

文章编号: 1001- 2486(2010) 04- 0131- 06

基于前景理论的 CGF Agent 决策建模研究*

胡记文, 尹全军, 冯 磊, 邓海军, 查亚兵

(国防科技大学 机电工程与自动化学院, 湖南 长沙 410073)

摘要: 前景理论(Prospect Theory, PT)描述了个体在不确定条件下的决策规律,较好地刻画了决策者的风险偏好。分析表明,PT在作战决策建模中具有一定的适用性和可行性。提出了一种基于PT的计算机生成兵力(Computer Generated Forces, CGF) Agent 决策建模方法,给出了CGF Agent的建模框架,对CGF Agent决策过程中方案选择的具体步骤和算法进行了详细阐述。实验表明,所提出的方法在一定程度上提高了CGF Agent决策行为的真实性,证明了所提出方法的有效性。

关键词: 作战决策; 前景理论; 计算机生成兵力; 不确定决策

中图分类号: TP391 **文献标识码:** A

Research of CGF Agent Decision-making Modeling Based on Prospect Theory

HU Ji-wen, YIN Quan-jun, FENG Lei, DENG Hai-jun, ZHA Ya-bing

(College of Mechatronics Engineering and Automation, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: Prospect Theory (PT) describes the rules of individual's decision-making under uncertainty, which can well characterize decision-makers' risk preference. PT has certain valuable feasibility and applicability in operational decision-making modeling. Based on PT, a method of CGF Agent decision-making modeling was proposed. The Computer Generated Forces (CGF) Agent framework was presented. The detailed description of concrete process and algorithm for course of action (COA) selection in Agent decision-making was also given. The experiment shows that the proposed method has good performance and can improve the authenticity of CGF Agent's decision-making behavior.

Key words: operational decision-making; prospect theory; computer generated forces; uncertain decision making

在作战仿真应用中, 一个重要的研究内容就是对战场上各类指挥实体的作战决策行为的建模。战场环境是一个复杂的、具有多种不确定因素的环境^[1]。在战场上, 指挥员很难准确预知对手下一步的行动方案, 也难以对自己的决策所造成的结果进行准确的预期, 而不得不在其决策过程中掺杂概然因素^[2], 所以作战决策是一种不确定条件下的决策。同时, 由于人的认知能力和认知容量的局限, 决策者无法在短时间内客观把握所要解决问题的空间, 无法找到所有备选方案, 决策者只能在分析可得信息基础上, 搜索可接受的选择方案, 也就是能满足其期望水平的选择, 而不是通过考察所有可能的方案之后再选择效用最大的最优方案^[2-3]。因此, 作战决策实际上是一种有限理性的行为。

目前, 在作战仿真领域研究不确定条件下的决策, 主流理论一直是期望效用理论(Expected Utility Theory, EU)。基于EU的决策以下假设为基础^[3]: (1) 决策者可以构建备选方案的完备集; (2) 每个备选方案的结果及其概率是已知的; (3) 决策目的是使一个给定的效用函数的期望值达到最大。在西蒙等^[3]看来, 这些假设高估了人类的认知能力和理性——认为决策者能够获得有关决策问题的全部信息, 并能加以客观分析, 其决策结果是最优的。事实上, 人类在不确定性条件下的判断决策是存在偏差和有限理性的, 因此基于EU的最优决策模型有时不能很好地描述人类在不确定条件下的决策特点。

本文试图将PT引入作战决策建模中, 探索不确定条件下CGF Agent决策行为建模理论和方法, 构

* 收稿日期: 2009- 12- 27

基金项目: 国家自然科学基金资助项目(60904055); 国家部委基金资助项目(513040104)

作者简介: 胡记文(1978—), 男, 博士生。

建能体现不确定条件下人类决策行为特点的 CGF Agent 建模框架。该框架以通用 Agent 框架^[9]为设计基础,以 PT 为决策工具,通过刻画决策者在“收益”和“损失”情况下不同的风险偏好,描述决策者在不确定战场环境中的决策过程,实现对决策者“有限理性”决策行为的建模,在一定程度上提高 CGF Agent 决策行为的真实性。

1 前景理论

与传统理性决策理论相比,PT^[4-5]更多地从心理学角度研究人们决策所依据的准则和经验,着重于从人自身的心理特质、行为特征出发,揭示了影响行为选择的非理性因素,为研究不确定情况下的人类判断和决策方面提供了新的思路。PT 有以下基本观点:

