Stochastic Combinatorial Optimization from TSPs to MSTs via Caterpillars and Dogerpillars

J. Michael Steele
University of Pennsylvania, Wharton School

A Path Through the TSP, MST, and Interpolations

A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP

A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP
  - (Computational Differences, Stochastic Similarities, and Consequences of Both.)

A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP
  - (Computational Differences, Stochastic Similarities, and Consequences of Both.)
- Part II: Seeing Structures and Framing Conjectures

A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP
  - (Computational Differences, Stochastic Similarities, and Consequences of Both.)
- Part II: Seeing Structures and Framing Conjectures
  - (Some News you can Use and a TSP conjecture that fails.)

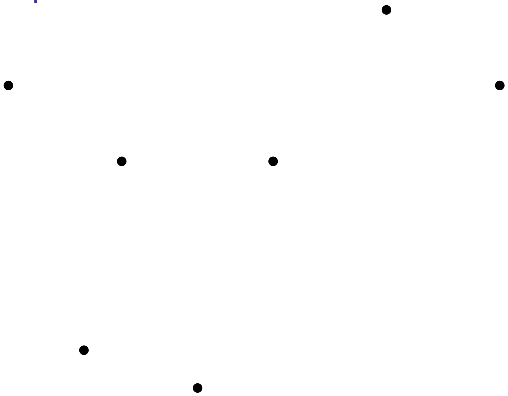
A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP
  - (Computational Differences, Stochastic Similarities, and Consequences of Both.)
- Part II: Seeing Structures and Framing Conjectures
  - (Some News you can Use and a TSP conjecture that fails.)
- Part III: Interpolation — The Real Theme

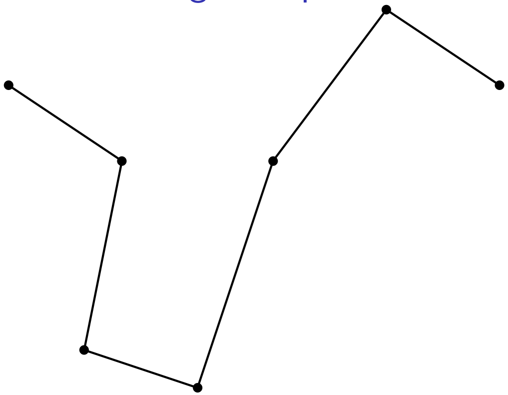
A Path Through the TSP, MST, and Interpolations

- Part 1: Getting Started with the MST and TSP
  - (Computational Differences, Stochastic Similarities, and Consequences of Both.)
- Part II: Seeing Structures and Framing Conjectures
  - (Some News you can Use and a TSP conjecture that fails.)
- Part III: Interpolation — The Real Theme
  - (Wherein the famous Dogerpillars are introduced and explored.)

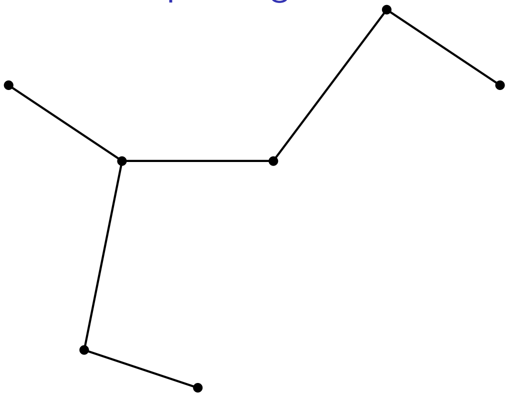
Consider a set of points ...



Now, a Shortest Path through the points ...



Now, consider a minimal spanning tree ...



Computational Theory of the TSP

Computational Theory of the TSP

- The TSP Problem is NP-Complete

Computational Theory of the TSP

- The TSP Problem is NP-Complete
- As a practical mater, solving large TSPs exactly is not possible

Computational Theory of the TSP

- The TSP Problem is NP-Complete
- As a practical mater, solving large TSPs exactly is not possible
- There are recent (circa 1995) ϵ-approximation algorithms

Computational Theory of the TSP

- The TSP Problem is NP-Complete
- As a practical mater, solving large TSPs exactly is not possible
- There are recent (circa 1995) ϵ-approximation algorithms
- These are still impractical — $O(n^{p(1/\epsilon)})$

Computational Theory of the TSP

- The TSP Problem is NP-Complete
- As a practical mater, solving large TSPs exactly is not possible
- There are recent (circa 1995) ϵ-approximation algorithms
- These are still impractical — $O(n^{p(1/\epsilon)})$
- Bottom Line:
  Even the ϵ approximation to the TSP is "essentially" impossible

Computational Theory of the TSP

- The TSP Problem is NP-Complete
- As a practical mater, solving large TSPs exactly is not possible
- There are recent (circa 1995) ϵ-approximation algorithms
- These are still impractical — $O(n^{p(1/\epsilon)})$
- Bottom Line:
  Even the ϵ approximation to the TSP is "essentially" impossible
- On the Other Hand:
  There is a $O(n)$ time ϵ algorithm (Karp-Steele (1985)) if you assume a probability model for the points.

