

doi: 10.7690/bgzd.2013.04.010

基于多 Agent 构建认知主导决策模型

邵伟, 陈春, 陶金

(海军蚌埠士官学校信息技术系, 安徽 蚌埠 233012)

摘要: 为了加强军事仿真系统的决策能力, 建立基于多 Agent 构建认知主导决策模型。分析决策认知过程, 抓住人认知经验中的关键要素进行描述建模, 通过这些要素生成一个计算模型, 模拟认知主导决策中决策者的决策过程, 并以一个登陆战役的例子进行仿真计算。仿真结果表明: 认知主导决策模型相对于基于规划的决策模型更符合高度不确定、信息部完全、时间压力下的作战决策行为的特点和规律。

关键词: 决策模型; 认知主导决策; 多 Agent

中图分类号: TJ03 **文献标志码:** A

Recognition-Primed Decision-Making Model Based on Multi-Agent

Shao Wei, Chen Chun, Tao Jin

(Department of Information Technology, Bengbu Naval Petty Officer Academy, Bengbu 233012, China)

Abstract: To strengthen the decision-making capability of military simulation systems, build the recognition-primed decision based on multi-Agent. Analyze decision and recognition process, describe and establish model by capturing the key factors of person's cognitive experience. Create the calculation model by these factors, simulate the decision process of RPD, and simulate and calculate based on a landing campaign example. The simulation results show that the recognition-primed decision-making model is more suitable for the decision-making behavior with the uncertainty, un-complete information and the time pressure than the decision-making model based on regular.

Key words: decision-making model; recognition-primed decision; multi-Agent

0 引言

战争问题具有较强的复杂性和不确定性。影响战争决策行为的因素很多, 不仅有理性的, 而且有非理性, 其中, 人是最重要也是最活跃的非理性因素。传统的决策建模和仿真方法(诸如还原论方法以及归纳推理方法)不能很好地刻画战争决策行为, 还基本停留在对客观数据的分析与使用上, 几乎没有涉及人的因素, 如决策人员的认知能力、经验知识等, 缺乏对个人决策行为、组织决策行为等非理性因素成分的计算机分析, 其模拟结果并不能完全反映真实的决策过程, 故其模拟结果的可信度较低^[1]; 因此, 战争决策问题必须考虑那些影响决策行为的人的因素。为了对决策系统中人的因素进行定量描述, 就要建立起人的各种相应的数学模型。

将战争决策行为作为研究对象, 从人的因素角度出发, 基于 Agent 对其进行建模与仿真, 是目前军事决策建模与仿真领域最具活力、最具影响的方法之一。建立决策模型的关键是选择合适的决策方法表达决策者的决策认知过程。在实际作战中, 指挥员面临复杂、动态、高风险的战场环境, 通常需要对突发事件做出快速决策。这些决策大多是在不

确定的环境条件、不完备的信息、紧张的时间压力下, 通过对战场态势评估, 在有限的时间内从自己的知识和经验中选择一种可行而不一定是最优的行动方案, 即基于自然决策理论进行决策。

在基于自然决策理论的决策模型中, 最典型的是 Klein 提出的认知决策中的认知主导决策模型, 即 RPD 决策模型。RPD 决策模型重点是在对环境的评估而不是比较诸多决策方案。当完成对环境评估后, RPD 运用以往的经验生成一个决策方案, 当该方案基本满足决策者的要求时, 将付之实施; 否则, 将进行迭代直至决策者基本满意^[2]。要了解 RPD 决策模型如何模拟人的经验, 首先需要了解 RPD 的相关概念。决策者理解具体决策态势并选择与之适合的行动方案, 主要取决于过去的经验。在某个具体领域内的经验越丰富, 其经验与当前的态势就越匹配^[3]。在 RPD 概念框架下, 人的决策在很大程度上受到过去经验的影响。即 RPD 就是描述专家怎样制定决策。通过计算的形式表述专家经验对于生成模拟决策处理过程的模型非常关键。基于以上分析, 笔者建立一个基于多 Agent 的 RPD 决策模型, 称之为 RPD Agent, 通过利用多 Agent 系统仿真技术来