(1) 决策者并不特别在意决策所带来的效益值的绝对水平,而是在意决策所造成的效益值相对于某一参考点的变化,参考点反映了决策者对决策结果的心理预期。

(2) 参考点是随着环境、态势的变化而变化的。

(3) 面对收益,决策者是风险规避的;面对损失,决策者是风险偏好的。

(4) 人们对收益和损失都具有敏感性递减特征,且对损失比对收益更敏感。

具体而言,假设现有方案 a 和方案 b 供决策者选择,方案 a 将会使结果 x_i 以 p_i 的概率出现,而方案 b 将使结果 x_i 以 q_i 的概率出现,按照 PT,决策者在方案 a 和 b 之间偏好前者,当且仅当:

$$\sum_i \pi(p_i) v(\Delta x_i) > \sum_i \pi(q_i) v(\Delta x_i)$$

其中: $\pi(p_i)$ 、 $\pi(q_i)$ 为决策权重函数,反映决策者对结果概率的主观感知;

$v(\Delta x_i)$ 为价值函数,反映决策者对于决策结果的价值评估;

$\Delta x_i = x_i - x_0$, 是结果 x_i 相对于某一参考点 x_0 的偏离值。

2 前景理论在作战决策中的适用性与可行性分析

根据对作战决策和 PT 的前期研究,认为结合 PT 的思想进行作战仿真决策建模,可以很好地描述战场实体在不确定条件下风险决策的本质。

从战场环境的特点来看:决策者决策时不仅要当前态势进行判断与估计,还要对未来敌方行为和战场可能出现的态势进行判断与估计。由于一般情况下难以准确预知敌方行为,因而战场态势的发展具有不确定性。PT 是经济学中的前沿理论,特别侧重于研究人在不确定情况下进行判断和决策的过程,而且其在研究不确定经济活动中人的决策上,已经取得了很大成功。

从人类认知能力来看:人类由于自身认知能力和认知容量的限制,难于获得解决问题的所有信息,也难以对备选方案结果进行准确预测,因此其决策不一定总是最优的。PT 作为西蒙“有限理性”理论的发展,承认人类认知能力上的缺陷,并在其理论中对“有限理性人”的决策特点进行了具体体现。

从人类决策过程中的心理历程来看^[7]:认知心理学认为,人类决策过程受其心理因素影响,不仅决策者的即时情绪对决策产生影响,而且决策者可能的决策结果所导致的期望情绪也会影响决策。PT 是卡尼曼将心理学的研究引入经济学领域的产物,不论是 PT 中的决策者的决策偏好,还是决策的决策权重,无不体现了决策者的心理因素对决策的影响。

基于以上分析,笔者认为 PT 适用于作战决策,并能确定较好结果。另外,从可行性上来讲,PT 在管理学、经济学、社会学领域都已有成功的应用。Mcdermott^[8], Levy^[9], Taliafero^[10] 等将 PT 应用到外交决策的研究上,取得了不少优秀成果;Schultz^[11] 研究了二战中艾森豪威尔在“市场花园”行动中的决策过程,对 PT 进行了实验验证;美国 New Mexico 大学的 Ganderton^[12] 等应用 PT 研究了面对灾害风险时的保险购买问题。近年来国内在 PT 的研究与应用上也取得了一定成就。国防科技大学孙多勇^[13] 教授将 PT 应用于突发事件与行为决策的研究,探讨了危机条件下人们行为决策的特征与规律;北京交通大学赵凛^[14] 博士将 PT 应用于出行路径决策,建立了基于 PT 的出行路径选择模型;林明旺^[15] 将 PT 用于研究外交决策,也取得了一定进展。

基于以上研究成果, 笔者认为将 PT 用于作战决策是可行的。

3 基于前景理论的 CGF Agent 作战决策

3.1 基于前景理论的 CGF Agent 框架

在 CGF Agent 决策建模中, 引入 PT 可以更真实地描述不确定战场条件下决策者的决策行为, 本文在通用 Agent 框架的基础上, 设计了如图 1 所示的基于 PT 的 CGF Agent 框架。该框架在组成上与通用 Agent 框架基本相同, 包含信息输入、工作存储器、长期存储器、态势评估、规划、推理、学习、行为输出等功能模块。所不同的是, 框架中行动方案(Course Of Action, COA)的分析与选择基于 PT 实现。这种决策方式, 由于要进行备选方案前景值的计算, 本质上是一种基于效用的决策。与难以进行不确定知识表示的基于知识的逻辑智能体相比, 基于 PT 的 Agent 在不确定推理能力方面具有明显的优势。与一般的基于效用的智能体相比, 基于 PT 的 Agent 还具备了能够描述决策者决策过程中的决策偏误和有限理性的能力。

