第11卷 第4期　　自 动 化 学 报　　Vol. 11, No. 4
1985年10月　　ACTA AUTOMATICA SINICA　　Oct., 1985

多人多级递阶决策的几个问题——鼓励性对策及模型简化

郑 应 平
（中国科学院自动化研究所）

摘　要

本文讨论了按鼓励性对策法则进行的二人、二级决策问题. 利用可诱导区的概念对这一问题解的结构给予了完整的描述并总结出将其化为单人决策问题的等价原理. 这一原理可以推广到一般情况,并可用于简化一类多人、多级决策问题. 一些特殊情形也可以利用这一途径求解. 文中给出了数值例子.

一、引　言

许多社会、经济、环境等大型复杂决策系统的定量研究已日益引起重视. 具有多决策人的问题在概念和方法上都有许多新的进展[1-4]. 特别是系统具有多级递阶结构时，Stackelberg 对策可能是个合适的模型. 以最简单的二人二级对策问题为例，处在较高层级的"领导"首先宣布自己的策略. 然后另一决策人"随从"就以这种策略作为约束，对自己的目标函数进行最优化并确定自己的决策. 领导则必须预见到随从的这种反应而设计好自己的策略，以期得到最优的目标函数值. 这种利用策略来导致对领导自己最有利的结果的方法称为"鼓励性控制"（Incentive Control)，而领导的策略则称为"鼓励性策略".

这里有两个前提假定. 1）随从是"理性"的决策人,他的理性准则就是按自己的目标函数最优化行事而没有别的考虑；2）领导则遵守自己宣布的诺言. 他不能欺骗对方,在对方采取行动以后改变方针以牟取更好的结果. 否则，问题将大为复杂而需要专门的讨论.

由于领导所能采用的任一可能的鼓励性策略均对应于随从的特定响应并导致完全确定的结果——两人的决策量和由此决定的目标函数值. 这些可能达到的"决策对"的全体构成两人决策量的乘积空间中的一个子集合,称为"可诱导区"(IR = Inducible Region)[5]. 换言之,可诱导区中的任一决策对,只要领导愿意,总可以通过选择适当的鼓励性策略,对随从进行诱导而实现之. 从而可以把这种二人决策等价为由领导在其可诱导区中选择一对最优决策的单人决策问题. 这种想法为多人复杂决策系统提供了一种可能的简化原则. 本文则旨在阐明这一原则并说明其在具体情形下的应用.

本文于1983年7月21日收到.

二、二人、二级鼓励性 Stackelberg 对策的可诱导区

考察由一个领导和一个随从进行的鼓励性对策问题，他们分别选择自己的决策量 $u_L \in U_L$ 和 $u_F \in U_F$（其中 U_L 和 U_F 分别为他们的可允决策量集合），而力图使各自的目标函数 $J_L = J_L(u_L, u_F)$ 或 $J_F = J_F(u_L, u_F)$ 达到极小．对策的规则仍然是领导先宣布自己的鼓励性策略 $u_L = \gamma_L(u_F)$，随从以此作为约束而解决自己的最优化问题，然后领导遵守自己的策略并执行由其决定的决策量．问题则为如何通过选择 $\gamma_L(\cdot)$ 而使最后得到的 J_L 为最小．

研究这一问题时，首先要弄清在所有可能的结果 $(u_L, u_F) \in U_L \times U_F$ 之中究竟能达到怎样的最优值 J_L；为此又应选用怎样的 $\gamma_L(\cdot)$．这又相当于对任一特定结果 $(\bar{u}_L, \bar{u}_F)$ 如何判断能否通过选择 $\gamma_L(\cdot)$ 诱导随从采用 $u_F = \bar{u}_F$．把以上概念形式化，引入

