A Collaborative Proposal to the
NSF Experimental Expeditions Program

MCAI 2.0

# Next-Generation Model Checking and Abstract Interpretation with a Focus on Systems Biology and Embedded Systems


Edmund M. Clarke (Lead PI CMU)
Rance Cleaveland (PI U. Maryland)
Patrick Cousot (Co-PI NYU)
Bud Mishra (Co-PI NYU)









# Our Vision

MCAI 2.0

To gain fundamental new insights into the **emergent behaviors** of complex biological and embedded systems through the use of **revolutionary**, highly **scalable**, and fully **automated** modeling and analysis techniques.

# Our Goals

MCAI 2.0

- **Develop MCAI 2.0** - Next-Generation Model Checking and Abstract Interpretation
- **Apply MCAI 2.0** to Challenges Problems in complex biological and embedded systems
- **Create IMDECS** - The **I**nstitute for **M**odel **D**iscovery and **E**xploration of **C**omplex **S**ystems

# Model Checking

MCAI 2.0

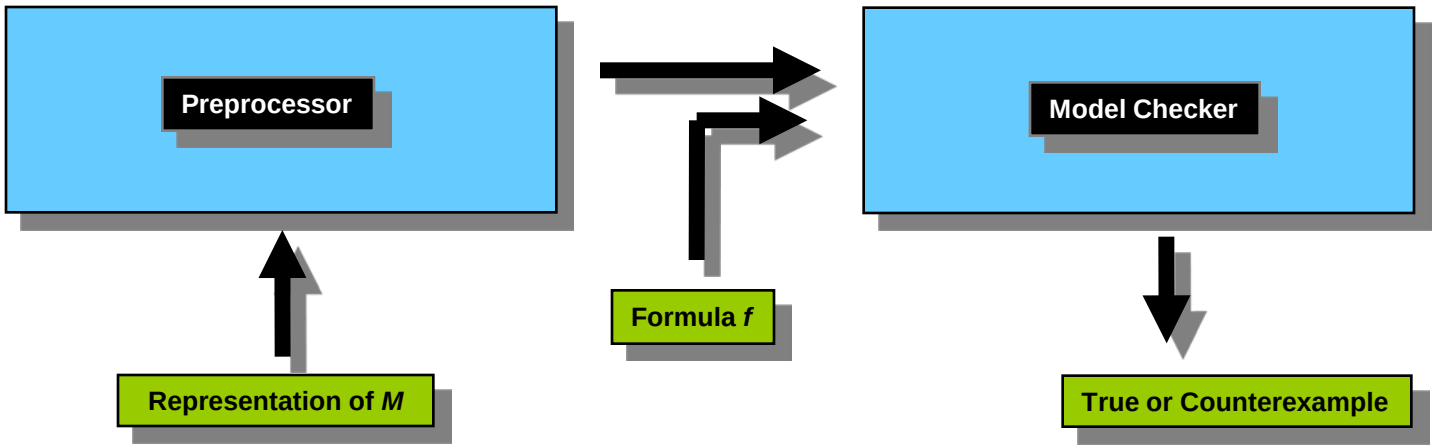
The Model Checking Problem (**Clarke** and Emerson 81):

Let ***M*** be a **state-transition graph**

Let ***f*** be a **formula of temporal logic**

e.g., a U b means "**a** holds true Until **b** becomes true"


Find all states ***s*** of ***M*** such that $M, s \models f$



# Advantages of Model Checking

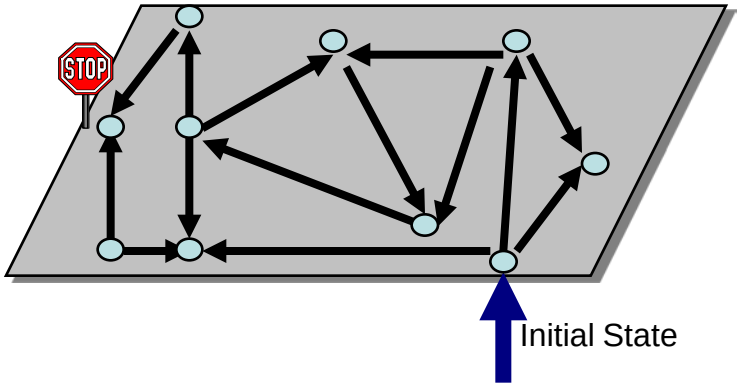
MCAI 2.0

- No Proofs!  (Algorithmic not Deductive)
- Fast  (compared to other rigorous methods)
- No problem with partial specifications
- Diagnostic counterexamples

# Advantages of Model Checking

MCAI 2.0

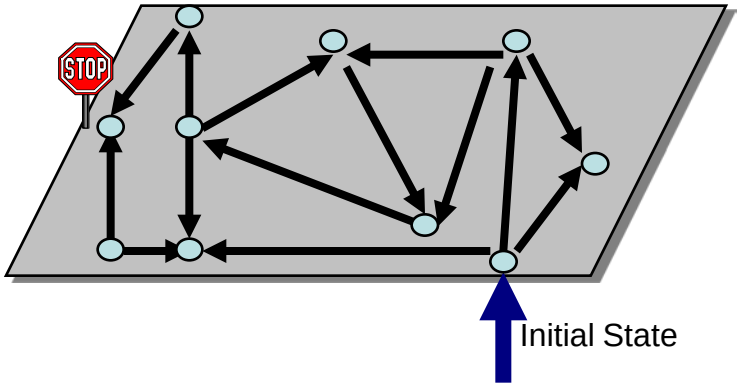
- No Proofs!  (Algorithmic not Deductive)
- Fast  (compared to other rigorous methods)
- No problem with partial specifications
- Diagnostic counterexamples



# Advantages of Model Checking

MCAI 2.0

- No Proofs!  (Algorithmic not Deductive)
- Fast  (compared to other rigorous methods)
- No problem with partial specifications
- Diagnostic counterexamples




Safety Property:

