David Aldous (with J. Michael Steele)

Maximum Partial Matchings on Random Trees: An Illustration of the Local Weak Convergence Methodology

[Slides 2 – 8] Work through a specific problem in some detail.

[Slides 9 – 10] So what was the technique?

[Slides 11 – 14] More sophisticated problems.

These slides + survey paper are on my web site.

n-vertex graph G
positive weights ξ_e on edges e
partial matching S is a vertex-disjoint set of edges, and its weight is $\mathrm{wt}(S) = \sum_{e \in S} \xi_e$.
So associated with the weighted graph G is a number

$$M(G) = \max\{\mathrm{wt}(S) : \ S \text{ a partial matching } \}.$$

We study this under the **Probability Model**:

- Edge-weights (ξ_e) are i.i.d.
  $\xi > 0, \ E\xi < \infty, \ \ \mathrm{dist}(\xi)$ non-atomic.

- G is a random tree T_n uniform on all n^{n-2} trees on vertex-set $[n]$.

It takes 10 seconds to conjecture
Theorem $n^{-1} EM(T_n) \to c$ as $n \to \infty$
where the limit constant c depends on $\mathrm{dist}(\xi)$.

We will outline a proof and a fairly explicit characterization of c.

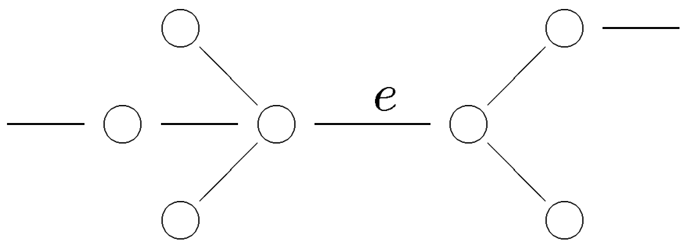
On a rooted tree T define the bonus $B(T)$ by

$$\begin{aligned} B(T) &= \text{max weight of matching} \\ &- \text{max weight of matching not using root.} \end{aligned}$$

An arbitrary edge e of T defines two rooted subtrees – call them $T^{big}(e)$ and $T^{small}(e)$. An elementary argument gives an explicit criterion for whether e is in the max-weight matching $\mathcal{M}$:

$$e \in \mathcal{M} \text{ iff } \xi_e > B(T^{small}(e)) + B(T^{big}(e))$$

(assuming cannot have =).



The maximum weight $M(T)$ is obtained by summing ξ_e over $e \in \mathcal{M}$. Rephasing in terms of a uniform random edge $\mathbf{e}$ leads to

$$\frac{EM(T)}{n-1} = E\left[\xi\ 1(\xi > B(T^{small}(\mathbf{e})) + B(T^{big}(\mathbf{e})))\ \right].$$

So if we can prove joint convergence

$$\Big(B(T_n^{small}(\mathbf{e})),\ B(T_n^{big}(\mathbf{e}))\Big) \ \stackrel{d}{\to}\ (Y,Z) \qquad (*)$$

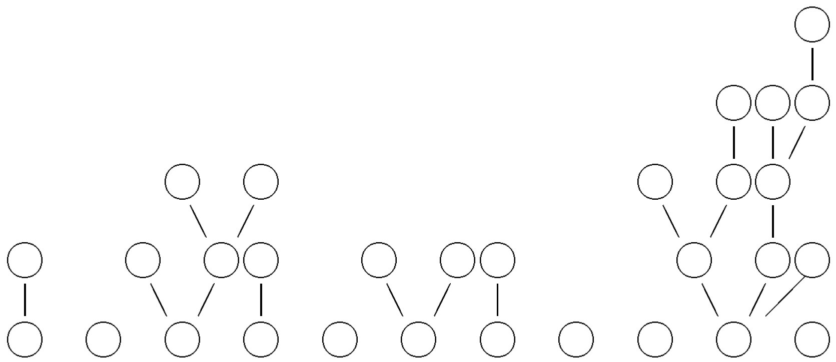
and identify the limit distribution, then we have proved the Theorem:

$$\lim_n n^{-1} EM(T_n) = E[\xi\ 1(\xi > Y + Z)].$$

Note we have not yet appealed to a particular model of random n-vertex trees T_n.

But are we getting anywhere? We have just reduced study of <u>expectation</u> of $M(T_n)$ to study of <u>distribution</u> of $B(T_n^{big}(\mathbf{e}))$ etc which is a (small) difference of two (large) random quantities. These look like harder questions!

Let us forget "bonus" for a minute; think about the subtrees themselves. Grimmett (1980): for a random edge e of the uniform random tree on $[n]$, there is the following $n \to \infty$ limit of (T_n^{small}, T_n^{big}).



