Gestalt Perception

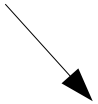
15-494 Cognitive Robotics
David S. Touretzky &
Ethan Tira-Thompson

Carnegie Mellon
Spring 2009

What's a Gestalt?

- German word meaning “whole” or “form”.
- A complete pattern or configuration.
- A configuration or pattern of elements so unified that...

  The whole is greater than the sum of its parts.



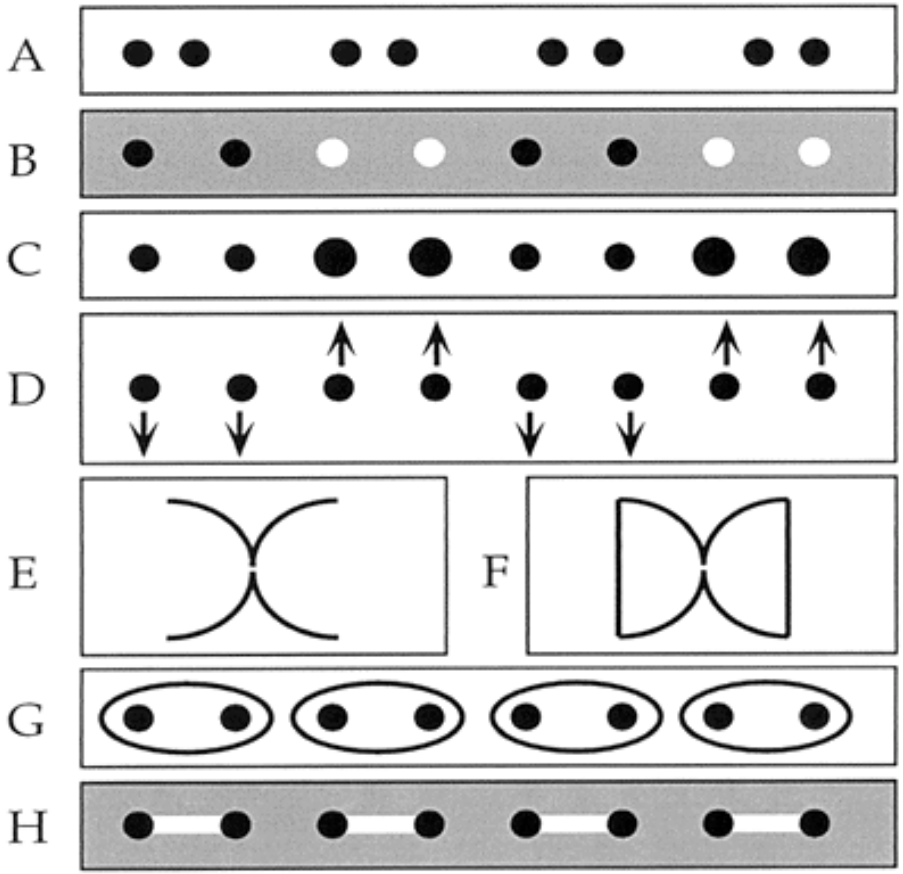
- The term “sum,” in German *”Zusammenfassung,”* means “summing up” or “synopsis”, not arithmetic sum.

Gestalt Perception

- Describes perceptual phenomena and theoretical principles governing *perceptual organization*.
  - Perceptual grouping.
  - Figure-ground organization.
  - Frames of reference.
  - *Pragnanz*: tendency toward regular, ordered, stable, balanced states.
- A holistic view of perception, rather than purely bottom-up.

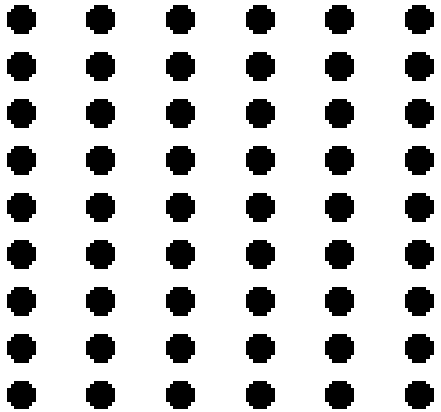
Principles of Grouping

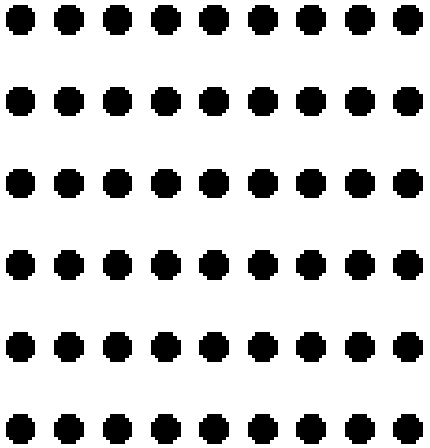
- A: proximity
- B: similar color
- C: similar size
- D: common fate
- E: good continuation
- F: closure
- G: common region
- H: element connectedness



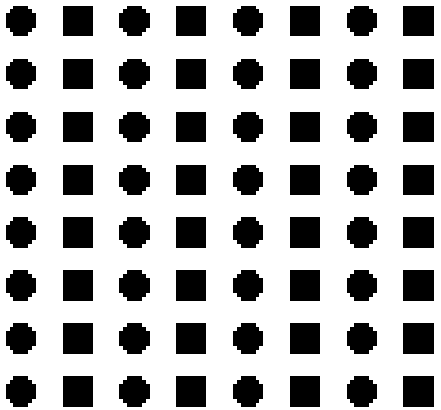
From S. Palmer, “Gestalt Perception”, MIT Encyclopedia of Cognitive Science

