Artificial Intelligence

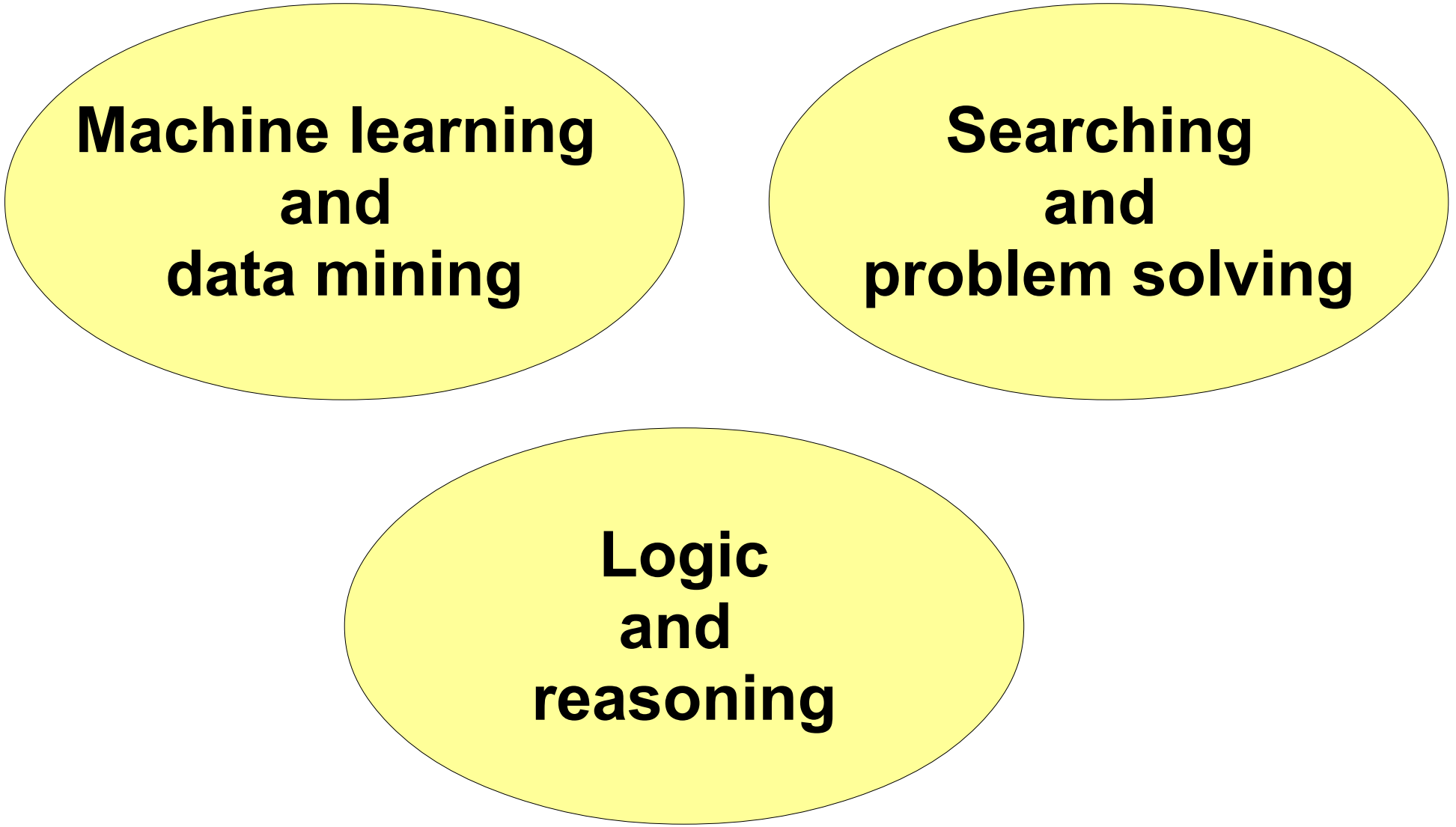
LECTURE 1

Artificial Intelligence (AI)

- There is no standard definition of exactly what artificial intelligence is.
- One of the definitions:

The study of making computers do things that the human needs intelligence to do.

