博弈控制（三）

多重均衡与优化

吴建设

多重均衡与优化

1、 占优策略与智猪博弈

2、博弈的多重纳什均衡

3、帕累托最优均衡

4、帕累托最优均衡与纳什均衡的关系

5、如何得到帕累托最优均衡

1、占优策略与智猪博弈

- 在博弈论（Game Theory）中，“智猪博弈”是一个著名的例子。假设猪圈里有一头大猪、一头小猪。猪圈的一头有猪食槽，另一头安装着控制猪食供应的按钮，按一下按钮会有10个单位的猪食进槽，但是谁按按钮就会首先付出2个单位的成本，若大猪先到槽边，大小猪吃到食物的收益比是9：1；同时到槽边，收益比是7：3；小猪先到槽边，收益比是6：4。那么，在两头猪都有智慧的前提下，最终结果是怎样的？。

智猪博弈

按一下按钮会有10个单位的猪食进槽，但是谁按按钮就会首先付出2个单位的成本，大小猪吃到食物的收益比是9：1；同时到槽边，收益比是7：3；小猪先到槽边，收益比是6：4。

大猪 \ 小猪	按	等待
按	**5，1**	**4，4**
等待	**9，-1**	**0，0**

1、占优策略与智猪博弈

占优策略：　又一个例子：

两家公司，A和B，在考虑是否通过广告促销。它们的利润额将依赖于哪一家公司做广告，或者两家公司都做广告，或者两家公司都不做广告。这些可能性和相应的利润额被总结在旁边的矩阵里。

		厂商B	
		做广告	不做广告
厂商A	做广告	5,5	15,0
	不做广告	0,15	10,10

厂商A和厂商B的广告博弈

1、占优策略与智猪博弈

- 对A，　无论B怎么做，做广告都是最优的。所以做广告是A的占优策略。
- 对B：无论A怎么做，做广告也都是最优的。 所以做广告也是B的占优策略。
- 结论：两家厂商都应该做广告。

1、占优策略与智猪博弈

- 定义：在参与人各自的策略集中，如果存在一个与其他竞争对手可能采取的策略无关的最优选择，则称其为占优策略(Dominant Strategy)，与之相对的其他策略则为劣势策略。占优策略是博弈论（game theory）中的专业术语，所谓的占优策略就是指无论博弈对手如何行动都属于本人最佳选择的策略。

1、占优策略与智猪博弈

智猪博弈的其它例子：

- “搭便车”现象
- 穷人和富人修路博弈
- 大股东对管理者的监督
- 俗语“天塌下来有大个子顶着”

占优策略和纳什均衡的比较

1. 占优策略:“不管你怎么做，我所做的都是我能做得最好的。”
2. 纳什均衡:
   – “给定你的做法后，我所做的是我能做得最好的。”
   – 如果你有占优策略, 你可以使用此策略, 以不变应万变;
   – 如果你没有占优策略, 你必须随机应变。在达到了纳什均衡之后, 所有参与者都没有动机想再变了。

2、博弈的多重纳什均衡

交通博弈

	靠左行	靠右行
靠左行	$\underline{1}$，$\underline{1}$	-1，-1
靠右行	-1，-1	$\underline{1}$，$\underline{1}$

2、博弈的多重纳什均衡

约会博弈

	芭蕾舞	足球场
芭蕾舞	1，2	0，0
足球场	0，0	2，1

2、博弈的多重纳什均衡

进门博弈

	先进	后进
先进	-1，-1	<u>2</u>，<u>1</u>
后进	<u>1</u>，<u>2</u>	-1，-1

多重纳什均衡：产品标准

- 许多博弈可能有多个纳什均衡

	3.5“	5.5“
3.5“	8, 8	3, 2
5.5“	2, 3	6, 6

这个博弈被称为“协调博弈”：有两个纯战略纳什均衡，一个混合战略均衡。哪一个将出现呢？

资源争夺博弈（鹰鸽博弈)

	Hawk	Dove
Hawk	-1，-1	<u>10</u>，<u>0</u>
Dove	<u>0</u>，<u>10</u>	5，5

在只有鸽子一个苞谷场里，突然加入的鹰将大大获益，并吸引同伴加入。但结果不是鹰将鸽逐出苞谷场，而是一定比例共存，因为鹰群增加一只鹰的边际收益趋零时（鹰群发生内斗），均衡将到来。

