Bayesian **Maximum A Posteriori** (MAP) maximises prior *density* multiplied by likelihood
This is not statistically invariant.
It also suffers the inconsistency and other problems of Max Likelihood.

Minimum Message Length (**MML**) is statistically invariant and has general statistical consistency properties (which Maximum Likelihood and Akaike's Information Criterion (AIC) don't have).

– MML is also far more efficient than Maximum Likelihood and AIC
– MML is always defined, whereas for some - or many - problems AIC is either undefined or poor

Conjecture (1998, ...) that only MML and very closely-related Bayesian methods are in general both statistically consistent and invariant.

Back-up Conjecture: If there are any such non-Bayesian methods, they will be far less efficient than MML.

Turing Machine
$f : States \times Symbols \rightarrow \{L, R\} \cup Symbols.$

With binary alphabet,
$f : States \times \{0, 1\} \rightarrow \{L, R\} \ \cup \ \{0, 1\}.$

Any known computer program can be represented by a Turing Machine.

Universal Turing Machines (UTMs) are like a compiler and can be made to emulate *any* Turing Machine (TM).

Recalling from information theory that an event of probability p_i can be encoded by a binary code-word of length $l_i = \log_2 p_i$, and recalling from MML that choosing H to maximise $Pr(H|D)$ is equivalent to choosing H to minimise the length of a two-part message,

$$-\log Pr(H) \quad -\log Pr(D|H),$$

```
 -------------------------------------------
|    H1   |           Data given H1        |
 -------------------------------------------

 ---------------------------------------------
|    H2        |      Data given H2          |
 ---------------------------------------------
```

we can see the relationship between MML, (probabilistic) Turing machines and 2-part Kolmogorov complexity.

In principle, can infer *any* computable function. Relevant in *all* analysis domains - including bioinformatics and non-standard models of computing.