定义. $U_L \times U_F$ 中的一个特定结果$(\bar{u}_L, \bar{u}_F)$称为“可诱导的”，是指存在 $u_L = \gamma_L(u_F)$ 使得随从的响应

$$u_F = \arg\min_{u_F} J_F(\gamma_L(u_F), u_F) = \bar{u}_F, \tag{1}$$

这里要求 $\bar{u}_L = \gamma_L(\bar{u}_F)$．所有可诱导的 (u_L, u_F) 的全体构成的集合 $IR \subset U_L \times U_F$ 称为“可诱导区”．对此可得到以下有用的结果：

命题 1. 一组结果 $(\bar{u}_L, \bar{u}_F)$ 可诱导的充分必要条件是对任何 $u_F \in U_F$，$u_F \neq \bar{u}_F$，均可找到 u_L 使得 $J_F(u_L, u_F) > J_F(\bar{u}_L, \bar{u}_F)$．

证明. 先证充分性．由命题条件可以定义

$$u_L = \gamma_L(u_F) = \begin{cases} \bar{u}_L, & \text{当 } u_F = \bar{u}_F, \\ \text{任何 } u_L \text{ 使 } J_F(u_L, u_F) > J_F(\bar{u}_L, \bar{u}_F), & \text{当 } u_F \neq \bar{u}_F. \end{cases} \tag{2}$$

显然这时 u_F 的最好值只能是 $\bar{u}_F$．

反之（必要性），由 $(\bar{u}_L, \bar{u}_F)$ 的可诱导性存在 $u_L = \gamma_L(u_F)$ 使得 $u_F = \arg\min\limits_{u_F} J_F\ (u_L = \gamma_L(u_F), u_F) = \bar{u}_F$，亦即对任何 $u_F \neq \bar{u}_F$ 均有

$$J_F(u_L = \gamma_L(u_F), u_F) > J_F(\bar{u}_L = \gamma_L(\bar{u}_F), \bar{u}_F).$$

从而命题 1 的条件满足．证毕．

命题 1 表明只要随从的行动不符合要求，$u_F \neq \bar{u}_F$，领导就有能力使他受损（J_F 变大）以作惩罚．命题 1 的主要飞跃是把复杂的鼓励性对策规则吸收进去，而用一种易于数学处理的形式表现出来．从这一命题容易推出以下有用的结果：

命题 2. 令

$$M = \inf_{u_F} \sup_{u_L} J_F(u_L, u_F), \tag{3}$$

若 $J_F(\bar{u}_L, \bar{u}_F) < M$，则 $(\bar{u}_L, \bar{u}_F)$ 可诱导；若 $J_F(\bar{u}_L, \bar{u}_F) > M$，则 $(\bar{u}_L, \bar{u}_F)$ 不可诱导．

证明．$J_F(\bar{u}_L, \bar{u}_F) < M$ 时，由M之定义，必存在 $\delta > 0$，对任何 u_F 均存在 u_L 使 $J_F(u_L, u_F) \geqslant M - \delta > J_F(\bar{u}_L, \bar{u}_F)$，从而命题 1 条件满足而本命题的第一部分得证．

命题的第二部分，仍由M的定义对任何 $\delta > 0$，均存在 $\tilde{u}_F$ 使得对任何 u_L，$J_F(u_L, \tilde{u}_F) <$

$M+\delta$。这表明当 $J_F(\bar{u}_L,\bar{u}_F)>M$ 时，若令 $\delta=J_F(\bar{u}_L,\bar{u}_F)-M>0$，则命题 1 的条件被破坏从而 $(\bar{u}_L,\bar{u}_F)$ 是不可诱导的。命题证毕。

这里量M是随从的"最后防线"，即领导无法强迫随从去达到比 $J_F=M$ 更坏的结果。从而 $J_F=M$ 自然成了"可诱导区"的边界。对于边界上的点是否可诱导的问题有以下结果:

