6125021: Introduction to Combinatorics

Dr. Zichen Xu (徐子晨)
Department of Artificial Intelligence
Department of Computer Science and Technology
The Nanchang University

About Me

- Dr. Zichen Xu（徐子晨）
- Professor, School of Information Engineering
- Chair, Department of Artificial Intelligence
- Currently working in Ministry of Education
- Former Googler, Buckeye, Gator, and Bull

- Main field:
  - Distributed database management systems,
  - Data analysis, and AI-enabled data intensive Computing,
  - Green computing, Data storage

Generic Operational and Optimal Data Group

- Course page: http://good.ncu.edu.cn/~xuz/ic/homepage.html

- Email: zichenxu@outlook.com (Preferred contact info)

- ~~Come and meet me in my office during open office hour~~
  - ~~Time: Every Thursday Afternoon from 5PM to 5:30PM~~
  - ~~Office Location: IEB-608b~~

SCAN ME

Content of The Day

- Syllabus
- The style of this class
- Why you shall take full attention to Combinatorial Mathematics
- Some early combinatorial mathematics problem

Syllabus: 6125021 Introduction to Combinatorial Mathematics

- Class time: Saturday (Sat) 1:45PM- whenever I would like to finish

- Location: ~~Graduate school building 218~~
  - If something comes up and I need to reschedule, you will get notified

- Textbook:
  - Applied Combinatorics, online edition, by Mitchel T. Keller and William T. Trotter
  - Applied Combinatorics, 6th edition, by Alan Tucker
  - Both are online download available or you can purchase them at 50 Euro each.

Syllabus

- Description
  - Combinatorics and discrete mathematics are increasingly important, particularly for their applications in computer science. This course will give a brief overview of this subject.

- Topic List
  - Introduction to Combinatorics: first-order logic, factorials, permutations and combinations, binomial coefficients, Stirling numbers, double counting;
  - Induction and Comb. Basics: bijections, binomial theorem, generating functions;
  - Graph theory: bridges of Konigsberg, Eulerian circuits, trees, edge coloring, vertex coloring, planar graphs, Kempe's proof of the 5-color theorem;
  - Network flows: sphere packing bound, Hamming codes;
  - Graphic algorithms: Djkstra's algorithm for minimum spanning tree, depth first and breadth first algorithms for trees, greedy algorithm for graph coloring.

Syllabus

- How to pass this course?
  - Take at least 80% of the class and finish all homework = Pass the course
  - Take 100% of the class and finish all homework && (projects || do well in the tests) = B+
  - Take 100% of the class and finish all homework && do well projects && in tests = A
  - ∃ The best student in this class || No one = A+, and one recommendation letter

  - All the grades will be scaled based on normal distribution (bell shape)

Syllabus

- How many pieces of homework?
  - Approximately once per two weeks = roughly 4 homework
- How many projects?
  - Only one throughout the entire semester
  - More details in the later page
- How many tests?
  - No more than 4 tests, each test includes one or two questions
  - More details in the later page

Homework

- Roughly one or two questions will be asked randomly at the end of one lecture in two weeks
- All homework must be sent to ncuhomework@outlook.com
- The homework will be given a deadline to submit, late homework will get 0 pts, no exceptions
- I will randomly select a subset of all submitted homework, and give grades to you using multi-secretary algorithm
  - E.g., If the first three fail, everyone will get a relatively low grade
  - If not, then vice versa
  - If you did not submit, you will get a 0
  - If you submit late, you will get a 0

Projects

- Five in one group
  - Male + female group gets 5% more grades than all female or all male

- Content
  - One opening report: 7 mins talk and 2 mins for one/two questions, approximately 5 hours to present
  - One final report: everything as a full technical report

Projects

- Grading
  - The projects will be delivered with a list of hot topics and possible related combinatorics issues, you will come up with the paper idea/title
  - A good idea = 100 pts, else gets 90 pts
  - Each presentation will be graded by quality of the presentation and the slides
  - Each report will be graded by writing, typos, and logic flow
  - NOTE, PLAGIARISM WILL FAIL YOU IN THIS CLASS, I guarantee you this!
  - Reports will always have a Q&A round, each question asked will gain questioner 10 pts, each question answered will gain questionee 10 pts, each question unanswered will take 10 pts from questionee to questioner
  - If no question asked after one talk, every group gets 20 pts off, presenter gets 50 pts
  - At last, all grades will be summed and normalized

Test

- Tests are distributed randomly
- Test time will be selected randomly, at the beginning of the class
- Test contains one or two question(s) and give you 10 mins to finish
- Only the top n-1 test grades are taken into account

Your final grade

- Your final grade = homework (30%) + Projects (50%) + Tests/Attendances (20%)

- Second-chance policy
  - To those who are about to fail this course in the later stage of the class, you will have a revive project to finish (one week)
  - If you finished, you are saved
  - If not, you fail the course

Time Schedule

- Class begins in September
  - Once per week, on Sat
- Report process begins in mid October
  - One class in this month will be the presentation day
  - Idea list will be delivered by the end of Sep, before the holiday
  - The final week will be the final presentation week

The Style of the Class

- A loose-tight policy on grading
  - It is very hard to fail the course, if you apply the rule
  - It is very hard to get the top grade, you have to be very excellent

- The class would be very versatile in the tempo-manner
  - I will start on time
  - I may not give you any rest intervals
  - I may extend the class for a certain amount of time if necessary
  - If you wanna take the rest room, go and take it
  - If you wanna leave the class early, you will take one attendance off

The Style of the Class

- The time manner of the class (cont.)
  - If you feel sick and need to absence, give me the proof from University hospital
  - If you have something else more important, you have to notify me at least one week before, otherwise you are counted as absence
  - Missing one-third of the class will fail you

- This is an English and Mandarin combined-teaching class
  - Basically most of the knowledge will be taught in English
  - If you cannot understand the class, go and learn your English well

  (I will explain why later)

The Style of the Class

- This course will use multi-media materials
  - All materials will be delivered in PDF (non copy right violation)
- Feedback on the course is strongly encouraged
  - Every useful point will be added as bonus pts in your grade
- Strongly recommend you combine what you learn in the course into your research
  - Even if some of you do not do research or do not have research publication requirement in your graduation

The Only Requirement for the Class

Plagiarism is NOT ALLOWED in any sense!

Plagiarism is NOT ALLOWED in any sense!

Plagiarism is NOT ALLOWED in any sense!

What is Combinatorics

So there arise two kinds of *variation: complexion and situs.* And viewed in themselves, both complexion and situs belong to metaphysics, or to the science of whole and parts.