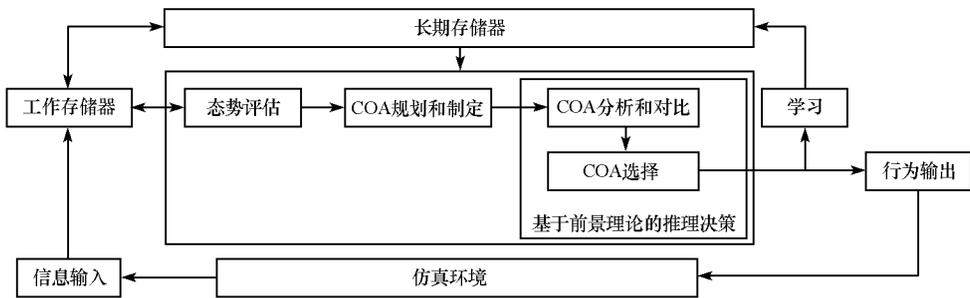


图 1 基于前景理论的 CGF Agent 框架

Fig. 1 Framework of CGF Agent based on prospect theory

3.2 基于前景理论的 CGF Agent 推理决策

在所提出的 CGF Agent 框架中, 推理决策模块对备选 COA 进行分析与对比, 计算各备选 COA 的前景值, 最终选择前景值最大的 COA 加以执行。

首先, Agent 根据态势分析的结果, 判断当前态势结果的位置(态势有利, 还是态势不利), 建立决策参考点, 并在此基础上确定 Agent 的风险态度: 风险规避还是风险追求。

其次, Agent 根据上一步的分析结果, 分析各备选 COA 的可能结果及其概率, 并对分析结果进行初步加工。根据 PT, 这一过程主要涉及 4 个步骤, 即结果的编码、合并、分解、消除。

最后, Agent 对经过初步加工过的备选 COA 分析结果进行评估, 按前景值最大化原则进行 COA 选择。图 2 给出了详细的基于 PT 推理决策流程。在该流程中, 关键是决策参考点的确定和各方案结果前景值的计算。前者直接决定决策者的风险偏好; 后者影响具体方案的选择。

参考点的确定通常有 3 种情况: (1) 大多数情况下, 参考点由作战目标决定; (2) 有些情况下, 参考点由决策者对作战行动的预期决定; (3) 有时参考点由当前态势决定。

备选 COA 前景值的计算: 涉及了该 COA 可能导致的每种结果的前景值以及决策权重的计算, 决策权重反映可能产生的某种结果对 COA 价值的影响力。决策权重是 p 的函数, p 表示产生某种结果的概率。

3.3 备选 COA 前景值的计算

各备选 COA 的前景值按下列方法与步骤计算:

(1) S : 代表决策者假定某 COA 实施后, 对战场态势的一种估计, 称为状态集; S 的子集代表由于某个 COA 的实施, 而可能发生的事件;

