Bridging the gap between vision and language:

A morphodynamical model of spatial cognitive categories

René Doursat

Brain Computation Laboratory
Department of Computer Science
University of Nevada, Reno

Jean Petitot

CREA
Ecole Polytechnique, Paris

A Morphodynamical Model of Spatial Cognitive Categories

1. **Spatial categorization**
2. **Cellular automaton model**
3. **Spiking neural model**
4. **Discussion**

A Morphodynamical Model of Spatial Cognitive Categories

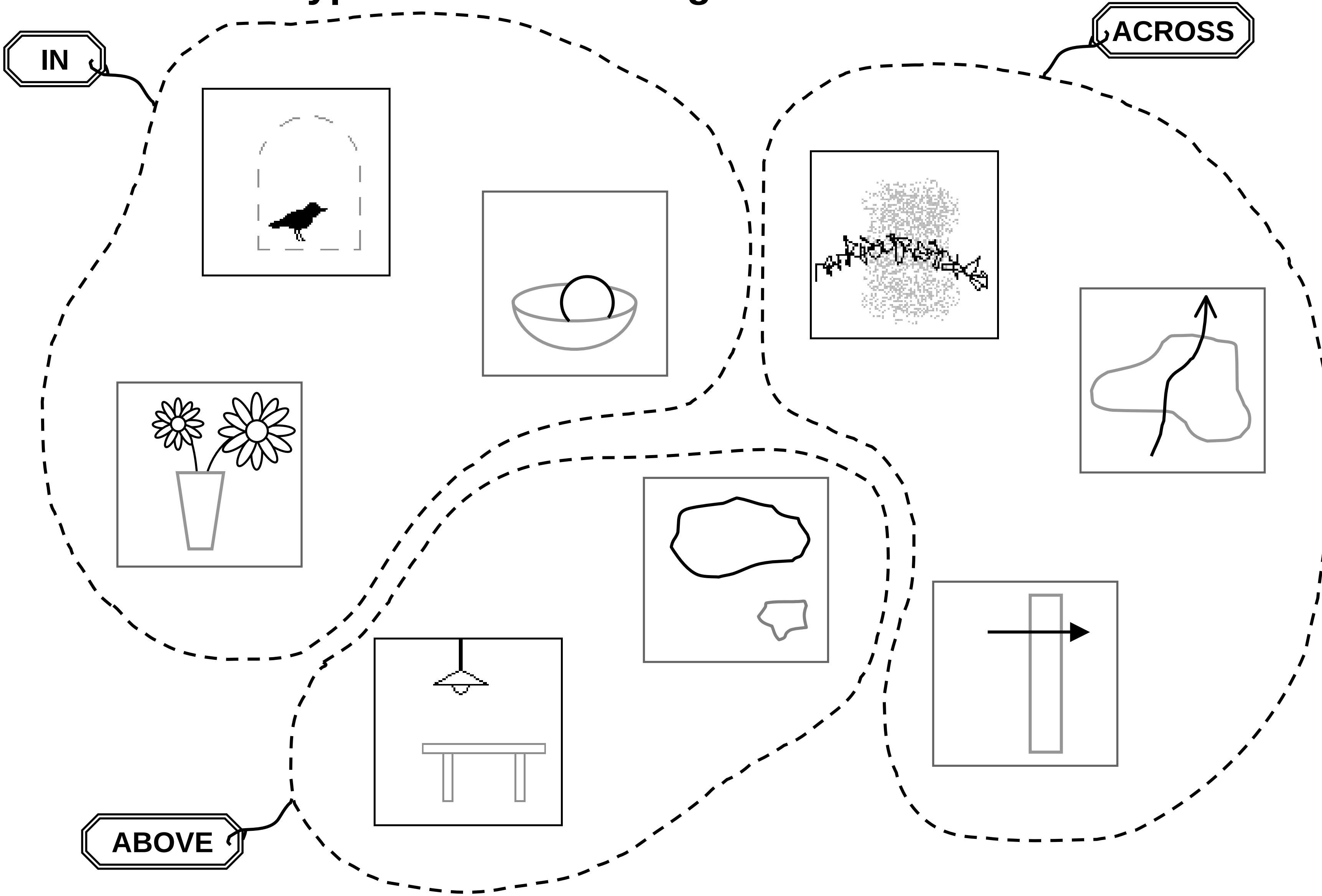
1. **Spatial categorization**
    - Object vs. scene categorization
    - Breaking up the categorical landscapes into protosemantic islands
    - Cognitive linguistics' collection of topological invariants
    - What is the "tolopogy of language"?
2. **Cellular automaton model**
3. **Spiking neural model**
4. **Discussion**

Object vs. scene categorization

Prototypes of object shapes are relatively “rigid”

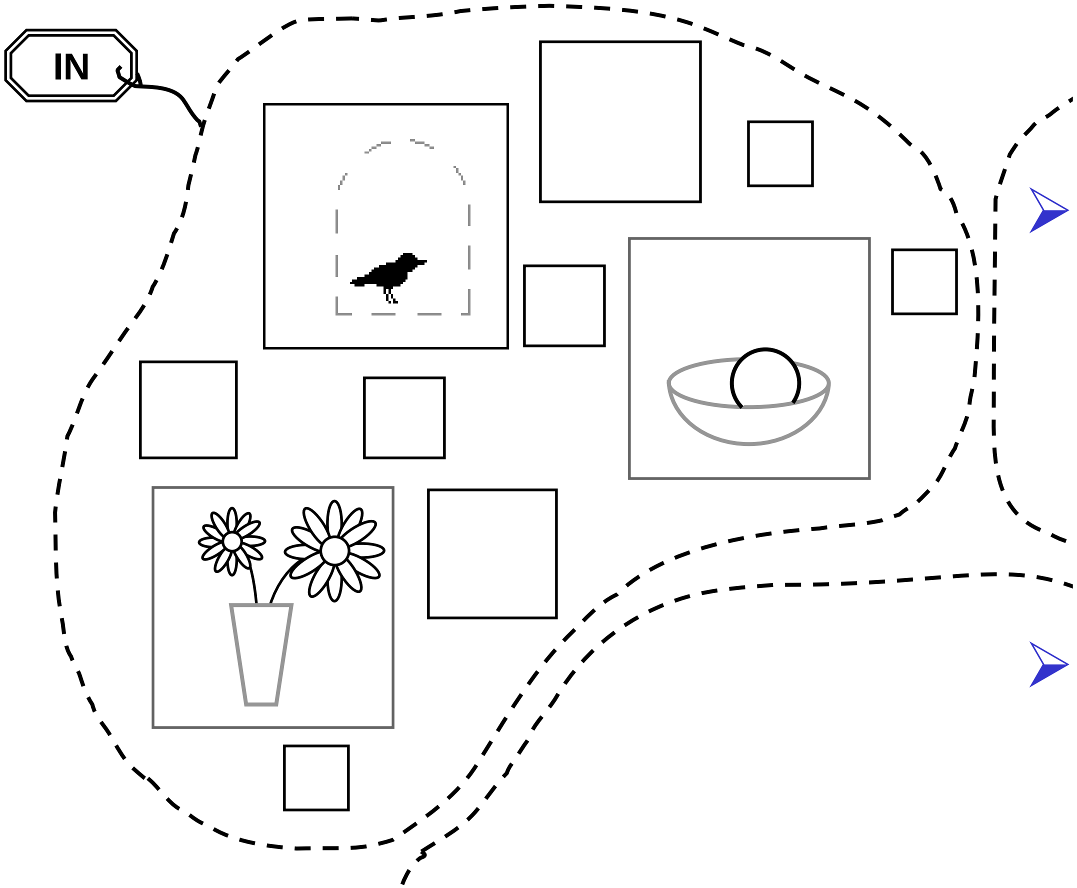
Object vs. scene categorization

Prototypes of scene configurations are "flexible"



Object vs. scene categorization

Prototypes of scene configurations are "flexible"



- *How can the infinite diversity of scenes be **categorized** under just a few linguistic elements?*

- *Equivalently, how can a single linguistic element encompass such a wide topological variety?*

Breaking up the categorical landscapes

The structure of one complex category: ‘in’

IN

(1) (a) the cat **in** the house

(b) the bird **in** the garden

(c) the flowers **in** the vase

(d) the bird **in** the tree

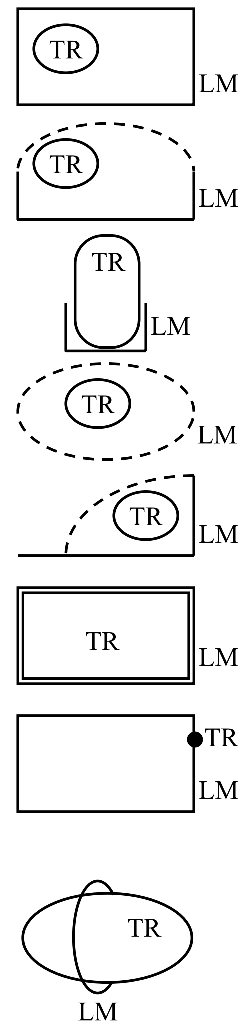
(e) the chair **in** the corner

(f) the water **in** the vase

(g) the crack **in** the vase

(h) the foot **in** the stirrup

(i) ?the finger **in** the ring



prototype

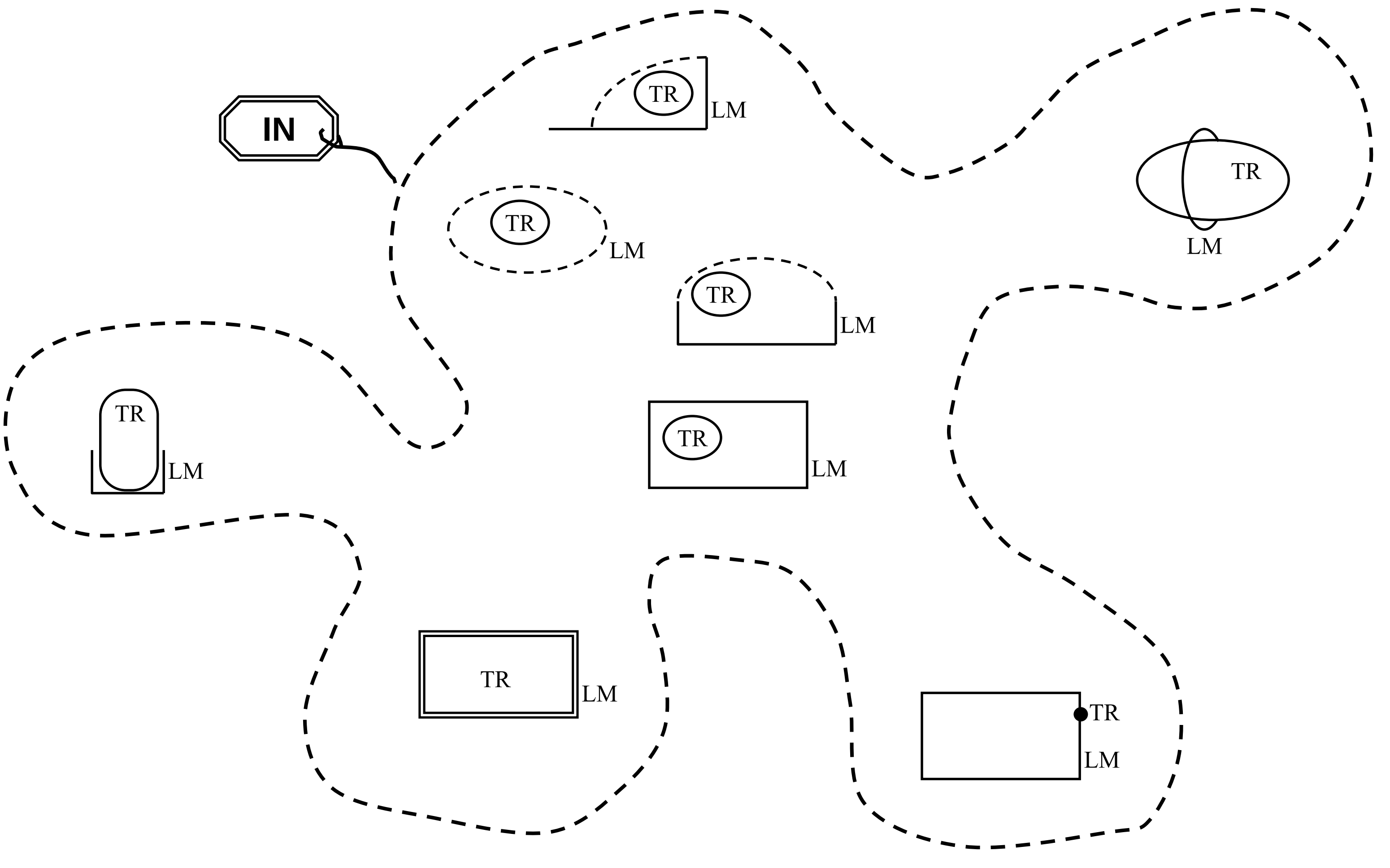
TR metonymy: flowers = stems

LM metonymy: vase = surface of vase

adapted from Herskovits (1986)

