Inferential Theory of Learning and Natural Induction

Ryszard Michalski
Machine Learning and Inference Laboratory
George Mason University

www.mli.gmu.edu/michalski
MLI: www.mli.gmu.edu

PLAN

1. Foundations
2. Role of inference in learning
3. Basics of the Inferential Theory of Learning (ITL)
4. Natural induction
5. Demo of natural induction

Representation Levels

COMPUTERS | **BRAINS**

Physical level:

Integrated circuits (chips, connections, electrical processes)

Neural structures (neurons, dendrites, axons synapses, electro - chemical processes)

***Algorithmic level*:**

Algorithms and data structures

Information processing in the neural networks of the brain

Knowledge level:

Knowledge representation and inference
Knowledge operators and processing

Knowledge representation and inference
High level cognitive processes

What is Knowledge?

- Recorded efforts to define knowledge go back at least to Socrates, who in Plato's Dialogues is credited with a view that

"Knowledge was said by us to be true opinion"

- Antoine Arnauld wrote in "The Port-Royal Logic" (1662):

"Logic is the art of directing reason to knowledge of things for the instruction of both ourselves and others."

- These characterizations are consistent with a "computational definition" of knowledge proposed in ITL:

"Knowledge is inference-enriched and validated information."

Symbol: Any observable entity that represents a choice from a set of predefined choices

Data: An ordered set of symbols

Information: Interpreted data; data given meaning

Knowledge: Inference-enriched and validated information. The inference can be inductive, analogical, or deductive.

Computational Definition of Knowledge

"Knowledge is inference-enriched and validated information."

Inference can be deductive, inductive or analogical.

Three Aspects of Knowledge

- Content
- Organization
- Certainty

WHAT IS LEARNING?

Learning is a process by which a system (a person, an animal, or a computer) increases its knowledge.

Or, briefly,

Learning is a process of increasing knowledge in a system.

Knowledge can be declarative, procedural, and a combination of both forms.

Characteristics of Human Learning

- Humans can learn from inputs in a vast array of forms, using all kinds of inference, generate many kinds of knowledge, and can represent this knowledge in a boundless number of ways.
- When learning knowledge, humans are able to apply diverse learning strategies* and employ them in a flexible, integrated, and goal-oriented manner
- Given a learning goal, humans are able to determine and apply a learning strategy, or their combination that is most suitable to achieving this goal.

*** By a learning strategy we mean here a combination of the primary type of inference, a knowledge representation, and a computational method that is employed in a given learning process (Michalski, 1987, 1993).**

Machine Learning

Is concerned with

-- developing computational methods for learning regardless whether they are realized in nature or not

-- building systems that implement these methods, and

-- applying these systems to real-world problems

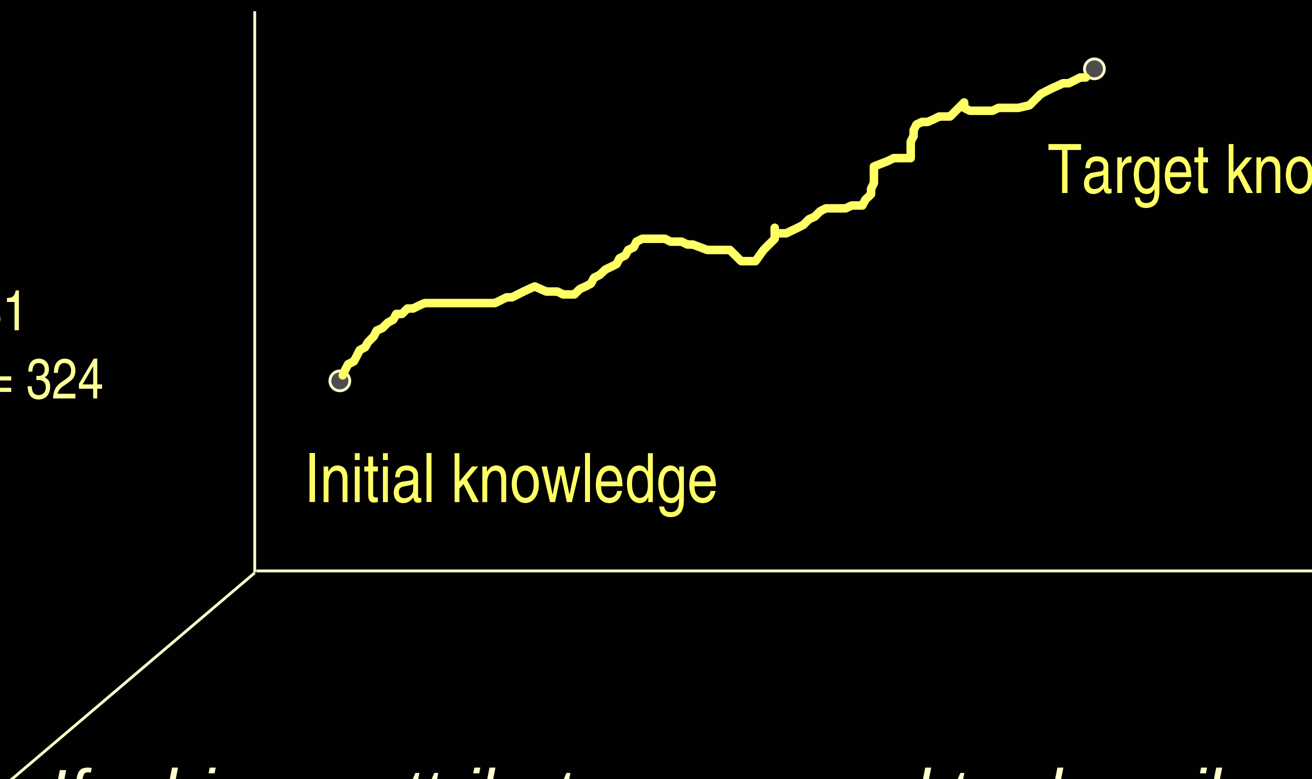
Characteristics of Machine Learning

- Most ML programs execute a single inferential strategy, that employs a specific representational and computational method
- For example, a decision tree learning program can, given examples of different classes, build a decision tree that inductively generalizes these examples.
- The input examples must be in the form of attribute-value vectors. They cannot be in the form of decision trees, graphs, nor complex relational descriptions.
- The output from the program is a decision tree. It cannot be a structural description, nor a semantic net, nor an analogy, nor a deductive consequence of the examples, etc.
- The program does not know what its learning goal is. Its learning goal is defined implicitly by the way the program operates and by its output.
- Similar limitations apply to other programs as well (e.g., rule learners, support vector machines, Bayesian learning, neural nets)

