Why Work on Derivative-Free Optimization? Because the Problems are Important and Cool

John Dennis, Rice University

July 2013

Thanks to: My 91 Different Collaborators and my 35 PhD students

Presentation Outline

1. Introduction
   - Target nonlinear optimization problems
2. Examples: Applications of our MADS algorithm with surrogates
3. Things that need to be done that I think we can do
   - Multiobjective Optimization
4. Summary

Presentation Outline

Target optimization problems - 1965-85

There are many practical problems of the standard type, and the NLP community has effective algorithms and software for their solution

BUT

I began to notice that some interesting problems I met in industrial settings were much smaller and nastier than the textbook ones
Greg Shubin, retired Head of Mathematics and Engineering Analysis at Boeing, estimated that 90% of Boeing's problems were...

Target optimization problems - the other 90%

$$
(NLP) \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \Omega, \end{array}
$$

where $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ may be discontinuous, Ω is any subset of $\mathbb{R}^n$ and:

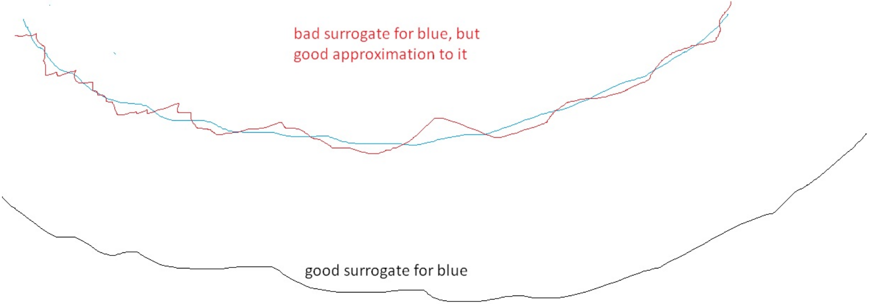
- evaluation of f and determination of membership in Ω are usually the result of a black box computer code with some '`if`'s and '`goto`'s
- the functions are expensive black boxes - secs, mins, days
- the functions may fail unexpectedly even for $x \in \Omega$
- only a few correct digits are ensured
- accurate approximation of derivatives is problematic
- the constraints defining Ω may be nonlinear, nonconvex, nonsmooth and may simply return '`yes/no`'.

Presentation Outline

1. Introduction
2. Examples: Applications of our MADS algorithm with surrogates
3. Things that need to be done that I think we can do
4. Summary

Surrogates are used in all the following examples

A surrogate function is any function used as a contextual stand-in for the actual function. We had much rather have an optimization surrogate that has a minimizer near the actual function minimizer than a surrogate that is a good approximation to the function.



All you need to know about surrogates for this talk

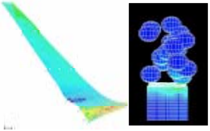
- Sometimes polynomials or RBFs are used as surrogates, but, mostly DACE interpolants are used.
- We use surrogates to suggest *where* to evaluate the actual function, not what value it will have.
- My way to construct convergence theory for methods with surrogates is to say what must be accomplished in each step of an algorithm, not how to accomplish it - and then give a fallback for failures.
- A *theory* should work for surrogates that are just random guesses

Boeing DE Applications

Design Explorer Applications

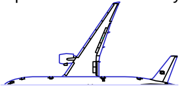
Helicopter Rotor Design

Space Station Power System



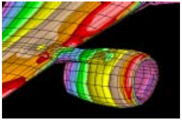
Shot peen forming of wing skins

Aerospike Nozzle



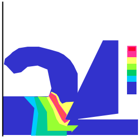
Multidisciplinary wing
planform design

3-D Fighter Aerodynamics



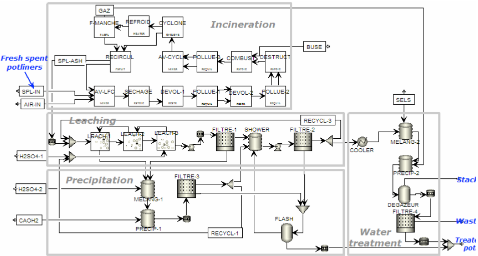
Engine Nozzle
Performance

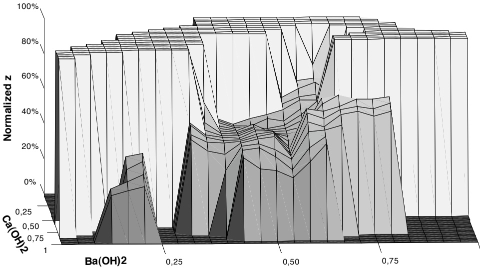
777 Engine Duct Seals



Machining,riveting,and drilling database

Evaluation of f and Ω requires running the ASPEN chemical engineering process simulation software.