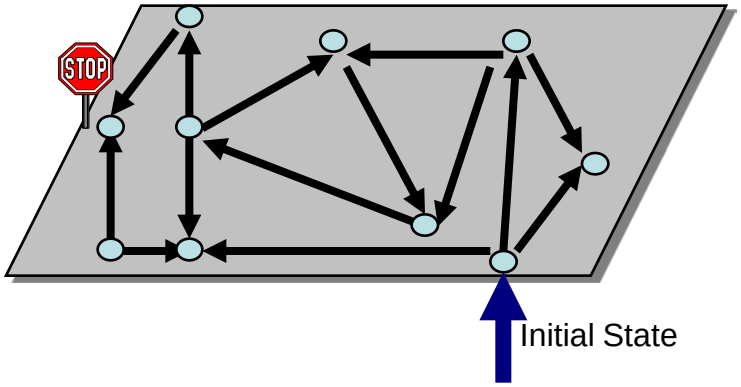
bad state 

unreachable

# Advantages of Model Checking

MCAI 2.0

- No Proofs!  (Algorithmic not Deductive)
- Fast  (compared to other rigorous methods)
- No problem with partial specifications
- Diagnostic counterexamples




Safety Property:

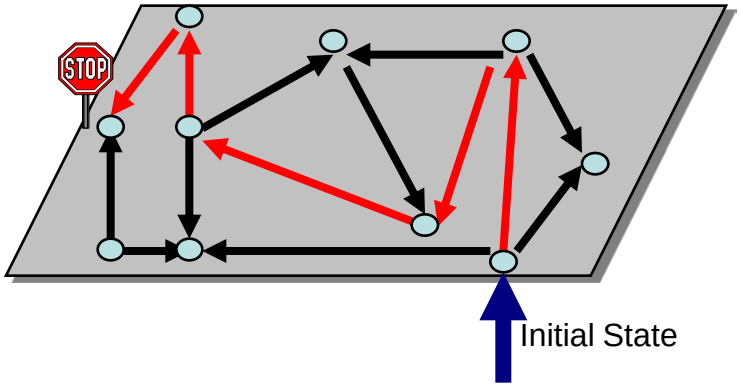
bad state  unreachable


Counterexample

# Advantages of Model Checking

MCAI 2.0

- No Proofs!  (Algorithmic not Deductive)
- Fast  (compared to other rigorous methods)
- No problem with partial specifications
- Diagnostic counterexamples




Safety Property:

bad state  unreachable


Counterexample

# Many Industrial Successes

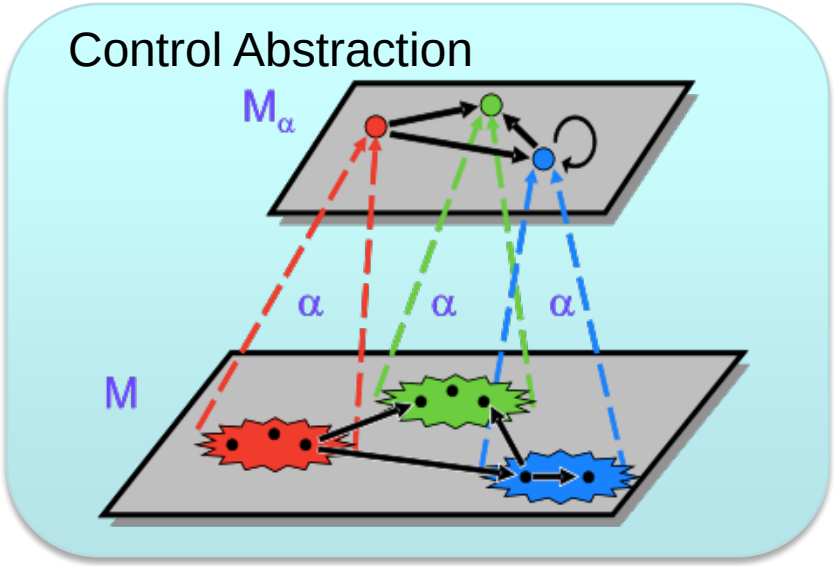
MCAI 2.0

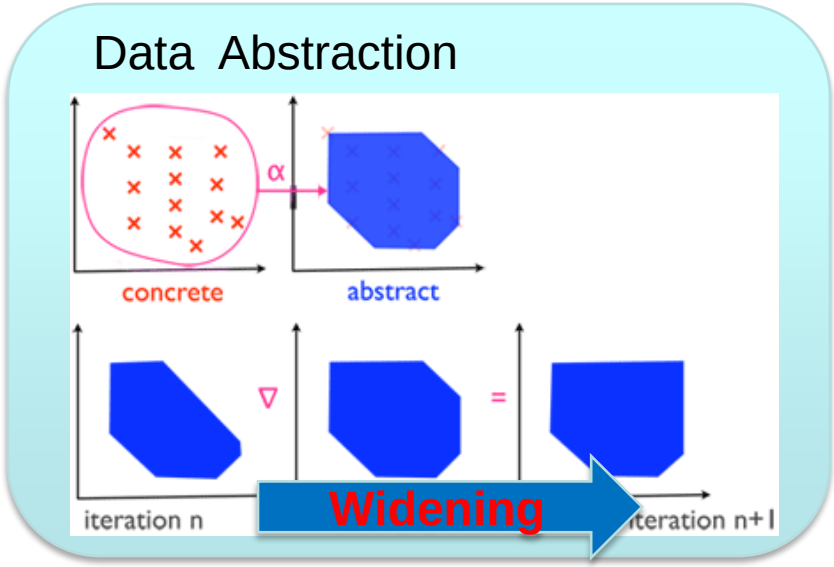

- Try 4195835 – 4195835 / 3145727 * 3145727.
  - In 94' Pentium, it doesn't return 0, but 256.
- Intel uses the SRT algorithm for floating point division. Five entries in the lookup table are missing.
- Cost: $500 million
- Xudong Zhao's Thesis on Word Level Model Checking

# Abstract Interpretation

MCAI 2.0

- Algebraic theory of **sound approximation** of the fixpoints of functionals obtained from computer programs
- Developed by **Cousot** & Cousot in 1977
- Abstracts the **concrete semantics** of a system into a simpler **abstract semantics**



# Features of Abstract Interpretation

MCAI 2.0

- **Automatic extraction of correct information** about the possible executions of complex systems
- Yields **precise** infinite abstract models
- **Accelerate convergence** to the fixpoint of a functional using a widening operator
- **Scalability!** e.g., **A380** primary flight control system:
  - 1 million lines of C code
  - 34 hours to analyze
  - 0 false alarms

