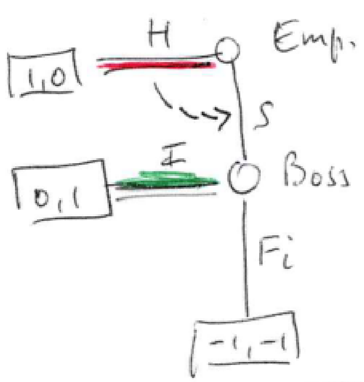
Boss Em.



Emp. wants to change T ⟶ Q
(payoff : -1 ⟶ 0)

Hence : NOT NE !

(**) (W, I, Ft ; HQ) in SG3



Emp. wants to change H ⟶ S
(payoff : 0 ⟶ 1)

Hence not NE !

Conclusion : (W I F_t , SQ) is SGPE.

Figure 8: Amplification: induced NE

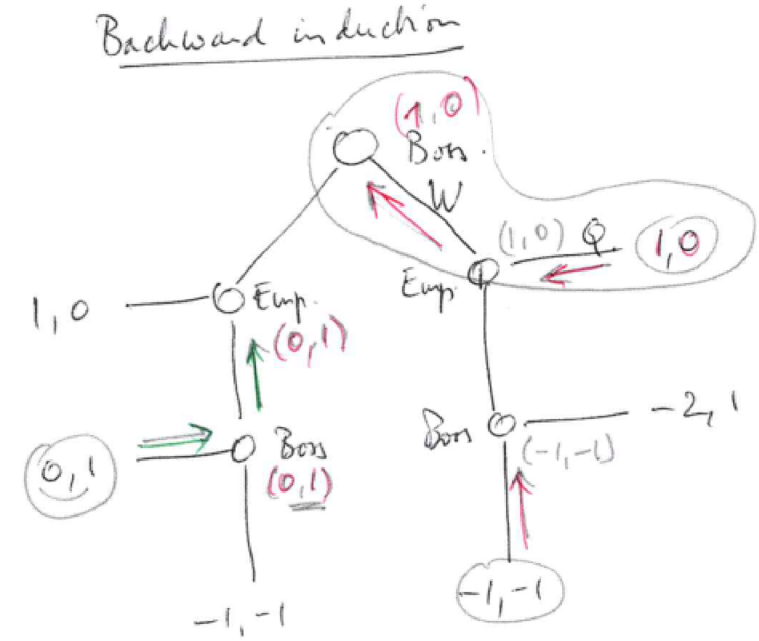


Figure 9: Backward induction solution: boss issues a warning, and employee quits stealing.