

# Knowledge Acquisition Evaluation Using Simulated Experts

Tri Minh Cao and Paul Compton  
School of Computer Science and Engineering  
University of New South Wales, Sydney  
Australia

# KA evaluation

- Reports on individual applications
- Sisyphus experiments
- Simulated experts
  - ML KBS, UC Irvine data set

# KA evaluation

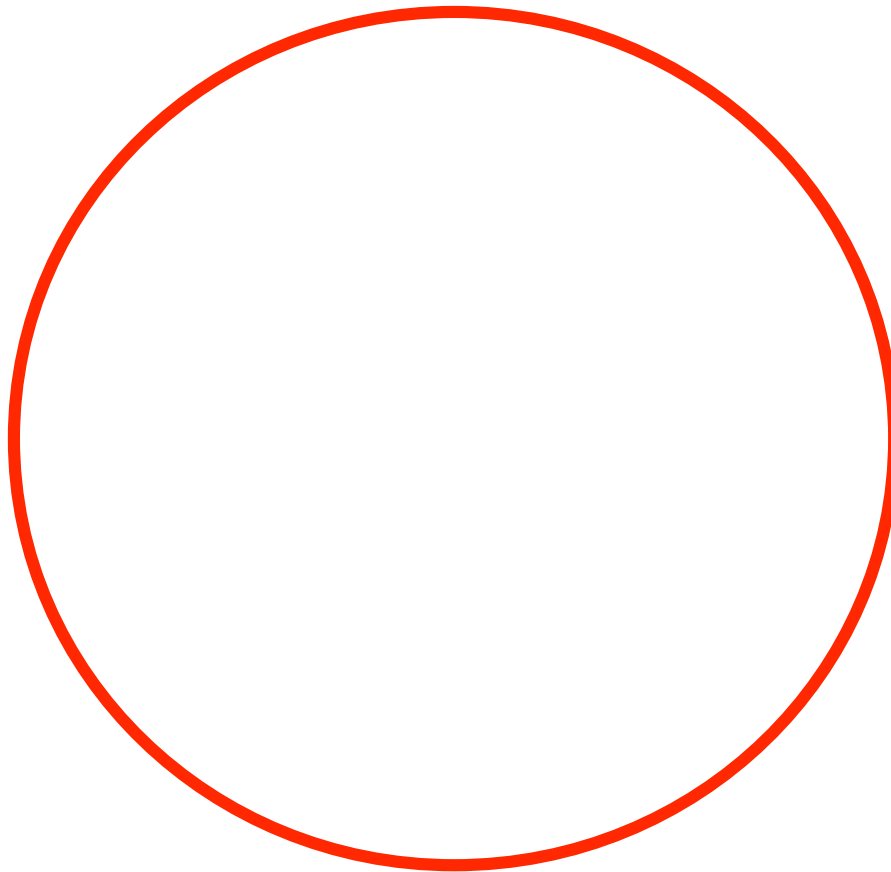
- Reports on individual applications
- Sisyphus experiments
- Simulated experts
  - ML KBS, UC Irvine data set
- Complexity analysis for KA ??
  - does the KA required vary with different knowledge structures?

# Assumptions

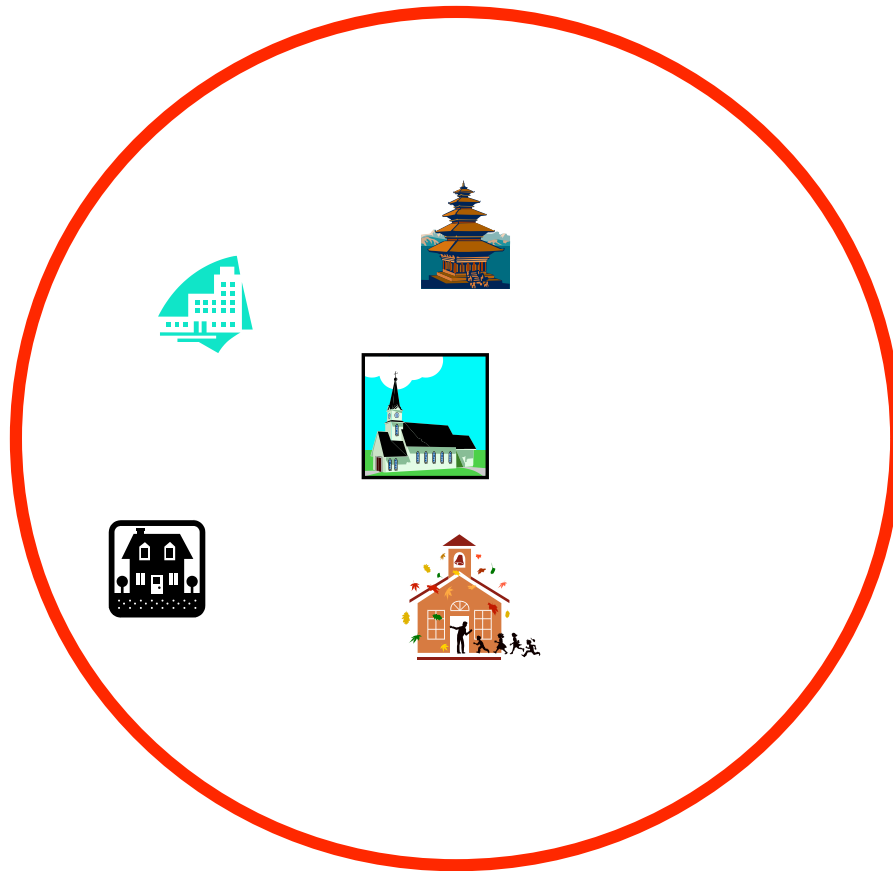
- Expertise can be characterised as
  - overgeneralised
  - overspecialised
  - or both

