

Talking Technical: Tricks of the Trade

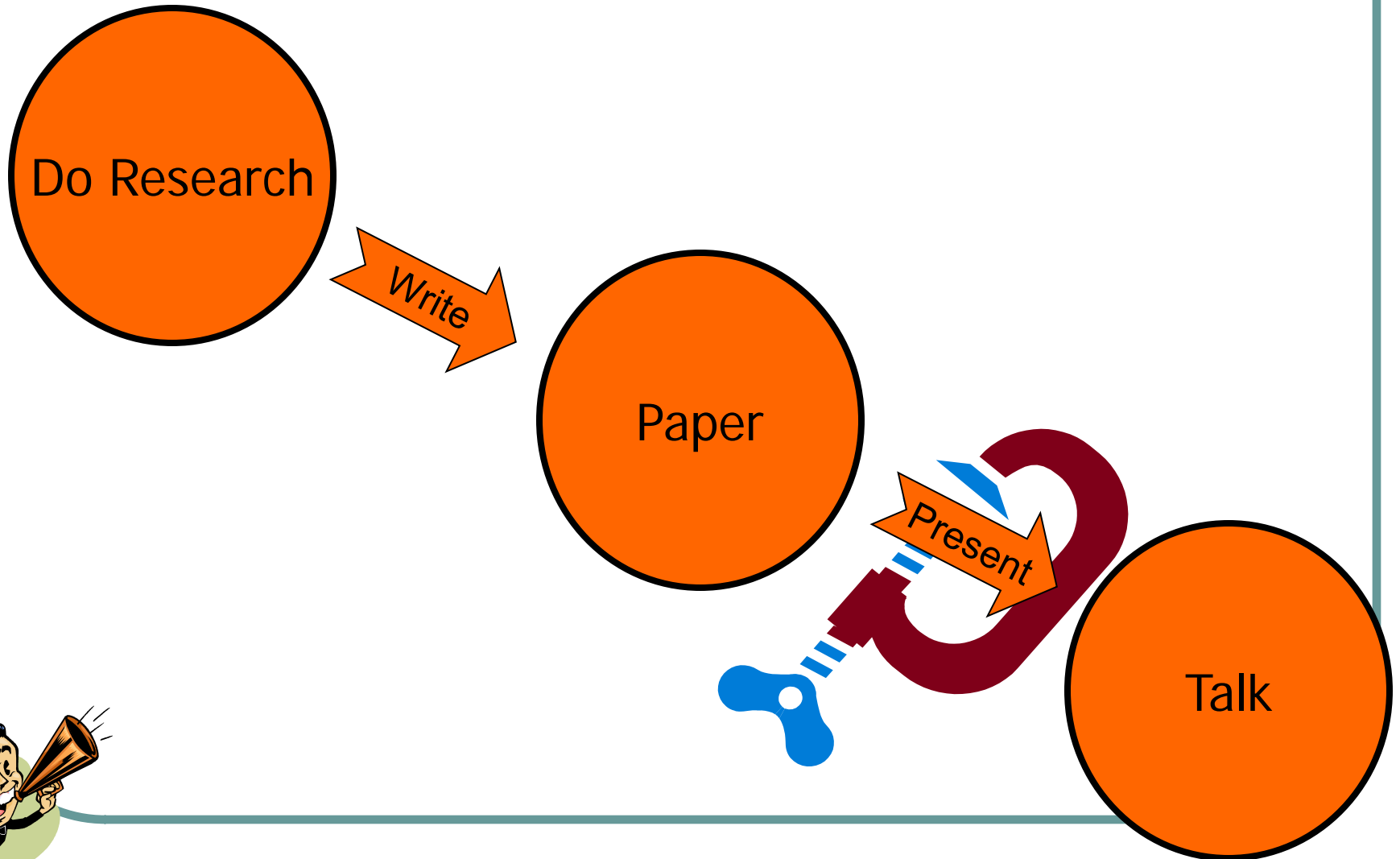
Terence Sim

10 Apr. 2012,
6 Apr. 2011,
28 Aug. 2008,
21 Mar. 2006

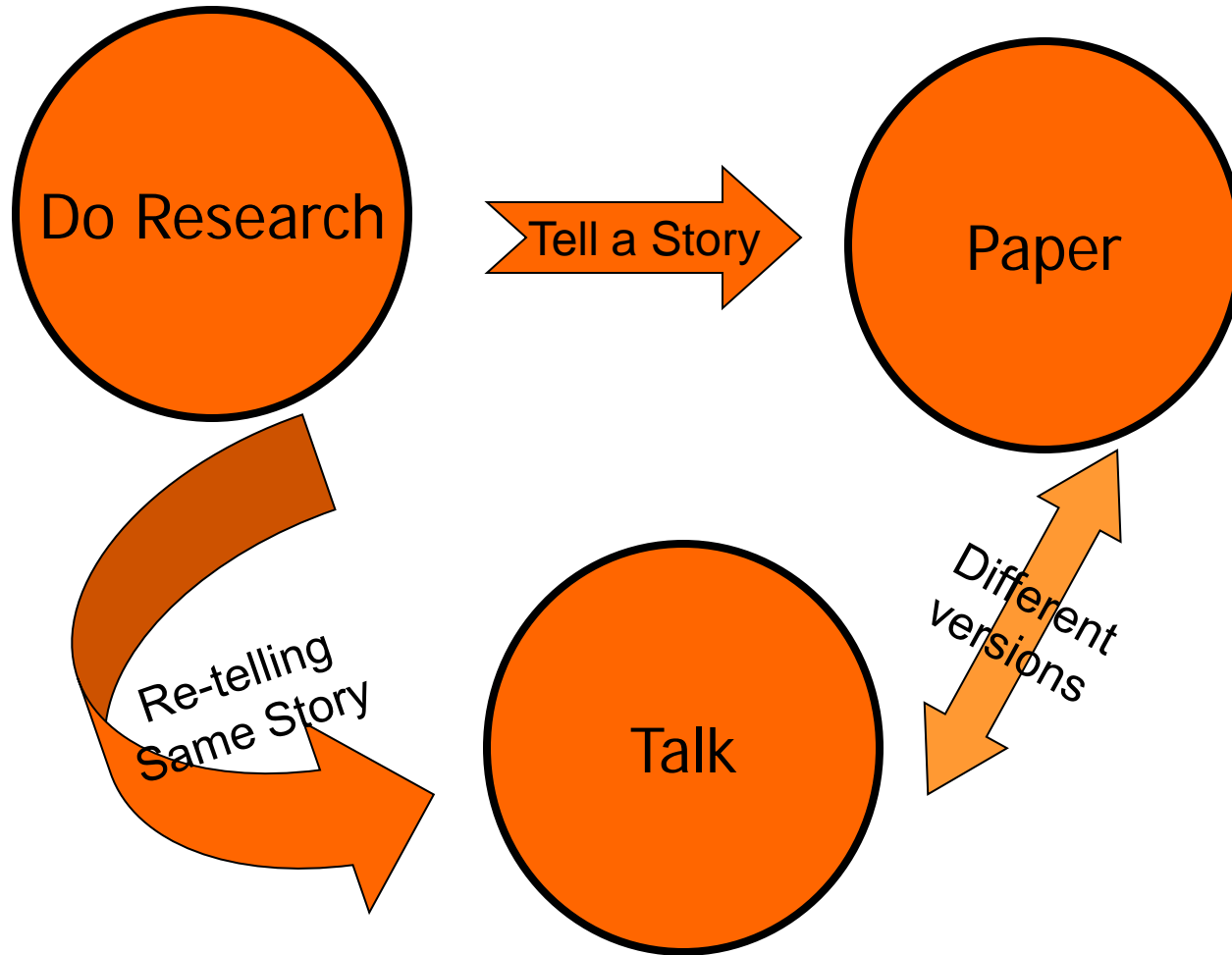
School of Computing
