



国防科技大学学报

Journal of National University of Defense Technology

ISSN 1001-2486, CN 43-1067/T

## 《国防科技大学学报》网络首发论文

题目：多方危机博弈的“遗传—动态时间规整”策略偏好拟合  
作者：刘万里，陈俞舟，刘鑫，傅曼青，李知，唐世平  
收稿日期：2024-04-03  
网络首发日期：2025-12-26  
引用格式：刘万里，陈俞舟，刘鑫，傅曼青，李知，唐世平. 多方危机博弈的“遗传—动态时间规整”策略偏好拟合[J/OL]. 国防科技大学学报.  
<https://link.cnki.net/urlid/43.1067.t.20251225.1008.002>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

doi: 10.11887/j.issn.1001-2486.24040006

## 多方危机博弈的“遗传-动态时间规整”策略偏好拟合

刘万里<sup>1</sup>, 陈俞舟<sup>1</sup>, 刘鑫<sup>1</sup>, 傅曼青<sup>2</sup>, 李知<sup>2</sup>, 唐世平<sup>2\*</sup>

(1. 中国电子科技集团公司第十研究所, 四川 成都 610036; 2. 复旦大学 复杂决策分析中心, 上海 200433)

**摘要:** 为实现复杂多方危机博弈场景下各方行为决策的分析预测, 本研究提出了结合行为体建模与“遗传-动态时间规整 (dynamic time warping, DTW)”的策略偏好拟合方法。通过拟合多方危机博弈中的智能行为体的历史时序数据, 模拟了国际行为体的危机决策过程, 并预测了系统状态变化趋势。通过将专家知识与机器学习相结合, 本文模型在多方博弈行为体策略偏好拟合方面取得了初步成果, 在历史案例状态序列预测任务中取得平均 DTW 距离 9.35。本文的方法为多方危机博弈问题提供了全新的研究路径。本文提出的建模方法与模型校准方法能够被广泛应用到多边多轮谈判、多边经济制裁、区域冲突等其他多方复杂博弈场景。

**关键词:** 行为体建模; 计算机模拟; 遗传算法; 动态时间规整; 非常规复杂博弈; 国际冲突管理

中图分类号: N945.25 文献标志码: A

## Genetic-dynamic time warping strategy preference fitting in multi-party crisis games

LIU Wanli<sup>1</sup>, CHEN Yuzhou<sup>1</sup>, LIU Xin<sup>1</sup>, FU Manqing<sup>2</sup>, LI Zhi<sup>2</sup>, TANG Shiping<sup>2</sup>

(1. The Tenth Research Institute of China Electronics Technology Group Corporation, Chengdu, Sichuan 610036, China; 2. Center for Complex Decision Analysis, Fudan University, Shanghai 200433, China)

**Abstract:** To analyze and predict decision-making in complex multi-party crisis game scenarios, a novel strategy preference fitting method that integrates agent-based modeling with the “Genetic-DTW (genetic-dynamic time warping)” algorithm is proposed. By fitting historical time-series data from intelligent agents in multi-party crisis game, this study simulates the decision-making processes of international actors during crises and predicts system state transitions. Combining expert knowledge with machine learning, our model achieves promising results in capturing strategy preferences of actors in multi-party games, obtaining an average DTW distance of 9.35 in historical case state sequence prediction tasks. This approach provides an innovative research path for understanding multi-party crisis decision-making. The proposed modeling and calibration methods can be widely applied to other complex multi-party game scenarios, including multilateral multi-round negotiations, multilateral economic sanctions, and regional conflicts.

**Keywords:** agent-based modeling; computer simulation; genetic algorithm; dynamic time warping; unconventional complex games; international conflict management

收稿日期: 2024-04-03

基金项目: 中电天奥创新理论技术群基金资助项目 (2022-1427-04-01)

第一作者: 刘万里 (1989—), 男, 四川南充人, 高级工程师, 博士, E-mail: liu.wli@163.com;

\*通信作者: 唐世平 (通信作者), 男, 湖南郴州人, 教授, 博士, 博士生导师, E-mail: twukong@fudan.edu.cn

引用格式: 刘万里, 陈俞舟, 刘鑫, 等. 多方危机博弈的“遗传-动态时间规整”策略偏好拟合[J]. 国防科技大学学报.

Citation: LIU W L, CHEN Y Z, LIU X, et al. Genetic-dynamic time warping strategy preference fitting in multi-party crisis games[J]. Journal of National University of Defense Technology.

从国际关系理论上来说,危机和冲突都是“讨价还价问题”(Bargaining)或者说是“博弈问题”<sup>[1]</sup>。常规博弈场景可以包含两到三个行为体,并在享有“共同知识”(Common knowledge)的情景下进行有限轮次的博弈。而在危机场景下的非常规多方战略博弈具有行为体众多、行为及行为规则多样、收益模糊、系统特性复杂等特点,且缺乏“共同知识”。在复杂多方的博弈场景下,分析并推断各方应对策略是极为困难的——同样处于危机冲突条件下,为什么某些行为体选择协商妥协,而某些行为体选择对抗升级?具体而言,应当如何量化国际行为体在复杂博弈场景下的行为逻辑?又如何根据多方博弈中的行为体博弈偏好预测危机走向?只有解决了上述问题,才能更加深入地理解当下和未来的国际危机冲突,实现对冲突进程的预测与管控。

现有的危机决策分析框架主要分为效用分析和计算模拟两大类。效用分析的主要研究方向包括经典博弈论与多元启发理论(poliheuristic theory)。基于经典博弈论的研究主要集中于简化的两方至三方博弈<sup>[1]</sup>,而多元启发理论通过将认知决策研究与理性效用分析结合,完成了对多个复杂现实案例的解释分析<sup>[2]</sup>。但该类模型受限于效用分析数学框架,很难用于分析超过三方的博弈场景,且模型构建基本依靠专家知识,可靠性与可迁移性不佳。

计算模拟方面,现有研究主要使用基于行为体的建模(agent-based modeling, ABM)来捕捉复杂的社会动态<sup>[3]</sup>。尽管 ABM 在描绘社会演化的涌现性结果上具有优势,但在多方冲突博弈领域,ABM 的应用大多仍局限于经验性的模型配置,研究目的局限于验证个别的国际关系理论,缺乏模型与现实实证数据的验证校准,很难将研究发现迁移至现实决策过程<sup>[4]</sup>。

总体而言,以往的研究通常依赖于专家构建的博弈论模型,这些模型往往将多方危机博弈视为简单的理性选择问题,或者是用讨价还价模型等数理模型来简化描述复杂的博弈决策过程。这类方法过分简化了国际危机博弈的复杂性和多样性,模型输入输出难以与现实数据进行比对校准,对现实博弈决策过程缺乏解释力。

为填补现有研究在复杂多方危机博弈问题上的空白,本文提出了基于行为体建模与“遗传—动态时间规整 DTW (dynamic time warping, DTW)”策略拟合的多方危机博弈推演模拟方法,在历史时序数据上训练符合实际国际行为体决策