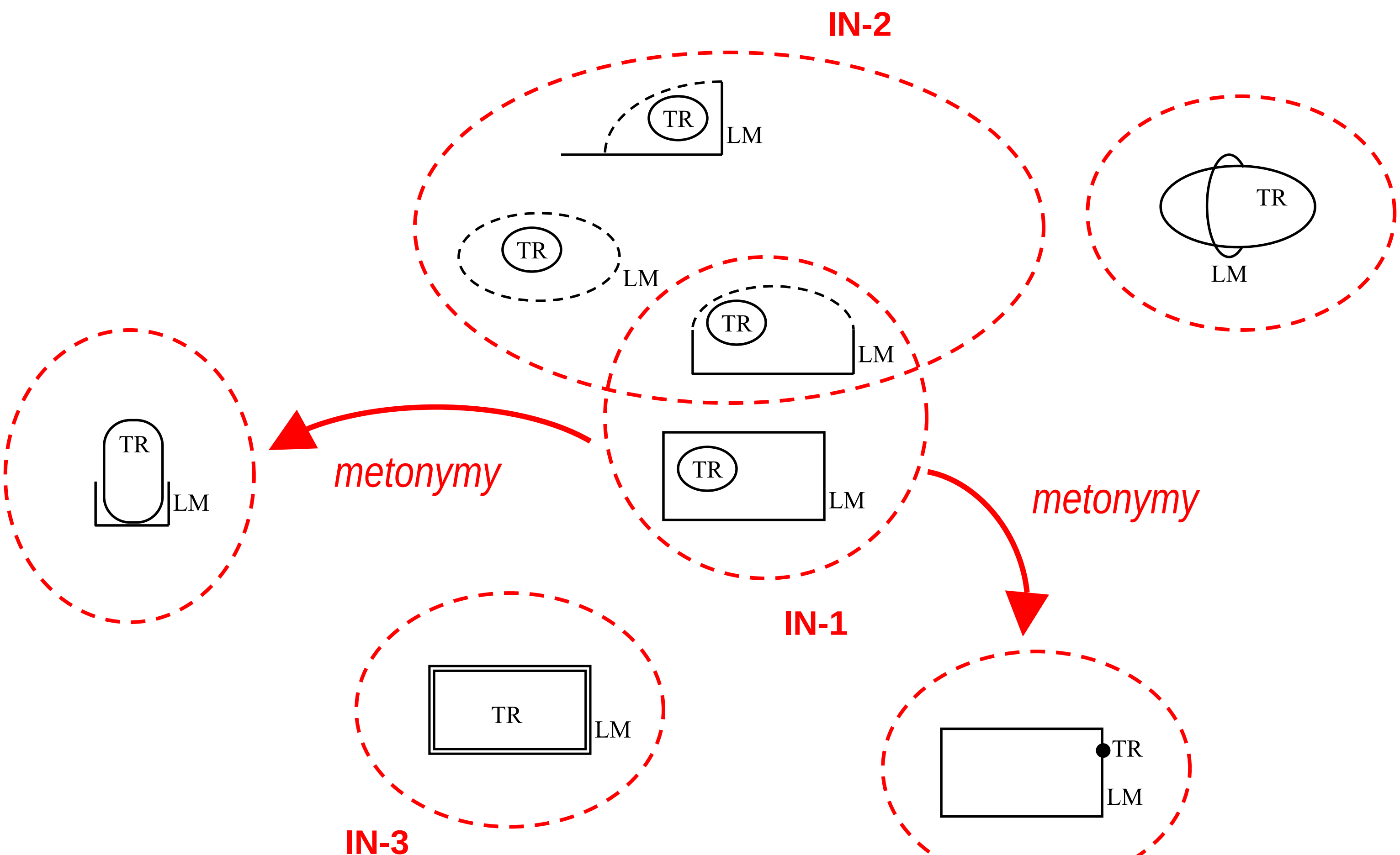
Breaking up the categorical landscapes

Prototype-based, radial category



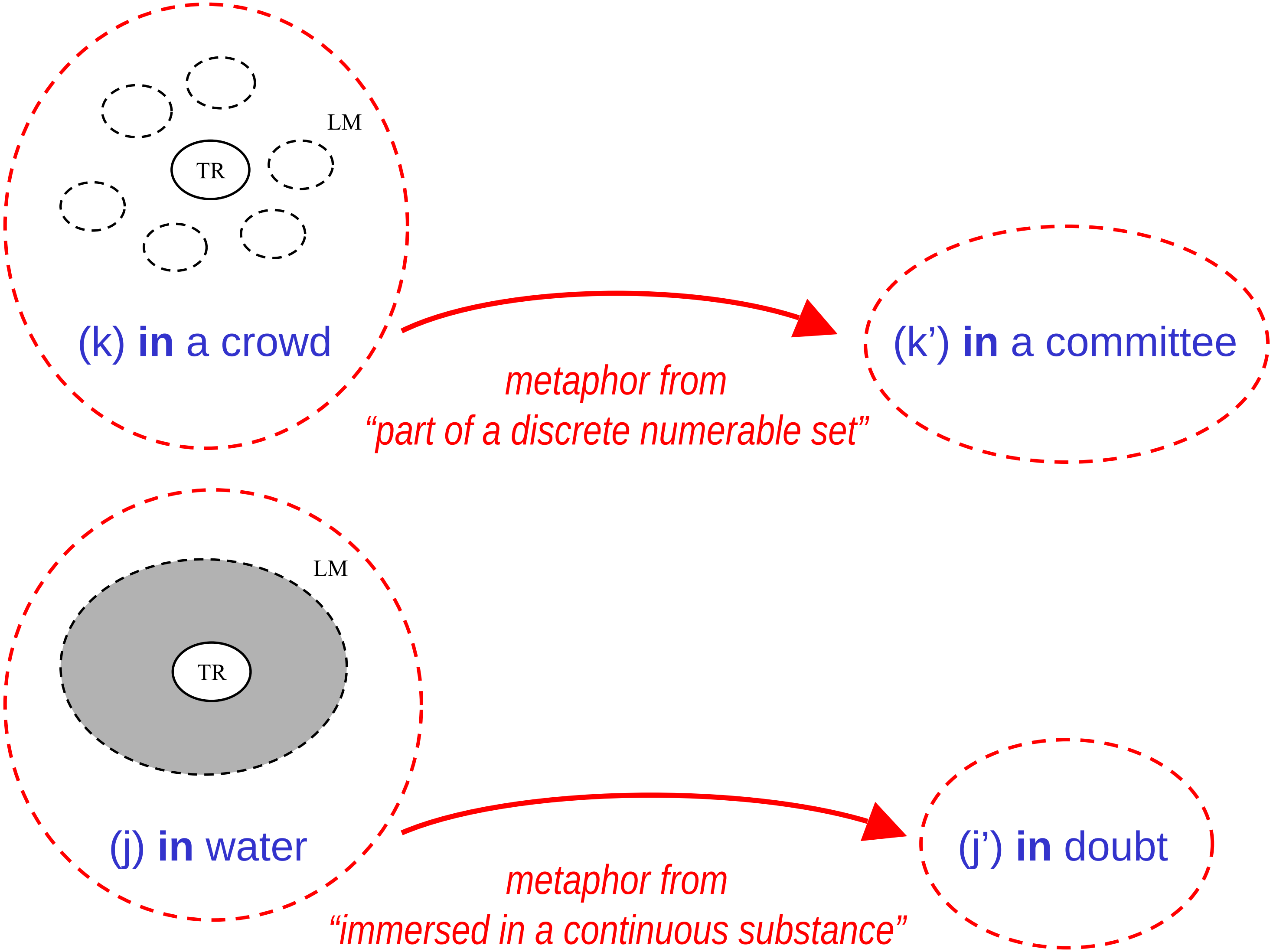
Breaking up the categorical landscapes

Protosemantic islands (with bridges)



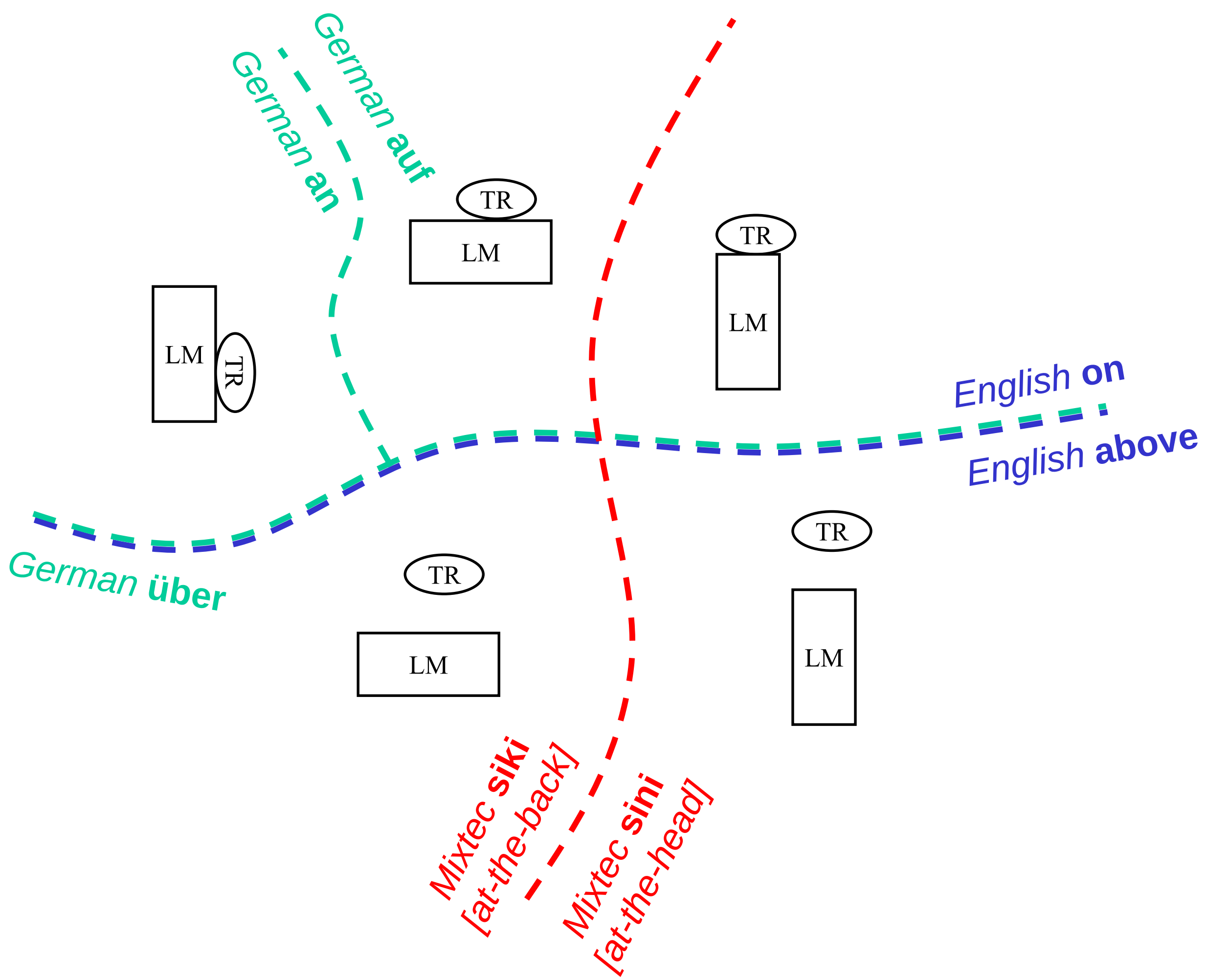
Breaking up the categorical landscapes

Further extensions by metaphorical mapping



Breaking up the categorical landscapes

More protosemantic segmentation: cross-linguistic variations



adapted from Regier (1986)

Breaking up the categorical landscapes

Summary

- *a semantic category is a cluster of* ***protosemantic*** *subcategories*

  \+ *metonymic effects*

  \+ *metaphorical mappings*

  \+ *categories do not overlap across languages*

- *we restrict our study to protosemantics: there is no unique classification criterion covering IN-1, IN-2, etc.*

- *. . . however, even focusing on a single protosemantic category, we are still facing a huge topological diversity*

Cognitive linguistics

Principles

- *what is central to language is **meaning**, not syntax*
- *but meaning is not about logical truth conditions*
- *meaning is **construals**, conceptualization, mental representations, schematization, categorization*
- *there is a common level of representation where language, perception and action become compatible*
- *language is not an autonomous functional set of syntactic rules that create meaning as a by-product*
- *syntax, semantics and pragmatics are not independent*

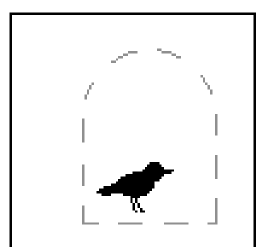
Filmore. Talmy, Langacker, Lakoff, . . .

Cognitive linguistics

Gestalt & mereology

- *traditional logical atomism (set theory): “things” are already individuated symbols and “relations” are abstract links connecting these symbols*

- *by contrast, in the Gestaltist or mereological conception, things and relations constitute analogic wholes: relations are not taken for granted but emerge together with the objects through segmentation and transformation*



Cognitive linguistics

Properties of construals

- *cognitive linguistics identifies semantic construals to abstract iconic scenes ("theater stage")*

- *one can view construals from different angles and study their properties:*

  - *figure (TR) and ground (LM)*
  - *perspective / viewpoint*
  - *profiling / salience*
  - *frames / context*
  - *etc.*

Cognitive linguistics

Collection of invariants

- *bulk invariance*

  (3) (a) The caterpillar crawled up along the filament.
  (b) The caterpillar crawled up along the flagpole.
  (c) The caterpillar crawled up along the redwood tree.
  $\rightarrow$ *'along' is insensitive to the girth of LM*

- *continuity invariance*

  (4) (a) The ball is in the box.
  (b) The fruit is in the bowl.
  (c) The bird is in the cage.
  $\rightarrow$ *'in' is insensitive to discontinuities in LM*

- *shape invariance*

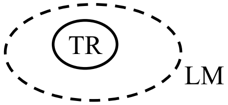
  (5) (a) I zigzagged through the woods.
  (b) I circled through the woods.
  (c) I dashed through the woods.
  $\rightarrow$ *'through' is insensitive to the shape of TR's trajectory*

adapted from Talmy

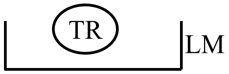
What is the "topology of language"?

- *language topology (LT) it is not the same as mathematical topology (MT)*
- *LT is sometimes **less** constrained than MT, as with the various examples of 'IN':*

closed container

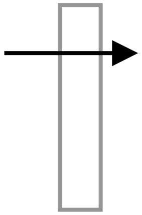


leaky container

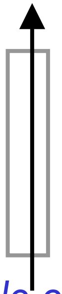


open container

- *LT is sometimes **more** constrained than MT, as with the metric ratios of 'ACROSS':*