Proximity

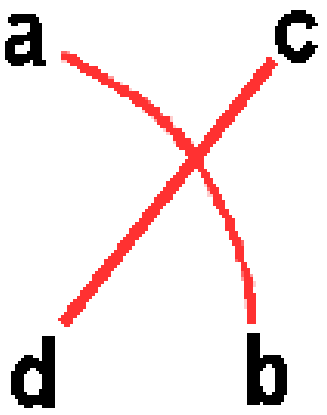


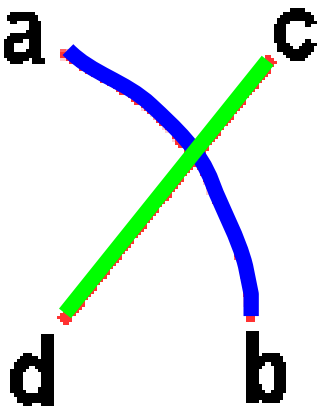


Similarity



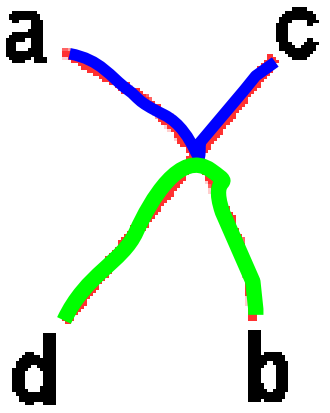
Good Continuation





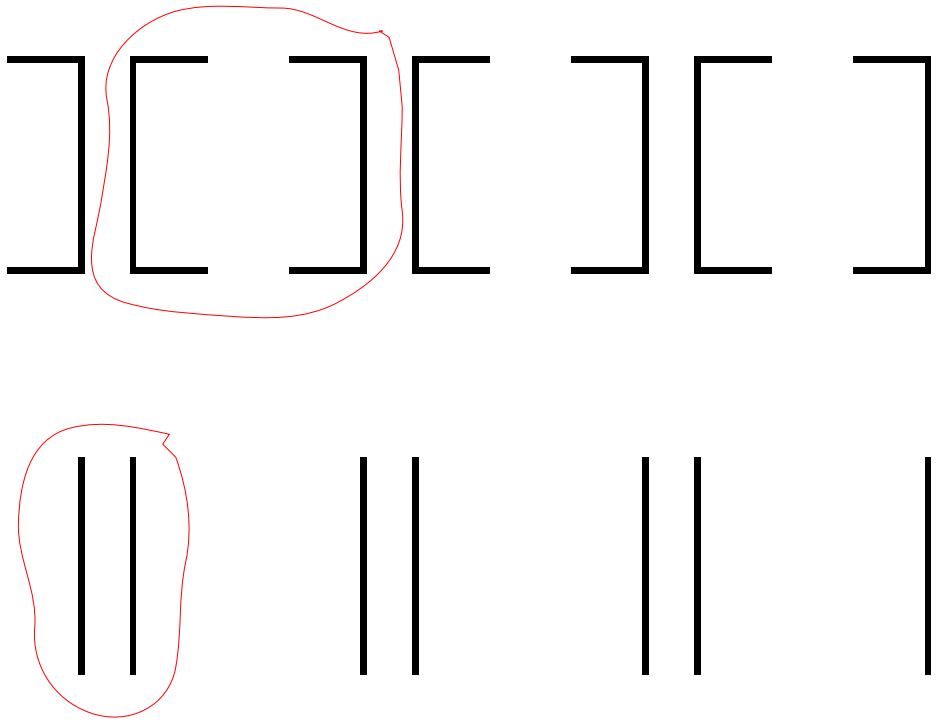
- Two smooth lines: a-b and c-d, that cross:

- Not two sharp curves a-c and b-d meeting at a point:



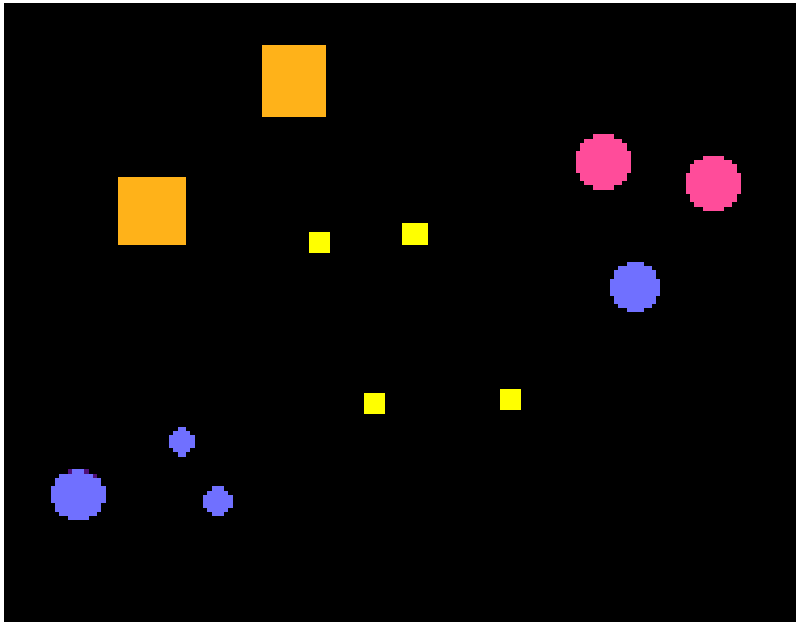
Closure

Closure and proximity can conflict:



Simulated Perceptual Grouping

- K. R. Thorisson (1994) model of perceptual grouping:

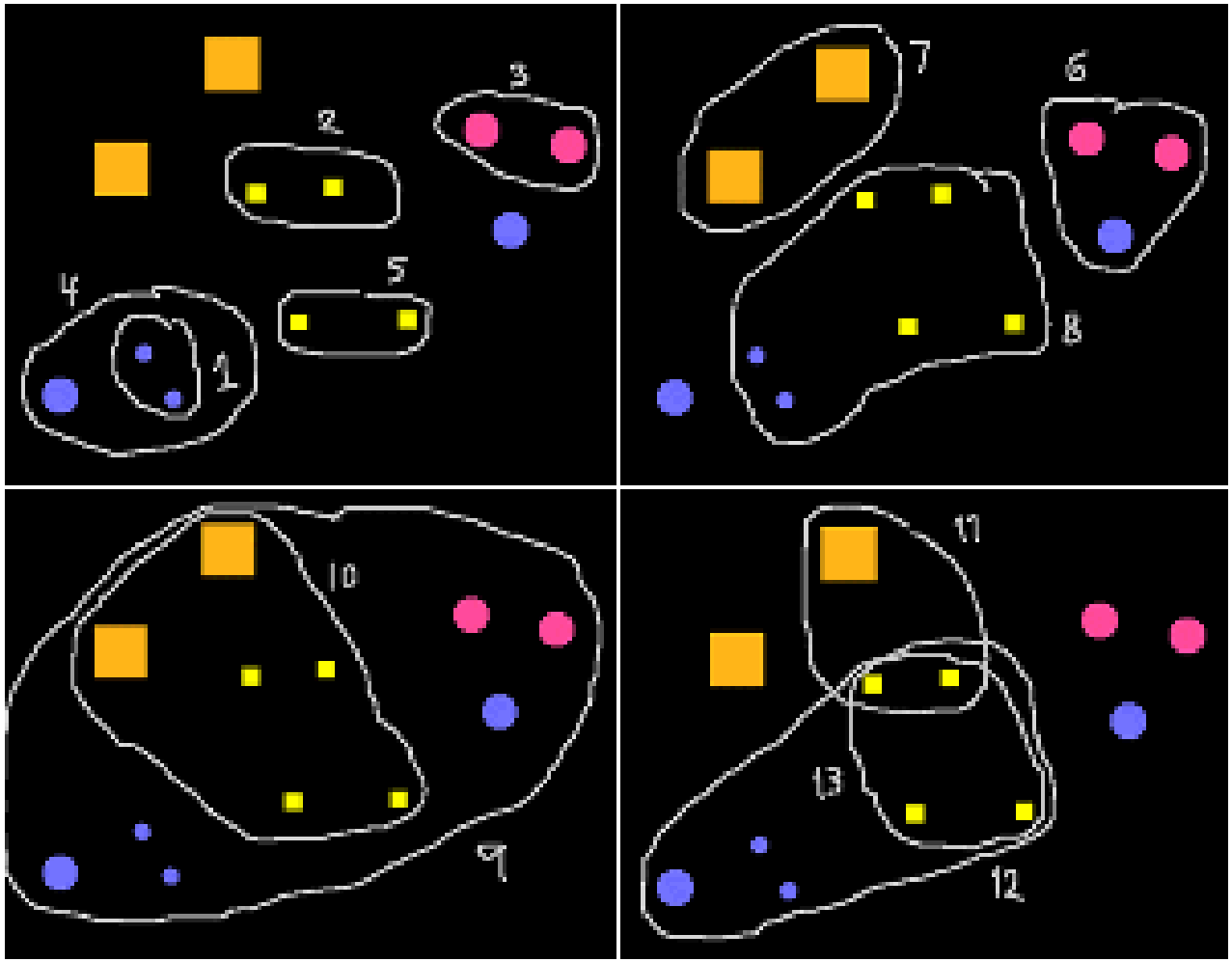


- How would you group these shapes?