方式的智能行为体,并在国际关系理论指导的环境框架中运行 ABM 模拟,解释和预测行为体在危机和冲突中的决策过程。经过不断完善,本研究提出的建模方法将有望为现实多方危机博弈决策提供支持。

## 1. 基于行为体建模的危机冲突模拟技术

在危机冲突研究领域,西方国际关系学者自 20 世纪 70 年代起就已开始运用 ABM 建模方法验证国家间冲突的相关理论。1977 年斯图尔特·布雷默(Stuart Bremer)与迈克尔·米哈尔卡(Michael Mihalka)构建的“马基雅维利自动机”(Machiavelli in Machina)通常被认为是将计算机模拟方法应用到国际冲突领域的奠基性研究<sup>[5]</sup>。

基于布雷默与米哈尔卡的建模框架,西方学者不断拓展模型规模与复杂度,并基于计算机技术的进步持续改进模拟程序底层架构,借助计算机模拟探究现实主义视角下各影响因素如何影响国家宏观行为决策。目前,ABM 建模方法已经被用于探索包括恐怖网络和毒品网络在内的众多不同的复杂社会问题<sup>[6-7]</sup>。表 1 简述了 1990 以来一些有代表性的研究。

表 1 1990 以来代表性的国际冲突 ABM 模型

Tab. 1 Selected ABM models of international conflict, post-1990s

模型名称	研究意义	参考文献
EARTH	通过简单行为规则和局部互动分析不同因素条件下的国家冲突行为及结果。	[8]
GeoSim	在传统国际冲突模型中纳入央地关系模拟,以更好地捕捉国内政治与结盟策略对边界变化的影响。	[9]
RealLand	使用 NetLogo 框架,额外纳入地形、资源、贸易等建模要素,使模型更加贴近现实。	[10]
WarAgent	不再使用专家制定的模拟规则,而是直接将大语言模型作为博弈行为体。	[11]

但是,这类现有的 ABM 系统存在以下问题:

1)绝大多数现有模型都是基于简化的博弈场景实现的,很少能做到与现实数据进行对比验证。这一问题在很大程度上是因为这些系统均没有来

自微观层面决策数据的验证支持<sup>[12-13]</sup>。

2) 大多数针对冲突的计算机模拟模型都因过分简化而忽略了结盟以外的一切外交博弈。现实世界中, 战争是政治的延续, 外交博弈在冲突爆发与缓和的过程中扮演着不可或缺的角色<sup>[1, 14-15]</sup>。

3) 要真正实现推演预测多方危机博弈进程, 需要对国家战略意图、国家政策偏好进行准确刻画, 并考虑冲突与外交的紧密交互关系。

本文提出的面向危机博弈的多行为体多回合博弈的 ABM 模拟方法旨在突破国内外研究在上述各方面的局限。

首先, 首次尝试从微观层面获取决策数据, 对模型进行训练和校准, 实现了基于现实数据的 ABM 建模和危机博弈状态预测。这一方法能够有效解决现有模型过于简化、缺乏现实数据验证支持的问题, 提高模型的准确性和可靠性。

其次, 提出的 ABM 模型涵盖多种外交博弈行为和战略决策, 不再局限于结盟与冲突行为。行为体的行为策略包括提出、接受、拒绝、执行等, 涵盖经济制裁、军事介入等冲突外交的主要议题。通过模拟行为体在多方谈判中的多维度互动, 以及冲突与外交的相互作用, 使得模型更加契合国际安全研究的核心理论和实证进展。

最后, 选择从战略层面而非战术层面入手探究多方危机博弈过程。模型将战场胜负等战术效果作为行为体观察状态值纳入模拟过程, 结合行为体行为偏好模型与议题偏好模型指导行为体决策, 能够模拟行为体在不同战术条件下采取的战略决策和互动策略, 实现以宏观战略博弈为主体的多方危机博弈理论建模与校准优化。

## 2. 基于马尔可夫决策过程的多方危机博弈建模

### 2.1 多方危机博弈建模要素

多方多回合的博弈决策过程通常具有极强的不确定性和动态性。行为体需要在未知或部分已知的环境中, 根据当前状态和可观测信息, 作出最优或次优决策。为了有效地模拟这种复杂的决策过程, 采用马尔可夫决策过程 (Markov decision process, MDP) 作为主要的建模框架:

首先, MDP 能够有效刻画不确定性和动态环境。多方博弈的决策过程充满了不确定性, MDP 恰恰能够有效地描述这种在不确定环境下的决策过程, 将复杂难以计算的博弈树转化为便于计算的马尔可夫决策问题, 借助状态转移概率刻画环

境反馈。

其次, MDP 天然适用于多轮决策问题, 而危机博弈恰恰就是一个多回合决策过程。MDP 框架能够帮助模拟行为体以序列方式执行行动, 兼顾行为体在不同阶段的策略演化和决策变化。

最后, MDP 框架能够为策略优化和学习算法提供良好基础。通过参数化行为体价值函数, 行为体可以通过学习算法不断改进策略, 不断逼近现实国家行为体的决策偏好。由于难以使用一个奖励函数量化行为体行为与历史行为模式的异同, 传统强化学习难以发挥作用, 本文转而采取遗传算法作为行为体策略优化的主要手段。

本文采用的 MDP 建模框架主要包括  $S, P, A, O$  四大组成要素:

1. 环境状态 ( $S$ ): 系统整体所处的状态。在本文重点刻画的多行为体危机冲突博弈中, 环境状态可以被描述为“危机状态”, 代表冲突各方所处环境的紧张程度或冲突强度。

2. 状态转移概率 ( $P$ ): 给定行为后环境状态转移至其他状态的概率分布。换句话说, 在马尔可夫假设下, 系统的下一状态直接依赖于当前系统状态及行为体动作。基于此, 在历史危机博弈数据上训练了状态转移概率矩阵  $P$ , 用于在模拟中向行为体提供环境状态变化的反馈。

3. 冲突各方的动作空间 ( $A$ ): 冲突各方所有可能的行动或决策。例如, 一方可能选择提出博弈诉求、接受博弈诉求、拒绝博弈诉求, 或采取行动。上述行为选择还可进一步依据博弈行为议题领域具体划分, 如行为体可以在战俘议题领域提出诉求, 也可以要求对方行为体停火撤军。通过灵活组合博弈行为与议题领域, 本文模型能够在有限动作空间中刻画复杂的博弈决策, 且支持通过修改议题领域来适应不同的危机冲突场景。

4. 观察状态 ( $O$ ): 行为体各自观测到自身所处的状态。在多方危机博弈场景下, 行为体的行为决策不仅受环境整体状态影响, 还受到行为体自身处境的制约。因此, 本文采用的 MDP 模型额外纳入了观察状态集合  $O$ , 从国内国外两个维度刻画行为体决策时所处的内部与外部条件。

