When Semantic Language Resources Meet Cognitive Systems

Saturday 31 May 2014

Presenters:

Katerina Pastra and Dimitris Mavroeidis

Tutorial Overview

Semantic language resources are increasingly being used beyond language technology applications to computer vision ones (e.g. large scale object recognition in Images-augmented WordNet, ImageNet) and cognitive robotics (for verbal interaction with humans and for verbalisation of visual scenes). This is the modern manifestation of a long-standing quest in Artificial Intelligence, regarding the integration of language with other modalities (images, gestures, body movements), or to put it more generally, the integration of symbolic and sensorimotor representations. Multimedia ontologies, collections of labelled images or video keyframes and knowledge-bases have appeared in different strands of Artificial Intelligence (AI) research. The automatic correlation of language and the denoted sensorimotor experiences has been a major challenge which is commonly known as the Semantic Gap problem.

On the other hand, there is growing experimental evidence that language is tightly related to perception and action. From Quillian's view of semantic memory as a lexical network accessed through a spreading activation of knowledge, modern neuroscience provides new evidence on the structure of semantic memory and points to the fact that semantic information is multisensory, multimodal and distributed. Intelligent multimedia systems, become more and more informed by experimental research on how the human brain works, with the aspiration that a simulation or transfer of mechanisms from the human brain to artificial agents will be more promising in terms of scalability and generalisation. In such research landscape, semantic language resources need to inform and be informed systematically by Cognitive Systems Research.

This tutorial aims to provide a comprehensive overview of semantic language resources, from a new, interdisciplinary perspective: that of cognitive science. In doing so, the tutorial will relate semantic language resources with the evolving field of Cognitive Systems, pointing to needs, challenges and future directions of research. Furthermore, it will familiarise the audience with new types of semantic resources that integrate language with vision and action, i.e. resources that correlate language with images, and motoric representations of actions. The cognitive underpinnings of semantic language resources and their integration with non-verbal modalities will be ellaborated through reference to the latest theories and experimental findings on how the human semantic memory works. A case study of a multimodal semantic network for cognitive systems will be presented (the PRAXICON), whose structure is corroborated by experimental findings on how the human brain works and a practical, hands on experience with the resource will be provided to the participants.

Tutorial Description - Outline

In the first part of the tutorial, we will position semantic language resources within intelligent multimedia systems and cognitive systems, elaborating on their current and potential contribution and presenting the challenges one faces in employing them in cognitive robotics, cognitive vision, and other intelligent multimedia system applications.

In the second part, we will give an overview of state-of-the-art semantic language resources, ranging from computational semantic lexicons to common-sense knowledge-bases. We will provide a comparative view of a number of semantic language resources that will comprise:

- profiling of the resources (developers, dates, languages involved, size, interfaces, links to other resources, applications)
- methodology used for their development, and
- contents: semantic relations covered (ranging from lexical semantic relations to conceptual relations such as temporal inclusion, cause, effect, goal, entailment), inclusion of facts or common sense assertions, instance vs. class distinctions, terms, domain, affect, word sense distinctions, figurative language coverage, links to Ontologies.

Furthermore, verbal and non-verbal information coupling in semantic language resources for addressing the different challenges in Cognitive Systems research will be presented. This coupling goes beyond labelled image collections (e.g. the Pascal Images Database), small scale labelled motion capture databases, multimedia ontologies, multisensory and multimedia corpora (e.g. the POETICON corpus) and has taken the form of an extension of known semantic language resources (e.g. the ImageNet resource which couples an image database with WordNet).

In the third part of the tutorial, we will present the cognitive underpinnings of semantic resources, starting from Quillian's lexical semantic networks and the underlying model on how semantic memory works, to state of the art theories and experimental findings on the structure and contents of semantic memory. The neuroscience perspective will point to directions in developing semantic resources for cognitive agents, which has been materialized through the PRAXICON, a multisensory semantic network. A live demonstration of the PRAXICON and a hands-on training session will conclude the tutorial.

Part I. Introduction to Cognitive Systems from a Language Perspective

- From Intelligent Systems to Multimedia Systems, to Cognitive Systems
- Applications and Needs
- The role of Semantic Language Resources in Cognitive Systems
- The Semantic Gap Problem

Part II Profiling Semantic Language Resources from a Cognitive Perspective

- Types (Semantic Lexica, Common Sense Knowledge Bases, Ontologies)
- Methodologies used for their development
- Contents: focus on semantic relations
- Extension trends & Cross-Resource Interfacing trends

- Verbal and Non-verbal Symbiosis in Semantic Resources

Part III. The Cognitive Underpinnings of Semantic Resources

- From Semantic Networks to Semantic Memory
- How can Neuroscience inform semantic language and/or multimodal resource development?
- A case study & hands-on exploration of a computational semantic memory for cognitive systems: The PRAXICON

When Semantic Language Resources Meet Cognitive Systems

Katerina Pastra and Dimitris Mavroeidis

Cognitive Systems Research Institute (CSRI) &
Institute for Language and Speech Processing (ILSP)
Athens, Greece

Tutorial Outline and Schedule

9:00 – 9:30	Introduction to Cognitive Systems • From Intelligent Multimodal Systems to Cognitive Systems • Applications, Needs and Challenges • The role of Semantic Language Resources (SLRs)
9:30 – 10:30	Profiling SLRs from a Cognitive Perspective • Types – Methodologies – Contents – Trends - Interfacing
10:30 – 11:00	Coffee Break
11:00 – 11:15	Verbal and non-Verbal Symbiosis in SLRs
11:15 - 12:30	Cognitive Underpinnings of SLRs • From Semantic Networks to Semantic Memories • How could Neuroscience inform SLR development? • Case Study: The PRAXICON
12:30 - 13:00	The PRAXICON – hands on session

Introduction to Cognitive Systems

- From Intelligent Multimodal Systems to Cognitive ones
- Applications, Needs and Challenges
  → the Semantic Gap problem
  → the Symbol Grounding problem
- The role of Semantic Language Resources (SLRs)
  → reaching towards Perception and Action

The AI quest for … Intelligence

- The two-fold objectives of Artificial Intelligence (AI):

a) The Engineering Objective:
construction of machines that do *intelligent* things

b) The Cognitive Objective:
use of computational modeling for studying the human brain (mental faculties)

Note the interrelation: the definition of intelligence and identification of mechanisms involved, determines the methodology to be followed in constructing an intelligent machine

Intelligence as approached by AI paradigms (1)

- Intelligence is achieved through

- operations on symbolic structures (Symbolic AI)

Related to Newell's and Simon's 1979 *physical symbol system hypothesis* that considered a symbolic system to be the necessary and sufficient condition for exhibiting intelligence (see review in Luger, 2002).

Explicit representations, search algorithms and heuristics for choosing among alternative solutions are all basic components of symbolic approaches. Strong AI approaches hold that a symbol system can provide a full account of intelligence regardless its implementation medium.

Intelligence as approached by AI paradigms (2)

Intelligence Mechanisms involve:

- Adaptation & learning

(Emergent or Biologically inspired AI, see review in Boden 1995)

intelligence emerges from dynamic patterns of activity and interaction with the real world

Cf. Connectionism, fuzzy logic, evolutionary computation…

Intelligence as approached by AI paradigms (4)

- Situated and Embodied AI

Essences of Intelligence (Brooks 1991, Brooks et al. 1998):
social interaction
sensorymotor experience
perceptual integration

Sensors and physical coupling of the machine with the world through interaction are the *sine qua non* features of an intelligent system, while representation amounts to the accumulation of the system states, which is "meaningless without interaction with the world" (Brooks et al. 1998).

Though essential for complex behaviours and tasks, symbols are just part of the intelligence story which needs to incorporate embodied AI notions too (Anderson, 2003; Chrisley, 2003)

The AI quest for…Intelligence

"We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to ***provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc.*** *Again I do not know what the right answer is, but I think both approaches should be tried."*

(Turing 1950, p.460). [emphasis not in the original text]

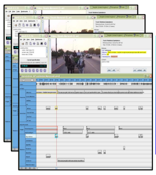
AI paradigm evolution and Cognitive Science

- From Cognitivism: Cognition as Information processing, symbolic computation--rule-based manipulation of symbols.

