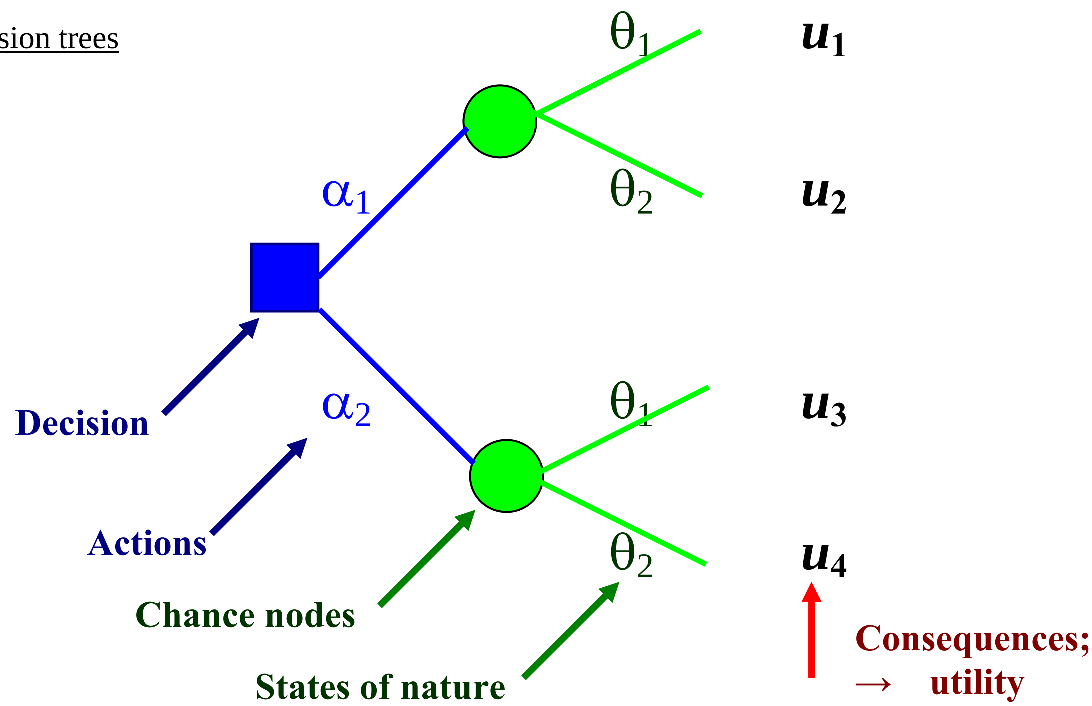
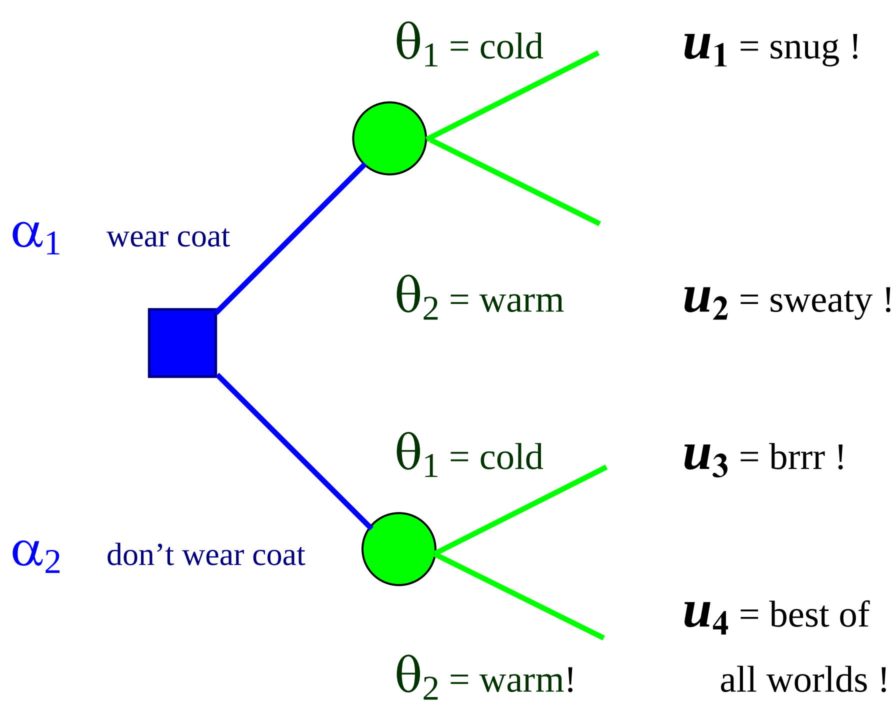
Probability

