Cognitive Architectures for Virtual Humans

Paul S. Rosenbloom | 6/15/2011

The projects or efforts depicted were or are sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM) Simulation Training and Technology Center (STTC). The content or information presented does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

Outline

- **Desiderata**
- **Dichotomies**
- **Techniques**
- **Results**

Desideratum I
Broad Spectrum Architectures

- **Easy to create simple virtual humans**
  - *Data driven*, like NPCEditor (as in SimCoach and other systems)
- **Can create very sophisticated virtual humans**
  - *Model based*, like SASO and MRE
- **Incrementally extend to arbitrary points in between**

Combining Paradigms

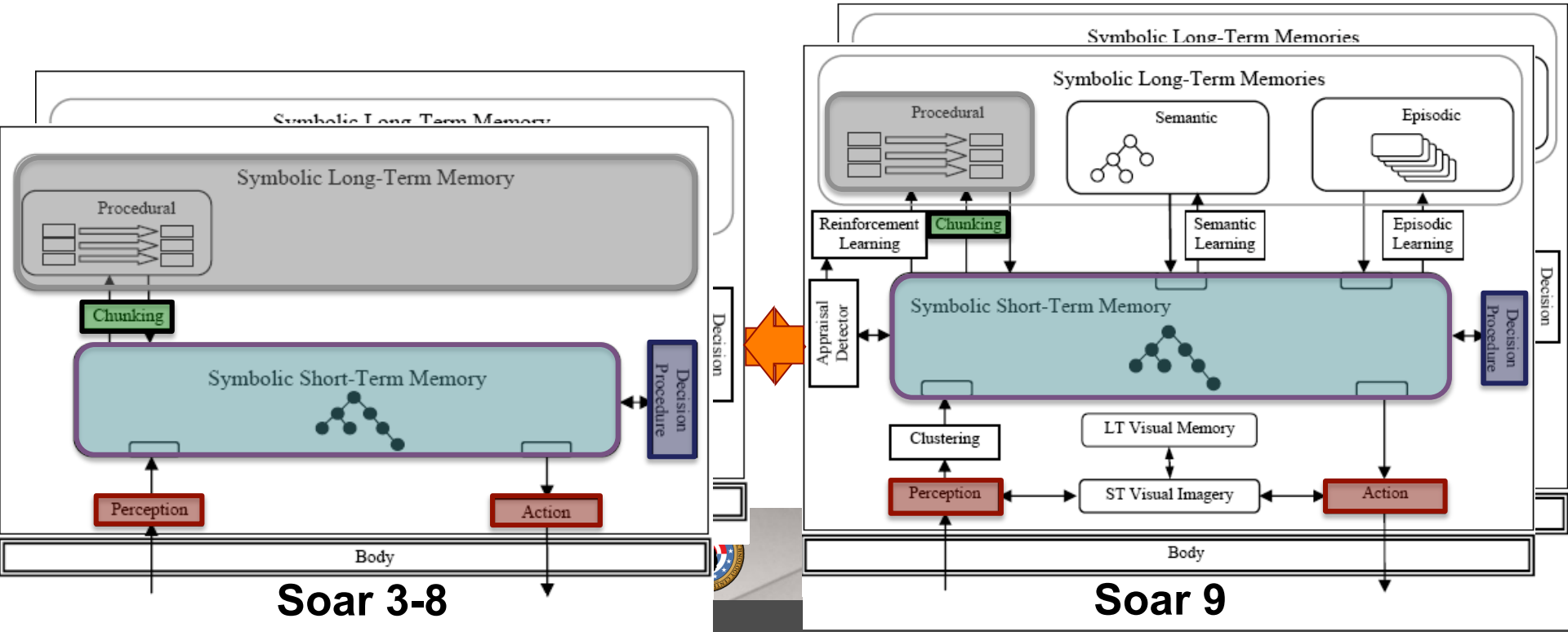
- **Data driven (*shallow*)**
  - *Simple statistical architecture in combination with large amounts of (uncertain) data/knowledge*
  - Achieves robustness through breadth of data/knowledge and focus on statistical regularities
- **Model based (*deep*)**
  - *Sophisticated symbolic reasoning architecture in combination with articulated models of domains of interest*
  - Achieves robustness via combinatoric flexibility of first principles reasoning over comprehensive models
- **The ideal solution is a *mixed* approach**
  - Probabilistic (statistical) + symbolic (relational)
  - Each provides strengths, and can counterbalance other's weaknesses, in authoring, learning, reasoning, perception, etc.
  - Multiplicative effect on robustness