(2) X : 代表结果集, 指某个 COA 实施后可能导致的所有结果的集合。 X 中的结果可以分为 3 类, 即: 中性结果; 收益型结果; 损失型结果。

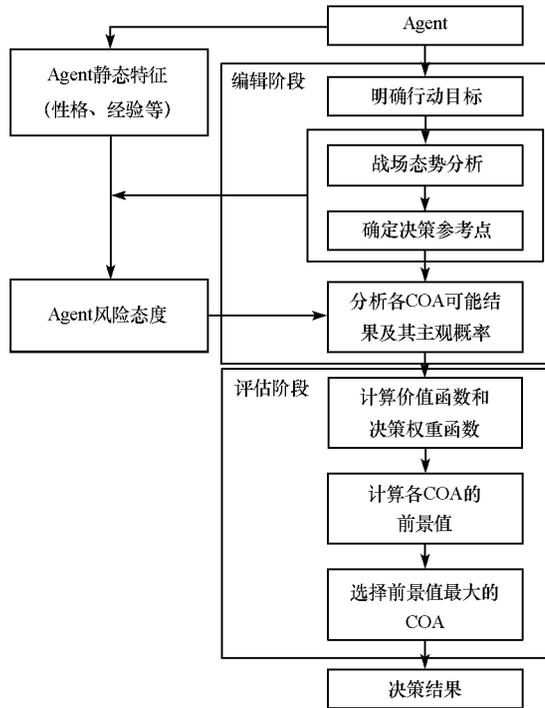


图2 基于前景理论的作战决策流程

Fig. 2 The flow chart of decision based on prospect theory

(3) 函数 $f: S \rightarrow X$ 表示一个不确定的前景, 即 $\forall s_i \in S, \exists x_i \in X$, 使得 $f(s_i) = x_i$ 。通常把 f 表示为一个序对 (x_i, A_i) , 其中 A_i 是 S 的子集, 称之为事件, (x_i, A_i) 表示事件 A_i 发生时导致结果 x_i 。

(4) 把前景 f 的结果 x_i 按升序排列, 即 $x_{-m} \leq \dots \leq x_{-1} \leq x_0 \leq x_1 \leq \dots \leq x_n$, 其中参考点的结果 $x_0 = 0$, 表示既不收益也不损失; 对于收益型结果有 $x_i > 0$; 对于损失型结果有 $x_i < 0$ 。

(5) 对于前景 $f = \{(x_i, A_i) | -m \leq i \leq n\}$, 收益型前景和损失型前景有不同的计算方法, 故令:

$$\begin{aligned} \text{收益型前景: } f^+ \left(\begin{matrix} s \\ s \end{matrix} \right) &= f \left(\begin{matrix} s \\ s \end{matrix} \right) & \text{if } f \left(\begin{matrix} s \\ s \end{matrix} \right) > 0; & f^+ \left(\begin{matrix} s \\ s \end{matrix} \right) = 0 & \text{if } f \left(\begin{matrix} s \\ s \end{matrix} \right) \leq 0; \\ \text{损失型前景: } f^- \left(\begin{matrix} s \\ s \end{matrix} \right) &= f \left(\begin{matrix} s \\ s \end{matrix} \right) & \text{if } f \left(\begin{matrix} s \\ s \end{matrix} \right) < 0; & f^- \left(\begin{matrix} s \\ s \end{matrix} \right) = 0 & \text{if } f \left(\begin{matrix} s \\ s \end{matrix} \right) \geq 0; \end{aligned}$$

(6) 由于一个前景可能有多个结果, 为了符合随机占优, 并方便实现对多个结果概率的累积, 依据累积泛函理论, 引入容量函数 W , 对于 $A \subset S$, 都有一个值 $W(A)$, 满足 $W(\emptyset) = 0, W(S) = 1$, 且当 $A \supset B$ 时, $W(A) > W(B)$ 。

(7) 令 $V(f)$ 表示前景值函数; $v(x)$ 表示价值函数, 则有

$$V(f) = V(f^+) + V(f^-); \text{ 其中 } V(f^+) = \sum_{i=0}^n \pi_i^+ v(x_i), V(f^-) = \sum_{i=-m}^0 \pi_i^- v(x_i)$$

决策权重 $\pi_i^+(f^+) = (\pi_0^+, \dots, \pi_n^+)$, $\pi_i^-(f^-) = (\pi_{-m}^-, \dots, \pi_0^-)$, 使用容量函数计算决策权重, 且假设前景 $f = \{(x_i, A_i) | \text{中事件 } A_i \text{ 的发生概率 } p(A_i) = p_i\}$, 则有

$$\begin{aligned} \pi_n^+ &= W^+(p_n), \pi_{-m}^- = W^-(p_{-m}); \\ \pi_i^+ &= W^+(p_i + \dots + p_n) - W^+(p_{i+1} + \dots + p_n), 0 \leq i \leq n-1; \\ \pi_i^- &= W^-(p_{-m} + \dots + p_i) - W^-(p_{-m} + \dots + p_{i-1}), -m+1 \leq i \leq 0; \end{aligned}$$

若当 $i \geq 0$ 时令 $\pi_i = \pi_i^+$, 当 $i < 0$ 时令 $\pi_i = \pi_i^-$, 则 $V(f) = \sum_{i=-m}^n \pi_i v(x_i)$ 。

在统计大量决策案例的基础上, 通过非线性回归, 得到目前大多数研究者认可的 $v(x)$ 、 $W^+(p)$ 及 $W^-(p)$ 的函数形式如下:

$$v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x < 0 \end{cases}; \quad W^+(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}}; \quad W^-(p) = \frac{p^\delta}{[p^\delta + (1-p)^\delta]^{1/\delta}}$$