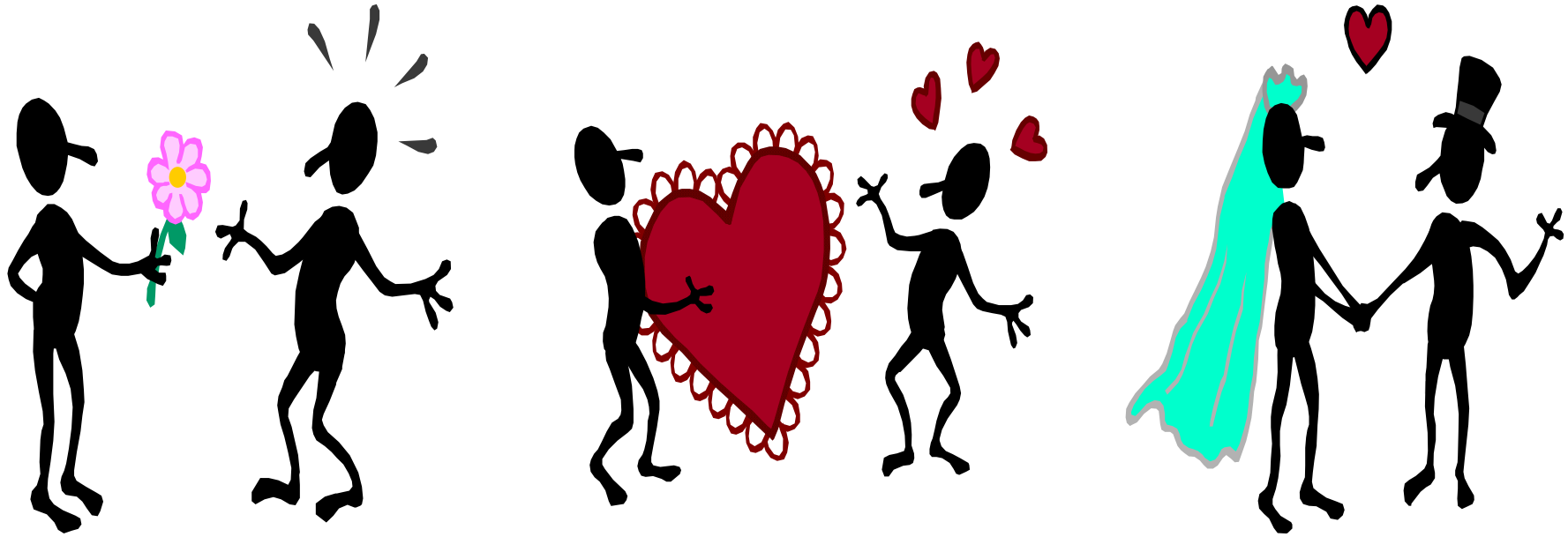
- Gottfried Leibniz

(on his dissertation when he is 20 yrs old)

Why you shall take full attention to Combinatorial Mathematics

- This is a must-take basic course for computer science student
- This is a must-take course for anyone who wants to graduate on time
- This is a very useful application mathematics
  - You may benefit from it for the entire of your life

Question: Stable Marriage




WARNING: This lecture contains mathematical content that may be shocking to some students

Dating Scenario

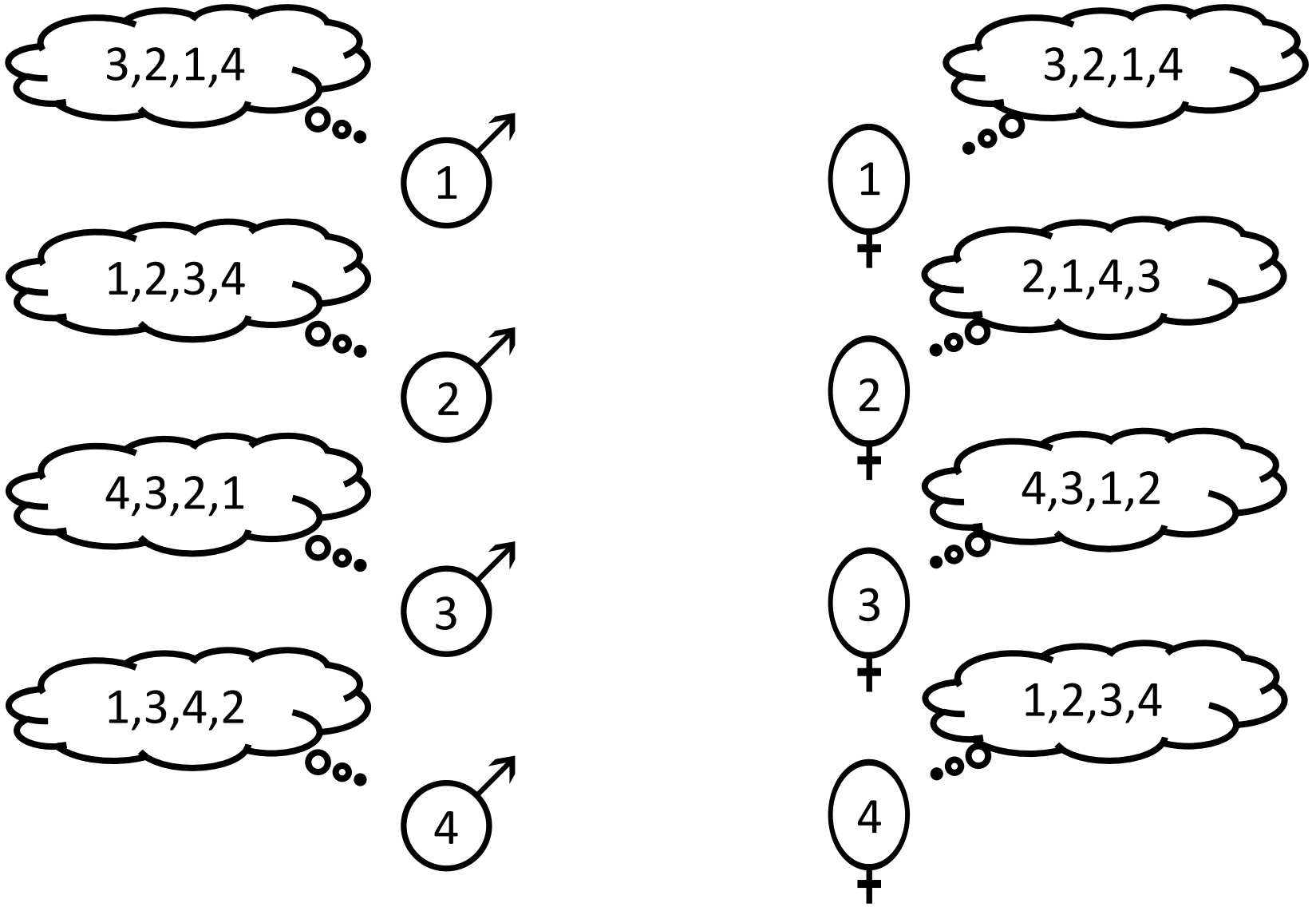
There are n boys and n girls

Each girl has her own ranked preference list of all the boys

Each boy has his own ranked preference list of the girls

The lists have no ties

Question: How do we pair them off?


3,2,1,4
1
1,2,3,4
2
4,3,2,1
3
1,3,4,2
4
3,2,1,4
1
2,1,4,3
2
4,3,1,2
3
1,2,3,4
4

More Than One Notion of What Constitutes A “Good” Pairing

Maximizing total satisfaction

Hong Kong and to an extent the USA?

