Introduction

- The presentation is divided into three main parts:

- A. Introducing the concept Computational Linguistics
- B. Establishing the concept of Artificial Intelligence (NLP)
- C. Will show how Computational Linguistics is used in various technological devices as well as other spheres of human life.

What is Computational Linguistics?

Computation al Linguistics is NOT about :

a. The general computer-use for Languages

b. The functional use of language on computer system

c. A general-purpose programming languages of computer , C++, JAVA, Pascal, Fortran, Cobol etc.

Computation Linguistics in layman's term is :

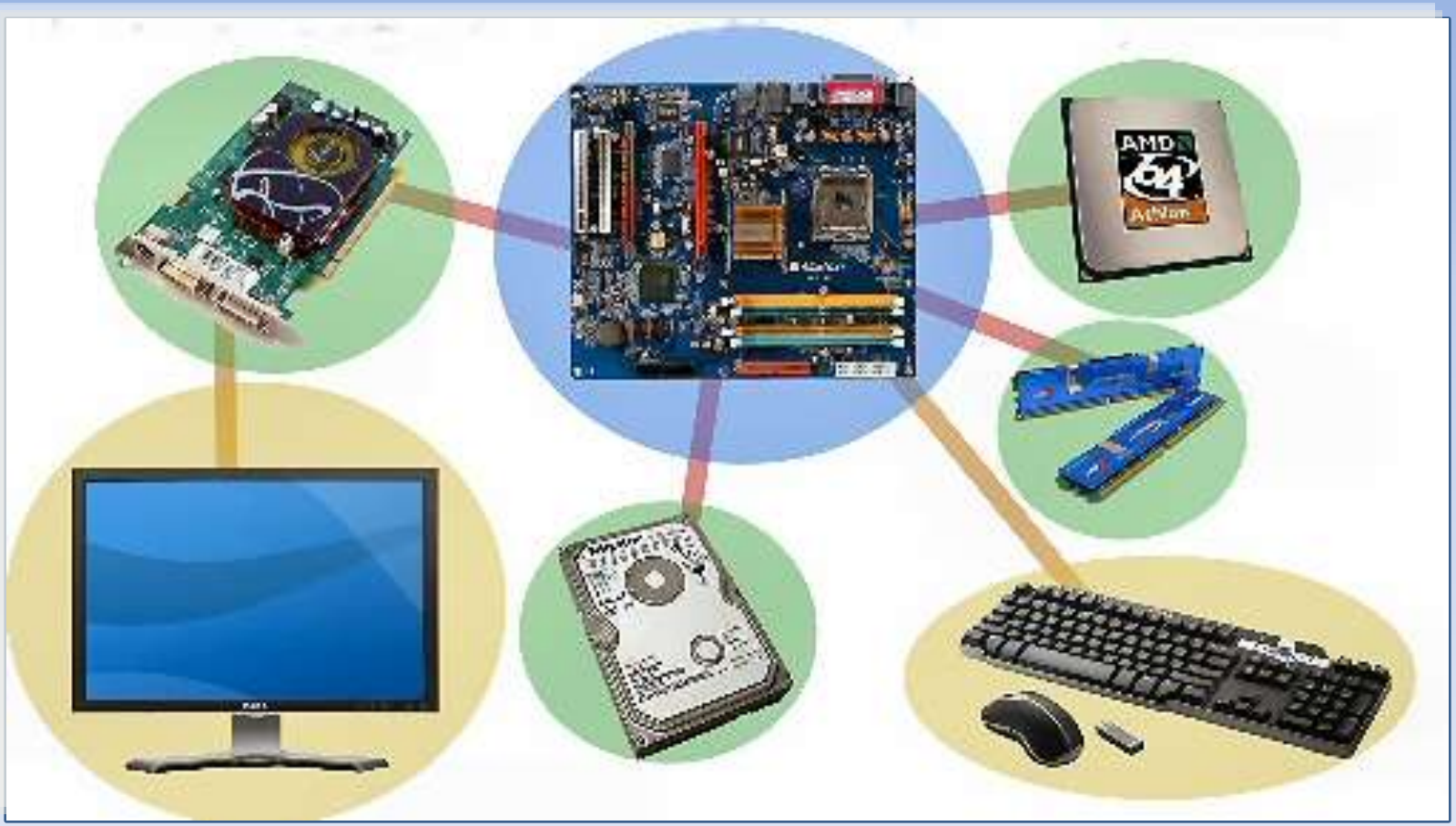
a. To make the computer learn Natural Languages (languages of human)

b. To make of the algorithms (logical formulations) and embed it into human language so that computer UNDERSTAND us (the human)

c. A simple version of the above would be to make computer like OUR BABY and make it learn our language

The heart and brain of Computer

From left to right (clock wise): 1. Graphic Card 2. Mother Board 3. CPU 4. RAM 5. Key-board, mouse 6. HD Hard-Disk 8. Monitor



The genesis of Computer

- **a. Assembly languages :** ARM, MIPS, x86
- Assembly languages are a family of low-level languages for programming in computers. It implements a symbolic representation of the numeric machine codes and other constants needed to program any electronic chip based hardware, particularly in CPU architecture.
- **b. High level languages :** BASIC, C, C++, C#, COBOL, Fortran, Java, Lisp, Pascal, Object Pascal
- In computing, a high-level programming language is written with strong abstraction from the details of the computer. It uses natural language elements with different permutation and combination of mathematical symbols.
- **c. Scripting languages :** JavaScript, Python, Ruby, PHP, Perl
- A scripting language or extension language is something that allows us to control one or many software application(s) which are already written in high-level language. Languages chosen for scripting purposes are often much higher-level than the language used by the host application. The process of labeling Unicode to existing mathematical numerical value is one good example of encrypting process.

The Knowledge based system:

Computers having reached a very high level of syntheses between Hardware and Software, are now ready for many other challenging works, and **Artificial Intelligence** is one of them.

