Power Indices and Game Theory (Applications to Bioinformatics)

Stefano MORETTI
stefano.moretti@dauphine.fr
LAMSADE (CNRS), Paris Dauphine

von Neumann and Morgenstern, 1944
basic formal language for
modeling economic phenomena.

GAME THEORY

NON-COOPERATIVE THEORY

- Games in extendive form (tree games)
- Games in strategic form (normal form)

COOPERATIVE THEORY

- Games in c.f.f. (TU-games or coalitional games)
- Bargaining games
- NTU-games

➢ Dominant strategies
➢ Nash eq. (NE)
➢ Subgame perfect NE
➢ NE & refinements
…

➢ Core
➢ Shapley value
➢ Nucleolus
➢ τ-value
➢ PMAS
….

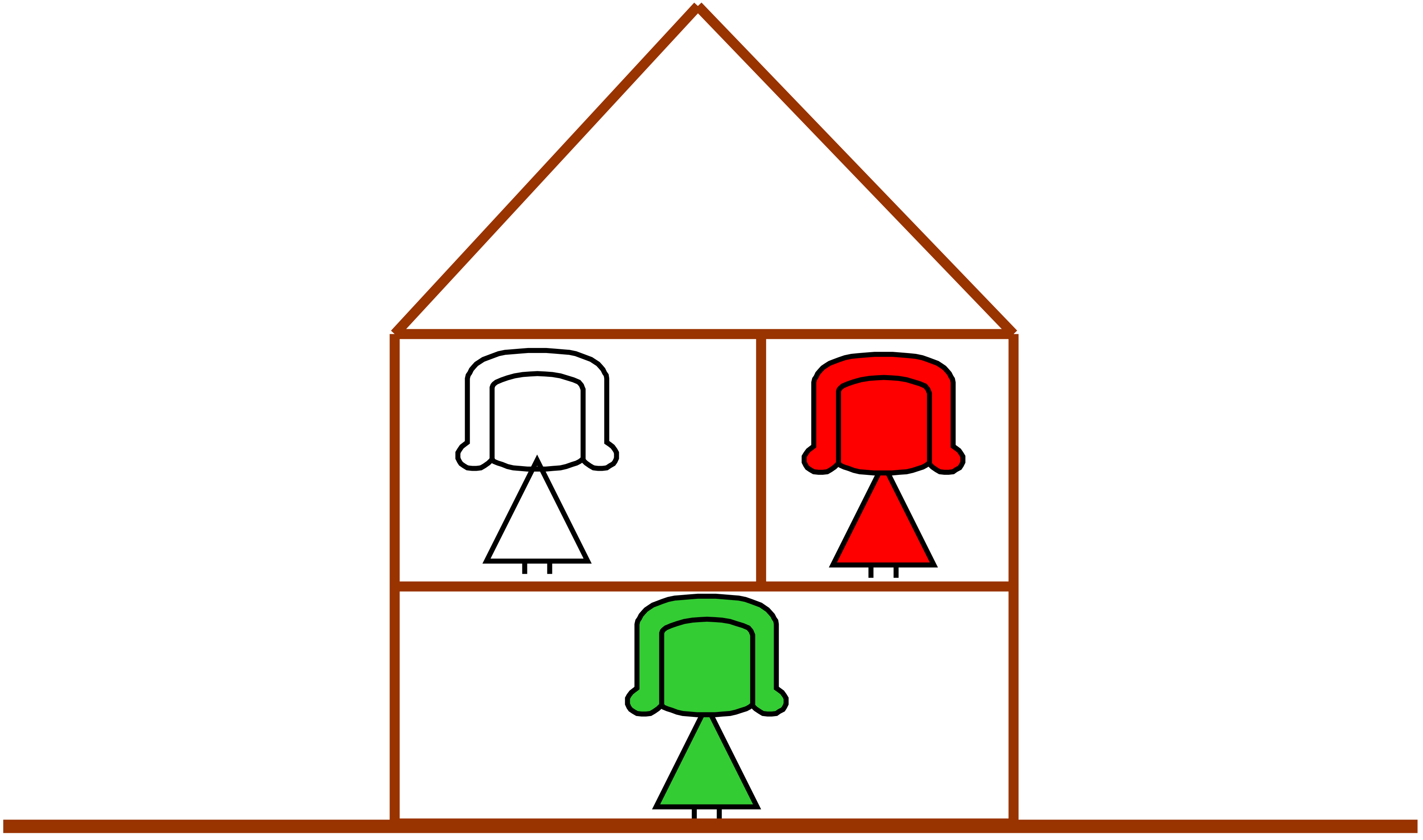
➢ Nash sol.
➢ Kalai-Smorodinsky
….

➢ CORE
➢ NTU-value
➢ Compromise value
…

No binding agreements
No side payments
Q: Optimal behaviour in conflict situations

binding agreements
side payments are possible (sometimes)
Q: Reasonable (cost, reward)-sharing

A building with three owners

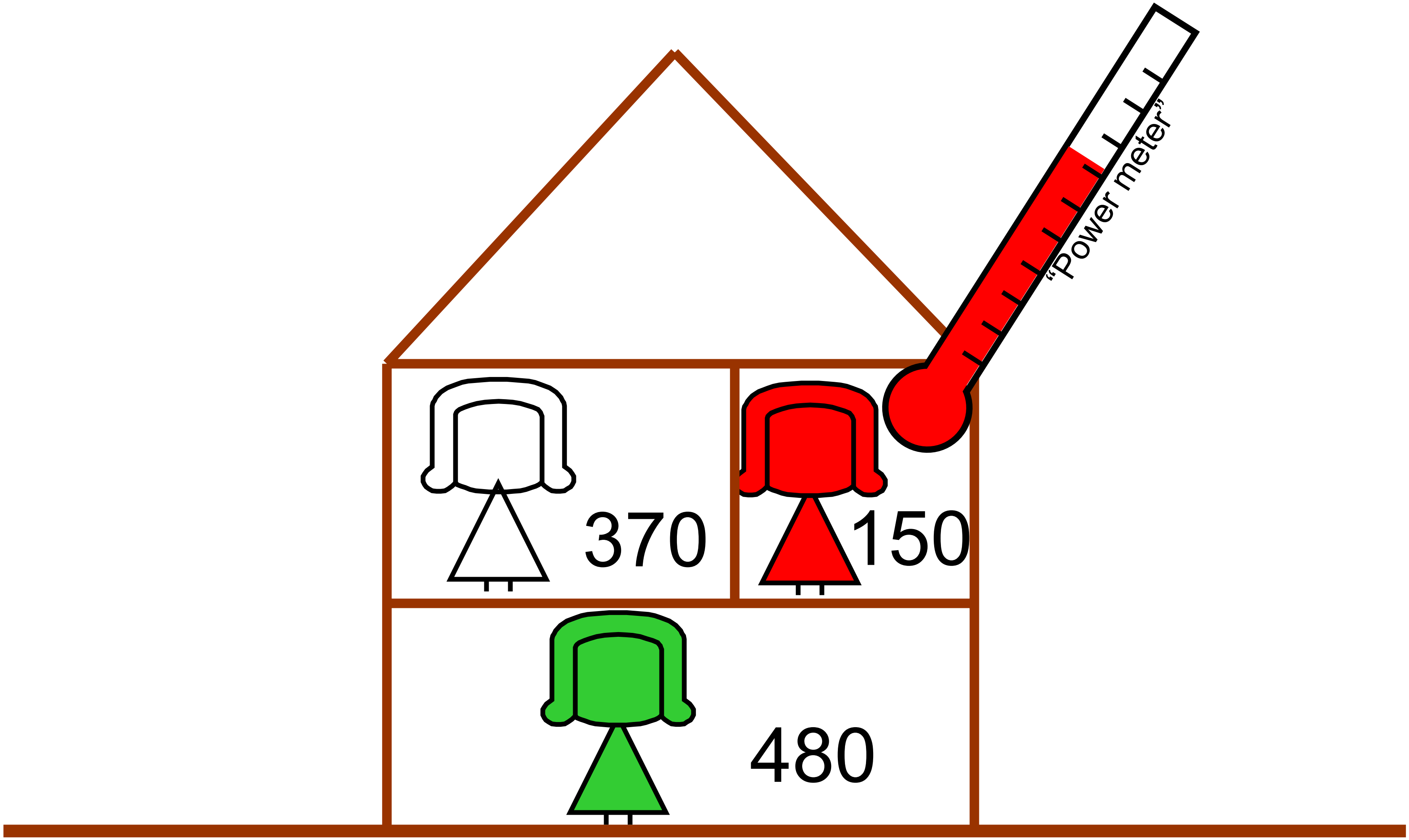


- Each owner has a weight (in thousandths)

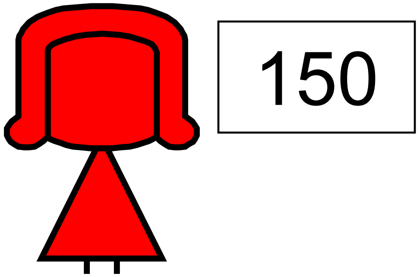
- Decision rule: a group of owners with at least 667 thousandths is *winning* → they may force a decision concerning common facilities (e.g., "to construct an elevator")

- Q: How to measure the power of each owner?