Decision trees



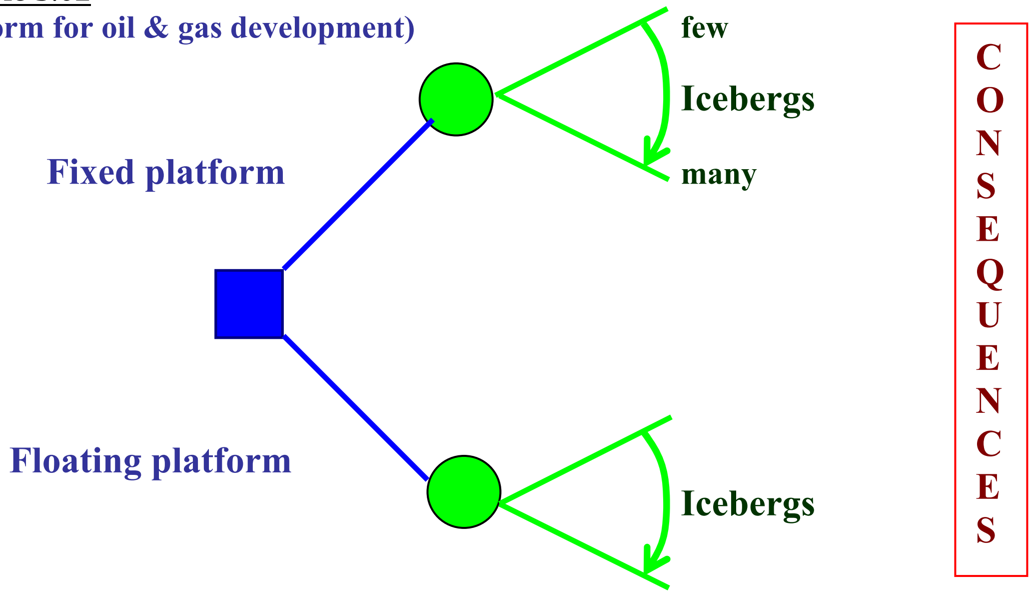
Which action will be taken is a decision completely controlled by the user.
Which state of nature will occur is beyond the user's control.

Example 3.01



The decision depends on the assessment of the probabilities of the states of nature and on the values of the utilities.

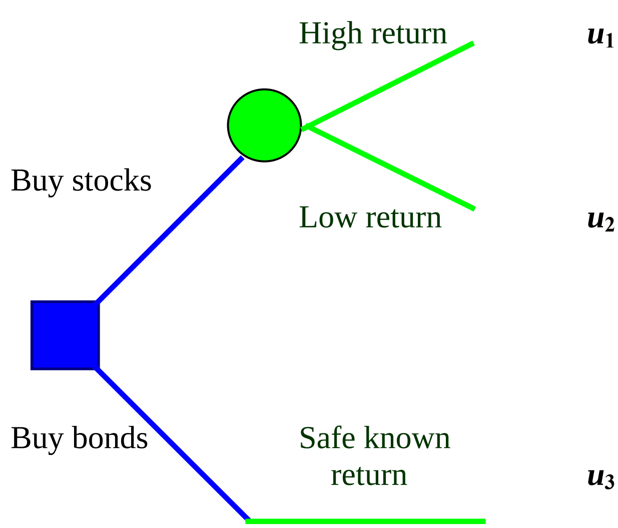
Example 3.02
(platform for oil & gas development)



This was one of the factors in the design decisions for the Hibernia and Terra Nova oil fields.

Example 3.03

Investment



The decision here depends not only on the average return but also on one's tolerance for risk. Risk aversion is a real life factor that we will not have the time to explore in this course.

