Dissect an ML algorithm for potential forms of interactivity

Krzysztof Gajos

HARVARD
School of Engineering
and Applied Sciences

Dissect interactivity for potential ML algorithm

Krzysztof Gajos

HARVARD
School of Engineering
and Applied Sciences

Hall of Fame?

Hall of Shame?

CueTIP

A Mixed-Initiative Interface for Correcting Handwriting Recognition Errors

Michael Shilman, Desney S. Tan, Patrice Simard

http://research.microsoft.com/en-us/um/people/desney/publications/UIST2006-CueTIP.mov

Today’s Mantra

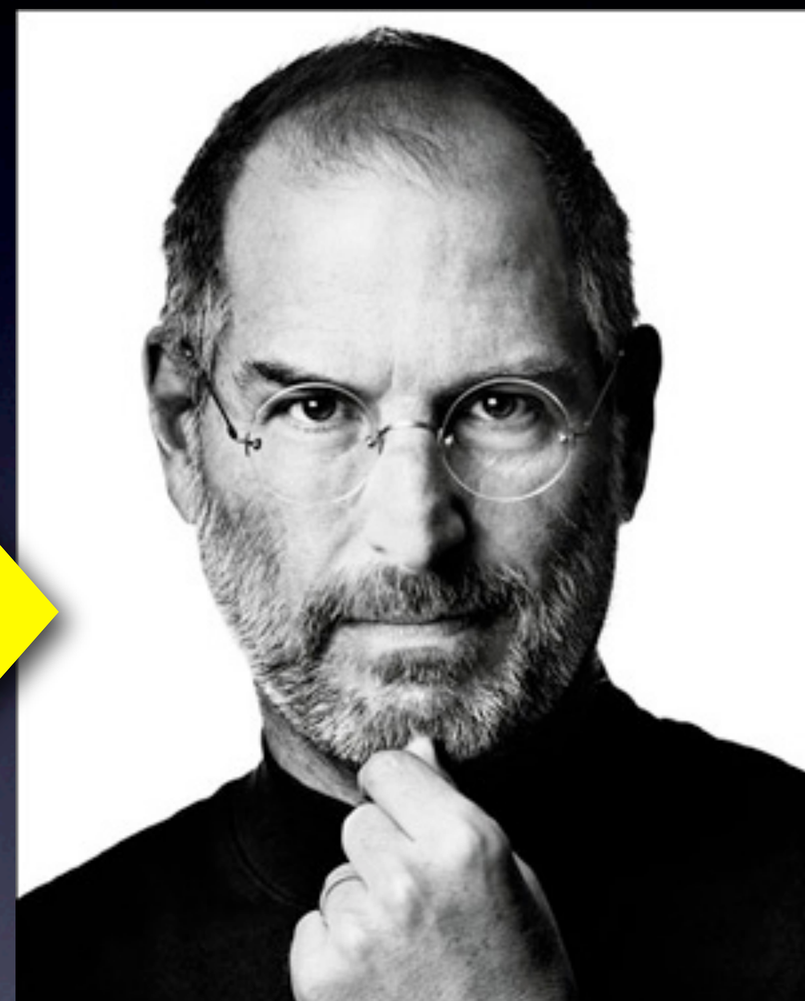
- Pick appropriate interaction for the task
- Understand the semantics of the interaction
- Develop/pick an algorithm that uses available information correctly and efficiently

