Artificial Intelligence AND Distributed Systems!

Applied Philosophy of Science's Research On A Networked Science Education System

By Joe Phillips, Jasmine Farley
and Ismael Cuevas
2021 November 10/11
for the Computer Science Society
of DePaul University

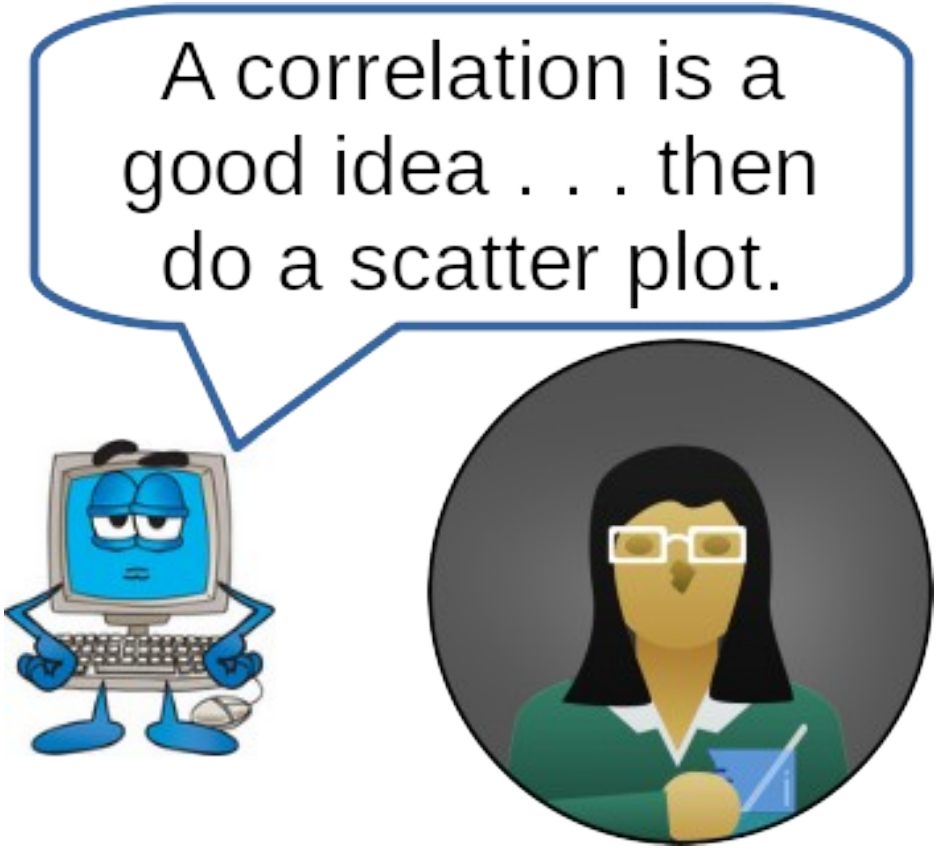
- This is Tiffany

- Tiffany loves science . . .

- Tiffany loves doing research, but wants feedback

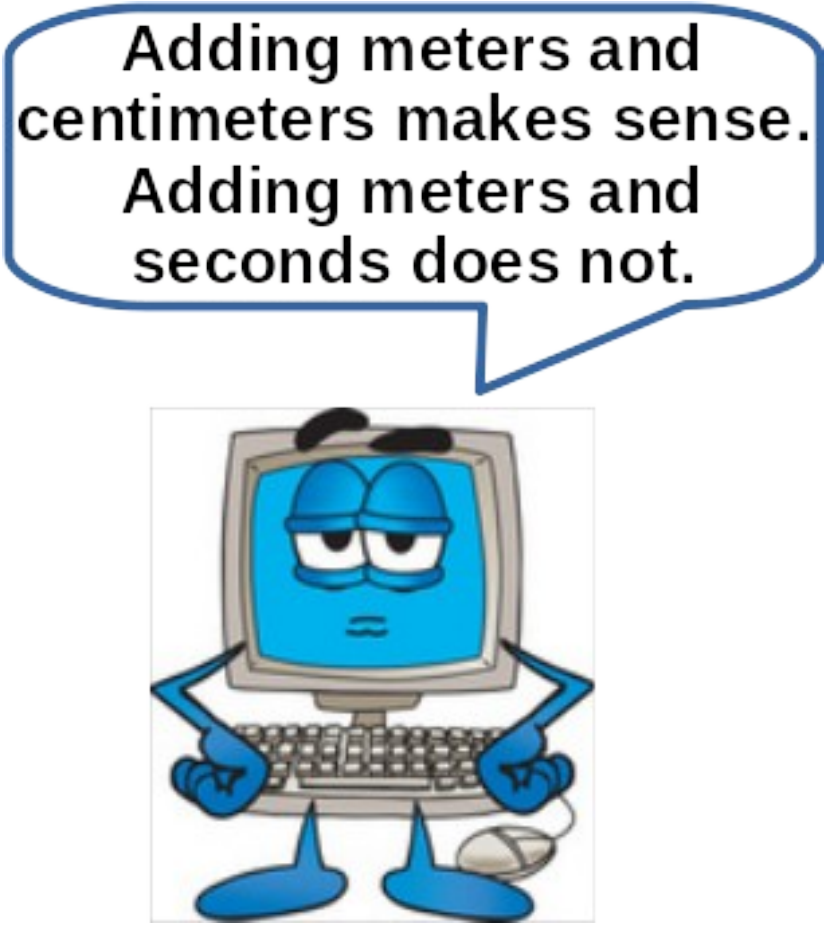
- Ordinary lab notebook websites let Tiffany record her findings.

- But the Scienceomatic gives Tiffany feedback and suggestions.

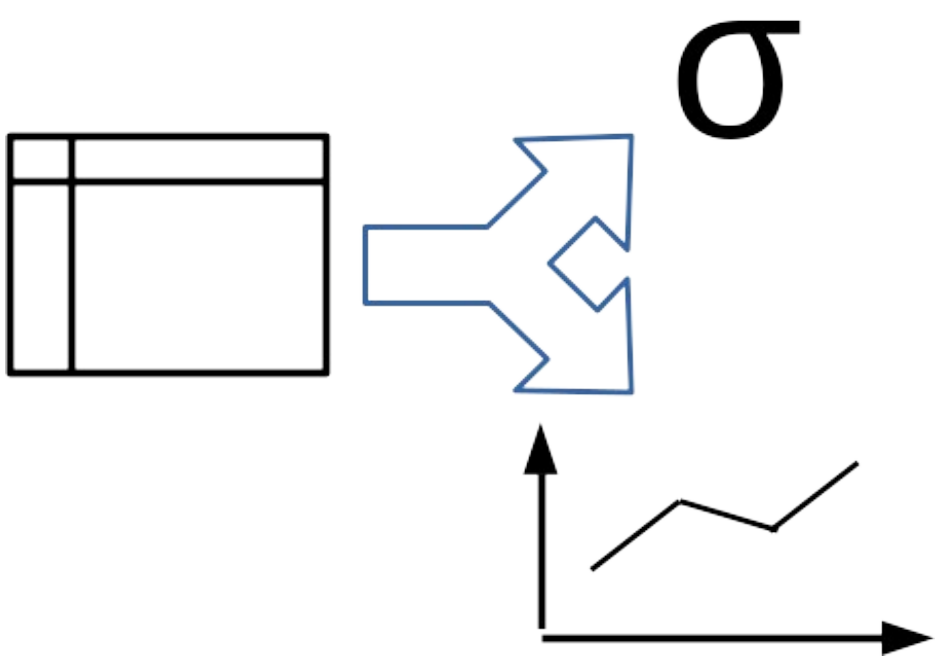


- The Scienceomatic can give feedback because it uses knowledge bases.
- Knowledge bases (kbs) are consistent, computable views of the natural world.
- Kbs hold the scientific data and knowledge on particular topics.