good example of ACROSS

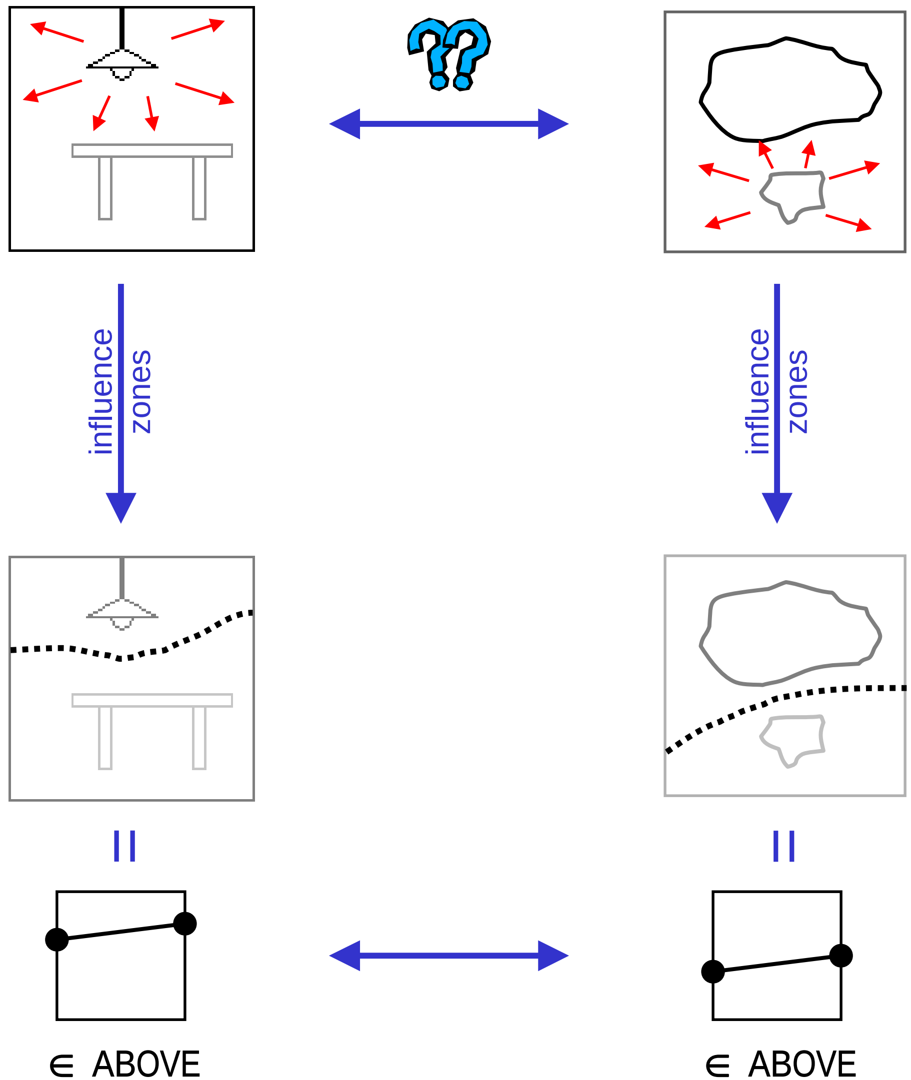


bad example of ACROSS

A Morphodynamical Model of Spatial Cognitive Categories

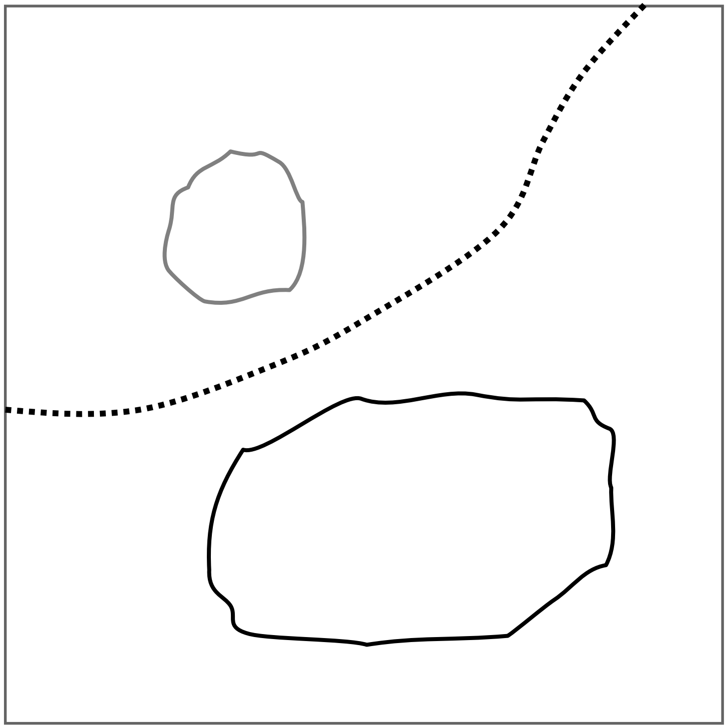
1. Spatial categorization
2. **Cellular automaton model**
   - Key to invariance: drastic morphological transforms
   - Perceptual-semantic classifier
   - Objects (a) expand and (b) collide
   - Singularities reveal the characteristic “signature” of the scene
3. Spiking neural model
4. Discussion

Key to invariance:
Drastic morphological transforms



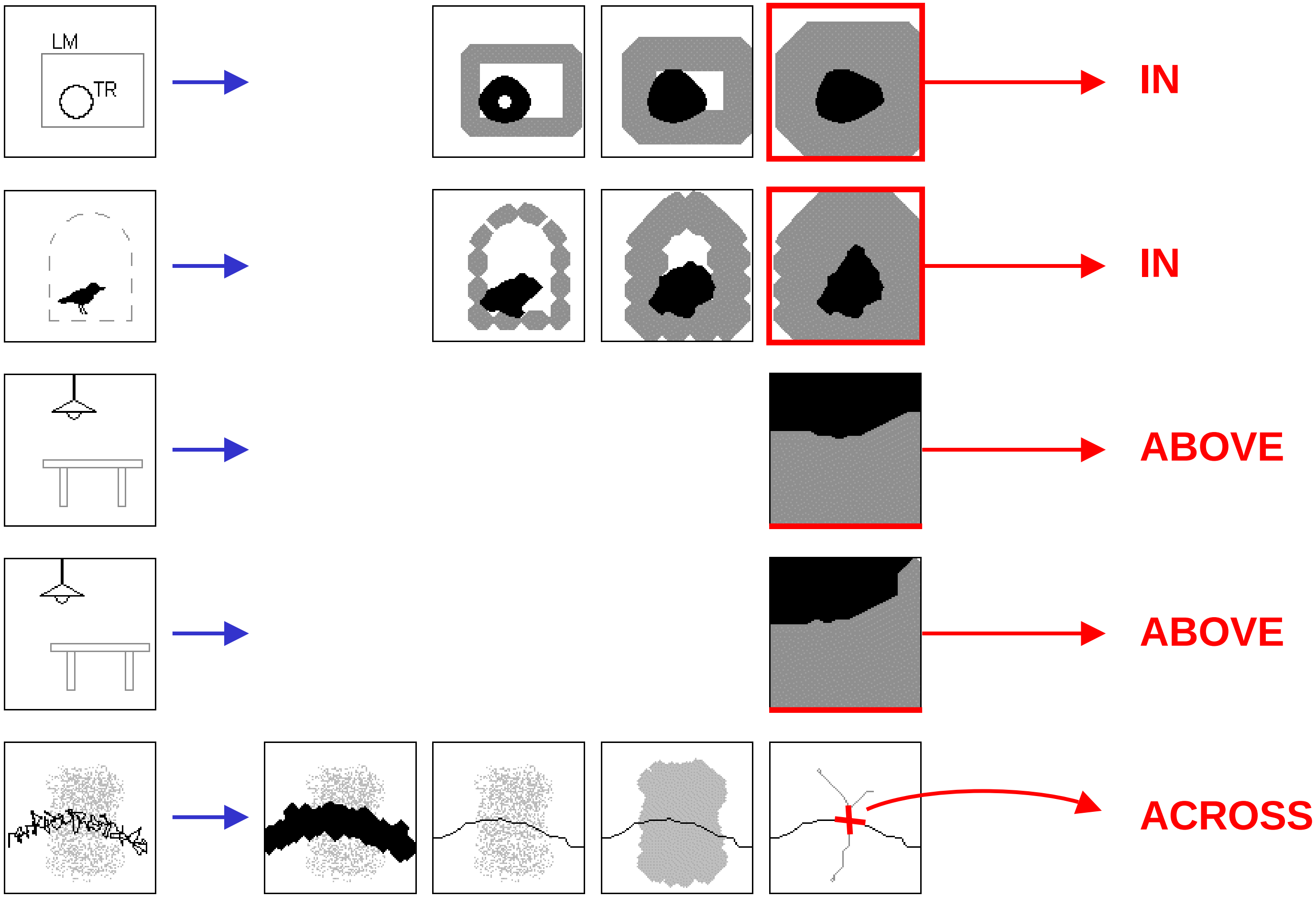
- *scenes representing the same* ***spatial class*** *are not directly similar*
- *what can be compared, however, are virtual structures generated by morphological transforms*

Skeleton by influence zones (SKIZ)



- *SKIZ, a.k.a. . .*
  - *medial axis transform*
  - *cut locus*
  - *stick figures*
  - *shock graphs*
  - *Voronoi diagrams, etc.*

Perceptual-semantic classifier



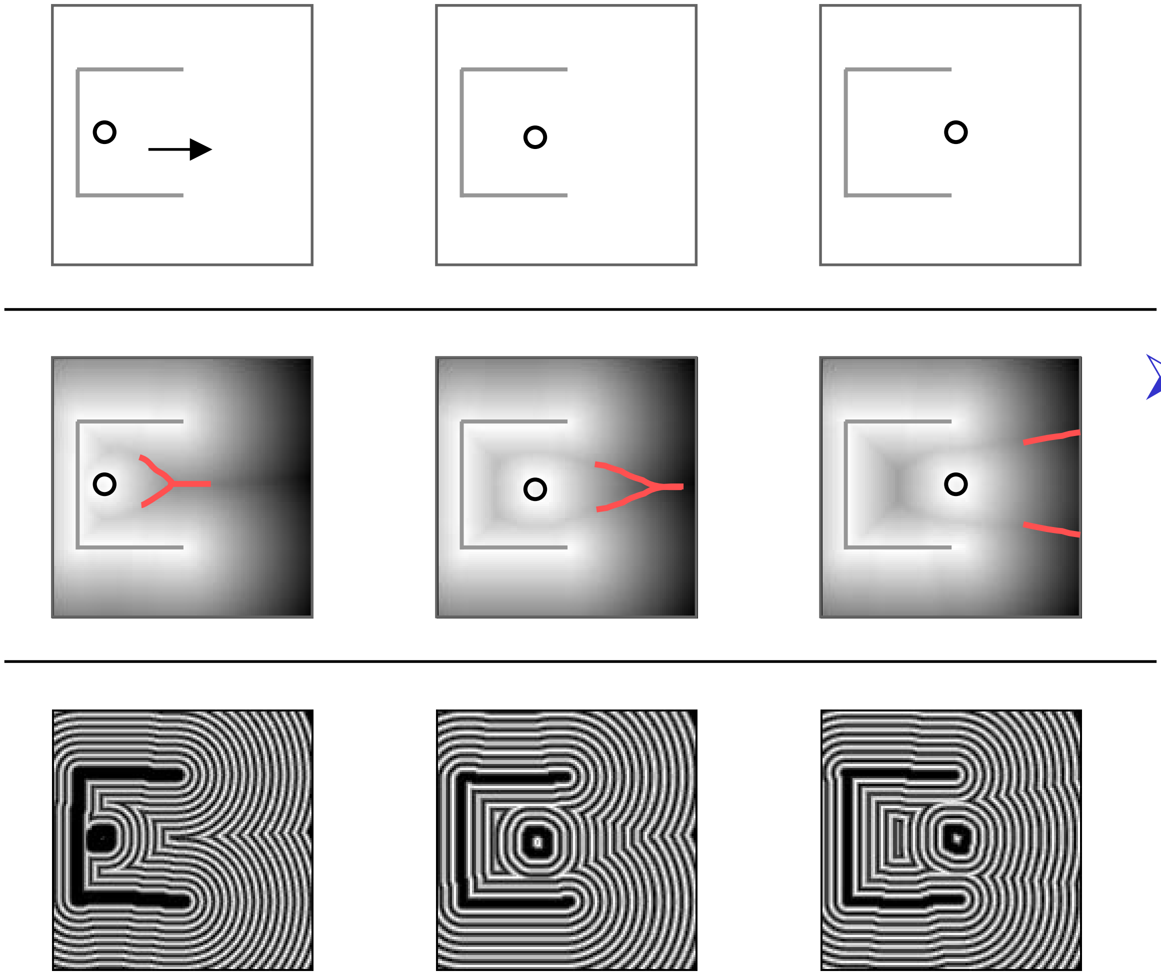
Principles of “active semantics”

a) *objects have a tendency to expand and occupy the whole space around them*

b) *objects are obstacles to each other’s expansion*

- *this creates virtual structures and singularities (e.g., SKIZ = skeleton by influence zones), which constitute the characteristic “signature” of the spatial relationship*
- *transformation routines considerably reduce the dimensionality of the input space, “boiling down” the input images to a few critical features*
- *singularities encode a lot of the image’s geometrical information in a compact and localized manner*

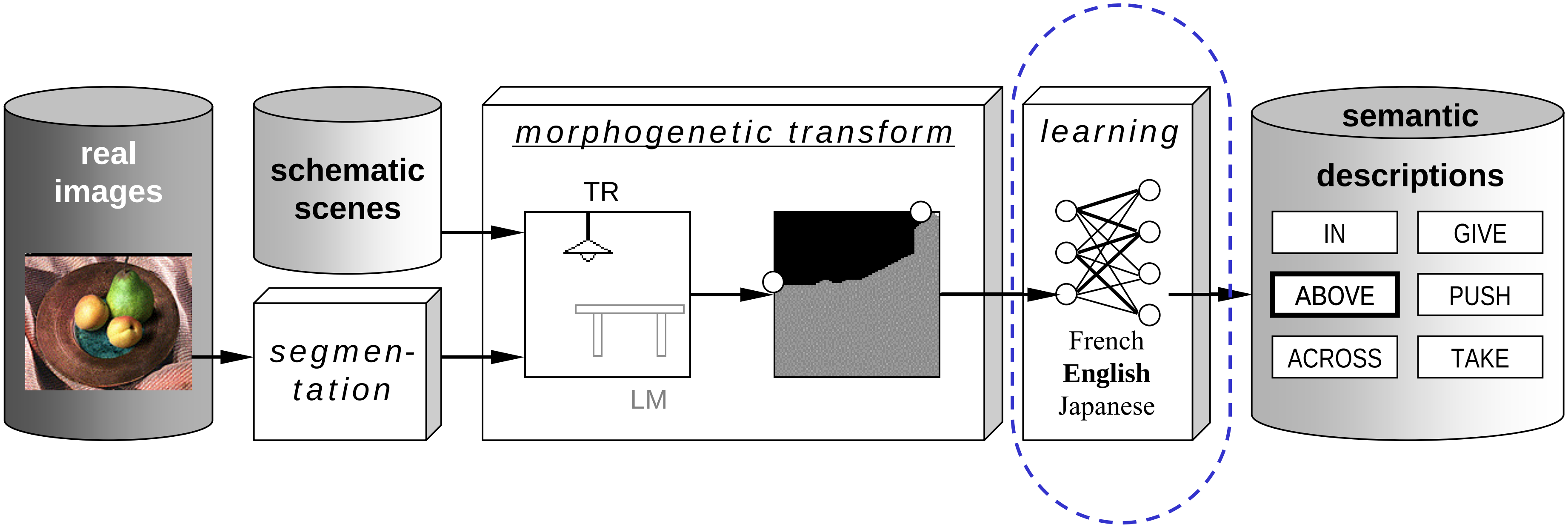
Dynamic evolution of singularities



- ***phase transition:*** *the singularity disappears as the TR exits the interior of the LM (robust phenomenon)*