Supple Project: Automatically Generating User Interfaces

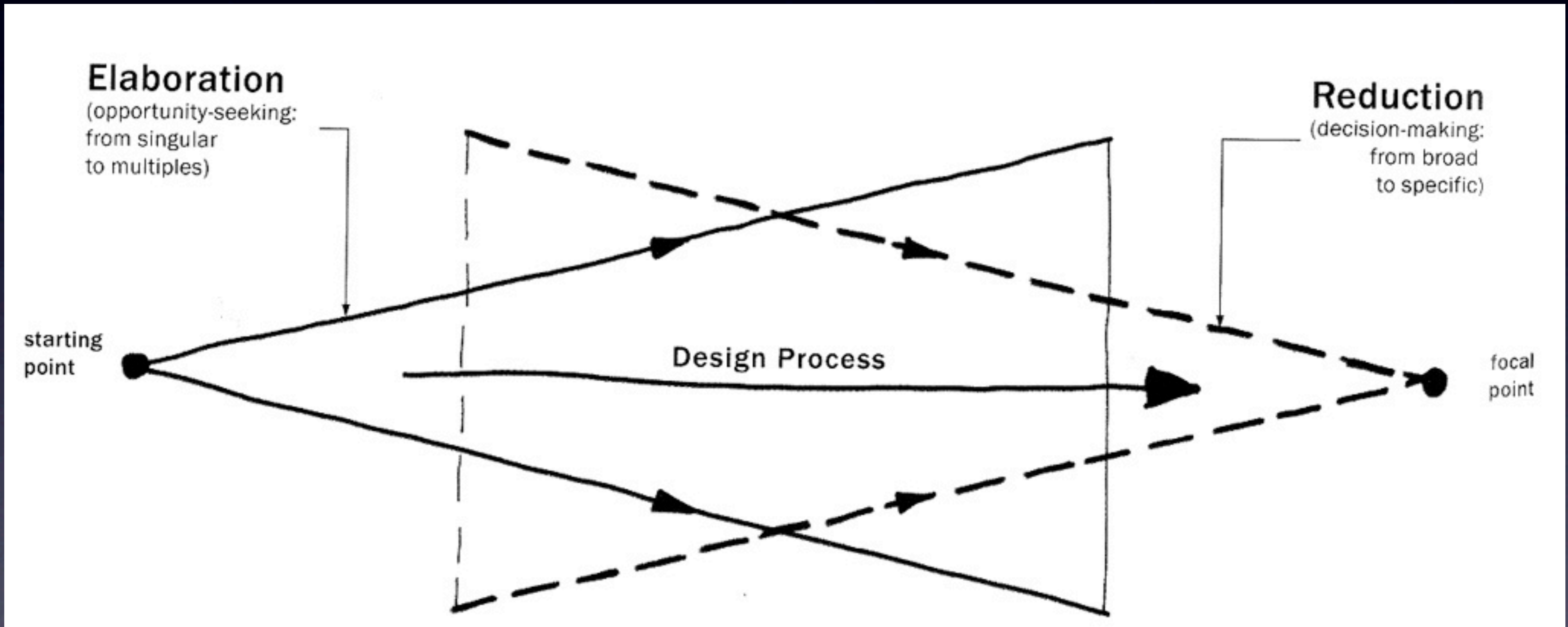
DESIGN

Design by Genius

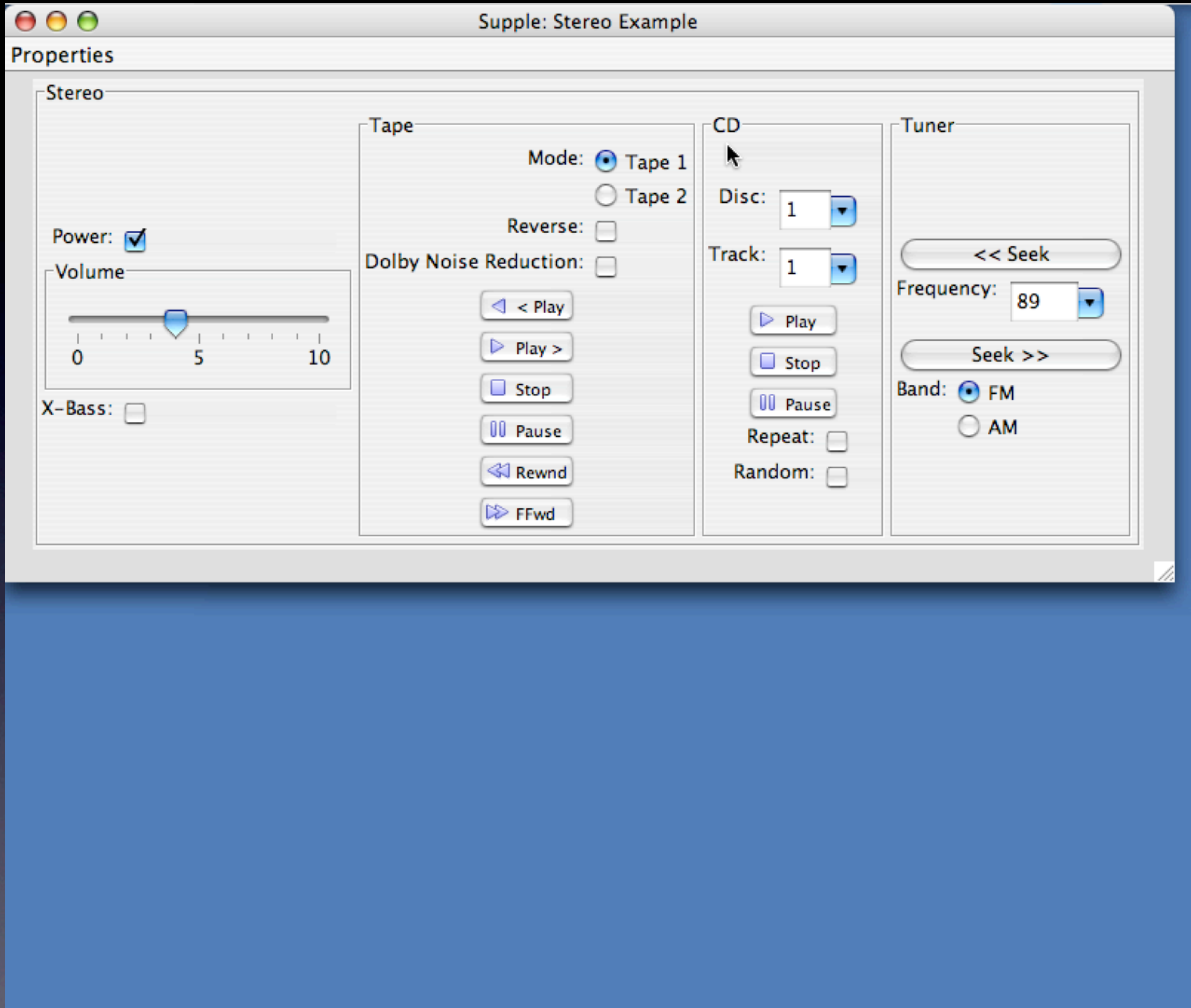
Specification



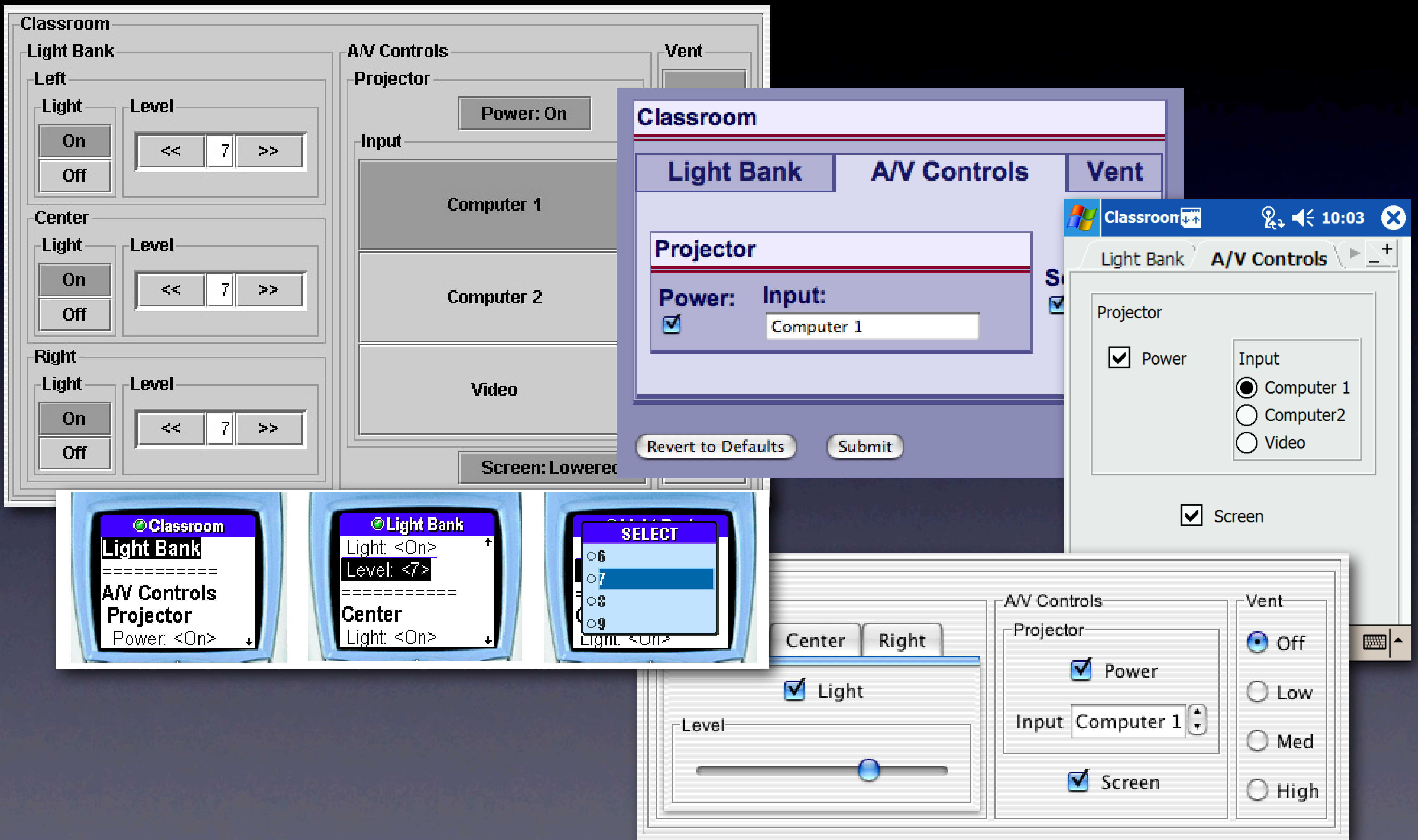
Design by Exploration

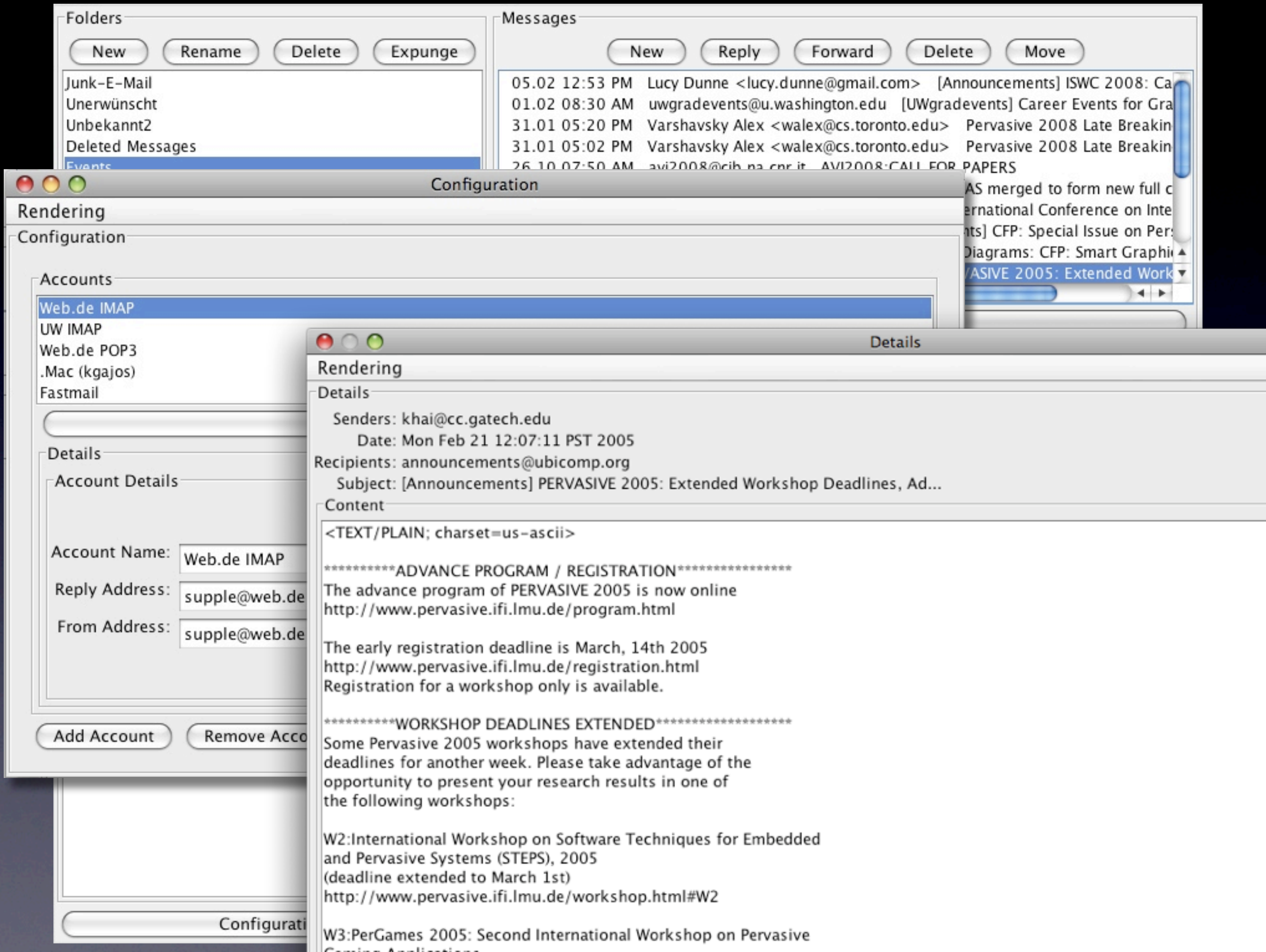


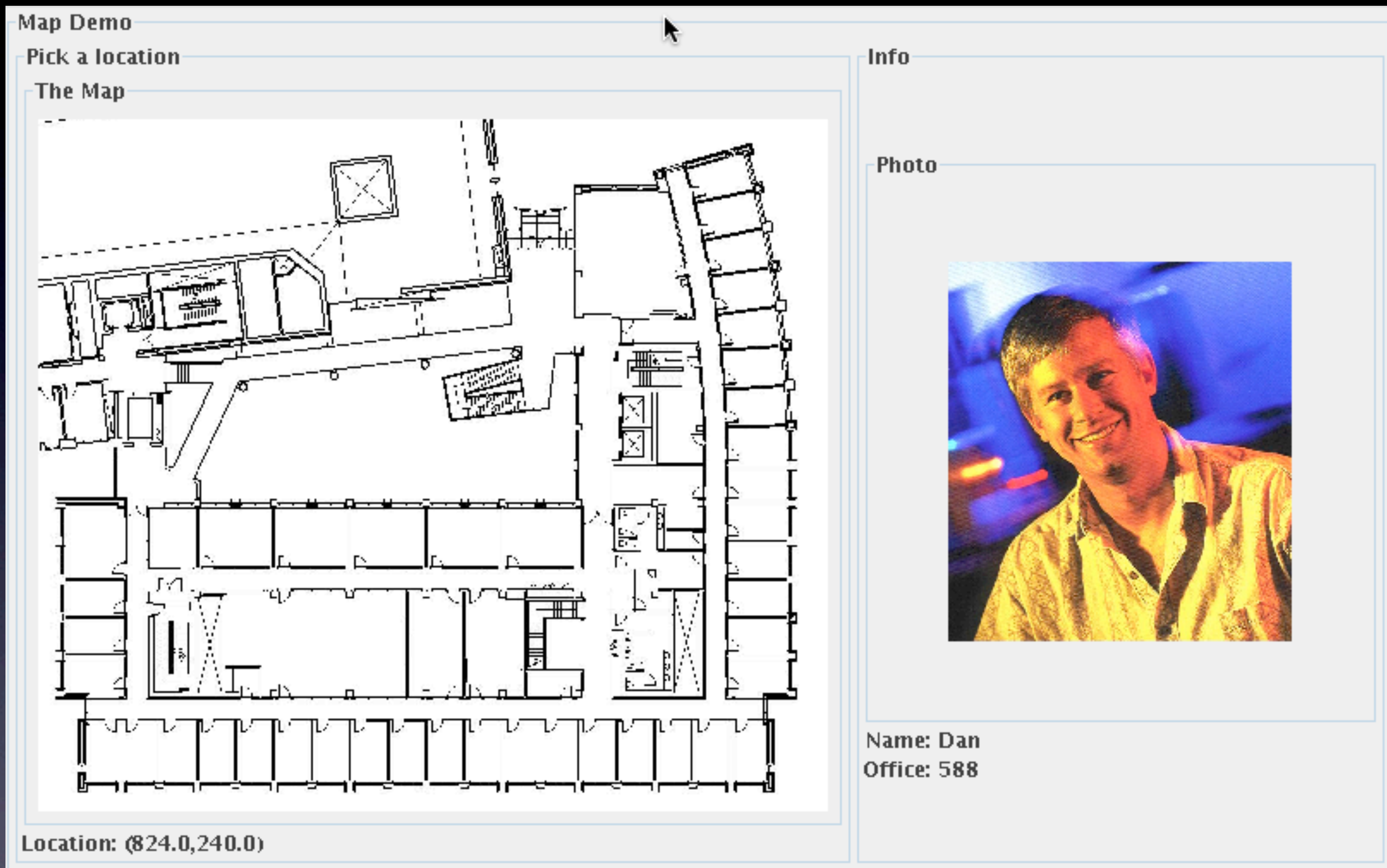
[Buxton, Sketching User Experiences]


Supple: Stereo Example
Properties
Stereo
Power:
Volume
0
5
10
X-Bass:
Tape
Mode:
Tape 1
Tape 2
Reverse:
Dolby Noise Reduction:
< Play
Play >
Stop
Pause
Rewnd
FFwd
CD
Disc:
1
Track:
1
Play
Stop
Pause
Repeat:
Random:
Tuner
<< Seek
Frequency:
89
Seek >>
Band:
FM
AM

Adaptation to Devices




Folders
New
Rename
Delete
Expunge
Junk-E-Mail
Unerwünscht
Unbekannt2
Deleted Messages
Messages
New
Reply
Forward
Delete
Move
05.02 12:53 PM Lucy Dunne <lucy.dunne@gmail.com> [Announcements] ISWC 2008: Ca
01.02 08:30 AM uwgradevents@u.washington.edu [UWgradevents] Career Events for Gra
31.01 05:20 PM Varshavsky Alex <walex@cs.toronto.edu> Pervasive 2008 Late Breakin
31.01 05:02 PM Varshavsky Alex <walex@cs.toronto.edu> Pervasive 2008 Late Breakin
Configuration
Rendering
Configuration
Accounts
Web.de IMAP
UW IMAP
Web.de POP3
.Mac (kgajos)
Fastmail
Details
Account Details
Account Name: Web.de IMAP
Reply Address: supple@web.de
From Address: supple@web.de
Add Account
Remove Acco
Configurati
Details
Rendering
Details
Senders: khai@cc.gatech.edu
Date: Mon Feb 21 12:07:11 PST 2005
Recipients: announcements@ubicomp.org
Subject: [Announcements] PERVASIVE 2005: Extended Workshop Deadlines, Ad...
Content
<TEXT/PLAIN; charset=us-ascii>
**********ADVANCE PROGRAM / REGISTRATION****************
The advance program of PERVASIVE 2005 is now online
http://www.pervasive.ifi.lmu.de/program.html
The early registration deadline is March, 14th 2005
http://www.pervasive.ifi.lmu.de/registration.html
Registration for a workshop only is available.
**********WORKSHOP DEADLINES EXTENDED*******************
Some Pervasive 2005 workshops have extended their
deadlines for another week. Please take advantage of the
opportunity to present your research results in one of
the following workshops:
W2:International Workshop on Software Techniques for Embedded
and Pervasive Systems (STEPS), 2005
(deadline extended to March 1st)
http://www.pervasive.ifi.lmu.de/workshop.html#W2
W3:PerGames 2005: Second International Workshop on Pervasive


Map Demo
Pick a location
The Map
Info
Photo
Name: Dan
Office: 588
Location: (824.0,240.0)

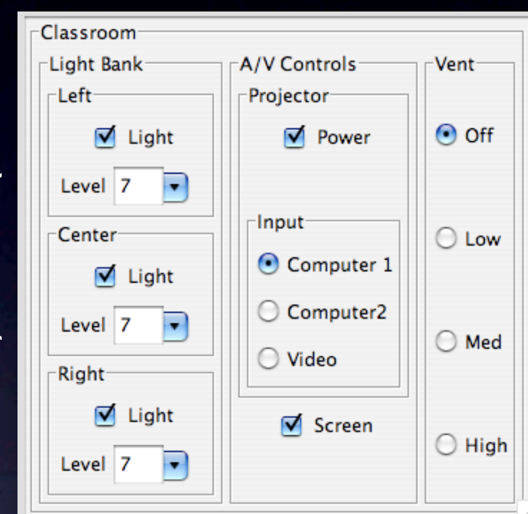
Design as Optimization

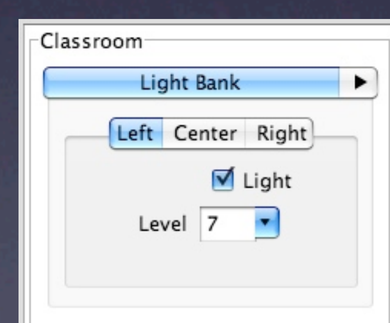
- How to enumerate the design space?
- How to evaluate solution quality?
- How to find the optimal solution efficiently?

Design as Optimization

- How to enumerate the design space?
- How to evaluate solution quality?
- How to find the optimal solution efficiently?

Evaluating Solution Quality

cost() = 4

cost() = 12

Factoring Cost Function

Cost of a particular rendering of a functional specification S

$$\boxed{\mathrm{cost}(rend(S))} = \boxed{\sum_{e \in S}} \boxed{\mathrm{cost}(rend(e))}$$

Sum over all elements of the functional specification

Cost of a rendering of an element e

Factoring Cost Function

$$\mathrm{cost}(rend(S)) = \sum_{e \in S} \boxed{\mathrm{cost}(rend(e))}$$

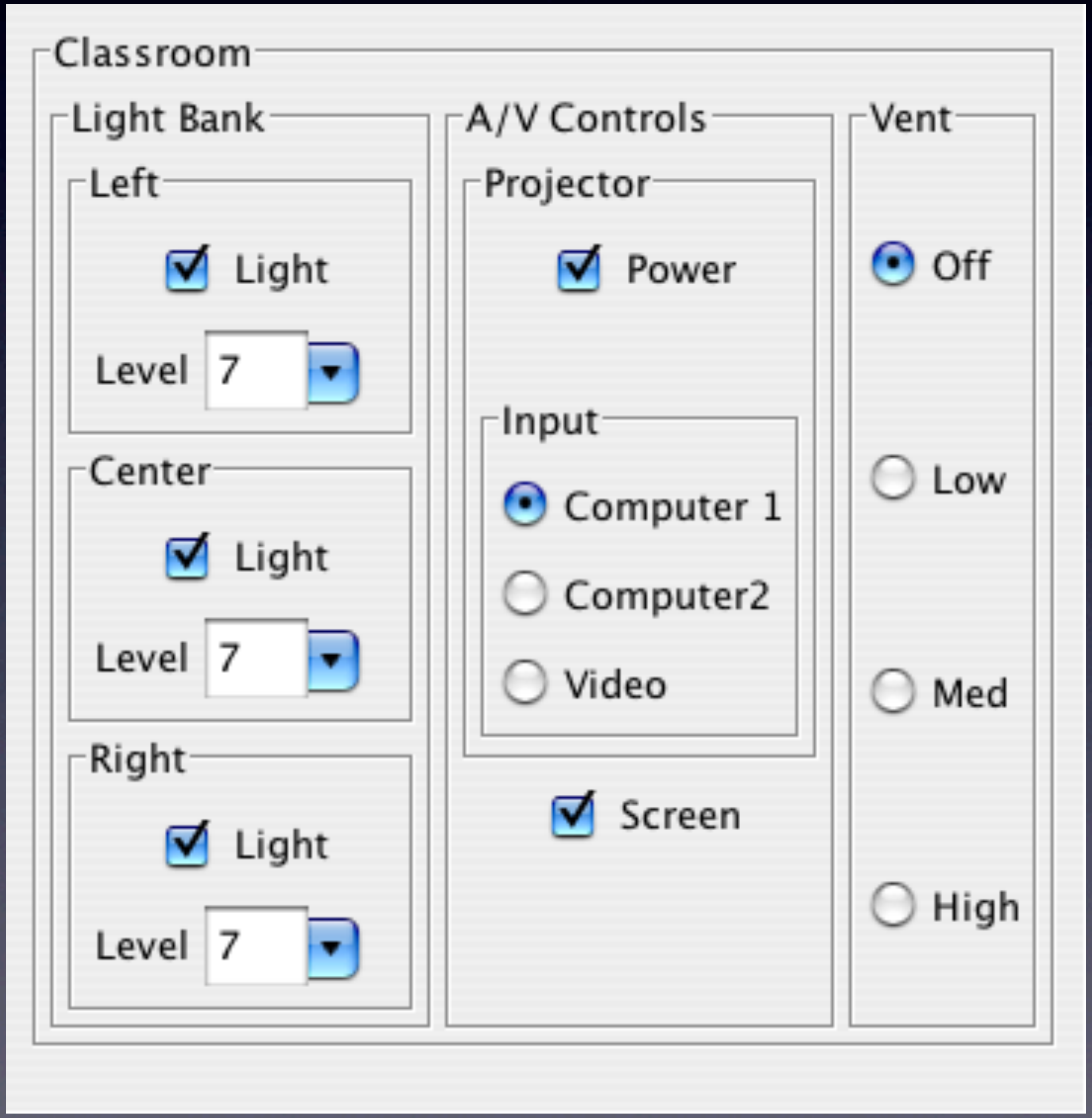
$$\mathrm{cost}(rend(e)) = \sum_{k=1}^{K} \boxed{u_k} \boxed{f_k(rend(e))}$$

A ***weight*** associated with a factor

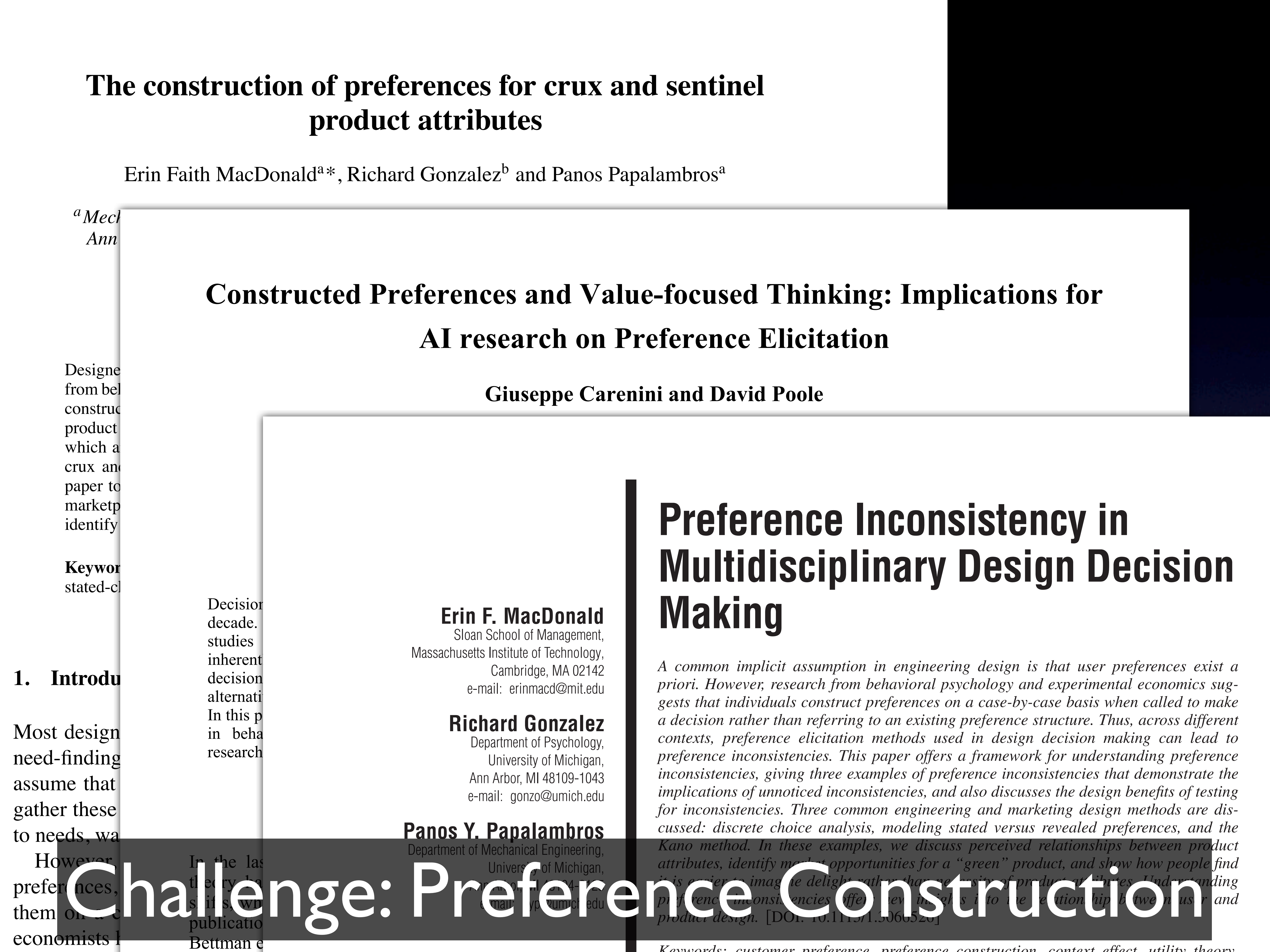
A ***factor*** reflecting presence, absence or intensity of some property

Container factor weight: 0.0
Tab Pane factor weight: 100.0
Popup factor weight: 1.0
Spinner for integers factor weight: 5.0
Spinner (domain size) factor weight: 49.5238
Spinner for non-integers factor weight: 6.0
Slider factor weight: 45.7143
Progress bar factor weight: 0.0
Checkbox factor weight: 0.0
Radio button factor weight: 0.5
Horizontal radio button factor weight: 10.0
Radio button (>=4 values) factor weight: 0.0
Radio button (>=8 values) factor weight: 74.2857

How Good Is This Interface?



How Good Is This Movie?


The construction of preferences for crux and sentinel product attributes

Erin Faith MacDonald[a]*, Richard Gonzalez[b] and Panos Papalambros[a]

Constructed Preferences and Value-focused Thinking: Implications for AI research on Preference Elicitation

Giuseppe Carenini and David Poole

Preference Inconsistency in Multidisciplinary Design Decision Making

Erin F. MacDonald
Sloan School of Management,
Massachusetts Institute of Technology,
Cambridge, MA 02142
e-mail: erinmacd@mit.edu

Richard Gonzalez
Department of Psychology,
University of Michigan,
Ann Arbor, MI 48109-1043
e-mail: gonzo@umich.edu

Panos Y. Papalambros
Department of Mechanical Engineering,
University of Michigan,

A common implicit assumption in engineering design is that user preferences exist a priori. However, research from behavioral psychology and experimental economics suggests that individuals construct preferences on a case-by-case basis when called to make a decision rather than referring to an existing preference structure. Thus, across different contexts, preference elicitation methods used in design decision making can lead to preference inconsistencies. This paper offers a framework for understanding preference inconsistencies, giving three examples of preference inconsistencies that demonstrate the implications of unnoticed inconsistencies, and also discusses the design benefits of testing for inconsistencies. Three common engineering and marketing design methods are discussed: discrete choice analysis, modeling stated versus revealed preferences, and the Kano method. In these examples, we discuss perceived relationships between product attributes, identify market opportunities for a "green" product, and show how people find

Challenge: Preference Construction

Here or There
Preference Judgments for Relevance

Ben Carterette[1], Paul N. Bennett[2], David Maxwell Chickering[3], and Susan T. Dumais[2]

[1] University of Massachusetts Amherst
[2] Microsoft Research
[3] Microsoft Live Labs

Abstract. Information retrieval systems have traditionally been evaluated over absolute judgments of relevance: each document is judged for relevance on its own, independent of other documents that may be on topic. We hypothesize that preference judgments of the form "document A is more relevant than document B" are easier for assessors to make than absolute judgments, and provide evidence for our hypothesis through a study with assessors. We then investigate methods to evaluate search engines using preference judgments. Furthermore, we show that by using inferences and clever selection of pairs to judge, we need not compare all pairs of documents in order to apply evaluation methods.