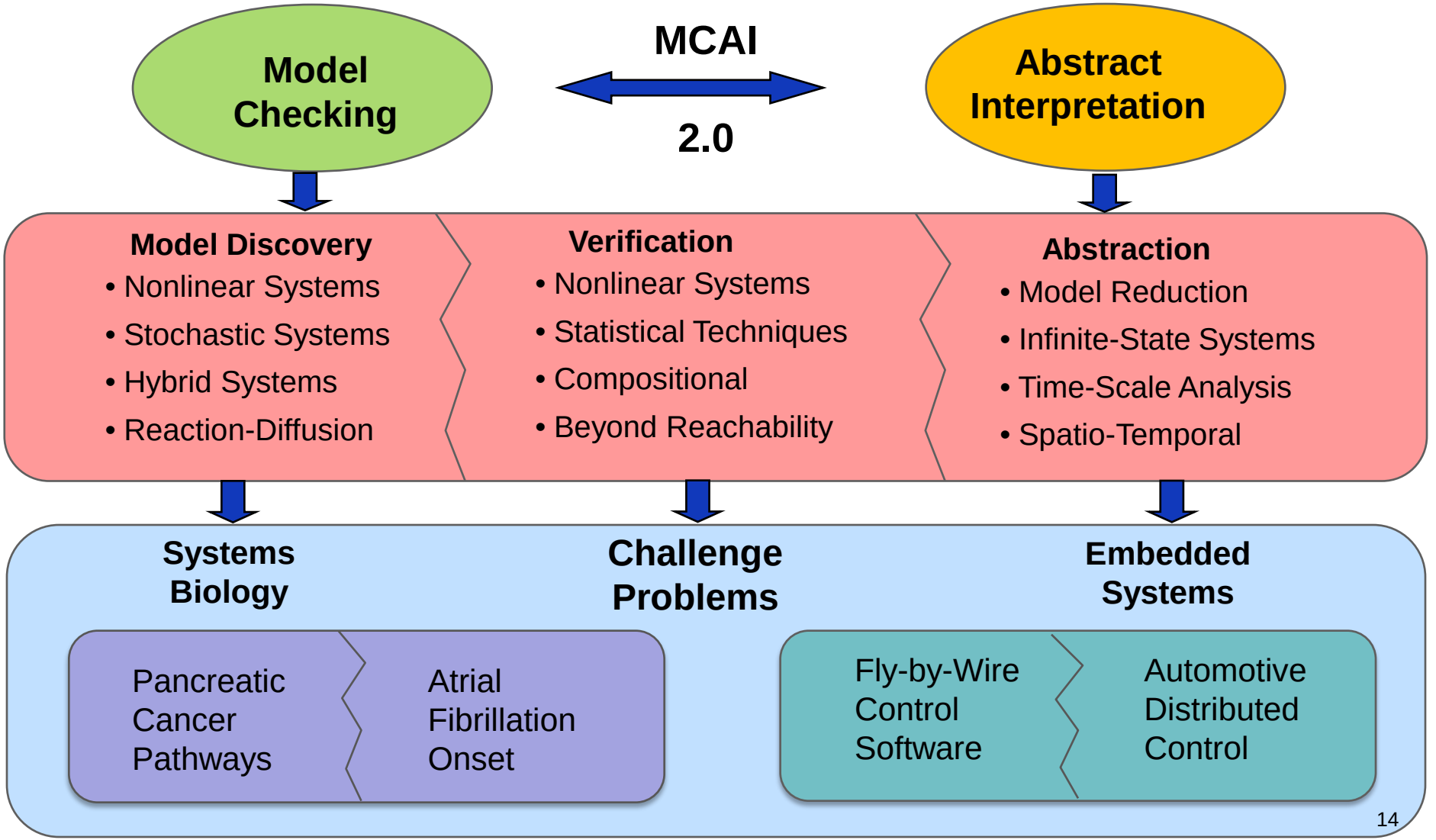
# This Expedition …

MCAI 2.0

- **Rethinking** and **deep integration** of Model Checking and Abstract Interpretation
- Driven by the **centrality of computational modeling** **in** science & engineering
- Application to complex **biological** and **embedded** systems
- Many of the same modeling, analysis, and verification techniques applicable to one domain **applicable to the other as well**

# MCAI 2.0 Strategic Plan

MCAI 2.0



# Primary Challenge: Scalability

MCAI 2.0

## Key Scalability Issues:

**Spatial Distribution**

**Stochastic Behavior**

**Highly Nonlinear Behavior**

**Mixed (Hybrid) Continuous-Discrete Behavior**

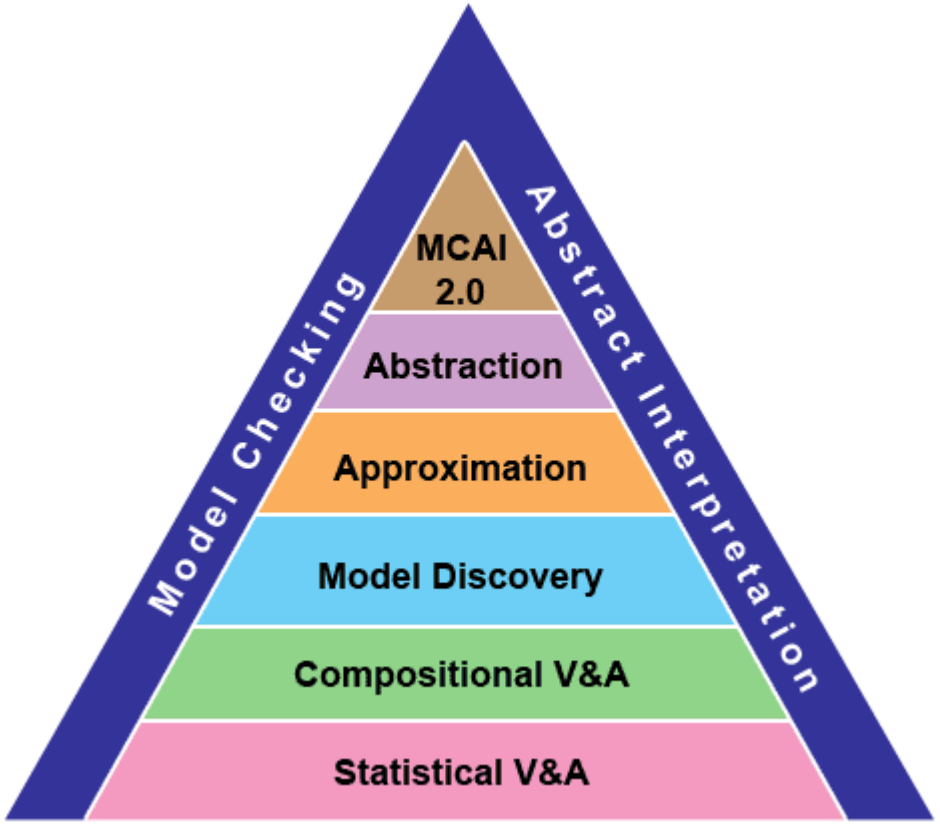
**Vast Numbers of System State Variables & Components**