Desideratum II
Tightly Integrated

- **Within the cognitive system**
  - Typical focus within work on cognitive architecture
- **Between cognition and perceptuomotor system**
  - Needed for virtual humans, intelligent robots, etc.
  - Implies *hybrid* systems
    - Combining discrete and continuous representations and processing
  - Also can benefit from mixed systems
    - Supporting general reasoning in presence of uncertainty

Desideratum III
Functionally Elegant

- **Broad scope of capability and applicability**
  - Embodying a superset of existing VH capabilities (cognitive, perceptuomotor, emotive, social, adaptive, …)
- **Theoretically elegant, maintainable, extendible**



Summary of Desiderata

- **Broadly and incrementally functional**
- **Theoretically elegant and simple for simple things**
- **Mixed and hybrid**
- **Supporting truly robust systems**
- **Maintainable and extendible**

Dichotomies Faced by Cognitive Architectures

- Data-driven versus model-based
- Probabilistic versus logical
- Central versus peripheral
- Discrete versus continuous
- Uniform versus diverse
- Explicit versus implicit
- Symbolic versus subsymbolic/neural
- Procedural versus declarative
- Goal-based versus utility-based
- Reactive versus deliberative
- …

Resolving Dichotomies

- **Choose a side**
  - Can work for some, particularly until challenges get too diverse
  - But usually inadequate over the long run
- **Bridge the dichotomy**
  - *Addition*: Add a box for each side
    - Yields two points on broad spectrum, but not full spectrum
    - Neutral on tight integration
    - Supports functional side of functional elegance, but not elegance
  - *Reduction*: Extract commonality that exists across dichotomy
    - Can yield full spectrum
    - Can provide leverage in tight integration based on what is shared
    - Can add elegance to functionality

    + *May uncover deep scientific results*

    - *May require compromise or yield residual*

Reduction Methods

- **Create generalization that subsumes both sides**
  - Markov logic yields a generalization over logic and probability
    - Also generalizes over other dichotomies
  - Traditional shallow rule systems can be thought of as generalizing over data-driven and model-based
    - Compromises both ends of dichotomy for simplicity and efficiency
- **Implement one side via other**
  - Soar implements deliberation via reactivity (plus decision proc.)
  - Data chunking tried to implement declarative via procedural
  - Graphical architecture implements diversity via uniformity
  - *Requires level/time-scale difference and non-peer integration*
- **Generalize implementation level beneath dichotomy**
  - Factor graphs implement both procedural and declarative

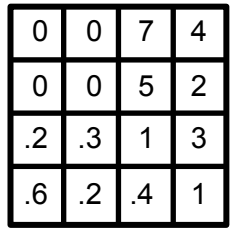
Techniques

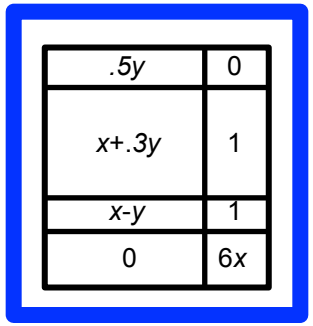
- **Piecewise continuous functions**
  - *Subsumption generalization* for representational primitives
    - N-ary predicates become N-dimensional functions
  - Embodies aspects of both discrete and continuous functions
    - Exact for discrete and symbolic functions
    - Can represent some continuous functions exactly and approximate others as closely as needed
- **Factor graphs w/ summary product algorithm**
  - *Implementation generalization* for complex reps. and processing
  - Generalizes over algorithms underlying many capabilities
    - Implement memories, decisions, etc.