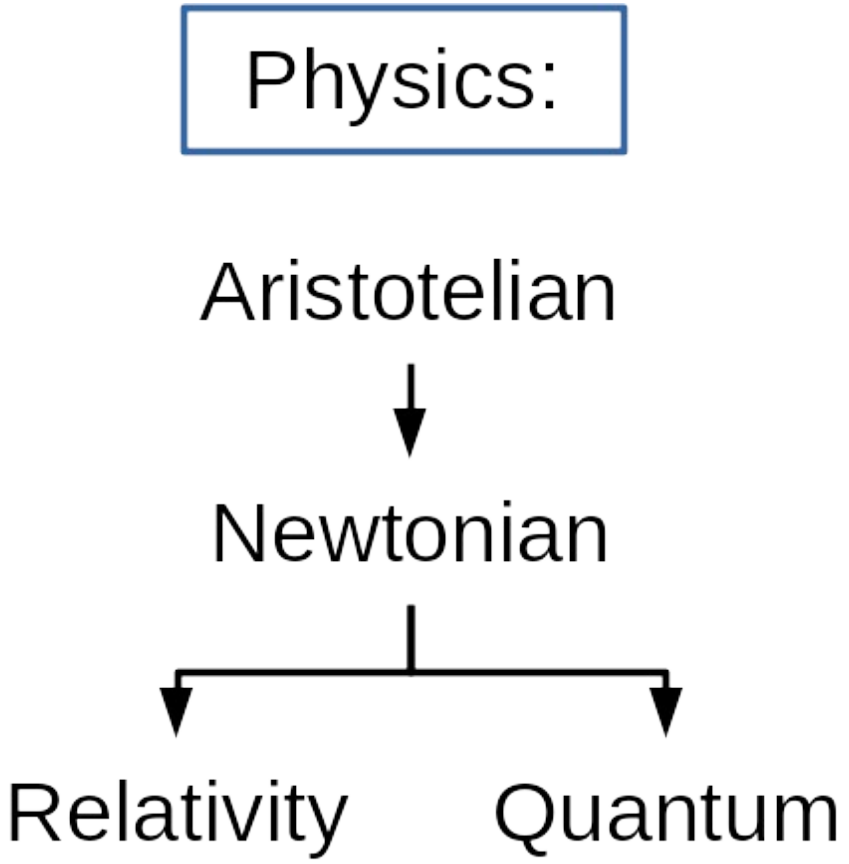
- Knowledge bases can check things like:
  - units and dimensions
  - arithmetic and equation usage
  - statistics and the strength of results
  - overall consistency



- Knowledge bases can also create
  - tables
  - graphs
  - charts

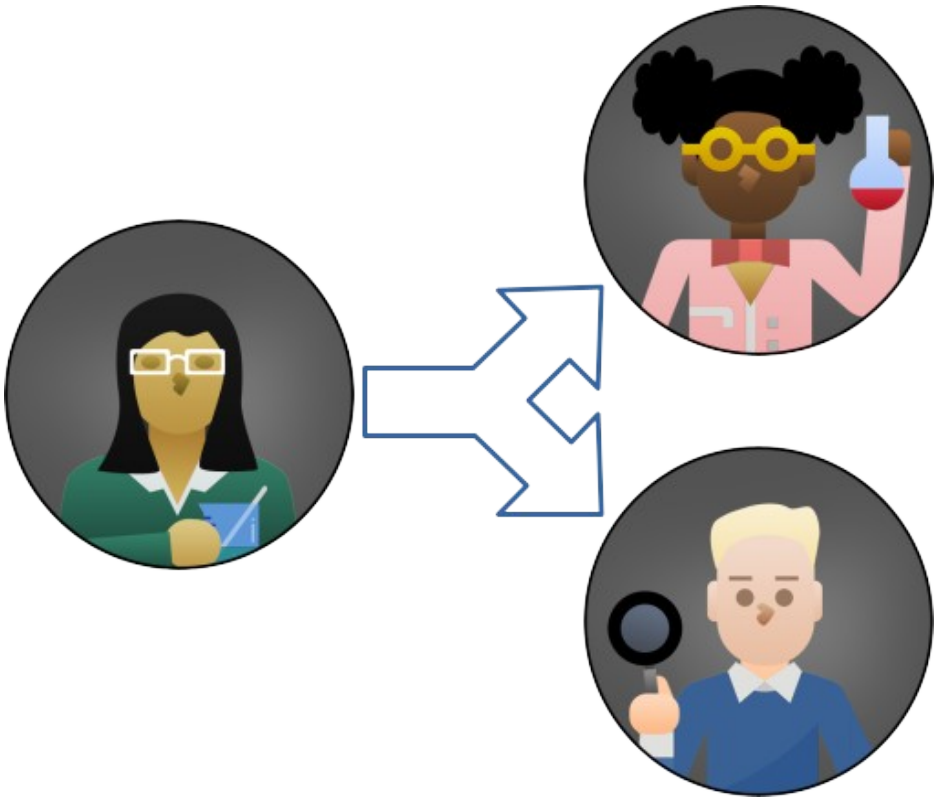


- Knowledge bases cumulatively grow on prior knowledge bases
- Competing worldviews are represented by competing knowledge bases

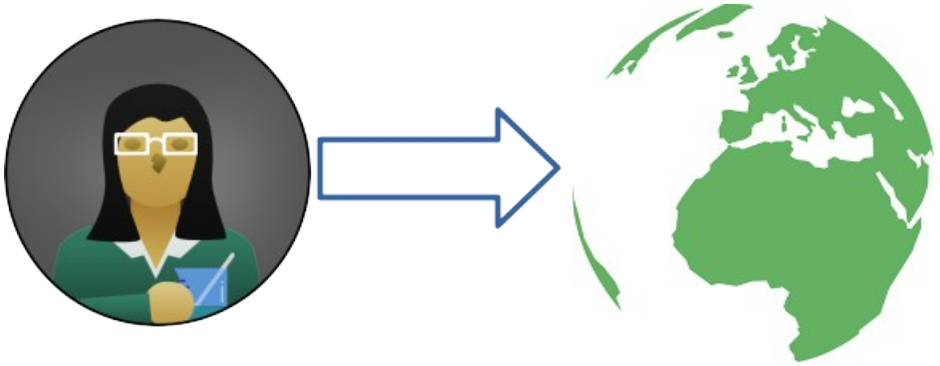


- Now Tiffany wants to share her results

- She can share her revised knowledge base with her close friends Teddy and Shakiya . . .

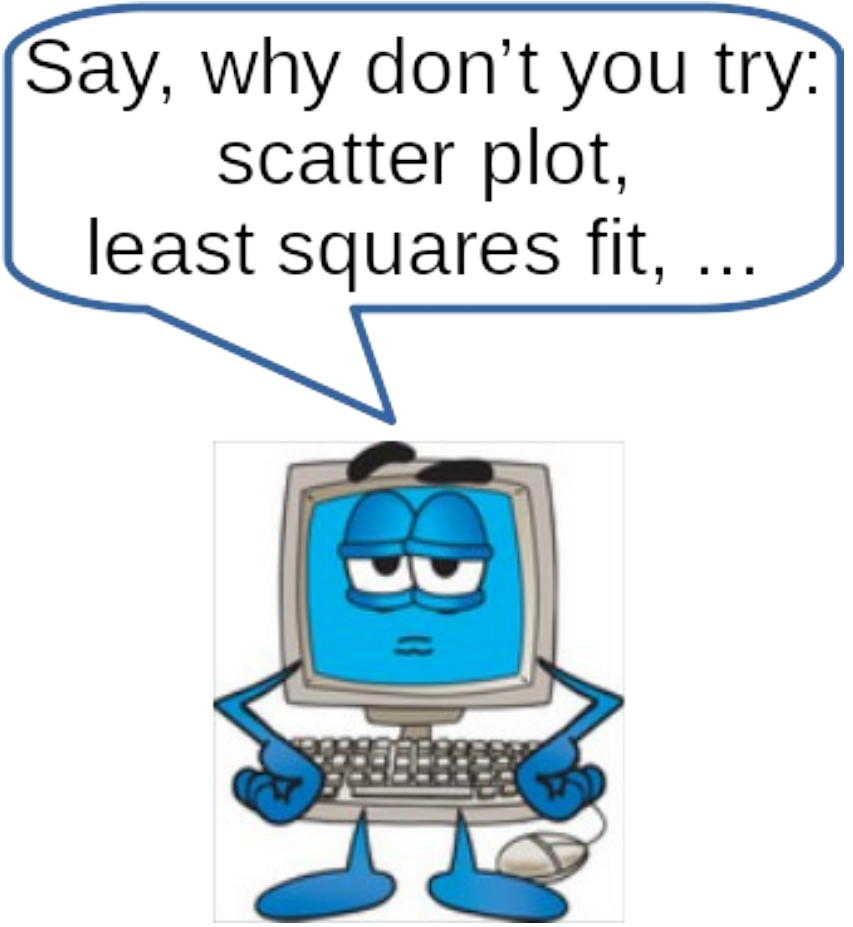


- . . . and if she wants, she can publish her knowledge base for others to use and extend.