National University of Singapore



Talking Technical



A Better Picture



Same Story, Different Retelling

Paper

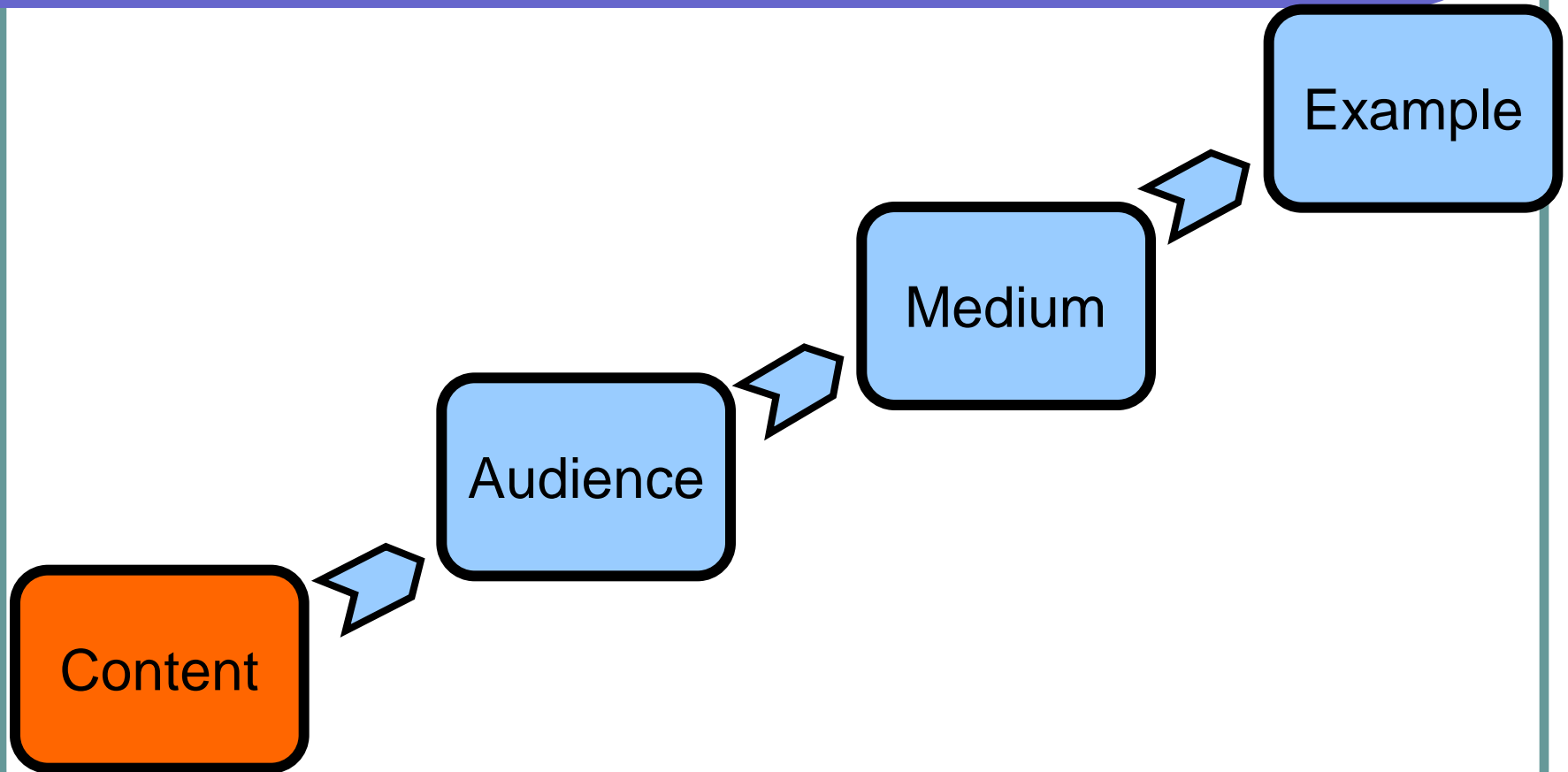
- Details
- Equations/Proofs
- Algorithms
- Experiments
- Charts/Figures/Table