- What principles apply?

Thorisson Model

Multiple runs produce different results:



Thorisson (cont.)

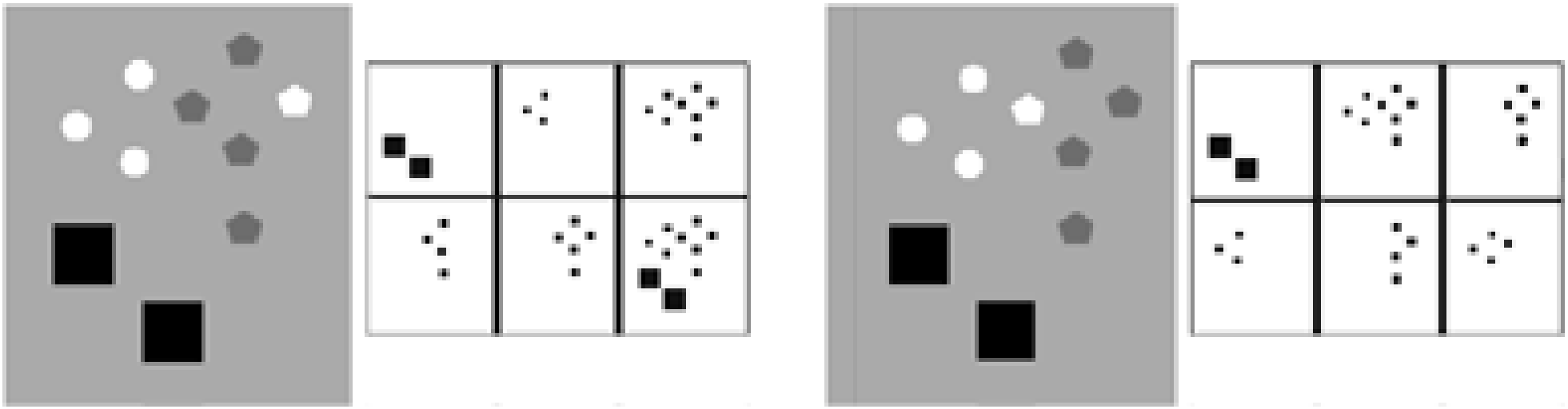
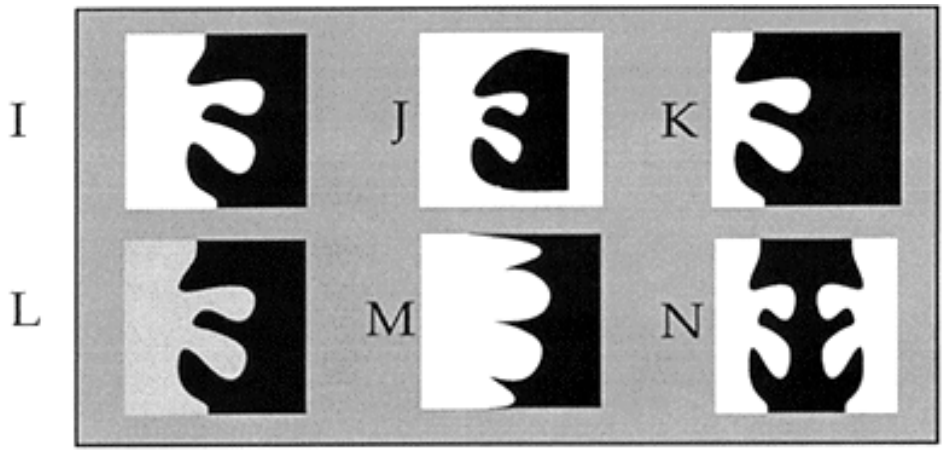


Figure-Ground Principles

- I: boundary belongs to the figure
- J: surroundedness
- K: size (figure is smaller)
- L: figure is higher contrast
- M: convexity
- N: symmetry



From S. Palmer, “Gestalt Perception”, MIT Encyclopedia of Cognitive Science

Not a quantitative theory: no way to predict how these principles combine or interact.

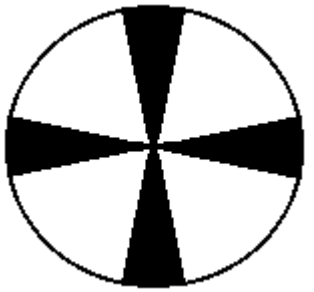
Which Part is the Figure?

- Smaller
- Higher contrast
- Surrounded

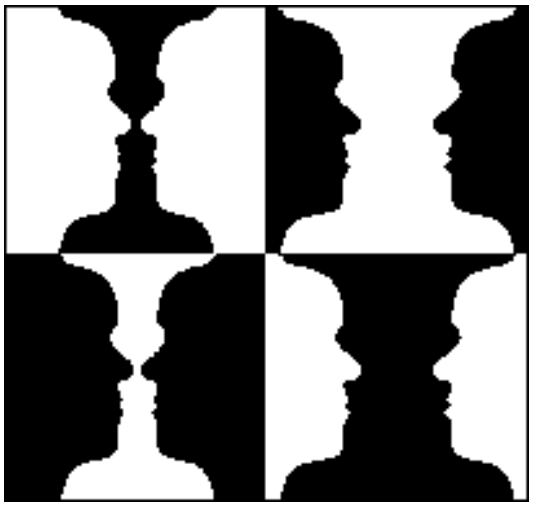
- Contrast

Smallness

- Figure tends to be smaller than ground.

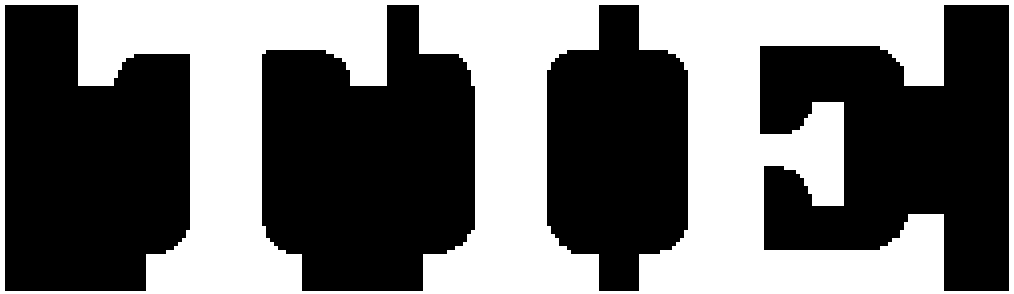


- The Rubin Vase:
  Vase is favored on the left; faces on the right.



