Expeditions in Computer Augmented Program Engineering

http://excape.cis.upenn.edu/

Cornell, Maryland, Michigan, MIT, Penn, Rice, UC Berkeley, UCLA, UIUC

NSF Site Visit, August 2013
Software Design Methodology

What has changed:
- Programming languages
- Libraries
- Verification technology

What has not changed:
- Programming is done by experts
- Fully specified by conventional programming
- Verification phase is distinct from design

Can we leverage modern analysis tools and increased computing power to revolutionize the task of programming?
Synthesis: A Plausible Solution?

- Classical: Mapping a high-level (e.g. logical) specification to an executable implementation
  - Derivation of programs from constructive proofs
  - Synthesis from temporal logic specifications
  - Refinement in model-based design

- Emerging trends:
  - Integrating different styles of specifications in a consistent executable (e.g. Program Sketching)
  - Programming by examples (e.g. Flashfill for Excel macros)
  - Programmer interaction and feedback (e.g. Program repair)
  - Model Based Design for embedded systems (e.g. Rhapsody for SysML)
ExCAPE Vision

Harnessing computation to transform programming:
Programming made easier, faster, cheaper
Synthesis Tool: Intelligent Assistance

- Designer expresses “what”, possibly using multiple input formats
- Synthesizer discovers new artifacts via integration and completion
- Synthesizer solves computationally demanding problems using advanced analysis tools
- Interactive iterative design
- Integrated formal verification
# Research Organization

<table>
<thead>
<tr>
<th>Tools and Evaluation</th>
<th>Design Methodology</th>
<th>Computational Engines</th>
<th>Education and Knowledge Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Challenge Problems</td>
<td>Apps for Mobile Platforms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multicore Protocols</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Networked Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Robotic Systems</td>
</tr>
</tbody>
</table>
Theme: Computational Engines

1. Reactive Synthesis: From logical specs to finite-state controllers
   - Compositional synthesis (Vardi)
   - Discrete-event systems theory for concurrency control (Lafortune)
   - Synthesis in presence of identifiers/data (Kress-Gazit, Seshia)
   - Automatic generation of environment assumptions (Alur)

2. Syntax-directed synthesis of code snippets
   - Modularity for scalability in Sketch (Solar-Lezama)
   - Synthesis from concrete examples + symbolic constraints (Alur, Martin)

3. Hybrid systems: Control of the physical world
   - Optimal performance of continuous-time controllers (Kress-Gazit)
   - Controllers for linear systems from LTL specs (Tabuada)
   - Theory of robustness for discrete/hybrid systems (Tabuada)
   - Handling nonlinear dynamics for hybrid systems & LTL specs (Kavraki, Vardi)
   - Platform and Contract Based Design for distributed embedded systems (Sangiovanni)

4. Core verification technology
   - Uniform sampling of satisfying assignments (Vardi)
   - Regular functions for string transformations and quantitative analysis (Alur)
   - Natural proofs for automatically verifying data structures (Parthasarathy)
ExCAPE Inspired Activities

- Bridging the gap between research communities in discrete event systems and reactive synthesis (Lafortune, Tripakis, Vardi)

- Application of multiple computational tools in the robotics challenge problem (Kavraki, Kress-Gazit, Pappas, Sangiovanni, Seshia, Tabuada, Vardi)

- Formalization of core computational problem in syntax-directed synthesis of program fragments (Alur, Bodik, Martin, Seshia, Solar-Lezama)

- Syntax-directed synthesis of finite-state reactive controllers and cache-coherence protocols (Alur, Bodik, Martin, Seshia)

- Contracts for automatic synthesis from specifications for embedded systems (Sangiovanni)
Syntax-Directed Synthesis Modulo Theories

Based on input format for SMTLib 2
Problem: Given a formula $\phi$ in an SMT theory with an extra function symbol $f$, and context-free language $L$ for templates, find an expression $e$ in $L$ such that $\phi[f/e]$ is valid.
Theme: Design Methodology

1. Multimodal specifications: Design expressed using multiple formats
   - Sketching: Program with holes + Correctness assertions (Solar-Lezama)
   - Protocol = Skeleton + Invariants + Example behaviors (Alur, Martin)

2. High-level logical specifications: Specify “what” and not “how”
   - From structured English to LTL to controllers (Kress-Gazit)
   - Logical expressions to DFA and back (Alur, Hartmann)
   - Robustness and stability constraints for control of hybrid systems (Pappas, Tabuada)
   - Contracts for requirement capture and elicitation (Sangiovanni)