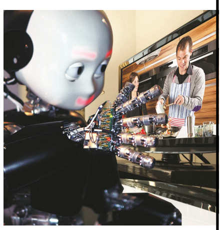
  - to Emergence: Cognition as the result of dynamic Interaction with the world

  - to Enactivism: (extending Situated Cognition, & Embodied Cognition); Cognition affects & is affected by sensorimotor interaction with the environment; knowledge is constructed this way

From Intellimedia to Cognitive Systems

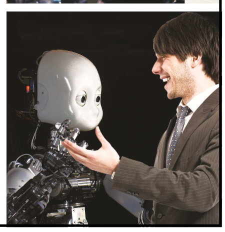


- from SHRDLU (Winograd ´72) to conversational robots of the new millennium (e.g. Roy et al. 2003)



→ diverse AI areas and applications in which a number of cognitive skills and abilities are needed and actually integrated, from Multimedia Information Retrieval to Robotics (Pastra and Wilks 2004), e.g.

- Audiovisual processing
- Human Machine/Robot Interaction
- Cognitive Robotics



Challenges

→ How does language relate to sensorimotor interaction with the world? What is its role in knowledge construction?

←Cf. the Symbol Grounding Problem (Harnad, 1990) and

←Cf. the Semantic Gap Problem (Hauptman, 2008)

Any Role for SLRs?

→ SLRs provide information on lexical concepts. Are they sufficient for representing embodied concepts?

If one was to bridge the semantic gap between sensorimotor experiences and language, and ground one to another, would state of the art SLRs be useful?

If not, what kind of changes would be necessary?

Issues in using SLRs in grounding

→What will grounded lemmas be like?
1-word? Multi-word? Word-centric? Sensorimotor representation – centric?

→ How will they be organised?

→ How specific or general should they be?

→ What kind of relations between entries/lemmas should be captured?

Profiling SLRs from a Cognitive Perspective

Types

Methodologies

Contents

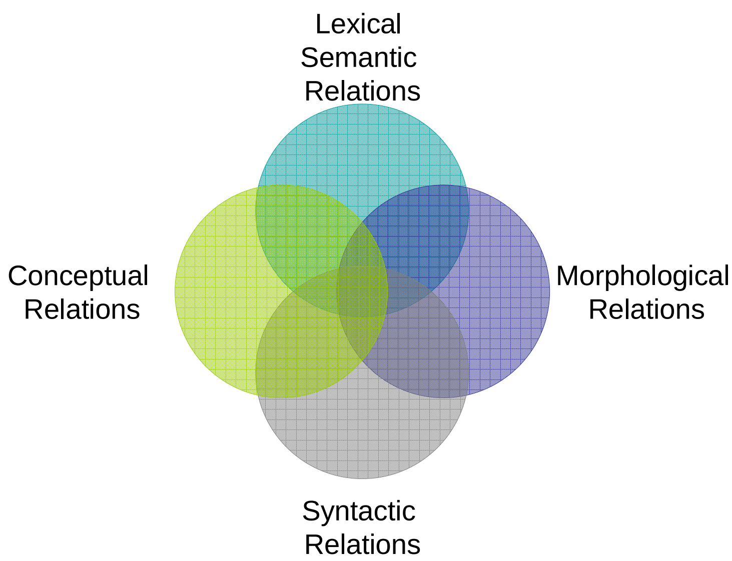
Trends

Interfacing

} Booklet with detailed tables per resource available at:

http://www.csri.gr/downloads

Types of SLRs



Types of SLRs (2)

- traditional dictionaries
- computational lexicons
- computational semantic lexicons
- common-sense knowledge bases (facts)
- ontologies and domain models

Categorization and story-telling…to learn / organise the world…

Profiling (1)

Resources	Main Developers	Institution	Date	Languages	Size	Interfaces
WordNet	George Miller, Christiane Fellbaum	Princeton University	started in 80s but documented in full in 1993 - still extended	EN	155.287 unique tokens - 117.659 synsets	Own Search interface: allows morpho stripping (base form of word stored) deals with irregularity. Treats mw expressions too. Many other e.g. JMWN lib to access all WNs (Pazienza:08)
EuroWordNet	Piek Vossen (coordinator)	Vrije University, Amsterdam	1996-1999, 2000 (final deliverable)	DE, NL, IT, ESP, FR, CZ, EST	Variable (7-44K synsets per language)	Pazienza above (map btw EWN and WN relations), Periscope (the project's) etc.
BalkaNet	D. Christodoulakis (coordinator)	Patras University, Greece	2001-2004	EL, BG, RO, TU, SR, CZ-extension	24K synsets (BG) approx.	VisDic (ML browser and editor), Cilix, Wordnet Management System (distributed network of servers)
EDR	T. Yokoi	Japan Key Technology Centre (+8 Manufacturers)	1986-1994 (2004 Final update of the Technical Guide)	Japanese, EN	Word-dictionary = 250K words (JP), 190K words (EN), Bilingual-dictionaries = 230K words (JP-EN), 190K words (EN-JP), Concept dictionary = 400K concepts, EDR corpus = 220K words (JP), 160K words (EN)	
SIMPLE	A. Lenci, N. Calzolari et al.	Uni of Piza (et al.)	1998-2000	Catalan, Danish, Dutch, EN, Finnish, FR, DE, GR, IT, Portuguese, ESP, Swedish	10K senses	ILSP interface for the Greek
PropBank	M. Palmer	University of Colorado Boulder	2002-2005 (main documentation)	EN	3300 verbs (4500 distinct framesets) (lexicon of verb syntactic framesets + annotated corpus - PennTreebank)	resource available from LDC

Profiling (2)

Resources	Main Developers	Institution	Date	Languages	Size	Interfaces
NomBank	A. Meyers	New York University	2004 - interim report, 2007 the release of version 1.0	EN	114.576 propositions annotated - 76.618 noun instances, no clear indication of number of unique common nouns included (lexicon of noun syntactic framesets + pointers to phrases in PennTreebank II Wall Street Journal)	resource available for downloading
VerbNet	K. Kipper	University of Colorado Boulder	approx. (1999 - main documentation 2005 and subsequent publications)	EN	4100 verb senses, Classification in Levin classes (191 first classes + 74 subclasses) and extended with another 55 classes by Kohronen and Briscoe 2004, and another 46 by Kohronen-Ryant unpublished	resource available for downloading and Unified Verb Index interface
FrameNet	C. Fillmore	Berkeley University of California	1996-2002 (two NSF grants - keeps being developed) - went through 3 releases	EN (but also other FrameNets are being developed for Spanish, Japanese, German)	10K word senses (6K fully annotated), 800 hierarchically related semantic frames, 135K example sentences annotated from BNC 100M word corpus, US newswire text from LDC and 2006 plans to use American National Corpus too	FrameSQL (Sato 2006) interfacing FrameNets in all languages
VerbOcean	T. Chklovski and P. Pantel	Uni of Southern California	2004 (paper documentation)	EN, DE (Regnieri - ppt only - VerbOzean - 4824 verb pairs - happens bf relation only)	29.165 Verb pairs	online interface
Mikrokosmos	S. Nirenburg	New Mexico State University	1993-1999	Spanish-Japanese-Russian-English Lexicons, corpus Spanish, Ontology common	Lexicon=20K words Spanish-35K distinct word senses, Ontology = 6K concepts, depth in ontology = 10 covers company merging and acquisition domain, 400 Spanish articles analysed	online exploration - broken link

Profiling (3)

Resources	Main Developers	Institution	Date	Languages	Size	Interfaces
Open Mind Project	P. Singh	MIT - Media Lab	2000-2004 pick, but still ongoing	EN, Brazilian OMCS corpus (2005-2007, 160K statements), GlobalMind Project 2006 to collect similar knowledge for Korean, Japanese, Chinese <-- users asked to translate also among these languages and English	Open Mind Common Sense Corpus (OMCS): 700K assertions, unedited contributions from 14K people (2004)	
Open Mind - ConceptNet	P. Singh	MIT - Media Lab	2004 documented, v3.0 (Havasi et al. 2007) - architecture related change, ConceptNet built on top of modular architecture that keeps data and processing/inferencing separate - stemming also added in NL module.	EN	OMCS net KB: assertions in binary relational format. 300K concepts, 1.6M links between them, 20 links per relation (~). 1.25M assertions dedicated to generic conceptual connections (K-Lines)	available online and resource can also be downloaded. API for searching the net allows to find path between nodes of network, get context (related concepts according to their distance from the source node and number and strenght of all paths that connect them), get analogous concepts (2 concept nodes are analogous if their incoming edges overlap / structural analogy e.g. apple/cherry, same properties e.g red, sweet and isA fruit - affective similarity predicted from propertyOf, isa, usedFor), projection (graph traversal from source node following single transitive relations e.g. locOf, isA, partOf, madeOf, subeventOf, effectOf), guessConcept (input novel concept in text and get potential analogies), guessMood (textual affect sensing some concepts classified into 6 emotions and then for any concept find paths to each of 6 categories and judge their strength and frequency). Accepts NL expressions (use of NLP system to process: MontyLingua: text normalisation, common sense informed POS, sem recogntion, chunking, shallow parsing, lemmatization, thematic role extraction, pr
Open Mind - LifeNet	P. Singh	MIT - Media Lab	2003 documented	EN	80K egocentric (1st person human experience) propositions inter-linked with 415K temporal and a-temporal Links e.g. I put on seat belt --> I drive the car.	

