HARMONY SEARCH

Music Inspired Meta Heuristic Algorithm

OVERVIEW

"A New Heuristic Optimisation Algorithm: Harmony Search"

Zong Woo Geem, Joong Hoon Kim and G.V. Loganathan

SIMULATION 2001 - SAGE

- Original Concept
- Example Applications
- Evaluation and Comparison
- Applying to Feature Selection
- Possible Improvements

FANCY DEFINITION

Harmony Search is the improvisation process of musicians.

During which, each **musician** plays a **note** for finding a **best harmony** all together.

REWORDED VERSION

Harmony Search is a meta heuristic algorithm trying to find a vector that minimises a certain cost function.

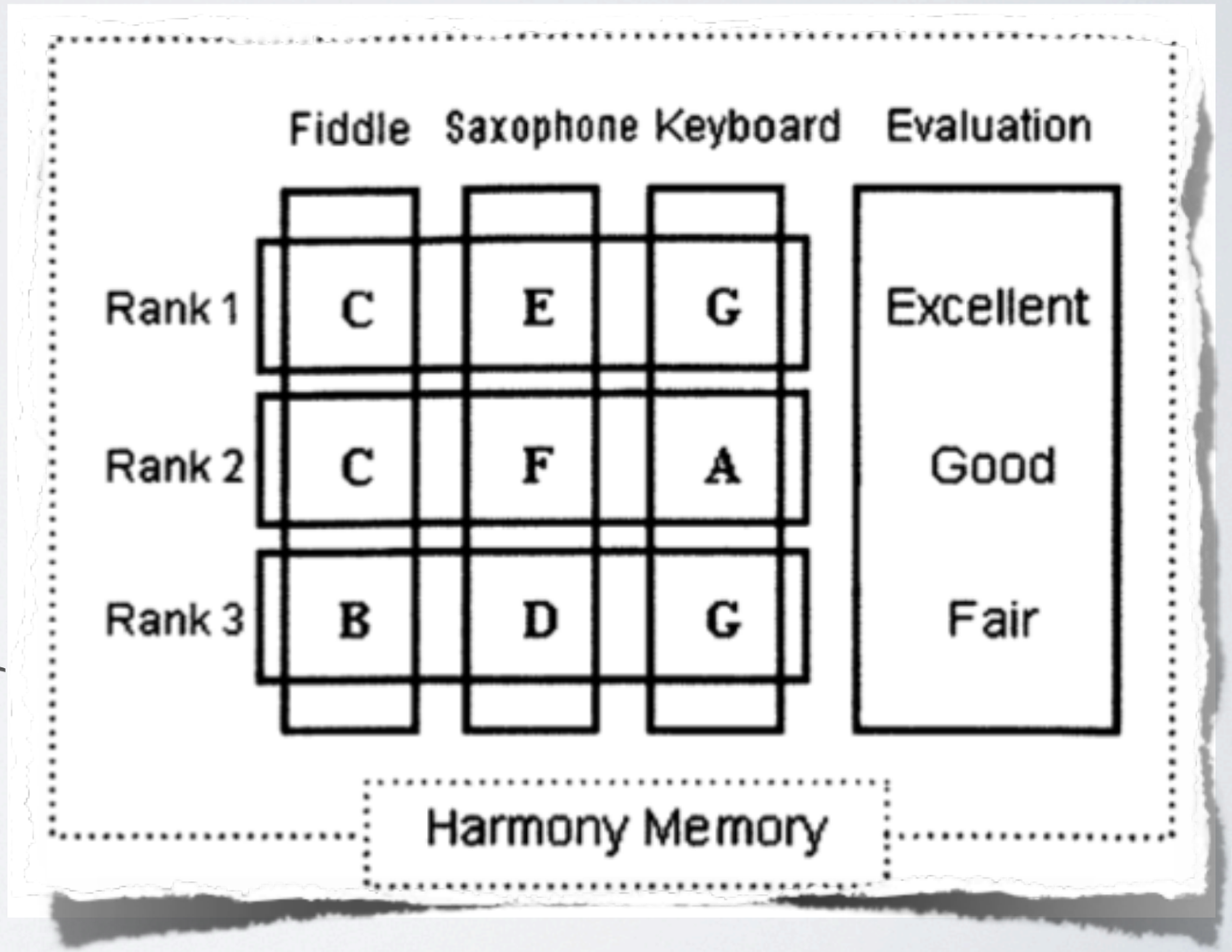
In the process, each **decision variable** generates a **value** for finding a **global optimum** all together.

