

# 论文写作指导

## (How to Write a Qualified Scientific Paper)

### 第五周：仿真或实验

王闻博 (wenbo\_wang@kust.edu.cn)

课件制作：杨春曦、王闻博

论文写作指导课题组  
机电工程学院  
昆明理工大学

2023 年 04 月 03 日

## Section 1

### 提纲

# 提纲

## 提纲

## 前课回顾

## 第五讲：仿真或实验

仿真简介

仿真和实验章节的写作

小结

## Section 2

### 前课回顾

# 课程目标

- **课程目标 (Goal) :** 学习学术期刊论文撰写的规范和方法
  - 掌握学术论文写作的基本常识和规范.
  - 学习学术论文的基本范式和写作技巧.
  - 学习相关写作辅助工具的使用方法.
  - 了解学术论文投稿流程及相关注意事项.
- 参考教材和资料
  - (美) 凯特·L·杜拉宾著, 雷蕾译, 《芝加哥大学论文写作指南 (第八版)》. 北京: 新华出版社, 2015, ISBN:978-7-5166-1599-7.
  - (美) 张俊东、杨亲正著, 《SCI 论文写作和发表: You Can Do It (第二版)》. 北京: 化学工业出版社, 2016, ISBN:978-7-122-26238-3.
- 线上同步
  - 腾讯会议号: 566-9282-5223.
  - 地址: <https://meeting.tencent.com/dm/Hjkr5Nig2309>.
- 任课教师联系方式: wenbo\_wang@kust.edu.cn

# 课程安排和考查形式

- 第一周：科技论文总体结构串讲.
- 第二周：论文标题、摘要的写作规范.
- 第三周：引言的写作.
- 第四周：从问题建模到理论结果阐述.
- **第五周：仿真或实验.**
- 第六周：结论、参考文献与投稿须知.
- 第七周或第八周：开卷考试.
  - 允许带除电子设备或有网络功能设备外的一切资料.

# 前课要点

- 严守学术道德与规范.
- 科研活动的创新模式.
- 一般的文献阅读方法.
- 学术论文的宏观框架结构.
- 学术论文标题撰写要点和模板.
- 学术论文摘要结构和模板.
- 引言的作用、结构和模板.
- 问题建模和理论结果阐述的写作原则.

## Section 3

### 第五讲：仿真或实验

## Subsection 1

### 仿真简介

# 前置问题 1：什么是（计算机）仿真？

## 计算机仿真（摘自：Simulation Modeling and Arena）

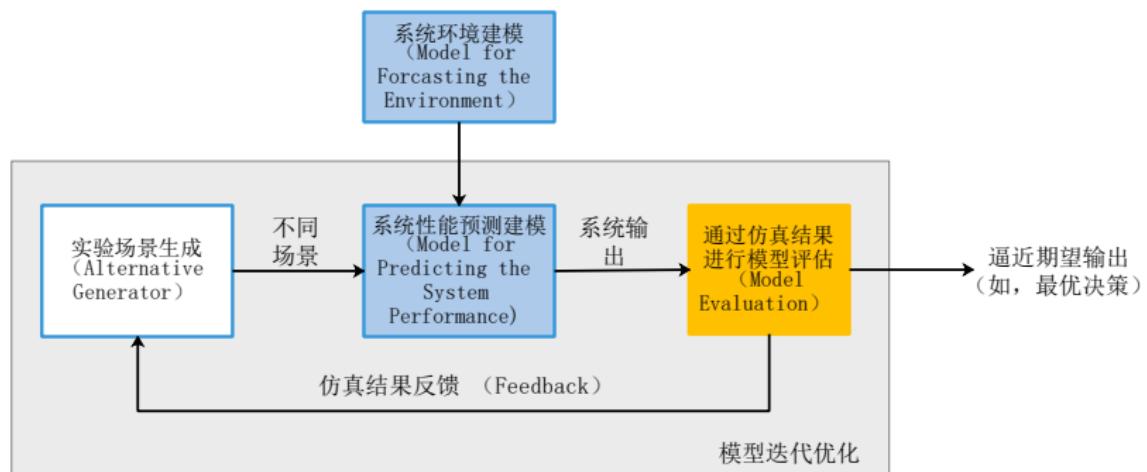
- Definition: A numerical technique for conducting experiments on a digital computer which involves logical and mathematical relationships that interact to describe the behavior of a system over time and space.
- 定义：“计算机仿真”是一种基于数字计算机的实验技术，以真实的物理系统的内在逻辑和数学关系为基础，描述系统随时间-空间变化的行为特征。

## 相比实物

- 仿真是对实物系统的抽象化（Abstraction）模型表达（Model Representation）。
- 计算机仿真使得系统的演进和交互过程能够被完全可控地模拟。
- 计算机仿真往往够提供无法纳入物理模拟的场景构建（如，飞行器模拟中的紧急/危险情况）。

# 仿真和系统建模/算法设计的关系

- 系统建模、理论分析和算法设计是仿真实验的基础.
- 仿真实验的目的在于通过描述性建模（Descriptive Modeling）展示系统的特性（Characteristics）和行为（Behaviors）**符合**之前章节中陈述的理论预测.
- “仿真实验-反馈-系统设计”的迭代过程是优化系统解决方案的必要途径.



# 计算机仿真的要件

- ① 基于系统分析和算法设计的**仿真模型**.
- ② 系统的主要控制因素 (Factors), 即**输入变量**.
- ③ 系统的响应 (Responses), 即**输出变量**.
- ④ 系统输入输出要素的**水平** (Levels of Factors).

## 计算机仿真的主要目的

- 确定 (或验证) 何种控制要素**显著决定**输出响应的水平.
- 验证或进一步解释理论分析得到的**期望结果**.
- 帮助理解系统运行中的**随机性特征**.
- 比较**所提出的算法**和一系列**参考算法**在给定环境中的性能指标优劣.

