Transfer of Knowledge in Cognitive Systems

Pat Langley

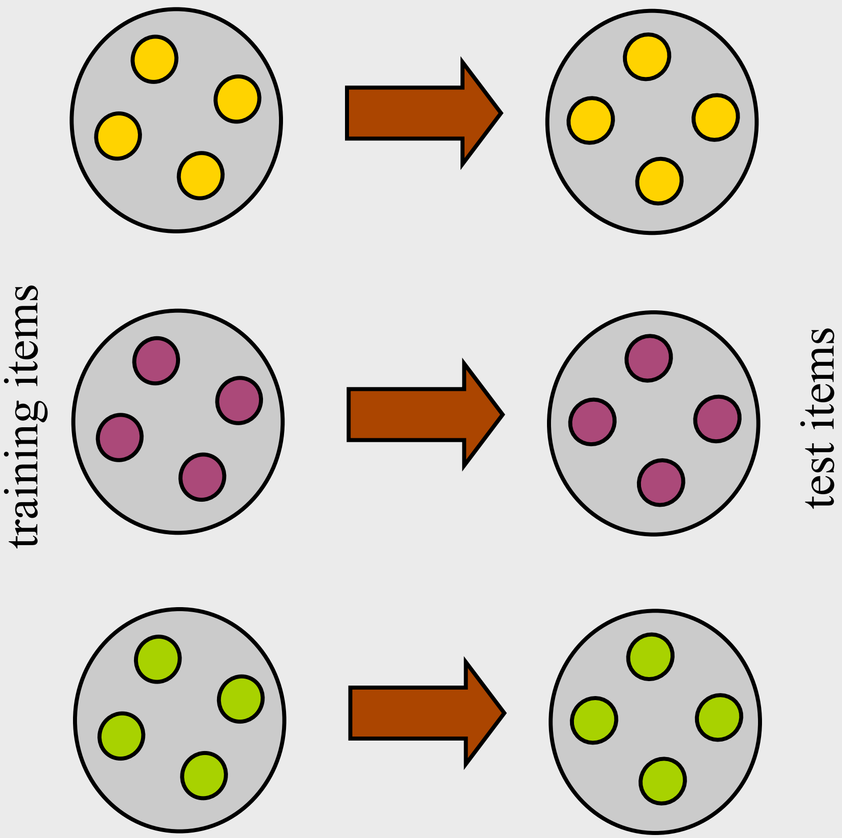
Computational Learning Laboratory
Center for the Study of Language and Information
Stanford University, Stanford, California USA

http://cll.stanford.edu/

Thanks to D. Aha, D. Choi, K. Forbus, J. Laird, S. Rogers, and T. Senator for useful dicsussions. This talk reports research funded by a grant from DARPA and AFRL, which are not responsible for its contents.

Generality in Learning

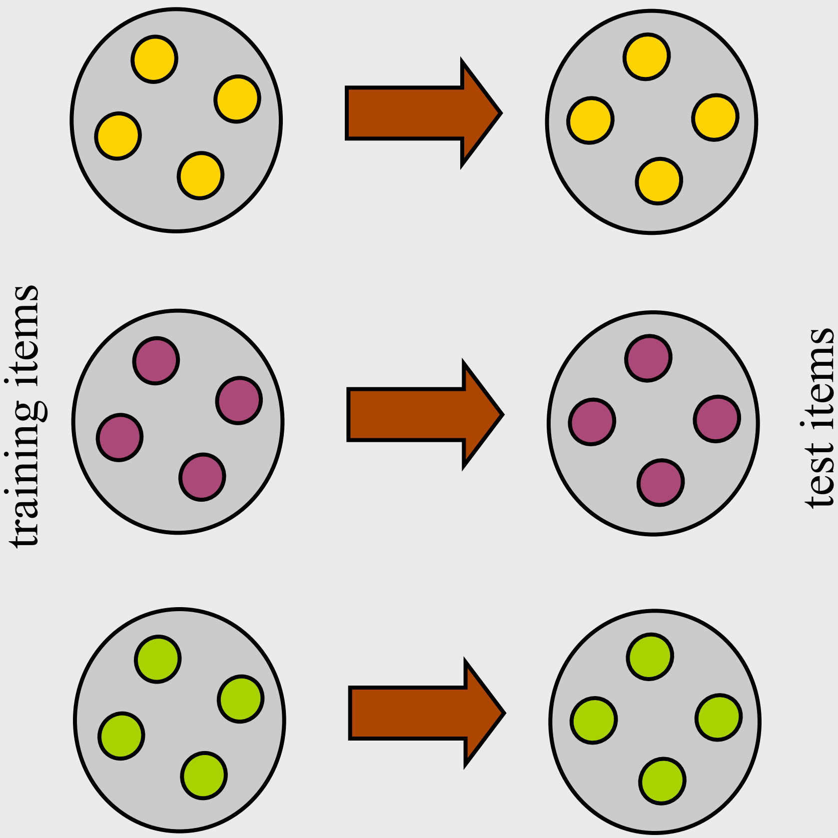
general learning in
multiple domains



Humans exhibit general intelligence by
their ability to learn in many domains.

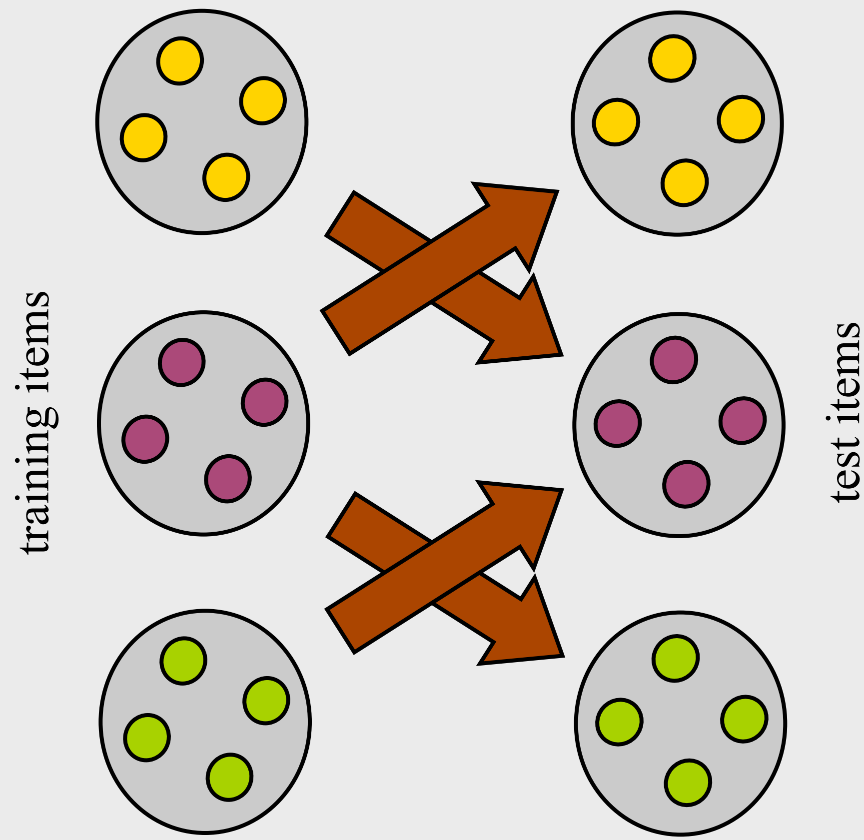
Generality and Transfer in Learning

general learning in
multiple domains



Humans exhibit general intelligence by
their ability to learn in many domains.

