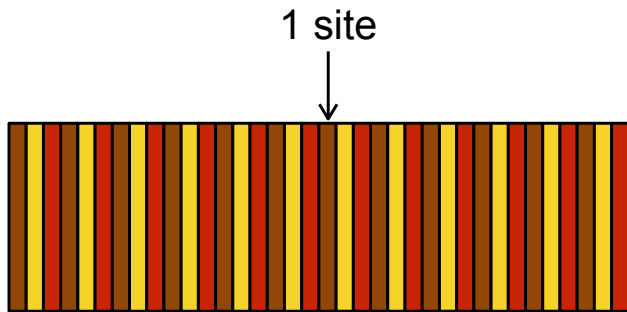


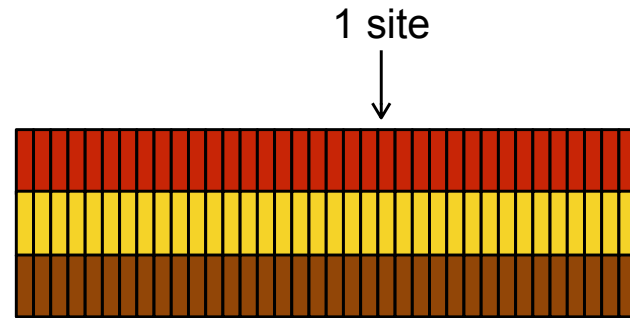
Dirichlet Process Prior

Two ways to model rate heterogeneity



site-specific rates

each site assigned to 1
of 3 rate categories



mixture model

each site has probability
 $1/3$ of being in each of
the 3 rate categories

Dirichlet process priors provide a third option...

Dirichlet Process Priors

A Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process

Nicolas Lartillot and Hervé Philippe

Lartillot and Philippe (2004)

PhyloBayes

<http://www.atgc-montpellier.fr/phylobayes>

A Dirichlet process model for detecting positive selection in protein-coding DNA sequences

John P. Huelsenbeck^{*†}, Sonia Jain[‡], Simon W. D. Frost[§], and Sergei L. Kosakovsky Pond[§]

Huelsenbeck et al. (2006)

BUCKy

<http://www.stat.wisc.edu/~ane/bucky/>

Bayesian Estimation of Concordance among Gene Trees

Cécile Ané,† Bret Larget,*† David A. Baum,† Stacey D. Smith,‡ and Antonis Rokas§*

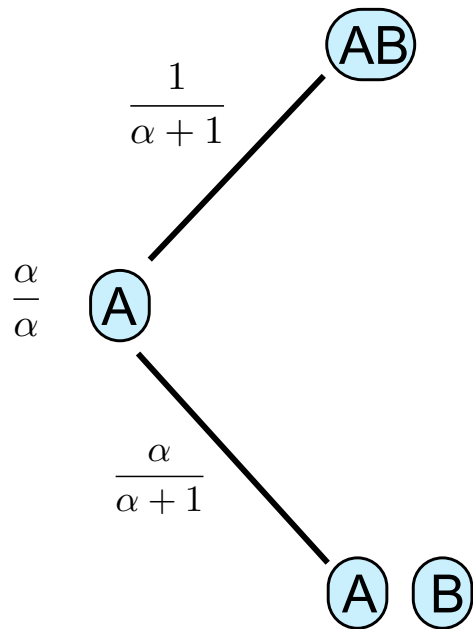
Ané et al (2007)

A Nonparametric Method for Accommodating and Testing Across-Site Rate Variation

JOHN P. HUELSENBECK,¹ AND MARC A. SUCHARD^{2,3,4}

Huelsenbeck and Suchard (2007)