3.4 决策参数的确定方法

α 、 β 分别反映决策者对“收益”和“损失”敏感性递减的速度; λ 反映损失厌恶程度; γ 、 δ 分别反映决策者对“收益”和“损失”结果的关注程度。对于不同的决策者, 由于其性格、经验等不尽相同, 上述参数值也会略有不同, 这反映了决策者之间的一种差异。由于这些参数表现的是决策者面对风险时的心理趋势和态度, 属于定性的量, 难以用公式直接求解。目前本研究借鉴文献[1]的方法, 采用如下方法确定决策参数:

- (1) 构造测试集, 即针对要模拟的指挥员的任务环境, 构造若干典型的决策样本;
- (2) 意向性测试, 针对决策样本集中的决策样本, 由指挥员进行决策, 并记录决策结果;
- (3) 初步设定决策参数, 由基于 PT 的 CGF Agent 对决策样本集中的想定进行决策;
- (4) 对比第(2)、(3)步的决策结果, 如果结果一致, 则说明所设定的决策参数基本能反映要仿真的指挥员的决策特点和心理趋势; 否则, 返回第(3)步, 继续试探所设定的决策参数。

4 实验验证及结果分析

实验想定如下: 数量不明的恐怖份子劫持若干人质, 被包围于一栋建筑物, 反恐力量的任务是解救人质, 且要求伤亡最小。在仿真开始前, 我们设置了两种不同的初始态势。初始态势 S1: 从情报得知, 到目前为止并无无人质伤亡; 初始态势 S2: 从情报得知, 到目前为止, 至少有一名人质伤亡。

当前态势是: 建筑物内出现短暂骚动, 且有枪声传出, 随后恐怖份子要求谈判。在此当前态势下, 指挥员 Agent 可采用的 COA 有:

- COA1: 同恐怖份子谈判, 避免局势失控, 力求和平解决;
- COA2: 对恐怖份子进行攻心战, 如武力展示、严密包围等, 让其知道不可能达成目的;
- COA3: 向恐怖份子进攻, 武力解救人质。

本实验构建了 3 种类型的指挥员 Agent。第一种指挥员 Agent 是基于产生式规则的 Agent, 在 Soar 体系结构基础上构建, 称为 Soar-Agent; 第二种指挥员 Agent 是基于效用的 Agent, 基于 EU 决策, 称为 EU-Agent; 第三种指挥员 Agent 基于本文所提出的框架构建, 称为 PT-Agent。

分别以 S1 和 S2 为初始态势, 让每一种指挥员 Agent 在相同的当前态势下, 各重复决策 20 次, 总共得到 120 个决策结果。图 3、图 4 分别记录了每种指挥员 Agent 在 S1、S2 下的决策结果。

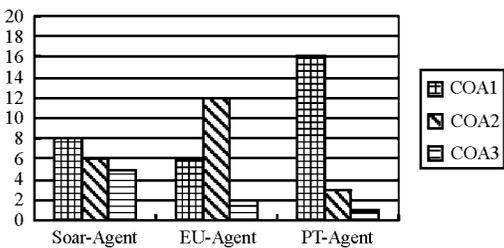


图 3 不同 Agent 在初始态势 S1 下的 COA 选择

Fig. 3 COA selection of each agent type in initial situation S1

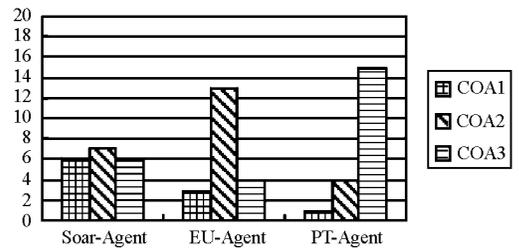


图 4 不同 Agent 在初始态势 S2 下的 COA 选择

Fig. 4 COA selection of each agent type in initial situation S2

从图中可以看出, 无论在 S1 下, 还是在 S2 下, Soar-Agent 对于各 COA 趋向于随机选择, 即每种 COA 的被选择次数差别不大。因为 Soar 作为一种产生式规则系统, 其决策依据当前态势与决策规则的匹配来进行。当前态势的发展存在多种可能性, 而 Soar 系统无法处理这种概然性推理, 只能对所有可能事实平等对待, 从而根据 COA 激发规则, 在各种 COA 间随机选择。