Fair bet

Example 3.04

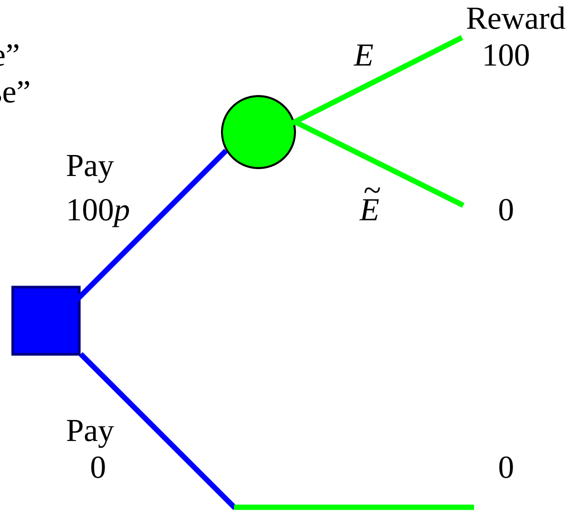
A client gives a \$100 reward iff (if and only if) the contractor's circuit board passes a reliability test. The contractor must pay a non-refundable deposit of \$100p with the bid. What is a fair price for the deposit?

Let E = (the event that the circuit board passes the test)
and $\tilde{E} = \sim E$ = not-E = (the event that the circuit board fails the test)
then $\tilde{E}$ is known as the **complementary event** to E.

Let $E = 1$ represent "E is true"
and $E = 0$ represent "E is false"
then

$$\boxed{E + \tilde{E} = 1}$$

Decision tree:



$$p = \frac{\text{deposit}}{\text{reward}}$$

The contractor has a free choice:
to take the contract or not.
If the contract is taken (= upper branches of decision tree):

(Gain if $\tilde{E}$) = $\mathbf{-100}\boldsymbol{p}$

(Gain if E) = $\mathbf{-100}\boldsymbol{p} + \mathbf{100} = \mathbf{100(1-}\boldsymbol{p}\mathbf{)} = \mathbf{100}\,\boldsymbol{\tilde{p}}$

Therefore Gain = $\mathbf{-100}\boldsymbol{p} + \mathbf{100}\boldsymbol{E}$

where E is random, (= 0 or 1; E is a **Bernoulli random quantity**).

If the contract is not taken (= lowest branch of decision tree):

Gain $\equiv 0$

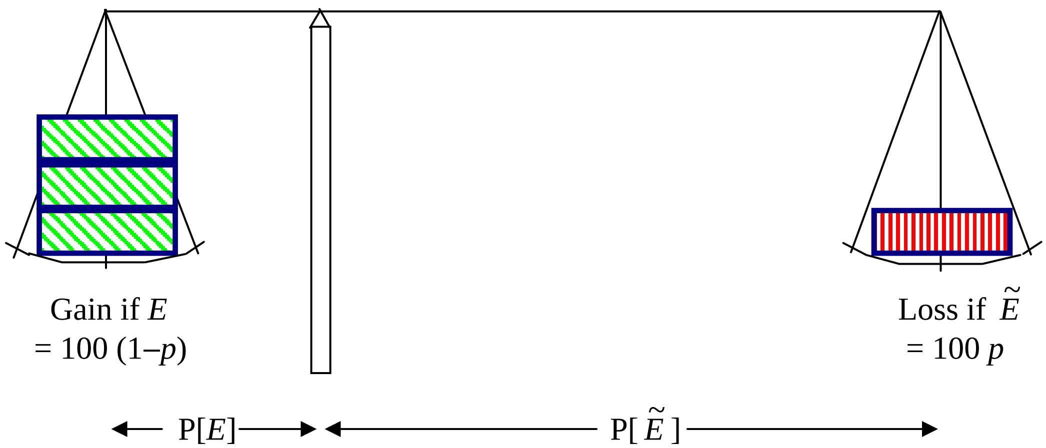
A fair bet $\Rightarrow$ **indifference** between decisions

$\Rightarrow \quad \mathbf{-100}\boldsymbol{p} + \mathbf{100}\boldsymbol{E} \approx \mathbf{0}$