Profiling (4)

Resources	Main Developers	Institution	Date	Languages	Size	Interfaces
CYC	D. Lenat	CycCorp	1984-to date (more than 20 years of development, more than 900 person years of work)	EN	CyC Knowledge Base (328K concepts, 3.5M explicit assertions; in more detail: 16K Predicates, 11K Event types, 400 relations between Events-Participants, 120 types of Emotion 40 relations between Temporal Things, 50 Propositional Predicates) and Inference Engine (more than 700 specific reasoners), CyC EN lexicon (15K root words, 20K proper names, 43K entries in lexicon mapped to CYC ontology)	OpenCyc = CYC ontology, 47K concepts and 306K facts in KB open for all, ResearchCyc = as in OpenCyc + more facts for concepts, large EN lexicon, parsing and generation tools and Java based interfaces for knowledge editing and quering - available for research. Cyc-NL module translates NL queries into CyCL and the other way (complete, efficient NLP suite). Semantic Knowledge Source Integration (SKSI) provides a declarative mapping between external sources of structured knowledge and the Cyc Knowledge Base.
Open Mind - StoryNet	P. Singh	MIT - Media Lab	2004 documented	EN	Scripts of general situations (e.g. eating at restaurant) and only have sequence of events; info on roles, emotions etc not included, could be inferred from CN or LN.	
Open Mind - Indoor mobile robots corpus (OMICS)	R. Gupta	MIT - Media Lab	2004 documented	EN	Corpus of assertions about objects/properties/actions in home or office environment. 26K propositions (screened out of 29K), 400 users, 400 photos of objects	online interface
EventNet	H. Lieberman	MIT - Media Lab	2005	EN	temporal reasoning toolkit that uses LifeNet temporal links to predict events. Output: 10K nodes and 30K links - not fully connected graph	resource available for download

Methodologies

Resource	Language theory-based (manually crafted)	Corpus-based (auto or semi-auto extraction)	Experimental Psychology Findings
WordNet (also EWN, Balkanet etc)			
SIMPLE			
VerbOcean			
VerbNet			
FrameNet			
ConceptNet			
EventNet			
MindNet			
ThoughtTreasure			
Mikrokosmos			
CYC			

SLRs - Contents

Resource	Morphological Rel.	Syntactic Relations	Lexical Sem. Relations	Conceptual Relations	Facts/Common Sense Kn.
WordNet					
Parole -SIMPLE					
VerbOcean					
VerbNet					
FrameNet					
ConceptNet					
EventNet					
MindNet					
Thought Treasure					
Mikrokosmos					
CYC					

SLR Content Analysis Example: WordNet (WN)

Morphological Relations (MorphoSem)	**Syntactic Relations (incl. SyntacticoSem)**	**Lexical Semantics Relations**	**Conceptual Relations**	**Facts**
Derivational (2003) e.g. build-builder	Minimal subcatgz frames + “**thematic role” + selectional restriction like relations in** Derivational links (2007)	synonymy antonymy meronymy attribution	Temporal, cause etc.	Instances (distinguished as such in 2006)

Strict POS distinction in organisation

Organised in synsents and relations among them

Noun Categorization in WN

{*act, action, activity*}
{*animal, fauna*}
{*artifact*}
{*attribute, property*}
{*body, corpus*}
{*cognition, knowledge*}
{*communication*}
{*event, happening*}
{*feeling, emotion*}
{*food*}
{*group, collection*}
{*location, place*}
{*motive*}
{*natural object*}
{*natural phenomenon*}
{*person, human being*}
{*plant, flora*}
{*possession*}
{*process*}
{*quantity, amount*}
{*relation*}
{*shape*}
{*state, condition*}
{*substance*}
{*time*}

Where does this categorisation come from? Predication of nominal concepts studies

Noun/Entity Features in WN

Features in WN:

- perceptual features e.g. small, yellow, round

  → in glosses

- parts e.g. wings, legs etc.

  → meronymy relations

- affordances e.g. fly, sit etc.

  → in glosses (and multiple super-ordinates depending on structural or functional perspective e.g. ribbon-cloth, ribbon-adornment)

Adjective/Feature Organisation in WN

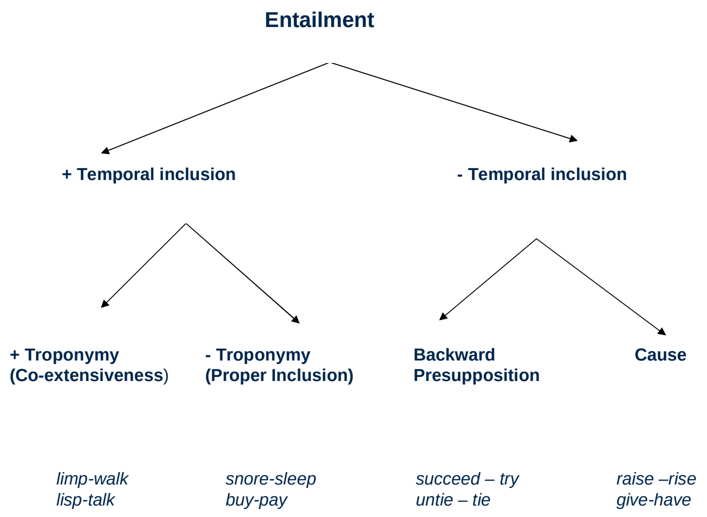
Organising principle: Antonymy

- descriptive
- reference modifying (old friend i.e. old friendship)
- social relation or function (e.g. presidential)
- temporal status (former)
- evaluative (e.g. good)
- action denoting (e.g. passive)
- epistemological (e.g. reputed) …

Exception: relational ADJ → e.g musical instrument
Linked to Nouns they are related to
(same concept – similar or different root)

Where does this categorisation come from? Word-Association Tests

Action/Verb Categorisation in WN



Action/Verb Categorization in WN

VERB classes 15:

- bodily care and functions (e.g. faint)
- **change** (e.g modify - diff subclasses of change e.g. change state, change shape, etc. + troponymy of these)
- cognition (e.g. judge)
- communication (e.g. beg)
- competition (e.g. campaign, fight)
- consumption (e.g drink)
- **contact** (troponyms of few base verbs : fasten, attach, cover, cut, touch, hold)
- creation (e.g. print, illuminate, shew…)
- emotion (e.g. fear)
- motion (make movement: e.g. shake, travel – locomotion e.g. run
- perception (e.g. watch)
- possession (e.g. hold, rip)
- social interaction (e.g. franchise)
- weather verbs (e.g rain)

=> states (suffice, belong, resemble – they share no sem props as others above they just refer to states – small sem clusters and org sim to adj)

Criteria for such categorization?

WN evolution line

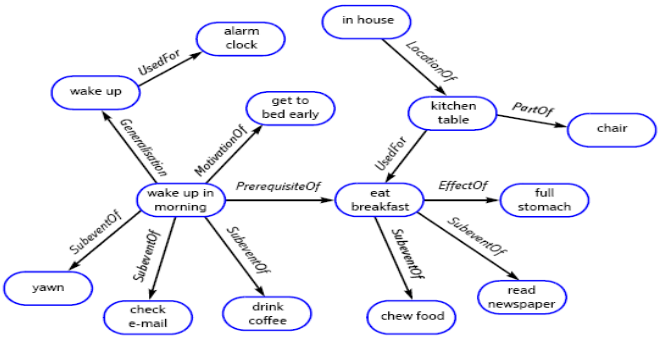
- initial aspiration to simulate how children acquire Language (how mental lexicon works)
- NLP applications (WSD, IR, etc.)
- Recently: entailment, emotion recognition…

- Yago (Suchanek 2007): fact inclusion in WN from Wikipedia

- WN Affect (Strapparava 2006): label WN affective synsets as ➔

Emotional (eg. anger), non-emotional affective e.g. mood, non-affective mental state (e.g. confusion), personality trait (e.g. competitive), behaviour (e.g. cry), attitudes (e.g. skepticism), physical or bodily states/feelings (e.g. pain, pleasure etc.)

