



**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

# ICPC Educate

Monday, September 1, 2025





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## ICPC Educate

- **ICPC Journal – The Journal of Competitive Learning** (Erich Baker)
- **Use of AI in Programming Competitions** (Arup Guha)
- **Using sports analogies, metaphors and quotes in coaching competitive programmers and teams** (Jerzy Jaromczyk)
- **Computers at War** (Bob Roggio)
- **User Simulation for Agent Evaluation** (Zhou Yu)





**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

# ICPC JOURNAL OF COMPETITIVE LEARNING

<http://journals.u.icpc.global/jcl>





**ICPC 2025**

WORLD FINALS BAKU

HOSTED BY ADA UNIVERSITY

## We have a passion for CL!

Characterized by the integration of challenge-based activities and contests, CL demonstrates particular efficacy in areas such as computer science and cybersecurity.

Competitions like ICPC are pivotal in cultivating technical skills and fostering creativity and inherently thrive on problem-solving, critical thinking, and rapid innovation.



ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Why do we need our own Journal?

- **Original Research**
  - Algorithmic problem solving, solutions testing, distributed software.
- **Academic**
  - It will serve as a peer-reviewed archive for novel designs, solutions, findings produced in the course of competitive pedagogy.
- **Collaboration**
  - Journal publications allow passionate educators and researchers to share their experiences and novel research with a like-minded community.
- **Repository**
  - JCL will archive data, problem sets, and solutions, and for educators, coaches, and the next generation of researchers.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## JCL Audience

Just within ICPC there are over 3450 coaches and teams, +75,000 volunteers and +400000 participants that would benefit from archiving our collective knowledge.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## JCL Audience

### MATHEMATICS

International Mathematics  
Modeling Challenge,  
International Mathematics  
Olympiad & Competitions

### CYBERSECURITY

Cyberpatriot, CSAW CTF,  
Panoply, NCDCC - National  
Collegiate Cyber Defense  
Competition

### BUSINESS

Blue Ocean, Wharton Global,  
Global Youth  
Entrepreneurship, GENIUS  
Olympiad Business

### CS

International Collegiate  
Programming Contest,  
International Olympiad of  
Informatics, Imagine Cup

### ROBOTICS

World Robot Olympiad, VEX  
Robot Competition, BotBall,  
RoboCup

### DEBATE

World Universities Debating  
Championship, COSMOS  
Global Debate on Demand,  
IPDC





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## JOURNAL SECTIONS

### RESEARCH

Accepting full length and brief manuscripts, JCL will publish novel research in areas related to research in pedagogy design, empirical studies linking experience and educational design, unique preparation strategies, and other quantitative sociological, psychological, and educational phenomena related to competitive learning.

### REVIEWS

Review articles focused on synthesizing or collating materials related to competitive learning will be considered.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## JOURNAL SECTIONS

### CURRICULUM

Vetting curriculum enhancement is an important piece for pushing the boundaries of competitive learning.

### EDITORIALS

Insights from experts and diverse perspectives on competitive learning.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## JOURNAL SECTIONS

### PROBLEM SETS & SOLUTIONS

JCL will feature problem sets from domains associated with competitive learning, including vetted problem set solutions. Problem sets will provide practical resources for educators, and allow them to implement competitive learning strategies.

### DATA

Curated data sets are a critical piece for creating shared approaches for developing competitive learning experiences. This includes datasets for research and analytical tools for educators and practitioners.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## JCL Logistics

<http://journals.u.icpc.global/jcl>

- Publication fees supported by ICPC
- Online & Open Source
- DOIs for Data Sets
- Indexed by Google Scholar
  - In process: DOAJ, Scopus, Web of Science

### Programming Contests, Competitive Programming Curricula, and how Everybody Benefits

Arup Guha  
Department of Computer Science  
University of Central Florida  
Orlando, Florida  
dmarino@ucf.edu

Jerzy W. Jaromczyk  
Department of Computer Science  
University of Kentucky  
Lexington, Kentucky  
jurek@g.uky.edu

Christian Yongwhan Lim  
Department of Computer Science  
Columbia University  
New York, New York  
yongwhan.lim@columbia.edu

Benjamin Reed  
Department of Computer Science  
San Jose State University  
San Jose, California  
ben.reed@sjsu.edu

Etienne Vouga  
Department of Computer Science  
University of Texas - Austin  
Austin, Texas  
paul.vouga@austin.utexas.edu

*Abstract*—The panelists, all of them competitive programming team coaches and college educators, share their experiences in teaching and working with competitive programming students and teams. The panel uses as the model the International Collegiate Programming Competition (ICPC) but additionally will describe other formats and resources including activities for K-12. We compare various competitive programming courses and different ways of incorporating components of computing courses into them. We discuss the payoffs for participating students of varying backgrounds and interests, aptitudes, and

That perspective also guides the design of materials and pedagogies for courses, practices, or extracurricular activities: by optimizing the desired payoffs, instructors ensure that every student can make capital out of participation. The payoffs for students include:

- Strengthening algorithmic problem-solving;
- Reinforcing program development techniques with testing and disciplined coding that produces correct and efficient





# ICPC 2025

WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

2025

**November**

Journal kick off. First issue published.

2026

**May**

The second issue is scheduled to be published in early 2026.

2026

**November**

Submissions will be accepted on a rolling basis for a third edition.

2027

**??**

JCL has an opportunity to shift to a quarterly publication.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

**JCL** | The Journal of Competitive Learning

CURRENT ARCHIVES ANNOUNCEMENTS ABOUT - Register Login

**Current Issue**  
Vol. 1 No. 1 (2025)  
Published May 21, 2025

**Special Issue**

We are excited to announce the publication of **Volume 1, Issue 1 of the *Journal of Competitive Learning (JCL)***, a new peer-reviewed journal dedicated to exploring the role of competition as a dynamic strategy for engaging students in learning. While rooted in the fields of computer science and technology—where competitive approaches such as coding contests, hackathons, and gamified assessments are widely used—JCL welcomes interdisciplinary contributions that examine competitive learning across diverse educational contexts. This inaugural issue showcases innovative research

[More](#)





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Editorial Team

- **Erich Baker:** Editor-in-Chief
- **Christian Lim:** Senior Editor
- **Bill Booth:** Senior Editor
- **Tomas Cerny:** Senior Editor
- **Bill Poucher:** Senior Editor

**[erich.baker@icpc.global](mailto:erich.baker@icpc.global)**





**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

# Use of AI in Programming Competitions

Wednesday, September 3, 2025  
by Arup Guha (dmarino@ucf.edu)  
University of Central Florida





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

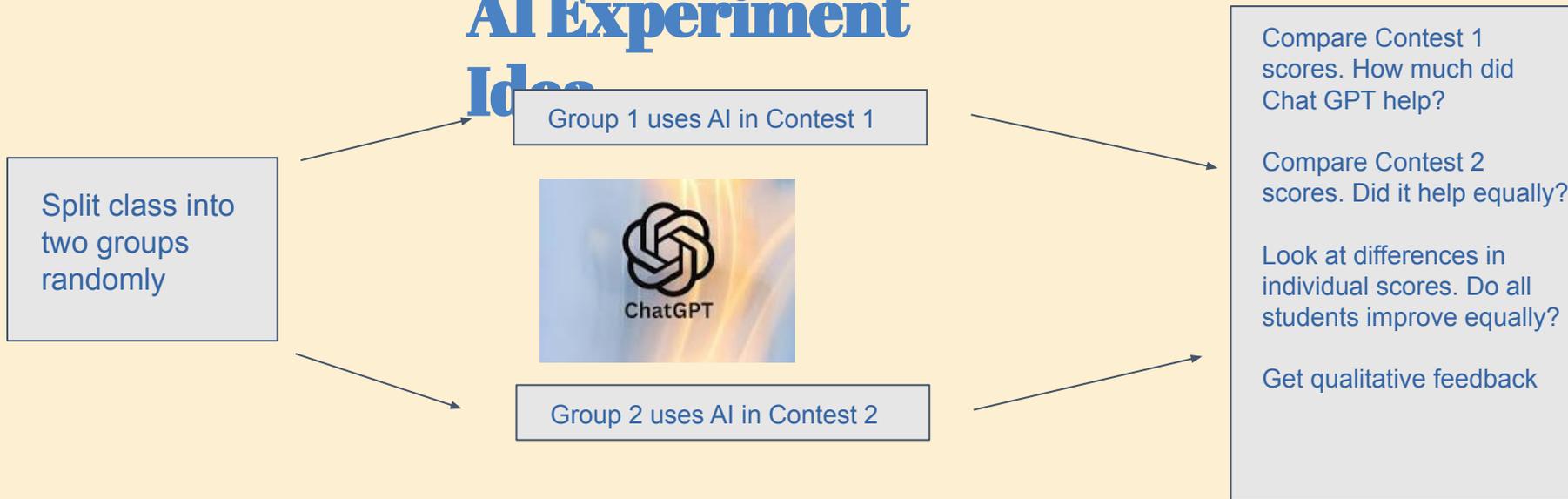
31 August – 5 September 2025 // hosted by ADA University

## COP 4516: Team Dynamics Course

- Course for students not on the programming team
- First seven Fridays: Three-hour individual contests with four questions each.
- Usually one question is “easy” requiring no special knowledge
- Two questions are usually based on the week’s lecture topic
- One question is random usually based on an old lecture topic.



## AI Experiment Idea





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Group Selection

### Registrar's Office

- Students Randomly placed into two different rooms due to space limitations.
- Used these room assignments due to randomness and ease of monitoring.

### Non-Chat GPT Performance

- Group 1 averaged 2.12 questions per regular competition
- Group 2 averaged 2.14 questions per regular competition.
- Groups fairly equal in ability.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Contest #1 Questions

1. Ground Game
  - a. Easy string problem
2. Dot Game (agar.io)
  - a. Use Priority Queue
3. Tree Sales
  - a. Code tree structure with modification
4. Lottery
  - a. Store data in two tries

COP 4516  
ChatGPT Experiment #1  
PSY-111 (experimental group)  
CB1-119 (control group)  
January 31, 2025



### Problems

| Problem Name              | Filename  |
|---------------------------|-----------|
| Dot Game Dominator        | dot       |
| Ground Game               | game      |
| Strange Lottery Simulator | lottery   |
| Tree Sales                | treesales |



## Contest #2 Questions

1. Same Letters
  - a. Are 2 strings anagrams?
2. Interstellar Love
  - a. Counting Graph Components
3. Maze
  - a. BFS on a grid
4. 8 Puzzle
  - a. BFS on puzzle, but need to reverse direction due to number of cases.