Perceptual-semantic classifier

Architecture



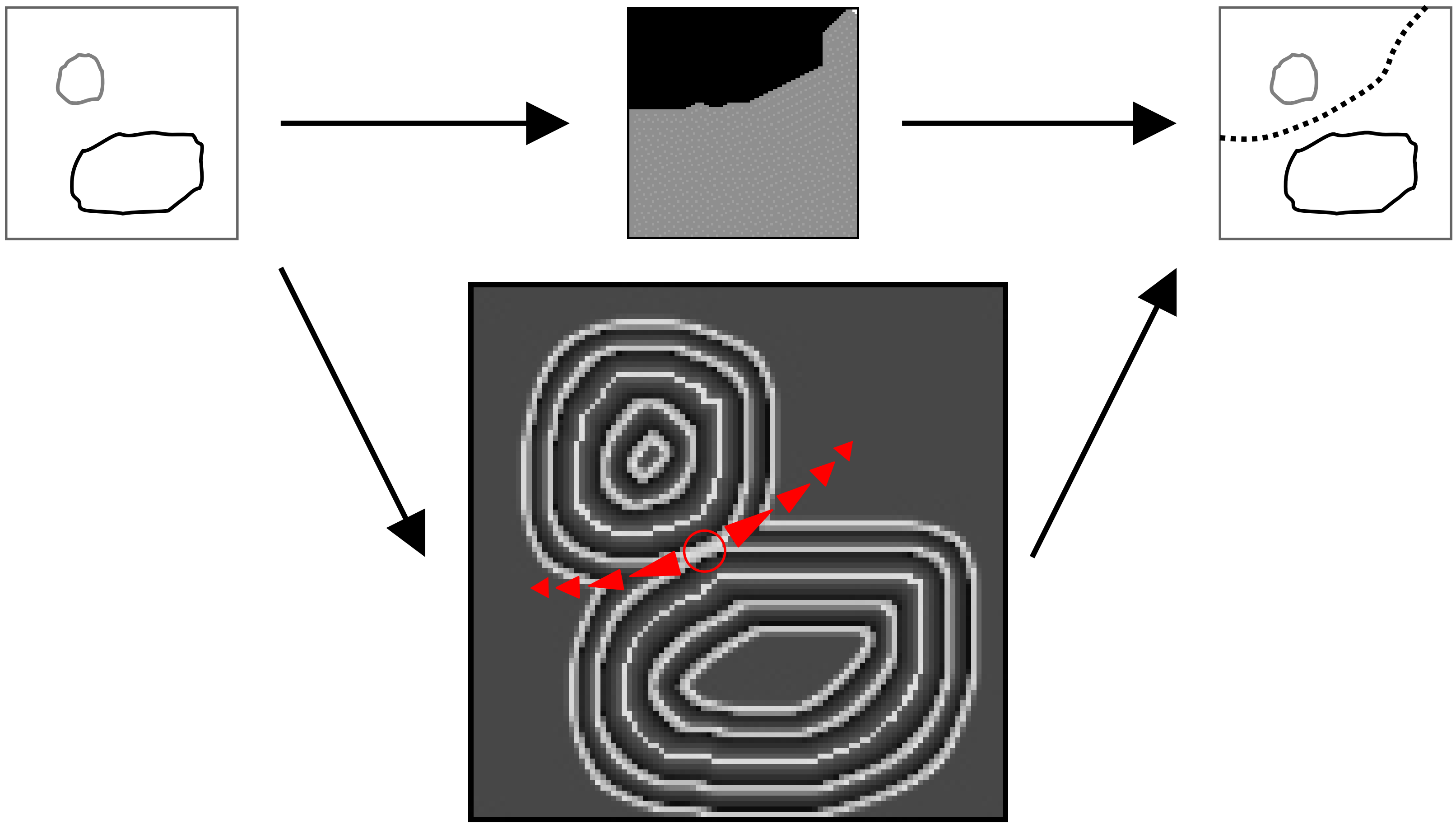
- ***later: introduce a learning module to combine protosemantic concepts into language-specific complex categories***

A Morphodynamical Model of Spatial Cognitive Categories

1. **Spatial categorization**
2. **Cellular automaton model**
3. **Spiking neural model**
   - Temporal coding
   - Oscillators and excitable units
   - Instead of group synchronization: traveling waves
   - Model 1: cross-coupled waves + border detection
   - Model 2: independent waves + complex cells
4. **Discussion**

Spiking neural model (preview)

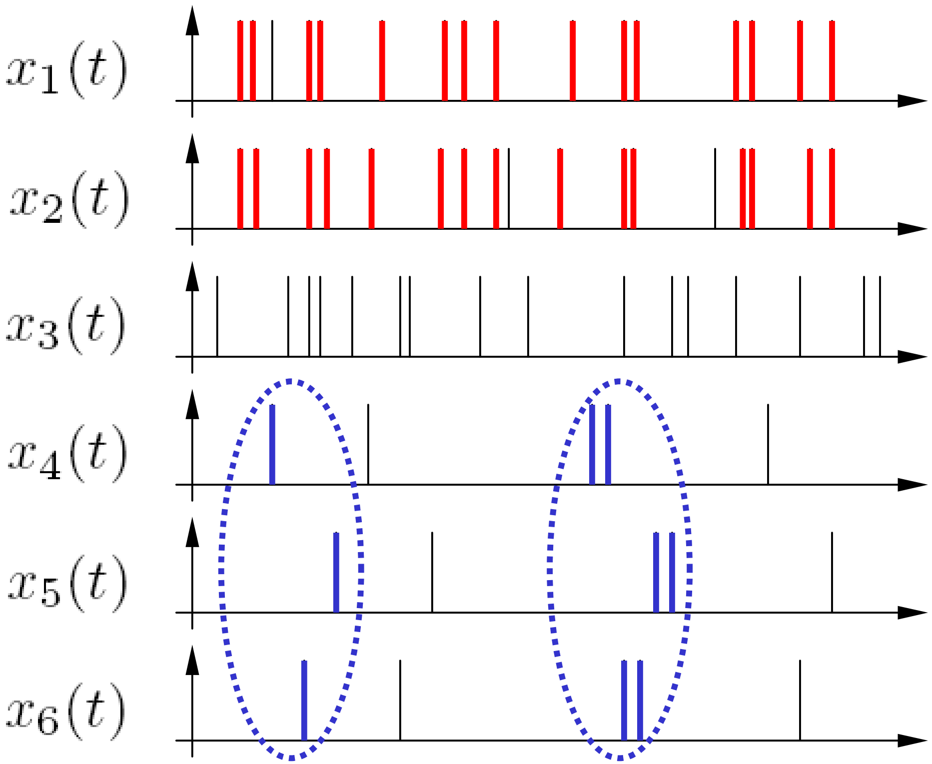
Replace discrete binary transforms with . . .



. . . real-valued, continuous dynamical system

Temporal coding

Synchronization vs. delayed correlations



$\langle\, x_1(t) \,\rangle = \bullet$ high activity rate

$\langle\, x_2(t) \,\rangle = \bullet$ high activity rate

$\langle\, x_3(t) \,\rangle = \bullet$ high activity rate

$\langle\, x_4(t) \,\rangle = \circ$ low activity rate

$\langle\, x_5(t) \,\rangle = \circ$ low activity rate

$\langle\, x_6(t) \,\rangle = \circ$ low activity rate

$$\boxed{\langle\, x_1(t)\, x_2(t) \,\rangle \gg \langle\, x_1(t)\, x_3(t) \,\rangle}$$

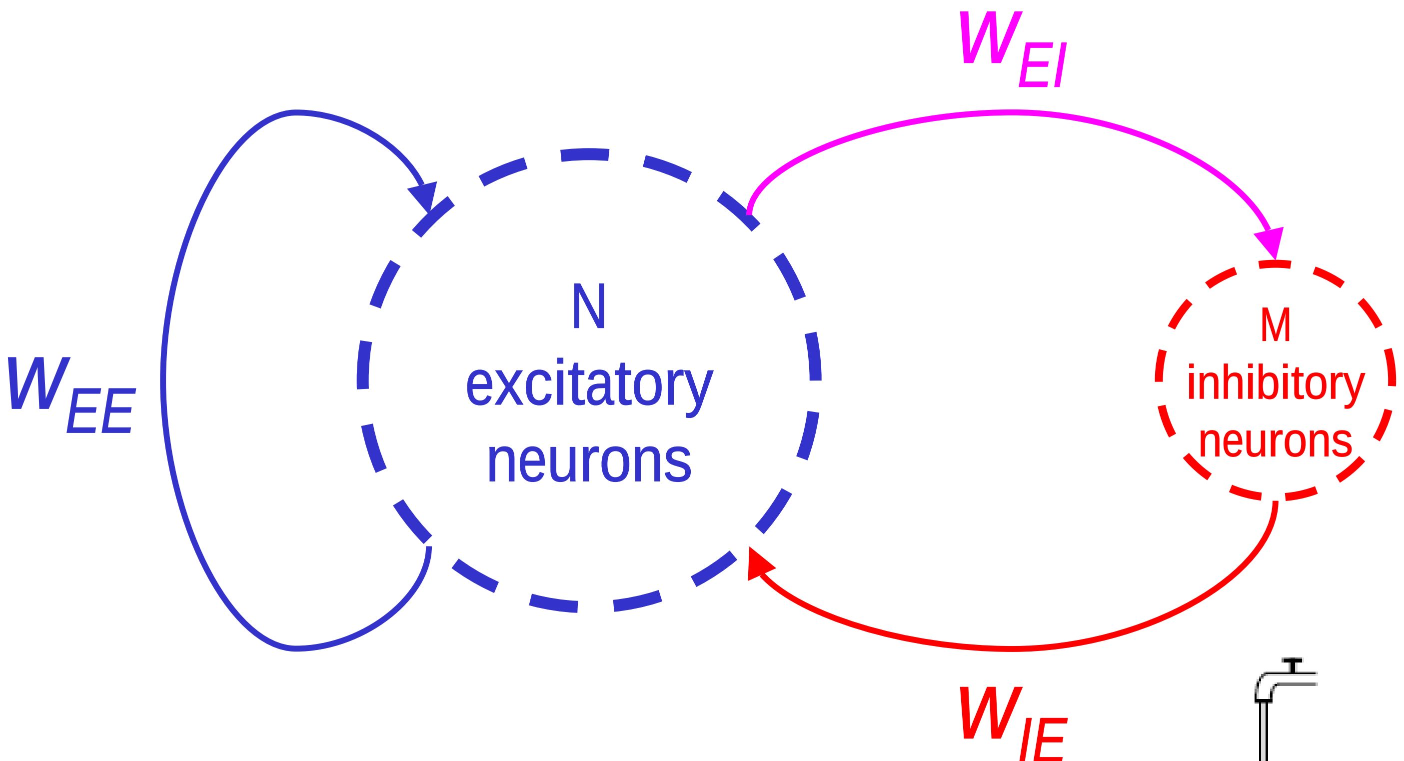
➢ 1 and 2 more in sync than 1 and 3

$$\boxed{\langle\, x_4(t)\, x_5(t - \tau_{4,5})\, x_6(t - \tau_{4,6}) \,\rangle}$$

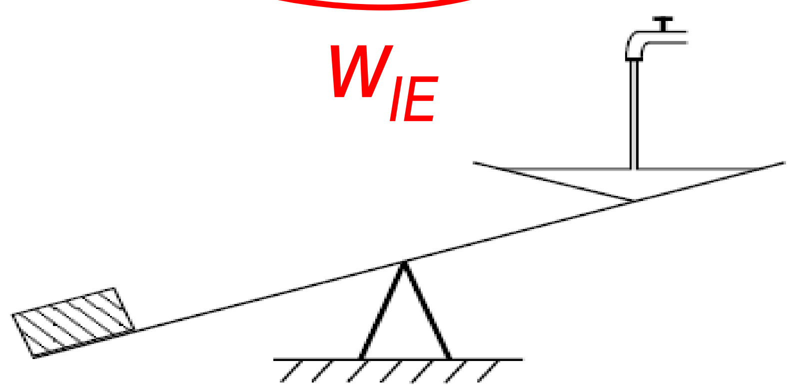
➢ 4, 5 and 6 correlated through delays

Oscillators and excitable units

Excitatory-inhibitory relaxation oscillator



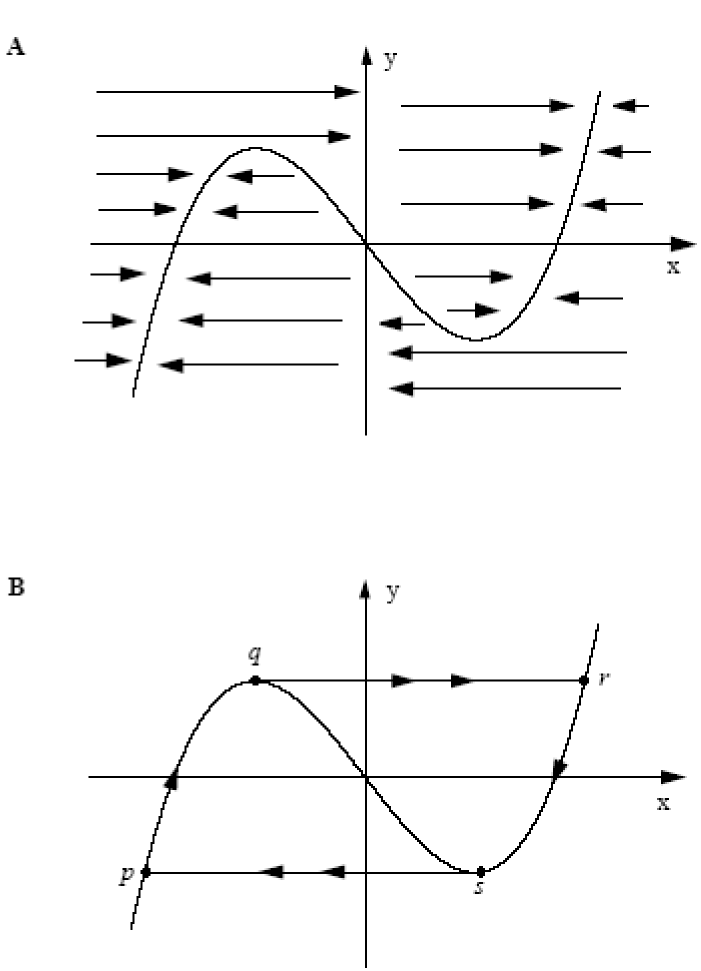
- relaxation oscillators exhibit discontinuous jumps
- different from sinusoidal or harmonic oscillations