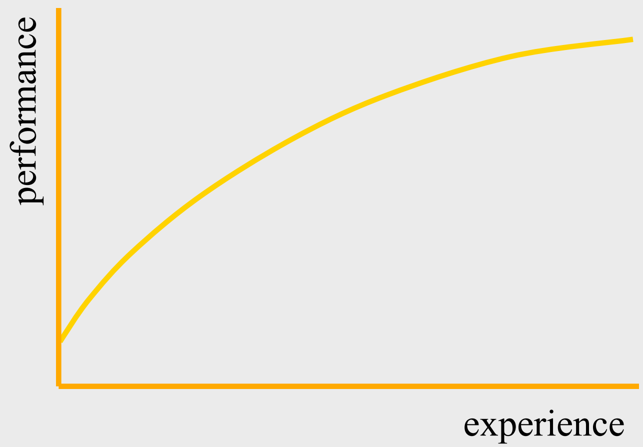
transfer of learning
across domains



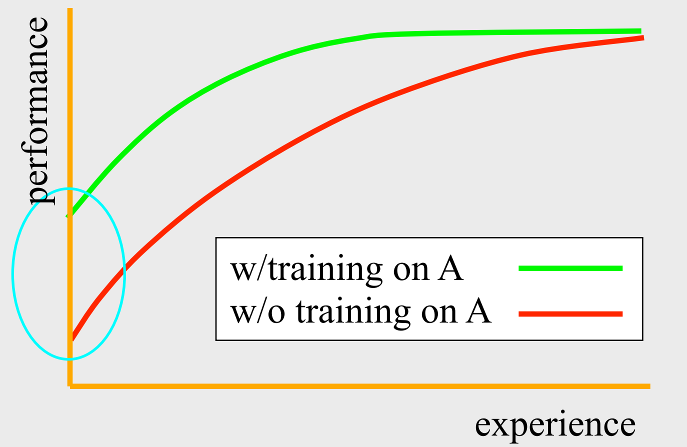
Humans are also able to utilize knowledge
learned in one domain in other domains.

What is Transfer?

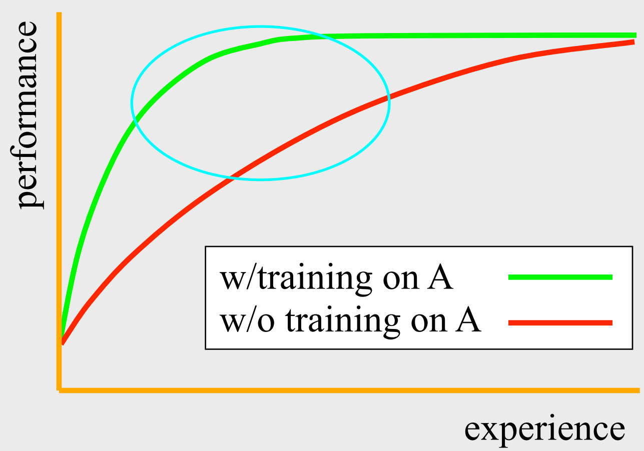
A learner exhibits transfer of learning from task/domain A to task/domain B when, after it has trained on A, it shows improved behavior on B.



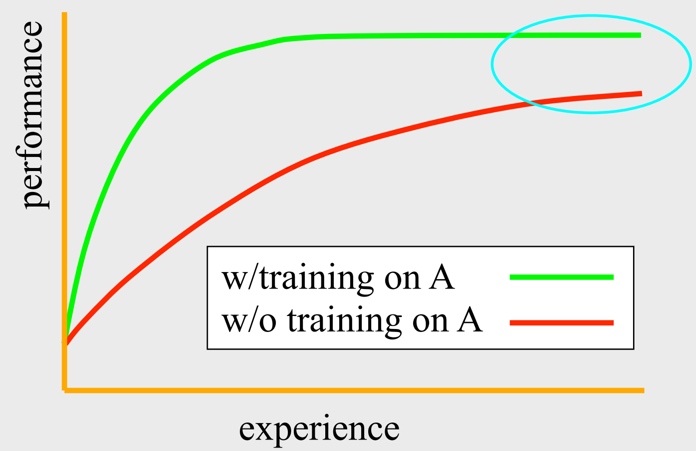
learning curve for task A



better intercept on task B



faster learning rate on task B



better asymptote on task B

What is Transfer?

- Transfer is a *sequential* phenomenon that occurs in settings which involve on-line learning.
  - Thus, multi-task learning does not involve transfer.
- Transfer involves the reuse of knowledge *structures*.
  - Thus, it requires more than purely statistical learning.
- Transfer can lead to improved behavior (*positive* transfer).
  - But it can also produce worse behavior (*negative* transfer).
- Transfer *influences* learning but is not a *form* of learning.
  - Thus, "transfer learning" is an oxymoron, much like the phrase "learning performance".

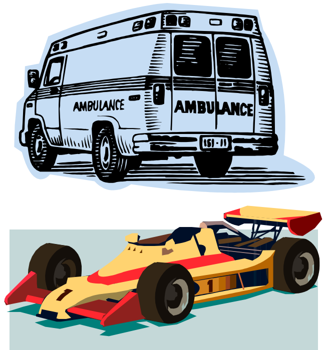
Roots of Transfer in Psychology

The notion of transfer comes from psychology, where it has been studied for over a hundred years:

- benefits of Latin (Thorndike & Woodworth, 1901)
- puzzle solving (Luchins & Luchins, 1970)
- operating devices (Kieras & Bovair, 1986)
- using text editors (Singley & Anderson, 1988)
- analogical reasoning (Gick & Holyoak, 1983)

Some recent studies have included computational models that predict the transfer observed under different conditions.

Domain Classes that Exhibit Transfer



From: tsenator@darpa.mil
To: langley@csli.stanford.edu
Subject: site visit next week
Date: Nov 14, 2004

Pat – I am looking forward to hearing about your progress over the past year during my site visit next week. - Ted

From: noname@somewhere.com
To: langley@csli.stanford.edu
Subject: special offer!!!
Date: Nov 14, 2004

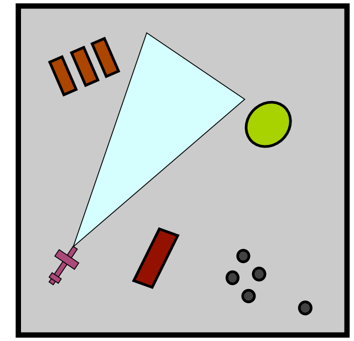
One week only! Buy v*i*a*g*r*a at half the price available in stores. Go now to http://special.deals.com

Which is an emergency vehicle? Which email is spam?

Classification tasks that involve assigning items to categories, such as recognizing types of vehicles or detecting spam. These are not very interesting.

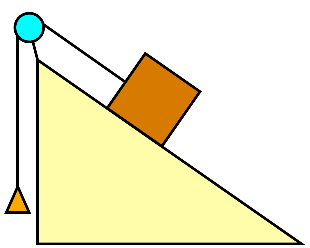
$$\begin{array}{r} 654 \\ -\ 321 \\ \hline \end{array} \quad \begin{array}{r} 456 \\ -\ 237 \\ \hline \end{array} \quad \begin{array}{r} 821 \\ -\ 549 \\ \hline \end{array}$$

$$\begin{array}{r} 940 \\ -\ 738 \\ \hline \end{array} \quad \begin{array}{r} 601 \\ -\ 459 \\ \hline \end{array} \quad \begin{array}{r} 400 \\ -\ 321 \\ \hline \end{array}$$

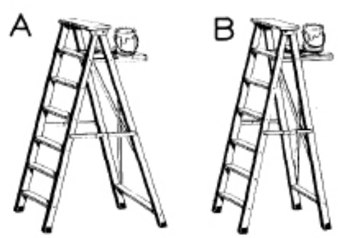


What are the problem answers? What path should the plane take?

Procedural tasks that involve execution of routinized skills, both cognitive (e.g., multi-column arithmetic) and sensori-motor (e.g., flying an aircraft).

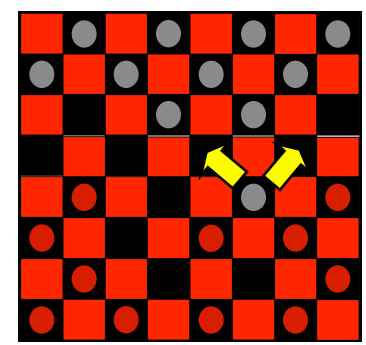


A block sits on an inclined plane but is connected to a weight by a string through a pulley. If the angle of the plane is 30 degrees and . . .



Which ladder is safer to climb on?

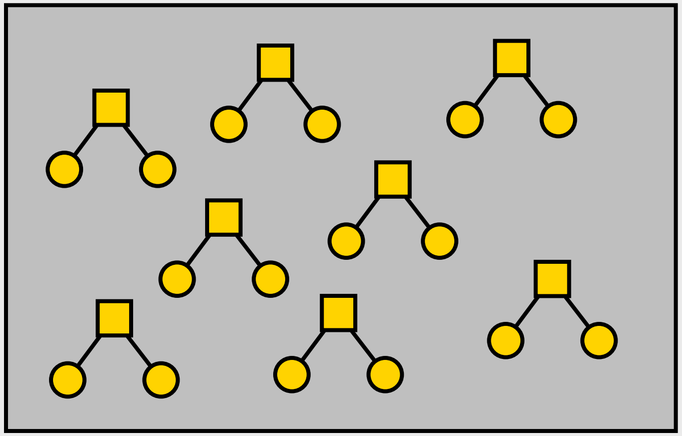
Inference tasks that require multi-step reasoning to obtain an answer, such as solving physics word problems and aptitude/achievement tests.