命题 3. 设 $J_F(\bar{u}_L,\bar{u}_F)=M$，则关于$(\bar{u}_L,\bar{u}_F)$的可诱导性有以下准则：1）若不存在任何 u_F 使 $\sup\limits_{u_L} J_F(u_L,u_F)=M$ 真正实现，那么 $(\bar{u}_L,\bar{u}_F)$ 可诱导；2）若存在唯一的 u_{FM} 使 $\sup\limits_{u_L} J_F(u_L,u_{FM})=M$，则当 $\bar{u}_F=u_{FM}$ 时，$(\bar{u}_L,\bar{u}_F)$ 可诱导；当 $\bar{u}_F\neq u_{FM}$ 时，$(\bar{u}_L,\bar{u}_F)$ 不可诱导；3）若存在两个或更多的 $u_{FM}=u'_{FM},u''_{FM},\cdots$均使 $\sup\limits_{u_L} J_F(u_L,u_{FM})=M$，则 $(\bar{u}_L,\bar{u}_F)$ 必不可诱导。

证明. 按三种情形分别证明.

1）对任何 u_F，$\sup\limits_{u_L} J_F(u_L,u_F)>M$，亦即总存在 u_L 满足 $J_F(u_L,u_F)>M$，从而 $(\bar{u}_L,\bar{u}_F)$ 是可诱导的；

2）同样的推理可知，任何使 $J_F=M$ 的 $(\bar{u}_L,\bar{u}_F=u_{FM})$ 是可诱导的。同时若 $\bar{u}_F\neq u_{FM}$，领导无法迫使随从不采用 $u_F=u_{FM}$，从而 $(\bar{u}_L,\bar{u}_F)$ 不可诱导；

3）当领导希望 $u_F=u'_{FM}$ 时，无法强迫随从不采用 $u_F=u''_{FM}$，因此不可诱导。证毕。

当然，这里隐含了"随从是不大愿意和领导合作的"一个行为假定。就是说虽然 u'_{FM}，u''_{FM} 将导致对随从而言相同的结果 $J_F=M$，随从仍不愿意采用领导所希望的决策量 $\bar{u}_F=u'_{FM}$。如果采用别的行为假定，例如"随从在不影响其自身利益时愿意合作"，则本命题亦只须作一些显易的修改。

总之，对比较简单的二人、二级的鼓励性对策问题，领导可以用他的鼓励性策略在可诱导区 $IR\subset U_L\times U_F$ 中实现对其最有利的结果。例如若 $\min\limits_{(u_L,u_F)\in IR} J_L$ 存在。这就是对策问题的解。

为了把这些结果用于解决更复杂的问题，可以用另一形式概括以上结果。

等价原理 1. 一个由领导和随从按鼓励性对策法则进行的二人、二级决策问题，等价于一个由领导来代表的单人决策问题，其目标函数不变（仍为 J_L），而决策量变为可以在 $IR\subset U_L\times U_F$ 中取值的任何决策对 (u_L,u_F)。

利用这一原理可以简化多人、多级的复杂问题。

三、一般多人多级递阶决策问题的简化

首先把前节讨论的鼓励性 Stackelberg 对策的概念稍作推广，考察$N+1$ 人的对策问题。设他们的决策量分别为 $u_0,u_1,\cdots,u_N$；目标函数分别为 $J_0(u_0,\cdots,u_N),\cdots,J_N(u_0,\cdots,u_N)$。假定:

1）其中有一个决策人，例如做决策 u_0 的人，在决策时拥有其他人的决策量的信息。同时他可以在对策开始宣布其策略

$$u_0 = \gamma_0(u_1, \cdots, u_N). \tag{4}$$

事实上这个人就是这种情形下的领导.

2）假定在领导宣布了策略（4）之后，其余N个人所面临的多人对策问题具有明确的解概念，并导致一个唯一的平衡解（$\bar{u}_1, \cdots, \bar{u}_N$）.

