

U N I V E R S I T Y O F C A L I F O R N I A , M E R C E D

How to Get Your CVPR Paper Rejected?

Ming-Hsuan Yang

UCMERCED

www.ucmerced.edu

Outline

- Conferences
- Journals
- Writing
- Presentation
- Lessons

Conferences

- CVPR – Computer Vision and Pattern Recognition, since 1983
 - Annual, held in US
- ICCV – International Conference on Computer Vision, since 1987
 - Every other year, alternate in 3 continents
- ECCV – European Conference on Computer Vision, since 1990
 - Every other year, held in Europe

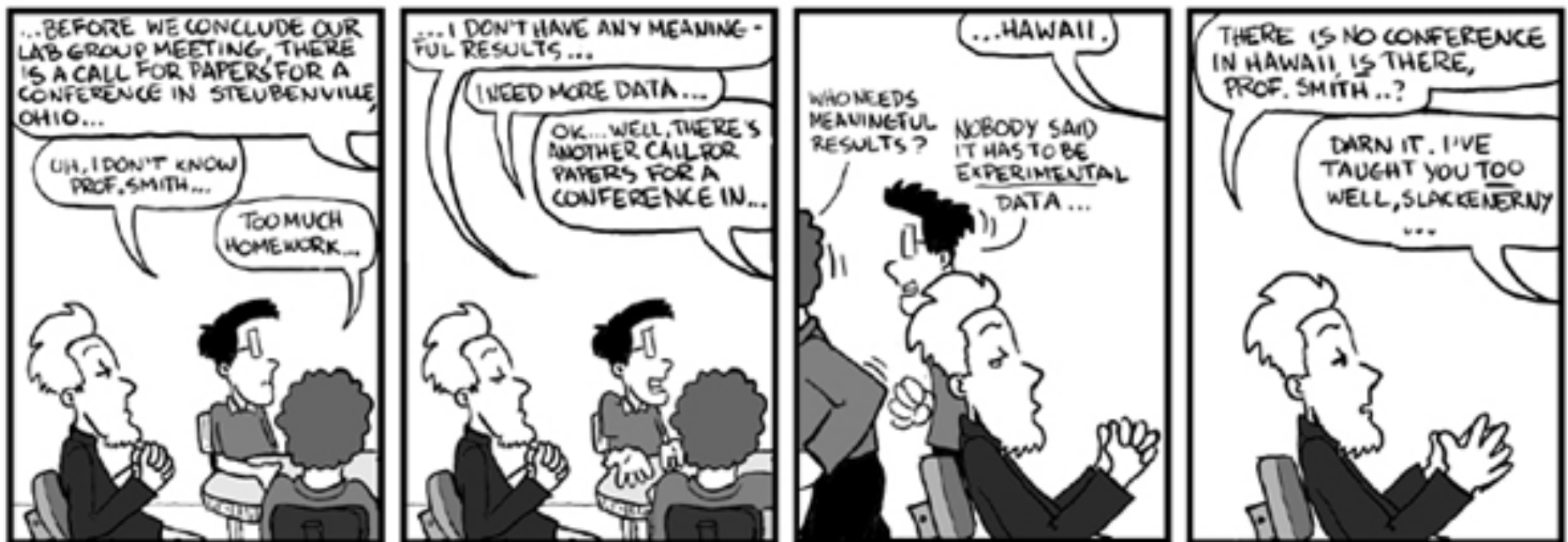
Conferences

- ACCV – Asian Conference on Computer Vision
- BMVC – British Machine Vision Conference
- ICPR – International Conference on Pattern Recognition
- SIGGRAPH
- NIPS – Neural Information Processing Systems

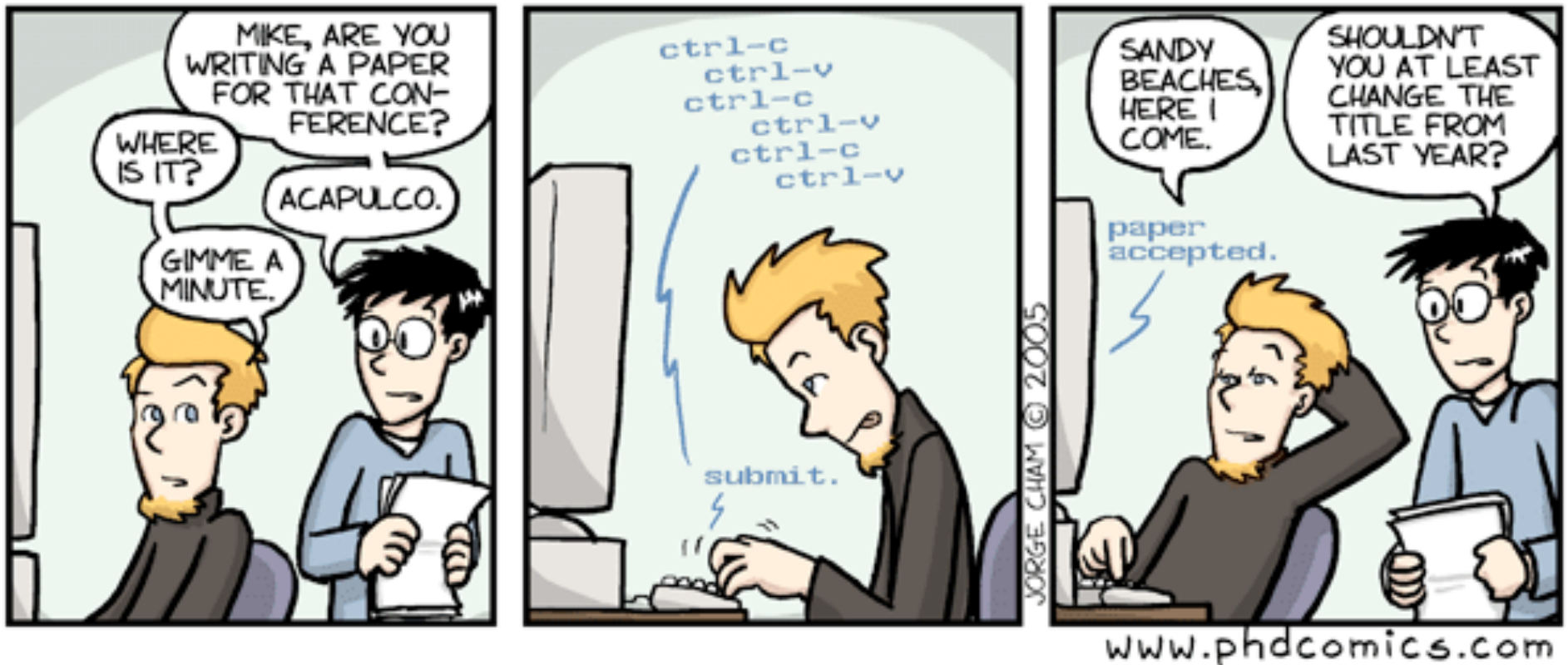
Conferences

- MICCAI – Medical Image Computing and Computer-Assisted Intervention
- FG – IEEE Conference on Automatic Face and Gesture Recognition
- ICCP – IEEE International Conference on Computational Photography
- ICML – International Conference on Machine Learning
- IJCAI, AAAI, MVA, ICDR, ICVS, DAGM, CAIP, ICRA, ICASSP, ICIP, SPIE, DCC, WACV, 3DPVT, ACM Multimedia, ICME, ...

Conference Location



Conference Location



- Me and confernece I want to attend
(location vs. reputation)

Conference Organization

- General chairs: administration
- Program chairs: handling papers
- Area chairs:
 - Assign reviewers
 - Read reviews and rebuttals
 - Consolidation reports
 - Recommendation
- Reviewers
- Authors

Review Process

- Submission
- CVPR/ECCV/ICCV
 - Double blind review
 - Program chairs: assign papers to area chairs
 - Area chairs: assign papers to reviewers
- Rebuttal
- Results

Area Chair Meetings

- Each paper is reviewed by 2/3 area chairs
- Area chair make recommendations
- Program chairs make final decisions
- Virtual meetings
- Onsite meetings
 - Several panels
 - Buddy/triplet