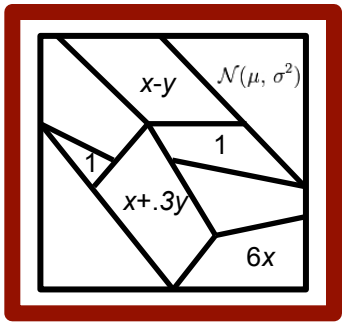
Both are relevant to bridging all listed dichotomies

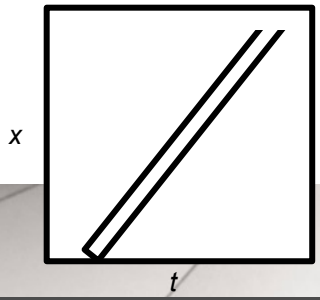
Space of Piecewise Continuous Functions

- **Types of regions**
  - Hypercubes (squares, cubes, etc.)
  - Hyperrectangles/orthotopes (rectangles, etc.)
  - Polytopes (polygons, etc.)
- **Types of functions over regions**
  - Constant, linear, polynomial, exponential, Gaussian, wavelet, ...
- **Additional sources of variation**
  - Axially aligned or not (for hypercubes/orthotopes)
  - Totally explicit or inactive regions implicit
  - Local borders or space-wide slices









USC

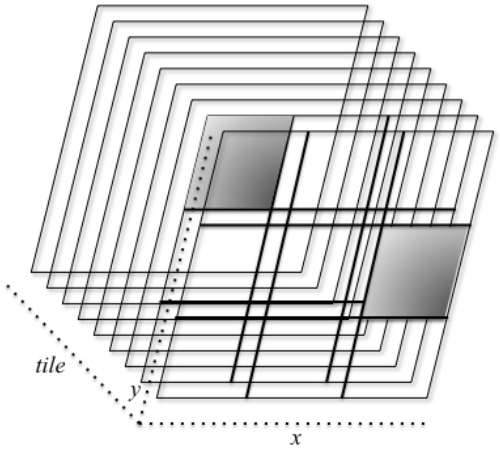
ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

Examples

- **Working memory**
  - `(O1 ^color Green) (O2 ^color Yellow) (O3 ^color Yellow) (O4 ^color Red)`

	Red	Green	Yellow	Blue
O1	0	1	0	0
O2	0	0	1	0
O3				
O4	1	0	0	0

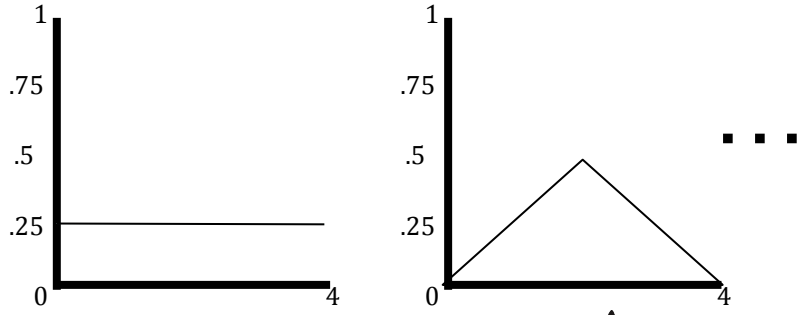
- **Mental imagery**



- **Episodic memory**

operator \ time	1	2	3	4	5
Left	1	0			0
Right	0	0			0
Up					
Down	0	1			0

- **Probability densities**



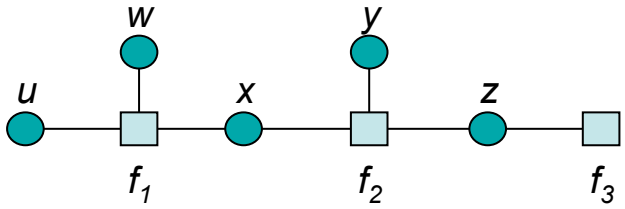
USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

Factor Graphs w/ Summary Product

- **Factor graphs are the most expressive form of GM**
  - More complex rep. + inference