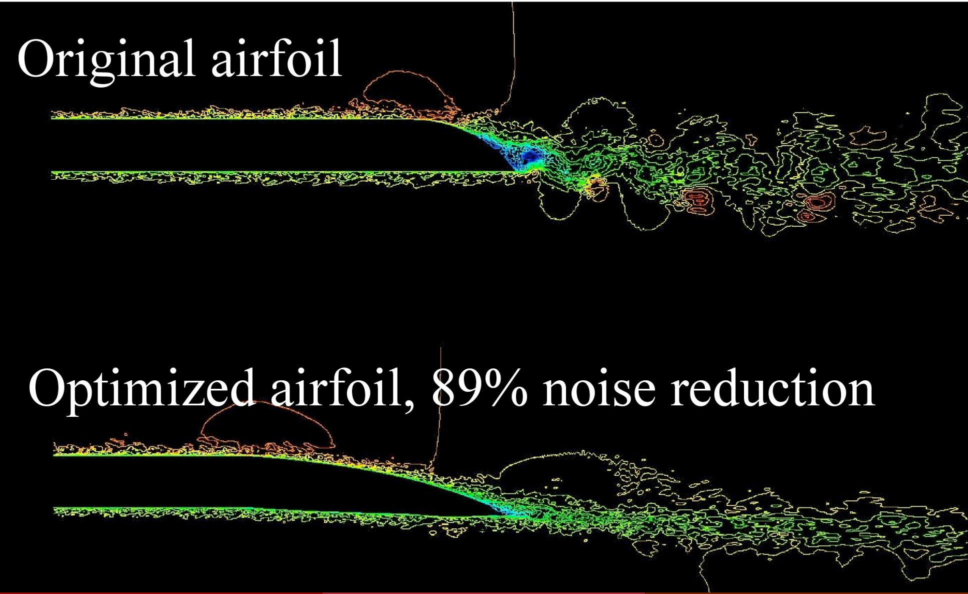


7 variables, 4 black box constraints.

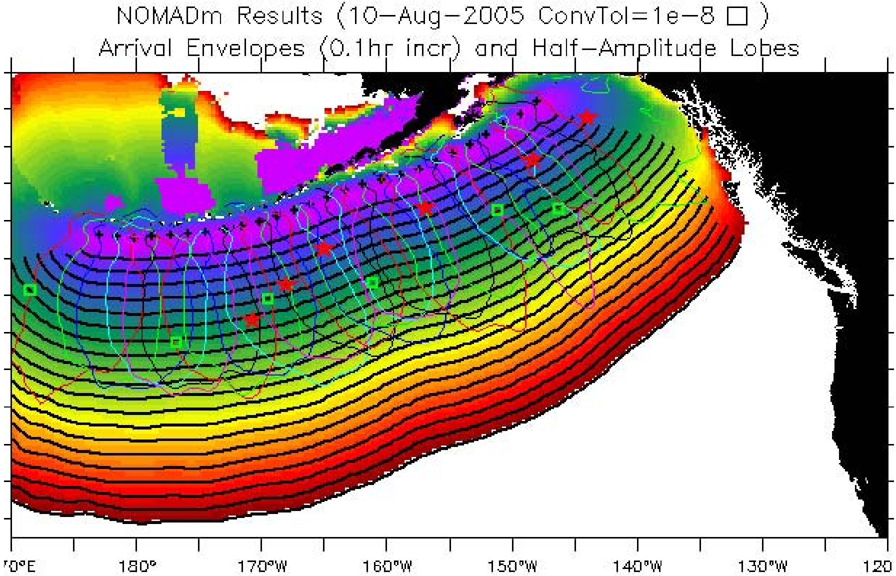
ASPEN fails on 43% of function calls (floor in graph).

Example: Trailing noise reduction - Marsden et al.

A very expensive objective function requiring 3d turbulent flow.