Complex Biological & Embedded Systems can exhibit any combination of these features

# Research Pillars to Achieve Scalability

MCAI 2.0

- **Abstraction:** Counterexample guided abstraction refinement + Abstract Interpretation + spatial abstraction + …
- **Approximation:** time-scale separation + differential invariants + hybrid systems + ...
- **Model Discovery:** state estimation + hybrid system identification + sequential Monte Carlo + …
- **Compositional V&A:** system decomposition + assume-guarantee + time-scale decomposition + …
- **Statistical V&A:** semi-exhaustive simulation + Bayesian model checking + Monte Carlo model checking + …



# Challenge Problems

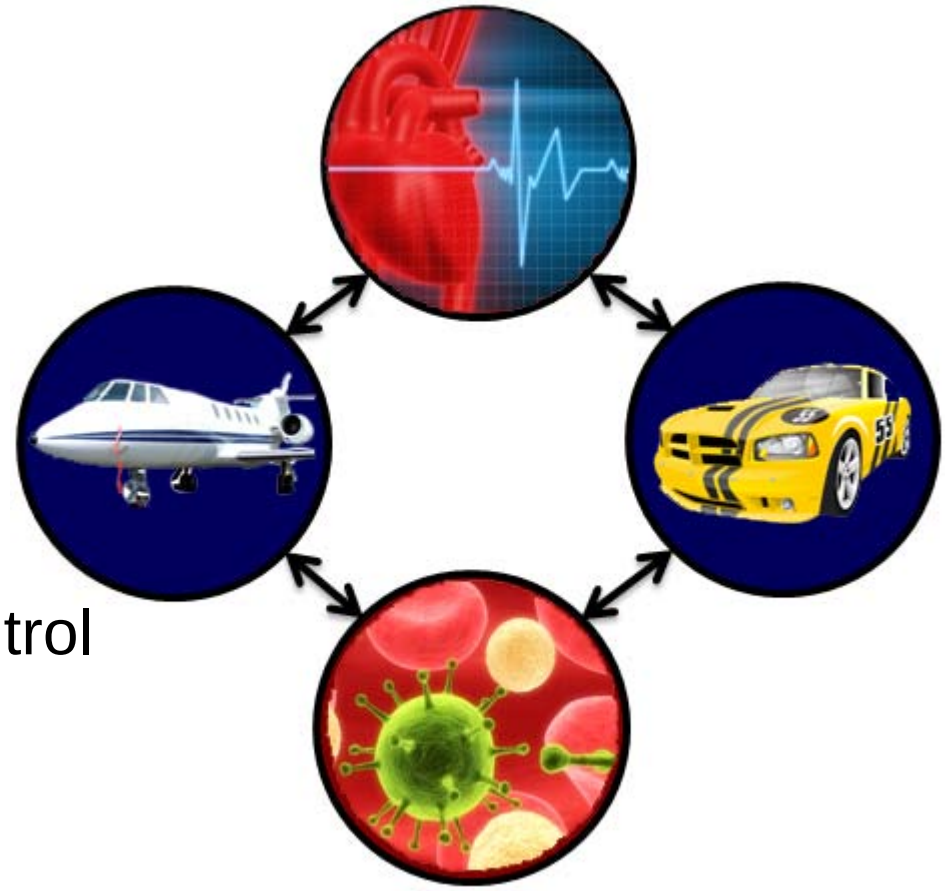
MCAI 2.0

***Systems Biology***

- Pancreatic Cancer
- Atrial Fibrillation

***Embedded Systems***

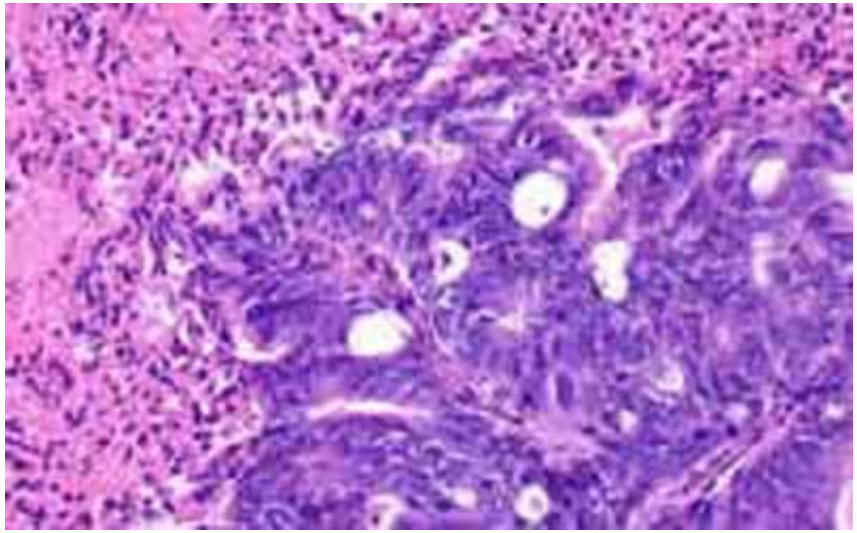
- Distributed Automotive Control
- Aerospace Flight Software



# Pancreatic Cancer

MCAI 2.0

- Cancer is a **failure mode** of the interactions between signaling pathways
- 4th leading cause of cancer death in the US and Europe
- Almost no progress in diagnosis and treatment in the past 30 years