An Imaginary Learning Process

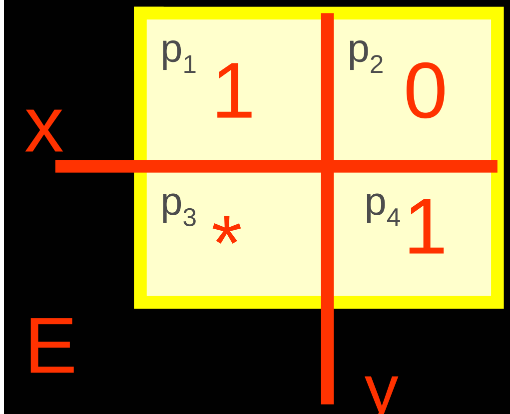
Viewing learning as a search through a knowledge space.

n =2
functions = 81
expressions = 324



If n binary attributes are used to describe entities, the space has 3^{2^n} points for learning one concept, and $2^n \times 3^{2^n}$ expressions.

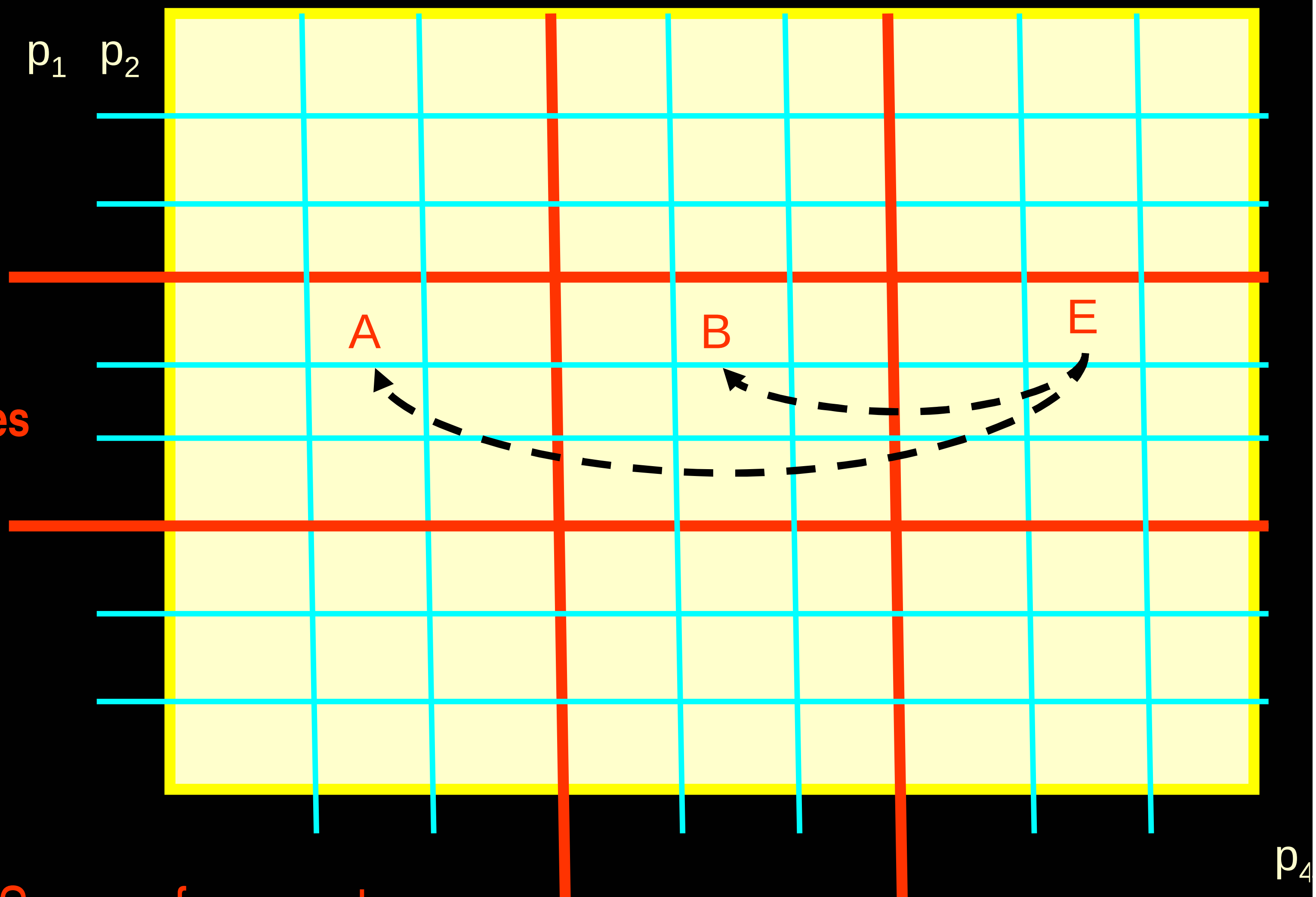
Illustration



Space of instances

A: * is 0

B: * is 1



Space of concepts

Inferential Theory of Learning (ITL)

- ITL aims at developing a unifying framework for characterizing and analyzing diverse forms of learning at the *knowledge* level. It views any act of learning as a process of acquiring or improving knowledge.
- The learned knowledge can be in any form or of any type, declarative or procedural, attributional or structural, demonstrative or hypothetical, conscious or subconscious.
- A learning process is goal-oriented, that is, it strives to acquire knowledge that satisfies some externally given criteria, e.g., helps to make a decision, to carry out a desired task, to solve a problem, to avoid a perceived danger, etc.