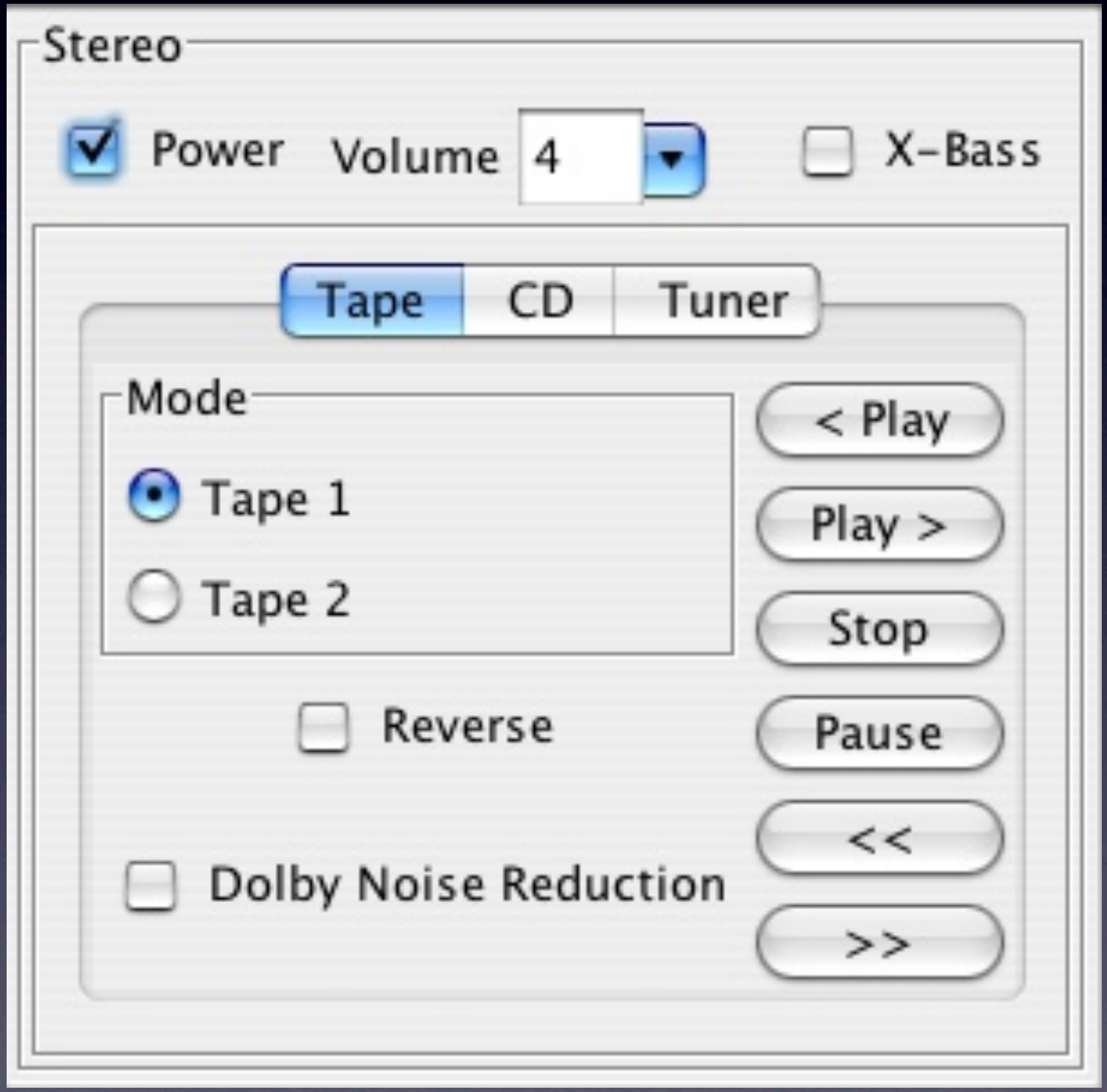
1 Introduction

Relevance judgments for information retrieval evaluation have traditionally been made on a binary scale: a document is either relevant to a query or it is not. This definition of relevance is largely motivated by the importance of *topicality* in tasks studied in IR research [1].

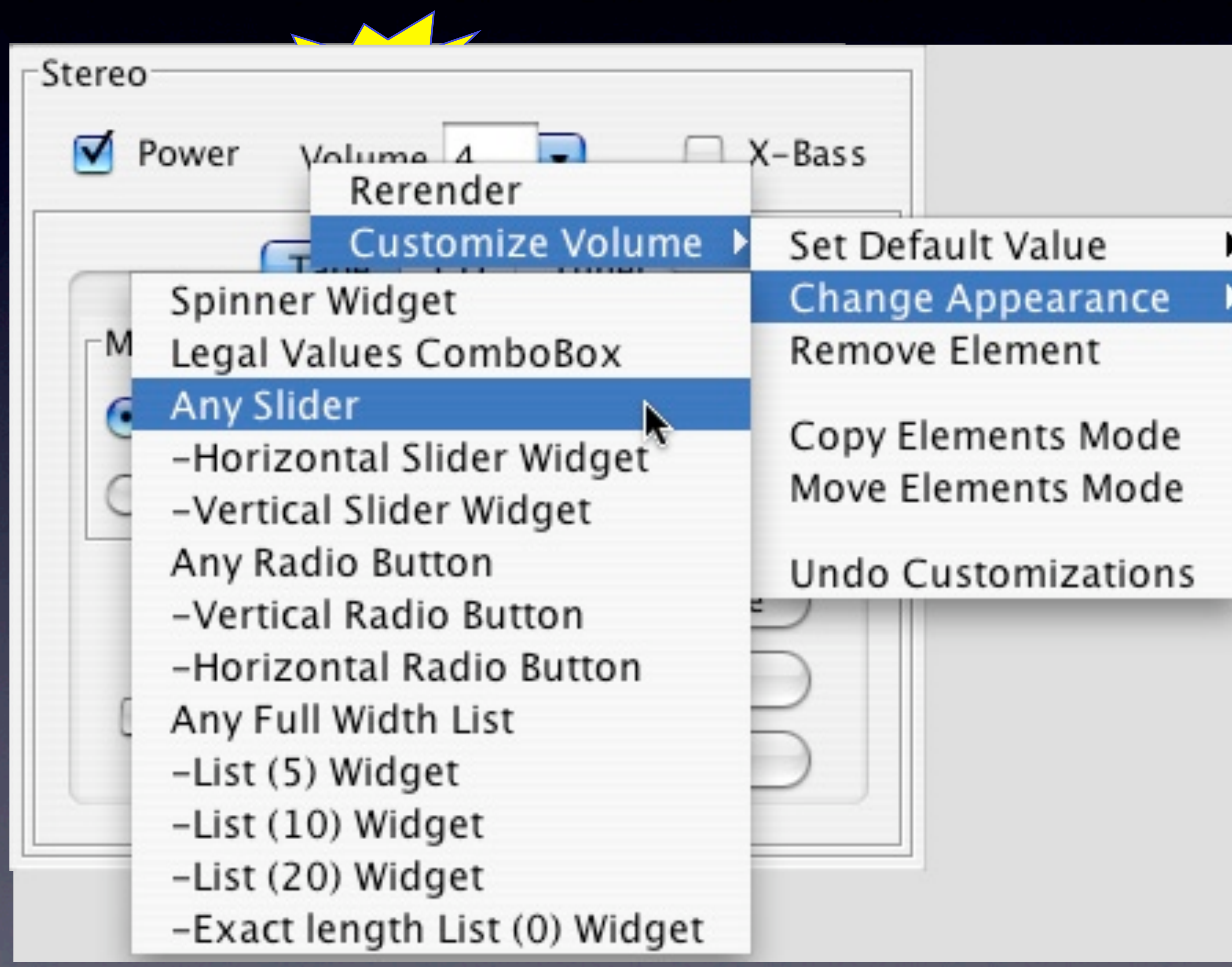
The notion of relevance can be generalized to a graded scale of absolute

Challenge: Ratings Not Reliable

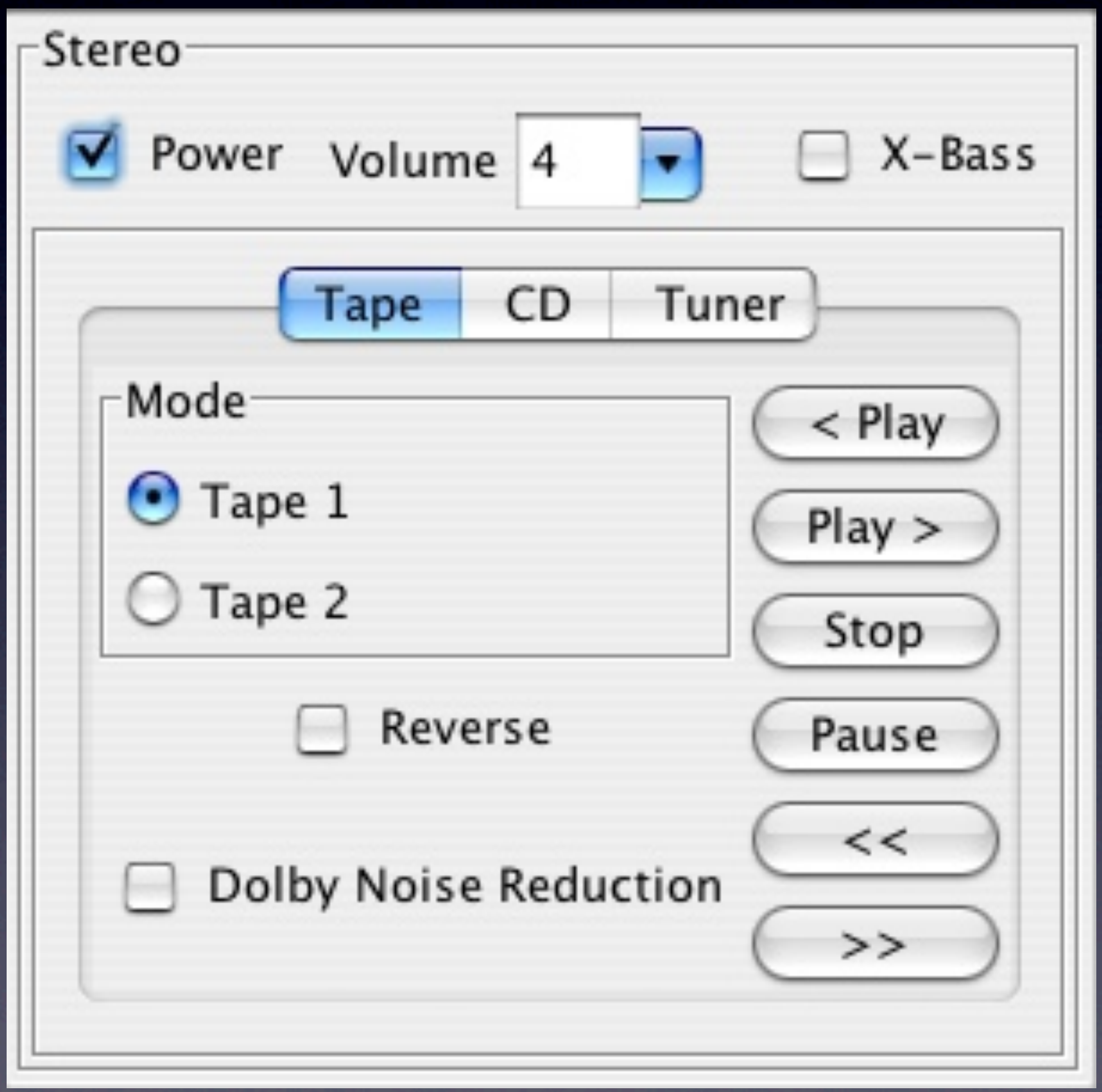
Preference Statements Through Example Critiquing



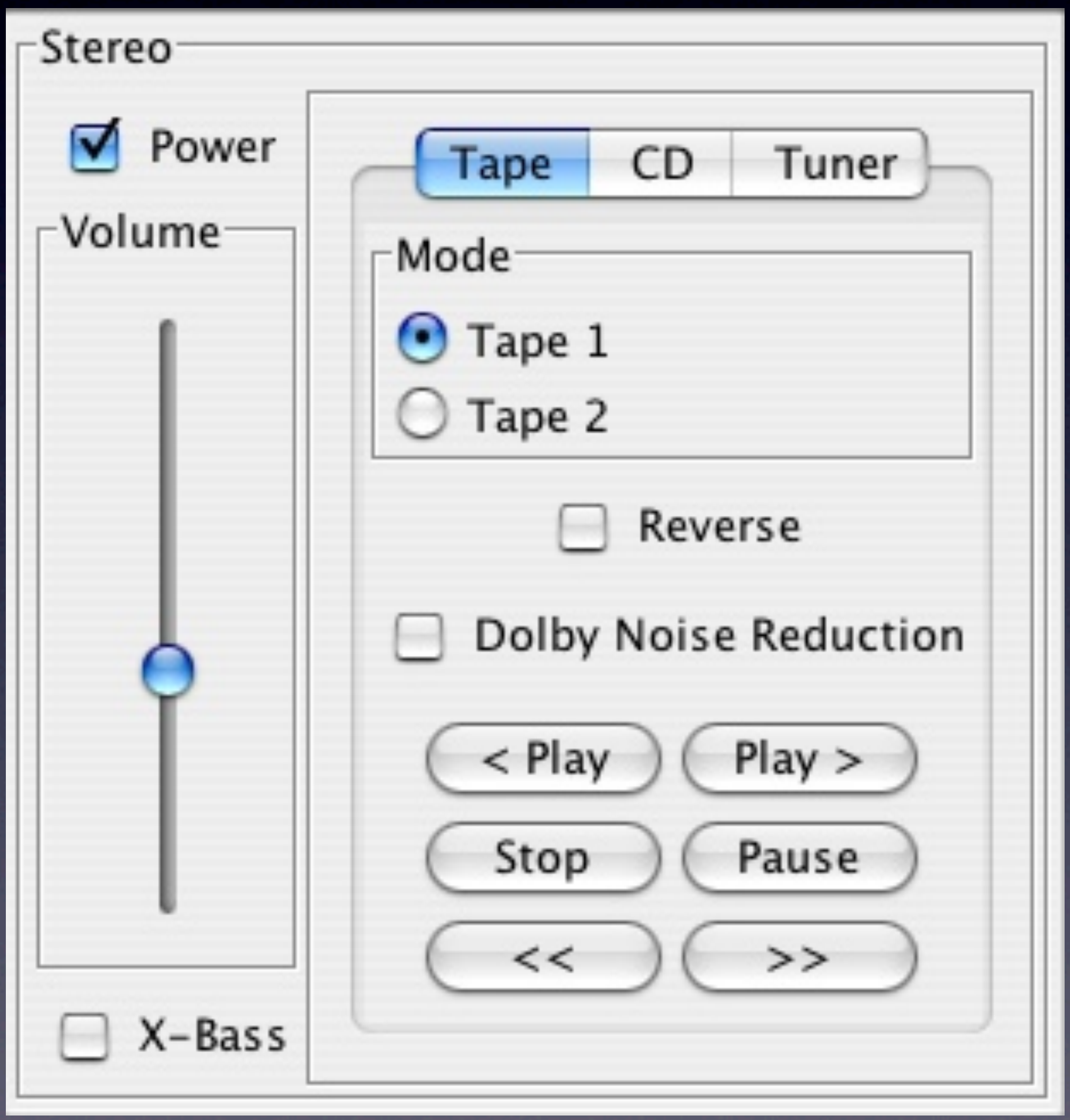
Preference Statements Through Example Critiquing



