Advanced Artificial Intelligence Revision

Inference Using Full Joint Distribution

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576
A full joint distribution for *toothache*, *cavity*, *catch* world.				

Probabilistic inference is a computation from observed evidence of **posterior probabilities** for **query propositions**. One task involved for accomplishing this is to extract the distribution over a subset of variables or a single variable, hence one way to do this is through **marginal probability**.

Marginalisation/Summing Out - posterior probabilities are summed out, therefore providing the following general **marginalisation rule**:

$$\boldsymbol{P}(\boldsymbol{Y}) = \sum_{z} \boldsymbol{P}(\boldsymbol{Y}, \boldsymbol{z})$$

Otherwise can be described as, a distribution over **Y** can be calculated by summing out the posterior probabilities from any joint distribution which contains **Y**.

This rule can also be applied through **conditional probabilities** instead of **joint probabilities**.

$$\boldsymbol{P}(\boldsymbol{Y}) = \sum_{z} \boldsymbol{P}(\boldsymbol{Y}|\boldsymbol{z}), P(\boldsymbol{z})$$

This rule is known as **conditioning**. **Marginalisation** and **conditioning** are useful for all kinds of derivations involving probability expressions. In most cases, computation of **conditional probabilities** of some variables given others is preferred.

Conditional probabilities can be found by first calculating the **product rule** to obtain an expression in terms of **unconditional probabilities** and then evaluating the expression from the **full joint distribution**.

For example, the probability of a cavity given evidence of a toothache can be calculated like the following:

$$P(cavity|toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)}$$

To check, the **complement** of this rule would also be calculated to check the validity of the results. The results of which should add to 1.0. Otherwise, a **normalisation constant** may be required.

$$P(\neg cavity|toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

This can also be written with a **normalisation constant** α, to ensure that the final probabilities will sum to 1.0. The previous equation can be rewritten with a normalisation constant like the following:

$$\boldsymbol{P}(X|\boldsymbol{e}) = \alpha\boldsymbol{P}(X,\boldsymbol{e}) = \alpha\sum_{y}\boldsymbol{P}(X,\boldsymbol{e},\boldsymbol{y})$$

Where $\boldsymbol{X}$ *== the query variable; cavity,* $\boldsymbol{E}$ *== set of observed evidence variables; toothache,* $\boldsymbol{Y}$ *== remaining unobserved variables; catch.*

For example:

$\boldsymbol{P}(cavity|toothache) = \alpha\ \boldsymbol{P}(cavity|toothache)$

$= \alpha\ [\boldsymbol{P}(cavity, toothache, catch) + \boldsymbol{P}(cavity, toothache, \neg catch)]$

$= \alpha\ [< 0.108, 0.016 > + < 0.012, 0.064 >] = \alpha < 0.12, 0.08 > = < 0.6, 0.4 >$