收稿日期: 2012-10-10; 修回日期: 2012-11-04

作者简介: 邵伟(1968—), 男, 安徽人, 学士, 副教授, 从事电子对抗、信息作战研究。

实现模拟决策者的决策过程。

1 决策认知过程分析

根据半个多世纪以来有关决策行为学的研究成果认为，影响决策者决策行为最突出的因素是决策环境和经验知识等。决策环境是指决策者决策时所有外界条件的总和。经验知识是指决策者的观念、经历和体验的总和，为决策行为提供了丰富的材料来源。决策者在进行决策活动时，首先要对决策环境进行认知，对环境进行认知之后运用以往的经验知识生成一个决策方案。经验知识主要包含 4 个要素：感知、目标、行动和期望。这 4 个要素系统地描述了决策者的认知过程。

感知：表示决策者理解和监控战场态势过程中所出现的重要的物质和精神反应，是指挥员对战场态势的一种感知反应。在理解和监控战场态势过程中，所有与态势发展相关的信息片断都汇集到决策者脑海之中。除此之外，作为一个专业的决策者，还需要知道哪些感受因素需要重点监控。感知通常是由决策者脑海中积聚的所有信息片断综合而成。

目标：是认知过程的重要组成部分。它表述了决策者试图达到的最终状态。决策者试图制定几个目标并选择付诸合适的行动，可以最大程度地实现这些目标。

行动：表示的是来自于决策者可供选择的一系列可能的决策。决策者凭借自身的经验，直观地判断哪个行动对决策实施有利，在脑海中先预估此行动是否适合当前态势的具体背景，如果适合，这项行动就变成决策；否则，就预估下一个行动，直到生成一个满意的结果。需要注意的是，此时的决策并不一定是最优方案。

期望：在决策过程中担当的是控制功能。表示的是评估标准，决策者依据这些标准评估态势的发展状况，以便确定由于态势变化或者决策效果不佳时需要做出哪些修正处理。

文献[4]中的基于 RPD 的海战仿真作战实体决策模型流程图与文献[5]中的 RPD 流程图都强调感知态势信息后对比是否有相似态势，如果没有相似态势，就需要获取更多的战场态势信息进行比对。在实际作战中，指挥员面临复杂、动态、高风险的战场环境，通常需要对突发事件做出快速决策。这些决策大多是在不确定的环境条件、不完备的信息、紧张的时间压力下，通过对战场态势评估，在有限的时间内从自己的知识和经验中选择一种可

行而不一定是最优的行动方案，所以有时候根本没有更多的时间和条件来获取更多的战场态势信息就需要做出决策；因此，笔者对此做出改进，流程图如图 1 所示。

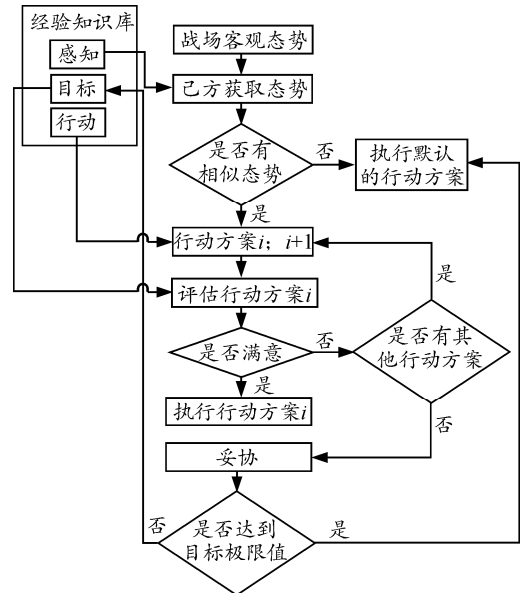


图 1 RPD 决策认知流程

2 经验知识数据结构

要利用 RPD Agent 对决策者的经验进行建模，必须建立相应的经验建模数据结构。文献[6]中第一次提出用一个数据结构来包含人的环境信息。基于文献[6]的研究成果，笔者选择一个数据结构来囊括 RPD Agent 经验的主要组成。RPD Agent 中的每个数据结构对应一条经验，包含了所有的感知、目标、行动等描述经验的因素。定义成如下的结构形式：

