Success as a Student Researcher: A Discussion of Best Practices

Fall 2018 – Part 1

Dr. Scott Ferguson
Dr. Kate Saul
Dr. Mark Pankow
Why do we do this seminar?
Doing good research is hard

• There are many new skills to learn
  – Understanding the literature
  – Setting up your experiments/simulations
  – Data analysis
  – Interpretation of your results
  – Writing and presenting

• There is often a long time-scale on rewards

• At many levels, you are responsible for directing your own education
Your strategic plan

• A strategic plan has three parts:
  – Where are you today? Graduate school should be part of your strategic plan
  – Where do you want to be in the future (5, 10, 20 years from now)
  – How do you get from here to there?

A strategic plan is a roadmap for your (professional) life
What is research?
What IS research?

• The process of finding out something that we (everyone) do not already know

• Builds upon the extant knowledge base

• It is *methodical*, repeatable and verifiable
  – You can specify, in advance, a procedure to accomplish your stated objective
Why is problem choice so important?

- You will be spending a lot of time on your problem
- Personal interest will keep you motivated

““When one can achieve self-expression in science, work becomes revitalizing, and laden with personal meaning.”” (Alon)
Finding an advisor

• Ideally, you have selected your school by identifying faculty you would like to work with
  – If not, start looking NOW

• A good advisor will serve as a mentor **AND** a source of technical assistance
  – They should help you set and achieve short-term and long-term goals
A research topic must ...

- be research
- not have been done before
- be significant
- have a greater than zero probability that you can do it
- lend itself to a viable research plan
- be accomplished with the facilities you have available
- fit into your strategic plan
Factors to consider

• Feasibility: How hard or easy it is
  – Problems are always more difficult than they seem

• Importance: Impact on the community and beyond
  – Who cares?
  – What will they do when they see your work?
  – How long will the answer be important?
Factors to consider

• **Interest**
  – Both internal and external
  – Do you have a passion for the topic?
  – What will keep you working on it?

• **Competence**
  – Why are you qualified?
  – Do you have an advantage?
Establishing a research objective / direction
Heilmeier questions

• What are you trying to do?

• How is it done today? What are the limitations of current practices?

• What is new in your approach and why do you think it will succeed?

• Assuming success, what differences does this make to us and society?
Heilmeier questions

• What are the risks, and what can we do about them?

• How long will it take?

• How much will it cost?

• What is the timeline and what are the deliverables that we should expect throughout the project?
Establishing a research objective

• Concise statement of what you intend to find out that we don’t already know

• Do not use words that mean “not research” to define your objective
  – Develop
  – Design
  – Optimize
  – Control
  – Manage
How to do it right!

- The research objective of this project is to account for uncertainty in engineering design decision making through the application of utility theory.

- The research objective of this project is to measure the cross-section of the muon-neutrino interaction at 5 GeV accurate to 10%.
Responsible research conduct
Research responsibility

• Your obligation:
  – Clear records
  – Honest results
  – Appropriate acknowledgment
    • authorship, citations, acknowledgments, funding

• Responsibility to:
  – Your advisor and group
  – The university
  – The sponsor
  – The law
  – The scientific community
  – The public-at-large
Considerations related to research and scholarly integrity

- Personal honesty, ethics, and morals
- Expertise in the field of study – doing good science
- Professional codes of conduct and research practice, including publication policies established by professional journals
- Data ownership and control
- Institutional policies and regulations
- Governmental policies and regulations

Do the right thing!