Harmony Search is the improvisation process of musicians.

During which, each **musician** plays a **note** for finding a **best harmony** all together.

KEY COMPONENTS

- Musicians (Variables)
- Notes (Values)
- Harmony (Set of Values)
- Harmony Memory (Set of Harmonies)



ITERATION STEPS

- Initialise a Harmony Memory
- Improvise a New Harmony
- Include Better Harmony and Exclude Worse
- Repeat until Stop Criteria is Satisfied

ADDED TWISTS

- Harmony Memory Considering Rate
  - ranges from 0 to 1
  - determines if a musician should find notes randomly within all possible range or pick from notes in the harmony memory

- Pitch Adjusting Rate
  - ranges from 0 to 1
  - determines if a musician should shift to neighbouring notes or play as normal

A SIMPLE EXAMPLE

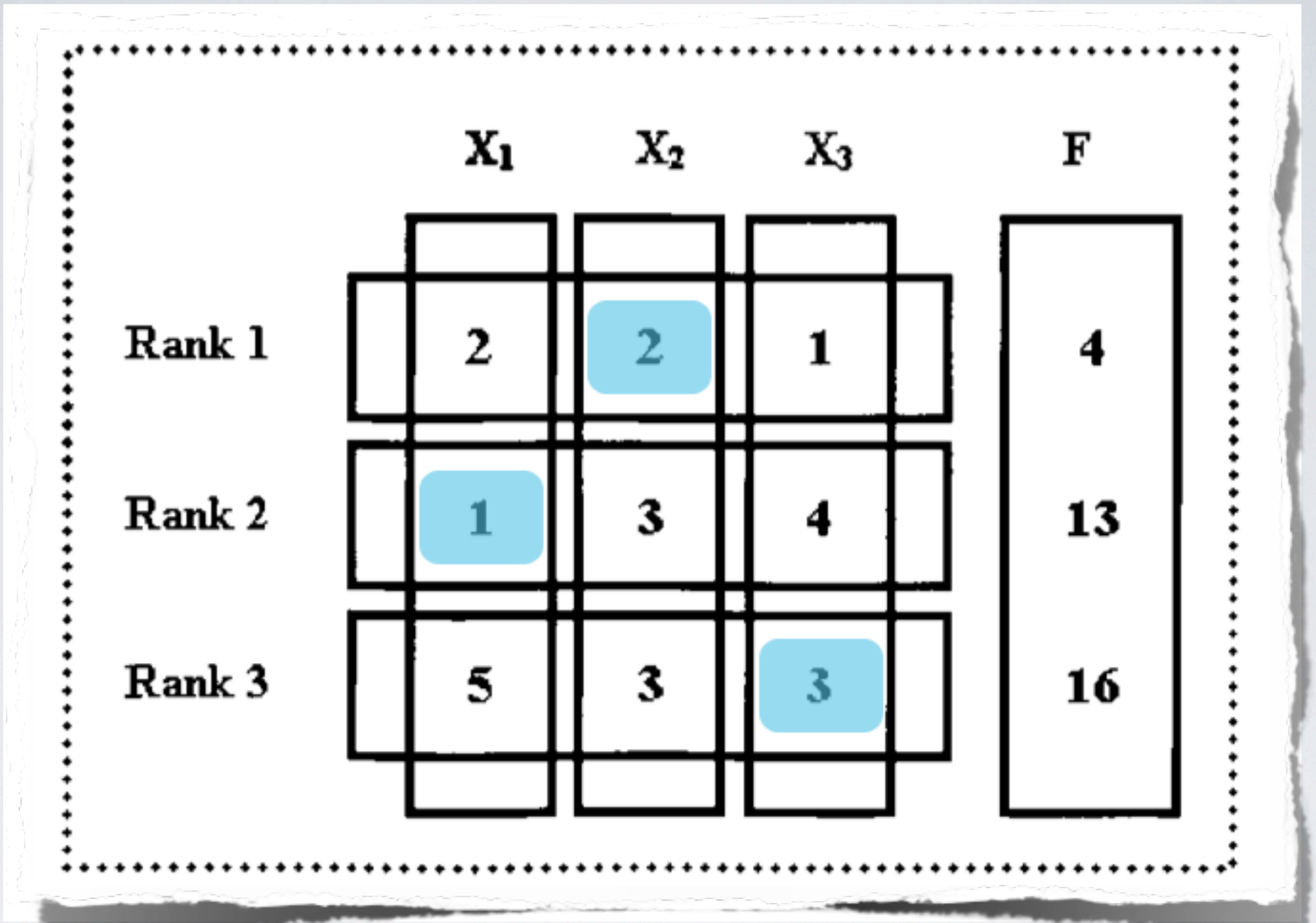
- Minimise function: (a-2)^2 + (b-3)^4 + (c-1)^2 + 3

