DSO 699: Workshop on Mathematical Writing

The Pattern Perception Hypothesis

Vishal Gupta

Course Logistics

This is a workshop!

- Limited lecture from me

- YOU **MUST** PARTICIPATE FREQUENTLY
  - Be a respectful listener (Cameras on, mic's muted when not needed)
  - Be constructive, not critical

- You will have lots of writing assignments!
  - Use LaTeX
  - Reuse work from your research – summer paper

- Participate on Slack often
  - Give/solicit feedback
  - Share examples of good/bad writing

I'm not an expert. (Don't tell anyone.)

- Everyone has scope to improve.
  - I'm still improving in my own writing.
  - We will use some of my drafts to revise in class.

- Most of this course is us teaching each one another through feedback/discussion
  - Seeing where a peer is confused highlights *why* that writing ineffective
  - Reading someone else's phrasing opens up possibilities

Is effective writing simply a matter of taste?

Is what counts as ``effective" language dependent?

What is the pattern perception hypothesis?

What is lucid writing?

- "Human perceptual processing has remarkable properties, the properties that enabled our ancestors to survive. Lucid writing exploits those properties." - McIntyre

- Examples of perceptual processing properties:
  - Sensitivity to "organic change"
  - Gap-Filling
  - Grouping

- Examples of writing techniques that exploit these properties:
  - Lucid Repetition
  - Grouping and Coherent Ordering
  - Explictness and Adhering to Convention

Is ``good" writing simply a matter of taste?

"[The Pattern Perception Hypothesis] suggests that much of the real experts' advice is not an arbitrary matter of style or culture but, rather, a reflection of how the human brain works – the result of biological as well as social evolution." -- McIntyre

This philosophy underlies the entire curriculum for this workshop.

Are second language speakers at a disadvantage for lucid writing?

Walking Lights

Walking Lights: Deconstructed

- Think about you perceived as you watched this video
  - Your brain picked a particular model, with particular linkages between the dots
  - Why? Why does your brain prefer that model?

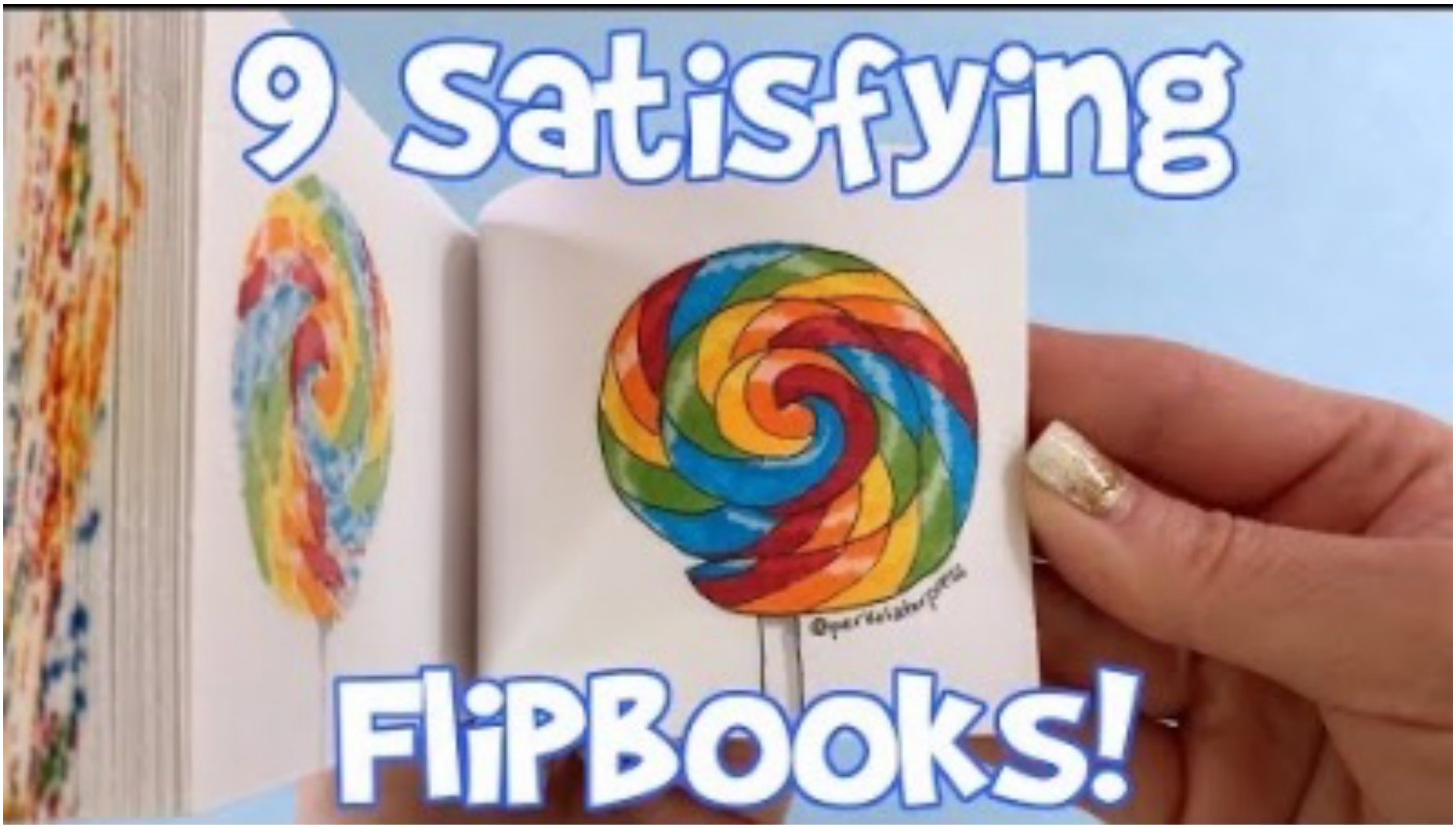
- We are perceptually sensitive to patterns in which