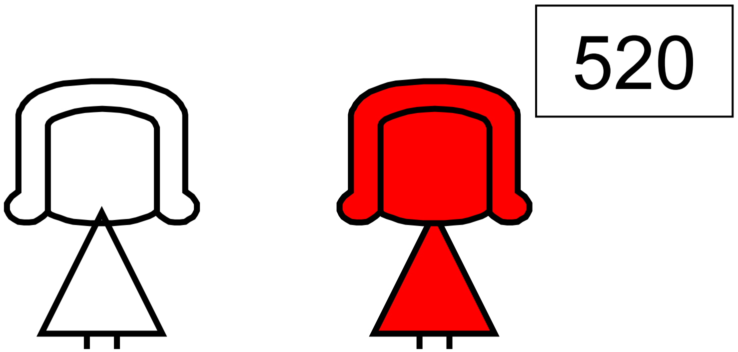
Power index



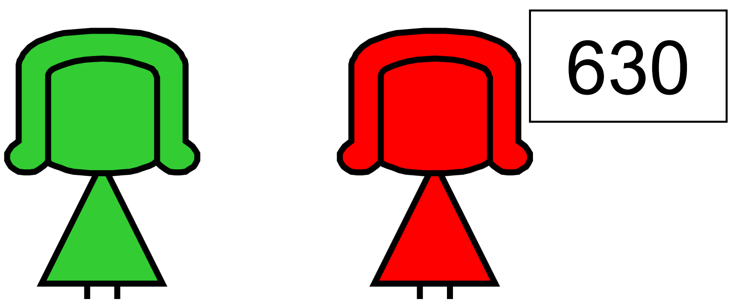
Which properties should a power index satisfy?