Responsible Conduct of Research, Scholarship, and Creative Activities
Michigan State University Graduate School, 2010  http://grad.msu.edu/
Considerations related to research and scholarly integrity

<table>
<thead>
<tr>
<th>Continuum</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research integrity</td>
<td>Best practices</td>
</tr>
<tr>
<td>Questionable research practices</td>
<td>Sloppy work, lack of expertise, or ignorance of policies and regulations</td>
</tr>
<tr>
<td></td>
<td>(other possibilities exist!)</td>
</tr>
<tr>
<td>Unacceptable research practices</td>
<td>Failure to correctly observe applicable policies and regulations</td>
</tr>
<tr>
<td>Research misconduct</td>
<td>Deliberate efforts to plagiarize, fabricate, or falsify research data</td>
</tr>
</tbody>
</table>
Research misconduct

• Misconduct that distorts scientific knowledge
  – Fabrication – reporting of non-existent data
  – Falsification – selective reporting of data, misrepresentation

• Misconduct that misleads the scientific community
  – Plagiarism
  – Inappropriate authorship
  – Duplicate publication
  – Abuse of peer-review
Research misconduct

• Misconduct relating to human subjects
  – Consent issues
  – Exploitation issues (e.g. financial, power…)

• Other
  – Conflict of interest
  – Poor record-keeping
  – IRB and IACUC approval issues

• HONEST, unintentional error is not misconduct. But BE CAREFUL.
Serious deviation: possible scenarios

- Violation of confidentiality in proposing, performing, reviewing, or reporting research
- Misrepresenting credentials in proposing or presenting research
- Stealing, destroying, or damaging the research property of others with intent to alter the research record
- Serious or continuing noncompliance with federal regulations or University policies
SURVEY of 2,212 Researchers

– Observed 201 instances of misconduct
– E.G.

► "A post doc changed the numbers in assays in order to 'improve' the data."

► "A colleague duplicated results between three different papers but differently labeled data in each paper."

► "A co-investigator on a large, interdisciplinary grant application reported that a postdoctoral fellow in his laboratory falsified data submitted as preliminary data in the grant. As principal investigator of the grant, I submitted supplementary data to correct the application."

► "A colleague used Photoshop to eliminate background bands on a western blot to make the data look more specific than they were."
Tuskegee Syphilis Study (1932-27)

- Poor African-Americans with advanced syphilis recruited to trial to study their “bad blood”

- Deliberately untreated to follow natural history of syphilis infection, even after penicillin available

- Over 400 men and their families involved
- Breach of Human Rights
- Lead to Belmont Report and protection of Human Subjects laws
Why does this happen?

• Sloppiness when conducting the literature review (e.g., cut & paste & forget the original citation)
• Inadequate knowledge of research literature
• Inadequate expertise in research methods
• Pressure from others to produce data quickly
• Time crunch
• Malfunctioning equipment
• Poor mentoring
• Personal problems
• Cultural differences
Consequences

• Investigation at institutional to federal level
• Withdrawal or correction of all pending and published papers and abstracts affected by the misconduct
• Reprimand, removal from project, rank and salary reduction, dismissal
• Restitution of funds to the granting agency
• Ineligibility to apply for Federal grants for years (debarment)
• Criminal prosecution
• I.E. the end of your research career!
Who is accountable

• Investigated
  – All authors that are involved in the specific data in question

• Held accountable
  – Primary author
  – Other authors whose results are found culpable
  – The PI
Real example: Jan Hendrik Schon

Jan Hendrik Schon

4 years after his Ph.D. (1997→2001), Schon was listed as an author on a new paper every 8 days. But others couldn’t reproduce his results, and then found suspicious things in Schon’s papers, like 2 curves with the same noise.
Real example: Jan Hendrik Schon

Bell Labs Fires Star Physicist Found Guilty of Forging Data

Like the mythical Faunus, whose razor-sharp hooves allowed him to flee to the sea, the reeling career of Jan Hendrik Schon came crashing down to earth last week. Schon, a 33-year-old physicist at Bell Labs, refuses to offer a coherent explanation for forging data in at least 17 published papers, according to a report released 25 September by a panel of independent investigators. Schon had been fired from Bell Labs last month, and the paper generated widespread media coverage.

Schon’s story is one of the most striking cases of scientific misconduct in modern history, and it is a cautionary tale for all those who work in the field of nanotechnology. By the time Schon was found guilty of forging data, his work had earned him an international reputation for his groundbreaking research in the field of nanotechnology.

In his career at Bell Labs, Schon was a prolific researcher, publishing over 50 papers in leading scientific journals. His work was widely cited and had a significant impact on the field of nanotechnology. However, it was his research on graphene nanoribbons that caught the attention of the scientific community.