Computational Theory of the MST

Computational Theory of the MST

- The MST problem is computationally easy!

Computational Theory of the MST

- The MST problem is computationally easy!
- Greedy algorithms of several kinds give $O(n^2)$ algorithms

Computational Theory of the MST

- The MST problem is computationally easy!
- Greedy algorithms of several kinds give $O(n^2)$ algorithms
- With fancy data structures you can even do better

Computational Theory of the MST

- The MST problem is computationally easy!
- Greedy algorithms of several kinds give $O(n^2)$ algorithms
- With fancy data structures you can even do better
- Current records are faster than $O(n \log^{\epsilon} n)$

Computational Theory of the MST

- The MST problem is computationally easy!
- Greedy algorithms of several kinds give $O(n^2)$ algorithms
- With fancy data structures you can even do better
- Current records are faster than $O(n \log^{\epsilon} n)$
- Bottom Line:
  The MST and TSP may LOOK like similar problems ....

Computational Theory of the MST

- The MST problem is computationally easy!
- Greedy algorithms of several kinds give $O(n^2)$ algorithms
- With fancy data structures you can even do better
- Current records are faster than $O(n \log^{\epsilon} n)$
- Bottom Line:
  The MST and TSP may LOOK like similar problems ....
- BUT:
  Their computational theory tells us that they are wildly different

Probability Theory of the TSP, MST, and More

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.
- Let L_n^{MST} and L_n^{TSP} denote the lengths of the optimal tree and optimal path

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.
- Let L_n^{MST} and L_n^{TSP} denote the lengths of the optimal tree and optimal path
- A moments thought suggests that these should be $O(\sqrt{n})$

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.
- Let L_n^{MST} and L_n^{TSP} denote the lengths of the optimal tree and optimal path
- A moments thought suggests that these should be $O(\sqrt{n})$
- We actually have a precise limit theorem in each case:

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.
- Let L_n^{MST} and L_n^{TSP} denote the lengths of the optimal tree and optimal path
- A moments thought suggests that these should be $O(\sqrt{n})$
- We actually have a precise limit theorem in each case:
- $$lim_{n\to\infty}\frac{L_n}{\sqrt{n}} = C\int_{R^2}\sqrt{f(x)}\,dx. \qquad \text{with probability one}$$

Probability Theory of the TSP, MST, and More

- Make life simple and consider the case of the TSP and MST for sets of n points chosen at random from a density f with compact support in R^2.
- Let L_n^{MST} and L_n^{TSP} denote the lengths of the optimal tree and optimal path
- A moments thought suggests that these should be $O(\sqrt{n})$
- We actually have a precise limit theorem in each case:
- $$lim_{n\to\infty} \frac{L_n}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}\, dx. \qquad \text{with probability one}$$
- For the TSP this is the famous Beardwood-Halton-Hammersly theorem of 1959. For the MST the result is from Steele (1988). The constants C_{TSP} and C_{MST} are not known exactly. The natural analogs hold in $d \geq 2$.

Idea of the Modern Proof of BHH

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization
  - Get the limits by an extension of Fekete's lemma

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization
  - Get the limits by an extension of Fekete's lemma
  - Back out to EL_n via a Tauberian theorem

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization
  - Get the limits by an extension of Fekete's lemma
  - Back out to EL_n via a Tauberian theorem
- This method is easier than that used by BHH, but it is close in spirit

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization
  - Get the limits by an extension of Fekete's lemma
  - Back out to EL_n via a Tauberian theorem
- This method is easier than that used by BHH, but it is close in spirit
- The modern approach to the variance is radically different from that used by BHH

Idea of the Modern Proof of BHH

- Three Steps: Reduction to Uniforms, Control of the Mean, and Control of the Variance
- Method for the Means
  - Dissect the square to k^2 subsquares
  - Get subadditivity and smooth it by Poissonization
  - Get the limits by an extension of Fekete's lemma
  - Back out to EL_n via a Tauberian theorem
- This method is easier than that used by BHH, but it is close in spirit
- The modern approach to the variance is radically different from that used by BHH
- The modern package is much more robust to changes in "problem" and "model".