Maximizing the minimum satisfaction

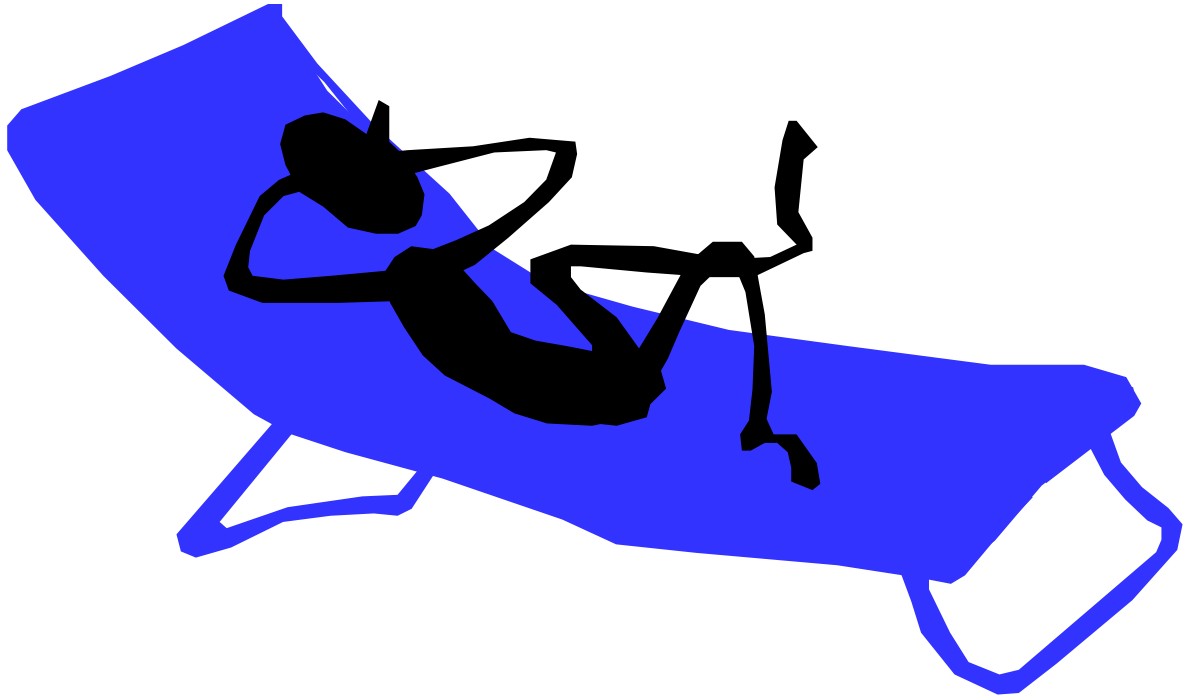
Western Europe, Canada?

Minimizing maximum difference in mate ranks

Sweden?

Maximizing people who get their first choice

Barbie and Ken Land?



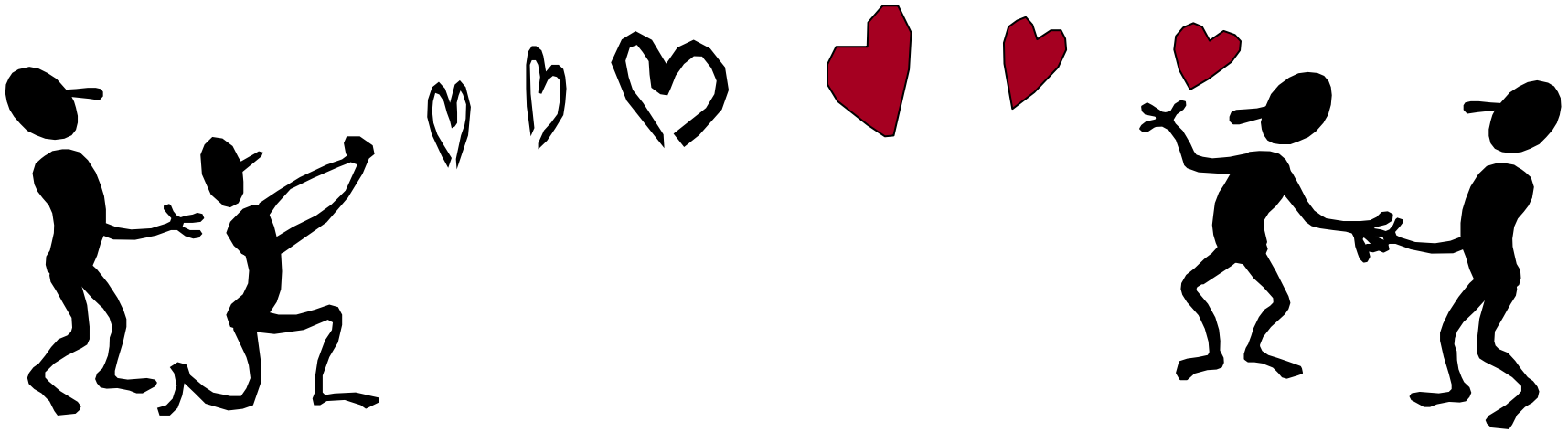
We will ignore the issue of what is “equitable”!

Rogue Couples

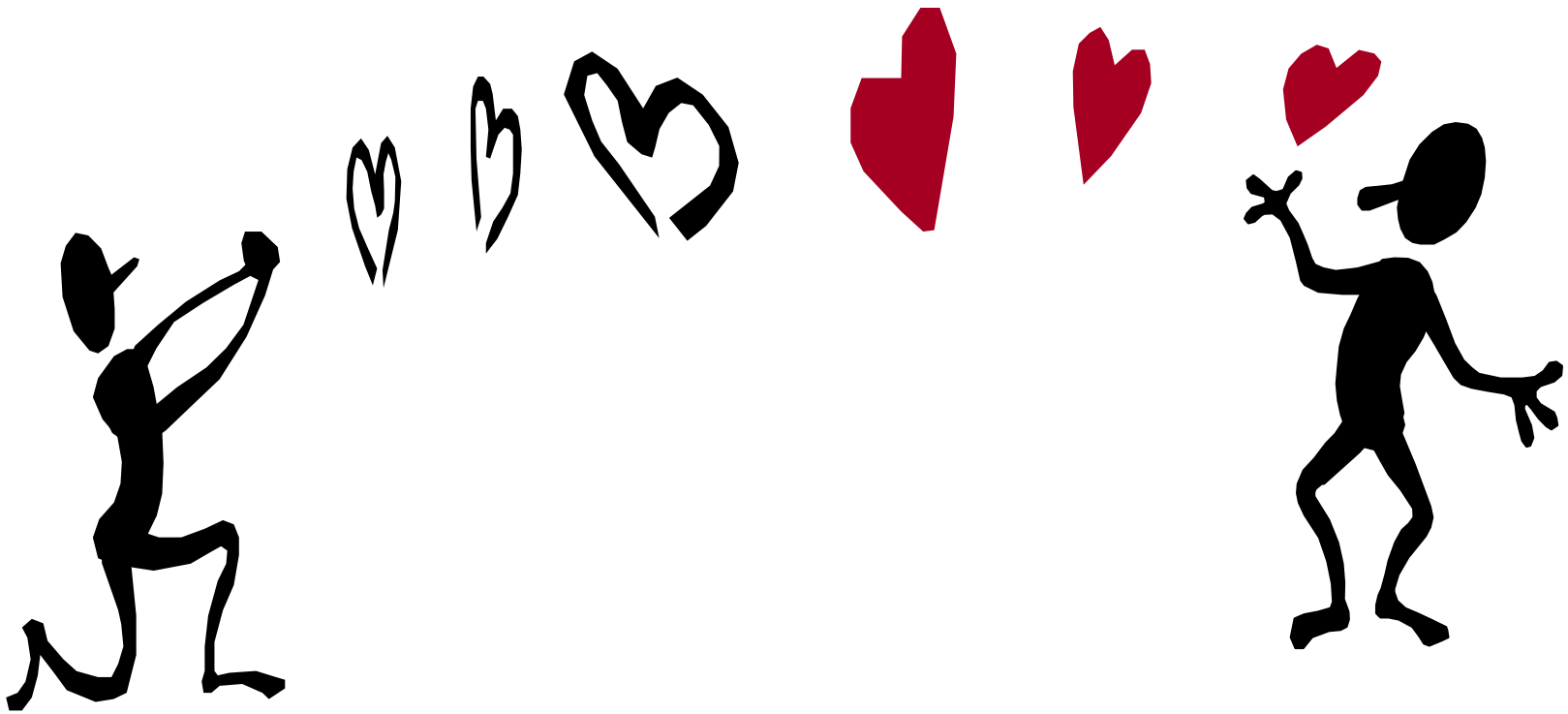
Suppose we pair off all the boys and girls

Now suppose that some boy and some girl prefer each other to the people to whom they are paired

They will be called a rogue couple



Why be with them when we can be with each other?