- We know the answer is (a=2, b=3, c=1)
- Let the available choices be (1,2,3,4,5) for each variables
- Harmony Memory size = 3

	X_1	X_2	X_3	F
Rank 1	2	2	1	4
Rank 2	1	3	4	13
Rank 3	5	3	3	16

Random Selected Notes

INITIALISATION



Each Musician Picks a Note from Harmony Memory

PICK NOTES

	X_1	X_2	X_3	F
Rank 1	2	2	1	4
Rank 2	1	2	3	9
Rank 3	1	3	4	13

Include Better Harmony and Exclude Worst (5,3,3)

NEW HARMONY

LATER IMPROVEMENTS

"Improved Harmony Search from Ensemble of Music Players"

Zong Woo Geem

Knowledge-Based Intelligent Information & Engineering Systems (KES) Part-2 2006 - Springer

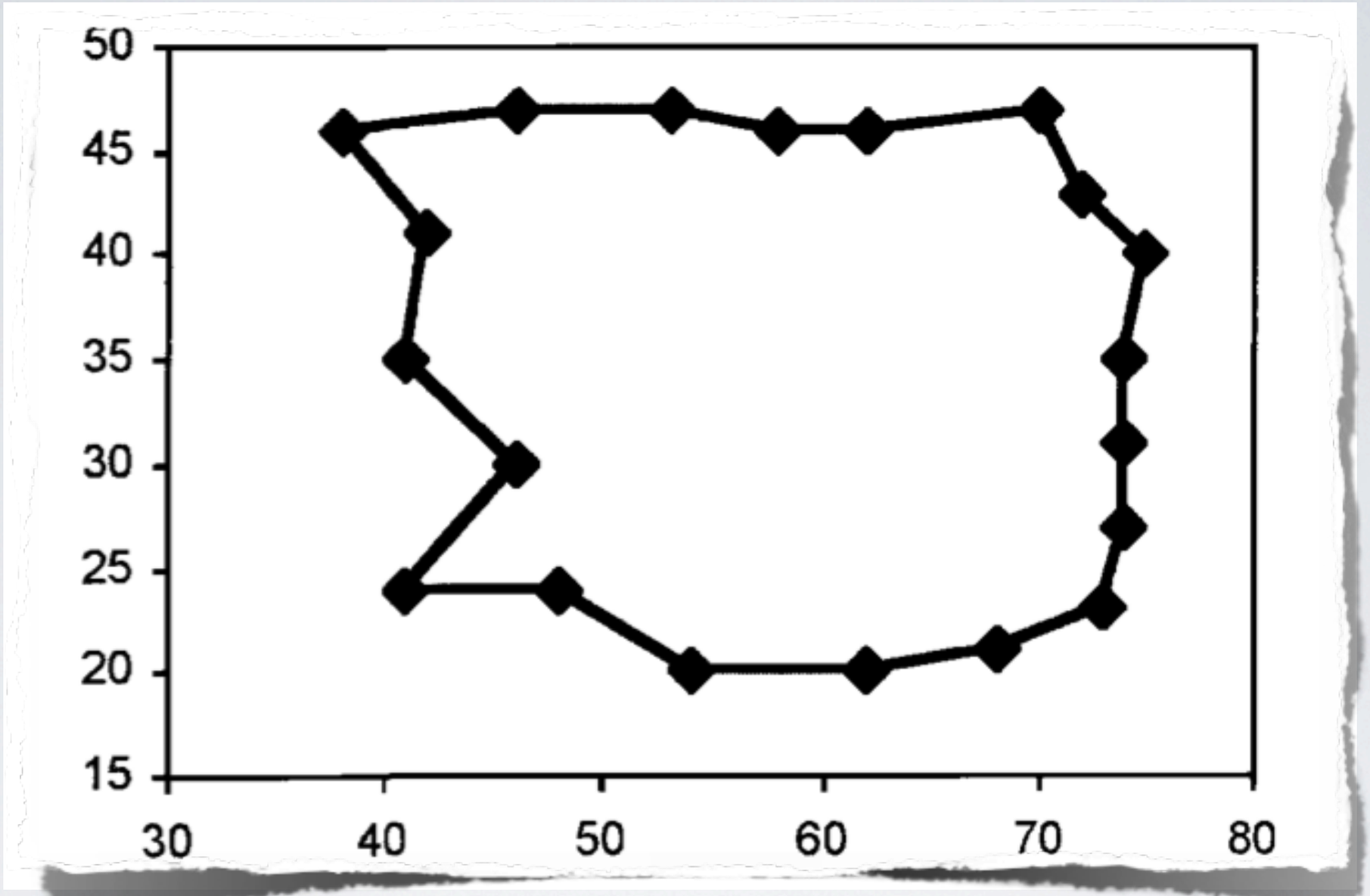
- Ensemble Consideration
  - enforce a pair wise relationship between musicians
  - combine closely related variables together