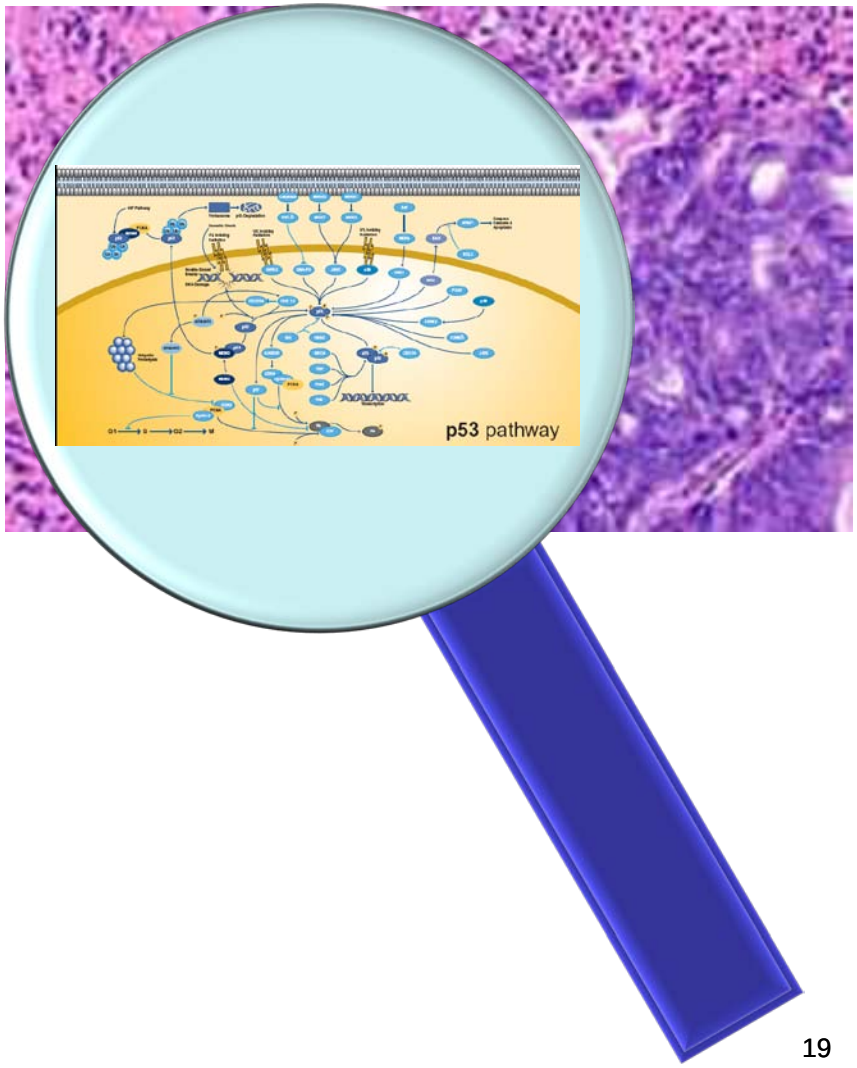
*Healthy and diseased pancreas cells*

**New insights into the dynamics of this deadly disease are urgently needed!**

# Pancreatic Cancer

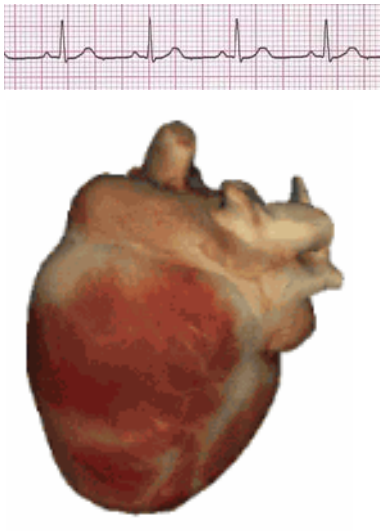
MCAI 2.0

- No animal model, so computational models are necessary
- Models provided by cancer experts at **TGEN** (Translational Genomics)
- We will build new analysis and verification tools
- TGEN collaborators will use tools to better understand cancer dynamics



# Atrial Fibrillation

MCAI 2.0



Normal Rhythm

ECG SIMULATION

Atrial Fibrillation

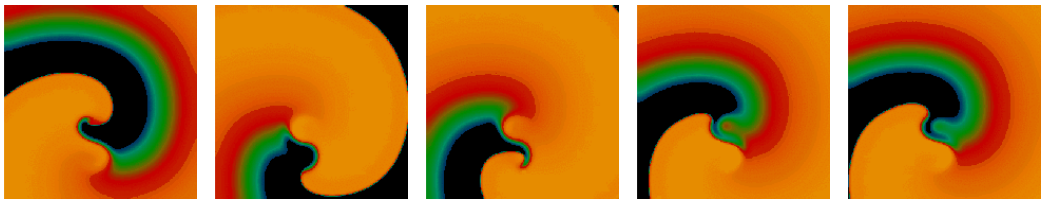
- **Most commonly diagnosed** cardiac arrhythmia
- **Risk increases with age:** >20% for people over 80 years old
- **10 million projected** to have AF by 2050
- **AF is responsible for** 15-20% of all strokes

# Atrial Fibrillation: Cardiac Modeling Needs MCAI 2.0

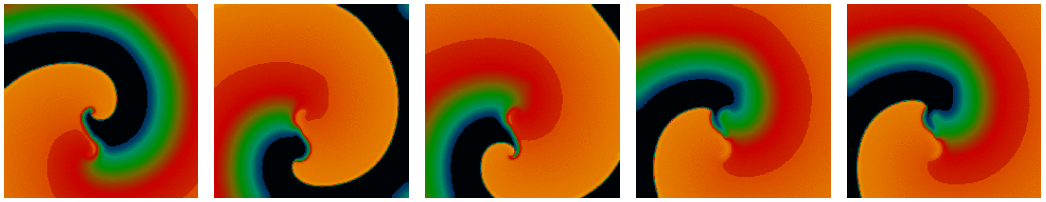
MCAI 2.0

**MCAI 2.0** can lead to reduced models that **improve understanding**, **enhance robustness**, and **avoid non-physiological behavior** while reproducing virtually the same dynamics

- Two electrophysiological models of cardiac cells
- Full model has 17 variables (hundreds of parameters)
- Reduced model has 4 variables (20 parameters)
- MCAI 2.0 can be used to show that both models have virtually the same spiral-wave dynamics