# 根据输入输出要素对仿真的分类

- 根据输出对输入的响应

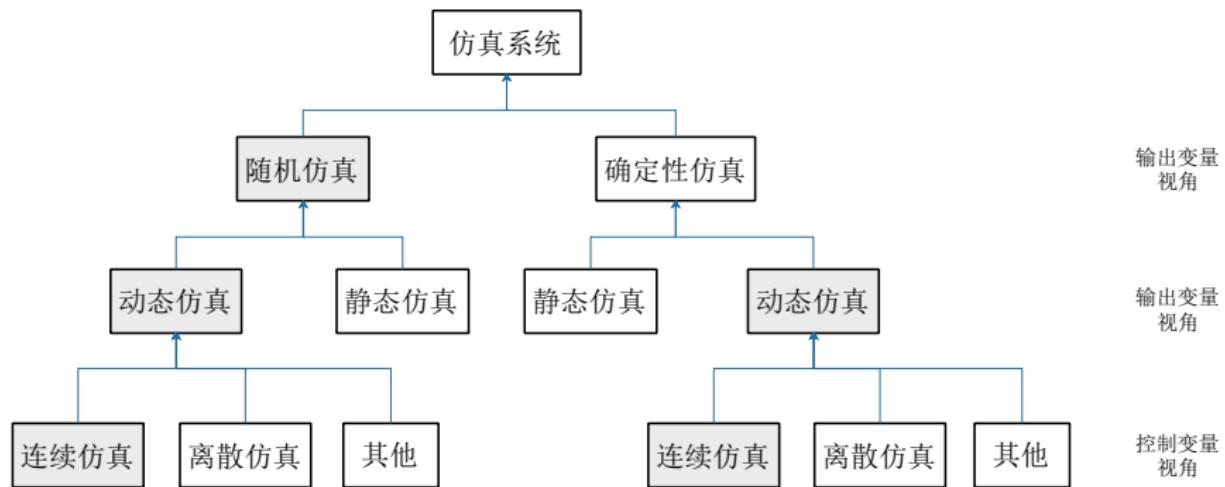
- 确定性仿真 (Deterministic Simulation) : 系统输入相同时, 总得到同样的输出 (如, 一般微分方程的解).
- 随机性仿真 (Stochastic Simulation) : 同样的输入下, 得到的系统输出由于系统的随机性总不同.
- 动态或静态仿真 (Dynamic or Static Simulation) : 系统响应是否随时间变化.

- 根据系统输入的性质

- 定量或定性仿真: 定性仿真中, 系统控制因素往往不以可衡量的数量形式出现 (如, 监督学习中 one-hot 编码的样本特征).
- 连续或离散 (Continuous or Discrete) 仿真: 连续控制变量 (如汽车的速度) 一般取有界实数值; 离散控制变量则取某些特定水平的数值 (如 0-255 表示的灰度图).
- 二元变量 (Binary or not) 仿真: 控制变量只有 0-1 两种水平/状态 (如, Gilbert-Elliott Markov 模型).

# 根据输入输出要素对仿真的分类（续）

## ● 仿真实验的分类

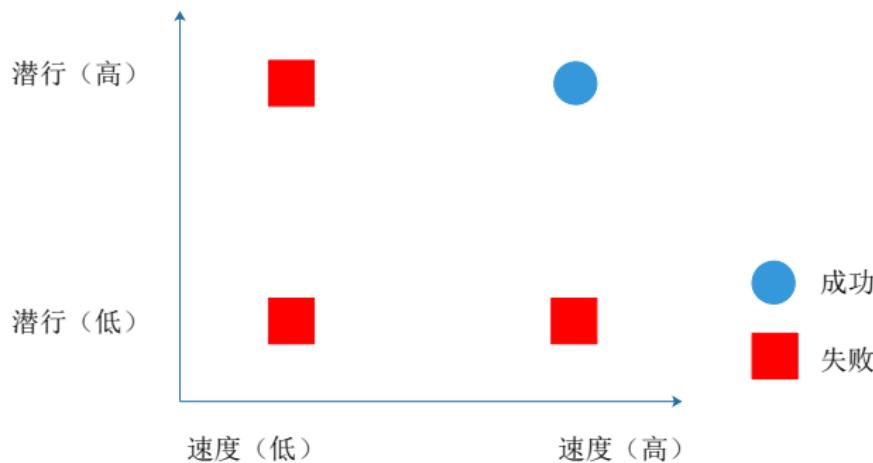


# 仿真实验的设计，以及输入输出要素的选取原则

- 慎重对待所有有关独立变量（Independent Variable）的假设
  - 在复杂仿真模型中，多个控制变量对系统的作用经常在统计学意义上是“混杂的（Cofounded）”。
  - 在缺乏关于多个变量对输出共同影响的先验知识的情况下，对“固定其他变量而研究单一变量对输出的作用”一事要谨慎——往往得到的结果是“不显著”。
  - 尽量选取“正交（Orthogonal）”的控制因素或控制因素族作为输入对象分别分析。

# 仿真实验的设计，以及输入输出要素的选取原则（续）

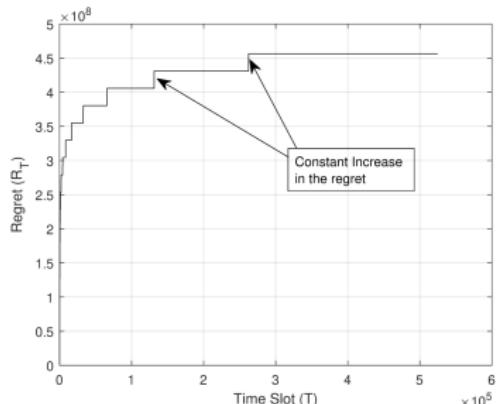
- 一个简单的双二元变量例子：“夺旗游戏”



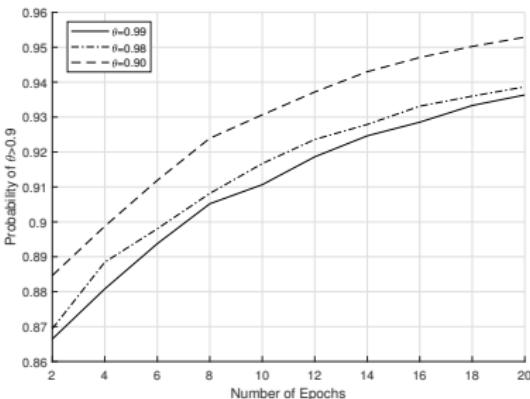
- 二维度输入：速度 (Speed)，潜行 (Stealth).
- 当固定一个维度变量的基准为“低”时，观察其他变量对输出的影响时，结果为“不显著”.

