

Knowledge Graph

A Material Graph Database

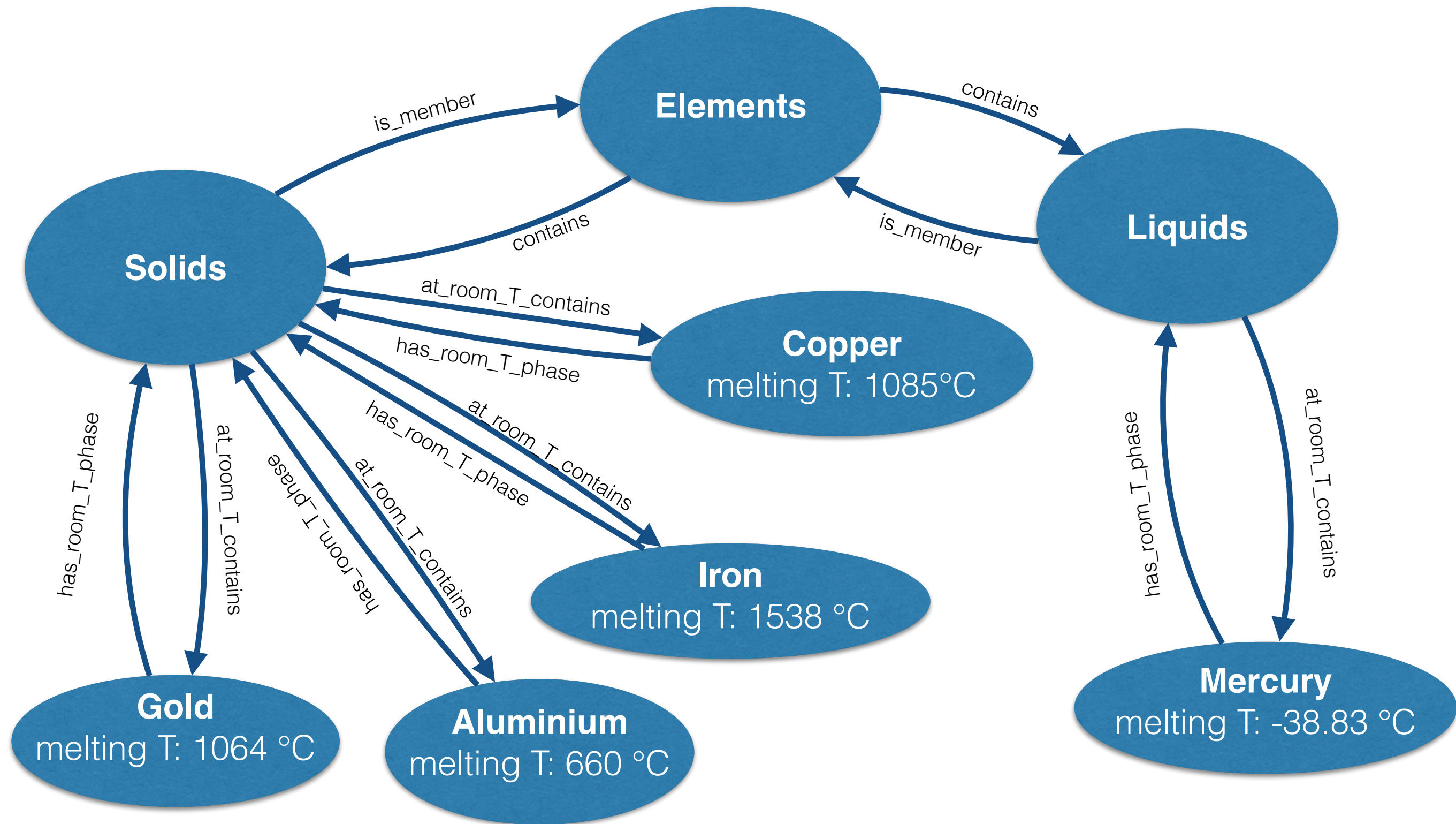
Different structures

- Database
- Graph

Database

ID	NAME	MELTING T	BOILING T	MASS	...
1	Copper	1085 °C	2560 °C	63.546u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
5	Gold	1064 °C	2970 °C	196.97u	...

Graph



Searching

- Database approach
- Graph approach

Database

ID	NAME	MELTING T	BOILING T	MASS	...
1	Copper	1085 °C	2560 °C	63.546u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
5	Gold	1064 °C	2970 °C	196.97u	...