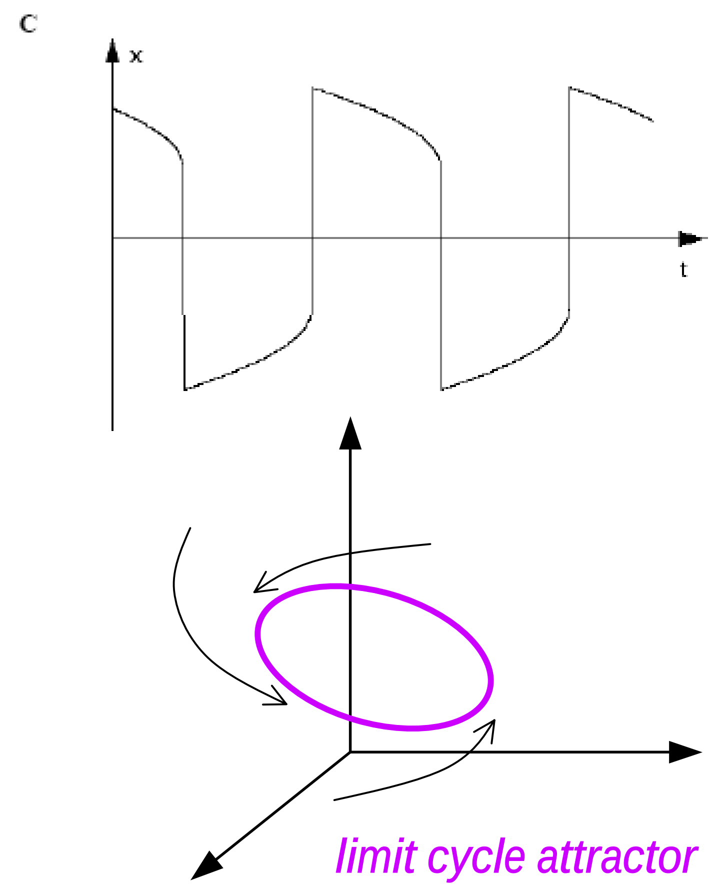


Wang, DeLiang (http://www.cse.ohio-state.edu/~dwang/)

Oscillators and excitable units

Van der Pol relaxation oscillator





Van der Pol relaxation oscillator

Wang, DeLiang (http://www.cse.ohio-state.edu/~dwang/)

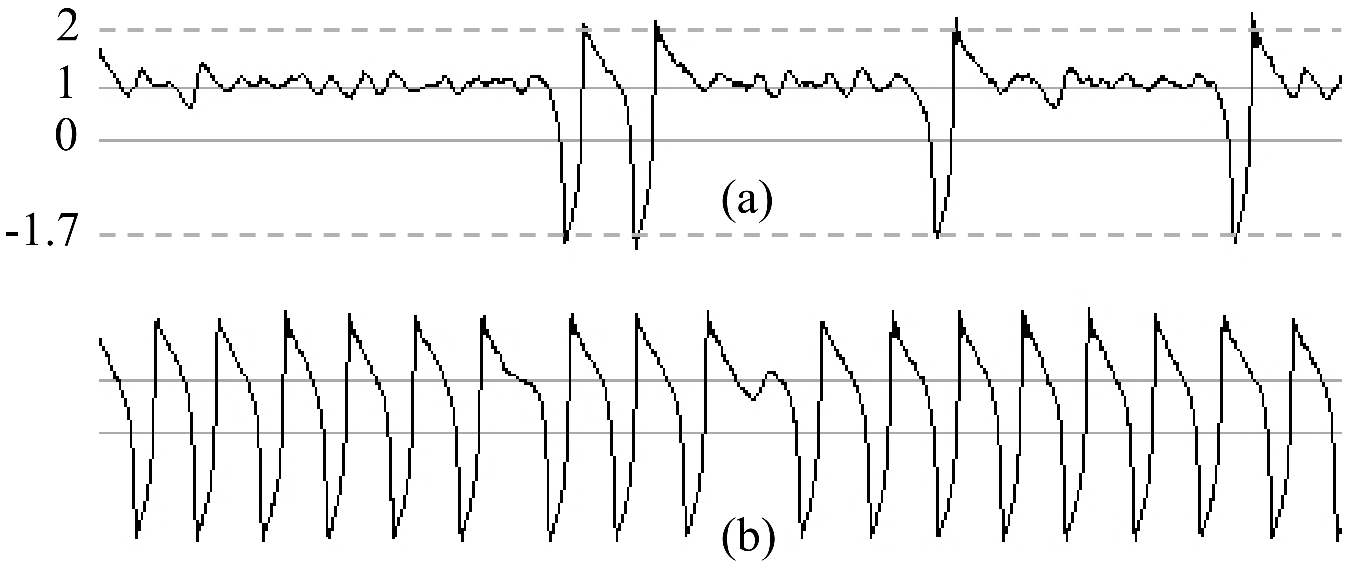
$$\ddot{x} + x = c(1 - x^2)\dot{x} \quad \Leftrightarrow \quad \begin{cases} \dot{x} = c(y - f(x)) \\ \dot{y} = -x \,/\, c \end{cases}$$

Oscillators and excitable units

Bonhoeffer-Van der Pol (BVP) stochastic oscillator

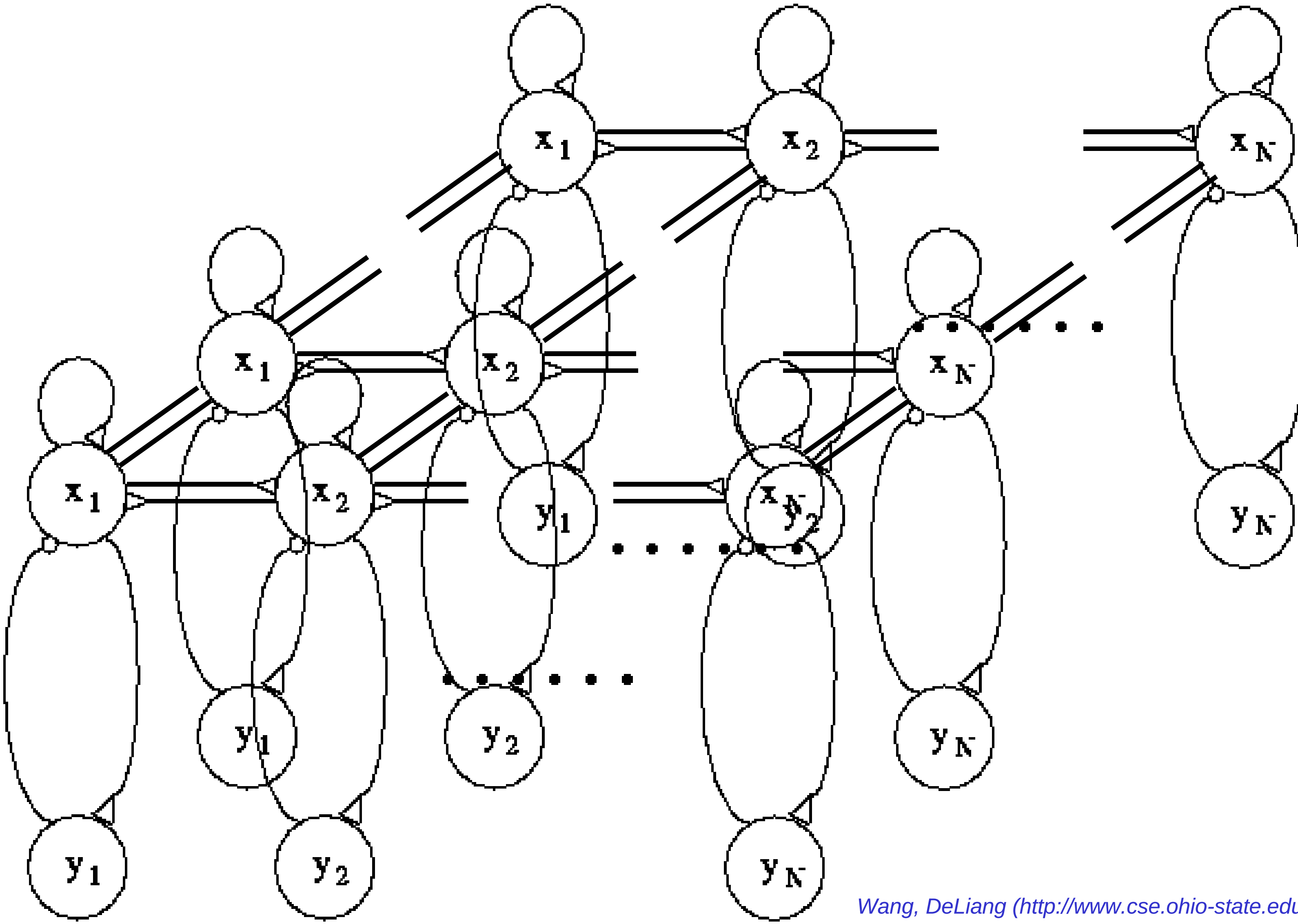
$$\begin{cases} \dot{u}_i = c\left(u_i - u_i^{\,3}/3 + v_i + z\right) + \eta + k\sum_j \left(u_j - u_i\right) + I_i \\ \dot{v}_i = \left(a - u_i - bv_i\right)/c + \eta \end{cases}$$

- *two activity regimes: (a) sparse stochastic and (b) quasi periodic*



Group synchronization

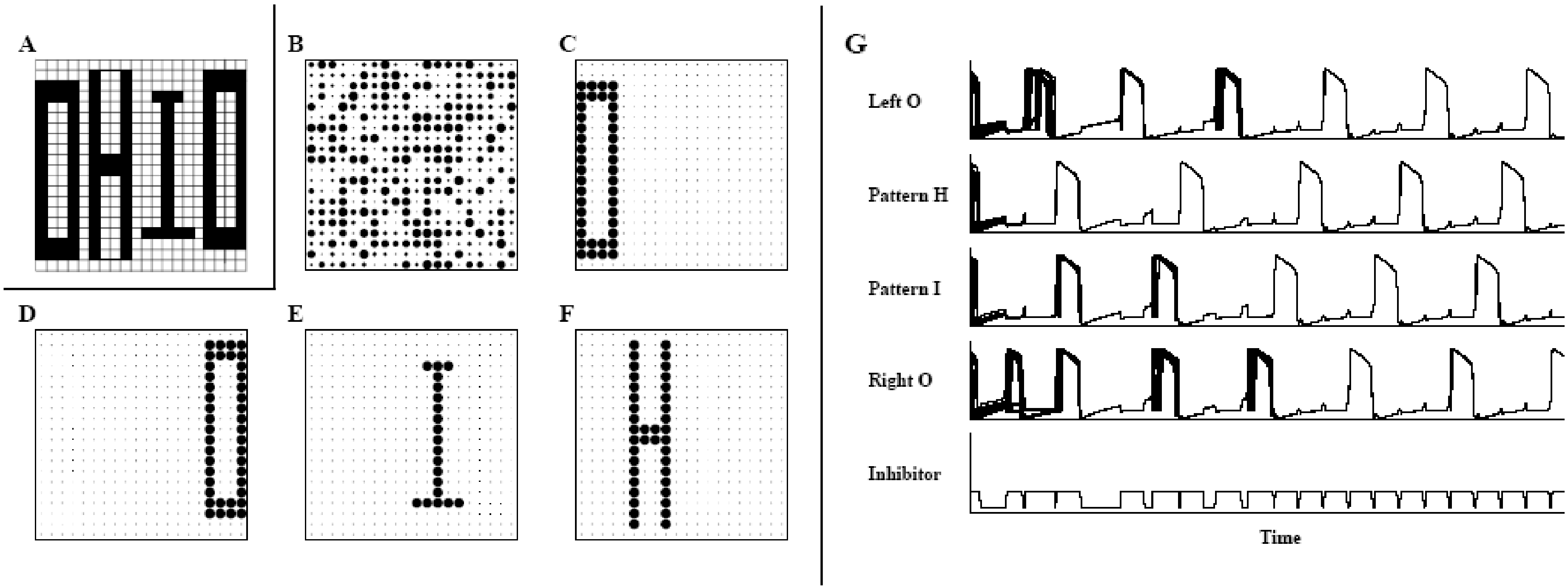
Networks of coupled oscillators



Wang, DeLiang (http://www.cse.ohio-state.edu/~dwang/)

Group synchronization

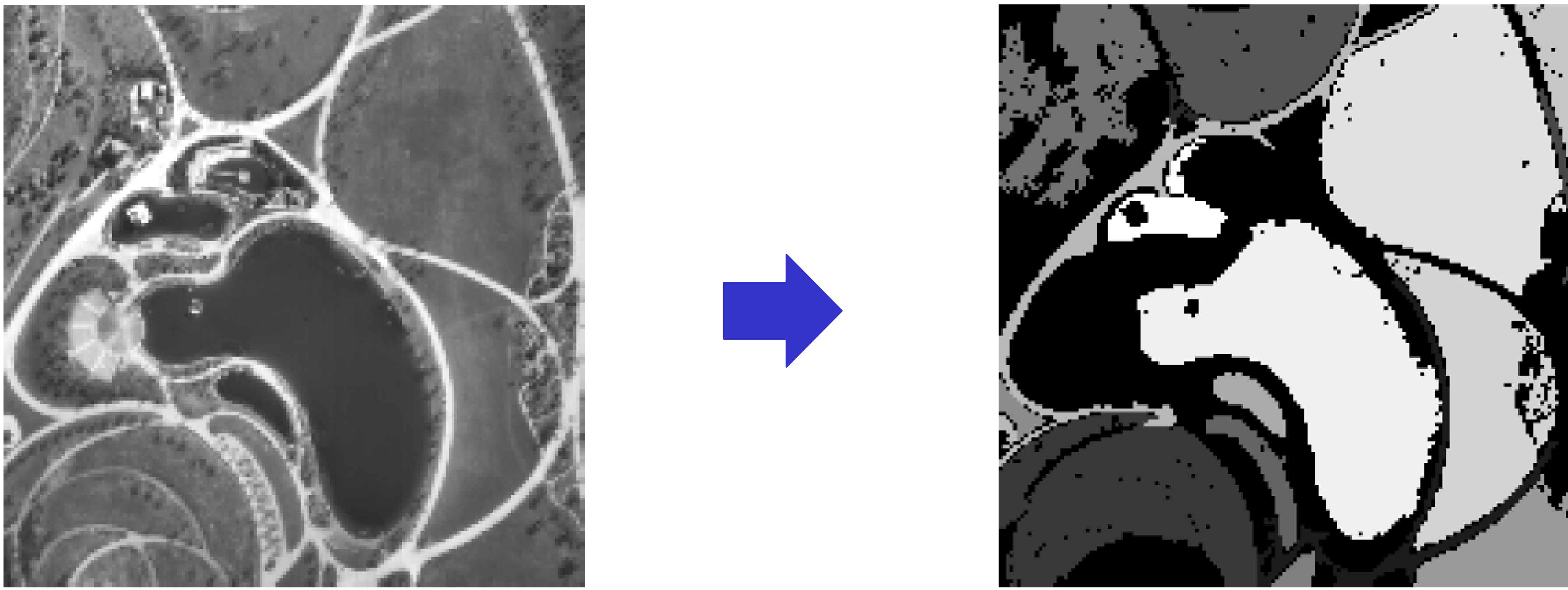
A model of segmentation by sync: LEGION



Wang, D. L. & Terman, D. (1995) Locally excitatory globally inhibitory oscillator networks. *IEEE Trans. Neural Net.*, 6: 283-286.