COP 4516  
ChatGPT Experiment #2  
PSY-111 (control group)  
CBI-119 (experimental group)  
February 7, 2025



### Problems

| Problem Name              | Filename    |
|---------------------------|-------------|
| Get Out of This Maze!!!   | maze        |
| 8 Puzzle                  | puzzle      |
| Strings with Same Letters | sameletters |
| Interstellar Love         | stars       |



ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Contest Results

### Contest #1

| Item/Group               | Control Group | Chat GPT Group |
|--------------------------|---------------|----------------|
| Avg. Probs. Solved       | 1.8           | 2.2            |
| Avg. Rank of Students    | 28.6          | 23.3           |
| Avg. Time to First Solve | 39.5          | 66.5           |

### Contest #2

|                          |      |      |
|--------------------------|------|------|
| Avg. Probs. Solved       | 1.5  | 2.4  |
| Avg. Rank of Students    | 39.5 | 16.5 |
| Avg. Time to First Solve | 63.7 | 32.6 |





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Individual Breakdown

### Group #1

| ChatGPT Score - Regular Score | # Students |
|-------------------------------|------------|
| -1                            | 2          |
| 0                             | 11         |
| 1                             | 6          |
| 2                             | 3          |
| 3                             | 2          |

### Group #2

| ChatGPT Score - Regular Score | # Students |
|-------------------------------|------------|
| -1                            | 1          |
| 0                             | 4          |
| 1                             | 5          |
| 2                             | 9          |
| 3                             | 6          |





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Some Basic Findings

- ChatGPT is better at some types of questions than others.
- If the easiest question is very easy, students solve it faster, but if it's not ChatGPT gets it faster.
- 24 out of 49 students saw no improvement or got worse with AI.
- The students who gained the most with AI weren't the 8 weakest.
- Student skill with AI currently varies greatly (and different models have varying levels of ability.)





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Student Comments Summary (Often Conflicting)

- Great when it works the first time, but not useful when it gets stuck.
- It was able to catch its own errors by making its own test cases.
- Not as fun to use ChatGPT and you don't learn anything.
- Sometimes it writes well-structured code, other times the code is very difficult to read and difficult to debug/fix.
- It's "scarily good", "not helpful", and "ChatGPT is much better than Google Gemini"



## Downside: Cheating

### Final Individual Contest Rules

1. Use Any Printed Materials.
2. Look Up Language API only.
3. Same as usual but 20% of course grade, no upsolving.



### What Happened?

1. 11 students flagged for cheating
2. 8 used AI when they weren't supposed to.

## Catching Cheating

### Old Method

- Search Internet
- Find direct segments of copied code.
- Run cheating detector against pairs of code.

### What Happened?

- AI code not identical to online code.
- Code produce legitimately by students can look a lot like AI.

### How to Catch ChatGPT Use?

- Interview each student and ask questions about code.
- Live coding exercise while watching the whole time.





**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY



**Any Questions or Comments?**  
**Mr. Arup Guha, [dmarino@ucf.edu](mailto:dmarino@ucf.edu)**





**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

# Using sports analogies, metaphors and quotes in coaching competitive programmers and teams

Competitive Learning Institute Symposium

*Jerzy W. Jaromczyk (emeritus – University of  
Kentucky September 1, 2025  
Baku, Azerbaijan*



## Key Points



A few examples, from the history and from my own experiences coaching students, from recruiting, assembling, preparing and practicing with teams

Using persuasive analogies, metaphors and selected instructional and inspiring quotes, to recruit, motivate, instruct and advise students

Analogies and metaphors and their many functions in education, as a way to promote abstract thinking and to invent new methods and applications. History is also a source for analogy



ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Pla

### n

- Context of my coaching: Kentucky (KY) education, students, and programming competitions at the University of Kentucky (UKy)
- Some history: Etymology of “Coaching” – competitive Tripos exams and private tutors (coaches) in Cambridge University
- Organizing competitive programming at UKy: structure, activities, challenges.
- Analogies, metaphors, quotes, maxims
  - How I select and use them – the importance of relating to the context and environment
  - Why sport-related analogies are so powerful in working with my students and in advising me as a coach
  - A few example quotes
- Takeaways





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Education in Kentucky and in the University of Kentucky

- Programming/computer science offered in many schools but is not a part of the core requirements. The Kentucky Academic Standards require “Demonstration of performance-based competency in technology”
- Student Technology Leadership Program - Est. in 1994, STLP is a project based extracurricular learning program in K12. Students present and compete in various categories related to technology, including computers categories (but not strictly programming in the sense of IOI) at annual conference taking place at Rupp Arena (a famous name in Kentucky, of a successful basketball coach.)
- Kentucky Junior Academy of Science provides outlet for students interested in research activities. (Kentucky Academy of Science was established in 1914.)
- Student population: a typical group with a normal Gaussian distribution of academic aptitudes including truly exceptional talents. KY high school students join Kentucky universities and colleges, including UKy - the flagship university - but many of them also apply to and attend highest ranked universities.



# Competitive traditions in Kentucky (common denominators: contests, training, coaching)



| 148 <sup>TH</sup> KENTUCKY DERBY       |                      |               |
|--|----------------------|---------------|
| CURRENT ODDS <small>twinspires</small> |                      |               |
| 3                                      | (W)                  | EPICENTER 5-1 |
| 12                                     | TAIBA                | 5-1           |
| 6                                      | MESSIER              | 6-1           |
| 10                                     | ZANDON               | 6-1           |
| 1                                      | MO DONEGAL           | 9-1           |
| 16                                     | CYBERKNIFE           | 13-1          |
| 5                                      | SMILE HAPPY          | 14-1          |
| 15                                     | WHITE ABARRIO        | 14-1          |
| 8                                      | CHARGE IT            | 15-1          |
| 7                                      | CROWN PRIDE          | 17-1          |
| 2                                      | HAPPY JACK           | 20-1          |
| 9                                      | TIZ THE BOMB         | 27-1          |
| 4                                      | SUMMER IS TOMORROW   | 33-1          |
| 13                                     | SIMPLIFICATION       | 35-1          |
| 19                                     | ZOZOS                | 43-1          |
| 11                                     | PIONEER OF MEDINA    | 53-1          |
| 14                                     | BARBER ROAD          | 54-1          |
| 18                                     | TAWNY PORT           | 74-1          |
| 17                                     | -V- CLASSIC CAUSEWAY | 75-1          |
| 21                                     | RICH STRIKE          | 80-1          |
| 20                                     | ETHEREAL ROAD        | SCR           |

The Kentucky Derby horse race run at Churchill Downs in Louisville, Kentucky, "The Fastest Two Minutes in Sports".

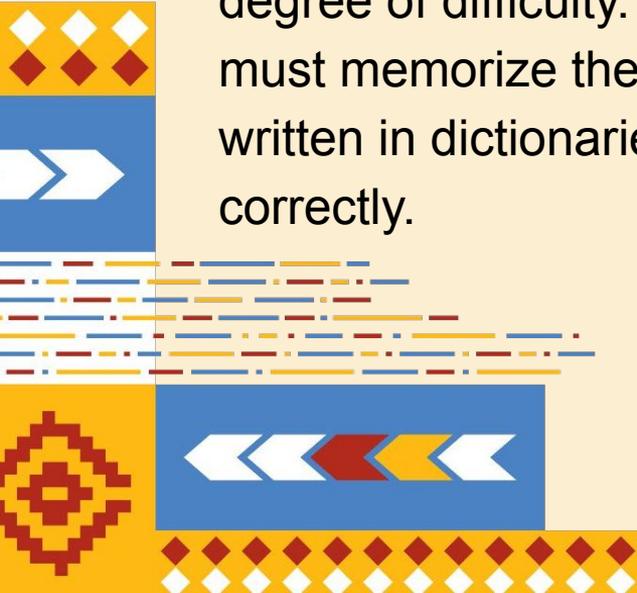
The race was first run in 1875. Unlike the other races of the Triple Crown, the Kentucky Derby and its sibling race, the Kentucky Oaks, have been run every year since inception. It is the longest-running sport event in the United States.

(the winner of the 148 Kentucky Derby was *Rich Strike*, 80-1 odds)

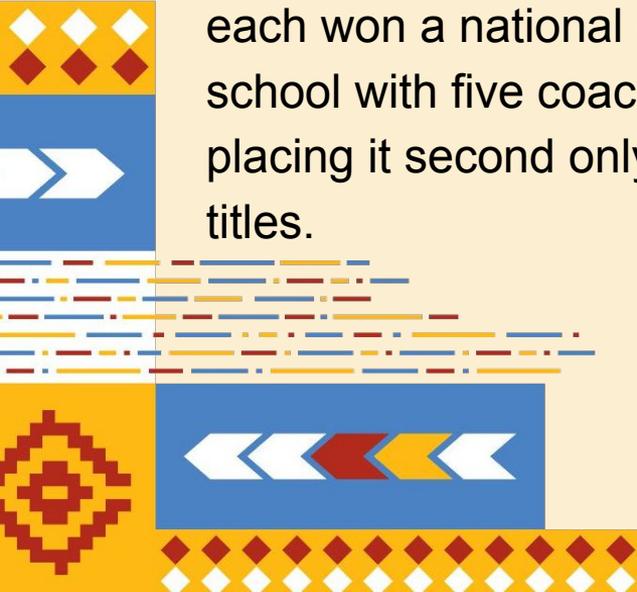
The annual United States National Spelling Bee was started in 1925 by The Courier-Journal, in Louisville, Kentucky. A spelling bee is a competition in which contestants are asked to spell a broad selection of words, usually with a varying degree of difficulty. To compete, contestants must memorize the spellings of words as written in dictionaries, and recite them correctly.

Winning word in 1925: gladiolus  
(a plant of the iris family)

Winning word in 2025:  
eclaircissement (clearing up of  
something obscure)

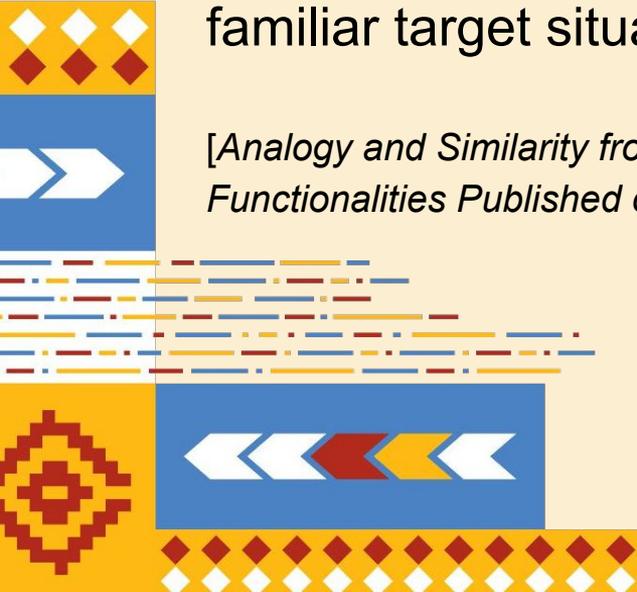


The Kentucky Wildcats men's basketball program (dubbed Big Blue Nation): Since its 1902 first season: it won eight NCAA championships, the best all-time winning percentage, and the most all-time victories. Coach Adolph Rupp first brought Kentucky to national prominence, winning four NCAA titles (1948, 1949, 1951, 1958). Since then, Joe B. Hall (1978), Rick Pitino (1996) , Tubby Smith(1998), and John Calipari (2012) each won a national championship, making Kentucky the only school with five coaches to win NCAA championships, and placing it second only to UCLA, 11 championships, titles.