这当然是个很强的假定，而且常需要有附加的行为假定才能得到保证，但毕竟有许多情形还是满足这些条件的. 本文的主要目的在于避开细节困难而说明下面的基本思想.

以上两个假定实际上已经规定了一个从策略γ_0到决策量组（$\bar{u}_0 = \gamma_0(\bar{u}_1, \cdots, \bar{u}_N)$, $\bar{u}_1, \cdots, \bar{u}_N$）的一个完全确定的映射. 遍取所有可能的$\gamma_0$，可以定义"所有可能结果"$(u_0, \cdots, u_N) \in U_0 \times \cdots \times U_N$的集合，该集合称为这个鼓励性对策的可诱导区$IR$.

仿照第二节的推理，可得以下更普遍一点的简化原则.

等价原理 2. 上述的$N+1$人决策问题等价于一个单人决策问题，其决策人可以用题中的领导代表，而他的可允决策量就是IR中的所有可能的决策组（$u_0, \cdots, u_N$）. 事实上，可经选用适当的策略γ_0来诱导而实现IR中的任何一组决策.

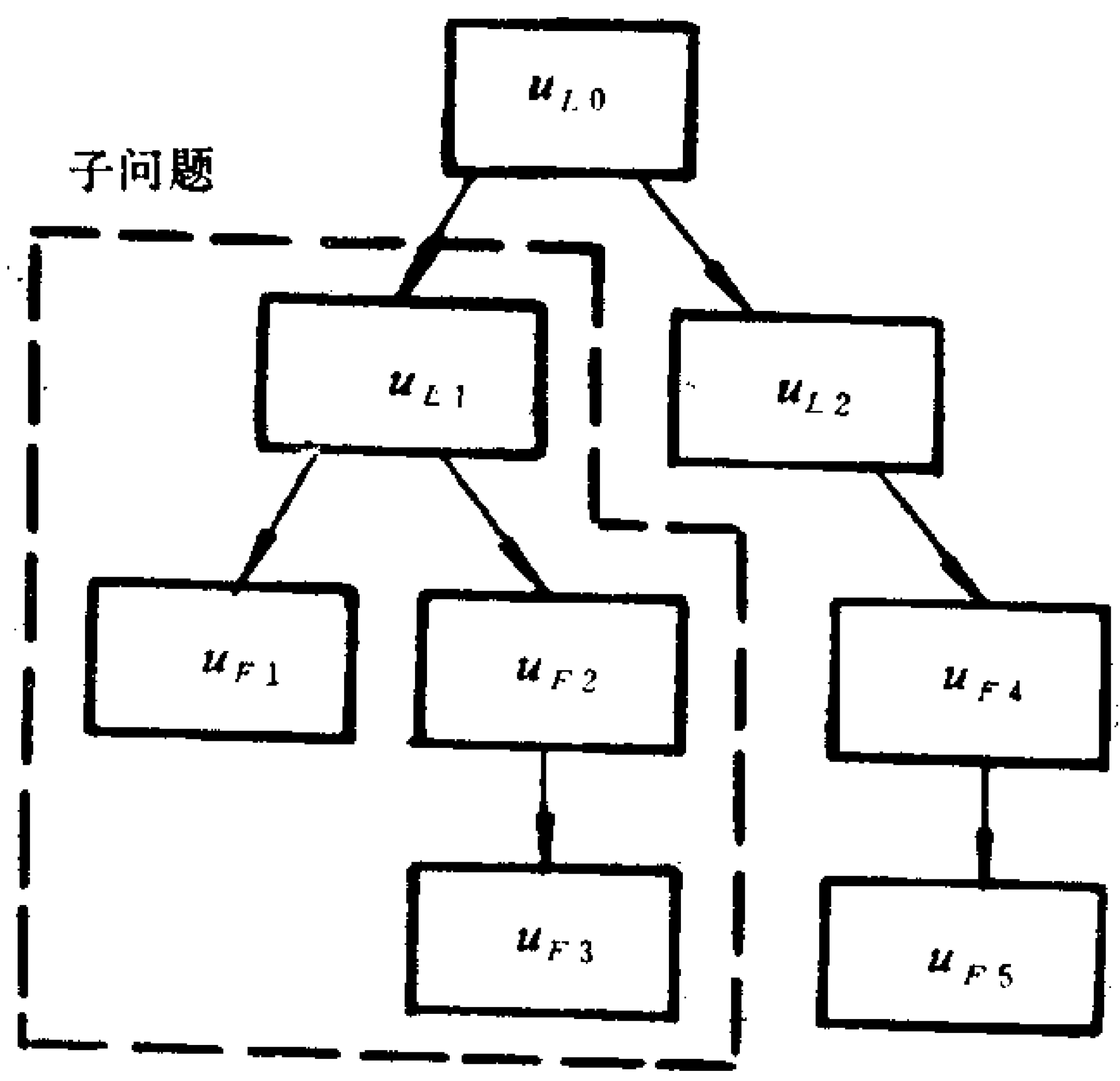

图 1　一个多人多级决策系统

为说明这一原理的应用，考察图 1 所示的一个复杂的多人多级决策问题. 图中的箭头表示"领导关系"，即上级可以知道下级的决策量以作为其策略中依据的信息. 其中一些"子问题"(例如图中所示包括（$u_{L1}, u_{F1}, u_{F2}, u_{F3}$）的一个小问题）在满足一定条件时可以从原问题中孤立出来而加以简化. 这就是下述的孤立条件：