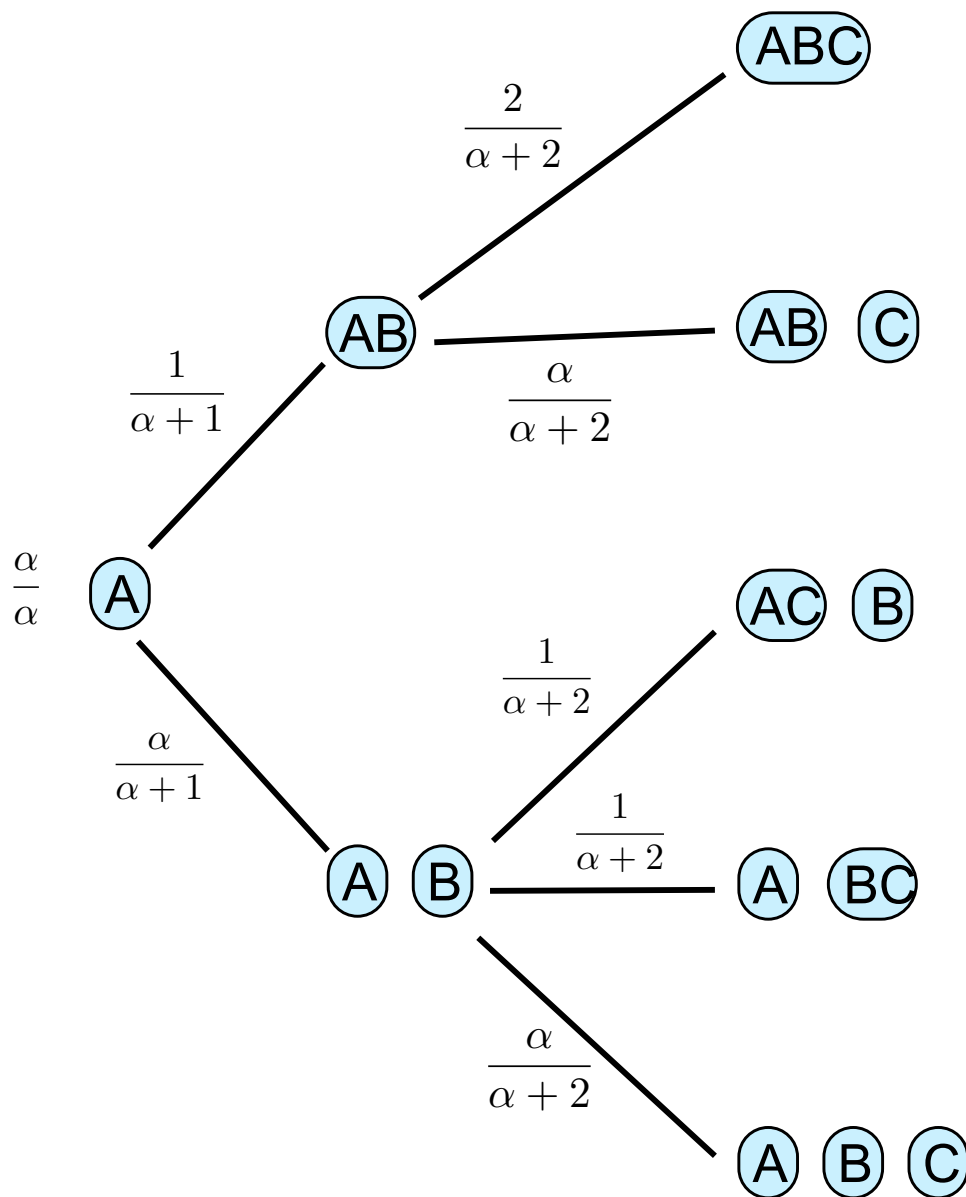
Dirichlet Process Prior



Imagine you have a collection of objects (e.g. sites) labeled A, B, C, ...