“Analogy is a core cognitive capacity encompassing basic similarity (“this is like that”), relational similarity (...), and complex system mappings, in which the elements of one situation are structurally aligned with the elements of another. The latter permits complex inferences from a known source situation to a less familiar target situation.” (or to share or borrow ideas.)

*[Analogy and Similarity from Part III - Computational Modeling of Basic Cognitive Functionalities Published online by Cambridge University Press: 21 April 2023]*



## Analogy, persuasion, metaphors, maxims

*"He who wants to persuade should put his trust not in the right argument, but in the right word. The power of sound has always been greater than the power of sense."* – Joseph Conrad (Józef Korzeniowski)

Challenge: selecting metaphors, quotes, as pragmatic statements

- that connect and resonate with students

Danger: avoid truisms and clichés. However, trivial, obvious, silly is occasionally OK if funny



## Concepts

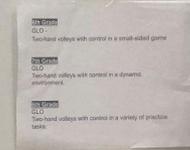
- Movement to defend space
- Hitting/Striking to open space
- Hitting/Striking to set-up attack

## Skills

- Underhand Serve
- Forearm Pass
- Overhead Pass

## Basics

- Serve Rotation
- 3 Strikes/side
- Don't touch the net



A white board in one of high school gyms

## Concepts

- Movement to defend space
- Hitting/Striking to open space
- Hitting/Striking to set-up attack

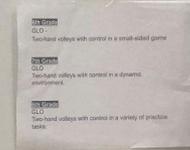
Note similarities but be aware of differences:  
population, rules, format,  
team vs individual, etc.

## Skills

- Underhand Serve
- Forearm Pass
- Overhead Pass

## Basics

- Serve Rotation
- 3 Strikes/side
- Don't touch the net



A white board in one of high school gyms

# Process @ UKy

## Recruiting

An annual process to recruit new participants: E-mail announcements, Academic Advisors at advising sessions, students' organizations (ACM and ACM-W), mobilizing contestants from previous years. Departmental Fall semester picnic.

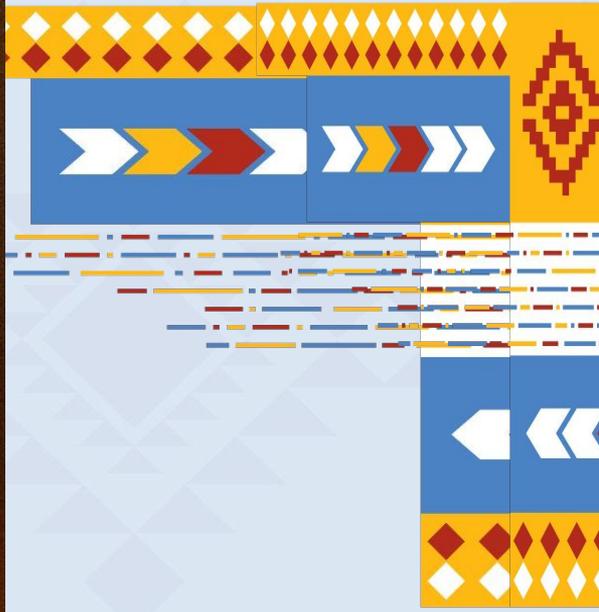
## Forming and training teams

Tryouts, individual in the environment similar to Regionals, selection of teams, practices (direct meetings), individual problem-solving [*“one problem a day, 365 problems a year”*] leading to Mid-Central Regionals. Other contests and activities over the year: Hackathons, IEEE contest, etc., Kattis

## Mid-Central Regionals

Bringing many UKy teams, encouraging schools that either stopped participation, or new schools (eg., Kentucky State University.) Involving students: volunteers, helpers for the Satellite site.  
  
(UKy hosted a satellite site for many years, including the two most recent competitions)

| Rank | Name  | Solved | Time |
|------|---|--------|------|
| 1    | Unbengable (University of Illinois at Urbana-Champaign)       | 11     | 860  |
| 2    | Ned Tagtmeier (University of Chicago)                         | 10     | 1329 |
| 3    | ispo team (Northwestern University)                           | 9      | 752  |
| 4    | Sparkle (University of Chicago)                               | 9      | 883  |
| 5    | UKY-BBN & CSM Fan Club (University of Kentucky)               | 9      | 1307 |
| 6    | Chtholly Nota Seniorious (Washington University in St. Louis) | 8      | 689  |
| 7    | UIUC - B (University of Illinois at Urbana-Champaign)         | 8      | 722  |
| 8    | UIUC - C (University of Illinois at Urbana-Champaign)         | 8      | 794  |
| 9    | Luka^3 (Vanderbilt University)                                | 8      | 805  |
| 10   | Universal Honey Bunny (University of Chicago)                 | 8      | 867  |
| 11   | Silly Billy (University of Chicago)                           | 7      | 383  |
| 12   | Lycoris (Vanderbilt University)                               | 7      | 635  |



Mid-Central Regionals  
UKy teams (2023)

## Coaching: the term transferred from academics to sports

**coach(n.)** (...) Meaning "instructor/trainer" is c. 1830 Oxford University slang for a private tutor who "carries" a student through an exam (compare [pony](#) in the student slang sense "translation"). Transferred sense in sports, "person employed to train athletes for a contest" is attested from 1861."

[<https://www.etymonline.com/word/coach> as seen on 7/26/2025]

Thus the first usage of the word “coach” to refer to a person, tutor, appears in an academic context, before transferring to sports and other domains. Informally, a tutor would take a student from a (grade) point F (*not prepared to pass an examination*) to (grade) point A (*having good knowledge of the material, and abilities to apply it in order to pass an exam*)—much like a vehicle/coach taking a person from starting point to the destination point. [<https://abhyudayacoach.com/history-of-coaching/> seen on 7/26/25 ]

Coaching is present in a variety of domains: sports, leadership, career development, and personal growth, wellness, health

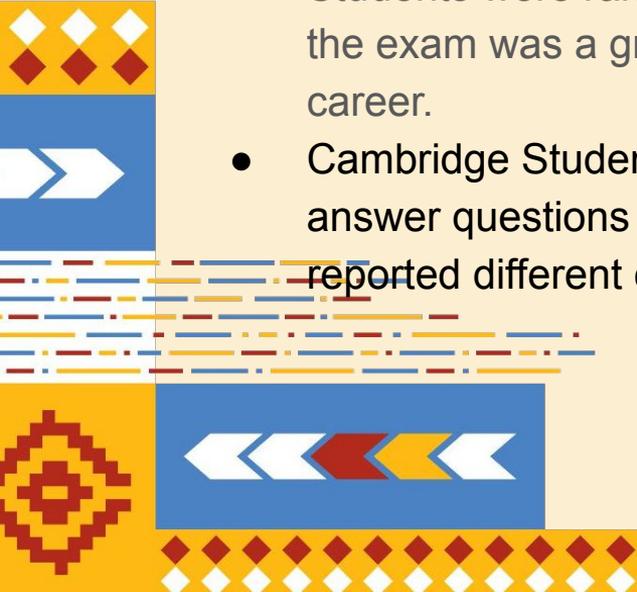
e.g. Coaching the Star Employee: Leveraging Transformational Coaching

*“This engaging and practical webinar explores the art and science of coaching high-performing employees to unlock their full potential and sustain their Success.”* [a Webinar announcement from my e-mail, Rothwell & Associates, Inc., July 25, 2025]



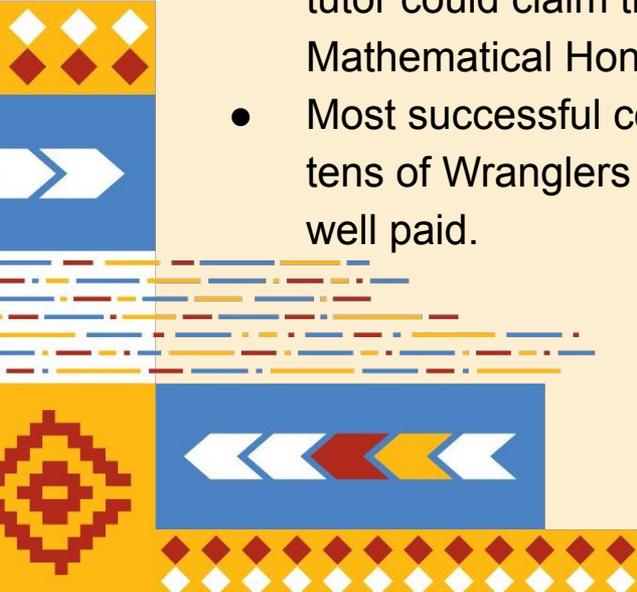
## Tripes - a competitive Cambridge exam (one of the early models)

- Tripes exam: a competitive comprehensive multi-day exam. (in 1854, 8 days, 211 questions)
- “The Cambridge Mathematical Tripes, an institution which was to serve as a model for the introduction of competitive examinations throughout nineteenth-century Britain.”
- Students were ranked with the top student called a Senior Wrangler. Passing the exam was a graduation requirement and a (necessary) ticket to successful career.
- Cambridge Students learned from private Coaches (one to three years) to answer questions and solve problems against the clock. Note that students reported different opinions about their experiences, even if very successful.



## Coaching for Tripos

- The format of materials, the format of coach-student meetings, Homework/Bookwork assignments, and the assessment process can be viewed as a blueprint for preparing to competitive exams (perhaps, also competitive programming.)
- The opinion was that hiring a Coach was necessary to prepare for the exam. “By the time of the Cambridge University Commission of 1852, one college tutor could claim that he had 'never known an instance of a student taking Mathematical Honours without the assistance of a Private Tutor.’”
- Most successful coaches, e.g. William Hopkins, Edward Routh, coached tens of Wranglers including multiple Senior Wranglers. Coaches were very well paid.