The paper, published in Science in 2007, claimed to present evidence of a new type of nanomaterial that could potentially revolutionize the field. However, upon closer inspection, it became clear that the data presented in the paper was fabricated.

The investigation into Schon’s work was led by a panel of independent scientists who were tasked with verifying the accuracy of his research. After a thorough review of the evidence, the panel concluded that Schon had indeed fabricated data in at least 17 papers.

The implications of Schon’s actions are significant. They highlight the importance of rigorous peer review and the need for transparency in the scientific community. It also raises questions about the culture of scientific publishing, where the pressure to publish can sometimes lead to unethical behavior.

Schon’s case is a stark reminder of the importance of ethical conduct in science. It serves as a warning to all those who work in the field and underscores the need for a continued focus on fostering a culture of integrity and accountability within the scientific community.
Real example: Jan Hendrik Schon

Duplicate Publication: for real!

- Two papers published in top-level journals, within one year
- Significant duplication of figures (data) and results
- Same authors, different order, different corresponding author
- Paper #1 was published before paper #2 was submitted
- Paper #2 does not cite paper #1

Paper #1

For the CNT-reinforced sandwich beam, the analysis shows good correlation with test data for a damping ratio ($Z_d$) of 0.3 and a cross-sectional stiffness (EI) of $87.7 \times 10^{-3}$ N m$^2$. Therefore, we conclude that carbon nanotube reinforcement results in a 200% increase in the baseline structural damping (due to the frictional energy dissipation during the movement of individual nanotubes in the film) and a 30% increase in the baseline bending stiffness (due to stiffening of the 2 mm adhesive sub-layer). Based on the observed increase in cross-sectional stiffness of the laminate from $65.1 \times 10^{-3}$ N m$^2$ to $87.7 \times 10^{-3}$ N m$^2$, the modulus of the carbon nanotube film was estimated using the Bernoulli–Euler theory as $41.2 \times 10^9$ psi (284 GPa). Table 2 compares the properties and operating conditions of the carbon nanotube film with commercially available viscoelastic damping polymers such as 3M1SD-112 and Soundcoat Dyad 606.

Paper #2

For the CNT-reinforced sandwich beam, the analysis shows good correlation with test data for a damping ratio ($Z_d$) of 0.3 and a cross-sectional stiffness (EI) of $87.7 \times 10^{-3}$ N m$^2$. Therefore, we conclude that carbon nanotube reinforcement results in a 200% increase in the baseline structural damping (due to energy dissipation during the deformation of nanotube clusters within the film) and a 30% increase in the baseline bending stiffness (due to stiffening of the 0.05 mm epoxy sub-layer). These results are summarized in Table 2. Based on the observed increase in cross-sectional stiffness of the laminate from $65.1 \times 10^{-3}$ to $87.7 \times 10^{-3}$ N m$^2$, the modulus of the carbon nanotube film was estimated using Bernoulli–Euler theory as 284 GPa. Table 3 compares the properties and operating conditions of the carbon nanotube film with commercially available viscoelastic damping polymers such as Soundcoat Dyad-606 [4] and 3M 1SD-112 [5]. While the damping properties of all three films are comparable at room temperature, for high temperature applications the carbon nanotube based films are expected to provide superior performance and reliability.
What is a difficult situation?

- Conflicts with your mentor about expectations for degree completion or research assignments
- Conflicts about issues such as authorship, data management, lab safety, protection of human subjects, animal welfare, or conflict of interest
- Problems with communication or collegiality in the research environment
What is suspected misconduct?

