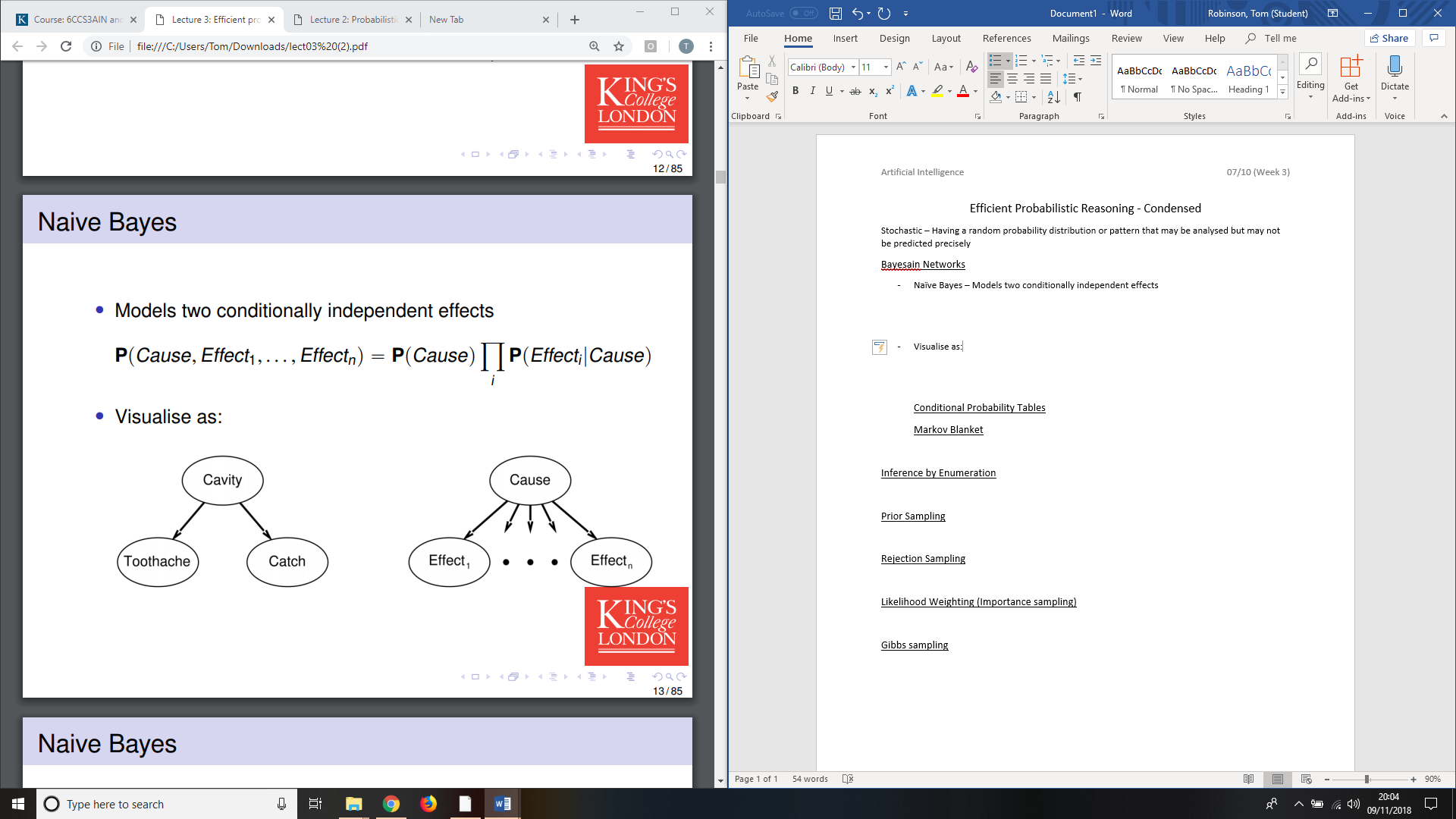
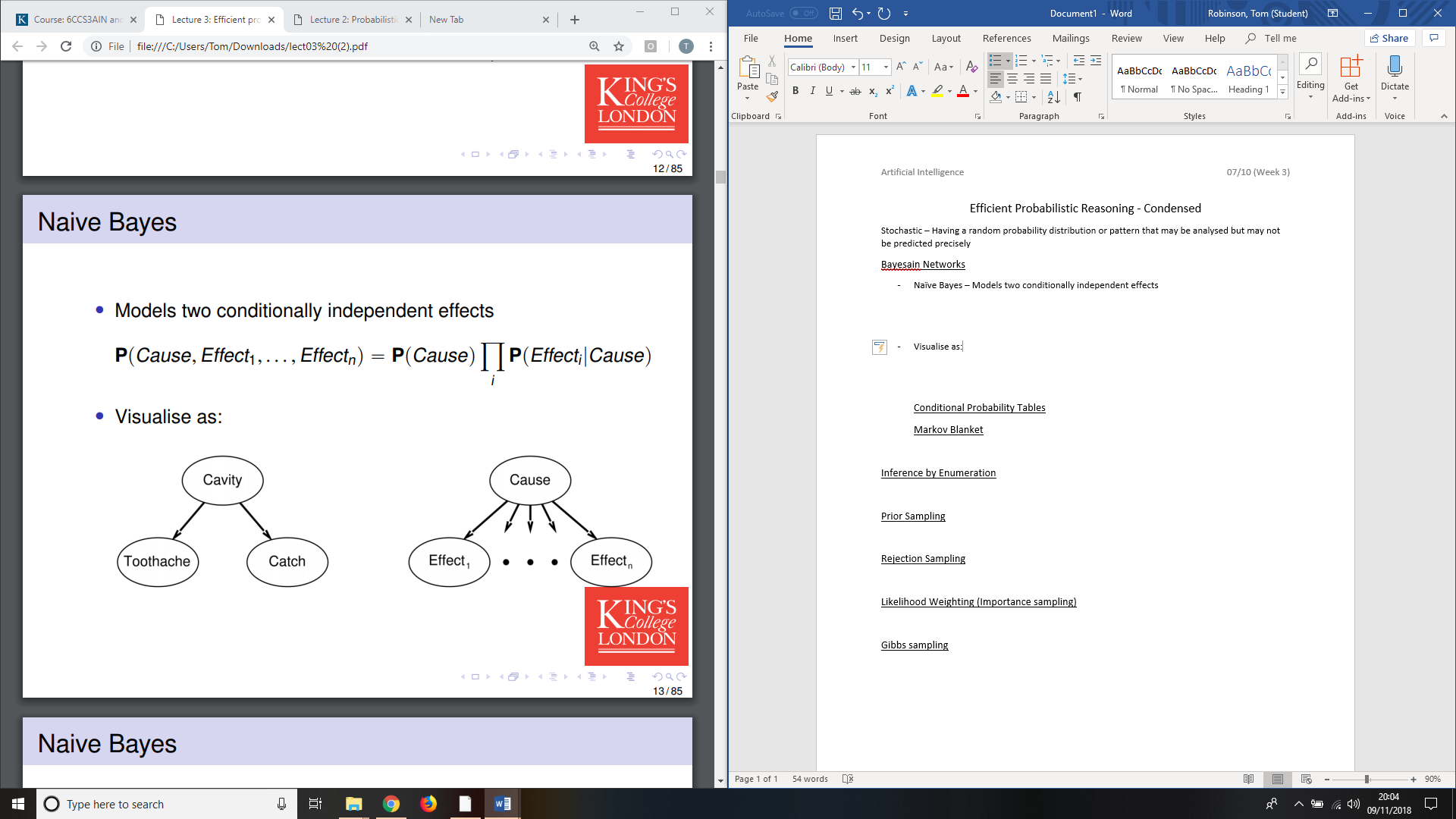
Efficient Probabilistic Reasoning - Condensed