B can either be added to A's group or form its own group

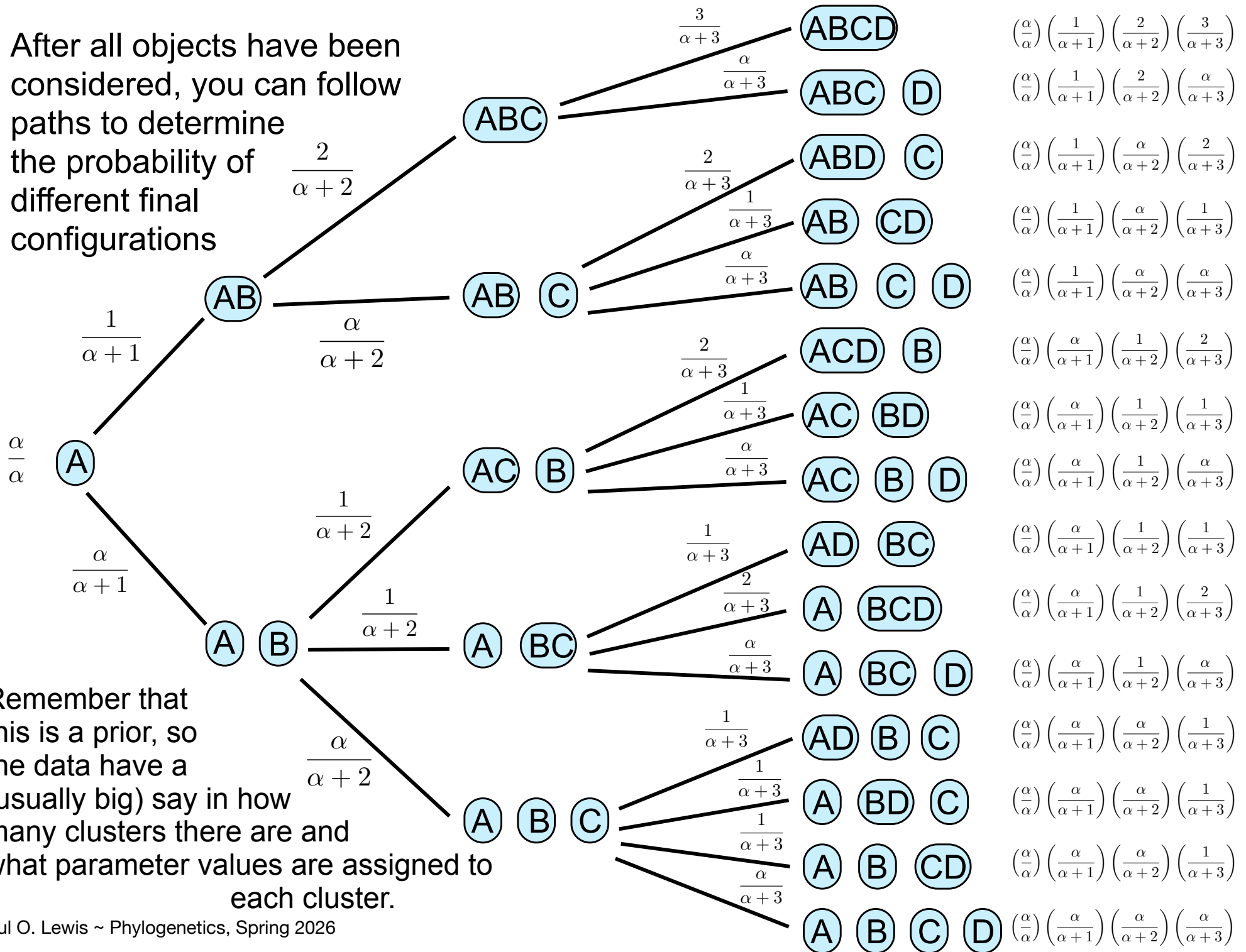
The parameter α determines the propensity for forming a new group



The third object C can either be added to an existing group...