• Fabrication of data

• Falsification

• Plagiarism

• Or a serious deviation from commonly accepted practices in your discipline
If you face a difficult situation

• Make notes about your observations
  – Focus on what you know or have observed
  – Do not rely upon hearsay
  – Avoid emotional reactions

• Get the big picture
  – Discuss the matter with the people involved in the situation – understand different perspectives
  – Ask for clarification, e.g., “I don’t understand this” or “what is the right thing to do in this situation” or “what did you mean by”
  – Do not make the situation worse – avoid rumors
If you face a difficult situation

• Do not take unilateral action – talk to someone such as your mentor, advisor, or lab director
  – Ask for advice about resolving the problem
  – Seek education about applicable procedures, policies, and rules
  – Be a responsible adult – if you are part of the problem, admit to your mistakes and ask for advice about how to help rectify the situation

• If the matter is not resolved amicably, seek advice from the NCSU Research Integrity Officer

Responsible Conduct of Research, Scholarship, and Creative Activities
Michigan State University Graduate School, 2010 http://grad.msu.edu/
NCSU resources

- [http://research.ncsu.edu/sparcs/training/training-rcr-courses/](http://research.ncsu.edu/sparcs/training/training-rcr-courses/)
- Preparing Future Leaders Blog: [http://pfl.grad.ncsu.edu/category/blog/](http://pfl.grad.ncsu.edu/category/blog/)
- Research integrity officer:
  Richard Best
  919-515-0158
  richard_best@ncsu.edu
Confidentiality

• Be aware of restrictions for your specific research project

• Proprietary information
• Personal health records
• Unpublished research
Understanding the work presented in the literature
How do I...

• ...know what and where to search?
• ...do the search?
• ...know if what I’m reading is important?
• ...keep track of what I read?
• ...keep up to date on the latest “news” after my preliminary search?
• ...maintain my own records?
Breadth

Depth

Interdisciplinary journals, review papers

Domain expert journals, long papers with lots of equations
Searching a new topic

- Identify **keywords** for your topic and browse ALL the articles you find
- Identify key **sources** and browse them
- Identify key **people** (leaders! heroes!) and look up their work
- Identify **review papers**
- Follow the **references**
- Revise the **keywords**
- Create **alerts**
What sources should you read?

• Journals
  – Inside your field
  – Interdisciplinary (e.g. Nature, Science)
  – Outside your field but encompassing complementary topics

• Conference proceedings (abstracts or papers)
  – Newest research
  – Check out webpages for specific conferences

• Theses/Dissertations
  – More detail than papers

• Patents and industry reports
  – If your work has commercial applications, patents can give design/process details.
  – Can understand state of intellectual property
Conducting a literature review

Trying to find a New Research Problem

Starting with a Clearly Defined Research Problem

Start with 1-3 Research Areas

Google
10th Anniversary Series - John Sack: Helping Researchers See Farther Faster

My updates: recommended based on My Citations  Learn more

In vivo pediatric shoulder muscle volumes and their relationship to 3D<em>in vivo</em> strength
H Soo Im, KE Alter, S Brochard, C Pons, FT Sheehan - Journal of Biomechanics, 2014

EMG Driven muscle Force Estimator (EMG-FE) Version 1.0 r7
distribution 1 User Guide
LL Manegaldo, LF de Oliveira - 2014

See all updates

Stand on the shoulders of giants
Google Scholar Results

An overview of evolutionary algorithms in **multiobjective optimization**
CM Fonseca, PJ Fleming - Evolutionary computation, 1995 - MIT Press
(2013) Evolutionary **multiobjective optimization** in water ... (2013) BSTBGA: A hybrid genetic algorithm for constrained **multi-objective optimization** problems. ... (2012) Optimal design of actively controlled adjacent structures for balancing the mutually conflicting **objectives** in design ...
Cited by 2308 Related articles All 17 versions Cite Save More

Physical programming-Effective **optimization** for computational **design**
A Messac - AIAA journal, 1996 - arc.aiaa.org
(2012) Condition based maintenance **optimization** considering multiple **objectives**. ... Visualization (HRV) method with range-based preferences for **multi-objective** decision making. ... (2010) Integration of Preferences in Hypervolume-Based **Multiobjective** Evolutionary Algorithms by ...
Cited by 429 Related articles All 7 versions Cite Save

Survey of **multi-objective optimization** methods for engineering
RT Marler, JS Arora - Structural and multidisciplinary **optimization**, 2004 - Springer
Cited by 1591 Related articles All 9 versions Cite Save