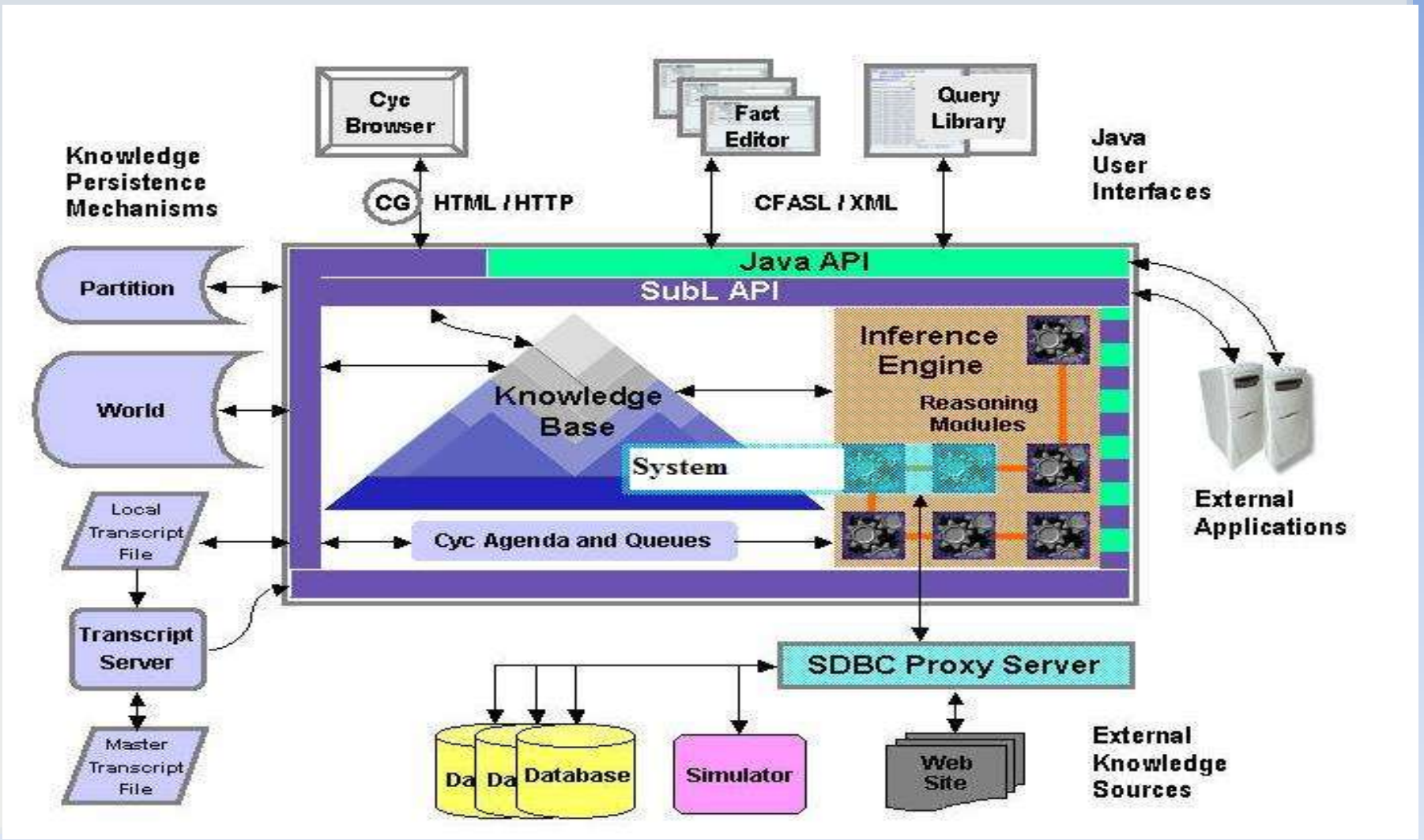
In one word, AI is about making computer **INTELLIGENT** by teaching how to master the 'mechanisms' in computation process for NATURAL languages with all its applications inbuilt in.

1 x 8 + 1 = 9
12 x 8 + 2 = 98
123 x 8 + 3 = 987
1234 x 8 + 4 = 9876
12345 x 8 + 5 = 987 65
123456 x 8 + 6 = 987654
1234567 x 8 + 7 = 9876543
12345678 x 8 + 8 = 98765432
123456789 x 8 + 9 = 987654321