$$f(u,w,x,y,z) = f_1(u,w,x)f_2(x,y,z)f_3(z)$$

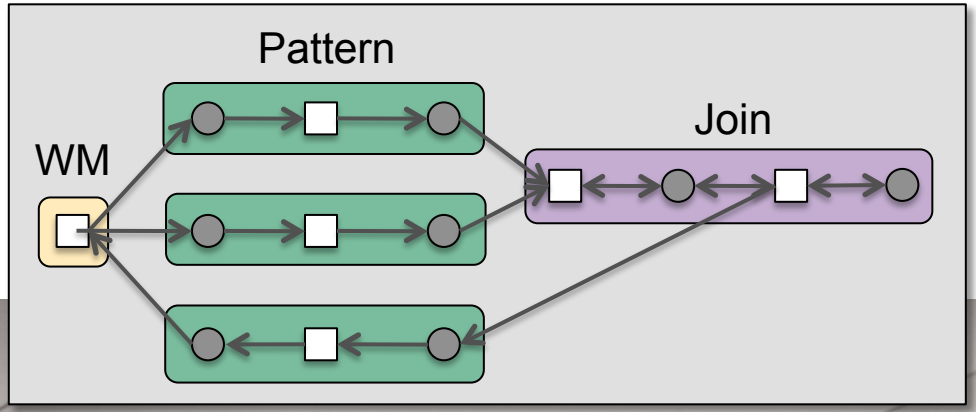


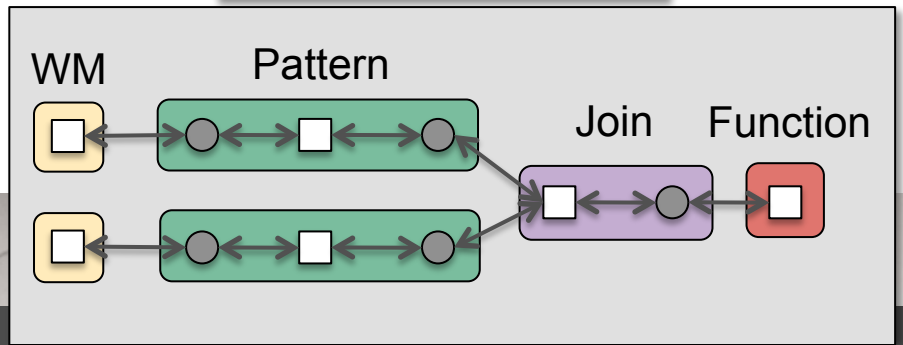
- **Summary product processes messages on links**
- **Implements a generalized conditional language**

```
CONDITIONAL Transitive
    conditions: (Next ob1:a ob2:b)
                (Next ob1:b ob2:c)
    actions:    (Next ob1:a ob2:c)
```

```
CONDITIONAL Concept-Weight
    condacts:  (concept object:O1 class:c)
               (weight object:O1 value:w)
    function:
```

w\c	Walker	Table	…
[1,10>	.01w	.001w	…
[10,20>	.2-.01w	"	…
[20,50>	0	.025-.00025w	…
[50,100>	"	"	…





Some Recent Results

- **Decision making**
- **Mental imagery**
- **Episodic learning**
- **Statistical question answering**
- ***Prediction-based supervised learning***

Decision Making

- **Preferences encoded via actions and functions**

```
CONDITIONAL goal-best ; Prefer operator that moves a tile into its desired location
     :conditions (blank state:s cell:cb)
                 (acceptable state:s operator:ct)
                 (location cell:ct tile:t)
                 (goal cell:cb tile:t)
     :actions (selected states operator:ct)
     :function 10

CONDITIONAL previous-reject ; Reject previously moved operator
     :conditions (acceptable state:s operator:ct)
                 (previous state:s operator:ct)
     :actions (selected - state:s operator:ct)
```

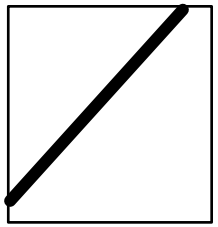
- **Most processing happens in graph via SP algorithm**
- **Complete implementation of Eight Puzzle**
  - 747 nodes (404 variable, 343 factor)
  - Solves a simple problem in 9 decisions
    - 113 messages/decision, 2.5 seconds/decision
- **Also initial implementation of reflection, but slow(er)**