A **multi-objective optimization** for green supply chain network **design**
... However, any "**design**" in nature is usually involving trade-offs among different incompatible **objectives**. ... The **multi-objective** model explicitly considers the environmental issues by introducing a new category of decision variables: the ... d i p the demand of customer for **product**. s i p ...
Papers Behind Paywalls
(1) Use University Credentials
(2) Try other versions on main Scholar page

Multiobjective evolutionary algorithms: Analyzing the state-of-the-art
DA Van Veldhuizen, GB Lamont - Evolutionary computation, 2000 - MIT Press
Abstract Solving optimization problems with multiple (often conflicting) objectives is, generally, a very difficult goal. Evolutionary algorithms (EAs) were initially extended and applied during the mid-eighties in an attempt to stochastically solve problems of this ...
Cited by 768 Related articles Cite Save

[PDF] Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art
DA Van Veldhuizen, GB Lamont - Evolutionary Computation - Citeseer
Abstract: Solving optimization problems with multiple (often conflicting) objectives is, generally, a very difficult goal. Evolutionary algorithms (EAs) were initially extended and applied during the mid-eighties in an attempt to stochastically solve problems of this ...
Cite

Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art
DA Van Veldhuizen, GB Lamont - 2000 - citeseer.ist.psu.edu
Abstract: Solving optimization problems with multiple (often conflicting) objectives is, generally, a very difficult goal. Evolutionary algorithms (EAs) were initially extended and applied during the mid-eighties in an attempt to stochastically solve problems of this ...
Cite

Multiobjective evolutionary algorithms: analyzing the state-of-the-art.
DA Van Veldhuizen, GB Lamont - Evolutionary computation, 1999 - europepmc.org
Solving optimization problems with multiple (often conflicting) objectives is, generally, a very difficult goal. Evolutionary algorithms (EAs) were initially extended and applied during the mid-eighties in an attempt to stochastically solve problems of this generic class. During the ...

...or, Google the author
The library

The James B. Hunt Jr. Library
Down the rabbit hole

• Pull the cited papers

• **READ** the cited paper itself before citing it – don’t rely on another paper’s interpretation. Sometimes they aren’t right! Don’t propagate error!

• Look at the citation numbers of papers to gauge importance (date of publication and subfield will influence this, so this is imperfect)
  – Science Citation Index, Web of Science, Sci Search
Understand your community

• Identify the research groups that are most relevant
  – Follow their work
  – Suggest as reviewers
  – Make contact at meetings

• Flag important methods
  – Understand best practices
  – Identify different information obtained from different approaches
  – Incorporate into your work

• Keep track of next steps suggested in discussions or raised in your reading

• Identify ways in which your work builds on other groups
Research social networks

- LinkedIn
- ResearchGate
- ResearcherID
- ORCID
- GoogleScholar

- Unique identifier to link you to research, distinguish you from other researchers
- Provides citation information
- Links to co-authors
The process is iterative

1. Find enough papers to create a foundation for your work, then do your work.
2. Find papers as you’re working.
3. Find papers to help solidify and motivate the work you’ve already done.
4. Check for papers early and often (Google scholar has an alert system).
Reading papers
Questions to ask while you are reading...

• Who are these authors? Do they have the relevant expertise?

• Is the source credible?

• Do I understand the terminology that is being used in the paper?
Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two bar truss and an I-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA
Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two bar truss and an I-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA
Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two bar truss and an I-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA
Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two bar truss and an I-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA
Start with the abstract

Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two bar truss and an I-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA
One strategy...