从内部的角度来看, 一个关键的观察要素是内部稳定性。例如, 领导人的更换、政府或组织的结构变动、大规模的社会动荡, 这些都是可能影响行为体决策的关键因素。

从外部的角度来看, 军事优劣势是模型使用的另一个观察维度。这一观察值描述了行为体在战场上的实际表现。例如, 如果行为体在重大战

役中取得了胜利，那么其军事优劣势的观察值将被设定为高；如果遭遇失败，则设定为低。

总体而言，MDP 为本文提供了一个可靠有效的建模框架，使模型能够模拟复杂情境下的行为体决策，为多方博弈场景下的行为体偏好拟合提供了必要的建模基础。

## 2.2 多方危机博弈时序数据结构定义

基于 MDP 的基本架构与上述模型框架设计，本文提出的博弈行为序列数据结构如表 2 所示，包括系统状态事件时间、行为体双方、行为策略、行为类别与环境特征变量向量。

表 2 结构型案例数据编码标准

Tab. 2 Coding standards for structural event data

列变量	编码含义
时间	动作发生的时刻标记。时间表达式要求符合 ISO-8601 标准，一个有效的数据示例：“2022-01-02”。
环境状态	当前系统状态参数。“-1”表示危机结束，“0”表示危机，“1”表示重大危机，“2”表示只谈，“3”表示边谈边打（低烈度），“4”表示边谈边打（高烈度），“5”表示打（低烈度），“6”表示打（高烈度）。
行为体 A	动作的主动方。通常为国家的 ISO 3166-1 三位字母代码，如“USA”。
行为体 A 角色	行为体的类别。“0”第一交战方及其盟友；“1”第二交战方及其盟友，“2”中立第三方，“3”霸权国。
行为	行为体 A 的博弈行为。可以扩展。“0”表示提出诉求，“1”表示接受诉求，“2”表示拒绝诉求，“6”表示执行手段。
议题领域	行为体 A 的博弈行为具体行为议题领域。见表 4。
行为体 B	动作的被动方。即行为体 B 的独立唯一标识 (Unique ID)，通常为国家的 ISO 3166-1 三位字母代码，如“USA”。
行为体 B 角色	行为体的类别。“0”第一交战方及其盟友，“1”第二交战方及其盟友，“2”中立第三方，“3”霸权国。
观察值 1	行为体内部稳定性变化。包括精英凝聚力、社会凝聚力、领导人安全、领

导人更替等。“-1”变差，“0”不变，“1”变好。

观察值 2 战场优劣势的变化。“-1”变差，“0”不变，“1”变好。

其中，系统状态值 (State) 对应 MDP 中的环境状态  $S$ ，代表了危机的总体紧张程度。当冲突状态发生变化时，行为体可能会决定更改其策略，例如由防守转为进攻，或由激进转为温和，或由拒谈转向妥协。本文试图通过状态转移模型  $P$  与行为偏好模型的结合来刻画此类深层次的行为体行为逻辑。

在数据编码过程中，本文将系统状态值分为三个主要分量指标：危机紧张程度、谈判有效程度和冲突剧烈程度。为将以上抽象概念操作化为量化指标，首先统计过去一个月内的历史事件，再基于各分量指标相关的历史事件数量编码每一时刻的危机紧张程度、谈判有效程度和冲突剧烈程度三项指标值，并最终合成总体系统状态值。行为事件数据筛选标准如表 3 所示：

表 3 系统状态值指标分量行为数据对应表

Tab. 3 Mapping of system state value indicators to behavioral data components

指标分量	事件数据筛选条件
危机紧张程度	交战方提出或拒绝诉求
谈判有效程度	交战双方之间达成妥协
冲突剧烈程度	交战双方之间发生军事冲突

行为 (Action) 与议题 (Issue) 的组合对应 MDP 中的动作空间  $A$ 。针对博弈议题领域，基于剑桥大学与联合国提出的和平语言条约数据库 (Language of Peace) 条款类别树进行补充修改，结合历史危机博弈案例实际议题领域分类，确定了 13 类主要的博弈议题领域 (表 4)，通过与博弈行为策略 (如提出、接受、拒绝、行动) 结合，共支持 52 类博弈行为。

表 4 博弈议题领域定义表

Tab. 4 Definition for game issue areas

编码	行为类别	定义与描述
0	交换战俘	敌对方交换战俘
1	重定边界线	确认或修改争议地区边界
2	停战	暂时停止敌对行为
3	赔偿	对战争损失提供金钱或资

源补偿		
4	谈判议程	确定和平谈判的主题和条件
5	撤军	军队退出争议或冲突地区
6	军事介入	军事干预
7	独立与国际承认	寻求并获得独立国家地位认可
8	经济制裁	实施经济措施以施加政治压力
9	封锁	限制敌国交通与通信
10	对外宣传	影响国际观点的信息传播
11	战后政权	建立或调整战后政权结构
12	战后援助	提供战后经济和技术援助

最后，行为体内部稳定性变化（观察值 1）与战场优劣势变化（观察值 2）的结合代表了 MDP 的观察状态  $O$ 。这些外部条件同样深刻影响了行为体的行为决策，因此本文的行为偏好模型也将这些条件考虑在内，用于调控行为体在不同状态下的策略选择。

依据上述数据结构与编码规范，本文基于历史档案与二手专著文献对二战后共 15 个重要的危机冲突事件案例进行了数据编码。数据来源包括历史学与国际关系学专著、外交官回忆录，以及历史新闻。这些数据被用于模型训练和检验。

### 2.3 多方危机博弈 ABM 模型构建

基于 ABM 建模方法，设计实现了针对多方复杂博弈的模拟推演模型。

首先，相较传统博弈模型，本文模型认为国际关系中的行为体遵循有限理性，即行为体更多地依赖于过去的经验和可获得的信息进行决策，而不是在完全信息上进行收益最大化分析。本文假设行为体的政策具有一定的连续性和惯性，不会在短时间内发生剧烈改变。也即历史经验对当前决策有显著影响。

其次，在国际危机博弈中，不同行为体的政策偏好和议题关注点存在差异。一方面，行为体对不同行动（如提出、接受、拒绝、执行）的倾向性不同，这取决于行为体自身的意图与决心。另一方面，在不同的谈判场景下，行为体对各个议题领域（如停战、撤军、赔偿）的重视程度不同。这与其战略需求密切相关。为此，本文为每个行为体设置了两套参数：行为偏好矩阵  $Q$  用于刻画行为体在不同状态和观察下选择各个行动的倾向性，通过历史行为数据进行拟合和更新；议