3. Interaction between user and synthesis tool
   - How to explain infeasible specifications (Bodik, Kress-Gazit)

4. Platform constraints integrated in system specification and design
   - Platform-based design: framework for expressing high-level algorithms and low-level constraints, with support for abstraction, composition... (Sangiovanni)
   - Modeling and coping with uncertainty of physical environment in design and synthesis of controllers (Kavraki, Vardi)
Challenge Problem: Robot Programming

Visit all rooms

Feasible specification

Unsynthesizable specification

Specification text

Proposition lists

Log window
Challenge Problem: Distributed Protocols

- Design challenging due to asynchronous model of communication
- Cache coherence protocols, Distributed coordination algorithms
- Successful application domain for formal verification / model checking
- Correctness involves both safety and liveness properties
- Proposed solution: Allow programmers flexibility

Protocol = Skeleton based on Extended-Finite-State-Machines
  + High-level requirements
  + Example behaviors
Challenge Problem: Networked Systems

- **Goal:** Automate resource management in networked systems in a safe manner using formal verification and synthesis tools.

- **Internet routing protocols**
  - Emerging new platform: Software Defined Networks (SDN)
  - Manual configuration management error-prone

- **Wireless Control Networks**
  - New challenge: Codesign control and routing to ensure stability and performance
Challenge Problem: Programming for Mobile Platforms

- New problem domain as a replacement for “Concurrent Programming”

- **Goal:** Improve programmer productivity for development of apps
  - Need to adapt to new platforms supported by mobile devices
  - Programmability by end-users can have huge impact

- Ongoing work ([Foster](#) and [Solar-Lezama](#))
  Automatic extraction of executable models of Android platform using Sketch synthesis tool
Synthesis for Computer-Aided Education

- **Emerging opportunity: MOOCs**

- **Challenge: Personalized feedback on assignments**
  - Manual feedback by TAs (not scalable)
  - Grading by peers (not reliable)
  - Evaluation on test cases (how to translate failed tests to errors?)

- **Application for ExCAPE tools for synthesis**
  - Introductory programming assignments *(Solar-Lezama)*
  - Modeling and Scheduling problems in Embedded Systems course *(Seshia)*
  - DFA construction in Theory of Computation *(Alur, Hartmann)*

  see automatatutor.com
How to integrate many tools being developed by ExCAPE researchers
Sketch, Transit, LTLMoP, ...
and by researchers around the world
RATSY, Comfusy, ...

1. Create a catalog and repository of open-source tools (to-do)

2. Infrastructure to aid design of synthesis tools: Rosette (Bodik)

3. VELLVM: Formal framework for reasoning about LLVM intermediate representation and compiler transformations (Martin, Zdancewic)

4. Exchange format to share computational engines and benchmarks
   - Syntax-directed synthesis modulo theories (ongoing)
   - Reactive synthesis (collaboration with outside groups)
Evaluation

- SynthLib format will help to compare computational capabilities of back-end engines for program synthesis

- Challenge: How to evaluate usability of synthesis tools?

  - **Proton**: Declarative framework for multi-touch gestures
    - Extensive user study to evaluate effectiveness (Hartmann)

  - **Ongoing work**: Is feedback from AutomataTutor helpful?
    - User study planned for Fall semester classes (Alur, Hartmann)
Collaborations across Disciplines
Collaboration Success Story: Summer School

- **ExCAPE Summer School**
  - June 12 – 15, 2013 at UC Berkeley

- Participation of ~ 100 students from at least 3 areas
  - Programming languages
  - Formal methods and verification
  - Control and cyber-physical systems

- **Organizers:** Bodik, Lafortune, Zdancewic

- **Tutorials:** Bodik/Torlak, Tabuada, Vardi

- **One-hour talks:** Alur, Lafortune, Seshia, Solar-Lezama, Tripakis
Collaboration Success Story: AutomataTutor

- **Automated Grading and Feedback for DFA Constructions**
  
  Alur, D’Antoni, Gulwani (MSR), Hartmann, Kini (UIUC), Viswanathan (UIUC)

  Paper in Proc. IJCAI, 2013
  Another submission planned for ACM CHI

  Underlying technology: logic (MSO) - automata connection and translations

- Web-site automatatutor.com public and free

- In-class use and controlled experiment for understanding the effectiveness of feedback planned for Fall 2013
  
  - Theory of Computation course at Penn and UIUC
  - Total sample of 350 students
Collaboration Success Story: Bridging The Gap

- Bridging the gap between reactive synthesis and supervisory control
  - Ehlers, Lafortune, Tripakis, Vardi
  - Draft paper available to be submitted to WODES 2014