Selected Areas of AI



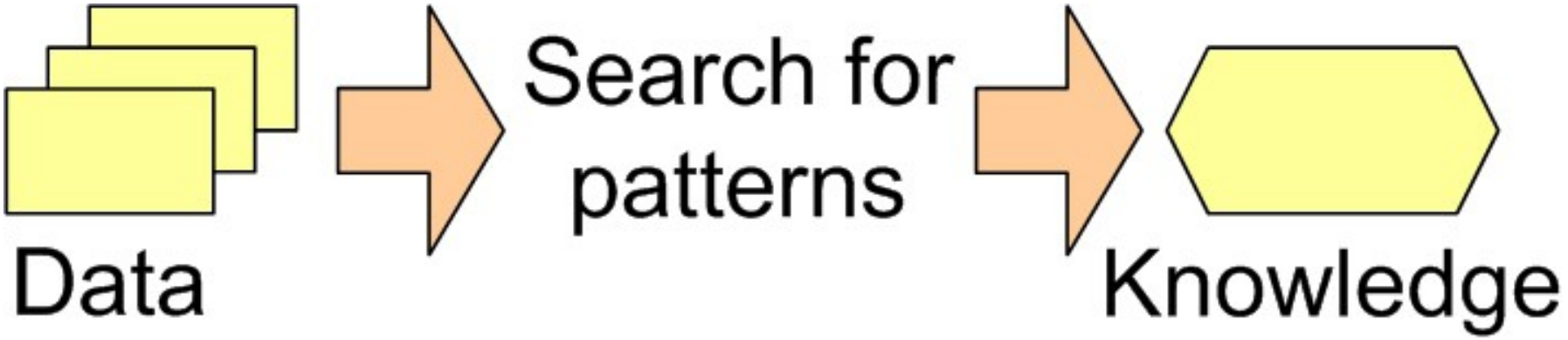
Machine Learning

- Machine Learning (ML) is the study of computer algorithms that improve automatically through experience.
- Supervised learning is the machine learning task of inferring a function from labeled training data (i.e., data with desired output values), e.g.:
  - **Data classification**
- Unsupervised learning is the machine learning task of inferring a function to describe hidden structure from "unlabeled" data, e.g.:
  - **Data clustering**

Data Mining and Knowledge Discovery

- Data Mining - the task of a learning machine to extract knowledge from training data.
- Data mining has been defined as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from data".
- The term pattern is used in a broad sense. It may be understood as relationships, correlations, dependencies, rules, trends, etc.

General Scheme of Knowledge Discovery



Data Mining Methods Used for Knowledge Discovery

- Statistics methods.
- Fuzzy sets.
- Rough sets.
- Machine learning methods.
- Evolutionary computing.
- Neural networks.

Data Representation

- Data can be presented as a table.
- Each row of a table represents an object (a case, an event, a patient).
- Each column of a table represents an attribute (a variable, an observation, a property) that can be measured for each object. It can also be suplied by a human expert or user.
- Entries of the table are attribute values.
- A table represents the so-called information system.

Information Systems

$$S=(U,A)$$

U – a nonempty set of objects,
A – a nonempty set of attributes.

Each attribute $a \in A$ is a function $a: U \to V_a$,
where V_a is a set of values of A.

Let $u \in U$ and $a \in A$.
$a(u)$ denotes the value of attribute a for object u.

Information System - Example

U/A	e	q	c	r	t
s1	high	good	yes	yes	no
s2	high	good	no	yes	no
s3	medium	good	yes	yes	yes
s4	low	avg	yes	yes	yes
s5	low	good	yes	yes	yes
s6	high	avg	no	no	yes

We have six stores (objects): *s*1, *s*2, *s*3, *s*4, *s*5, *s*6

Each store is characterized by five attributes:

- *e* – empowerment of sales personnel
- *q* – perceived quality of merchandise
- *c* – segmented customer-base
- *r* – good refund policy
- *t* – high traffic location

Information System – Example (*cont.*)

U/A	e	q	c	r	t
s1	high	good	yes	yes	no
s2	high	good	no	yes	no
s3	medium	good	yes	yes	yes
s4	low	avg	yes	yes	yes
s5	low	good	yes	yes	yes
s6	high	avg	no	no	yes

$$e(s1)=high$$
$$e(s2)=high$$
$$e(s3)=medium$$
$$e(s4)=low$$
$$e(s5)=low$$
$$e(s6)=high$$

$$U=\{s1,s2,s3,s4,s5,s6\}$$
$$A=\{e,q,c,r,t\}$$
$$V_e=\{low,medium,high\}$$
$$V_q=\{avg,good\}$$
$$V_c=V_r=V_t=\{no,yes\}$$

Decision Systems

- Sometimes we distinguish in an information system two classes of attributes, called ***condition*** and ***decision*** attributes.