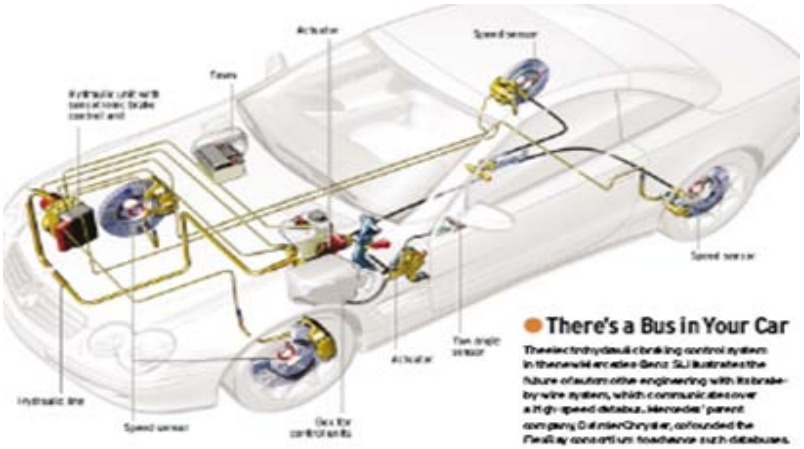
***Full Model**: 5 snapshots during one period of the Ten Tusscher et al. model (17 variables)*

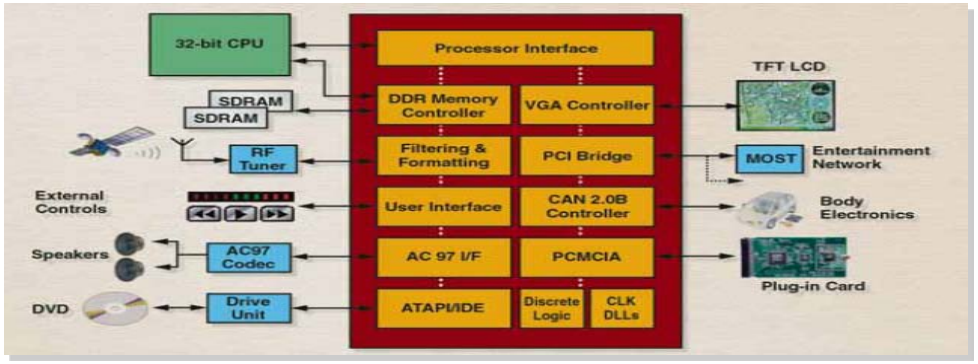


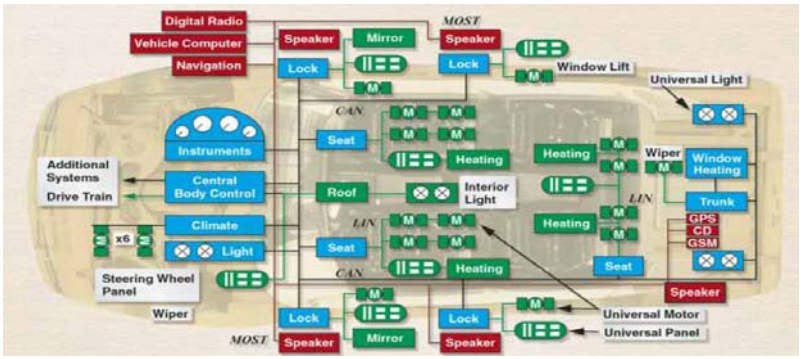
***Reduced Model**: 5 snapshots during one period of the Bueno et al. model (4 variables)*

# Automotive Embedded Systems

MCAI 2.0




Do you trust your car?

# Aerospace Systems: Software Driven!

MCAI 2.0

Mars Polar Lander (1999)
landing-logic error

**Mission Loss**

Spirit Mars Rover (2004)
file-system error

Airbus A380 Flight Deck


Do you trust flight software?

# Embedded Systems Need MCAI 2.0

MCAI 2.0

- **Scalability:** each new Mars mission employs more software than all previous Mars missions together
- **Often no models, only code**: software written in C, sometimes without the help of formal models
- **MCAI 2.0 can be used to** extract abstract models from source code, analyze generated models, drive C-code testers, …
- **MCAI 2.0 will also be used to** analyze specifications and models for design

# Multi-institutional, Multi-disciplinary Research Team: 7 Universities, 2 Colleges, 14 Departments/Schools

MCAI 2.0

Gerard Holzmann
LaRS
NASA JPL

Bud Mishra
CS & SOM
NYU CSHL

Patrick Cousot
CS
NYU

Amir Pnueli
CS
NYU

Ed Clarke
CS
CMU

Bruce Krogh
ECE
CMU

Chris Langmead
CS
CMU

Andre Platzer
CS
CMU

James Faeder
SOM
U. Pittsburgh

Klaus Havelund
LaRS
NASA JPL

## Institute for Model Discovery and Exploration for Complex Systems (IMDECS)

Nancy Griffith
Math & CS
CUNY

Radu Grosu
CS
SUNYSB

Scott Smolka
CS
SUNYSB

James Glimm
Applied Math
& Statistics
SUNYSB

Flavio Fenton
Biomedical Sci.
Cornell

Robert Gilmour
Biomedical Sci.
Cornell

Rance Cleaveland
CS
U. Maryland

Tongtong Wu
Public Health
U. Maryland

Steve Marcus
ECE ISR
U. Maryland

# Team Member Highlights

MCAI 2.0

- **Edmund Clarke:** co-inventor of Model Checking, co-recipient of 2007 ACM Turing Award, 1998 ACM Paris Kanellakis Theory and Practice Award, member of National Academy of Engineering
- **Amir Pnueli:** recipient of the 1996 ACM Turing Award for introducing temporal logic into computer science, many honorary degrees
- **Patrick Cousot:** co-inventor of Abstract Interpretation, received 2008 Humboldt Research Award, 1999 Laureate of the CNRS silver medal, 2006 EADS Scientific Grand Prix
- **Gerard Holzmann:** recipient of the 2001 ACM Software System Award and 2006 ACM Paris Kanellakis Theory and Practice Award, member of National Academy of Engineering
- **Jim Glimm:** awarded 2002 National Medal of Science for his work in shock wave theory & other cross-disciplinary fields in mathematical physics, member of National Academy of Sciences

# Institutional Support

MCAI 2.0

Strong institutional support from all member institutions:

- From top-level management (President, Provost)
- Organizational support & **matching funds ($1,200,000- $1,500,000)**
- Researchers eager to share data, problems and techniques, and anxious to take down scalability barriers posed by CPs

IMDECS to be housed in the new Gates Hillman Center at CMU

# Integrated Management for IMDECS

MCAI 2.0

| | | |
|---|---|---|
| **Director** | Edmund M. Clarke | (Lead PI, Carnegie Mellon University) |
| **Associate Director** | Amir Pnueli | (PI, New York University) |
| **Deputy Director** | Rance Cleaveland | (Co-PI, University of Maryland) |
| **Deputy Director** | Nancy Griffeth | (Co-PI, Lehman College, CUNY) |
| **Deputy Director** | Scott Smolka | (Co-PI, Stony Brook University) |

Plan, coordinate, and evaluate exciting and aggressive research, education, and outreach program. Disseminate MCAI 2.0 technology.