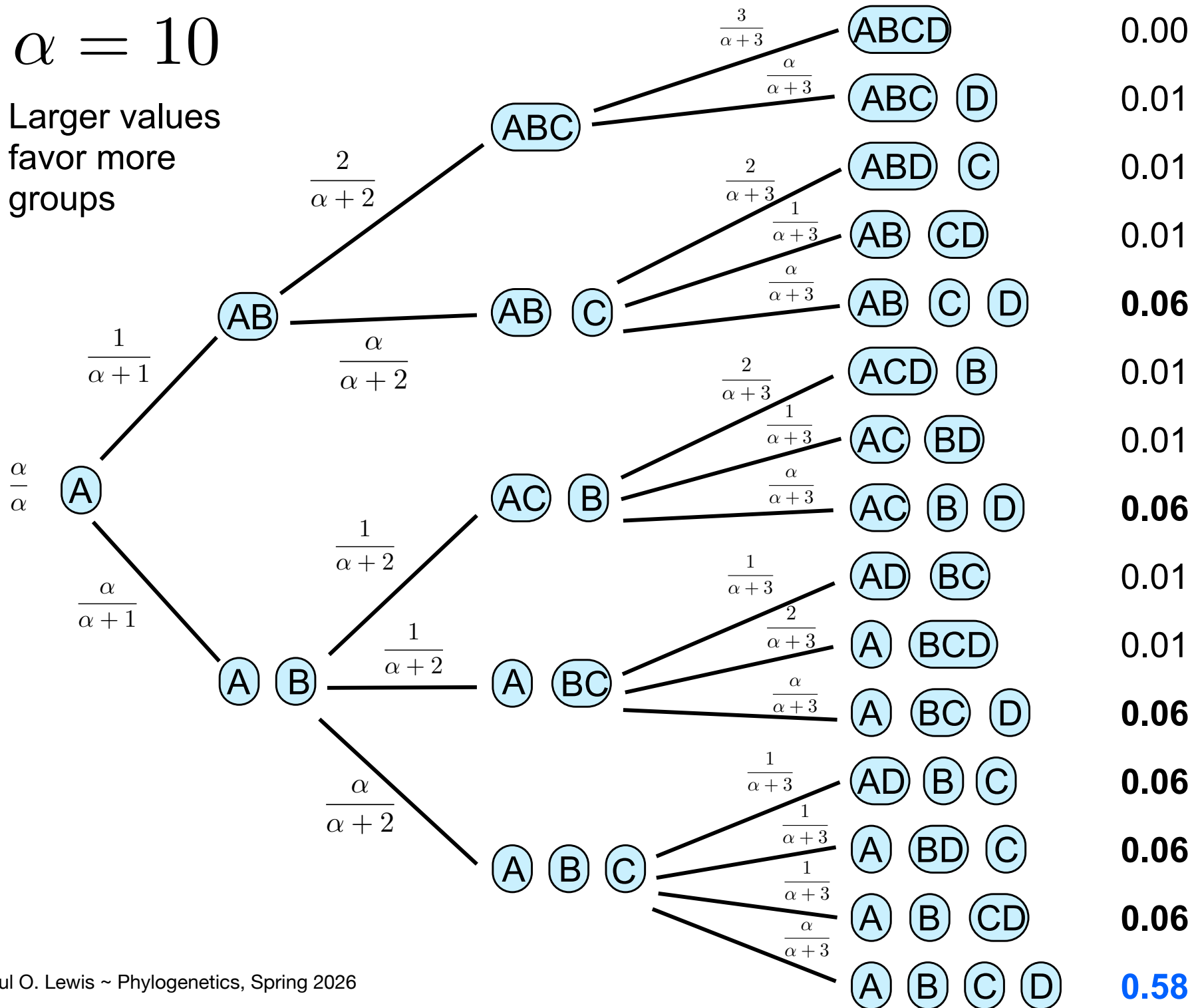
...or form its own group

After all objects have been considered, you can follow paths to determine the probability of different final configurations



$$\alpha = 10$$

Larger values favor more groups



Expected Number of Groups

$$\sum_{i=0}^{n-1} \frac{\alpha}{\alpha + i}$$

For example, if $n = 3$ and $\alpha = 1$:

$$\frac{\alpha}{\alpha + 0} + \frac{\alpha}{\alpha + 1} + \frac{\alpha}{\alpha + 2} = 1 + \frac{1}{2} + \frac{1}{3} = 1.83$$

Restaurant Analogy