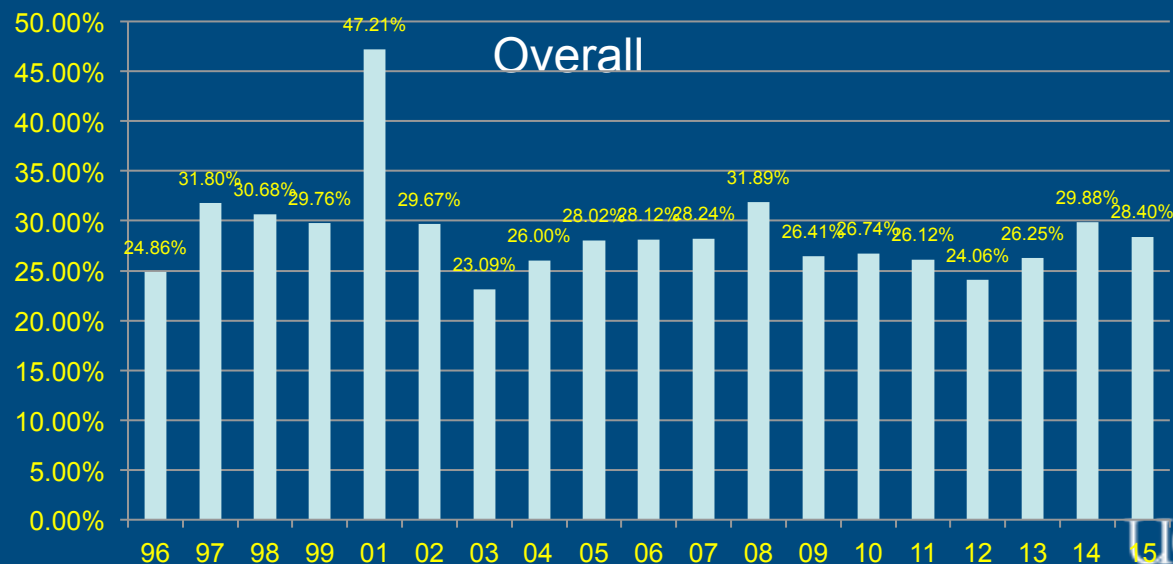
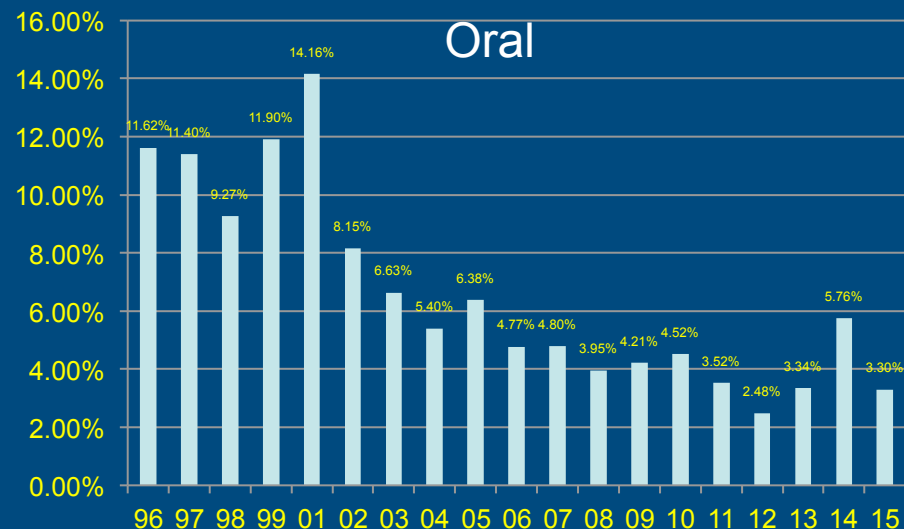
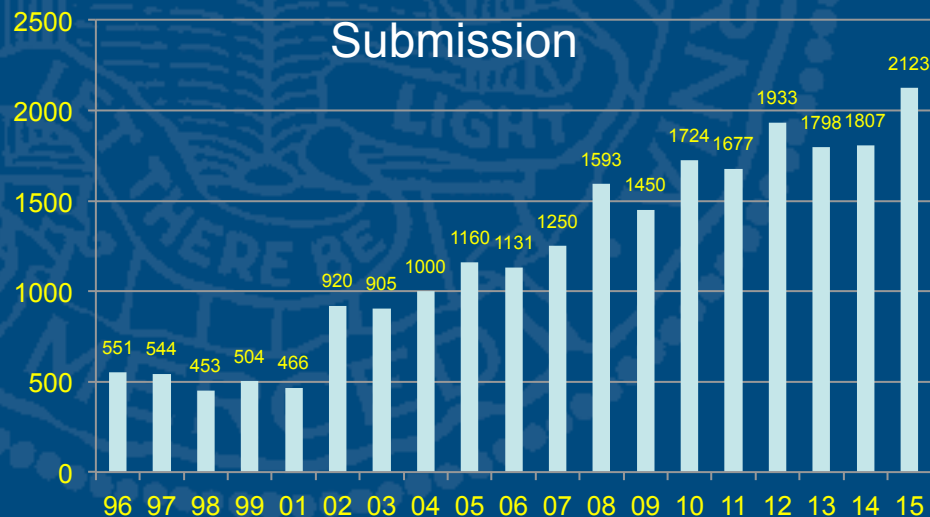
Triage

- Area chairs know the reviewers
- Reviews are weighted
- Based on reviews and rebuttal
 - Accept: (decide oral later)
 - Reject: don't waste time
 - Go either way: lots of papers
- Usually agree with reviewers but anything can happen as long as there are good justifications

Conference Acceptance Rate

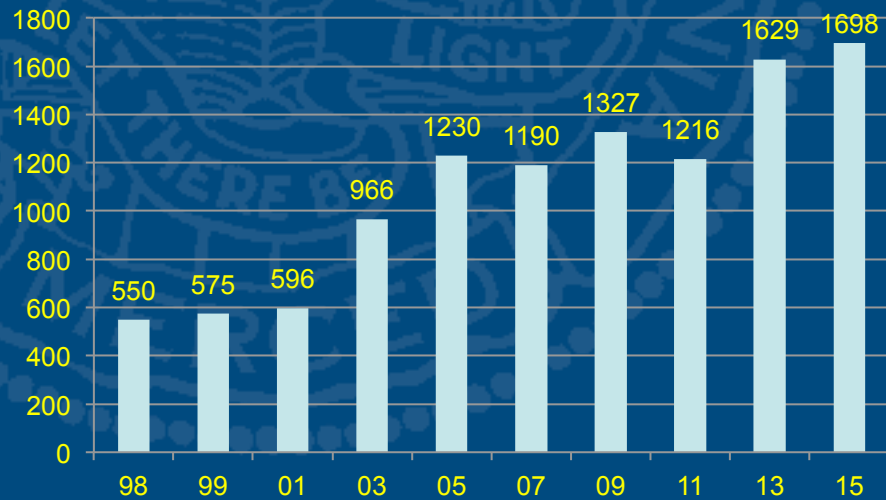
- ICCV/CVPR/ECCV: ~ 25%
- ACCV (2009): ~ 30%
- NIPS: ~ 25%
- BMVC: ~ 30%
- ICIP: ~ 45%
- ICPR: ~ 55%
- Disclaimer
 - low acceptance rate = high quality?