Talk

- Talk \neq Compress(paper)
- Main ideas
- Motivation

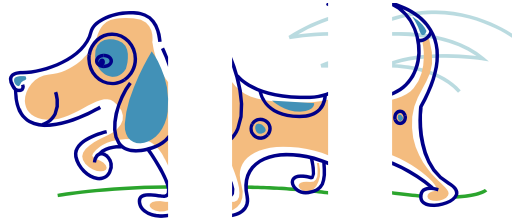


Road Map



Talk: Content

- Story:



- Main ideas of your research

- Details depend on type of talk

- Use mathematics sparingly!

- Avoid abbreviations unless commonly known

- SSFX vs. FSXF ???

- Enough details for people to understand complete story



Talk: Content

Besides describing your method,
talk about

- Motivation

- Why did you engage in this research?
- Why did you make certain choices?

- Surprises

- Any surprising discovery? Why, or why not?



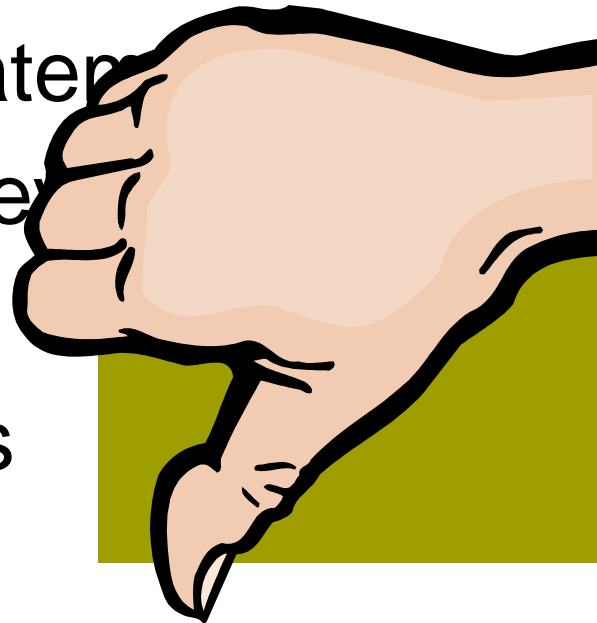
Prepare additional slides

- Hide them at the end
- These can contain additional proofs, experiments, diagrams, charts, etc.
- Anticipate some questions
 - Better still, lead audience to ask a question you have prepared for



Outline

- Introduction
- Problem Statement
- Literature Review
- Our Method
- Experiments
- Conclusion