Result of a Critique

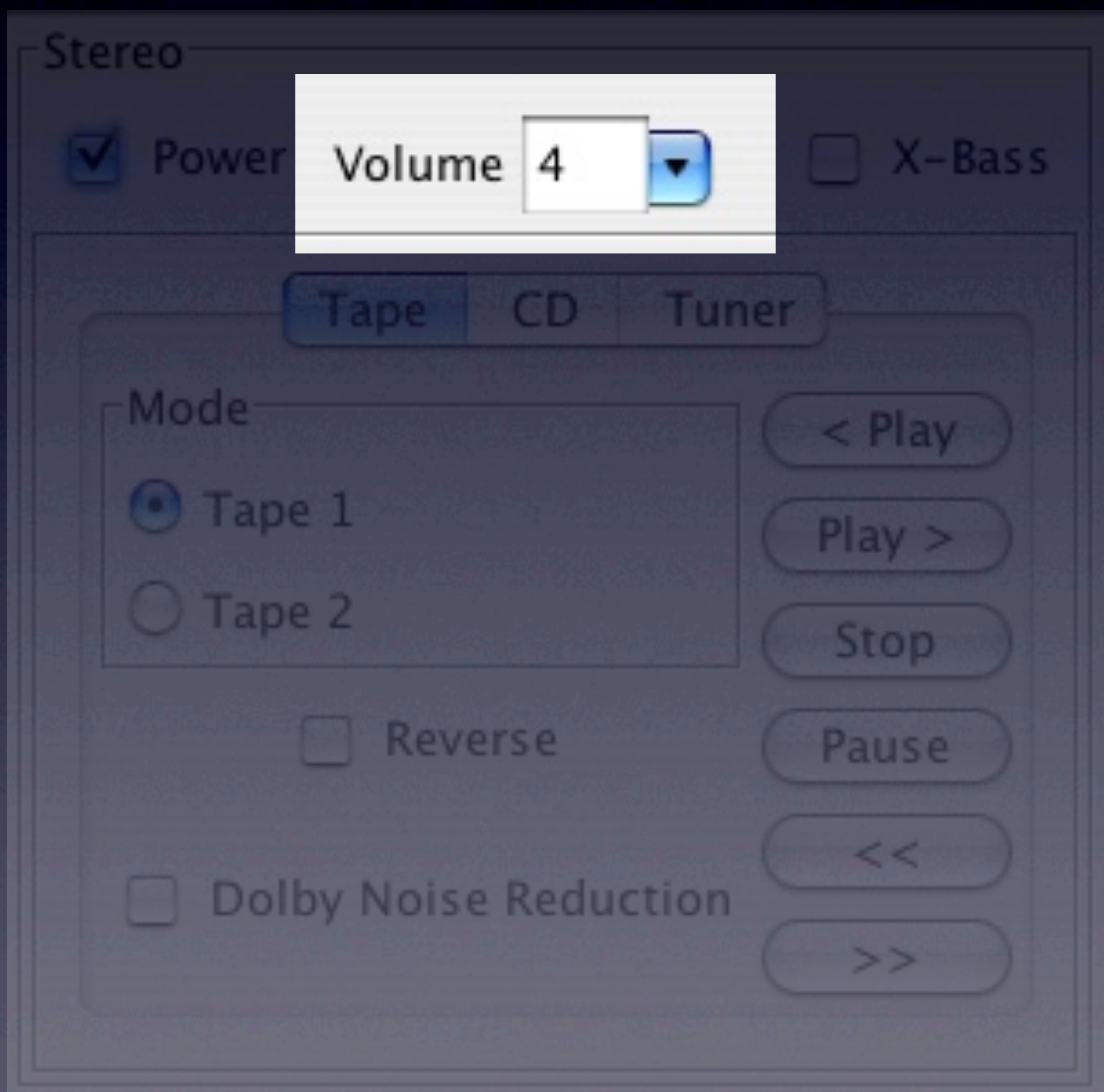


before



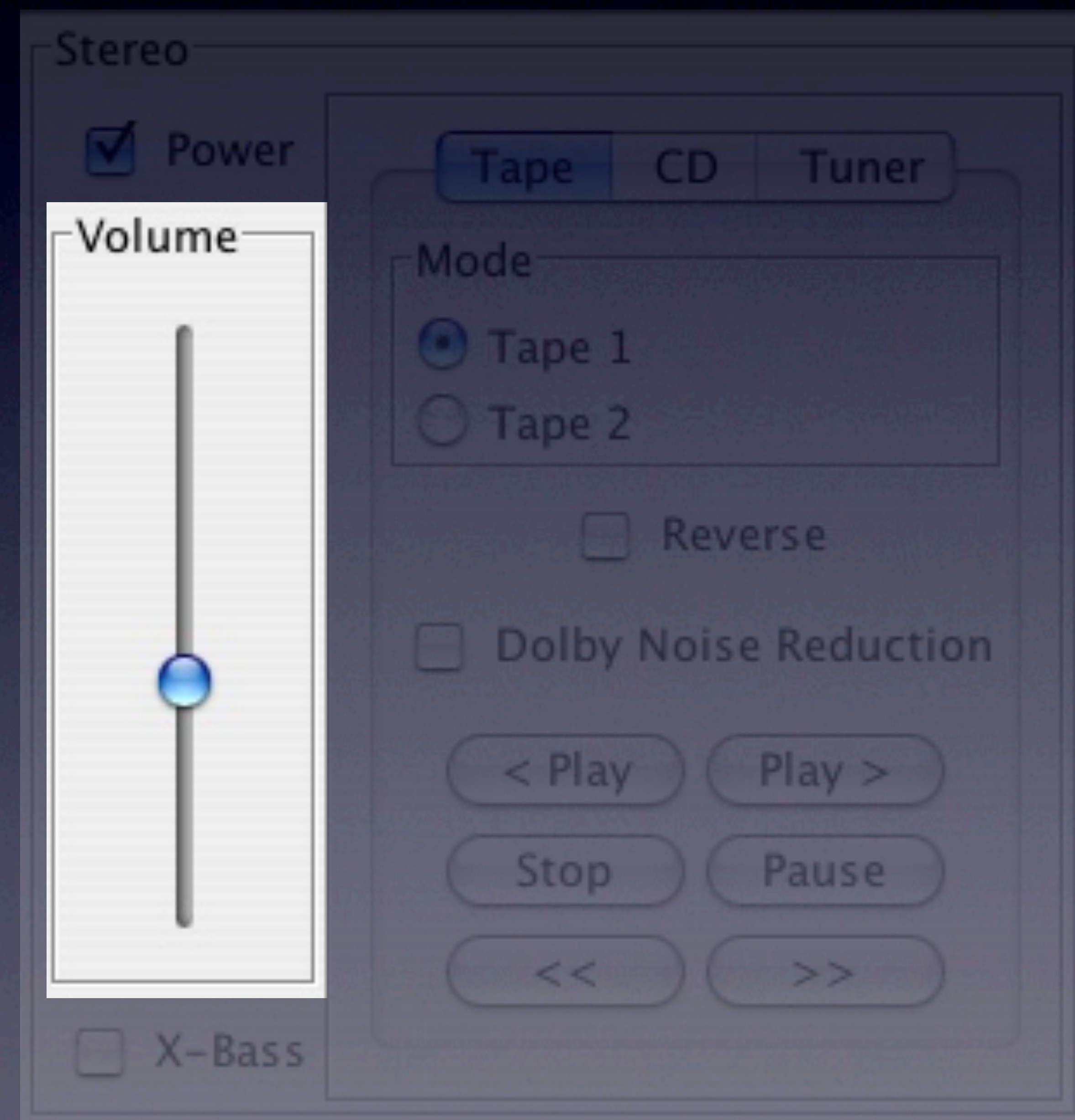
after

Result of a Critique
Provides feedback to the system



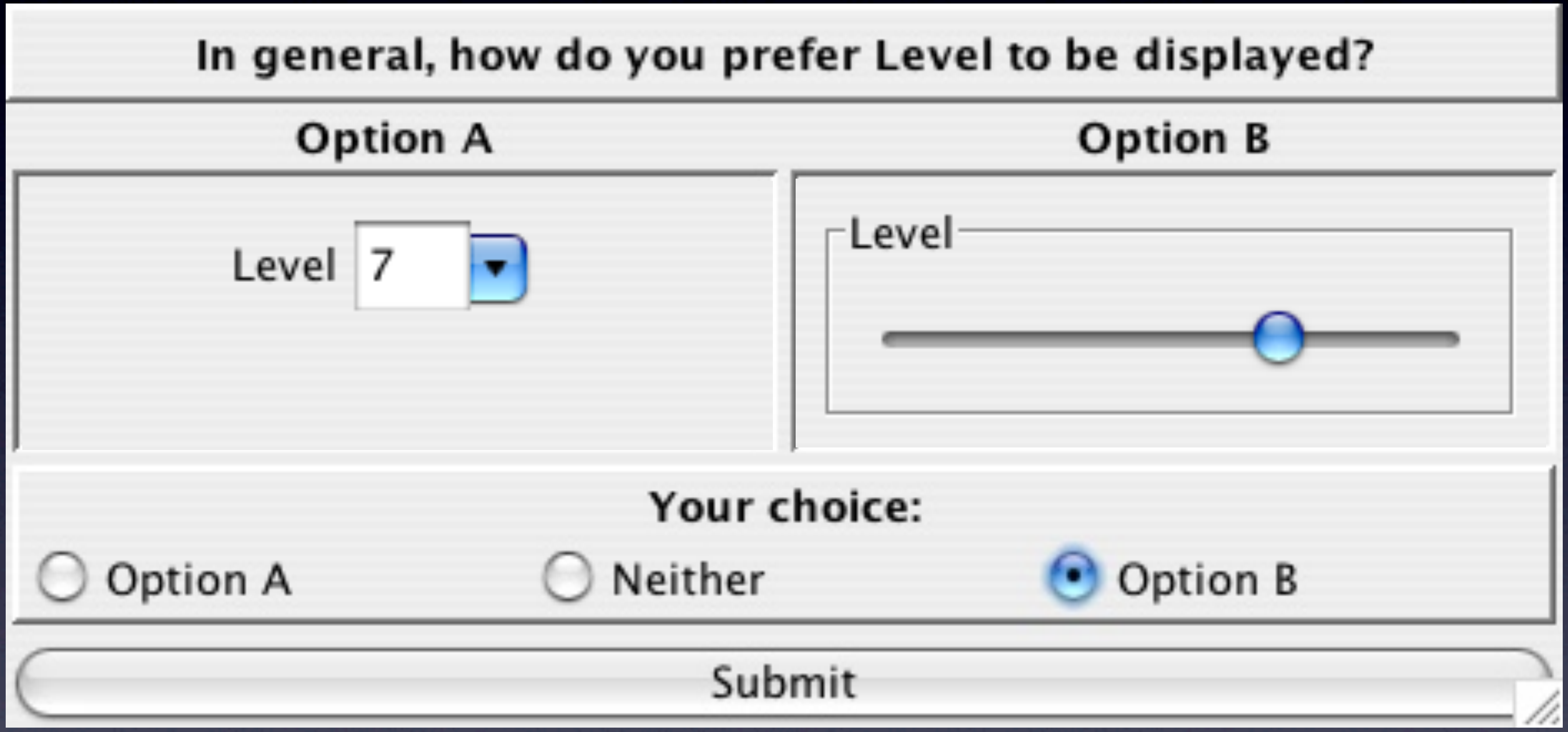
before

<

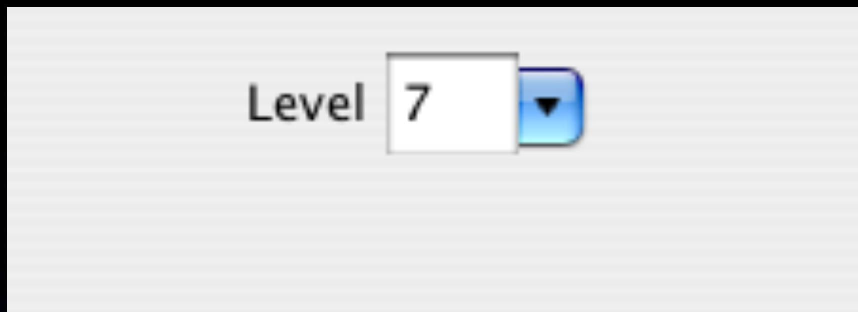


after

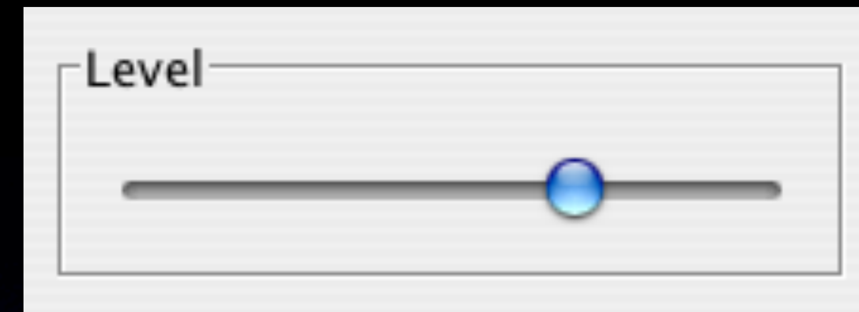
Active Elicitation via Pairwise Comparisons