# Expertise

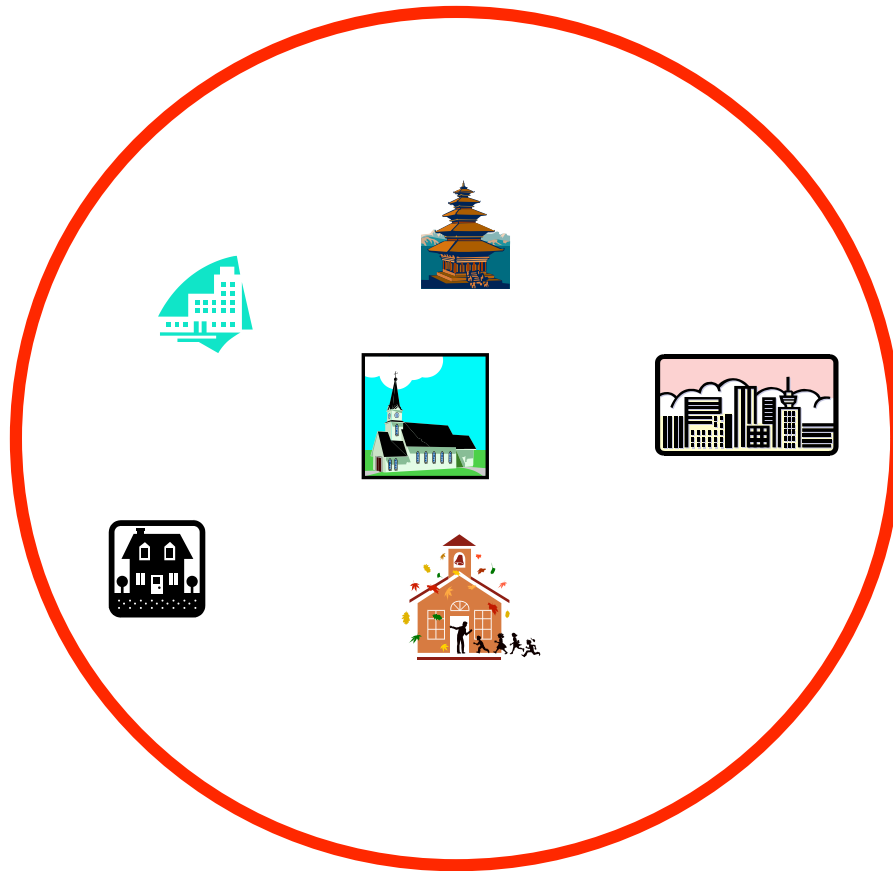
# Expertise



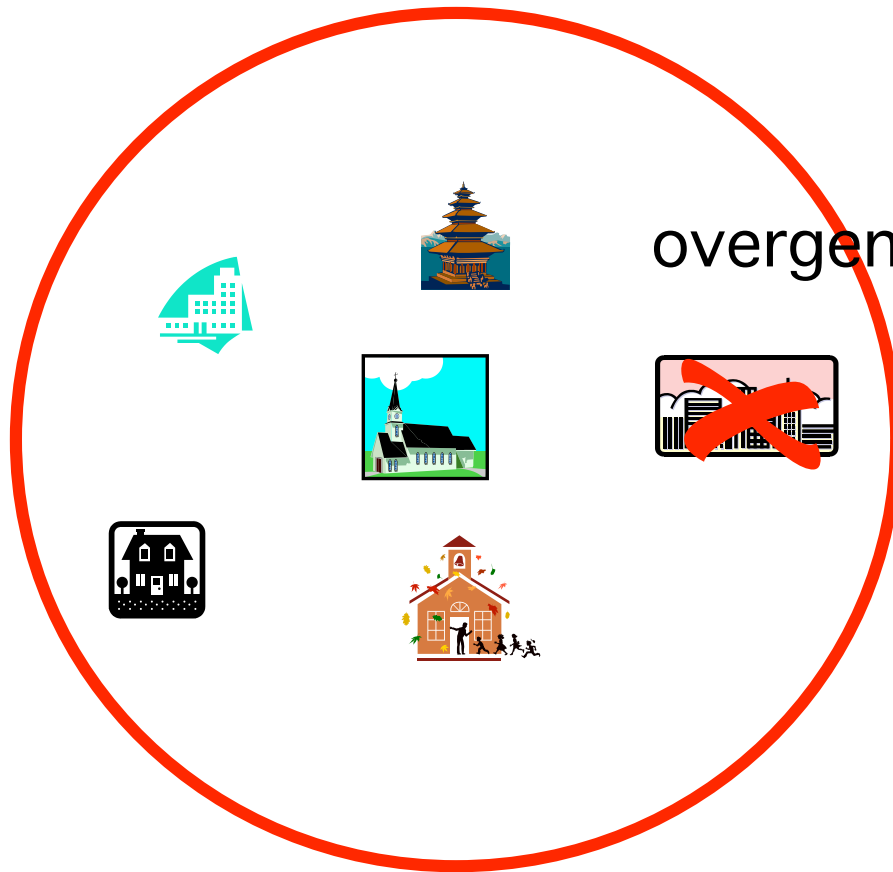
# Expertise



# Expertise

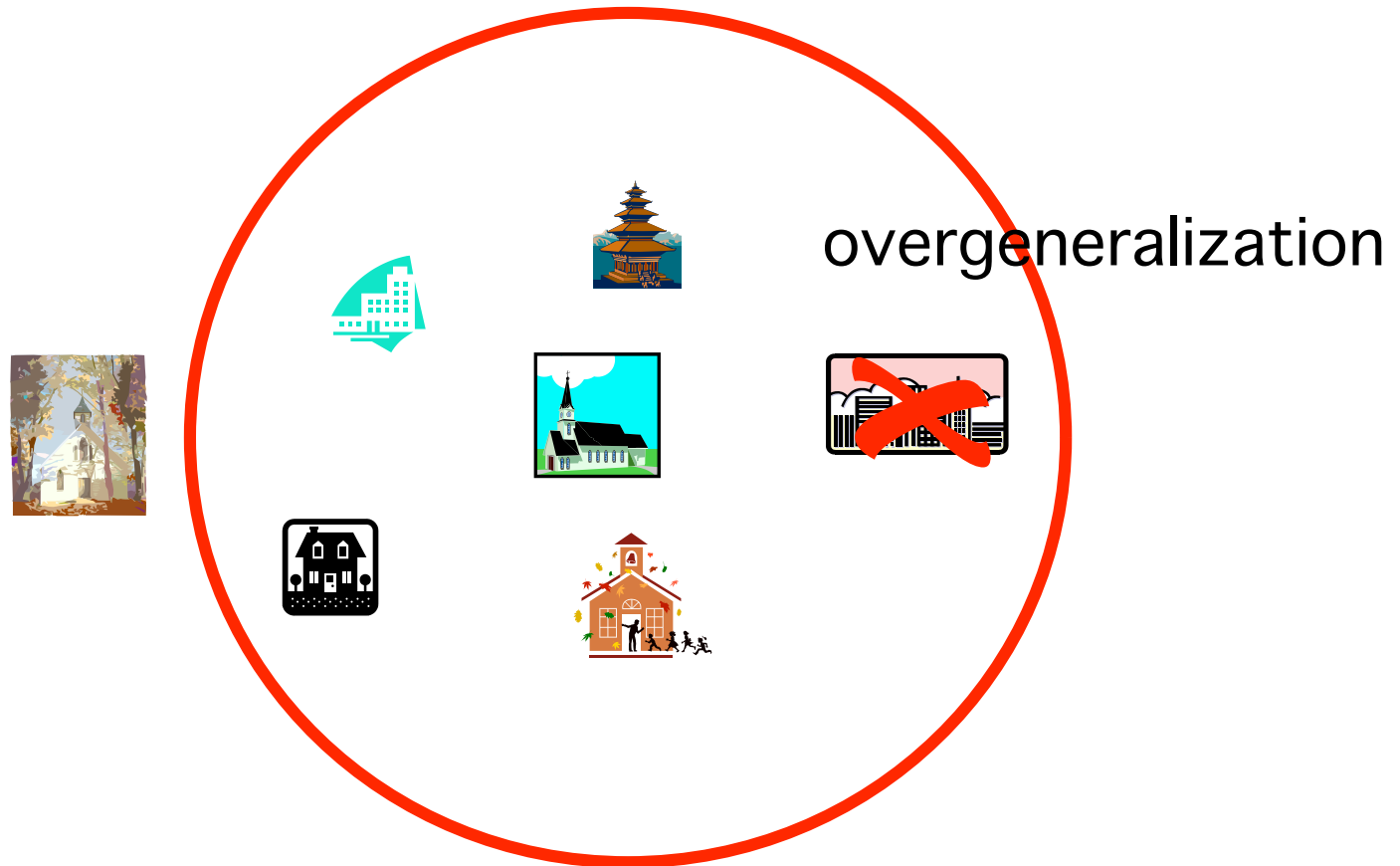


# Expertise



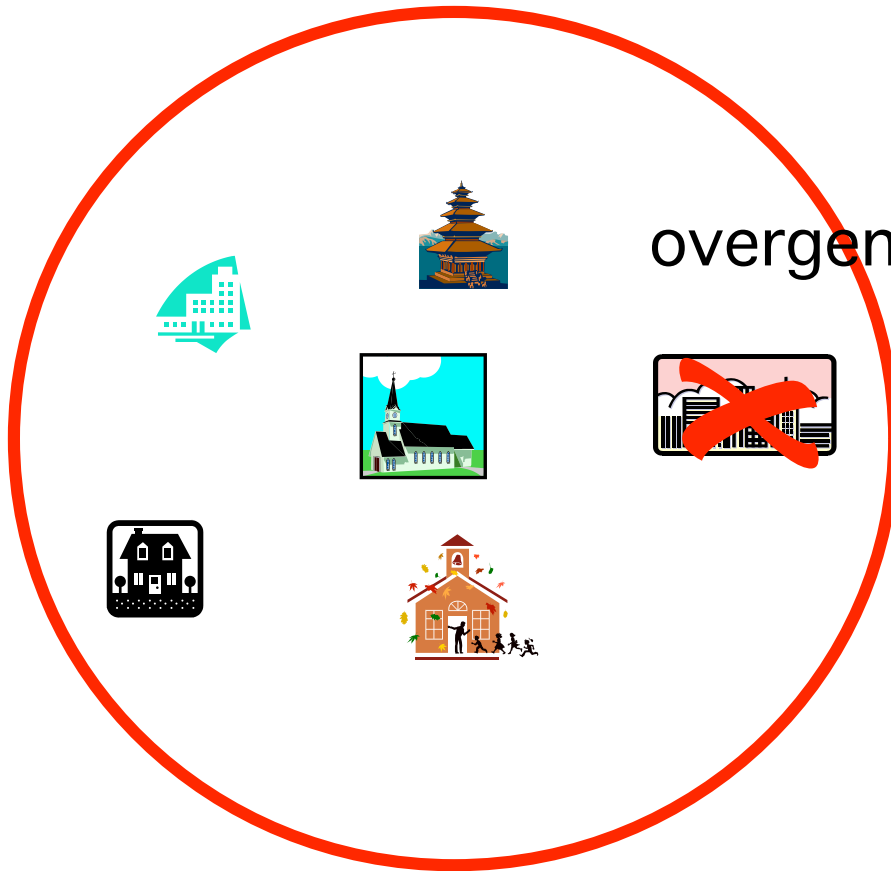
overgeneralization

# Expertise



# Expertise

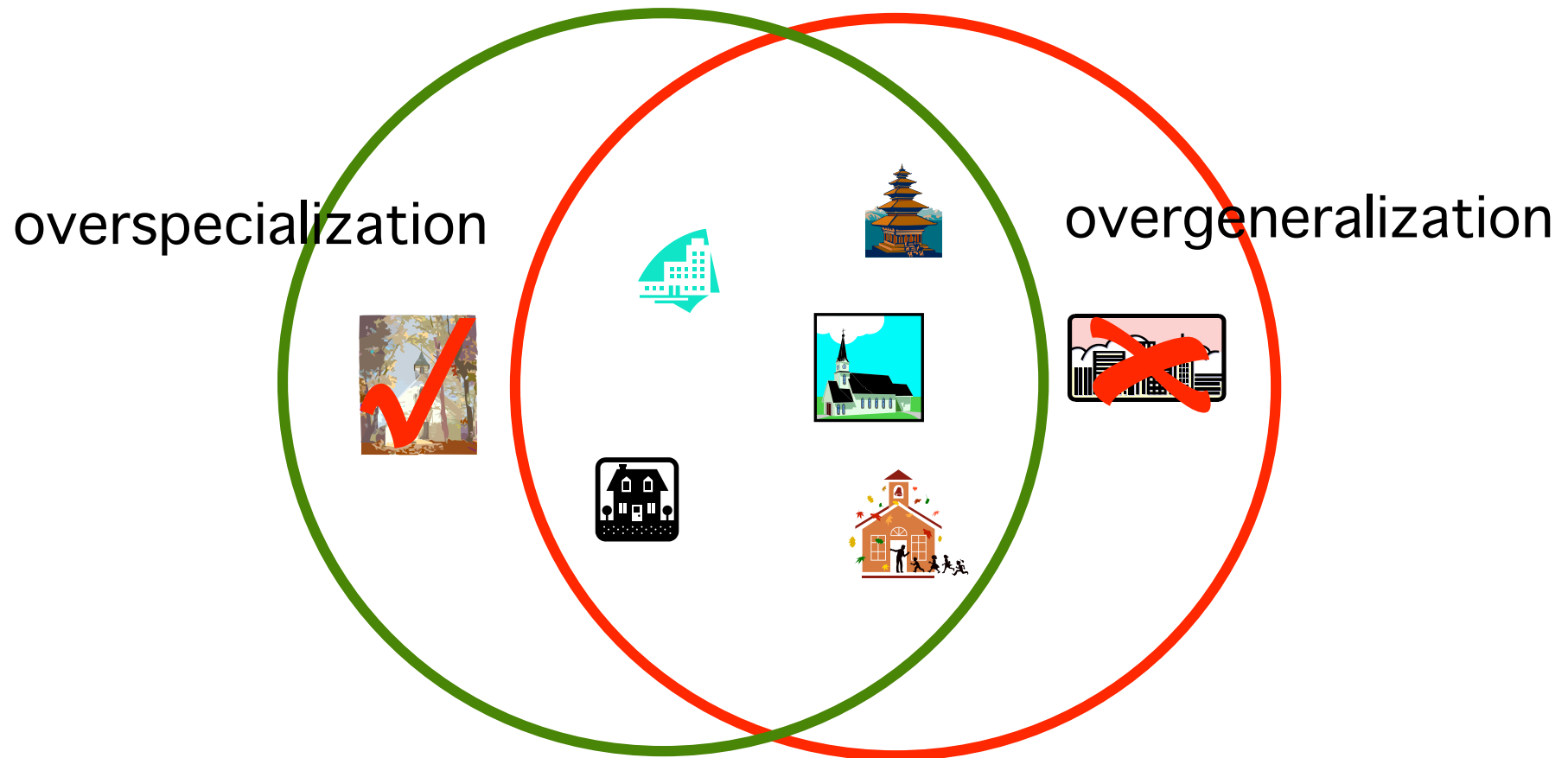
overspecialization



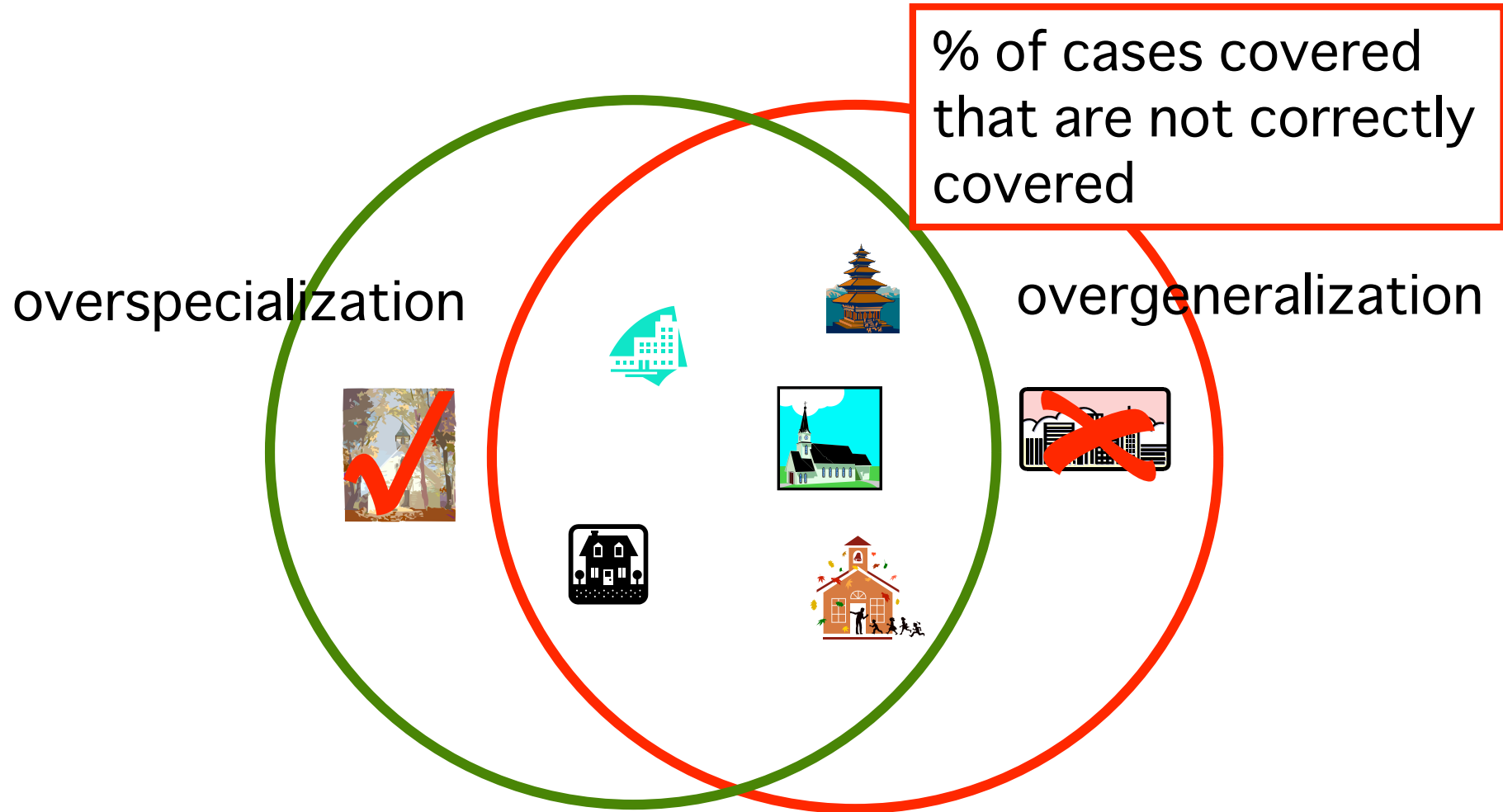
overgeneralization



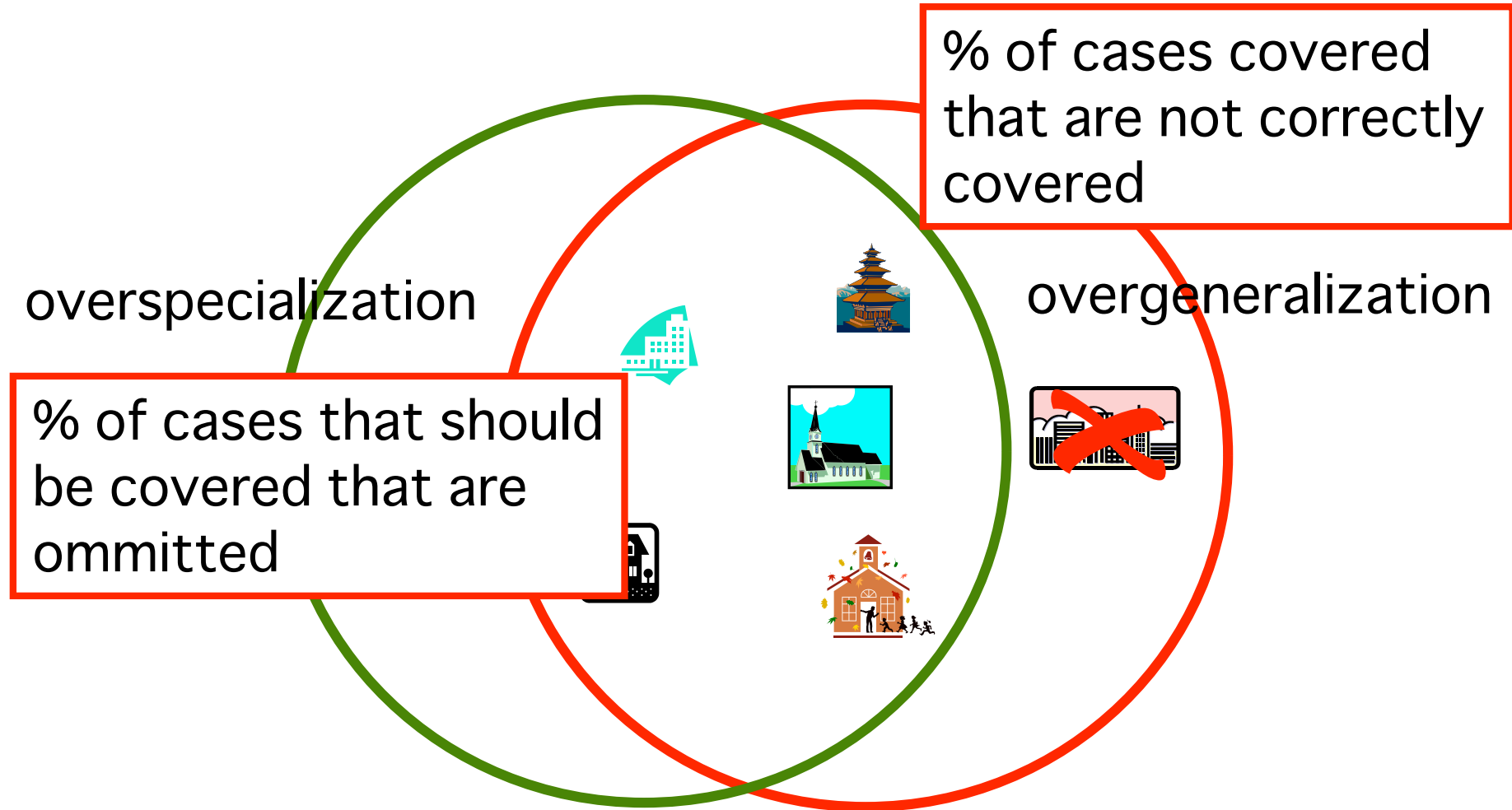
# Expertise



# Expertise



# Expertise



# Assumptions

# Assumptions

- Expertise can be overgeneralised, overspecialised or both