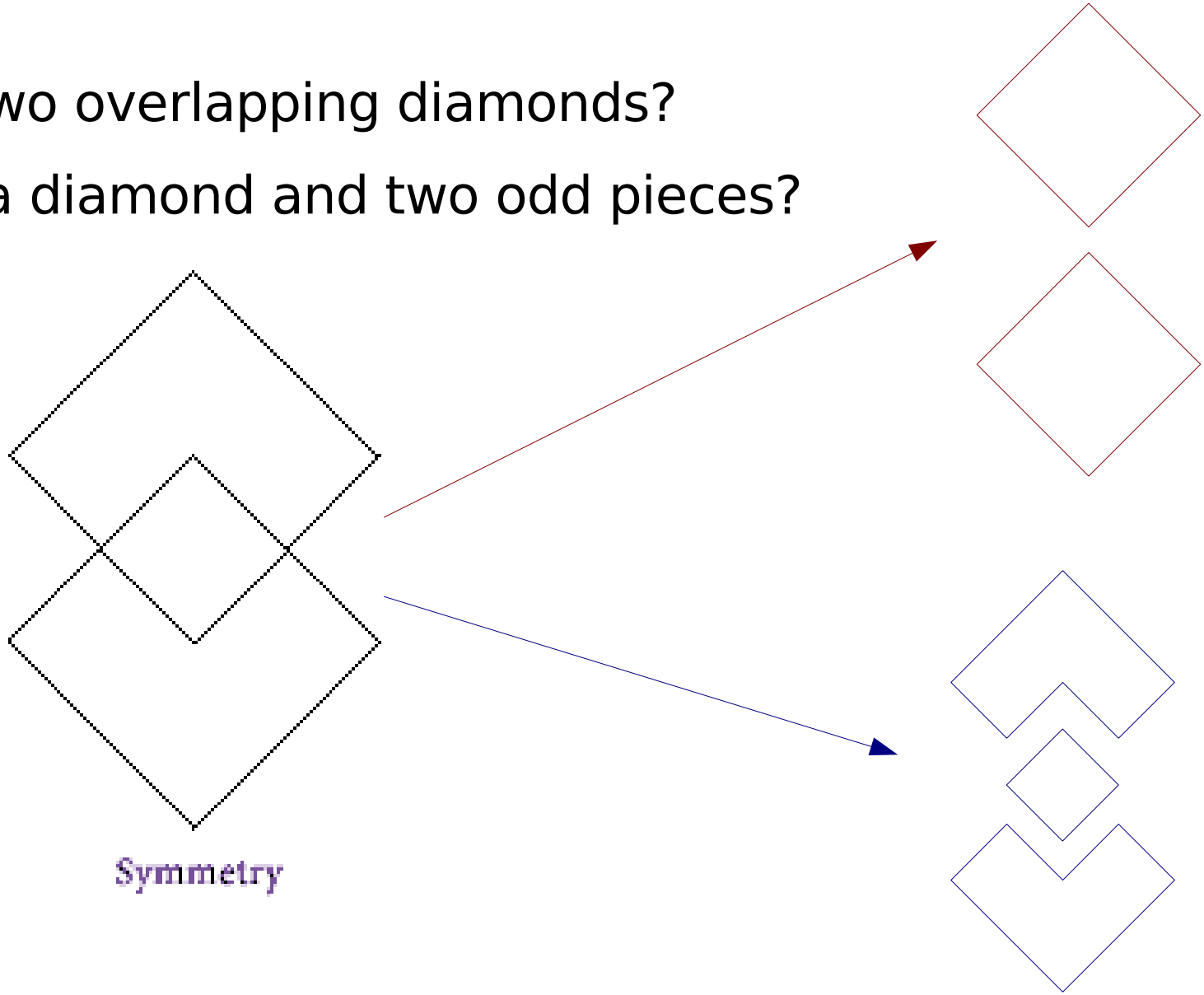
Surroundedness

- Areas surrounded by others tend to be seen as figure.
- This can be misleading. What do you see below?



Symmetry

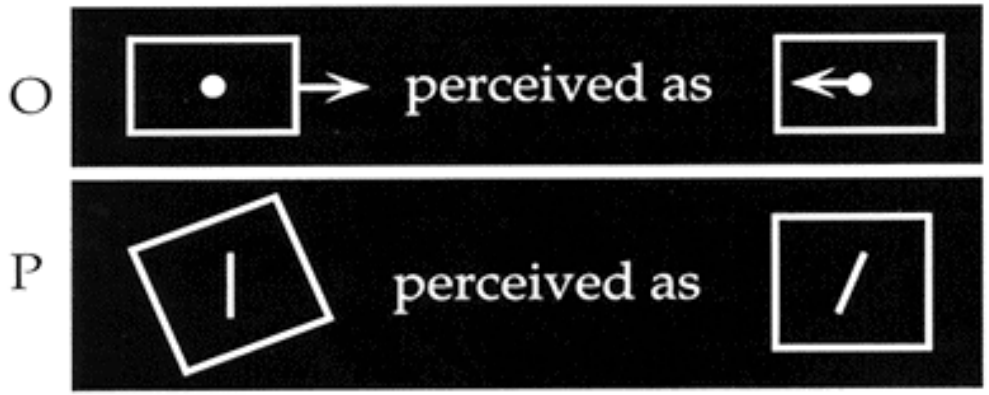
- Is this two overlapping diamonds?
- Or is it a diamond and two odd pieces?



Frame of Reference

- O: induced motion (figure moves relative to reference frame, not vice versa)