- Violated Harmony Consideration
  - include harmony that violates evaluation constraints
  - with penalty applied to score

- Very little detail in the paper

EXAMPLE APPLICATIONS

- Traveling Salesman Problem
- Minimise Continuous Variables
- Water Supply Pipeline Optimisation



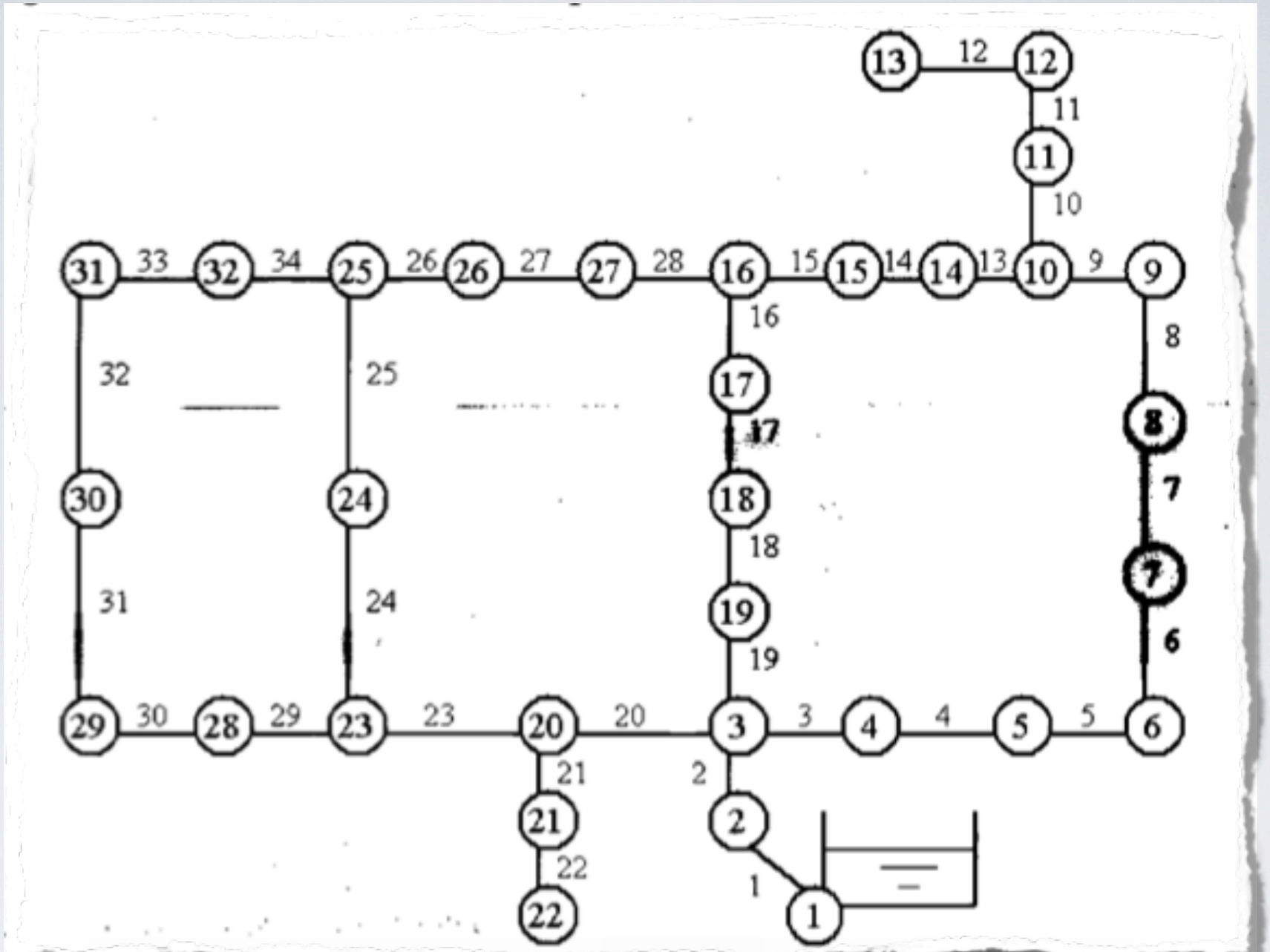
Global Optimum in 7 out of 20 runs - 20,000 iterations with added heuristics choosing better routes*

