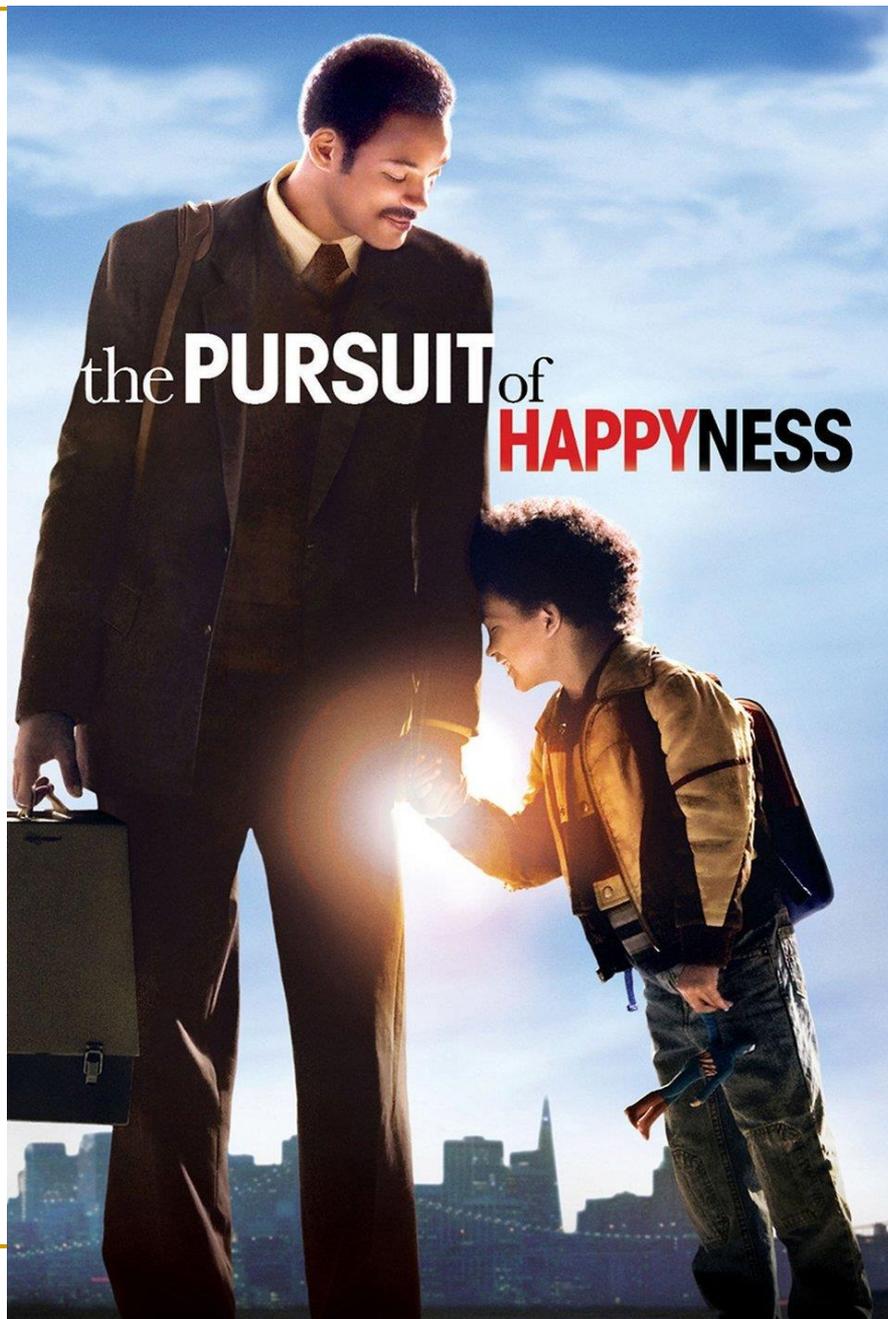


ISSRE2019 Doctoral Symposium Keynote
2019/10/28

The Pursuit of Happiness in Academia

Michael Rung-Tsong Lyu

Computer Science & Engineering Department
The Chinese University of Hong Kong



the **PURSUIT** of
HAPPYNESS

**“This part of my life...
this little part.... is
called Happiness.”**

Chris Gardner (Will Smith)

Motivation: Why Should I Do Research? (Where does my happiness come from?)