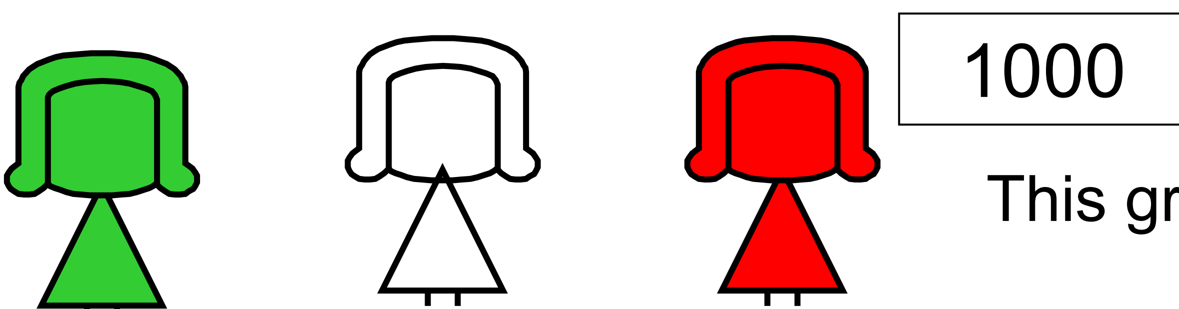
This group has less than 667 thousands



This group has less than 667 thousands



This group has less than 667 thousands

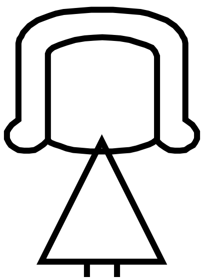


This group has more than 667 thousands

0

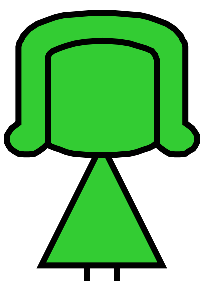
This group has less than 667 thousands

370

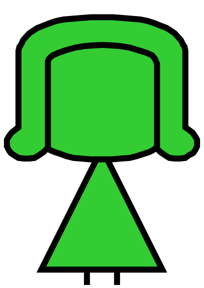


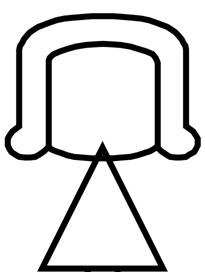
This group has less than 667 thousands

480



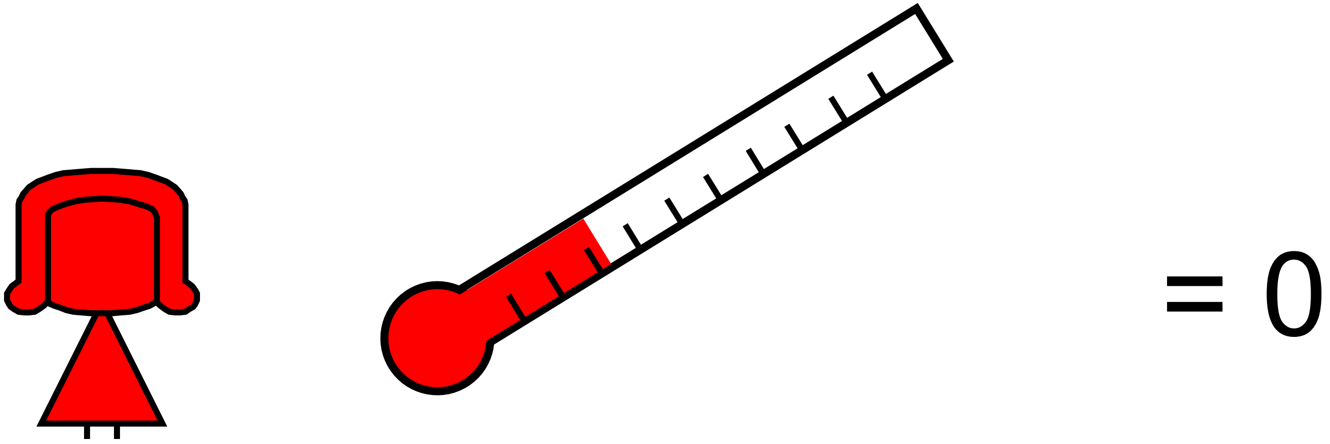
This group has less than 667 thousands





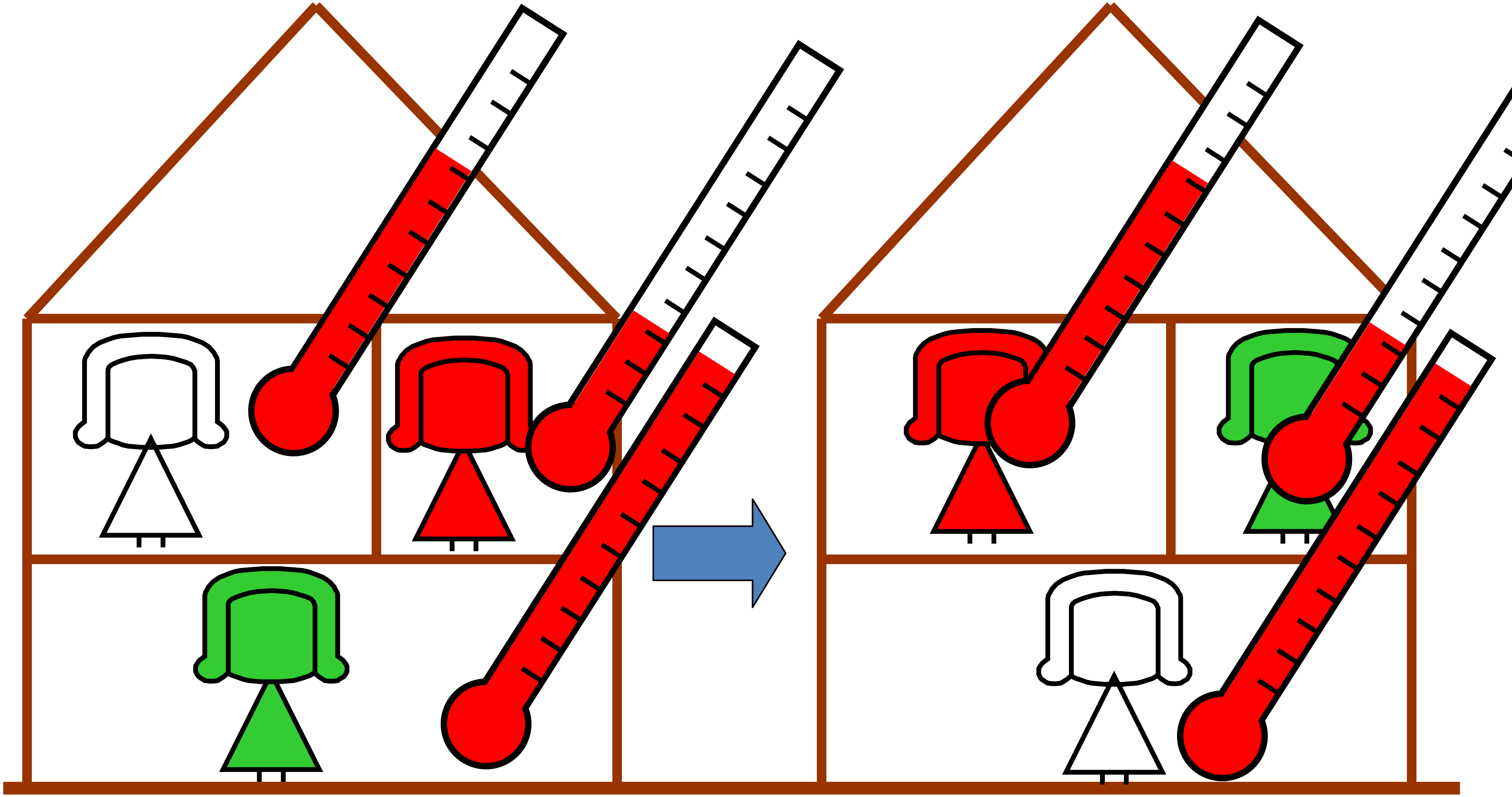
850

This group has more than 667 thousands



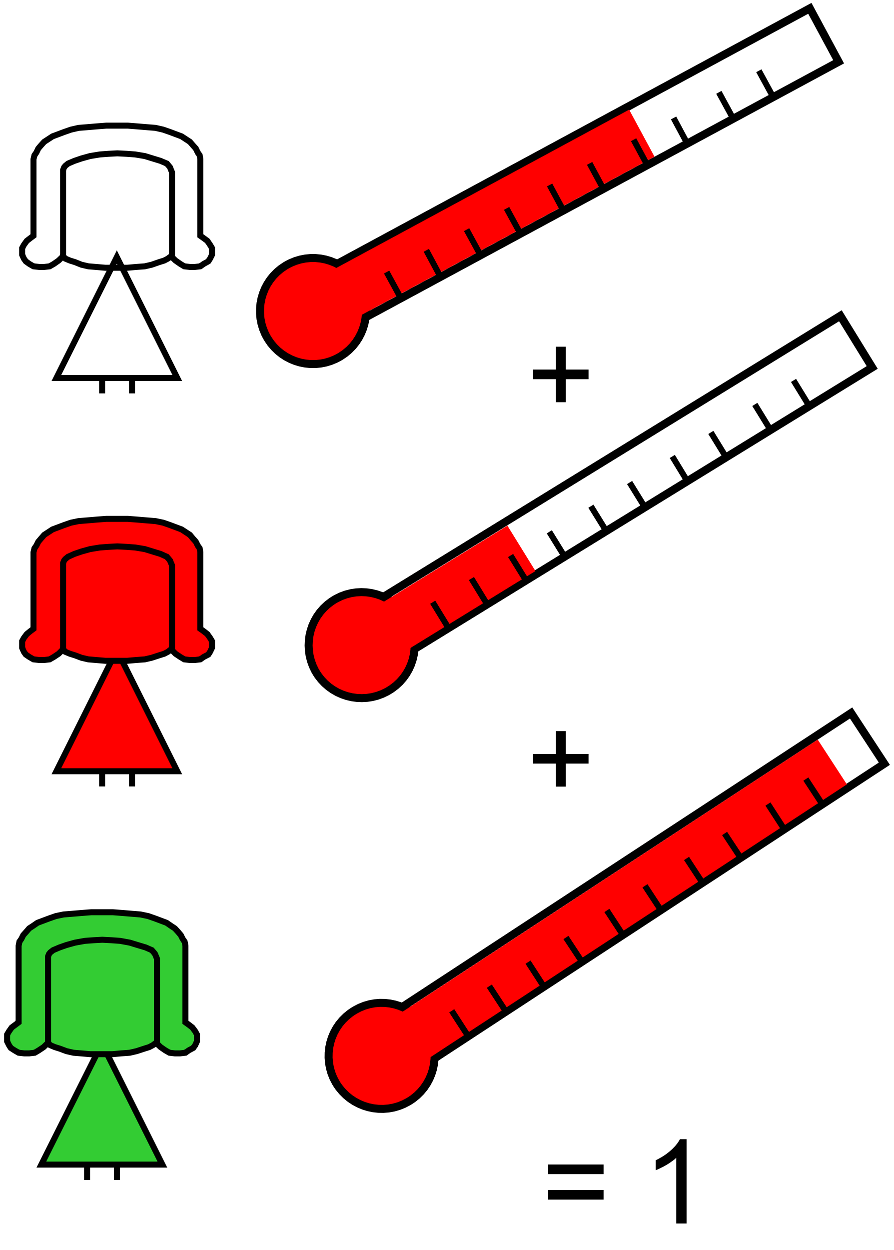
<u>Null player property</u>:

The power of the owners who never contribute to make a winning group must be zero.

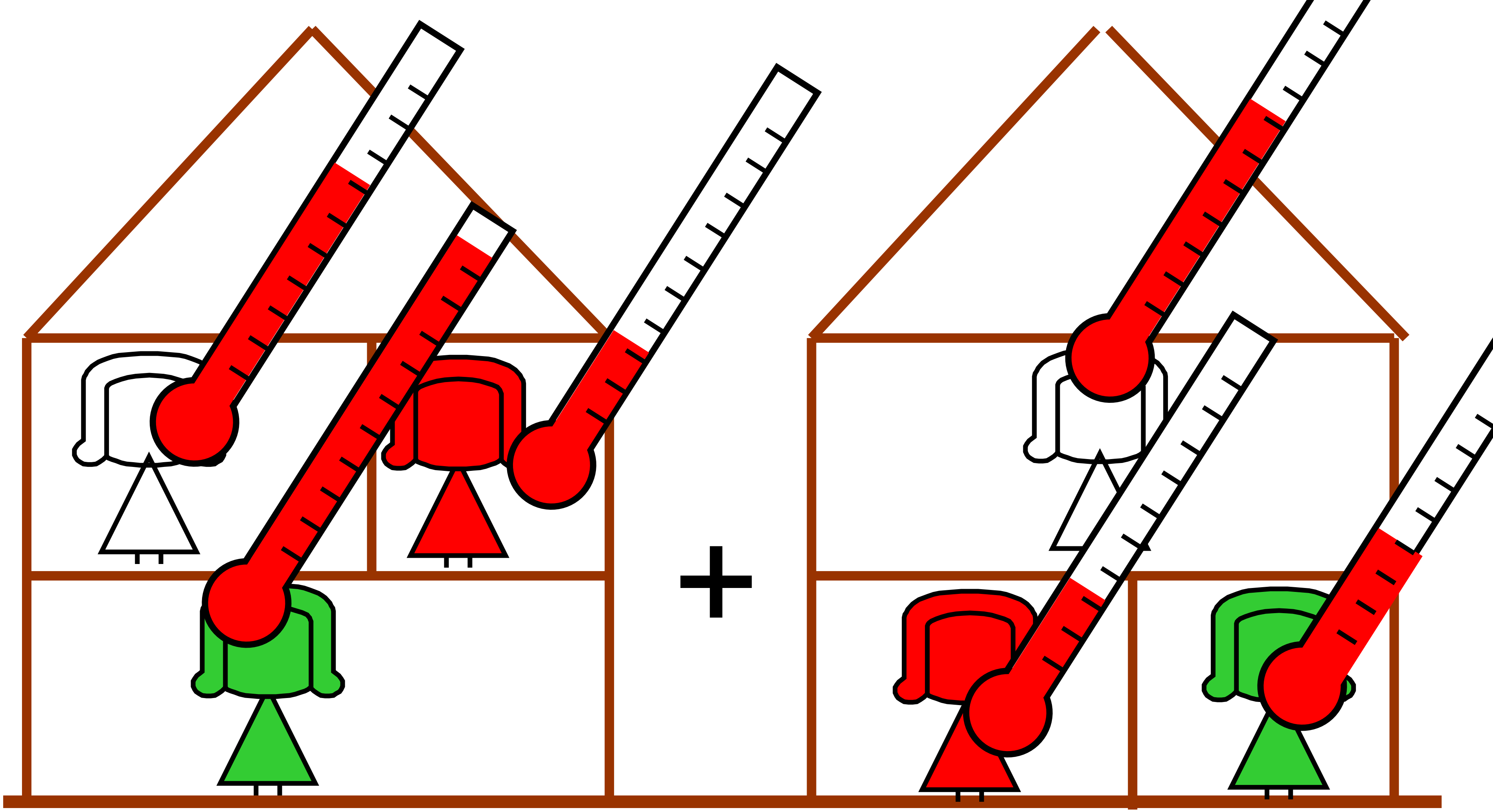


<u>Anonimity property</u>:

The power index should not depend on the names of the owners

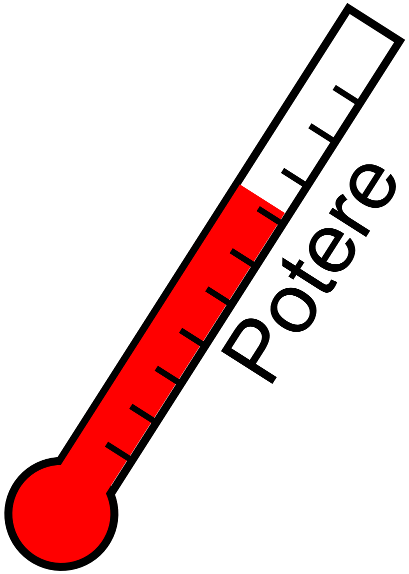


Efficiency property: the sum of the powers must be 1



Transfer property:

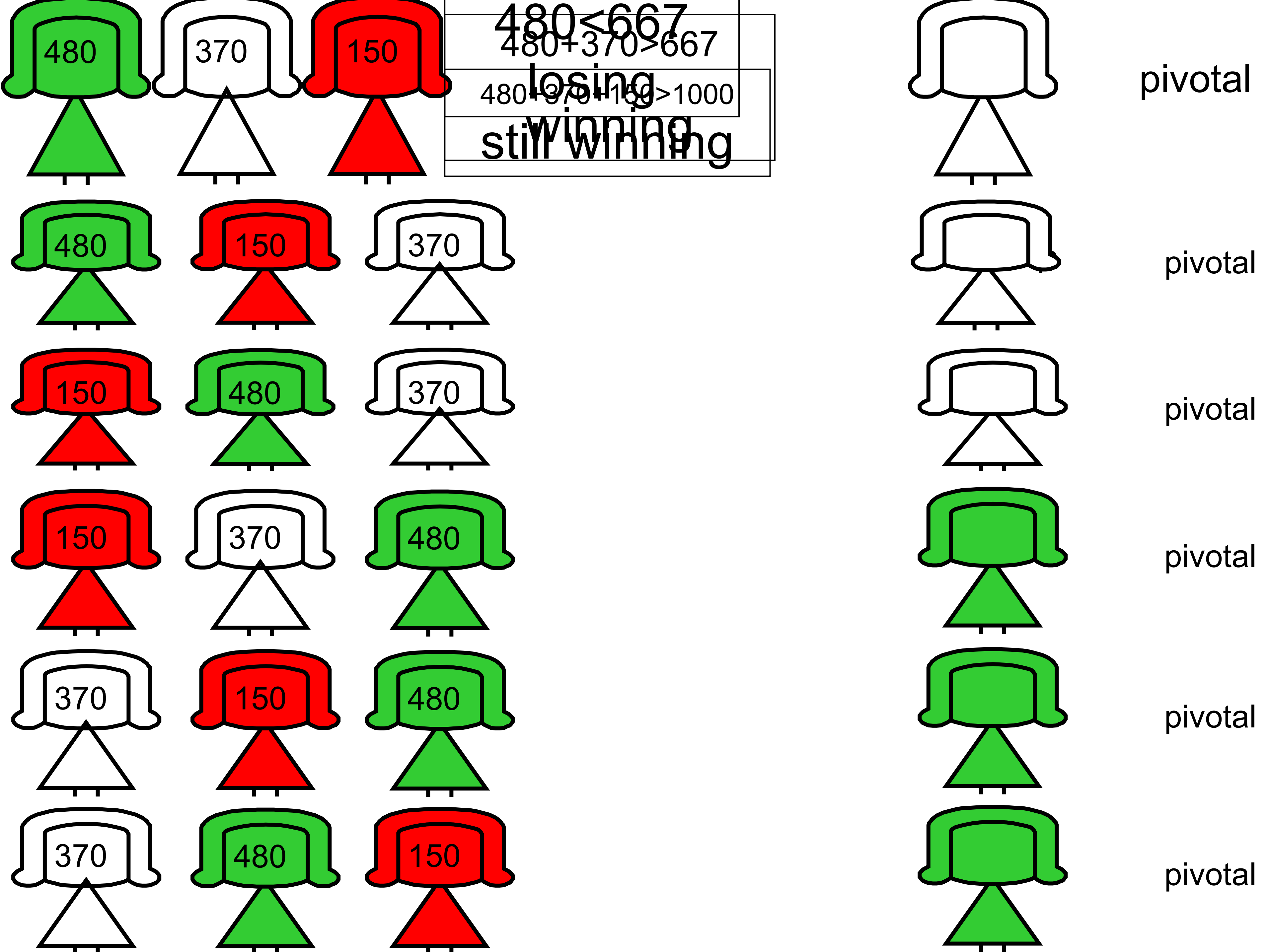
How to sum the power between two different interactive situations…(*see later*)