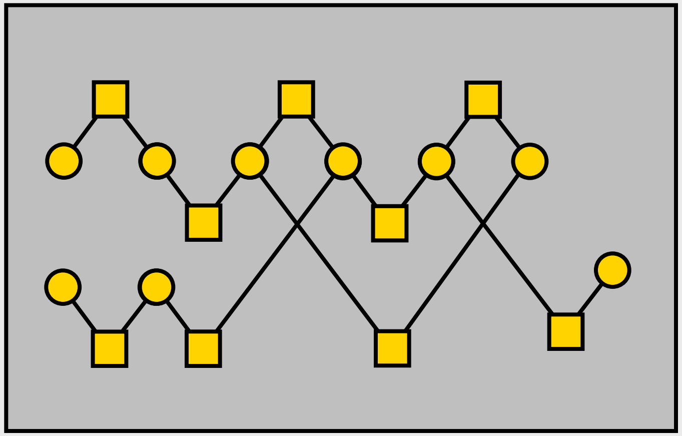
What should the blue team do? Which jump should red make?

Problem-solving tasks that benefit from strategic choices and heuristic search, such as complex strategy games.

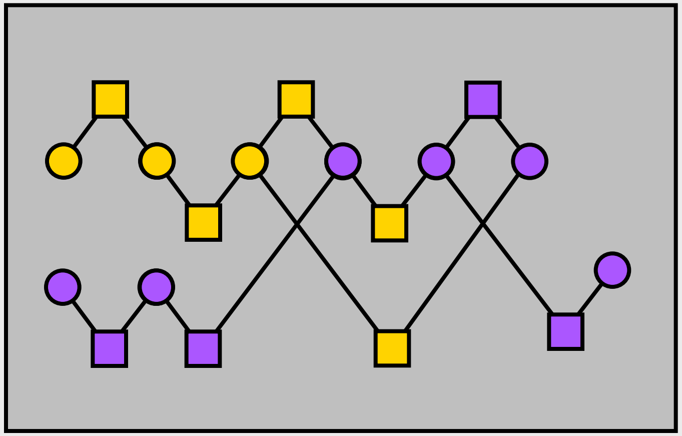
Claims About Transfer



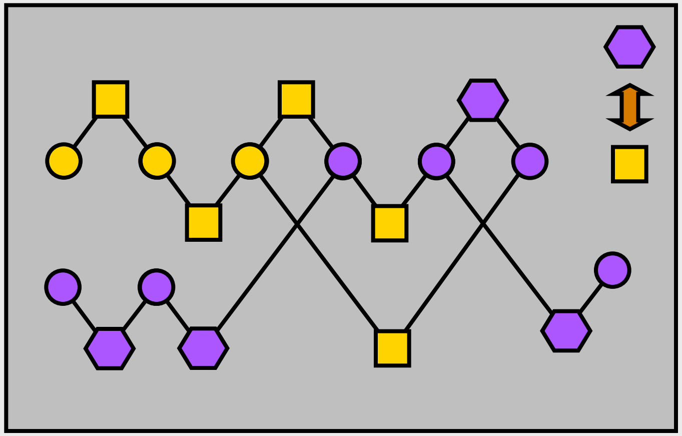
Transfer requires that knowledge be represented in a *modular* fashion.



Transfer requires the ability to *compose* these knowledge elements dynamically.



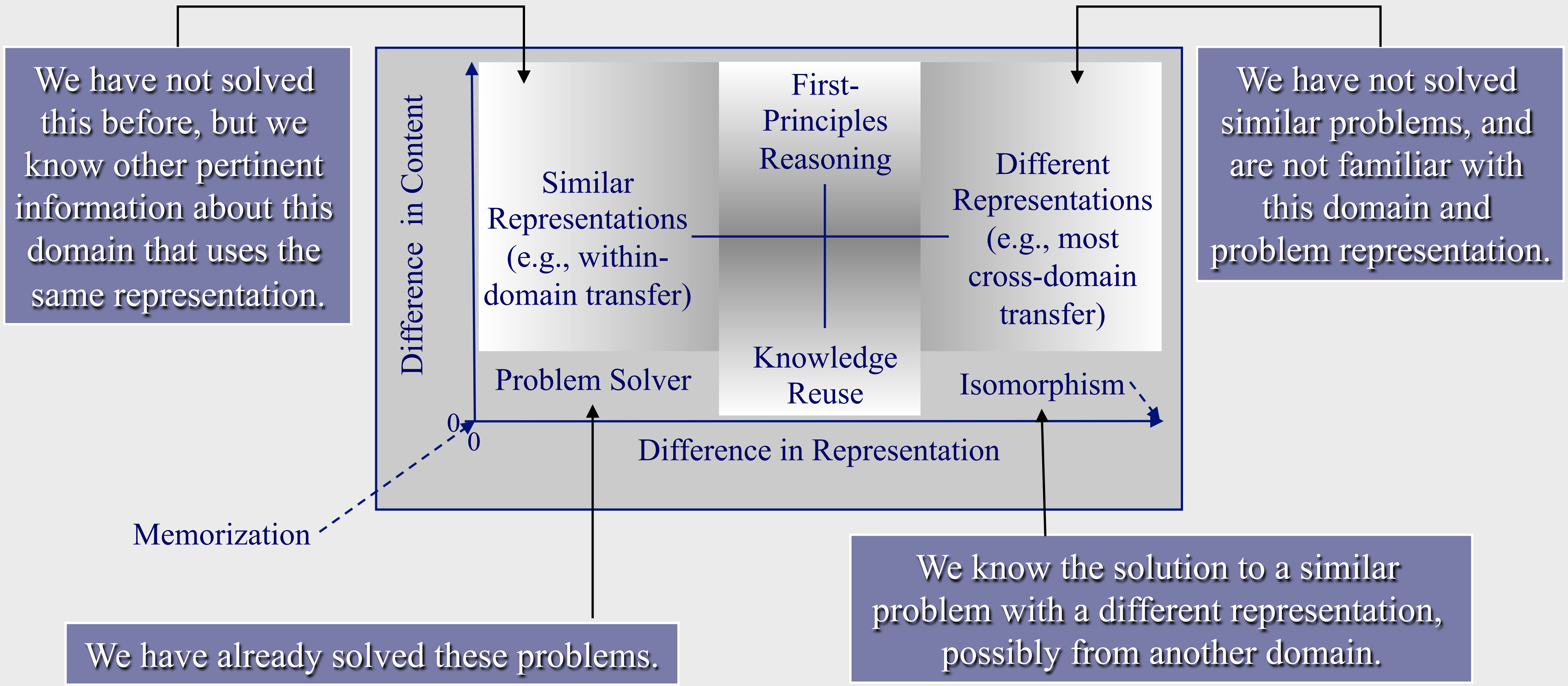
The degree of transfer depends on the *structure shared* with the training tasks.



Transfer across domains requires abstract relations among representations.

