SHORTENING A TEST FOR SKILL ASSESSMENT

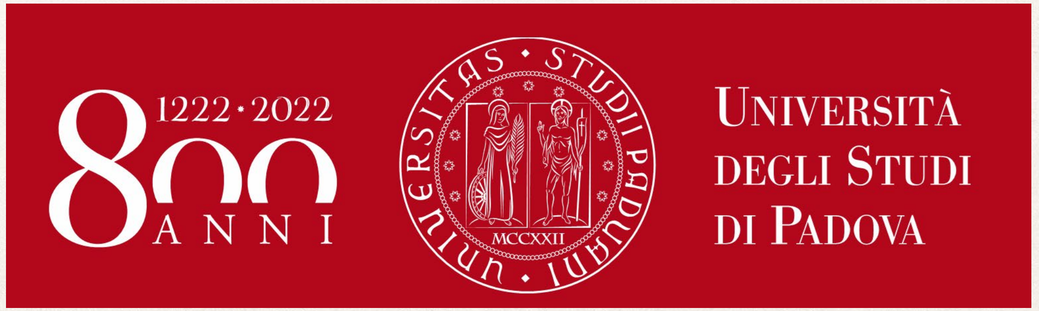
Students: Anna Panzeri, Cheyenne Cavender, Laura Van Hove, Teresa Gehrig, Christopher Schaar

Supervisor: Prof Pasquale Anselmi

KU LEUVEN

JACOBS
UNIVERSITY

UNIVERSITY OF AMSTERDAM



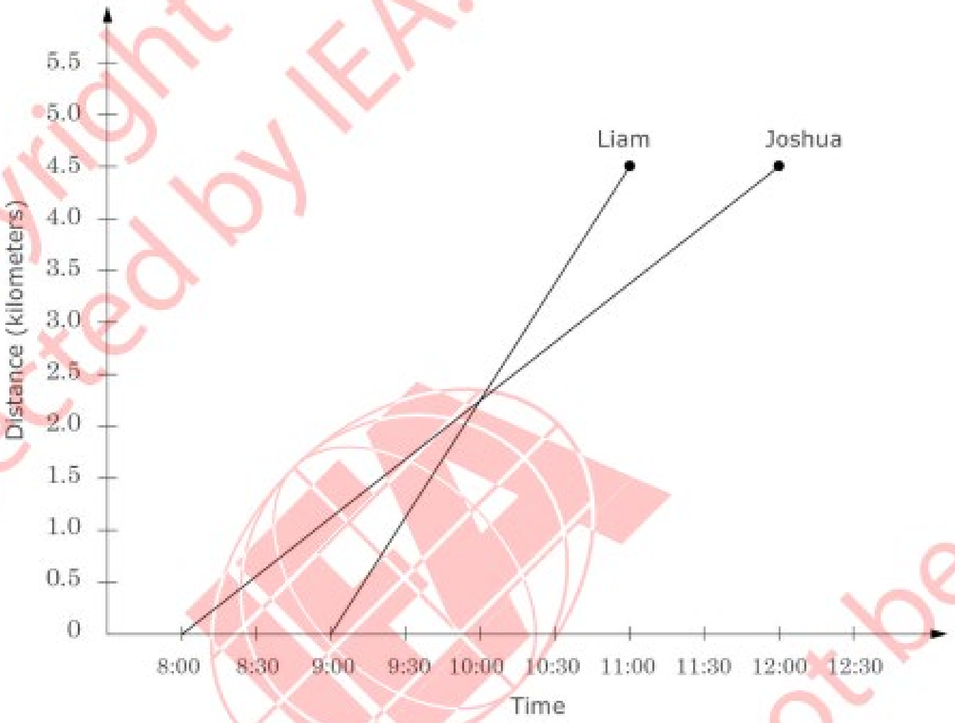
TIMSS, 2003

Trends In International Mathematics And Science Study (TIMSS), is an international assessment tool, that measures students' math, science and reading performance

(https://timssandpirls.bc.edu/)

Excerpt —->

The graph represents the distance and time of a hike taken by Joshua and Liam.



If they both started from the same place and walked in the same direction, at what time did they meet?