Step 3: The Variance and “News You can Use”

Step 3: The Variance and "News You can Use"

- There is a VERY GENERAL trick for bounding

$$\mathrm{Var} F(X_1, X_2, ..., X_n)$$

Step 3: The Variance and "News You can Use"

- There is a VERY GENERAL trick for bounding

$$\mathrm{Var} F(X_1, X_2, ..., X_n)$$

- Steele (1981), generalizing Efron and Stein (1980), showed it is bounded by

$$\frac{1}{2} \sum_{1 \le i \le n} E(F(X_1, X_2, ..., X_i, ..., X_n) - F(X_1, X_2, ..., \hat{X}_i, ..., X_n))^2$$

Step 3: The Variance and "News You can Use"

- There is a VERY GENERAL trick for bounding

$$\mathrm{Var} F(X_1, X_2, \ldots, X_n)$$

- Steele (1981), generalizing Efron and Stein (1980), showed it is bounded by

$$\frac{1}{2}\sum_{1\le i\le n} E(F(X_1, X_2, \ldots, X_i, \ldots, X_n) - F(X_1, X_2, \ldots, \hat{X}_i, \ldots, X_n))^2$$

- For the TSP and MST this can be used to prove that (in $d = 2$) there is a constant C such that for

$$\mathrm{Var} L_n \le C \qquad \text{for all } n$$

Step 3: The Variance and "News You can Use"

- There is a VERY GENERAL trick for bounding

$$\mathrm{Var} F(X_1, X_2, \ldots, X_n)$$

- Steele (1981), generalizing Efron and Stein (1980), showed it is bounded by

$$\frac{1}{2} \sum_{1 \le i \le n} E(F(X_1, X_2, \ldots, X_i, \ldots, X_n) - F(X_1, X_2, \ldots, \hat{X}_i, \ldots, X_n))^2$$

- For the TSP and MST this can be used to prove that (in $d = 2$) there is a constant C such that for

$$\mathrm{Var} L_n \le C \qquad \text{for all } n$$

- Even now this may seem surprising. Here, and in many other cases, it gives an very pleasing path to the desired strong laws.

Sharper Concentration

Sharper Concentration

- After intermediate results by several authors, Talagrand proved with his convex distance inequality that the TSP and MST in $d = 2$ have Gaussian tail bounds:

$$P(|L_n - EL_n| \geq x) \leq Ae^{-Bx^2}.$$

Sharper Concentration

- After intermediate results by several authors, Talagrand proved with his convex distance inequality that the TSP and MST in $d = 2$ have Gaussian tail bounds:

$$P(|L_n - EL_n| \geq x) \leq Ae^{-Bx^2}.$$

- The proof of this inequality can be somewhat simplified by using the Spacefilling Curve Heuristic to get a "certificate" for the size of Talangrand's distance. Results of the same nature can be obtained using Ledoux's Log-Sobolev inequality.

Sharper Concentration

- After intermediate results by several authors, Talagrand proved with his convex distance inequality that the TSP and MST in $d = 2$ have Gaussian tail bounds:

$$P(|L_n - EL_n| \geq x) \leq Ae^{-Bx^2}.$$

- The proof of this inequality can be somewhat simplified by using the Spacefilling Curve Heuristic to get a "certificate" for the size of Talangrand's distance. Results of the same nature can be obtained using Ledoux's Log-Sobolev inequality.
- The problems in $d > 2$ dropped briefly off the radar. Here we only know

$$\mathrm{Var} L_n \leq C n^{(d-2)/d} \qquad \text{for all } n$$

Sharper Concentration

- After intermediate results by several authors, Talagrand proved with his convex distance inequality that the TSP and MST in $d = 2$ have Gaussian tail bounds:

$$P(|L_n - EL_n| \geq x) \leq Ae^{-Bx^2}.$$

- The proof of this inequality can be somewhat simplified by using the Spacefilling Curve Heuristic to get a "certificate" for the size of Talangrand's distance. Results of the same nature can be obtained using Ledoux's Log-Sobolev inequality.
- The problems in $d > 2$ dropped briefly off the radar. Here we only know

$$\mathrm{Var} L_n \leq Cn^{(d-2)/d} \qquad \text{for all } n$$

- We do not know if this is the "truth" when $d > 2$. Lower bounds on variance are hard to come by. We have similar open issues with respect to sharp concentration in $d > 2$.