Underlying Assumption

Any form of inference may create knowledge that is useful for some purpose.

If that knowledge is memorized and useable for further reasoning, then we can say that the agent (human or machine) has learned this knowledge.

EQUATION FOR LEARNING

Learning = Inferencing + Memorizing

Types of Inference

The Fundamental Equation of Inference

P	$\cup$	BK	$\models$	C
P remise	with	**B** ackground **K** nowledge	Entails	**C** onsequent

where P, BK and C are set of sentences

Deduction: Given P and BK derive C (truth-preserving)

Induction: Given C and BK hypothesize P (falsity preserving)

Analogy: If P' is similar to P, hypothesize C' similar to C (a combination of induction and deduction)

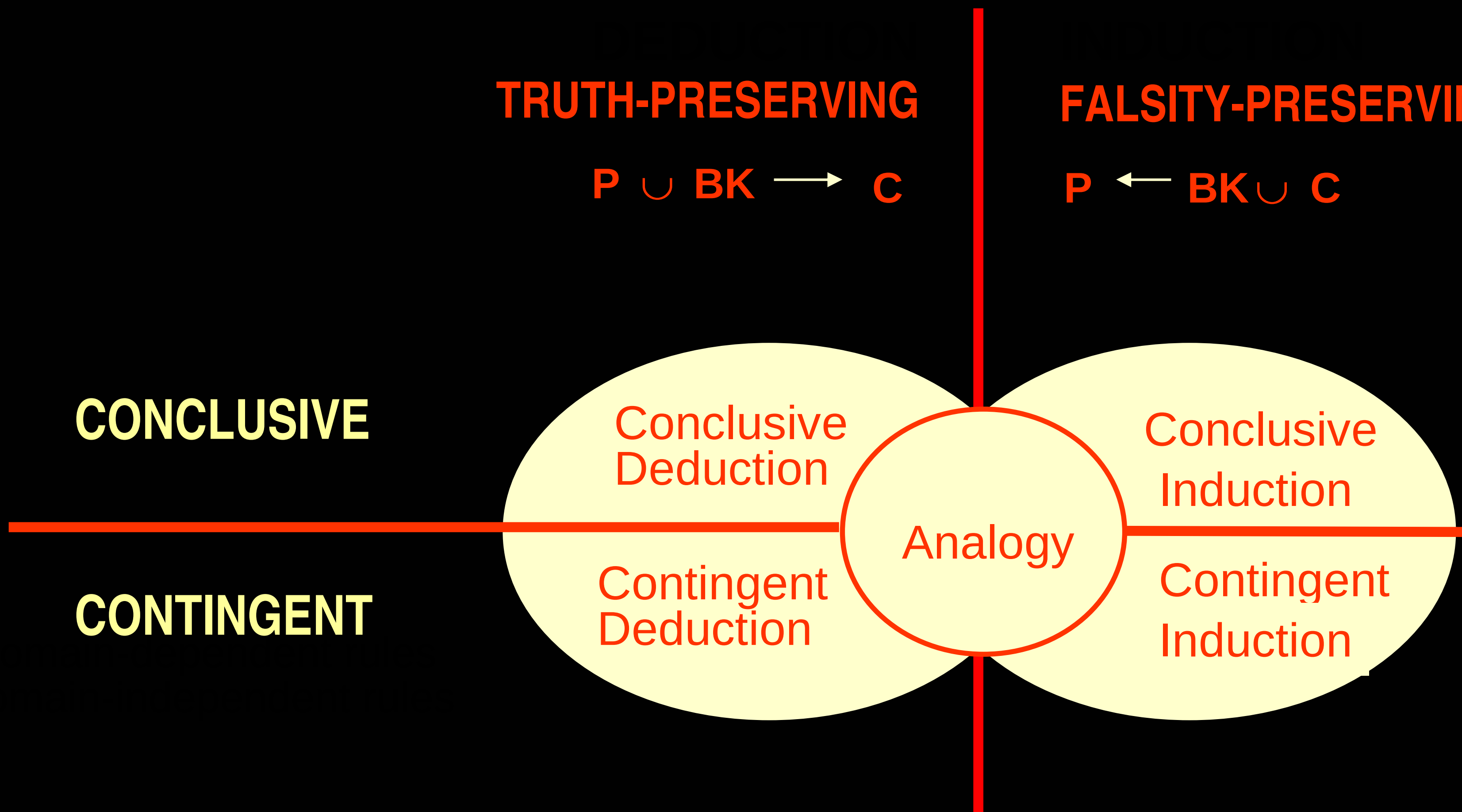
Basic Classification of Inferences

Types of Inference

Preserving

Truth	Falsity	Neither Truth nor Falsity
Deduction	***Induction***	***Analogy***

Types of Inference



Classification of Inductive Inferences

Induction

Inductive Generalization

Abductive Explanation

Inductive Specialization

Inductive Concretion

Inductive Generalization

C: Montreal is an attractive city.
Toronto is an attractive city.
Vancouver is an attractive city.

BK: Montreal, Toronto, Vancouver are cities in Canada

P: Maybe all cities in Canada are attractive.

Test of Inductive Inference:

$$P \cup BK \models C$$

Inductive generalization is frequently identified with induction. ITL, this is just a very important form of induction.

Examples of Other Forms of Inductive Inference

• Inductive specialization

Input: Lives(John, Québec) (John lives in Montréal.)

BK: Montreal $\subset$ Québec (Montreal is a part of Canada.)

$\forall x,y,z,\ y \subset z \& Lives(x,y) \Rightarrow Lives(x,z)$ (Living in x implies living in superset of x.)