What use is fairness, if it is not stable?

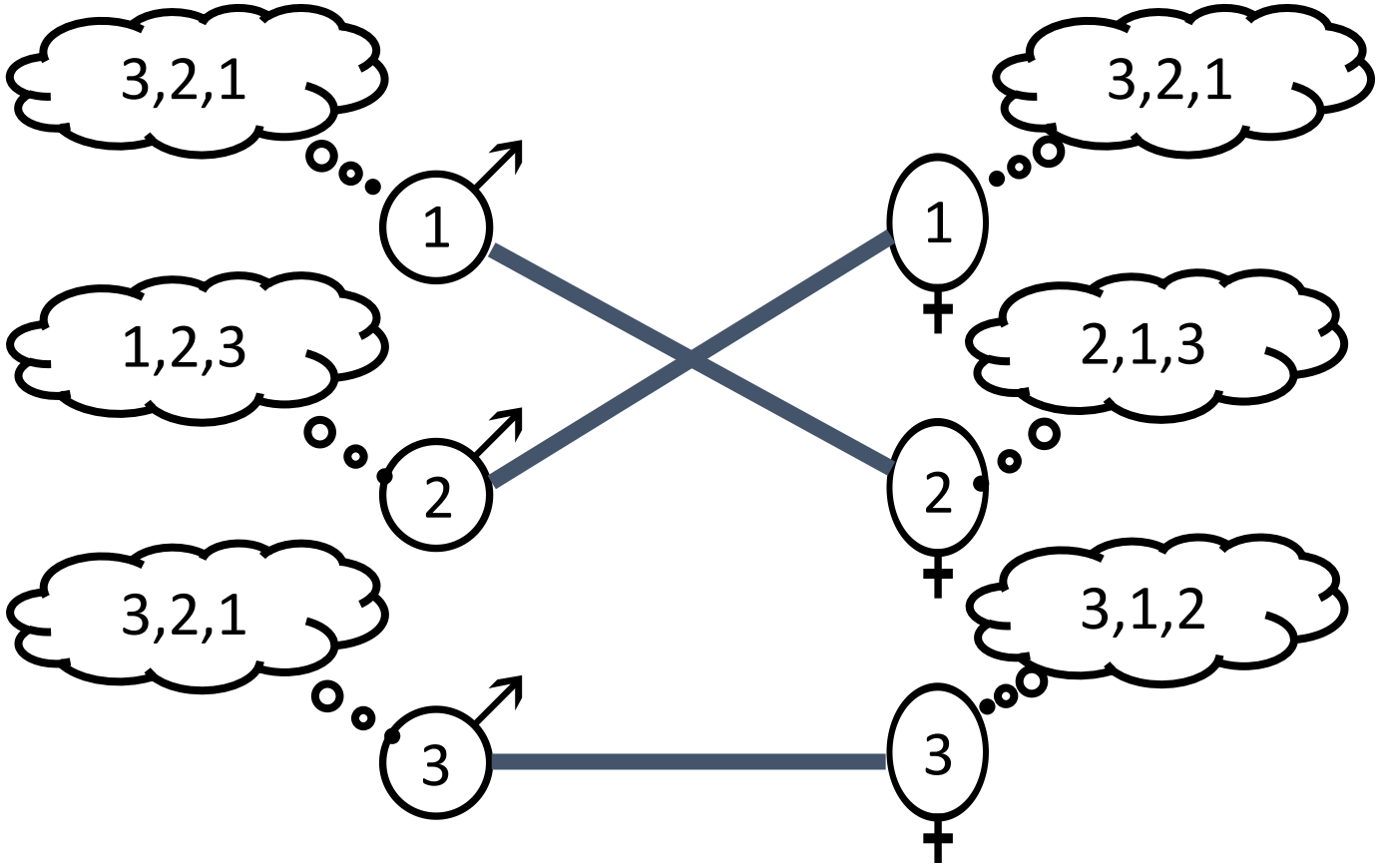
Any list of criteria for a good pairing must include stability. (A pairing is doomed if it contains a rogue couple)

Stable Pairings

A pairing of boys and girls is called stable if it contains no rogue couples

Stable Pairings

A pairing of boys and girls is called stable if it contains no rogue couples



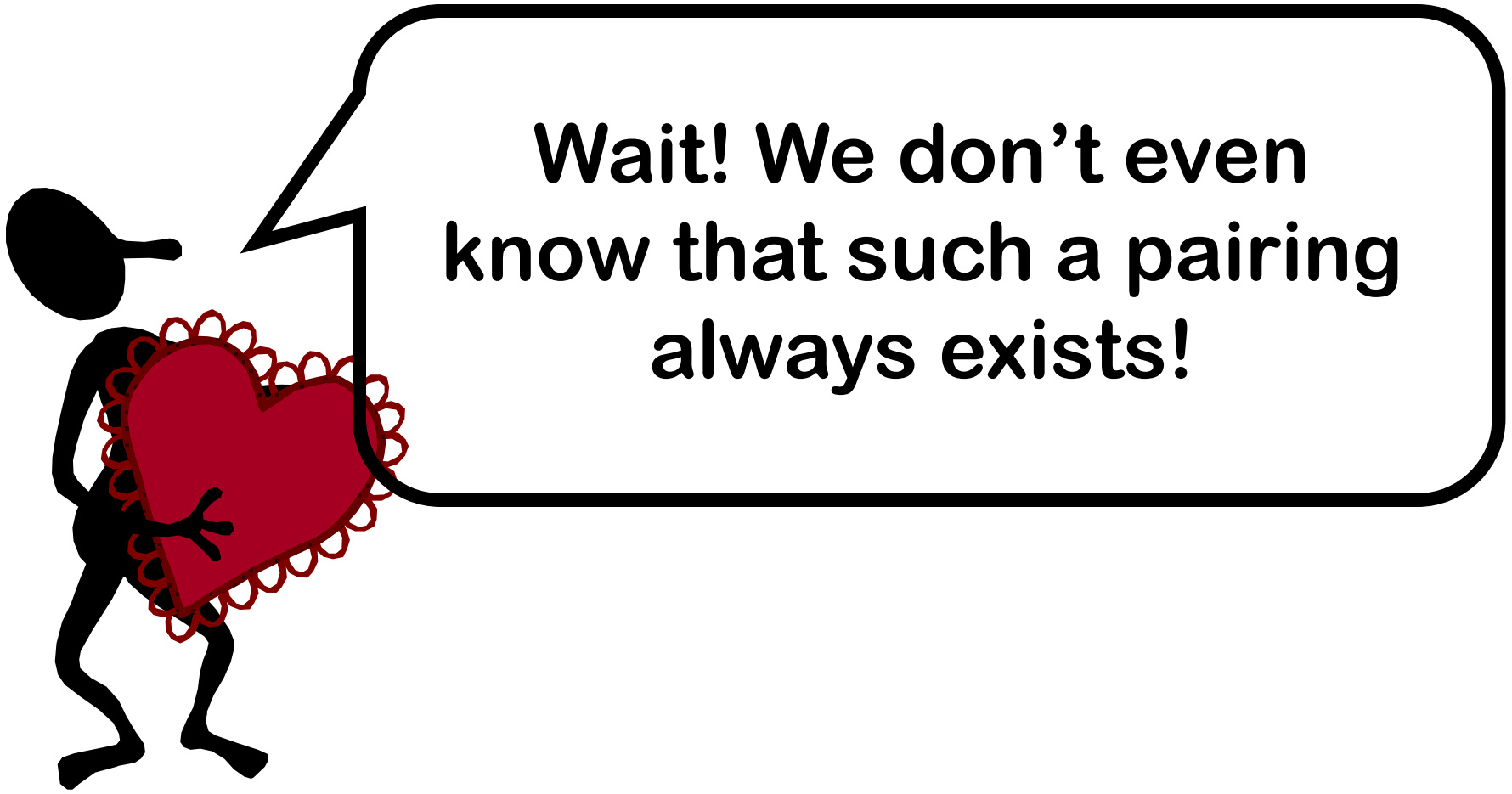
The study of stability will be the subject of the rest of the lecture