题偏好矩阵  $M$  用于描述行为体对其他行为体在各个议题上的诉求强度，数值化地体现行为体在不同议题上的优先级。

最后，多方危机博弈是一个多行为体多回合的交互过程，与双边或三边博弈不同，多方博弈中多个个体间的多轮交互完全可能导致整体系统出现复杂非线性的模式变化。

基于上述理论假设，可以将多方危机博弈过程建模为一个动态的多智能体交互过程。根据 MDP 框架，模型设置行为体集合  $I$ 、危机状态集合  $S$ 、状态转移矩阵  $P$ 、动作空间  $A$ 、观察集合  $O$ ，在每一回合行为体将根据行为偏好模型  $Q$  与议题偏好模型  $M$  决定本回合采取的行动与行动所在的议题领域，随后系统对行为体行为做出反馈，依据状态转移矩阵  $P$  更新系统状态。模型输出包括系统状态序列  $\hat{y}_S = \{s_0, s_1, s_2, \dots, s_n\}$  与博弈事件序列  $\hat{y}_E$ 。模型输出中，每个事件可以表示为一个四元组：

$$e_t^k = (i_t^k, j_t^k, a_t^k, d_t^k) \quad (1)$$

其中， $i_t^k \in I$  为发起行动的行为体， $j_t^k \in I$  为行动对象， $a_t^k \in A$  为行为类型， $d_t^k$  为行为议题领域。

模型在每个时间步  $t$  生成一个事件集合：

$$E_t = \{e_t^1, e_t^2, \dots, e_t^{K_t}\} \quad (2)$$

其中， $K_t$  是时间  $t$  发生的事件总数。整个模拟过程中，事件序列可表示为：

$$\hat{y}_E = \{E_0, E_1, E_2, \dots, E_T\} \quad (3)$$

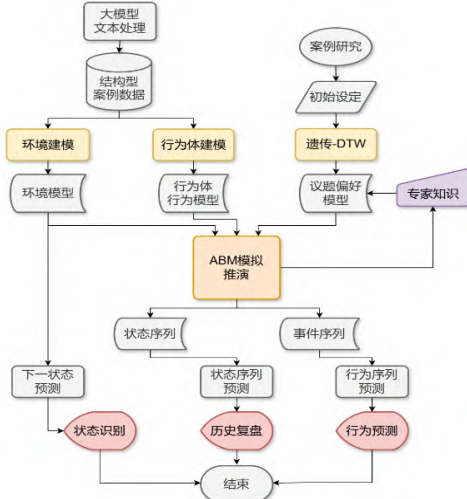


图 1 ABM 系统架构图

Fig. 1 ABM system architecture

### 2.3.1 行为体设定

模型中的行为体均为独立的博弈智能体，主要包括在危机博弈进程中具有自主决策能力的国家、地区武装组织、国际组织及其他第三方。

每个行为体都具有数个初始参数，如收益期望，取值在 0~10 之间，用于表示该行为体期望在博弈中获得的总收益；冲突立场，取值为-1、0 或 1，由于绝大多数多方危机冲突通常仅包含两大对立阵营，该参数取值范围完全能够刻画复杂危机事件中行为体的阵营分布。

在每个危机博弈场景中，行为体数量依据实际情况确定，数量通常不超过 10 个。每个行为体都具有独立的行为逻辑与行为能力，通过在前述收集的历史时序数据上进行机器学习训练与偏好拟合，行为体将基于历史案例同阵营行为体的行为规律选择行为类别（如提出、接受、拒绝、执行等），再基于当前事件案例的议题偏好矩阵选择行为议题领域（如交换战俘、重定边界线、停战、赔偿等）。

### 2.3.2 行为、行为规则与交互规则设定

模型的每一时刻即一个博弈回合，在一个时刻内，每个行为体将基于历史行为决策的访问次数表决定行动的行为类别，再基于当前事件时序数据拟合得到的议题偏好矩阵决定行动所在领域。

#### 1. 行为偏好模型

首先，系统通过行为策略访问次数模型对行为体进行行为策略拟合，使模拟行为体在行为

(Action) 选择上模仿历史博弈行为决策。

模型定义状态  $s$  为危机状态，状态集合  $S = \{s_1, s_2, \dots, s_n\}$ ，定义行动  $a$  为行为体的行动，行动集合  $A = \{a_1, a_2, \dots, a_m\}$ ，定义观察  $o$  为行为体对国内政治与战场形势变化情况的观察，观察集合  $O = \{o_1, o_2, \dots, o_k\}$ 。

访问次数表  $Q \in R^{n \times m \times k}$ ， $n$  为系统状态取值数量， $m$  为动作取值数量， $k$  为观察取值数量。矩阵每个元素  $Q(s_i, a_j, o_l)$  代表在状态  $s_i$ 、观察  $o_l$  下采取行动  $a_j$  的访问次数。

状态为  $s_1$  时的  $Q^{(s_1)}$  可以表示为：

$$Q^{(s_1)} = \begin{bmatrix} Q(s_1, a_1, o_1) & Q(s_1, a_1, o_2) & \dots & Q(s_1, a_1, o_k) \\ Q(s_1, a_2, o_1) & Q(s_1, a_2, o_2) & \dots & Q(s_1, a_2, o_k) \\ \vdots & \vdots & \ddots & \vdots \\ Q(s_1, a_m, o_1) & Q(s_1, a_m, o_2) & \dots & Q(s_1, a_m, o_k) \end{bmatrix} \quad (4)$$

在状态为  $s_1$ 、观察为  $o_1$  的条件下，行为体将执行该条件下历史访问次数最多的行动：

$$a^*(s_1, o_1) = \operatorname{argmax}_{a_j} Q(s_1, a_j, o_1) \quad (5)$$

#### 2. 议题偏好模型

不同危机事件场景下各方关注的议题不同，因此行为体对议题 (Issue) 的选择将进一步由行为体在当前场景下的议题偏好矩阵调控。

每个行为体的议题偏好矩阵  $M \in Z^{n \times m}$ ，其中， $n$  为行为体数量， $m$  为议题领域取值数量。矩阵元素  $a_{ij}$  代表了该行为体对第  $i$  个行为体在第  $j$  个议题上的诉求强度，取值 0 代表诉求较弱，10 代表诉求强烈。

行为体在选取议题领域时，将从议题偏好矩阵中寻找一个目标行为体  $i^*$  和一个议题  $j^*$ ，使  $a_{i^*j^*}$  取得  $a_{ij}$  中的最大值：

$$(i^*, j^*) = \operatorname{argmax}_{i,j} a_{ij} \quad (6)$$

#### 3. 信息交互规则

在选定行为与行为领域后，行为体的行动将对其他行为体产生交互作用。每个行为体都维护一个诉求提出列表，一个诉求接收列表，以及一个宣战列表，用于管理行为体交互状态 (表 5)。

表 5 行为体行为规则设定表  
Tab. 5 Behavior rules for agents

行为	描述
提出诉求	将诉求添加到自身提出列表末尾；同时将该诉求置入对方的接收列表末尾，如果该诉求已存在，将该诉求置顶。
接受诉求	从接收列表中取出顶端诉求，表示接受。如接受的诉求为重定边界线、谈判议程、国家独立或解除经济制裁，则对方获得数值等同于偏好矩阵中诉求值的收益。如果接受的是停战诉求，则从宣战列表中移除交战双方。
拒绝诉求	从接收列表中取出顶端诉求，表示拒绝。从对方提出列表中删除该诉求。
执行行动	行为体向自身偏好矩阵中诉求最高的敌方阵营行为体宣战，将对方添加到自身宣战列表，将自身添加到对方宣战列表。

