What about larger-scale representations?

Schemas: Essential properties

- Schemas have variables
  - Slots have restrictions (e.g., AGENT must be animate)
  - Default values (values in absence of more specific information)
    - But must be context-sensitive (agent in breaking window vs. bubble)
- Schemas can embed
  - BREAK contains DO and CAUSE
  - Not always simpler (e.g., room with picture of room)
- Schemas range across levels of abstraction
  - Original focus on lexical level (like GIVE, BREAK)
  - Also intended to span larger “events” (e.g., restaurant “script”)
- Schemas represent knowledge rather than definitions
  - Not “definitional” but what is “normal”

Challenges for traditional theories of schemas

- How to select relevant schemata (best-match problem)
- How to integrate multiple schemata (birthday party in restaurant)
- How to create new schemata
  - Specialize/generalize existing ones? Hybrids?
  - Transition from single instance to “general” knowledge
  - Proliferation makes selection problem more difficult

Schemas in constraint satisfaction networks

- Situations composed of primitive “features”
- A schema consists of knowledge about what features go with other features (i.e. constraints between features)
- Certain subpatterns tend to act in concert
  - Support each other and inhibit same sets of other units (“stable coalitions”)
- Good interpretations are goodness maxima / energy minima
- **No structure corresponds to a schema**
  - more like a description of structured/systematic behavior of system
Schema model (Rumelhart et al., 1986)

- Two subjects each imagined 8 different versions of 5 room types:
  - kitchen, office, bathroom, bedroom, living room

- For each imagined room, subject decided which of 40 descriptors applied to it

- Network has 40 units (one per descriptor); fully connected

- **Weights** on connections between units were set based on the likelihoods, across rooms, that the two descriptors agreed (both on or both off)

- **Biases** of units were set based on the likelihoods, across rooms, that individual descriptors were included

\[
\begin{align*}
    w_{ij} &= -\ln \frac{p(x_i = 0 \land x_j = 1)p(x_i = 1 \land x_j = 0)}{p(x_i = 1)p(x_j = 0)} \\
    bias_i &= -\ln \frac{p(x_i = 0)}{p(x_i = 1)}
\end{align*}
\]
Goodness surface in state space

**State space**
- A high-dimensional space with a dimension for each of $n$ units in the network
- Each unit’s activity (state) can be interpreted as a coordinate along its corresponding axis/dimension
- At any instant in time, the current pattern of activity over the entire network corresponds to a particular $n$-dimensional point in the space
- As units update their states, the point moves in state space

**Goodness surface**
- Each instantaneous pattern of activity has a corresponding “goodness” value
- Add an additional ($n+1$)st dimension to state space so that the goodness of each point (activity pattern) can be plotted “above” it
- The set of goodness values corresponding to all possible activity patterns forms a continuous surface “above” state space