We will:

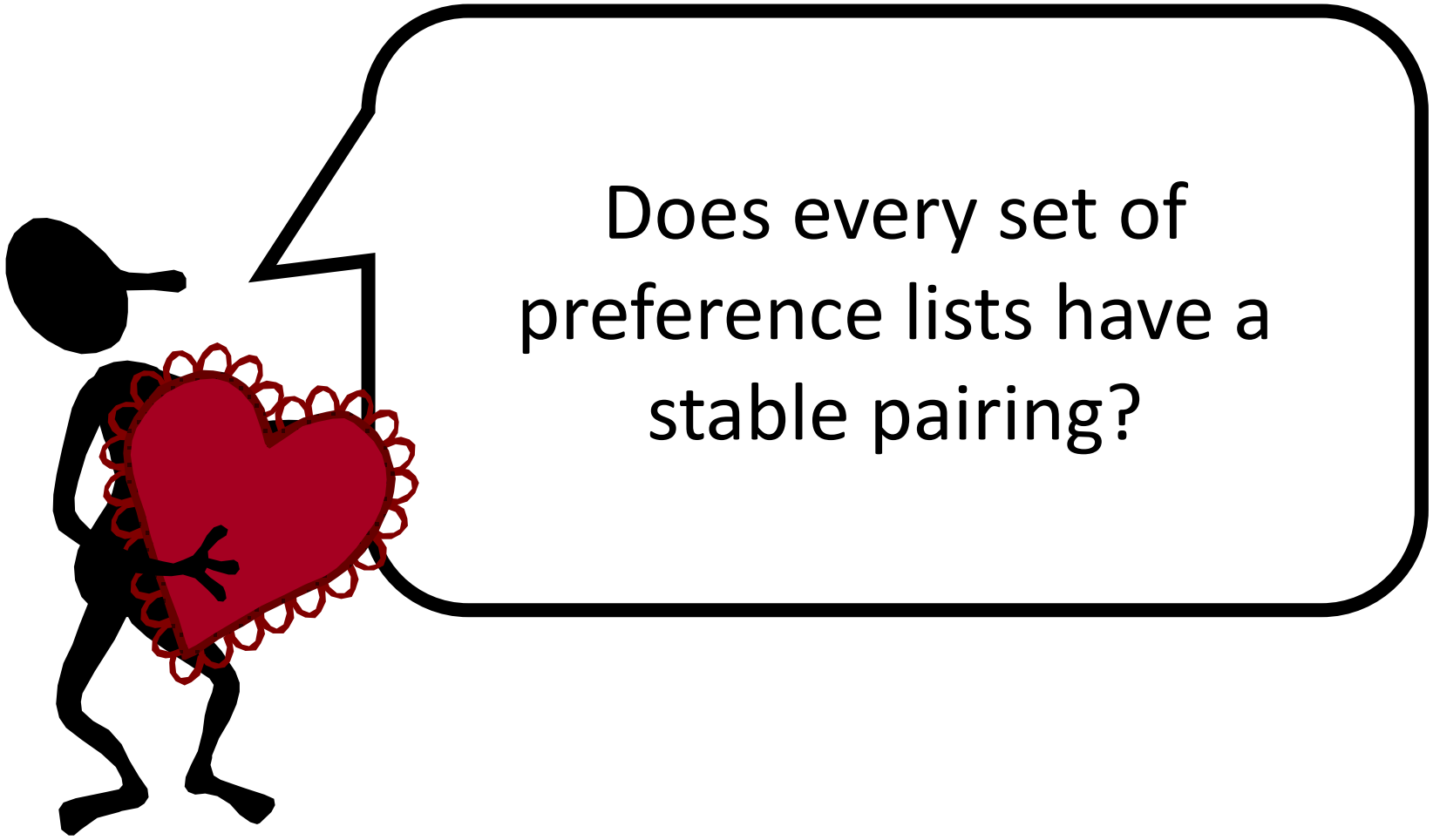
Analyze various mathematical properties of an algorithm that looks a lot like 1950s dating

Discover the naked mathematical truth about which sex has the romantic edge

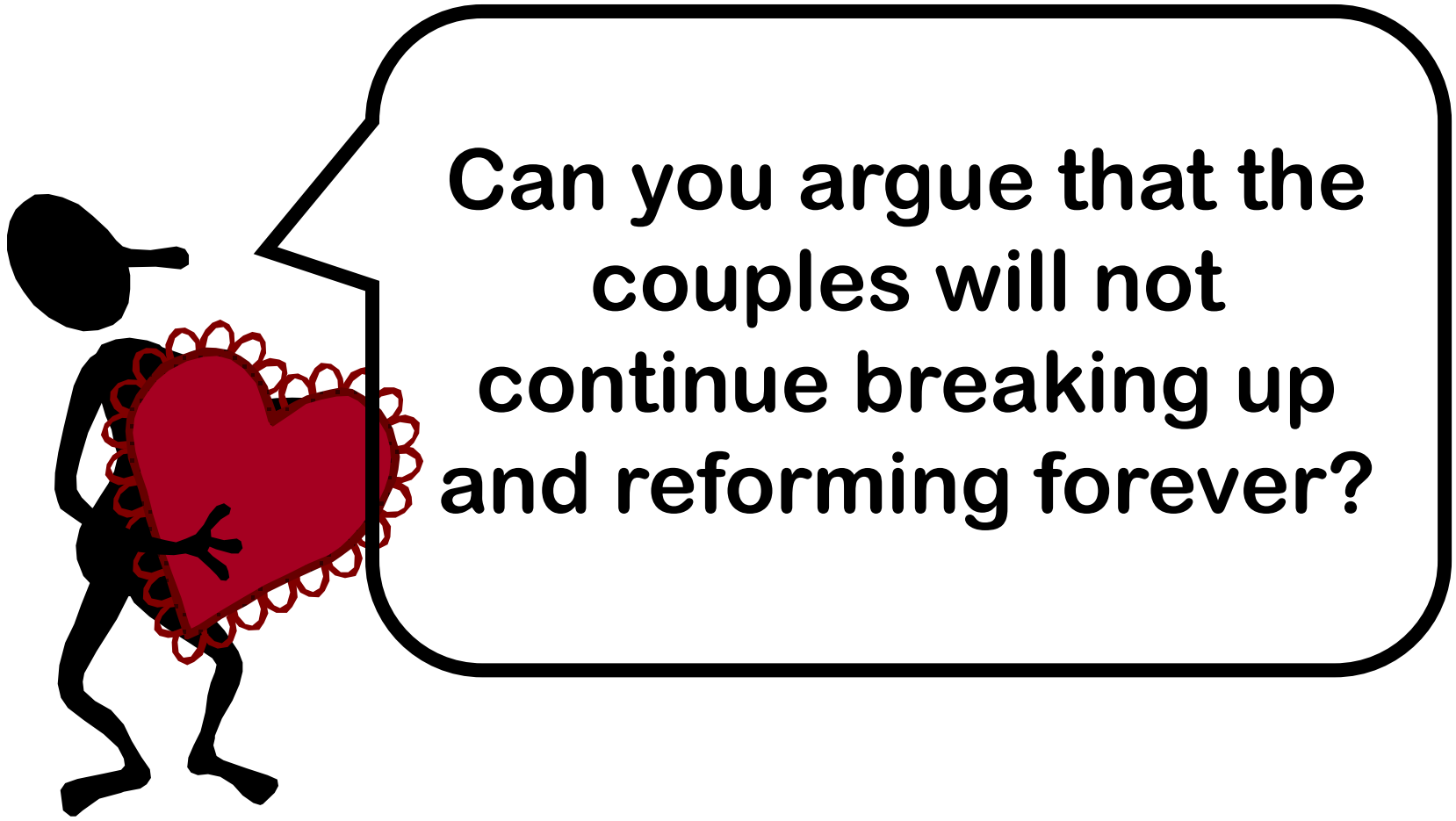
Given a set of preference lists, how do we find a stable pairing?