# 仿真实验的设计，以及输入输出要素的选取原则（续）

- 控制要素和输出要素的两种常见选择模式
  - 选择直接的“输入/输出变量（Input/Output Parameters）”。
  - 选择间接变量，如选择输入变量的“分布函数的决定参数（Distribution Parameters）”。
- 论文实例



(a)



(b)

**图：**(a) 选择直接的输入-输出变量：时间 vs. 算法效益（Regret）。(b) 输出选择间接变量：测量系统效率（即输出变量）的累积分布函数（CDF）。

## Subsection 2

### 仿真和实验章节的写作

# 总览：仿真和实验章节中所应包含的元素

- 仿真和实验环境的描述.
- 仿真和实验环境中基本参数的选择与表述.
- 仿真和实验步骤的表述.
- 仿真和实验结果的呈现：以图表形式为主.
- 仿真和实验结果的分析.

# 描述仿真和实验环境注意事项

- 仿真和实验环境的描述一般放在章节开始处，总领整章.
- 主要目的
  - 实验环境的描述在于保证实验的**可重复性和可验证性**.
- 需要涵盖的细节等级
  - 应包括主要环境（如仿真设备、程序库等）的细节，即**详细描述相关实验使用什么完成了什么**.
  - 在部分计算机相关研究领域，往往鼓励作者提供开源仿真代码、数据集等的网址（如 Github 地址）.
- 特殊试验（如涉及动物和人员参与的实验）还要做**道德申明**.

# 仿真和实验环境的描述举例 I

- 原文：K. Hong et al., “SpiderRadio: A Cognitive Radio Implementation Using IEEE 802.11 Components,” in IEEE Transactions on Mobile Computing, 2013.
- Section 7.1 Testbed Setup （侧重对实验环境/平台的描述）
  - (点明室内实验环境) Each node of the **indoor group** is a (仿真硬件和软件环境) **standard desktop PC running Linux 2.6 operating system**. They were all equipped with **Orinoco 802.11 a/b/g PCMCIA wireless card** (Fig. 15a). There is no PCMCIA slot for the desktop PC; (对硬件的特殊处理) so we make use of an **ENE-CB1410 PCMCIA to PCI adapter card** for converting the PCMCIA devices to operate on the desktop PC.
  - (点明室外实验环境) ...In the **outdoor group**, SpiderRadio is (仿真硬件和软件环境) **deployed on two laptops running Linux 2.6 operating system: Compaq NC4010 and Dell Inspiron 700m**. Both of them were equipped with **Orinoco 802.11 a/b/g PCMCIA wireless card**. (配置细节) The TX powers of these wireless devices were set to 100 mW. Another laptop running Windows VISTA and equipped with Wi-Spy 2.4x, acted as a monitor in the testbed. These Orinoco devices are equipped with **Atheros 5212 (802.11 a/b/g) chipsets..**

# 仿真和实验环境的描述举例 II

- 原文：V. Mnih et al., “Asynchronous Methods for Deep Reinforcement Learning,”（深度强化学习的经典论文）.
- Section 5 Experiments (侧重描述“完成了什么”)
  - (点明在什么平台上完成了什么) We use four different platforms for assessing the properties of the proposed framework. We perform most of our experiments using the [Arcade Learning Environment](#), which provides a simulator for Atari 2600 games. This is one of the most commonly used benchmark environments for RL algorithms. We use the Atari domain to compare against state of the art results, as well as to carry out a detailed stability and scalability analysis of the proposed methods. We performed further comparisons using [the TORCS 3D car racing simulator](#). We also use two additional domains to evaluate only the A3C algorithm —[MuJoCo](#) and [Labyrinth](#). MuJoCo is a physics simulator for evaluating agents on continuous motor control tasks with contact dynamics. Labyrinth is a new 3D environment where the agent must learn to find rewards in randomly generated mazes from a visual input. (受篇幅所限在补充材料中进一步给出细节) [The precise details of our experimental setup can be found in Supplementary Section 8.](#)

# 仿真和实验环境的描述举例 II (续)

- 原文：V. Mnih et al., “Asynchronous Methods for Deep Reinforcement Learning,” (深度强化学习的经典论文).
- Section 5 Experiments (侧重描述“完成了什么”)
  - (继续说明完成了什么) ... We then evaluated asynchronous advantage actor-critic on 57 Atari games... (说明使用了什么结构的神经网络) We trained both a feedforward agent with the same architecture as as well as a recurrent agent with an additional 256 LSTM cells after the final hidden layer... (说明了仿真的硬件环境和仿真时间) We trained our agents for four days using 16 CPU cores, while the other agents were trained for 8 to 10 days on Nvidia K40 GPUs.