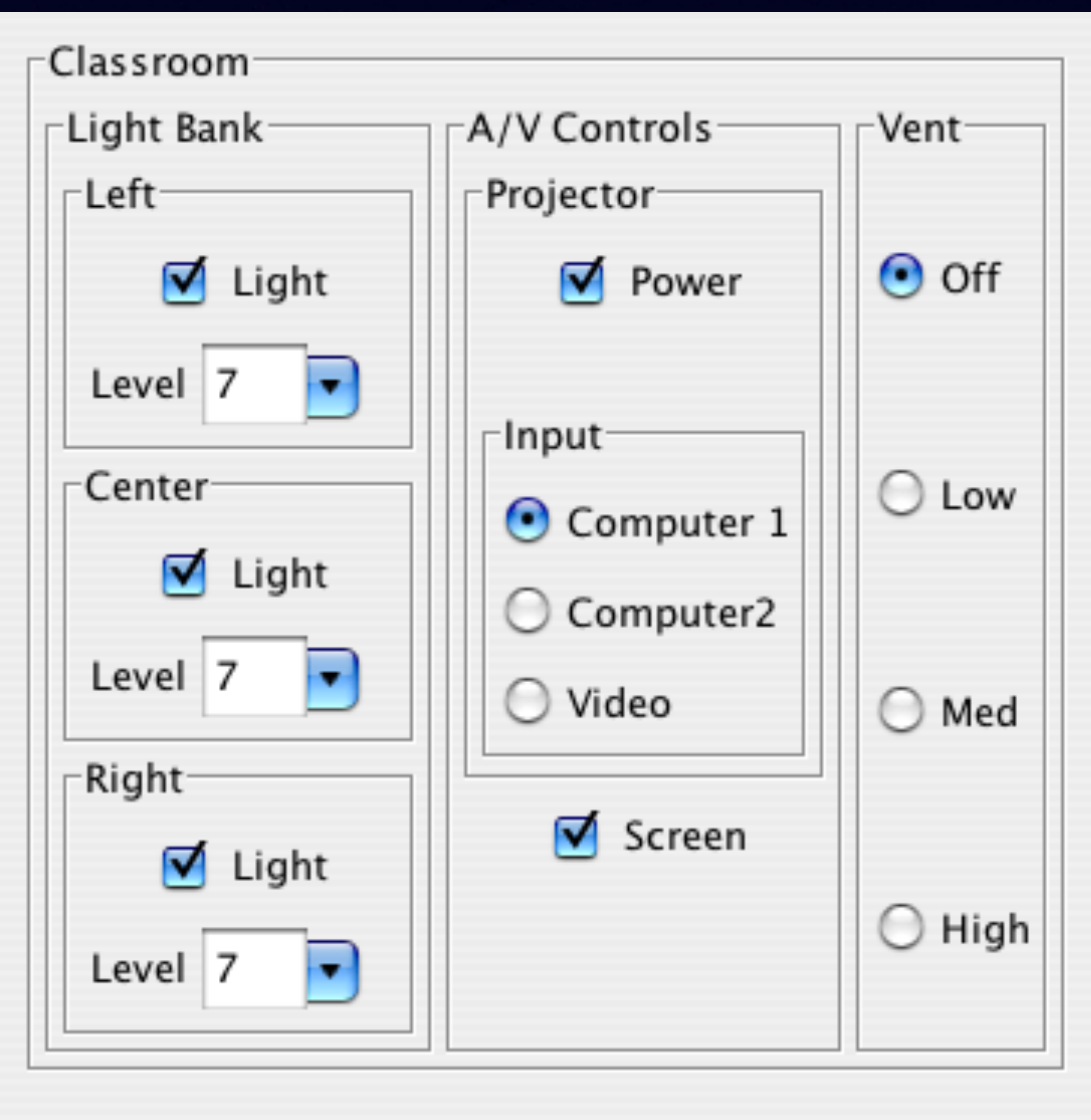
In isolation, sliders are preferred



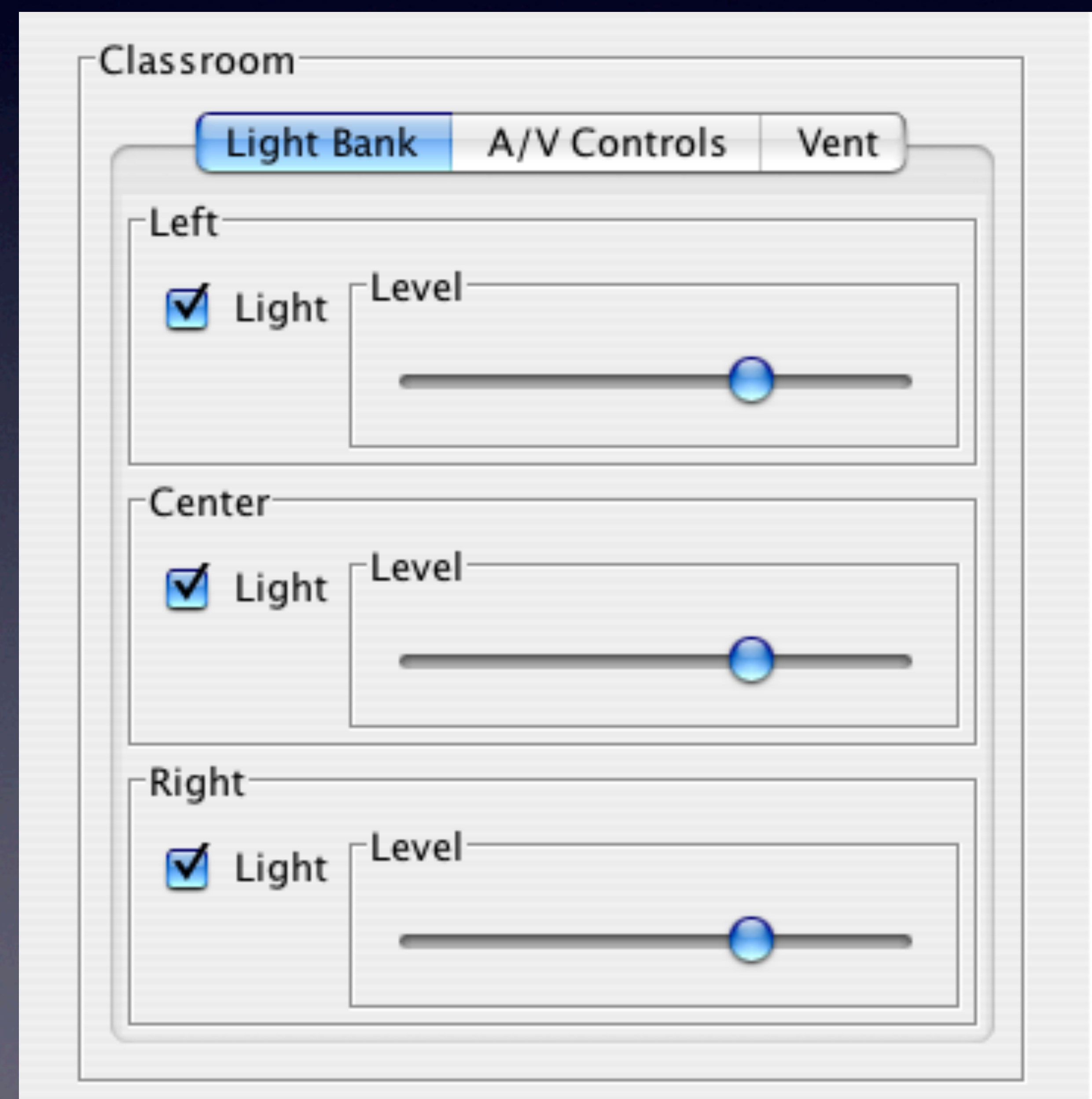
<



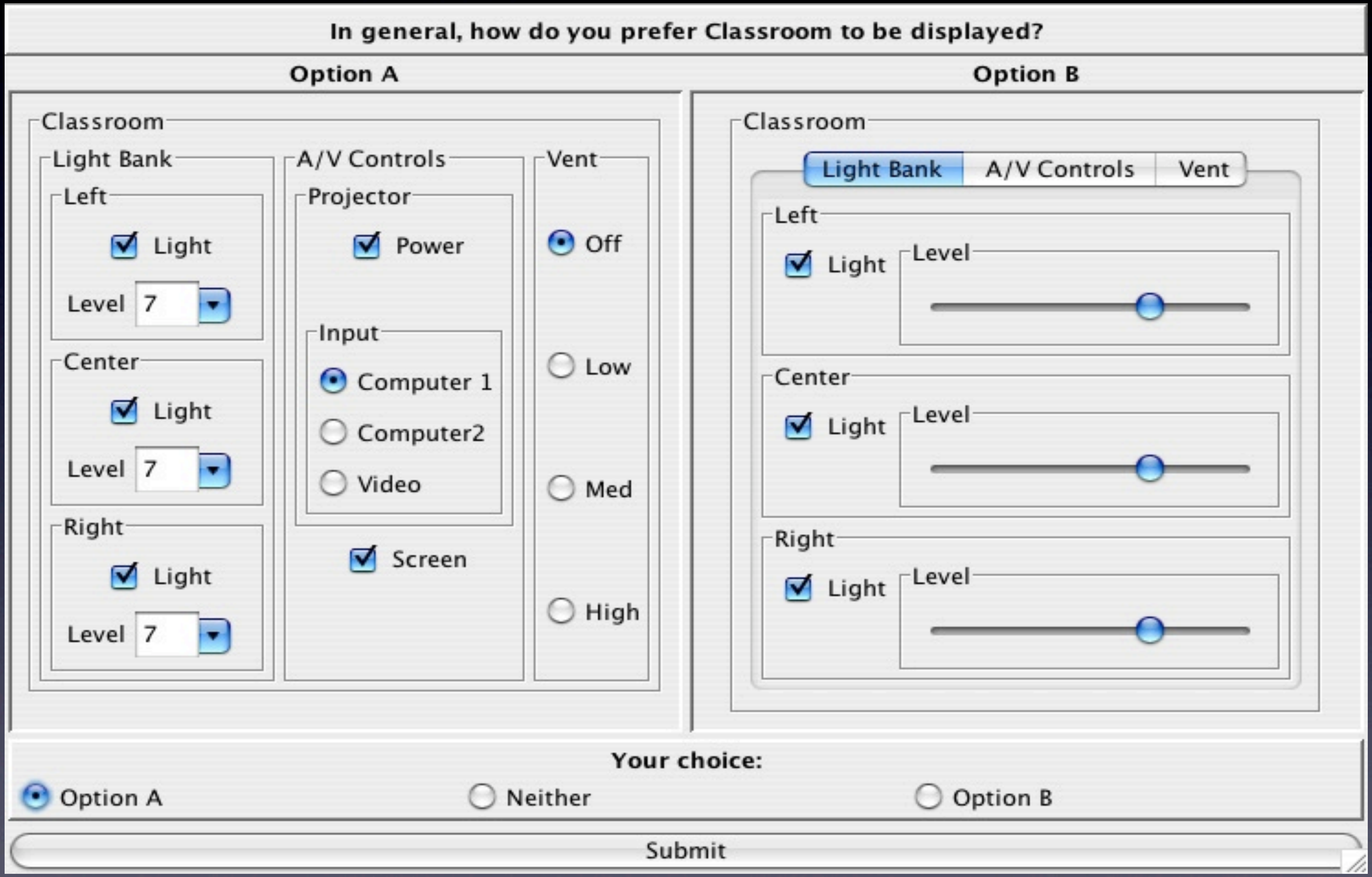
But in some contexts combo boxes may be better



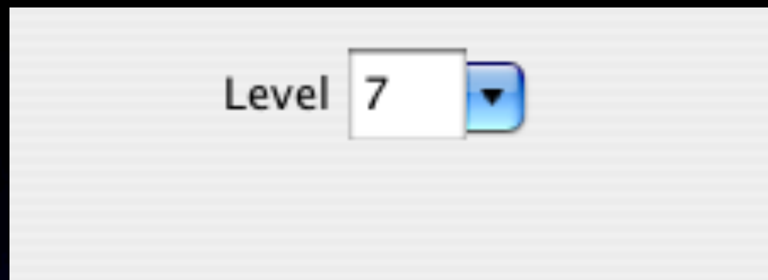
>



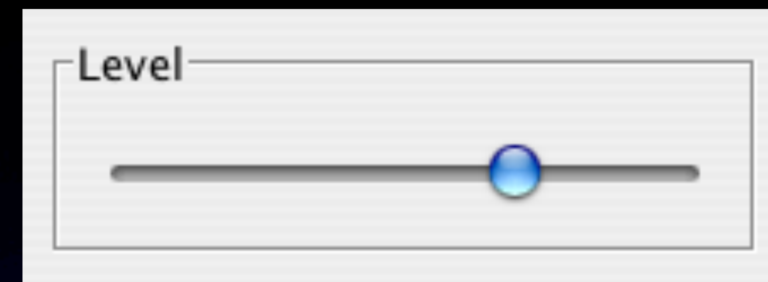
Situated Feedback
with Active Elicitation



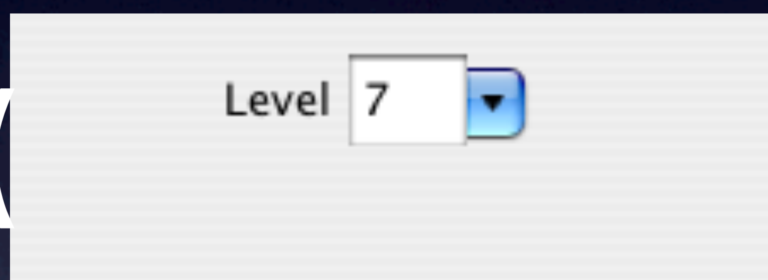
The Semantics of the Interactions



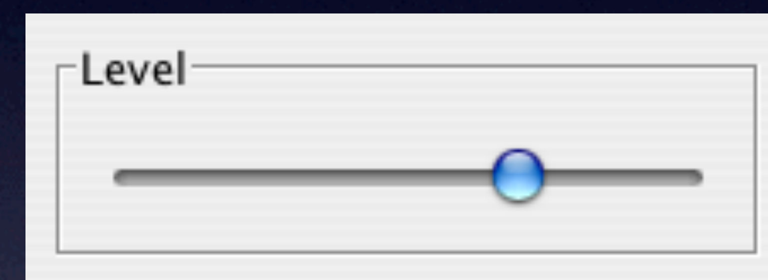
<



cost(



) ≥ cost(



)

$$\sum_{e \in S} \sum_{k=1}^{K} u_k f_k(rend_1(e)) \geq \sum_{e \in S} \sum_{k=1}^{K} u_k f_k(rend_2(e))$$

Formalizing the Learning Problem

Set up an optimization problem that maximizes:

$$margin - \sum_i slack_i$$

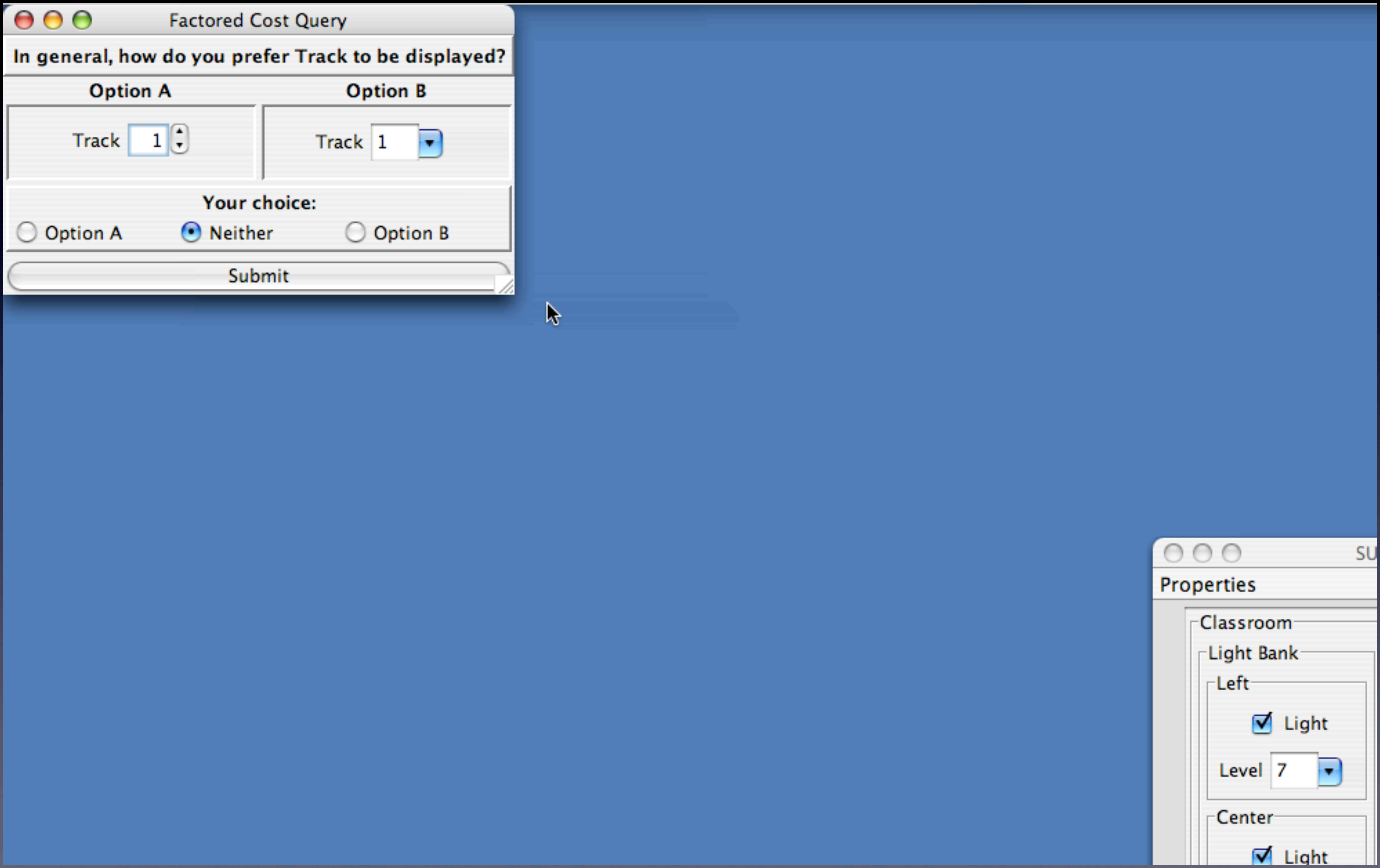
Subject to the constraints:

$$\sum_{e \in S} \sum_{k=1}^{K} u_k f_k(\mathrm{rend}_1(e)) - \sum_{e \in S} \sum_{k=1}^{K} u_k f_k(\mathrm{rend}_2(e)) \geq margin - slack_i$$

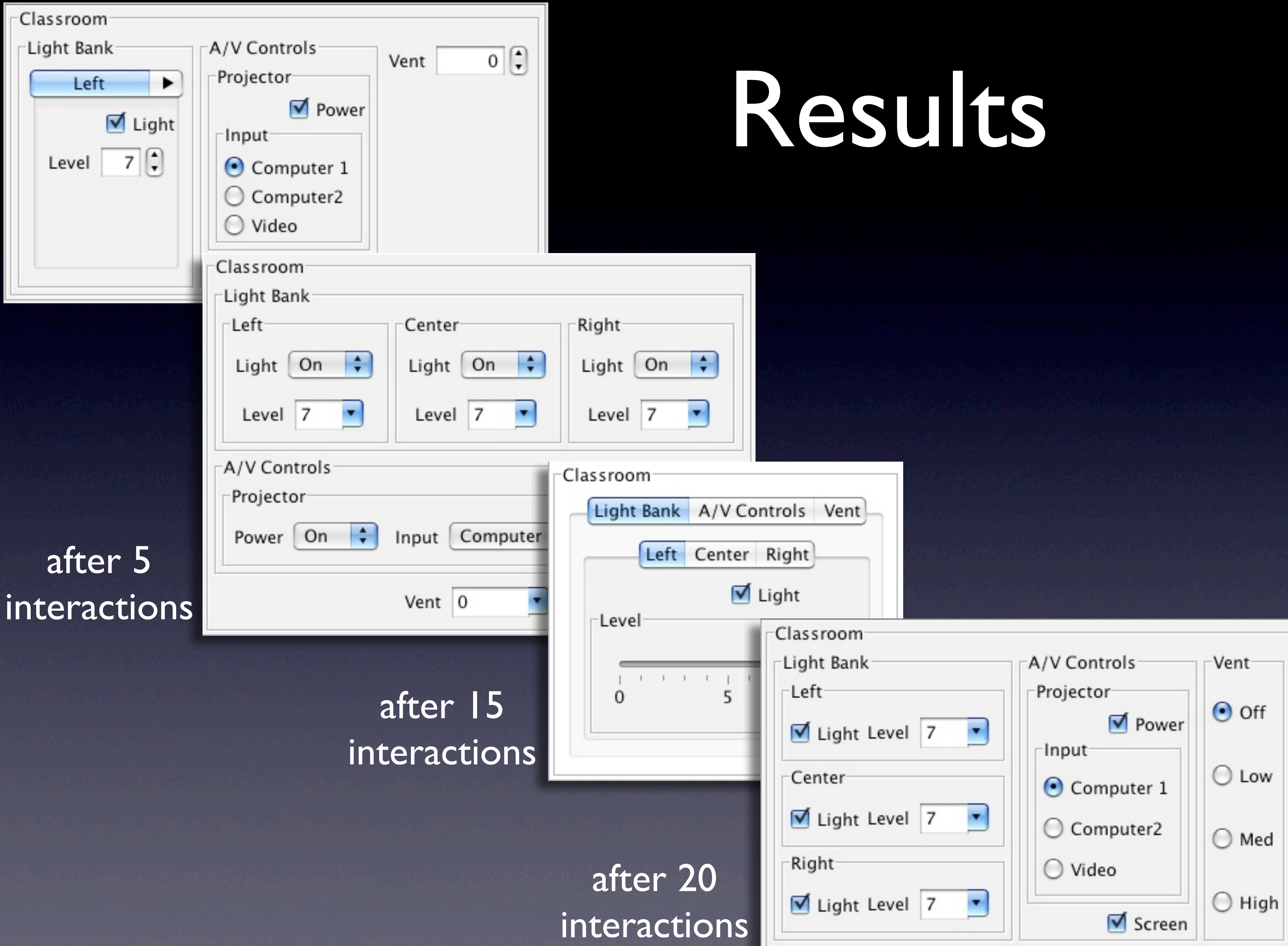
$$u_k \geq 0$$

$$u_k \leq C$$

Results


Factored Cost Query
In general, how do you prefer Track to be displayed?
Option A
Option B
Track 1
Track 1
Your choice:
Option A
Neither
Option B
Submit
Properties
Classroom
Light Bank
Left
Light
Level 7
Center
Light