Output: Lives(John, Montreal): α (*Maybe* John lives in Montreal.)

• Inductive concretion

Input: Going-to(John, Washington, New York) (John is going from Washington to New York.)

BK: Likes(John, driving) (John likes driving.)

$\forall x,y$, Driving(x,y) $\Rightarrow$ Going-to(x,y) ("Driving to" is a special case of "going to.")

$\forall x,y$, Likes(x,driving)$\Rightarrow$Driving(x,y) (Liking to drive m-implies driving to places)

Output: Driving(John, New York): α (*Maybe* John is driving to New York.)

• Abductive derivation

Input: In(House, Smoke) (There is smoke in the house.)

BK: Fire $\Leftrightarrow$ Smoke: α, β (Fire usually indicates smoke &reverse.)

Output: In(House, Fire): β (*Maybe* there is fire in the house.)

• Constructive inductive generalization (generalization plus abduction)

Input: In(John'sApt, Smoke) (Smoke is in John's apartment.)

BK: Fire $\Leftrightarrow$ Smoke: α, β (Fire usually indicates smoke &conversely.)

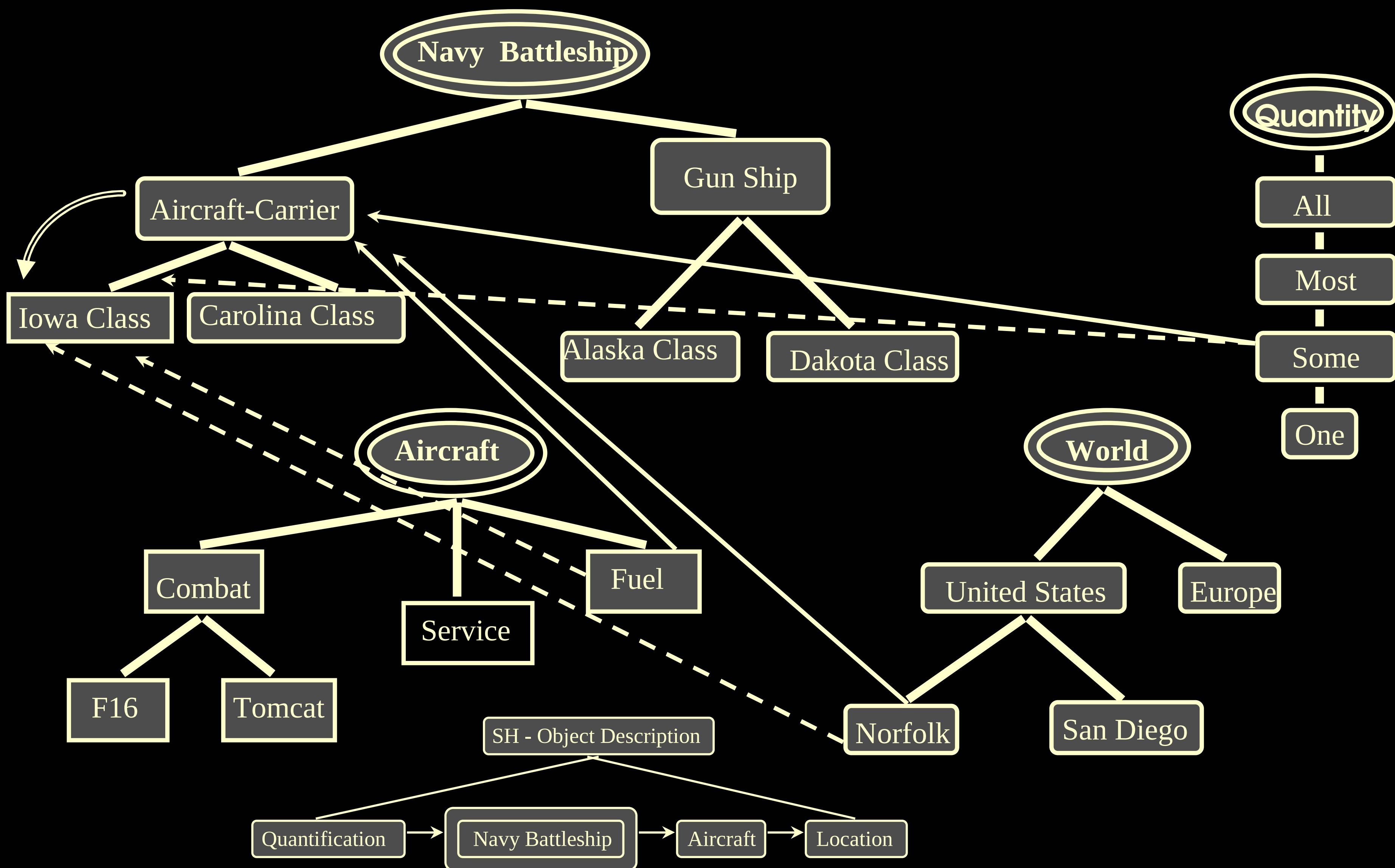
John'sApt $\subset$ GKBld (John's apt. is in the Golden Key building.)

Output: In(GKBld, Fire): β (*Maybe* there is a fire in the Golden Key bld.)

INDUCTIVE SPECIALIZATION

Input: *Some Norfolk aircraft-carriers* have fuel aircraft
BK: Iowa class is a popular type of Navy Battleship

Output: Maybe *some Norfolk aircraft-carriers of Iowa class* have fuel aircraft



Inferential Theory of Learning Views Learning as a Transformation

Given:

- Input information $I = \{ I_i \}$
- Initial knowledge $K = \{ K_i \}$
- Goal specification $G = \{ G_i \}$
- Transmutations $T = \{ T_i \}$

Determine:

- New knowledge, K', that satisfies goal G, by applying knowledge transmutations, T, to K and I.