(A) 8:00
(B) 8:30
(C) 9:00
(D) 10:00
(E) 11:00

M012025

1.
Create a skill map:
29 items * 14 skills

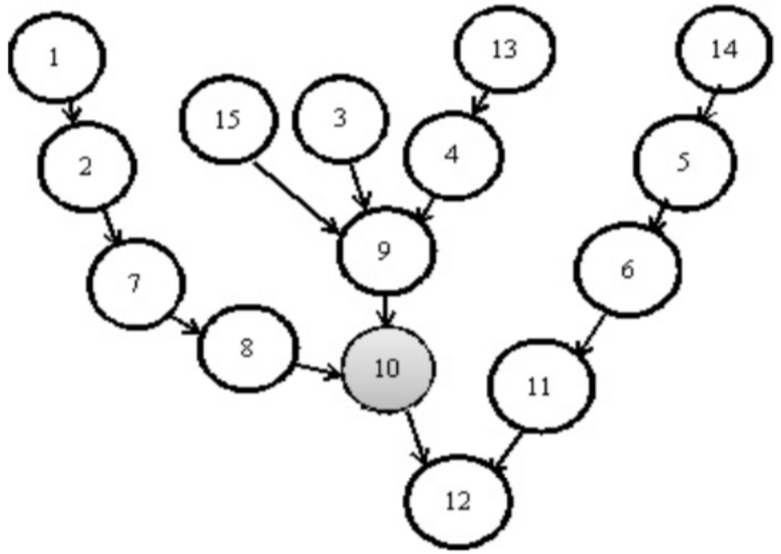


Figure 6: Hierarchical relationship among the attributes for booklet 1

```
map <- matrix( c(
 1,  1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1,
 2,  0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
 3,  0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0 ,
 4,  0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0 ,
 5,  1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 6,  0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0 ,
 7,  0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 8,  0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0 ,
 9,  0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0 ,
 10, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0 ,
 11, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 12, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 13, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0 ,
 14, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0 ,
 15, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 16, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 17, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1 ,
 18, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0 ,
 19, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 20, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0 ,
 21, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0 ,
 22, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1 ,
 23, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 24, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ,
 25, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0 ,
 26, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0 ,
 27, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0 ,
 28, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0 ,
 29, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0) ,
          29, 15, byrow = TRUE)
map
```

The Procedure (1)

1.
Create a skill map

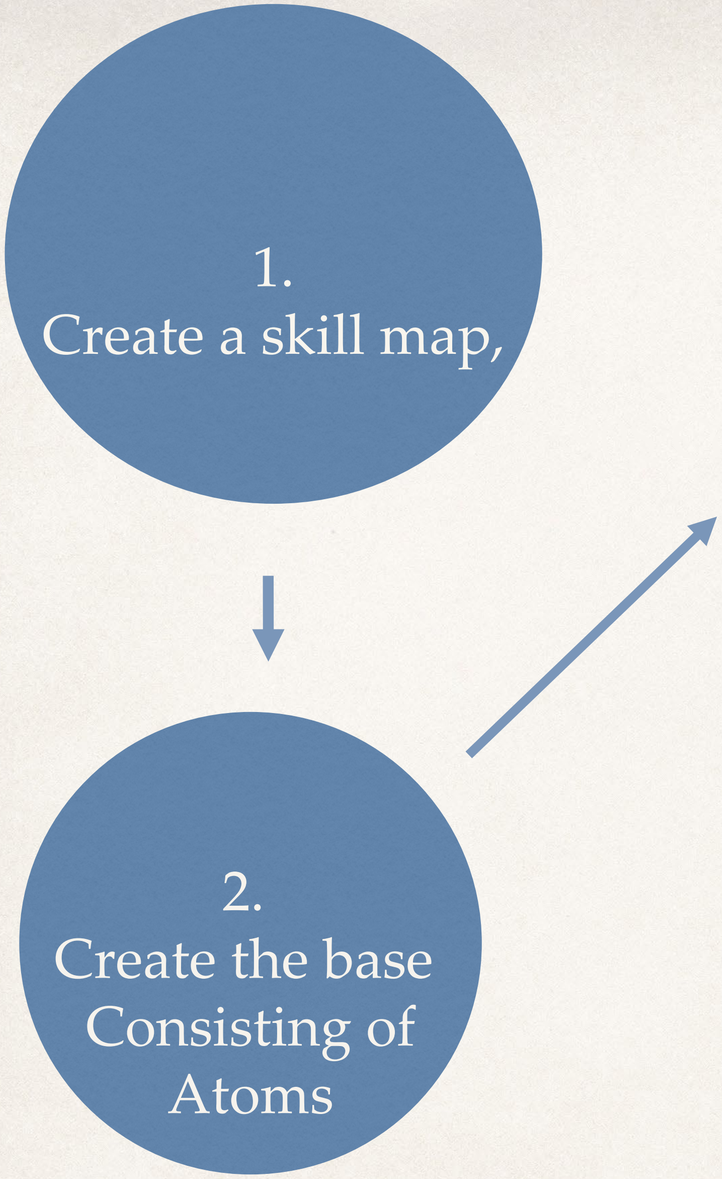
2.
Create the **base** of the knowledge Space: by looking at the graph with the 14 skills
The Base Consists of Atoms

```
#  1  2  3  4  5  6  7  8  9 11 12 13 14 15
base <- matrix (c(
   1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
   1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
   1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
   1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0,
   1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0,
   0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,
   0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1,
   0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1,
   0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
   0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0,
   0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0,
   0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
   0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
   0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0,
   0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0,
   0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
   0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0,
   0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0,
   0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0,
   0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0),
   ncol=14, byrow= TRUE)
```