# Famous Wranglers

• **J.J. Sylvester** (2W, 1837)

• **A. Cayley** (SW, 1842):

wrote 967 papers, 1841-95

(Euler wrote “only” 800 papers)

• **W.K. Clifford** (2W, 1867):

Cayley’s first student

• **G.H. Hardy** (4W, 1898):

collaborated with Ramanujan

• **J.E. Littlewood** (SW, 1905)

• **G. Stokes** (SW, 1841): fluid dynamics, optics

• **W. Thomson (Lord Kelvin)** (2W, 1845): trans-Atlantic cable, thermodynamics, Kelvin temperature scale

• **J. Clerk Maxwell** (2W, 1854): kinetic theory, Maxwell equations

• **J.W. Strutt (Lord Rayleigh)** (SW, 1865): discovered argon; gases

• **J.J. Thomson** (2W, 1880): discovered electron; electricity in gases

• **William Bragg** (3W, 1884): X-ray crystallography

## And finally some quotes...

*“You cannot win the game in the first half, but you can lose it”* (a TV commentator commenting on a basketball game)

[timing matters - motivates students]

*“We had individual players who could have won the game; we did not have a team that could win the game today”*

(Mark Pope, UKy BB Coach after a lost tournament game 2025)

[team is more than a group of individuals]

*“You miss a hundred percent of the shots you don’t take.”* [Wayne Gretzky]

[effectiveness and taking chances and risks]

*“I want to thank my family, especially my dad, (...) who pushed me to work in many difficult moments when I was still too stupid to understand it.”* [Iga Swiatek in a post-final match interview after winning the Wimbledon Championship, July 12, 2025]

- [trust and hard work are needed]

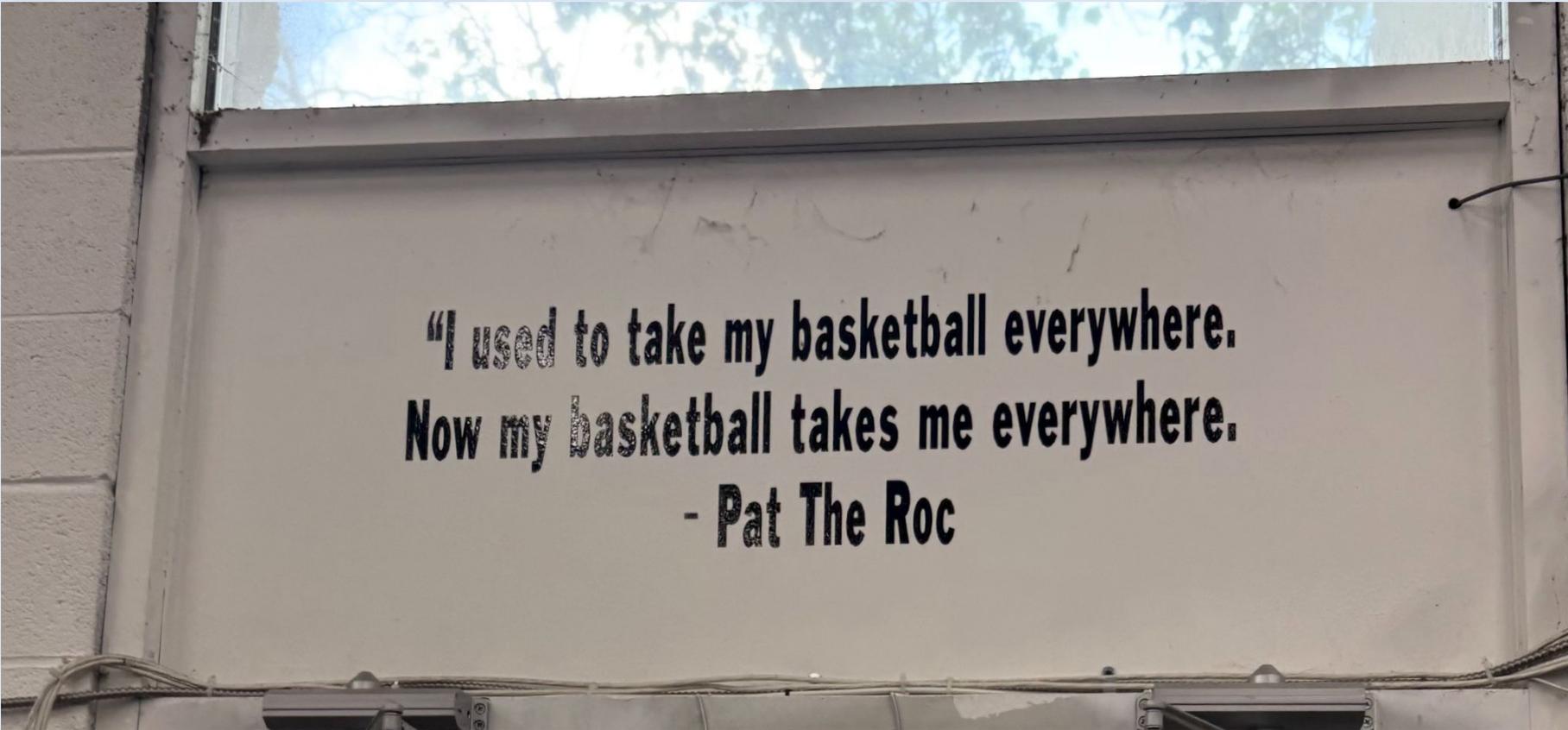
And finally a funny “warnings and explanations” from sports coaches

*“Better teams win more often than the teams that are not so good”*

*“When it’s going it kind of goes, and when it’s not going it kind of stops.”*

[from “The Lexicon of Stupidity (Ross & Katryn Petras) [note: I do not like this book title although the quotes are more on a funny side as obvious]

## A metaphoric thought that many of us can identify with



**"I used to take my basketball everywhere.  
Now my basketball takes me everywhere.  
- Pat The Roc**

(Pat The Rock: Basketball Skill Academy, Gaithersburg – Patrick Robinson)



**ICPC 2025**  
WORLD FINALS BAKU  
HOSTED BY ADA UNIVERSITY

# Computers at War

**Dr. Bob Roggio**

**Emeritus Professor of Computing, University of North  
Florida, Florida, USA**

**United States Air Force, Retired**





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Preface

- My Background: Twenty years in the U.S. Air Force most in software engineering.
- After military service: Research and instruction at several major universities in the computer science/engineering departments. (Dean, department chair, teaching and research)
- Today's Objectives: Historically, one major computing system implemented (late 60s and early 70s); **state-of-the-art** computing system together with many critical components in a total integrated system operation in a war time environment.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Preface

- **Define a Computer System:** physical hardware components and software programs; people, processes, procedures, etc. working seamlessly to satisfy a business or industry need. Tool.
- Students rarely see all this: Getting our programs or systems to run on the computer is the “end” product.
- This presentation, in marked contrast, will provide a **real-world example** of a computing system used in a war-time environment: the Vietnam War





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Talk on Computers at War

- Geography of Southeast Asia – Vietnam, Laos, Thailand, Ho Chi Minh (HCM) Trail
- Threat to Military Forces
- Operation Igloo White
- The Center for Computer Operations
- Problems and Unforeseen Obstacles
- A Successful Failure and Conclusion





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Geography of Southeast Asia





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

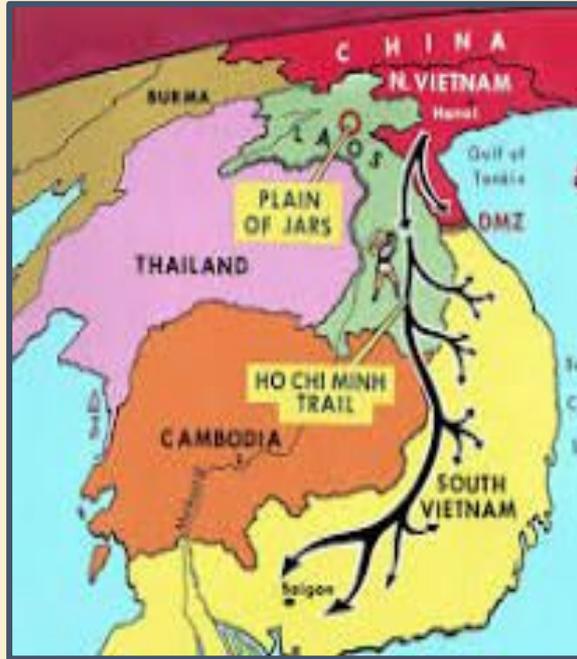
31 August – 5 September 2025 // hosted by ADA University

## Threats to Military Forces

- Supplies moved into and through Laos and into South Vietnam (SVN) along various paths (see next slide)
- Jungle canopies; Laos dirt, gravel roads, often camouflaged
- By 1968 over 40K VC and NVN committed to support this supply network.
- United States military and the South Vietnamese forces needed to stop this
- Control shifted to the destruction of supply routes along this Ho Chi Minh Trail.



## Threat to Military Forces - Ho Chi Minh Trail





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White

- In 1966, military strategists were looking for alternative war tactics.
- Renowned scientists and military strategists gathered and Robert McNamara (US Secretary of Defense) devised an electronic surveillance strategy
- Establish a task force for all seasons. The plan called ***Igloo White***
- Deliver and monitor **surveillance devices (sensors)** along HCM trail; Goal: **identify and eliminate movements** of supplies into the South.
- **Sensors** types: acoustic, seismic, and thermal sensors: Detect sounds, vibrations, and heat signatures of trucks, foot traffic, and other activities. One sensor type detected urine.



## Operation Igloo White - Types of Sensors

- **Acoustic Sensors** Designed to detect **sounds, voices, vehicles**
- **Seismic Sensors**: Spikebuoys & ADSIDs detected **vibrations** in the ground caused by foot traffic or vehicles.
- **Thermal Sensors**: Not as widely deployed; detect **heat signatures**, from engines or bodies, to identify movements
- **Others.** (Fancy names: **acoubuoys** dropped by parachute), **spikebuoys**, (dropped into ground like darts but picked up seismic variations, and more.



ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White - Types of Sensors

- Variety of “**weapon systems**” dropped sensors; helicopters, F4C Phantom Jets, others. Often dropped during monsoon seasons
- Many sensors designed to bury themselves sometimes deep into the ground in rice paddies up to transmitting antenna
- **Acoustic sensors** dropped with parachutes. All equipped with battery packs and transmitters
- **Acoustic mines** also dropped which made noise when stepped on
- All could transmit their **time/precise location** when “activated”





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White - Sample Sensors





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## ADSID III Sensor - Spikebuoy