# Assumptions

- Expertise can be overgeneralised, overspecialised or both
- Knowledge is acquired in chunks

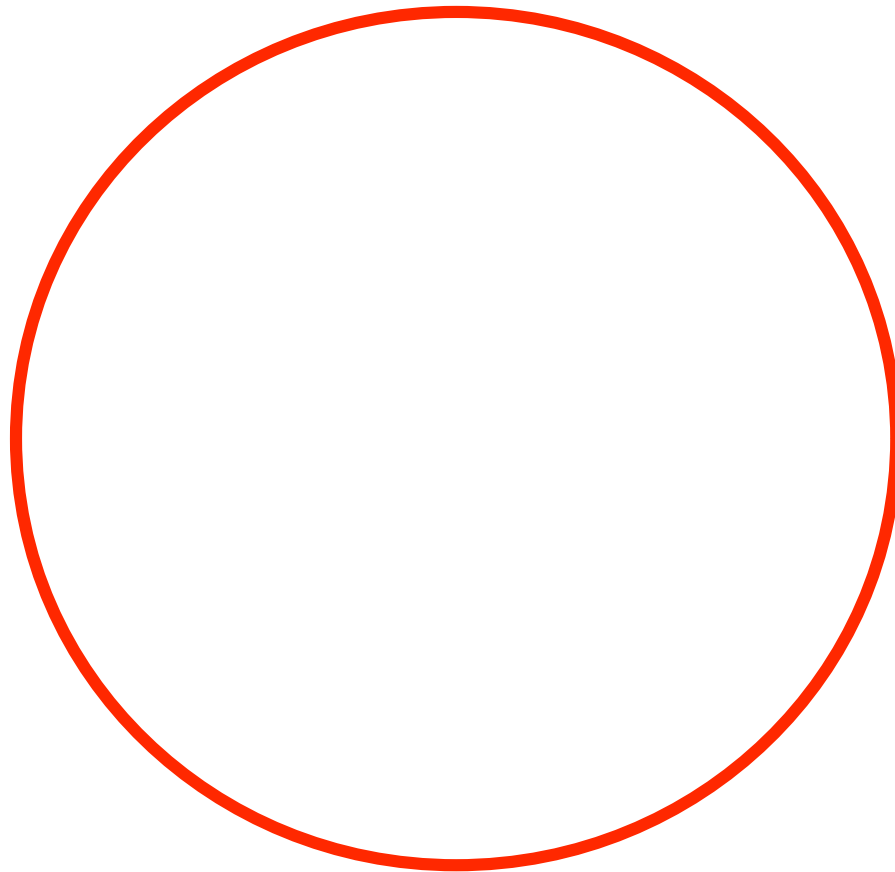
# Assumptions

- Expertise can be overgeneralised, overspecialised or both
- Knowledge is acquired in chunks
- KA performance is characterised by the number of interactions needed

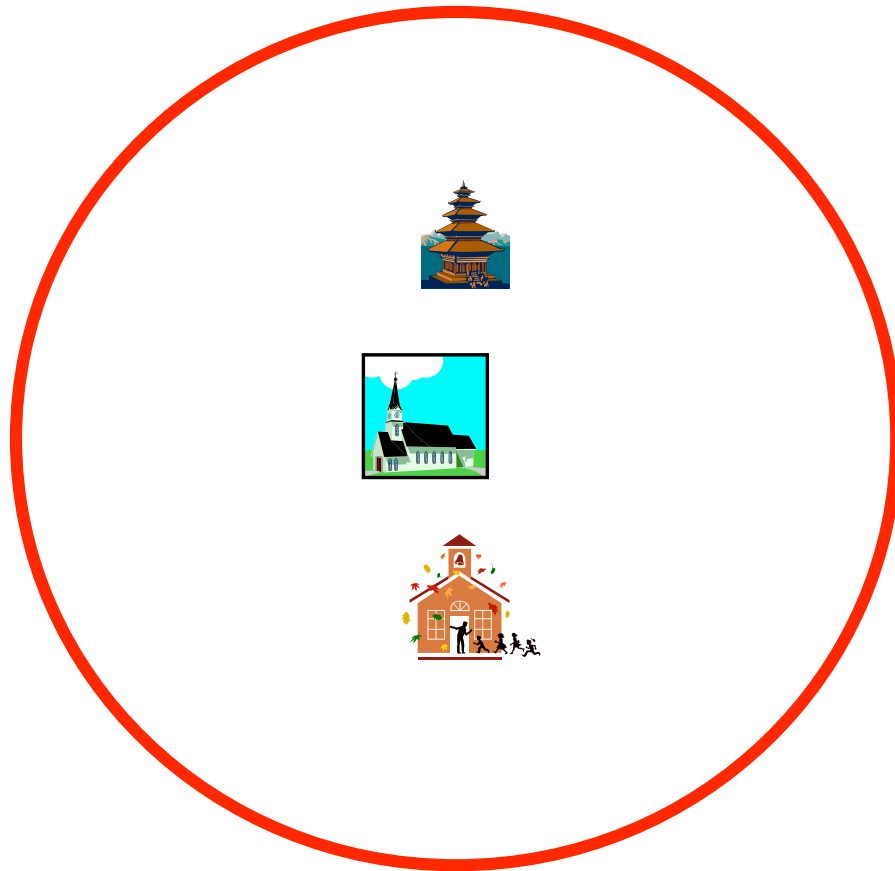
# Assumptions

- Expertise can be overgeneralised, overspecialised or both
- Knowledge is acquired in chunks
- KA performance is characterised by the number of interactions needed
- *Stream of random cases; chunks added when needed*