Another Example: ConceptNet



Verbal and Non Verbal Symbiosis in LRs

- Types:

Multimedia Thesauri (e.g. Benitez et al. 2000)
Multimedia Ontologies (e.g. Zinger 2005 – OntoImage)
Multimedia Taxonomies (e.g. Hauptmann 2007 – LSCOM)
Multimedia Corpus (e.g. Pastra et al. 2010 – POETICON corpus)
Labeled Image Databases (see review in Torralba 2011)

- Long History

Ad hoc links of various types in AI systems since the late seventies (see review in Pastra and Wilks 2004)

Verbal and Non Verbal Symbiosis in SLRs

- Large scale object recognition using SLRs:

The ImageNet Case (www.image-net.org)

14+ Million Images manually indexed to ~ 21K WN Synsets

~ 150K Images have bounding box around the object of interest

Images linked to Synsets at any level of the taxonomy; inheritance applies.

Verbal and Non Verbal Symbiosis in SLRs (2)

- Dang et al. 2010:

Use of semantic hierarchies in Object recognition for:
→Going large scale
→Filtering visual similarity with semantic similarity
→Use hierarchical cost in miss-classification error metrics

- Russakovsky et al. 2010:

Extending WN noun synsets with visual attribute information (colour, shape etc.) → 384 synsets x 25 images per synset x 20 attributes annotated per image

The Cognitive Underpinnings of SLRs

- From Semantic Networks to Semantic Memories
  - → What is a Semantic Memory?
  - → Which applications need a Semantic Memory?
- How could Neuroscience inform SLR development?
  - → Some important findings
  - → The Minimalist Grammar of Action
- A Case-Study: The PRAXICON Semantic Memory
  - → The structure of the PRAXICON
  - → Concepts and relations in the PRAXICON
  - → Examples

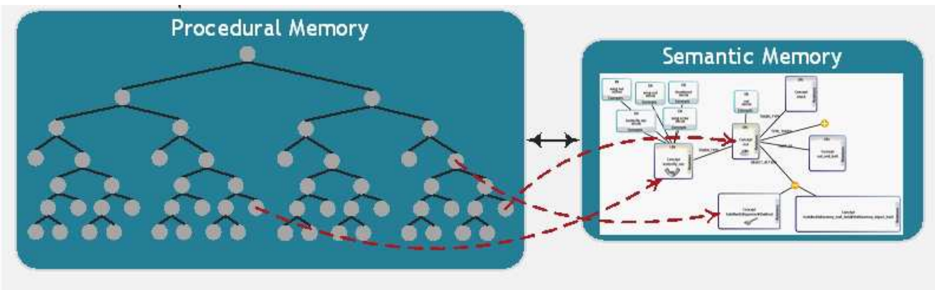
Semantic Memories

- Long term Memory (see Tulvig 1972)

  → episodic (tied to specific learning experiences)

  → semantic (general knowledge of the world, and related generalisation and reasoning abilities - see also Quillian 1968 on semantic networks)

  → procedural (related to single action & action sequence learning, created through repeated learning)



Semantic Memories (2)

- Issues

  → type of knowledge stored

  → structure of memory space

  → use/activations (in memory search, retrieval, decision making)

Theories on Semantic Memory

Many theoretical accounts on structure & neural basis of SM
(cf. extensive reviews in Kiefer and Pulvermueller 2012, McNorgan et al. 2011, Meteyard et al. 2012)

(1) Concepts are flexible, distributed representations; they comprise modality-specific conceptual features (latter stored in distinct sensorymotor brain areas) [Kiefer and Pulvermueller, in press]

(2) Much of the semantic memory content is related to perception and action and is represented in a brain region that overlaps with or corresponds to regions responsible for perception and action (Patterson et al. 2007)

Basic Level Categories (1)

- Verbal Categorization:

Basic Level = category of maximum information gain for similarity-based categorisation (category distinctive enough and homogeneous) (Rosch et al. 1978)

- *Most general categories whose members :*
  - *possess significant numbers of attributes in common*
  - *participate in common motor sequences*
  - *have similar shapes (identifiable from averaged shapes of members of the class)*
- *Most inclusive categories:*
  - *For which an image as a whole can be formed*

Basic Level Categories (2)

- Basic Level advantage = faster and more accurate categorisation at that level (Jolicoeur et al. 1984)

  → not confounded by:

  - word frequency, length of word, joint image-word frequency, order of word learning…

- Literature on basic level effects when recalling information from semantic memory (in healthy subjects and patients of e.g. semantic dementia, Alzheimer, PTSD);

  → theories on activation of concepts in semantic memory:

  e.g. Rogers and Patterson 2007

Rogers and Patterson 2007

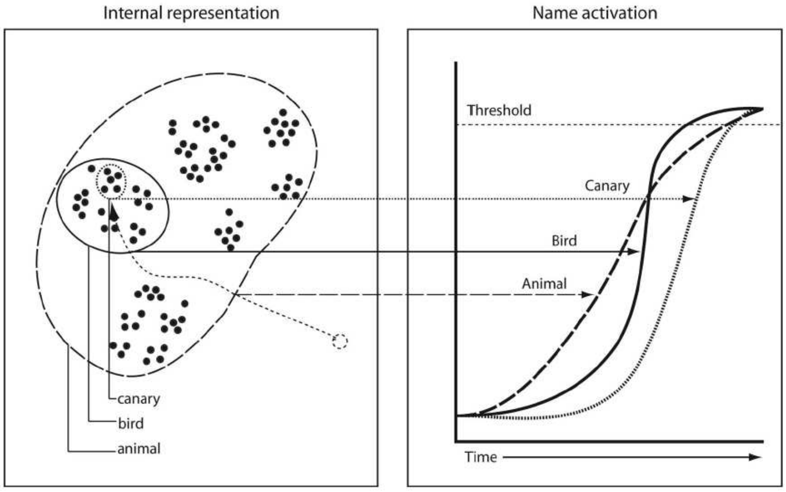


Figure 5. Schematic representation showing the basis for the parallel distributed processing theory's prediction about the time course of activation for names at different levels of specificity. When a visual stimulus appears, the model's semantic representation state begins to move from some neutral point toward the appropriate

PDP theory

The mechanism by which conceptual knowledge structure results in a basic-level advantage **depends upon the pattern of generalization fostered by conceptual representations as the network learns to <u>name</u>** (Rogers & McClelland, 2004).

(a) the more frequent the label, the more quickly and strongly it will become activated, all else being equal; and

(b) these frequency effects interact with the similarity structure of the semantic representations, so that (again, all else being equal) names are more slowly acquired and more difficult to activate when they apply across sets of items with very different representations, or there are items with different names that have very similar representations. Exemplars of basic-level categories are represented as similar to one another and as distinct from other items, and so basic-level names get the most benefit and the least interference from similarity-based generalization.

Theories on Semantic Memory (2)

How could it be implemented?

McClelland → neuroscience evidence suggests SM to be implemented as a separate memory not subsumed to episodic memory. Suggestion that hippocampal formation and the neocortex form complementary learning system. Former facilitates auto and hetero-associative learning which is used to reinstate and consolidate gradually learned info in the neocortex.

Semantic Memory & Language

Traditional representation of semantic knowledge through:

- **Semantic Networks (hierarchical or non)** (see Collins and Quillian 1969, Collins and Loftus 1975) and/or **Feature Bundles**

NOTE:

- **all such knowledge is represented through LANGUAGE only, and carries all idiosyncrasies of language…(i.e. the semantic gap to the sensorimotor space lurks behind these resources)**

Semantic Memory & Language (2)

A number of knowledge bases around (of different types):

- WordNet (hierarchical lexical resource) (Fellbaum 1998)
- Common sense knowledge bases (e.g. ConceptNet, CYC) etc.

A number of cognitive architectures with recently incorporated semantic memory modules:

- SOAR (Laird et al. 2009)
- ACT-R (Anderson et al. 2004)
- ICARUS (Langley 2009)

Semantic Memory & Language (3)

Common ASSUMPTION in such networks that agents have :

(a) sensorimotor experiences related directly or indirectly to what the language representations denote, and

(b) mechanisms for performing such link between language, perception and action

Aka: These modules/resources are NOT embodied, they are tied to language idiosyncrasies and lack structure that will unify language-perception-action.

Note: linking robots/intelligent systems to the web and interconnecting the knowledge they acquire through a cloud, can only be useful if…

Why Needed for Artificial Agents? (1)

Currently, robots have
episodic and procedural memory ONLY

ONE SHOT learning ← need for Generalisation

- Semantic memories (SM) in Robots usually generated directly by perceptual systems (for object/action recognition) ← reasoning?
- Sometimes indirectly present through association strength information in episodic memory

We envision: Self-exploration models for gathering information, input to episodic/procedural memory, and then updating of Semantic Memory → generalization

Why Needed for Artificial Agents? (2)

Currently, intelligent systems have disembodied semantic memories...