$$F = (C^*, G^*, A^*)$$

式中：F 表示数据结构；C*、G*、A* 都是一个结构，C* 包含了经验中所有的感知，G* 包含了经验中所有的目标，A* 包含了经验中所有的行动。一个具体的数据结构中，感知、目标、行动的数量和类型都是经验的具体体现。

当提出决策请求时，RPD Agent 搜索自身的数据结构表，寻找与之匹配的数据结构。如果没找到匹配的数据结构，就说明模型针对此种态势无经验可循来做出决策。如果找到匹配的数据结构，就会找到与之相关的感知、目标、行动等。以上的行为描述了 RPD 的认知过程。

一旦确定了数据结构，RPD Agent 必须开发对决策态势的感知，类似于人基于过去经验而产生的对外部环境的主观反应。为了实现这种感知，RPD

Agent 要将描述外部环境的环境变量聚合成对环境的内在主观判断的感知。感知由描述基本物质和精神参数的低层环境变量通过高层抽象聚合而成。

3 RPD 决策模型

3.1 感知模型

在 RPD Agent 中, 环境变量被分配了数值来表示变量变化的特征。值越高, 变量对感知的影响越大。感知值通过与之相关的环境变量值进行加法运算计算而得, 如式 (1) 所示:

$$cv_j = \sum_{i=1}^n e_{i,j} \quad (1)$$

式中: $e_{i,j}$ 是与第 j 个感知相关的第 i 个环境变量的值; cv_j 是第 j 个感知的值; n 是与第 j 个感知相关的环境变量的数目。这个表达式适用于所有的感知。

一旦所有的感知值计算出来, 进一步被转化成模糊值。这样处理的目的是因为人倾向于通过模糊判断来表述环境而不是通过离散值来表述。RPD Agent 基于每个感知值为每个感知生成一个模糊值。每个感知值有 3 种估计, 分别是: 不满意的模糊估计、基本满意的模糊估计和满意的模糊估计。感知值越高, 越易落入满意区域。三级模糊估计可以很好地评估估计值的状态以及估计值的变化, 参见文献[7]中讨论的模糊变量和模糊估计。这些模糊估计在 RPD Agent 中是描述经验的一个重要元素。

模型用户根据对 RPD Agent 所赋予的经验类型, 通过变换模糊估计的形状和范围, 来调节 RPD Agent 对某个感知的主观描述。基于过去的经验, 如果需要对某个感知进行不同的模糊描述, 可以生成不同的模糊估计。

3.2 行动模型

数据结构内的每个行动表示的是过去或者现在的决策选择。可以将行动看成是选择, 通过环境变量来表征。因此, 在某种意义上, RPD Agent 可以通过对每个行动求取感知值的和来确定行动的满意程度。计算公式如下:

$$AV_i = \sum_{j=1}^n cv_{j,i} \quad (2)$$

式中: $\sum_{j=1}^n cv_{j,i}$ 表示与第 i 个行动相关的所有的感知值的和; n 是与第 i 个行动相关的感知的数目; AV_i 表示第 i 个行动的行动值。行动值越大, 行动越被赞同。RPD Agent 通过这种方法来模拟决策者在进

一步评估之后, 对最赞成的行动的直觉认同。计算中用到的是感知数值而不是感知模糊值, 是因为这个计算仅仅是作为对最赞成的行动的一个直觉指标。RPD Agent 在选择行动之前必须进行进一步评估, 以选出最适合态势的行动方案。