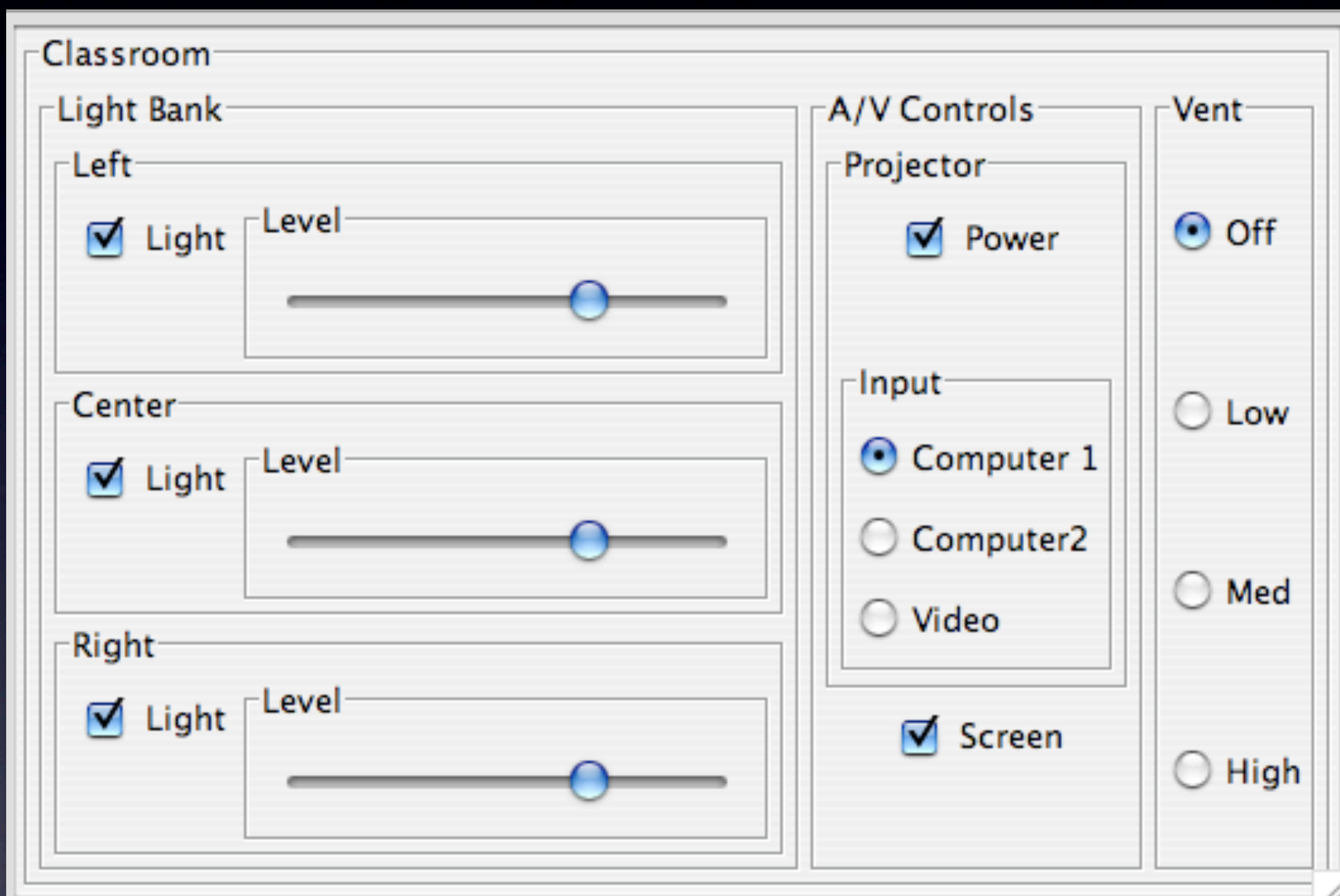
Results

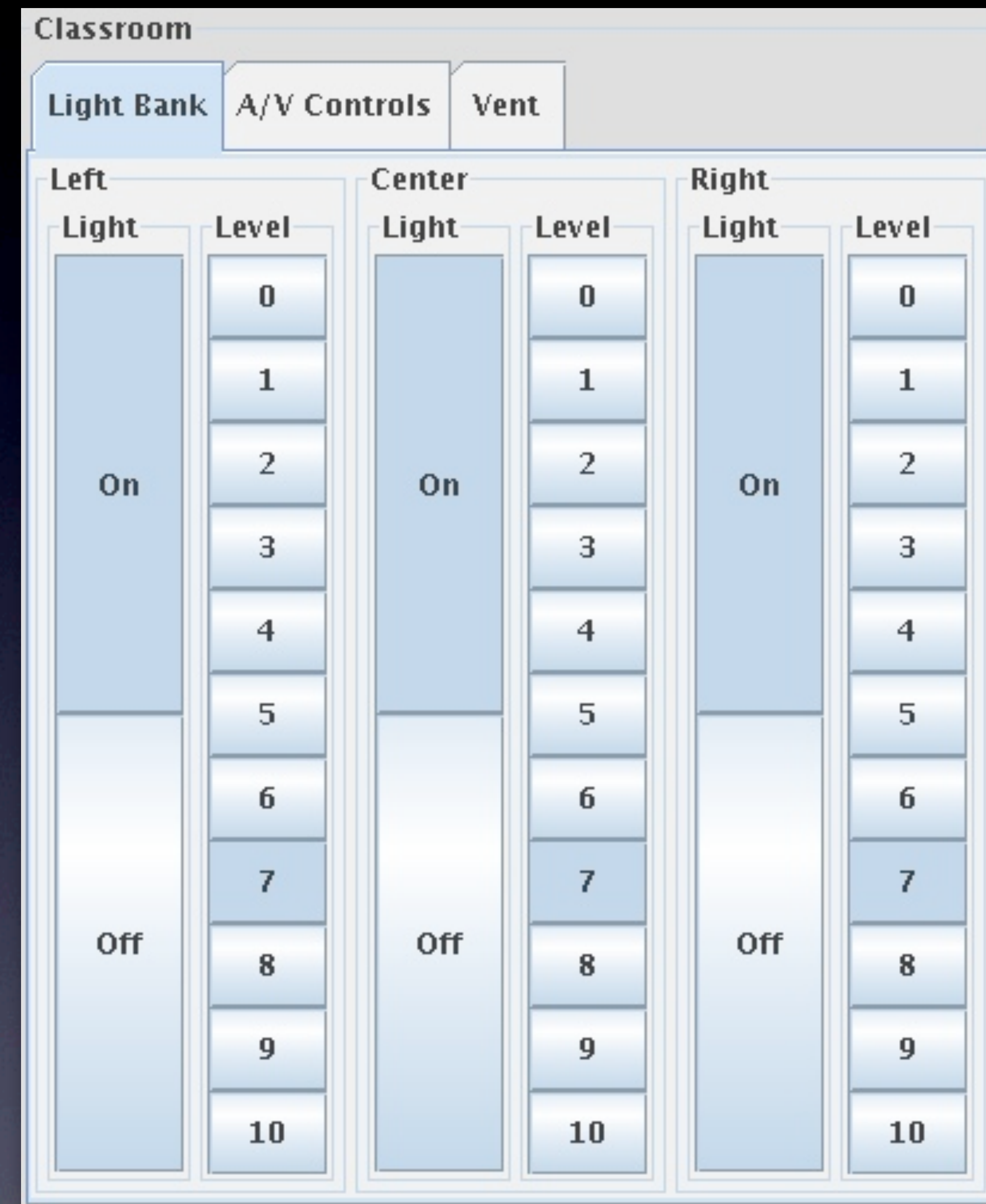


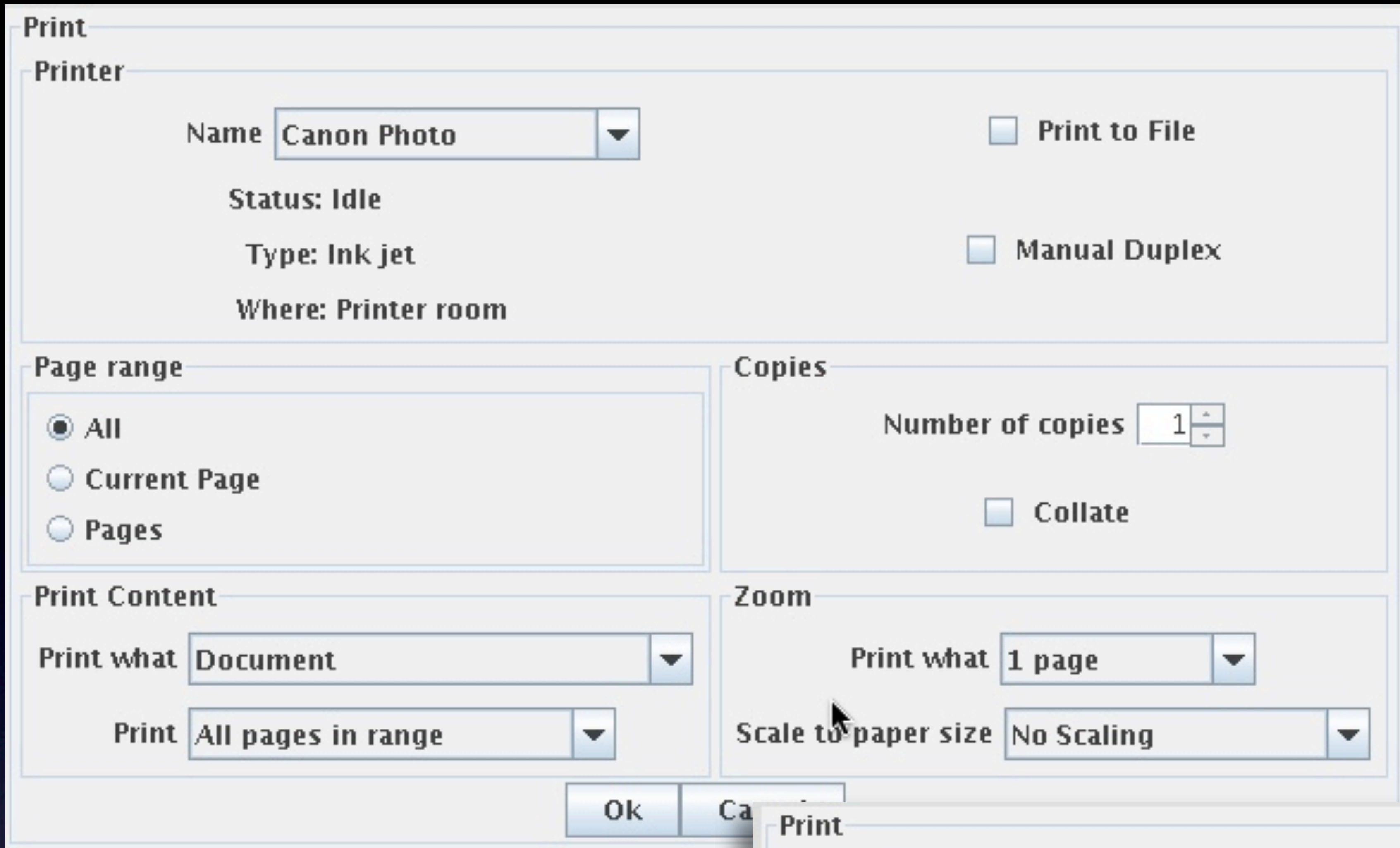
after 5 interactions

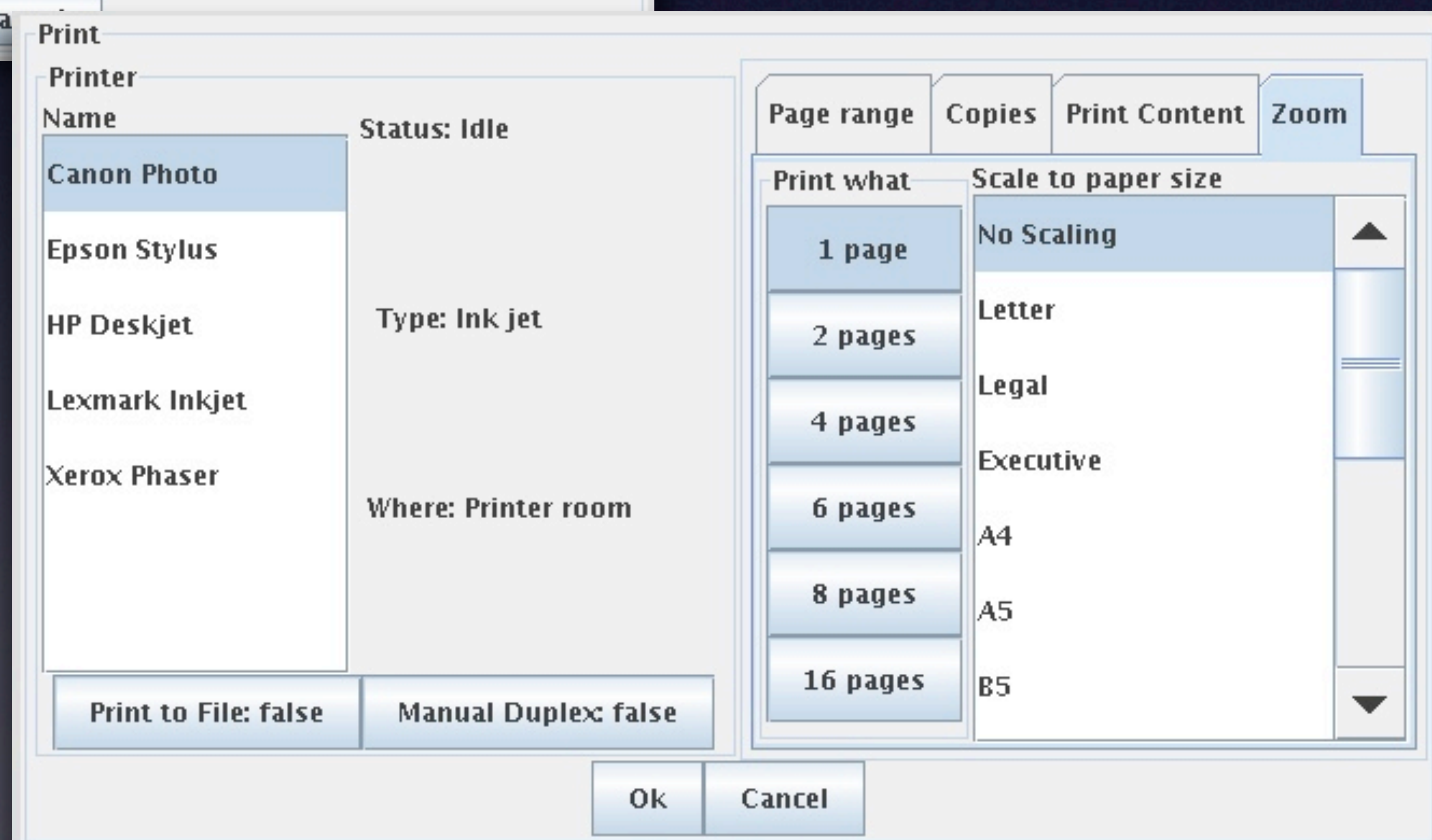
after 15 interactions

after 20 interactions


Classroom
Light Bank
Left
Light
Level
Center
Light
Level
Right
Light
Level
A/V Controls
Projector
Power
Input
Computer 1
Computer2
Video
Screen
Vent
Off
Low
Med
High


Classroom
Light Bank
A/V Controls
Vent
Left
Light
Level
On
Off
0
1
2
3
4
5
6
7
8
9
10
Center
Light
Level
On
Off
0
1
2
3
4
5
6
7
8
9
10
Right
Light
Level
On
Off
0
1
2
3
4
5
6
7
8
9
10


Print
Printer
Name
Canon Photo
Print to File
Status: Idle
Type: Ink jet
Manual Duplex
Where: Printer room
Page range
All
Current Page
Pages
Copies
Number of copies
1
Collate
Print Content
Print what
Document
Print
All pages in range
Zoom
Print what
1 page
Scale to paper size
No Scaling
Ok


Print
Printer
Name
Canon Photo
Epson Stylus
HP Deskjet
Lexmark Inkjet
Xerox Phaser
Status: Idle
Type: Ink jet
Where: Printer room
Page range
Copies
Print Content
Zoom
Print what
1 page
2 pages
4 pages
6 pages
8 pages
16 pages
Scale to paper size
No Scaling
Letter
Legal
Executive
A4
A5
B5
Print to File: false
Manual Duplex: false
Ok
Cancel

Learning a Distance Metric from Relative Comparisons

Matthew Schultz and Thorsten Joachims
Department of Computer Science
Cornell University
Ithaca, NY 14853
{schultz,tj}@cs.cornell.edu

Abstract

This paper presents a method for learning a distance metric from relative comparison such as "A is closer to B than A is to C". Taking a Support Vector Machine (SVM) approach, we develop an algorithm that provides a flexible way of describing qualitative training data as a set of constraints. We show that such constraints lead to a convex quadratic programming problem that can be solved by adapting standard methods for SVM training. We empirically evaluate the performance and the modelling flexibility of the algorithm on a collection of text documents.