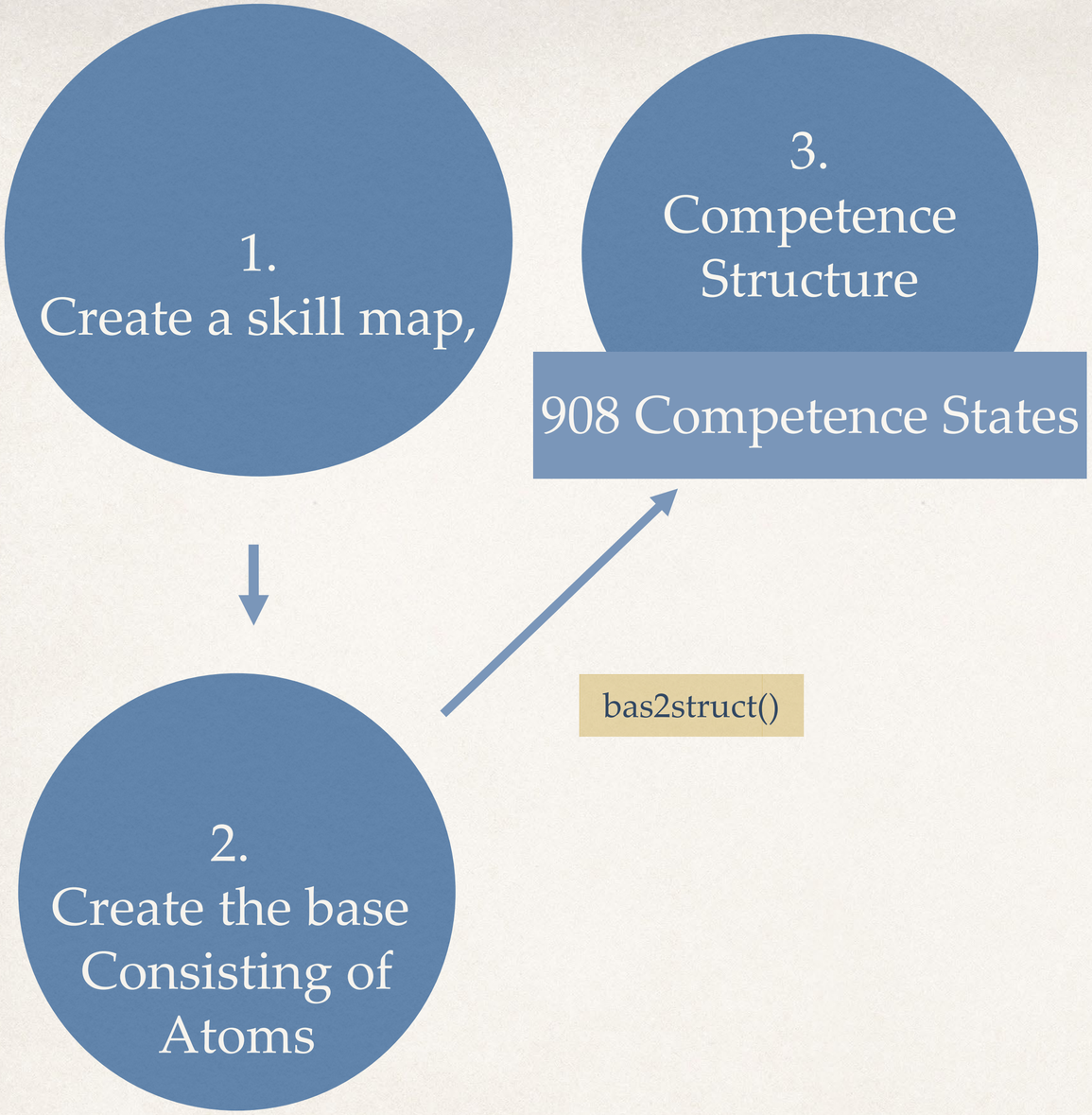
The Procedure (2)