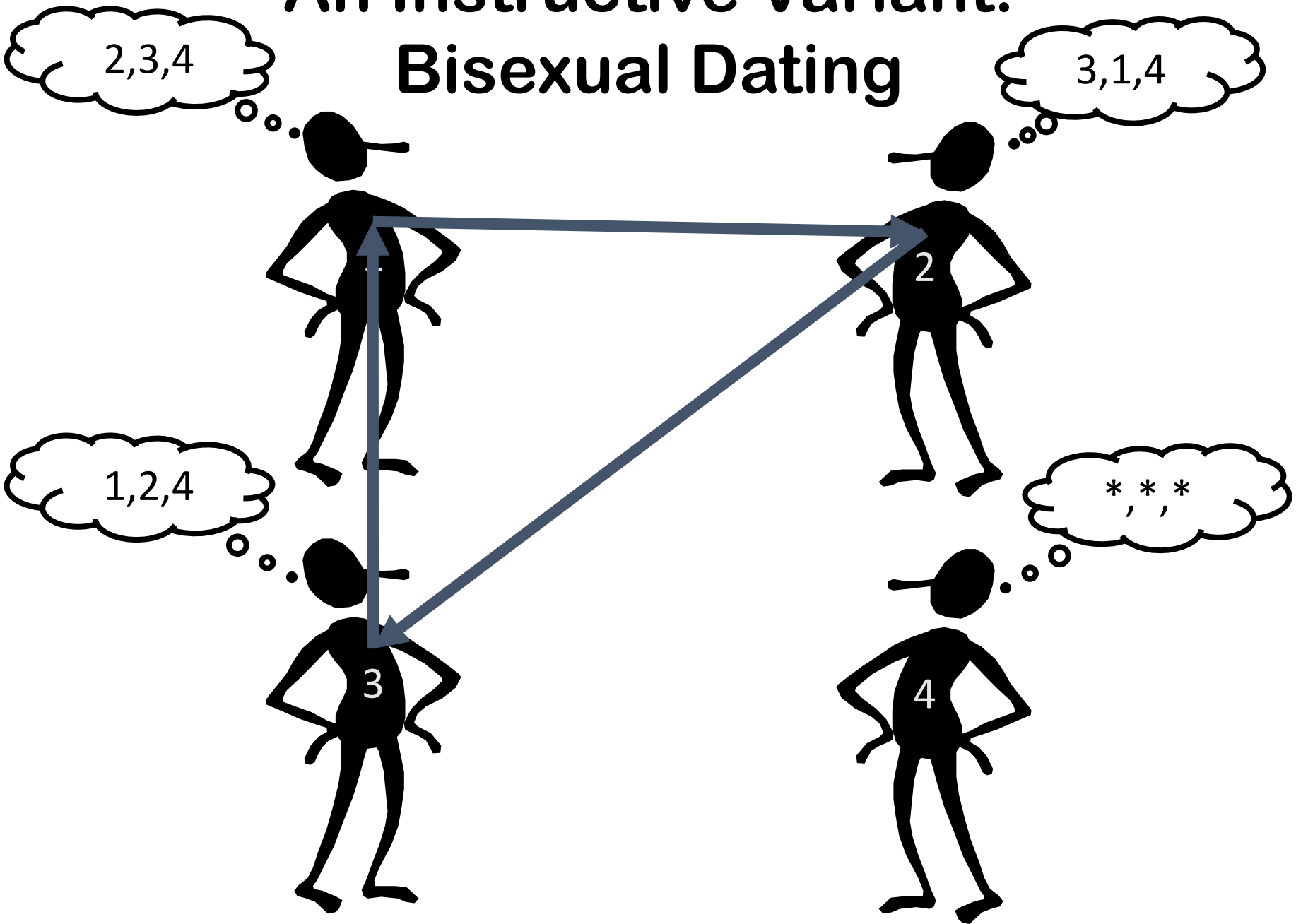
Better Question:



Idea: Allow the pairs to keep breaking up and reforming until they become stable


Can you argue that the couples will not continue breaking up and reforming forever?