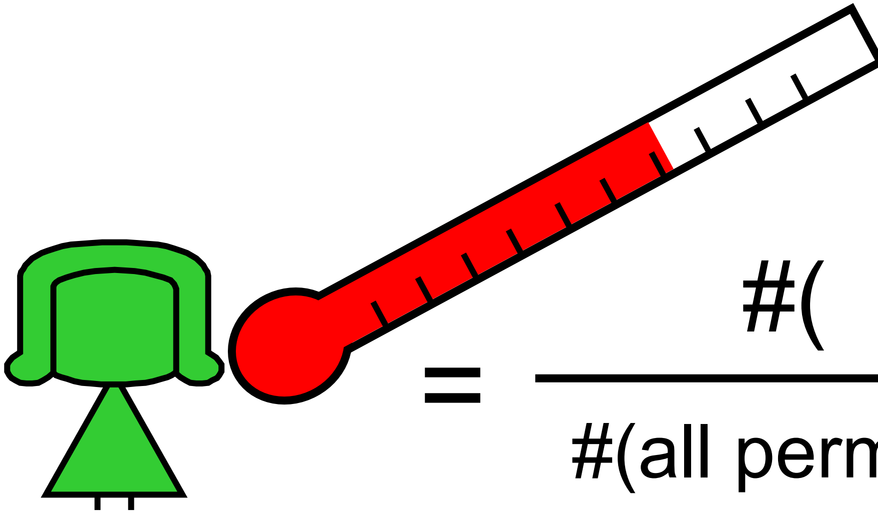
Shapley&Shubik power index (1954)

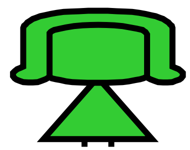
Satisfies *anonymity*, *efficiency*, *null player* and *transfer* properties

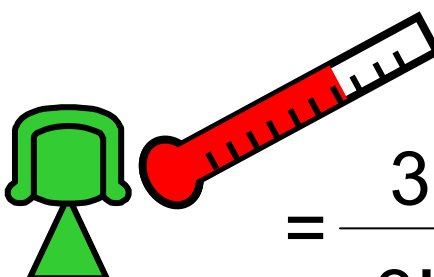
… it is the unique power index which satisfies such properties on the class of simple games…


480
370
150
480<667
480+370>667
losing
480+370+150>1000
winning
still winning
pivotal
480
150
370
pivotal
150
480
370
pivotal
150
370
480
pivotal
370
150
480
pivotal
370
480
150
pivotal

Shapley&Shubik power index (1954)

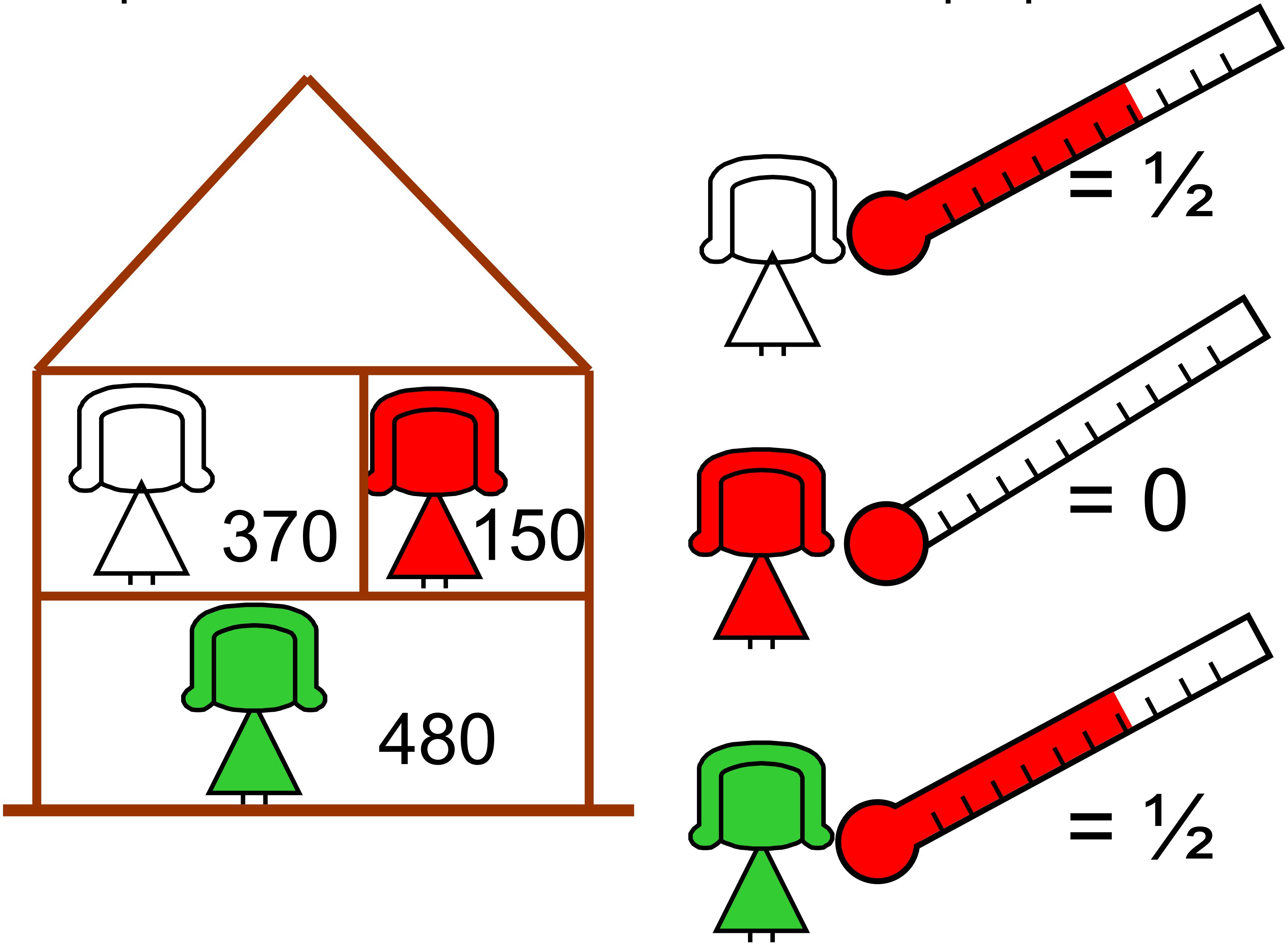


= #( pivotal) / #(all permutations of players)



$$= \frac{3}{3!} = \frac{3}{6} = ½$$

… a power index which satisfies such properties…



Simple games

A simple game is a (voting or similar) situation in which every potential coalition (set of players/voters) can be either *winning* or *losing*.

DEF. A *simple* game is a pair (N,v) where

- N is a finite set (*players set*) and
- v is map (*characteristic function*) defined on the power set 2^N such that
  - $v(S) \in \{0,1\}$ for each *coalition* $S \in 2^N$
  - By convention $v(\emptyset)=0$. We will assume $v(N)=1$.

Example (weighted majority game)

- Three owners Green (G), White (W), and Red (R) with 48%, 37% and 15% of weights, respectively.

- To take a decision the 2/3 majority is required.

- We can model this situation as a simple game({G,W,R},w) s.t.:

$$
\boxed{\begin{array}{c}
w(G) = 0 \\
w(W) = 0 \\
w(R) = 0 \\
w(G,W) = 1 \\
w(G,R) = 0 \\
w(W,R) = 0 \\
w(G,W,R) = 1
\end{array}}
$$

Transfer property

A solution Φ is map assigning to each simple game (N,v) an n-vector of real numbers. For any two simple games $(N,v),(N,w)$, Φ satisfies the transfer proeprty if it holds that $\Phi(v \vee w)+\Phi(v \wedge w) = \Phi(v)+\Phi(w)$.

Here $v \vee w$ is defined as $(v \vee w)(S) = (v(S) \vee w(S)) = \max\{v(S),w(S)\}$, and $v \wedge w$ is defined as $(v \wedge w)(S) = (v(S) \wedge w(S)) = \min\{v(S),w(S)\}$,

EXAMPLE

Two TU-games v and w on N={1,2,3}.