Gain > 0 : too generous to contractor
Gain < 0 : contract too expensive

But gain is a random quantity

<u>Balance of Judgement</u>:
[Simple mechanics can provide analogies for much of probability theory]



The bet is fair iff gain and loss balance:

Taking moments: $100\,(1-p) \times \text{P}[E] = 100\,p \times \text{P}[\tilde{E}]$
But $\tilde{E} = 1 - E$ and $\text{P}[\tilde{E}] = 1 - \text{P}[E]$
$\Rightarrow (1 - p + p) \times \text{P}[E] = p$

Therefore the fair price for the bid deposit occurs when $\boxed{p = \text{P}[E]}$
and the fair price is (deposit) = (contract reward) × P[E].

<u>Example 3.04</u> (continued)

Suppose that past experience suggests that E occurs 24% of the time.
Then we estimate that P[E] = .24 and the fair bid is $100 \times .24 =$ **<u>$24</u>.**

<u>Odds</u>

Let s be the reward at stake in the contract (= $100 in example 3.04).
The odds on E occurring are the ratio r, where

$$r = \frac{\text{loss if } \tilde{E}}{\text{gain if } E} = \frac{s\,p}{s\,(1-p)} = \frac{P[E]}{P[\tilde{E}]} = \frac{p}{\tilde{p}} = \frac{p}{1-p} \quad \Rightarrow P[E] = \frac{r}{r+1}$$

In example 3.04,

$$r = \frac{p}{1-p} = \frac{.24}{.76} = \frac{6}{19} = \text{“19 to 6 against” (or “6 to 19 on”, but usually larger \# first)}$$

and $r = \dfrac{6}{19} \quad \Rightarrow \quad p = \dfrac{r}{r+1} = \dfrac{\frac{6}{19}}{\frac{6}{19}+1} = \dfrac{6}{6+19} = \dfrac{6}{25} = .24$

"Even odds" $\Rightarrow$ $\boldsymbol{r = 1{:}1}$ $\Rightarrow$ $\boldsymbol{p = \tfrac{1}{2} = 50\%}$
"Odds on" when $\boldsymbol{p > .5}$, "Odds against" when $\boldsymbol{p < .5}$.

<u>Incoherence</u>:

Suppose that no more than one of the events $\{E_1, E_2, \ldots, E_n\}$ can occur. Then the events are **incompatible** (= **mutually exclusive**).

If the events $\{E_1, E_2, \ldots, E_n\}$ are such that they exhaust all possibilities, (so that at least one of them must occur), then the events are **exhaustive**.

If the events $\{E_1, E_2, \ldots, E_n\}$ are both incompatible and exhaustive, (so that *exactly one* of them must occur), then they form a **partition**, and

$$E_1 + E_2 + \ldots + E_n = \mathbf{1}$$

<u>Example 3.05</u>

A bookmaker accepts bets on a partition $\{ E_i \}$.
You place a non-refundable deposit (a bet) $p_i s_i$ on E_i occurring.
The bookie pays a stake s_i to you (but still retains your deposit) if E_i occurs.
If E_i does not occur, then you lose your deposit $p_i s_i$.
Assume that one bet is placed on each one of the events $\{ E_i \}$.
The bookie's gain if E_h is true is

(gain) = (revenue) – (payout)

$$g_h = \left(\sum_{i=1}^{n} p_i s_i \right) - s_h \qquad \text{(for } h = 1, 2, \ldots, n\text{)}$$

The bookie can arrange $\{ p_i \}$ such that $\sum g_h > 0$ $\rightarrow$ unfair bet! (= incoherence)

Suppose $s_i = s$ $\forall i$ **(for all values of *i*)**, then

$$g_h = s\left[\left(\sum_{i=1}^{n} p_i \right) - 1 \right] \qquad \text{(for } h = 1, 2, \ldots, n\text{)}$$

A fair bet is then assured if

$$\boxed{\left(\sum_{i=1}^{n} p_i \right) = 1}$$

(**total probability theorem**);

the probabilities are then **coherent**.