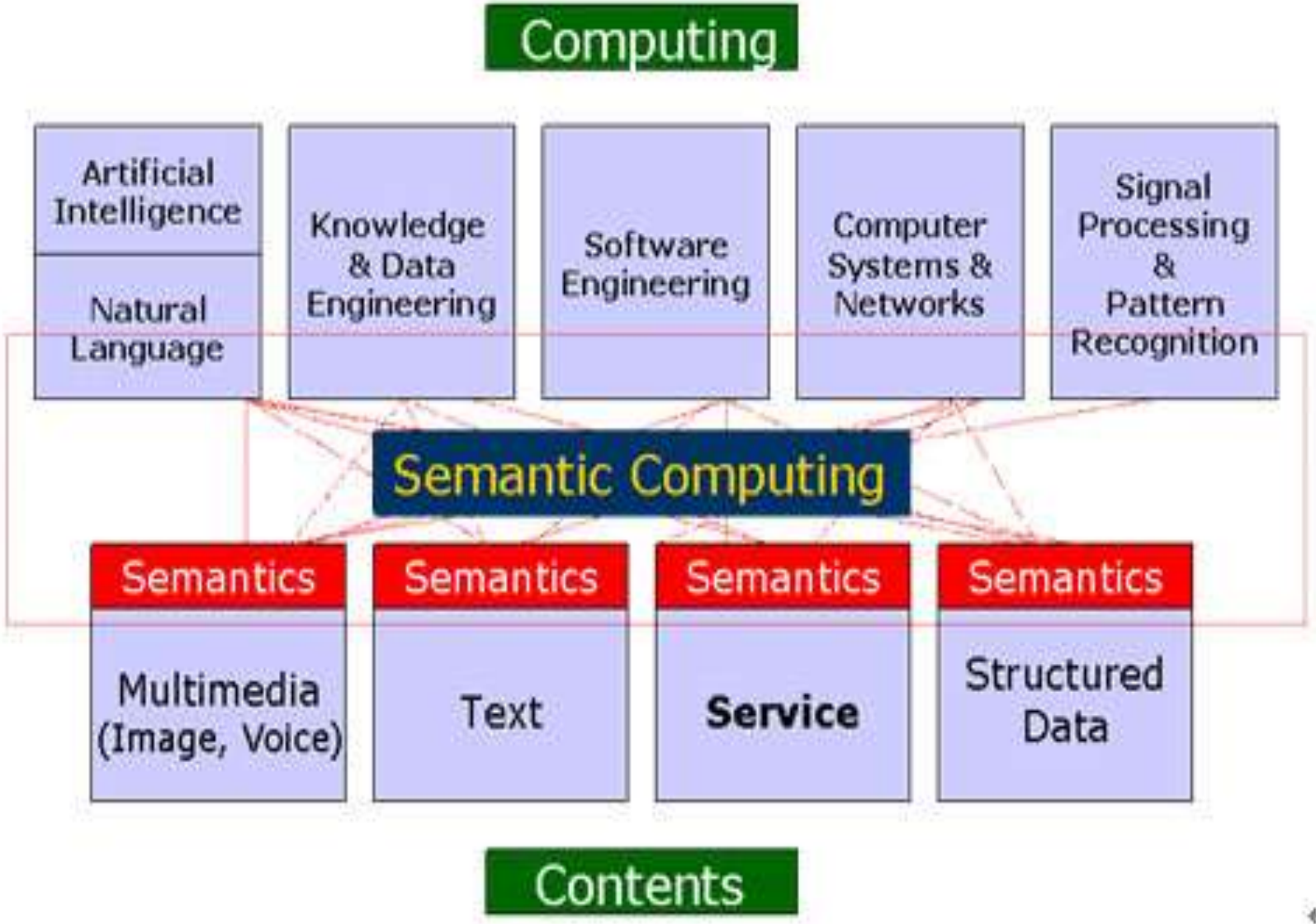
1 x 9 + 2 = 11
12 x 9 + 3 = 111
123 x 9 + 4 = 1111
1234 x 9 + 5 = 11111
12345 x 9 + 6 = 111111
123456 x 9 + 7 = 1111111
1234567 x 9 + 8 = 11111111
12345678 x 9 + 9 = 111111111
123456789 x 9 +10= 1111111111

1 x 1 = 1
11 x 11 = 121
111 x 111 = 12321
1111 x 1111 = 1234321
11111 x 11111 = 123454321
111111 x 111111 = 12345654321
1111111 x 1111111 = 1234567654321
11111111 x 11111111 = 123456787654321
111111111 x 111111111 = 12345678987654321