**Executive Committee**: PI's + Co-PI's meet once a month.

**Advisory Board**: Prominent researchers in MCAI-related fields, both from the application side (i.e., Biology and Engineering: 2 people) and the computer science side (2 people). Will also invite an NSF representative. Meets once a year.

**Collaboration Plan**: Through interdisciplinary Challenge Problems (CPs, shared personnel), executive committee meetings, exchange students and faculty, three-day annual meetings, meetings at regular CS conferences.

# Challenge Problems: The Driving Forces Behind our Expedition

MCAI 2.0

| | CP Name | Faculty Team |
|---|---|---|
| 1. | **Signaling Pathways in Pancreatic Cancer** | Clarke, **Faeder**, **Langmead**, Mishra |
| 2. | **Fibrillation Onset in Cardiac Tissue** | **Fenton**, Gilmour, Glimm, Grosu, Mishra, Smolka |
| 3. | **Distributed Automotive Control** | Clarke, **Cleaveland**, Griffeth, Krogh, Marcus, Platzer, Pnueli, Smolka |
| 4. | **Aerospace Control Software** | Clarke, Cleaveland, Cousot, **Havelund**, Holzmann, Marcus, Platzer |

# Nurturing Synergy

MCAI 2.0

## Application/Synthesis Facilitators

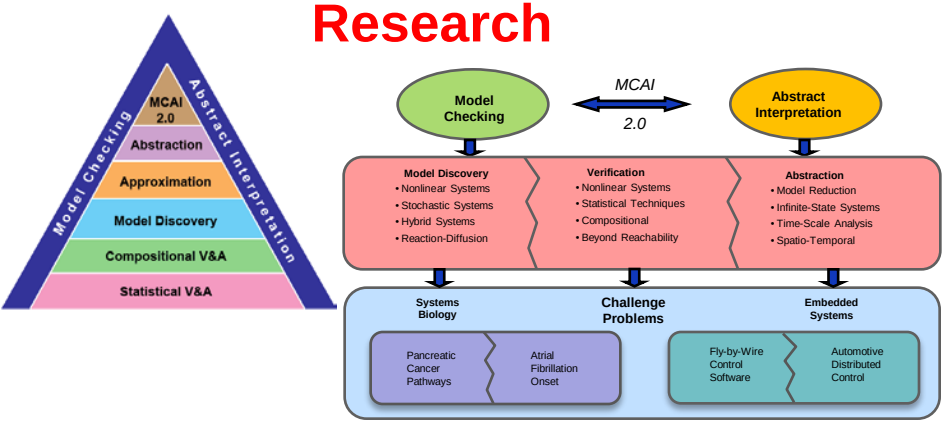
| | | MCAI 2.0 Themes | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Model Checking | Abstract Interpretation | Model Discovery | Compositional V&A | Statistical V&A | Dissemination, Outreach, & Diversity | |
| Application Areas | Computational Biology and Cancer | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Faeder, Mishra, Langmead, Wu |
| | Excitable Cells and Arrhythmia | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Fenton, Gilmour, Glimm |
| | Embedded Systems and Software | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Havelund, Holzmann, Krogh |
| | | Clarke, Cleaveland, Holzmann | Cousot, Platzer, Pnueli | Griffeth, Grosu | Cleaveland, Holzmann | Langmead, Marcus, Smolka | Griffeth, Mishra, Smolka | |
| | | Synthesis Facilitators | | | | | | |

Application Facilitators

# Institute Activities

MCAI 2.0

## Research




## Building Research Community

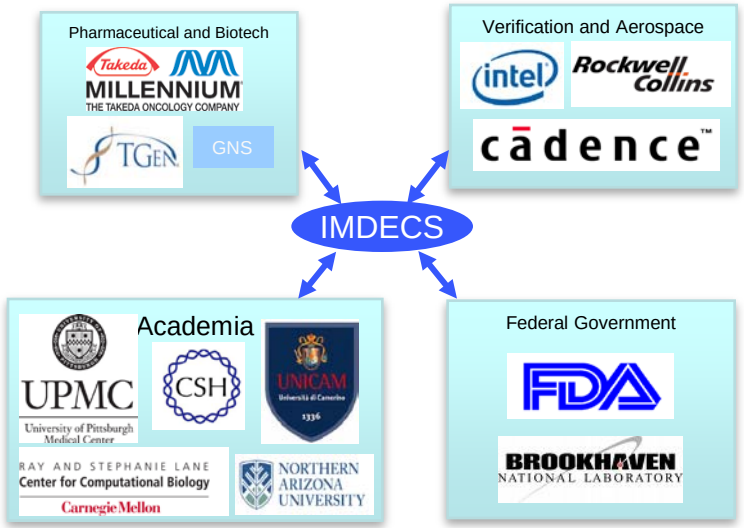
- Web Portal
- Panels & tutorials at major conferences
- Models, Tools, Data
- Annual symposium
- Host visiting Scientists

IMDECS

## Education

- Postdocs
- Doctoral students
- Undergraduate Research Opportunities (OurCS)
- Minority-Focused Intersession Workshops
- Complex Systems Science & Engineering Educational Program