TRAVELING SALESMAN

	EXACT	GRG	GA	EP	HS(1)	HS(2)
f(x)	1.3935	1.3934	1.4339	1.3772	1.3771	1.3965
%	0.0000	-0.0072	+2.8992	-1.1697	-1.1769	+0.2153
x_1	0.82288	0.8229	0.8080	0.8350	0.8348	0.8290
x_2	0.91144	0.9115	0.8854	0.9125	0.9124	0.9080
g_1	7.05×10^{-9}	1.0×10^{-4}	3.7×10^{-2}	1.0×10^{-2}	1.0×10^{-2}	1.3×10^{-2}
g_2	1.73×10^{-8}	-5.2×10^{-5}	5.2×10^{-2}	-7.0×10^{-3}	-6.7×10^{-3}	3.7×10^{-3}

Min (a-2)^2 + (b-1)^2
where a-2b+1=0, -a^2/4-b^2+1>=0

CONTINUOUS DATA



Least Cost $6,056 - 162,636 iterations

Non-linear Programming $6,320 Genetic $6,073

WATER PIPELINES

EVALUATION PROS

- Very simple algorithm
- Easy to implement
- Global best heuristic
- Produce new solution based on all existing solutions found

EVALUATION CONS

- No good stop criteria
  - max iteration bound

- Hard to set parameters
  - data gives no hint to good settings

- Costly evaluations
  - for entire solution all together

- Can but not handling continuous data well
  - continuous range being discretised into small steps

COMPARISON TO GENETIC ALGORITHM

- Many similar concepts
  - random selection v.s. mutation
- Considers all existing vectors rather than only two parents
- Not sensitive to initial values
- Faster convergence

APPLYING TO FEATURE SELECTION

- Existing implementations
  - Particle Swarm
  - Genetic Algorithm
  - Ant Colony
  - Simulated Annealing

- Can already see analogies in feature selection

- Plug and play approach, a search strategy only

HORIZONTAL APPROACH CONCEPT MAPPING

- Map musicians to attributes
- Map notes to 0 or 1
- Harmony is represented as a series of bits
-

A	B	C	D	E	F
0	1	1	0	0	0

= (B,C)

HORIZONTAL APPROACH ITERATION STEPS

A	B	C	D	E	F	Score
I	I	I	0	0	0	0.9
I	0	I	0	0	I	0.7
I	I	0	0	0	0	0.3
I	***0***	***I***	***I***	***0***	***I***	***0.4***

to be excluded

- Initialise Harmony Memory
- Pick a new subset (A, C, D, F)
- Evaluate, include if better and exclude the worst
- Repeat until max iteration

AntColonyOptimisation

GeneticAlgorithm

ParticleSwarmOptimisation

HorizontalHarmonySearch

Data

WEKA

FuzzyRoughSubsetEvaluator

TEST ENVIRONMENT

	Ant	*HA1*	*HA2*	*PSO*	*GA*
heart	7/13	*8/13*	*8/13*	7/13	7/13
ionosphere	8/34	*10/34*	*11/34*	7/13	10/34
cleveland	8/12	*8/12*	*8/12*	8/12	8/12
olitos	5/13	*5/13*	*6/13*	5/13	6/13
web	out of memory	*1012/2557*	*1119/2557*	187/2557	352/2557
3-completed	6/38	*not responding*	*8/38*	7/38	8/38
wineScaled	5/13	*5/13*	*5/13*	5/13	5/13