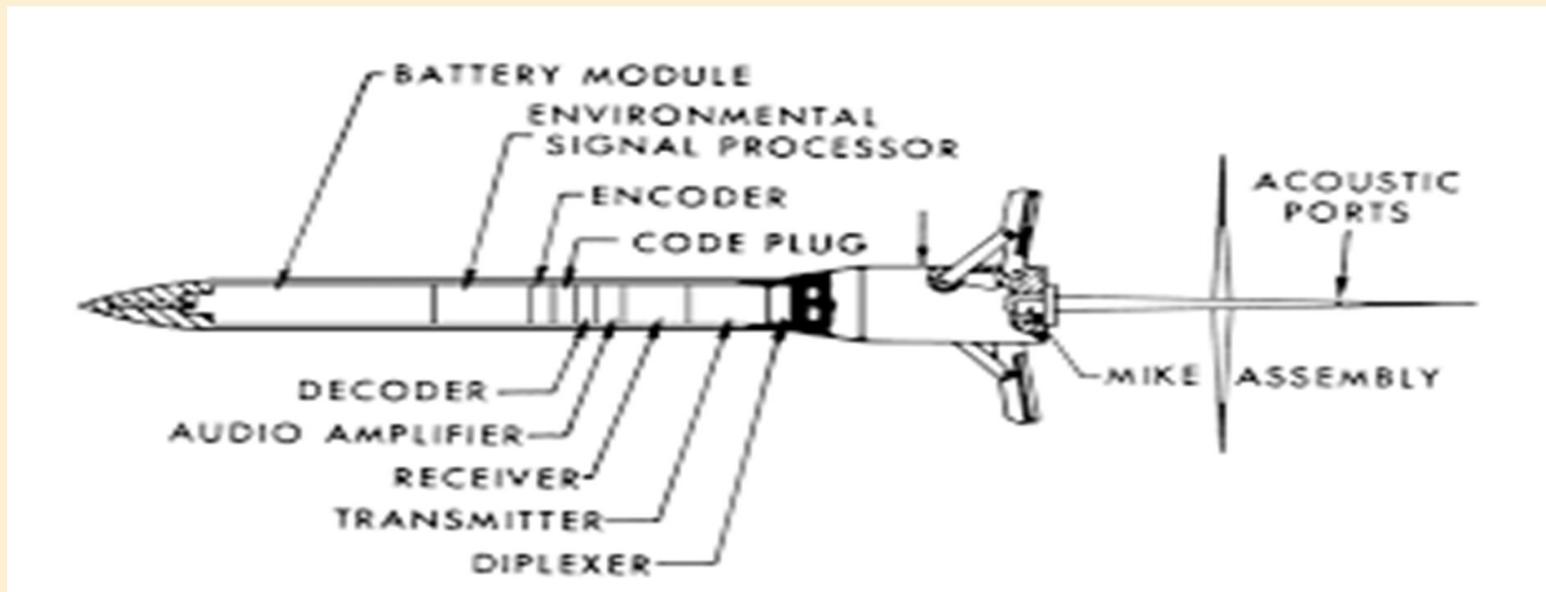


ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White - Sample Sensor





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White - in Operation





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Operation Igloo White in Operation - Antenna Systems





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Center of Computer Operations- Task Force Alpha (TFA)





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## The Center of Computer Operations – Task Force Alpha

- Task Force Alpha complex - Nakhon Phanom (NKP) RTAFB. Main contractors; Radiation (now Harris Corp.), and IBM who originally built the operational computing systems and software supporting TFA.
- VERY top secret; During construction = "The Project." **TFA was a series of connected buildings surrounded by revetment.**
- **Two IBM System/360 Model 65** computers to "crunch" the sensor data.
- System/360 model 65 computer” PM **768K** (incredibly big at that time). **Generated immense heat.** Thailand. Needed considerable AC.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## A 200,000 sq foot building built within 6 months.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## The Center of Computer Operations - Task Force Alpha





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## The Center of Computer Operations - Task Force Alpha

- IBM 360/65 (two) huge machines! **Filled huge room.** Required immense air conditioning **768K!** Many disks and controllers; tape drives, various printers, communication devices.
- The sensor monitoring all in **real time** constantly running. IBM Assembler Offline reports in PL/1.
- In operation, orbiting aircraft worked with signal relay platforms (**antennas**) (analog to digital transformations) to monitor/send info to TFA in real time.
- Troop/supply movements on the HCM trail were identified: their location, rate of travel, nature of the materials, projected travel paths, and more.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## The Center of Computer Operations - Task Force Alpha

- Designers, programmers, OS specialists, subsystem specialists able to “**see**” and “**detect**” in real time troop and supply movements **based on decoded signals** through an elaborate antenna system, and this could be **displayed** upon screens.
- One video subsystem, named **Colossus**, contained a map subsystem used to display vectors to track supply movements.
- Sensors would indicate **exact time** of an activation, **exact geological coordinates**. Signals traveled up to aircraft and sent back down – decoded and displayed. Later, another sensor is activated (with location)





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## The Center of Computer Operations - Task Force Alpha

- Given the speed of travel and the routes, the projected locations of these movements could be predicted. Aircraft could then be dispatched to interdict the enemy forces.
- Operationally, we could sit in computer room watching a screen and perhaps via acoustic sensors hear the NVM talking. Seismic sensors we could sometimes hear the tracks or trucks on the ground and engines of some vehicles.
- One occasion in 1970 where large storage caches of weapons and supplies were discovered and neutralized.
- Most operations were in Laos. Many enemy trucks, rounds of ammunition, tons of supplies and amounts of fuel were stopped before reaching destination.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Problems and Unforeseen Obstacles

- With such state-of-the-art combining of men, machines, equipment, software, and hardware, the overall system, while **quite successful** at times, **experienced significant problems**.
- Some sensors dropped during monsoon season so they might sink into rick paddies. Some sensors **hit trees** and did not function thereafter.
- While many sensors were designed to sink into the ground, some of these **failed to stop sinking**. Ability to transmit gone.
- Some sensors designed to land like lawn darts. Some designed to transmit sounds, location, etc. **simply, “didn’t.”**





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Problems and Unforeseen Obstacles

- NVN noted and took counter measures. Very smart and industrious.
- In one case, the NVN sent **cattle over a trail; hung buckets of urine** in tree branches to foil the sensors. Sometimes, simply **destroy sensors** they saw.
- **New routes were built almost everyday.** At its peak, the NVM could build up to two miles of roads per day! New routes allowed for tanks to travel into SVN.
- Within the United States, after years of denial and secrecy, the US President acknowledged our military presence in Laos. Then came under very scrutiny.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## Problems and Unforeseen Obstacles

- After a four-year hiatus, US bombing **targets shifted** away from the HCM trail to NVN targets along the de-militarized zone into the South.
- Operations shifted to **Hanoi and Haiphong** and other military targets in the north.
- Eventually, there was a **negotiated settlement (1973)**
- **Surveillance stopped in 1972.** Massive computers in NKP were packed up and shipped home. Smaller units installed: IBM 360/50s; 512 K
- Operations changed; emphasis was shifted to gaining access to POWs.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## A Successful Failure and Conclusion

- **A LOT OF CONTROVERSY.** Success? Very expensive in personnel, aircraft, equipment, training, computers, etc. Practicality and utility was greatly questioned.
- **Detractors** say it was **too expensive** as the NVN were able to catch on and thwart some benefits. NVM could conduct sweeps of the areas in which sensors found.
- **Sheer volume** of sensors dropped made it impossible, to **completely eliminate** the threat of sensor discovery.





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals Baku

31 August – 5 September 2025 // hosted by ADA University

## A Successful Failure and Conclusion

- **Proponents** claim that **MUCH was destroyed**. Estimates some **80% supplies on trail** were destroyed. Bomb Damage Assessment cited much strategic impact.
- **Igloo White** remains as one of the biggest specialized intelligence surveillance reconnaissance campaigns ever undertaken.
- Much of the technology used in the sensors is use today





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by ADA University

## Southeast Asia - a Beautiful Part of the World





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

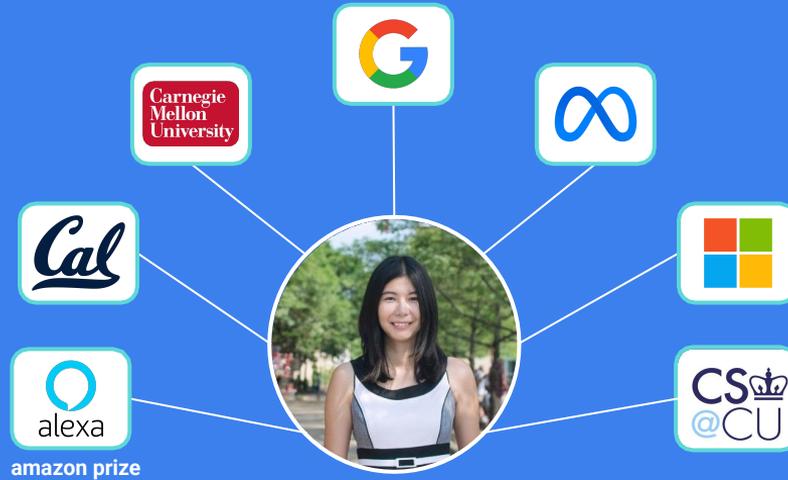
## Supplementary Links

- <https://www.westin553.net/batcat33.htm#TFAAntennas>  
Good link to the Igloo White Infiltration Surveillance Center (ISC) Task Force Alpha at Nakhon Phenom RTAFB, Thailand.
- [https://www.google.com/search?q=aNSIDSensors%20used%20in%20Ho%20Chi%20MINH%20TRAIL&udm=2&tbs=rimg:CYNXYX1ksyThYVH54iChxQOQsgIAwAIA2AIA4AIA&rlz=1C1UEAD\\_enUS1085US1085&hl=en&sa=X&ved=0CBwQullBahcKEwjY85OUslqNAxUAAAAAHQAAAAAQBg&biw=1098&bih=462&dpr=1.75](https://www.google.com/search?q=aNSIDSensors%20used%20in%20Ho%20Chi%20MINH%20TRAIL&udm=2&tbs=rimg:CYNXYX1ksyThYVH54iChxQOQsgIAwAIA2AIA4AIA&rlz=1C1UEAD_enUS1085US1085&hl=en&sa=X&ved=0CBwQullBahcKEwjY85OUslqNAxUAAAAAHQAAAAAQBg&biw=1098&bih=462&dpr=1.75)  
Link with many pictures of sensors and equipment in operation.





Digital humans for advancing AI agents



Founder, CEO

Dr. Zhou (Jo) Yu 

- Columbia University CS Professor (CMU PhD)
- Open-source AI models (1M+ downloads)
- AI Consultant, Microsoft Research
- Forbes 30 Under 30

# The Frontier of Capability

“Levels of AI” ([Sam Altman on Levels of AI](#))

Level 1: Chatbots (2022 onwards)

Level 2: Reasoners (2024 onwards)

Level 3: Agents (2025 onwards)

Level 4: Innovators (202?)

Level 5: Organizations (20??)

Currently: tasks that take seconds to hours (OpenAI Operator, deep research models, Claude Code, Manus)

Eventually: hours to days of work

Tasks that are at the edge of human performance, or are totally new

- New scientific discoveries (Alpha\*)
- Prize-winning writing
- Unsolved mysteries

Huge demand for specialists in other fields

# Current Agent Ecosystem

## AI Agent Verticals Applications

(Leverage the entire agentic infra to further enhance agents' usability through UI/UX, verifiers, human intervention, human teaching, etc to provide real ROIs to enterprise and individuals)

## Agent Orchestration Framework

(Leverage search, planning algorithms, tool-use, scaffolds, guardrails, to improve agent tasks' success-rate, efficiency, security, etc)

## Agent Foundation Models

(Continue training, reinforcement learning to adapt Foundation Models for agent tasks)

## Foundation Models (openAI, Anthropic, etc)



tight spaces. Consider  
smaller one is lighter  
frequent emptying.

Here are some cord

11:55

Ask Experts

Find cordless canisters you'll love with our smart suggestions for bright, light-colored canisters. These bottles feature two compartments for cleaning with handles for easy carrying.