- Connecting two different strands of research
  - Reactive synthesis from CAV community
  - Supervisory control from Discrete-Event-Systems/Control community

- Understanding common threads and differences
  - Theoretical expressiveness
  - Decision procedures and complexity bounds
  - Applications
Collaboration Success Story: Synthesis for Robotics

- Working group focused on making robot programming easier
  Kavraki, Kress-Gazit, Lafortune, Pappas, Sangiovanni, Seshia, Tabuada, Vardi
  + multiple post-docs, students

- Connecting robotics with reactive synthesis, model-based design, hybrid systems
  - Joint papers in diverse conferences: CAV, HSCC, ICRA, RSS

- Organization of RSS workshop on formal methods for robotics and automation, Berlin, June 2013
  Kavraki, Kress-Gazit

- Working group meeting planned for November at Rice
Collaboration Success Story: Protocol Synthesis

- **TRANSIT**: Synthesis of distributed protocols
  - Project at Penn started in early 2012
  - Collaboration between Alur (Formal Methods) and Martin (Comp Architecture)

- **Design methodology**: Multi-modal specifications
  - Inspired by Sketch (Bodik, Solar-Lezama)

- **Computational engine**
  - Inspired by Counter-example guided inductive synthesis (Seshia, Solar-Lezama)
  - Found new instantiation of CEGIS

- **Emerging collaboration**: Tripakis (synthesis of state in alternating-bit protocol)

- **New project**: SynthLib to connect Rosette, Sketch, TRANSIT
Collaboration Success Story: SynthLib

- Syntax Directed Synthesis Modulo Theories
  - Alur, Bodik, Juniwal, Martin, Raghothaman, Seshia, Singh, Solar-Lezama, Torlak, Udupa
  - To appear in Proc. FMCAD, Oct 2013, with an accompanying tutorial

- Forms the basis for synthesis competition
  - Interchange format being finalized on top of SMTLib 2
  - Benchmarks being collected from multiple sources
  - Competition planned for July 2014 at CAV/FLoC in Vienna
Broader Impacts, Management, Collaboration
Broader Impacts

- Workshops, special sessions at conferences: CAV'13, ACC'13, PLDI'13, RSS'13
- Summer school (June 12-15, 2013, at UC Berkeley)
- Graduate course at Berkeley: Program synthesis for everyone (Fall'12)
- Technology support for online education: Autograder, AutomataTutor
- Open source library: OMPL (see ompl.kavrakilab.org)
- Many invited talks and honors for PIs
- Publications at a broad range of conferences
- Emerging synthesis community: Synthlib competition (July'14)
Breadth of Conferences

Formal Verification
CAV, TACAS, FMCAD...

Control
ACC, CDC ...

Robotics
ICRA, RSS, ...

Programming Languages
POPL, PLDI,...

Cyber-Physical Systems
HSCC, EMSOFT, ...

Logics & Automata
LICS, ICALP, ...

PLDI 2013: 34th ACM Conf. on Programming Languages Design and Implementation

Papers:
1. Specifying protocols with concolic snippets (Alur, Martin)
2. Optimizing database-backed applications with program synthesis (Solar-Lezama)
3. Automated feedback generation for intro. Programming assignments (Solar-Lezama)
4. Formal verification of SSA-based optimizations for LLVM (Martin, Zdancewic)
5. Natural proofs for structure, data, and separation (Parthasarathy)

Workshop: PLOOC 2013
1st Workshop on Prog. Lang. Technology for Massive Open Online Courses
Co-organizer: Solar-Lezama
Collaboration with Industry

- Industrial Advisory Board
  - Fix (Intel), Godbole (Honeywell), Godefroid (Microsoft)
  - Gupta (NEC), Kuehlmann (Coverity), Mosterman (Mathworks)
  - Wegman (IBM), Zave (AT&T)

- Research collaborations with industry researchers
  - HP Labs, Intel, Microsoft, Samsung, Mozilla, GreenArrays

- SRC/DARPA Research Center TERRASWARM

- iCyPhy center at Berkeley (IBM and United Technologies)

- Student internships during Summer’13: Intel, MSR, NEC Labs...
Collaboration with Govt Agencies and Programs

- DARPA HACMS program for design of attack-resilient control systems
- DOE: Compilers for Exascale machines
- NSF Workshop on Future Directions in Formal Methods (Dec 2012)
- Collaboration with other Expeditions: CMACS (CMU), PPM (MIT)
- Collaboration with RiSE center in Austria
Management and Collaboration