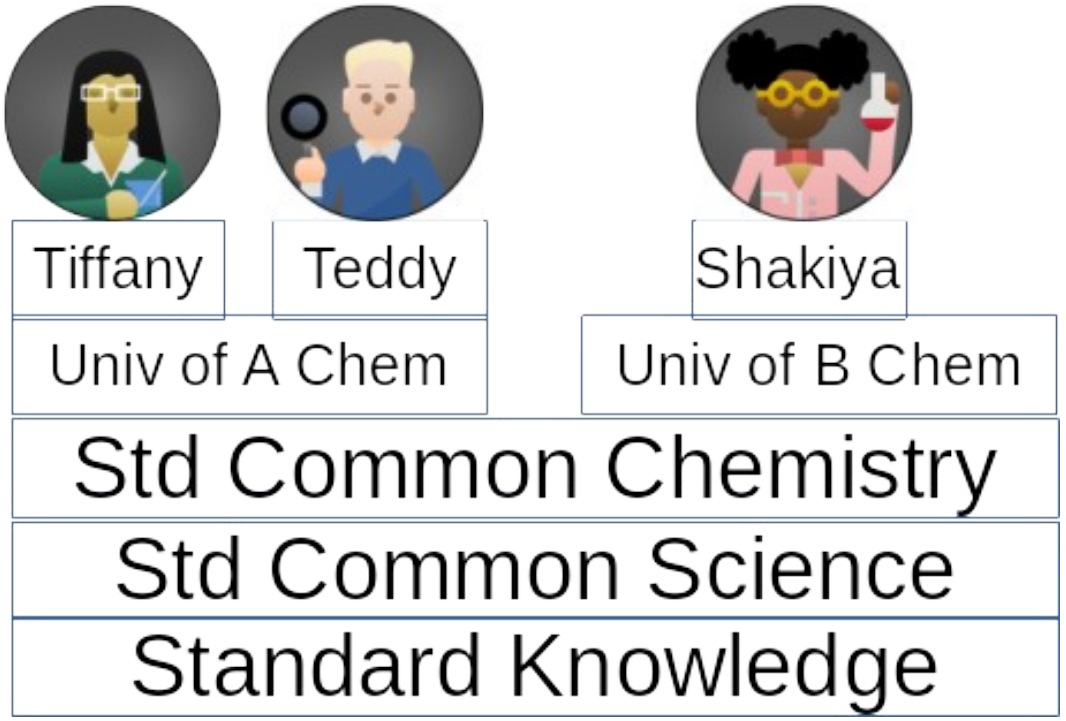


- When they do so, Tiffany will automatically get credit

- Later, the Scienceomatic can Tiffany suggesting operators to try.



Knowledge bases are hierarchical

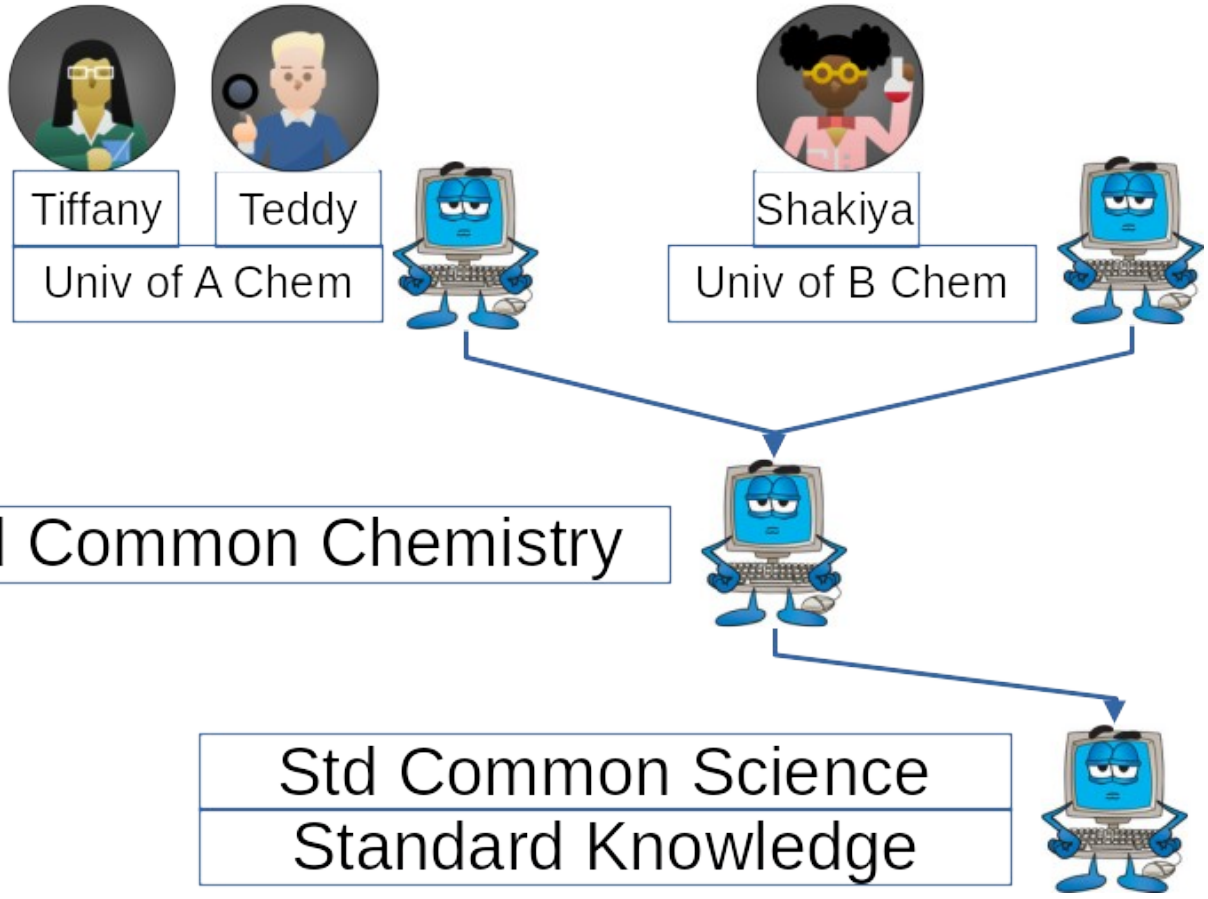


- Tiffany, Teddy and Shakiya all share the same
  - Standard knowledge
  - Common science
  - Common chemistry
- Tiffany and Teddy share:
  - Their university's chemistry kb

Knowledge bases can be networked

Knowledge bases can be distributed across multiple computers

- Different institutions are responsible for debugging and improving different parts of kb

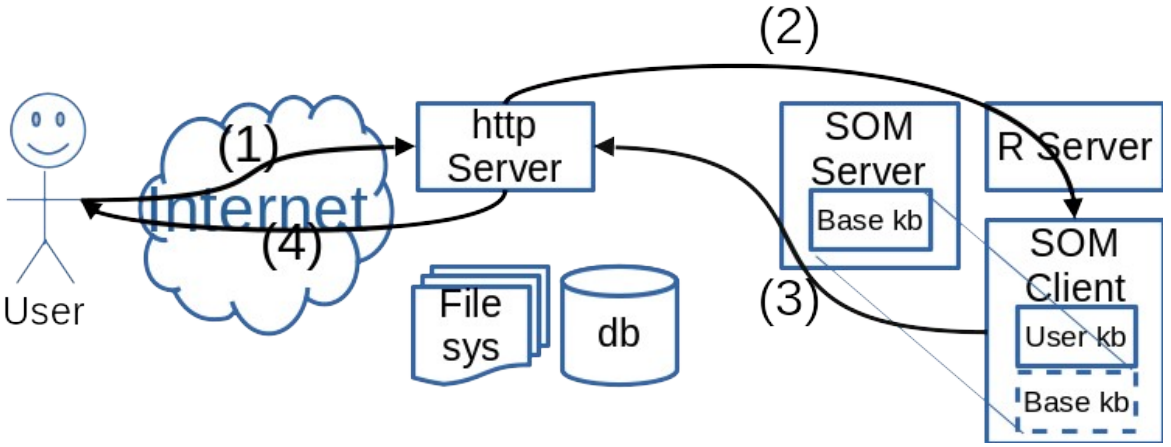


Check Out the Website!

The Front End

- Technologies:
  - Angular
  - Bootstrap
- Making RESTful calls to the server side to perform C.R.U.D. commands and update the client accordingly