Organic Change

Some things change *slightly*,
while other things stay the same.

- Human beings not only PREFER to see organic change, but it aids their understanding

Example: Perceptual Accuity to Organic Change



A Mathematical Example: Organic Change

$$
\begin{aligned}
\mathbb{E}\left[X_i f(\boldsymbol{X})\right] &= \mathbb{E}\left[\mathbb{E}\left[X_i f(\boldsymbol{X}) \mid \boldsymbol{X}^{-i}\right]\right] && \text{(Tower Rule)} \\
&= \mathbb{E}\left[\mathbb{E}\left[\lambda_i f(\boldsymbol{X} - \boldsymbol{e}_i) \mid \boldsymbol{X}^{-i}\right]\right] && \text{(Stein-Chen Lemma)} \\
&= \mathbb{E}\left[\lambda_i f(\boldsymbol{X} - \boldsymbol{e}_i)\right] && \text{(Tower Rule).}
\end{aligned}
$$

- How many steps do you do per line? (But, really… how many?)
  - Why?
  - *"Two trivialities omitted can add up to an impasse."* *-- J.E. Littlewood*
- Why do we align the equations like this?

Organic Change and Adult Learners

- Organic change also relates to the way in which adults learn new things
  - (Different from how children/toddlers learn)

- A TON of educational research suggests:
  - Adults learn by anchoring new concepts in relation to concepts they already know
  - The more connections and anchoring, the more easily the adult attains mastery of the concept
  - The fewer connections and anchoring, the more likely the adult will forget the idea

- How does this FACT about adult learning processes relate to organic change?

How do I

exploit perceptual sensitivity to organic change in my writing?

Exploiting Organic Change I: Lucid Repetition

- What is “lucid repetition”?

- *Lucid repetition* is
  - Using the same words or phrases for similar things
  - Using different words or phrases for different things

By contrast,

- *Gratuitous variation* is
  - Using different words or phrases for the same thing

- *Incongruous repetition* is
  - Using the same words or phrases for different things

Examples: Lucid Repetition

- Example:
  - “Whereas the spectral method engenders Gibbs fringes, no discretization oscillations are manifested by the TVD algorithm.”
  - “Whereas the spectral method produces Gibbs fringes, the TVD methods produces no Gibbs fringes.”