A More General Case of Example 3.05:

An operator accepts deposits on a partition $\{ E_i \}$.
k_i people each place a non-refundable deposit $p_i s_i$ on E_i occurring.
Note that p_i is a measure of the likelihood of E_i occurring.
(The more likely E_i is, the greater the deposit that the operator will require).
If E_i occurs, then the operator pays a stake s_i to each of the k_i contractors, (but still retains all of the deposits).
If E_i does not occur, then each of the k_i contractors loses the deposit $p_i s_i$.

The operator's gain if E_h is true is

(gain) = (revenue) − (payout)

$$g_h = \left(\sum_{i=1}^{n} k_i p_i s_i \right) - k_h s_h \qquad \text{(for } h = 1, 2, \ldots , n\text{)}$$

Now assume a more common situation, not of equal stakes, but of equal deposits:

$$p_1 s_1 = p_2 s_2 = \ldots = p_n s_n = b$$

Then

$$g_h = \left(b \sum_{i=1}^{n} k_i \right) - k_h \left(\frac{b}{p_h} \right) \qquad \text{(for } h = 1, 2, \ldots , n\text{)}$$

The number p_h is a measure of how likely the gain g_h is to occur.
Therefore use p_h as a weighting factor, to arrive at an expected gain:

$$\mathrm{E}[G] = \sum_{h=1}^{n} p_h g_h = b \sum_{h=1}^{n} p_h \left(\left(\sum_{i=1}^{n} k_i \right) - k_h \left(\frac{1}{p_h} \right) \right)$$

$$= b \left(\left(\sum_{i=1}^{n} k_i \right) \left(\sum_{h=1}^{n} p_h \right) - \left(\sum_{h=1}^{n} k_h \right) \right) = b \left(\sum_{i=1}^{n} k_i \right) \left(\left(\sum_{i=1}^{n} p_i \right) - 1 \right)$$

$$\mathrm{E}[G] = 0 \quad \text{if and only if} \quad \boxed{\sum_{i=1}^{n} p_i = 1}$$

Notation:

$A \wedge B$ = events A and B both occur; $A \wedge B = A \times B = A\ B$
$A \vee B$ = event A or B (or both) occurs;
(but $A \vee B \neq A + B$ unless A, B are incompatible)

Some definitions:

Experiment = process leading to a single outcome

Sample point (= simple event) = one possible outcome (which precludes all other outcomes)

Event E = set of related sample points

Possibility Space = universal set = Sample Space S =

{ all possible outcomes of an experiment }

By the definition of S, any event E is a subset of S: $E \subseteq S$

Classical definition of probability (when sample points are equally likely):

$$\mathrm{P}[E] = \frac{n(E)}{n(S)} \quad ,$$

where $n(E)$ = the number of [equally likely] sample points inside the event E.

More generally, the probability of an event E can be calculated as the sum of the probabilities of all of the sample points included in that event:

$$\boxed{\mathrm{P}[\boldsymbol{E}] = \Sigma\, \mathrm{P}[X]}$$

(summed over all sample points X in E.)

Empirical definition of probability:

P[E] = **(limit as # exp'ts $\to \infty$ of) { relative frequency of E }**

Example 3.06 (illustrating the evolution of relative frequency with an ever increasing number of trials):

`http://www.engr.mun.ca/~ggeorge/3423/demos/cointoss.exe`

or import the following macro into a MINITAB session:

`http://www.engr.mun.ca/~ggeorge/3423/demos/Coins.mac`

Example 3.07: rolling a standard fair die. The sample space is

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

$n(S) = 6$ (the sample points are equally likely)

$$P[1] = 1/6 = P[2] = P[3] = \ldots$$

$P[S] =$ **1**

(S is absolutely certain)

The **empty set** (= null set) = **Ø** = {} [Note: this is *not* { 0 } !]

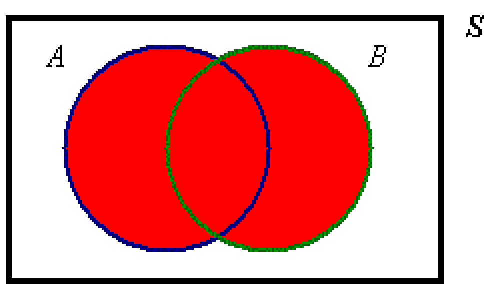
P[**Ø**] = **0** **(Ø is absolutely impossible)**

The **complement** of a set A is A' (or $\tilde{A}$, A^*, A^c, NOT A, ~A, $\overline{A}$).