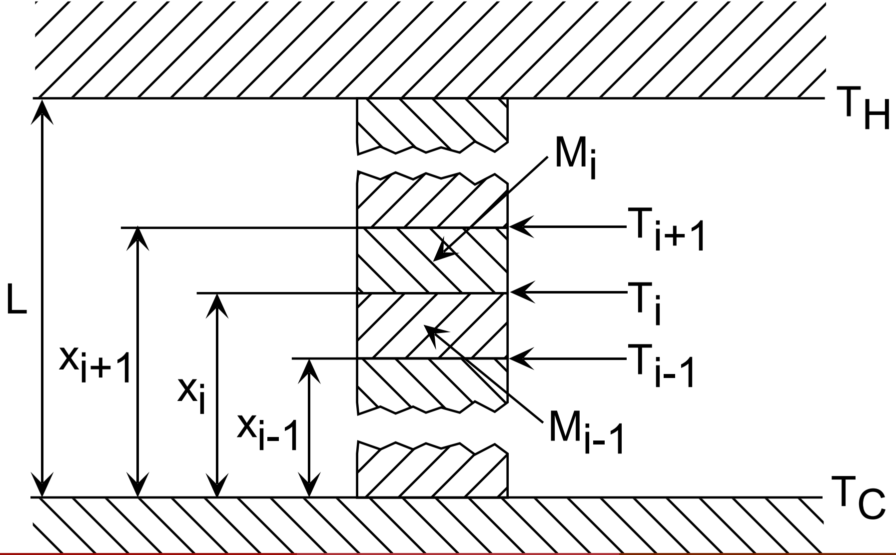


Weird constraints caused by the ocean floor



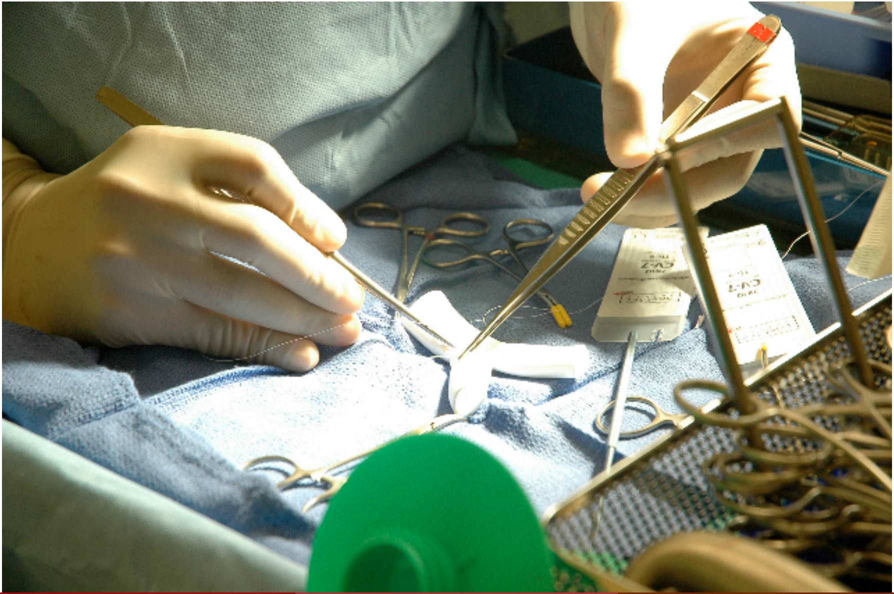
Example:Heat shields

Categorical variables - unrelaxable discrete variables - branch and bound is not practical for these variables in an application setting



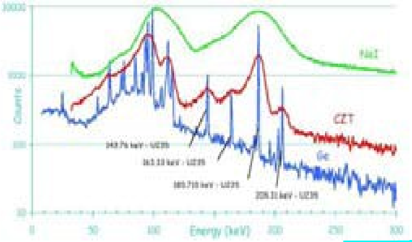
Example:Y Grafts for pediatric cartiology - Alison Marsden

The result will say much about the quality of a llittle person's life

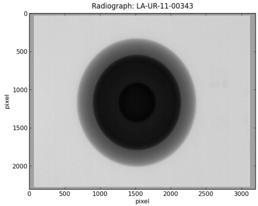


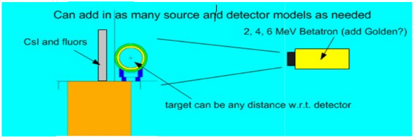
Nuclear Weapon and Nuclear Material Detection

Mesh Adaptive Direct Search (MADS) has been applied to solve inverse nuclear detection problems.