CVPR

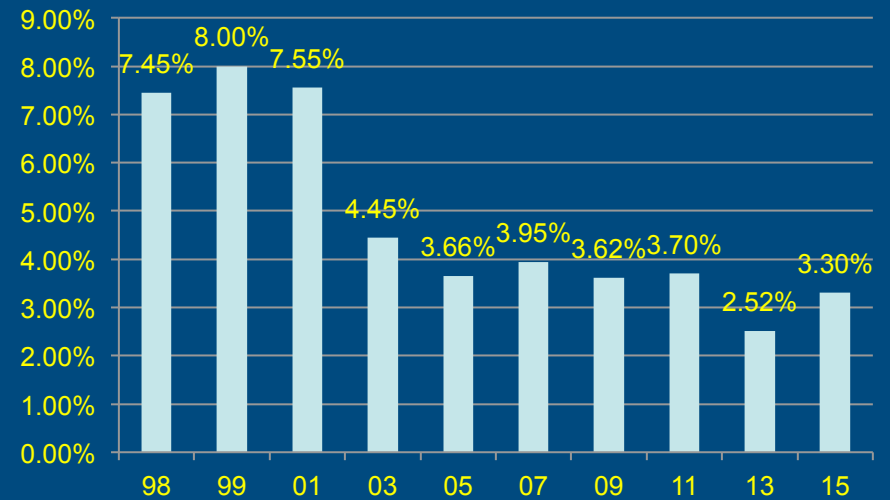


ICCV

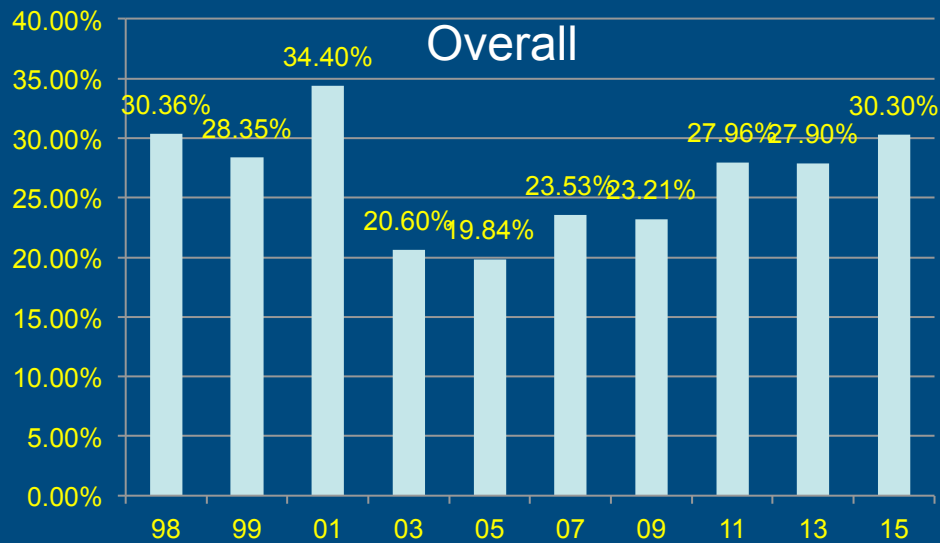
Submission



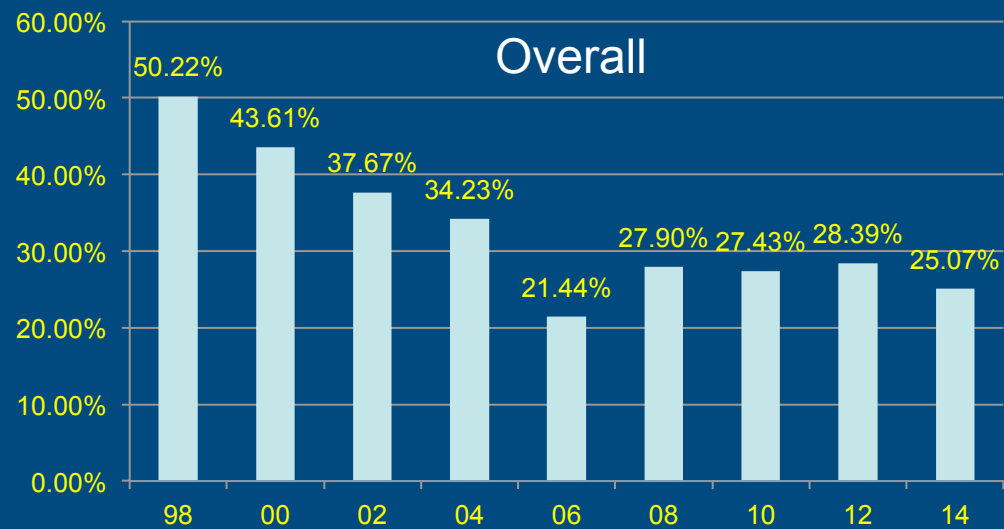
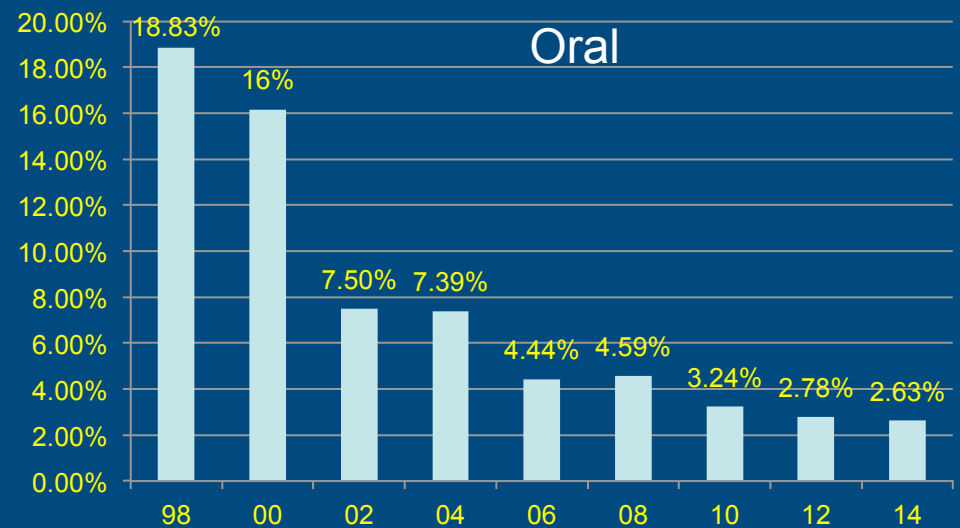
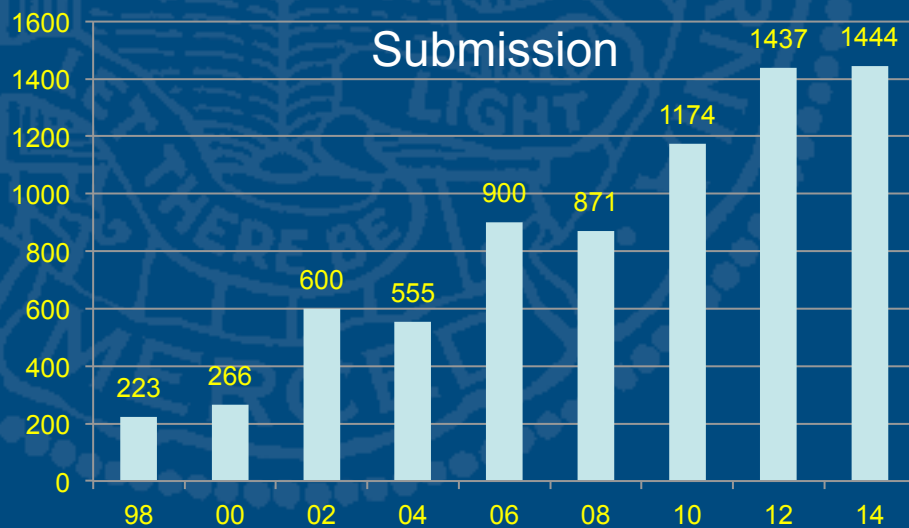
Oral



Overall



ECCV



Top 100 Publications - English

- For what it is worth (h5 index by Google Scholar)

1. Nature

2. The New England Journal of Medicine

3. Science

...

55. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

...

Top Publications - E&CS

1. Nano Letters

...

8. IEEE Conference on Computer Vision
and Pattern Recognition (CVPR)

...

16. IEEE Transactions on Pattern Analysis
and Machine Intelligence

...

Reactions

- Top journal papers
- Workshops vs conferences
- Waiting for the review or final results
- Acceptance
- Reject
- Mixed feeling
- Finding an error
- Resubmit?
- This time, it will go through
- Paper finally accepted
- Registration
- Oral presentation
- Poster presentation

Database Community

- Jeffrey Naughton's ICDE 2010 [keynote](#)
- What's wrong with the reviewing process?
- How to fix that?

Journals

- PAMI – IEEE Transactions on Pattern Analysis and Machine Intelligence, since 1979 (impact factor: 5.96, #1 in all engineering and AI, top-ranked IEEE and CS journal)
- IJCV – International Journal on Computer Vision, since 1988 (impact factor: 5.36, #2 in all engineering and AI)
- CVIU – Computer Vision and Image Understanding, since 1972 (impact factor: 2.20)

Journals

- IVC – Image and Vision Computing
- TIP – IEEE Transactions on Image Processing
- TMI- IEEE Transactions on Medical Imaging
- MVA – Machine Vision and Applications
- PR – Pattern Recognition
- TMM – IEEE Transactions on Multimedia
- ...

PAMI Reviewing Process

- Associate editor-in-chief (AEIC) assigns papers to associate editors (AE)
- AE assigns reviewers
- First-round review: 2-4 months
 - Accept as is
 - Accept with minor revision
 - Major revision
 - Resubmit as new
 - Reject

PAMI Reviewing Process

- Second-round review: 2-4 months
 - Accept as is
 - Accept with minor revision
 - Major revision (rare cases)
 - Reject
- EIC makes final decision
- Overall turn-around time: 6 to 12 months
- Rule of thumb: 30% additional work beyond a CVPR/ICCV/ECCV paper

IJC/CVIU Reviewing Process

- Similar formats
- Slightly longer turn-around time

Journal Acceptance Rate

- PAMI
 - 2013: 151/959: 15.7%
 - 2014: 160/1018: 15.7%
- IJCV: ~ 20% (my guess, no stats)
- CVIU: ~ 25% (my guess, no stats)

From Conferences to Journals

- How much additional work?
 - 30% additional more work for PAMI?
 - As long as the journal version is significantly different from the conference one
- Novelty of each work
 - Some reviewers still argue against this
 - Editors usually accept paper with the same ideas

How to Get Your CVPR Paper Rejected?