  - “Girls are made of sugar, spice and everything nice, while slugs, snails and puppy dog tails make up boys.”
  - “Girls are made of sugar, spice and everything nice; boys are made of of slugs, snails and puppy dog tails.”

  - “Human perceptual processing has remarkable properties, the properties that enabled our ancestors to survive. Lucid writing exploits those properties.” -- McIntyre

Do not underestimate the power of lucid repetition!

- *"Ask not what your country can do for you. Ask what you can do for your country."* – John Fitzgerald Kennedy.
  - Compare to: "Don't ask what your country can do for you. Focus on what you can do for your country."

- "*Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that every word tell.*'' -- Strunk and White, *Elements of Style*

Exploiting Organic Change II: Avoid Gratuitous Variation

- "Formerly, science was taught by the textbook method, while now the laboratory method is employed."
- "Formerly, science was taught by the textbook method; now it is taught by the laboratory method."

Taken from Strunk and White, *Elements of Style*

- "Whereas the spectral scheme produces Gibbs fringes, the TVD method gives rise to no discretization oscillations."
- "Whereas the spectral method produces Gibbs fringes, the TVD methods produces no Gibbs fringes."

Taken from McIntyre

Exercise: (Alone)

- The author seemed to want to either define Gibbs fringes or remind the reader what they were in this sentence.
  - “Whereas the spectral scheme produces Gibbs fringes, the TVD method gives rise to no discretization oscillations.”

- Rewrite the sentence to
  - Use lucid repetition
  - Explicitly define Gibbs fringes/remind the reader what it is.

Avoid repetition that does not convey meaning

- Repetition hides the organic change

```
df <- tibble::tibble(
  a = rnorm(10),
  b = rnorm(10),
  c = rnorm(10),
  d = rnorm(10)
)

df$a <- (df$a - min(df$a, na.rm = TRUE)) /
  (max(df$a, na.rm = TRUE) - min(df$a, na.rm = TRUE))
df$b <- (df$b - min(df$b, na.rm = TRUE)) /
  (max(df$b, na.rm = TRUE) - min(df$a, na.rm = TRUE))
df$c <- (df$c - min(df$c, na.rm = TRUE)) /
  (max(df$c, na.rm = TRUE) - min(df$c, na.rm = TRUE))
df$d <- (df$d - min(df$d, na.rm = TRUE)) /
  (max(df$d, na.rm = TRUE) - min(df$d, na.rm = TRUE))
```

- Rewrite to emphasize organic change

```
rescale01 <- function(x) {
  rng <- range(x, na.rm = TRUE)
  (x - rng[1]) / (rng[2] - rng[1])
}
```

```
df$a <- rescale01(df$a)
df$b <- rescale01(df$b)
df$c <- rescale01(df$c)
df$d <- rescale01(df$d)
```

Let's summarize

- ***Defined*** lucid writing is writing which exploits the properties of human perception

- A first property: Sensitivity to Organic Change
- How Exploit?
  - Use lucid repetition
  - Avoid gratuitous variation
  - Avoid repetition that obscures the changing elements

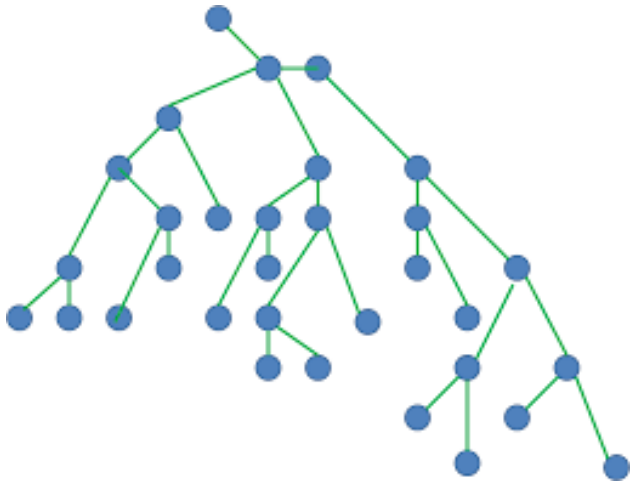
Exploiting Organic Change III: Coherent Ordering