Back End

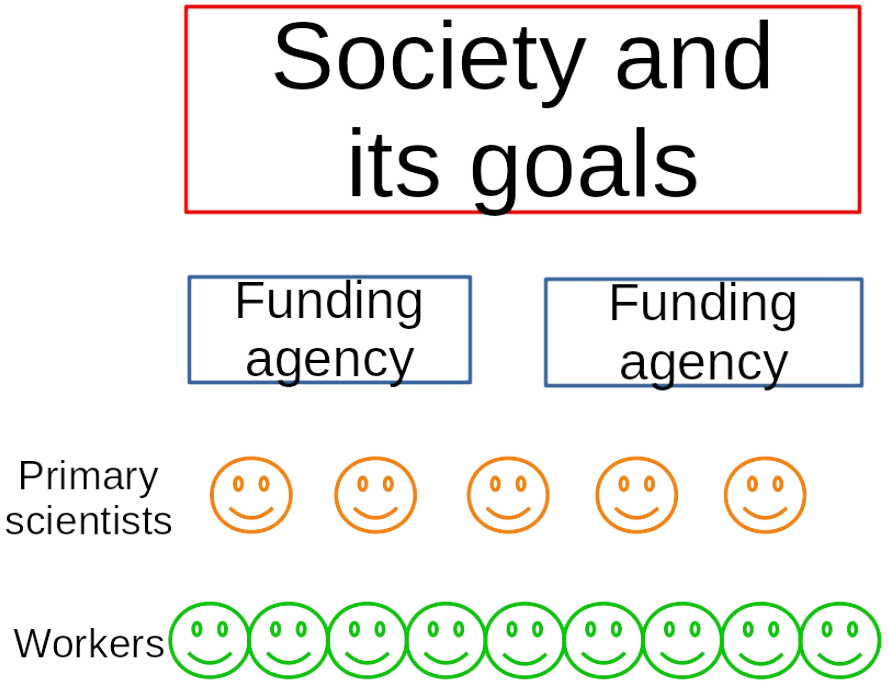


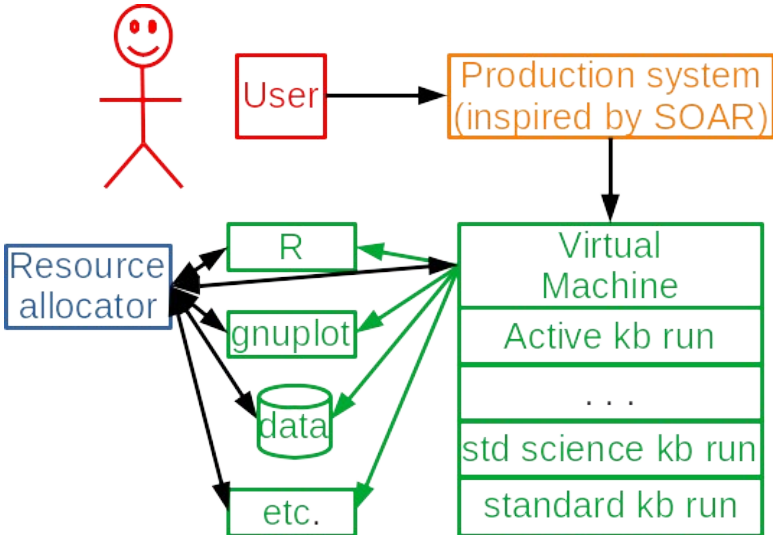
1. Client requests a page
2. HTTP Server sends request to SOM process dedicated for that client
3. SOM process sends JSON response
4. HTTP server formats response as HTML

How to Collaborate in the 21st Century?

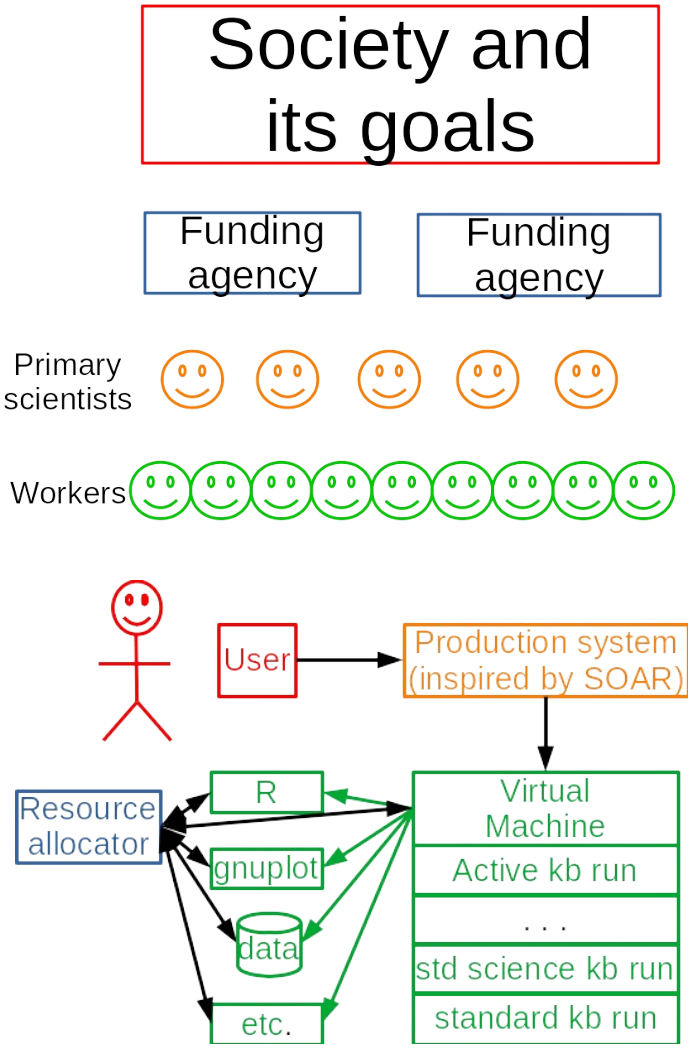
- How technology can help us
  - Use strengths of computers
    - Accuracy
    - Exhaustive search
      - Mechem
    - Ability to use lots of data
    - Ability to use lots of knowledge
      - Networked environments
  - Ability to connect humans
    - across geographic separation
    - across temporal separation
- Not trying to supplant humans!
  - Humans and computers have different strengths

The Analogy Between How Humans and Our System Does Science



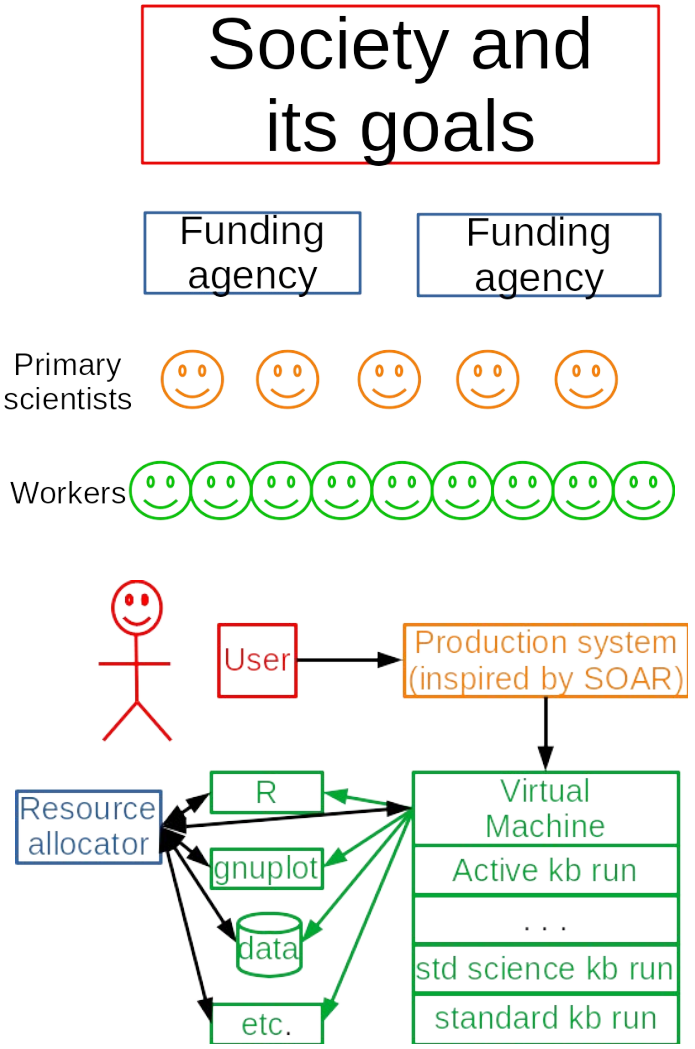


The Workers: basic scientific computation



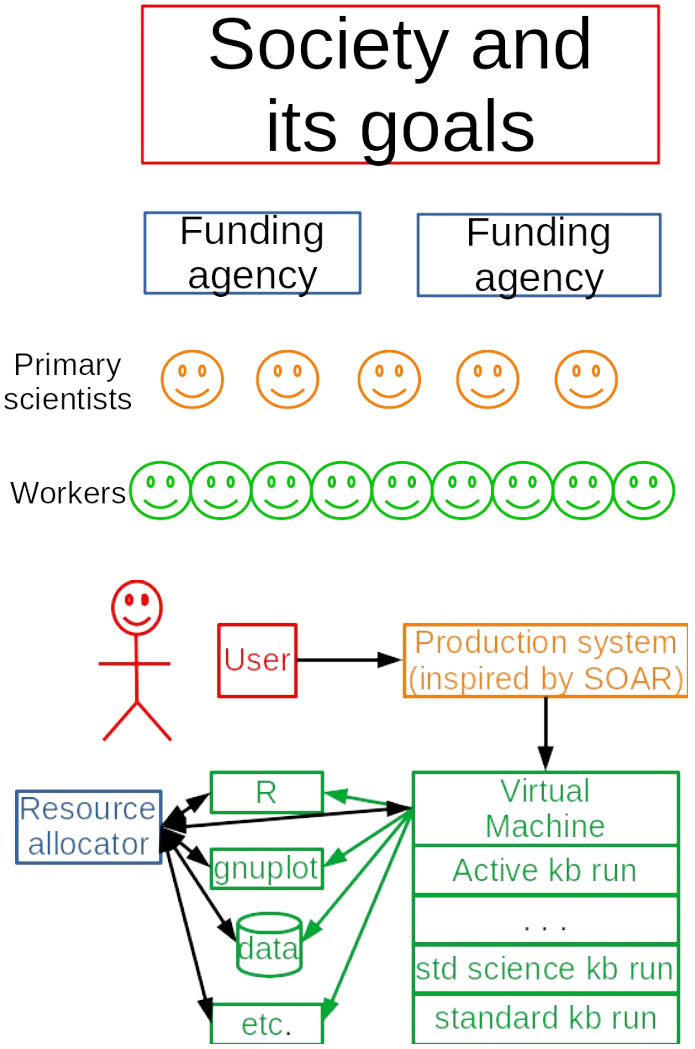
- The Virtual Machine:
  - **Serves as:** Knowledge of ***how*** to do basic reasoning (e.g. modus ponens, arithmetic)
  - **Serves as**: Textbook procedural knowledge: ***when*** to use algebra, statistics, etc.
- Knowledge base
  - **Purpose:** hold declarative knowledge
    - charge of electron
    - mammalian phylogenetic tree(s)
  - **Serves as:** "factual" textbook knowledge
  - Composed of kb runs that cumulatively build on each other
- Auxiliary programs
  - **Purpose:** Specialize algorithm running
  - **Serves as:** Knowledge of ***how*** to do algebra, statistics, etc.