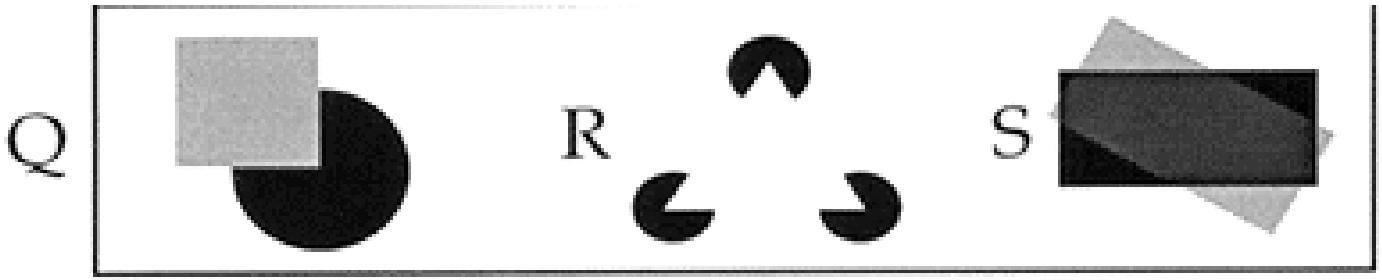
- P: rod-and-frame effect



From S. Palmer, “Gestalt Perception”, MIT Encyclopedia of Cognitive Science

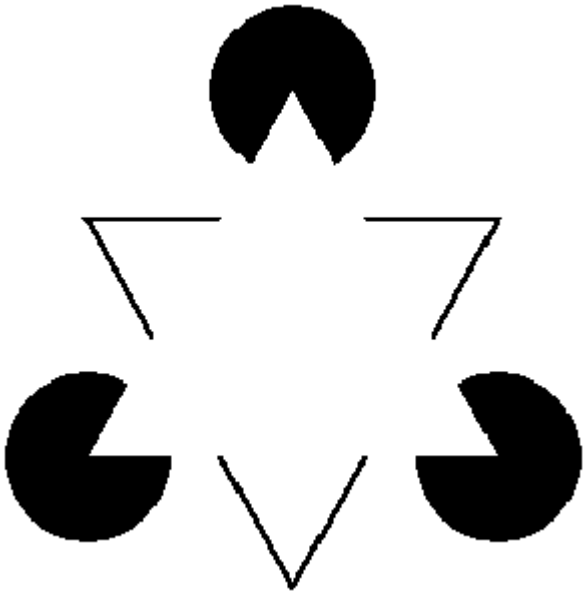
Other Organizational Phenomena

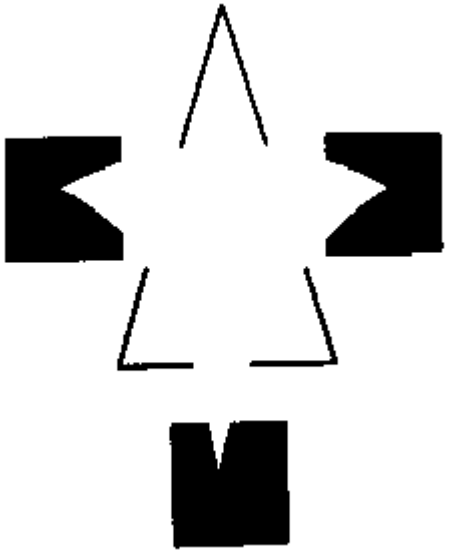
- Q: amodal completion.
- R: illusory contours
- S: color scission



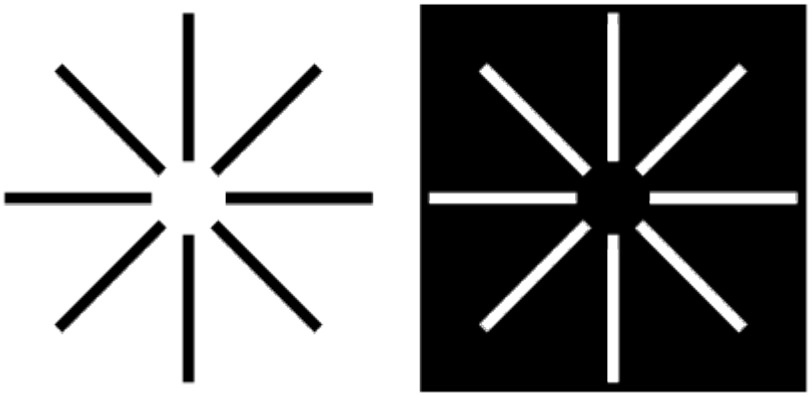
From S. Palmer, “Gestalt Perception”, MIT Encyclopedia of Cognitive Science