CLOSURE UNDER UNION: QHELP EXAMPLE

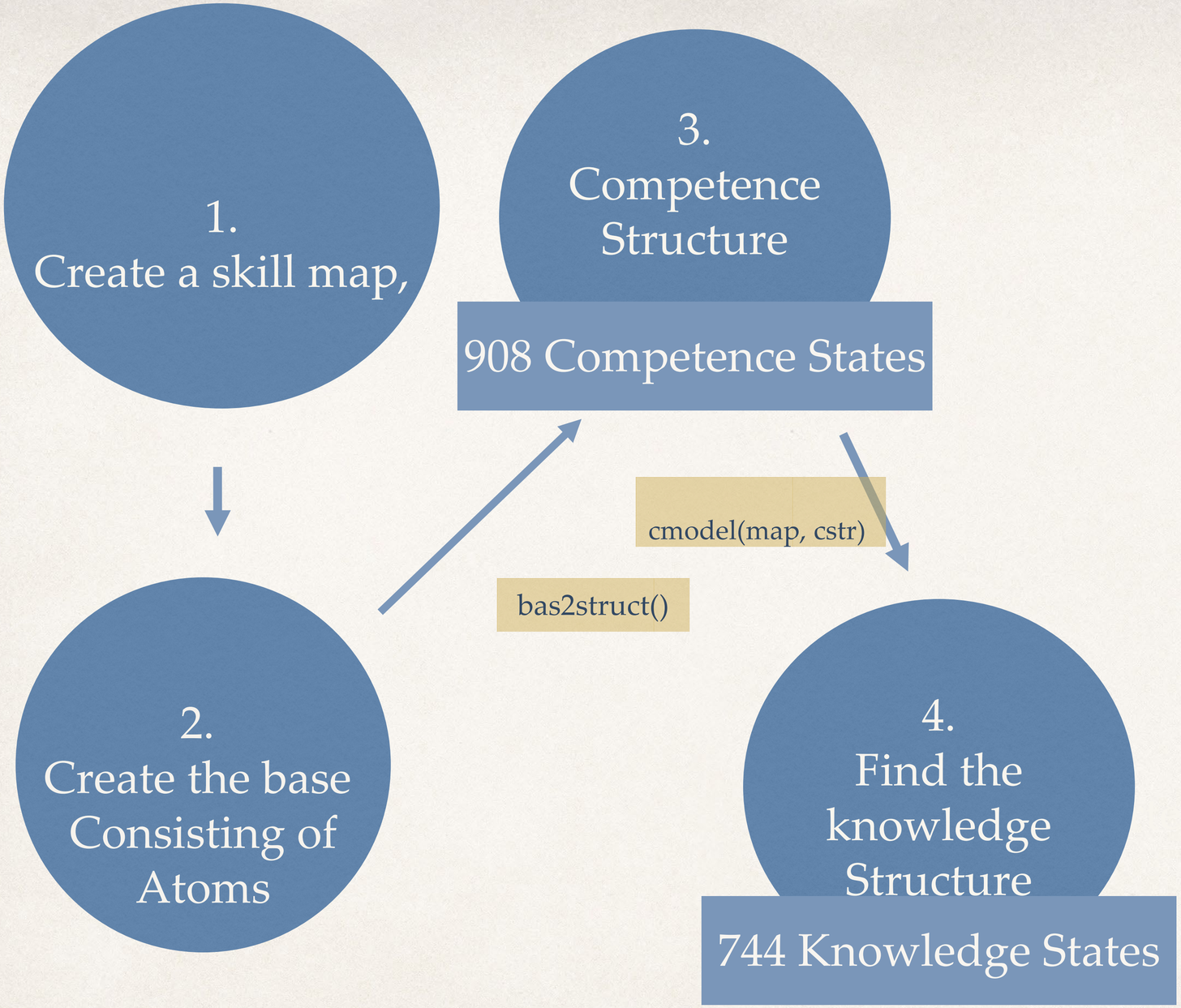
Not closed under union	Closed under union
R	R
R, JASP	R, JASP
Psychophysics	Psychophysics
-	R, Psychophysics
-	R, Psychophysics, JASP