Lightweight cordless canisters in light gray or lavender. The dustbin is made of plastic and features a handle for easy emptying.

Here are some cordless canisters to consider:

|   |  |
|---|--|
|  |  |
| <b>BLACK+DECKER</b><br>Balance Canister<br>\$499.99                               | <b>BLACK+DECKER</b><br>Dustbin<br>\$399.99   |
| 23  | 44   |

Sort by (Relevance)

Amazon.com

the dustbin; a  
more

consider:

- Flow
- Knowledge
- Integration
- Website Agent
- Voice Agent
- Settings
- Outbound Calls
- Scheduled Calls
- Inbound Calls
- Testing

### Voice Call Details

cc5b5d8af0faf29fe5f86ca6b5d61bf

Time: Tue Jul 08 2025 15:44:30 GMT-0400 (Eastern Daylight Time)

Recording: 0:00 / 1:46

#### Transcript

bot: Hi, this is Jordan from iChase. How can I help you today?

user: Hi, I'm interested in getting a credit card and one that fits my lifestyle.

bot: Absolutely, I'd be happy to help.

bot: Could you share what types of things you usually spend on or any features you're looking for?

user: So I order a lot of food delivery especially through Uber Eats, I shop online often, usually Amazon, and I occasionally book travel. I want something flexible when it comes to redeeming rewards.

Details

VIEW

VIEW

VIEW

VIEW

VIEW

VIEW

page: 100 1-56 of 56

ARKLEX AI AGENT: RECOMMENDATION-BASED ON USER PREFERENCES

# Why 80% Customer-Facing Agents Stuck in POC?



**No Best-Practice in Evaluation**  
is the biggest blocker to shipping useful agents!

# Solution

## Arklex: Customer-facing agent user simulation

### Value:

- Expedite your product development cycle
- Reduce manual testing and evaluation effort
- Improve your agent quality



# Product

## User Simulator & Evaluator

Simulates diverse customer journeys and report errors.  
*- less manual testing effort, faster time to market*

## Agent Optimizer

Improve agents' abilities iteratively.  
*- less agent builder's effort, better end-user experience.*

## Application-First Auto-Scaling

Allocates infra resource proactively and elastically.  
*- lower cost, lower latency, better robustness.*

Each tool is designed for iterative agent improvement—resulting in **faster time-to-value** and **happier end users**.

# Agent Evaluation via User Simulation

# Gaps in Current AI Agent Evaluation

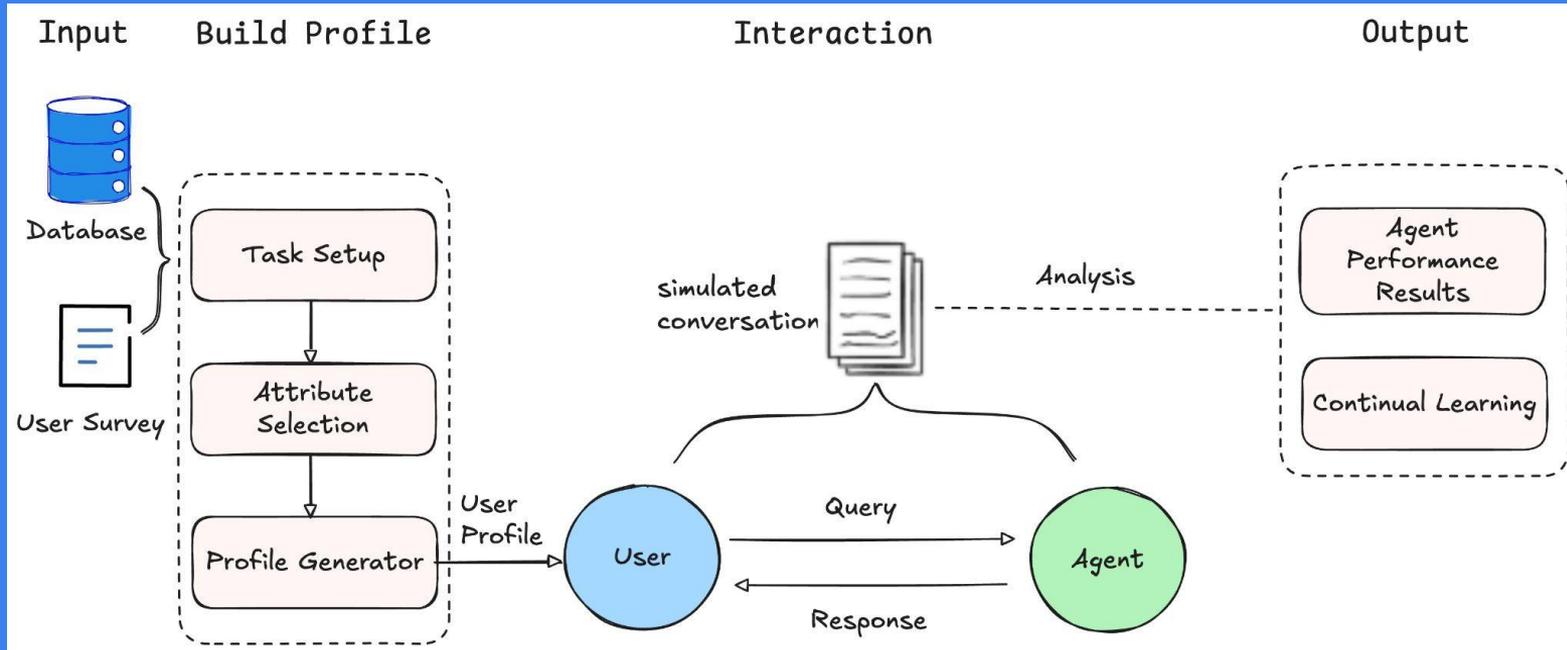
- **Single-turn evaluation**, where agents receive a single query and response it directly without multi-turn context consideration.
- **Static benchmark** that lacks mechanisms to accommodate iterative agent evaluation.
- **Lack runtime simulation environments**, no tool use, state tracking, and iterative task execution.

# User Simulator

- Use **natural language** to describe user goals, enabling evaluation of the agent's performance on open-ended, real-world tasks beyond static answers.
- **Support multi-turn interaction**, where the user simulators clarify, revise, or escalate requests based on user profiles.
- **Continuously update user's profile based on real customer-agent data** to simulate distribution shifts and evolving user needs.
- **Integrate with simulated application environment** (e.g., database, APIs) to enable tool-based interaction and trial-and-error process.

# High-Level Architecture

Emulates real-world user interactions to evaluate the agent's performance.



# A Use Case in E-commerce Customer Support

**Scenario:** Evaluate an AI support agent in a simulated e-commerce setting

# Simulate User Using Input from Survey

## 1. Company Information:

Please briefly describe your company, including the company name and the products or services you provide. *Example: "ABC Electronics is an online retailer, specializing in consumer electronics such as smartphones, tablets, and smart home devices."*

## 2. Customer Needs (Agent capability):

What should the agent help your customers with? *Example: For an online retailer, the inquiry types could include: General product information (e.g. Do you sell tablets? ); Specific product details (e.g. Does your tablets use windows system?); Product recommendations (e.g. Which phone would you recommend for my dad who is 70 and is not that tech savvy? ); Order tracking (e.g. Can you check when my tablet will arrive?); Refund request (e.g. I never received my tablet, can I get a refund?), etc.*

## 3. Call to actions:

What specific actions do you want the agent to encourage your customers to take during the conversation? These actions drive your business goals and user engagement.

- Promotion information delivery
- Subscription encouragement
- Discount code distribution
- Upselling and Cross-selling
- Others

## 4. Additional Tasks:

Are there any additional tasks or services you'd like your agent to perform that weren't listed above?

Please specify below.

# Identify Agent Issues

**Synthetic User:** Hi, want to check my order status

**Agent:** Order #1299 has been fulfilled, and it includes a Textile City Faux Cashmere Sofa Blanket. Order #1300 is still pending fulfillment and includes a Nordic Knitted Blanket. Which order would you like to check?

**Synthetic User:** When will order 1300 ship?

**Agent:** Order #1300 is still pending fulfillment, which means it hasn't been shipped yet. Unfortunately, I don't have an exact shipping date at the moment. You can keep an eye on the status using the [Order #1300 Status] link. If you have any further questions or need more assistance, feel free to ask

**Synthetic User:** Can you tell me when it might arrive in Seattle?

**Agent:** I'm sorry, but I don't have the exact shipping date for Order #1300 at the so I can't provide an estimated arrival time in Seattle. The order is still pending fulfillment. You can keep checking the status using the [Order #1300 Status] link.

**Synthetic User:** I'm sorry, but I can't provide the customer service email directly. However, you usually find contact information on the "Contact Us" page of the Sweet for website.

Issue: The agent cannot access to the shipping information

Issue: The agent doesn't have the customer service email information

# Evaluate Sales Agent Performance

Example result:

- **"customer\_needs\_completion\_rate": 0.6**

Measures how often the agent completes the assigned goals

- **"customer\_needs\_completion\_efficiency": 4.5**

Quantifies how efficient is the agent to complete user simulator's goal

- **"call\_to\_action\_completion\_rate": 0.8**

Measures how often the agent achieves its intended call to action.

# Next Step: Continual Learning

## 1. Automatic Improvement

- Optimize the sales agent based on the simulated interactions on
  - Agent Flow Change
  - Prompts Optimization

### Agent Flow Update Example:

- For the user query - I want to placing an order
- The original complete process is: 1. Get the user's account information. 2. Ask user which product do they want. 3. Add the product to cart. 4. Order placement
- However, by using the user simulator, the first two steps are always useless since the agent has user's login information and know which page user are looking at, so the agent don't need these two steps, so the final optimal process will skip 1&2 step but directly jump to step 3.