- **Goal**: Foster collaboration across disciplines and institutions

- **Executive committee**: Alur, Bodik, Lafortune, Sangiovanni, Vardi

- **Project manager**: Liz Ng

- **Associate Director**: Dana Fisman

- **Frequent meetings**
  - Monthly webinar
  - Face-to-face meeting of all PIs every year
  - Telecons for individual themes/projects
  - Visits by individual PIs to other institutions
Rotating Postdoc Program

- Each ExCAPE postdoc has two mentors, at two different institutions

- Year 2012-13:
  - **Ruediger Ehlers** (Robotics)
    Mentors: Kress-Gazit (Cornell), Seshia (UC Berkeley)

- For the upcoming year:
  - **Xiaokang Qiu** (PhD UIUC), Apps for mobile platforms
    Mentors: Foster (Maryland), Solar-Lezama (MIT)
  - **Indranil Saha** (PhD UCLA), Robotics
    Mentors: Pappas (Penn), Seshia (UC Berkeley)
  - **Christos Stergiou** (PhD UC Berkeley), Multicore protocols
    Mentors: Martin (Penn), Tripakis (UC Berkeley)
New Research Collaborations
Connections Across Themes
Collaboration Success Story: Summer School

- ExCAPE Summer School
  June 12 – 15, 2013 at UC Berkeley

- Participation of ~ 100 students from at least 3 areas
  - Programming languages
  - Formal methods and verification
  - Control and cyber-physical systems

- Organizers: Bodik, Lafortune, Zdancewic

- Tutorials: Bodik/Torlak, Tabuada, Vardi

- One-hour talks: Alur, Lafortune, Seshia, Solar-Lezama, Tripakis
Collaboration Success Story: AutomataTutor

- **Automated Grading and Feedback for DFA Constructions**
  
  Alur, D’Antoni, Gulwani (MSR), Hartmann, Kini (UIUC), Viswanathan (UIUC)

  Paper in Proc. IJCAI, 2013
  Another submission planned for ACM CHI

  Underlying technology: logic (MSO) - automata connection and translations

- Web-site automatatutor.com public and free

- In-class use and controlled experiment for understanding the effectiveness of feedback planned for Fall 2013
  
  Theory of Computation course at Penn and UIUC
  Total sample of 350 students
Collaboration Success Story: Bridging The Gap

- Bridging the gap between reactive synthesis and supervisory control
  
  Ehlers, Lafortune, Tripakis, Vardi

  Draft paper available to be submitted to WODES 2014

- Connecting two different strands of research
  
  - Reactive synthesis from CAV community
  - Supervisory control from Discrete-Event-Systems/Control community

- Understanding common threads and differences
  
  - Theoretical expressiveness
  - Decision procedures and complexity bounds
  - Applications
Collaboration Success Story: Synthesis for Robotics

- Working group focused on making robot programming easier
  Kavraki, Kress-Gazit, Lafortune, Pappas, Sangiovanni, Seshia, Tabuada, Vardi
  + multiple post-docs, students

- Connecting robotics with reactive synthesis, model-based design, hybrid systems
  - Joint papers in diverse conferences: CAV, HSCC, ICRA, RSS

- Organization of RSS workshop on formal methods for robotics and automation, Berlin, June 2013
  Kavraki, Kress-Gazit

- Working group meeting planned for November at Rice
Collaboration Success Story: Protocol Synthesis

- **TRANSIT**: Synthesis of distributed protocols
  - Project at Penn started in early 2012
  - Collaboration between Alur (Formal Methods) and Martin (Comp Architecture)

- **Design methodology**: Multi-modal specifications
  - Inspired by Sketch (Bodik, Solar-Lezama)

- **Computational engine**
  - Inspired by Counter-example guided inductive synthesis (Seshia, Solar-Lezama)
  - Found new instantiation of CEGIS

- **Emerging collaboration**: Tripakis (synthesis of state in alternating-bit protocol)

- **New project**: SynthLib to connect Rosette, Sketch, TRANSIT
Collaboration Success Story: SynthLib

- Syntax Directed Synthesis Modulo Theories
  - Alur, Bodik, Juniwal, Martin, Raghothaman, Seshia, Singh, Solar-Lezama, Torlak, Udupa
  - To appear in Proc. FMCAD, Oct 2013, with an accompanying tutorial

- Forms the basis for synthesis competition
  - Interchange format being finalized on top of SMTLib 2
  - Benchmarks being collected from multiple sources
  - Competition planned for July 2014 at CAV/FLoC in Vienna