Uranium spectra [P. Peerani]





Experimental set-up scheme of radiography systems

Target optimization problems - the other 90%

$$(NLP) \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \Omega, \end{array}$$

where $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ may be discontinuous, Ω is any subset of $\mathbb{R}^n$ and:

- evaluation of f and of the functions defining Ω are usually the result of a black box computer code with some '`if`'s and '`goto`'s
- the functions are expensive black boxes - secs, mins, days
- the functions may fail unexpectedly even for $x \in \Omega$
- only a few correct digits are ensured
- accurate approximation of derivatives is problematic
- the constraints defining Ω may be nonlinear, nonconvex, nonsmooth and may simply return '`yes/no`'.

Presentation Outline

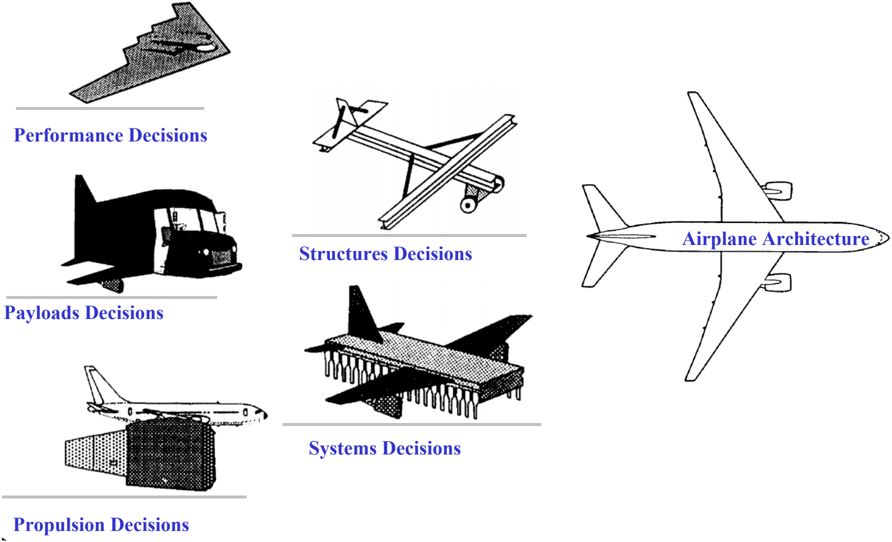
What needs to be done

We have done some useful things, and we are proud of that, BUT....

- What if there is uncertainty in the underlying evaluations? How sure are we of our answer?
- How to make the solution robust: If we design a wing, the properties of the material used to make it will not be known exactly, and the manufacturing process will not be executed exactly as we specify.
- All the problems we see are multiobjective optimization problems.

Multiobjective Design

This slide has been at Boeing so long that no one knows its origin:



Approaches to Multiobjective Optimization

- Goal programming: Set goals for all the objectives except one. Use the goals to turn the other objectives into constraints
- Do "trade studies". Plot the results of goal programming with various goals to find a "Pareto" point where the trade off of one objective against another seems reasonable. Avoid points where a slight improvement in one objective requires a much larger degradation in another. Balanced trade off points are called the knee of the Pareto set.
- Have all the decision makers gather round, hold hands, and meditate to a weighting of the objectives. Favored by academics, but seldom works on real problems because it tends to miss the knee.
- Finding the Pareto knee is everyone's choice

Hubris or Sanguinity?

Both types of uncertainty can be modelled for optimization by incorporation into f, Ω

$$\begin{aligned} \text{minimize} \quad & \hat{f}(x) \\ \text{subject to} \quad & x \in \hat{\Omega}, \end{aligned}$$

where $\hat{f}(x), \hat{\Omega}$ incorporate uncertainty modeled by the user's scheme of choice

Still belongs to target class,but $\hat{f}(x)$ and $\hat{\Omega}$ will be much more expensive

We have come a long way for this class of problems, but we have a ways to go to be really able to do what is needed at reasonable cost

I am sure that the new developments will have to involve simple surrogates to lessen the number of actual function evaluations. Surrogates are not accurate function approximations

Finally...

Happy Birthday Philippe!!