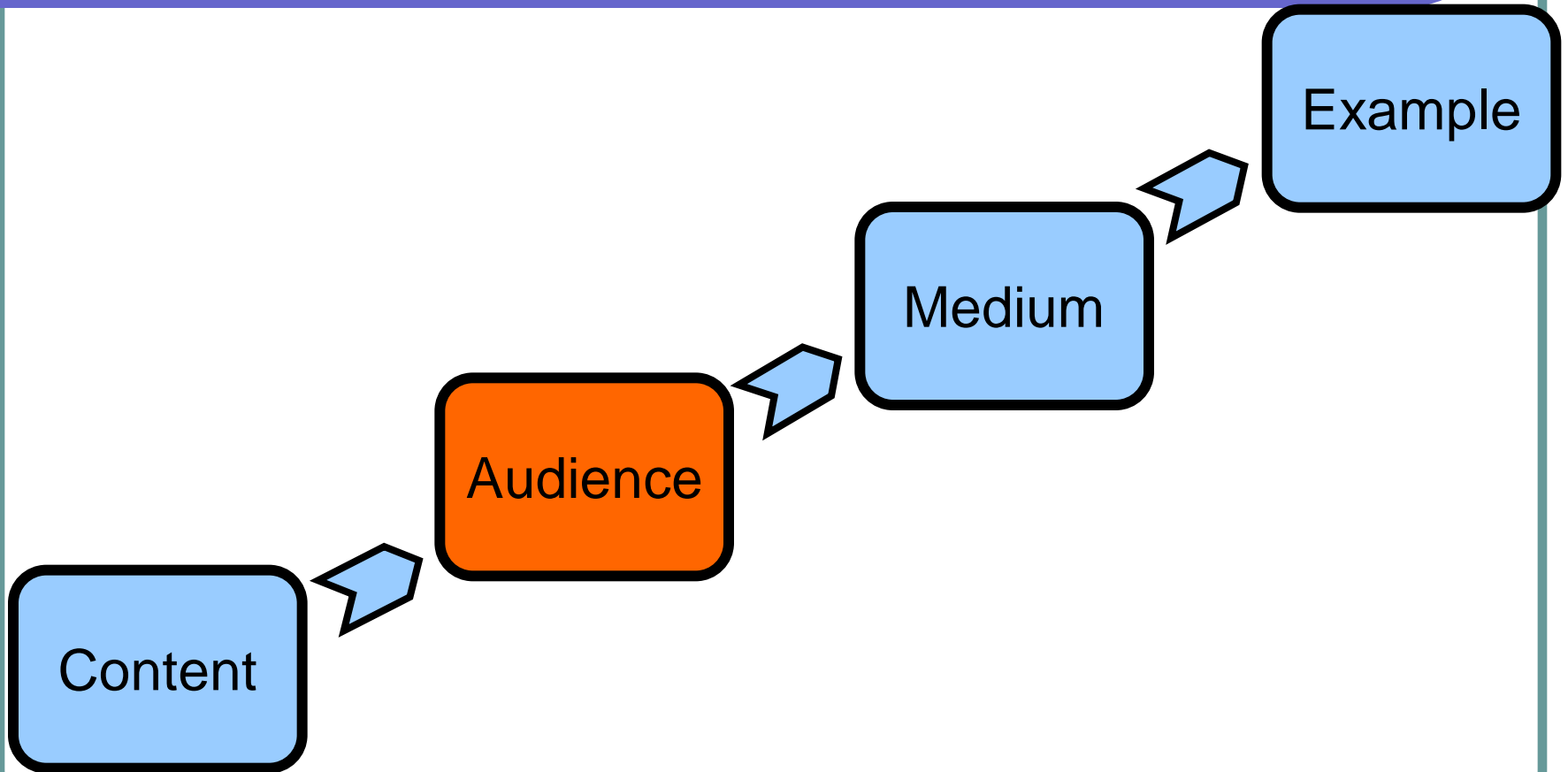


Meta-content

- Outline is meta-content,
 - a road map to navigate the talk
- Unnecessary if talk is short
 - Just start with the problem statement
- If used, simply let audience read
 - Don't insult audience
- If used, repeat it at appropriate places



Road Map



Talk: Audience

- Human psychology
- Put humans in a dimly lit, cosy room, with a constant background drone
 - What happens?



Human Psychology

- Limited short-term memory
 - Remembers 7 ± 2 things
- Short attention span
 - “Tunes out” quickly if nothing interesting
- Visual-Aural receptiveness
 - Responds to Visual + Aural stimuli
 - Responds to eye contact



5 ways to put audience to sleep

- Speak inaudibly: mumble
- Maintain monotonous voice
- Fill slides with lots of equations and text
- Avoid eye contact
 - Look at floor or ceiling
- Hide behind rostrum
 - Do not appear until talk is over



5 ways to engage audience

- Dress smartly and conservatively
- Speak clearly
 - project voice, pronounce words
 - vary pitch and pace of voice
- Avoid visual overload
 - Minimize symbols, use icons/images instead
- Look at audience: left, back of room, right
- Move around, gesture, smile!
 - But not too much!



Repetition

- Tell them what you're going to tell them
- Tell them
- Tell them what you told them



Repetition

- Tell them what you're going to tell them
 - In your Introduction
- Tell them
 - In your main body
- Tell them what you told them
 - In your summary



Handling Q & A

- No questions?
 - Usually means boring talk
- Listen to question carefully, make sure you understand, then answer it
- Repeat/rephrase question
 - Clarifies your understanding
 - Allows other people to hear question
- Don't get defensive!
 - Okay to admit ignorance, failure

