DDDAS: Dynamic Data Driven Application Simulations

Craig C. Douglas  
University of Kentucky and Yale University  
douglas-craig@cs.yale.edu  
http://www.dddas.org

and a supporting cast of thousands from two projects:

Martin Cole, Yalchin Efendiev, Richard Ewing, Victor Ginting, Chris Johnson, Greg Jones, Raytcho Lazarov, Chad Shannon, Jenny Simpson  
Janice Coen, Leo Franca, Robert Kremens, Jan Mandel, Anatolii Puhalskii, Anthony Vodacek, Wei Zhao

Supported in part by the National Science Foundation (ITR-DDDAS)
Shasta-Trinity National Forest 1999 Fire
*(only 142,000 acres)*
Data to Drive Application

Where is the fire?

- Use remote sensing data to locate fires, update positions, and find new spot fires.
  - Satellite: thermal wavelengths
  - Airborne
    - AIMR (NCAR operated): Airborne Imaging Microwave Radiometer – clouds cannot hide a fire from one of these.
    - EDRIS (USFS/NASA operated): Visible, near IR, and IR downward scanning – shows fire with respect to topography
    - IR Video cam: look through smoke to find fire clearly.
Data to Drive Application (cont.)

What is the fuel?

- Geographic Information System (GIS) fuel characterization data to specify spatial distribution of fuel.
- Landsat Thematic Mapper (TM) bands -> NDVI (Normalized Difference Vegetation Index) - related to the quantity of active green biomass.
- AIMR - already used for fire mapping. Testing use as a biomass mapper: difference in vertical and horizontal polarizations gives emissivity, vegetation geometry and biomass.
Data to Drive Application (cont.)

◆ What is the terrain like in that area? What small-scale features are there?
  – New topography sets give world topography at 30 arcsec (~1 km), US at 3 arcsec (~100 m).
  – Better local sources might be available.
    ◆ Need to know where possible fire breaks are.

◆ What are the changing weather conditions?
  – Large-scale data (current analyses or forecasts) used for initial conditions and for updating boundary conditions.
How a DDDAS Might Work (Research Mode)

- Use simulations: first use all available data for past (and eventually current) experimental fires to direct collection at crucial times and places.
- Attempt to prove that the prediction of large fire behavior can be far more effective than the traditional method of tracking and intuition.
How a DDDAS Might Work (Operational Mode)

- Human or a sensor (possibly on a satellite) determines a fire has started near locality X.
- Need to determine severity and possible expansion.
- Produce a 48 hour prediction and post it on a public, known web site.
  - While running model at large-scale over a region…
  - Use latest satellite data (or dispatch reconn aircraft with scanners and/or Thermacam) to locate fire boundary.
- Determine communication methods for firefighters.
- Offer advice where to attempt to halt fire spread.
How a DDDAS Might Work (Operational Mode; cont.)

Have application

- Seek out fuel classification data and recent greenness data.
- Collect recent large-scale data (analyses and forecast) for atmosphere-fire model initial and boundary conditions.
- Initialize and spawn smaller-scale domains, telescoping down to the fire area.
- Ignite a fire in the model at observed location.
- Simulate the next Y hours of fire behavior.
- Dispatch forecast to Web site.
Leaky Underground Storage Tanks

Need to develop monitoring and clean up methods.
Bioremediation Strategies

**Macroscale**
- Injection
- Recovery

**Microscale**
- Flow
  - Growth Mechanisms:
    - Attachment
    - Detachment
    - Reproduction
    - Adsorption
    - Desorption
    - Filtration
    - Interaction

**Mesoscale**
- Input:
  - Substrate
  - Suspended Cells
  - Oxygen
Savannah River Site

- Difficult topography
- Highly Heterogeneous Soils
- Saturated and Unsaturated Flows
- Reactions with disparate time scale
- Transient/Mixed Boundary Conditions
Saturated/Unsaturated GCT 1.2 Simulation of the Pressure Field beneath the Old Burial Ground at the Savannah River Site
Need for Simulation

- Develop better understanding of nonlinear behavior
  - Computational Laboratory ↔ Experiments
  - Understand sensitivities of parameters
  - Isolate phenomena then combine