Transmutations

- Transmutations are generic types of knowledge change
- They change one or more aspects of knowledge, i.e., its contents, organization, and/or its certainty.
- Formally, a transmutation takes as arguments a set of sentences (S), a set of entities (E), and background knowledge (BK), and generates a new set of sentences (S'), and/or new set of entities (E'), and/or new background knowledge (BK'):

T: S, E, BK ---> S', E'

Knowledge Generation Transmutations

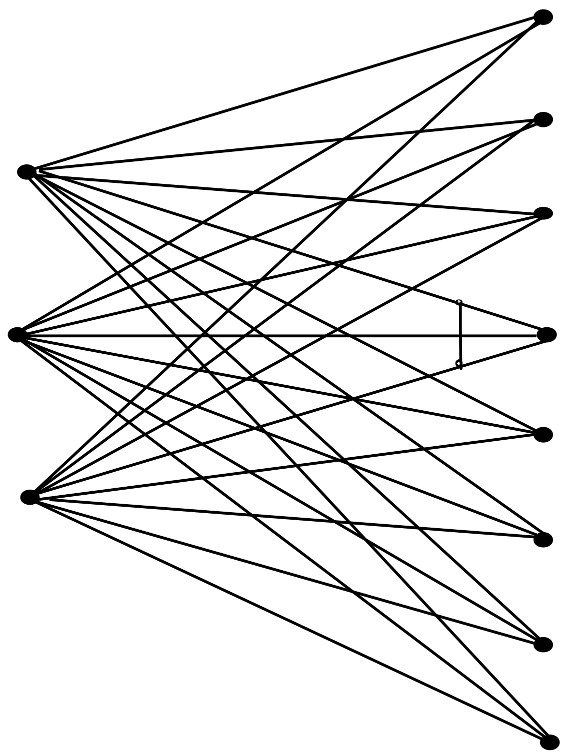
A Selection

Inference Type

DEDUCTION

ANALOGY

INDUCTION



Transmutation

Generalization
Specialization

Abstraction
Concretion

Explanation
Prediction

Similization
Dissimilization

Selection
Generation

Agglomeration
Decomposition

Characterization
Discrimination

Association
Disassociation

Generalization vs. Abstraction

Definition:

Reference set ---the set of entities being described by or referred to in a set of sentences. The set of sentences is called a *description* of the reference set.

- *Generalization* (*specialization*) increases (decreases) the reference set
- *Abstraction* (*concretion*) decreases (increases) the amount of detail specified in the description of the reference set.

Example of Knowledge Transmutations

INPUT: “color(my-office-pencils, light-blue)”

GENERALIZATION:

“color(all-my-pencils, light-blue)“

SPECIALIZATION:

“color(my-desk-pencils, light-blue)”

ABSTRACTION:

“color(my-office-pencils, blue)”

CONCRETION:

“color(my-office-pencils, light-sky-blue)”

GENERALIZATION and ABSTRACTION:

“color(all-my-pencils, blue)“

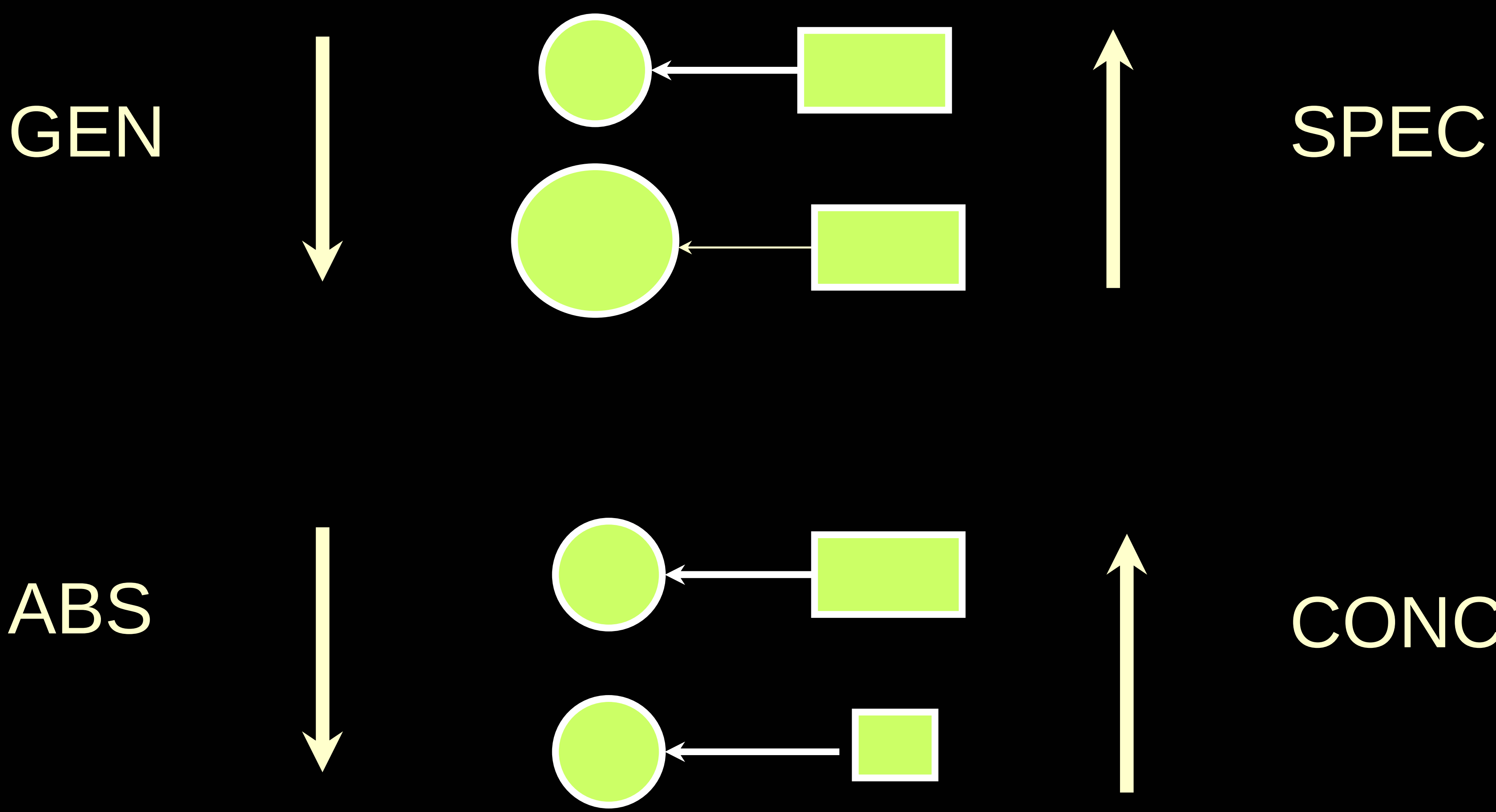
PACKET OF KNOWLEDGE (PAK)