						$\wedge$			$\vee$	
$v(1) = 0$	Φ		$w(1) = 1$	Φ		$v\wedge w(1) = 0$	Φ		$v\vee w(1) = 1$	Φ
$v(2) = 1$			$w(2) = 0$			$v\wedge w(2) = 0$			$v\vee w(2) = 1$	
$v(3) = 0$			$w(3) = 0$			$v\wedge w(3) = 0$			$v\vee w(3) = 0$	
$v(1, 2) = 1$		$+$	$w(1, 2) = 1$		$=$	$v\wedge w(1, 2) = 1$		$+$	$v\vee w(1, 2) = 1$	
$v(1, 3) = 1$			$w(1, 3) = 0$			$v\wedge w(1, 3) = 0$			$v\vee w(1, 3) = 1$	
$v(2, 3) = 0$			$w(2, 3) = 1$			$v\wedge w(2, 3) = 0$			$v\vee w(2, 3) = 1$	
$v(1, 2, 3) = 1$			$w(1, 2, 3) = 1$			$v\wedge w(1, 2, 3) = 1$			$v\vee w(1, 2, 3) = 1$	

Real applications of simple games

- Voting by disciplined party groups in multi-party parliaments (probably elected on the basis of proportional representation);
- USA President election
- UN Security Council
- voting in the EU Council of Ministers
- voting by stockholders (holding varying amounts of stock).
- lawmaking power of the United States
- …

Weighted majority example

- Suppose that four parties receive these vote shares:
  - Party A, 27%;
  - Party B, 25%;
  - Party C, 24%;
  - Party D 24%.
- Seats are apportioned in a 100-seat parliament:
  - Party A: 27 seats   Party C: 24 seats
  - Party B: 25 seats   Party D: 24 seats
- Seats (voting weights) have been apportioned in a way that is precisely proportional to vote support, but voting power has not been so apportioned (and cannot be).

Weighted majority example (2)

A:27 seats;

B:25 seats;

C:24 seats;

D:24 seats

- Party A has *voting power* that greatly exceeds its slight advantage in *seats*. This is because:
  - Party A can form a winning coalition with any one of the other parties; and
  - the only way to exclude Party A from a winning coalition is for Parties B, C, and D to form a three-party coalition.

A:27 seats; B:25 seats; C:24 seats; D:24 seats; Quota: 51

A:2 seats; B:1 seats; C:1 seats; D:1 seats; Quota: 3

…

w(A) =1
w(B) =0
w(C) = 0
W(D)=0
w(A, B) =1
w(A, C) = 1
w(A, D) = 1
w(B, C) = 0
w(B, D) = 0
w(C, D) = 0
w(A, B, C) = 1
w(A, B, D) = 1
w(A, C, D) = 1
w(B, C, D) = 1
w(A, B, C, D) = 1

Power Indices

- Several *power indices* have been proposed to quantify the share of power held by each player in simple games.
- These particularly include:
  - the Shapley-Shubik power index (1954);
  - And the Banzhaf power index (1965).
- Such power indices provide precise formulas for evaluating the voting power of players in weighted voting games.

The Shapley-Shubik Index

- Let (N,v) be a simple game (assume v is *monotone*: for each $S,T \in 2^N$. $S \subseteq T \Rightarrow v(S) \leq v(T)$)
- "Room parable": Players gather one by one in a room to create the "grand coalition",
- At some point a winning coalition forms.
- For each ordering in which they enter, identify the *pivotal* player who, when added to the players already in the room, converts a losing coalition into a winning coalition.

A	A	A	A	A	A	B	B	B	B	B	B
B<=	B<=	C<=	D<=	C<=	D<=	A<=	A<=	C	D	C	D
C	D	B	B	D	C	C	D	A<=	A<=	D<=	C<=
D	C	D	C	B	B	D	C	D	C	A	A

C	C	C	C	C	C	D	D	D	D	D	D
D	D	A<=	B	A<=	B	C	C	A<=	B	A<=	B
A<=	B<=	D	D<=	B	A<=	A<=	B<=	C	C<=	B	A<=
B	A	B	A	D	D	B	A	B	A	C	C

The Shapley-Shubik Index (cont.)

- Player *i*'s *Shapley-Shubik power index value* is simply

$$\frac{\text{Number of orderings in which the voter } i \text{ is pivotal}}{\text{Total number of orderings}}$$

- Power index values of all voters add up to 1.
- Counting up, we see that A is pivotal in 12 orderings and each of B, C, and D is pivotal in 4 orderings. Thus:

Voter	*Sh-Sh Power*
A	1/2
B	1/6
C	1/6
D	1/6

- So according to the Shapley-Shubik index, Party A has 3 times the voting power of each other party.

The Banzhaf Index

- The *Banzhaf power index* works as follows:
  - A player i is *critical* for a winning coalition if
    - i belongs to the coalition, and
    - the coalition would no longer be winning if i defected from it.
- Voter i's *Banzhaf power Bz*(i) is

$$\frac{\text{Number of winning coalitions for which } i \text{ is critical}}{\text{Total number of coalitions to which } i \text{ belongs.}}$$

The Banzhaf Index (2)

- Given the seat shares before the election, and looking first at all the coalitions to which A belongs, we identify:

  {A},{A,B},{A,C}, {A,D}, {A,B,C}, {A,B,D}, {A,C,D}, (A,B,C,D}.

- Checking further we see that A is critical for all but two of these coalitions, namely
  - {A} (because it is not winning); and
  - {A,B,C,D} (because {B,C,D} can win without A).
- Thus: $Bz(\text{A}) = 6/8 = .75$

The Banzhaf Index (3)

- Looking at the coalitions to which B belongs, we identify:

  {B},{A,B}, {B,C}, {B,D}, {A,B,C}, {A,B,D}, {B,C,D}, (A,B,C,D}.

- Checking further we see that B is critical to only two of these coalitions:
  - {B}, {B,C}, {B,D} are not winning; and
  - {A,B,C}, {A,B,D}, and {A,B,C,D} are winning even if B defects.
- The positions of C and D are equivalent to that of B.
- Thus: $Bz(\mathrm{B}) = Bz(\mathrm{C}) = Bz(\mathrm{D}) = 2/8 = .25$

Power indices: a general formulation

- Let $p_i(S)$, for each $S \in 2^N \setminus \{\emptyset\}$, $i \notin S$, be the probability of coalition $S \cup \{i\}$ to form (of course $\sum_{S \subseteq N : i \notin S} p_i(S) = 1$)

- A power index $\psi_i(v)$ is defined as the probability of player i to be critical in v according to p:

$$\psi_i(v) = \sum_{S \subseteq N : i \notin S} p_i(S) \, [v(S \cup \{i\}) - v(S)]$$

Power indices: a general formulation (2)

- According to the Banzhaf power index, every coalitions has the same probability to form: $p_i(S)=1/(2^{n-1})$, for each $S \in 2^N \setminus \{\emptyset\}$, $i \notin S$
- According to the Shapley-Shubick power index, compute $p_i(S)$ according to the following procedure to create at random from N a subset S to which i does not belong:
  - Draw at random a number out of the urn consisting of possible sizes 0,1,2,…,n-1 where each number has probability 1/n to be drawn
  - If size s is chosen, draw a set out of the urn consisting of subsets of $N \setminus \{i\}$ of size s, where each set has the same probability, i.e. 1/combinations(n-1,s)
  - indeed, $p_i(S)=(s!\,(n-s-1)!)/n!$

UN Security Council

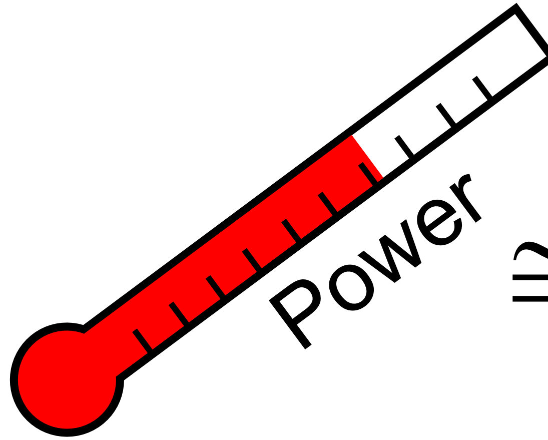
- 15 member states:
  - 5 Permanent members: China, France, Russian Federation, United Kingdom, USA
  - 10 temporary seats (held for two-year terms) (http://www.un.org/)

UN Security Council decisions

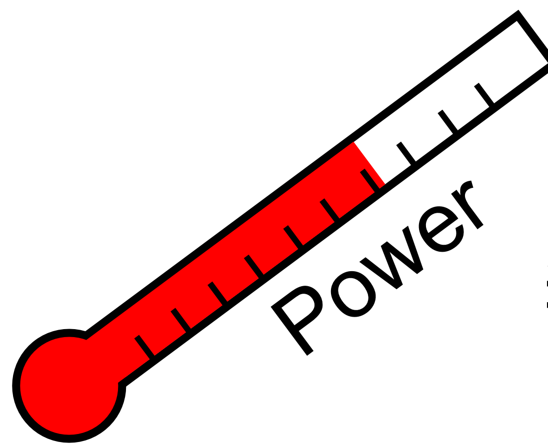
- Decision Rule: **substantive resolutions** need the positive vote of at least nine Nations but…

  …it is sufficient the negative vote of one among the permanent members to reject the decision.
- **Q**: quantify the power of nations inside the ONU council to force a substantive decision?
- Game Theory gives an answer using the Shapley-Shubik power index:

Shapley-Shubik power index



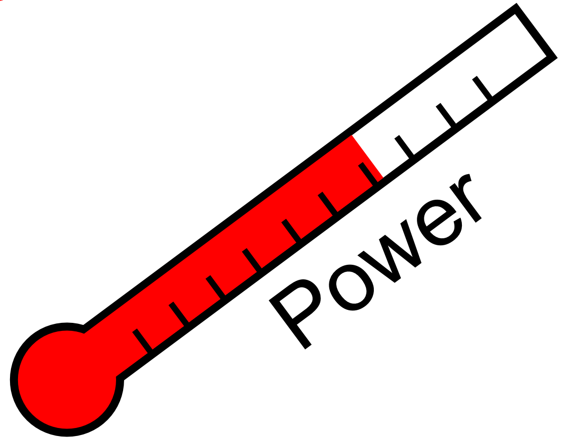
≅ 19.6%



≅ 0.2%

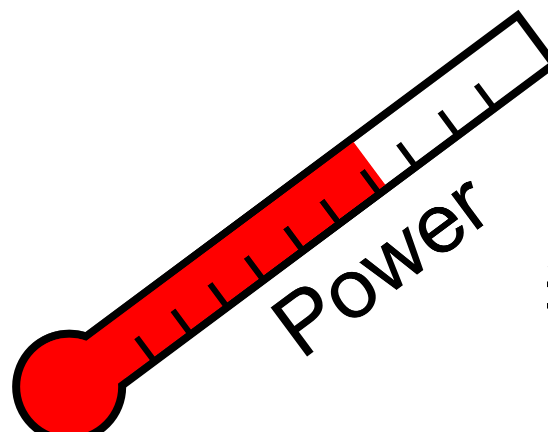
temporary seats since January 1st 2007
until January 1st 2009

Banzhaf power index



$\cong$ 5%

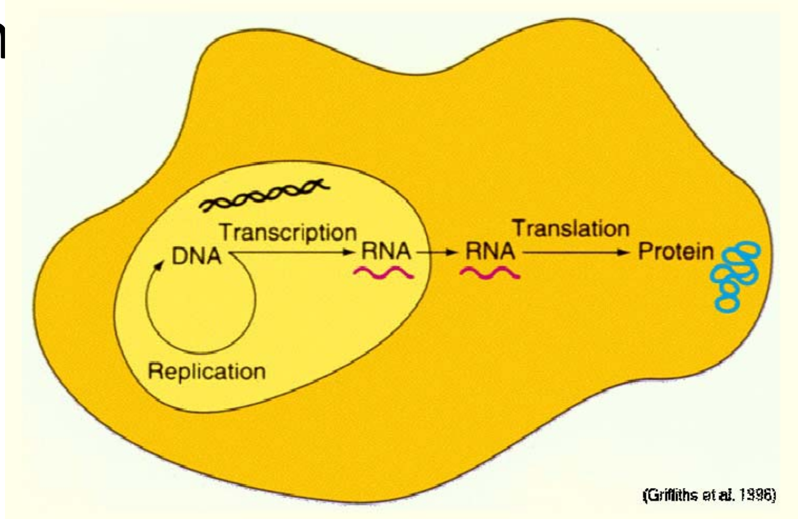
temporary seats since January 1st 2007
until January 1st 2009

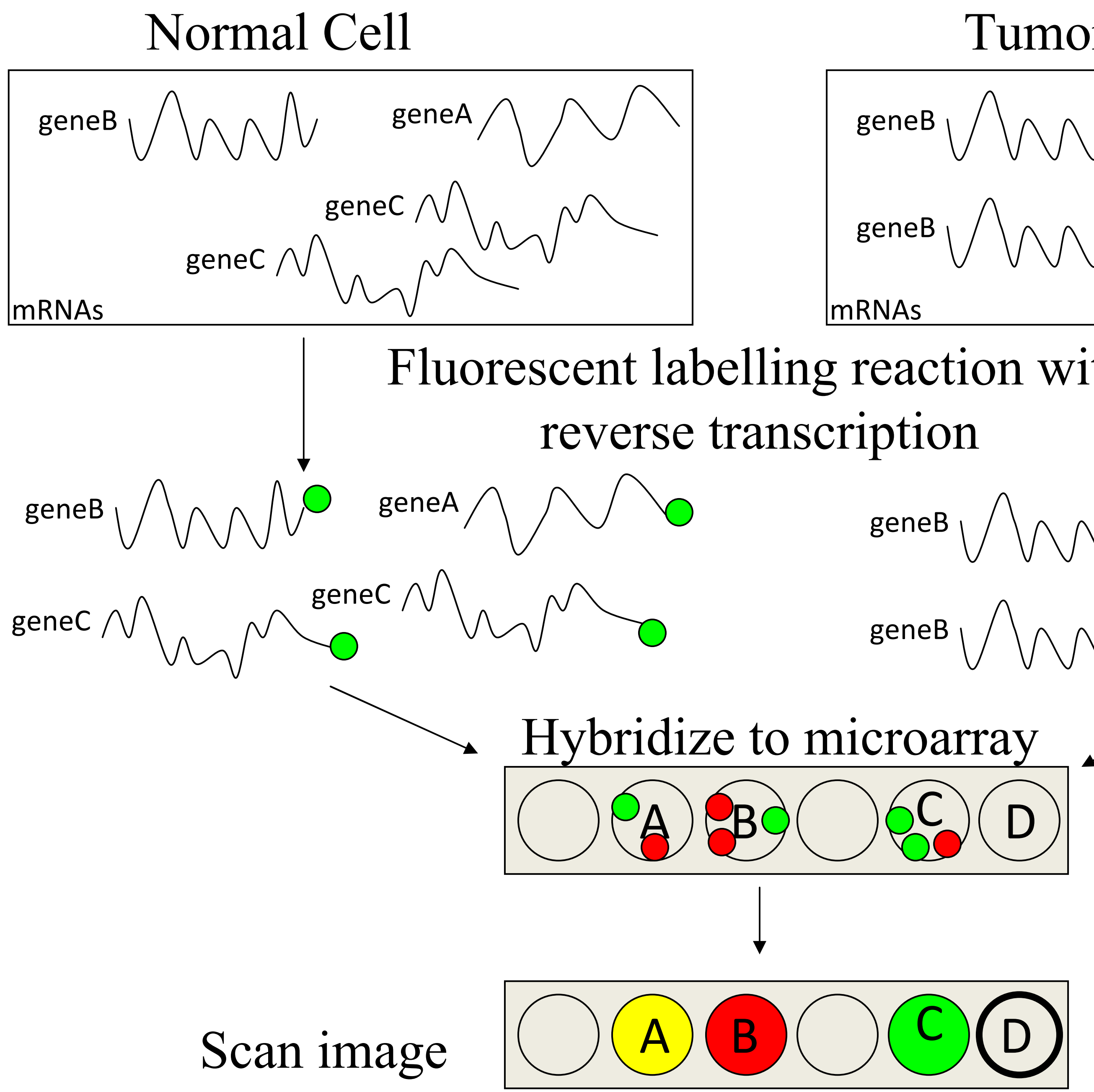


$\cong$ 0.5%

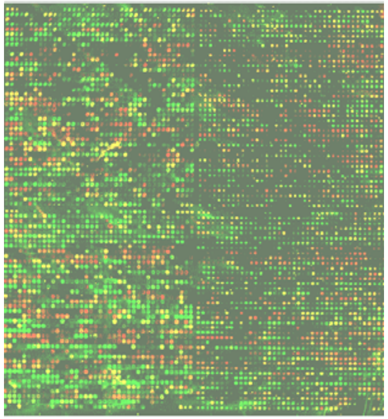
Central “dogma” of molecular biology (Crick (1958)

- *Gene expression* occurs when genetic information contained within DNA is *transcripted* into mRNA molecules and then *translated* into the proteins.
- Nowadays, microarray technology is available for taking “pictures” of gene expressions. Within a single experiment of this sophisticated technology, the level of expression of thousands of genes can be estimated in a sample of cells under a given con

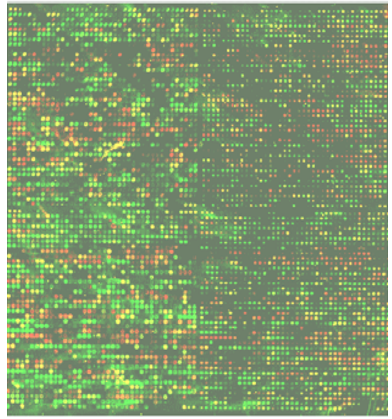




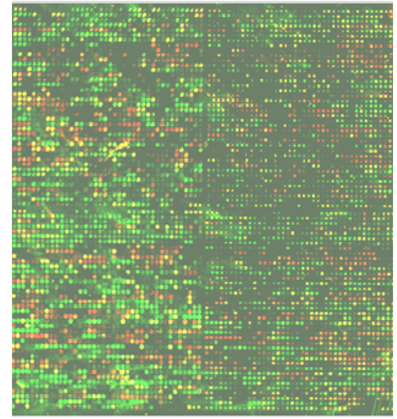
	Normal	Tumor
Gene A	1	1
Gene B	1	2
Gene C	2	1
Gene D	0	0



Array1



Array2



Array3

...

Arrays

	Array1	Array2	Array3	Array4	Array5	...
Gene1	0.46	0.30	0.80	1.51	0.90	...
Gene2	0.10	0.49	0.24	0.06	0.46	...
Gene3	0.15	0.74	0.04	0.10	0.20	...
Gene4	0.45	1.03	0.79	0.56	0.32	...
Gene5	0.06	1.06	1.35	$\boxed{1.09}$	1.09	...
...	...	...	...	...	...	

Expression level of gene 5 in array 4

The dimension of information

- A typical experiment: a table of numbers with more than 22000 rows (genes) e 60 of arrays (samples).
- If we would print the entire table with a character of 12pt, it would be necessary almost 3700 pages A4…
- …a surface of almost 220 square meters!