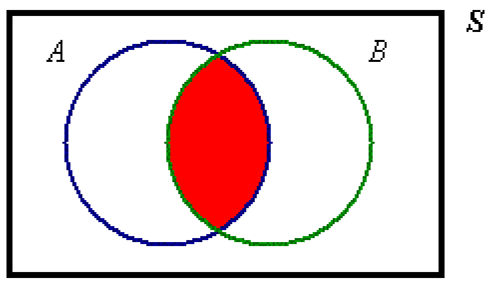
$n(\sim A) = n(S) - n(A)$ and

$P[\sim A] = 1 - P[A]$

The **union** $A \cup B = (A \text{ OR } B) = A \vee B$



The **intersection** $A \cap B = (A \text{ AND } B) = A \wedge B = A \times B = A\,B$

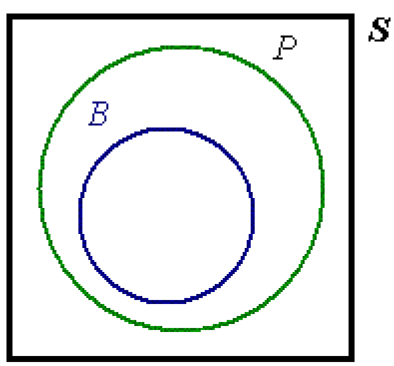


For *any* set or event **E** :

$\varnothing \cup \mathbf{E} =$	$\mathbf{E}$	$\mathbf{E} \cap \sim\mathbf{E} =$	$\varnothing$
$\varnothing \cap \mathbf{E} =$	$\varnothing$	$\mathbf{E} \cup \sim\mathbf{E} =$	S
$S \cup \mathbf{E} =$	S	$\sim(\sim\mathbf{E}) =$	$\mathbf{E}$
$S \cap \mathbf{E} =$	$\mathbf{E}$	$\sim\varnothing =$	S

The set B is a **subset** of the set P : $B \subseteq P$.

Read the symbol "$\subseteq$" as "is contained entirely inside"



If it is also true that $P \subseteq B$, then $P = B$ (the two sets are identical).

If $B \subseteq P$, $B \neq P$ and $B \neq \varnothing$, then $B \subset P$ (B is a **proper subset** of the set P).

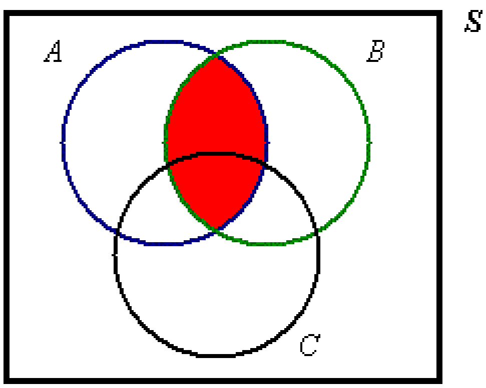
$B \cap P = B$ For *any* set or event **E** : $\varnothing \subseteq \mathbf{E} \subseteq S$

$B \cup P = P$ Also: $B \cap \sim P = \varnothing$

Example 3.08

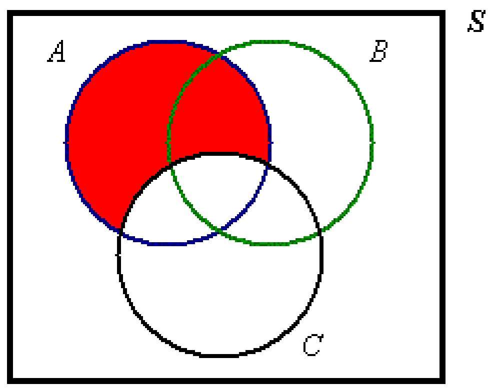
Examples of Venn diagrams:

