



西安电子科技大学  
XIDIAN UNIVERSITY

# 智能控制

## 基于博弈的控制

吴建设

jshwu@mail.xidian.edu.cn

人工智能学院

PPT下载地址：

<https://web.xidian.edu.cn/jshwu/teach.html>



## 引子

- ▶ 20 世纪著名的诺贝尔经济学奖得主萨缪尔森 (Paul A. Samuelson, 1915-2009) 曾经说过：

要想在现代社会做一个有文化的人，  
必须对博弈论有一个大致的了解。
- ▶ 萨缪尔森于1970年（55岁）获得诺贝尔经济学奖，是第一个获得诺贝尔经济学奖的美国人。

# 博弈论简介

## 引子

- 20 世纪著名的诺贝尔经济学奖得主萨缪尔森 (Paul A. Samuelson, 1915-2009) 曾经说过：

要想在现代社会做一个有文化的人，  
必须对博弈论有一个大致的了解。

- 萨缪尔森于1970年（55岁）获得诺贝尔经济学奖，是第一个获得诺贝尔经济学奖的美国人。

# 博弈论简介

## 1、什么是博弈论

一个经典的例子：囚徒困境

- 两个共谋犯罪的人被关入监狱，不能互相沟通情况。如果两个人都不揭发对方，则由于证据不确定，每个人都坐牢一年；若一人揭发，而另一人沉默，则揭发者因为立功而立即获释，沉默者因不合作而入狱十年；若互相揭发，则因证据确实，二者都判刑五年。
- 由于囚徒无法信任对方，因此倾向于互相揭发，而不是同守沉默。最终导致纳什均衡仅落在非合作点上。

# 博弈论简介

## 1、什么是博弈论 (Game Theory)

一个经典的例子：囚徒困境 (Prisoner's Dilemma)。

	D	C
D	P P	T S
C	S T	R R

	D	C
D	-5 -5	0 -10
C	-10 0	-1 -1

$$T > R > P > S, \quad 2R > T + S$$

# 博弈论简介

## 2、囚徒困境的启示

- (D, D) 这个状态，谁也不能通过单方面改变自己的策略增加自己的收益。因此，谁也不会主动离开这个状态，即改变自己的策略为C。
- 当然，这两个人可以订立攻守同盟，即合作，这时的博弈状态是 (C, C)。

# 博弈论简介

## 2、囚徒困境的启示（续）

- (C, C)这个状态是不稳定的，任何人都能通过单方面改变自己的策略增加自己的收益。因此，每个人都有意愿离开这个状态，即改变自己的策略为D。
- 囚徒困境反映了个人理性与集体利益的矛盾。

### 3、纳什均衡

假设有 $n$ 个玩家 (agent, player) 参与博弈，给定其他人策略的条件下，每个玩家选择自己的最优策略（个人最优策略可能依赖于也可能不依赖于他人的策略），从而使自己效用最大化。所有局中人策略构成一个策略组合 (Strategy Profile)。

各个参与人的在某个时间点的策略的有序集称为策略组合，比如囚徒困境中的 (D, D), (C, C), (C, D) 都是策略组合。

### 3、纳什均衡

纳什均衡指的是这样一种策略组合，即在给定别人策略的情况下，没有人愿意单方面改变自己的策略，从而打破这种均衡。

## 4、纳什均衡的数学表示

博弈的数学表示：

$$G = \{ S_1, \dots, S_n : u_1, \dots, u_n \}$$

其中， $S_i$ 表示参与人 $i$ 的策略集合，也叫策略空间。

比如： $S_i = \{C, D\}, i=1, \dots, n$ 。

$u_i$ 是给定各个参与人的策略的情况下，参与人 $i$ 的收益的计算函数，也叫收益函数

$$u_i(s_1, \dots, s_n) = f(s_1, \dots, s_n),$$

其中 $s_i$ 是参与人 $i$ 所采用的策略。

## 4、纳什均衡的数学表示

纳什均衡的定义：在博弈

$$G = \{ S_1, \dots, S_n : u_1, \dots, u_n \}$$

中，如果由各个博弈方的各一个策略组成的某个策略组合  $(s_1^*, \dots, s_n^*)$  中，任一博弈方  $i$  的策略  $s_i^*$  都是对其余博弈方策略的组合  $(s_1^*, \dots, s_{i-1}^*, s_{i+1}^*, \dots, s_n^*)$  的最佳对策，即

$$u_i(s_1^*, \dots, s_i^*, \dots, s_n^*) \geq u_i(s_1^*, \dots, s_i, \dots, s_n^*)$$

对任意  $s_i \in S_i$  都成立，则称  $(s_1^*, \dots, s_i^*, \dots, s_n^*)$  为  $G$  的一个纳什均衡。

# 纳什均衡的一致预测性质

**一致预测：** 如果所有博弈方都预测一个特定博弈结果会出现，所有博弈方都不会选择与预测结果不一致的策略，即没有哪个博弈方有偏离这个预测结果的愿望，因此预测结果会成为博弈的最终结果。

- ▶ 只有纳什均衡才具有一致预测的性质。
- ▶ 一致预测性是纳什均衡的本质属性。
- ▶ 一致预测并不意味着一定能准确预测，因为有多重均衡，预测不一致的可能。