1）子问题中有一人（u_{L1}）为局部领导，他拥有所属各u_{Fi}的信息，并可在决策开始之前宣布形如 $u_{L1} = \gamma_{L1}(u_{F1}, u_{F2}, u_{F3})$ 的鼓励性策略；

2）各个局部随从的目标函数J_{Fi}可以与u_{L1}有关（这表示局部领导对他们的影响和控制)，但与该子问题以外的各u_i无关. 即更高级的领导或兄弟部门的决策只有通过对J_{L1}的影响来控制该子问题的结果，他们不能"越级"对下面的人员（u_{Fi}）直接进行控制. 当然这里毫不排斥下面任何一个人的决策可以影响到领导甚至是最高领导的目标函数（J_{L0}）.

3）假定在宣布γ_{L1}之后，各u_{Fi}可按完全确定的解概念导致唯一的决策结果. 如前所述，这时该子问题的可诱导区 $IR_{L1} \subset U_{L1} \times U_{F1} \times U_{F2} \times U_{F3}$ 是完全确定的.

由等价原理 2，可以用一个单决策人代替该子问题中的四个人. 其决策量是四元组（$u_{L1}, u_{F1}, u_{F2}, u_{F3}$）$\triangleq u'_L$，取值于可诱导区$IR_{L1}$(图 2).

只要能不断找出满足"孤立条件"的子问题，用上述办法就可不断进行简化. 最后甚至可能化为单人决策问题，虽然其可允决策构成一个很复杂的高维集合，但概念上并没有什么困难.

用图 3 的例子说明方法的应用. 这是一个$N+1$人组成的"单线联系"的决策系统.