- McIntyre refers to coherent ordering at a micro-level
  - "Rule 23: Establish terminology at first occurrence" – Strunk and White
  - Example:

$$\text{Let } x \text{ be the order quantity so that the } t\text{-period cost is } c(x, d_t).$$

$$\text{The } t\text{-period cost is } c(x, d_t), \text{ where } x \text{ is the order quantity.}$$

- But coherent ordering is a deeper idea that also applies at a maro-level
  - How should you structure the overall sequence of ideas to convey meaning?
  - An analogy: Tree Traversal.
    - How many ways to do this while respecting organic change?



Other Thoughts on Coherent Ordering

- We will talk a lot about coherent ordering
  - Particularly when we discuss outlining
- How can you identify *incoherent* ordering? Pay attention to the questions people ask?
  - My advisor's line margins….
- The benefits of coherent ordering are often conciseness and clarity, both at micro and macro scales.

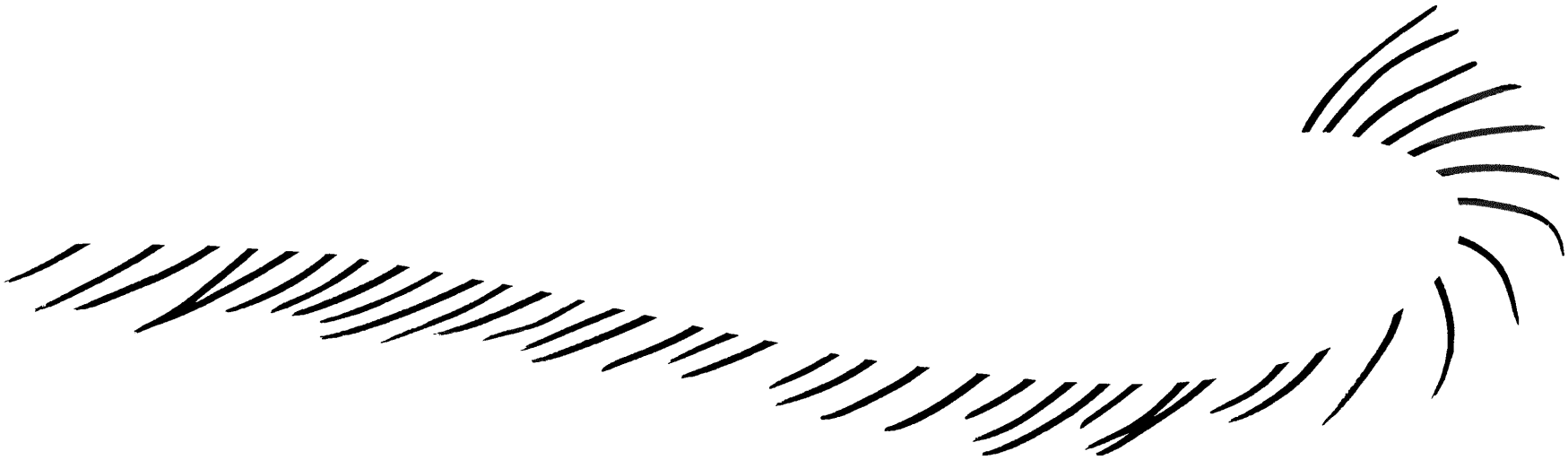
Quick Break: BB King and Organic Change

Let's summarize

- ***Defined*** lucid writing as writing which exploits the properties of human perception

- A first property: Sensitivity to Organic Change
- How Exploit?
  - **Use lucid repetition**
  - **Avoid gratuitous variation**
  - **Strive for coherent ordering**

A Second Perceptual Property: Gap-Filling



Can you see the smooth curve? Why?