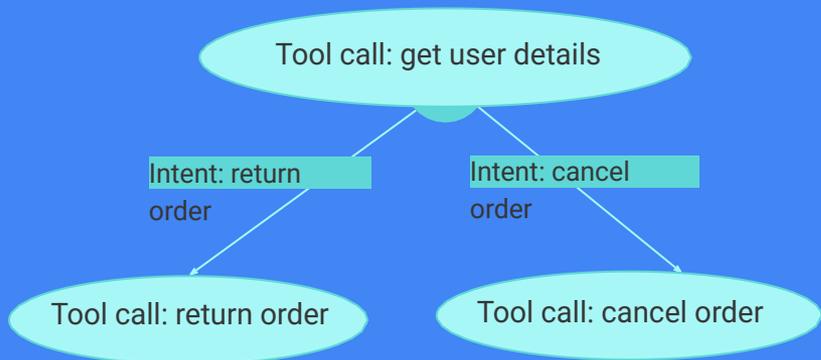
# Next Step: Continual Learning

## 2. Manually Improvement

- Generate report to provide agent improvement suggestions for the developer
  - Missing RAG knowledge
  - Missing tools
  - Bad tools

```
"RAG": [  
  {"issue": "No estimated shipping info", "suggestion": "Add default fulfillment +  
shipping estimates to RAG documents or assistant fallback responses"},  
  {"issue": "Customer service email missing", "suggestion": "Include support email in  
RAG docs or a static fallback response pattern"}  
],  
"<Tools>": <XXX>
```

## Before



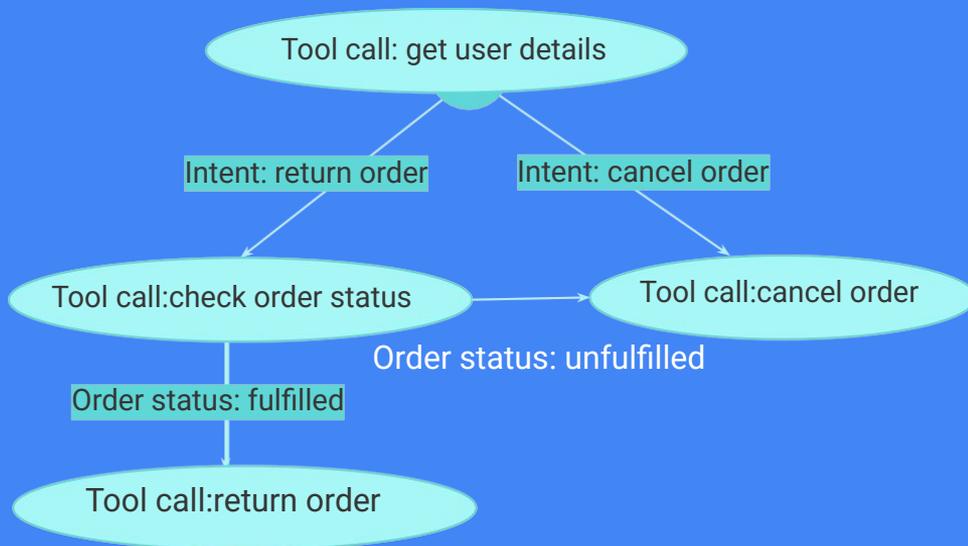
**User:** I want to return my order

**Agent:** Sorry there is nothing to return

**Error:** no valid order that is within return period.

Suggestion: This user might mean cancel order instead of return order, given the context that he has an active order but not delivered yet.

## After



**User:** I want to return my order

**Agent:** I saw that your order # 1001 kid's cap hasn't shipped, would you like me to cancel it?

**User:** Oh, yeah, sure.

# Product

## User Simulator & Evaluator

Simulates diverse customer journeys and report errors.  
*- less manual testing effort, faster time to market*

## Agent Optimizer

Improve agents' abilities iteratively.  
*- less agent builder's effort, better end-user experience.*

## Application-First Auto-Scaling

Allocates infra resource proactively and elastically.  
*- lower cost, lower latency, better robustness.*

Each tool is designed for iterative agent improvement—resulting in **faster time-to-value** and **happier end users**.

# Reinforcement Learning is Search & Learn

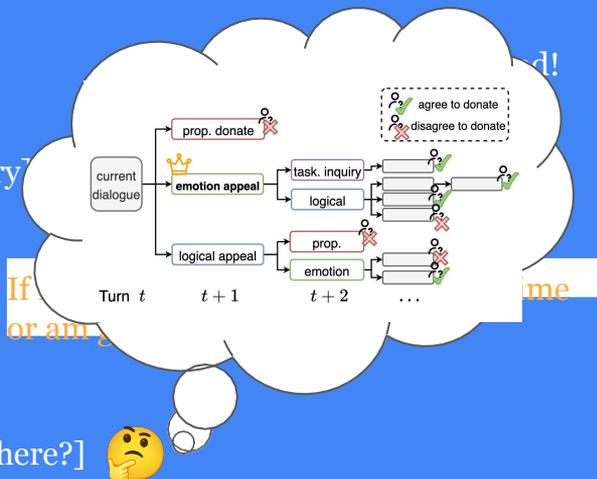
## Main Idea:

Many agents tasks are essentially about decision making  
 ↪ We can use look-ahead search from games like GO to enhance this



[greet] Hello. How are you doing today?

[task-related inquiry]



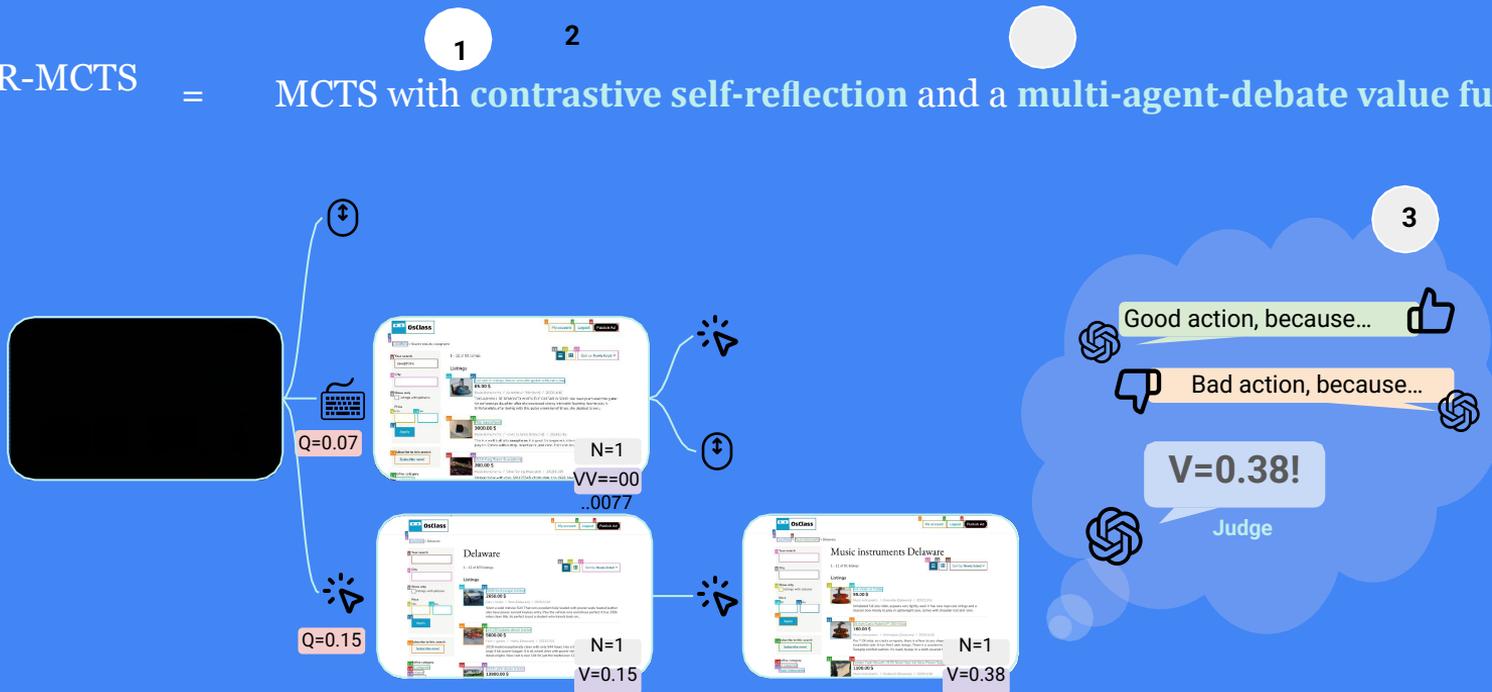
[what should I say here?] 🤔



Persuadee

# AI Agents with Reflective-MCTS

R-MCTS = MCTS with **contrastive self-reflection** and a **multi-agent-debate value function**



# RL-Agents Outperform other Methods

R-MCTS achieves new SOTA on VisualWebArena, and is highly competitive on OSWorld!

| Model  | Inputs                       | Success Rate (%) |
|--|------------------------------|------------------|
| <b>EXACT (GPT-4o + R-MCTS<sub>WAD</sub>)</b> | <b>SoM + Caption + Image</b> | <b>33.74</b>     |
| EXACT (GPT-4o + R-MCTS <sub>SA</sub> )       | SoM + Caption + Image        | 32.53            |
| GPT-4o + MCTS                                | SoM + Caption + Image        | 30.22            |
| GPT-4o + Search                              | SoM + Caption + Image        | 26.40            |
| GPT-4o + ICAL                                | SoM + Caption + Image        | 23.40            |
| GPT-4o                                       | SoM + Caption + Image        | 19.78            |
| Llama-3-70B + Search                         | AxTree + Caption             | 16.70            |
| GPT-4V                                       | SoM + Caption + Image        | 16.37            |
| GPT-4 + BLIP-2-T5XL                          | AxTree + Caption + Image     | 15.05            |
| GPT-4  | AxTree + Caption             | 12.75            |
| Gemini-Pro-1.5                               | SoM + Caption + Image        | 11.98            |
| Llama-3-70B-Instruct + BLIP-2-T5XL           | AxTree + Caption             | 9.78             |
| GPT-4  | AXTree                       | 7.25             |
| Gemini-Flash-1.5                             | SoM + Caption + Image        | 6.59             |
| Gemini-Pro                                   | SoM + Caption + Image        | 6.04             |
| Gemini-Pro                                   | SoM + Caption + Image        | 5.71             |
| Gemini-Pro + BLIP-2-T5XL                     | AxTree + Caption             | 3.85             |
| GPT-3.5 + BLIP-2-T5XL                        | AxTree + Caption             | 2.97             |
| GPT-3.5 + LLaVa-7B                           | AxTree + Caption             | 2.75             |
| GPT-3.5                                      | AXTree                       | 2.20             |
| Gemini-Pro                                   | AXTree                       | 2.20             |
| Mixtral-8x7b + BLIP-2-T5XL                   | AxTree + Caption             | 1.87             |
| Mixtral-8x7b                                 | AXTree                       | 1.76             |
| Llama-2-70B                                  | AXTree                       | 1.10             |
| IDEFICS-80B-Instruct                         | SoM + Caption + Image        | 0.99             |
| IDEFICS-80B-Instruct                         | AXTree + Caption + Image     | 0.77             |
| Llama-2-70B + BLIP-2-T5XL                    | AxTree + Caption             | 0.66             |
| CogVLM                                       | SoM + Caption + Image        | 0.33             |
| CogVLM                                       | AXTree + Caption + Image     | 0.33             |