# 仿真和实验环境的描述举例（中文 I）

- 原文：鞠默然等，“改进的 YOLO V3 算法及其在小目标检测中的应用,” 光学学报, 2019.
- 第四节. 实验结果与分析（侧重描述实验数据集和实验平台环境）
  - (描述所用数据集) ... 利用 VEDAI 航拍图像作为数据集, 对 YOLO V3 算法和改进的 YOLO V3 算法进行对比实验. (描述数据集的细节) VEDAI 数据集是将原始大视场卫星图像分割成  $1024 \text{ pixel} \times 1024\text{pixel}$  的图像, 包含各种各样的车辆, 背景和混淆对象. 数据集中包含可见光和红外两种图像. 所有的图像都采用了与地面相同的距离, 是典型应用于监视侦察的场景. 此外, 数据集中还包含了对  $1024 \text{ pixel} \times 1024\text{pixel}$  图像进行下采样得到的分辨率为  $512\text{pixel} \times 512\text{pixel}$  的图像, 为评估提供更小、更具挑战性的目标图像.VEDAI 数据集中平均每张图像包含约 5.5 个车辆目标, 并且它们只占图像总像素的 0.7%, 是典型的小目标检测数据集. 数据集中共有 9 类对象, 包括:airplane, boat, car, truck, tractor, camping, pickup, vans, others...  
● (描述软硬件环境) 实验条件为: 操作系统为 Ubuntu 14.04, 深度学习框架为 Darknet; CPU 为 i7-5930K, 内存为 64GB, GPU 为 NVIDIA GeForce GTX TITAN X .

# 仿真和实验环境的描述举例（中文 II，对比用）

- 例 1

- 本文使用 Windows10 64 位操作系统, 硬件环境为 Intel(R) Core(TM) i5-8400 CPU @ 2.80 GHz 处理器, 8 GB 内存, GPU 为 NVIDIA GTX 1060 Ti 6 GB 的笔记本电脑; 深度学习框架为 TensorFlow 1.8.0.

- 例 2

- 本次实验的电脑型号为神州战神 Z7-KP7EC, 内存为 16G, 处理器为 Intel core i7-8750H, 显卡为 NVIDIA GTX 1060, 系统为 Win 10. 使用 Darknet 深度学习框架, AlexeyAB 版本的 YOLO v3 模型, 并配置 CUDA 和 cuDNN 进行加速计算.

# 仿真和实验环境参数的描述

- 仿真和实验的环境参数 (System/Environment Settings) 不同于实验中的待研究输出参数 (Input/Output Factors under Investigation) .
- 环境参数一般指仿真和实验过程中一次性设定的被所有后续仿真采纳的公用参数设定集.
- 环境参数的表述应承接对仿真和实验环境的描述.
- 建议对环境参数的描述以文字叙述和表格混排为主.

# 仿真和实验环境参数的描述：英文论文举例 (表格文字混排)

- ... Throughout the simulation, the nodes of the network are placed inside a disk with a radius of 100m. External interferers are randomly placed on a ring with an inner radius of 100m and an outer radius of 200m. To create channel heterogeneity, we place a strong interferer, affecting half of the frequency bands in half of the inner disk. 20% of the remaining resource blocks are randomly interfered by external links... (列表描述主要仿真参数) [The main simulation parameters for the network environment is summarized in Table 1.](#)

表: Simulation parameters of the wireless environment.

Network parameter	Value
Carrier frequency	2GHz
Total bandwidth	40MHz
Sub-channel bandwidth	5MHz
Path loss exponent $\alpha$	4
No. of uniformly distributed tap delays	7
Rayleigh Variance	$10^{-2}$
Log-mean of shadowing	0
Log-variance of shadowing	0.01

# 仿真和实验环境参数的描述：中文论文举例

- 原文：杨志鹏等，“基于深度强化学习的四旋翼无人机航线跟随,”指挥与控制学报, 2022.
- (描述基础神经网络结构) 本实验构建的 A(ctor) 网络、C(critic) 网络及其目标网络为 BP 神经网络, 其中, 输入层一层、隐藏层一层, 设定隐藏层中神经元个数为 150. (描述实验对象统一设定参数) 实验中, 设定无人机每个电机的最大转速为 319.15rad/s, 无人机悬停转速为 225.67 rad/s. 在 DDPG 算法的应用中, 设立 var 表示探索变量. 当每次和环境交互时,A 网络会生成动作值 a, 系统再从以 a 为均值、var 为方差的正态分布中随机取出新的值, 有利于无人机学习未知动作. (描述神经网络统一训练参数和优化方法) 设定探索变量初始值为 3, 随着四旋翼无人机不断地训练学习,var 值进行指数衰减. 设定 A 网络和 C 网络学习速率分别为 0.001 和 0.002, 奖励折扣系数为 0.97. 当设定的经验池存满数据后, 神经网络中采用自适应矩估计 (adaptive moment estimation, Adam) 的梯度下降优化方法更新参数, 每次从经验池抽取的经验数量为 32.

# 仿真和实验环境参数的描述：其他例子

- 原文：X. Lu et al., “Video Object Segmentation with Episodic Graph Memory Networks,” ECCV 2020.
- (介绍数据预处理过程) During pre-training, we randomly crop  $384 \times 384$  patches from static image samples, and the video clip length is  $N+1=4$ . During the main training, we crop  $384 \times 640$  patches from real training videos. We randomly sample four temporal ordered frames from the same video with a maximum skip of 25 frames to build a training clip. Data augmentation techniques, like rotation, flip and saturation, are adopted to increase the data diversity. For O-VOS, we select a saliency dataset, MSRA10K and semantic segmentation dataset, COCO, to synthesize videos. (介绍训练集) We use all the training videos in DAVIS<sub>17</sub> and Youtube-VOS for the main training. For Z-VOS, we use two image saliency datasets, MSRA10K and DUT, to generate simulated videos. These two datasets are selected following conventions for fair comparison. After that, we take advantage of the training set of DAVIS<sub>16</sub> to finetune the network.

# 仿真和实验步骤的表述要点

- **注意：**仿真和实验结果的说服力，很大程度上取决于实验步骤表述的逻辑性.
- 要点 1: 完整描述仿真和实验输入因素和待观察的输出因素/指标.
- 要点 2: 详细描述仿真和实验过程中的其他关键性的一致被控（不变）因素以及相同环境设置.
- 要点 3: 详细描述用作系统性能比较基准（Baseline）的参考算法、同本文所提出算法的比较结果，并做综合分析.
- 要点 4: 基于输入-输出要素做分组实验设计.