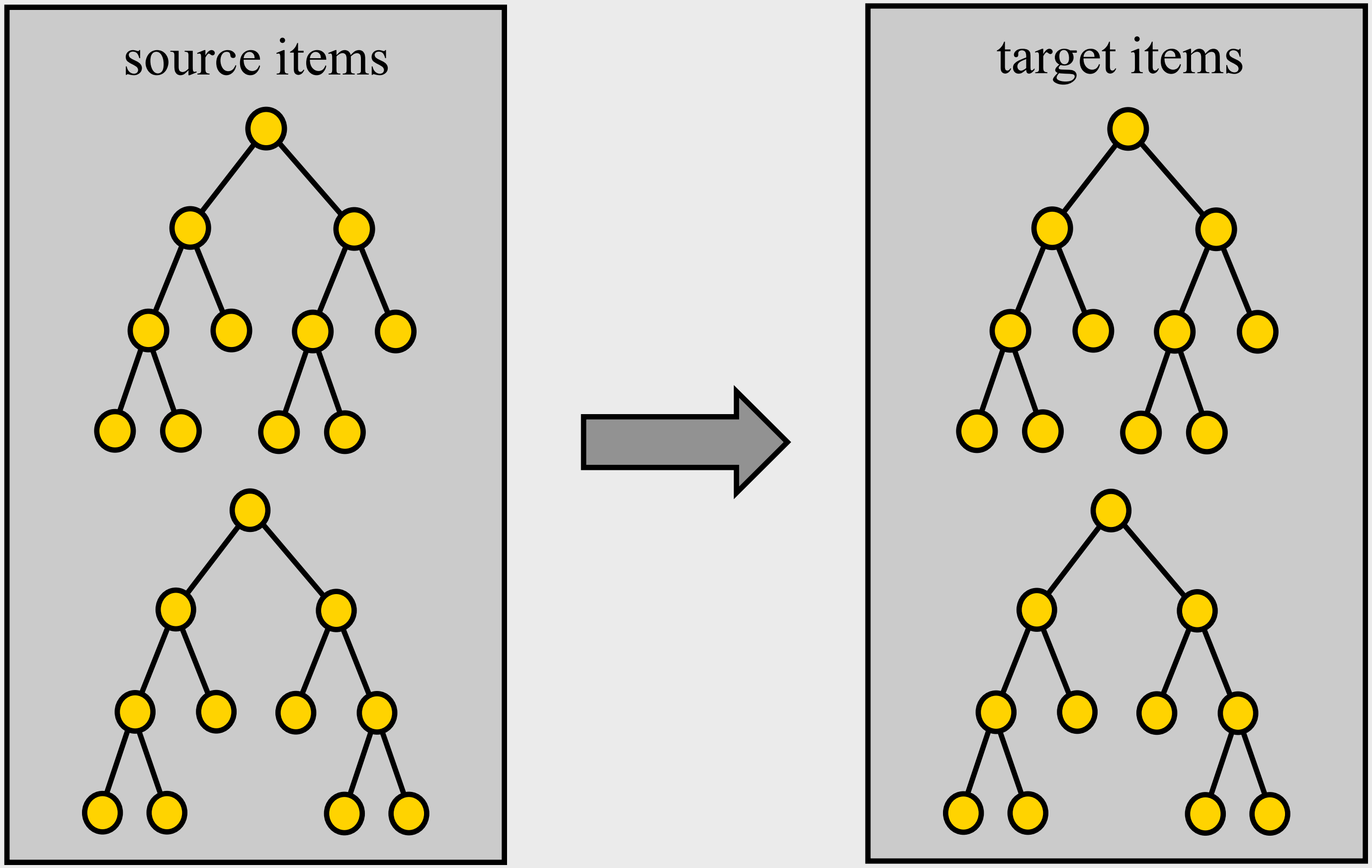
Dimensions of Knowledge Transfer

Knowledge transfer complexity is determined primarily by differences in the knowledge content and representation between the source and target problems.



Memorization

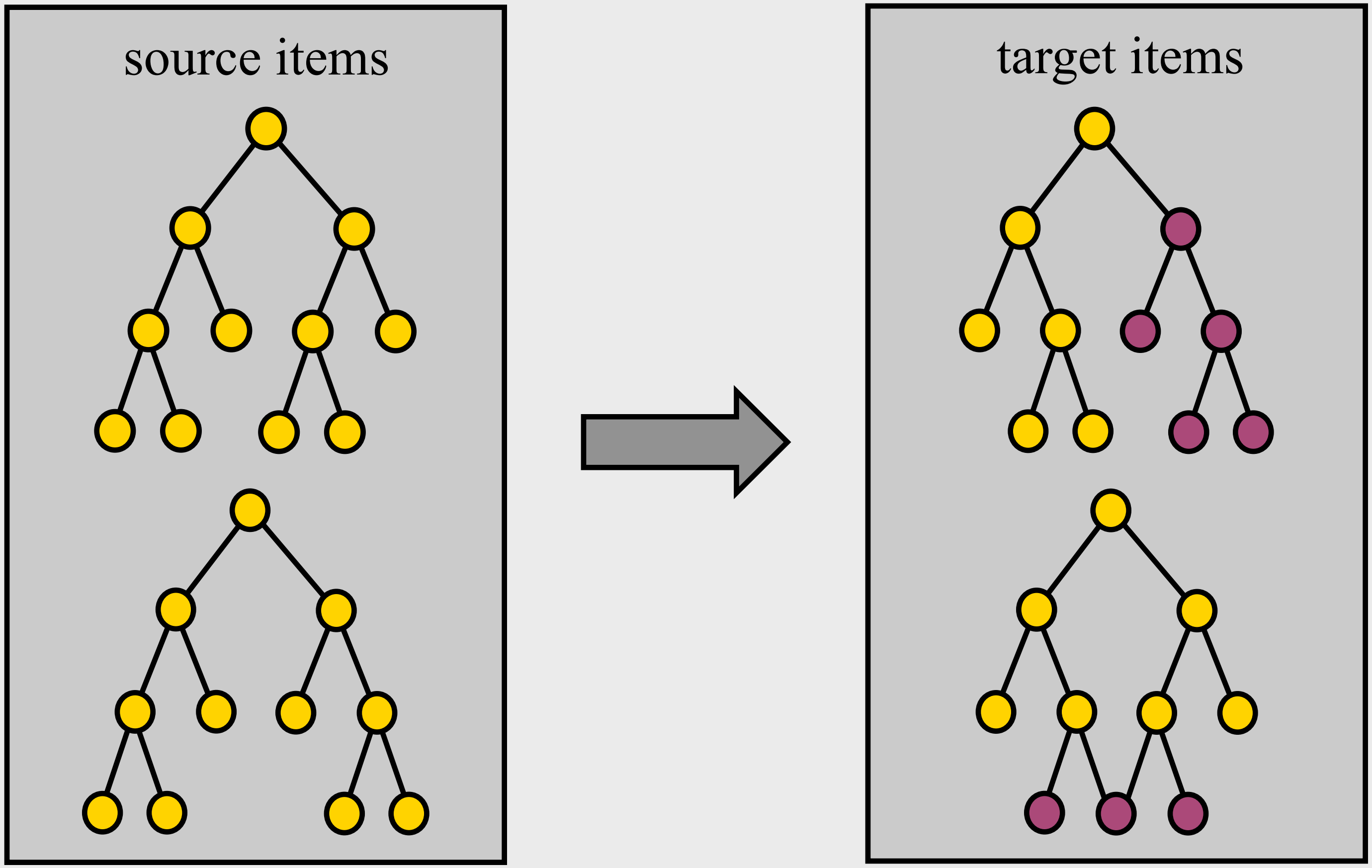
Improvement in which the transfer tasks are the same as those encountered during training.



E.g., solving the same geometry problems on a homework assignment as were presented in class. This is not very interesting.

Within-Domain Lateral Transfer

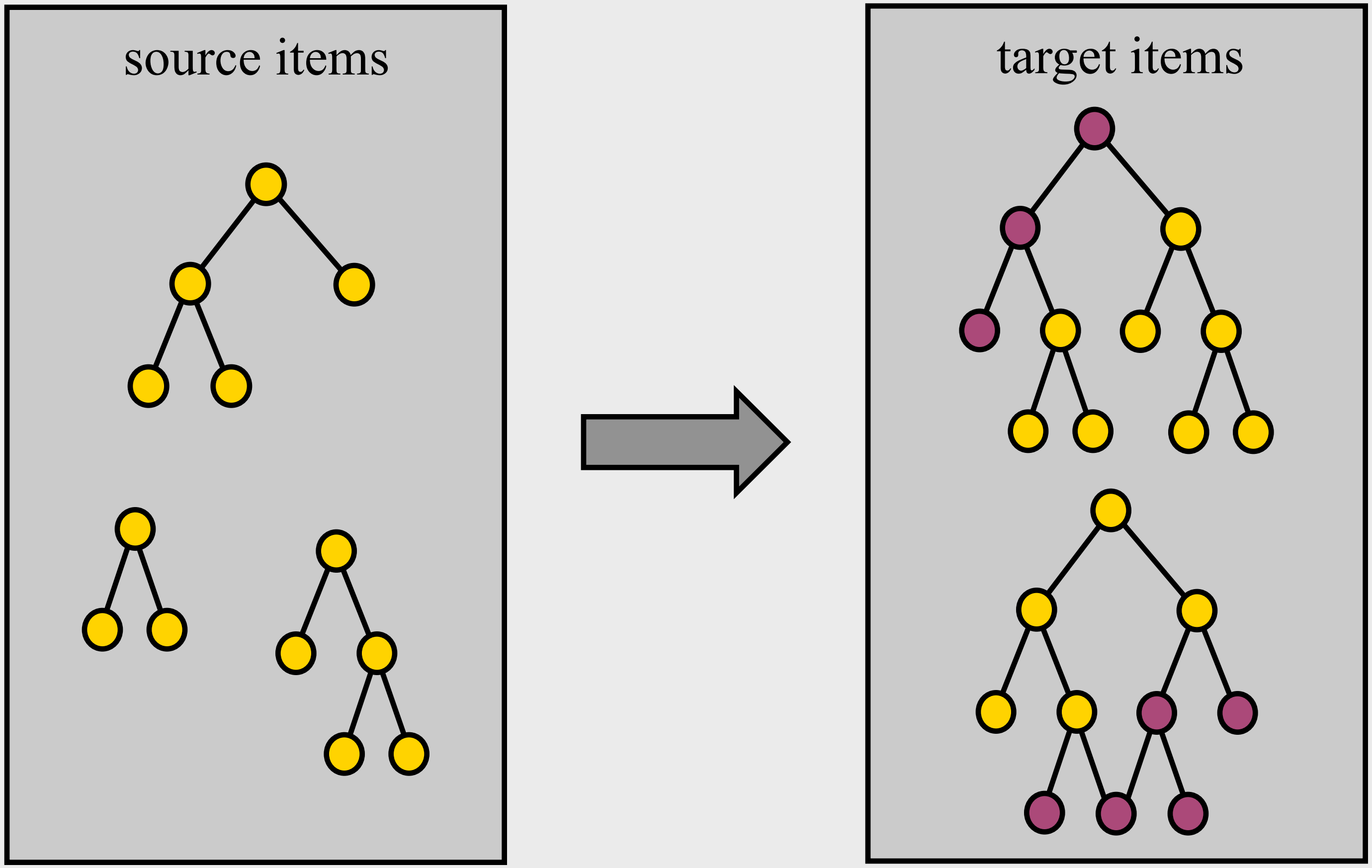
Improvement on related tasks of similar difficulty within the same domain that share goals, initial state, or other structure.