- Scale up information and understanding
  - Microscale → Laboratory → Field

- Obtain bounding calculations

- Develop predictive capabilities
  - Optimization and Control
Identification (Inverse) Problem

- **Determine Suitable Mathematical Model**
- **Estimate Parameters Within Mathematical Model**
Large Scale Interactive Applications on Remote Supercomputers

- Model Development and Formulation
- Coupled Codes with Complex Boundary Conditions
- Numerical Discretization and Parallel Algorithm Development
- MPP Code Development
- Field Testing and Production Runs
- User Environments and Visualization Tools

Need for Interactive tracking and steering and possibly elimination of Human in the Loop
Graphics Pre-Processing

- 3D grid creation and editing
- Material properties
- Initial conditions
- Time dependent boundary conditions
- Multiple views
Graphics Post-Processing

- Multiple vector/scalar fields
- Time animation
- Multiple slices/Iso-surfaces
- Stereo rendering, lighting models
- Overlay images for orientation
- Volume rendering

Hierarchical Representations
Dynamic Data-Driven Application Systems

**Context:**

Dynamic  →  Immediacy, Urgency, Time-Dependency
Data-Driven  →  Feedback loop between applications, algorithms, and data (measured and computed)
Algorithms  →  (focused context) differential-algebraic equations simulation

**Assumptions:**

Need time-critical, adaptive, robust algorithms
Adaptive Dynamic Algorithms

- Optimization/Inverse Problems
- Incorporate Uncertainty
- Data Assimilation (interpolation)
  - Feedback for experimental design
  - Global influence of perturbations
- Sensor embedded algorithms
- Algorithm automatically restarts as new data arrives
  - Pipelining, systemic computation
  - Warm-started algorithms
Adaptive Dynamic Algorithms (cont.)

- Multiresolution capabilities
  - down-scaling / up-scaling
  - model reduction
- Quick, interactive visualization
- Data Mining / Analysis
  - on input as well as output
- Adaptive gridding
- Parallel Algorithms
- Mathematical analysis for problems in which location of boundary conditions is unknown.
Issues of Perturbations from On-Line Data Inputs

◆ Solve:

\[ F(x + \Delta x(t)) = 0 \iff \text{Choice of new approximation for } x \]

- Do not need a precise solve of equation at each step
  ◆ Incomplete solves of a sequence of related models
  ◆ Effects of perturbations (either data or model)
  ◆ Convergence questions?
- Premium on quick approximate direction choices
  ◆ Lower-rank updates
  ◆ Continuation methods
- Interchanges between algorithms and simulations

◆ Fault-tolerant algorithms
Incorporating Statistical Errors

- Are data perturbations within statistical tolerance?
- Sensitivity analysis
- Filters based upon sensitivity analysis
- Data assimilation
- Bayesian methods
- Monte-Carlo methods
- Outliers (data cleaning)
- Error bars for uncertainty in the data

Difficult for coupled, non-linear systems
Knowledge Based Systems

◆ Intelligent Interfaces
  – Intuitive (no manuals needed)
  – Platform Independent
  – Hidden Algorithmic Details
  – Advanced Graphical Object Representation
  – Visualization

◆ Multiple Scales
  – Knowledge detail
  – Adaptive
System Support

- Parallel/Distributed Platforms (including I/O)
- Embedded systems (e.g., programmable logical arrays)
- Quality of Service
  - Fault tolerant computational environment
  - Fault tolerant networking
  - Data vouching
- Prioritization of resources based upon time criticality
  - Resource Brokerage (e.g., National Security)
Parallel Multi-...

- Model
  - Mathematical
  - Physical
- Scale
- Level
- Error analysis

Significant open question: Is there a technique for analyzing problems similar to generalized solutions and Sobolev spaces with our boundary condition lack of knowledge?
From http://www.cnn.com

◆ June 25, 2002. President Bush declares disaster areas. He arrived in Arizona after declaring parts of the state federal disaster areas in the wake of a devastating wildfire that has burned more than 351,000 acres, freeing up $20 million in emergency federal aid. Bush planned to meet with firefighters and area residents and get an aerial view of the massive Rodeo-Chediski fire, which has destroyed at least 375 homes and 16 businesses and displaced 30,000 people. Numerous Arizona residents requested that the U.S. Forest Service be declared a target in the U.S. War on Terrorism.

Picture courtesy CNN