USC

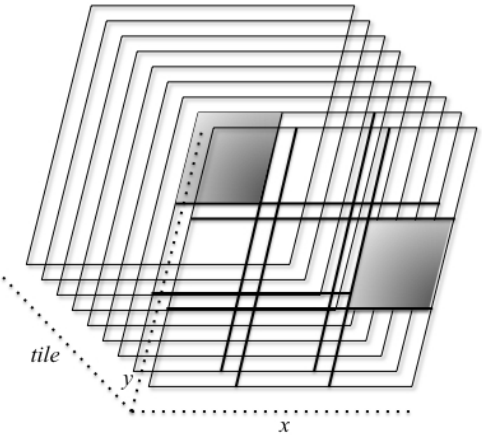
INSTITUTE FOR CREATIVE TECHNOLOGIES

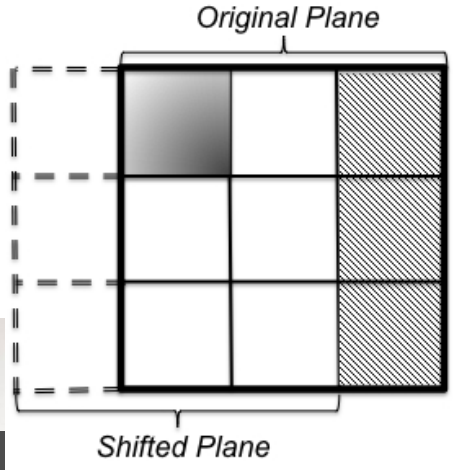
Mental Imagery

- **Beginnings of mental imagery**
  - 2D imagery with translation operation
- **Translation requires an angled, shifted delta function**
  - Need extended functional form for efficiency in uniform rep.
  - Implemented a special purpose optimization: *offset* factors
    - Also currently important in reflection and may be relevant to EM
- **Need 3D, time, scaling, rotation, …**
- **Need more focus on predicate extraction**



```
CONDITIONAL Move-Right
   :conditions (selected state:s operator:o)
               (operator id:o state:s x:x y:y)
               (board state:s x:x y:y tile:t)
               (board state:s x:x+1 y:y tile:0)
   :actions (board state:s x:x+1 y:y tile:t)
            (board – state:s x:x y:y tile:t)
            (board state:s x:x y:y tile:0)
            (board – state:s x:x+1 y:y tile:0)
```





Funded by AFOSR/AOARD

Episodic Learning

- **Initialize LTM with a temporal prior and an EM conditional for each predicate that includes state**

```
CONDITIONAL Time
 Condacts: (Time value:t)
 Function: [1,2) — .6667t
```

```
CONDITIONAL Time-Selected
 Condacts: (Time value:t)
           (Selected state:0 operator:op)
 Function: [1,∞)×[Left,Right,Up,Down] — 1
```

- **History of top-level state in WM is recorded in temporal slices of functions in EM conditionals**

operator

	1	2	3	4	5
Left	1	0			0
Right	0	0			0
Up					
Down	0	1			0

time

Final region extends to ∞, implicitly extrapolating to future

- **Scope & slope of temporal prior updated each cycle**

```
Function: [1,5) — .0833t
```

- **Retrieve best previous state given cues by SP/max**

USC

ICT INSTITUTE FOR CREATIVE TECHNOLOGIES

Statistical Question Answering

- **The NPCEditor learns to choose appropriate answers to questions from statistics gathered over pairs of questions and answers**
  - Also has additional dialogue components that can affect choice
- **Implemented Bayesian computation of language model of answers given question**

$$P(a,Q) = \sum_{s} \varphi(a,s) \prod_{i} \phi(s,q_i)$$

  - Compiled sentence-pair statistics into semantic memory
  - Can be used directly to choose best answer
- **Extending to full Kullback-Liebler divergence**

$$D\big(P(Q)||P(A)\big) = \sum_{a} P(a|Q) \log \frac{P(a|Q)}{\pi_A(a)}$$

- **Also looking to further extend capabilities and run scale-up experiments**

Plans

- **Continue with mental imagery**
  - Including extending function representation
- **Pervasive prediction**
  - Decisions choose next operator *and* predict next situation
  - Support perception, understanding, learning, appraisal, …
- **Implement more complete learning capability**
  - Based on predictions, actuals and dependencies
  - Chunking, reinforcement, supervised and unsupervised
- **Pursue further capabilities**
  - Theory of Mind, behavior understanding, speech and natural language, perceptuomotor behavior (SLAM), …
- **Evaluate, optimize and apply architecture**

Gold

- On path to bridge dichotomies
- Decisions, reflection and beginnings of imagery with little additional code
  - Continued promise of *functional elegance*
  - Step towards *tight integration*
- Getting experience with data-driven statistical processing
  - A significant step towards *broad spectrum*
- First bit of learning
- Lots of exciting projects starting

Coal

- Still little learning and no true perception
- Function representation needs significant rethinking
- Speed of code becoming an issue