- Condition attributes describe the objects in terms of available information.
- Decision attributes (in many cases, one decision attribute) partition these objects into groups (the so-called decision classes).

Decision Systems (*cont.*)

$$S=(U, A\cup D)$$

$$U-\ a\ nonempty\ set\ of\ objects,$$
$$A-\ a\ nonempty\ set\ of\ condition\ attributes,$$
$$D-\ a\ nonempty\ set\ of\ decision\ attributes.$$

$$In\ the\ particular\ case:$$
$$S=(U, A\cup\{d\})$$

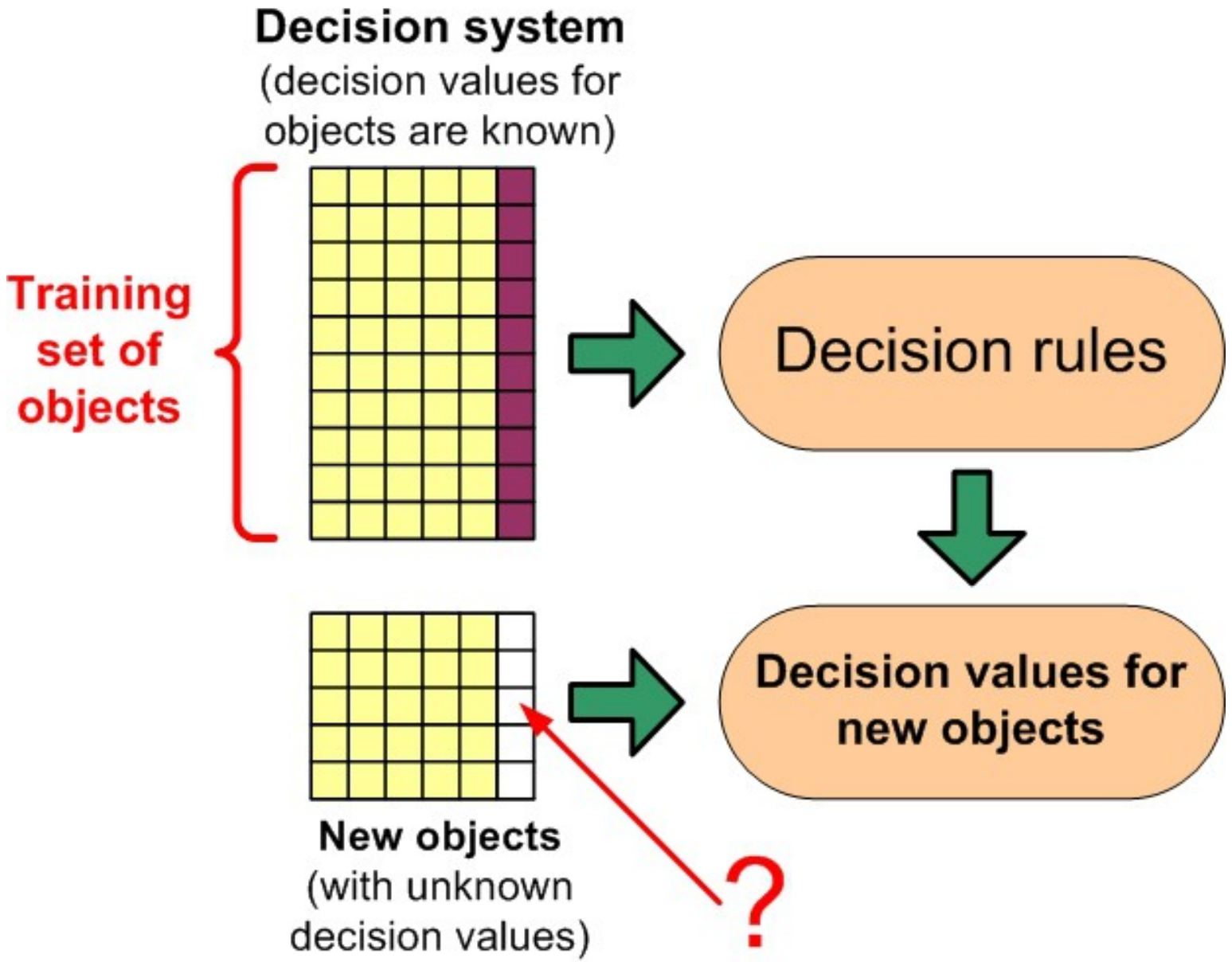
Decision System - Example

U/A	e	q	c	r	t	p
s1	high	good	yes	yes	no	yes
s2	high	good	no	yes	no	no
s3	medium	good	yes	yes	yes	yes
s4	low	avg	yes	yes	yes	yes
s5	low	good	yes	yes	yes	yes
s6	high	avg	no	no	yes	no