The Primary Scientists: the idea generators



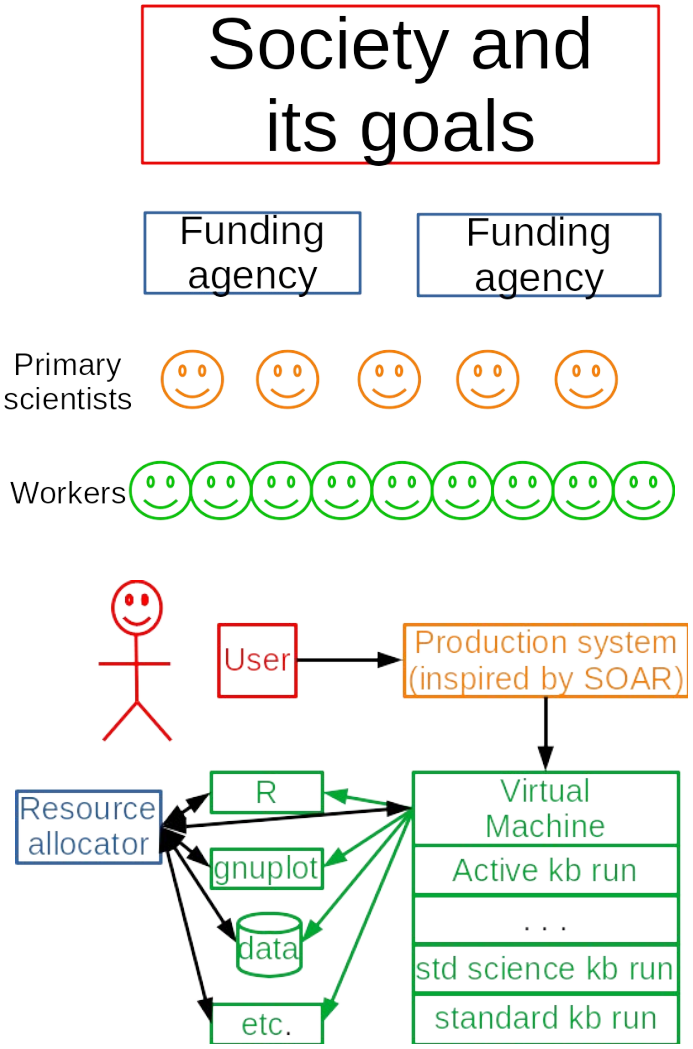
- Production System
  - **Purpose**: Question asking
  - **Serves as**: "*the literature*" + heuristics of what to try next
- Bypass-able
  - Can passively watch user, records results
  - In passive mode can say "*You've already tried that! These are the results . . .*"

The Funding Agency: the resource allocators



- Resource allocator
  - **Purpose:** Allocates (scarce?) computing resources
  - **Serves as:** Funding agency
- Authenticates user processes
- Allocates resources
  - Computational time
  - Memory
  - Access to data
  - Network access to remote resources

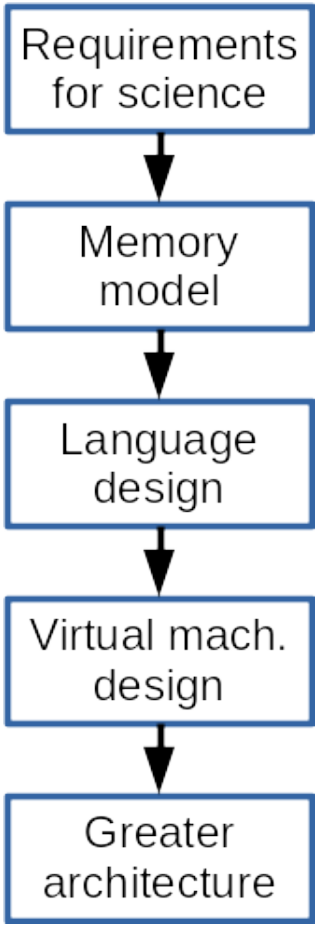
Society and Its Goals: Telling What is Important



- User
  - **Purpose:** Sets goals/policy for production system
  - **Serves as:** Society
- Can choose what to do under direct control

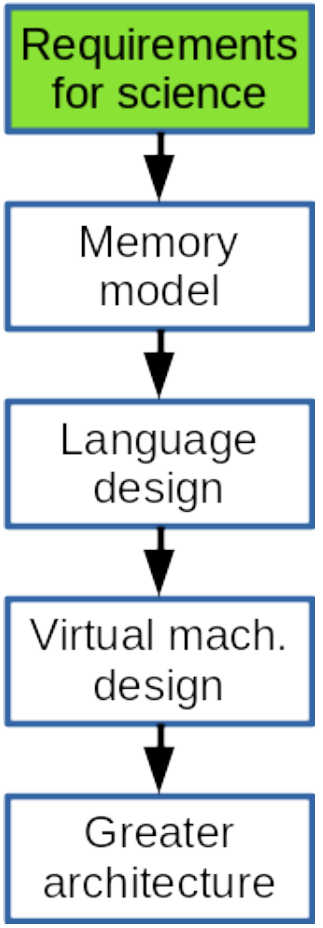
Towards a Better Architecture

- Circa 2012 – present
- A historically accurate account
  - rational way to do it
  - but I lucked upon it
- Design trajectory
  1. Requirements for science
  2. Memory model
  3. Language
  4. Virtual Machine
  5. Overall architecture



Requirement: Annotated Values

- Have values
  - units
  - dimensions
  - Limiting domains
- Examples
  - 9.8(*metersPerSecSqr*)
  - 299792458(*metersPerSecond*)
  - 6.022140e+23 (*inverseMol*)
  - 273.2(*kelvin*)
    - ***Can't be less than 0!*** (limit on domain)



Requirement: Justified Values

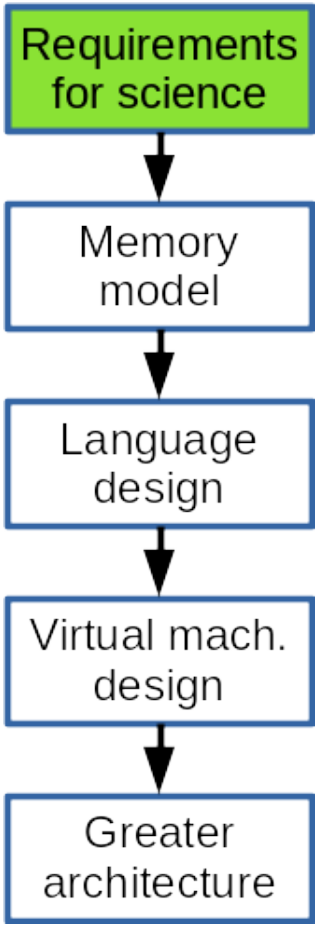
- Justifications keep track of where values came from:
  - Observation (e.g. "*What is Joe's mass?*")
  - By definition (e.g. 100 cm = 1 meter)
  - Calculation
- Calculation:
  - Truth preserving: (*e.g*. modus ponens, arithmetic)
  - Non-truth preserving (*e.g*. abduction)

```
JoeTellsJoesMass2020Jan22
[*ByMeasurement|
 `Joseph Phillips`,
 Mass,
 `Joseph Phillips`,
  ^Date{*2020,1,22*},
  `Joe's master bathrm`,
`Conair Corp Model WW404GD
scale`*];


80.51(*kgDomain*) <~
joeTellsJoesMass2020Jan22;
```