A Second Perceptual Property: Gap-Filling

Gap-Filling

- Gap-filling is an unconscious process
- You see some things, and your brain INFERS that other things are there
- As you read, your brain does the same thing.
  - This is basically “Model-fitting” or thought-extrapolation

- How can we exploit this?
  - Encourage the reader to (unconsciously) make the RIGHT inferences
  - Prevent the reader from (unconsciously) making the WRONG inferences

- In other words, help the reader predict what’s coming next
  - Don’t surprise them or frustrate them

How do I exploit gap-filling in my writing?

Exploiting Gap-Filling I: Preventing wrong inferences

- Mathematics has conventions. Respect them.
  - Lower case Greek letters – typically parameters
  - Boldfaced lower case letter – vector
  - Boldfaced capital letter – matrix
  - Use lucid repetition: similar symbols for similar things

- Avoid the Stroop Effect - delay in reaction because of incongruous stimuli

Green Red Blue

Purple Red Purple

Mouse Top Face

Monkey Top Monkey

Can you (mic muted) read the **COLOR** of each of these words?

Is top list or bottom list easier?

Exercise

- Go through each example and
  1. Identify the gap-filling that is taking place. What does the reader expect?
  2. Be specific about how the writing frustrates the expectation
  3. Rewrite the example (change variables if needed)

- Take a, S, $\mathbb{P}$, and γ to be functions.

- Let $N(\cdot)$ be a linear function and $L(\cdot)$ be a non-linear function.

- Claim: If $x_1 \in x_4 \subseteq x_3 \setminus x_5$, then $x_1 \notin x_5$.

- Let s and t be two times with $t_0 > s > t > T$.

More Subtle Examples

- "Skilled writers distinguish what helps the reader versus indulging themselves."
  - Where did we frustrate the reader's expectation?
  - Use "distinguish … **from**" instead of "distinguish … **versus"**

- "Skilled writers distinguish what helps the reader from what indulges the writer.''

- Strunk and White has a collection of these types of examples where subtle violations of parallelism sound "odd." (See Item 19 of *Elements of Style*.)
  - Example: "Both … and"
  - "It was both a long lecture and very tedious."
  - "It was both a long and tedious lecture."

Exploiting Gap-Filling II: Be Explicit and Plain

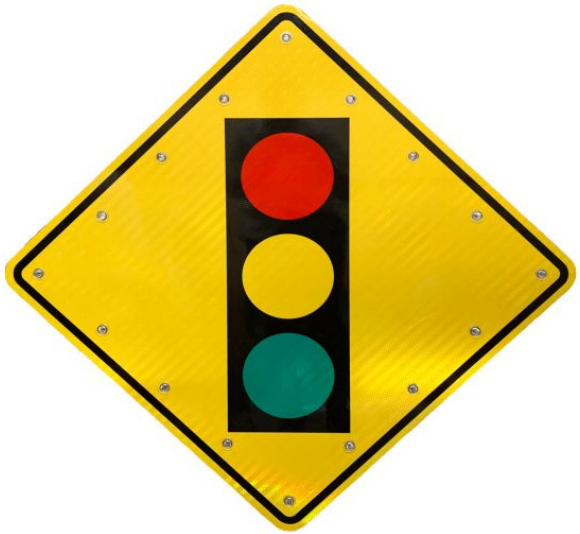
- Be explicit! More explicit than you think:
  - *"Two trivialities omitted can add up to an impasse." J.E. Littlewood*

- Typical proofs follow a sandwich structure
  - We are going to prove XXXX
  - ** Prove XXX ***
  - This completes the proof of XXX.

- In fact
  - Claims within a proof follow this structure
  - Sections of a paper follow this structure
  - Entire paper follows this structure!

How explicit?

“Lucid, informative writing minimizes the computational load on the reader's pattern-perception machinery. It tries to save the reader's time, not the writer's. It is like good road signposting, boringly explicit and unvaried from the writer's viewpoint. The lie of the land is what interests the reader, not the signposts as such.” -- McIntyre



Exercise:

- Go through the first 2 pages of your longer written work.
- Identify all the “sign-posting” that you used.

- Come back and we will discuss.

Quick Break?