An Instructive Variant: Bisexual Dating

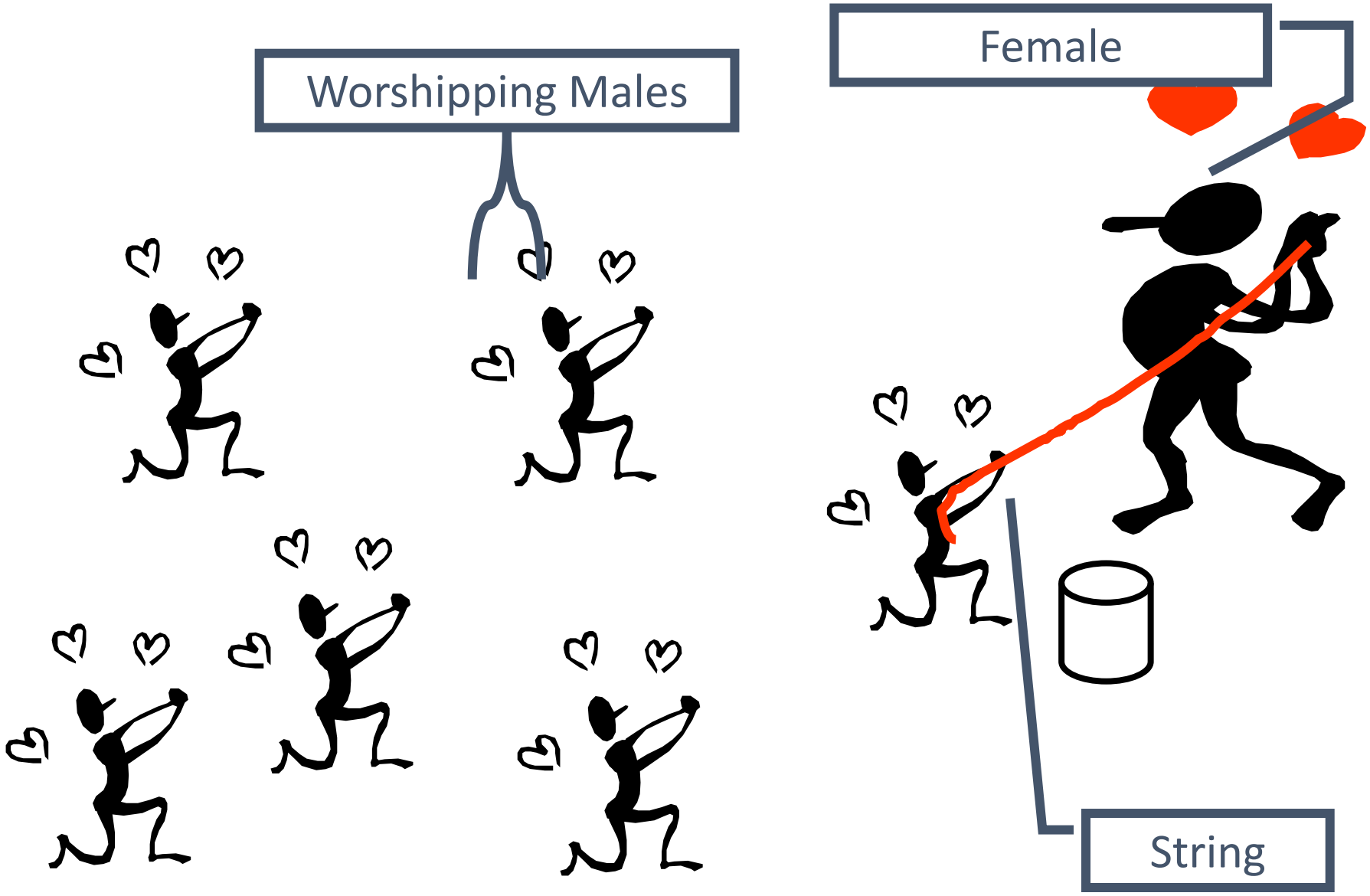


Insight

Any proof that heterosexual couples do not break up and re-form forever must contain a step that fails in the bisexual case

If you have a proof idea that works equally well in the hetero and bisexual versions, then your idea is not adequate to show the couples eventually stop

The Traditional Marriage Algorithm



The Traditional Marriage Algorithm

For each day that some boy gets a "No" do:

Morning

- Each girl stands on her balcony
- Each boy proposes to the best girl whom he has not yet crossed off

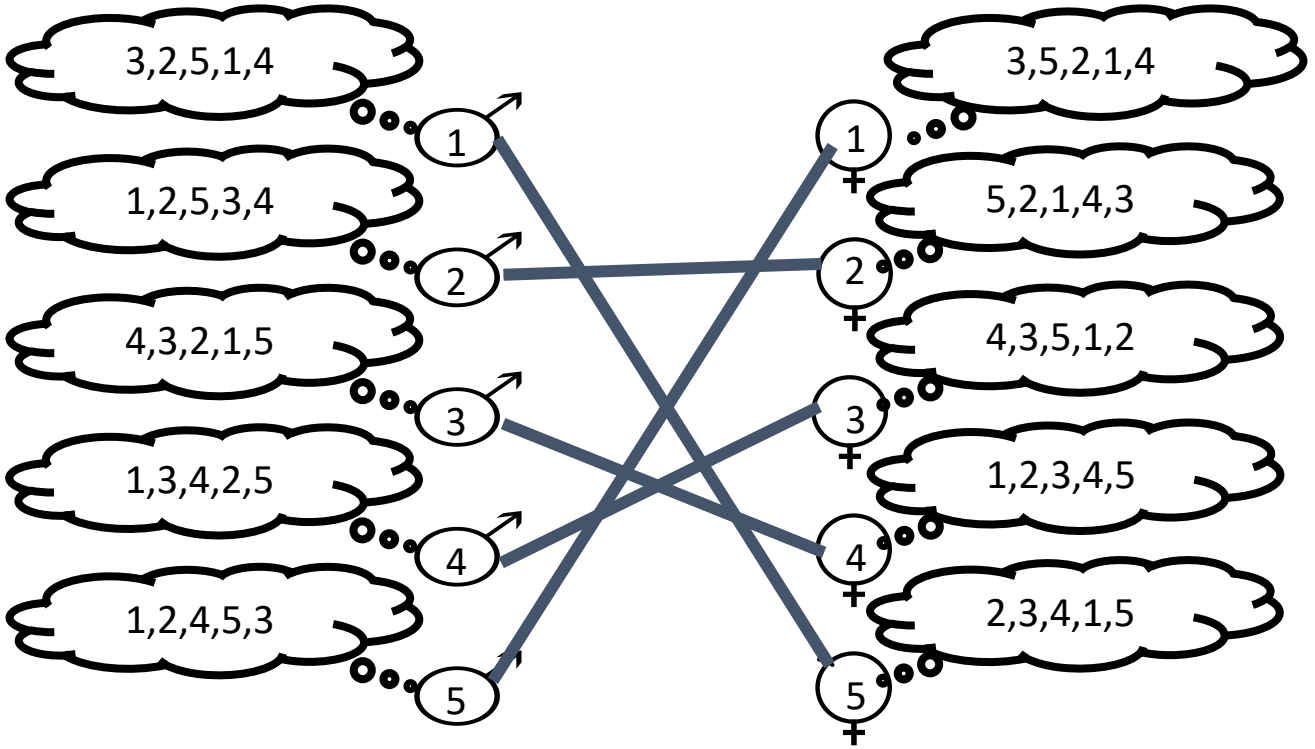
Afternoon (for girls with at least one suitor)

- To today's best: "Maybe, return tomorrow"
- To any others: "No, I will never marry you"