Stochastic – Having a random probability distribution or pattern that may be analysed but may not be predicted precisely

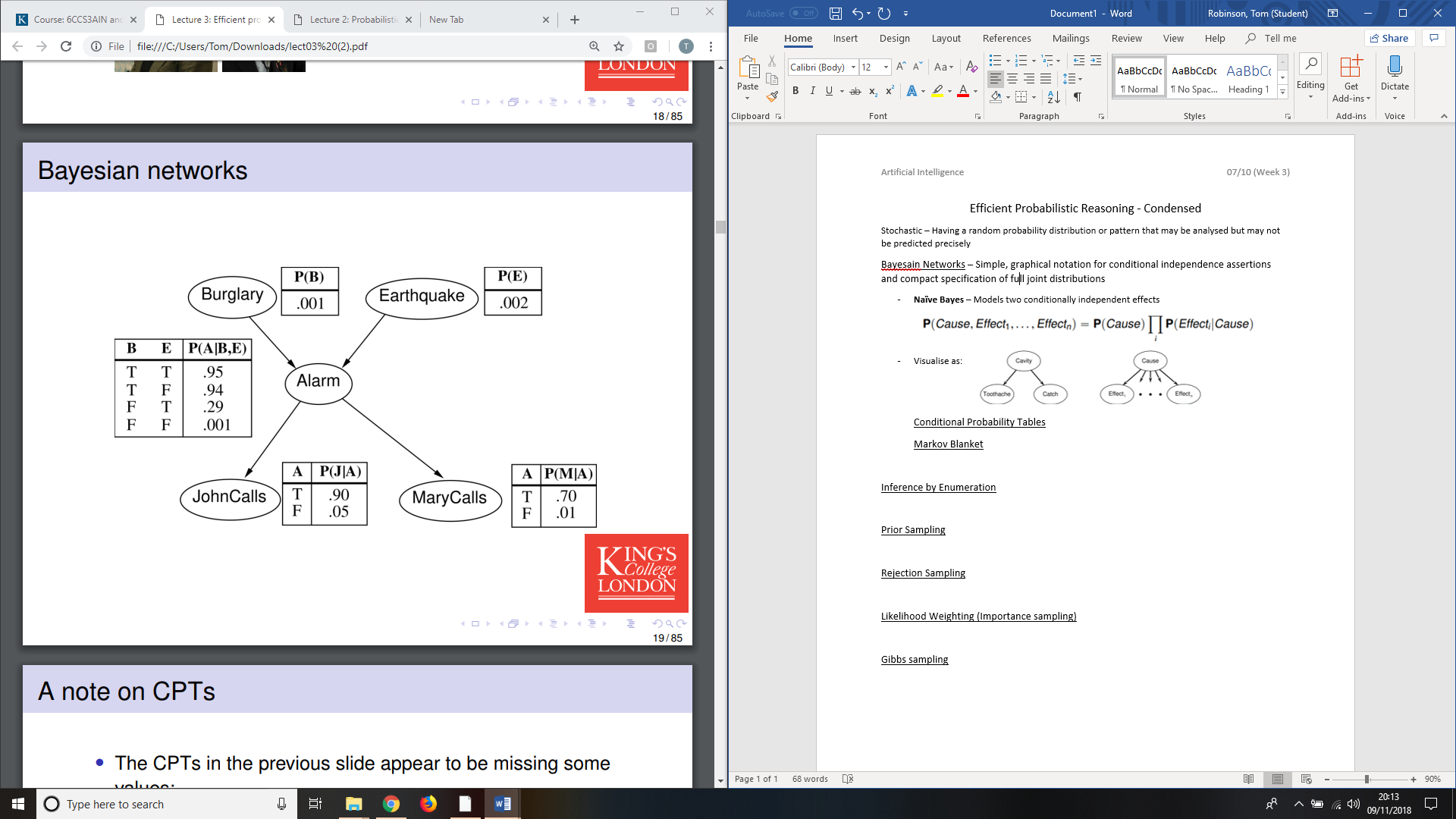
Bayesain Networks – Simple, graphical notation for conditional independence assertions and compact specification of full joint distributions