3.3 目标模型

一旦确认了行动方案, RPD Agent 就会利用经验来评估所选行动方案是否满足态势的目标。在认知主导决策时, 决策者会通过精神模拟来评估所选行动方案是否满足当前态势。与每个目标相关的一系列感知, 通过这些感知来预估这些目标在何种程度上能够实现。目标与感知之间的联系是模型用户控制经验的另一个要素。目标同样存在模糊估计, 这些模糊估计主要是描述目标的实现程度。再一次使用不满意、基本满意和满意 3 个模糊判断指标。目标值的评估是通过计算与目标相关的感知得到的。公式如下:

$$GV = \sum_{i=1}^n (GV_c^i cw_i) \quad (3)$$

式中: GV_c^i 是与目标 j 相关的第 i 个感知模糊值的数值表示; cw_i 是第 i 个感知的权重; n 是与目标 j 相关的感知的数目; GV 是第 j 个目标的目标值。计算中涉及到感知的权重, 主要是因为决策者在考虑多个感知时, 存在轻重之分。上式适用于所有目标值的计算。

目标模糊值可以通过在目标值中加入模糊值函数处理而成, 与前面介绍的感知模糊值的处理类似。目标模糊值函数的描述是 RPD Agent 中经验描述的又一个重要部分。目标模糊值表示的是 RPD Agent 对每个目标通过行动实现目标程度的评估。

3.4 行动模型与目标模型的关系

需要指出的是, 行动模型与目标模型之间的关系是初评与二次评估的关系。笔者采取“两步评估”的形式来评估行动方案的满意度。第一步就是决策者凭直觉和经验对某个行动产生一个认同值, 即传统意义上的印象值。这个认同值不会以精确的方式表达出来, 而是形成一个大概的评估结果, 如图 2、图 3 所示; 第二步就是利用经验来评估所选行动方案是否满足态势的目标。决策者会通过精神模拟来评估所选行动方案是否满足当前态势。与每个目标相关的是一系列感知, 通过这些感知来预估这些目标在何种程度上能够实现。第二步评估是较为详细

的评估。



图 2 决策者在决策响应之前的心理状态

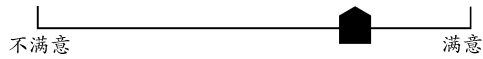


图 3 决策者对某个行动方案的直观认同值

如果第一步决策者就不认同所选的行动方案，就不会产生第二步的详细评估。第二步必须依赖第一步的评估数据。这个条件满足之后，就可以基于与每个目标相关的感知来评估这些目标在何种程度上能够实现。

在评估时，借助一种工具，在“满意”与“不满足”之间呈现出基本问题的连续集。响应者可以沿着连续集拖动指针(使用鼠标点击拖动功能)来显示自己的评估结果，当指针到达正确的位置时，再点击固定住按钮。开始位置如图 2 所示，拖动的位置如图 3 所示。由于决策者只需要在 2 种描述符之间指定一个位置，而不需要在一个模糊的刻度上给出一个精确的点，故响应将会很快。

3.5 期望模型

如果所有的目标都是评估满意，则此时被评估的行动方案就作为模型的决策。如果 RPD Agent 评估出一个或者多个目标在满意度上比其他目标更满意，就要看是否能通过妥协达到期望值。期望值通过妥协函数来实现。

一般情况下，决策者很难发现一种行动方案能完全满足其所有目标。经验告诉决策者，针对具体的目标，需要在多大程度上做出妥协仍然能达成一个满意的决策。RPD Agent 就是通过这种方式进行模拟仿真。在数据结构中为每个目标配备了一个修正 Agent。每个修正 Agent 负责评估其初始目标的实现程度。