Intermezzo: A Natural Conjecture

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations
- How about processes that are stationary with ergodic density f?

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations
- How about processes that are stationary with ergodic density f?
- This question is motivated by recent progress by Nobel (2008) who showed results in Vapnik-Chervonikis theory continue to hold in the stationary ergodic case.

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations
- How about processes that are stationary with ergodic density f?
- This question is motivated by recent progress by Nobel (2008) who showed results in Vapnik-Chervonikis theory continue to hold in the stationary ergodic case.
- As it happens, for the TSP and MST one DOES NOT have the strong law for stationary ergodic processes.

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations
- How about processes that are stationary with ergodic density f?
- This question is motivated by recent progress by Nobel (2008) who showed results in Vapnik-Chervonikis theory continue to hold in the stationary ergodic case.
- As it happens, for the TSP and MST one DOES NOT have the strong law for stationary ergodic processes.
- The construction of the counter-example in Steele (2009) is too complicated to give in detail, but it is actually pretty simple given some conceptual hints.

Intermezzo: A Natural Conjecture

- It is natural to ask if one really needs independent, identically distributed observations
- How about processes that are stationary with ergodic density f?
- This question is motivated by recent progress by Nobel (2008) who showed results in Vapnik-Chervonikis theory continue to hold in the stationary ergodic case.
- As it happens, for the TSP and MST one DOES NOT have the strong law for stationary ergodic processes.
- The construction of the counter-example in Steele (2009) is too complicated to give in detail, but it is actually pretty simple given some conceptual hints.
- Three Hints Do It: Cyclic processes, $O(n^{-(1+\epsilon)/2})$ shifts, and subsequent scales.

Part III: Interpolation between the TSP and MST

Part III: Interpolation between the TSP and MST

- There are several ways the TSP is hard and the MST is easy:

Part III: Interpolation between the TSP and MST

- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier

Part III: Interpolation between the TSP and MST

- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier
  - Probabilistic Differences — E.g. the CLT of Kesten and Lee.

Part III: Interpolation between the TSP and MST

- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier
  - Probabilistic Differences — E.g. the CLT of Kesten and Lee.
- Here is a puzzle: Are there "structures" that fall between the TSP and MST?

Part III: Interpolation between the TSP and MST

- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier
  - Probabilistic Differences — E.g. the CLT of Kesten and Lee.
- Here is a puzzle: Are there "structures" that fall between the TSP and MST?
- There is at least on reasonably well-studied class of graphs does fall in between

Part III: Interpolation between the TSP and MST

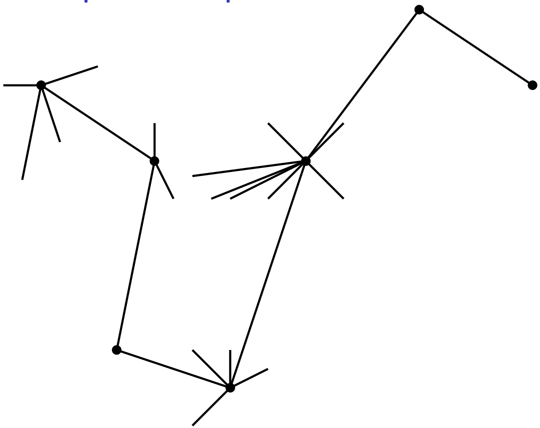
- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier
  - Probabilistic Differences — E.g. the CLT of Kesten and Lee.
- Here is a puzzle: Are there "structures" that fall between the TSP and MST?
- There is at least on reasonably well-studied class of graphs does fall in between
- In graph theory, a CATERPILLAR is a graph G with two properties:
  - It is itself a tree (i.e. connected and without cycles)
  - It contains a path P such that if P is deleted from G the resulting graphs is a union of disjoint stars.