Group synchronization

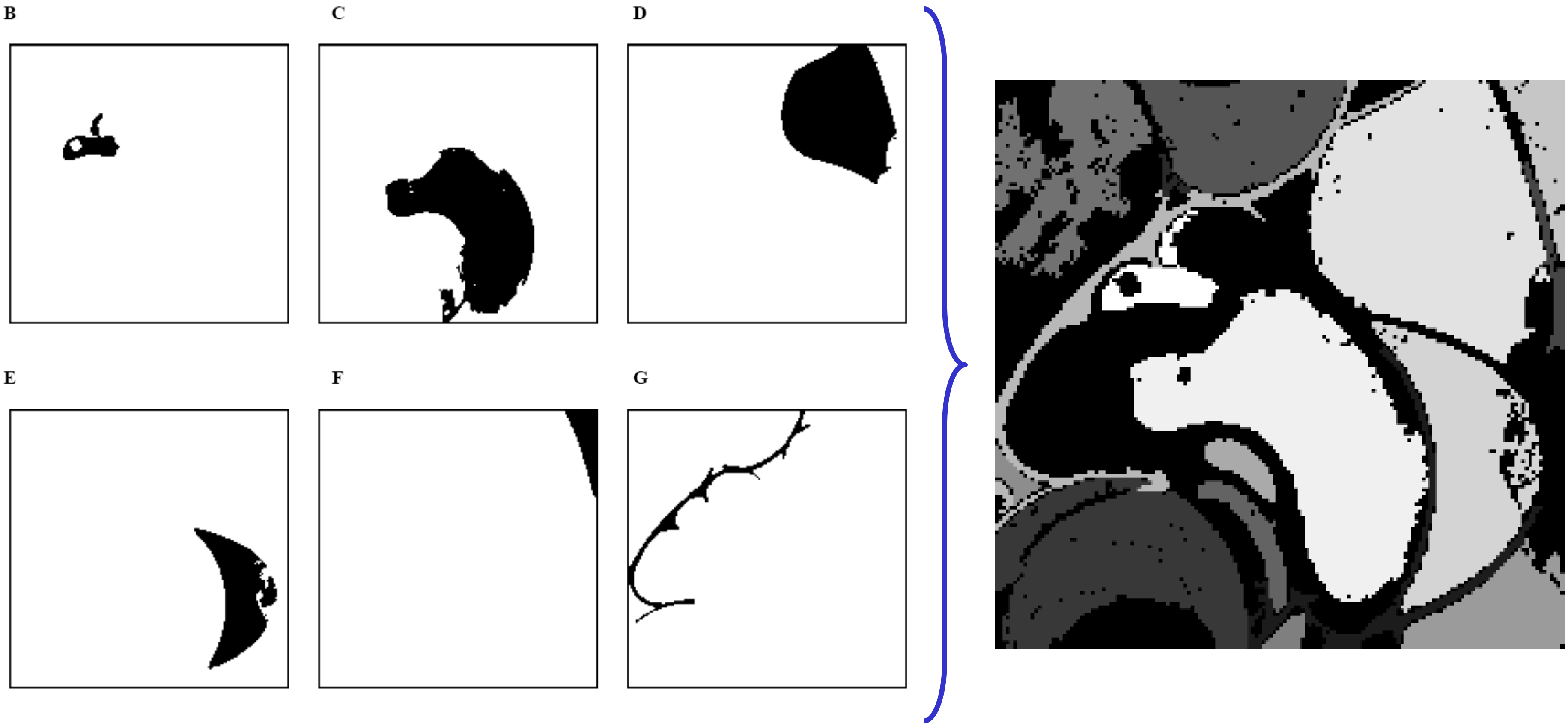
A model of segmentation by sync: LEGION



Wang, D. L. & Terman, D. (1997) Image segmentation based on oscillatory correlation. *Neural Computation*, 9: 805-836,1997

Group synchronization

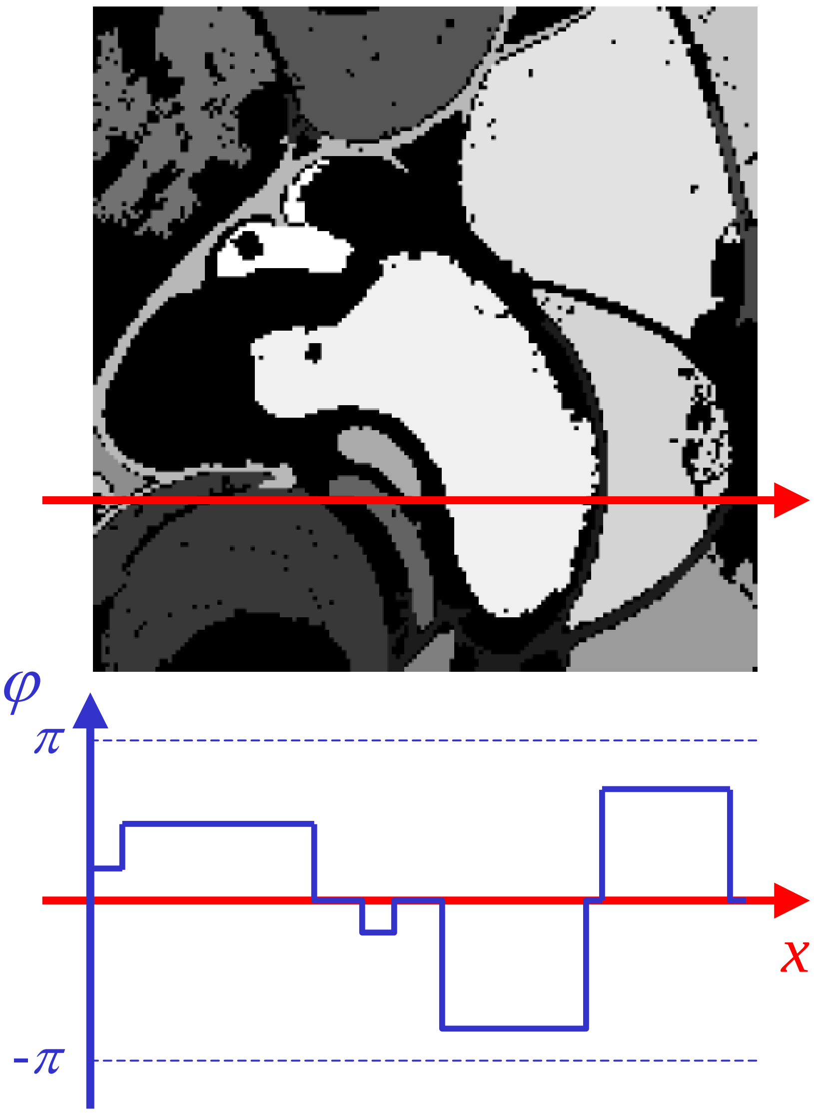
A model of segmentation by sync: LEGION

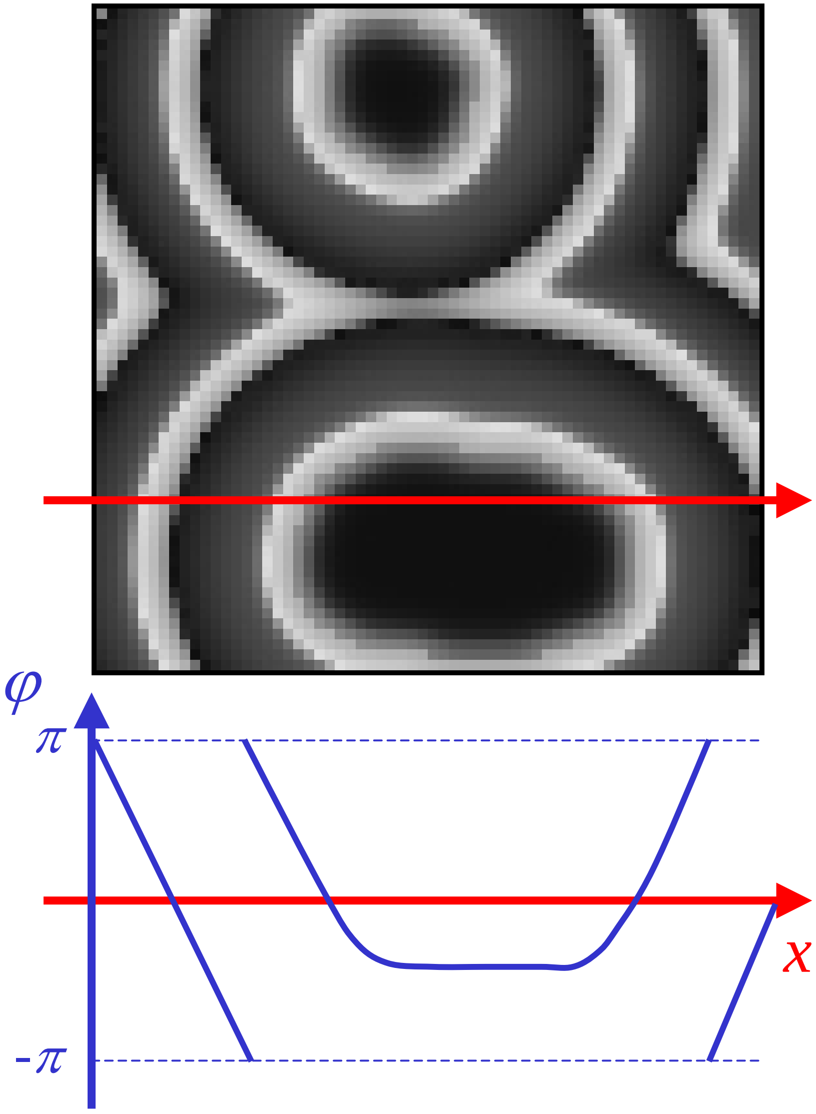


Wang, D. L. & Terman, D. (1997) Image segmentation based on oscillatory correlation. *Neural Computation*, 9: 805-836,1997

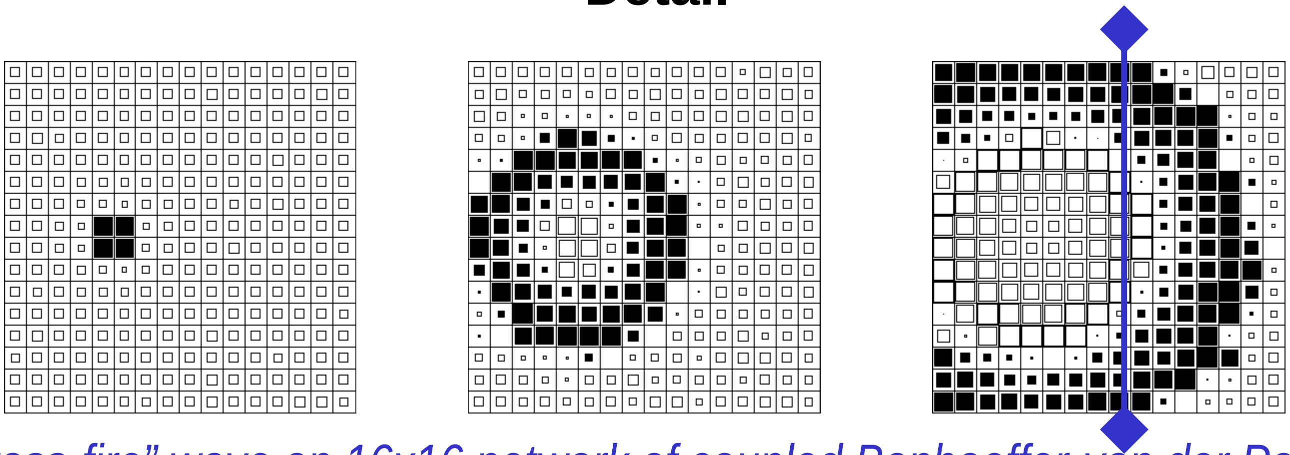
Instead of group synchronization: traveling waves