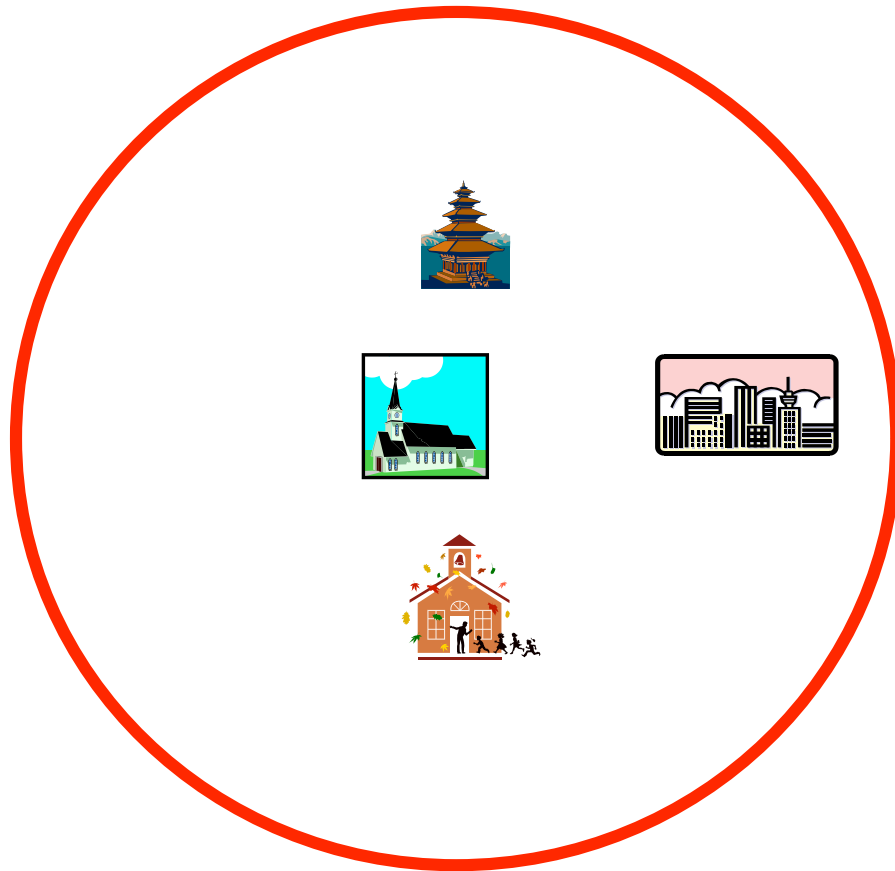
# Adding chunks



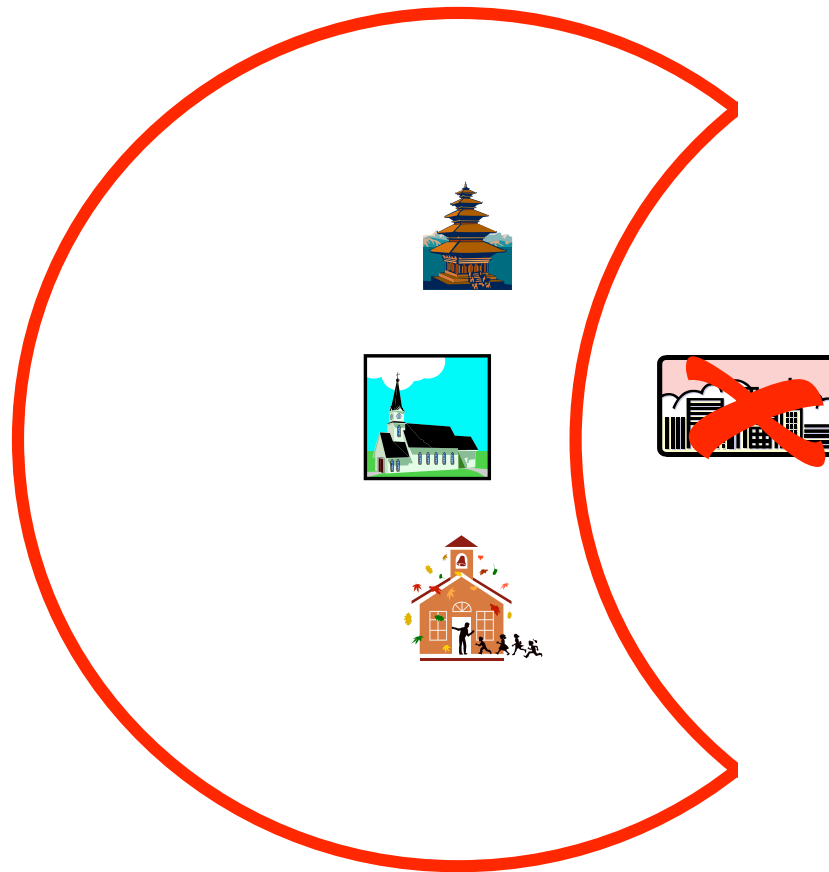
# Adding chunks



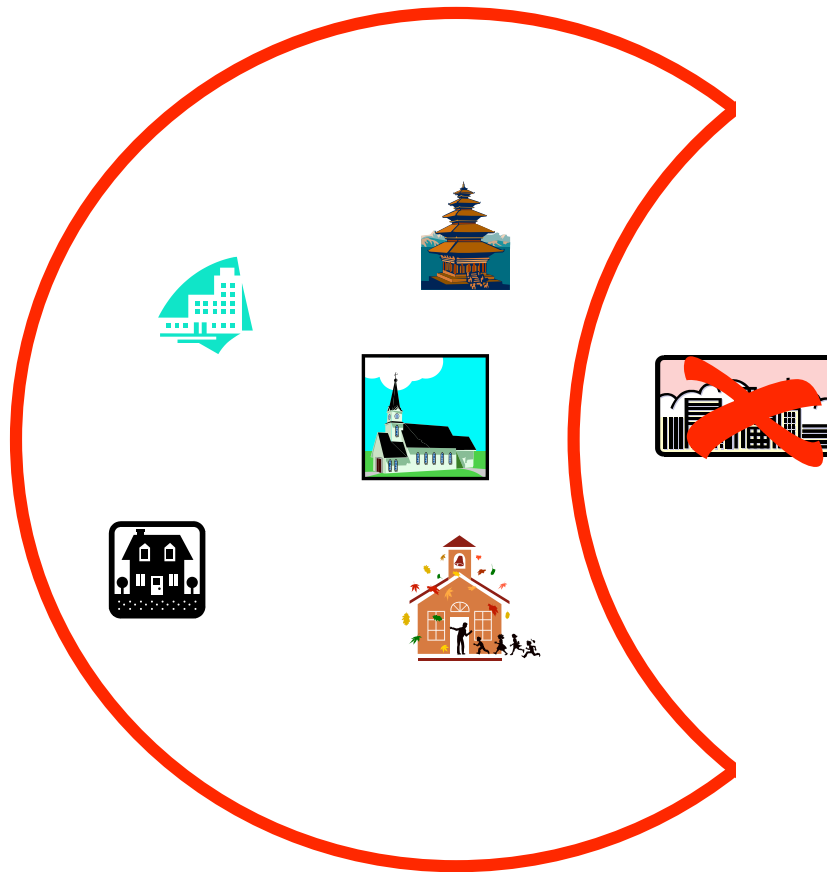
# Adding chunks



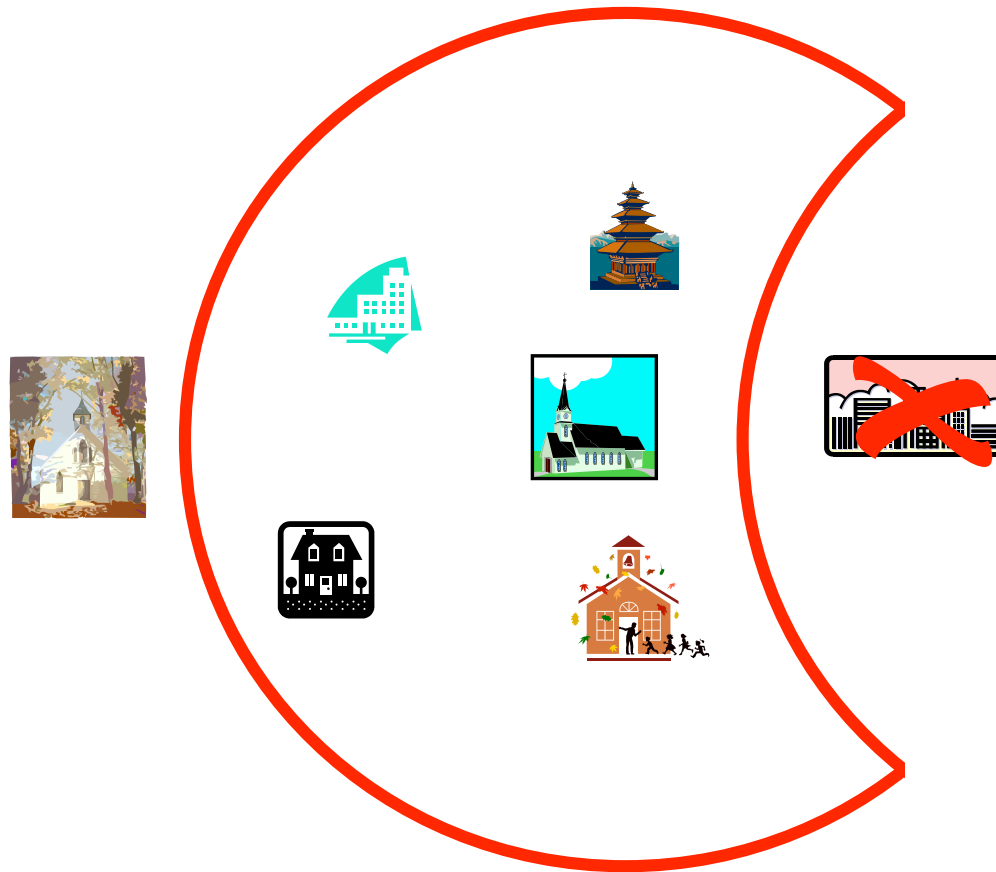
# Adding chunks



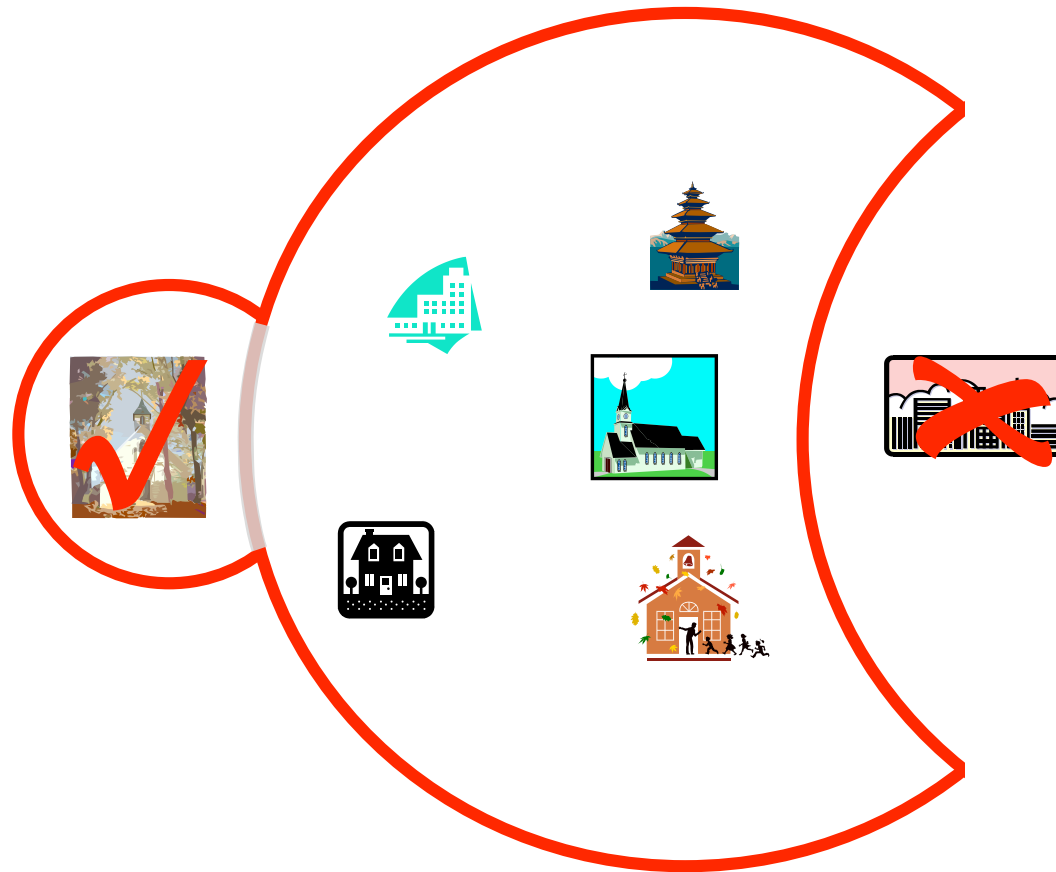
# Adding chunks



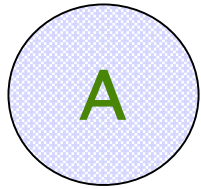
# Adding chunks



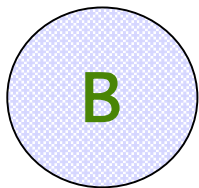
# Adding chunks



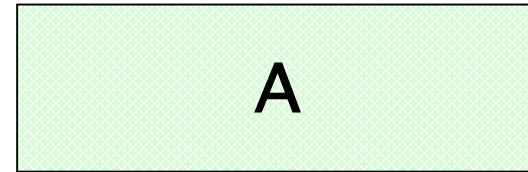
# Simulation 1



case

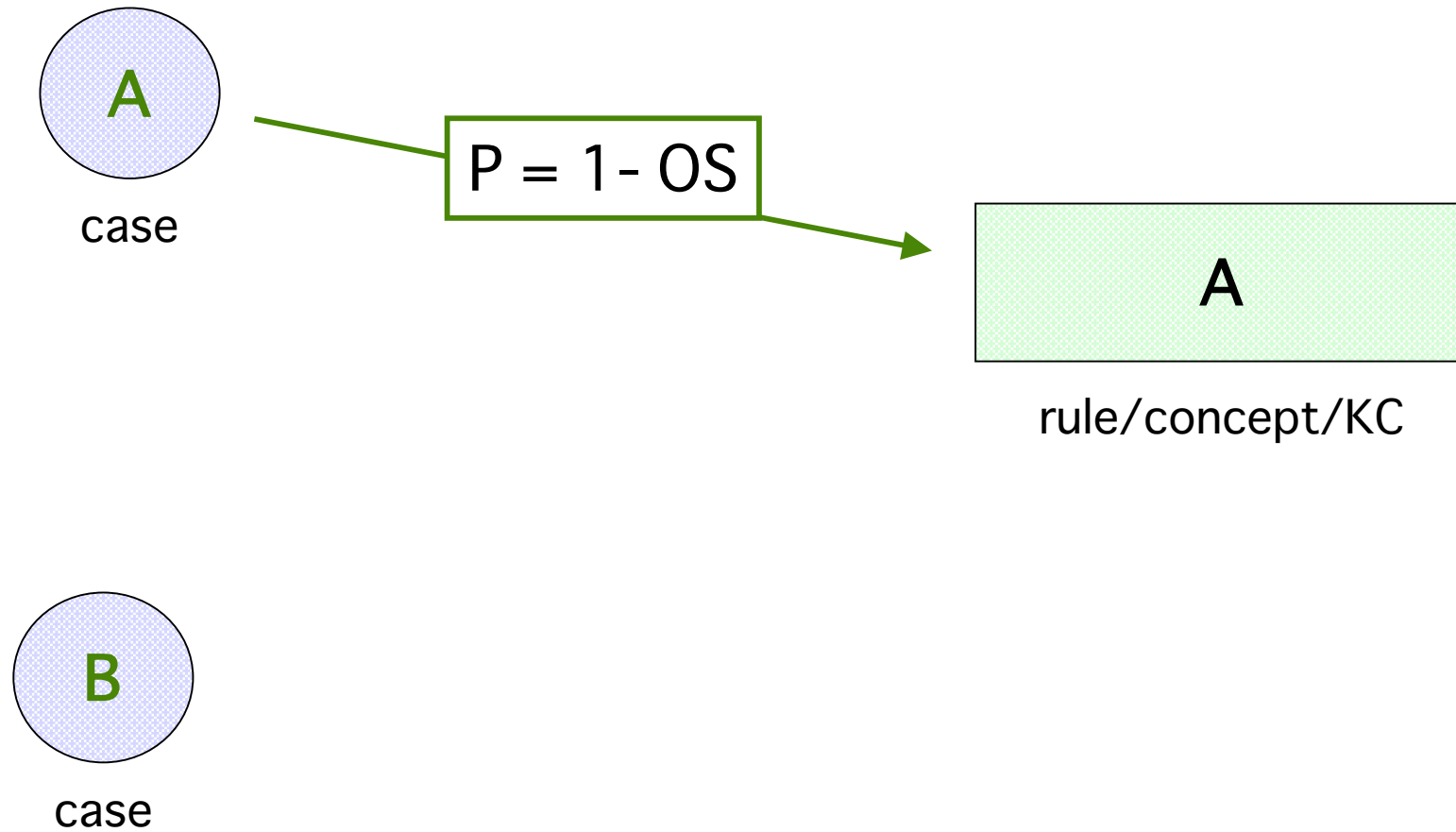


case