- Jim Kajiya (SIGGRAPH 93 papers chair):
[How to get your SIGGRAPH paper rejected?](#)
- Bill Freeman:
[How to write a good CVPR submission](#)
- Do not
 - Pay attention to review process
 - Put yourself as a reviewer to exam your work from that perspective
 - Put the work in right context
 - Carry out sufficient amount of experiments
 - Compare with state-of-the-art algorithms
 - Pay attention to writing

Review Form

- Summary
- Overall Rating
 - Definite accept, weakly accept, borderline, weakly reject, definite reject
- Novelty
 - Very original, original, minor originality, has been done before
- Importance/relevance
 - Of broad interest, interesting to a subarea, interesting only to a small number of attendees, out of CVPR scope

Review Form

- Clarity of presentation
 - Reads very well, is clear enough, difficult to read, unreadable
- Technical correctness
 - Definite correct, probably correct but did not check completely, contains rectifiable errors, has major problems
- Experimental validation
 - Excellent validation or N/A (a theoretical paper), limited but convincing, lacking in some aspects, insufficient validation
- Additional comments
- Reviewer's name

Learn from Reviewing Process

- Learn how others/you can pick apart a paper
- Learn from other's mistakes
- Get to see other reviewers evaluate the same paper
- See how authors rebut comments
- Learn how to write good papers
- Learn what it takes to get a paper published

Put Yourself as Reviewer

- Reviewer's perspective
- How a paper gets rejected?
- What are the contributions?
- Does it advance the science in the field?
- Why you should accept this paper?
- Is this paper a case study?
- Is this paper interesting?
- Who is the audience?

Novelty

- What is new in this work?
 - Higher accuracy, significant speed-up, scale-up, ease to implement, generalization, wide application domain, connection among seemingly unrelated topics, ...
- What are the contributions (over prior art)?
- Make a compelling case with strong supporting evidence

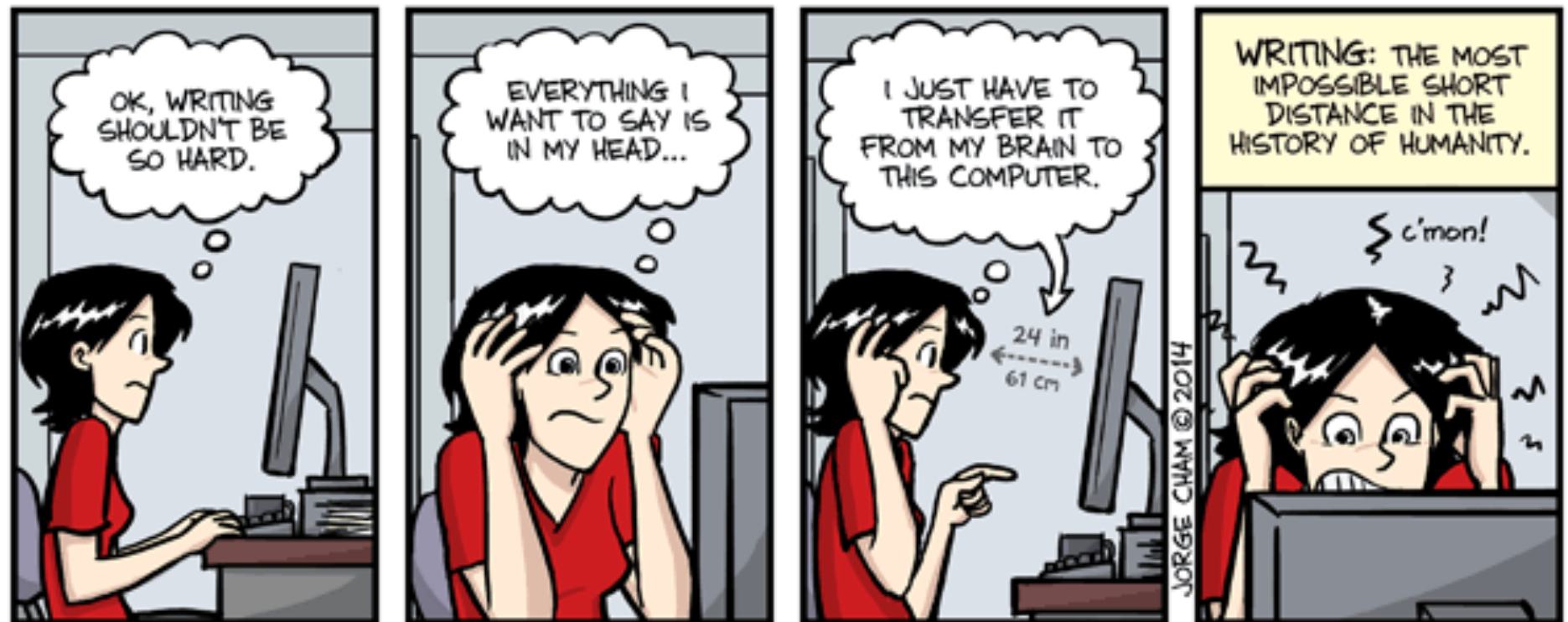
Experimental Validation

- Common data set
- Baseline experiment
- Killer data set
- Large scale experiment
- Evaluation metric
- Realize things after submission
- Friendly fire

Compare With State of the Art

- Do your homework
- Need to know what is out there (and vice versa)
- Need to show why one's method outperforms others, and in what way?
 - speed?
 - accuracy?
 - sensitive to parameters?
 - assumption
 - easy to implement?
 - general application?

Writing



WWW.PHDCOMICS.COM

Writing



WWW.PHDCOMICS.COM

Writing

- Reviewing a poorly written paper
- Clear presentation
- Terse
- Careful about wording
- Make claims with strong evidence

Writing

- Matt Welsh's blog on [scientific writing](#)
- Sharpen your mental focus
- Force you to obsess over every meticulous detail – word choice, word count, overall tone, readability of graphs (and others such as [font size](#), layout and spacing, and [page limit](#))

Writing

- Crystalizing the ideas through the process of putting things together
- Hone the paper to a razor-sharp, articulate, polished work

Writing

- Write the paper *as early as possible*, sometimes *before* even starting the research work
- Will discover the important things that you have not thought about
- The process of writing results in a flood of ideas

Writing

- Even if a paper is not accepted, the process is energizing and often lead to new ideas for the next research problems
- Submitting the paper is often the start of a new line of work
- Riding on that clarity of thought would emerge post-deadline (and a much-needed break)

Tell A Good Story

- Good ideas and convincing results
- But not too much (vs grant proposal)



Presentation

- Good artists copy, great artists steal
- Not just sugar coating
- Not just a good spin
- Tell a convincing story with solid evidence
- Present your ideas with style
- Q&A
- Real stories

Interesting Title

- Cool titles attract people
- Grab people's attention
- Buzz word?
- But don't be provocative

Math Equations

- Minimal number of equations
 - No more, no less
 - Too many details simply make a paper inaccessible
- Too few equations
- Many good papers have no or few equations
 - CVPR 13 best paper
 - CVPR 05 HOG paper

Figures

- Be clear
- Sufficient number of figures

Theoretical or Applied?

- Computer vision is more applied, at least nowadays
- Theory vs real world
- More high impact papers are about how to get things done right

Common Mistakes

- Typos
- Unsupported claims
- Unnecessary adjectives (superior!)
- “a”, “the”
- Inanimate objects with verbs
- Inconsistent usage of words
- Laundry list of related work (or worse copy sentences from abstracts)
- Bad references
- Laundry list of related work
- Repeated boring statements

Get Results First than Writing?

- Conventional mode
 - Idea-> Do research -> Write paper
- “How to write a great research paper” by Simon Peyton Jones
 - Idea -> Write paper -> Do research
 - Forces us to be clear, focused
 - Crystallizes what we don't understand
 - Opens the way to dialogue with others: reality check, critique, and collaboration
- My take
 - Idea -> Write paper -> Do research -> Revise paper -> Do research -> Revise paper -> ...

Supplementary Material

- Important
- Add more results and large figures
- Add technical details as necessary (don't miss important details)
- Derivation details, e.g., proof of a theorem

Most Important Factors

- Novelty
- Significant contributions (vs. [salami publishing](#))
- Make sure your paper is non-rejectable (above the bar with some error margin)

Reviews

- Me: Here is a faster horse
- R1: You should have used my donkey
- R2: This is not a horse, it's a mule
- R3: I want a unicorn!

Rebuttal or Response

ADDRESSING REVIEWER COMMENTS

BAD REVIEWS ON YOUR PAPER? FOLLOW THESE GUIDELINES AND YOU MAY YET GET IT PAST THE EDITOR:

Reviewer comment:

"The method/device/paradigm the authors propose is clearly wrong."

How NOT to respond:

✗ "Yes, we know. We thought we could still get a paper out of it. Sorry."

Correct response:

✓ "The reviewer raises an interesting concern. However, as the focus of this work is exploratory and not performance-based, validation was not found to be of critical importance to the contribution of the paper."