From political and social science to genomics…

- Players are genes
- Who knows the **decision rule** in this context?
- IDEA: we can make a rule on microarray gene expression profiles.
- Example: we define a criterion to establish which genes have abnormal expressions on each array

	array1
gene1	1.121
gene2	2.453
gene3	3.586

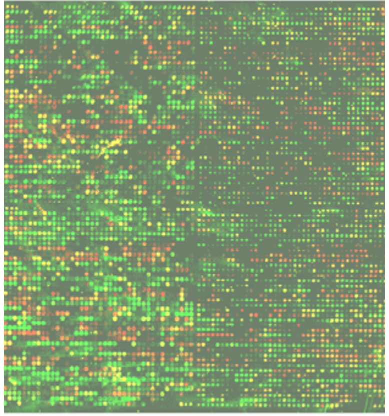
	array1
gene1	0
gene2	1
gene3	1

Decision rule

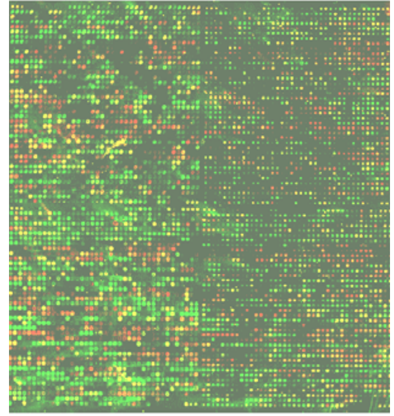
A group of genes is *winning* on a single array if all genes that have abnormal expressions belong to that group

	array1
gene1	0
gene2	1
gene3	1

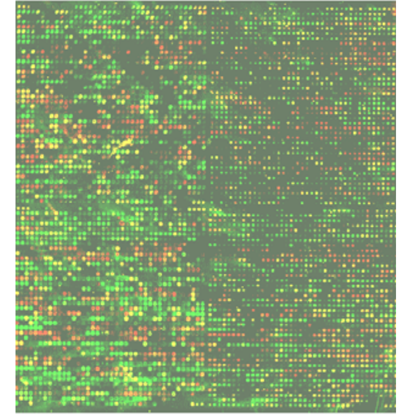
Both groups {gene2, gene3} and group {gene1, gene2, gene3} are winning.



Array1



Array2



Array3

...

	array1	array2	array3
gene1	0	1	0
gene2	1	1	0
gene3	1	0	1

- coalition {gene2, gene3} is winning two times out of three;
- coalition {gene1, gene2} is winning one time out of three;
- And so on for each coalition...

Example

	Array1	Array2	Array3
g_1	0	1	0
g_2	1	1	0
g_3	1	0	1

The corresponding *microarray game*

$<\{g_1,g_2,g_3\},v>$ tale che

$v(\emptyset)=v(\{g_1\})=v(\{g_2\})=0$

$v(\{g_1,g_3\})=v(\{g1,g2\})=v(\{g_3\})=1/3$

$v(\{g_2,g_3\})=2/3$

$v(\{g_1,g_2,g_3\})=1.$

The Shapley value is

$Sh_{g1}=1/6$ $Sh_{g2}=1/3$ $Sh_{g3}=1/2$

Axioms for the Shapley value on microarray games

Property 1: Null Gene (NG)

A gene which does not contribute to change the worth of any coalition of genes, should receive zero power.

Prop.2:Equal Splitting (ES)

Each sample should receive the same level of reliability. So the power of a gene on two samples should be equal to the sum of the power on each sample divided by two.

g1
g2
g3

s1	
0	ψ_1
0	ψ_2
1	ψ_3

$+$

s2	
1	ψ'_1
1	ψ'_2
0	ψ'_3

$=$

	s1	s2	
g1	0	1	$(\psi_1+\psi'_1)/2$
g2	0	1	$(\psi_2+\psi'_2)/2$
g3	1	0	$(\psi_3+\psi'_3)/2$

Partnership of genes

A group of genes S such that does not exist a proper ($\subset$) subset of S which contributes in changing the worth of genes outside S.

Example

These two sets are partnerships of genes in the corresponding Microarray game

	s1	s2	s3
g1	0	1	1
g2	0	1	1
g3	1	0	1

Property 3: Partnership Monotonicity (PM)

(N,v) a microarray game. If two partnerships of genes S and T, with $|T|\geq|S|$ are such that they are

-*disjoint* ($S\cap T=\emptyset$),

-*equivalent* ($v(S)=v(T)$)

-*exhaustive* ($v(S\cup T)=v(N)$),

then genes in the smaller *partnership* S must receive more relevance then genes in T.

Example

		s_1	s_2
ψ_1	g1	0	1
ψ_2	g2	0	1
ψ_3	g3	1	0
ψ_4	g4	1	0
ψ_5	g5	1	0

$\psi_i \geq \psi_k$

For each $i\in\{1,2\}$ $k\in\{3,4,5\}$

Property 4: Partnership Rationality (PR)

The total amount of power index received from players of a partnership S should not be smaller than v(S)

Property 5: Partnership Feasibility (PF)

The total amount of power index received from players of a partnership S should not be greater than v(N)

Theorem (Moretti, Patrone, Bonassi (2007))**:**

The Shapley value is the unique solution which satisfies NG, ES, PM, PR, PF on the class of microarray games.

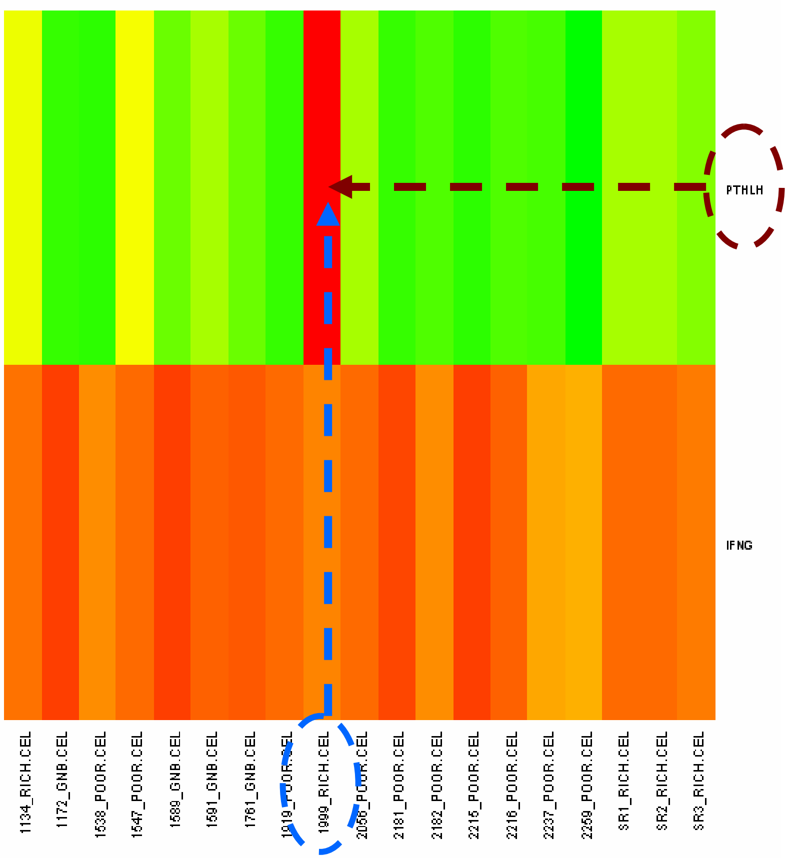
Real data analysis

Application (1): Neuroblastic Tumors data

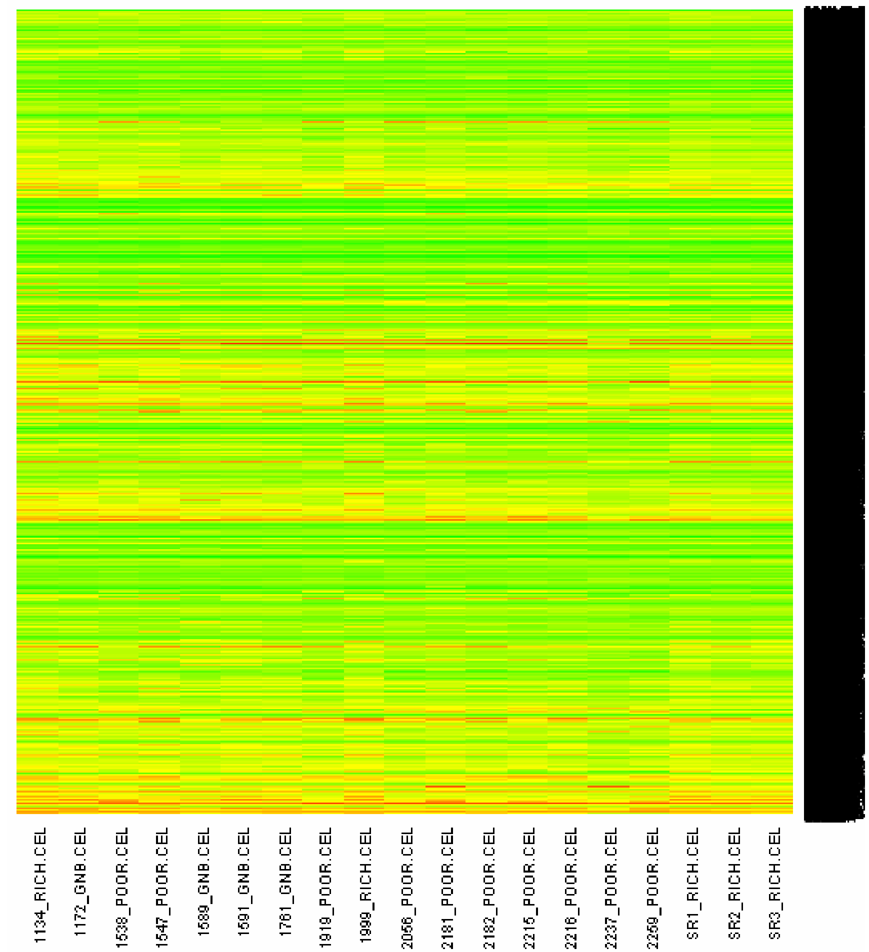
(*Cancer,* 113(6), 1412 – 1422)

- Neuroblastic Tumors (NTs) is a group of pediatric cancers with a great tissue heterogeneity.
    - Most of NTs are composed of undifferentiated, poorly differentiated or differentiating neuroblastic (Nb) cells with very few or absence of Schwannian Stromal (SS) cells: these tumors are grouped as Neuroblastoma (Schwannian stroma-poor) (labeled as NTs-SP).
    - The remaining NTs are composed of abundant SS cells and classified as Ganglioneuroblastoma (Schwannian stroma-rich) intermixed or nodular and Ganglioneurom (labeled as NTs-SR).
    - The evolution of the disease is strongly influenced by the istology of the tumor and children with NTs-SR have a better prognosis w.r.t, NTs-SP.