REFERENCE SET   DESCRIPTION

α –– certainty, BK-Background Knowledge

John is 7’2’’ tall. This is a good book.
Silver Queen is a gondola in Aspen.
Trees have leaves. $E = mc^2$

Generalization vs. Abstraction



Natural Induction and AQ Learning

- Natural induction is a process of inducing hypotheses in the forms that appear natural to people, and by that easy to interpret and understand
- Such forms include logic-style rules, natural language-like descriptions, and various graphical forms and visualizations.
- AQ learning is a progressive covering algorithm (a.k.a "separate-and-conquer") that uses a more *attributional calculus*, a logic system that integrates elements of propositional, predicate, and multi-valued logics
- Unlike conventional rules that use attribute-value or *attribute-rel-value* conditions, attributional calculus rules can involve more elaborate conditions, such as

  blood-type = A, weight > 200 Lb, color = red or blue or green

  x = 2..8   x1 = x2   x1 & x3 ≥ x5, Count{x1,x2,x4,x6,x7 EQ 2} ≥ 3
- By using a more expressive description language, AQ learning can discover compact and understandable regularities in data
- Descriptions can be optimized according to task-oriented criteria
- Descriptions can be evaluated using an exact match or a flexible match

Attributional Calculus

Attributional calculus combines elements of propositional logic, predicate logic and multiple-valued logic in order to provide a simple representation system supporting *natural induction*---a process of inducing human-type knowledge from computer data.

An important construct is an attributional rule:

CONDITION => DECISION

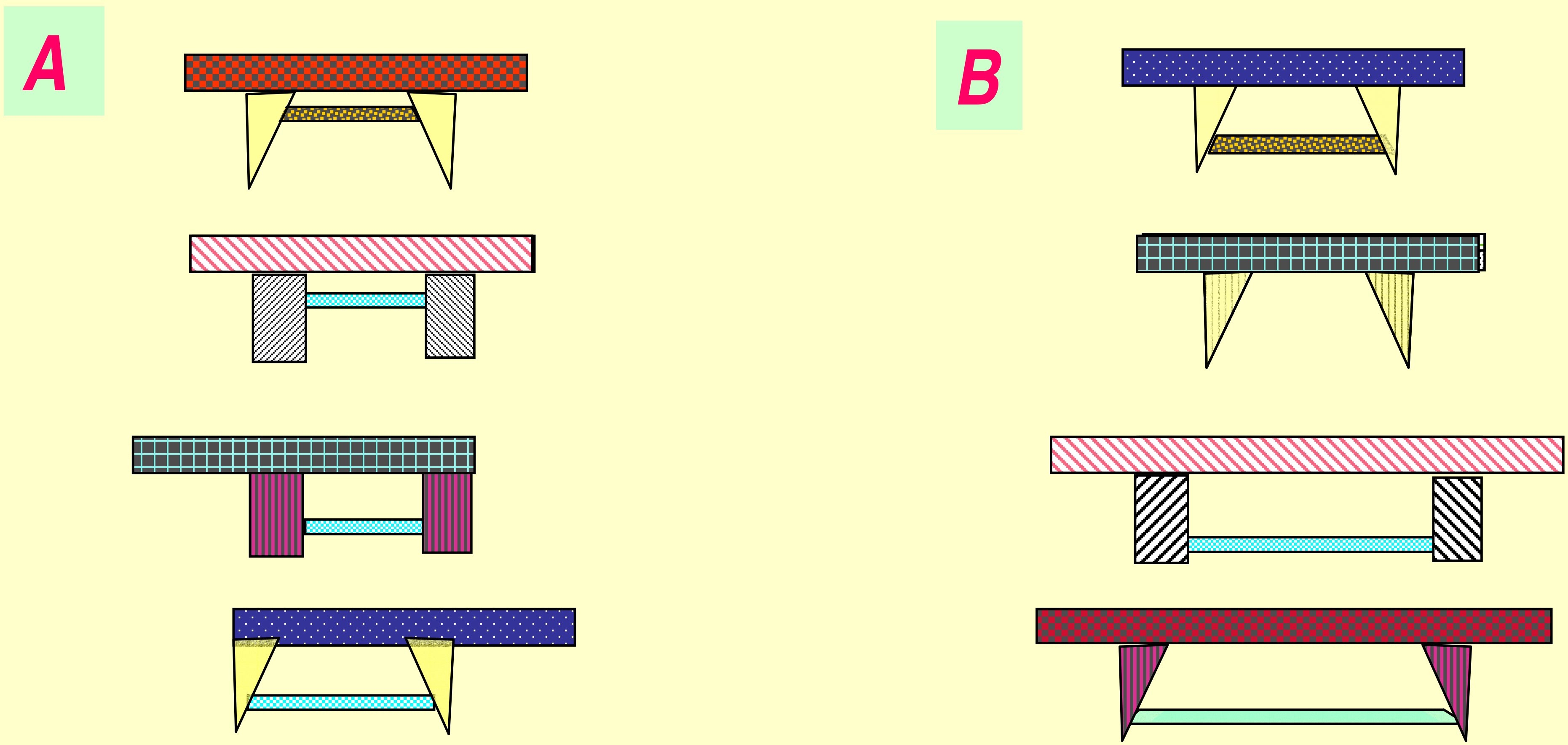
where ***CONDITION*** is a conjunction of *attributional conditions*, and ***DECISION*** is an elementary *attributional condition*. An attributional condition is in the form:

[L rel R]

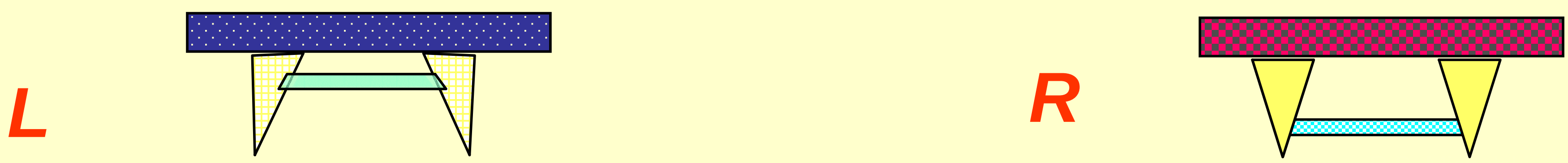
where ***L*** is one or more attributes joined by "&" or "v"
R is a value, a list of values joined "v", a pair of values joined by ".." or an attribute
rel is a relational symbol from the set $\{ =, \neq, \geq, >, <, \leq \}$.

An Illustration of Natural Induction in VINLEN

Given are samples of tables produced by companies A and B



Determine which company might have produced these tables



A Result of Natural Induction

1. Hypothesis creation (by AQ learning)

Company A ← If top is asymmetrical, or cross-bar is high

Company B ← If top is symmetrical and cross-bar is low, or there is no cross-bar

2. ***Hypothesis matching:*** The left table matches hypothesis for company A, and the right table matches hypothesis for company B.

Data Mining

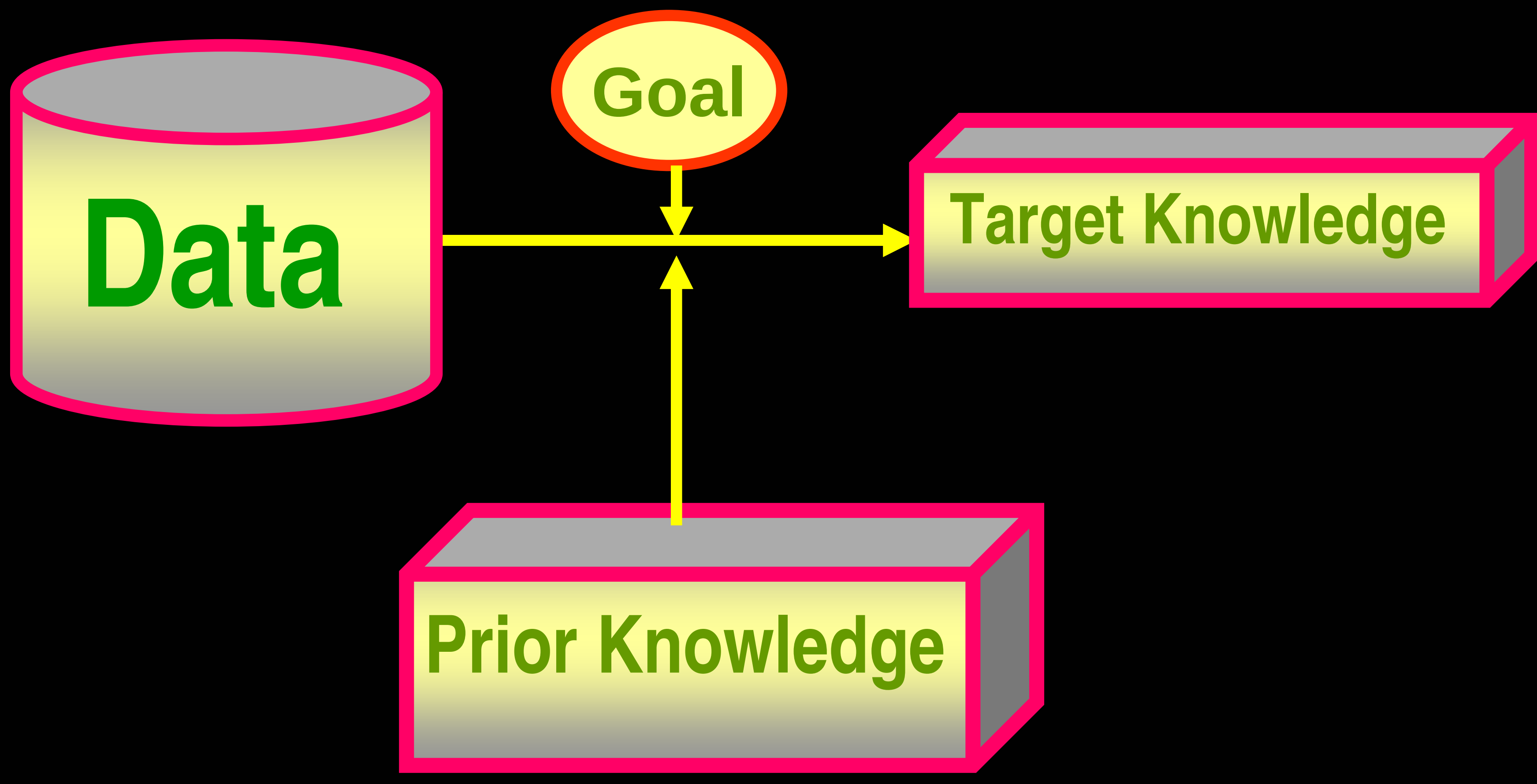
DATA $\rightarrow$ PATTERNS

- Association Rules
- Decision Trees
- Statistical Summaries
- Bayesian Networks
- Evolutionary Algorithms
- Nearest Neighbors
- Neural Nets