Kanizsa Triangle: Illusory Contours

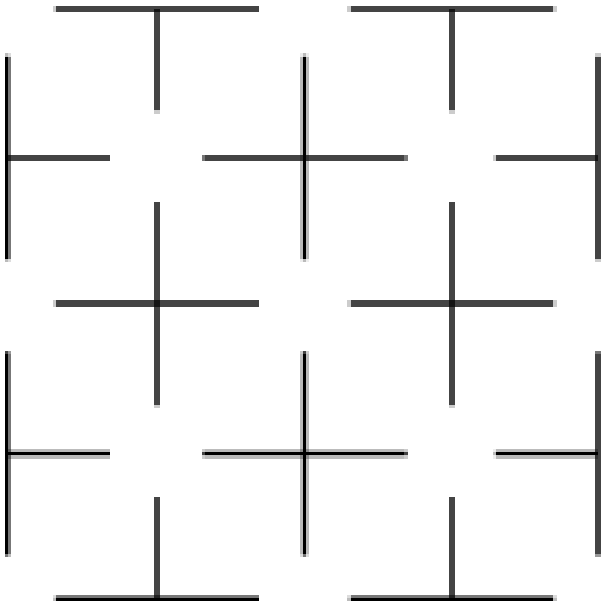




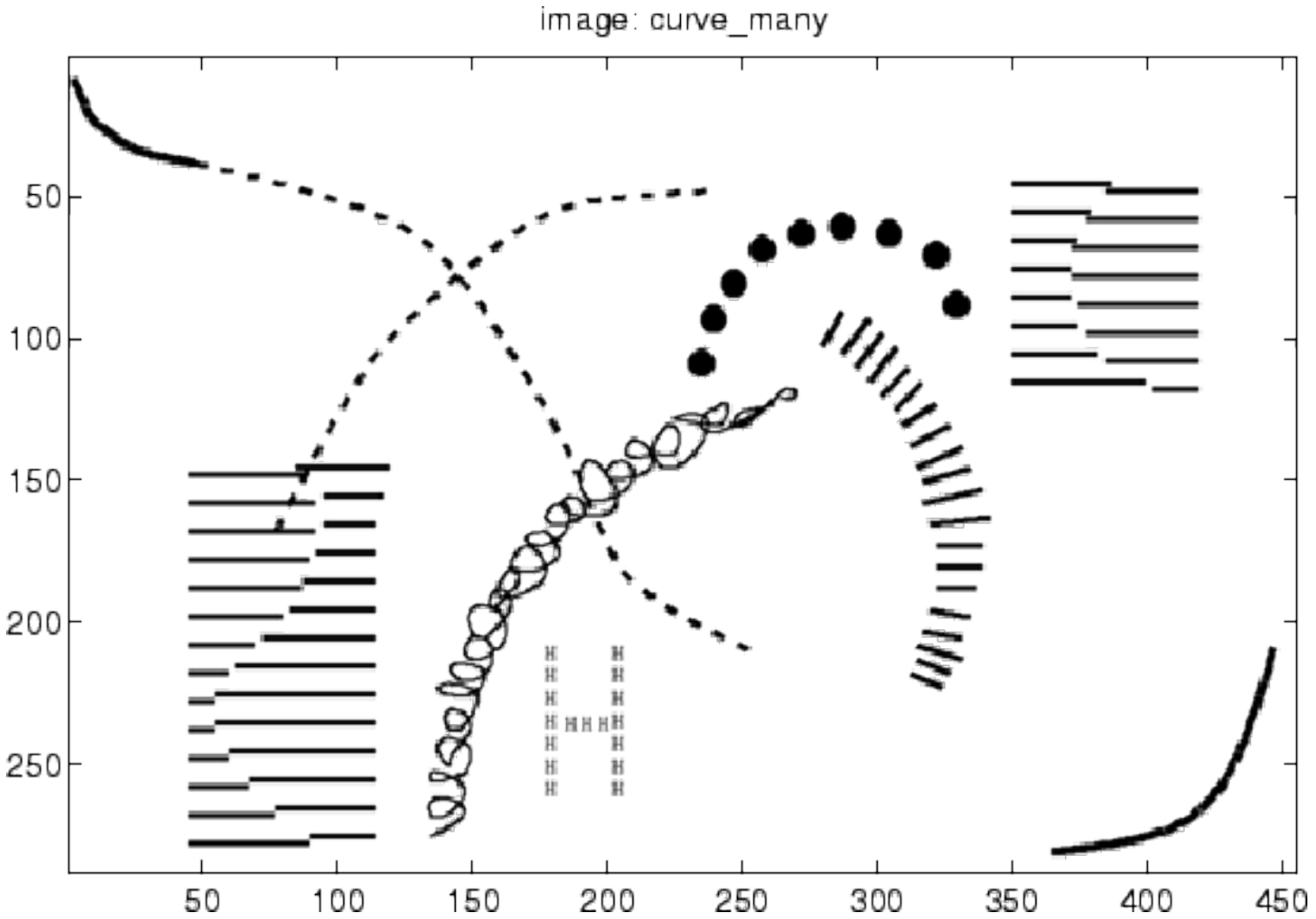
Illusory Contours



Ehrenstein Effect



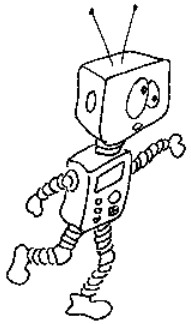
Simple Curves

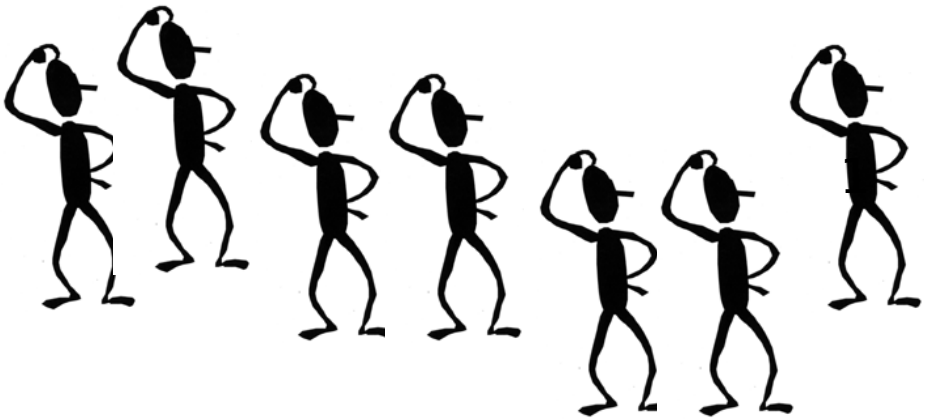


From L. Brassard, “The Visual World”

Implications for Robotics

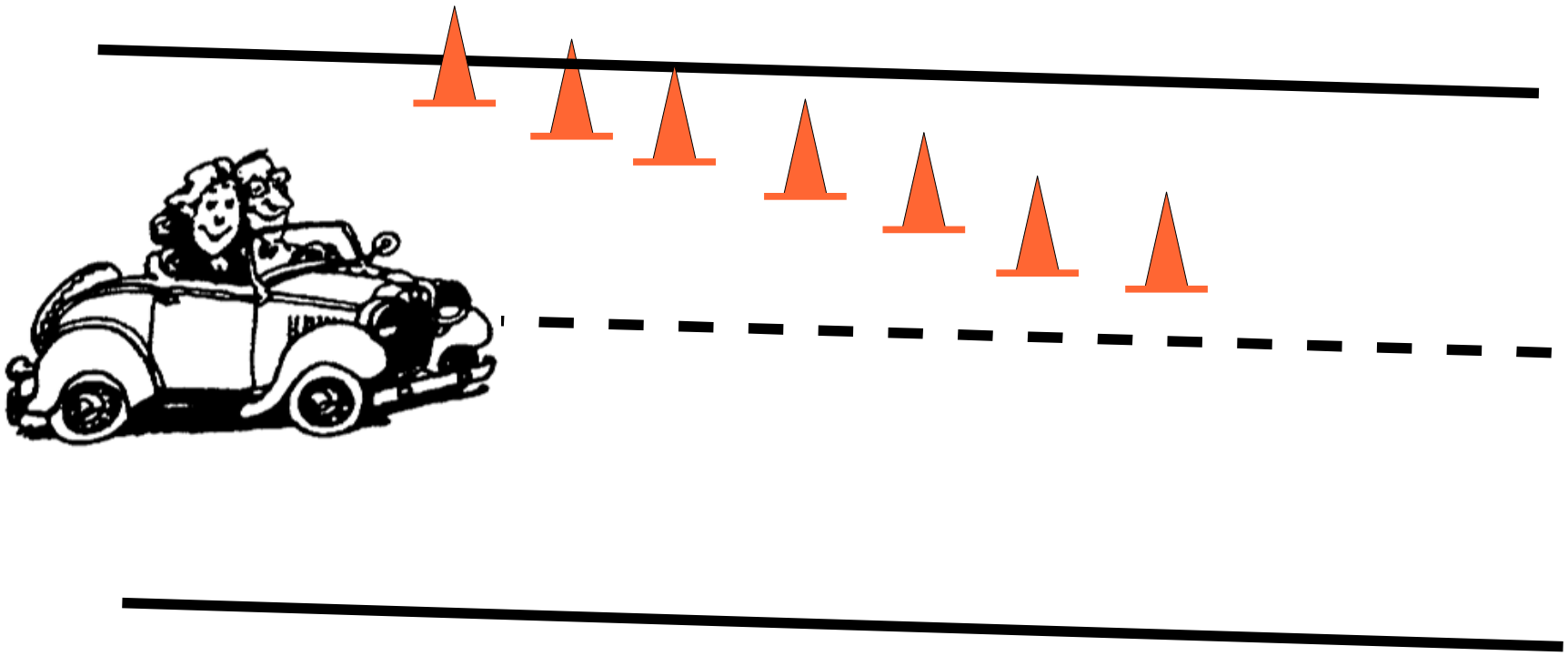
- Robots should “see” the world the way people do.
- Recognize spatial relationships such as boundaries, contours, groups of things.
- How do you “stand in line”?