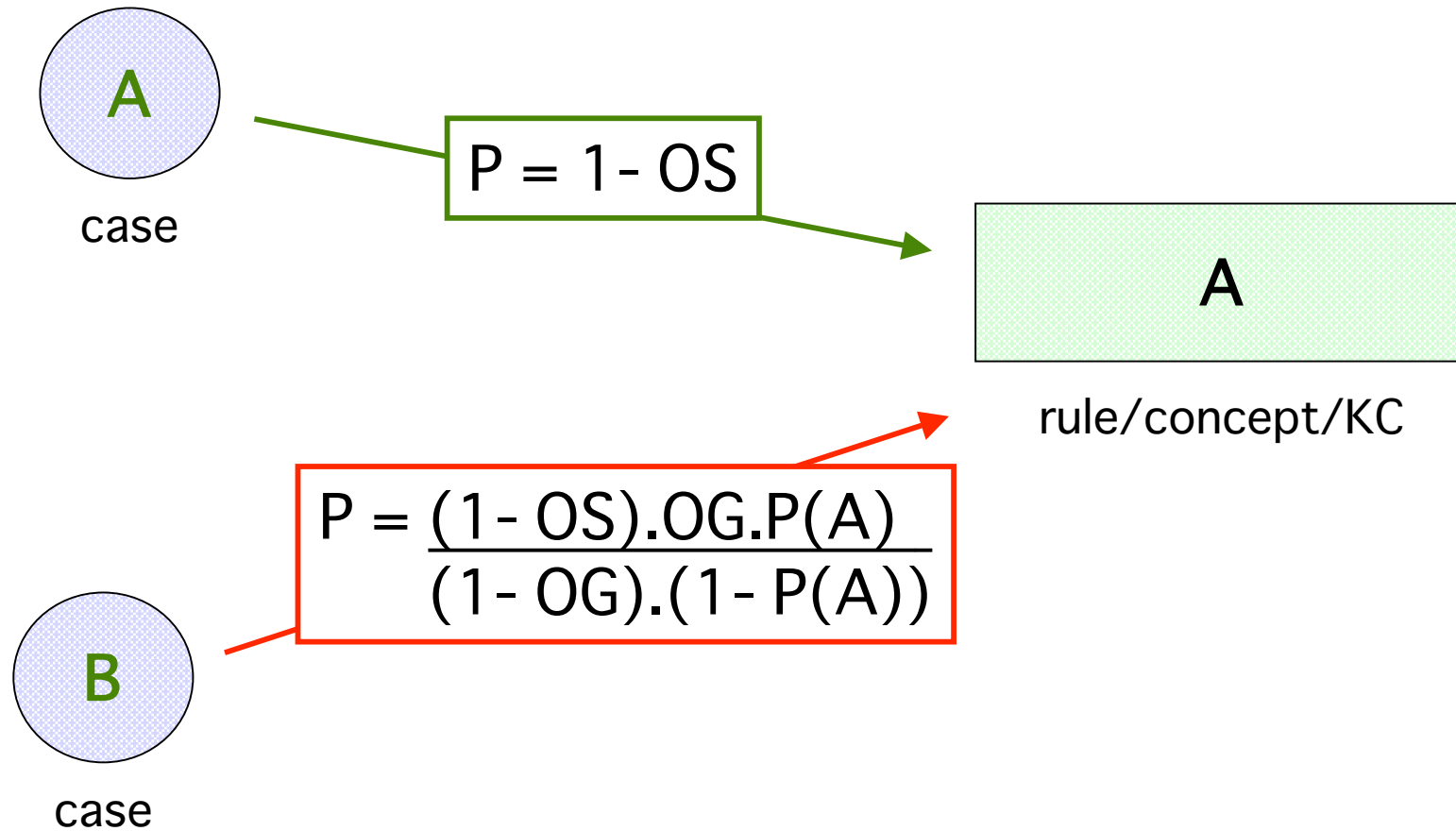


rule/concept/KC

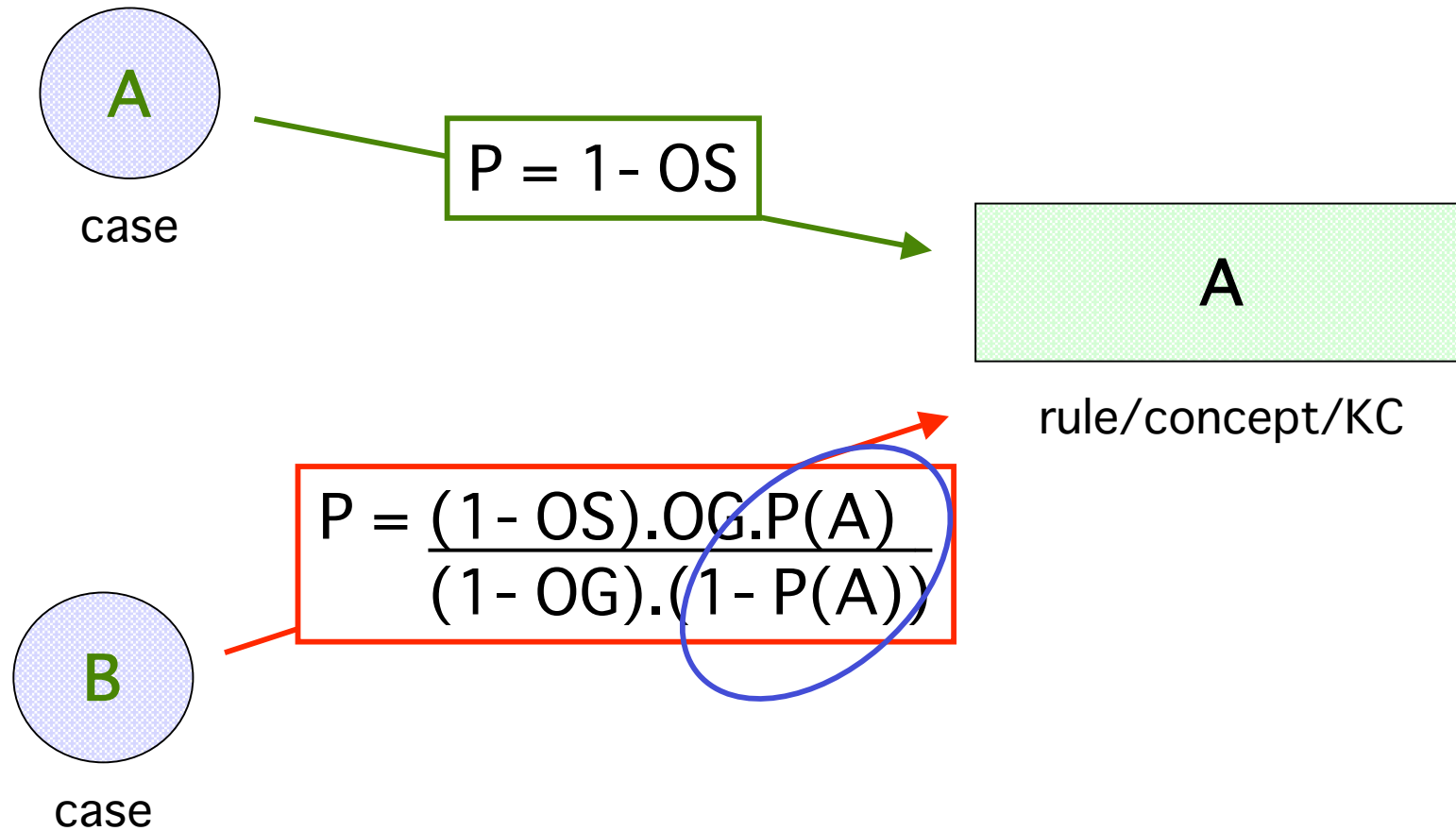
# Simulation 1



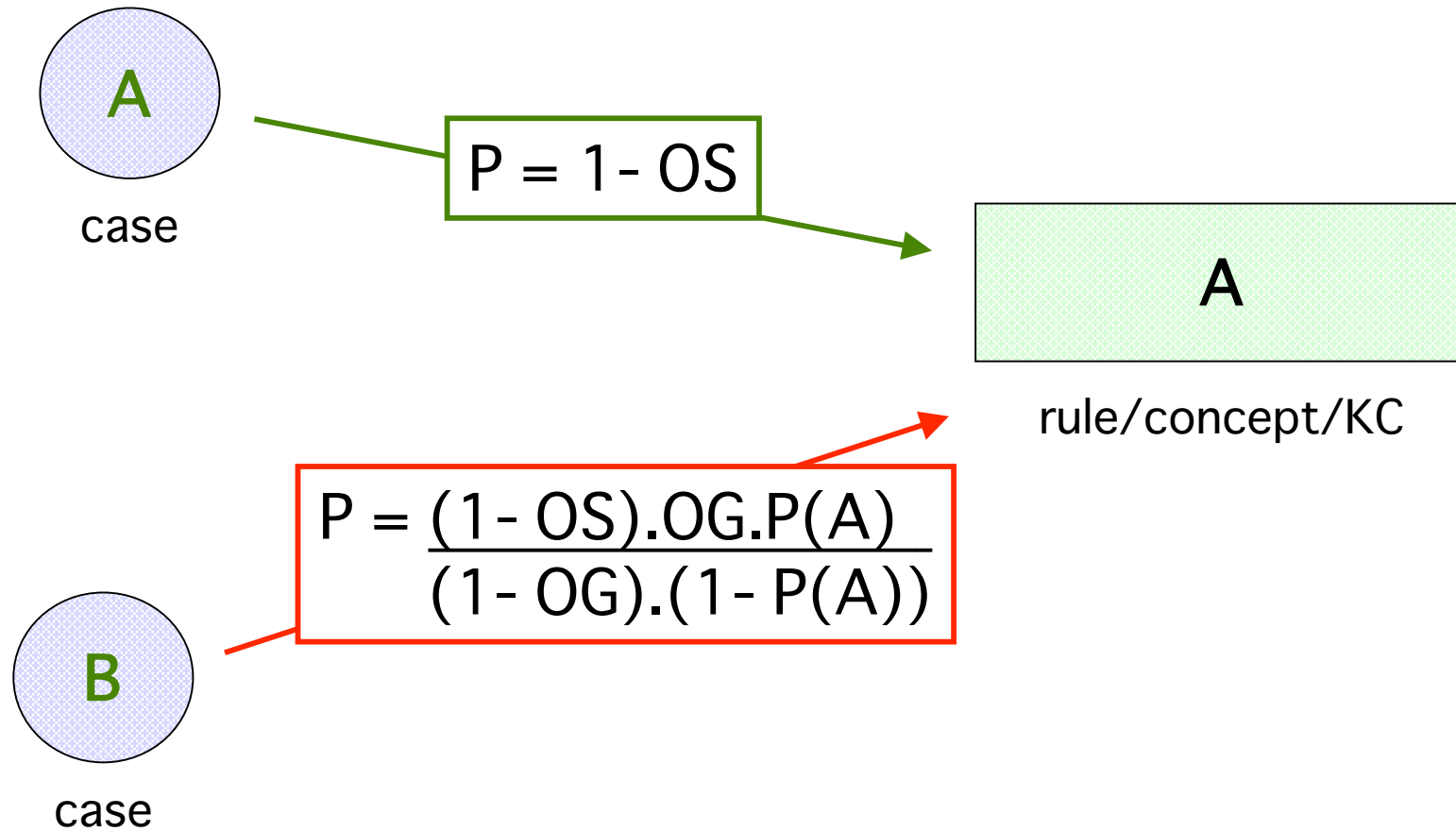
# Simulation 1



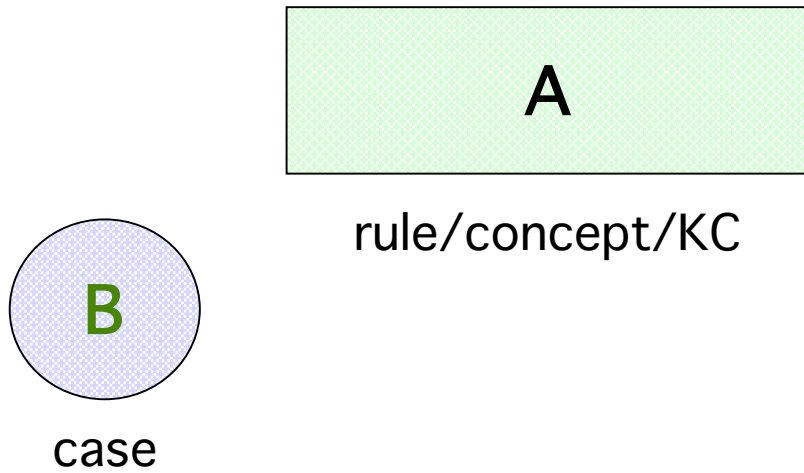
# Simulation 1



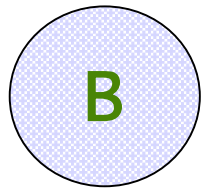
# Simulation 1



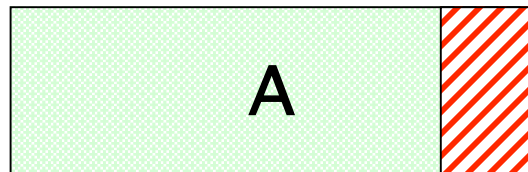
# Simulation 2



# Simulation 2



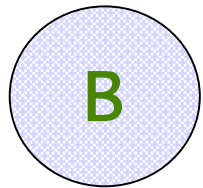
case



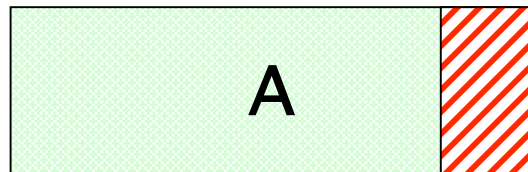
rule/concept/KC

$$P = \frac{(1 - OS) \cdot OG \cdot P(A) \cdot OS}{(1 - OG) \cdot (1 - P(A))}$$