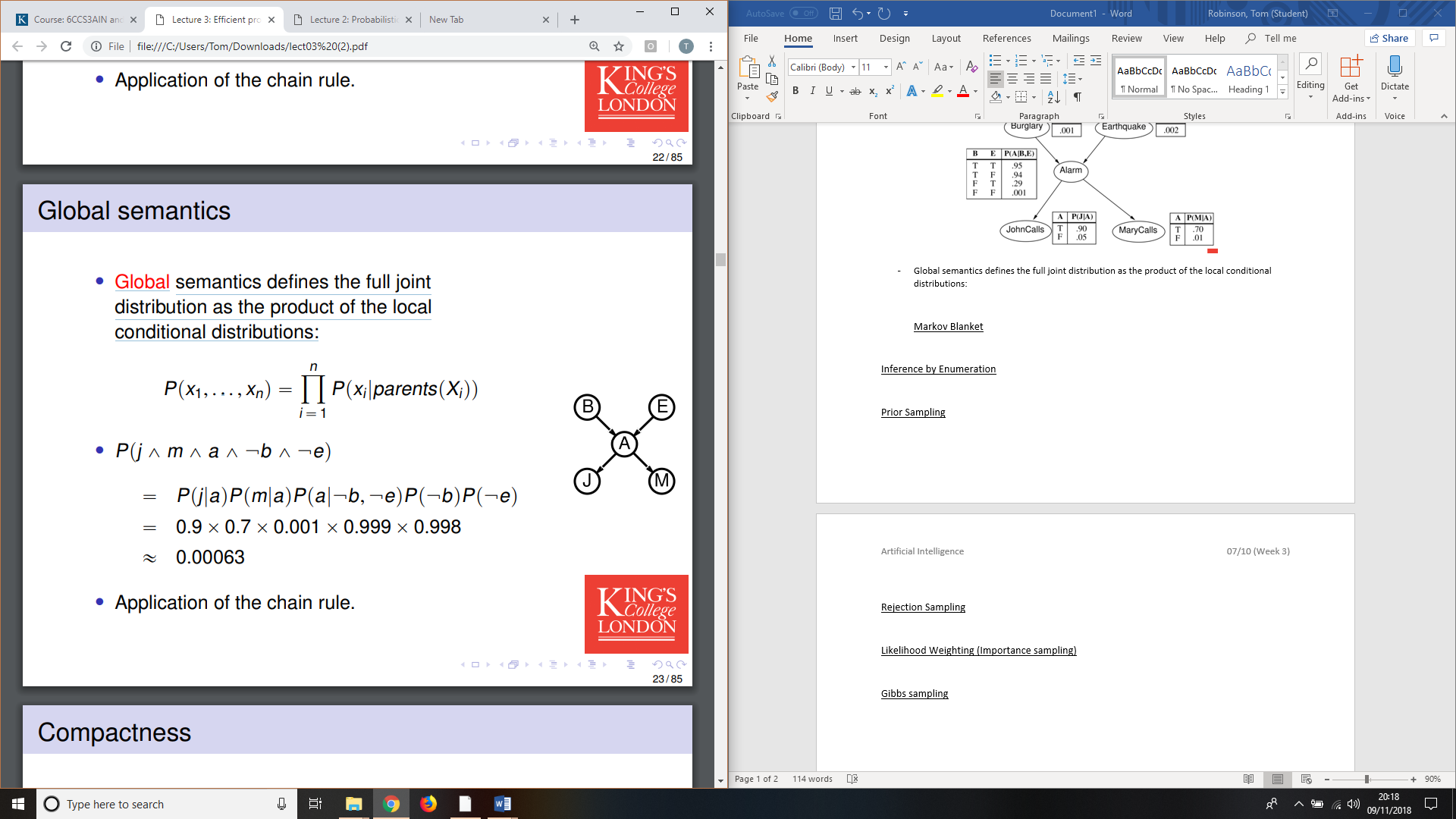
* **Naïve Bayes** – Models two conditionally independent effects