Part III: Interpolation between the TSP and MST

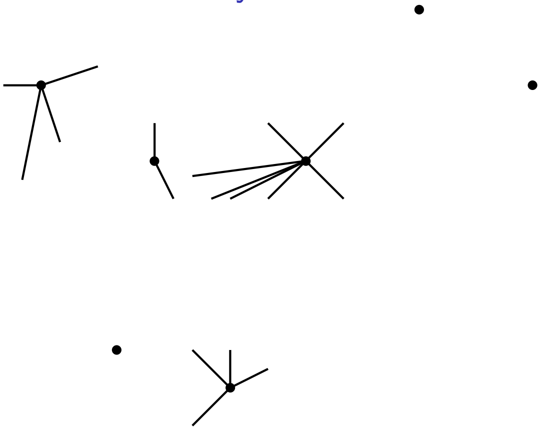
- There are several ways the TSP is hard and the MST is easy:
  - The computational issues we mentioned earlier
  - Probabilistic Differences — E.g. the CLT of Kesten and Lee.
- Here is a puzzle: Are there "structures" that fall between the TSP and MST?
- There is at least on reasonably well-studied class of graphs does fall in between
- In graph theory, a CATERPILLAR is a graph G with two properties:
  - It is itself a tree (i.e. connected and without cycles)
  - It contains a path P such that if P is deleted from G the resulting graphs is a union of disjoint stars.
- The minimal spanning caterpillar is well defined and we naturally have

$$L_n^{\text{MST}} \leq L_n^{\text{CAT}} \leq L_n^{\text{TSP}}$$

A Picture of a Simple Caterpillar ...



Remove a path and have only stars ...



Strong Law for Caterpillars

Strong Law for Caterpillars

- One gets the asymptotics of the means just as one does in the modern proof of the BHH

Strong Law for Caterpillars

- One gets the asymptotics of the means just as one does in the modern proof of the BHH
- For the variance, one again uses the Jackknife inequality.

Strong Law for Caterpillars

- One gets the asymptotics of the means just as one does in the modern proof of the BHH
- For the variance, one again uses the Jackknife inequality.
- This time there are some technical difficulties:
  - So far I only get

$$\mathrm{Var} L_n \leq C_\epsilon n^\epsilon \quad \text{for all } n$$

  - One suspects this can be improved to universal boundedness as for the TSP and MST

Strong Law for Caterpillars

- One gets the asymptotics of the means just as one does in the modern proof of the BHH
- For the variance, one again uses the Jackknife inequality.
- This time there are some technical difficulties:
  - So far I only get

$$\mathrm{Var} L_n \leq C_\epsilon n^\epsilon \quad \text{for all } n$$

  - One suspects this can be improved to universal boundedness as for the TSP and MST
- Still, this is good enough. One gets the strong law for minimal spanning caterpillars.

From Caterpillars to Dogerpillars

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete
  - Nevertheless, a non-trivial fraction of the length of a caterpillar will be in the stars.

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete
  - Nevertheless, a non-trivial fraction of the length of a caterpillar will be in the stars.
- One would like a richer class than caterpillars — something that would eventually catch up with the spanning tree.

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete
  - Nevertheless, a non-trivial fraction of the length of a caterpillar will be in the stars.
- One would like a richer class than caterpillars — something that would eventually catch up with the spanning tree.
- This leads one to the notion of a Dogerpillar.

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete
  - Nevertheless, a non-trivial fraction of the length of a caterpillar will be in the stars.
- One would like a richer class than caterpillars — something that would eventually catch up with the spanning tree.
- This leads one to the notion of a Dogerpillar.
- Note: If you Google "caterpillar" you will find much irrelevant information.

From Caterpillars to Dogerpillars

- One can ask: Are Caterpillars more like MSTs or TSPs?
  - Computation of the minimum spanning caterpillar is NP complete, because ...
  - Just getting the order of the Stars right is NP complete
  - Nevertheless, a non-trivial fraction of the length of a caterpillar will be in the stars.
- One would like a richer class than caterpillars — something that would eventually catch up with the spanning tree.
- This leads one to the notion of a Dogerpillar.
- Note: If you Google "caterpillar" you will find much irrelevant information.
- Ironically, if you Google "dogerpillar" you still find much that irrelevant. It's hard to come up with a neologism these days.

Introducing Dogerpillars

Introducing Dogerpillars

- Definition: A graph G is a dogerpillar (more precisely a $k = k(n)$-dogerpillar) if
  - there is a path P in G such that
  - if you delete P from G the resulting graph is a collection of **disjoint trees** such that each tree has no more than k vertices.

Introducing Dogerpillars

- Definition: A graph G is a dogerpillar (more precisely a $k = k(n)$-dogerpillar) if
  - there is a path P in G such that
  - if you delete P from G the resulting graph is a collection of **disjoint trees** such that each tree has no more than k vertices.
- Specializations:
  - Taking $k = n$, we see that the Dogerpillar is a tree
  - Taking $k = 0$, we see that the Dogerpillar is a path
  - The most interesting cases are $k = o(n)$, especially $k = O(\sqrt{n})$.