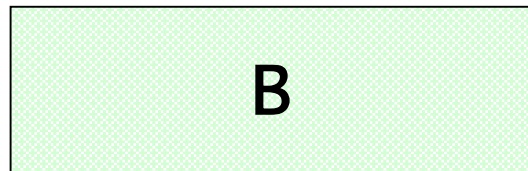
# Simulation 2



case



rule/concept/KC

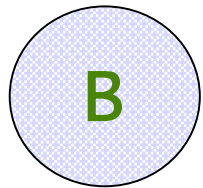


rule/concept/KC

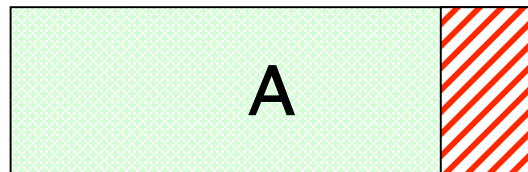
$$P = \frac{(1 - OS).OG.P(A).OS}{(1 - OG).(1 - P(A))}$$

$$P = \frac{(1 - OS).OG.P(A).(1 - OS)}{(1 - OG).(1 - P(A))}$$

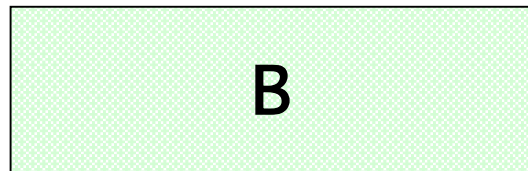
# Simulation 2



case



rule/concept/KC



rule/concept/KC

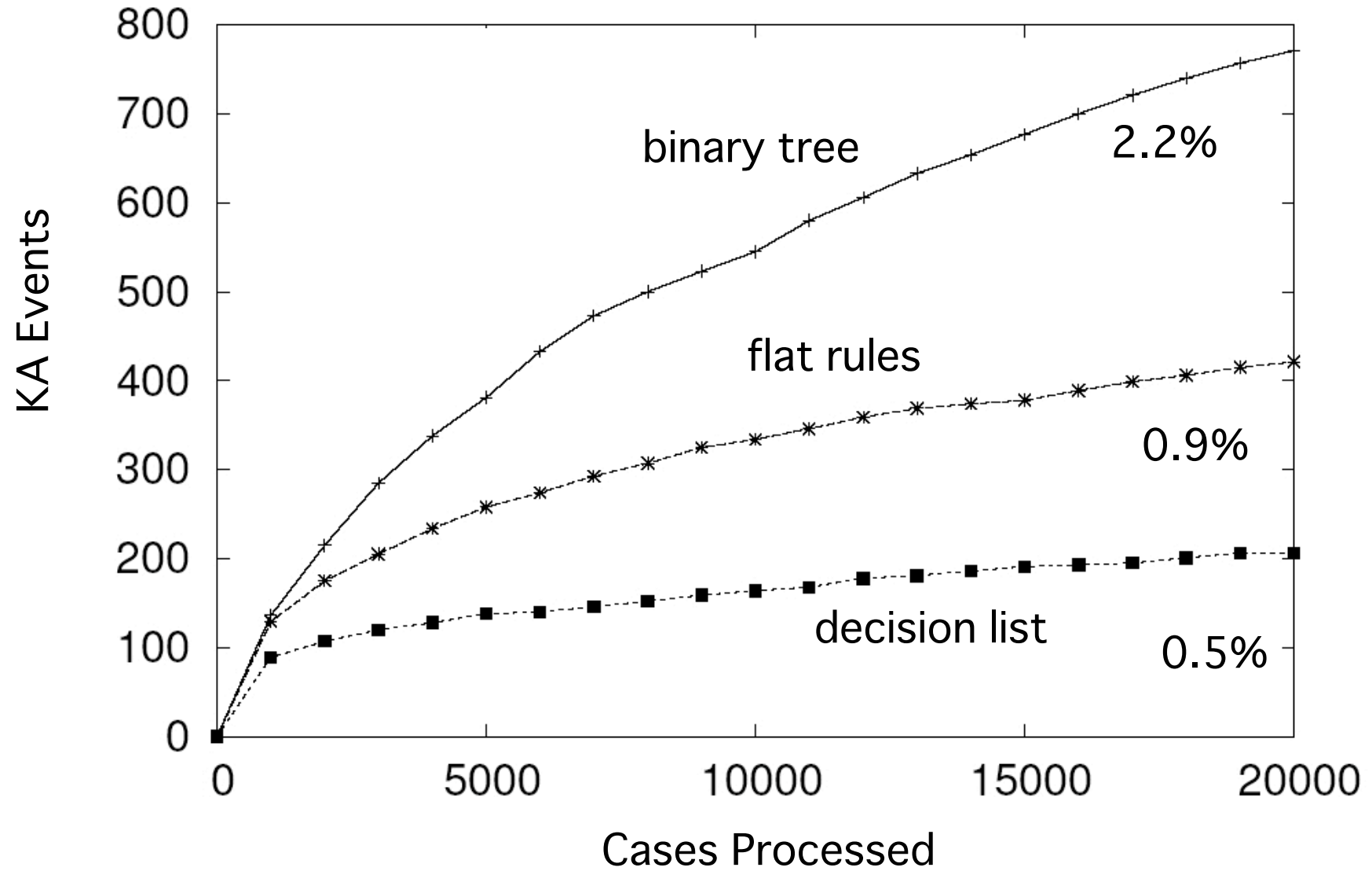
$$P = \frac{(1 - OS).OG.P(A).OS}{(1 - OG).(1 - P(A))}$$