* Visualise as:

Conditional Probability Tables

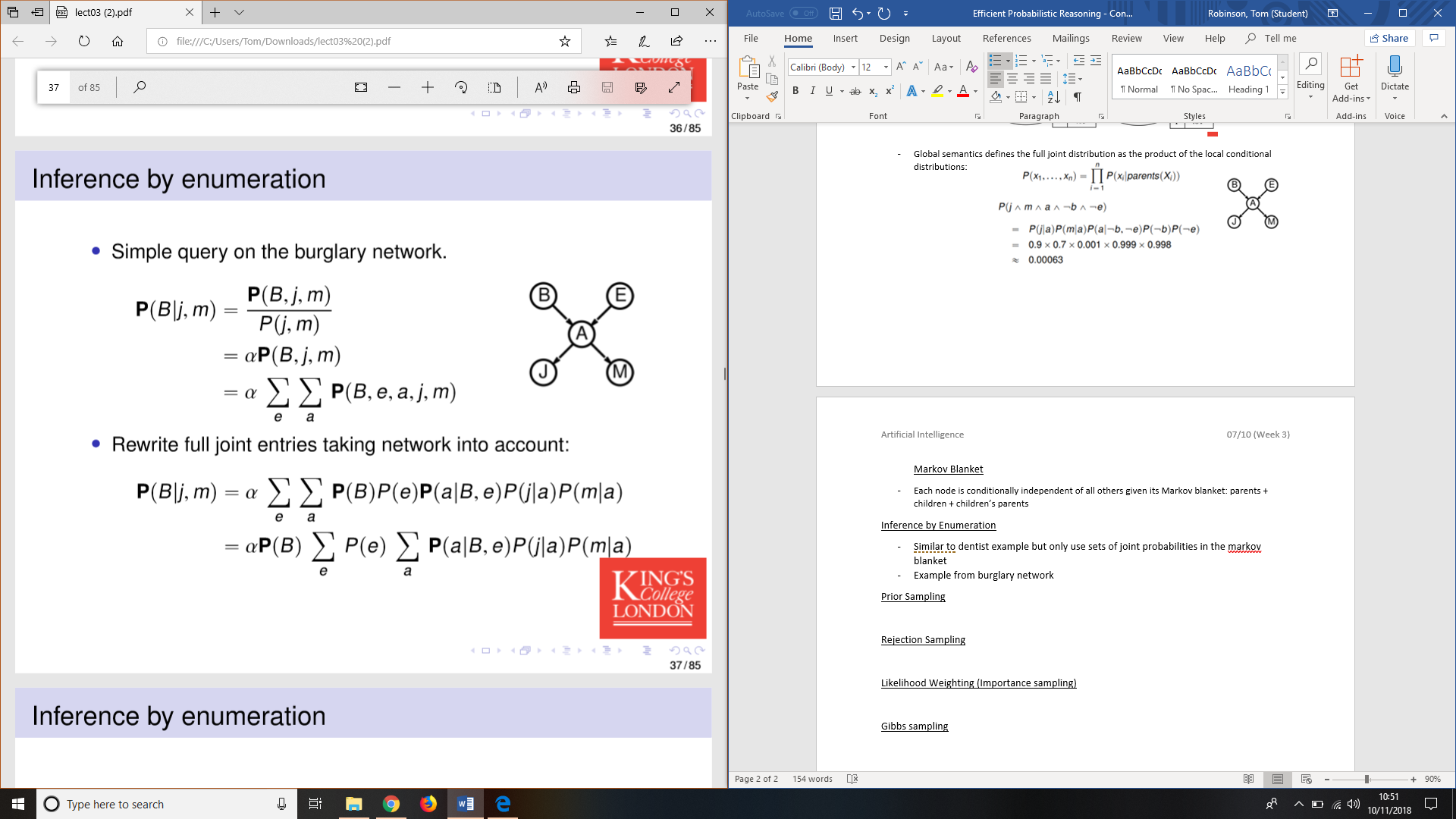
* Example: I’m at work, neighbor John calls to say my alarm is ringing, but neighbor Mary doesn’t call. Sometimes it’s set off by minor earthquakes. Is there a burglar?