3、帕累托最优均衡

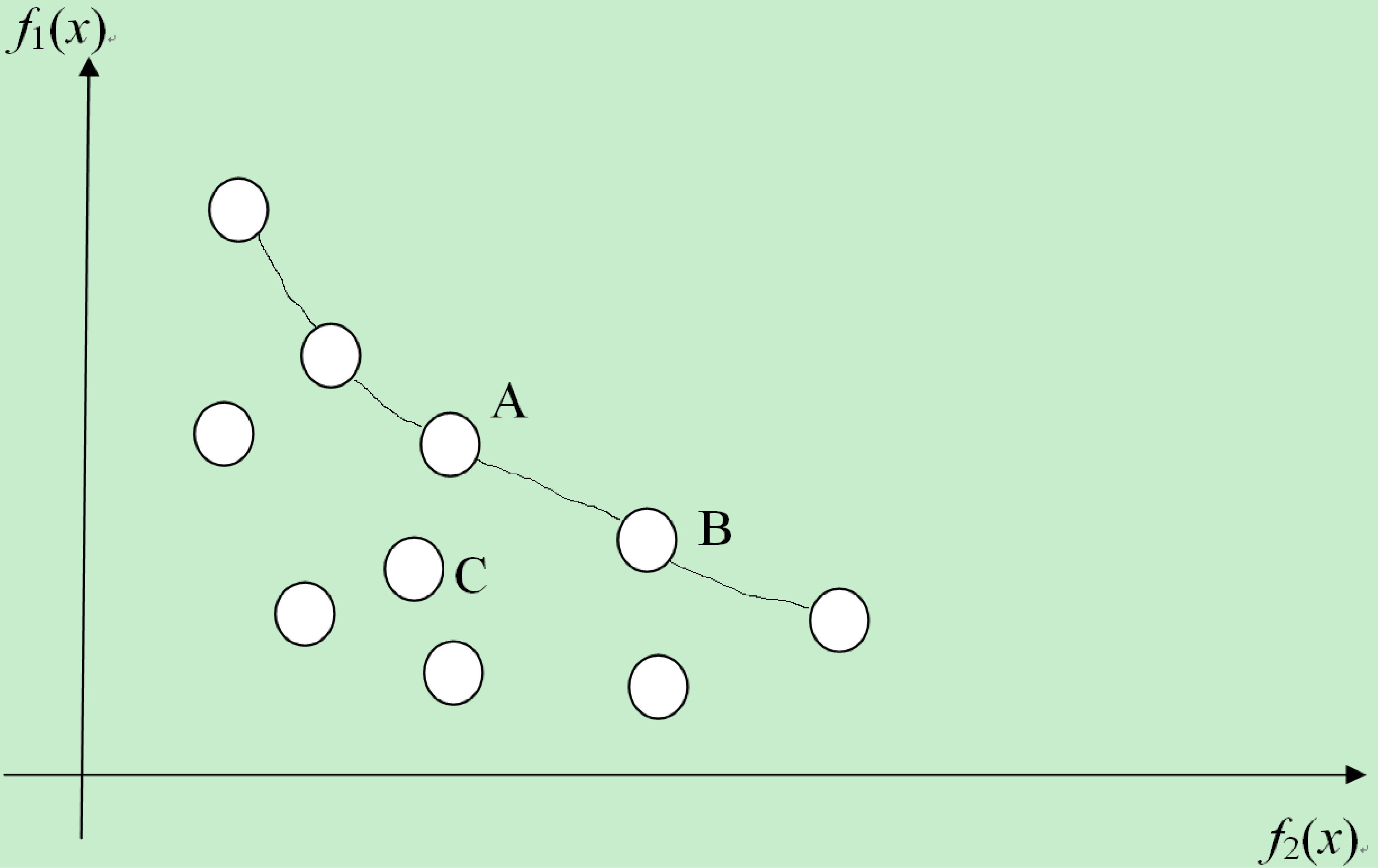
- 帕累托（ 1848年7月15日- 1923年8月19日）：经济学家、社会学家。洛桑大学政治经济学教授，论著有《政治经济学讲义》、《普遍社会学》、《社会主义体制》、《事实与理论》、《民主制的变革》，当过意大利铁路公司的总经理，曾出任（B.墨索里尼的）意大利政府驻国联代表。

3、帕累托最优均衡

- 如果从一种策略组合到另一种策略组合的变化中，在没有使任何人境况变坏（收益变少）的前提下，使得至少一个人变得更好，这就是帕累托改善。
- 帕累托最优的状态就是不可能再有更多的帕累托改善的策略组合；换句话说，不可能再改善某些人的境况，而不使任何其他人受损。