1. Events A and B both occur.



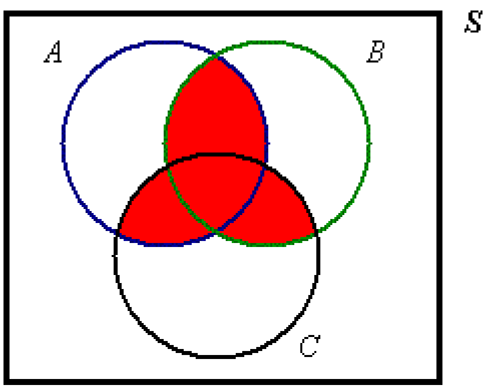
$\boldsymbol{A \cap B}$

2. Event A occurs but event C does not.



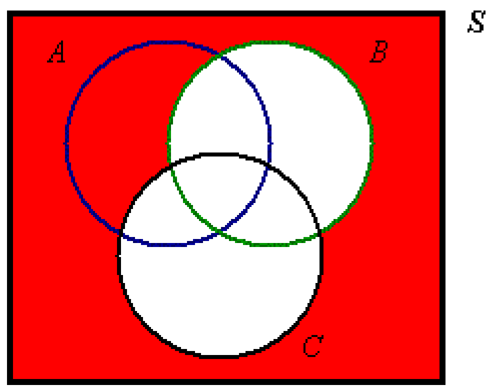
$\boldsymbol{A\,C\,'}$

3. At least two of events A, B and C occur.



$(AB) \vee (BC) \vee (CA)$

4. Neither B nor C occur.



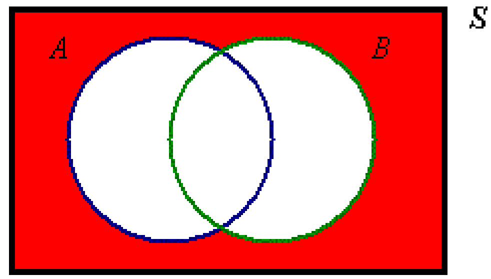
$\sim B \wedge \sim C = \sim(B \vee C)$

Example 3.08.4 above is an example of **DeMorgan's Laws**:

$\sim(A \cup B) =$

$$\tilde{A} \cap \tilde{B}$$

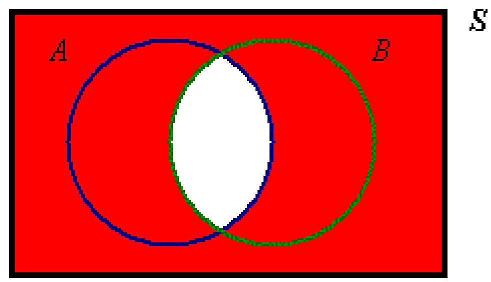
"neither *A* nor *B*"



$\sim(A \cap B) =$

$$\tilde{A} \cup \tilde{B}$$

"not both *A* and *B*"

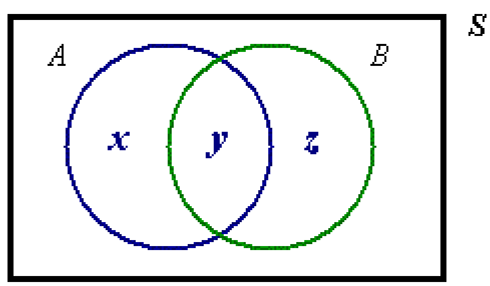


General Addition Law of Probability

$\mathbf{P}[A \vee B] = x + y + z$

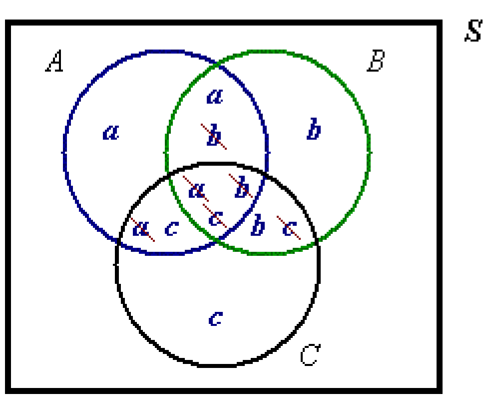
$\mathbf{P}[A] = x + y$
$\mathbf{P}[B] = y + z$