* Global semantics defines the full joint distribution as the product of the local conditional distributions:

Markov Blanket

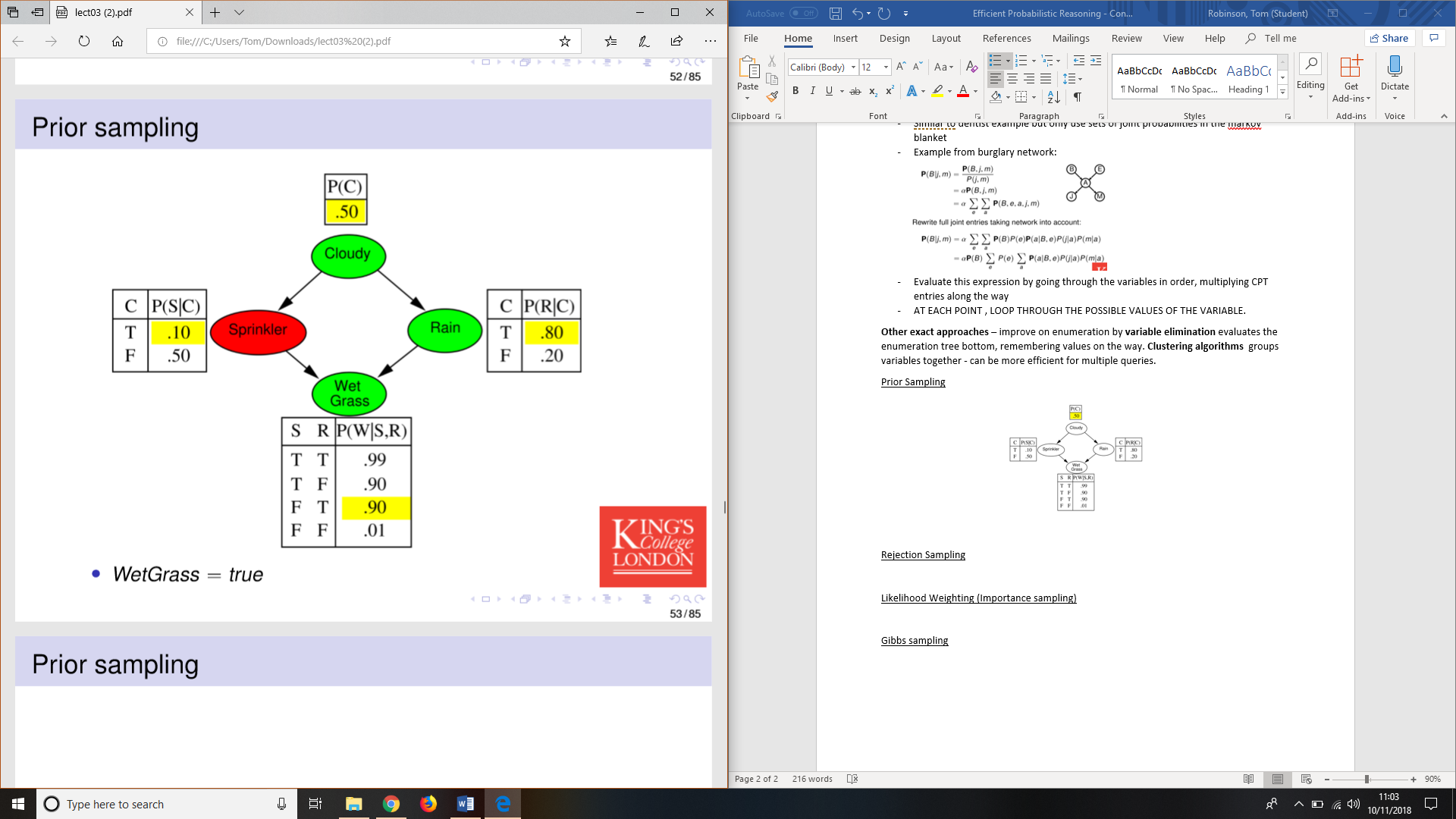
* Each node is conditionally independent of all others given its Markov blanket: parents + children + children’s parents