3、帕累托最优均衡

- 帕累托最优是指优化问题（例如资源分配问题）的一种理想状态，即假定固有的一群人和可分配的资源，从一种分配状态到另一种状态的变化中，在没有使任何人境况变坏的前提下，也不可能再使某些人的处境变好。换句话说，就是不可能在不损害一些人的利益的情况下，去增加另一些人的利益。



在图示的例子中，f1是参与人1的收益，f2是参与人2的收益，状态 C 不在帕累托前沿面(Pareto Frontier)因为存在帕累托改进使它改进到 A 或者 B 点。

帕累托最优均衡与纳什均衡的关系

- **帕累托最优均衡：满足帕累托最优条件的均衡**
- **纳什均衡是不是均为帕累托最优？**
  - 单从逻辑上来分析，能不能得出纳什均衡（博弈的一种结果）都是帕累托最优的?
  - 事实是，纳什均衡不一定是帕累托最优。产品标准博弈和囚徒困境博弈的结果就是例子（见下页）

3、帕累托最优均衡与纳什均衡的关系

（3.5”, 3.5”) 帕累托优于(5.5”, 5.5”)

	3.5”	5.5”
3.5”	8, 8	3, 2
5.5”	2, 3	6, 6

协商可以
帮助协调到一个
帕累托最优均衡

帕累托最优均衡与纳什均衡的关系

- 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，2R>T+S

- 帕累托最优指的是：在给定现有资源条件下，不存在任何其他配置结果使某些人情况更好，而又不使任何其他人处境更坏。显然，在上述囚徒困境模型当中，甲乙两个人都从理性的角度出发，追求自身效用的最大，结果是双方不合作，都认罪了。也就是说，实现了纳什均衡。
- 但是，如果甲与乙合作，产生的结果要比双方不合作好得多。也就是说，如果双方合作，就存在帕累托改进（帕累托改进是指一种变化，在没有使任何人境况变坏的前提下，使得至少一个人变得更好。帕累托改进是达到帕累托最优的路径和方法）。

3、帕累托最优均衡与纳什均衡的关系

- 结论

纳什均衡不一定是帕累托最优均衡

帕累托最优状态也不一定是纳什均衡

有一些纳什均衡是帕累托最优均衡

- 启示

博弈是求解帕累托优化解的一条途径：将博弈的纳什均衡设计成优化问题的纳什均衡。

锁定效应与路径依赖

	3.5’’	5.5’’
3.5’’	<u>8</u>, <u>8</u>	3, 2
5.5’’	2, 3	<u>6</u>, <u>6</u>

- 达成协议有难度，最初的非帕累托均衡可以被锁定（lock-in), 导致路径依赖（path dependence)：结果是整个社会被锁定在现有的产品（标准）（非帕累托最优）。

交通博弈与交通规则

	靠左行	靠右行
靠左行	1，1	-1，-1
靠右行	-1，-1	1，1

交通规则的演变

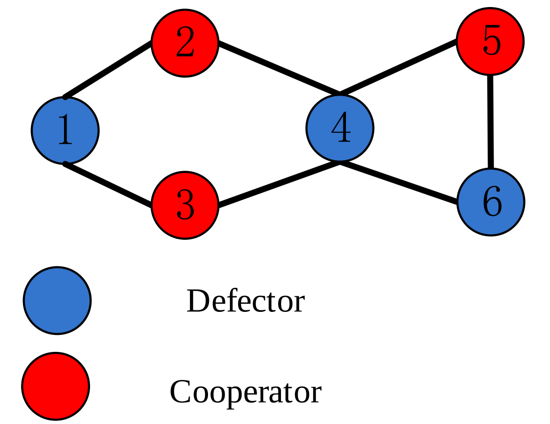
- 在多个纳什均衡之间不存在优劣之分时，偶然事件对选择具有重要意义；
- 从历史上来考察，许多交通规则一开始并不体现为法律，而是长期演化而来的。在欧洲大陆的早期，道路行走规范是非常地方化的，有些地方采用靠左走的习惯，有些地方采用靠右走的习惯，是不统一的。只是随着道路的增加和地区间交往的扩大，地方性的习惯才逐步演变为区域性的规范，然后有演变为全国性的规范。但直到１９世纪前，道路规则也仅仅是作为规范而得到遵守，而不是作为交通法律而得到执行。现在欧洲大陆的靠右走的规则是在法国兴起的。