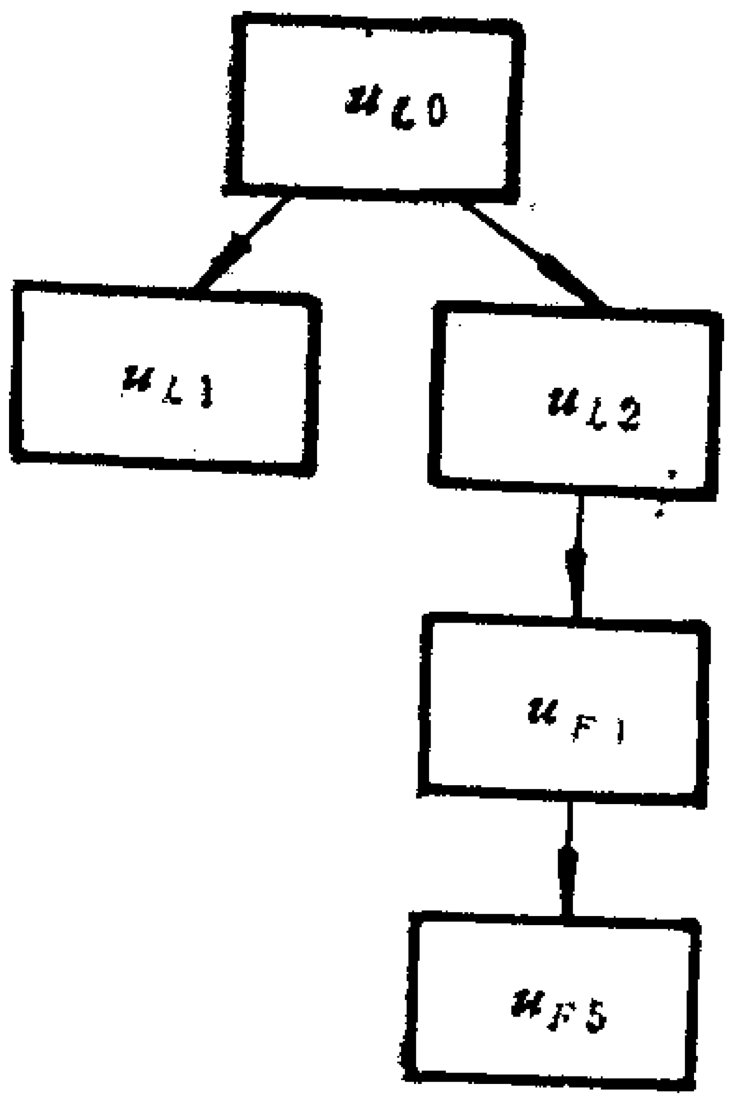

图 2 对图 1 所作的初步简化

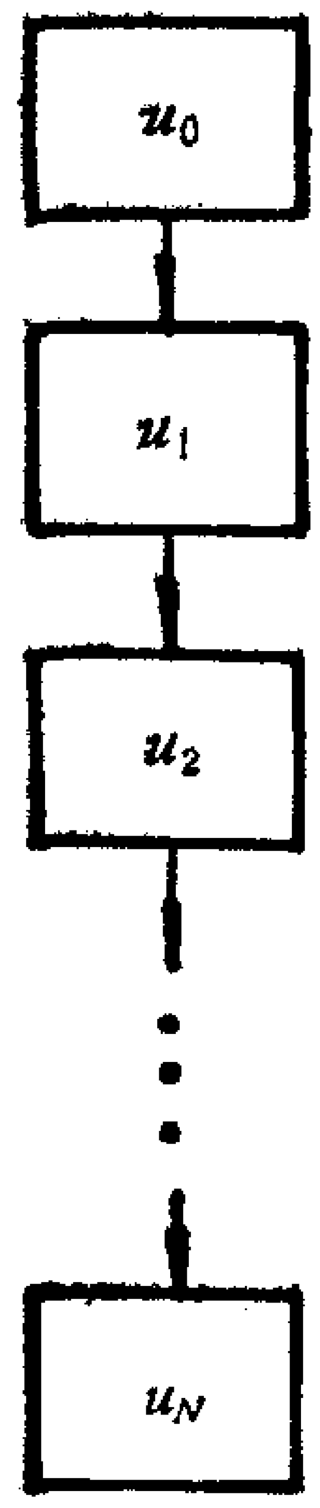

图 3 $N+1$ 决策人，$N+1$ 级的决策问题

这里

$$u_i \in U_i (i = 0, \cdots, N),$$

各人的品质函数为

$$\begin{aligned} J_0 &= J_0(u_0, u_1, \cdots, u_N), \\ J_1 &= J_1(u_0, u_1, \cdots, u_N), \\ J_2 &= J_2(u_1, \cdots, u_N), \\ &\cdots\cdots\cdots\cdots\cdots \\ J_N &= J_N(u_{N-1}, u_N). \end{aligned}$$