■ External drive

- ❑ Award, degree and diploma
- ❑ Parents, teachers, friends
- ❑ Peer pressure

■ Internal drive

- ❑ Sense of honor and responsibility
- ❑ Strong ambition (self-expectation)
- ❑ “Happiness” (sense of achievement/fulfillment, curiosity)

Research Problem Selection

- Good research largely depends on the selected problem
 - 90% of a research job is done when you find a good problem.
 - A good problem is difficult to find
 - Not too easy or too difficult
 - How to select a problem?
 - Is it an old problem or a new problem?
 - Usually, new problems have more opportunities
 - Is it a significant problem?
 - Practically important yet technically challenging
-

The Key Ingredients of Research: Contribution (Where is the beef?)

- One major contribution is better than many small ones
- What is the contribution type?
 - Knowledge discovery
 - Knowledge invention
 - Knowledge integration
 - Knowledge application
- Idea! Idea! Idea!
- Identify, describe, and demonstrate the big idea
- Asking the right problem and asking the problem right

Asking the Right Problem and the Right Way to Ask Problem

- Asking the problem right can lead to asking the right problem
 - Can we predict reliability?
 - Can we predict reliability with models?
 - Can we predict reliability of software?
 - Can we predict reliability with user experience?
-

How to Find Research Problems?

1. New Solution to Old Problems
2. New Problems
3. New Areas

1. New Solution to Old Problems

- New solution to a (reduced) problem
 - Fermat's last theorem
 - For all positive integers a , b , c and n , there is no solutions to
$$a^n + b^n = c^n \text{ when } n \geq 3$$
 - Java and C alias/pointer analysis (PLDI2013)
 - “DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation” (Gan and Tao), SIGMOD 2015

1. New Solution to Old Problems

- New solution from the same area
 - Coolstreaming (INFOCOM 2015 Test of Time Paper Award)
- New solution from other areas
 - Recommendation applied to Software Reliability Engineering (ICSE2010 Distinguished Paper Award)

2. New Problems (Usually with a Twist)

- Extension

- M⁴ (Maxi-Min Margin Machine) work

- Summarization

- “A Design Paradigm for N-Version Software”

- Variation

- Linear combination model

2. New Problems (Usually with a Twist)

■ Refinement

- Refine existing solution
 - Missing data prediction with Collaborative Filtering (2007 SIGIR paper)
- Refine existing problem (and solution)
 - Concept of social recommendation (using probabilistic matrix factorization)
- Revisit on “Asking the problem right can lead to asking the right problem”

Reliability Prediction of Web Services

- Approach 1: Neighborhood-based approach – to consider **users**
 - Approach 2: Model-based approach – to consider **data sparsity**
 - Approach 3: Time-aware approach – to consider **temporal factor**
 - Approach 4: Network coordinate based approach – to consider **spatial factor**
 - Approach 5: Ranking-based approach – to consider **ranking**
 - Approach 6: Reputation-aware approach – to consider **reputation**
-

Reliability Prediction of Web Services

- Approach 1: Neighborhood-based approach – to consider **users** [ICSE'10, ACM TOSEM]
 - Approach 2: Model-based approach – to consider **data sparsity** [IEEE TSC'13]
 - Approach 3: Time-aware approach – to consider **temporal factor** [ISSRE'11]
 - Approach 4: Network coordinate based approach – to consider **spatial factor** [ICWS'12]
 - Approach 5: Ranking-based approach – to consider **ranking** [TPDS'13]
 - Approach 6: Reputation-aware approach – to consider **reputation** [SCC'13]
-

2. New Problems (Usually with a Twist)

- Generalization
 - Machine learning, Deep learning
- Specialization
 - Online algorithms

3. New Areas

- Exploration
 - AI by Deep Learning
- Cross Disciplinary Research
 - AI and Software Engineering
- Disruptive Research
 - Quantum computing (*new theory*)
 - Internet and WWW (*new user experience*)
 - Apple's revolution (*new business model*)

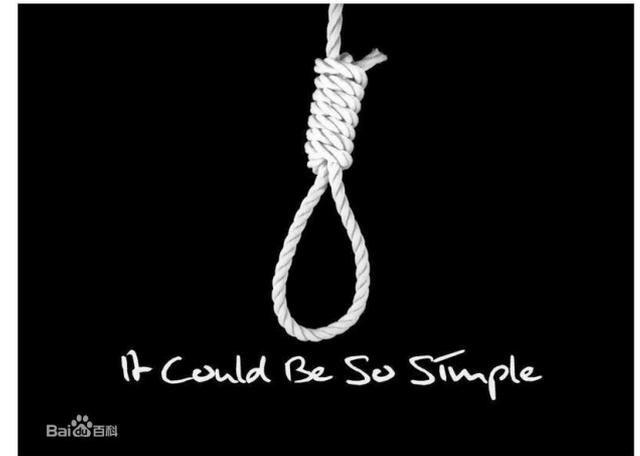
Good versus Bad Research Problems

- Good research problem
 - After the research, more people will be doing it – you opened the door
 - After the research, nobody can improve it – you closed the door
 - Bad research problem
 - Nobody will follow the research
 - Nobody really cares the research
 - Research subject hotspot index: $df(t)/dF(t)$
 $f(t)$ =No. of top papers and $F(t)$ =No. of total papers
-

How to Find the Solution?