$\mathbf{P}[A \wedge B] = y$



$$\mathrm{P}[A \vee B] = \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \wedge B]$$

Extended to three events, this law becomes



$$\mathrm{P}[A \vee B \vee C] = \mathrm{P}[A] + \mathrm{P}[B] + \mathrm{P}[C] - \mathrm{P}[A \wedge B] - \mathrm{P}[B \wedge C] - \mathrm{P}[C \wedge A] + \mathrm{P}[A \wedge B \wedge C]$$

If two events A and B are **mutually exclusive**
(= **incompatible** = have no common sample points), then
$A \cap B = \varnothing \;\Rightarrow\; \mathrm{P}[A \wedge B] = 0$ and the addition law simplifies to

$$\mathrm{P}[A \vee B] = \mathrm{P}[A] + \mathrm{P}[B] .$$

Only when A and B are mutually exclusive may one say " $A \vee B$ " = " $A + B$ ".

Total Probability Law

The total probability of an event A can be partitioned into two mutually exclusive subsets: the part of A that is inside another event B and the part that is outside B:

$$\boxed{\mathrm{P}[A] = \mathrm{P}[A \wedge B] + \mathrm{P}[A \wedge \sim B]}$$

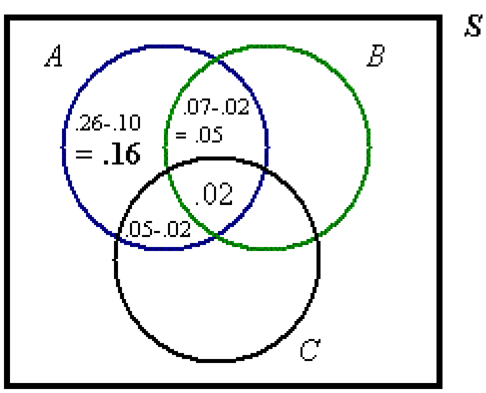
Special case, when $A = S$ and $B = E$:

$$\mathrm{P}[S] = \mathrm{P}[S \wedge E] + \mathrm{P}[S \wedge \sim E]$$
$$\Rightarrow \quad 1 = \mathrm{P}[E] + \mathrm{P}[\sim E]$$

Example 3.09

Given the information that $\mathrm{P}[ABC] = 2\%$, $\mathrm{P}[AB] = 7\%$, $\mathrm{P}[AC] = 5\%$ and $\mathrm{P}[A] = 26\%$, find the probability that, (of events A,B,C), *only* event A occurs.

A only $=$ ***AB'C'***



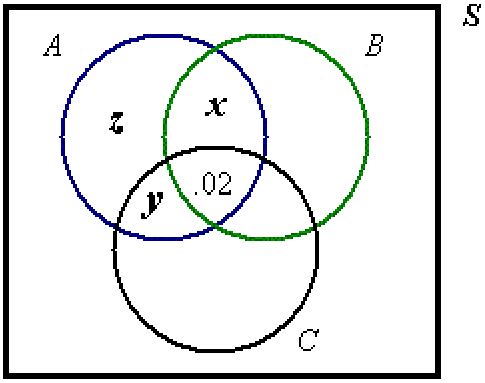
We know the intersection probabilities, therefore we start at the centre and work our way out.

The sum of the probabilities in the three lens regions illustrated is .05 + .02 + .03 = **.10** .

The required probability is in the remaining one region of A and is **.26 – .10 = .16** .

On the next page is a more systematic way to solve this example.

[Example 3.09 continued]



P[AB] = .07 = x + .02 ⇒ x = .05

P[AC] = .05 = y + .02 ⇒ y = .03

P[A] = .26 = z + .03 + .02 + .05

⇒ P[$AB'C'$] = z = .26 − .10 = <u>.16</u>

Alternatively, $P[AB'C'] = P[A] - P[AB] - P[AC] + P[ABC]$
$= .26 - .07 - .05 + .02 = \underline{.16}$

If the information had been provided in the form of unions instead of intersections, then we would have started at the outside of the Venn diagram and worked our way in, using deMorgan's laws and the general addition law where necessary.

[End of Section 3]

[Space for additional notes]