E.g., solving new physics problems that involve some of the same principles but that also introduce new ones.

Within-Domain Vertical Transfer

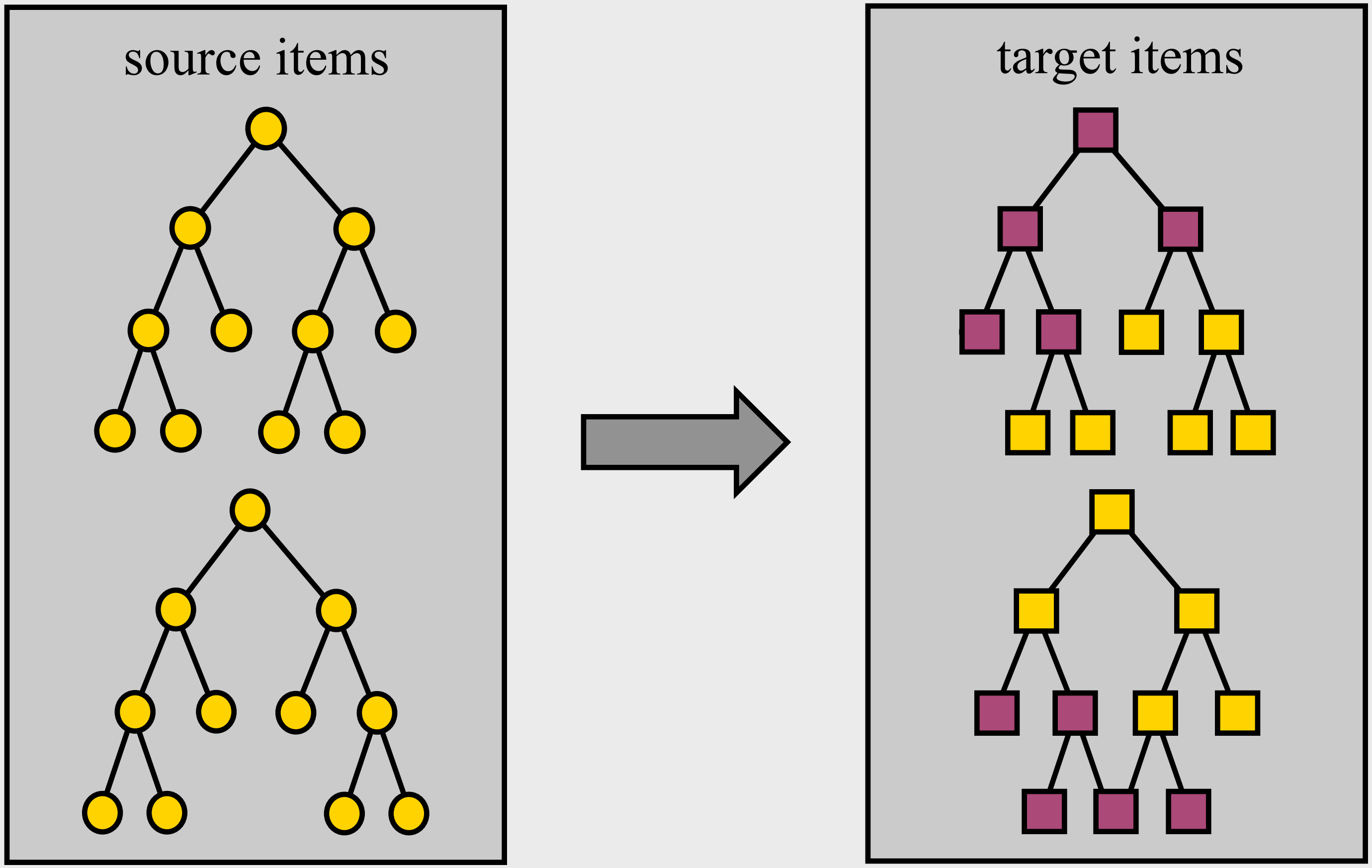
Improvement on related tasks of greater difficulty within the same domain that build on results from training items.



E.g., solving new physics problems that involve the same principles but that also require more reasoning steps.

Cross-Domain Lateral Transfer

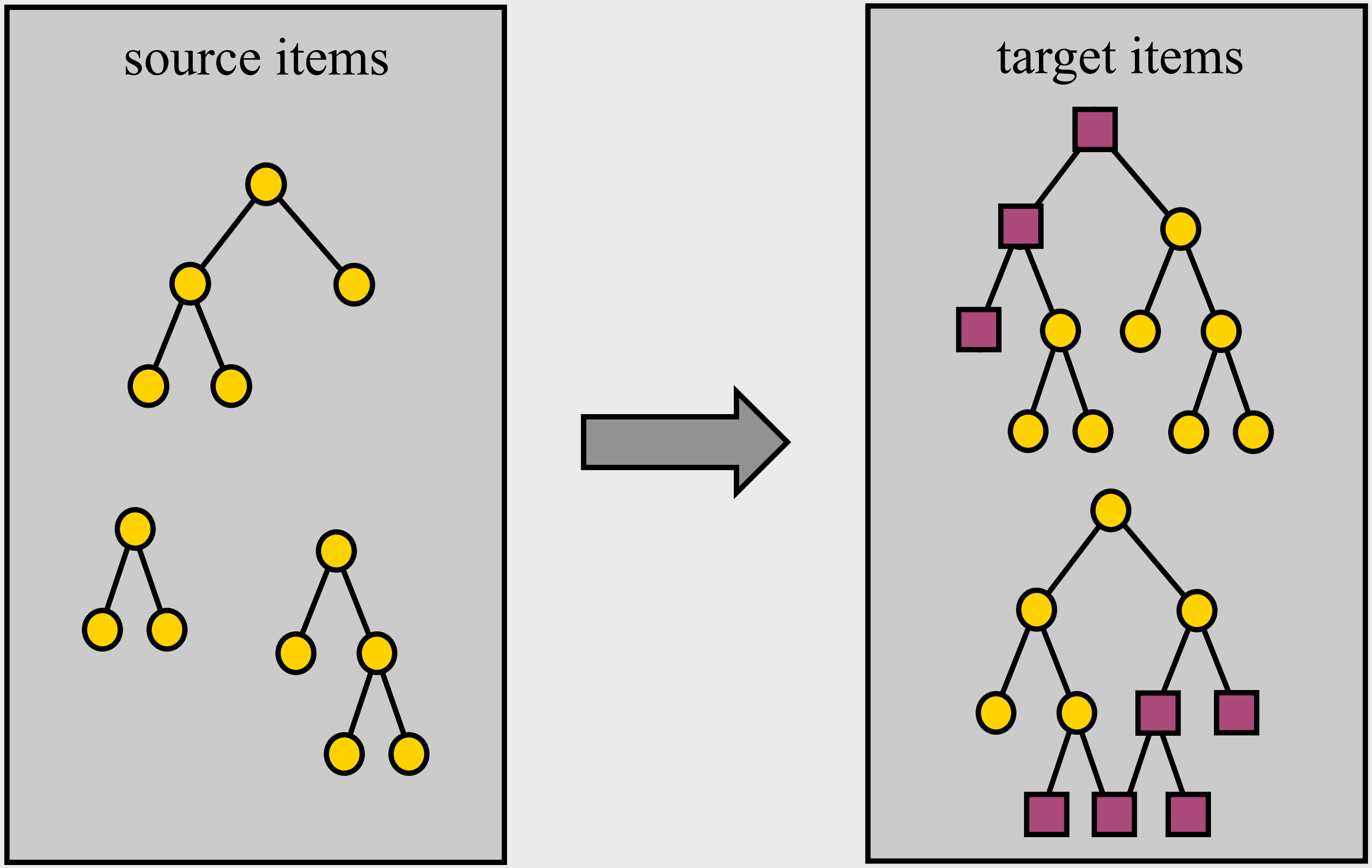
Improvement on related tasks of similar difficulty in a different domain that shares either higher-level or lower-level structures.