Inference by Enumeration

* Similar to dentist example but only use sets of joint probabilities in the markov blanket
* Example from burglary network:
* Evaluate this expression by going through the variables in order, multiplying CPT entries along the way
* AT EACH POINT , LOOP THROUGH THE POSSIBLE VALUES OF THE VARIABLE.

**Other exact approaches** – improve on enumeration by **variable elimination** evaluates the enumeration tree bottom, remembering values on the way. **Clustering algorithms** groups variables together - can be more efficient for multiple queries.

Prior Sampling – Sample the CPT randomly multiple times

* This time we get [Cloudy = true, Sprinkler = false, Rain = true, WetGrass = true]
* Write this: [true, false, true, true]
* If we repeat the process many times, we can count the number of times [true,false,true,true] is the result.
* The proportion of this to the number of total runs is P(c,¬s,r,w)
* The more runs, the more accurate the probability

To get values with evidence we need conditional probabilities: P(X|e). Could compute joint probability and sum out the conditionals but that is an inefficient method.

Rejection Sampling – Stop sampling when reach a variable you don’t want to sample

* Inefficient for unlikely outcomes

Likelihood Weighting (Importance sampling) – Fix evidence variables to true, so just sampling relevant events.

* Have to weigh them with likelihood they fit the evidence, use probabilities we know to weight the samples
* Consider we want to establish P(Rain|Cloudy = true, WetGrass = true) from
  + We want P(Rain|Cloudy = true, WetGrass = true). We pick a variable ordering: Cloudy, Sprinkler, Rain, WetGrass. as before. Set the weight w = 1 and we start.
  + Cloudy = true so w 🡨 w x P(Cloudy = true) 🡪 w =0.5
    - Cloudy = true, Sprinkler = ?, Rain = ?, WetGrass =?
  + Sprinkler is not an evidence variable, so we don’t know whether it is true or false.
    - Sample a value just as we did for prior sampling: P(Sprinkler|Cloudy = true) = <0.1,0.9>. Will probably return false. W will remain the same.
    - Cloudy = true, Sprinkler = false, Rain = ?, WetGrass =?
  + Rain is not an evidence variable, so we don’t know whether it is true or false.
    - Sample a value just as we did for prior sampling: P(Rain|Cloudy = true) = <0.8,0.2>. Will probably return true. W will remain the same.
    - Cloudy = true, Sprinkler = false, Rain = true, WetGrass =?
  + WetGrass is evidence variable with value true, so we set w 🡨 w x P(WetGrass = true | sprinker = false, Rain = true). W 🡨 0.45
    - Cloudy = true, Sprinkler = false, Rain = true, WetGrass = true
  + So we end with the event (true,false,true,true) and weight 0.45.
  + To ﬁnd a probability we tally up all the relevant events, weighted with their weights. This example: Rain= true

Gibbs sampling – Generate samples by randomly changing the previous sample

* Evidence variables are fixed, choose next state by randomly sampling non-evidence variable. This is CONDITIONAL ON MARKOV BLANKET
* EXAMPLE ON PAPER

SUMMARY

* Bayesian networks exploit conditional independence to create a more compact set of information.
* Reasonably efﬁcient computation for some problems.
* Five approaches to inference in Bayesian networks.
  + Exact: Inference by enumeration.
  + Approximate: Prior sampling
  + Approximate: Rejection sampling
  + Approximate: Importance sampling/likelihood weighting
  + Approximate: Gibbs sampling
* Can answer a simple query for any BN.