Link with Perception and Action (sensorimotor representations) will allow:

- their use in embodied cognition applications (robotics, human-robot interaction etc.) and large scale object/action recognition
- investigation of semantics and language (and in particular verbal categorization) from a cognitive perspective that may open up new directions in language research itself

Findings in Neuroscience

On the tight link between Language – Perception – Action:

- **Mirror neurons**: action perception and production activate the same brain areas
- **Visuomotor neurons**: visual object perception and action production tightly connected
- **Broca's area role**: the neural locus of (among others) language and action perception and production; suggestions for common syntactic (hierarchical and compositional) processes in language and action

Grammars for Action

- **Kirsch, 1964**: suggested a grammar of drawings analogous to text grammar;
- **Gregory 1974**: suggested grammar of vision analogous to language grammar;
- **Lashley 1974**: suggested that syntax may apply not only to language but also to other forms of behaviour, such as goal directed action

………

- **Fadiga 2005, 2009, 2011**: supramodal syntax hypothesis and experimental evidence that Broca's area is the neural locus of (among others) language and action perception and production; suggestions for common syntactic (hierarchical/dependency-based and compositional) processes in language and action

The minimalist grammar of action

Katerina Pastra and Yiannis Aloimonos (2012), *"The Minimalist Grammar of Action"*, Philosophical Transactions of the Royal Society B, 367(1585):103

The first **generative grammar of action** that employs the structure-building operations and principles of Chomsky's Minimalist Programme as a reference model

The grammar is based on a number of **basic findings in experimental research, and in that sense it has a biological basis**. It provides for an action-centric, embodied representation in SLRs.

Action Constituents (1)

Tool complement (t_c): the effector of a movement, this being a body part, a combination of body parts or the extension of a body part with a graspable object used as a tool. Syntactic feature.

Grasping with pliers vs. grasping with tweezers

Related Neuroscience Evidence:
Iriki 1996, Fadiga et al. 2000, Mantovani et al. 2011

Object complement (o_c): any object affected by a tool-use action. Syntactic Feature. E.g.. Confer Fadiga et al. 2000.

grasping a pencil with the hand vs. *grasping a glass with the hand*

Action Constituents (2)

Goal (g): the final purpose of an action sequence of any length or complexity. Inflectional feature!

1. *Same movement type, same tool and affected object, but different goal:*
   grasping a pencil in order to displace it vs. *grasping a pencil in order to write;*
2. *Same movement type, different tool or affected object, same goal:*
   grasping an apple to displace it vs *grasping a cube to displace it*
3. *Final goal of an action structure can be predicted from its first subactions*
   'extend hand towards pencil' (finger preshaping during hand transfer, cf. Jeannerod et al. 1995), (grasping neurons discharge before contact with object, cf. Fadiga et al. 2000, Fogassi et al. 2005), (Cattaneo et al. 2007)

The minimalist grammar of action (1)

Action Grammar Terminals: The simplest actions, i.e. perceptible movements carried out by an agent to achieve a goal, which have (one or more) body part tool-complements and no object-complements. Action terminals are further distinguished from each other through their perceptible motor features such as speed, force and direction

Action Grammar Non-Terminals: These are perceptible action phrases, that consist of action terminals (or other non-terminals) in certain temporal configuration; they may have both tool-complements and object complements. They involve interaction with objects beyond one's own body or with other agents, for attaining a particular goal/task

The minimalist grammar of action (2)

Production Rules	
4	A’’ → g A’
3	A’ → (m) A’
2	A’ → A’ (o_c)
1	A’ → A t_c

Effects/Results → the ‘static fingerprints’ of actions…

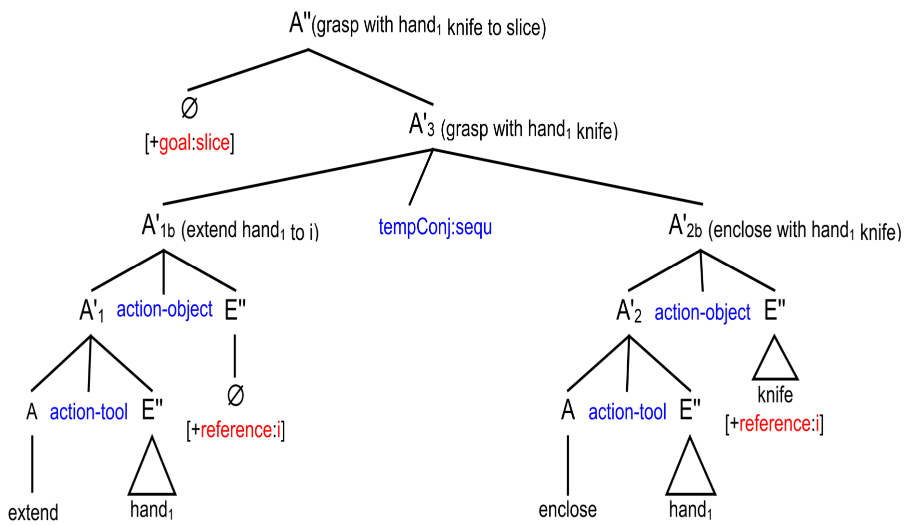
Features:

- **Tool** Complement (**t_c**)
- **Affected Object** Complement (o_c**)**
- **Physical Space** Modifier (m)
- **Goal** Modifier (g)

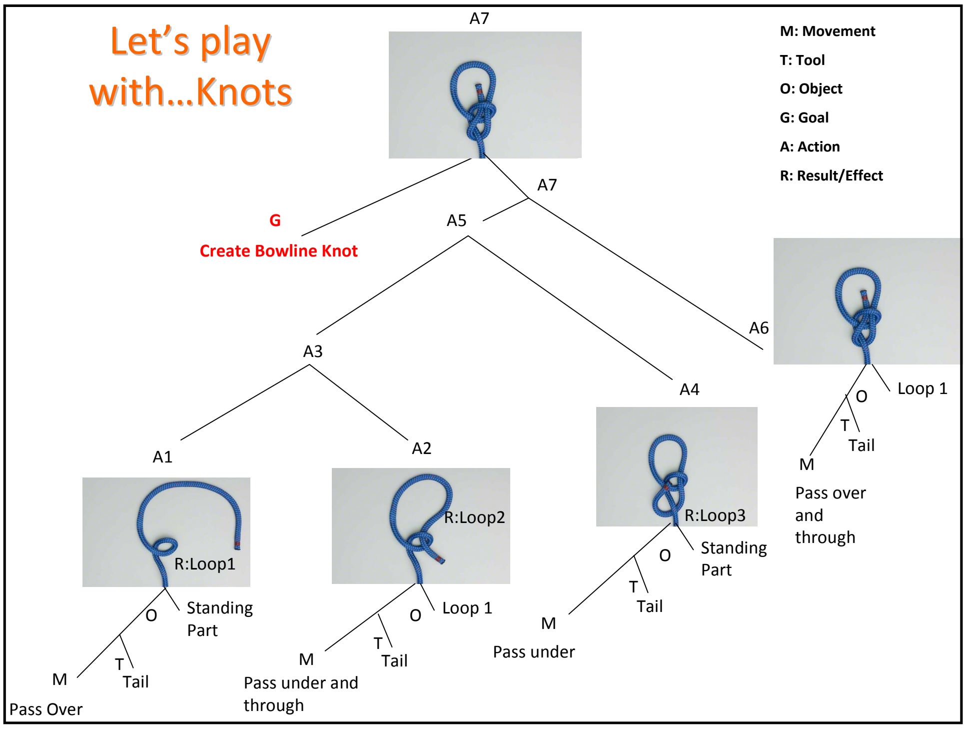
Minimalist operators driven by Features:
Merge and **Move**

The operators drive the application of the rules bottom-up

Action Grammar Example



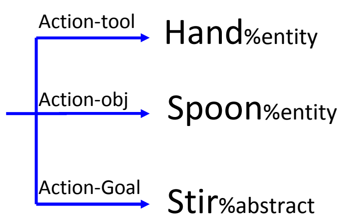
Let's play with…Knots



Action-Constituents

MOVEMENT_TOOL_AFFECTEDobject_GOAL

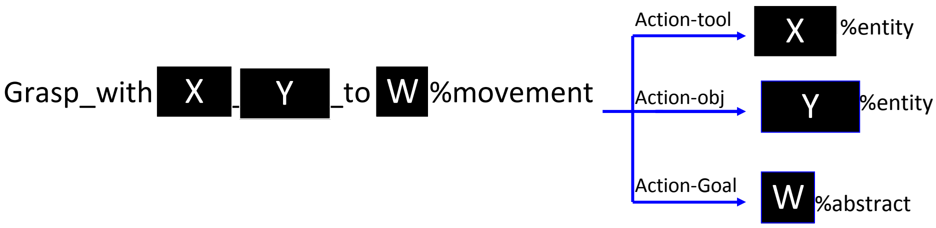
Grasp_with Hand_Spoon_toStir%movement



Movements are complex concepts: they have 3 **inherent constituents** that affect the execution of the motor program.