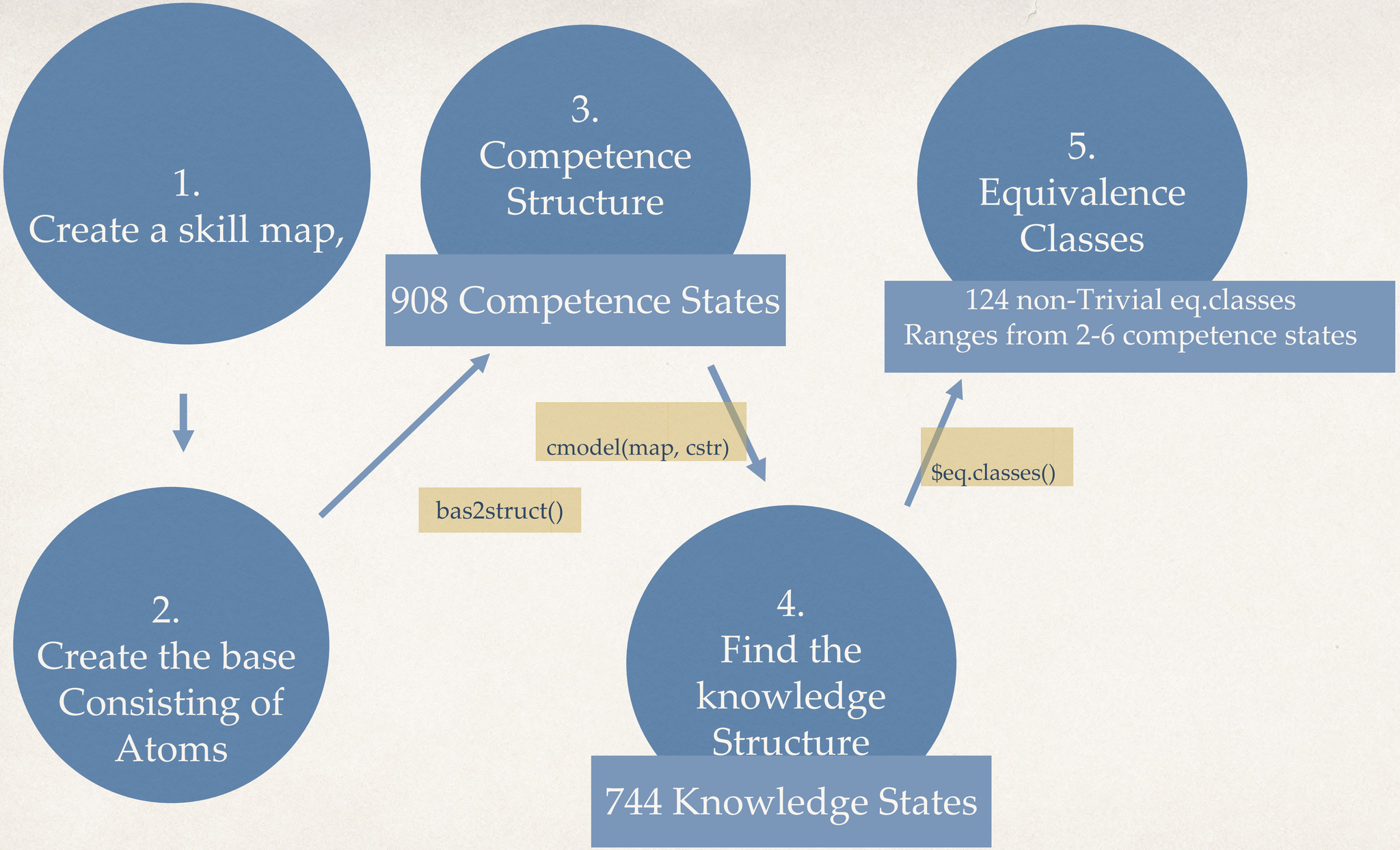
The Procedure (3)



The Procedure (3)



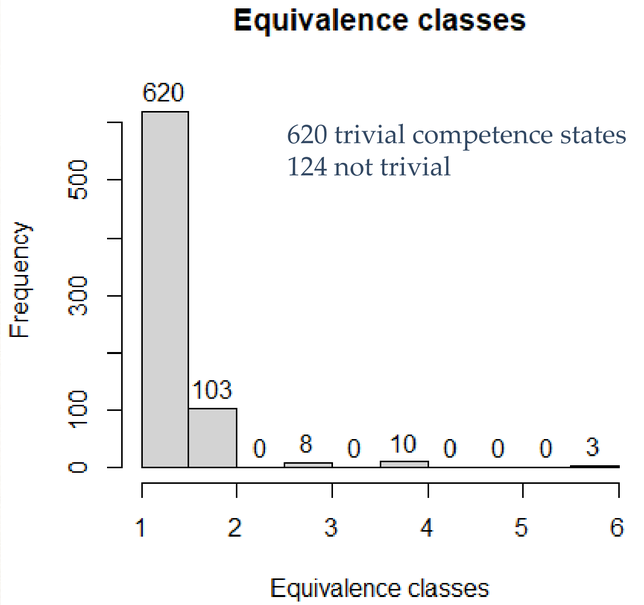
The Procedure (3)