Reviewer comment:

"The authors fail to reference the work of Smith et al., who solved the same problem 20 years ago."

How NOT to respond:

✗ "Huh. We didn't think anybody had read that. Actually, their solution is better than ours."

Correct response:

✓ "The reviewer raises an interesting concern. However, our work is based on completely different first principles (we use different variable names), and has a much more attractive graphical user interface."

Reviewer comment:

"This paper is poorly written and scientifically unsound. I do not recommend it for publication."

How NOT to respond:

✗ "You #&@*% reviewer! I know who you are! I'm gonna get you when it's my turn to review!"

Correct response:

✓ "The reviewer raises an interesting concern. However, we feel the reviewer did not fully comprehend the scope of the work, and misjudged the results based on incorrect assumptions."

www.phdcomics.com

JORGE CHAM © 2005

Good surprise

- One CVPR paper: BR, BR, DR
- Two ECCV paper: PR, PR, BR
- One CVPR 15 paper: BR, BR, WR -> poster, poster, poster
- One CVPR 15 paper: DR, WA, BR -> Poster, Poster, WR

Bad surprise

- Two ECCV papers: PA, PA, BR
- One CVPR 15 paper: WA, BR, BR -> Poster, Poster, WR
- One CVPR 16 paper: WR, WR, BR

Never Know What will Happen

Masked Meta-Reviewer ID: Meta_Reviewer_1

Meta-Reviews:

Question

Consolidation Report

All reviewers agree that this paper has moderate novelty of using partial and spatial information for sparse representation. However, they also concern about

- unclear presentation on technical details (eg. definitions, inference algorithm, pooling methods, template updating schemes, experimental settings etc.),
- not extensive experimental comparison (needs tests on more challenging videos),
- missing justification of the assumption (complementary nature of two kinds of pooling features) and the efficacy of each term.

The authors rebuttal addresses most issues, but is not sufficient to ease the main concerns of R1 and R2. So, the AC recommends the paper to be rejected as it is.

Decision

Definitely Accept

Challenging Issues

- Large scale
 - CVPR 2011 best paper: pose estimation
 - CVPR 2013 best paper: object detection
- Unconstrained
- Real-time
 - CVPR 2001: face detector
 - CVPR 2006: scalable object recognition
- Robustness
- Recover from failure

Interesting Stats

- Best papers and top cited papers in computer science
- Best papers = high impact?
- Oral papers are more influential?
- CVPR Longuet-Higgins prize
- ICCV Helmholtz award

Data Set Selection

- NIPS 02 by Doudou LaLoudouana and Mambobo Bonouliqui Tarare, Lupano Tecallonou Center, Selacie, Guana
- The secret to publish a paper in machine learning conferences?
- Read the references therein carefully!

Data Set Selection

Data Set Selection

Danilo LaLondeanu¹ and Manohar Bhandalapati Tarure²
Lupata Teleshon Center
Salerno, GI, USA
danilo@laldeanu.com, tarure@teleshon.com

Abstract

We introduce the community to a new construction principle whose practical implications are very broad. Central to this research is the idea to improve the presentation of algorithms in the literature and to make them more appealing. We define a new notion of capacity for data sets and derive a methodology for selecting from them. The experiments show that even for not so good algorithms, you can show that they are significantly better than all the others. We give some experimental results, which are very promising.

1 Introduction

Learning is a mind-blowing subject. A year spent in artificial intelligence is enough to make one believe in God. Unfortunately, so far it has been handled only from a particular one-sided point of view. The VC-theory reports only for some people does not offer what we would be in right to ask from such a theory: we want good bounds for our algorithms. We offer with this article a broad new approach that allows you to present your algorithms in a much more principled and rigorous way than has been done before. Many researchers, especially in publications at NIPS, have tried to show when their algorithms are better (in some sense) than some other given set of algorithms. It is always very hard to compare hypotheses of data set selection. It is strange then that learning theorists, as they call themselves, even the last 20 years have concentrated on the model selection problem and not the data selection problem which is what people actually do. The two problems can in some sense be seen as the dual of each other, but it is not obvious to you one that you solve for the other one. And vice-versa, in this article we lay down the foundations and introduce to the community of machine learning people a new subjective principle: structural dataset minimization. Essentially we begin to formalize the connection of how engineering approach of the selection procedure that everyone should provide. In doing so we find concrete bounds for when the data selected really is better than other datasets and implement this and how algorithms for finding such datasets. We then use approach-independent the classical approach.

The structure of the paper continues to its content follows a classical trend: section 1 presents some the bounds you can see in lots of situations, section 2 shows how to use these bounds by choosing your algorithm. Section 3 describes some experiments which, of course, are good. Section 4 concludes the article with must thoughts and future work we will do.¹

2 Bounds

Let us introduce some notations. Assume a researcher has invented an algorithm A^* and he wishes to show that this policy and \mathcal{D} are superior with respect to some bad policy \mathcal{C} to a certain set of other people, we have put all of our experiments in this paper, even the bad ones, but we did not get any bad ones.²

¹And only if the paper is accepted.

²And only if the paper is accepted.

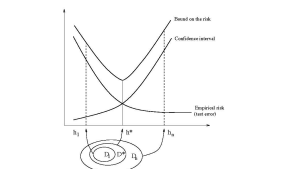


Figure 1: Graphical depiction of the structural data sets minimization construction principle.

The highly successful and efficient algorithm presented in the following. Once again, assume a researcher has invented³ an algorithm A^* and he wishes to show that it is superior with respect to some loss function ℓ to a given fixed set of algorithms A_1, \dots, A_n , that other researchers have made. Let us try to choose a fixed class of artificial datasets \mathcal{D} , for example artificial data that is generated by a mixture of a Gaussian with covariance matrices $\Sigma_1, \dots, \Sigma_n$ in a d -dimensional space with a noise model ϵ and ℓ is a loss function. Let us define $w = (w_1, \dots, w_n)$, we have the following theorem:

Theorem 1 The capacity $\mathcal{H}(\mathcal{D})$ is bounded by

$$\mathcal{H}(\mathcal{D}) \leq \min \{ \ell^2(w^T, w) \}$$

where n is the number of parameters (sum of dimensions of the vectors c, d, ϵ and f) and R is the largest possible value of any coordinate of any vector in the data sets.

Some people would argue that this theorem is valid only for data sets parameterized by gaussian distribution, but actually it is one kind of a real life problem. The interested reader can read the literature about gaussian processes.

Note, the task is to optimize these parameters such that the algorithm A^* appears much better than the other ones. Let us denote the loss function ℓ as $\ell(w, A_1, \dots, A_n, \Sigma_1, \dots, \Sigma_n, \epsilon, d, f)$ and we could this set of algorithms with a uniform distribution to choose our bias.

This can be done with the following minimization:

$$\min_{w \in \mathcal{D}} \ell(w, A_1, \dots, A_n, \Sigma_1, \dots, \Sigma_n, \epsilon, d, f)$$

which is equivalent to

$$\min_{w \in \mathcal{D}} \ell(w, A_1, \dots, A_n, \Sigma_1, \dots, \Sigma_n, \epsilon, d, f)$$

¹And only if the paper is accepted.

²And only if the paper is accepted.

³And only if the paper is accepted.

given fixed set of algorithms⁴ A_1, \dots, A_n , that other researchers have made. For this purpose, the researcher selects some data sets using what is called an empirical data of minimization method. The latter consists in taking some fixed set of data sets D_1, \dots, D_n and find a data set \mathcal{D} in D_1, \dots, D_n so that

$$\ell(A^*, \mathcal{D}) < \min_{i=1, \dots, n} \ell(A_i, \mathcal{D})$$

Note that this problem is ill-posed. A natural generalization would be to find more than one data set in which your algorithm performs well but this is a difficult problem that has not been solved so far by the community. Current efforts in solving this problem have focused on producing more artificial data sets rather than algorithms to achieve this goal.

We have the following theorem:

Theorem 2 Let \mathcal{D} be a set of training sets, then means that the given set of algorithms is robust with a fixed distribution P (which could be anything a priori), then with probability $1 - \eta$ over a sampling in the algorithm, and for all $i = 1, \dots, n$, we have:

$$\forall D \in \mathcal{D}, R_{\text{train}}(D) \leq R_{\text{test}}(D) + \sqrt{\frac{\mathcal{H}(\mathcal{D})}{m} \log(n)}$$

where $\mathcal{H}(\mathcal{D})$ is the capacity of the set of training sets defined as:

$$\mathcal{H}(\mathcal{D}) = \max_{\{A_1, \dots, A_n\} \in \mathcal{A}} \min_{D \in \mathcal{D}} \sum_{i=1}^n \ell(A_i, D) \leq \mathcal{H}(\mathcal{D})$$

We are proud now to supply the following elegant proof.