The architecture of Artificial Intelligence



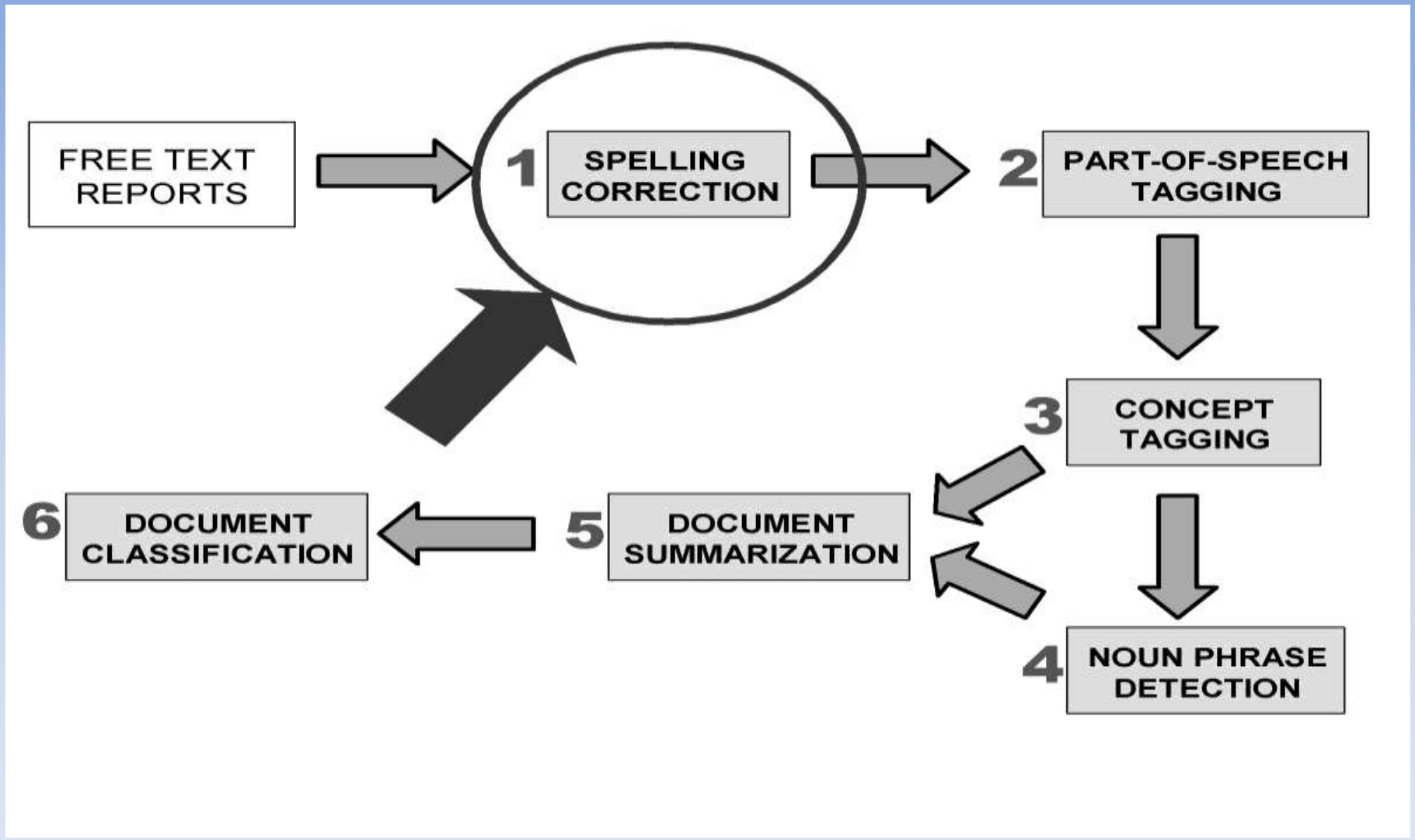
The computing process



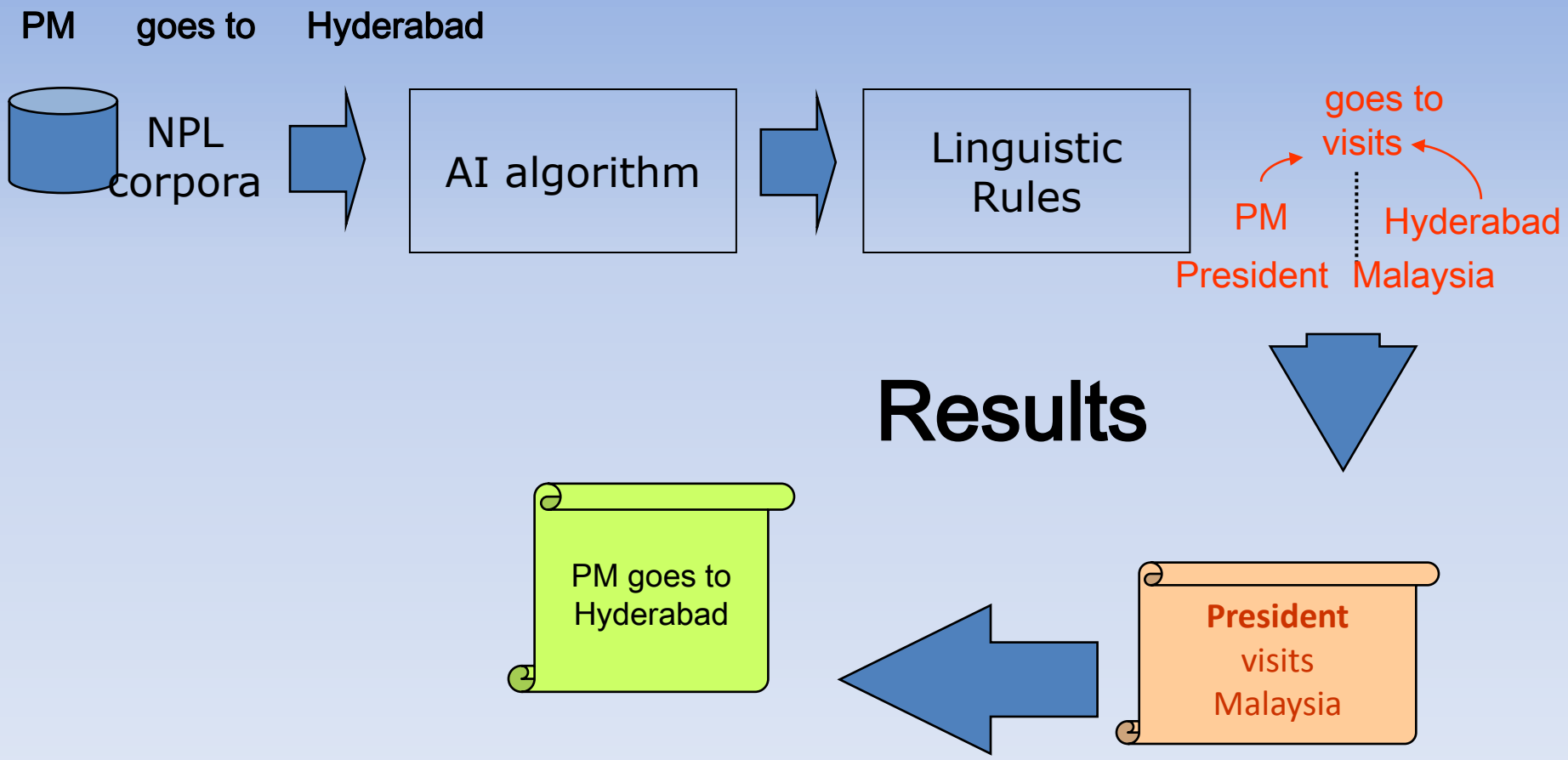
Branches of AI

- a) Logical AI
- b) Search
- c) Pattern recognition
- d) Representation
- e) Inference
- f) Common sense and reasoning
- g) Learning from experience
- h) Planning
- i) Epistemology
- j) Ontology
- k) Heuristics
- l) Genetic programming