注意这里任一 u_i 不能越级影响其"下级"的目标函数 $J_{i+2}, J_{i+3}, \cdots$. 假定决策 u_i 可以依据其所有"下级"的决策量信息 $u_{i+1}, u_{i+2}, \cdots$，并事先宣布形如 $u_i = \gamma_i(u_{i+1}, \cdots, u_N)$ 的鼓励性策略，而且逐级定义的对策问题都有唯一解. 这时可按前述孤立原则和等价原理自下而上逐级化简求解. 具体做法为:

第 1 步. 设

$$M_N = \inf_{u_N} \sup_{u_{N-1}} J_N(u_{N-1}, u_N),$$

则

$$IR_{N-1} = \{(u_{N-1}, u_N) | J_N(u_{N-1}, u_N) \leqslant M_N\}.$$

为简单计，假定 U_N 中不存在 u_N 使真正实现 $\sup\limits_{u_{N-1}} J_N(u_{N-1}, u_N) = M_N$，这时如命题 3 所证，$J_N = M_N$ 的点均包含在 IR_{N-1} 之中.

第 2 步. 设

$$M_{N-1} = \inf_{(u_{N-1}, u_N) \in IR_{N-1}} \sup_{u_{N-2}} J_{N-1}(u_{N-2}, u_{N-1}, u_N),$$

则

$$IR_{N-2} = \{(u_{N-2}, u_{N-1}, u_N) | J_{N-1}(u_{N-2}, u_{N-1}, u_N) \leqslant M_{N-1}\}.$$

为简单计，仍作了与第 1 步相类似的假定．依此类推最后达到

第N步．设

$$M_1 = \inf_{(u_1,\cdots,u_N)\in IR_1} \sup_{u_0} J_1(u_0, \cdots, u_N),$$

则

$$IR_0 = \{(u_0, \cdots, u_N) | J_1(u_0, \cdots, u_N) \leqslant M_1\}.$$

这里和前面一样仍做了简化的假定． 否则在定义 IR_i 时需要具体分析 $J_{i+1} = M_{i+1}$ 边界上的点，这可用命题 3 来进行．

至此，问题已化成为单人决策问题．

第$N+1$步．该复杂决策问题的解为

$$J_0^* = \inf_{(u_0,\cdots,u_N)\in IR_0} J_0(u_0, \cdots, u_N) \tag{5}$$

可能被真正达到，或者被逼近到一定程度．

数值例子[7]．

考察$N=2$ 情形，设

$$J_0 = u_0 + 4u_1 + 6u_2, \qquad J_1 = -\frac{1}{2}u_0 - u_1 - u_2, \qquad J_2 = -u_1 - u_2,$$

其中 $0 < u_0 \leqslant 3$，$0 < u_1 \leqslant 2$，$0 < u_2 \leqslant 1$．按上述步骤求解．

第 1 步．求出

$$M_2 = \inf_{u_2} \sup_{u_1}(-u_1 - u_2) = \inf_{u_2}(-u_2) = -1,$$

$$IR_1 = \{(u_1, u_2) | -u_1 - u_2 \leqslant -1\}.$$

第 2 步．求出

$$M_1 = \inf_{(u_1,u_2)\in IR_1} \sup_{u_0} \left(-\frac{1}{2}u_0 - u_1 - u_2\right) = \inf_{(u_1,u_2)\in IR_1} (-u_1 - u_2) = -3,$$

$$IR_0 = \left\{(u_0, u_1, u_2) \middle| \frac{1}{2}u_0 + u_1 + u_2 \geqslant 3\right\}.$$