Proof. Let us denote by m the number of points in the testing set, we can see that introducing a priori algorithm A^* :

$$P_A \left[\sup_{D \in \mathcal{D}} |R_{\text{train}}(D) - R_{\text{test}}(D)| > \epsilon \right] \leq P_A \left[\sup_{D \in \mathcal{D}} |R_{\text{train}}(D) - R_{\text{test}}(D)| > \epsilon \right]$$

which is trivially insensitive to permutations so that we can condition on the algorithm A^* and \mathcal{D} . Thus we also have the right to play with the swapping permutation so it has been done in the theoretical but practically we used VC framework, which means that we work only with the values of $\ell(A_i, \mathcal{D})$. After some more notes which we admit for brevity, this leads to the vanishing of the supremum. We are then left with a sum of two random variables whose sum can be controlled using the Bernstein-Hoeffding inequality. This is here where the tricky part begins. It is known that averaging over two random variables does not give you a good control of their expectation. But this can be overcome if we consider many exact copies of the first two variables. Then we have plenty of them, as much as we want. And we can then control the expectation because now the value of m is big. We call this trick the replica trick. Note the replica trick has been used more times in the literature of machine learning, the most of all possible algorithms, the same algorithm has been varied many times with multiple variations so that if you are an algorithmic function, these algorithms appear to be equivalent.⁵

The theorem we just proved should be considered as the dual of the theorem of Vapnik and Chervonenkis. And this should be the case because it is just the dual of it. And we believe this is the more natural setting for your every day design of algorithms. Maybe our approach is complementary to the one of Vladimir and Alexei but we are one step forward because we can infer/compute the probability over the data sets just by looking at the VCJ memory database. Just try to do the same with your set of functions and we will talk. Anyway, we insist that our approach allows a lot of common properties with the classical VC framework. One of them is that unfortunately we cannot say much but we try to, so we are different we have our own "no free lunch" theorem but we try to keep it. Here is our first theorem:⁶

Theorem 3 A result set \mathcal{D} that is as good as the best two experiments.⁷

¹And only if the paper is accepted.

²And only if the paper is accepted.

³And only if the paper is accepted.

⁴And only if the paper is accepted.

⁵And only if the paper is accepted.

⁶And only if the paper is accepted.

⁷And only if the paper is accepted.

⁸And only if the paper is accepted.

⁹And only if the paper is accepted.

¹⁰And only if the paper is accepted.

¹¹And only if the paper is accepted.

¹²And only if the paper is accepted.

¹³And only if the paper is accepted.

¹⁴And only if the paper is accepted.

¹⁵And only if the paper is accepted.

¹⁶And only if the paper is accepted.

¹⁷And only if the paper is accepted.

¹⁸And only if the paper is accepted.

¹⁹And only if the paper is accepted.

²⁰And only if the paper is accepted.

²¹And only if the paper is accepted.

²²And only if the paper is accepted.

²³And only if the paper is accepted.

²⁴And only if the paper is accepted.

²⁵And only if the paper is accepted.

²⁶And only if the paper is accepted.

²⁷And only if the paper is accepted.

²⁸And only if the paper is accepted.

²⁹And only if the paper is accepted.

³⁰And only if the paper is accepted.

³¹And only if the paper is accepted.

³²And only if the paper is accepted.

³³And only if the paper is accepted.

³⁴And only if the paper is accepted.

³⁵And only if the paper is accepted.

³⁶And only if the paper is accepted.

³⁷And only if the paper is accepted.

³⁸And only if the paper is accepted.

³⁹And only if the paper is accepted.

⁴⁰And only if the paper is accepted.

⁴¹And only if the paper is accepted.

⁴²And only if the paper is accepted.

⁴³And only if the paper is accepted.

⁴⁴And only if the paper is accepted.

⁴⁵And only if the paper is accepted.

⁴⁶And only if the paper is accepted.

⁴⁷And only if the paper is accepted.

⁴⁸And only if the paper is accepted.

⁴⁹And only if the paper is accepted.

⁵⁰And only if the paper is accepted.

⁵¹And only if the paper is accepted.

⁵²And only if the paper is accepted.

⁵³And only if the paper is accepted.

⁵⁴And only if the paper is accepted.

⁵⁵And only if the paper is accepted.

⁵⁶And only if the paper is accepted.

⁵⁷And only if the paper is accepted.

⁵⁸And only if the paper is accepted.

⁵⁹And only if the paper is accepted.

⁶⁰And only if the paper is accepted.

⁶¹And only if the paper is accepted.

⁶²And only if the paper is accepted.

⁶³And only if the paper is accepted.

⁶⁴And only if the paper is accepted.

⁶⁵And only if the paper is accepted.

⁶⁶And only if the paper is accepted.

⁶⁷And only if the paper is accepted.

⁶⁸And only if the paper is accepted.

⁶⁹And only if the paper is accepted.

⁷⁰And only if the paper is accepted.

⁷¹And only if the paper is accepted.

⁷²And only if the paper is accepted.

⁷³And only if the paper is accepted.

⁷⁴And only if the paper is accepted.

⁷⁵And only if the paper is accepted.

⁷⁶And only if the paper is accepted.

⁷⁷And only if the paper is accepted.

⁷⁸And only if the paper is accepted.

⁷⁹And only if the paper is accepted.

⁸⁰And only if the paper is accepted.

⁸¹And only if the paper is accepted.

⁸²And only if the paper is accepted.

⁸³And only if the paper is accepted.

⁸⁴And only if the paper is accepted.

⁸⁵And only if the paper is accepted.

⁸⁶And only if the paper is accepted.

⁸⁷And only if the paper is accepted.

⁸⁸And only if the paper is accepted.

⁸⁹And only if the paper is accepted.

⁹⁰And only if the paper is accepted.

⁹¹And only if the paper is accepted.

⁹²And only if the paper is accepted.

⁹³And only if the paper is accepted.

⁹⁴And only if the paper is accepted.

⁹⁵And only if the paper is accepted.

⁹⁶And only if the paper is accepted.

⁹⁷And only if the paper is accepted.

⁹⁸And only if the paper is accepted.

⁹⁹And only if the paper is accepted.

¹⁰⁰And only if the paper is accepted.

Theorem 2 (No Free Lunch!) The generalization error of two datasets for all algorithms is the same:

$$E(R_{\text{train}}(D)) = E(R_{\text{test}}(D))$$

so that there is no better dataset than the one you already have.