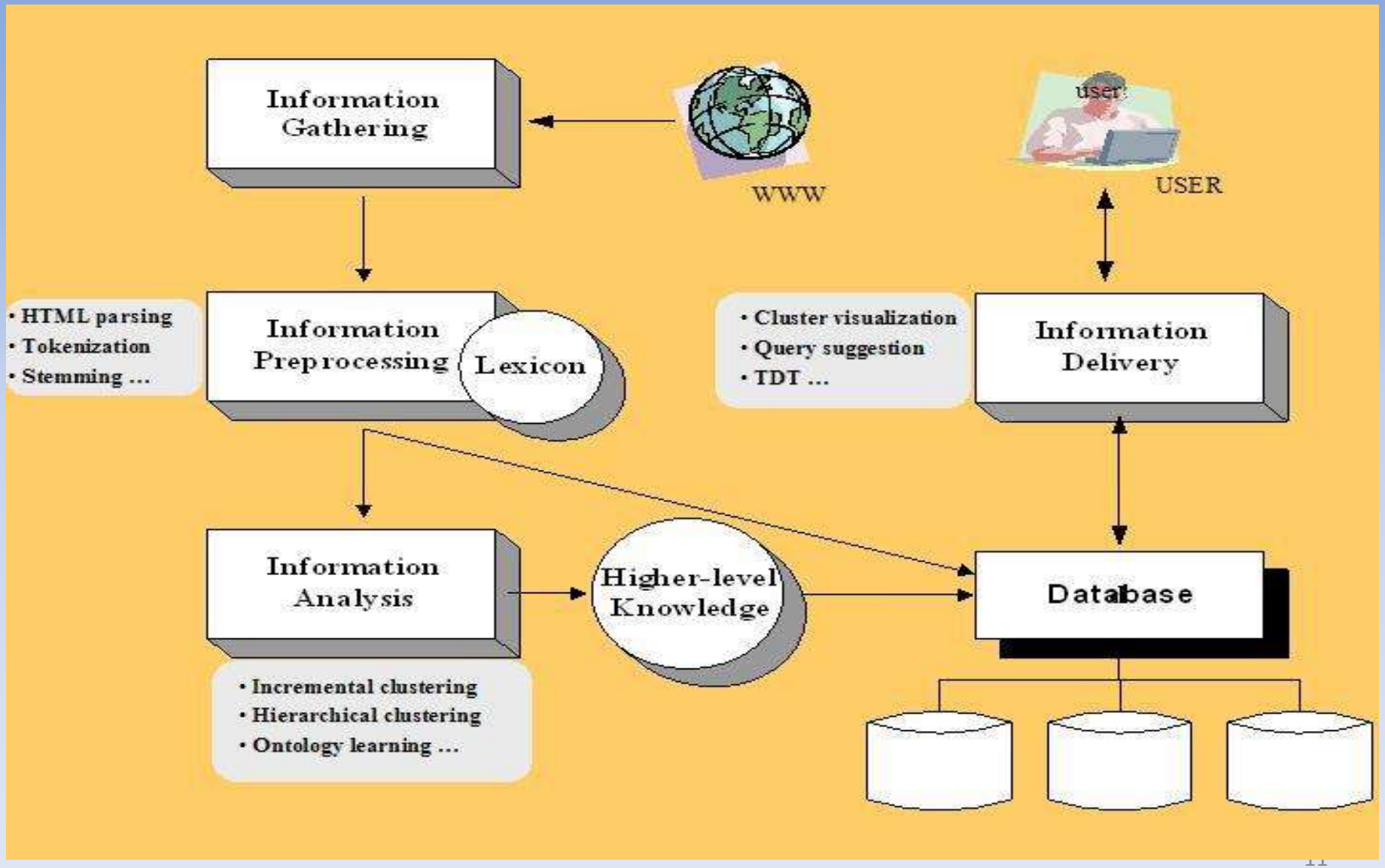
AI as the basis of CL



CL ! How does it work ?



CL, what does it do?



- **The domains of Computational Linguistics**
- **a) Search engines: Google et al.**
- **b) Web dictionaries**
- **c) Web translation/transliteration**
- **d) Text Messages of Cell phones**
- **e) Voice and speech recognition system**
- **f) Braille computer**
- **g) GPS with voice activation**
- **h) Automated bank and phone machines**
- **i) Mining of the legal documents**
- **j) Mining of medical records**
- **k) Market statistics for product promotion**

What is Computational Linguistics?

- Computational Linguistics is the computational analysis of natural languages.
  - Process information contained in natural language.
- Can machines understand human language?
  - Define ‘understand’
  - Understanding is the ultimate goal. However, one doesn’t need to fully understand to be useful.