## 4、纳什均衡的意义

- 纳什均衡首先对 “看不见的手” 的原理 提出挑战。
- “看不见的手” 的原理最早是由亚当·斯密在《国富论》中提出来的。
- 斯密认为，每个人都应利用好自已的资本，使之产生最大的价值。从主观上讲，这个人并不想增进公共福利，更不知道他实际上增加了多少公共福利，他所追求的仅仅是个人的利益所得，但他这样做的时候，有一只看不见的手在引导着他去帮助实现增进社会福利的目标，而这种目标并非是他本意想追求的东西。

## 4、纳什均衡的意义

- ▶ 通过追求个人利益，却无意识地增进了社会利益，其效果比真地想促进社会利益时所得到的效果要好。
- ▶ 斯密之所以提出了这一论断，他认为人们都有“利己心”，是“利己心”驱使着人们去获得最大利益，每个人都得到了利益，那么社会也就得到了利益，因为财富是所有国民对必需品和享用品的消费。这就是斯密“看不见的手”的实质。
- ▶ 但是我们可以从“纳什均衡”中引出“看不见的手”原理的一个悖论：从利己目的出发，结果损人不利己，既不利己也不利他。
- ▶ 从这个意义上说，“纳什均衡”实际上动摇了西方经济学的基石。

## 4、纳什均衡的意义

类似的思想也出现在我国战国时期儒家思想家荀子的作品中。荀子的《荀子·富国》：

“足国之道，节用裕民而善臧其余。节用以礼，裕民以政。彼裕民，故多余，裕民则民富，民富则田肥以易，田肥以易则出实百倍”

# 荀子



荀子 约公元前313—公元前238

名况，字卿，因荀与孙二字古音相通，故又称孙卿。中国战国时期赵国猗氏（今山西安泽）人，汉族。著名思想家、文学家、政治家，儒家代表人物之一，时人尊称“荀卿”。曾三次出齐国门下学宫的祭酒，后为楚兰陵（今山东兰陵）令。荀子对儒家思想有所发展，提倡性恶论，常被与孟子的性善论比较。

对重整儒家典籍也有相当的贡献。

治家國學



弘孝国学

## 4、纳什均衡的意义

“足国之道，节用裕民而善臧其余。  
节用以礼，裕民以政。彼裕民，故多  
余，裕民则民富，民富则田肥以易，  
田肥以易则出实百倍”

翻译：使国家富足的办法途径是，节约用度，使百姓富裕，并且善于储备那些节余的东西。用礼法制度节约用度，用政策使百姓富裕。因为民众富裕了，就会有节余。百姓丰衣足食，土地就肥沃而整治得好，土地肥沃而整治得好，收成就会增加百倍。

# 荀子



荀子 约公元前313——公元前238

名况，字卿，因“荀”与“孙”二字古

音相通，故又称孙卿。中国战国时期赵

国猗氏（今山西安泽）人，汉族。著名思

想家、文学家、政治家，儒家代表人物

之一，时人尊称“荀卿”。曾三次出齐

国稷下学宫的祭酒，后为楚兰陵（今山东

兰陵）令。荀子对儒家思想有所发展，提

倡性恶论，常被与孟子的性善论比较。

对重整儒家典籍也有相当的贡献。

成就国家

# 定义：什么是博弈论

- 博弈论又称对策论，它使用严谨的数学模型研究冲突对抗条件下的最优决策问题、是研究竞争的逻辑和规律的数学分支。
- 简单的说，博弈论是研究决策主体在给定信息结构条件下，如何决策以使自己的效用最大化，以及不同决策主体之间决策的均衡。

# 定义：什么是博弈论

- 博弈的字面意思就是赌博、下棋，赌博和下棋就是游戏了，所以也有人把博弈论叫做游戏理论。
- 赌博和游戏常常要千方百计的应付对手，所以博弈论有人也称为对策论。

## 5、博弈的要素

➤ **参与人**（玩家，Player，Agent）。

是博弈的参与主体，他通过选择自己的行动（策略）来使自己的收益最大化。参与人可以是自然人，企业，国家等。

在基于博弈的优化算法中，参与人代表算法中的某个元素。

## 5、博弈的要素

- ▶ 参与人的策略集。

是参与人可供选择的策略的集合。

比如囚徒困境中的欺骗和合作。

- ▶ 收益（效用，支付）

在给定的策略组合下，参与人得到的确定的效用水平，或是参与人得到的期望效用水平。

收益由收益函数来计算。

- ▶ 行动。

参与人在博弈的某个时间点的策略变量。

## 5、博弈的要素

### ➤ 行动的顺序

在动态博弈中，行动的顺序对博弈的结果有重要影响。同样的策略集，行动的顺序不同，每个人的最优策略就不同，博弈的结果就不同。

尤其在不完全信息博弈中，后行动者根据先行动者的行动来获取信息。

## 5、博弈的要素

### ➤ 信息

参与人有关博弈的知识，特别是关于其它参与人的特征和行动的知识，以及自然的选择

**完美信息**是指参与人对其它参与人的行动选择有准确的掌握。

## 5、博弈的要素

### ➤ 共同知识

所有参与人知道的每一步的信息集。

### ➤ 策略（战略）

- 参与人在给定信息集情况下的行动规则，它规定参与人在什么情况下选择什么行动。
- 策略与行动：策略是行动的规则而不是行动本身。
- 在静态博弈中，策略和行动是相同的。
- 策略必须是完备的：必须给出参与人在每一种可能的情况下的行动选择。

## 5、博弈的要素

### ➤ 均衡

指所有参与人的最优策略组合。