### 2.3.3 环境设定

博弈环境主要用于刻画危机博弈过程中各方所处的外部环境，包括危机状态、战场形势，以及国际局势（比如国家间关系）的变化。

#### 1. 环境状态转移模型

基于马尔可夫假设，系统状态将始终基于上一步的系统状态及上一步的行为体行为发生变化。为刻画这一变化过程，本文采用了一个在历史时序数据上训练得到的状态转移模型，用于在 ABM 模拟中指导系统完成危机状态转换。

假设系统有状态集合  $S = \{s_1, s_2, \dots, s_n\}$ ，及行动集合  $A = \{a_1, a_2, \dots, a_m\}$ 。系统状态  $s$  表示危机的不同发展状态，行动  $a$  代表在这些状态下可能采取的行动。

状态转移矩阵  $P^{(j)} \in R^{n \times n}$ ，其中  $n$  是系统状态的数量。矩阵的元素  $P_{ik}^{(j)} = P(s_k | s_i, a_j)$  表示在观察到行动  $a_j$  的情况下，从状态  $s_i$  转移到状态  $s_k$  的概率。

$$P^{(j)} = \begin{bmatrix} P(s_1 | s_1, a_j) & P(s_2 | s_1, a_j) & \cdots & P(s_n | s_1, a_j) \\ P(s_1 | s_2, a_j) & P(s_2 | s_2, a_j) & \cdots & P(s_n | s_2, a_j) \\ \vdots & \vdots & \ddots & \vdots \\ P(s_1 | s_n, a_j) & P(s_2 | s_n, a_j) & \cdots & P(s_n | s_n, a_j) \end{bmatrix} \quad (7)$$

矩阵中，每一行代表一个起始状态，每一列

代表一个可能的下一个状态。如果行动  $a_j$  在状态  $s_i$  下发生了  $N_{ij}$  次，并且在该条件下，转移到状态  $s_k$  的次数是  $M_{ijk}$ ，那么转移概率  $P(s_k | s_i, a_j)$  可以估计为：

$$P(s_k | s_i, a_j) = \frac{M_{ijk}}{N_{ij}} \quad (8)$$

在推演模拟开始时，模型初始化一个概率分布  $\pi = [\pi_1, \pi_2 \dots \pi_n]$ ，概率初始值

$$\pi_i = \frac{1}{n}, i \leq n, i \in N, n \text{ 为系统状态的数量。}$$

在每个时间步骤  $t$ ，遍历所有该时刻的行为  $a_t$ ，更新概率分布：

$$\pi' = \pi P^{(t)} \quad (9)$$

其中， $P^{(t)}$  是对于行动  $a_t$  的所有状态转移概率矩阵。

系统的下一状态  $s' = \operatorname{argmax}_i \pi_i$ 。

#### 2. 战争成本

除系统状态转移外，环境还将在每一时间步更新每个行为体的战争成本。环境设有一个战争成本因子  $\alpha = 1$ ，每一时间步，所有参战的行为体损失战争成本  $\alpha$ 。

此外，交战时间越长，战争成本越高，行为体崩溃的风险（Risk of collapse）越高，向敌方要求停战求和的可能性也就越高<sup>[7]</sup>。因此，各停战相关议题都设有一个停战诉求因子  $\beta$ ，每一时间步，参战行为体对敌方行为体在各停战相关议题上的诉求强度增加  $\alpha\beta$ 。停战相关议题的诉求因子列举如表 6：

表 6 停战诉求因子初始设定表

Tab. 6 Initial settings for ceasefire demand factors

议题	因子初始值
解除封锁	1
撤军	2
停战	3
交换战俘	0.5
赔偿	0.25

#### 3. 结合机器学习的博弈议题偏好拟合



不同博弈场景下所涉议题种类不同，各行为体对各议题的重视程度也不同，举例而言，某个问题在某个危机博弈中是核心议题，但在其他停战谈判中很可能只是一般性议题。为使模型能够捕捉行为体在特定博弈场景下的议题偏好，提出了结合机器学习的“遗传-DTW”行为体议题偏好拟合方法。

本文行为体议题偏好拟合方法的核心是结合了遗传算法与动态时间规整（dynamic time warping, DTW）算法的组合优化算法。其中，遗传算法被用于探索议题偏好参数空间，寻找可能的最优解，而 DTW 则用于评估历史博弈行为序列与模型生成的博弈行为序列之间的相似度，从而作为遗传算法的适应度函数。这种结合保证了在一个宽泛的参数空间中找到一个较好的初始解。

议题偏好拟合的对象是议题偏好参数矩阵  $M$ ，矩阵中的每个元素代表了某一行为体对另一行为体在某一议题上的诉求强弱。该参数定义一方面便于模型使用者基于专家知识评估模型对历史博弈场景的拟合效果，另一方面，在未来也能够支持借助专家知识对算法产生的策略偏好进行分析校准，并针对性地提出建议和调整。

### 3.1 基于遗传-DTW 算法的议题偏好拟合

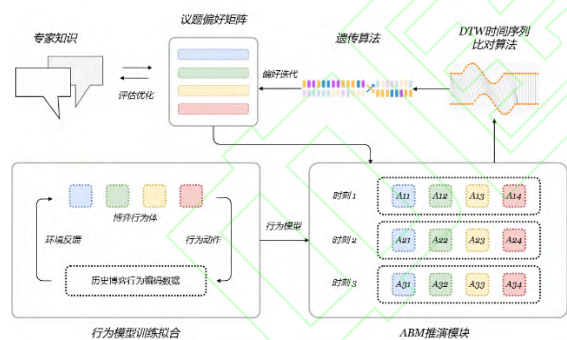


图 2 行为体策略拟合原理图

Fig. 2 Agent strategy fitting principle

#### 3.1.1 基于 DTW 算法的适应度函数

由于 ABM 系统本身的不确定性和随机性，ABM 系统得出的状态序列通常很难直接与真实值进行比较<sup>[16]</sup>。这一方面是因为 ABM 系统生成的预测数据庞杂且多噪声，另一方面是因为 ABM 系统的时刻概念很难与现实对应。通常 ABM 系统中每个行为体行动一次计一个时刻 (tick)，但一个时刻究竟对应现实的 1 小时还是 1 天并不确定，并且这一对应关系也完全可能是动态变化的。

为解决这一问题，将 DTW 算法与遗传算法结合，将 DTW 距离计算函数作为遗传算法的适应度函数 (fitness function, FF)，用于优化拟合行为体的偏好参数矩阵，最终达到校准 ABM 系统的目的。

DTW 算法是一种时间序列匹配算法，广泛应用于语音识别、手写识别和模式匹配等领域。其核心思想是寻找两个时间序列之间的最佳对齐路径，使得整体距离最小<sup>[17]</sup>。

DTW 算法的特点是可以将时间序列的某些片段进行伸长、缩短或扭曲来使两条时间序列尽量相互匹配，并根据匹配情况计算时序相似度。不同于传统的时间序列比较方法，DTW 可以弥补两个时间序列之间由于时间轴上的扭曲和变形所导致的不同步问题，能够在一定程度上解决 ABM 系统的时刻与真实数据的时刻不匹配的问题。