The Procedure (3)

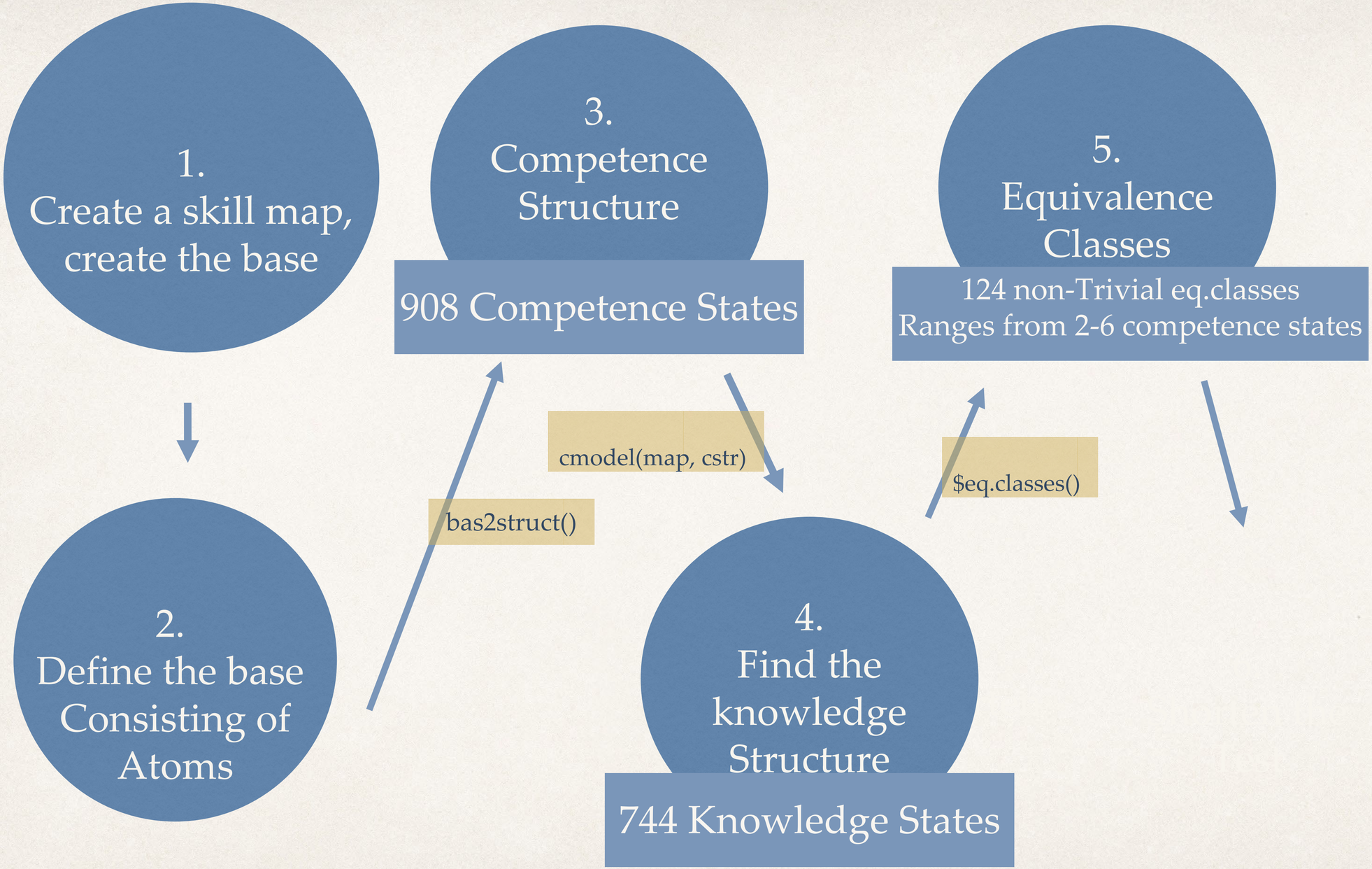
EQUIVALENCE CLASSES

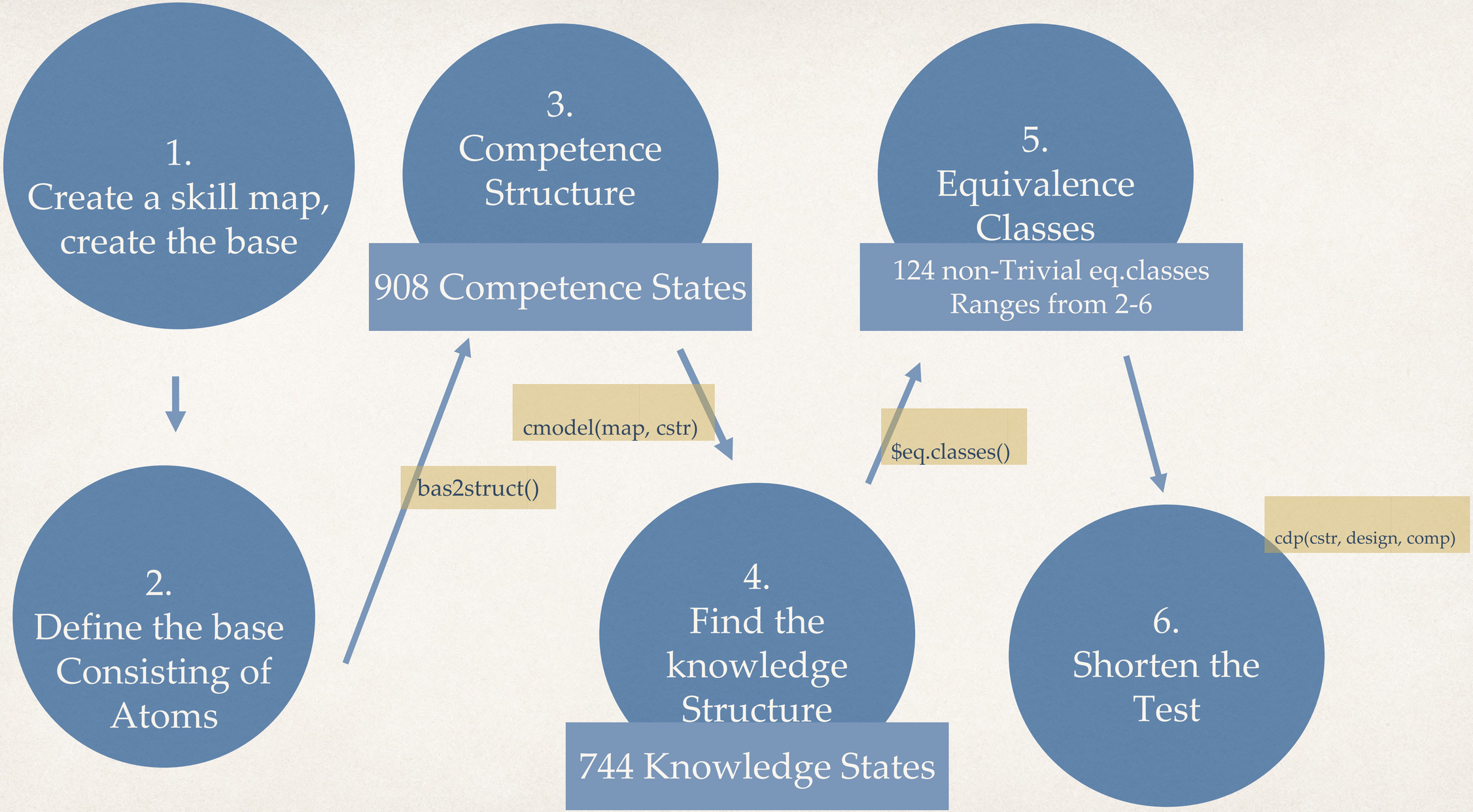
When different **competence states** delineate the same **knowledge state**.



EX:
5 students have
5 different set of **abilities**
(competence states)..
..but they all result in
solving the same **exercise**
(knowledge state)

So, we cannot know their real abilities!


1.
Create a skill map, create the base
2.
Define the base Consisting of Atoms
3.
Competence Structure
908 Competence States
4.
Find the knowledge Structure
744 Knowledge States
5.
Equivalence Classes
124 non-Trivial eq.classes
Ranges from 2-6 competence states
cmodel(map, cstr)
bas2struct()
$eq.classes()


1.
Create a skill map, create the base
2.
Define the base Consisting of Atoms
3.
Competence Structure
908 Competence States
bas2struct()
cmodel(map, cstr)
4.
Find the knowledge Structure
744 Knowledge States
5.
Equivalence Classes
124 non-Trivial eq.classes
Ranges from 2-6
$eq.classes()
6.
Shorten the Test
cdp(cstr, design, comp)

Conclusion

- We shortened the test **from 29 to 17 items**
- Examining **17 Competencies**, as **informative as the original test**
- From deterministic point of view the test is as **good as the original one**. In practice smaller tests are less reliable than longer ones.
- Because when careless errors and lucky guesses occur, shorter tests are less reliable than longer ones, meaning that it is more difficult to recover the true competence state of the individual from his/her item responses

Why Improve Tests?