Let's summarize

- ***Defined*** lucid writing as writing which exploits the properties of human perception

- A first property: Sensitivity to Organic Change

- A second property: Gap-filling
- How exploit?
  - **Sketch just enough of the curve to let reader fill it in**
  - **Abide by conventions to not frustrate reader expectations**
  - **Be explicit. More than you think you need to be.**

A Third Property: Perceptual Grouping

- How many dots are there in each group?

- Do you perceive this as 13 dots? Or as 4 groups of dots?
  - What about the first grouping? Isn't it really 2 groups of 2 dots?
  - What about the second?

- The human mind groups by space in space, time, color, shape, …

- Grouping affect our UNDERSTANDING of what we are seeing
  - People who design instrument panels know this deeply
  - Who knows the story of "Three Mile Island"?

Pattern Grouping in LaTeX

$$\begin{aligned}
\mathbb{E}\left[X_i f(\boldsymbol{X})\right] &= \mathbb{E}\left[\mathbb{E}\left[X_i f(\boldsymbol{X}) \mid \boldsymbol{X}^{-i}\right]\right] && \text{(Tower Rule)} \\
&= \mathbb{E}\left[\mathbb{E}\left[\lambda_i f(\boldsymbol{X} - \boldsymbol{e}_i) \mid \boldsymbol{X}^{-i}\right]\right] && \text{(Stein-Chen Lemma)} \\
&= \mathbb{E}\left[\lambda_i f(\boldsymbol{X} - \boldsymbol{e}_i)\right] && \text{(Tower Rule).}
\end{aligned}$$

- Uses whitespace to group logical units

```
\begin{align*}
\Eb{ X_i f(\bX)}
    & \ = \
    \Eb{ \Eb{ X_i f(\bX) \mid  \bX^{-i} } }
        &&(\text{Tower Rule})
    \\
    & \ = \
    \Eb{  \Eb{ \lambda_i  f(\bX - \be_i) \mid  \bX^{-i} } }
        &&(\text{Stein-Chen Lemma})
    \\
    & \ = \
    \Eb{  \lambda_i  f(\bX - \be_i) }
        &&(\text{Tower Rule}).
\end{align*}
```

- Line break don't coincide with logical units

```
\begin{align*}
\mathbb E[ X_i f(\bX)]  & \ = \ \mathbb E[ \mathbb E[ X_i f(\bX) \mid  \bX^{-i} ] ] &&(\text{Tower Rule}) \\
& \ = \ \Eb{  \Eb{ \lambda_i
f(\bX - \be_i) \mid  \bX^{-i} } } &&(\text{Stein-Chen Lemma})
 \\  & \ = \
   \Eb{
    \lambda_i  f(\bX - \be_i)
    }
       &&(\text{Tower Rule}).
\end{align*}
```

Exercise: Grouping Mathematical Formulas

- Use grouping to rewrite these equations to improve understanding

- Let $a_n = (-1)^{n \bmod 2}\ n^2$

- Let $\mathbf{p} \in \mathbb{R}^{4n}$ be such that $p_i = 1/i$ for $1 \leq i \leq n$, $p_i = 0$ for $n+1 \leq i \leq 3n$, and $p_i = 1/(4n - i + 1)$ for $3n < i \leq 4n$.

- $n^{-1/2}\sqrt{\log d + \log(1/\epsilon)} \cdot \dfrac{65}{m}$

Let's summarize

- ***Defined*** lucid writing as writing which exploits the properties of human perception
- A first property: Sensitivity to Organic Change
- A second property: Gap-filling
- A third property: Grouping
- How exploit?
  - Group similar things (formulas, ideas, terms) near one another
  - Separate dissimilar things (formulas, ideas, terms)

Final Words

- You can go very deep with this idea of linking perceptual properties to good writing style
  - Read “Lucidity Principles in Brief” for some brief notes

- The lines between the perceptual phenomena are blurry
  - What’s important is the thesis of lucid writing
  - This thesis has shaped (and continues to shape) my own writing

  - *“Lucid, informative writing is the only defensible kind of scientific writing. Lucid, informative writing minimizes the computational load on the reader's pattern-perception machinery. It tries to save the reader's time, not the writer’s.”* -- McIntyre

- Thoughts?