Marks Revisited: Beyond Bertin

Tamara Munzner
Department of Computer Science
University of British Columbia
Northeastern Vancouver
12 Jul 2024
@tamara@vis.social

Marks and channels: Foundational model
• decompose visual encoding into marks & channels
  - marks
    • geometric primitives
    • represent data items
  - channels
    • control appearance of marks
    • represent data distributions
widely used
- Bertin 1967
- Sensing of Graphics

Current Marks & Channels Model
- marks & channels model is a design space
  - descriptive power: ability to describe significant range of existing examples
  - evaluative power: ability to help assess multiple design alternatives
- generative power: ability to help designers create new designs
- many names: taxonomies, typologies, classifications, frameworks, models, grammars...
- design spaces help us reason
  - impose systematic & actionable structure on set of possibilities for specific problem
  - support reasoning about design choices
  - capture key variables of play
  - increase cognitive efficiency & technical effectiveness by grouping similar tasks together to facilitate meaning about design

Why analyze visual encodings?
- marks & channels model is a design space
- analyzing idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Talk outline
- explain current marks & channels model
- walk through many questions that arise when teaching it
- present preliminary ideas towards an alternative model

Visual encoding model
• analyze idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Visual encoding model: Spatial data
• marks for items of spatial data
  - idiom: choropleth map
  - marks: point for data items, rectangular bars for links

Visual encoding model: Temporal data
• marks for items of tabular data
  - idiom: table
  - marks: points for data items, rectangular bars for links

Visual encoding model: Multidimensional data
• marks for items of multidimensional data
  - idiom: matrix
  - marks: points for data items, rectangular bars for links

Design spaces in visualization: continuing theme
- exploring the design space of visual encodings
  - marks & channels model is a design space
  - analyzing idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Marks and channels: Foundational model
• decompose visual encoding into marks & channels
  - marks
    • geometric primitives
    • represent data items
  - channels
    • control appearance of marks
    • represent data distributions

Visual encoding model
• analyze idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Current Marks & Channels Model
- marks & channels model is a design space
  - descriptive power: ability to describe significant range of existing examples
  - evaluative power: ability to help assess multiple design alternatives

Why analyze visual encodings?
- marks & channels model is a design space
- analyzing idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Visual encoding model
• analyze idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point

Visual encoding model: Spatial data
• marks for items of spatial data
  - idiom: choropleth map
  - marks: point for data items, rectangular bars for links

Visual encoding model: Temporal data
• marks for items of tabular data
  - idiom: table
  - marks: points for data items, rectangular bars for links

Visual encoding model: Multidimensional data
• marks for items of multidimensional data
  - idiom: matrix
  - marks: points for data items, rectangular bars for links

Design spaces in visualization: continuing theme
- exploring the design space of visual encodings
  - marks & channels model is a design space
  - analyzing idiom structure as combination of marks and channels
  - 1 channel: vertical position
    - mark: line
  - 2 channels: vertical position, horizontal position
    - mark: point
Rethinking book design space: Visualization Analysis & Design 2e

Teaching Challenges

Quiz: Name marks/channels
• Shooting Media Coverage
  • A-points
  • B-lines
  • C-areas
  • A-position
  • B-color
  • C-length
  • D-area
  • E-angle

Quiz: Name marks/channels
• Tax Rates
  • marks
    – A-points
    – B-lines
    – C-areas
  • channels
    – A-position
    – B-color
    – C-length
    – D-area
    – E-angle

Mark/channel analysis: scope & limits
• model: one mark for one data item
  • model inherited from Bertin (Semiology of Graphics, 1967)
  • never questioned
  • geometric motivation
    – geometric primitives have dimensions
    – how could we argue with math?

Channels: Model evolves, heavily studied
• effectiveness rankings
• expressiveness matches, data & task

Encoding vs decode: Where do models diverge?
• idiom pie chart
  • encode: area marks with area channel
  • ordered: radius, uniform length
  • accuracy: area less accurate than rectangular aligned position/length
  • decode: not as tight, probably arc length, maybe also area

Constraints
• consider marks and channels as imposing constraints
  – when does mark type constrain channel use?
  – when does using one channel constrain another channel?

Alternative Ideas

Teaching: Bertini in-class exercises, catalyst for questions
• decoding marks & channels

Many, many questions
• so what?
  – evidence that this design space could be improved!

Quiz: Name marks
• points, lines, areas?

Encodings vs decoding models
• Encoding model: what should visualization designer do?
  – prescriptive model, providing guidance for design
  – predicting viewer response differs from inferring or reverse-engineering designer intent when encoding!

Many, many questions
• so what?
  – evidence that this design space could be improved!

Mark/Channel analysis: scope & limits
• model scope: one mark for one data item
  • model: one mark for one data item, glyphs, multiple views

Alternative Ideas

Teaching design space: analyze visual encoding & map to data
• assignment: analyze existing encoding with marks & channels
  – visual channels used?
    • channel X encodes attribute Y
  – marks used?
    • mark of type X encodes attribute Y

Many, many questions
• so what?
  – evidence that this design space could be improved!
Channel use: what does it mean?

- Does channel size encode directly to attributes?
  - yes? sizes differ
  - according to dog name in data
- not?: sizes differ not meaningful
- just emerges from choice of layout, radial or rectangular
- not a "real" attribute encoding

Can we use size channel to encode another attribute?
- no, not free
- it's "taken" already
- would change meaning
- Size channel is Unavailable

Channel Availability Model
- Encoded channels:
  - clear meaning
  - many channels can be directly used for redundant encoding
- Free: which channels free to encode another attribute?
  - without changing usability of existing encoding

Unavailable: which channels unavailable / precluded / taken?
- because of mark type
- because of idioms/algorithms design specifics
- because other channels used!