- Five typical steps in scientific research
 - ❑ Observation/Exploration
 - ❑ Assumption/Proposition
 - ❑ Theory/Methodology
 - ❑ Verification/Experimentation
 - ❑ Conclusion/Elaboration
- Thinking outside the box

The Goulding knot



Literature Survey

- Research lifecycle: imitating, remembering, analyzing and innovating
- Use tools
 - Trace backward
 - Tutorial paper and reference list
 - Trace forward
 - Use Google scholar to find papers that cite the current work
- Proactive vs. passive reading
 - Reading with a critical attitude
 - Reading according to your own agenda
 - Reading between lines (not only what was said but what was not said)
- Form a study group

Nurturing Good Taste

- There are many mediocre papers published
 - Do not ruin your taste by poor-quality papers
 - Read selectively
 - Highly cited papers and papers from first-tier journals and top top-ranked conferences
 - Classification of papers
 - Type A: 80% understanding (main idea, solution method and main results)
 - Type B: 50% understanding (idea & results)
 - Type C: 20% understanding (only introduction)
 - Learn to appreciate good papers and criticize poor papers
-

Monitoring Activities of Leading Research Group in Your Field

- Identify leading research groups in your field
- Find out their recent research focus

Writing

- Critical to the sale of your ideas/results
- Paper organization
 - Proper arrangement of texts, figures and tables
- Multi-pass writing style
 - 1st pass: Detailed outline
 - 2nd pass: Rapid writing
 - 3rd pass: Fine fine-tuning
 - 4th pass: Cross-reading

Writing Procedure

- Carefully determine the paper title
- Proper use of names and notations
- Tell them what you are going to do, tell them what you are doing, tell them what you have done.
- Motivation! Motivation! Motivation!

Title

- A title needs to be catchy but precise
 - “Like Likes Like”
 - “All Models Are Wrong; Some Are Worse Than Others”
 - “Do Developers Discover New Tools On The Toilet?”
 - “All Your Clicks Belong to Me: Investigating Click Interception on the Web”
 - Use acronyms
 - KEEP: The Knowledge and Education Exchange Platform
 - M⁴ (Maxi-Min Margin Machine)
-

Motivation! Motivation! Motivation!

- The **introduction** is by far the most important section in the entire paper, especially for conferences.
- Reviewers are always very busy.
- If a reviewer can reject your paper without reading it all, it saves time!
- The introduction is the first section they read, so make sure your paper does not get killed in Section 1.
- “5 years ago I used to write the **introduction** last. Now it is always the first section I write.”

Strong Statements Are Dangerous ...

- Be very careful when you make strong statements about some research issue: there are people who think otherwise.
- Be especially careful when taking position on some hotly debated topics in the community, like:
 - Supervised learning vs. non supervised learning
 - Parametric vs. non parametric
 - Statistical vs. analytical
 - Partitioned vs global multiprocessor scheduling
 - Hard real-time vs wireless
 - Testing vs static analysis
- Instead of saying “X is black”, say “X is usually black, but in some cases that are not considered in this paper it is white”.

...But If You Are Confident, Go For It!

- However, high impact papers are those that successfully challenge existing preconceptions.
 - “**DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation**”, SIGMOD 2015 best paper award.
- So do not be shy when you state the main contribution of your paper!
- Just be sure to prove your point well enough; the keyword here is “**successfully challenge.**”