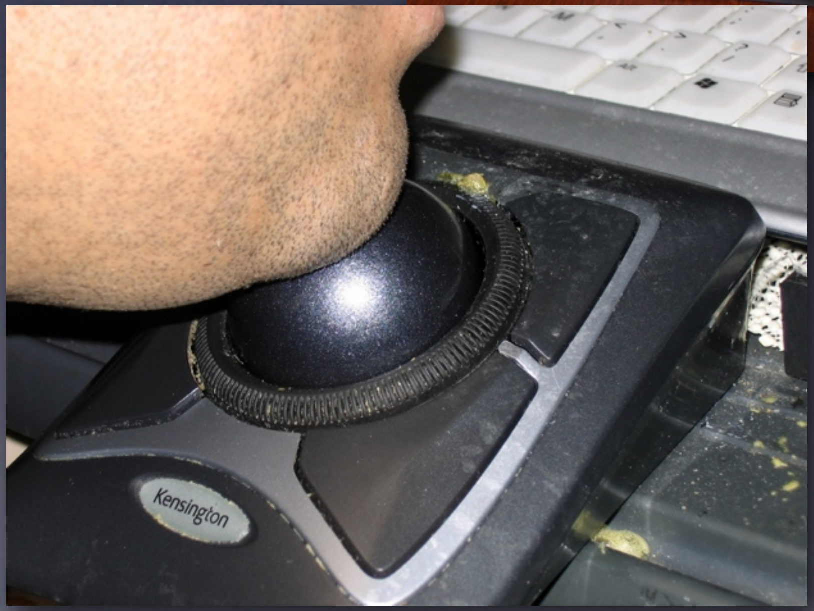
1 Introduction

Distance metrics are an essential component in many applications ranging from supervised

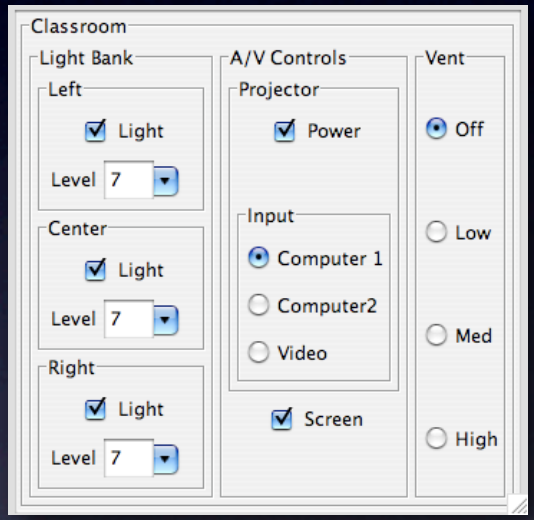
Semantics of bookmark / like / plus / star

Concerns in UI Design

- Perceptual effort
- Cognitive effort
- Motor effort
- Aesthetics

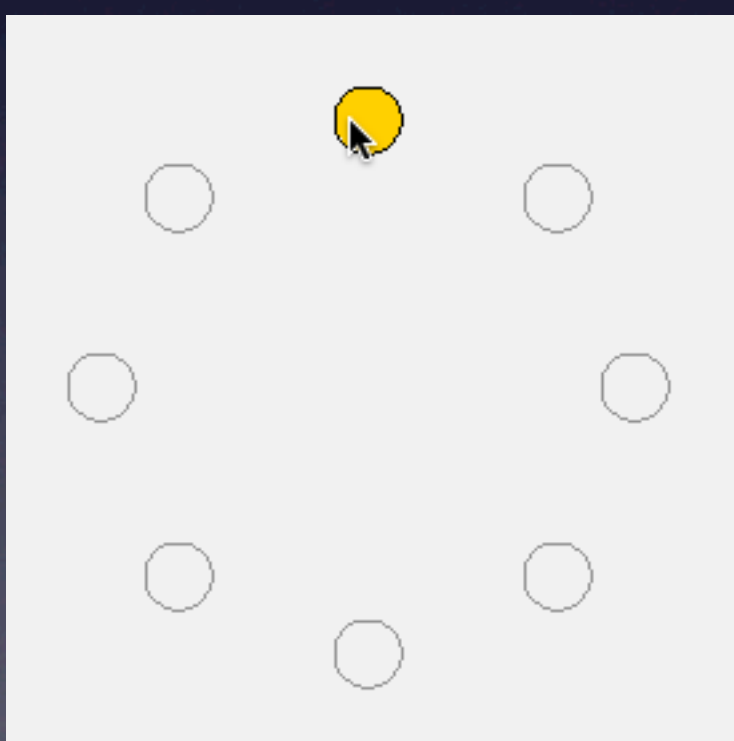


Adapting to Motor Abilities

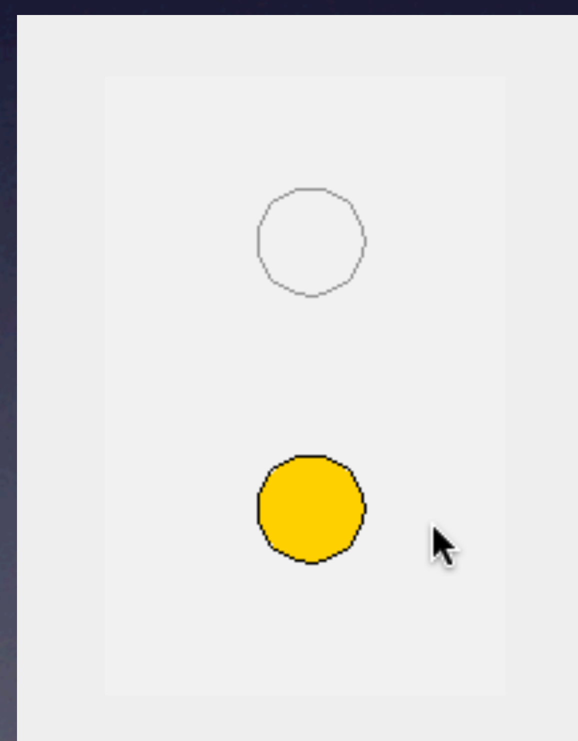
cost() = time

Collect Motor Performance Data

Pointing

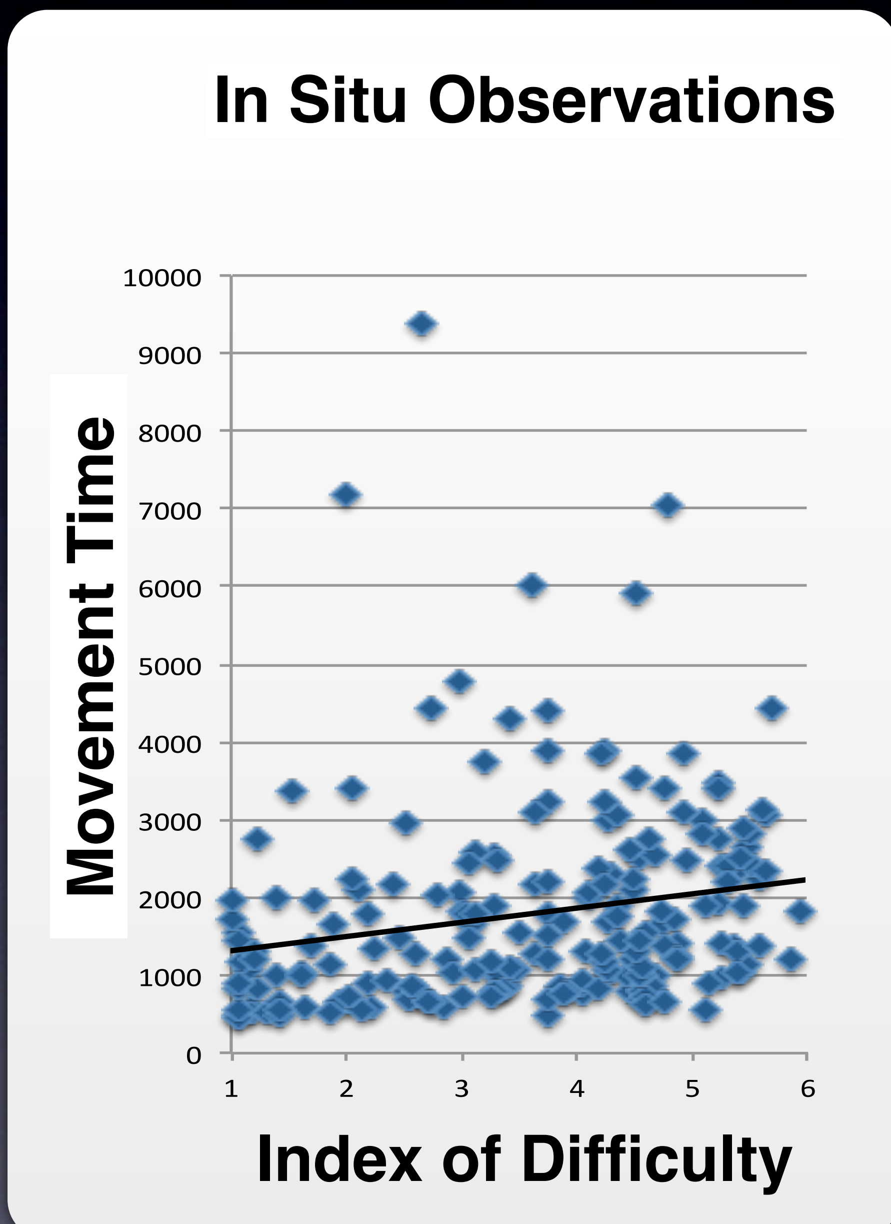


Dragging



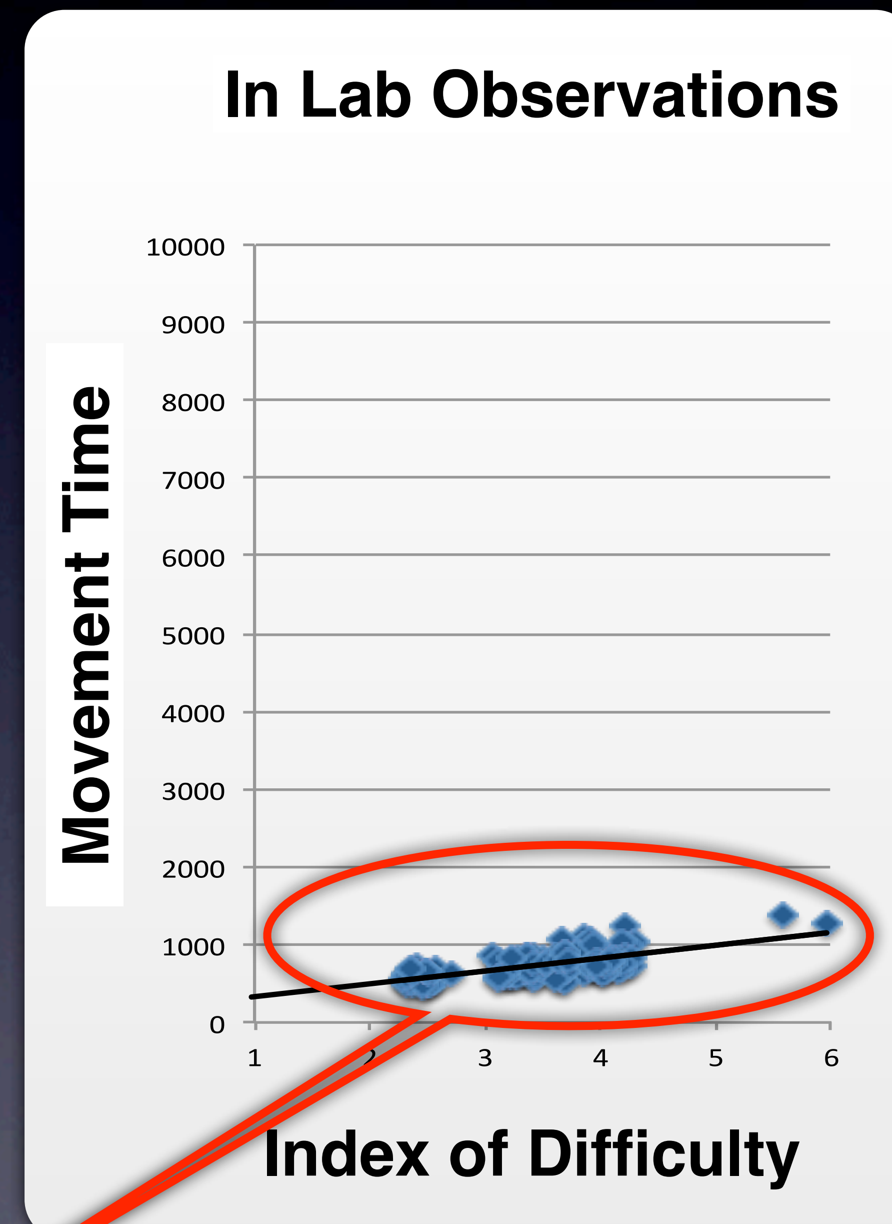
List Selection

What we get in the wild



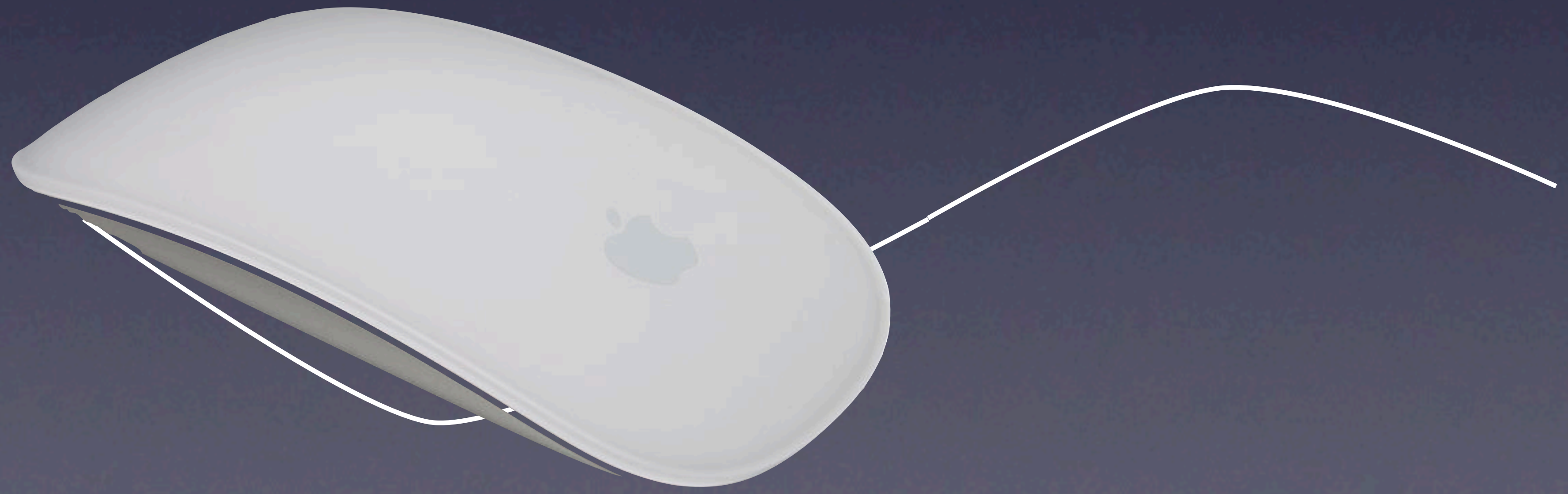
$\neq$

What we want



Deliberate, targeted movements

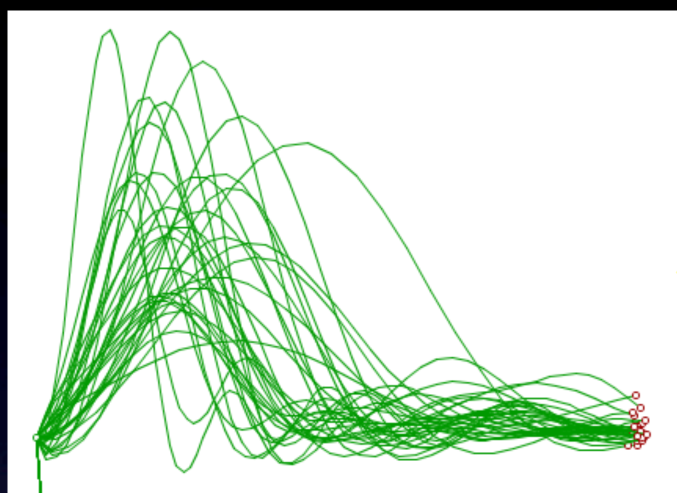
Deliberate, targeted movement?

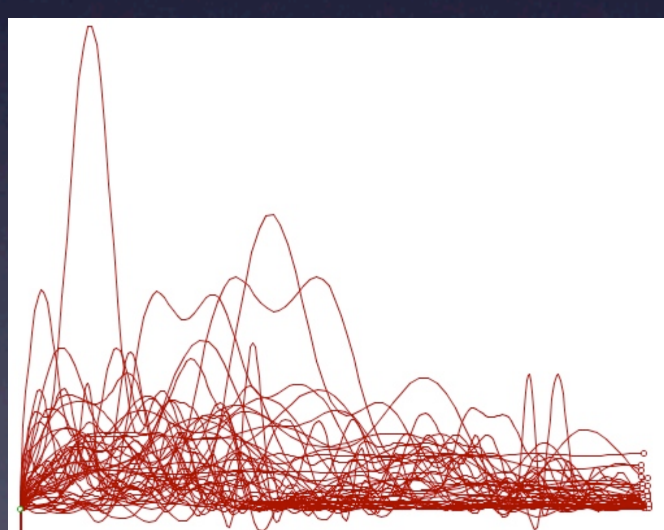



Machine
Learning

Data from a
Positive
experiment

Positive
examples



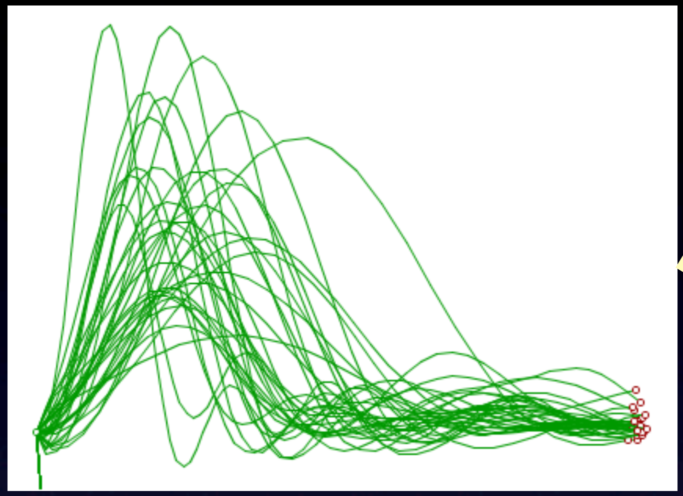


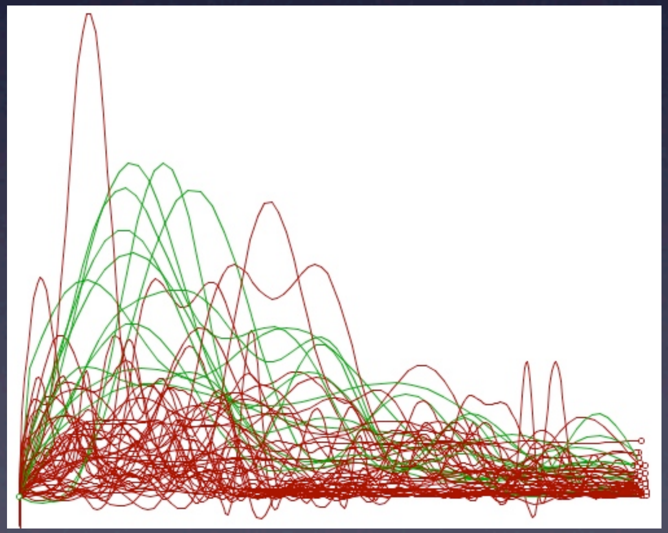
Negative
examples

???