Exploring the Effectiveness of Using Graveyard Data When Generating Design Alternatives

The objective of this paper is to demonstrate that unique alternative designs can be efficiently found by searching the discarded data (or graveyard) from a multiobjective genetic algorithm (MOGA). Motivation for using graveyard data to generate design alternatives arises from the computational cost associated with real-time design space exploration of multiobjective optimization problems. The effectiveness of this approach is explored by comparing (1) the uniqueness of alternatives found using graveyard data and those generated using an optimization-based search, and (2) how alternative generation near the Pareto frontier is impacted. Two multiobjective case study problems are introduced—a two-bar truss and an L-beam design optimization. Results from these studies indicate that using graveyard data allows for the discovery of alternative designs that are at least 70% as unique as alternatives found using an optimization-based alternative identification approach, while saving a significant number of functional evaluations. Additionally, graveyard data are shown to be better suited for alternative generation near the Pareto frontier than standard sampling techniques. Finally, areas of future work are also discussed. [DOI: 10.1115/1.4024913]

Keywords: design alternatives, multiobjective optimization, multiobjective genetic algorithms, design space exploration, MGA

1. Introduction

Product designers face the challenge of creating products in markets where customers have highly heterogeneous preferences. Improving performance in market-related objectives, such as market share of preference, requires a product line, i.e., a set of related products that are offered by a single company. This is different from a product family where commonality of features, components and subsystems is often explicitly enforced (Simpson, Maier, and Mistree 2001).

Customers also desire products that maximize their value for money (Prahalad and Maslach 2010). This leads to complex design problems that require: (1) advanced techniques to capture and model customer preferences for product features; and (2) optimization techniques capable of searching the expansive mixed-integer design space associated with the resulting combinatorial problem. The nature of many feature-packaging problems supports the use of heuristic optimization...
Why the introduction is important

• Tells you what is known in the field

• Explains the limitations of our current understanding

• Leads to a focal point – what question is going to be answered

PROVIDES PERSPECTIVE
Questions to ask while you are reading...

• Is there other work I should be looking at first (previous papers in this area)

• Who can I talk to about the confusing parts of this paper?

• Am I spending too much time reading the less important parts?
Three pass approach

• Pass 1: General idea

• Pass 2: Content, but not details

• Pass 3: Understand in depth
Pass 1

• **READ:**
  – Title, abstract, introduction
  – Section and subsection headings (if they exist)
  – Glance at math
  – Conclusions
  – Skim references

• **ANSWER:**
  – Category: What kind of paper is this? Clinical? Technical development?
  – Context: What is it related to?
  – Correctness: Assumptions valid?
  – Contributions: Main contributions.
  – Clarity: Well-written?

• Keep going? Peripheral to research area?
Pass 2

• **READ:**
  – Careful look at figures/graphs
  – Mark relevant references you haven’t read

• **ANSWER:**
  – Summarize main thrust with evidence – cover sheet/notes

• Keep going?
Pass 3

- **CLOSE ATTENTION:**
  - Thought experiment
  - Check all assumptions – any missing?
  - Check all details – any errors?
  - Limitations – any unidentified? How to address?
  - Important citations – any missing?
  - Think about future work
Questions to ask after reading the paper…

• Is the proposed approach a good one?

• Are the findings persuasive? Supported by enough evidence?

• Is there an alternative interpretation of the data not addressed by the authors?

• How does this relate to my work?
Takeaways from reading EVERY paper

• What did they actually do?

• What knowledge did we gain because of this?

• What assumptions did they make?

• What papers did they reference?

• What opportunities for future work exist?
Now that you’ve read the paper

• How do I store this paper?

• Should I use a template for my notes?

• How do I save my notes?
Sit down and write your notes

• As you’re reading, take notes (a few sentences) about each paper
  – Method
  – Major contribution

• Bin papers by common theme

• Write a paragraph (or more) about each theme, citing the papers you’ve found.
  – Tie to your work
What is a Literature Review?

Introduction

According to the U.S. Department of Energy (DOE), “Increased R&D efforts and innovation will be required to continue to expand the wind energy industry” in the United States, where wind power is still a maturing technology. This trend is evident in the number of wind farm layouts that the demand for energy will be met by by 2030, as reported in the American Wind Energy Association (AWEA) report, with the intent of supplying 20% of the total electricity demand utilizing wind power. In addition, outdated means of energy development have been attributed to climate change, pollution, and permanent depletion of natural resources. These concerns have led to public demand for cleaner, sustainable energy sources like wind turbine technology. To meet these demands it will be necessary to ensure new wind farm installations are developing as much power as possible, which depends on the wind farm layout's capability to account for local wind conditions and aerodynamic interaction between turbines.