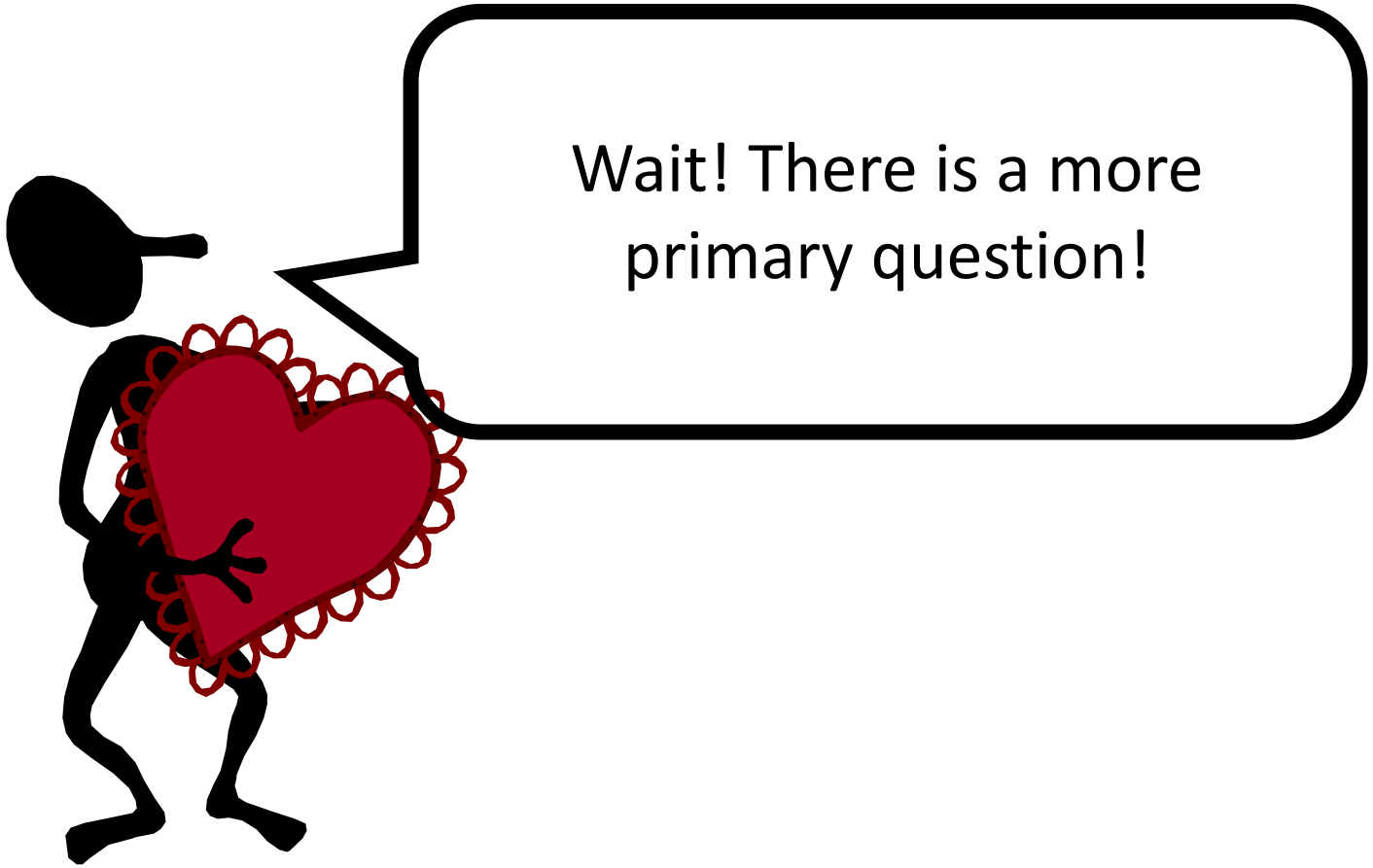
Evening

- Any rejected boy crosses the girl off his list

Termination: If no boys get a "No", each girl marries boy to whom she just said "maybe"


3,2,5,1,4
1,2,5,3,4
4,3,2,1,5
1,3,4,2,5
1,2,4,5,3
1
2
3
4
5
3,5,2,1,4
5,2,1,4,3
4,3,5,1,2
1,2,3,4,5
2,3,4,1,5
1
2
3
4
5

Does Traditional Marriage Algorithm always produce a stable pairing?



Does TMA Always Terminate?

It might encounter a situation where algorithm does not specify what to do next (e.g. “core dump error”)

It might keep on going for an infinite number of days

Improvement Lemma:

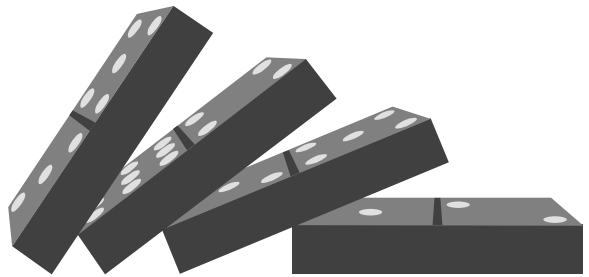
If a girl has a boy on a string, then she will always have someone at least as good on a string next (or for a husband)

She would only let go of him in order to "maybe" someone better

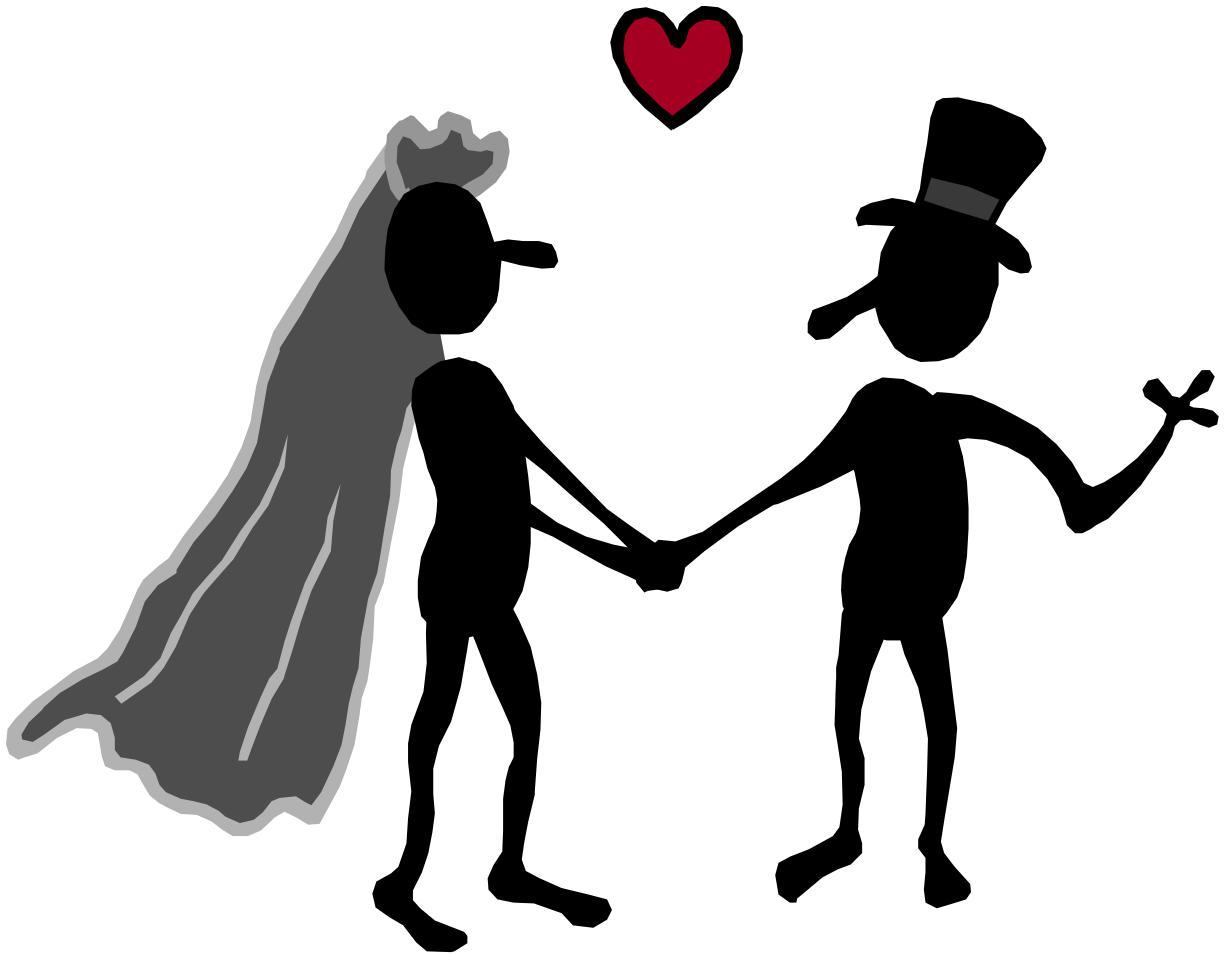
She would only let go of that guy for someone even better

She would only let go of that guy for someone even better

AND SO ON…



Corollary: Each girl will marry her absolute favorite of the boys who visit her during the TMA



Lemma: No boy can be rejected by all the girls

Proof (by contradiction):

Suppose boy *b* is rejected by all the girls

At that point:

Each girl must have a suitor other than b

(By Improvement Lemma, once a girl has a suitor she will always have at least one)

The *n* girls have *n* suitors, and *b* is not among them. Thus, there are at least n+1 boys

Theorem: The TMA always terminates in at most n^2 days

A “master list” of all n of the boys lists starts with a total of $n \times n = n^2$ girls on it

Each day that at least one boy gets a “No”, so at least one girl gets crossed off the master list

Therefore, the number of days is bounded by the original size of the master list

Great! We know that TMA will terminate and produce a pairing

But is it stable?

Conclusion of the Day

- We have met
- You have been notified with the syllabus
- We played a combinatoric game