VisualWebArena Leaderboard

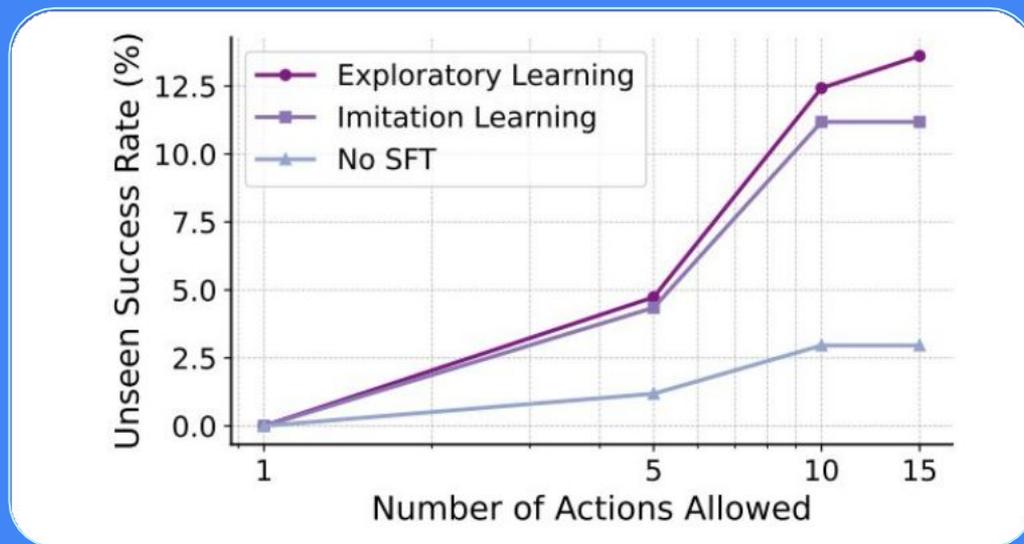
| Rank | Model   | Details                        | Score |
|------|---|--------------------------------|-------|
| 1    | learn-by-interact w/ Claude-3.5-sonnet<br>Anonymous authors<br>Under Review, '24                          | t=0                            | 22.5  |
| 2    | <b>R-MCTS w/ GPT-4o (max 60 tree nodes)</b><br>Columbia University & Microsoft Research<br>Yu et al., '24 | —                              | 16.6  |
| 3    | GPT-4<br>OpenAI<br>OpenAI, '23  | t=1.0, top-p=0.9<br>len = 128k | 12.24 |
| 4    | GPT-4o<br>OpenAI<br>OpenAI, '24   | t=1.0, top-p=0.9<br>len = 32k  | 11.36 |
| 5    | GPT-4 Vision (0409)<br>OpenAI<br>OpenAI, '23  | t=1.0, top-p=0.9<br>len = 32k  | 10.82 |
| 6    | learn-by-interact w/ Gemini-1.5-pro<br>Anonymous authors<br>Under Review, '24                             | t=0                            | 10.3  |
| 7    | Owen-Max<br>Owen<br>Owen Team, '24  | t=1.0, top-p=0.9<br>len = 32k  | 6.87  |
| 8    | GPT-4o-mini<br>OpenAI<br>OpenAI, '24  | t=1.0, top-p=0.9<br>len = 128k | 6.21  |
| 9    | Gemini-Pro-1.5<br>Google<br>Gemini Team, Google, '24  | t=1.0, top-p=0.9<br>len = 128k | 4.81  |
| 10   | Owen-vl-Max-0809<br>Owen<br>Owen Team, '24  | t=1.0, top-p=0.9<br>len = 32k  | 4.32  |

OSWorld Leaderboard

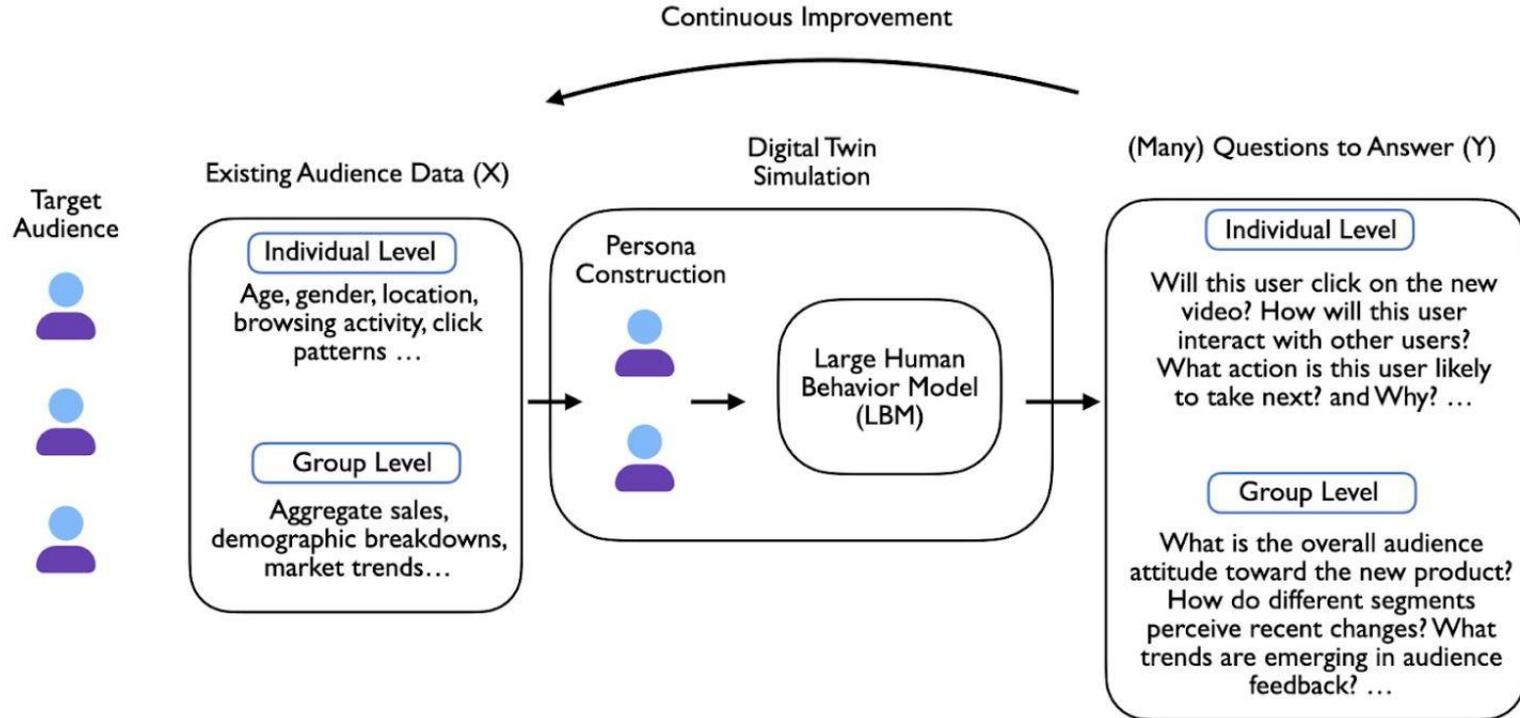
# AI Agents Further Improve after Distillation

GPT-4o after exploratory learning

exhibits compute **scaling properties** without augmenting with search algorithm!



# Digital Twins



# DAPLab: Data, Agents, Processes Lab

13 top faculty

- The best NYC AI lab across disciplines
- We chair the top conferences (NeurIPS, ICML, ICLR, AISTATS)
- We win best papers, top AI awards, and run national AI centers

5 year agenda

- New **agent-ready stack** for enterprise automation
- Open source systems, benchmarks, applications
- Work closely w/ handful of partners across industries

**Safe, Reliable, Efficient → Trust → Adoption**

HAI



Systems



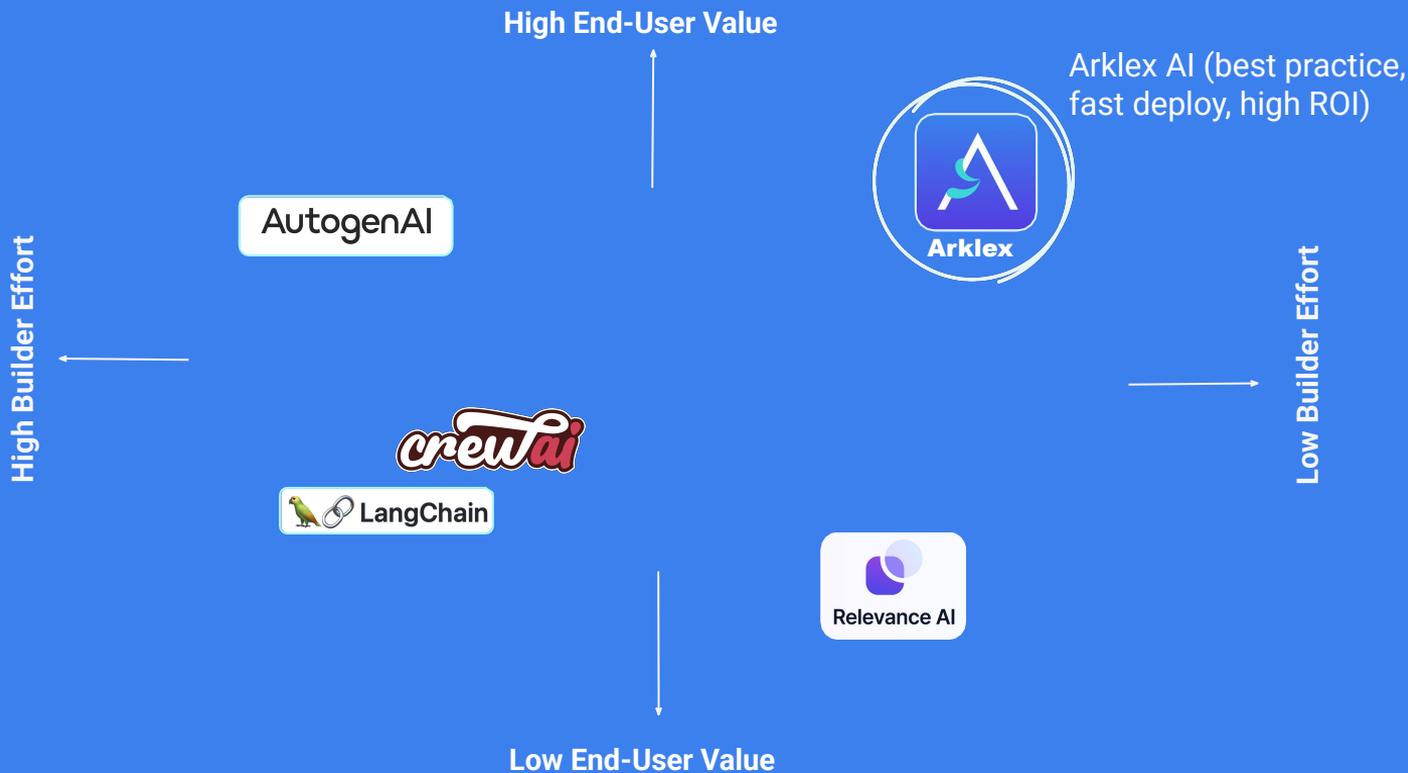
Applied AI



Algorithms



# Competition





ICPC International Collegiate Programming Contest

# The 2025 ICPC World Finals **Baku**

31 August – 5 September 2025 // hosted by **ADA University**

## Thank you!

- **Reminder: ICPC Library on Wednesday, September 3, 2025!**