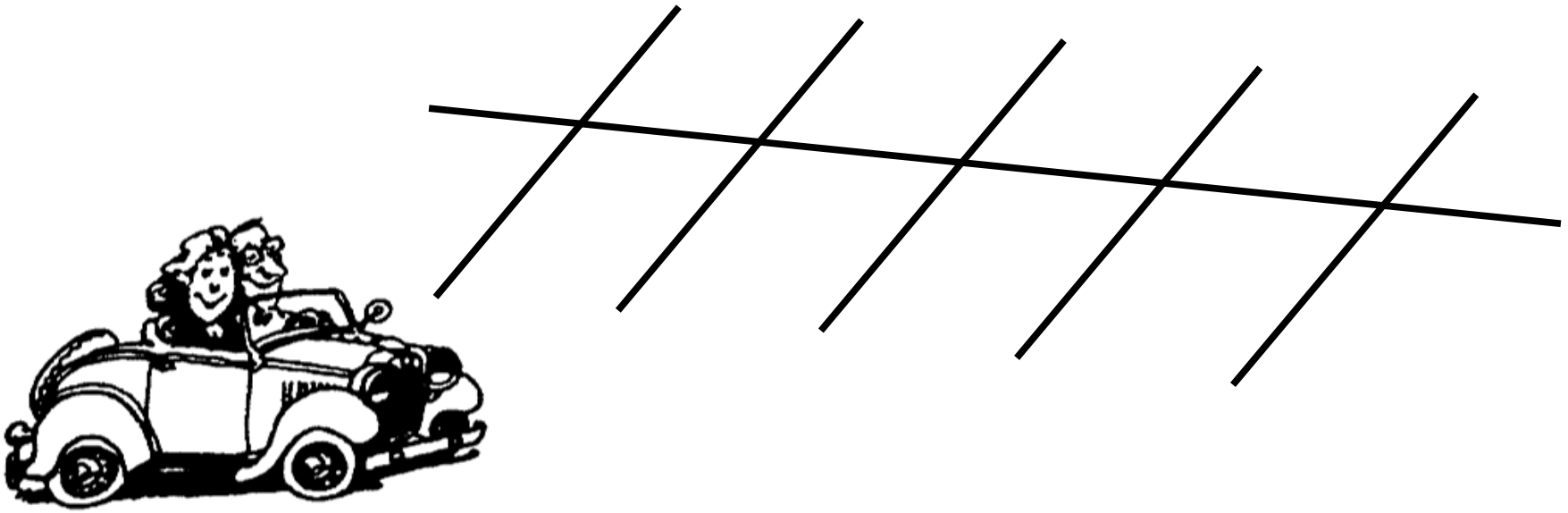
Implications for Robotics

- What's going on here?



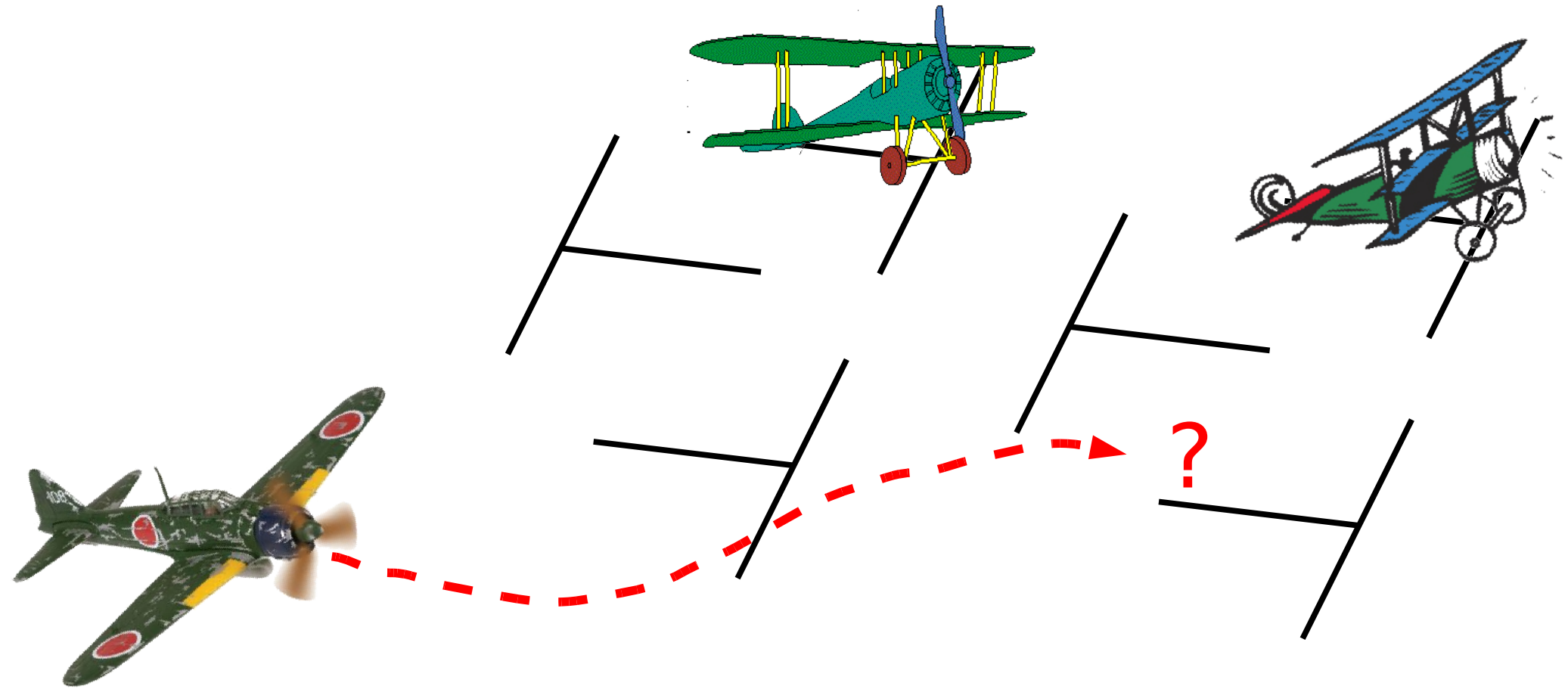
Implications for Robotics

- Where do you park the car?



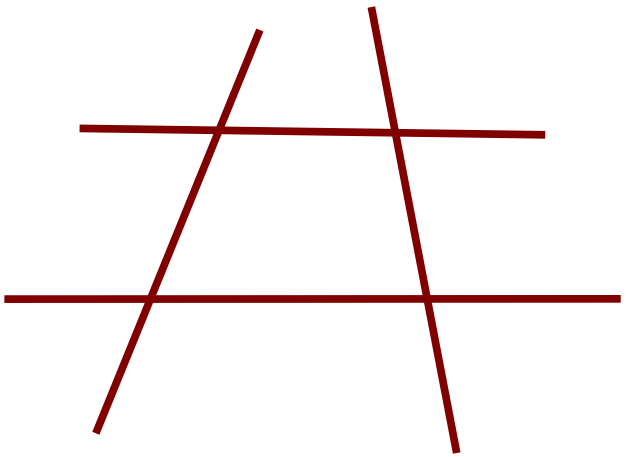
Implications for Robotics

- Where do you park the plane?



Why Was Parsing the Tic-Tac-Toe Board Difficult?

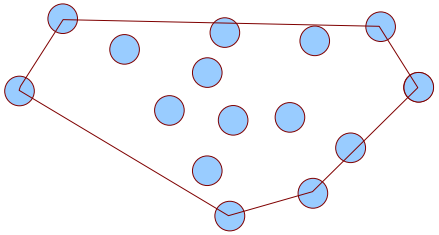
- Grid of lines defined regions that weren't fully enclosed.