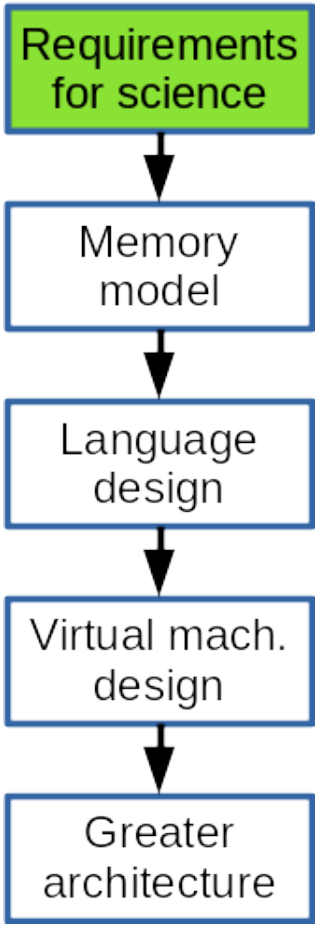
Requirement: Multiple Values

- Estimates of Age of the Earth
  - 6000 years (Ussher)
  - 75 Kya (Buffon)
  - “several billion” (de Maillet, Buffon)
  - $-\infty$? (Hutton, Lyell)
  - 100 Mya (Lord Kelvin)
  - 20-40 Mya (Lord Kelvin)
  - 3.4 Gya (Rutherford)
  - 4.6 Gya (Meyer)
  - 4.5 ± 0.3 Gya (Houterman)
- Potentially multiple answers per attribute
  - List from most believed to least so
  - Represent true multi-valued attributes as lists of lists

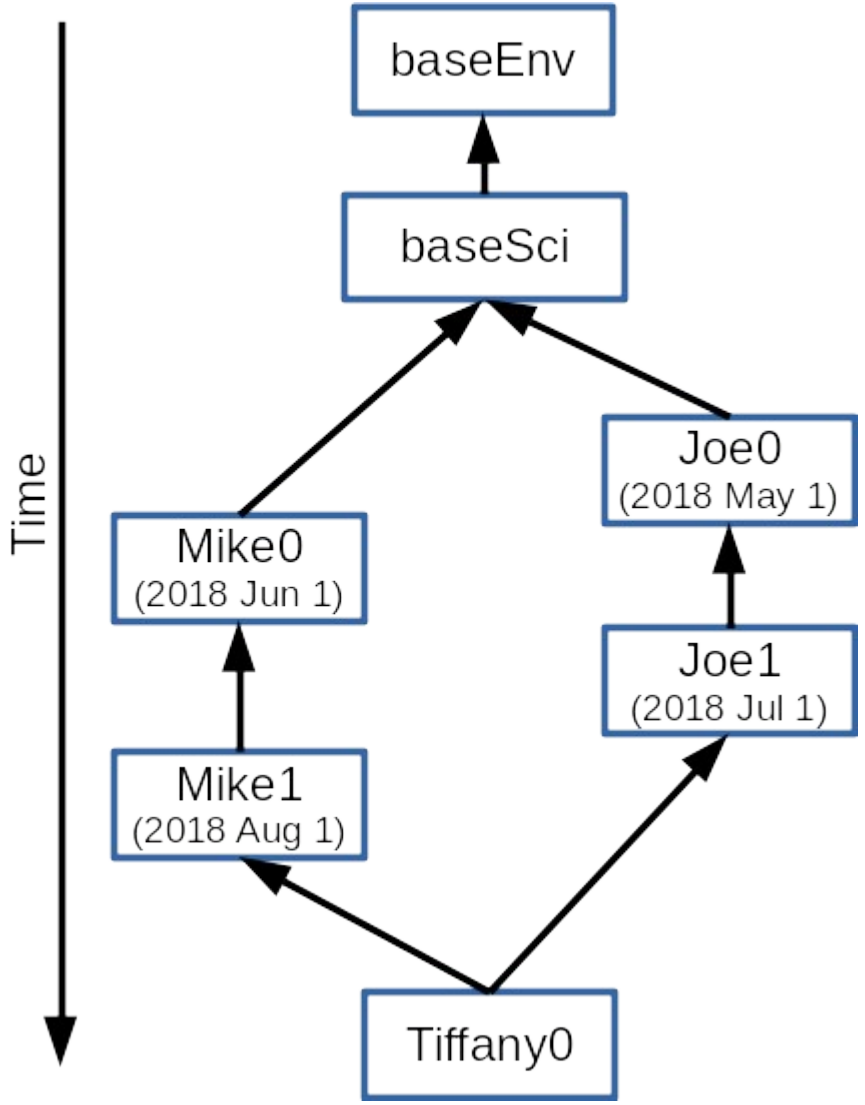


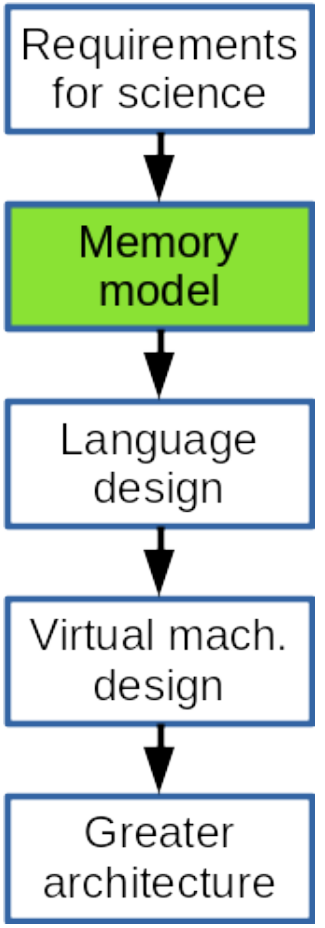
Requirement:
When in doubt, generalize

- Rationals > Integers
- Complex > Real
- Maps > Arrays
- Bags > Sets
- Iterators > Integer indices



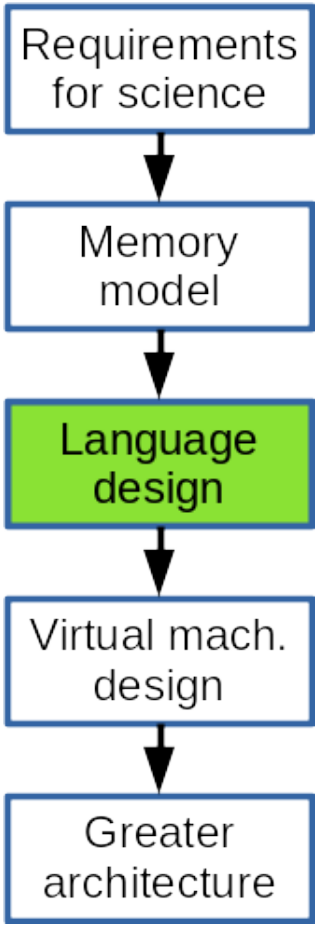
Memory model: Monotonic Knowledge Base





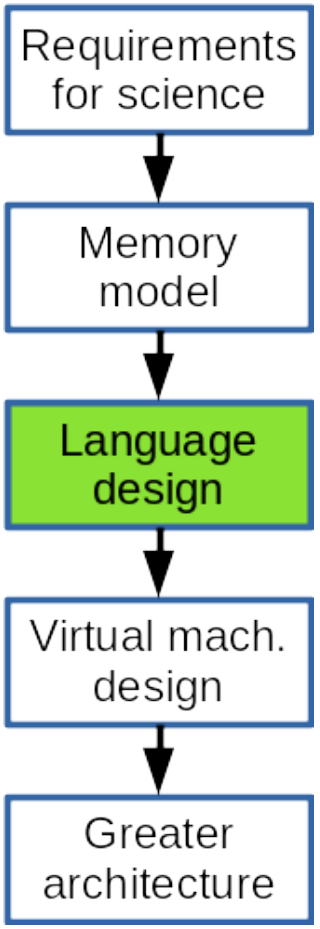
Language: Frame System

- In A.I. since 1970s
- Now really popular:
  - Object-Oriented Programming Languages
  - XML, JSON
- Even represents loops, conditionals and functions:
  - Need to represent anonymous objects



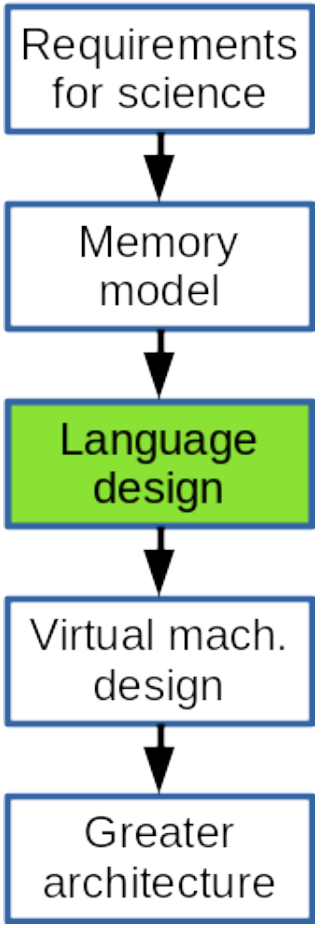
Language

```
`Mercury (planet)`
{*
  instanceOf->assertZ(Planet);
  ^SummaryTextA[*toEnglish*]->
    assertZ("Mercury (0.4 AU from the Sun)...");
  imageFilenameListA->
    assertZ([ "mercury.jpg"] );
  massA->
    assertZ(3.3022e+23(*kilograms*) );
  aphelionA->
    assertZ(69816900(*kilometers*));
  perihelionA->
    assertZ(46001200(*kilometers*));
  orbitalPeriodA->
    assertZ(87.9691(*days*));
*};
```