# 仿真和实验步骤的表述模板

- 描述实验方法、输入输出因素（观测指标）、对照方法设置、及数据处理方法。
  - ① **描述对照算法**: 对本文提出的方法给出对应的章节/伪代码/公式索引；对他人提出的方法给出参考文献.
  - ② **描述实验分组设计及对应的对照算法设置**: 说明分组方法、对应的分组逻辑及设置何种对照方法等.
  - ③ **描述（分组对应的）观测指标**: 说明输入输出指标名称、物理意义(实体实验中需说明使用的观测方法、仪器等).
  - ④ **描述数据处理和统计方法**.

# 仿真和实验步骤的表述要点：论文举例

- 对基准参考算法的描述：(原文) W. Wang and A. Leshem, “Non-Convex Generalized Nash Games for Energy Efficient Power Allocation and Beamforming in mmWave Networks”, in IEEE Transactions on Signal Processing.
  - ..., we introduce **three baseline algorithms** for joint power allocation and beamforming performance comparison with the proposed Algorithm 2. The first baseline algorithm, “**Global Iterative Zero Forcing (ZF)**”, is adapted from [10] to iteratively enforce both the Tx and the Rx filters of each link to be in the null space of the equivalent interfering channels with a modified SINR constraint<sup>1</sup>. The second baseline algorithm, “**Global Coordinated Tx-BeamForming (BF)**”, is adopted from [11] as a centralized Tx-beamforming procedure for sum-of-power minimization with SINR constraints through full link coordination using fixed Rx beamforming vectors. Likewise, it is assumed that the CSI of all interfering links is known to each transmitter. The third algorithm is a global iterative Rx-Tx coordinated MSE beamformer, “**Global Coordinated-MSE Tx-Rx BF**” where the transmit beamformers use MSE beamformers based on full CSI following (38).

---

<sup>1</sup>Exchanges CSI among links are assumed for global iterative ZF. Instead of integrating ZF into the MMSE-based Rx-beamforming procedure [10] the adopted scheme uses stand-alone ZF for both Tx- and Rx-beamforming.

# 仿真和实验步骤的表述要点：论文举例（续）

- 以输入变量为分组依据设计实验分组：(原文) W. Wang and A. Leshem, “Non-Convex Generalized Nash Games for Energy Efficient Power Allocation and Beamforming in mmWave Networks”, in IEEE Transactions on Signal Processing.
  - 第一组输入为“number of links”: Fig.3 (原图略) Performance comparison for different numbers of links through Monte Carlo simulations at an SINR threshold of 20dB for all the links. (a) Sum of supply power vs. numbers of links for different algorithms. (b) Sum of spectrum efficiency vs. numbers of links for different algorithms.
  - 第二组输入为“SINR thresholds”: Fig. 5. Monte-Carlo simulations for number of iterations to convergence vs. SINR thresholds... Fig. 6. Monte-Carlo simulations for total supply power vs. SINR thresholds. The proposed scheme with three different Tx-BF methods and the Global Coordinated Tx-BF scheme achieve the same level of energy efficiency.

# 图表的使用和规范化呈现方式

## 图表的作用

- 对系统指标的观测结果做定量化的分析与比较.
- 结合基于文字的分析，为前文提出的理论结果的正确性或提出的算法可靠性、效率等提供可视化证据.
- 省略文字，紧缩篇幅.

数据可视化工具：

- Matlab.
- Python (如 gnuplot 包).
- Julia、R 等.
- 不推荐使用：中低级语言如 C/C++ 等.

# 图示的使用和规范化呈现方式

## 图示的绘制原则

- 自明性 (Self-containedness)：读者只看图题 (Caption)、图面 (Figure) 和图例 (Legend)，就可基本无误地理解图意，而无需阅读正文。
- 写实性 (Truthfulness)：严格地忠实于所分析因素的原始数据和本来面貌。
- 清晰性 (Clearness)：对相关输入输出因素的信息都要清晰完整地呈现出来，符号、数字单位等在图中或图题中有明确说明。

次要原则：

- 不因追求美观而将图形做得过于花哨。
- 坐标轴的说明应清楚，量和单位缺一不可。
- 使用不同颜色、线型等区分不同的输入-输出曲线，一幅图中的曲线数量不宜过多。

# 图示的分类与举例

## 图示的分类 |

- 柱状图、直方图 (Histograms) : (原文) W. Wang, et al., "Decentralized Caching for Content Delivery Based on Blockchain: A Game Theoretic Perspective," IEEE ICC 2018.

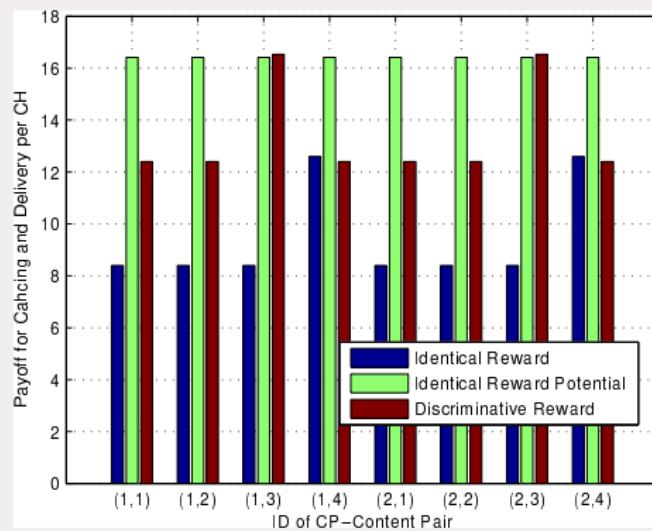


图: Payoff comparison of action groups with different rewarding schemes.