E.g., solving problems about electric circuits that involve some of the same principles as problems in fluid flow but that also introduce new ones.

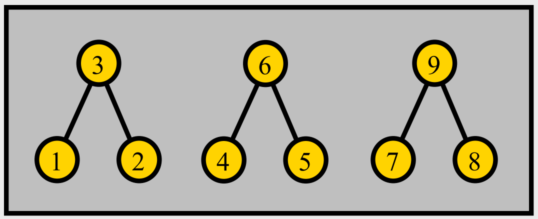
Cross-Domain Vertical Transfer

Improvement on related tasks of greater difficulty in a different domain that share higher-level or lower-level structures.

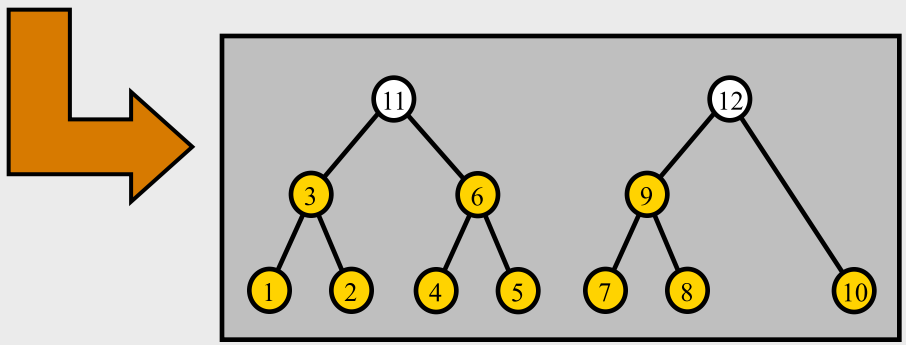


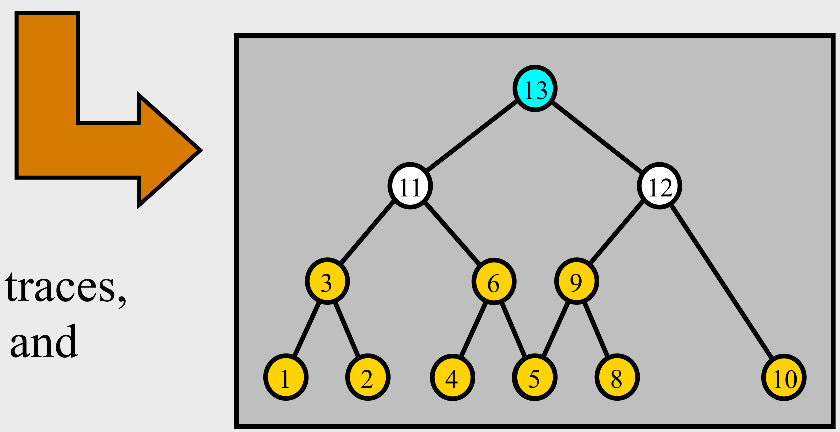
E.g., solving physics problems that require mastery of geometry and algebra or applying abstract thermodynamic principles to a new domain.

Approaches to Transfer: Cumulative Learning



Methods for cumulative learning of hierarchical skills and concepts define new cognitive structures in terms of structures learned on earlier tasks.

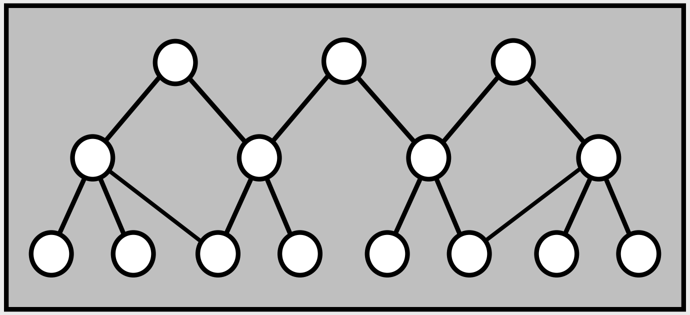




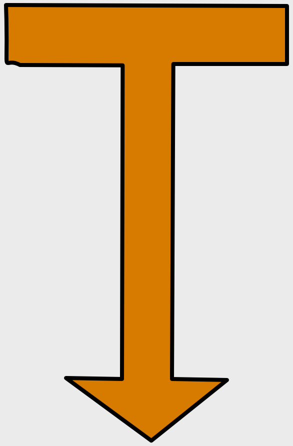
This approach is well suited to support vertical transfer to new tasks of ever increasing complexity.

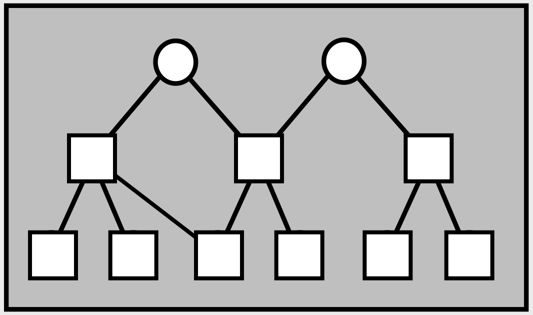
Learning can operate on problem-solving traces, observations of another agent's behavior, and even on direct instructions.

Approaches to Transfer: Analogical Reasoning

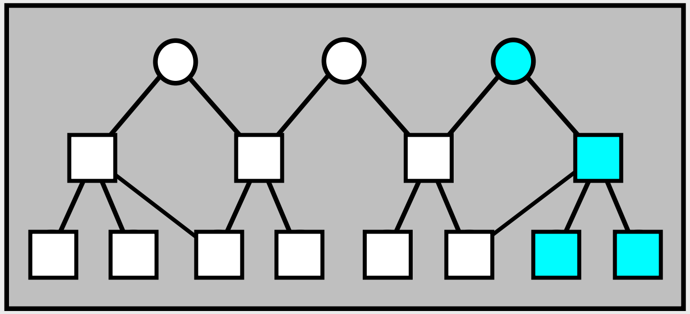


Methods for analogical reasoning store cognitive structures that encode relations in training problems.





Upon encountering a new problem, they retrieve stored experiences with similar relational structure.



Additional relations are then inferred based on elements in the retrieved problem.

Analogical reasoning can operate over any stored relational structure, but must map training elements to transfer elements, which can benefit from knowledge. This approach is well suited for lateral transfer to tasks of similar difficulty.