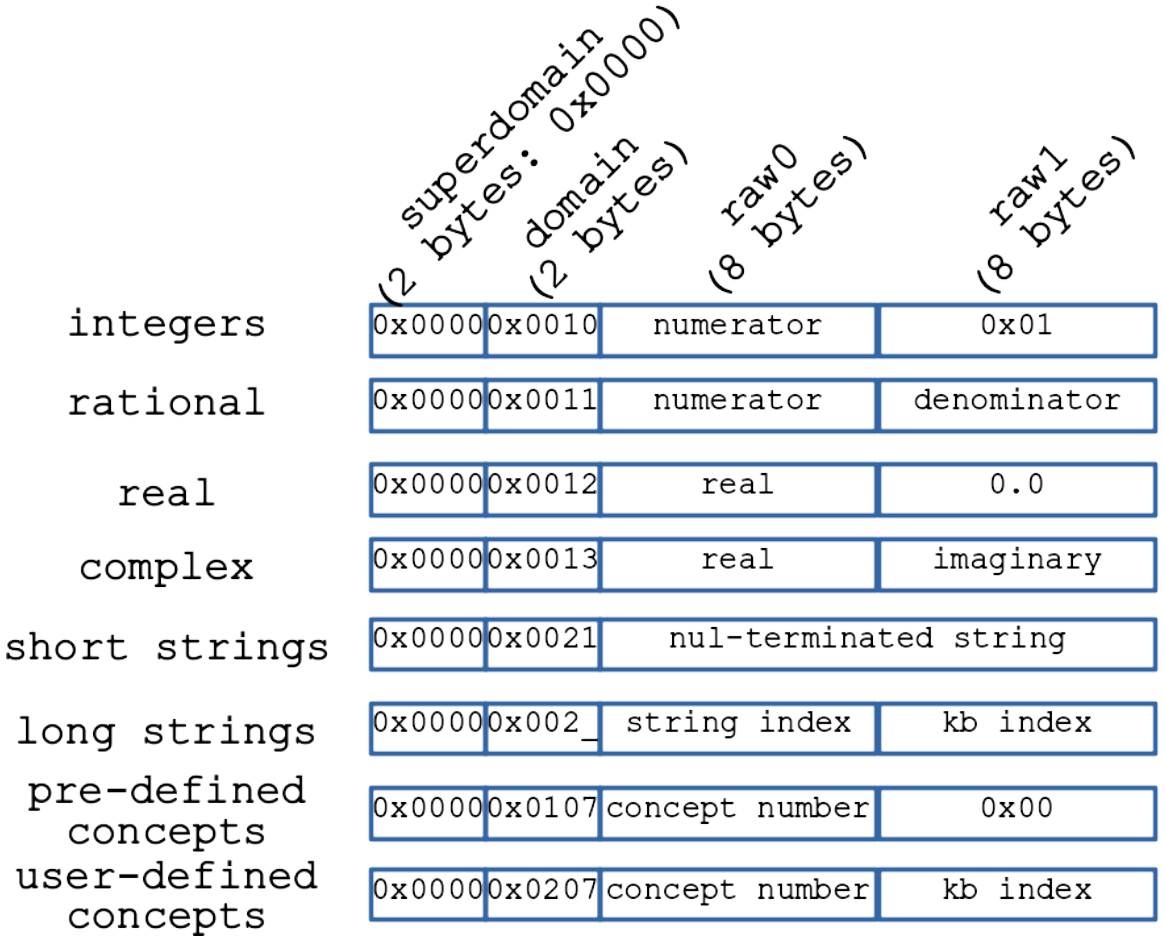


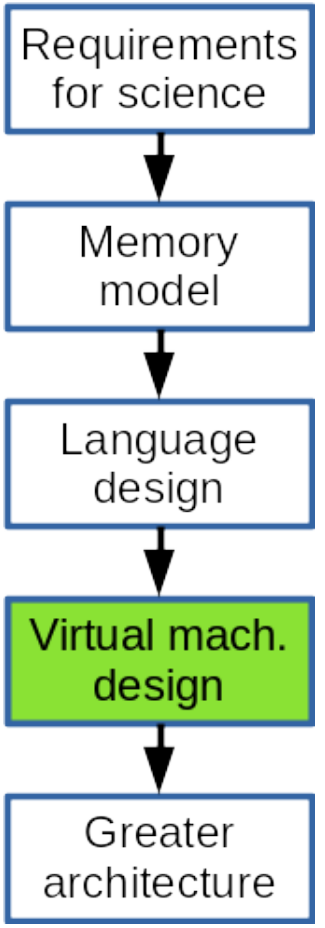
Language

```
^Do
[*[
   ^VarDecl[*@i,Rational*],
   ^For
   [*
      @i := 0,
      @i <  10,
      stdOut->printLn(@i),
      @i := @i + 1
   *]
]*];
```



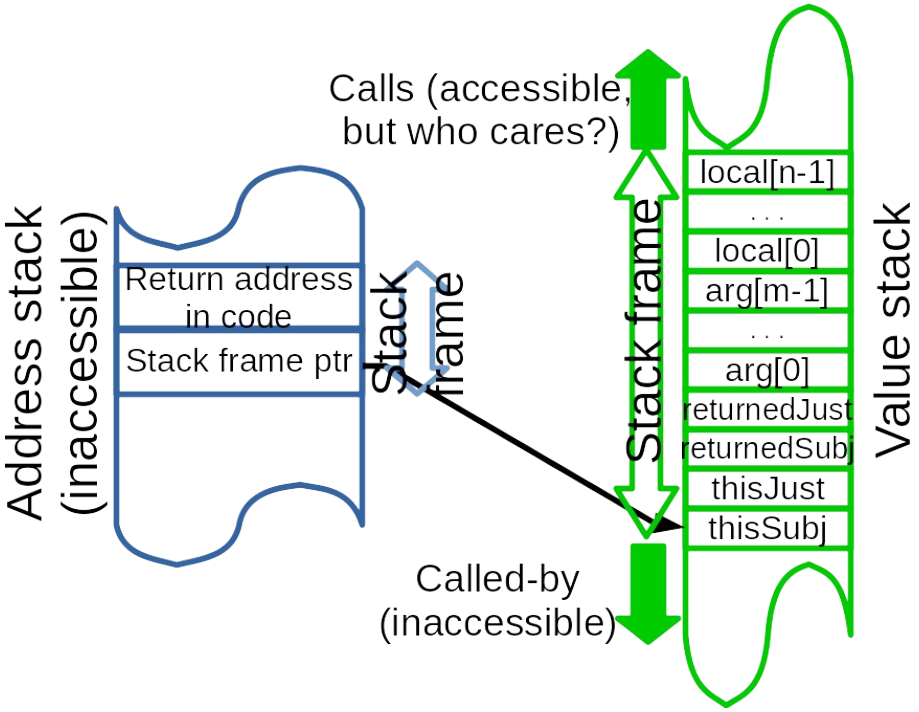
Virtual Machine: machine word





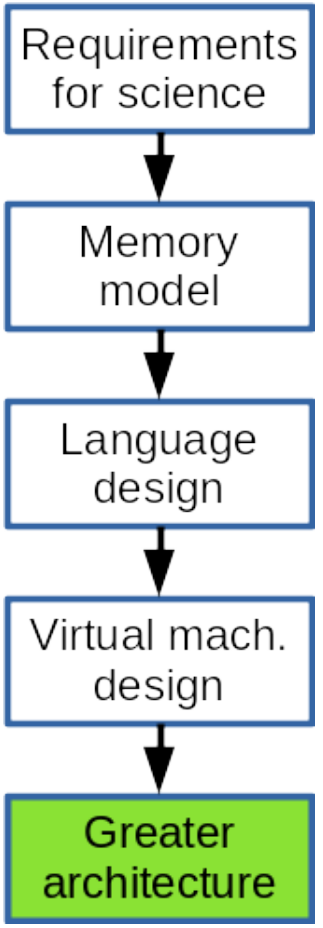
Virtual machine: stack frame

- Stack-based
  - Similar to Java Virtual Machine
- Value Stack
  - grows up
  - unaware of address stack
- Instructions can only get values above the stack
  - impossible to get data from who called you
- kb as a whole acts as heap