Instead of phase plateaus: phase gradients

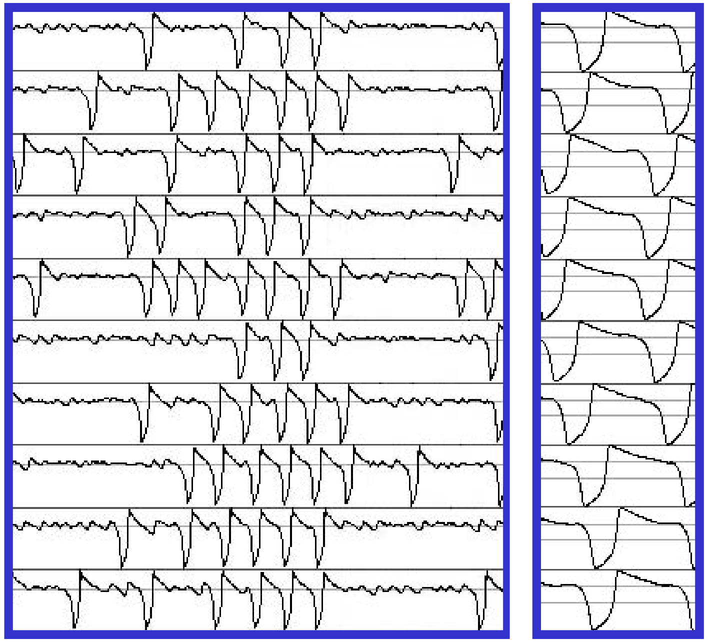




Traveling waves
Detail

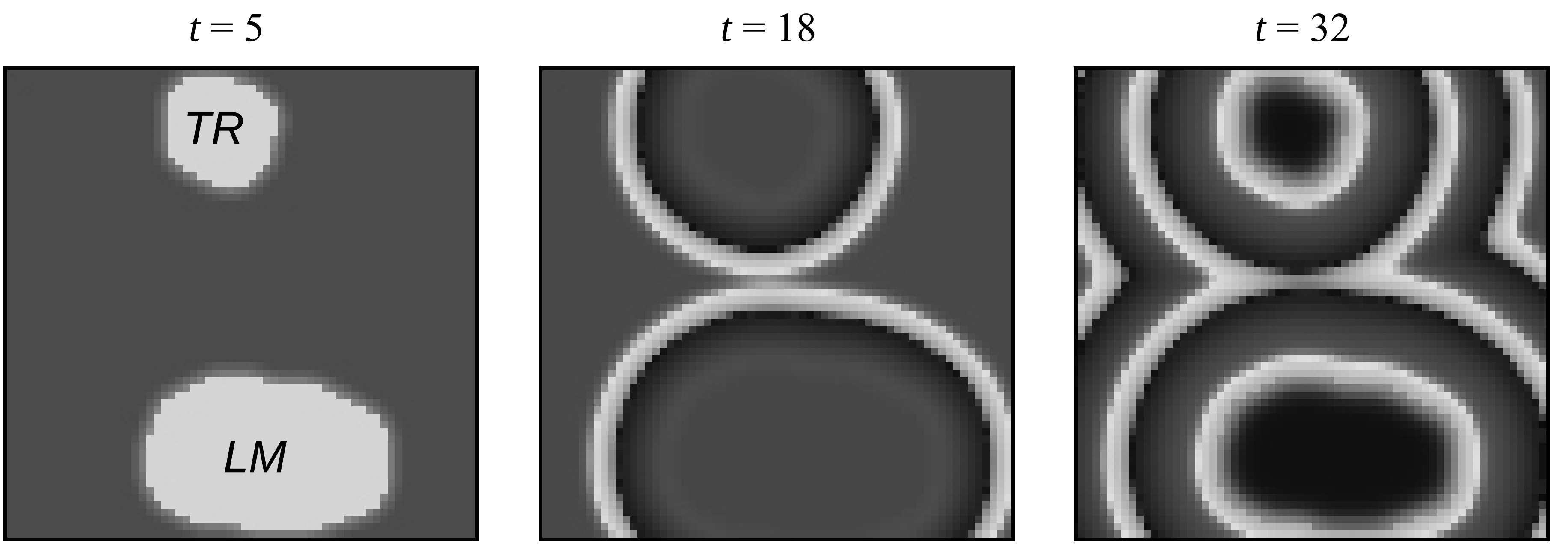


- *"Grass-fire" wave on 16x16 network of coupled Bonhoeffer-van der Pol units*



Traveling waves

Wave collision

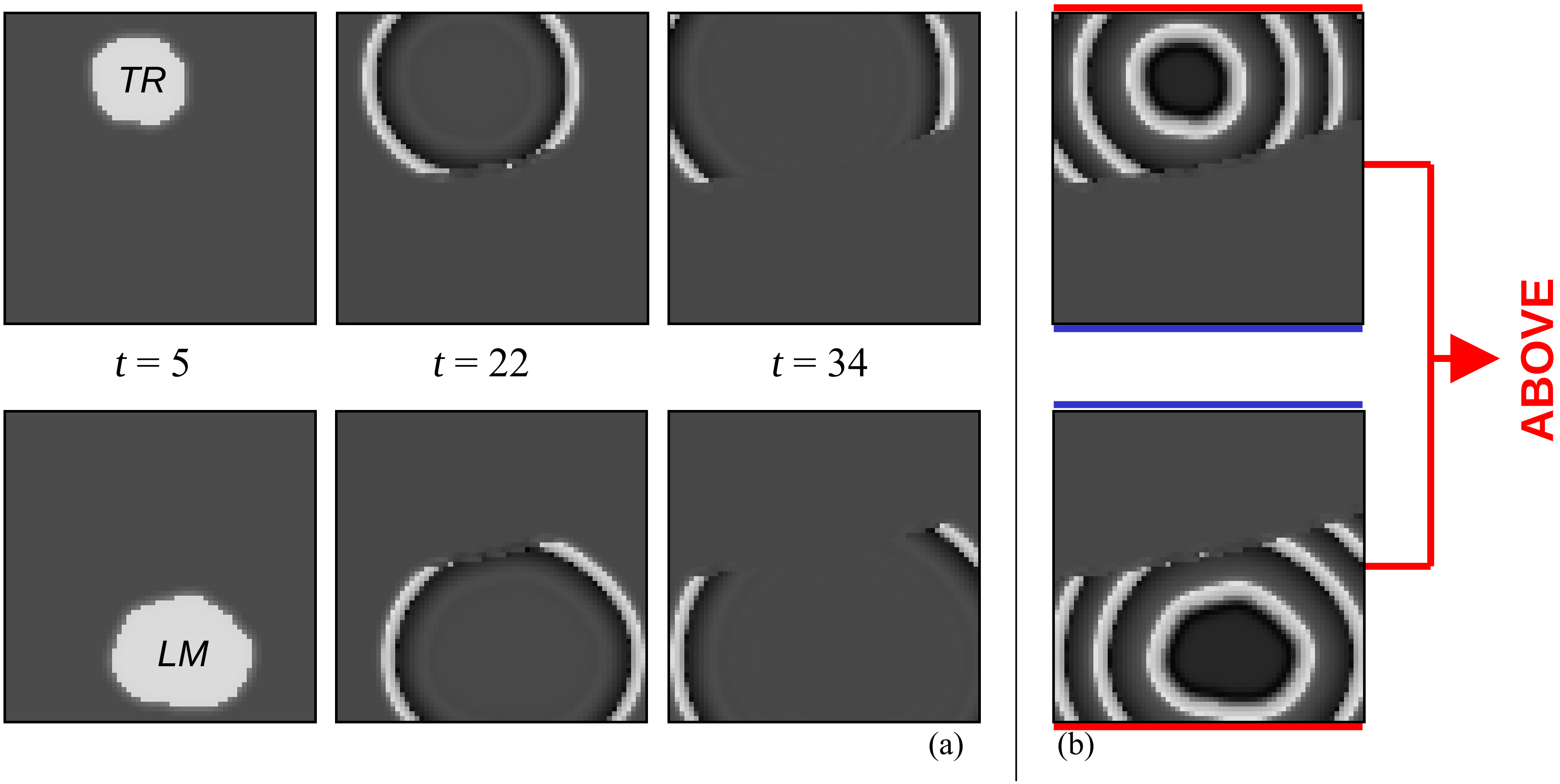


- *64 x 64 lattice of locally coupled Bonhoeffer-van der Pol oscillators*
- *. . . but how can we discriminate between activity coming from TR and LM?*

Doursat, R. & Petitot, J. (2005) Dynamical Systems and Cognitive Linguistics: Toward an Active Morphodynamical Semantics. To appear in *Neural Networks*.

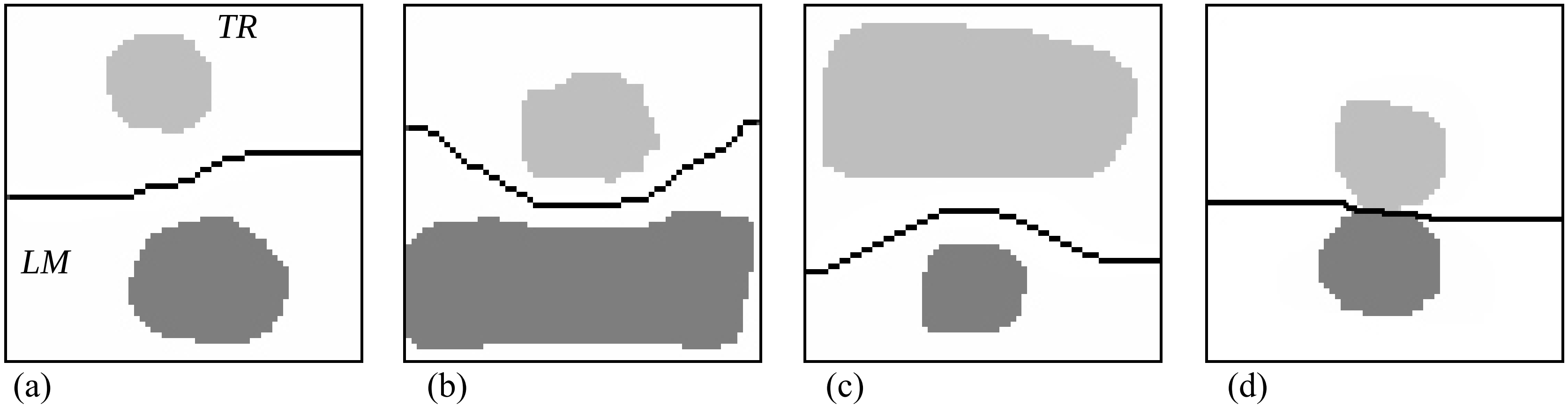
Traveling waves

Model 1: crossed-coupled waves + frame border detection



- *use two cross-coupled, mutually inhibiting lattices of coupled oscillators*

Frame border detection not enough



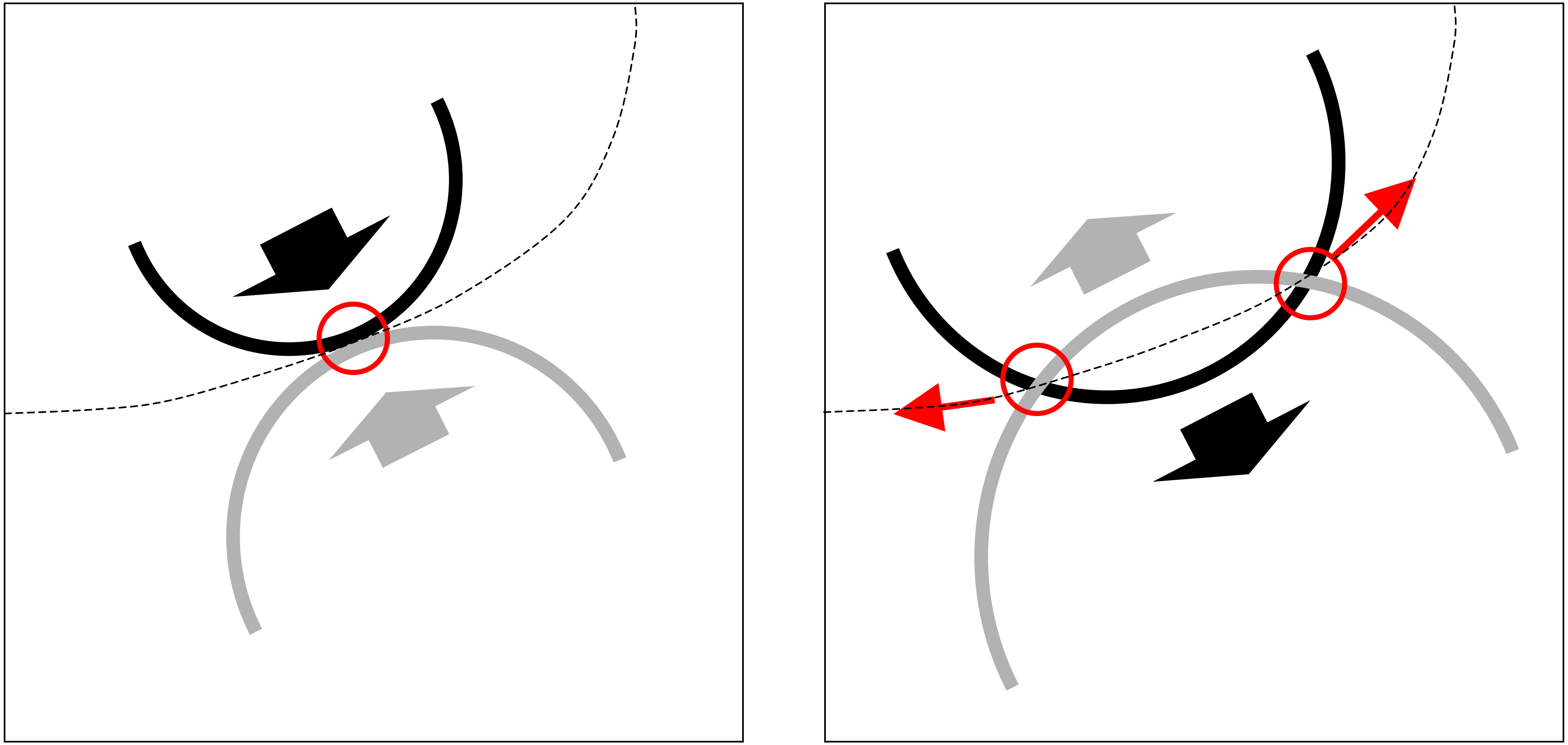
- *how to distinguish among:*
  (a-c) English 'above'
  (b) Mixtec 'siki': LM is horizontally elongated (Regier, 1996)
  (c) French 'par-dessus': TR is horizontally elongated and covers LM
  (d) German 'auf': TR is in contact with LM

- *problem: all yield the same type of frame border activity (upper half TR, lower half LM)*

- *need for a refined SKIZ-based signature*

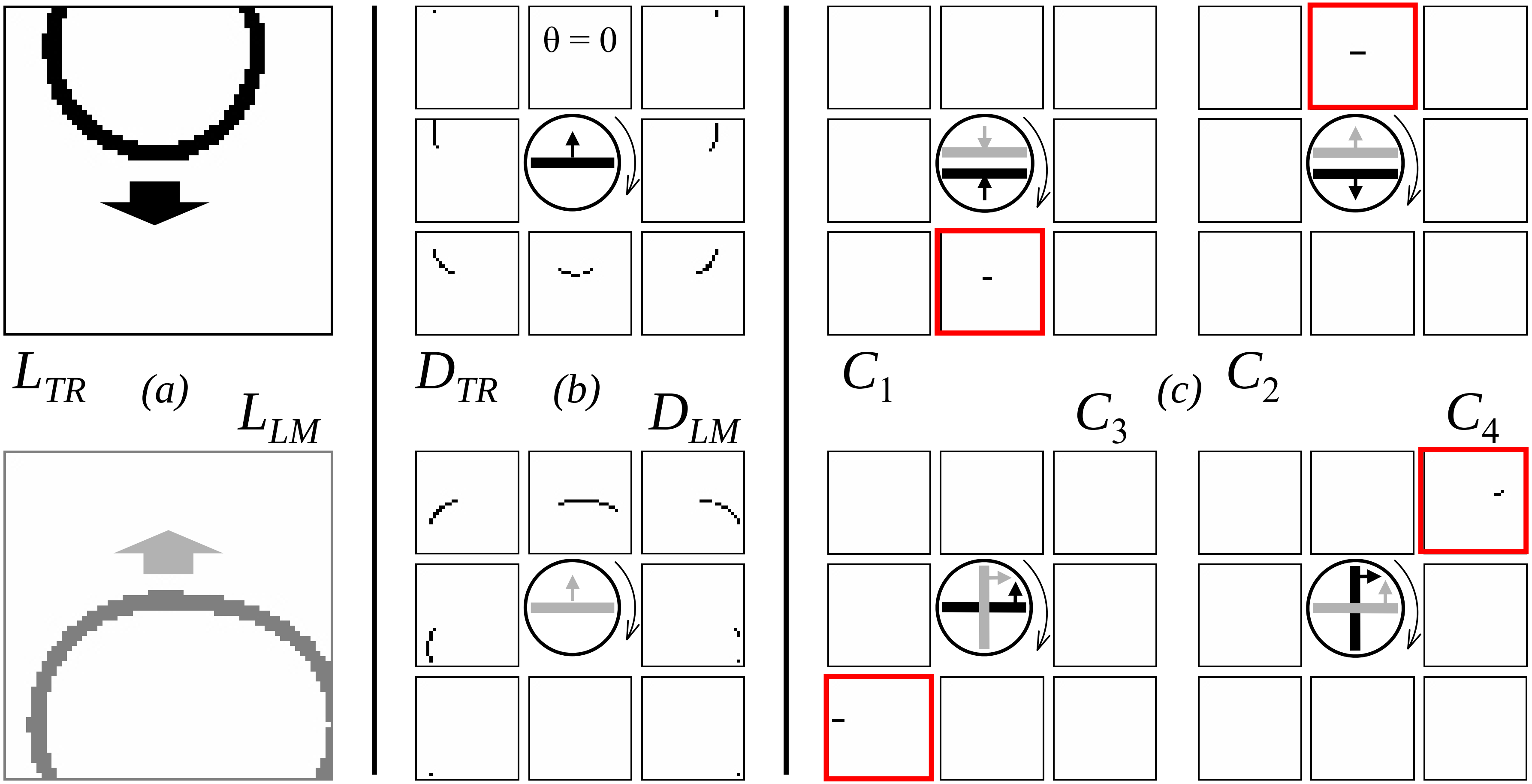
Traveling waves

Model 2: independent waves + complex readout cells



Traveling waves

Model 2: independent waves + complex readout cells



the activity in layers C provide a sparse signature of the scene specific of the SKIZ line