第 3 步．这个对策问题的解可求出为

$$J_0^* = \inf_{(u_0,u_1,u_2)\in IR_0} (u_0 + 4u_1 + 6u_2) = 9,$$

对应 $u_0 = 3$，$u_1 = 3/2$，$u_0 = 0$．当然，它实际上只能被逼近到需要的程度．

四、结 束 语

可诱导性和可诱导区的概念实际上由文献 [4] 提出并由文献 [5, 6] 赋予了这个名称．本文用文献 [9] 的表述方法对此给出了一个完整而严格的表述，进而又把它推广到更一般的鼓励性对策问题并导致一个可能把多人复杂对策问题化为单人决策问题的简化原则．对一些特殊情形和数值例子的处理，表明了这一简化方法可以解决问题．

当 U_i 为一般线性空间，J_i 满足一些凸性条件时，领导常具有“无限惩罚能力”从而可诱导区可以是全空间．这时有更多的问题可以解决[10]．

把这些概念推广而用于多维或动态对策问题时，可允决策量集合 U_i 往往是函数空间．利用函数空间中的运算也可对一大类问题给出解答．关于这方面的基本思想和主要做法可以参看文献 [9，10]．

参 考 文 献

[1] Cruz, J. B., Leader-Folower Strategies for Multilevel Systems, *IEEE Trans. Automatic Control*, **23** (1978), 244.

[2] Basar, T. and Olsder, G. J., Dynamic Noncooperative Game Theory, Academic Press, London, 1982.

[3] Gupta, N. K., An Overview of Differential Games. Advances in Control and Dynamic Systems, **17**, Academic Press, 1981.

[4] Y. C. Ho, Luh, P. B. and G. J. Olsder, A Control-theoretic View on Incentives, *Automatica*, **18** (1982), 167.

[5] Chang, T. S. and Luh, P. B., The Concept of Inducible Region in Stackelberg, Game, Proceedings of the 1982 American Control Conference, Arlington, Virginia, June 1982, 139—140.

[6] Chang, T. S. and Luh, P. B., A Complete Solution for Two-Person, Single-Stage, Deterministic Stackelberg Game, Preceedings of thte 21st IEEE Conference on Decision and Control, Orlando, Florida, Dec. 1982, 176.

[7] B. Tolwinski, Equilibrium Solutions for a Class of Hierarchical Games, Application of System Theory to Economics, Management and Technology, J. Gutenbaum and M. Niezgodka Eds., Warsaw, PWN, 581—600, 1981.

[8] Luh, P. B. Chang, T. S. and T. Ning, Three-le vel Hierarchical Desision Problems, *Trans. IEEE Automatic. Control (to appear)*, 1984.

[9] Zheng, Y. P., Basar, T. and Cruz, J. B., Stackelberg Strategies and Incentives in Multi-person Deterministic Decision Problems, *Trans. IEEE System, Man and Cybernetics*, SMC-14, 10—24, 1984.

[10] Zheng, Y. P. and Basar, T., Existence and Derivation of Optimal Affine Incentive Schemes for Stackelberg Games with Partial Information: A Geometric Approach, *Int. J. of Control*, **35** (1982), 997.

SOME PROBLEMS IN MULTI-PERSON, MULTI-LEVEL HIERARCHICAL DICISION-MAKING——INCENTIVE GAME AND MODEL SIMPLIFICATION

ZHENG YINGPING

(Institute of Automation, Academia Sinica)

Abstract

The problem of 2-person, 2-level decision making according to incentive game rule is discussed. The structure of its solution is described via the concept of inducible region. It then can be summerized as an "equivalance principle" for reducing the problem to that with only single decision maker. This principle can be further used to reduce a class of mult-person, multi-level decision problems. Some special cases with numerical examples are solved to show the applicability of the approach.