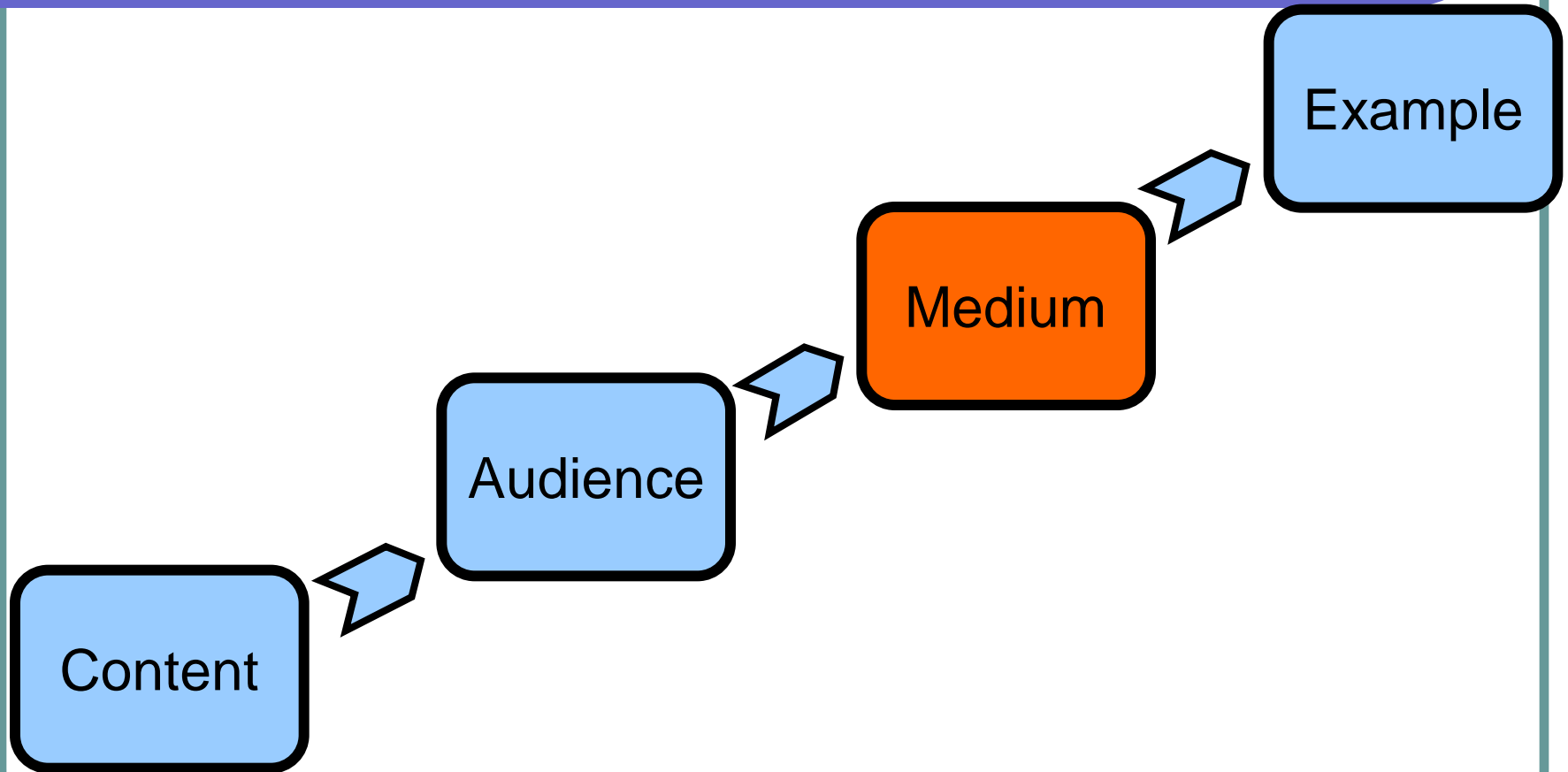


Handling Q & A

- Watch the clock!
 - Don't overrun your allotted time
 - Be flexible to adjust your pace
 - Don't let difficult questions derail your talk



Road Map



Talk: Medium

Paper

- Offline, passive
- No speaker; no sound
- Cross-reference possible
- Paper is *dead tree*, not interactive

Talk

- Real-time, active
- Speaker; guide
- Linear presentation
Limited X-ref
- Technological aids:
animation, interaction



Fonts

- Arial, Verdana
- Arial, Verdana
- Arial, Verdana



- Times Roman
- Times Roman
- Times Roman



Colors

- Dark background, white words, OR
- White background, black words
- Avoid gaudy colors

Colors

- Dark background, white words, OR
- White background, black words

- Avoid gaudy colors

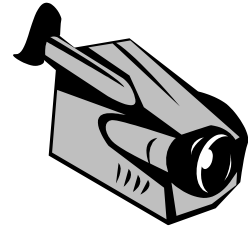
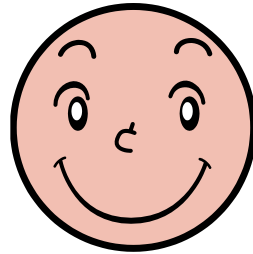


Animation + Video

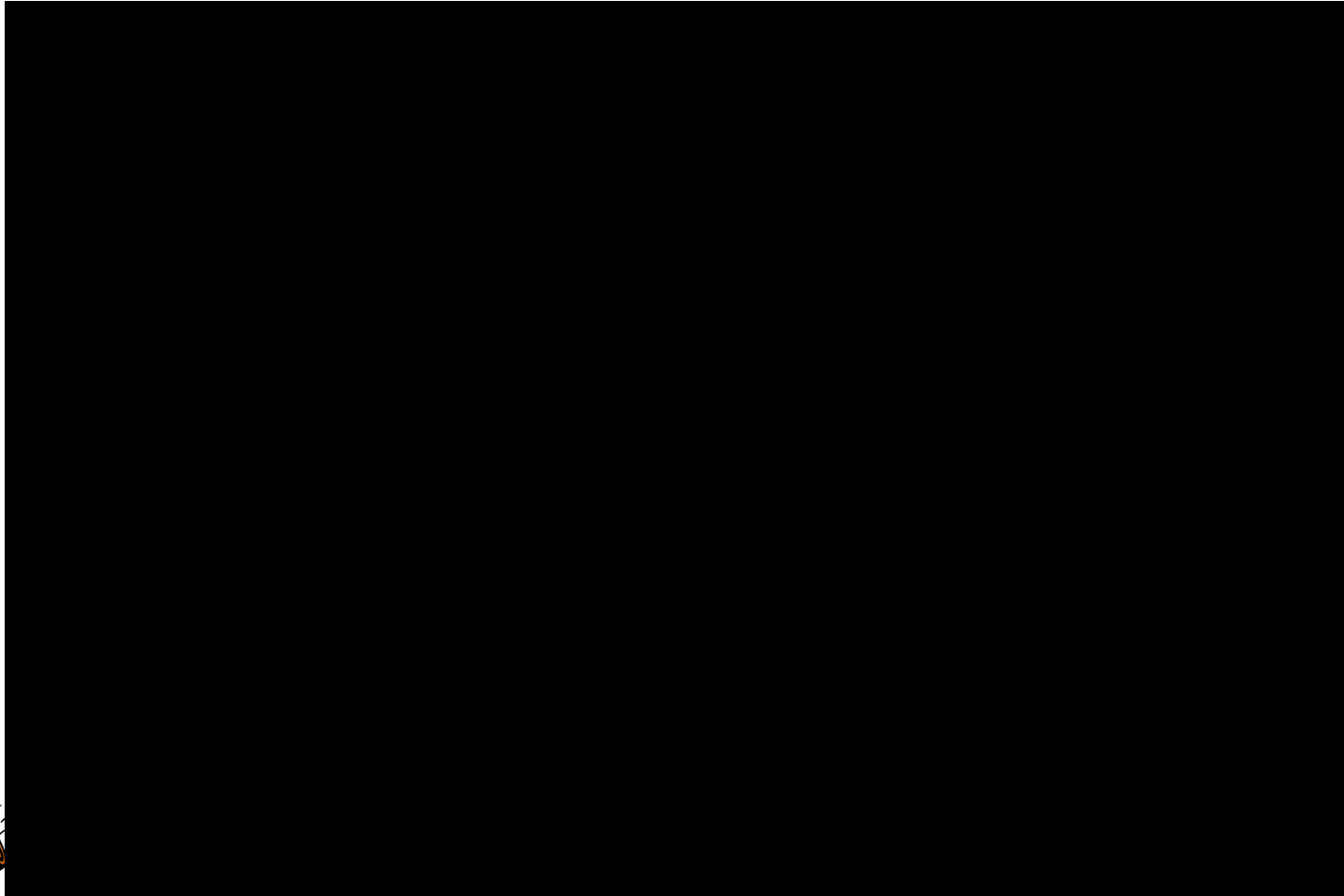
- We rendered each face under varying illumination and pose.
- Illumination: single light source placed from left to right at increments of 20° , and from bottom to top at increments of 20°
- Pose: camera placed from left to right at increments of 20° , and from bottom to top at increments of 20°



Animation + Video



Animation + Video



Example

Music Transcription Using an Instrument Model

Jun Yin, Terence Sim, Ye Wang and Arun Shenoy

ICASSP 2005

Music Transcription

1 Music score



A musical score on a single staff in 4/4 time, starting with a treble clef. The notation includes a series of eighth notes followed by a pair of beamed eighth notes and a final chord.