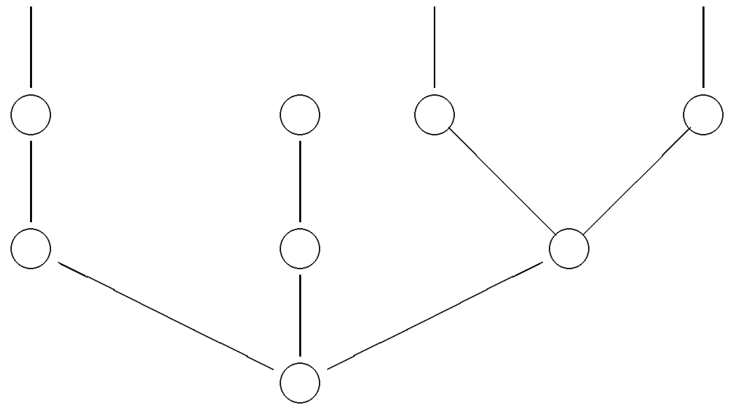
Each tree is distributed as PGW(1), that is a Galton-Watson branching process with Poisson(1) offspring.

T_∞^{small} is a single PGW(1) tree.

T_∞^{big} is a infinite sequence of PGW(1) trees attached to a baseline.

Look first at small subtrees. Certainly we have $B(T_n^{small}) \stackrel{d}{\to} B(T_\infty^{small})$. Although the realizations of T_∞^{small} can be any finite tree $\mathbf{t}$, we don't need to calculate $B(\mathbf{t})$ for every finite tree, because we have a recursion. For a finite tree $\mathbf{t}$ with first generation subtrees $\mathbf{t}_i$,

$$B(\mathbf{t}) = \max(0; \xi_i - B(\mathbf{t}_i), i \geq 1).$$



Apply to T_∞^{small}: bonus $Y = B(T_\infty^{small})$ satisfies

$$Y \stackrel{d}{=} \max(0, \xi_i - Y_i,\ 1 \leq i \leq \text{Pois}(1))$$

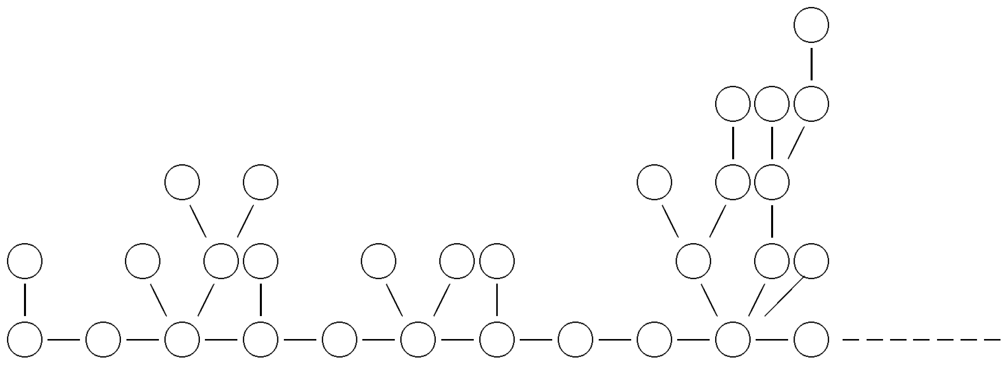
where $(\xi_i, i \geq 1)$ are i.i.d. with given weight distribution, and $(Y_i, i \geq 1)$ are i.i.d. with the unknown distribution of Y. This is a prototype of a **recursive distributional equation**. Easy: $\exists$ unique solution, depending on dist(ξ).

Recall $B(T_n^{big}) \stackrel{d}{\to} B(T_\infty^{big})$, where limit T_∞^{big} is a infinite sequence of PGW(1) trees attached to a baseline. What about the bonuses?

$$B(T_\infty^{big}) = \infty - \infty = \quad ?????$$

But suppose heuristically $B(\mathbf{t})$ makes sense for suitable infinite trees. Again we get a recursion

$$B(\mathbf{t}) = \max(\xi - B(\mathbf{t}^{infinite}), B(\mathbf{t}^{finite})).$$



Apply to T_∞^{big}: bonus $Z = B(T_\infty^{big})$ satisfies

$$Z \stackrel{d}{=} \max(\xi - Z', B(T_\infty^{small}))$$

where $Z' \stackrel{d}{=} Z$, unknown distribution. Easy: $\exists$ unique solution, depending on dist(ξ).

This <u>suggests</u> $B(T_n^{big}) \stackrel{d}{\to} Z$. Rigorous proof requires some thought; but can exploit the Markov chain $z \to \max(\xi - z, Y)$.

So we have proved

Theorem $n^{-1} EM(T_n) \to c = E[\xi \, 1(\xi > Y + Z)]$ where Y and Z are solutions of

$$Y \stackrel{d}{=} \max(0, \xi_i - Y_i, \ 1 \le i \le \text{Pois}(1))$$

$$Z \stackrel{d}{=} \max(\xi - Z', Y).$$

Can do explicit calculations in the special case where ξ has exponential(1) distribution:

$$\begin{aligned} P(Y \le y) &= \exp(-ce^{-y}), \quad y \ge 0 \\ P(Z \le z) &= (1 - be^{-z}) \exp(-ce^{-z}), \quad z \ge 0 \end{aligned}$$

where $c \approx 0.7146$ is the strictly positive solution of $c^2 + e^{-c} = 1$ and $b = \frac{c^2}{c^2+2c-1} \approx 0.5433$. So

$$\lim_n n^{-1} EM(T_n) =$$

$$\int_0^\infty s e^{-s} ds \int_0^s c(e^{-y} - be^{-s}) \exp(-ce^{-y} - ce^{-(s-y)}) \, dy$$

$$\approx 0.2396.$$

This problem illustrates two techniques.