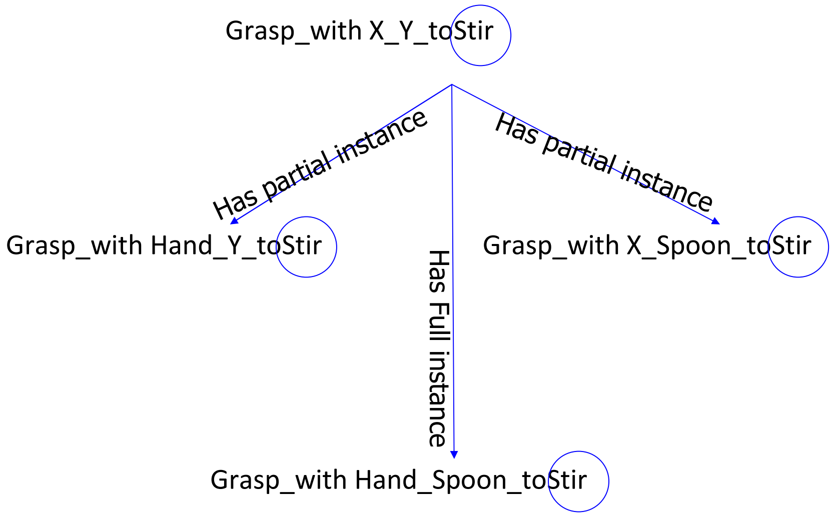
Conditions for distinguishing a new action concept

Motor Program Generators



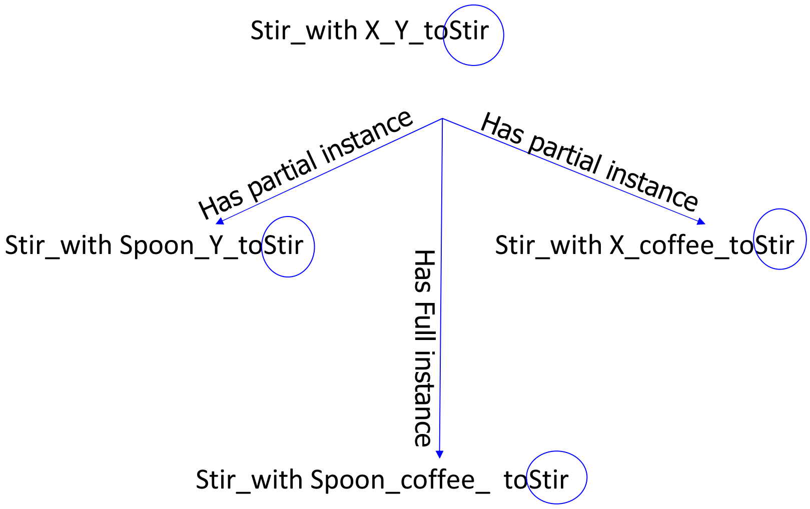
Action-related concepts

Family of generators



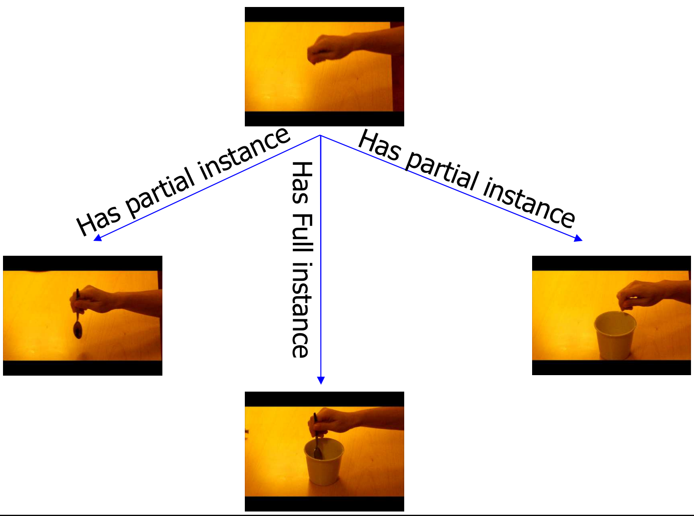
Action-related concepts

Family of generators (2)



Action-related concepts

Family of generators (3)



The PRAXICON

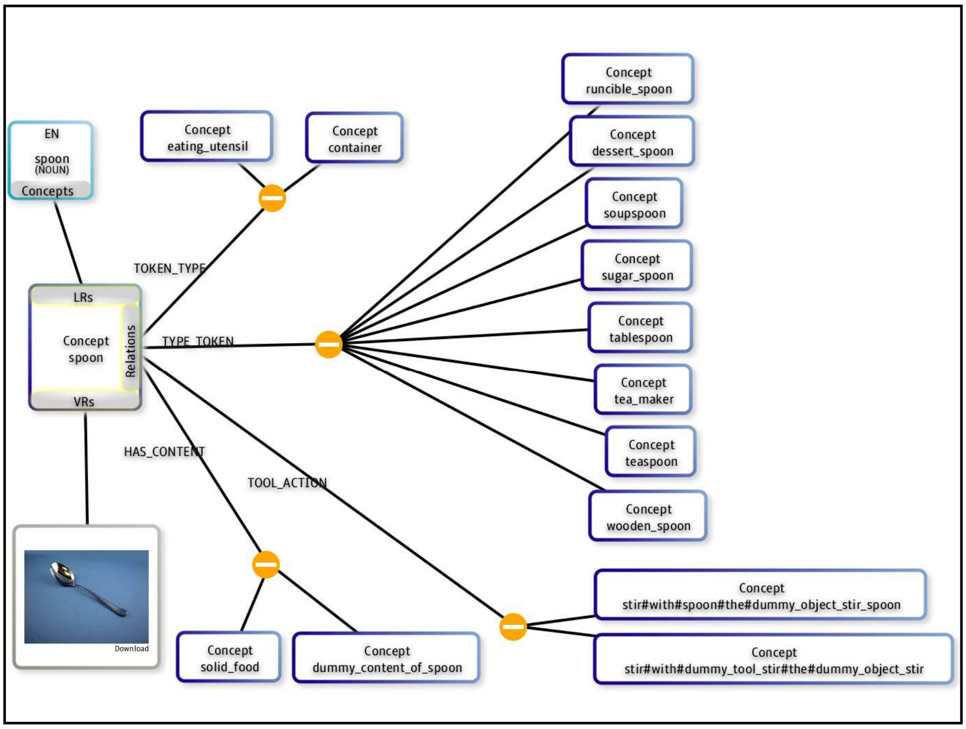
- PRAXICONs: From Liepman's (1908) input/output motor representations stored in memory, to…embodied-concept representations perceived and stored in memory for behaviour generation and understanding

