Maze games, proof games and some others

Wilfrid Hodges
Queen Mary, University of London
www.maths.qmul.ac.uk/~wilfrid/amsterdam05.ppt

Plan

1. The Dawkins question and its converse
2. Defeasible non-social aims
3. Some examples
4. Avoiding the Ignorance Requirement
5. Conclusions

1. The Dawkins question and its converse

"The whole purpose of our search ... is to discover
a suitable actor to play the leading role
in our metaphors of purpose.
We ... want to say, 'It is for the good of ...'.
Our quest in this chapter is for the right way
to complete that sentence."

Richard Dawkins, *The Extended Phenotype*,
OUP 1982 p. 81.

Generalised, the Dawkins question is:

Given a game G, what possible agents with what aims could be represented by the players in the game G?

For several well-known logical games,
the standard motivations for the agents
fail to answer the Dawkins question.
They don't match the rules of the game.

For Lorenzen proof games:

Wilfrid Hodges,
'Dialogue foundations: A sceptical look',
Proceedings of the Aristotelian Society, Supplementary volume 75 (2001) 17-32.

For Hintikka language games:

Wilfrid Hodges,
'The logic of quantifiers', *Schilpp Library of Living Philosophers volume on Jaakko Hintikka* (several years overdue).

I argued those cases in terms of the detailed structures of the games.

Today I start at the other end,
and aim for general principles.

DATA:

1. A set S of situations.
2. A set A of agents, each with an aim that makes sense in all situations in S.

PROBLEM:

To design a form of game G
with the agents in A as players, so that

For each situation s there is a game G(s);

For each agent α,

trying to win the game G(s)
equals
trying to achieve α's aim in situation s.

2. Defeasible non-social aims

We say that the aim of an agent is **indefeasible** if in every situation there is some way that the agent can achieve her aim.

Otherwise the aim is **defeasible**.

For simplicity I ignore degrees of achieving one's aim.

We say an agent's aim is **social** if it is in terms of interaction with other agents.

E.g. the agent aims to achieve maximal utility
in bargaining with other agents,
or aims to kill other agents,
or to prevent them achieving their aims.

Otherwise the agent's aim is **non-social**.

Even when the aim of an agent α is non-social and indefeasible,
actions of other agents may make it hard or impossible
for α to achieve her aim.

We begin with two-person games.

Since we are ignoring degrees of achieving an aim,
the payoff for each player is either win or lose.

In principle we allow both players to win,
though for the following maze games
this doesn’t happen.

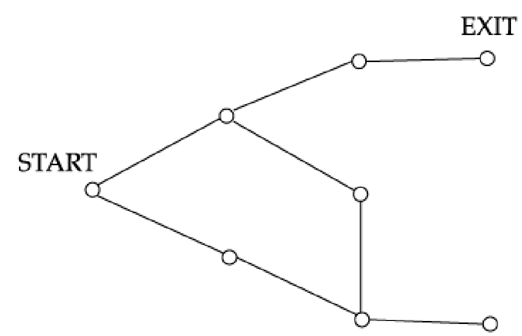
WORKING EXAMPLE: MAZE GAMES

We have a set M of mazes (the situations).
Each maze in M has a unique start;
some have one or more exits.

For each maze m,

the aim of $\exists$ is to show that there is a path from start to an exit in m;

the aim of $\forall$ is to show that there is no such path.

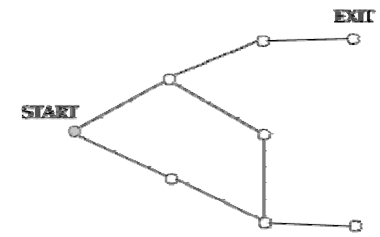


IGNORANCE REQUIREMENT

If an agent knows that her aim is unachievable, then no action of hers counts as a rational step towards achieving her aim.

But a player can’t react rationally to a situation if she has no information about the situation.

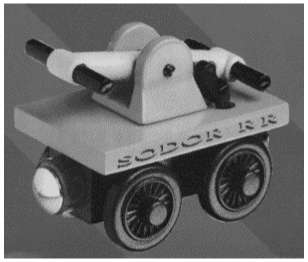
So we have to suppose that each player has partial information about the situation.
We must allow the players
to explore the situation as part of the game.



The players move alternately.
They both have the incentive to explore the maze.

Both players need exactly the same information.

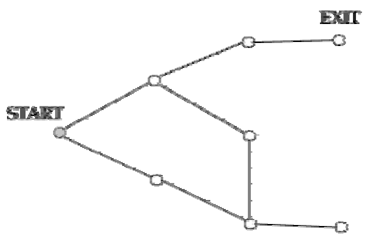
So the game is completely cooperative,
like a hand pump cart:



The players hope for different outcomes,
but this doesn’t affect their choices.

This is a game for two players.

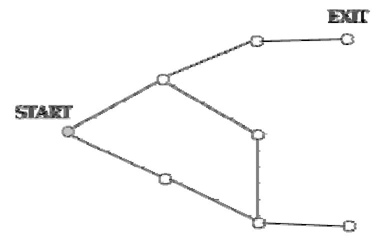
So each player is required to keep pursuing their aim
even if the other player is uncooperative.



Since the aims are non-social,
there seems no point in allowing one player to
frustrate the aims of the other player.
(It's entertaining but what does it tell us?)