The complexity of the wind farm layout problem lies in the dynamic conditions of the farm site and the modeling available to represent realistic wind patterns and wake interactions. A turbine in wind will develop a turbulent wake that decreases the wind speed downstream. On a wind farm, turbines are typically placed in close enough proximity that the effect of placement in a wake could drastically reduce the effective wind speed to downstream turbines and therefore decrease their power output. Essentially, the goal is to incorporate as many turbines as possible while minimizing land use, without compromising the efficiency of the farm due to wind speed decrement from wake interactions. This work seeks to develop an optimization algorithm that will consider previous approaches...

Previous Approaches

The first attempt at applying computational optimization algorithms to the wind farm layout problem was by Mosetti et al. [2], utilizing a genetic algorithm (GA). This preliminary study developed the framework that has been continuously used for comparison purposes, including the objective function formula and the use of the Jensen wake model [3]. Mosetti et al. used a discretized solution space of 100 × 100 cells and limited the placement of turbines to the center of each cell. The genetic algorithm considers each row of the grid as a binary string. A similar but improved GA approach was performed by Grady et al. [4] whose algorithm incorporated heuristic knowledge about wind farms and utilized more advanced computational resources. Two more recent studies, one by Wan et al. [5] and one by Mittal [6] improved on both previous genetic algorithm approaches by implementing a second discretization phase that allowed the turbines to be moved within their assigned cells. Though these works have advanced the study of wind farm layout optimization, these genetic algorithms may have been hindered by the inherent discretization of their binary...

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received March 14, 2011; final manuscript received May 2, 2012; published online July 23, 2012. Assoc. Editor: Wei Chen.
Why Write One?

You have to.

Shows the need and novelty of your work (the research gap)

Gives reasons why you’re looking at your research problem in a technical (not an impassioned) way

Your work needs a solid foundation in existing literature in order to be taken seriously.
# Software for building a personal library

<table>
<thead>
<tr>
<th>Compare Products</th>
<th>MENDELEY</th>
<th>EndNote</th>
<th>RefWorks</th>
<th>zotero</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic software package (includes all features listed below)</td>
<td>Free</td>
<td>$250</td>
<td>$100</td>
<td>Free</td>
<td>$79</td>
</tr>
<tr>
<td>Free web storage space (online backup of your papers)</td>
<td>2GB</td>
<td>1GB</td>
<td>NA</td>
<td>300MB</td>
<td>NA</td>
</tr>
</tbody>
</table>

## Reference/Document Management

<table>
<thead>
<tr>
<th>Feature</th>
<th>MENDELEY</th>
<th>EndNote</th>
<th>RefWorks</th>
<th>zotero</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization of PDFs and other documents</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Citation Plug-ins for Word</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Citation Plug-ins for LibreOffice</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Annotations/Highlighting in PDFs</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Cross-platform syncing across desktop, web and mobile devices</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>

## Knowledge Discovery

<table>
<thead>
<tr>
<th>Feature</th>
<th>MENDELEY</th>
<th>EndNote</th>
<th>RefWorks</th>
<th>zotero</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free and open database approaching 100 million documents</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Personalized paper recommendations</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Readership statistics &amp; community tags</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Open Web API</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Full text search across all your papers</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Search across external databases</td>
<td>Almost there!</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

## Collaboration

<table>
<thead>
<tr>
<th>Feature</th>
<th>MENDELEY</th>
<th>EndNote</th>
<th>RefWorks</th>
<th>zotero</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private groups</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Public groups</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Social network</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Collaboration newsletter</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>
Revisiting the research plan
Heilmeier questions

• What are you trying to do?

• How is it done today? What are the limitations of current practices?

• What is new in your approach and why do you think it will succeed?

• Assuming success, what differences does this make to us and society?
Heilmeier questions

• What are the risks, and what can we do about them?

• How long will it take?

• How much will it cost?

• What is the timeline and what are the deliverables that we should expect throughout the project?