$$P = \frac{(1 - OS).OG.P(A).(1 - OS)}{(1 - OG).(1 - P(A))}$$

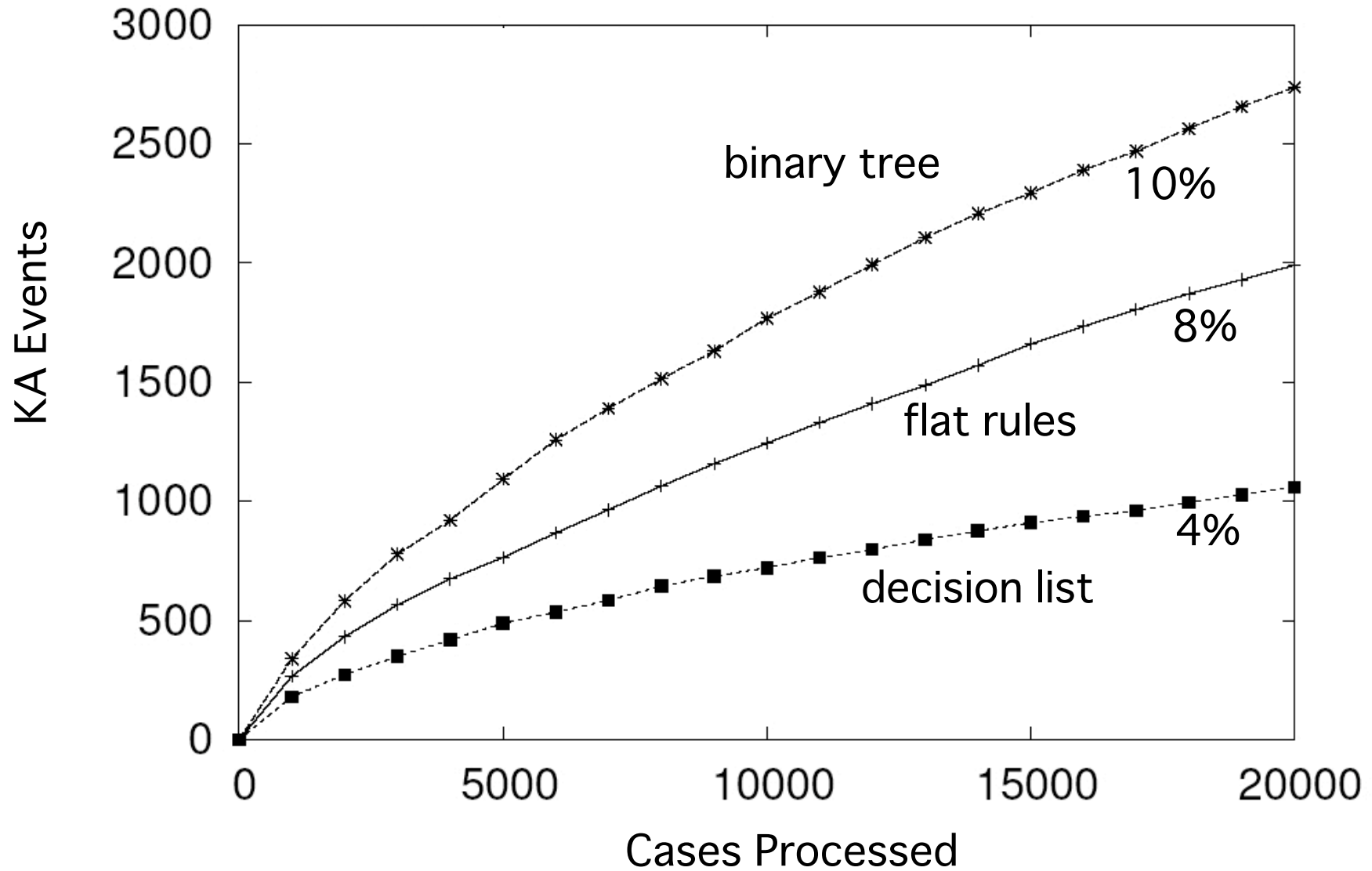
or

$$P = 1 - OS$$

10% overgeneralisation, 10% overspecialisation



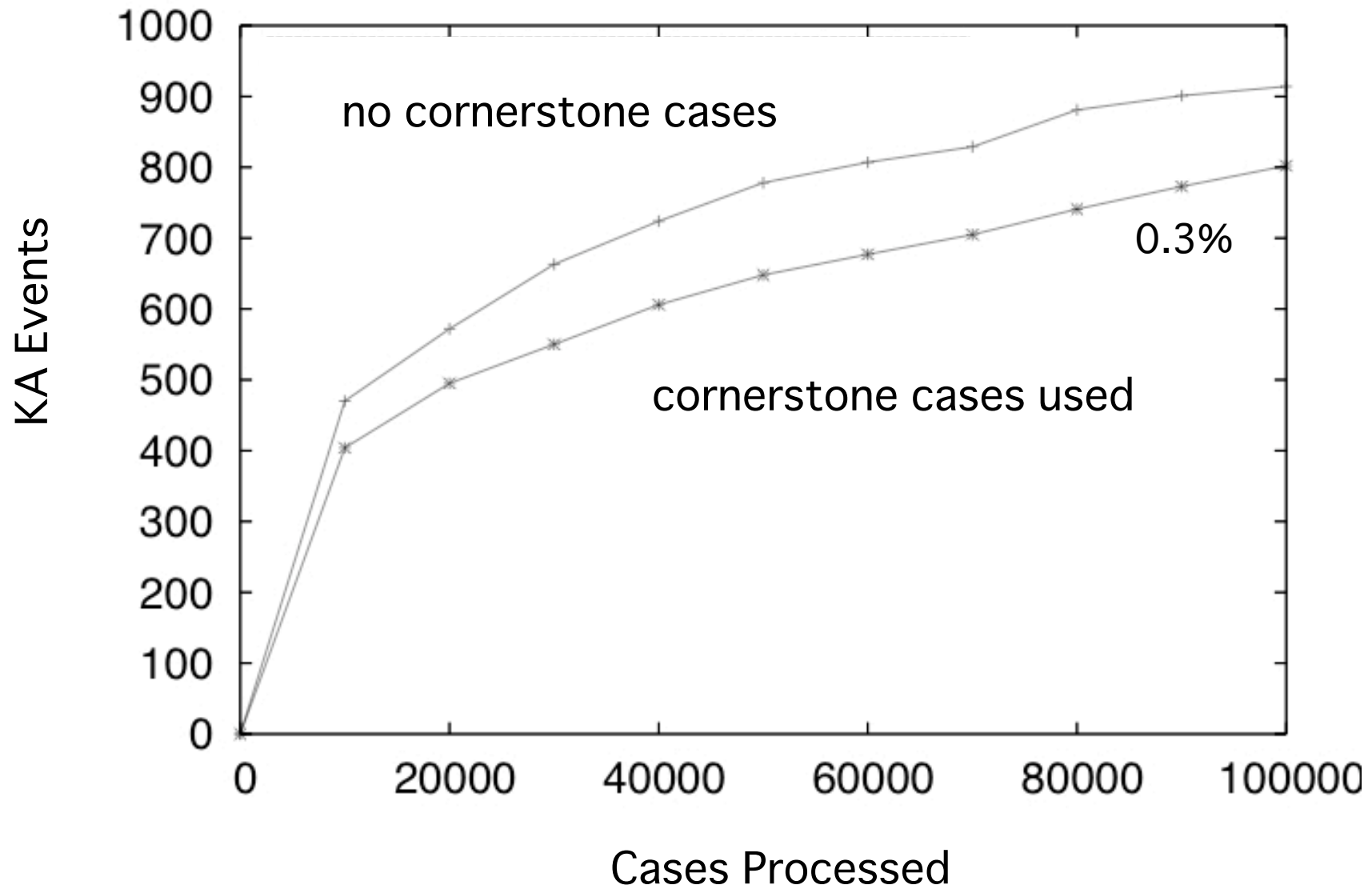
30% overgeneralisation, 10% overspecialisation



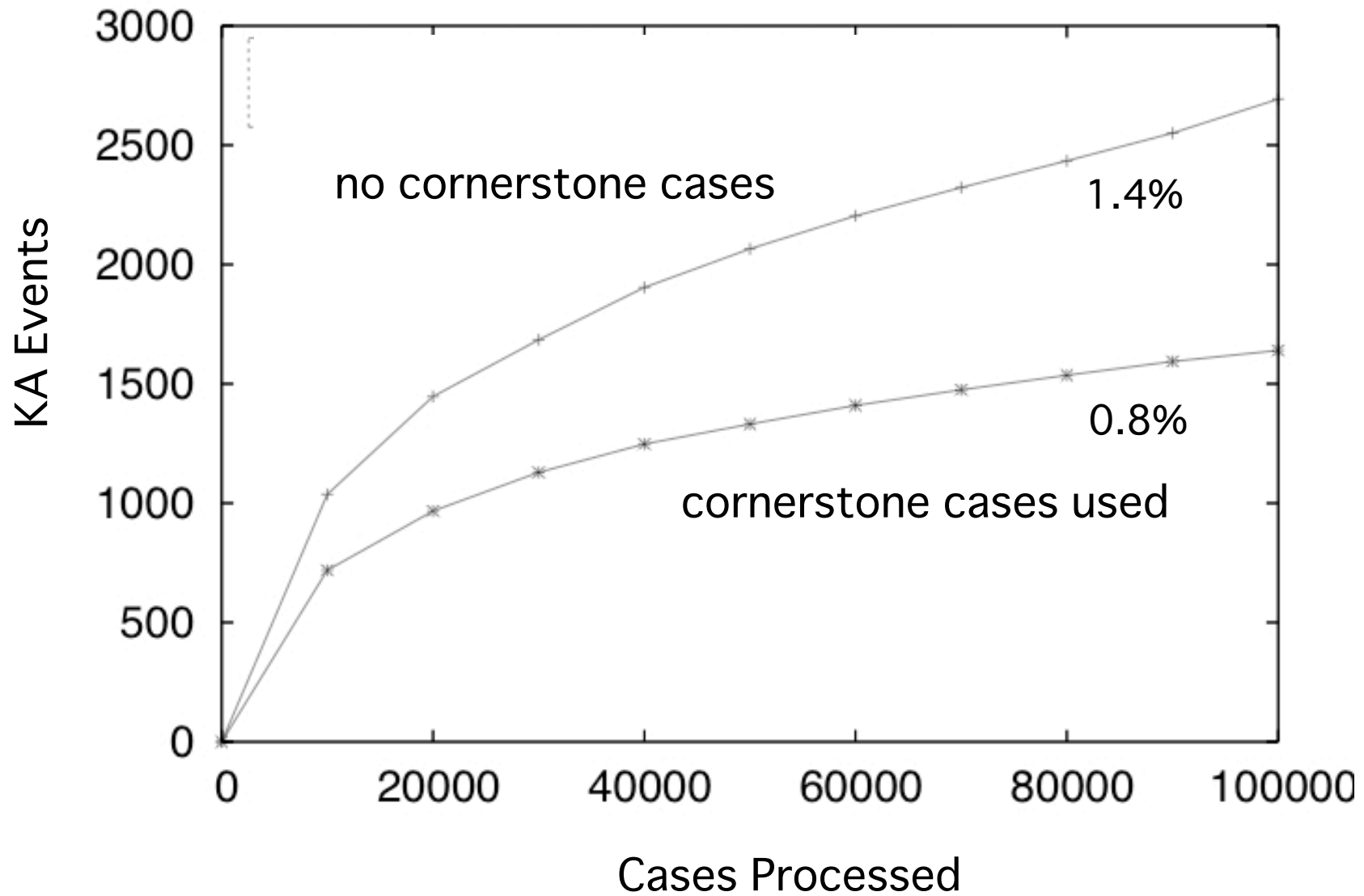
# Cornerstone cases

- A case for which a rule/concept/KC is added is stored
- New KC checked against these cases
- Expert improves the KB