Database

ID	NAME	MELTING T	BOILING T	MASS	...
1	Copper	1085 °C	2560 °C	63.546u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
5	Gold	1064 °C	2970 °C	196.97u	...

Database

ID	NAME	MELTING T	BOILING T	MASS	...
1	Copper	1085 °C	2560 °C	63.546u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
5	Gold	1064 °C	2970 °C	196.97u	...

linear search

Database

ID	NAME △	MELTING T	BOILING T	MASS	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
1	Copper	1085 °C	2560 °C	63.546u	...
5	Gold	1064 °C	2970 °C	196.97u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...

Database

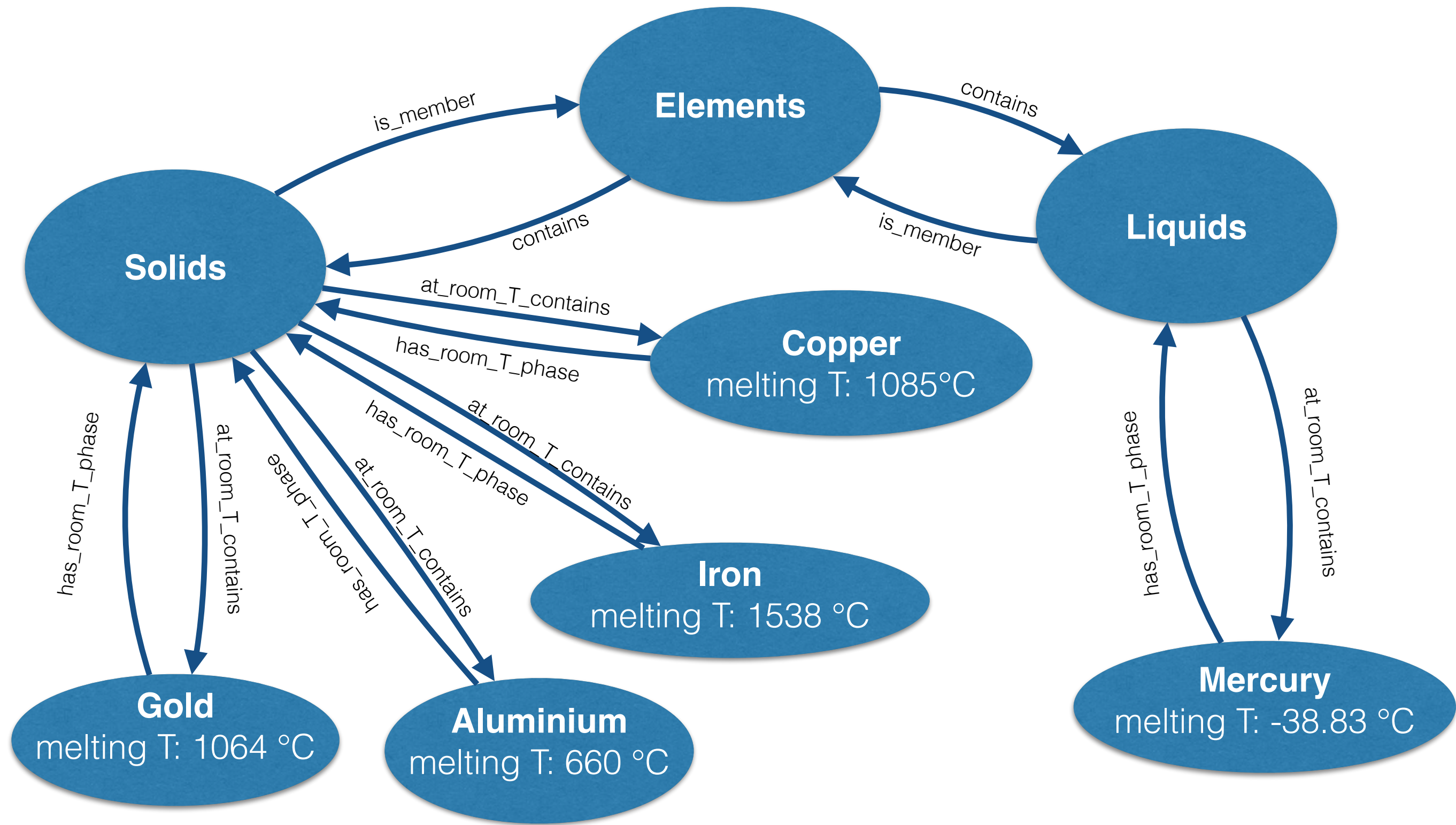
ID	NAME △	MELTING T	BOILING T	MASS	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
1	Copper	1085 °C	2560 °C	63.546u	...
5	Gold	1064 °C	2970 °C	196.97u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...