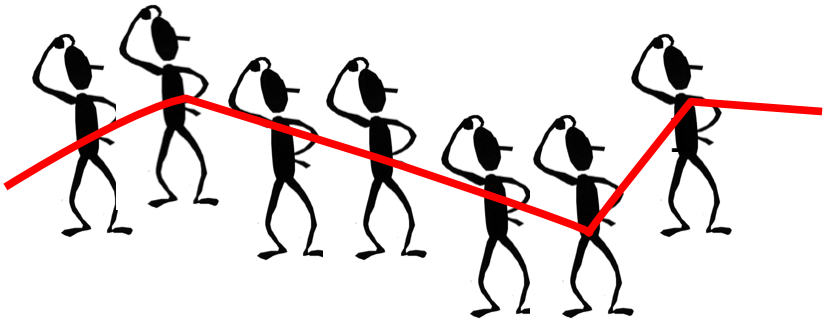
- We had to induce our own “completion” effect to extract the board regions.
- Working in camera space forced us to deal with perspective effects.

Can More Primitives Help?

- Convex hull for defining regions:



- Boundary detection and extrapolation:



Convex Hull in Tekkotsu

- The convex hull operator is a crude approximation to gestalt completion, but only for convex shapes.
- Defined in DualCoding::PolygonData:

```
Shape<PolygonData> convexHull(const Sketch<bool> &sketch)
```

- A polygon is a collection of edges and vertices (dual representation for efficiency).

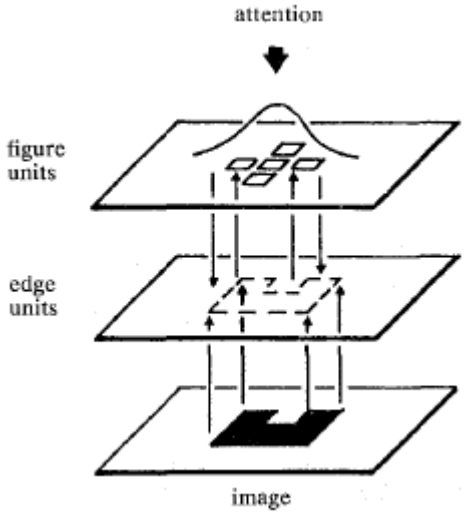
- Why isn't convex hull sufficient?
  - Some shapes aren't convex
  - Not a good representation of curved boundaries

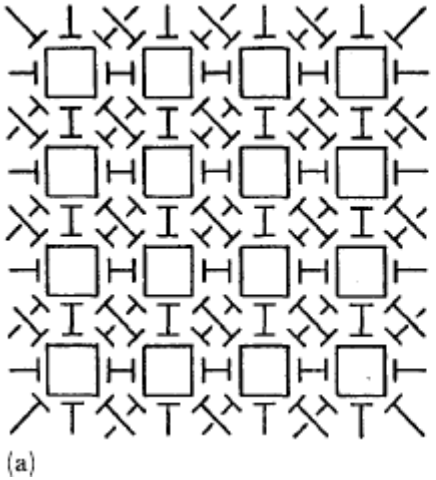
Pragnanz ("Good Form") Revisited

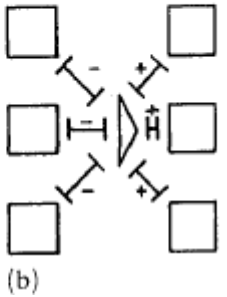
- Early gestalt theorists (1930s)` suggested that holistic perception came from interacting electrical fields in the brain.
  - This has since been disproved.
- More recently, recurrent neural networks such as the Hopfield network or Boltzmann machine have reproduced some aspects of holistic perception.
  - Perceptual principles act as sources of constraint on states of the network.
  - Network searches for "minimum energy" states that optimally satisfy the constraints.

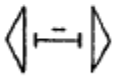
Kienker et al.(1986)

- Boltzmann machine doing figure-ground separation.
- Edge units (triangles) support some neighboring figure units (squares) and inhibit others.
- Opposite-orientation edges inhibit each other.



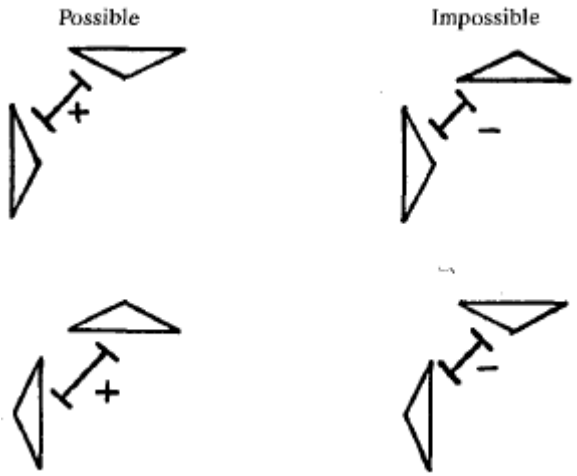






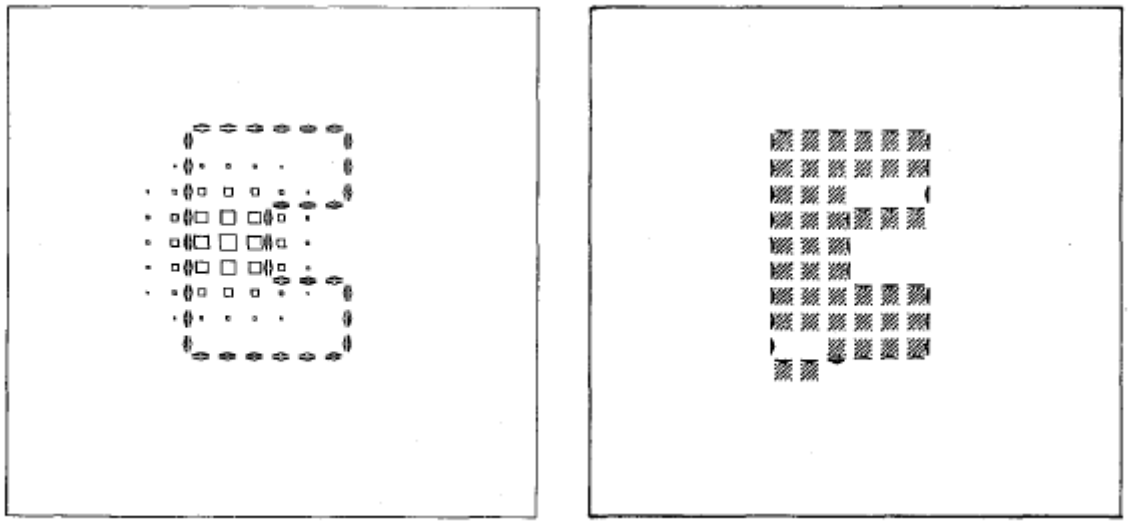
Edge Constraints

- Activation propagates among edge units.
- These local constraints will combine to produce a globally coherent solution.



Kienker et al.

- Start with an outline.
- Network probabilistically settles into a state in which figure is marked and ground is unmarked.



(a) (b)

Spiral Example