## Broader Impacts



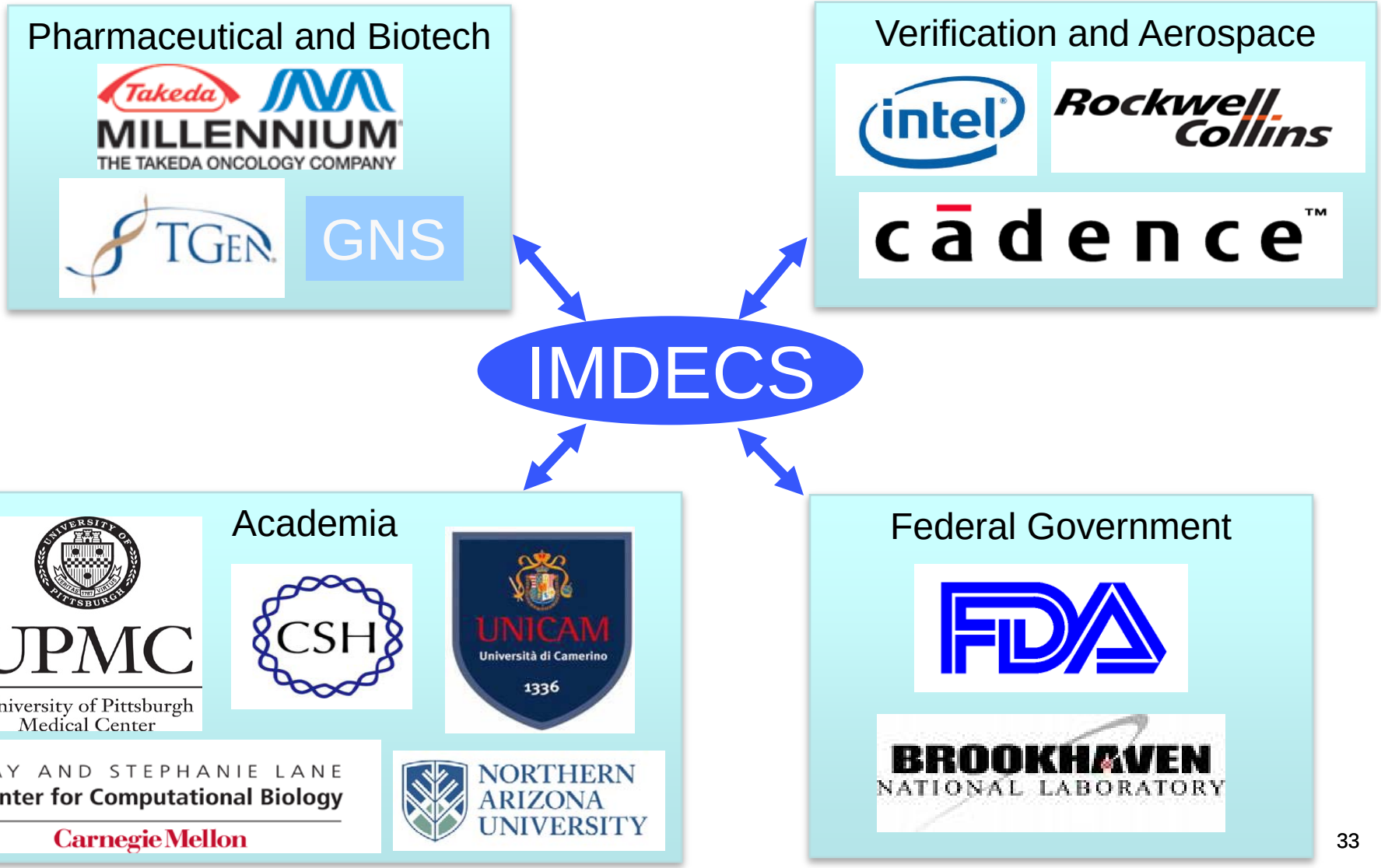
# Education & Outreach Program Highlights

MCAI 2.0

- Highly ambitious/cross-disciplinary **Complex Systems Science & Engineering** (CSSE) educational program
  - Subdisciplines in **BioSystems Science & Engineering** (BSSE) and **Embedded Systems Science & Engineering** (ESSE)
  - Prepare undergraduate & graduate students to **attack complex integrative problems using interdisciplinary approaches** and to work effectively as part of a multi-disciplinary team
  - **CSSE program will span a number of scientific disciplines** with course offerings in computer science, numerical methods, systems & electrical engineering, biomedical engineering, and physiology & biochemistry
- **Minority-Focused Intersession Workshops** for Undergraduates on Understanding and Analyzing Complex Embedded and Biological Systems
- OurCS: biannual conference providing hands-on **research opportunities for undergraduates**

# Broader Impact – Partners

MCAI 2.0



# Selected Quotes from Partner Institutions

MCAI 2.0

The proposed research directly addresses problems of great importance to the aviation industry.
*Dr. John Borghese, Rockwell Collins*

I believe the field [cardiac modeling] thus could benefit greatly from the tools that will be developed within your proposal. Model checking is well-positioned to facilitate development of reliable mathematical models that can provide valuable information about the action of drugs and devices on the electrophysiology of the heart.
*Rick A. Gray, FDA*

The verification technologies you will be developing promise to revolutionize the way scientists elucidate the molecular and cellular mechanisms that drive this [Pancreatic Cancer], and other cancers.
*Dr. Arijit Chakravarty, Millennium Pharmaceuticals*

My laboratory at Cold Spring Harbor is very excited by the goals of your proposed Expeditions in Computing project … we will be able to provide considerable help and support in the context of your [Pancreatic Cancer] Challenge Problem.
*Dr. Michael Wigler, Cold Spring Harbor*

Your proposed technologies are urgently needed to fully realize the potential of the computational models of pancreatic cancer being developed here at TGEN in our quest to make individualized therapeutics a reality.
*Dr. Daniel Von Hoff, TGEN*

# Value-Added as an Expedition

MCAI 2.0

- Fundamentally **new intellectual territory** for computer science
- Unique **societal benefits**
- Tremendous potential for **inspiring new groups** (including under-represented minorities) **to choose careers in computer science**
- MCAI 2.0 Research Plan & Challenge Problems requires **critical mass** and **visibility** that cannot be achieved with piece-meal efforts
- Our research proposal is fundamentally **cross-disciplinary** & **cross-pollinating**:
  - CPs require large teams involving both domain scientists and computer scientists
  - Transformative syntheses also across Model Checking & Abstract Interpretation and other verification & analysis approaches

# Summary

MCAI 2.0

**Expedition will lead to:**

1) **Fundamental intellectual and scientific contributions** driven by questions of scalability of MCAI 2.0 technology to complex systems.

2) **Societal impact** --- significant advances in treatment of pancreatic cancer & atrial fibrillation, and in development of automotive and aerospace control software.

3) **Integration of research, education, and outreach.** New educational program in *Complex Systems Engineering & Science* with BSSE and ESSE sub-disciplines. Research opportunities for undergraduates.

4) **Increasing diversity in computer science.** Bringing in a new generation of students, traditionally not drawn to CS; also, broadening the public image of CS.

**Thank You!**