Area marks: Rethinking
- area marks are a terrible name
  - other marks all have graphical area too
- showing us to encode with color
- comparative points of view: they're all just polygons
- there's also an "area" channel, which is confusingly different
- area is not the only channel in play with these marks!

Area marks: obvious example: choropleth maps
- what can we do to California? could we encode additional data?
  - cannot shrinkgrow (size channel)
  - cannot translate (position channel)
  - cannot rotate (orientation channel)

why not?
- would lose meaning of that mark boundary is the data
  - also lose meaning for other occluded marks

"area" mark is not specific enough
- AreaPositionOrientationShape mark? nah...
- idea: Interlocking

Interlocking marks: What does it mean?
- does channel size encode attribute?
  - yes? sizes differ
- according to dog name in data

Analysis marks
- what type of mark?
  - line
  - no, not length coded
  - point mark with rectangular shape?
  - 2025 yd
  - 2025 yd
  - cannot change position / size / orientation
  - area?
  - 2020 no, area/shapes do not convey meaning
  - 2025 yd
  - fully interlocking
    - position, size, shape, orientation all locked

Interlocking marks: Circle packings
- also are interlocking marks, not size-coded point marks
  - more like treemap than Cartesian
  - channel availability analysis: customized circle packing
    - occupied channels
      - horizontal position encodes tax rate
      - color: rate, redundant with horizontal position
      - size (2D area) market cap
    - Free channels
      - Unavailable channels
        - vertical position used by algorithm to avoid overlap & minimize gaps
        - shape & orientation equal and unavoidable can't just change, would need to redo layout

Interlocking marks: Non-spatial
- also are interlocking marks
  - example with non-spatial data?
  - treemaps
    - show hierarchy with containment, not connection
    - encode additional attributes with area/size
    - again, cannot change just one mark alone
    - could recompute layout to change all at once
    - combined layout of all marks together
      - carries meaning
      - unlike spatial data mark boundaries
      - individual mark boundaries lose intrinsic meaning

Interlocking marks: Tile heatmaps
- 2D matrix/grid as index
  - position in use as index
  - size/area & shape & orientation all equal (locked down)
- simplest possible case of interlocking marks!
  - more regular than choropleths or treemaps
  - but underdetermined similarities
  - full extent of cell used for color coding...different from using a point mark within the cell

Interlocking marks: Circle packings
- customized circle packings are special case
  - including beeswarm plots
  - general circle packing
    - algorithmic constraint: no overlaps, minimal gaps
    - position unavailable since used by algorithm
  - Doring cartogram
    - can treat as special case of circle packing, with additional constraints to maintain relative position from geographic location
    - throw away shape by regularizing to circles
    - add size coding

Quiz: Name that mark
- UFC fights: points? lines? areas?

Analyzing marks
- what type of mark?
- line
- no, not length coded
- point mark with rectangular shape?
- 2025 yd
- 2025 yd
- cannot change position / size / orientation
- area?
- 2020 no, area/shapes do not convey meaning
- 2025 yd
- fully interlocking
- position, size, shape, orientation all locked

Interlocking marks: Circle packings
- yes interlocking
- A: already covered
- B/C: equal-area algorithms
- E/F: multi-level
- top-level interlocking marks
- bottom-level squares units
- L: fanout for more
- 9. subsets: position-specific designing
- no point marks
- size: coded by area

Line marks: Rethinking
- do line charts use line marks?
  - many channels unavailable: size, position, shape, orientation
  - proposal: rename from "area" to "Interlocking"

Line marks: Naming two cases separately
- line segments showing single item, vs curved lines showing multiple items
  - should we reason about them separately instead of analyzing them together?
  - single line segment: express single qualitative attribute for one item with length
  - single mark represents single item of data
  - proposal call these "segments"
  - curved / complex lines
  - proposal call these "paths"
  - single mark represents many items of data

Distinguishing marks through constraints
- highly constrained interlocking marks
  - many channels unavailable: size, position, shape, orientation
  - proposal: rename from "area" to "Interlocking"

Unconstrained point marks:
- can encode more info in any channel at all!
  - size, position, shape, orientation
  - color, motion...
- does "point" imply circular shape?
- proposal is "unconstrained" a better / more concise name?

Line marks: Rethinking
- do line charts use line marks?
  - many channels unavailable: size, position, shape, orientation
  - proposal: rename from "area" to "Interlocking"

Distinguishing marks through constraints
- highly constrained interlocking marks
  - many channels unavailable: size, position, shape, orientation
  - proposal: rename from "area" to "Interlocking"

Unconstrained point marks:
- can encode more info in any channel at all!
  - size, position, shape, orientation
  - color, motion...
- does "point" imply circular shape?
- proposal is "unconstrained" a better / more concise name?

so... what about line marks?
From marks to glyphs: multiple marks/item
- glyphs: more than one mark per item
  - grouped bars
  - stacked bars
  - multiple views
  - bar chart small multiples

From marks to glyphs: multiple marks/item
- glyphs: more than one mark per item
  - grouped bars
  - stacked bars
  - multiple views
  - bar chart small multiples
More stuff

- this talk  
  http://www.cs.ubc.ca/~tmm/talks.html#northeastern24

-- more questions? thoughts on answers??

- book 
  http://www.cs.ubc.ca/~tmm/vadbook

- full courses, papers, videos, software, talks 
  http://www.cs.ubc.ca/group/infovis 
  http://www.cs.ubc.ca/~tmm

Visualization Analysis and Design. Munzner. 
@tamaramunzner 
@tamara@vis.social