So a better-designed game will allow each player
to add information at the edge of knowledge.

For example in maze games, the moves can be
to extend the surveyed part of the maze by one step.



The game does now represent the aims of the agents.

But their interaction is now so minimal
that the point of using a game to represent it
is not clear.

In short, games are not a helpful tool for
understanding agents with defeasible non-social aims.

This analysis assumed that the agents need the same information about the situation, and can seek it together.

This might not always apply, even in mazes. E.g. if ∃'s aim is to show that a Sudoku puzzle has a solution, and ∀'s aim is to show it doesn't, then ∃ might want to find a solution, whereas ∀ might want to do some mathematical calculations on the puzzle.

This would make the level of interaction between the players still less.

ANOTHER APPROACH

Give all the moves to ∃.
Then a winning strategy for ∀ records that no path through the maze reaches an exit.

This approach is mathematically helpful.
But it removes all connection with the agents' aims:

1. The aims of ∃ are irrelevant to the existence of a winning strategy for ∀.

2. ∀ does absolutely nothing, so his motivation can't be relevant.

3. Some examples

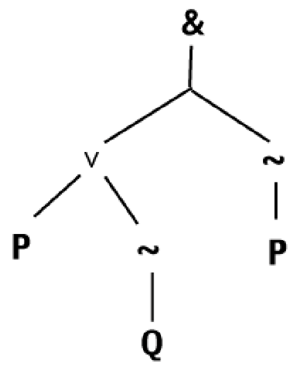
PROOF GAMES

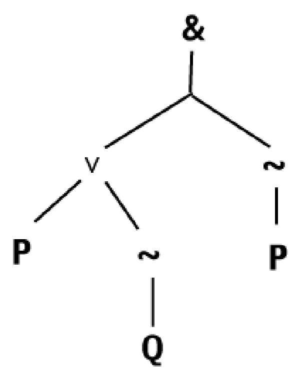
In a proof game, ∃ aims to show that a given sentence ϕ is provable, ∀ aims to show that it isn't.

Knowledge of ϕ can allow the players to tell whether their aims are achievable.
So by the Ignorance Requirement, we should assume ϕ is made available in steps.

Styles differ with the proof types.

In tableau proof games, ∃ wants to show a formula has no model, ∀ wants to show it has one.





This is a maze game:
the formula tree is the maze,
and the exits are unextendable open ‘branches’.

Who chooses at ∨?

By our previous analysis, it makes no difference.

Beware of arguments that assume ∀ knows which direction will lead to a model.

BACK AND FORTH GAMES

A situation consists of a pair of structures M and N.

Player ∀ ‘attempts to show that [M and N] can be distinguished’ and player ∃ ‘wishes to establish that they are equivalent’. (Stirling, *Modal and Temporal Properties of Processes* p. 57.)

The players take turns to choose appropriate items in M and N.

∀ wins as soon as the elements chosen from M don’t match those chosen from N.

With these motivations,
the players must not know what M and N are,
except for what they discover during the play.

So in general their choices will be random.

There are theorems saying that ∃ has a winning strategy in these games if and only if M and N are equivalent (in the relevant logic).

But these winning strategies are purely abstract; there is no way that ∃ could be led to them by her aim.

4. Avoiding the Ignorance Requirement

If the agents' aims are either social or indefeasible, or both, then the Ignorance Requirement lapses.

Then one can make some sensible games.

SOCIAL AIMS

Immerman, *Descriptive Complexity* p. 91,
describes the aims in a back and forth game thus:
∀ 'tries to point out a difference between the two structures', and
∃ 'tries to match his moves so that the differences between [the structures] are hidden'.

The aim of ∃ is no longer non-social.
Even if M and N are not equivalent,
∃ still has the task of matching ∀'s moves.
So there is no need for her to be ignorant of M and N.

However, ∀'s aim is irrelevant to the existence of a winning strategy for ∃.

∀ might as well be blind Nature.

But we keep assigning motives to ∀.

Presumably this is to give ∃ the social aim of showing up ∀'s incompetence?
This aim is usually not mentioned
(except by feminists and others who object to the personalising of mathematics).

Can proof games be rescued likewise?

Not that I know.
But one early inspiration for proof games was the medieval obligatio game.
In this game Respondens tries to maintain rationality in a very difficult situation, and
Opponens tries to trap Respondens into irrationality (by forcing him into a contradiction that he could have avoided).

The aim of Opponens is clearly social.
The rules of the game are only partly formalised, because rationality is not a formal notion.

INDEFEASIBLE AIMS

When an agent's aim is indefeasible,
we can imagine the agent pursuing the aim
in full knowledge of the situation
but having to adjust to actions of other agents.

One special and important case is where
all agents have indefeasible aims for which
they have winning strategies against each other.
Cf. the Banach-Mazur games in my
Building Models by Games.

5. Conclusions

CONCLUSIONS

1. We should be very suspicious of games
that are claimed to represent aims of the players,
when those aims are defeasible and non-social.

CONCLUSIONS

2. When the aims of agents are claimed to motivate the form of a game,
we need to say what information is available to the agents while they play the game.

Some motivations in the literature make no sense under perfect information.

CONCLUSIONS

3. When the real interest of the game lies in the winning strategies of one player,
the aims of the other players are an entertaining irrelevance.

Compare Wittgenstein's falling leaf
that says to itself
'Now I'll go this way, now I'll go that way'.