Approaches to Transfer: Mapping Representations

Voltage Drop

Electrical
Resistance R

Source domain: Electricity

Knowledge: Ohm's law

$$I = \frac{V_1 - V_2}{R}$$

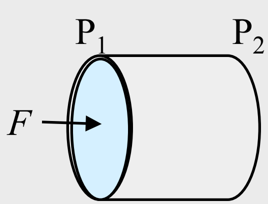
Mapping
Process

$$F = \frac{P_1 - P_2}{R}$$

Knowledge: Poiseuille's law

Target domain: Fluid Flow

Pressure
Drop

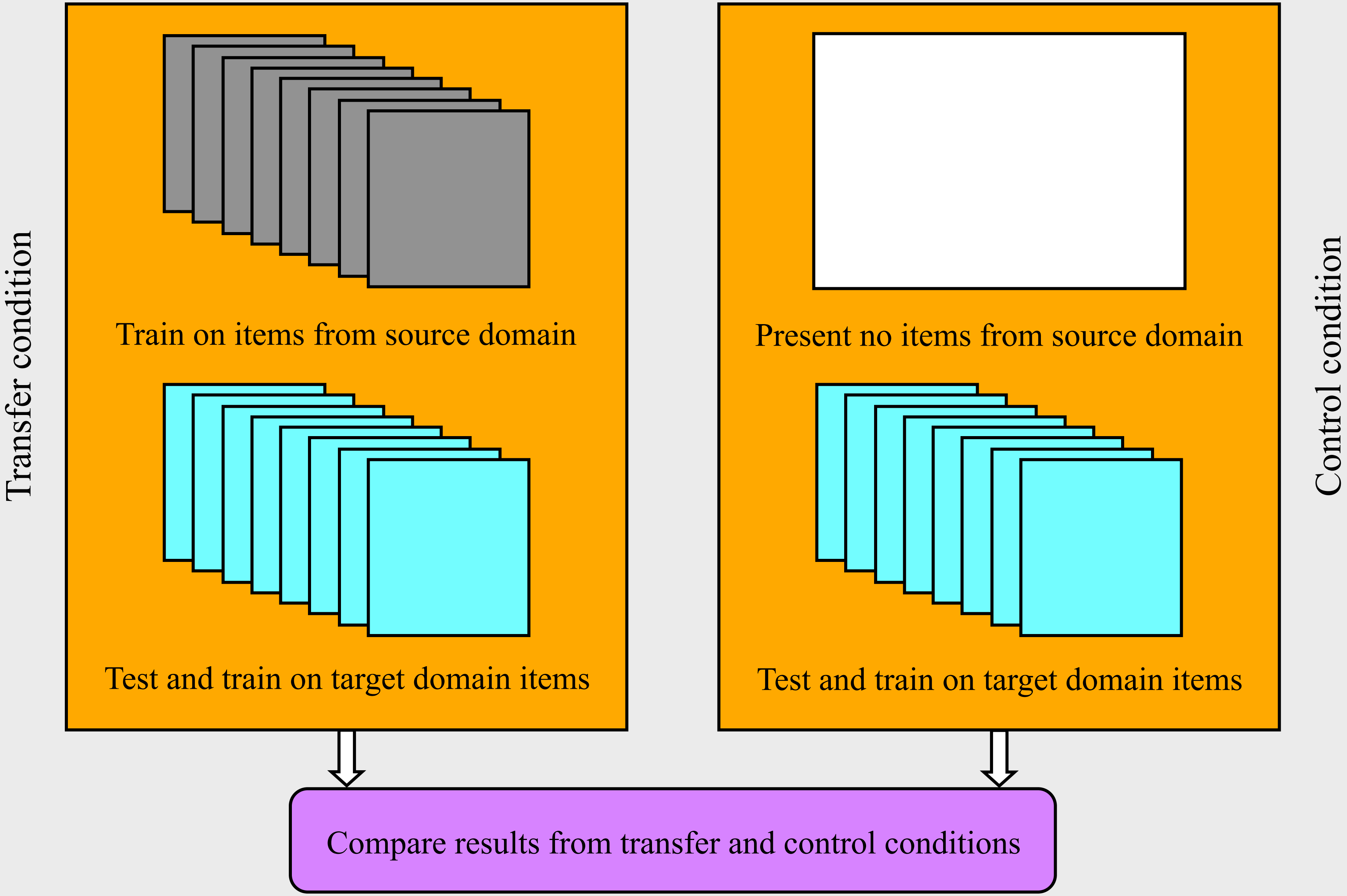


Resistance to
Flow R

Transfer of learned knowledge across domains may require mapping between their representations of shared content.

Q: If P_1=3, P_2=2, and R=2, then what force F is being applied, assuming we only know Ohm's law for electric currents?

Experimental Studies of Transfer



Dependent Variables in Transfer Studies

Dependent variables for transfer experiments should include:

- Initial performance on the transfer tasks
- Asymptotic performance on the transfer tasks
- Rate of improvement on the transfer tasks

These require collecting learning curves over a series of tasks.

Such second-order variables build on basic metrics such as:

- Accuracy of response or solutions to tasks
- Speed or efficiency of solutions to tasks
- Quality or utility of solutions to tasks

Different basic measures are appropriate for different domains.

The ICARUS Architecture

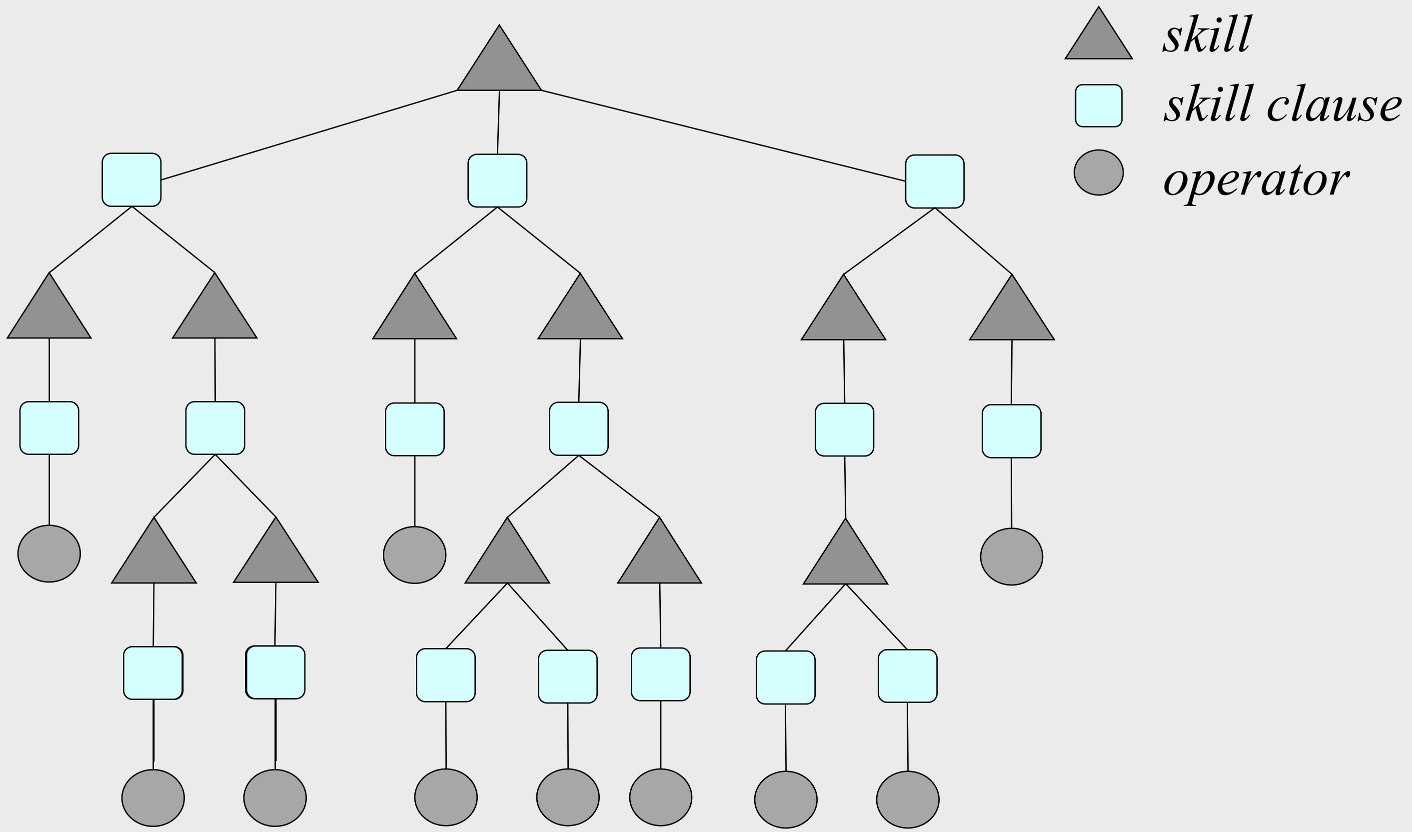
We have studied transfer in ICARUS, an architecture that incorporates some key assumptions from theories of human cognition:

1. Short-term memories are distinct from long-term stores
2. Memories contain modular elements cast as symbolic structures
3. Long-term structures are accessed through pattern matching
4. Cognitive processing occurs in retrieval/selection/action cycles
5. Cognition involves dynamic composition of mental structures

The last of these assumptions is central to ICARUS' account of structural knowledge transfer.