# 图示的分类与举例（续）

## 图示的分类 II

- 线(点)图 (Dotted Lines, 注意不同的线型) : (原文) W. Wang and A. Leshem, "Non-Convex Generalized Nash Games for Energy Efficient Power Allocation and Beamforming in mmWave Networks", in IEEE Transactions on Signal Processing.

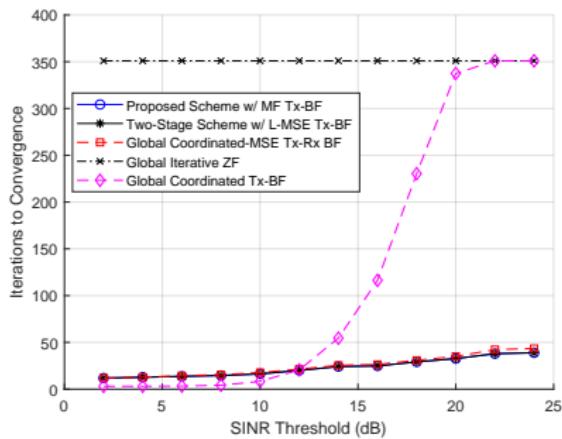
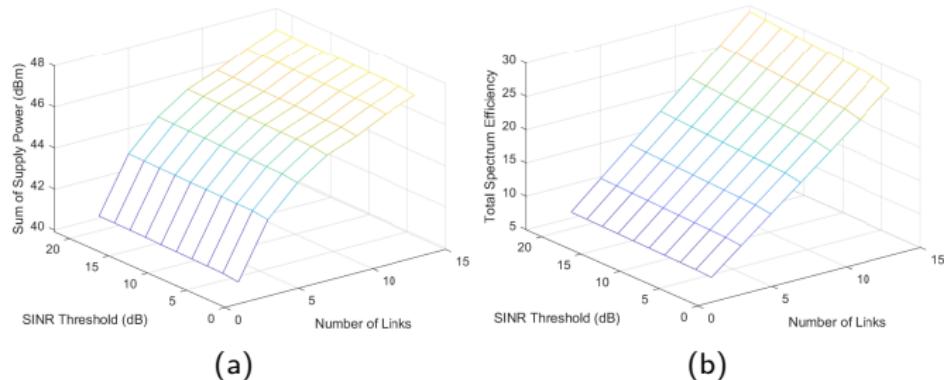


图: Monte-Carlo simulations for number of iterations to convergence vs. SINR thresholds.

# 图示的分类与举例（续）

## 图示的分类 III

- 三维图（3D Graphs, 注意不同坐标轴名称）：(原文) W. Wang and A. Leshem, “Non-Convex Generalized Nash Games for Energy Efficient Power Allocation and Beamforming in mmWave Networks”, in IEEE Transactions on Signal Processing.



**图：**Monte Carlo simulation results for Algorithm 2 using matched filters for Tx-beamforming. (a) Sum of supply power vs. different numbers of links and SINR thresholds. (b) Sum of spectrum efficiency vs. different link numbers and SINR thresholds.

# 图示的分类与举例（续）

## 图示的分类 IV

- 向量场图 (Vector Field) : (原文) X. Liu et al., "Evolutionary game for mining pool selection in blockchain networks." in IEEE Wireless Communications Letters.

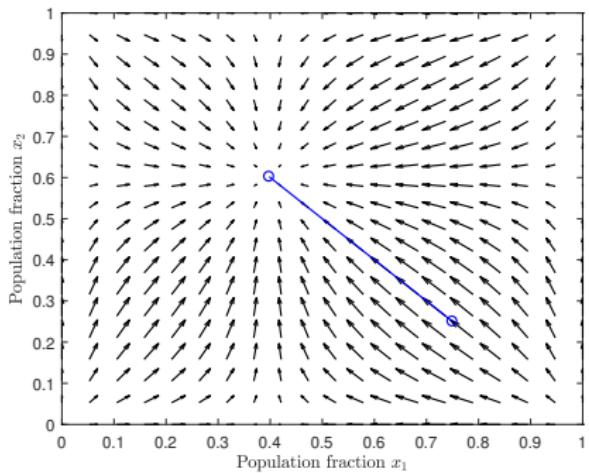
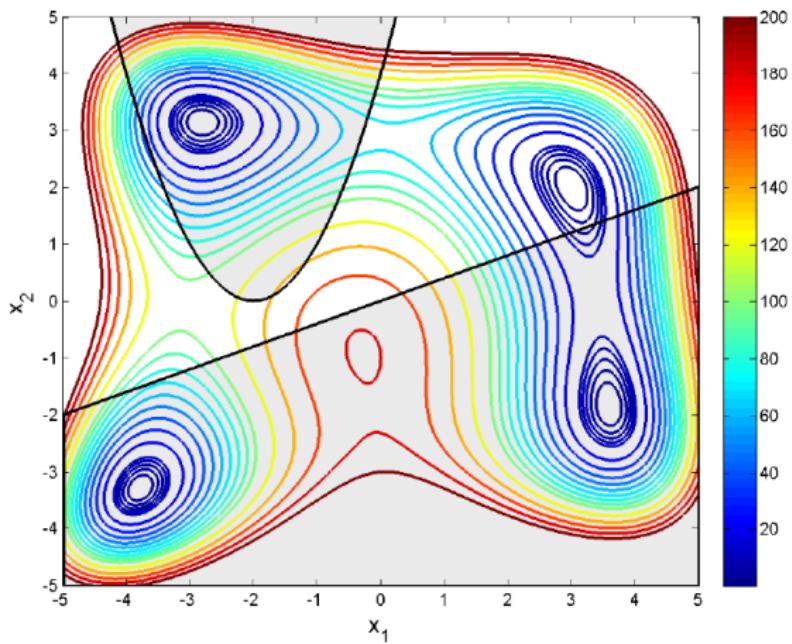


图: Replicator dynamics of the pool-selection strategies and the evolution trajectory starting from  $\mathbf{x}(0) = (0.75, 0.25)$ .