网络博弈：多参与人博弈

比如图中6参与人进行雪堆博弈，蓝色代表所选策略为D,红色代表所选策略为C,如何计算各参与人的收益。

雪堆博弈的收益矩阵为（*0<r<1*）：

	C	D
C	1,1	$1-r$, $1+r$
D	$1+r$, $1-r$	0, 0



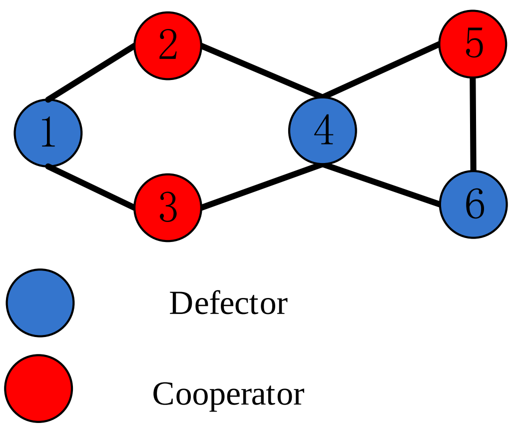
如果r>0.5, 上图是一个纳什均衡

网络博弈：多参与人博弈

演化博弈：

从一个随机的初始状态开始，博弈人依次调整自己的策略使自己的收益最大化，整个网络处于一个演化的过程中。

	C	D
C	1,1	$1-r$, $1+r$
D	$1+r$, $1-r$	0, 0



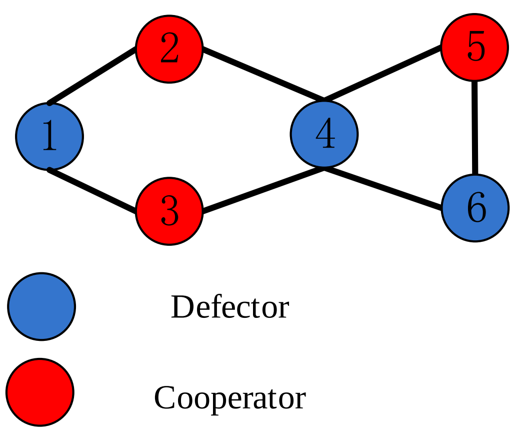
如果r>0.5, 上图是一个纳什均衡

网络博弈：多参与人博弈

如何找到网络博弈的纳什均衡：

1. 随机给每个节点初试状态设置为C或D
2. 依次对每个节点计算其采用C或者D的收益，改变其策略使其收益最大
3. 重复2的过程到每个节点的状态不再改变。

	C	D
C	1,1	$1-r$, $1+r$
D	$1+r$, $1-r$	0, 0



基于博弈的优化算法

优化算法的结果是得到帕累托最优。

- 方法一：求纳什均衡

  每次改变一个参与人的策略，增加自己的收益。

  - 适用问题：帕累托最优就是博弈的纳什均衡
  - 算法设计思路：恰当的设计博弈，使优化问题的解成为博弈的纳什均衡。
  - 举例：最小节点覆盖问题。

基于博弈的优化算法

优化算法的目标是得到帕累托最优

- 方法二：求帕累托改进。

  每次改变两个、三个参与人的策略，不减少其它人的收益的条件下，增加自己的收益。

  - **适用问题：** 帕累托最优虽然不是纳什均衡，但却是纳什均衡的改善。
  - 算法设计思路：分析帕累托最优与纳什均衡的关系。
  - 举例：囚徒困境

- 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， 2R>T+S

从博弈的角度，博弈人不愿意单独改变自己的策略，但从优化的角度，双方可以同时改变策略。因而从纳什均衡出发可以搜索到(C，C)。

基于博弈的优化算法

优化算法的结果是得到帕累托最优。

- 方法三：与其它优化方法相互结合：灵活多变、具体问题具体设计。

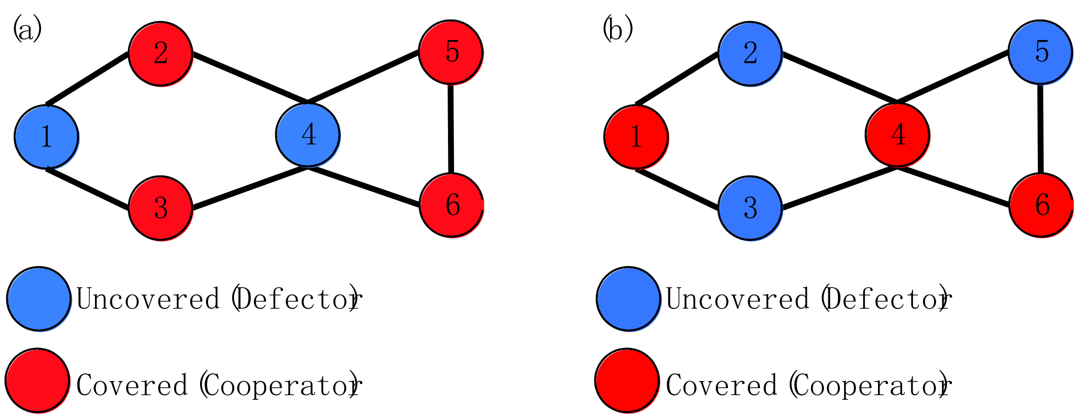
博弈优化实例：网络节点最小覆盖问题

网络节点最小覆盖问题:

是一个著名组合优化问题，其目的在于找出给定网络的最小节点集合以覆盖所有的边。一个节点覆盖其所有连边。

极小节点覆盖：节点集合中去掉任何一个点，就不能覆盖网络所有边。

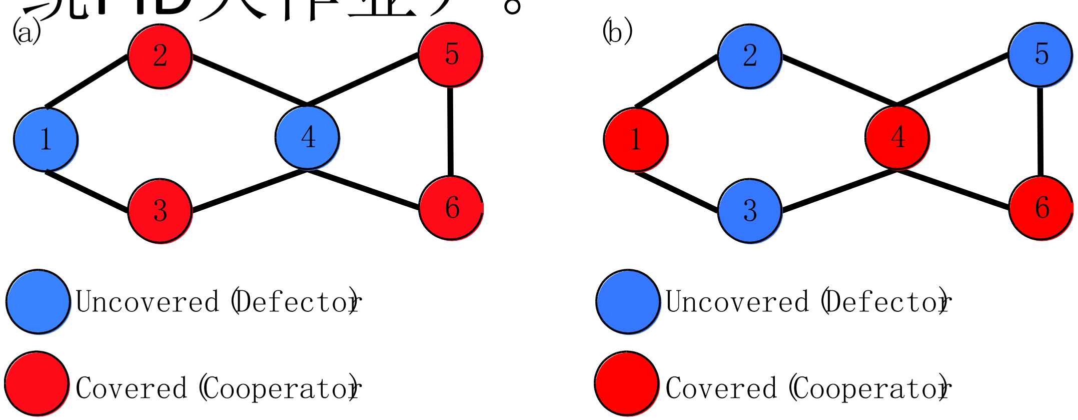
最小节点覆盖：极小节点覆盖中节点数目最少的。



红颜色节点集合为覆盖节点集合

博弈优化实例：网络节点最小覆盖问题

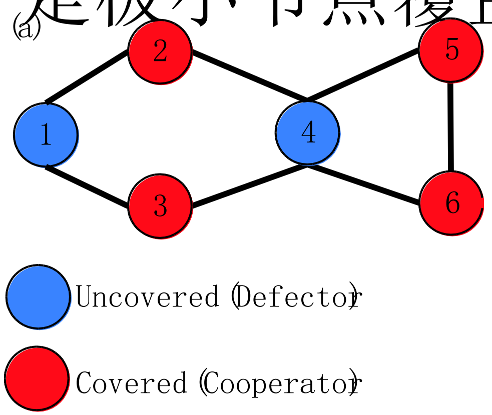
可以验证：当雪堆博弈满足$r<1/k_{max}$ 时（k_{max}为网络节点的最大度），网络博弈的纳什均衡中的采用合作策略的节点构成极小节点覆盖。（作业 3：自己编程序验证这个结论，网络可自定，节点数目不少于10）。作业 3 是选做题，可代替作业 2（专家系统PID大作业）。

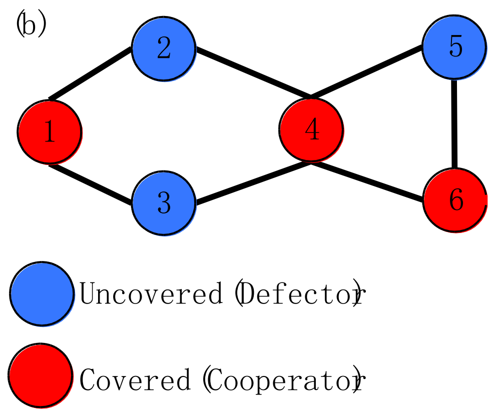


蓝颜色节点集合
为覆盖节点集合

博弈优化实例：网络节点最小覆盖问题

作业说明：节点的初始状态可随机定为Cooperator 或者 Defector,按照某种给定顺序（例如 1 ,2,...,10）依次检查每个节点，是否改变其状态可以获得更大收益，如果是则改变状态，否则不改变，直到所有节点都不再改变状态为止。验证合作的节点集合是否是极小节点覆盖。





蓝颜色节点集合为覆盖节点集合