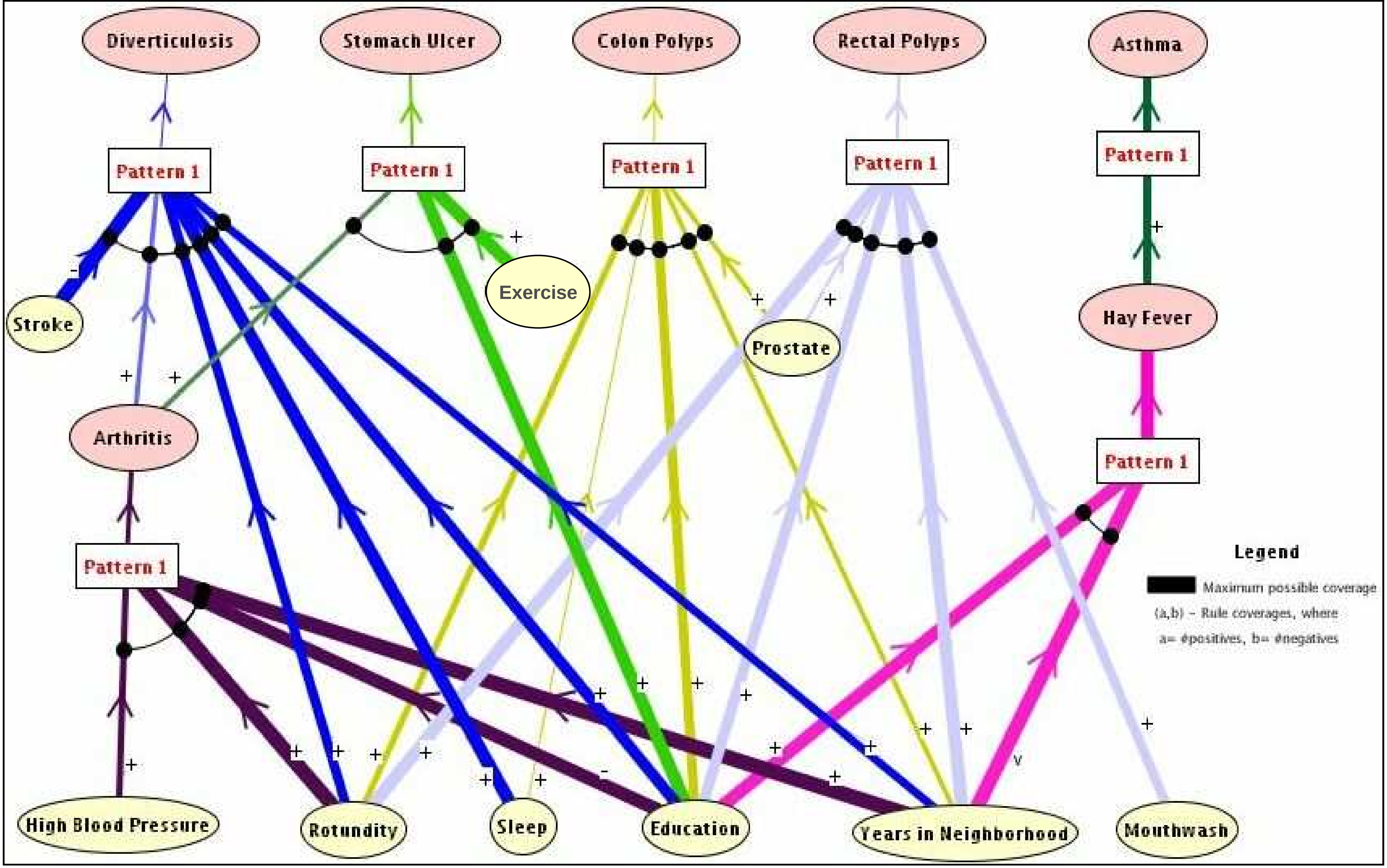
Knowledge Mining

Engages Data and Prior Knowledge to Create Target Knowledge

DATA + PRIOR _KNOWLEDGE + GOAL $\rightarrow$ TARGET_KNOWLEDGE



A Concept Association Graph for Medical Domain:
Link's Thickness Represents Relative Support



Potential Value and Applications of ITL and NI

- Can provide ideas and clues for developing more powerful cognitive models of learning and inference
  (E.g., Collins and Michalski, "A Core Theory of Human Plausible Reasoning," *Cognitive Science*, 1990; papers by Michalski
  www.mli.gmu.edu; select *Papers for 1990, and subsequent papers on plausible reasoning*

- Has many applications: multistrategy learning, data mining and knowledge discovery, knowledge mining, inductive databases, non-Darwinian evolutionary computation, optimization, engineering design, and others.
  www.mli.gmu.edu; select *Papers*

Summary

- ITL views any form of learning, in humans, animals, or machines as a process of increasing knowledge in the system
- Knowledge is an inference-enriched and validated information
- Natural induction is a process of generating inductive hypotheses in the forms natural for people, and by that easy to understand and interpret
- Desirable research directions: analysis of all forms transmutations, implementation of transmutations in humans, animals and computers, advancing methods for natural induction, measuring knowledge comprehensibility, exploring the relationship of natural induction to human learning and knowledge discovery, and other.

For more information,
relevant publications, and/or for
downloading MLI Laboratory's software,
visit:

http://www.mli.gmu.edu

If you have any questions, contact:

Ryszard Michalski michalski@gmu.edu
or my collaborator
Ken Kaufman kaufman@gmu.edu

Acknowledgement

This research has been supported in part by the National Science Foundation under Grants No. IIS-9906858 and IIS-0097476, and in part by the UMBC grant under LUCITE Task #32.

A Demo of Natural Induction

iAQ

To download:

http://www.mli.gmu.edu/mlisoftware.html