= classifier

Data from a formal experiment

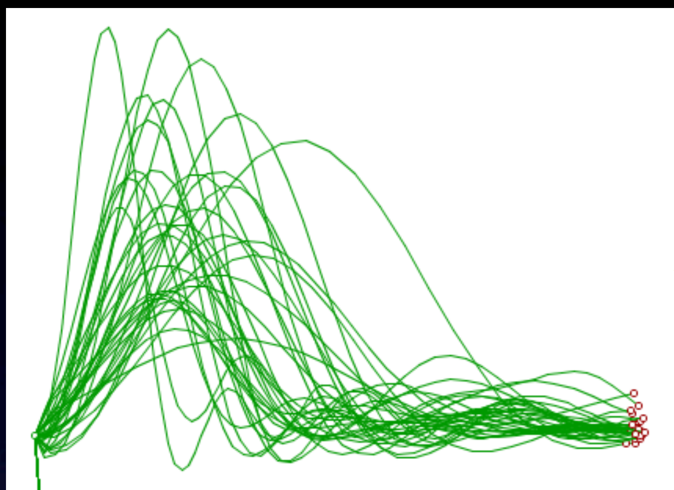


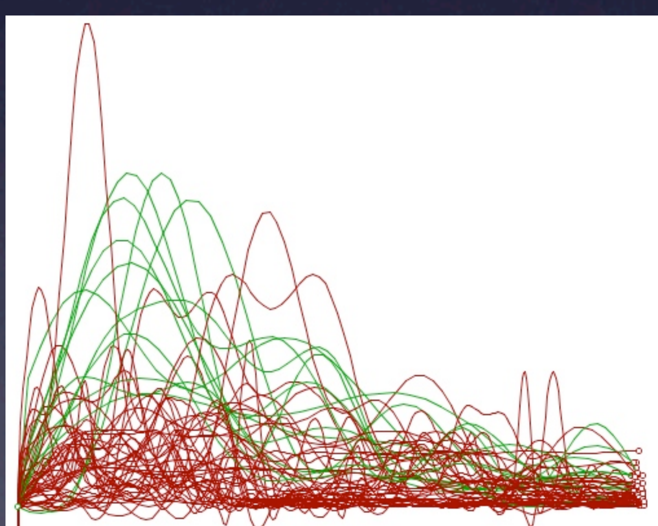


?? = classifier

Data from in situ observations = Mix of unlabeled +ve and -ve examples

Data from a formal experiment



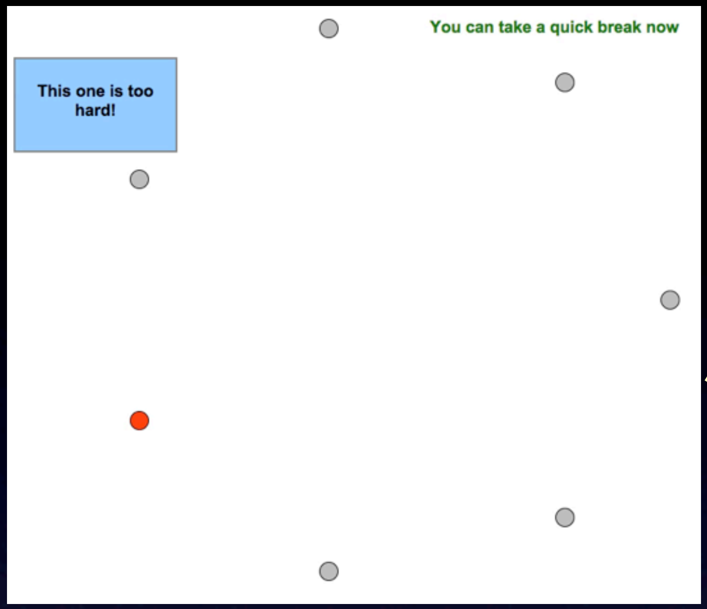


[Elkan & Noto, KDD'08]

= classifier

Data from in situ observations = Mix of unlabeled +ve and -ve examples

Positive
Examples



Unlabeled
Examples

[Elkan & Noto, KDD'08]

= classifier

Learning Classifiers from Only Positive and Unlabeled Data

Charles Elkan
Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093-0404
elkan@cs.ucsd.edu

Keith Noto
Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093-0404
knoto@cs.ucsd.edu

ABSTRACT

The input to an algorithm that learns a binary classifier normally consists of two sets of examples, where one set consists of positive examples of the concept to be learned, and the other set consists of negative examples. However, it is often the case that the available training data are an incomplete set of positive examples, and a set of unlabeled examples, some of which are positive and some of which are negative. The problem solved in this paper is how to learn a standard binary classifier given a nontraditional training set of this nature.

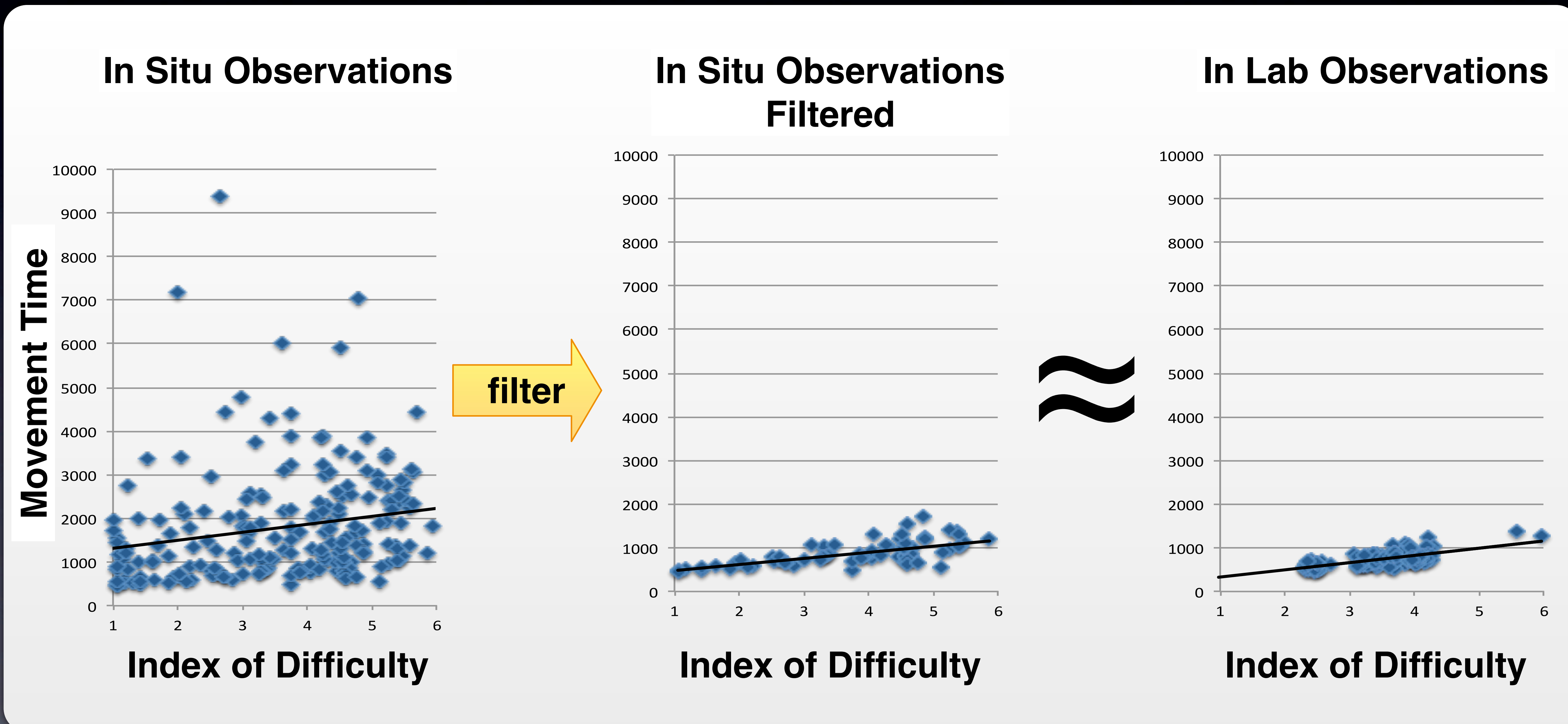
Under the assumption that the labeled examples are selected randomly from the positive examples, we show that a classifier trained on positive and unlabeled examples predicts probabilities that differ by only a constant factor from the true conditional probabilities of being positive. We show how to use this result in two different ways to learn a classifier from a nontraditional training set. We then apply these two new methods to solve a real-world problem: identifying protein records that should be included in an incomplete specialized molecular biology database. Our experiments in this domain show that models trained using the new methods perform better than the current state-of-the-art biased SVM method for learning from positive and unlabeled examples.

1. INTRODUCTION

The input to an algorithm that learns a binary classifier consists normally of two sets of examples. One set is positive examples x such that the label $y = 1$, and the other set is negative examples x such that $y = 0$. However, suppose the available input consists of just an incomplete set of positive examples, and a set of unlabeled examples, some of which are positive and some of which are negative. The problem we solve in this paper is how to learn a traditional binary classifier given a nontraditional training set of this nature.

Learning a classifier from positive and unlabeled data, as opposed to from positive and negative data, is a problem of great importance. Most research on training classifiers, in data mining and in machine learning assumes the availability of explicit negative examples. However, in many real-world domains, the concept of a negative example is not natural. For example, over 1000 specialized databases exist in molecular biology [7]. Each of these defines a set of positive examples, namely the set of genes or proteins included in the database. In each case, it would be useful to learn a classifier that can recognize additional genes or proteins that should be included. But in each case, the database does not contain any explicit set of examples that should *not* be included, and it is unnatural to ask a human expert to identify such a set. Consider the database that we are associated with, which is called TCDB [15]. This database contains

Results

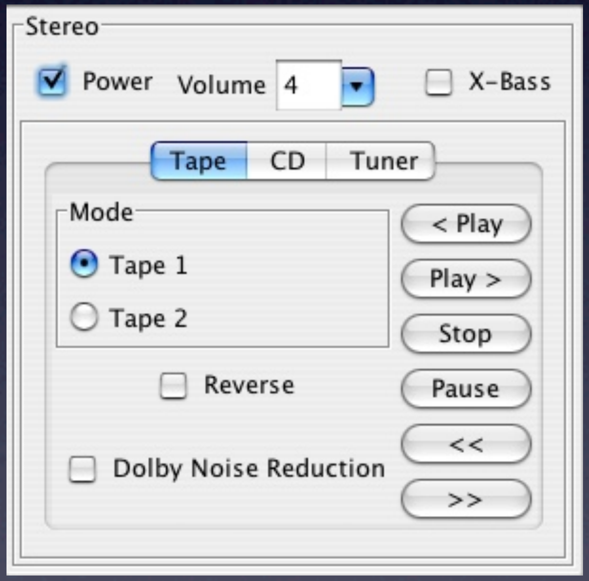


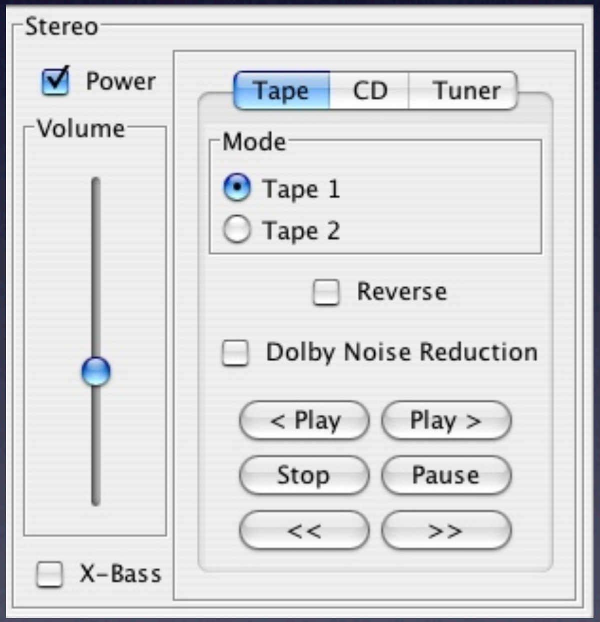
In Situ Observations
In Situ Observations Filtered
In Lab Observations
Movement Time
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
1
2
3
4
5
6
filter
≈
Index of Difficulty
Index of Difficulty
Index of Difficulty

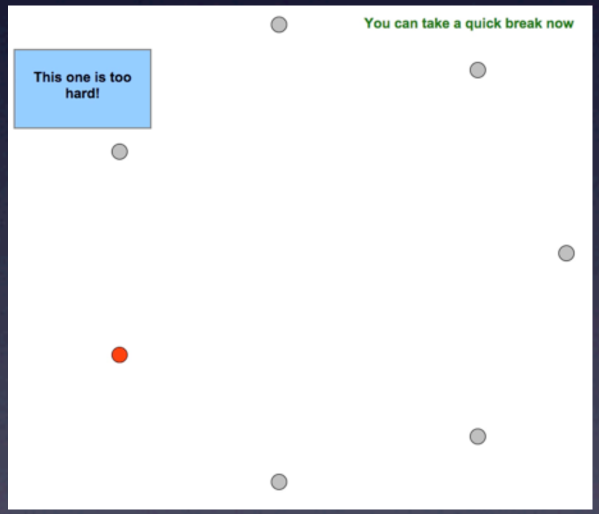
Conclusion

- Pick appropriate interaction for the task

Which is better?



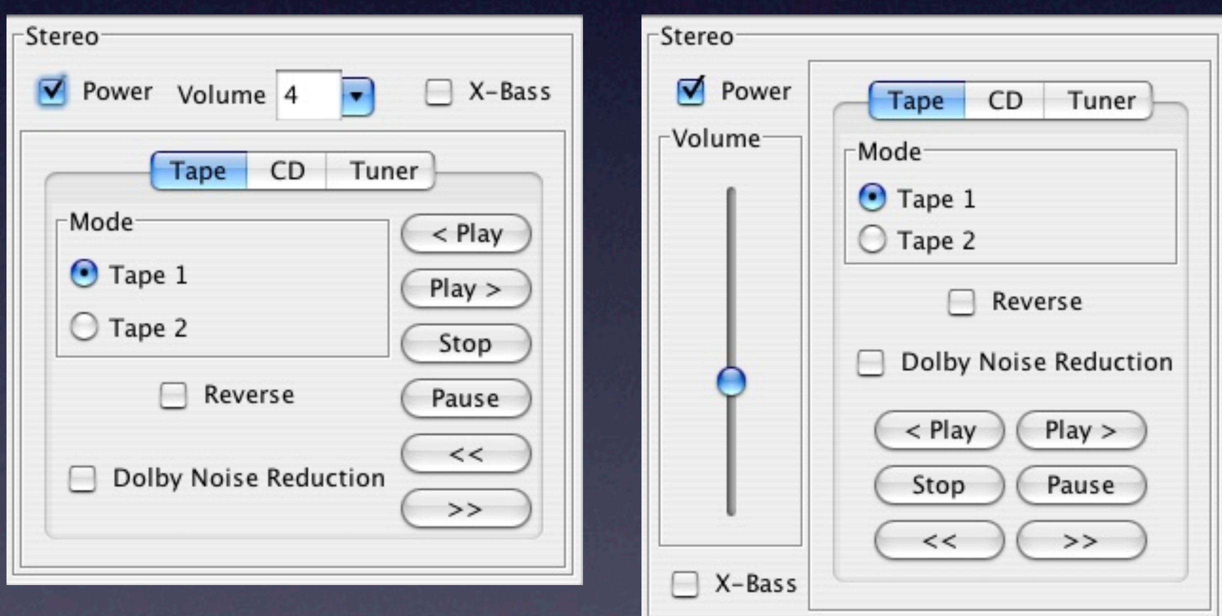


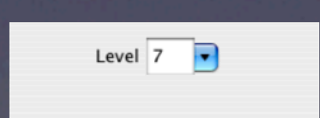


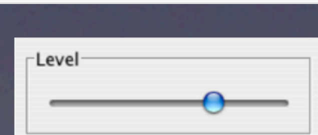
Conclusion

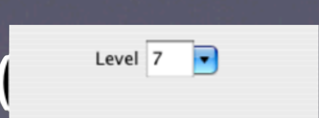
- Pick appropriate interaction for the task
- Understand the semantics of the interaction

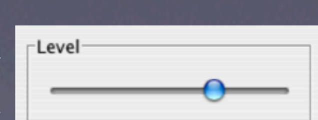
Which is better?



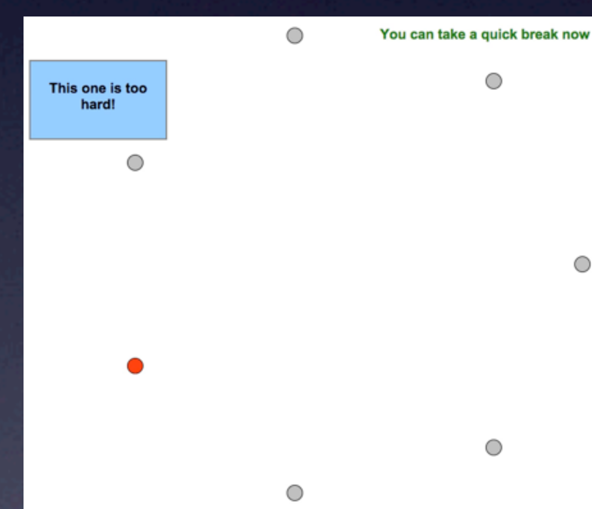


<


cost(

) ≥ cost(

)



Positive Examples

Unlabeled Examples

Conclusion

- Pick appropriate interaction for the task
- Understand the semantics of the interaction
- Develop/pick an algorithm that uses available information correctly and efficiently

Maximize:

$$margin - \sum_i slack_i$$

Subject to the constraints:

$$\sum_{e \in S} \sum_{k=1}^{K} u_k f_k(\mathrm{rend}_1(e)) - \sum_{e \in S} \sum_{k=1}^{K} u_k f_k(\mathrm{rend}_2(e)) \geq margin - slack_i$$

$$u_k \geq 0$$

$$u_k \leq C$$

[Elkan & Noto, KDD'08]

Conclusion

- Pick appropriate interaction for the task
- Understand the semantics of the interaction
- Develop/pick an algorithm that uses available information correctly and efficiently

Krzysztof Gajos
iis.seas.harvard.edu

HARVARD
School of Engineering
and Applied Sciences