Hierarchical Organization of Skills

ICARUS organizes skills in a hierarchical manner, which each skill clause referring to its component subskills.



Each subtree in the skill hierarchy has the potential for transfer to new problems in which these structures are useful.

Transfer in ICARUS

- What forms of knowledge does ICARUS transfer?
  - *Hierarchical/relational skill and concept clauses*
- Where does the transferred knowledge originate?
  - *It comes from experience on source problems and background knowledge*
- How does ICARUS know what to transfer?
  - *Skills are indexed by goals they achieve, with preference for more recently learned structures*

Synthetic Agents for 'Urban Combat'

Urban Combat is a synthetic environment, built on the Quake engine, used in the DARPA 'Transfer Learning' program.

Tasks for agents involved traversing an urban landscape with a variety of obstacles to capture a flag.

Our experiments with Urban Combat demonstrated that:

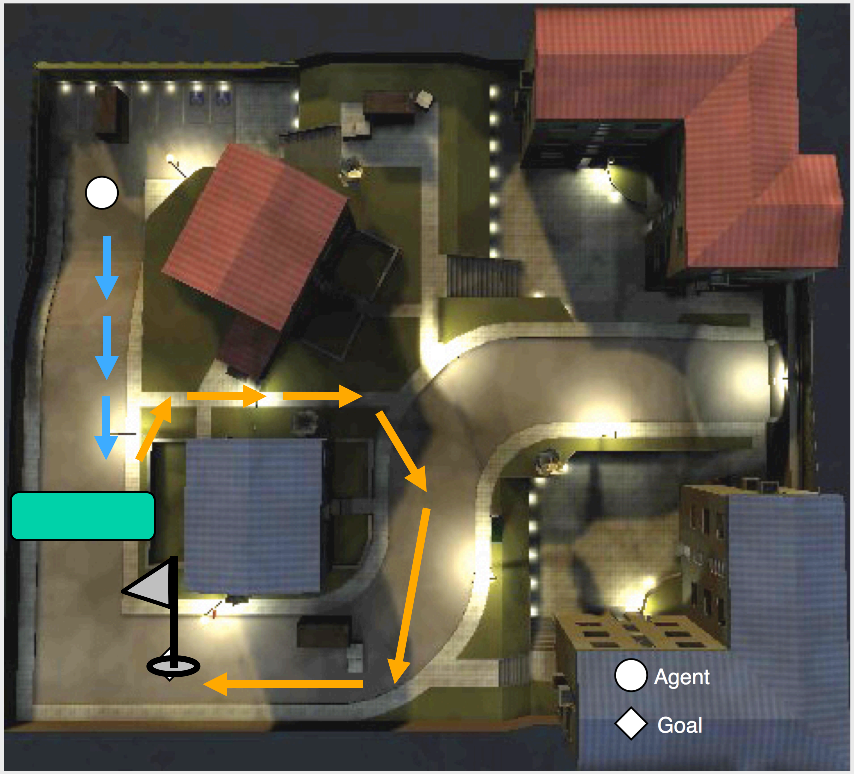
- ICARUS supports simple forms of transfer without modification;
- Deeper types of transfer require a form of analogical mapping.

Both mechanisms relied centrally on the reuse of hierarchical skills.

A Transfer Scenario from Urban Combat

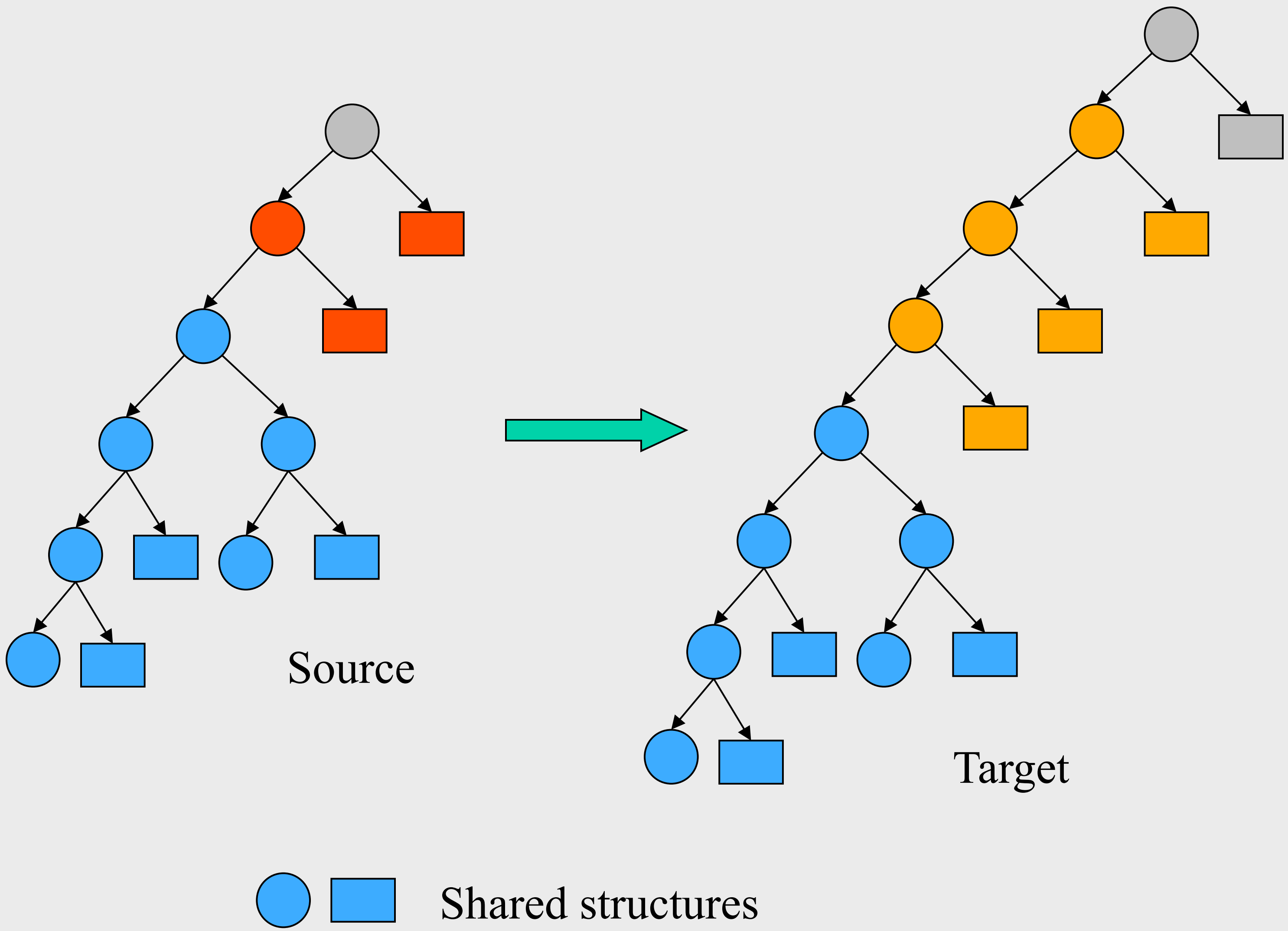
Source Problem

Target Problem



Here the first part of the source route transfers to the target, but the second part must be learned to solve the new task.

Structures Transferred in the Scenario



Key Ideas about Transfer in ICARUS

- The most important transfer concerns *goal-directed behavior* that involves sequential actions aimed toward an objective.
- Transfer mainly involves the reuse of knowledge *structures*.
- Organizing structures in a *hierarchy* aids reuse and transfer.
- Indexing skills by goals they achieve determines *relevance*.
- One can learn hierarchical, relational, goal-directed skills by analyzing traces of expert behavior and problem solving.
- Skill learning can build upon structures acquired earlier.
- Successful transfer benefits from knowledge-based inference to recognize equivalent situations.

Open Research Problems in Transfer

There remain many research issues that we must still address:

- Goal transfer - across tasks with distinct but related objectives
- Negative transfer - minimizing use of inappropriate knowledge
- Context handling - avoiding catastrophic interference
- Representation mapping
  - Lateral - Deep analogy that involves partial isomorphisms
  - Vertical - Bootstrapped learning that builds on lower levels

These challenges should keep our field occupied for some time.

Closing Remarks

Transfer of learned knowledge is an important capability that:

- involves the sequential reuse of knowledge structures
- takes many forms depending on source/target relationships
- has been repeatedly examined within psychology/education
- has received little attention in AI and machine learning
- requires a fairly sophisticated experimental method

Transfer originated in psychology, and it is best studied in the context of cognitive architectures, which have similar roots.