当一个或多个目标不能完全实现时，修正 Agent 试图通过妥协达成期望值，主要是降低评估目标实现的标准。由于修正 Agent 是自主决策，不受其他 Agent 评价的影响。存在一个极限值，低于此极限值，修正 Agent 将无法达成妥协，达到期望值。妥协的极限值都是在 RPD Agent 中描述经验的参数，它们表示一个人解决目标冲突的能力。

修正 Agent 在前面的目标值中加入一个倍增因数得到一个修正目标值，从而完成其妥协任务。计算公式如下：

$$GV^n = GV \times RF \tag{4}$$

式中： GV 是前面计算的某个具体目标的目标值； RF 是风险因数； GV^n 是妥协目标值。风险因数表示 RPD Agent 对风险的容忍度。

当计算出新的目标值，就确定了相应的新的目标模糊值。如果此时所有的目标都完全实现，就形成了最终的妥协期望值，此时的行动方案就是最后的决策。如果无法达到期望值，就会选择下一个最适合的行动方案进行评估，评估方式同上。如果 RPD Agent 最终无法找到一个满意的行动方案，就会调用与态势相关的默认决策。

4 应用实例

一般的登陆战役主要由集结上船行动、海上航渡行动、突击上陆行动等一系列战斗行动组成。其中突击上陆行动是决定登陆战役成败的关键行动，在突击上陆行动中，正确选择登陆地域是确保突击上陆行动成功的重要前提。

在某次登陆战役中，指挥员在决定哪里作为登陆的地点时，需要将登陆地域的水文地理作为一个感知因素。指挥员想知道可供选择的登陆地域(每个登陆地域都对对应着一种行动或决策方案)中哪个最适合突击登陆。

水文地理的评估需很多环境变量，如水深、潮汐、水流等。与水文地理相关的环境变量通过某种方式组合，生成某个具体登陆地域的水文地理模型。

表 1 描述了水文地理感知的一个结构。每个变量有 2~3 个描述值(基于经验给出的描述值)以及相应的数值。带星号的值表示变量值在这次登陆区域决策的考虑之列。

表 1 水文地理感知结构

感知	环境变量	描述	数值
水文地理	暗礁	没有	2
		部分	1*
		大量	0
	水深	浅	2
		中等	1
		深	0*
	锚地	没有	0*
		有	2
	潮汐	小	2
		中等	1*
		大	0
	水流	弱	2
中等		1	
强		0*	

利用式 (1) 进行计算得知，如果最佳区域是 10(10 表示所有环境变量都取最佳值)，此区域的

感知值是 2。再将这个值进行模糊化处理, 如图 4 所示。一个值是 2 的感知, 其不满意的模糊估计的高度是 0.6, 基本满意的模糊估计的高度是 0.4, 利用二者之和作分母, 进行归一化处理, 则不满意的主观概率就是 0.6/1.0, 基本满意的主观概率就是 0.4/1.0, 产生一个随机数来作选择。

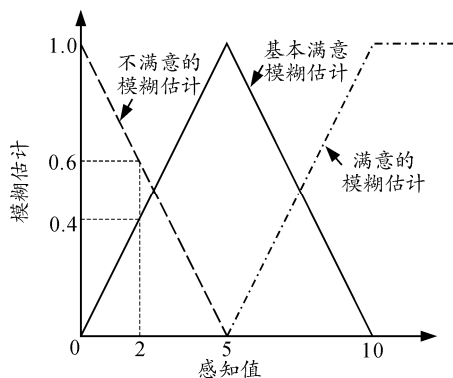


图 4 水文地理的模糊估计

以上是对某个登陆地域的其中一个感知因素进行模糊评估的一个应用实例, 在选择具体的登陆区域(即选择相应的登陆行动)时, 还需要考虑与之相关的多个感知因素, 如海滩地形、障碍物、集结区和路程等因素。行动模型就是将与之相关的感知因素的感知值求和, 得出行动值, 再结合具体的行动目标进行进一步的评估, 假设决策者对其中某个登陆地域候选方案的直觉认同值如图 5 所示, 此行动方案即遭到决策者否决, 决策者即开始预估另一种方案; 如果决策者对其中某个登陆地域候选方案的直觉认同值如图 6, 此行动方案留待在目标模型中结合具体的目标继续评估。