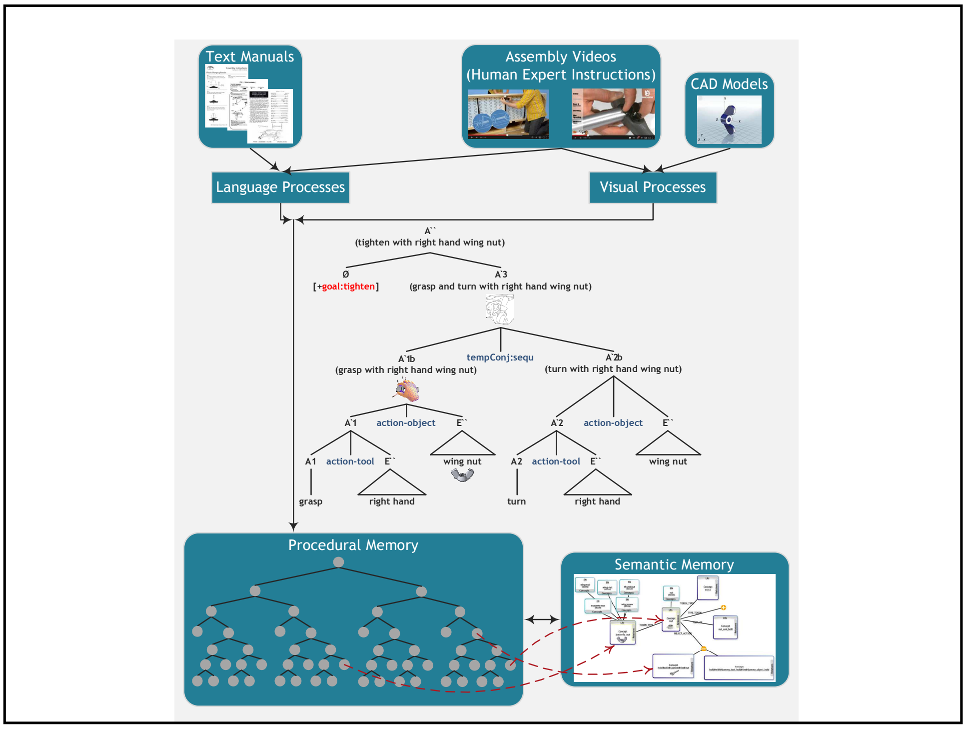
The PRAXICON is

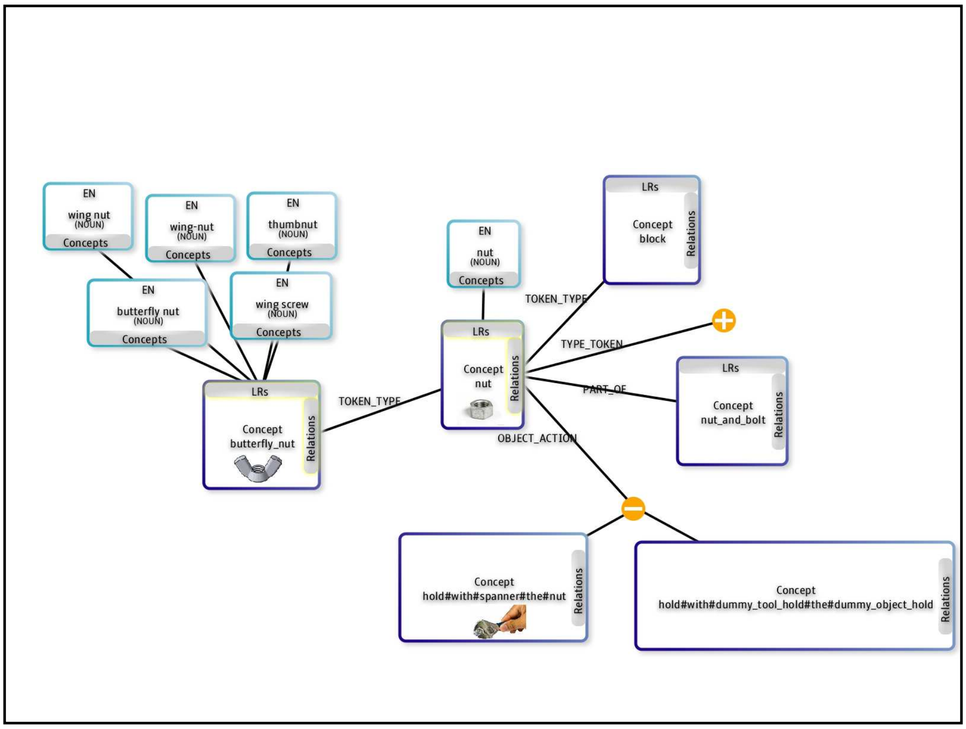
a) Action/Sensorimotor-centric SLR (Minimalist Grammar of Action used)

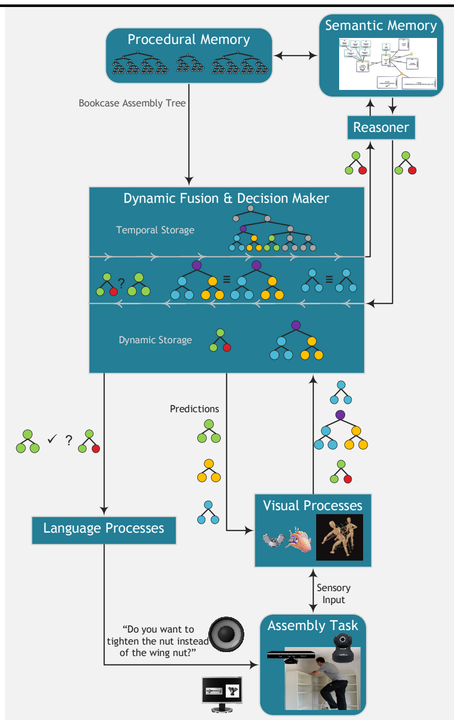
b) With Concept-Specificity indication (Basic Level Theory and first ever algorithm)

c) Driven by Neuroscience findings in all Knowledge Representation Decisions




Text Manuals
Assembly Videos
(Human Expert Instructions)
CAD Models
Language Processes
Visual Processes
A``
(tighten with right hand wing nut)
Ø
[+goal:tighten]
A`3
(grasp and turn with right hand wing nut)
A`1b
tempConj:sequ
A`2b
(grasp with right hand wing nut)
(turn with right hand wing nut)
A`1
action-object
E``
A`2
action-object
E``
A1
action-tool
E``
wing nut
A2
action-tool
E``
wing nut
grasp
right hand
turn
right hand
Procedural Memory
Semantic Memory


EN
wing nut
(NOUN)
Concepts
EN
wing-nut
(NOUN)
Concepts
EN
thumbnut
(NOUN)
Concepts
EN
butterfly nut
(NOUN)
Concepts
EN
wing screw
(NOUN)
Concepts
LRs
Concept
butterfly_nut
Relations
TOKEN_TYPE
EN
nut
(NOUN)
Concepts
LRs
Concept
nut
Relations
LRs
Concept
block
Relations
TOKEN_TYPE
TYPE_TOKEN
PART_OF
OBJECT_ACTION
LRs
Concept
nut_and_bolt
Relations
Concept
hold#with#spanner#the#nut
Relations
Concept
hold#with#dummy_tool_hold#the#dummy_object_hold
Relations



PRAXICON Structure (1)

- **Concepts** (nodes – multi-representational)
- **Relations** (edges – labeled, mostly bidirectional)

→ One concept may have many relations to many concepts BUT there is only one relation linking two specific concepts

→ Some relations are more important for a concept than others; they are denoted as '**inherent**' relations

PRAXICON Structure (2)

Concepts: Characteristics

TYPE: entity, movement, feature, abstract
STATUS: constant, variable, template
PRAGMATIC STATUS: literal, figurative
SPECIFICITY LEVEL: Basic Level, Superordinate, Subordinate

Abstract concepts – compare:
Poverty vs. Cutlery
Cutting instrument vs. knife vs. butterknife

PRAXICON Structure (2)

Relations: a finite set

ACTION_AGENT
ACTION_GOAL
ACTION_OBJECT
ACTION_RESULT
ACTION_TOOL
ASPECT_CONCEPT
COMPARED_WITH
ENABLES

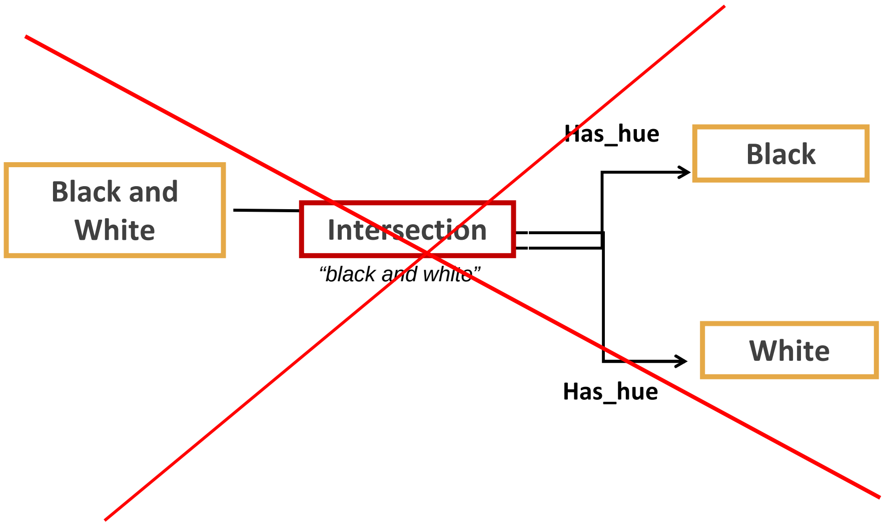
MORE
LESS
METAPHOR_OF
PRODUCER_OF
TYPE_TOKEN