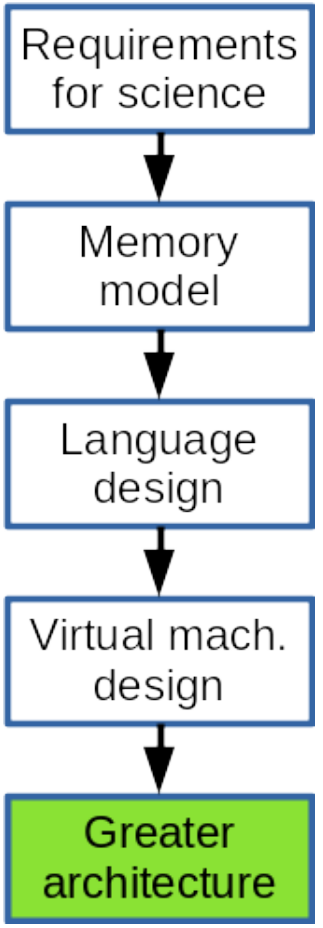
Specialized Programs: Follow the Procedure

- For use when this with established procedure
- Call specialized algorithm
  - Statistics
  - Algebra
  - Plotting



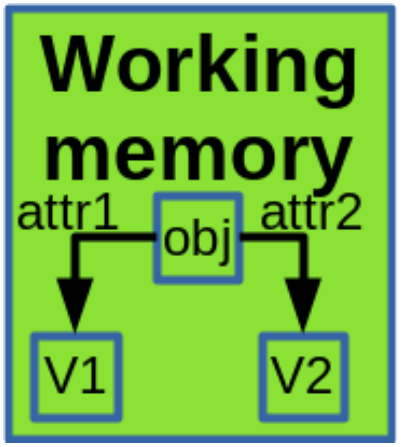
Production System: What questions follow from the last result?

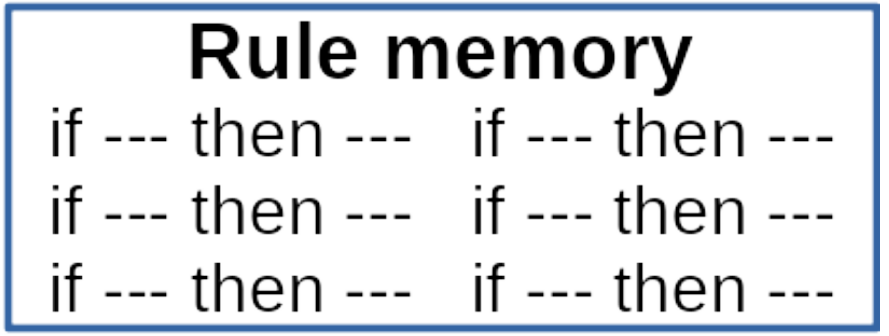
- For question asking
- Inspired by SOAR
  - Generalization of human (and robot) computation
- Our issues are a little different



Production System: What questions follow from the last result?

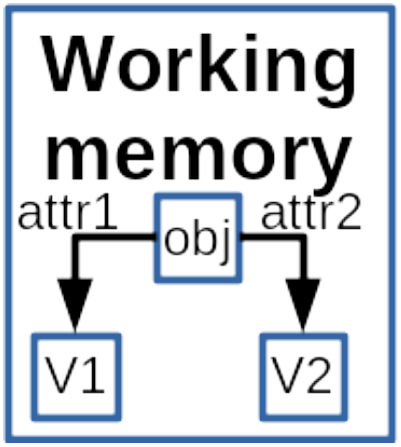
- **Working memory**
- **Traditionally**:
  - what you hold in your consciousness
- **For Us**: “the literature”
  - Memory of what has been tried, and how well or poorly it worked

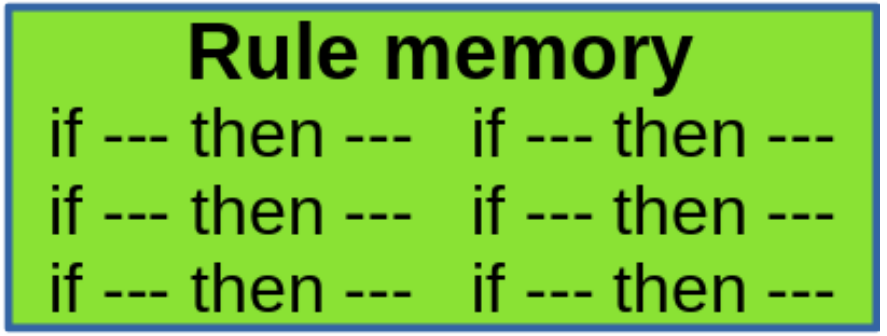




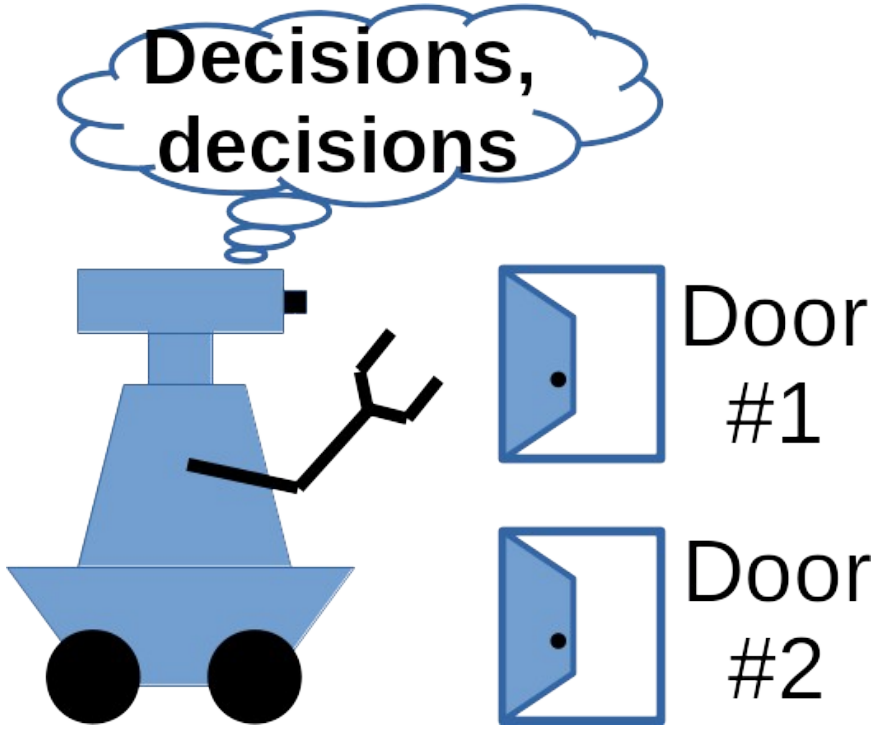
Production System: What questions follow from the last result?

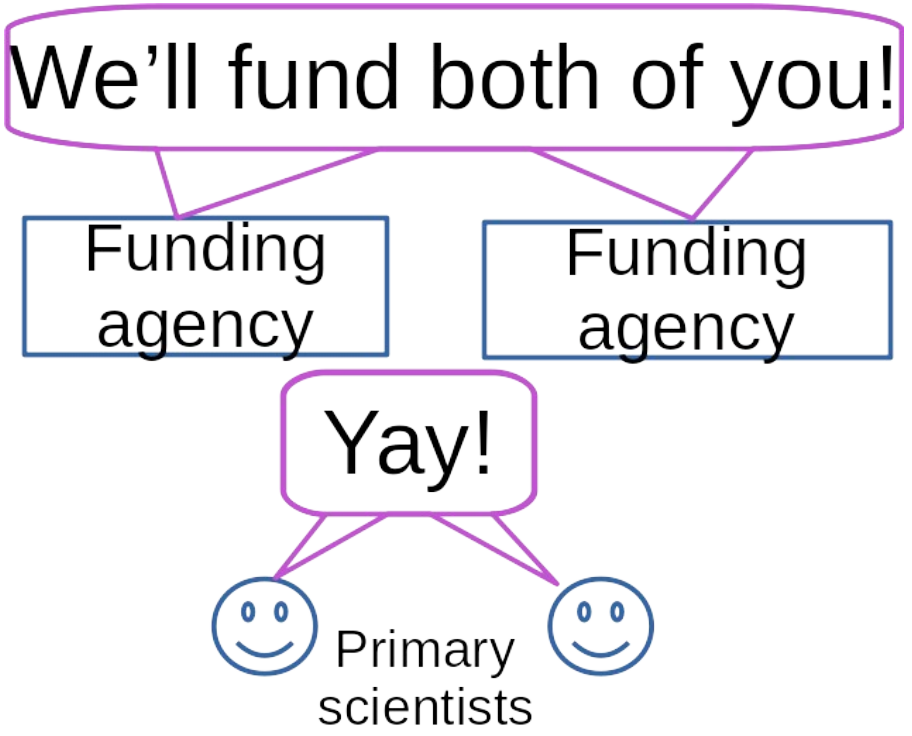
- **Rule memory**
- **Traditionally and For Us**: New idea generators
  - Heuristics about what is worth researching





Production System: What questions follow from the last result?





- A robot can only commit to one path at a time
- Funding agencies can (and do) try multiple paths simultaneously