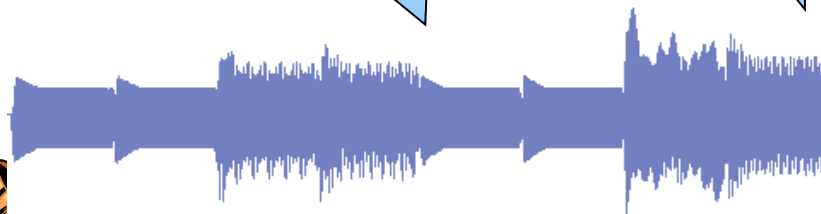


Synthesis

Easy!

Transcription

Hard!



Audio signal



Alternative notation

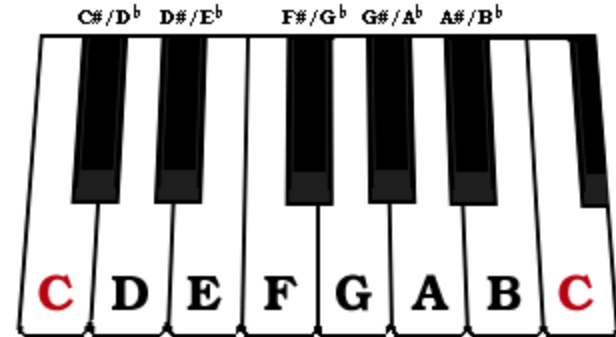
- MIDI format
 - Musical Instrument Digital Interface
 - Well-established “encoding”

Onset	Duration	Pitch	Loudness
1	29	20	1.5278
26	30	22	1.4738
52	30	20	1.4726
52	30	24	1.4952
77	31	22	1.4188
77	31	25	1.4322
103	30	27	1.4605
129	30	29	1.4593



Basic music terminology

- Musical Scale
 - A3=220 Hz
 - Exponentially Stepped
 - Semitone Step= $\sqrt[12]{2}$
 - Octave Step= 2



semitone

Note	Freq (hz)	Note	Freq (hz)
A3	$A3 \cdot 2^{(0/12)} = 220$	C#4	$A3 \cdot 2^{(4/12)} = 277$
A#3	$A3 \cdot 2^{(1/12)} = 233$	D4	$A3 \cdot 2^{(5/12)} = 294$
B3	$A3 \cdot 2^{(2/12)} = 247$	D#4	$A3 \cdot 2^{(6/12)} = 311$
C4	$A3 \cdot 2^{(3/12)} = 262$	E4	$A3 \cdot 2^{(7/12)} = 330$

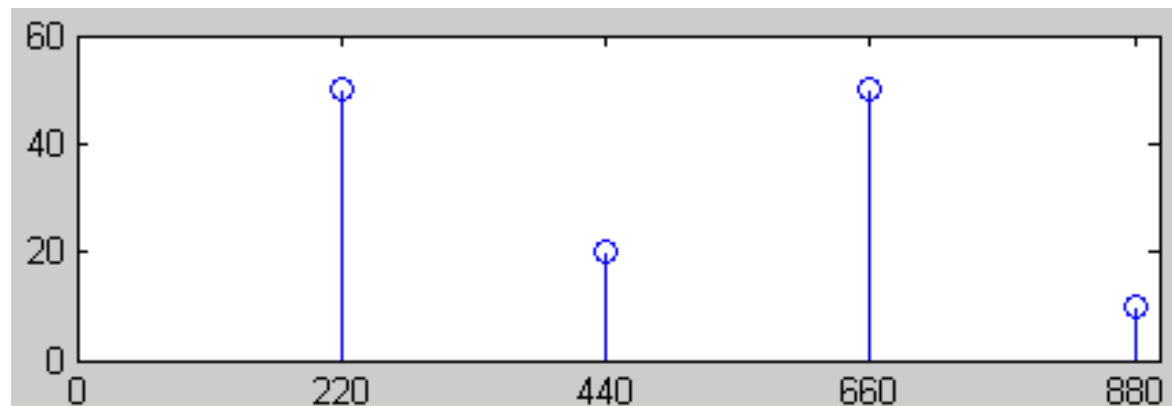


Basic music terminology

- Musical Sound

- Series of Sinusoid Waves
- Fundamental = F
 - Related to pitch
- Harmonics = kF , k integer
- Harmonic Structure: characterizes an instrument

Freq	Amp
220	50
440	20
660	50
880	10



Harmonic Structure: [1, 0.4, 1, 0.2]



Basic music terminology

- Monophonic: 1 note at a time
 - No simultaneous notes
 - Transcribing this is relatively easy
 - Polyphonic: many notes together
 - Harmonic structure overlap!
 - e.g. A3 + A4
 - (220, 440, 660, 880, ...) + (440, 880, ...)
 - e.g. C4 + E4 (some harmonics are close together)
- Hard to decipher