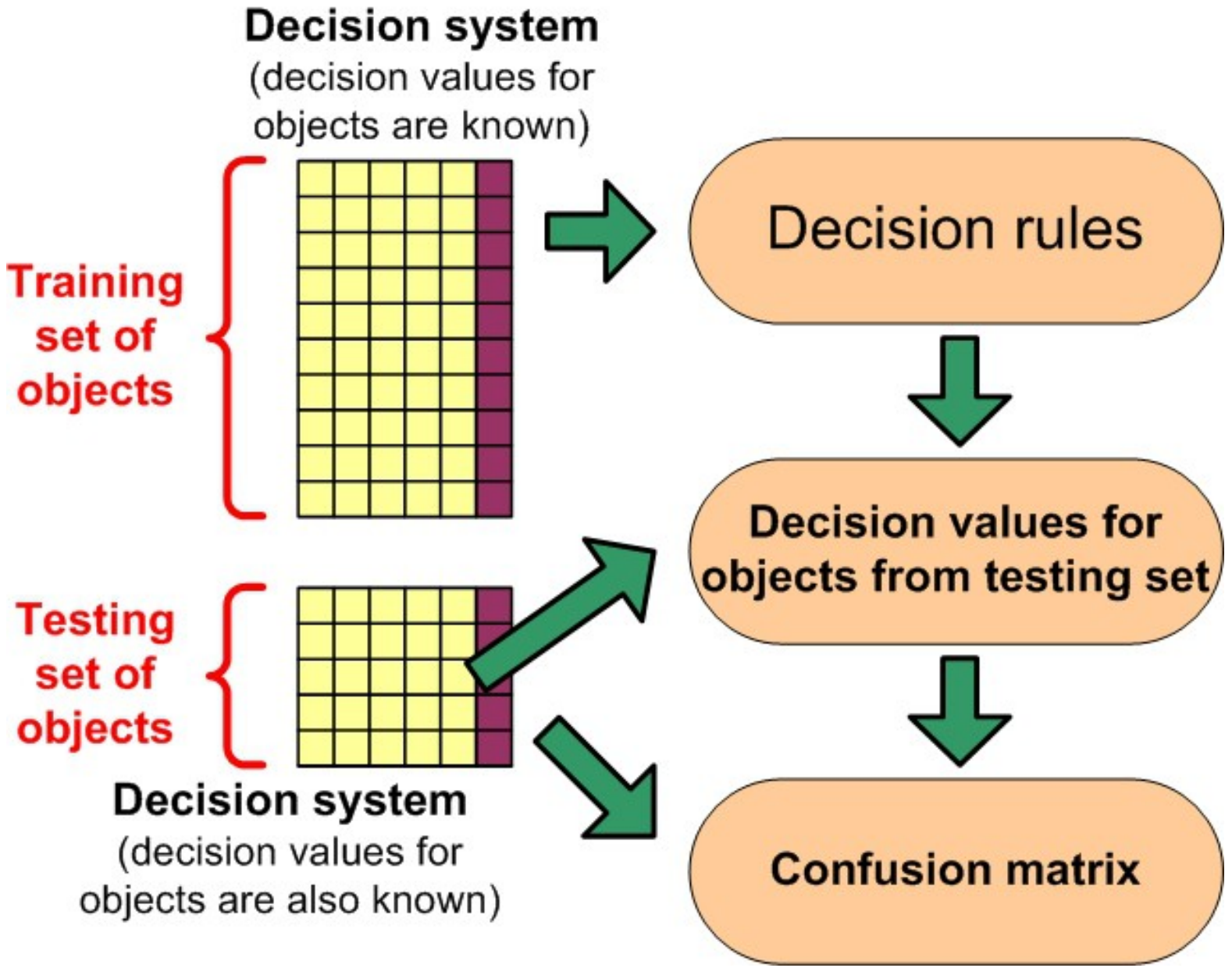
A decision attribute: *p* – store profit

$$S=(U, A\cup\{p\})$$
$$U=\{s1, s2, s3, s4, s5, s6\}$$
$$A=\{e, q, c, r, t\}$$
$$V_p=\{no, yes\}$$

Classification Problem



Testing Classification



Confusion Matrix