Why I like Dogerpillars

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP
- The suggest some questions that seem compelling:

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP
- The suggest some questions that seem compelling:
  - For "large enough k" the Dogerpillar should become the MST.

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP
- The suggest some questions that seem compelling:
  - For "large enough k" the Dogerpillar should become the MST.
  - It seems very interesting to know this critical rate

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP
- The suggest some questions that seem compelling:
  - For "large enough k" the Dogerpillar should become the MST.
  - It seems very interesting to know this critical rate
  - At the critical rate, we have the CLT, but we may be able to get the CLT at much lower critical rates.

Why I like Dogerpillars

- The minimal spanning dogerpillars do seem to be the most natural interpolation of the class of spanning trees and TSP paths
- They retain much of the tractability of the MST and TSP
- The suggest some questions that seem compelling:
  - For “large enough k” the Dogerpillar should become the MST.
  - It seems very interesting to know this critical rate
  - At the critical rate, we have the CLT, but we may be able to get the CLT at much lower critical rates.
- Let’s look at the progress to date...

Probability Theory of Dogerpillars — Circa March 2010

Probability Theory of Dogergpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\mathrm{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\mathrm{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\mathrm{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP
- Nevertheless, there are many questions that are open ...

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\mathrm{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)} dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP
- Nevertheless, there are many questions that are open ...
  - Relaxation of $k_n \sim n^\alpha$

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\mathrm{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP
- Nevertheless, there are many questions that are open ...
  - Relaxation of $k_n \sim n^\alpha$
  - Universal boundedness of $\mathrm{Var} L_n^{\mathrm{DOG}}$ when $d = 2$.

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n\to\infty} \frac{L_n^{\text{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP
- Nevertheless, there are many questions that are open ...
  - Relaxation of $k_n \sim n^\alpha$
  - Universal boundedness of $\text{Var} L_n^{\text{DOG}}$ when $d = 2$.
  - Sharp Tail Bounds as Talagrand found for the TSP and MST in $d = 2$.

Probability Theory of Dogerpillars — Circa March 2010

- For all $k_n \sim n^\alpha$ with $0 < \alpha < 1$ one has the strong law:

$$\lim_{n \to \infty} \frac{L_n^{\text{DOG}}}{\sqrt{n}} = C \int_{R^2} \sqrt{f(x)}dx$$

- The proof of this result required the development of many facts that are analogous to those that were useful for the MST and TSP
- Nevertheless, there are many questions that are open ...
  - Relaxation of $k_n \sim n^\alpha$
  - Universal boundedness of $\text{Var}L_n^{\text{DOG}}$ when $d = 2$.
  - Sharp Tail Bounds as Talagrand found for the TSP and MST in $d = 2$.
  - CLT for at least some interesting ranges of k_n, at least $k_n \sim n^{1/2+\epsilon}$

Concluding Remarks

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.
- We did not go into the widely diffused applications in computer science, communication theory, and networks of many kinds.

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.
- We did not go into the widely diffused applications in computer science, communication theory, and networks of many kinds.
- YOU WILL at some future time find a place in your work to apply the **jackknife inequality** that we reviewed our discussion of the surprising bound $\mathrm{Var} L_n \leq C < \infty$

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.
- We did not go into the widely diffused applications in computer science, communication theory, and networks of many kinds.
- YOU WILL at some future time find a place in your work to apply the **jackknife inequality** that we reviewed our discussion of the surprising bound $\mathrm{Var} L_n \leq C < \infty$
- We introduced Caterpillars and Dogerpillars

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.
- We did not go into the widely diffused applications in computer science, communication theory, and networks of many kinds.
- YOU WILL at some future time find a place in your work to apply the **jackknife inequality** that we reviewed our discussion of the surprising bound $\mathrm{Var} L_n \leq C < \infty$
- We introduced Caterpillars and Dogerpillars
- We covered what is known of their probability theory and suggested ways they may be useful in the future.

Concluding Remarks

- We've reviewed the probability theory of the TSP and MST as it has evolved over the last 25 years.
- We did not go into the widely diffused applications in computer science, communication theory, and networks of many kinds.
- YOU WILL at some future time find a place in your work to apply the **jackknife inequality** that we reviewed our discussion of the surprising bound $\mathrm{Var} L_n \leq C < \infty$
- We introduced Caterpillars and Dogerpillars
- We covered what is known of their probability theory and suggested ways they may be useful in the future.
- Thank You for Your Attention!