Idea

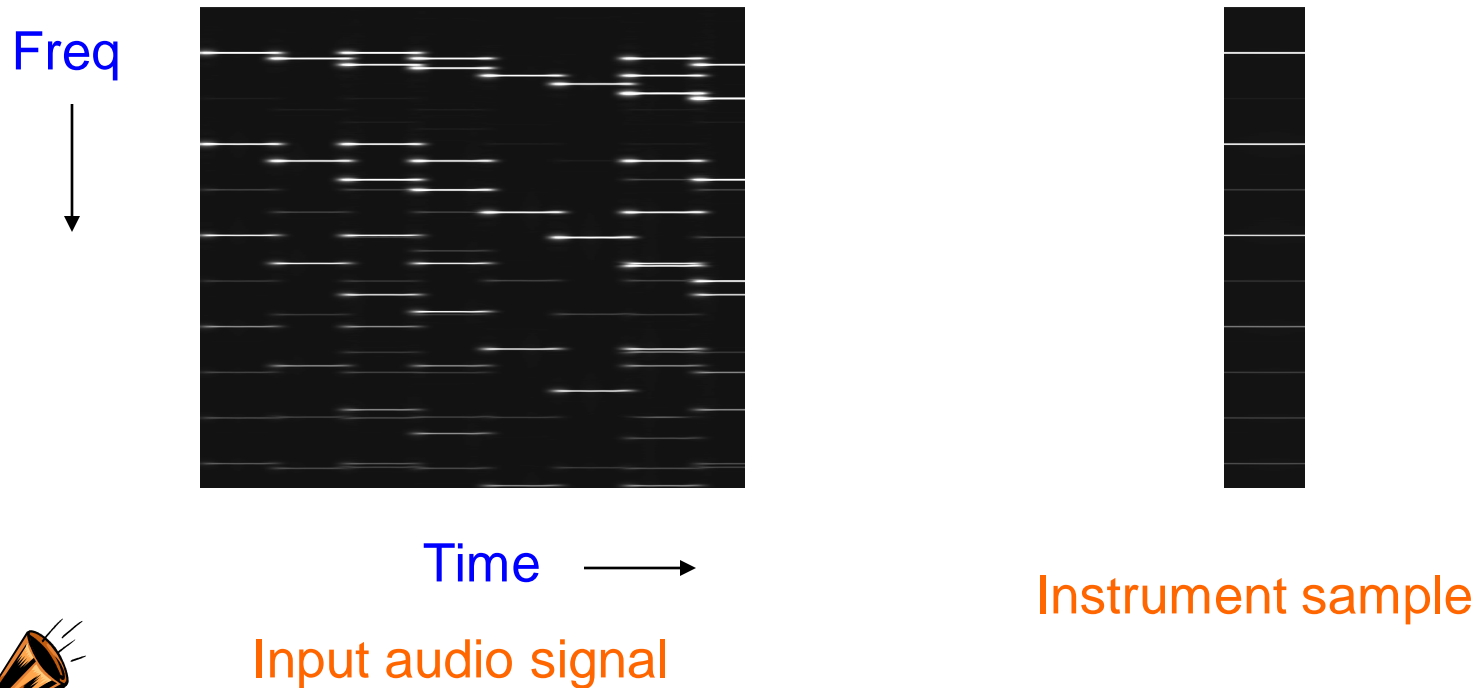
- Use model of instrument to disambiguate
- Assume harmonic structure
 - Constant across pitch
 - Constant over time
 - Only 1 sample required
 - True for certain instruments, e.g. piano

● Search for harmonic structure in audio signal



Method

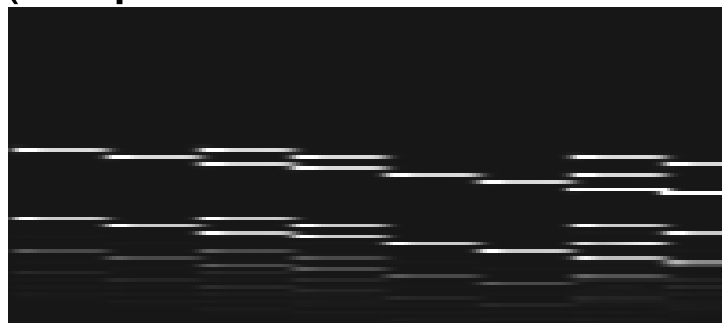
1. Create *frequency spectrum* from input audio and instrument sample



Method

2. Create *musical spectrum* from frequency spectrum

Discretize to 1496 bins
(88 pitches * 17 harmonics)