要计算两条时间序列的 DTW 距离，首先设推演输出的系统状态 (State) 序列为：

$$X = \{X_1, X_2, \dots, X_n\} \quad (10)$$

设历史场景的真实状态序列为：

$$Y = \{Y_1, Y_2, \dots, Y_n\} \quad (11)$$

其中， $n$  为序列  $X$  和  $Y$  的长度，即博弈推演的总时刻数； $X_i$  与  $Y_j$  分别表示在第  $i$  时刻推演结果的环境状态变量取值以及第  $j$  时刻真实环境状态序列的环境状态变量取值。

随后，构建一个  $n \times n$  的距离矩阵  $D$ ，其中  $D(i, j)$  表示序列  $X$  的前  $i$  个节点与序列  $Y$  的前  $j$  个节点之间的累计距离：

$$D = \begin{bmatrix} D(X_1, Y_1) & D(X_1, Y_2) & \dots & D(X_1, Y_n) \\ D(X_2, Y_1) & D(X_2, Y_2) & \dots & D(X_2, Y_n) \\ \vdots & \vdots & \ddots & \vdots \\ D(X_n, Y_1) & D(X_n, Y_2) & \dots & D(X_n, Y_n) \end{bmatrix} \quad (12)$$

借助动态规划填充距离矩阵，最终可以计算得到  $DTW = D(n, n)$ ，即历史序列  $X$  与真实序列  $Y$  之间的最短累计距离。

首先设定初始条件如下：

$$D(1,1)=d(X_1,Y_1) \quad (13)$$

$$D(i,1)=\sum_{k=1}^i d(X_k,Y_1) \quad (14)$$

$$D(1,j)=\sum_{k=1}^j d(X_1,Y_k) \quad (15)$$

使用动态规划方法填充距离矩阵  $D$ ：

$$\begin{aligned} D(i,\mathcal{Y}) &= d(X_i,\mathcal{Y}_j) + \\ \min(D(i-1,\mathcal{Y}), D(i,\mathcal{Y}-1), D(i-1,\mathcal{Y}-1)) \end{aligned} \quad (16)$$

其中， $d(X_i,Y_j)$  表示  $X_i$  与  $Y_j$  节点间的距离，距离函数  $d$  可以是欧几里得距离、曼哈顿距离或其他适合的度量方式。

由于环境状态变量取值较小，取值范围固定，方差较小，本文选择采用堪培拉距离 (Canberra Distance) 来衡量推演结果和历史场景的真实环境状态的差异，堪培拉距离的计算公式如下：

$$d(X_i,Y_j)=\frac{|X_i-Y_j|}{|X_i|+|Y_j|} \quad (17)$$

根据上述距离计算函数，DTW 距离取值为  $[0,|X|+|Y|]$ ，取值越小代表历史与真实情况的系统状态序列越相似。

涌现性 (Emergence) 是 ABM 系统的核心特征之一。涌现现象描述的是当多个独立的个体行为体相互作用时，通过局部的行为规则和互动产生出在宏观层面上不可直接预测的全局行为。这一方面意味着一个经过校准的 ABM 系统很可能能够捕捉系统的宏观状态变化，如多方博弈过程中冲突烈度的变化，但同时也意味着 ABM 系统生成的数据通常包含较多的噪声与随机性，尤其是在推演生成的事件序列与真实事件之间很难找到一一对应关系，导致推演出的状态序列与现实状态序列难以同步。

传统的时间序列比较方法很难处理此类缺乏直接对应关系的对齐问题，而 DTW 算法的核心特点恰恰在于它能够动态地拉伸、压缩时间序列的某些部分，以实现最优对齐，进而评估时间序列之间的差异，从而提高 ABM 系统的拟合度。

此外，相较传统 ABM 系统仅针对模型终止状态进行优化，本文通过将 DTW 距离作为遗传算法的适应度函数，把复杂的行为体偏好优化问

题转化为一个以  $D(n,n)$  为目标函数的最小化优化问题，将整个博弈进程的状态变化纳入行为拟合，使 ABM 系统在应对多样复杂的博弈场景时具有较高的稳定性与鲁棒性。

### 3.1.2 结合遗传算法的偏好矩阵参数拟合

遗传算法 (genetic algorithm, GA) 是模拟自然选择和遗传学机制的搜索算法。在多目标优化问题中，遗传算法往往可以找到全局最优或接近全局最优的解<sup>[18-19]</sup>。

遗传算法主要包含选择、交叉和变异这三个操作算子  $\varphi, \Gamma, \psi$ ，其主要组成要素包括：

$$GA = C, FF, P_0, M, \varphi, \Gamma, \psi, T \quad (18)$$

其中， $FF$  表示算法的适应度评价函数，即基于 DTW 算法的时间序列相似度计算； $P_0$  表示优化对象的初始值，每个初始偏好矩阵由取值为 0~10 的随机自然数组成； $M$  表示种群的大小，根据行为体的数量决定； $\varphi, \Gamma, \psi$  是三个操作算子参数，根据实际运行情况进行调整； $T$  是结束条件，本系统由于收敛性条件待验证，采用指定代数结束。 $C$  为算法的优化目标。本文所用遗传算法的伪代码如下：

算法 1 遗传-DTW 算法

Alg. 1 Genetic-dynamic time warping algorithm

输入：DTW,  $P_0, M, \varphi, \Gamma, \psi$

输出：C

1 设定参数 DTW,  $M, \varphi, \Gamma, \psi$ ;

2  $g = 1$ ;

3 While  $g \leq G$  do

4 使用  $\Gamma, \psi$  算子生成新偏好矩阵种群  $P_g$ ;

5 对每个偏好矩阵运行适应度函数 DTW;

6 使用  $\varphi$  算子选择下一代的父母;

7  $g = g + 1$ ;

8 End

9 返回种群  $P_T$  中适应度表现最好的偏好矩阵 C

## 3.2 基于遗传-DTW 算法的模型推理测试

为训练模型并进行模型测试评估，首先将数据区分为训练集与测试集。其中，测试集为某特定危机场景的历史数据，训练集为其余案例事件数据。在训练集上训练了状态转移模型与行为偏好模型，随后又使用“遗传-DTW”算法拟合了

各参与方的议题偏好矩阵。

拟合过程中设置遗传代数  $G$  为 10 代，种群的大小  $M$  为 80，父代的选择  $\phi$  采用轮盘赌选择 (Roulette Wheel Selection, RWS)，每代选择的父代数量为 12，保留 8 个父代用于下一代繁殖。交叉操作  $\Gamma$  使用双点交叉 (Two-point crossover)，变异操作类型为随机变异，变异操作  $\psi$  的概率设置为 15%，每次变异的参数值在  $[-3.0, 3.0]$  范围内随机变化。适应度函数  $FF$  通过重复运行 ABM 系统 300 次，计算模型输出的前 100 回合系统状态序列与历史数据的前 100 回合状态序列之间的 DTW 距离，取  $100 - DTW$  作为适应度，适应度越高越好。