The consequences of the theorem are very harsh: it means that if you don't do well, then you are not very skilled. We will have not worked out what the researchers had for breakfast (of course, in the remark we retain some liberty). This leads to the natural consequence that to say anything one should restrict the set of algorithms to some natural set, in a companion paper we prove that the set of datasets restricted to all Bayesian algorithms has no uniform improvement. The same is true for the set of all neural algorithms if you leave the following four parameters: loss function (linear loss, squared loss, quadratic loss), kernel (linear, polynomial, Gaussian, Laplace, Gaussian, etc.), regularization (L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11, L12, L13, L14, L15, L16, L17, L18, L19, L20, L21, L22, L23, L24, L25, L26, L27, L28, L29, L30, L31, L32, L33, L34, L35, L36, L37, L38, L39, L40, L41, L42, L43, L44, L45, L46, L47, L48, L49, L50, L51, L52, L53, L54, L55, L56, L57, L58, L59, L60, L61, L62, L63, L64, L65, L66, L67, L68, L69, L70, L71, L72, L73, L74, L75, L76, L77, L78, L79, L80, L81, L82, L83, L84, L85, L86, L87, L88, L89, L90, L91, L92, L93, L94, L95, L96, L97, L98, L99, L100, L101, L102, L103, L104, L105, L106, L107, L108, L109, L110, L111, L112, L113, L114, L115, L116, L117, L118, L119, L120, L121, L122, L123, L124, L125, L126, L127, L128, L129, L130, L131, L132, L133, L134, L135, L136, L137, L138, L139, L140, L141, L142, L143, L144, L145, L146, L147, L148, L149, L150, L151, L152, L153, L154, L155, L156, L157, L158, L159, L160, L161, L162, L163, L164, L165, L166, L167, L168, L169, L170, L171, L172, L173, L174, L175, L176, L177, L178, L179, L180, L181, L182, L183, L184, L185, L186, L187, L188, L189, L190, L191, L192, L193, L194, L195, L196, L197, L198, L199, L200, L201, L202, L203, L204, L205, L206, L207, L208, L209, L210, L211, L212, L213, L214, L215, L216, L217, L218, L219, L220, L221, L222, L223, L224, L225, L226, L227, L228, L229, L230, L231, L232, L233, L234, L235, L236, L237, L238, L239, L240, L241, L242, L243, L244, L245, L246, L247, L248, L249, L250, L251, L252, L253, L254, L255, L256, L257, L258, L259, L260, L261, L262, L263, L264, L265, L266, L267, L268, L269, L270, L271, L272, L273, L274, L275, L276, L277, L278, L279, L280, L281, L282, L283, L284, L285, L286, L287, L288, L289, L290, L291, L292, L293, L294, L295, L296, L297, L298, L299, L300, L301, L302, L303, L304, L305, L306, L307, L308, L309, L310, L311, L312, L313, L314, L315, L316, L317, L318, L319, L320, L321, L322, L323, L324, L325, L326, L327, L328, L329, L330, L331, L332, L333, L334, L335, L336, L337, L338, L339, L340, L341, L342, L343, L344, L345, L346, L347, L348, L349, L350, L351, L352, L353, L354, L355, L356, L357, L358, L359, L360, L361, L362, L363, L364, L365, L366, L367, L368, L369, L370, L371, L372, L373, L374, L375, L376, L377, L378, L379, L380, L381, L382, L383, L384, L385, L386, L387, L388, L389, L390, L391, L392, L393, L394, L395, L396, L397, L398, L399, L400, L401, L402, L403, L404, L405, L406, L407, L408, L409, L410, L411, L412, L413, L414, L415, L416, L417, L418, L419, L420, L421, L422, L423, L424, L425, L426, L427, L428, L429, L430, L431, L432, L433, L434, L435, L436, L437, L438, L439, L440, L441, L442, L443, L444, L445, L446, L447, L448, L449, L450, L451, L452, L453, L454, L455, L456, L457, L458, L459, L460, L461, L462, L463, L464, L465, L466, L467, L468, L469, L470, L471, L472, L473, L474, L475, L476, L477, L478, L479, L480, L481, L482, L483, L484, L485, L486, L487, L488, L489, L490, L491, L492, L493, L494, L495, L496, L497, L498, L499, L500, L501, L502, L503, L504, L505, L506, L507, L508, L509, L510, L511, L512, L513, L514, L515, L516, L517, L518, L519, L520, L521, L522, L523, L524, L525, L526, L527, L528, L529, L530, L531, L532, L533, L534, L535, L536, L537, L538, L539, L540, L541, L542, L543, L544, L545, L546, L547, L548, L549, L550, L551, L552, L553, L554, L555, L556, L557, L558, L559, L560, L561, L562, L563, L564, L565, L566, L567, L568, L569, L570, L571, L572, L573, L574, L575, L576, L577, L578, L579, L580, L581, L582, L583, L584, L585, L586, L587, L588, L589, L590, L591, L592, L593, L594, L595, L596, L597, L598, L599, L600, L601, L602, L603, L604, L605, L606, L607, L608, L609, L610, L611, L612, L613, L614, L615, L616, L617, L618, L619, L620, L621, L622, L623, L624, L625, L626, L627, L628, L629, L630, L631, L632, L633, L634, L635, L636, L637, L638, L639, L640, L641, L642, L643, L644, L645, L646, L647, L648, L649, L650, L651, L652, L653, L654, L655, L656, L657, L658, L659, L660, L661, L662, L663, L664, L665, L666, L667, L668, L669, L670, L671, L672, L673, L674, L675, L676, L677, L678, L679, L680, L681, L682, L683, L684, L685, L6

Data Set Selection

Data Set Selection

Doudou LaLoudouana* and Mambobo Bonouliqui Tarare

Lupano Tecallonou Center

Selacie, GUANA

doudoula3@hotmail.com, fuzzybear@yahoo.com

Abstract

We introduce the community to a new construction principle whose practical implications are very broad. Central to this research is the idea to improve the presentation of algorithms in the literature and to make them more appealing. We define a new notion of capacity for data sets and derive a methodology for selecting from them. The experiments show that even for not so good algorithms, you can show that they are significantly better than all the others. We give some experimental results, which are very promising.

Data Set Selection

References

- [1] R. Shapire, Y. Freund, P. Bartlett, and W.S. Lee. *Bushing the margin: A new explanation for the effectiveness of U.S voting methods.* *The Anals of Statistics*, 1998.
- [2] N. Cristianini and J. Shawe-Taylor. *Data Set Selection for Dummies.* Egham University Press, 2005.
- [3] P. Bartlett. *The Sample Complexity of Pattern Classification with Neural Networks: the Number of Citations is More Important than the Number of Readers.* *Biowulf Transactions* , 1998.
- [4] D. MacKay. *I did it my way.* *Self published*, 2002.
- [5] G. Fung and O. L. Mangasarian. Data Selection for Support Vector Machine Classifiers. *Proceedings of KDD'2000*, 2000.
- [6] J. Friedman, T. Hastie, and R. Tibshirani. *Data sets for additive models: a statistical view of bragging and boasting.* Technical report, Technical Report, Department of Statistics, Stanford University., 1998.
- [7] C.M. Bishop. *Neural Not works for Pattern Recognition*, Oxford University Press, 1995
- [8] D.H. Wolpert, *The lack of a priori distinctions between learning algorithm research,* Neural Computation, 1996.
- [9] V.N. Vapnik and A.Y Chervonenkis and S. Barnhill, *The VC Way: Investment Secrets from the Wizards of Venture Capital*, Biowulf Publishing (now out of print) , 2001.

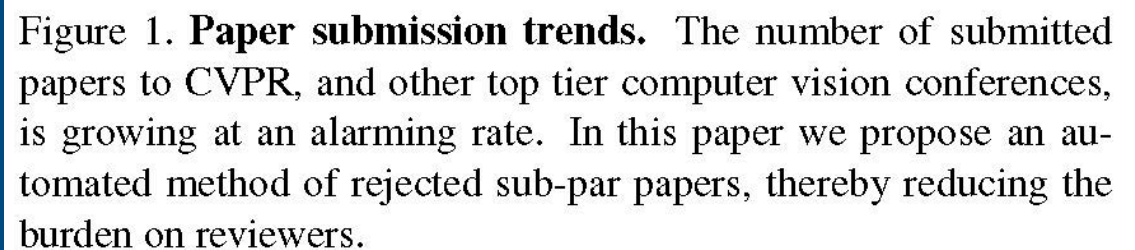
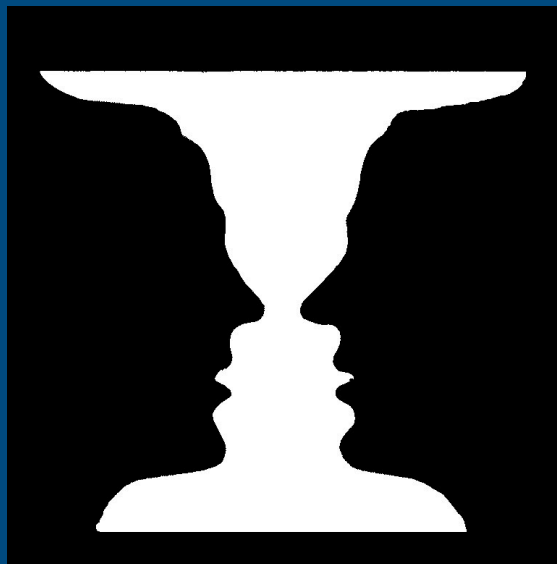
(originally) [6] ... a egotistical view of bragging and boasting..... UCMERGED

Where Is My Advisor?

Ask Someone to Proofread

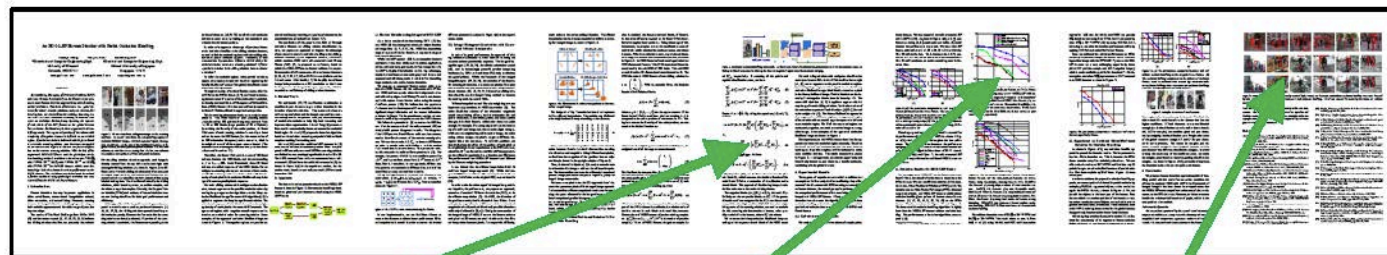
- Certainly your advisor
- Polish your work
- My story

A black and white photograph of a dalmatian dog standing in a field of tall grass or reeds. The dog is facing right, looking down at the ground. The background is a dense field of tall, thin stalks, possibly reeds or grass, creating a textured, somewhat abstract pattern. The lighting is bright, casting shadows on the ground.