Database

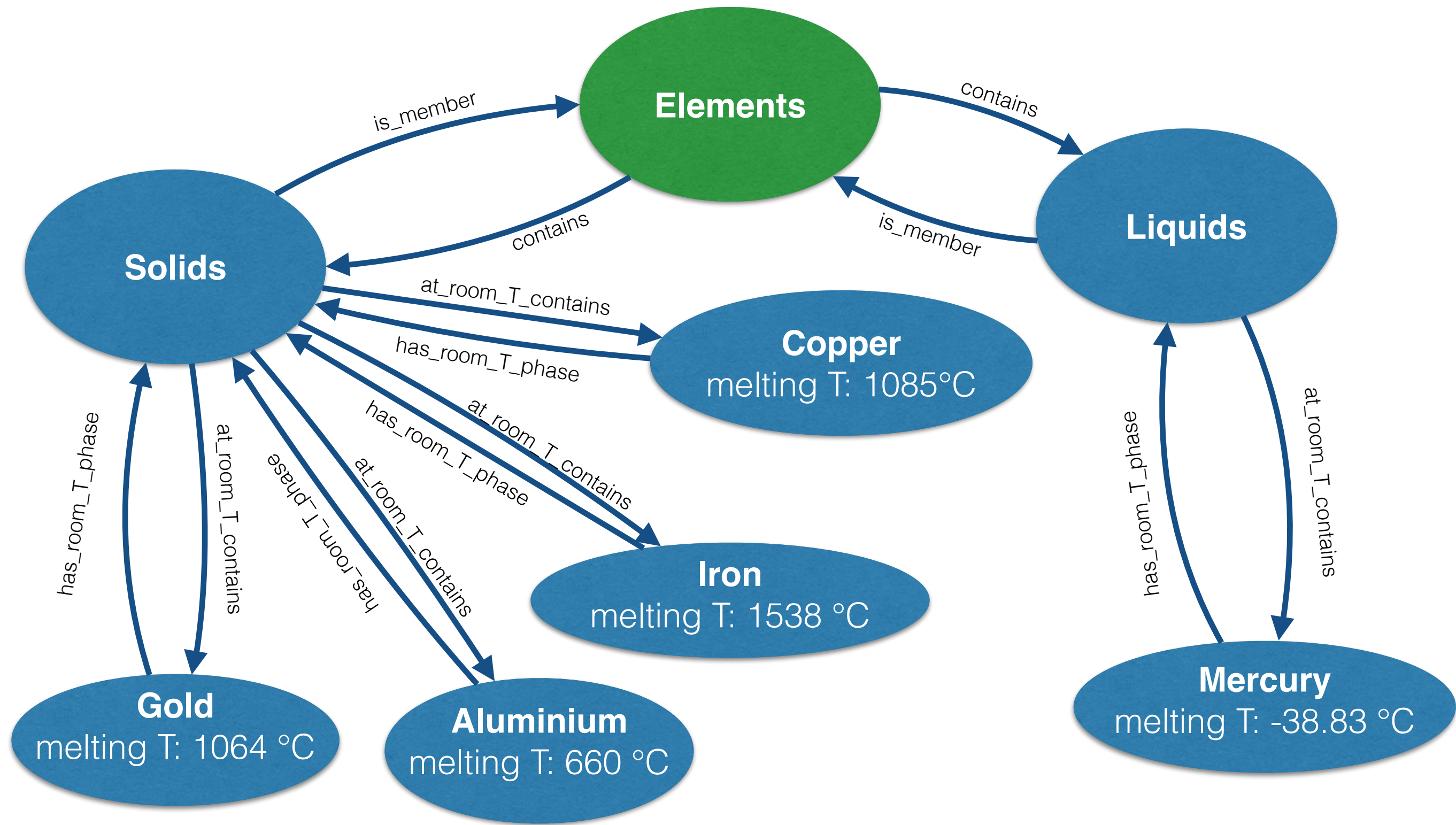
ID	NAME △	MELTING T	BOILING T	MASS	...
4	Aluminium	660.3 °C	2519 °C	26.98u	...
1	Copper	1085 °C	2560 °C	63.546u	...
5	Gold	1064 °C	2970 °C	196.97u	...
2	Iron	1538 °C	2862 °C	55.845u	...
3	Mercury	-38.83 °C	356.7 °C	200.59u	...