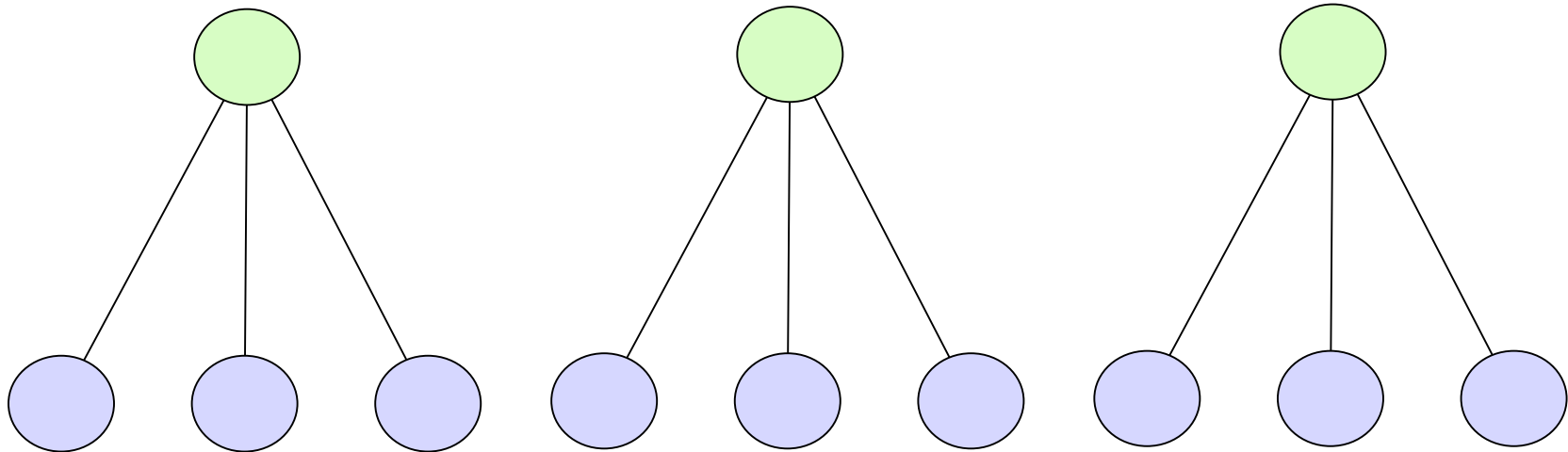
20% overgeneralisation, 20% overspecialisation



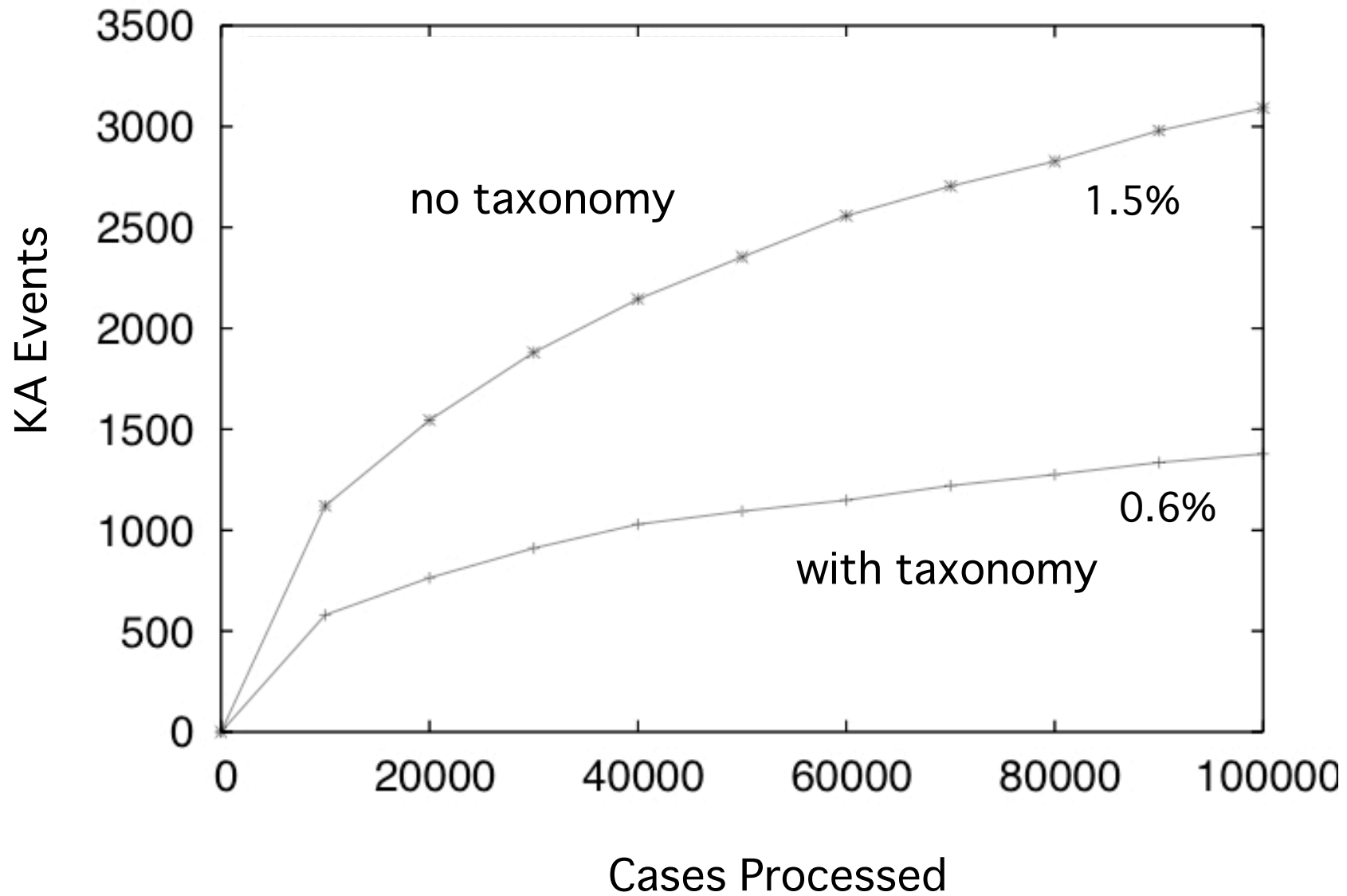
30% overgeneralisation, 30% overspecialisation



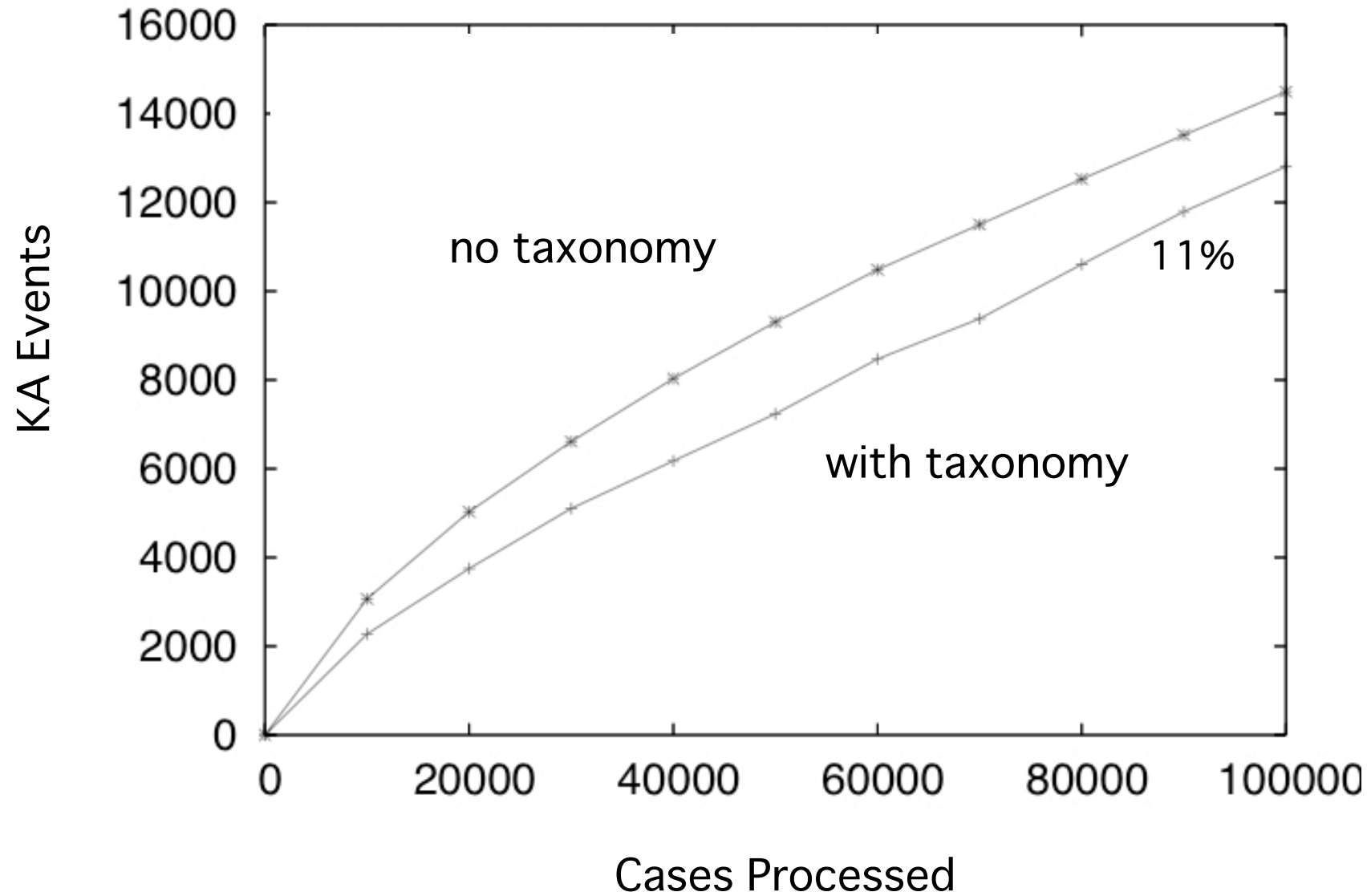
# Taxonomy / intermediate conclusion



30% overgeneralisation, 30% overspecialisation



40% overgeneralisation, 40% overspecialisation



# Conclusions

- Cornerstone cases are useful
- Taxonomies / ontologies are useful if the domain is well known (????)
- Is it possible to develop a complexity analysis for knowledge acquisition ?
- Simulation - further work
  - Order of instances
  - Class / concept distributions
  - Depth of correction
  - Ontology structures
  - Probabilities ~ human expertise?