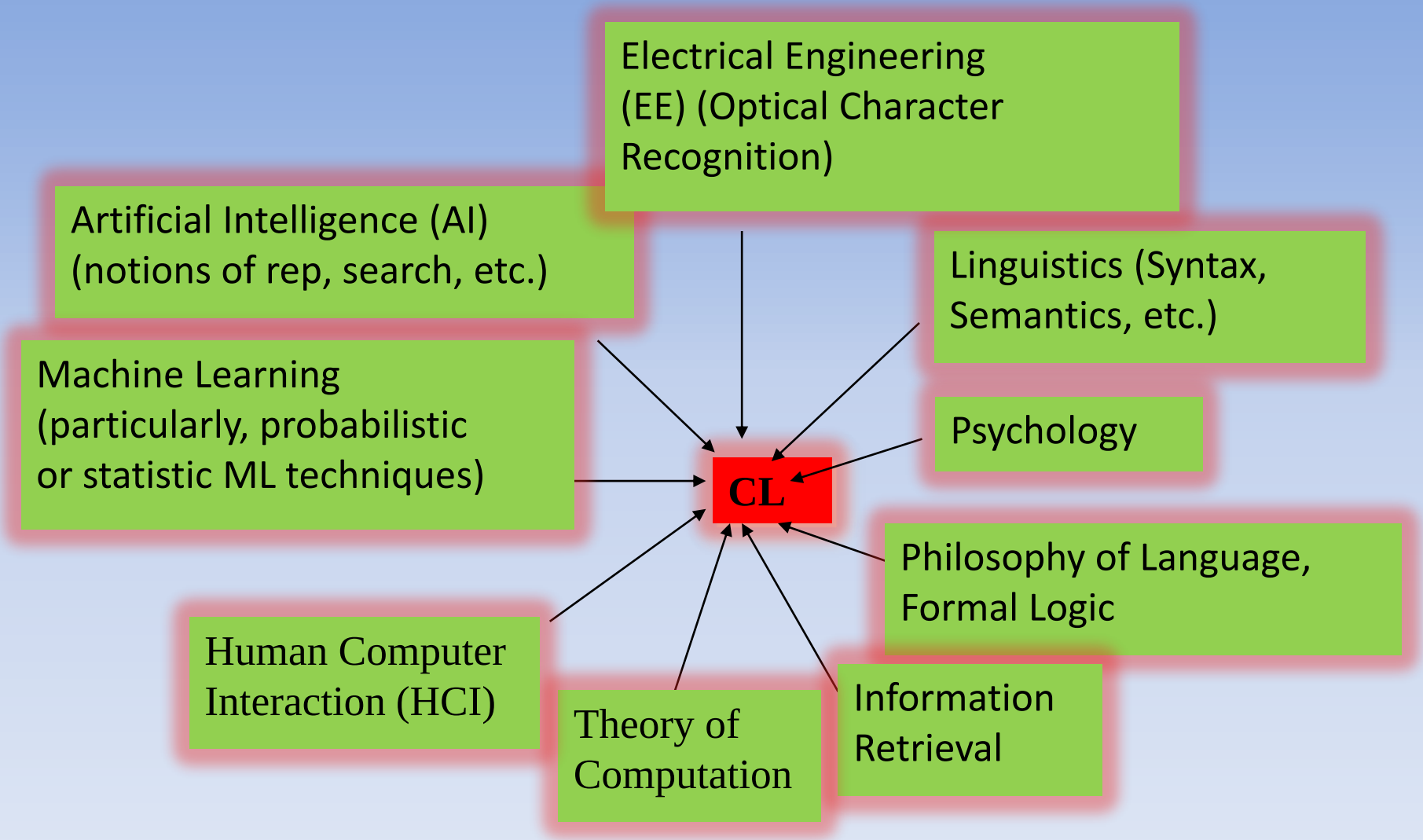
CL vs NLP

Why “Computational Linguistics (CL)” rather than “Natural Language Processing” (NLP)?

- Computational Linguistics
  - — Computers dealing with language
  - — Modeling what people do

- Natural Language Processing
  - —Applications on the computer side

Relation of CL to Other Disciplines



A click at key-board and screen display

A click at key-board and screen display

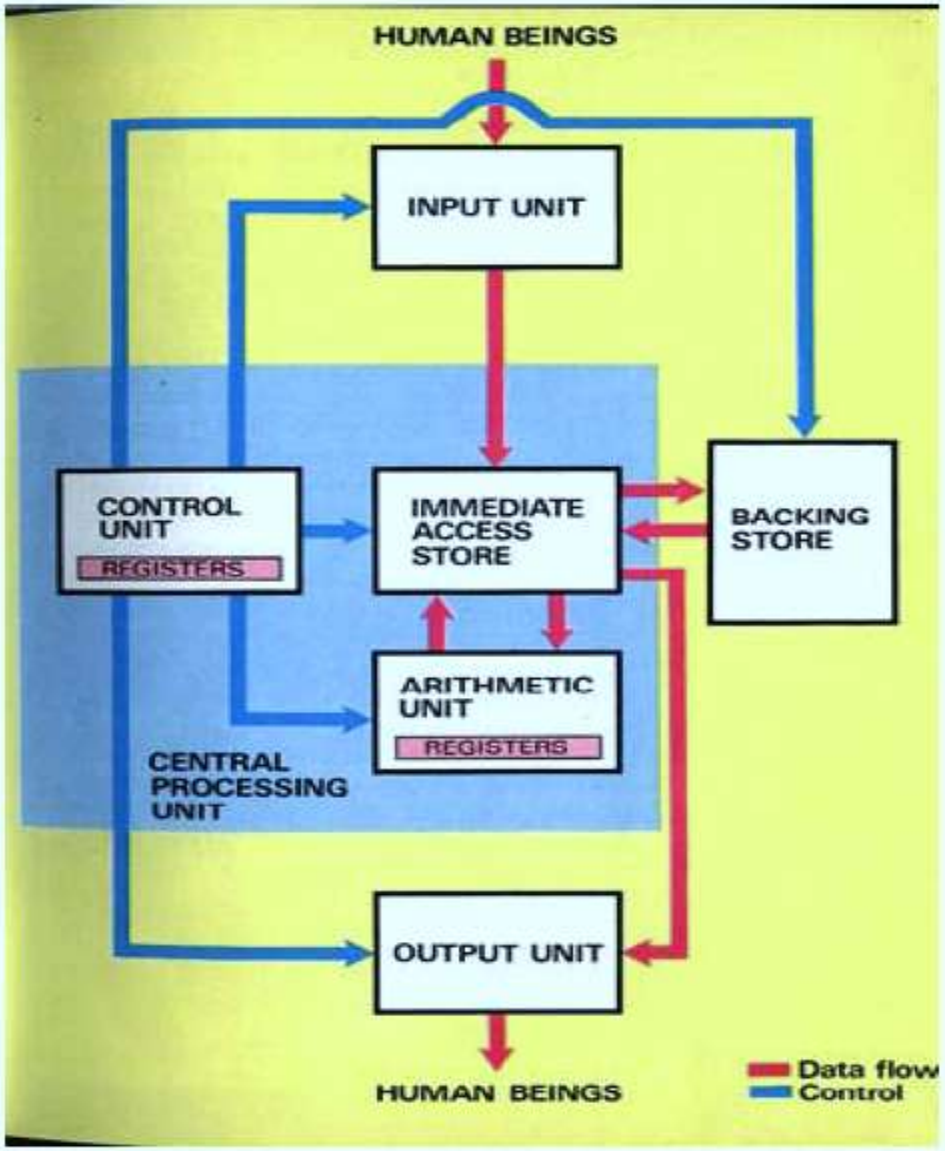
1.Keyboard to Mother-board (magnetic field of the key-board changes the power energy to electronic energy)