# 图示的分类与举例（续）

## 图示的分类：V

- 等高线图 (Contour): 常见于优化算法设计实验.



# 图示的分类与举例（续）

## 图示的分类：其他

- 饼图 (Pie): 示例略.
- 实验结果截图 (Screenshot) 或照片 (Photo): 一般为非量化的仿真、实验结果示意和对比图.
- 示例原文: J. Schulman, et al., "Proximal Policy Optimization Algorithms," 强化学习领域经典论文.

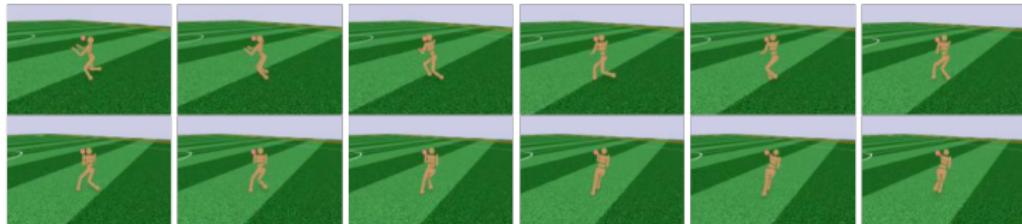


图: Still frames of the policy learned from [RoboschoolHumanoidFlagrun](#). In the first six frames, the robot runs towards a target. Then the position is randomly changed, and the robot turns and runs toward the new target.

# 图示的分类与举例（续）

## 图示的分类：其他（续）

- 照片 (Photo): 显微照片.



Fig. 16. A cytoplasmic lobe (arrowhead) forms from the caudal spermatid head and mitochondria sheath (arrow) and extends toward the lumen of the seminiferous tubules. Scale bar = 1  $\mu$ m.

# 表格的使用和设计要点

## 表格的设计原则

- 自明性、写实性、清晰性.
- 突出要点：表格要有明确的表达目的性.
- 简洁性：只列出重要的现象、算式、参数、数据等，排除分析或计算的中间步骤或环节.

次要原则：

- 数据表格的表身应尽可能采用三线表：表头、项目、数据.

# 表格设计举例 I

## ● 三线表

表：不同增强方法训练的 YOLO v3 模型精度、漏检率对比。

增强方法	mAP%	TP	FP	FN	漏检率%
原图	88.51	376	50	59	13.6%
暗通道去雾算法	89.91	429	93	47	9.9%
多尺度 Retinex 算法	90.63	438	74	59	11.9%
自适应对比度增强算法	93.92	409	72	26	6%

# 表格设计举例 II

- Latex 下编制的复杂表：(原文) I. Achituve, et al., “Communication Efficient Distributed Learning over Wireless Channels,” arXiv:2209.01682, 2022.

**表：**Test accuracy ( $\pm$  std) with 4 and 9 works on CIFAR-10 and CIFAR-100, respectively. The numbers in bold are the statistically significant best results.

No. of clients	CIFAR-10		CIFAR-100	
	4	9	4	9
Concat Workers Embed	84.35 $\pm$ 0.5	81.02 $\pm$ 0.1	53.06 $\pm$ 0.2	51.05 $\pm$ 0.5
Best Worker Pred	59.31 $\pm$ 0.6	44.87 $\pm$ 1.3	18.67 $\pm$ 1.4	7.97 $\pm$ 1.6
Avg. Workers Preds	76.03 $\pm$ 0.6	61.60 $\pm$ 0.6	31.14 $\pm$ 0.7	16.89 $\pm$ 1.5
Avg. Workers Embed	84.28 $\pm$ 0.2	80.58 $\pm$ 0.5	50.80 $\pm$ 0.3	45.78 $\pm$ 0.3
FedOCS	84.48 $\pm$ 0.7	80.52 $\pm$ 0.3	<b>53.98 <math>\pm</math> 0.7</b>	<b>49.79 <math>\pm</math> 0.2</b>

# 仿真和实验结果分析

## 仿真和实验结果分析的基本原则

- 理论上讲，仿真和实验结果分析是论文中最有创造性见解、对逻辑性要求最严格的部分（之一）。
- 以论文中的前述理论结果为基础，对以图表为载体的方针和实验结果进行理论分析和综合。
- 通过逻辑推理和理论分析，从结果中提出科学结论（一般为对本文提出的理论结果或算法有效性的验证）。

在结果分析中应着重回答的问题：

- 为什么出现这样的结果？
- 出现这样的结果意味着什么？

# 仿真和实验结果分析的分节/分段讨论

- 对同一组实验的图表展示和文字分析共同组成一个完整的结果分析段落或小节.
- **根据输入-输出（观测）指标分节（段）**：适用于对同一研究对象施以相同处理因素的分析情况.
- **根据不同处理因素分节（段）**：适用于对比几种处理因素实验结果的分析情况.
- **根据不同观察内容分节（段）**：适用于研究或观察某研究对象不同方面特点的分析情况.

# 材料组织的重点：通过仿真实验结果呈现本文理论结果的正确性或提出算法的优越性

- 对同一组仿真或实验结果的图表展示和文字分析**共同**组成一个完整的结果分析段落或小节.
- “正确性”来源于对本次实验或观察结果做出的**理论解释和讨论**.
- “优越性”来源于仿真中本文算法和其他算法的**性能比较**.
- 针对非仿真的实验结果（如生物化学领域的实验）的分析，应突出新发现现象和其原因的分析，如将本次结果，与过去及其他研究结果（如，不同时间、不同地点、相同或不同的研究对象中的研究结果）相比较，分析异同，解释产生差别的可能原因.