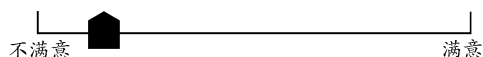


图 5 决策者对某个行动方案的直观认同值(一)

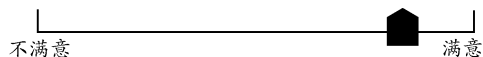


图 6 决策者对某个行动方案的直观认同值(二)

假设登陆区域决策的一个目标是完成登陆任务, 与之相关的感知有 5 个, 这些感知以及他们的模糊值、模糊值的数值表示见表 2。假设每个感知的权重相同, 目标值是 8(取 10 为最大值, 10 表示所有的模糊值都是满意状态)。

需要注意的是, 表 2 中得出的目标值与图 6 中的直观认同值, 尽管在具体数值形式上存在一定的差距(图 6 是粗略估算, 表 2 是具体计算), 但就其满意度而言基本保持一致, 即“两步评估”的结果基本保持一致。

表 2 目标评估实例

目标	感知	模糊值	数值
完成登陆任务	海滩地形	基本满意	1
	海滩水文地理	满意	2
	海滩障碍物	满意	2
	海滩集结区	基本满意	1
	目标路程	满意	2
目标值(GV)			8

5 结束语

RPD Agent模型用来改进军事仿真系统的决策能力, 尤其是作战层面的决策能力。构建RPD Agent有效模型的关键是通过数值计算的方式来描述人的知识和经验。RPD Agent抓住了人的认知经验中的关键要素进行描述建模, 主要包括: 1) 描述经验的行动、感知、目标的数量和类型; 2) 通过感知和模糊估计将人的外部环境转化为其内在感受; 3) 将目标评估与具体的感知相关联; 4) 使用模糊估计来评估目标的实现程度; 5) 通过妥协方法来描述RPD Agent的决策能力。

通过这些要素生成一个计算模型, 模拟认知主导决策中决策者的决策过程。仿真结果表明: 对已有的基于规则的决策模型, 需要较多的态势信息和较长的计算时间进行规则推理, 对不同的行动方案进行分析评估, 按照“最优原则”得出一个决策结果。而文中的决策模型按照“满意原则”, 直接从经验知识的相似态势中找出一个可行的行动方案作为决策结果。相对于前者, 文中决策模型更加符合高度不确定、信息不完备、时间压力下的作战决策行为的特点和规律。

参考文献:

- [1] 魏宾, 胡晓峰, 司光亚. 战争决策行为建模与仿真的研究[J]. 系统仿真学报, 2003(12): 1678-1682.
- [2] 魏宾, 胡晓峰, 司光亚. 浅谈战争决策行为多 Agent 建模[J]. 装备指挥技术学院学报, 2004(2): 5-9.
- [3] Klein G. Sources of Power: How People Make Decisions[M]. Cambridge: The MIT Press, MA, 1998: 35-47.
- [4] 吴扬波, 贾全, 朱一凡. 基于复合 Agent 的 RPD 决策模型[J]. 计算机工程, 2010(6): 288-290.
- [5] 孟庆操, 赵晓哲, 姜伟. 基于 Recognition-Primed Decision 模型的多智能体作战仿真[J]. 系统仿真学报, 2011(2): 294-299.
- [6] Minsky M L. A Framework for Representing Knowledge [M]. In Computation and Intelligence, Luger, G.F. Ed., Massachusetts Institute of Technology A.I. Laboratory: AAAI Press, Menlo Park, 1995: 163-189.
- [7] Zadeh L A. Fuzzy sets and their applications to cognitive and decision processes[M]. New York: Academic Press, 1975: 65-78.