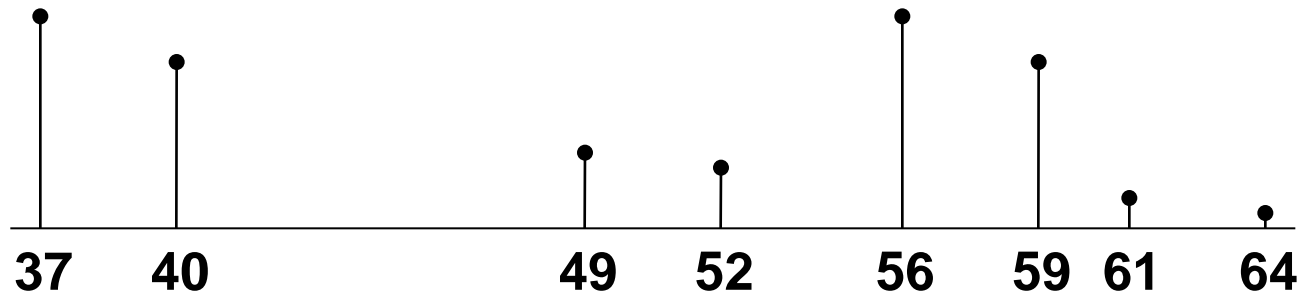


3. Match using *spectrum subtraction algorithm* -- estimates *pitch* and *loudness*

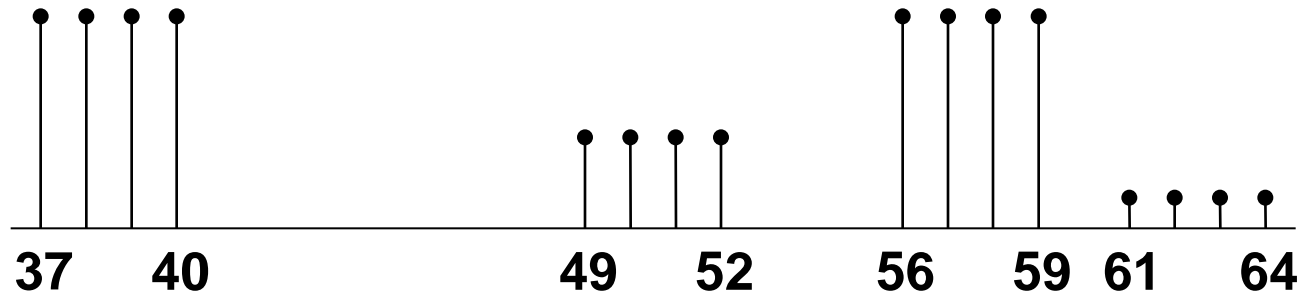


Spectrum Subtraction Algorithm

Input Z_M



**Ins.
model I**



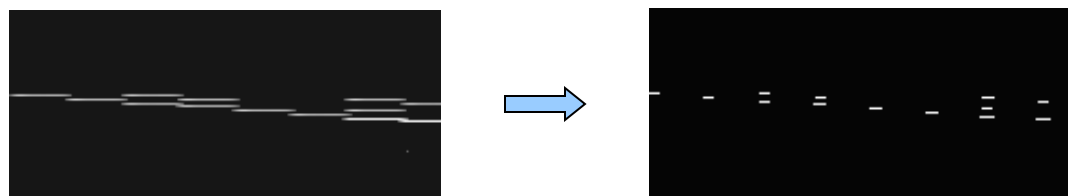
Slide Match Output

($a=1$, $p=37$)

($a=0.8$, $p=40$)

System Implementation

4. Detect *onset* and *duration*







5. Output table

Onset	Duration	Pitch	Loudness
1	29	20	1.5278
26	30	22	1.4738
52	30	20	1.4726
52	30	24	1.4952

6. Convert to MIDI file



Some Results

	Input	Output
● Segment 1		
● Minuet in G Major		



System performance

- Overall Precision: 0.96
Overall Recall: 0.98
- Performance not affected by
 - The duration of the note
 - The number of simultaneous notes
 - The instrument of the music, as long as the correct instrument model is used
- Performance degraded by
 - The pitch of the note is too low
 - The instrument harmonic structure differs from that in the music



Main Contributions

- Proposed to use Instrument Model for transcription.
 - Disambiguates between overlapping harmonics
 - Able to transcribe polyphonic music
- Developed *Spectrum Subtraction Algorithm* to estimate Pitch and Amplitude.
 - Efficient: linear in number of pitches
- (Not shown) Extended to multi-instrument transcription.



End of Example

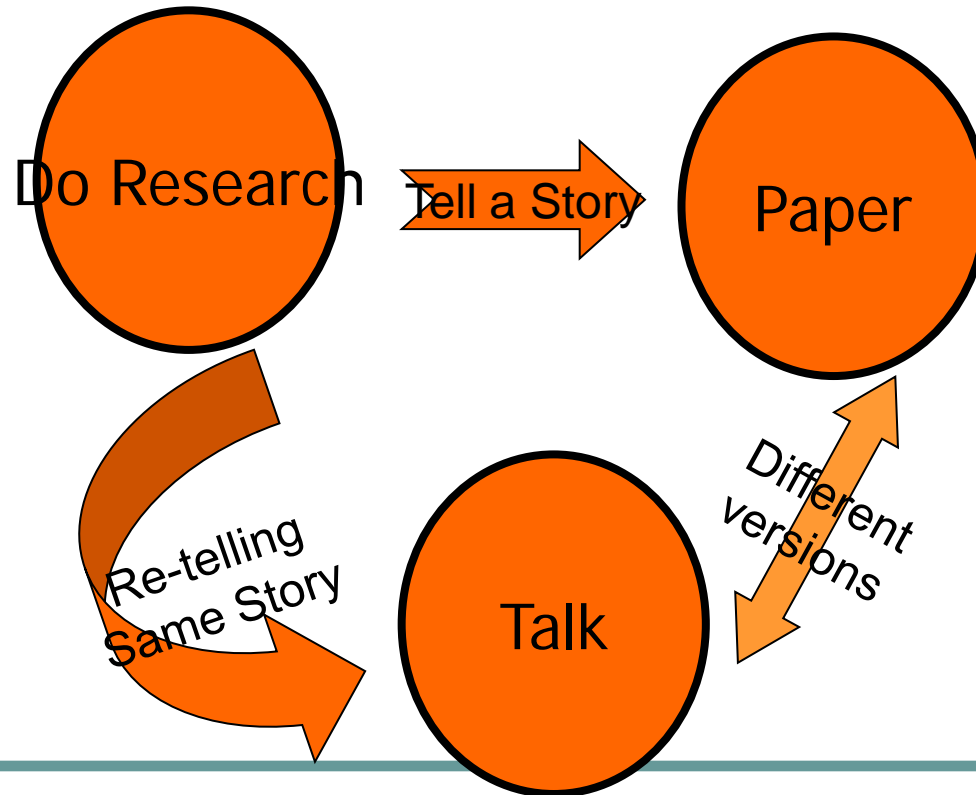
Critique

- How was the talk in terms of
 - Content
 - Audience
 - Medium ?
- How can it be improved?



Summary

- Technical Talk \neq Compress(paper)
- Pay attention to ***Content, Audience, Medium***



Thank You!