Paper Gestalt

- CVPR 10 by Carven von Bearnensquash,
Department of Computer Science,
University of Phoenix
- Main Point: Get your paper looking pretty with right mix of equations, tables and figures

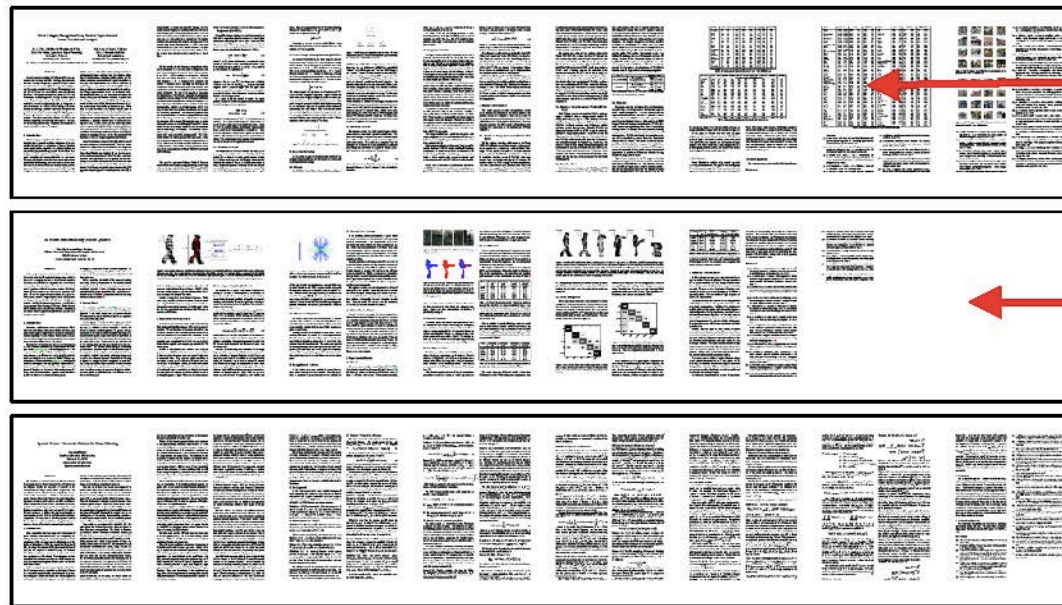


Math: Sophisticated mathematical expressions make a paper look technical and make the authors appear knowledgeable and “smart”.

Plots: ROC, PR, and other performance plots convey a sense of thoroughness. Standard deviation bars are particularly pleasing to a scientific eye.

Figures/Screenshots: Illustrative figures that express complex algorithms in terms of 3rd grade visuals are always a must. Screenshots of anecdotal results are also very effective.

Figure 6. Characteristics of a “Good” paper.



Large confusing tables.

Missing pages.

Lack of colorful figures.

Figure 7. Characteristics of a “Bad” paper.

Tools

- [Google scholar](#) h-index
- Software: [publish or perish](#)
- DBLP
- Mathematics genealogy
- Disclaimer:
 - h index = significance?
 - # of citation = significance?

Basic Rules

- Use LaTeX
- Read authors' guideline
- Read reviewers' guideline
- Print out your paper – what you see may NOT be what you get
- Submit paper right before deadline
 - Risky
 - Exhausting
 - Murphy's law
- Do not count on extension

Lessons

- Several influential papers have been rejected once or twice
- Some best papers make little impact
- Never give up in the process

Karma?

WHAT YOU THINK OF YOUR PROFESSOR vs. TIME



JORGE CHAM © 2014

WWW.PHDCOMICS.COM

Your Advisor and You

- Suggesting a research topic
- When your advisor presents your work
- When you explain your work
- Demos
- Good results

Start Working Early!

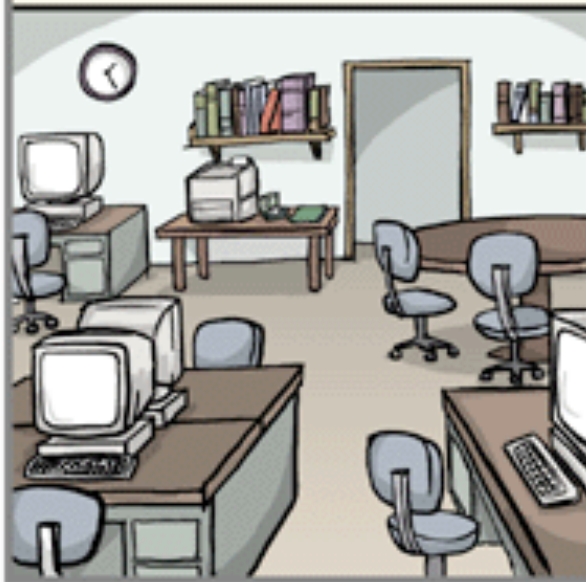
- Write, write, write...
- Ask others for comments

SUMMER DAYS...

THE LAB: 1 DAY AFTER ADVISOR LEAVES FOR VACATION.



THE LAB: 2 DAYS AFTER ADVISOR LEAVES FOR VACATION.



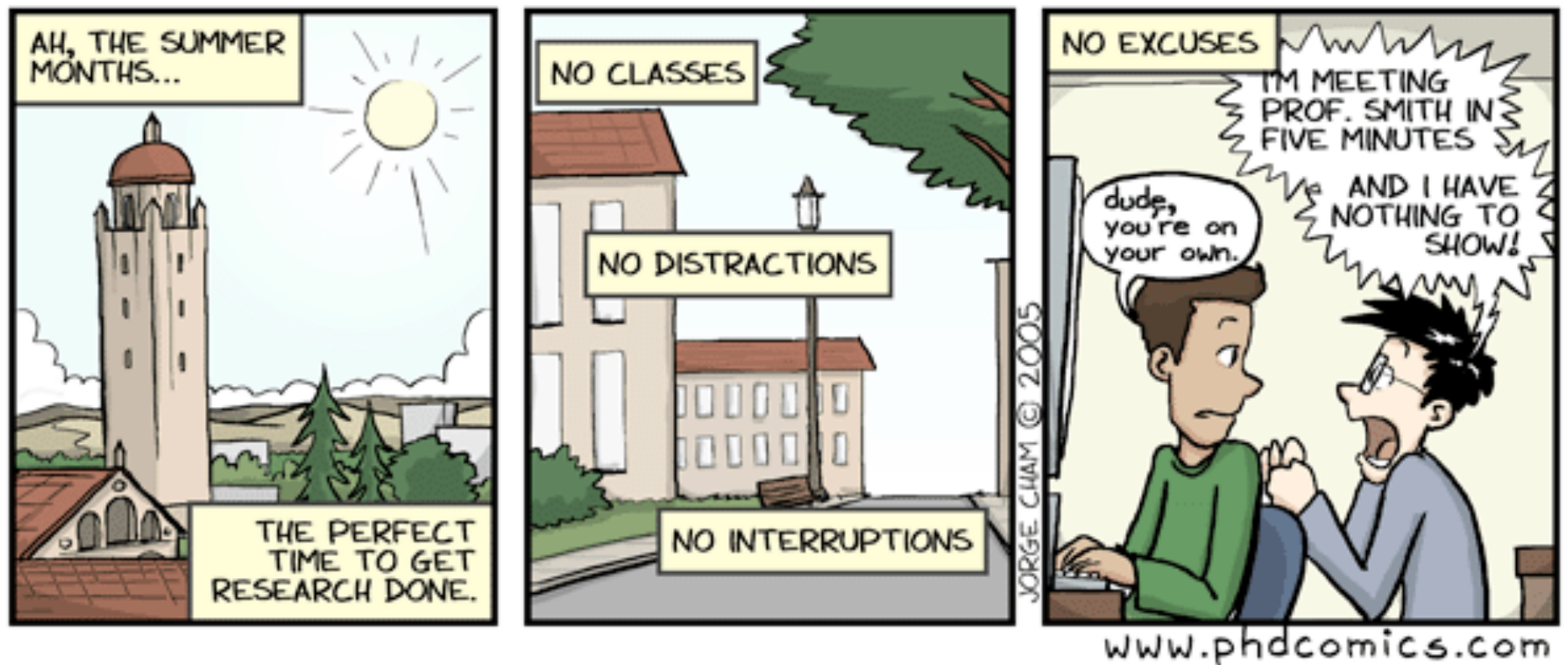
THE LAB: 1 DAY BEFORE ADVISOR COMES BACK FROM VACATION.



JORGE CHAM © 2004

www.phdcomics.com

Work Hard in the Summer



Quotes from Steve Jobs

- “ I'm convinced that about half of what separates successful entrepreneurs from the non-successful ones is pure perseverance. ”
- “ Creativity is just connecting things. When you ask creative people how they did something, they feel a little guilty because they didn't really do it, they just saw something. It seemed obvious to them after a while. ”