A Morphodynamical Model of Spatial Cognitive Categories

1. **Spatial categorization**
2. **Cellular automaton model**
3. **Spiking neural model**
4. **Discussion**
   - Future work
   - Originality
   - Appendix: pattern formation in excitable media

Future work

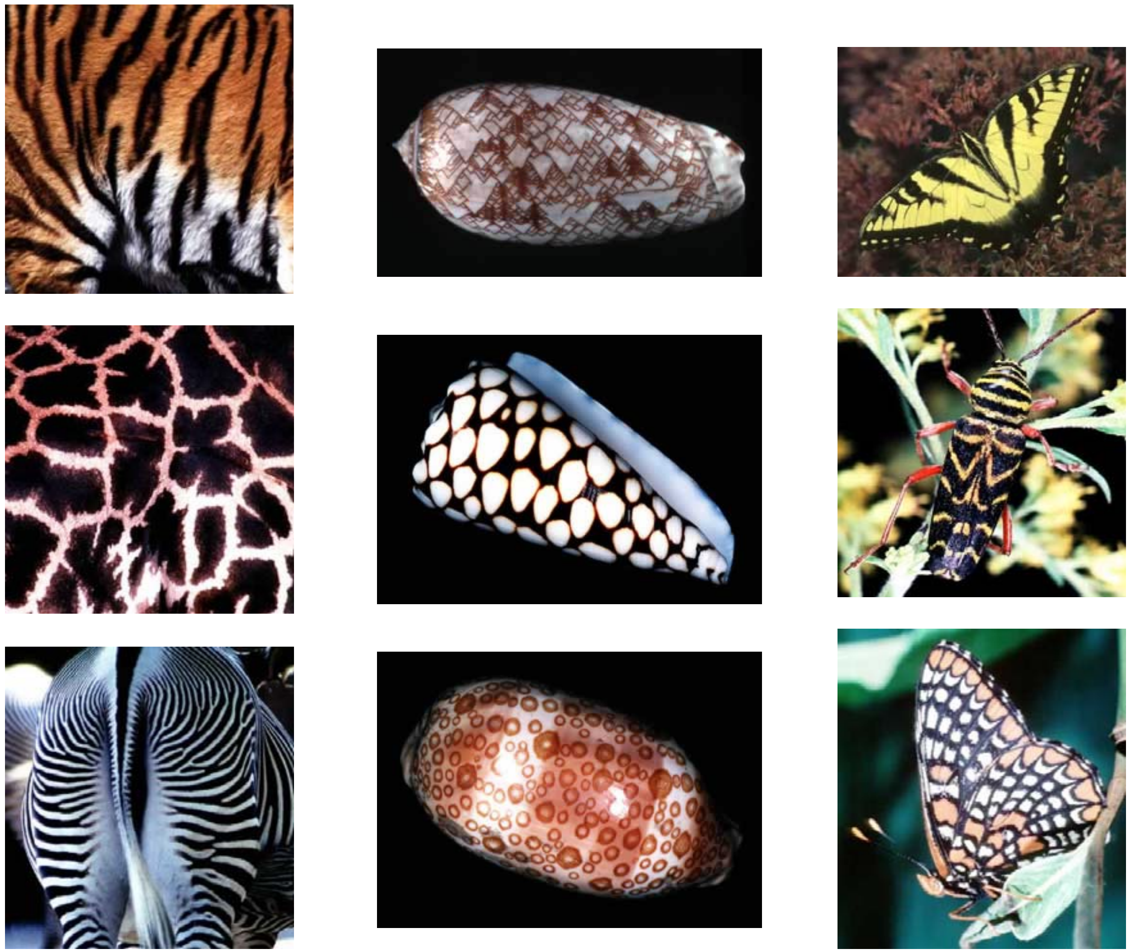
1. *wave dynamics and scene database*
   - *systematic investigation of morphodynamical routines using a database of image/label pairs*
2. *real images and low-level visual processing*
   - *start from real images via segmentation preprocessing*
3. *learning the semantics from the protosemantics*
   - *combine protosemantic features (IN-1, IN-2, etc.) into full-fledged cultural-linguistics categories (IN, AUF, etc.) using learning methods*
4. *verb processes and complex scenes*
   - *also investigate movies (bifurcation of singularities) and composition between schemas*

Originality

1. *bringing large-scale dynamical systems to cognitive linguistics*
   - *CL is lacking computational foundations — there were a few attempts, but mostly small "hybrid" ANNs*
2. *addressing semantics in cellular automata and neural networks*
   - *using large-scale network of coupled neural units for high-level semantic feature extraction — normally used for low-level image processing or visual cortical modeling (e.g., PCNNs, CNNs)*
3. *advocating pattern formation in neural modeling*
   - *many physical, chemical, and biological media exhibit pattern formation; as a complex system, too, the brain produces "forms" = spatiotemporal patterns of activity — yet, not a main field of research*
4. *suggesting wave dynamics in neural organization*
   - *waves open a rich space of temporal coding for mesoscopic neural modeling, between micro neural activities and macro mental objects*

Pattern formation

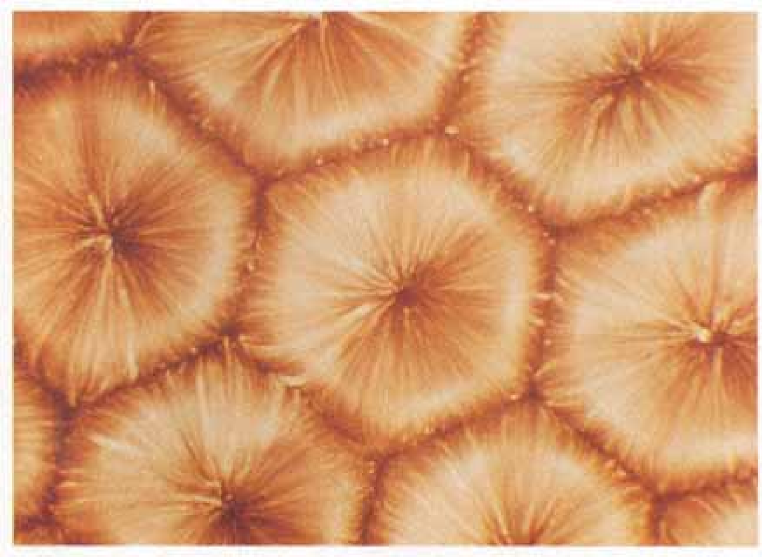
Stationary patterns



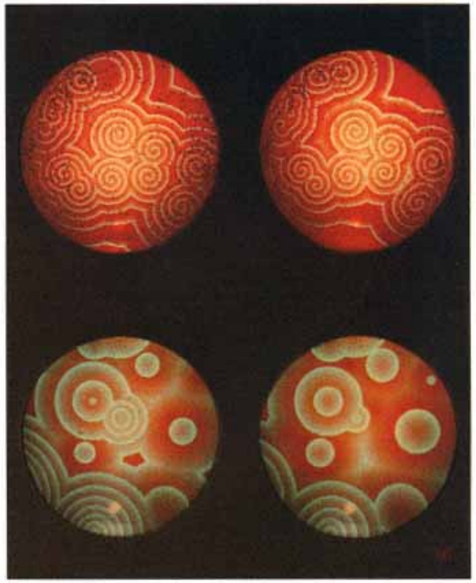
Mammal fur, seashells, and insect wings
(Scott Camazine, http://www.scottcamazine.com)

Pattern formation in excitable media

Physical-chemical media



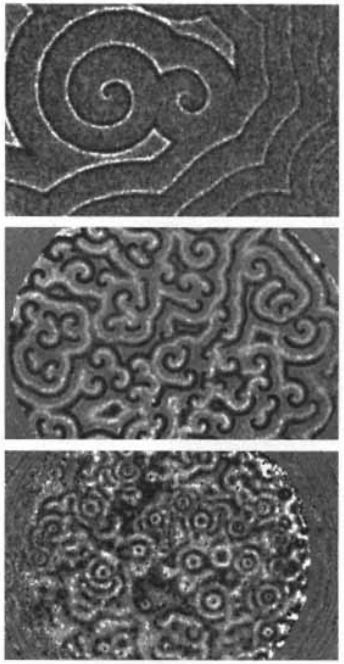
Rayleigh-Benard convection cells
in liquid heated uniformly from below
(Manuel Velarde, Universidad Complutense, Madrid.)



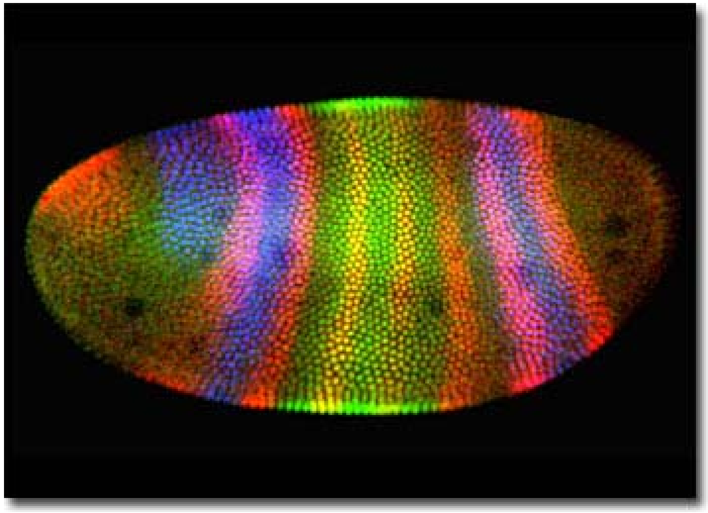
Circular and spiral traveling waves
in Belousov-Zhabotinsky reaction
(Arthur Winfree, University of Arizona.)

Pattern formation in excitable media
Multicellular structures

Spiral waves in the heart in a model of a dog heart
(James Keener, University of Utah.)



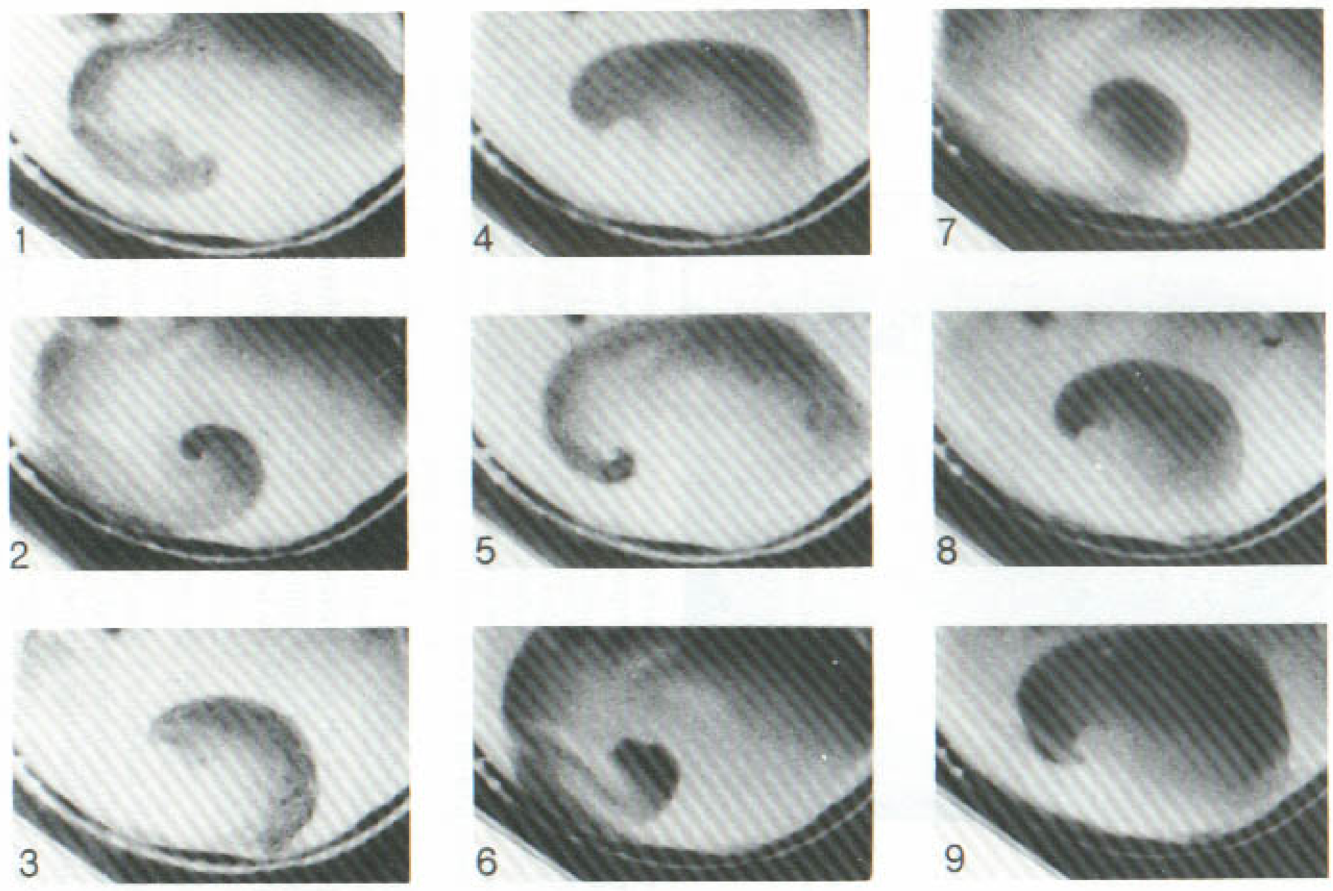
Wave patterns in aggregating slime mold amoebas
(Brian Goodwin, Schumacher College, UK.)



Differential gene expression stripes in fruit fly embryo
(Steve Paddock, Howard Hughes Medical Institute)

Pattern formation in excitable media

Retina of the chicken



Dark front of spreading depression rotating on the retina of a chicken
(40-second interval frames)
(Gorelova and Bures, 1983)

A Morphodynamical Model of Spatial Cognitive Categories

1. **Spatial categorization**
2. **Cellular automaton model**
3. **Spiking neural model**
4. **Discussion**