状态转移模型、行为偏好模型与议题偏好模型训练完成后，在后 100 回合历史数据上进行了模型推理测试。与传统 ABM 模型在推理时被动观察涌现性结果不同，本文模型在推理过程中将继续在当前事件既有数据上进行校准，以将当前案例的行为体议题偏好纳入模型考量。

具体而言，一次完整的模型推理将首先重复运行 ABM 模拟 1000 次，分别计算每次模拟前 100 回合的状态序列与前 100 回合训练数据的真实状态序列之间的 DTW 距离，取距离最小的 1 次模拟，截取后 100 回合系统状态序列作为模型预测值，也即基于当前场景的已知博弈事件序列预测后续危机发展态势。

基于上述方法，对测试案例的系统状态序列进行了 10 次预测测试，绘制 DTW 最优对齐路径图如图 3 所示，图中每一条折线路径对应一次推理预测结果。路径代表预测序列与真实序列的对齐情况，路径越平滑，越接近对角线，表明序列之间的相似度越高。其中，红色线为基线模型，即使用前 100 回合系统状态值的平均值 4 (高烈度边谈边打) 作为后 100 回合的系统状态预测值，基线模型的 DTW 距离为 14.60。蓝色线为 10 次推理的模型预测值，即模型预测的后 100 回合系统状态序列，蓝色线的深浅代表 DTW 距离的大小，颜色越深，代表距离越小，模型表现越好。经过实验，模型预测值的平均 DTW 距离为 9.35，显著好于基线模型。

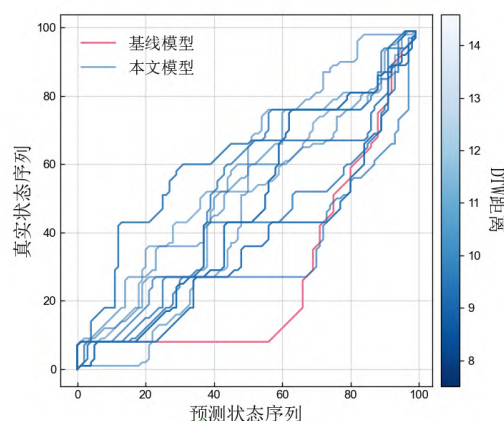


图 3 10 次模型推理的 DTW 最优对齐路径图

Fig. 3 DTW optimal warping path of ten model generations

图 4 所示是模型在某危机场景下的推理预测结果示意。可见，通过对微观决策的模拟，模型可以在宏观上推演得到较为合理的系统状态变化，模型在 0-40 时刻基本准确地描述了该危机进入第三年后开始“时谈时打”，系统状态在 2 至 4 之间来回波动。在 40-70 时刻捕捉到了一方对另一方的战术施压，进入长期“伴随谈判的高烈度冲突”，系统状态长期保持在 4。最终又在 70-100 时刻成功预测了危机某一方发动的最后一波攻势，推演出了冲突烈度在长期稳定后的再次急剧升级与降级，最终进入协议签署过程，回归“边谈边打”。

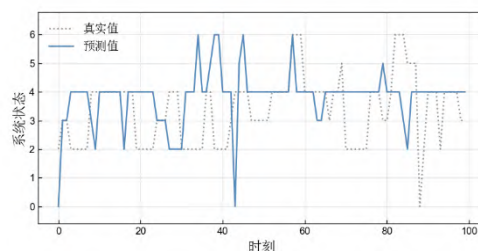
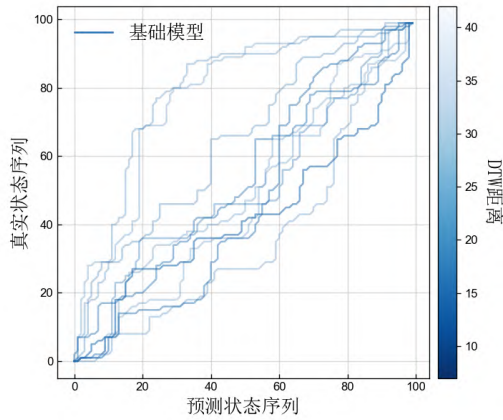


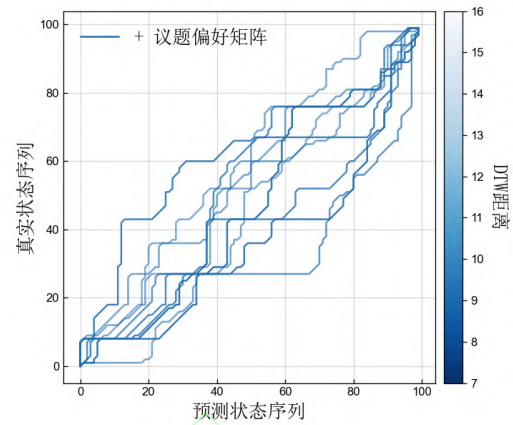
图 4 模型预测状态序列示意图

Fig. 4 Schematic diagram of the model predicted state sequence

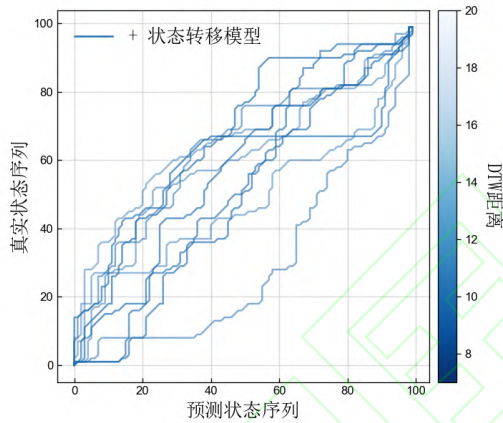
本文也对模型各组成部分进行了消融实验 (如图 5)。在基础模型中，每一回合系统状态将以相同概率转移至任一状态 (包括自身)，行为体以相同概率选择任意动作，议题偏好矩阵中的参数以均匀分布随机生成。



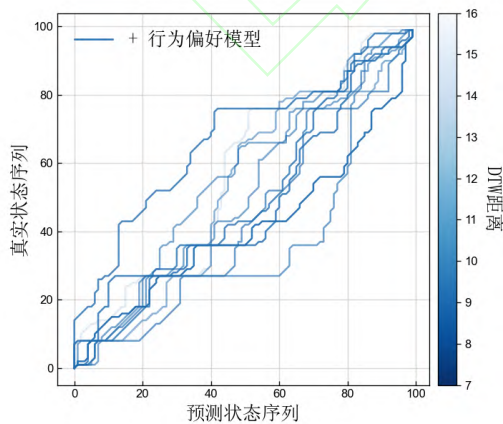
(a) 基础模型 DTW 最优对齐路径图  
(a) Optimal DTW alignments of the base model



(d) +议题偏好矩阵 DTW 最优对齐路径图  
(d) Optimal DTW alignments with issue preference matrix