2. Mother-board to Process (non-binary to binary process)

3. Processor to Hard Disk (mechanical signal to optical signal)

4. HD to RAM (static signal to digital and free-floating data)

5. Optics to Visual (Non visual to visual graphic)



http://pointlessmuseum.com/computer/007.html

A Sampling of “Other Disciplines”

- Linguistics: formal grammars, abstract characterization of what is to be learned.
- Computer Science: algorithms for efficient learning or online deployment of these systems in automation processes for computer.
- Engineering: stochastic(statistic) techniques for characterizing regular patterns for learning and ambiguity resolution.
- Psychology: Insights into what linguistic constructions are easy or difficult for people to learn or to use

Language and Intelligence: Turing Test

- Turing test:
  - machine, human, and human judge
- Judge asks questions of computer and human.
  - Machine's job is to act like a human, human's job is to convince judge that he's not the machine.
  - Machine judged "intelligent" if it can fool judge.
- Judgment of "intelligence" linked to appropriate answers to questions from the system.

Putting sounds together:

- A sub-discipline of linguistics called **phonology** helps us to put the sounds together and understand the structure of syllables.
- V a, I, etc.
- CV me
- VC use
- CCV dress
- CCCV strange [C1= S, C2= p/t/k, C3= l/r]
- *CCCCV …NPos.
- The reason for this phonological constraint is the maximum power of a vowel.
- A vowel usually can take the maximum load of three consonant sounds which can be produced without any problem.
- This is obeyed by almost every language of the earth.

Building words; a higher step than the syllables:

- I always cite Humpty Dumpty in making this point clear to anyone. So, let's see what Humpty Dumpty has to say about the words:

Humpty appears in Lewis Carroll's Through the Looking-Glass (1872), where he discusses semantics and pragmatics with Alice.

"I don't know what you mean by 'glory,' " Alice said.
Humpty Dumpty smiled contemptuously. "Of course you don't—till I tell you. I meant 'there's a nice knock-down argument for you!' "
"But 'glory' doesn't mean 'a nice knock-down argument'," Alice objected.
"When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less."
"The question is," said Alice, "whether you can make words mean so many different things."
"The question is," said Humpty Dumpty, "which is to be master that's all."

- I do not claim that we can enjoy the same freedom that Humpty Dumpty has in the above passage in terms of making up words, but if we pay attention to the 'smaller units' called 'morpheme' or 'word-parts', we can definitely accelerate the process of learning a language and make it comprehensible even for a machine such as computer.

Word-part: need more attention and care:

- 1. [V→N]
- ‘-al’ Meaning: ‘doing the act of X’

  Exp: **propose-proposal, refuse-refusal.**
- ‘-ation’ Meaning: ‘the result of ‘X-ing’

  Exp: **realize-realization, appreciate-appreciation**
- 2. [V→Adj]
- ‘-able’ Meaning: ‘able to be X-ed’

  Exp: **fix-fixable, move-movable.**
- ‘-ive’ Meaning: ‘having the property of doing X’

  Exp: **interrogate-interrogative, negate-negative.**
- 3. [N→V]
- ‘-ize’ Meaning: ‘to put in X’

  Exp: **hospital-hospitalize, capital -capitalize.**
- ‘-fy’ Meaning: ‘to make of X’
- Exp: **beauty-beautify, person-personify**.

Speech/Character Recognition

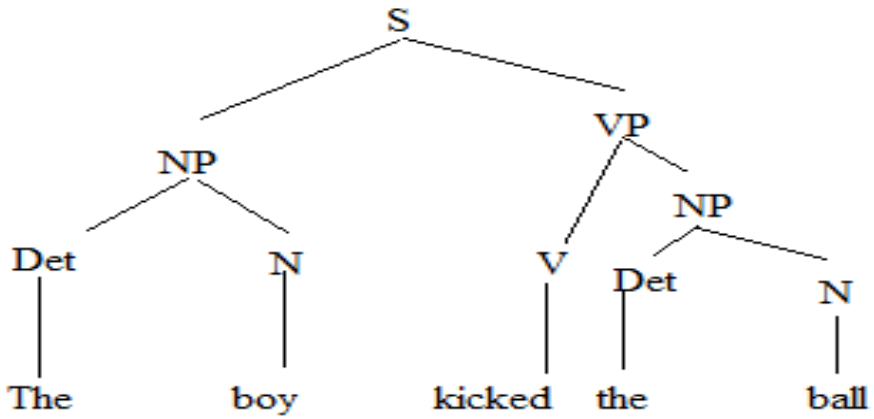
- Decomposition into words, segmentation of words into appropriate phones or letters
- Requires knowledge of phonological patterns:
  - night-rate == nitrate
  - grey day == grade A
  - why choose == white shoes
  - I scream == ice cream

Morphological Analysis

- Inflectional
  - duck + s = [N duck] + [plural s]
  - duck + s = [V duck] + [3rd person s]
- Derivational
  - kind, kindness
- Spelling changes
  - drop, dropping
  - hide, hiding

PS-Rules: Phrase-structure rules (Chomsky 1956)

- The most simple Ps-Rule of Chomsky 1956 which is still somewhere as the base of his latest ‘minimalism’ is the simplest rules of SYNTAX to understand. For example:
- **The boy kicked the ball**.
- The boy= subject =(NP_1, an article and a noun)
- kicked the ball= VP
- the ball= object = (NP_2, an article and a noun) and thus,



(1) $S \rightarrow NP; VP$

$VP \rightarrow V; NP$

…Ps-Rules

(4) S → NP; VP

VP → V ; NP

These two simple rules of GB Theory explain the organization of constituents in clauses. The first rule can be read as 'a sentence consists of a noun phrase (which is the subject) and a verb phrase'. The second rule states that 'a verb phrase consists of a verb and a noun phrase (object)'. The semicolons that occur in the right-hand side of the rules indicate that the two constituents may occur in either order. Together, the rules generate the following structures:

(5) NP_{sub} V NP_{obj} [=SVO]
NP_{sub} NP_{obj} V [=SOV]
V NP_{obj} NP_{sub} [=VOS]
NP_{obj} V NP_{sub} [=OVS]

What is the point?

