Wide-Area Knowledge Representation: A Foundation for The Noosphere

John C. Mallery

Artificial Intelligence Laboratory

Massachusetts Institute of Technology

Overview

- *Annotation, Indexation & Retrieval*
- *Distributed Knowledge Representation*
- *Constraint-Guided Systems*
- *Ternary Relation Representations*
  - *Bi-Directional Ternary Relations*
  - *Self-Indexation*
  - *Completeness*
  - *Time & Space Complexity*
  - *Equality*

Background

- Natural Language Understanding
  - Syntactic Parsing & Generation
  - Semantic Representation
- Open Meeting System
  - Hypertext based on Persistent Ternary Relations
  - Light-Weight Semantics
  - Taxonomic Information Access
  - Email & Web Form Processing Interfaces

Annotation, Indexation & Retrieval

- Information Access
  - Current, Complete, Correct, Felicitous
- Challenges
  - Hierarchical Walkers take Geologic Time
  - Resources & Data Streams are Opaque
  - High Latency, Intermittent Availability
- Solution
  - Assertional Substrate with Caching
  - Indexation in the Infrastructure

Distributed Knowledge Representation

- Interlinked Assertions
- Build on Distributed Object System
- Terms Resolved by Low Latency Name Service
- Relations Served by Owners
- Contexts for Different Link Semantics
- Infrastructure Caches
  - Indices
  - Predication Information

Constraint-Guided Systems

- Reference System
  - Complete, Efficient Graph Matching
- Semantic Inversion
  - Induction of Constraint Descriptions
  - Graphical Displays
  - Sentence Generation
- Lexical Classification
  - Hierarchical Recognition

Ternary Relation Representations

- *First Class Relations*
- *Bi-Directional Interpretation*
- *Pervasive Symmetries*
- *Arbitrarily Expressive*
- *Encodes & Decodes Natural Language*
- *No Epistemic Pre-commitments*

Bi-Directional Ternary Relation

- *Components*
  - *Subject*
  - *Relation Type*
  - *Object*
- *Constraints*
  - *Subject Relation <--> Object Relation*
  - *Subject <-> Object*
  - *Relation Type*

Self-Indexation

- Ternary Relations
  - Determinate Inverse
  - Objects Cache Relation Type Count
- Token Universals
  - Track Token Instances
  - Maintain Token Instance Count
- Context Unions
  - Abstract Space & Time

Completeness

- *All First Order Structures Indexed*
  - *Enter via Token Universal*
  - *Traversal via Bi-Directional Relations*
- *No Holes in Index*
  - *Representation is the Index*
- *Extending Indexation*
  - *By New Assertions*
  - *By Lexical Classification*

Time & Space Efficiency

- Time Complexity for Reference
  - Smallest Accessible Possibility Space
  - Constraint Ordering
    - By Complexity Class
    - By Bounded Possibility Count
  - Reformulation of Reference Specifications
- Space Complexity ~ = 2x Relations
  - Actual Relation
  - Instance Relation to Token Universal

Equality

- *Lexicalist, Surface Semantics*
- *Meaning Postulates Map Equivalences*
- *Context Organize*
  - *Regions Conforming to Specific Semantics*
  - *Temporal Structures*
  - *Belief Structures (e.g.., counterfactuals)*
  - *Semantic Transfer Postulates*

Middleware Requirements

- Low Latency Name Service
- Exporting Index Data
  - Token Instance Counts
  - Relation Inverse Counts
- Hierarchical Caching
  - Down to address space level
- Meaning Negotiation Protocol

Middleware Services

- Reference (Name Service)
  - by Constraints
  - by Generating Constraints from an Example
- Walking Structures under Constraints
  - Mapping Actions over Network Regions
  - Cache Prefetch
- Notification, Alerting, Pattern-Action Rules
- Translation Across Semantic Contexts

Conclusions

- Ternary Relations Preferred
  - Simplest Invertible Representation
  - Complete Self-Indexation
  - Excellent Computational Properties
  - Arbitrarily Expressive
  - No Epistemic Pre-commitment
- Simple Initial Usage as Property Lists
- Incremental Evolution of Services

Epilogue

- *Differentiate the implementation from the representation*
- *Name service is the starting point for assertional substrate.*
- *Constraint-guided retrieval*