HAS_ANTHROPOGENIC_EFFECT
HAS_COLOUR
HAS_CONDITION
HAS_CONTENT
HAS_DEPTH
HAS_FORCE
HAS_HEIGHT
HAS_HUE
HAS_INSTANCE
HAS_INTENSITY
HAS_LENGTH
HAS_LOCATION
HAS_LUMINANCE
HAS_MATERIAL

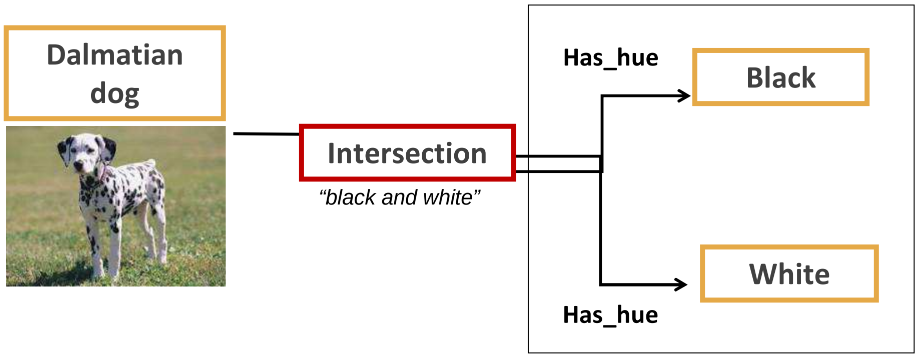
HAS_MEASUREMENT_UNIT
HAS_MEASUREMENT_VALUE
HAS_MEMBER
HAS_NATURAL_EFFECT
HAS_PART
HAS_PARTIAL_INSTANCE
HAS_SHAPE
HAS_SIZE
HAS_SPEED_RATE
HAS_STEP
HAS_TEMPERATURE
HAS_TEXTURE
HAS_TIME_PERIOD
HAS_VISUAL_PATTERN
HAS_VOLUME
HAS_WEIGHT
HAS_WIDTH

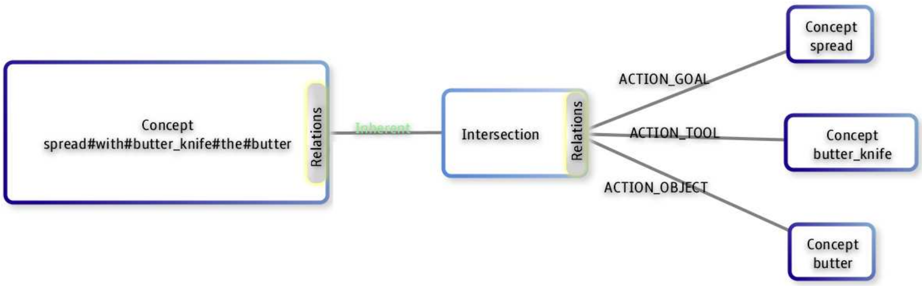
PRAXICON Structure (3)

Relations: Intersection



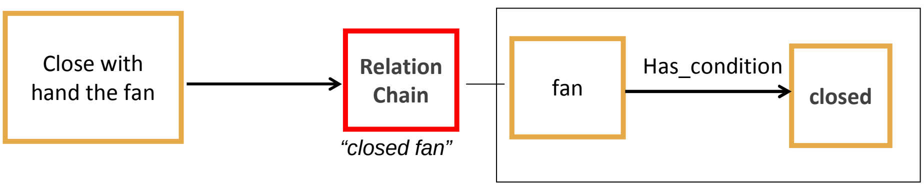
PRAXICON Structure (3b)





PRAXICON Structure (4)

Relations: Relation Chain

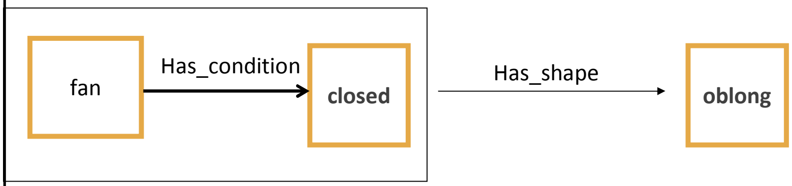


Why is such representation important?

Consider: "the fan is oblong"

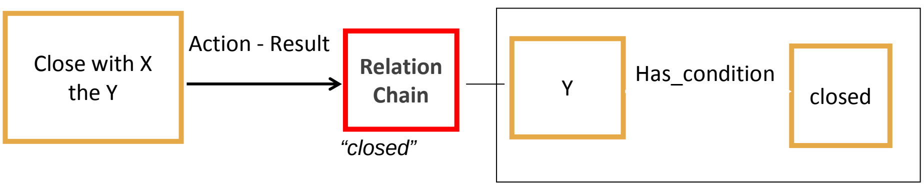
PRAXICON Structure (4b)

Relations: Relation Chain



PRAXICON Structure (4c)

Relations: Relation Chain

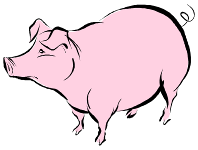


So, passive participles lexicalize systematically relation chain structures

PRAXICON Structure (5)

"pork", "χοιρινό"

pig —Aspect-concept— food



PRAXICON Structure (5)

"σφίγγω"

Hold_withX_theY —Has_force— tight

PRAXICON Structure (5)

Relations: Intersection

«πλουσιόπαιδο», “rich boy”

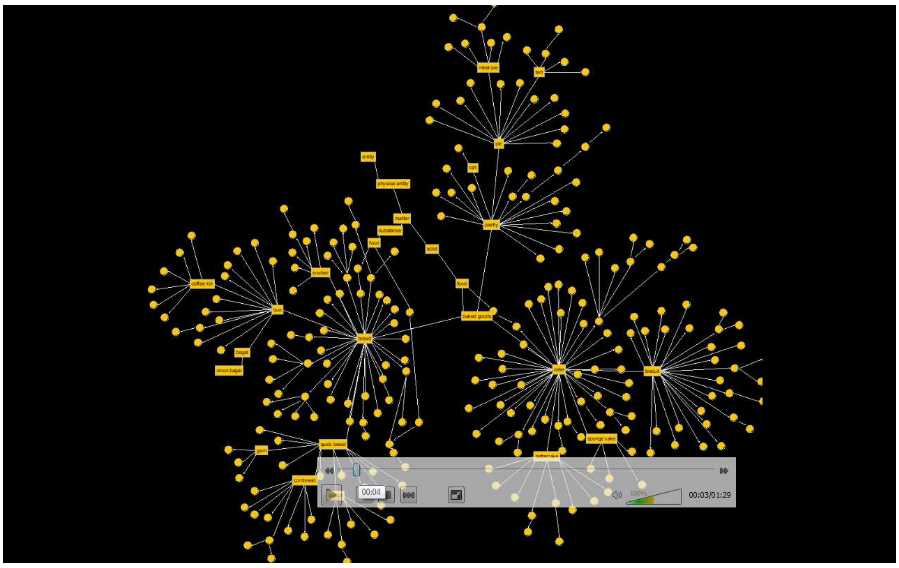
x

Intersection

Aspect-concept → Wealth “πλούτος”

Token-type → Young man, boy “παιδί”

Semantic Memory Activation in the PRAXICON



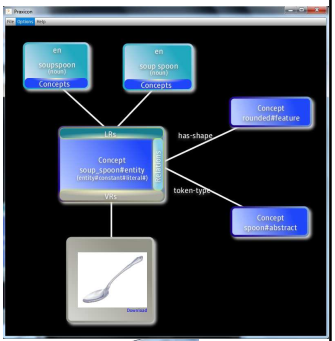
PRAXICON suite of resources and tools

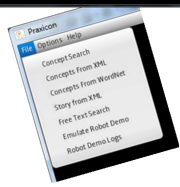
The PRAXICON Semantic Memory, its visual exploration interface (GUI) and the integrated language analysis and reasoning tools

In two forms:
- as a web service (database and game)
- as a downloadable

for local installation.

Contents:
- Embodied WordNet - Lexical Database (more than 100K concepts and relations) - Cognitive Experiments (5K)
- Corresponding visual representations from the ImageNet database.





From POETICON to…POETICON++

From Jan 2008 to Dec 2015

POETICON: The Poetics of Everyday Life

(2008-2011)

Grounding Resources and Mechanisms for Artificial Agents

POETICON++: Robots need Language

(2012-2015)

A computational mechanism for behaviour generalisation & generation in robots

Visit: www.poeticon.eu

Supplementary Material

http://www.csri.gr/downloads/SLRs.html

- Detailed Bibliography
- Videos shown in the tutorial
- Booklet with detailed profiling of SLRs (pdf)