# 仿真实验结果分析举例 |

## 算法性能对比

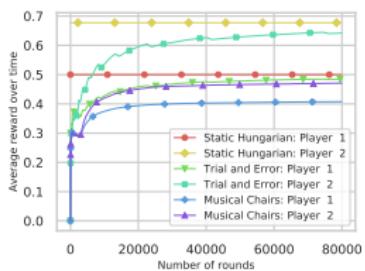
- (原文) W. Wang et al, “Decentralized Learning for Channel Allocation in IoT Networks over Unlicensed Bandwidth as a Contextual Multi-player Multi-armed Bandit Game,” in IEEE Trans. Wireless Comm.
- Figure 2a provides an intuitive illustration on the evolution of the players' average rewards as the bandit algorithms progress over time. In Figure 2a the curves marked as “Static Hungarian” indicate the expected rewards of each individual players (Figure 2(a)) when the true social-optimal allocation is adopted. Figure 2b shows the comparison of the average regret over different time horizon for the proposed trial-and-error learning algorithm, the GoT algorithm, the MC algorithm and the SIC-MMAB algorithm. In this experiment, for the same sequence of context-arm-value sampling with a total rounds of  $T$ , the bandit game is repeatedly played for 200 times with each algorithm. The shaded areas around the solid curves indicates the empirical performance variation during the Monte Carlo simulations. (分析异同) The gap between the regrets of trial-and-error learning and the other algorithms clearly indicates that the proposed algorithm is able to better adapt to the stochastic evolution of the contextual dimension.



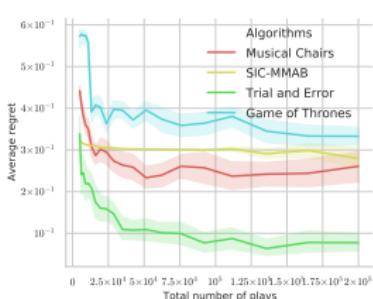
# 仿真实验结果分析举例 I (续)

## 算法性能对比 (续)

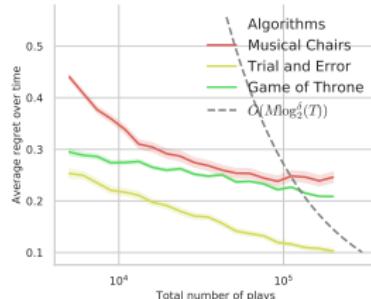
- (来源) W. Wang et al, "Decentralized Learning for Channel Allocation in IoT Networks over Unlicensed Bandwidth as a Contextual Multi-player Multi-armed Bandit Game," in IEEE Trans. Wireless Comm.



(a)



(b)



(c)

图: (a) Average payoff of individual players with respect to time. (b) Average regrets of plays vs. different time horizons  $T$ . (c) Regret evolution with respect to the total number of plays  $T$ .

# 仿真实验结果分析举例 I (续)

## 算法性能对比 (接上页)

- (原文) W. Wang et al, “Decentralized Learning for Channel Allocation in IoT Networks over Unlicensed Bandwidth as a Contextual Multi-player Multi-armed Bandit Game,” in IEEE Trans. Wireless Comm.
- (同理论上界和其他算法对比) In Figure 2c, we compare the evolution of the average regrets of trial-and-error learning, MC and the GoT algorithm [39] in a slightly larger problem with 5 arms, as the total number of plays (horizons)  $T$  increases. The dashed curve “ $O(M\log_2^\delta(T))$ ” represents a heuristic regret bound in the same form as (48), where the heuristic bound has a set of parameters as  $c_1 M = 200$ ,  $c_2 M = 40$  and  $\bar{C}_1 = 0$  for the considered game. Apart from the similar finding that the two learning-based algorithms achieves a better performance than MC, it is worth noting from the simulation result that the proposed trial-and-error learning algorithm has a much faster convergence rate than GoT.

# 仿真实验结果分析举例 II

## 算法性能对比

- 原文不详.
- 实验结果表明, 以扩张率  $1, 1, 1+1, 1, 1$  为基准模型结果, 此时由于扩张率都为 1, 不使用空洞卷积. 当扩张率为  $1, 1, 1+2, 2, 2$  时, 出现相同扩张率叠加态, 此时验证精度为 53.17%, 相比于基准模型结果, 提升了 0.81%. 当扩张率为  $2, 2, 2+3, 3, 3$  或更大时, 验证精度保持不变或下降. 此时基准模型的扩张率为  $1, 1, 1+1, 1, 1$ , 不使用空洞卷积, 精度为 53.67%; 使用相同扩张率叠加态之后, 验证精度分别提升了 0.74%, 0.63%, 1.13% 和 0.92%.

# 仿真实验结果分析举例：其他

- 原文：X. Lu et al., “Video Object Segmentation with Episodic Graph Memory Networks,” ECCV 2020.
- (通过对比论证提出算法优越性) ...Our approach obtains a final score of 80.2%, significantly outperforming state-of-the-arts. Compared to memory-based method S2S[64], our model achieves much higher performance (i.e., 80.2% vs 64.4%), which verifies the effectiveness of our external graph memory design. Moreover, our method performs favorably on both seen and unseen categories. Overall, our method achieves huge performance promotion over time-consuming online learning base methods without invoking online fine tuning. This demonstrates the efficacy of our core idea of formulating O-VOS as a procedure of learning to update segmentation network.

## Subsection 3

### 小结

# “仿真和实验”写作要点小结

- “仿真和实验”承接上文，为理论分析或所提出方法的验证提供必要证据.
- 仿真和实验的设计要求逻辑清晰、简洁.
- 行文要求：清晰（仿真/实验过程逻辑清楚严密）、简洁、有说服力.
- 仿真实验的可重复性和实验结果的可复现性是“仿真和实验”写作的第一要点.
- 注意仿真实验的设计（尤其是输入输出因素的选择）的合理性、代表性，以及行文中分段、分节的逻辑性.
- 仿真和实验结果的呈现应具有直观性，实验结果多以图、表形式呈现，力求实验过程和结果的表述逻辑清晰，结果一目了然.

# 结束页

## 开放式讨论

### 问题与讨论