Main References

Anselmi, P., Heller, J., Stefanutti, L., & Robusto, E. (2022). Constructing, improving, and shortening tests for skill assessment. Journal of Mathematical Psychology, 106, 102621.

Heller, J., Stefanutti, L., Anselmi, P., & Robusto, E. (2015). On the link between cognitive diagnostic models and knowledge space theory. Psychometrika, 80(4), 995-1019.

Spoto, A., Stefanutti, L., & Vidotto, G. (2010). Knowledge space theory, formal concept analysis, and computerized psychological assessment. Behavior Research Methods, 42, 342-350.

Su, Y. L., Choi, K. M., Lee, W. C., Choi, T., & McAninch, M. (2013). Hierarchical cognitive diagnostic analysis for TIMSS 2003 mathematics. Centre for Advanced Studies in Measurement and Assessment, 35, 1-71.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Pictures taken from: www.pexels.com (cc0 licensing).

Special thanks to Prof. Pasquale Anselmi!

University
of Pasquale

Table 4. Q-Matrix of Booklet 1 for the Eighth Grade TIMSS 2003 Mathematics Test

	Item\Attribute	1	2	3	4	5	6	7	8	9	11	12	13	14	15	Sum
1	M012001	1	0	0	0	0	0	0	0	0	0	0	1	0	1	3
2	M012002	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
3	M012004	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2
4	M012040	0	0	0	0	0	1	1	0	0	0	0	0	0	0	2
5	M012041	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
6	M012042	0	0	0	0	1	0	0	0	1	0	0	0	0	0	2
7	M032570	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
8	M032643	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2
9	M012016	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2
10	M012017	0	0	0	0	0	1	0	0	0	1	0	0	1	0	3
11	M022251	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
12	M022185	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
13	M022191	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
14	M022194	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2
15	M022196	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
16	M022198	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
17	M022199	0	0	1	0	0	0	0	0	1	0	0	1	0	1	4
18	M022043	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
19	M022046	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
20	M022050	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
21	M022057	0	1	0	0	0	0	0	1	0	0	0	0	0	0	2
22	M022066	0	0	1	0	0	0	0	0	0	0	0	0	0	1	2
23	M022232	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
24	M022234B	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
25	M032142	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2
26	M032198	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
27	M032640	0	0	0	0	0	1	0	0	0	1	0	0	1	0	3
28	M032755	0	1	0	0	1	0	1	1	0	1	0	0	0	0	4
29	M032163	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	Sum	3	7	3	3	5	6	4	2	4	3	1	3	3	3	

Table 5. Q-Matrix of Booklet 2 for the Eighth Grade TIMSS 2003 Mathematics Test

	Item\Attribute	1	2	3	4	5	6	7	8	9	10	11	13	14	15	Sum
1	M012016	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2
2	M012017	0	0	0	0	0	1	0	0	0	0	1	0	1	0	3
3	M022251	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
4	M022185	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
5	M022191	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
6	M022194	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2
7	M022196	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
8	M022198	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
9	M022199	0	0	1	0	0	0	0	0	1	0	0	1	0	1	4
10	M012025	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
11	M012027	1	1	0	0	0	0	0	1	0	0	0	0	0	0	3
12	M012029	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
13	M022139	1	1	0	0	0	0	0	1	0	0	0	0	0	0	3
14	M022144	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
15	M022253	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2
16	M022156	0	1	0	0	0	1	0	0	0	0	0	0	0	0	2
17	M022104	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
18	M022106	1	0	0	0	0	1	1	0	0	0	0	0	0	0	3
19	M022110	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
20	M032307	0	0	0	0	0	0	0	0	1	1	0	0	0	1	3
21	M032523	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
22	M032701	0	1	0	0	0	1	0	0	0	0	0	0	0	0	2
23	M032704	1	1	0	0	0	0	1	0	0	0	0	0	0	0	3
24	M032525	0	0	0	0	0	1	0	0	1	0	0	0	0	0	2
25	M032381	0	0	1	0	0	0	1	0	0	0	0	0	0	0	2
26	M032416	0	0	1	0	0	0	0	0	0	0	0	0	0	1	2
27	M032160	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2
28	M032540	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
29	M032698	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2
30	M032529	0	1	0	0	0	0	0	1	0	0	0	0	0	0	2
	Sum	4	8	4	3	5	9	5	3	4	2	1	1	4	4	

If you feel inspired improve the test further, by incorporating elements from booklet 2 - don't hesitate to slack us your findings ;)