HA1: iteration 1000 memorySize 50

HA2: iteration 500 memorySize 20

sub-optimal results

TESTING RESULT

EVALUATION PROS

- It works!
- Simple algorithm
- Easy to implement
- Require little CPU and memory

EVALUATION
CONS

- 3 Parameters to tune
  - Harmony Memory size, max iteration, random selection rate

- Whole subset evaluation at every iteration is expensive
- Impractical for large attribute data (gene)
- Takes longer than standard search
- Very limited note choices (0 and 1)

VERTICAL APPROACH

- Designed to solve some of the obvious problems
- Approach from a more natural and intuitive angle

EXPERT-DECISION ANALOGY

- Musician can be treated as an expert
- Note domain can be generalised into available choices
- Harmony Memory can be seen as a pool of combined decisions

VERTICAL APPROACH CONCEPT MAPPING

- Initialise experts (musicians)
- Initialise available choices to be all attributes (notes) & no vote
E1	E2	E3	E4	E5	E6
A	-	B	B	C	-

  = (A, B, C)

VERTICAL APPROACH ITERATION STEPS

E1	E2	E3	E4	E5	E5	Score
D	A	-	-	A	C	
-	A	-	B	B	C	
B	A	D	B	C	D	
B	***A***	**-**	***D***	***A***	***D***	

- Initialise decision space (Harmony Memory)
- Experts make decisions, combine into new subset (A, B, D)
- Evaluate, include if better and exclude worst
- Repeat until max iteration

	Ant	*HA1*	*VA1*	*VA2*	*PSO*	*GA*
heart	7/13	8/13	***8/13***	***7/13***	7/13	7/13
ionosphere	8/34	10/34	***8/34***	***8/34***	7/13	10/34
cleveland	8/12	8/12	***9/12***	***8/12***	8/12	8/12
olitos	5/13	5/13	***6/13***	***5/13***	5/13	6/13
web	out of memory	1012/2557	***452/2557***	***425/2557***	187/2557	352/2557
3-completed	6/38	not responding	***7/38***	***7/38***	7/38	8/38
wineScaled	5/13	5/13	***5/13***	***5/13***	5/13	5/13

VA1: iteration 100 memorySize 10

VA2: iteration 250 memorySize 25

quicker & better result

TESTING RESULT

EVALUATION PROS

- It still works!
- Need less iterations to find good results & faster to process
- Bound subset size
  - same as number of experts
  - find global best subset within a given number of attributes
- Max Iteration gives a bound to processing time
  - can roughly predict the scale of improvement
  - if more iterations was used

EVALUATION CONS

- Whole subset evaluation is expensive still
- Even more parameters! 5 now
  - number of experts
  - neighbour selection rate
- Experts are not really experts
- Reluctant in finding smaller subsets

POSSIBLE IMPROVEMENTS

- Auto-adjusted Parameters
- Intelligent Experts
- Weighted Decisions
- Multiple Phases
- Fuzzy and Rough

AUTO-ADJUSTED PARAMETERS

- Better Stop Criteria
  - no. of iterations since last best subset
  - last decision update
  - score improvement rate of new subset and entire space

- Dynamic Expert Size
  - more expert at start, less towards the end

- Dynamic Decision Space
  - more to consider in the beginning, less to choose later

INTELLIGENT EXPERTS

- Each experts can be distinct and independent
- Experts can also cooperate in pairs or groups
- Inject different ideas, preferences, strategies to each experts

WEIGHTED DECISIONS

- Each experts can have different weights
- Each decisions can have different weights
- The new subset is then a result of voting

MULTIPLE PHASES

- Divide into different phases, each having different search strategies and evaluation criteria
  - Initial: random decision, lose evaluation
  - Intermediate: wiser selection, finer evaluation
  - Final: picky experts with greedy selection, tight evaluation

MAKE IT FUZZY AND ROUGH

- There's ***nothing fuzzy*** or ***rough*** involved here
- What is going on?!

QUESTIONS AND IDEAS