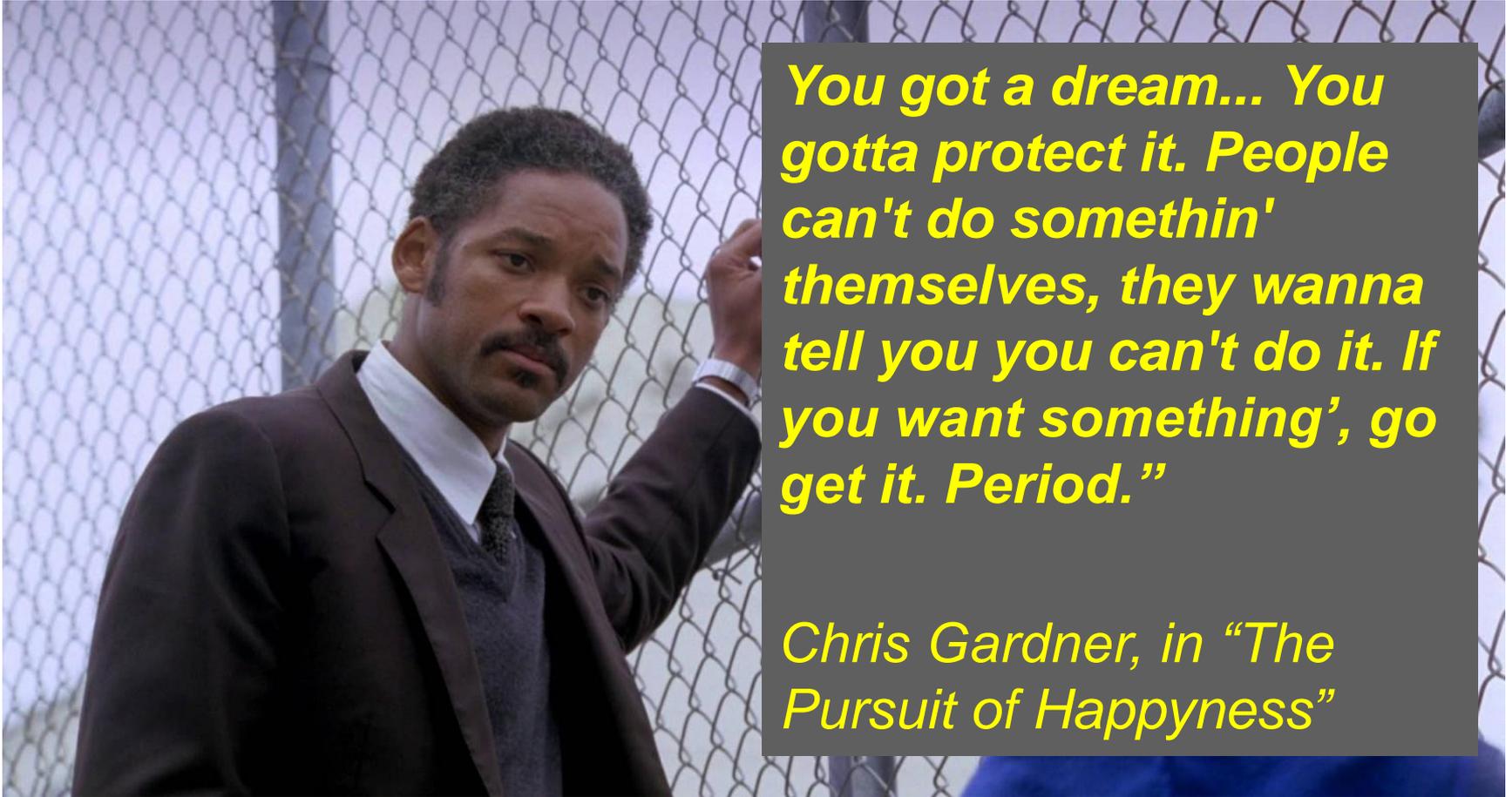
Criticize Your Writing in Reviewer's View

- A main factor of your success is to know how others think and feel
 - Reviewers' mindset: "You are assumed guilty until proven innocent"
 - Remind instead of assume, but don't humiliate their intelligence
 - Proper use of citation
 - Clearly and articulately indicate your contributions
 - Criticize yourself first, and leave reviewers no room for further criticism
 - Remember, the reviewers may be hostile
 - But they should not be malicious ...
-

You Cannot Make Everybody Happy

- Different people are looking for different things.
 - Also they are often biased.
 - You must accept that it is simply impossible to make everybody perfectly happy; you are forced to make trade-offs.
 - For the same reason, take all people's reaction with a grain of salt.
 - The key: two half glasses of water are better than one full and one empty glass here.
 - Just one negative review is enough to kill a conference paper.
 - The lesson: bad results can turn out good, so don't loss your heart.
-

Conclusions



You got a dream... You gotta protect it. People can't do somethin' themselves, they wanna tell you you can't do it. If you want something', go get it. Period."

Chris Gardner, in "The Pursuit of Happyness"

Thank you!