Local weak convergence. Some size-n random discrete structure, viewed relative to a "random root", converges in distribution to an infinite discrete random structure.

Recursive distributional equations. When the limit structure is tree-like, quantities of interests can often be expressed in terms of solutions of RDEs. That is, an "unknown" distribution X satisfies

$$X \stackrel{d}{=} f(A; X_i, 1 \leq i \leq N)$$

where the distribution (A, N) and the function $f(\cdot)$ are given, and $X;\ X_i, i \geq 1$ are i.i.d.

xxx table from A-B

TSP (traveling salesman problem) in the mean-field model of distance.

Model: Complete graph on n vertices; the $\binom{n}{2}$ edge-lengths are i.i.d. with exponential (mean n) distribution.

Let L_n = length of minimal TSP tour.

It takes 10 seconds to conjecture

$$n^{-1} EL_N \to c.$$

It turns out there is a formula for the limit:

$$c = \int_0^\infty xP(X_1 + X_2 > x) \ dx \approx 2.04$$

where X solves the RDE

$$X \ \stackrel{d}{=} \ \text{second minimum}_i(\xi_i - X_i)$$

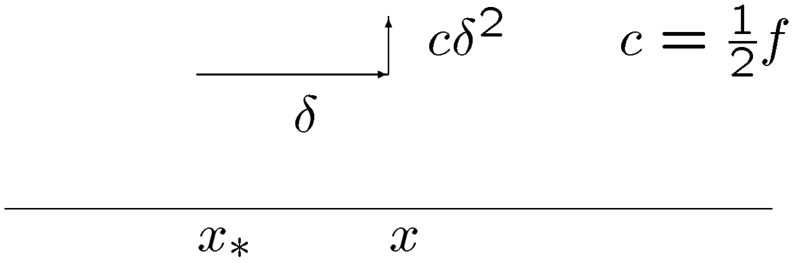
for $0 < \xi_1 < \xi_2 < \ldots$ a Poisson(rate 1) process. Mézard-Parisi (1985) argued this via replica method of statistical physics; Aldous (2001) outlines how to start a rigorous argument using local weak convergence.

The problem has similarities to "partial matching":

- model is locally tree-like; the ξ_i are distances to nearest points.
- X's have bonus interpretation.

Calculus meets Combinatorial Optimization

A minimum x_* satisfies $f'(x_*) = 0, \; f''(x_*) > 0.$



$$c = \tfrac{1}{2} f''(x_*)$$

In statistical physics language: typically the **scaling exponent** = 2 in all dimensions. That is, for $f : R^d \to R$ and a minimum x_*,

$$\min_{x:|x-x_*|=\delta} (f(x) - f(x_*)) \asymp \delta^2.$$

Consider a TSP or MST or . . . problem. Compare an arbitrary feasible solution $\mathbf{x}$ with the optimal (minimal) solution $\mathbf{x}_*$ by the two quantities

$\delta_n(\mathbf{x}) = \frac{1}{n}\times$ (number of edges in $\mathbf{x}$ but not in $\mathbf{x}_*$)

$\varepsilon_n(\mathbf{x}) = \frac{1}{s(n)}\times$ (cost difference between $\mathbf{x}$ and $\mathbf{x}_*$)

where $s(n)$ is the growth rate for optimal cost. Now define $\varepsilon_n(\delta)$ to be the minimum value of $\varepsilon_n(\mathbf{x})$ over all feasible solutions $\mathbf{x}$ for which $\delta_n(\mathbf{x}) \geq \delta$. Although the function $\varepsilon_n(\delta)$ will depend on n and the problem instance, we anticipate that for typical instances drawn from a suitable probability model it will converge in the $n \to \infty$ limit to some deterministic function. $\varepsilon(\delta)$. The *universality* paradigm suggests there might be scaling exponents α defined by

$$\varepsilon(\delta) \approx \delta^\alpha \text{ as } \delta \downarrow 0$$

and that the exponents should be robust under model details and should serve to categorize different combinatorial optimization problems into a small set of universality classes. If our analogy with freshman calculus is correct, the simplest combinatorial optimization problems should have scaling exponent 2.

This program is work-in-progress. What we believe is

• for MST, for any reasonable model (e.g. random points in R^d) the scaling exponent $= 2$.

• for TSP in the mean-field model of distance, the scaling exponent $= 3$.

The second result represents sophisticated use of the methodology of this talk.

Understanding the scaling exponent for TSP on random points in R^d seems far beyond current theory; but is amenable to simulation.