(b) +状态转移模型 DTW 最优对齐路径图  
(b) Optimal DTW alignments with state transition model



(c) +行为偏好模型 DTW 最优对齐路径图  
(c) Optimal DTW alignments with behavior preference model

图 5 消融实验不同阶段下的模型表现对比图  
Fig. 5 Comparison of model performance at different stages in ablation experiments

随后，依次加入训练得到的状态转移模型、行为偏好模型与偏好矩阵模型重复进行实验，每一阶段进行 10 次模型推理，计算 DTW 距离的平均值，结果如下表所示：

表 7 模型消融实验结果

模型	10 次预测平均 DTW 距离
基础模型	29.84
+状态转移模型	12.40
+行为偏好模型	10.09
+议题偏好矩阵	9.35

#### 4. 结论

多行为体多回合的非常规博弈问题由于其参与方数量众多、行为策略复杂、缺乏共同知识等特点，是传统博弈论难以解决的研究问题，在国际关系学领域也基本停留在理论探索与案例比较研究。有部分研究尝试使用基于 ABM 的模拟方法验证博弈中常见的理论或规律，但目前为止仍没有研究真正实现基于现实数据训练校准计算模型，并对现实危机博弈事件进行推演复盘。

针对上述问题，本文的首要贡献在于首次实现基于现实数据训练和校准多方博弈 ABM 模型，并用于复杂多方危机博弈问题的推演和复盘。通

过构建统一的 MDP 框架收集历史数据, 结合 ABM 校准方法, 实现了数据驱动的 ABM 建模和危机博弈状态预测, 并在历史数据集上进行了模型效果检验。这一尝试在多方复杂博弈的量化推演研究方面填补了空白。

其次, 提出并应用了结合专家知识与机器学习的混合建模方法。在传统博弈论模型难以应对多行为体多回合非常规博弈的背景下, 结合了专家领域知识和数据驱动的机器学习技术, 尝试对行为体的偏好和行动策略进行建模与优化, 确保了模型的一定预测能力和现实适用性。这一创新为多方博弈问题, 尤其是多方危机博弈提供了新的研究路径。

最后, 基于 ABM 的推演模拟方法也为国际关系研究提供了崭新的研究视角与研究方向。在未来, 通过对推演生成的博弈路径(状态-行为序列)进行进一步频次统计、相关挖掘、聚类分析, 国际关系学者将能够更加深刻地理解怎样的物质与行为条件将使一方更可能选择退让, 又有怎样的条件将使一方更可能保持强硬。

本文通过构建基于行为体的预测模型来尝试回答这一问题。总体而言, 模型已经基本具备了对博弈态势变化的部分解释能力。但必须承认, 由于无法将时刻与现实时间相对应, 现有模型只能用于理解战略博弈的总体发展趋势, 且拟合得到的行为体偏好参数仍然难以直接解释行为体偏好与行为之间的因果联系。期待未来的研究在这一方面有所针对性的突破。

在后续研究中, 期待通过继续增加数据规模、优化改进现有模型, 实现对历史博弈路径的大规模数据挖掘, 进一步回答多方博弈中的行为体策略选择问题, 期待最终实现为国际冲突和平解决提供理论支持与决策参考。

此外, 还将在下一步工作中进一步拓展现有模型的应用场景, 将本文提出的 ABM 建模框架与行为体策略偏好拟合方法迁移应用至其他多方复杂博弈场景中, 如多边多轮谈判、多边经济制裁、区域冲突等。

## 参考文献

[1] POWELL R. Bargaining theory and international conflict [J]. *Annual Review of Political Science*, 2002, 5: 1-30.  
[2] YE M. Poliheuristic theory, bargaining, and crisis decision making[J]. *Foreign Policy Analysis*, 2007, 3(4):

317-344.

[3] MASAD D P. Agents in conflict: comparative agent-based modeling of international crises and conflicts [D] Fairfax: George Mason University, 2016.

[4] AXTELL R, SHAHEEN J A E. Agent-based models with qualitative data are thought experiments, not policy engines: a commentary on Lustick and Tetlock 2021[J]. *Futures & Foresight Science*, 2021, 3(2): e87.

[5] BREMER S A, MIHALKA M. Machiavelli in machina: or politics among hexagons [M]//Deutsch K W, Fritsch B, Jaguaribe H, et al. *Problems of world modeling: political and social implications*. Massachusetts: Ballinger Publishing Company, 1977: 303-337.

[6] MAGLIOCCA N R, MCSWEENEY K, SESNIE S E, et al. Modeling cocaine traffickers and counterdrug interdiction forces as a complex adaptive system[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(16): 7784-7792.

[7] ILACHINSKI A. Modelling insurgent and terrorist networks as self-organised complex adaptive systems[J]. *International Journal of Parallel, Emergent and Distributed Systems*, 2012, 27(1): 45-77.

[8] CUSACK T R, STOLL R J. Exploring realpolitik: probing international relations theory with computer simulation[M]. Boulder: L. Rienner Publishers, 1990.

[9] CEDERMAN L E. *Emergent actors in world politics: how states and nations develop and dissolve*[M]. Princeton, N.J. Princeton University Press, 1997.

[10] KARL S. RealLand: a wargaming approach to computational international relations [D]. Fairfax: George Mason University, 2017.

[11] HUA W Y, FAN L Z, LI L Y, et al. War and peace (WarAgent): large language model-based multi-agent simulation of world wars[EB/OL].(2023-11-28)[2024-03-14]. <https://arxiv.org/abs/2311.17227>

[12] 唐世平. 计算社会科学和科学决策的未来[J]. *国际政治科学*, 2023, 8(3): 1-27.

TANG S P. Computational social science and the future of scientific decision-making[J]. *Quarterly Journal of International Politics*, 2023, 8(3): 1-27. (in Chinese)

[13] AXTELL R L. Short-Term opportunities, medium-run bottlenecks, and long-time barriers to

progress in the evolution of an agent-based social science [M]// FRANK B A, BARTELS E M. Adaptive engagement for undergoverned spaces: concepts, challenges, and prospects for new approaches. Santa Monica, California: RAND Corporation, 2022: 469-505.

[14] VAN BEEK M, LOPATE M Z, GOODHART A, et al. Hierarchy and war[J]. American Journal of Political Science, 2025, 69(1): 299-313.

[15] CEDERMAN L E, GIRARDIN L. Computational approaches to conflict research from modeling and data to computational diplomacy[J]. Journal of Computational Science, 2023, 72: 102112.

[16] DYER J, CANNON P, FARMER J D, et al. Calibrating agent-based models to microdata with graph neural networks[EB/OL]. (2022-06-15)[2024-03-16].

<https://arxiv.org/abs/2206.07570>.

[17] Dynamic time warping[M]//Information Retrieval for Music and Motion. Berlin, Heidelberg: Springer, 2007: 69-84.

[18] DOMKE J. Generic methods for optimization-based modeling [C]//Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS), 2012: 318-326.

[19] Chapter 6 - Genetic Algorithms [M]//YANG X S. Nature-Inspired Optimization Algorithms. 2nd ed. [S.l.]: Academic Press, 2021: 91-100.