binary search

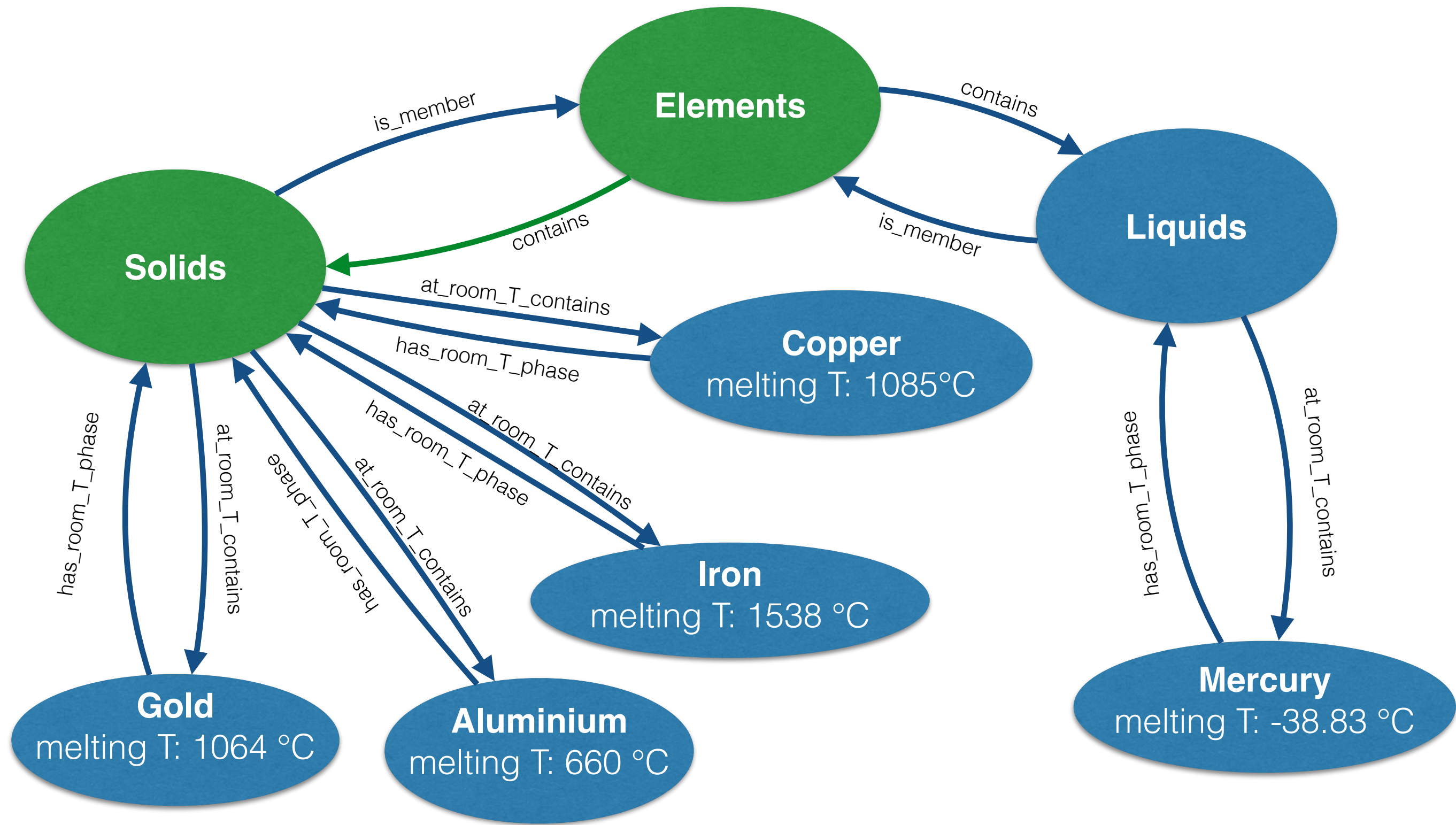
Graph



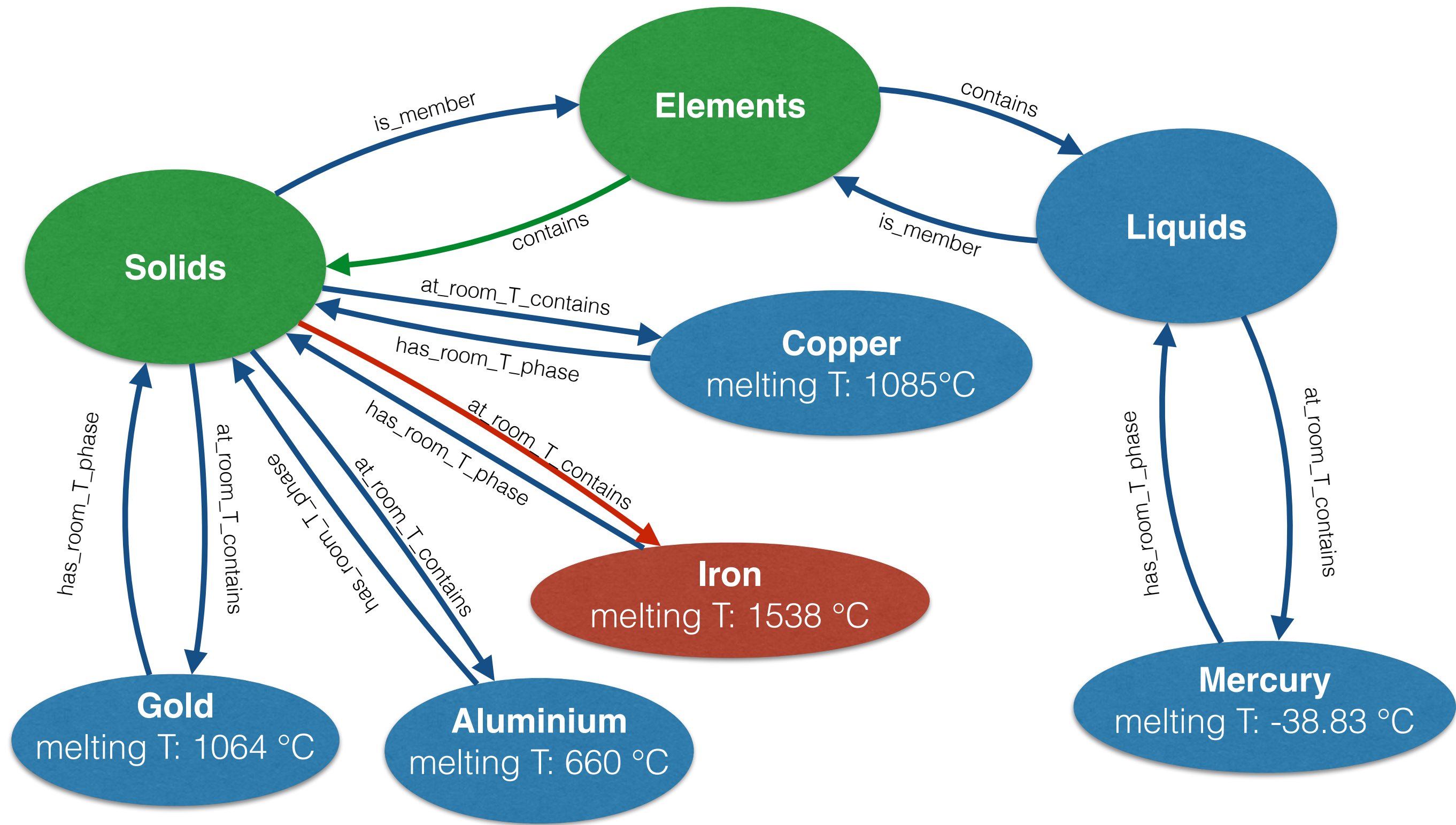
Graph



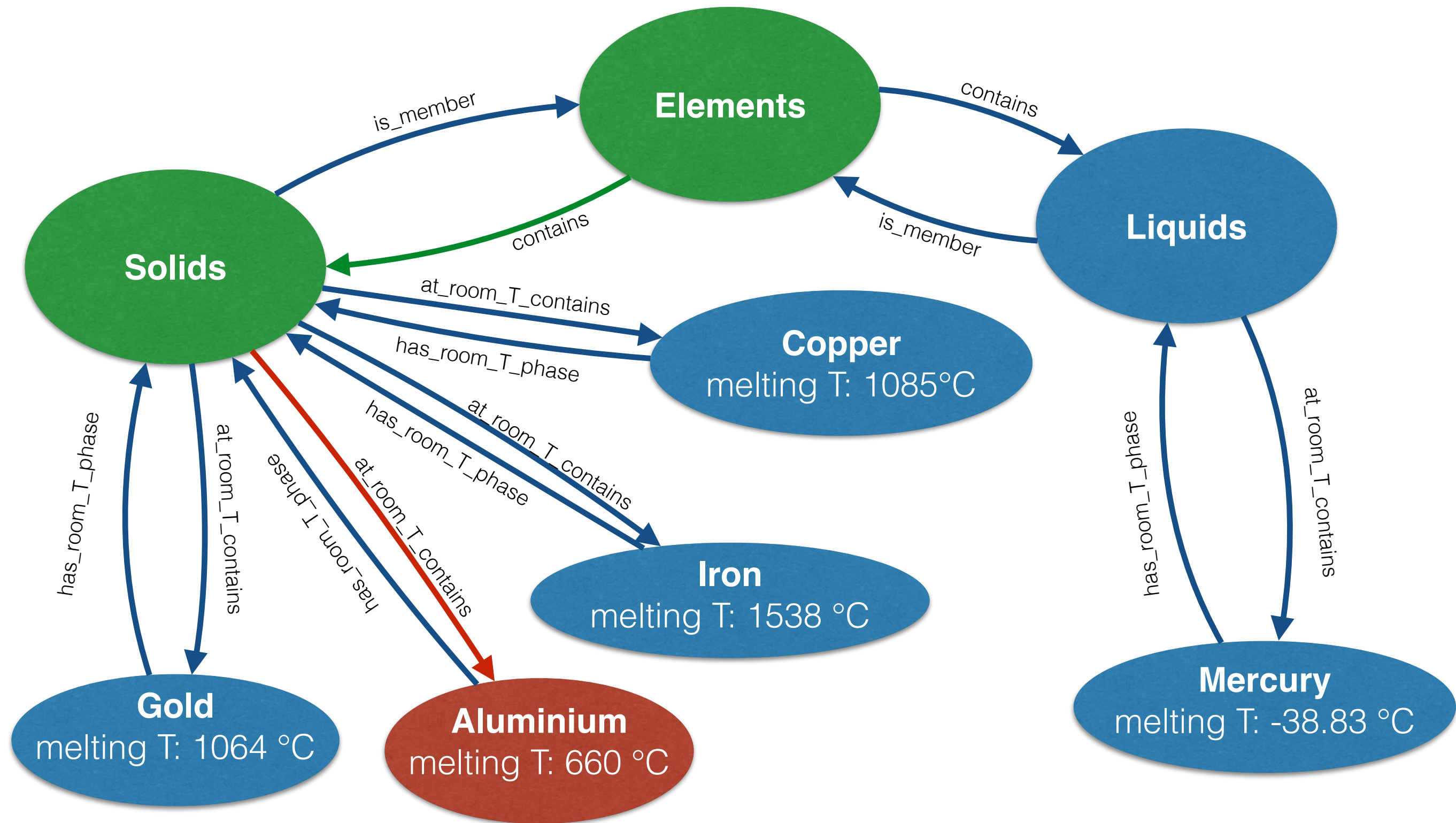
Graph



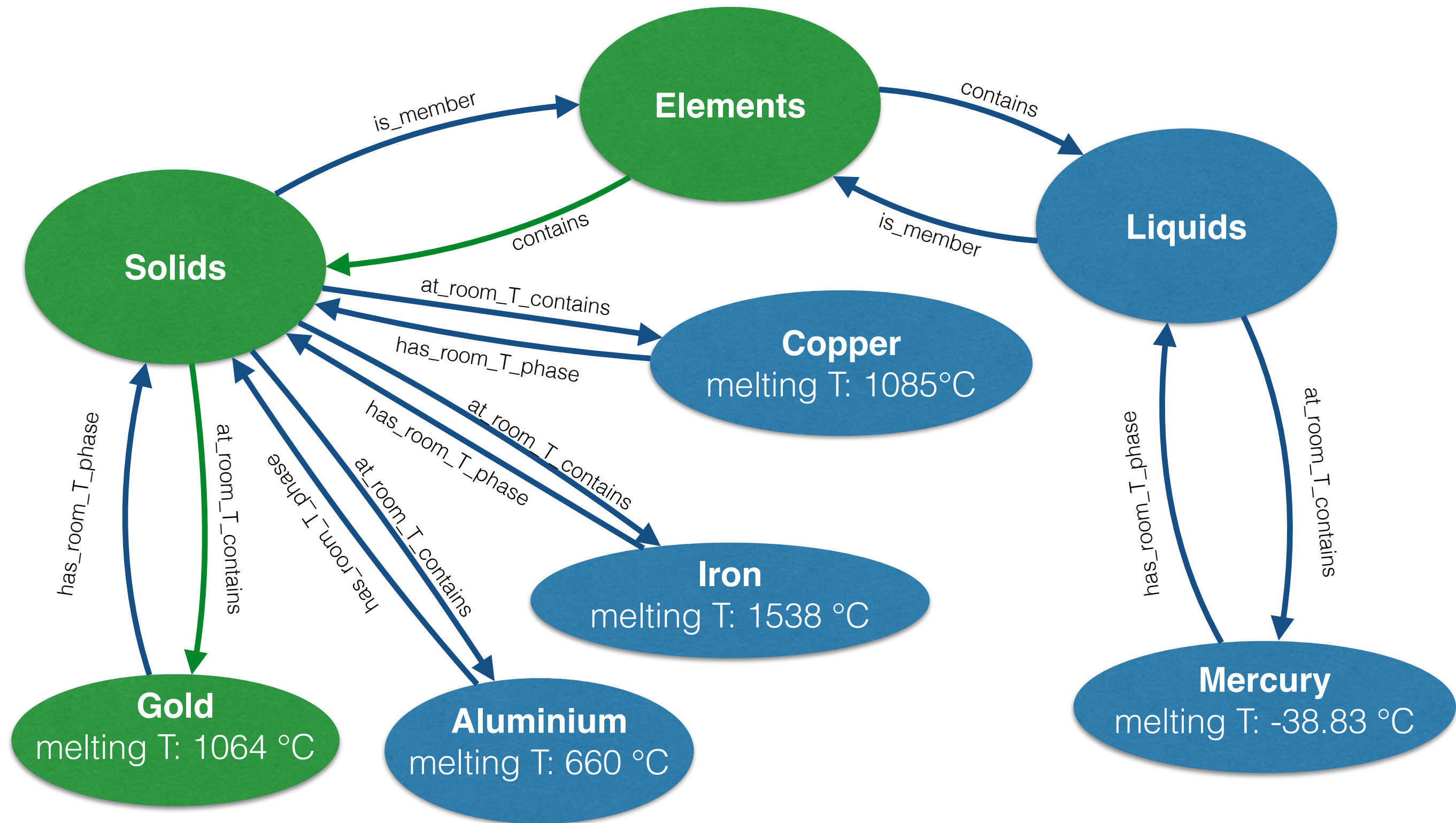
Graph



Graph



Graph



Searching

- Database approach
 - linear search
 - binary search only in indexed columns
- Graph approach
 - graph traversal : depends a lot on the graph structure!

Project plan

- Fetch material DB (FreeBase.com, Python)
- Implement graph library (C++)
- Visualize the graph
- Optimize graph search
- (if time allows) Parallelize graph traversal

Data description

C++ graphlib

- ...describe graph / node layout

Visualization

- Demo

Query

Find the N elements with melting point closer X

Graph traversal

exact

Element 1

Element 2

Element 3

Element 4

Element 5

Element 6

Element 7

Element 8

Element 9

Element 10

Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=3

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=3

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=4

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Similarity links

Graph traversal

exact

Element 1

Element 2

Element 3

Element 4

Element 5

Element 6

Element 7

Element 8

Element 9

Element 10

Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=3

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Graph traversal

exact

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=2

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=3

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

depth=4

Element 1
Element 2
Element 3
Element 4
Element 5
Element 6
Element 7
Element 8
Element 9
Element 10
Element 11

Parallelization

- ... load imbalance
- show scaling

Parallel random walk

- show convergence
- show scaling

TODO

Credits

- Data:
 - <http://www.freebase.com>
- Visualization:
 - <http://visjs.org>