对于 EU-Agent, 可以看出, 无论在 S1 下, 还是在 S2 下, 其选择明显偏向于 COA2。因为 EU-Agent 依据 EU 进行决策, 它计算各 COA 实施后所有可能结果的综合效用, 然后选择综合效用最大的那个 COA 执行。在本实验中, 大多数情况下 COA2 具有最大综合效用。

对于 PT-Agent, 可以发现, 该 Agent 的决策风险倾向在两种初始态势下明显不同, 在 S1 下, PT-Agent

趋向于选择 COA1; 在 S2 下, PF-Agent 趋向于选择 COA3。可以根据 PT 的原理来解释这种决策差异, 指挥员 Agent 的任务是以最小伤亡解救人质, 这也是其决策参考点。在 S1 中, 因为到目前为止并无人员伤亡, 态势有利, PT-Agent 趋向于风险规避。在三种 COA 中, COA1 最为保守, COA3 最为冒险。因此, 其决策趋向于选择 COA1, 拒绝 COA3。而在 S2 中, 因为已经确知有人质伤亡, 态势不利, PF-Agent 趋向于风险追求, 因此其决策趋向于选择 COA3, 拒绝 COA1。

5 结论

为了更真实地描述指挥员在不确定战场条件下的决策特征, 本文对 PT 在 CGF Agent 决策建模中的应用进行了探索性研究。介绍了 PT 的基本原理, 分析了 PT 用于作战决策的适用性和可行性; 在此基础上提出了一种基于 PT 的 CGF Agent 决策建模方法, 对于所提出的 CGF Agent 的模型框架、决策流程、决策方法以及决策参数的确定方法进行了详细阐述; 通过给出的决策实例, 对本文的研究进行了验证。

作为 PT 应用于不确定决策的探索性工作, 本研究还存在大量后续工作: 首先, PT 本身还有很多需要解决的问题, 如, 参考点确定、容量函数选择等; 其次, PT 本质上也是一种基于效用的理论, 在作战决策过程中, 如何合理量化各种决策指标, 也是下一步需要研究的问题; 最后, 在决策参数的确定上, 下一步可以考虑采用神经网络或是机器学习等自动化水平更高的方法。

参考文献:

- [1] 杨建池. Agent 建模理论在信息化联合作战仿真中的应用研究[D]. 长沙: 国防科技大学, 2007.
- [2] 厉新光, 马金生. 作战指挥心理学[M]. 北京: 长征出版社, 2005.
- [3] Simon H A. A Behavioral Model of Rational Choice[J]. Quarterly Journal of Economics, 1955, 25: 99- 118.
- [4] Kahneman D, Tversky A. Prospect Theory: An Analysis of Decision Making Under Risk[J]. Econometrica, 1979, 47: 263- 291.
- [5] Tversky A, Kahneman D. Advances in Prospect Theory: Cumulative Representation of Uncertainty[J]. Journal of Risk and Uncertainty, 1992, 9: 195- 230.
- [6] 郭齐胜, 样立功, 等. 计算机生成兵力导论[M]. 北京: 国防工业出版社, 2006.
- [7] 庄锦英. 情绪与决策的关系[J]. 心理科学进展, 2003, 11: 423- 431.
- [8] Medemott R. Risk-Taking in International Politics: Prospect Theory in American Foreign Policy[M]. Ann Arbor: The University of Michigan Press, 1998.
- [9] Levy J S. Prospect Theory, Rational Choice and International Relations[J]. International Studies Quarterly, 1997, 41: 42- 54.
- [10] Taliaferro J W. Balancing Risks: Great Power Intervention in the Periphery[M]. Ithaca, Cornell University Press, 2004.
- [11] Schultz J V. A Framework for Military Decision Making Under Risks [M]. Air University Press, 1997.
- [12] Ganderton P T, Brookshire D S, McKee M, et al. Buying Insurance for Disaster-type Risks: Experimental Evidence[J]. Journal of Risk and Uncertainty, 2000, 20(3): 271- 283.
- [13] 孙多勇. 突发事件与行为决策[M]. 北京: 社会科学文献出版社, 2007.
- [14] 赵凇. 基于“前景理论”的路径选择行为建模及实例分析[D]. 北京: 北京交通大学, 2007.
- [15] 林民旺. 前景理论与外交决策[J]. 外交评论, 2006, 5: 124- 129.