- Well you must be wondering as to am I going to teach you the outdated PS-Rules!
- The answer is no. Let me explain something important to you.
- A sentence has basic constituents like S, O and V and since these are three, the permutation and combination of these will bring us six different choices, such as:

SOV
SVO
VSO
OSV
VOS
OVS

- Now, I want to show you the application and
- utility of the so called PS-Rules to demonstrate
- an unknown fact as to why out of many word-orders
- that are available in the world's languages, some word-orders
- are dominant while others are just too rare to even attest.

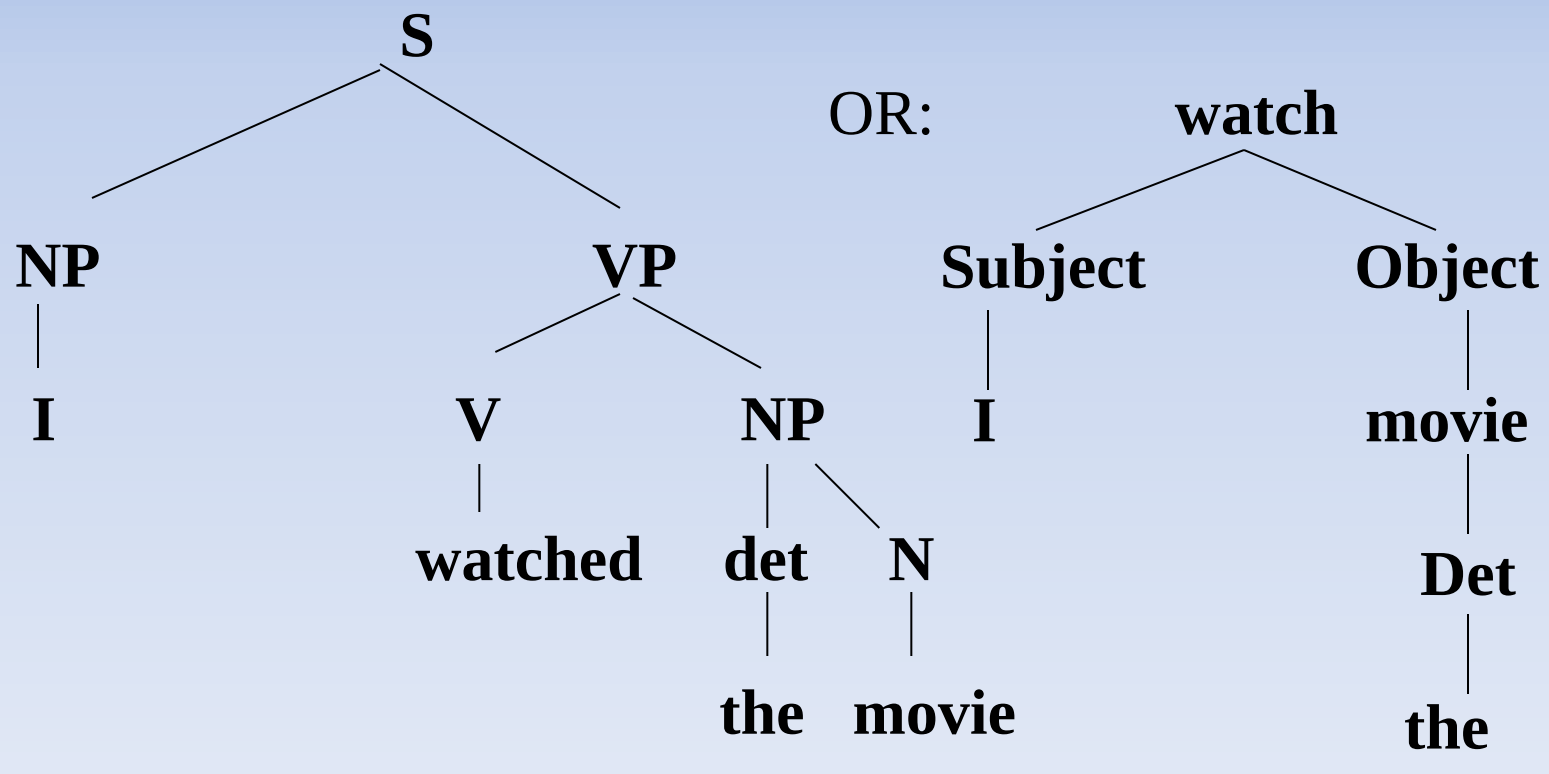
Ps-rules to word-order and then to correlation factors:

Word order	Correlation
VO	OV
Preposition + noun	Noun + preposition
Noun + genitive	Genitive + noun
Noun + adjective	Adjective + noun
Noun + relative clause	Relative clause + noun
Sentence-initial question word	Non-initial question word
Prefixes	Suffixes
Auxiliary verb + main verb	Main verb + auxiliary verb
Comparative adjective + standard of comp	Standard of comp + comparative adjective
Verb + adverb	Adverb + verb
Negative + verb	Verb + negative
Subordinator + clause	Clause + subordinator

Source: Whaley 1997; Pp86

Syntactic Analysis

- Associate constituent structure with string
- Prepare for semantic interpretation



Rules for Passive in English:

- 1. The first very rule in English to transform an Active to Passive is that the verb must be a transitive one in the clause.
- 2. Change the positions of the subject and the object.
- 3. The verb of the sentence, having been changed to its V1 form first, must be transformed into V3.
- 4. We must introduce a BE verb after the changed place of the Obj as the Subj of the passive.
- 5. This Be verb must be given the left over TENSE marker of the original verb of the clause and should agree with the changed Obj that is the subj of the passive sentence.
- 6. We can insert or place (optional, though) a BY phrase before the Obj of the Passive sentence.
- Go back