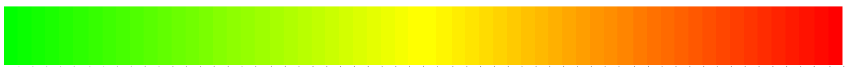
2 geni



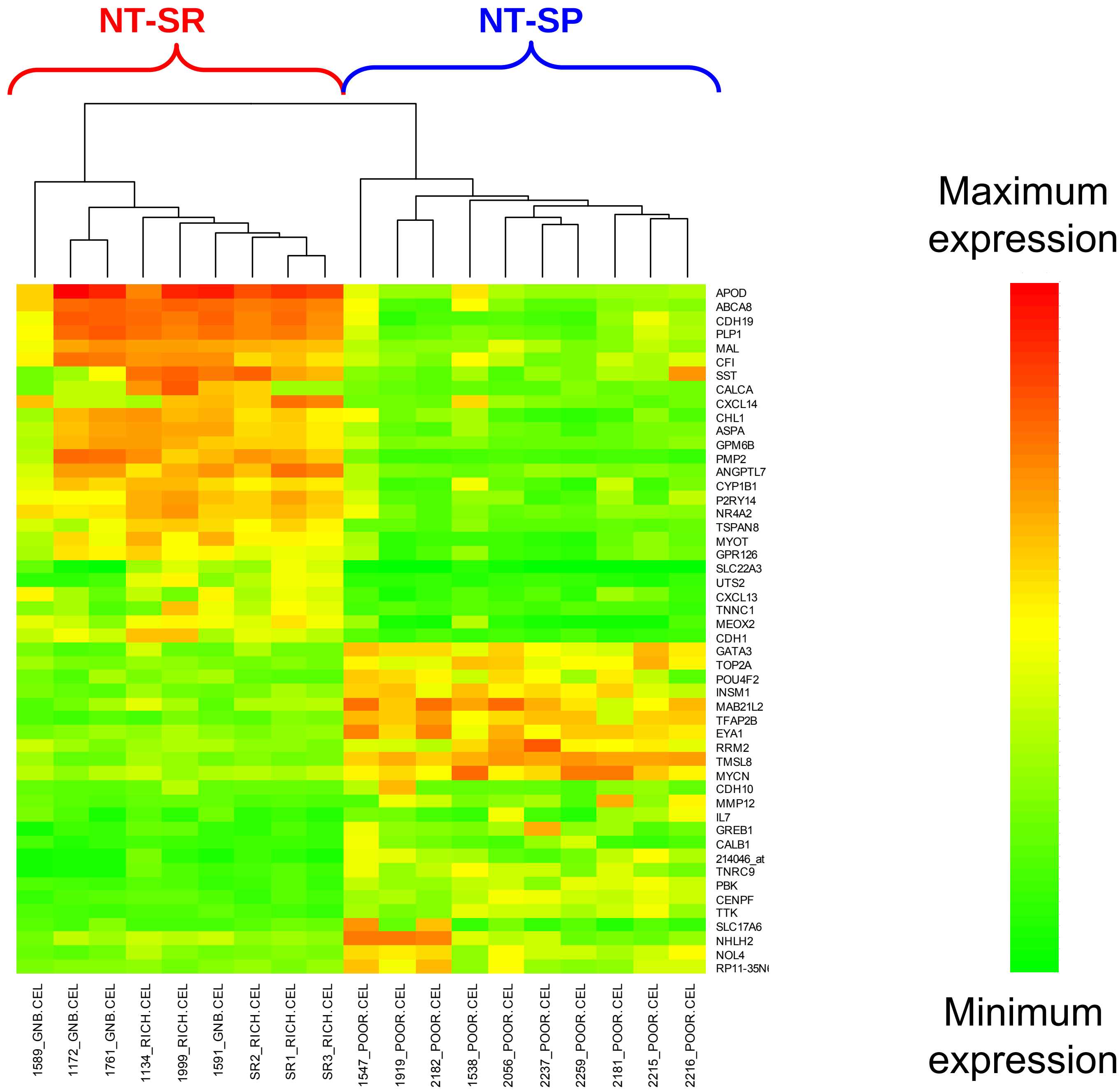
22283 geni



Minimum expression



Maximum expression



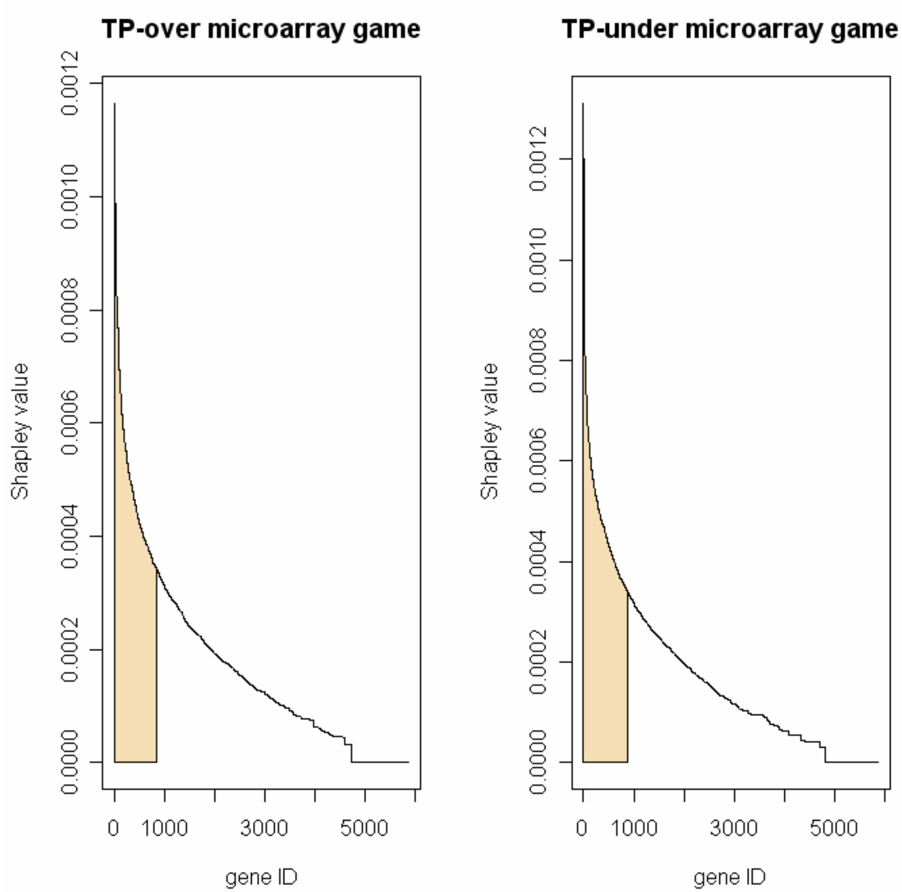
NT-SR
NT-SP
Maximum expression
Minimum expression
APOD
ABCA8
CDH19
PLP1
MAL
CFI
SST
CALCA
CXCL14
CHL1
ASPA
GPM6B
PMP2
ANGPTL7
CYP1B1
P2RY14
NR4A2
TSPAN8
MYOT
GPR126
SLC22A3
UTS2
CXCL13
TNNC1
MEOX2
CDH1
GATA3
TOP2A
POU4F2
INSM1
MAB21L2
TFAP2B
EYA1
RRM2
TMSL8
MYCN
CDH10
MMP12
IL7
GREB1
CALB1
214046_at
TNRC9
PBK
CENPF
TTK
SLC17A6
NHLH2
NOL4
1589_GNB.CEL
1172_GNB.CEL
1761_GNB.CEL
1134_RICH.CEL
1999_RICH.CEL
1591_GNB.CEL
SR2_RICH.CEL
SR1_RICH.CEL
SR3_RICH.CEL
1547_POOR.CEL
1919_POOR.CEL
2182_POOR.CEL
1538_POOR.CEL
2056_POOR.CEL
2237_POOR.CEL
2259_POOR.CEL
2181_POOR.CEL
2215_POOR.CEL
2216_POOR.CEL

Application (2): effects of air pollution

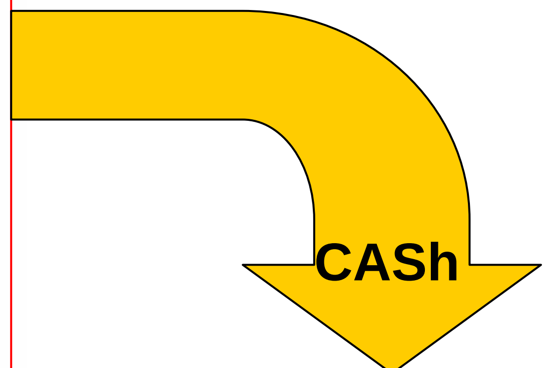
BMC Bioinformatics (IF 3.49), **9**:361).

- Study population:
  - 23 children from 12 families (2 siblings) from the areas of Teplice (TP) in Czech Republic
  - TP is infamous for air pollution
  - 24 children from the rural, less polluted are of Prachatice (PR)
  - Hybridization to Agilent Human 1A Oligo Microarray (v2) G4110B, containing over 22000 60mer probes
  - Individual samples were hybridized with a sample of the common reference (a pool of PR individuals)
  - Data have been normalized, condensed and filtered by *Genedata*, Basel (CH)

Selection based on two criteria: Shapley value and CASh



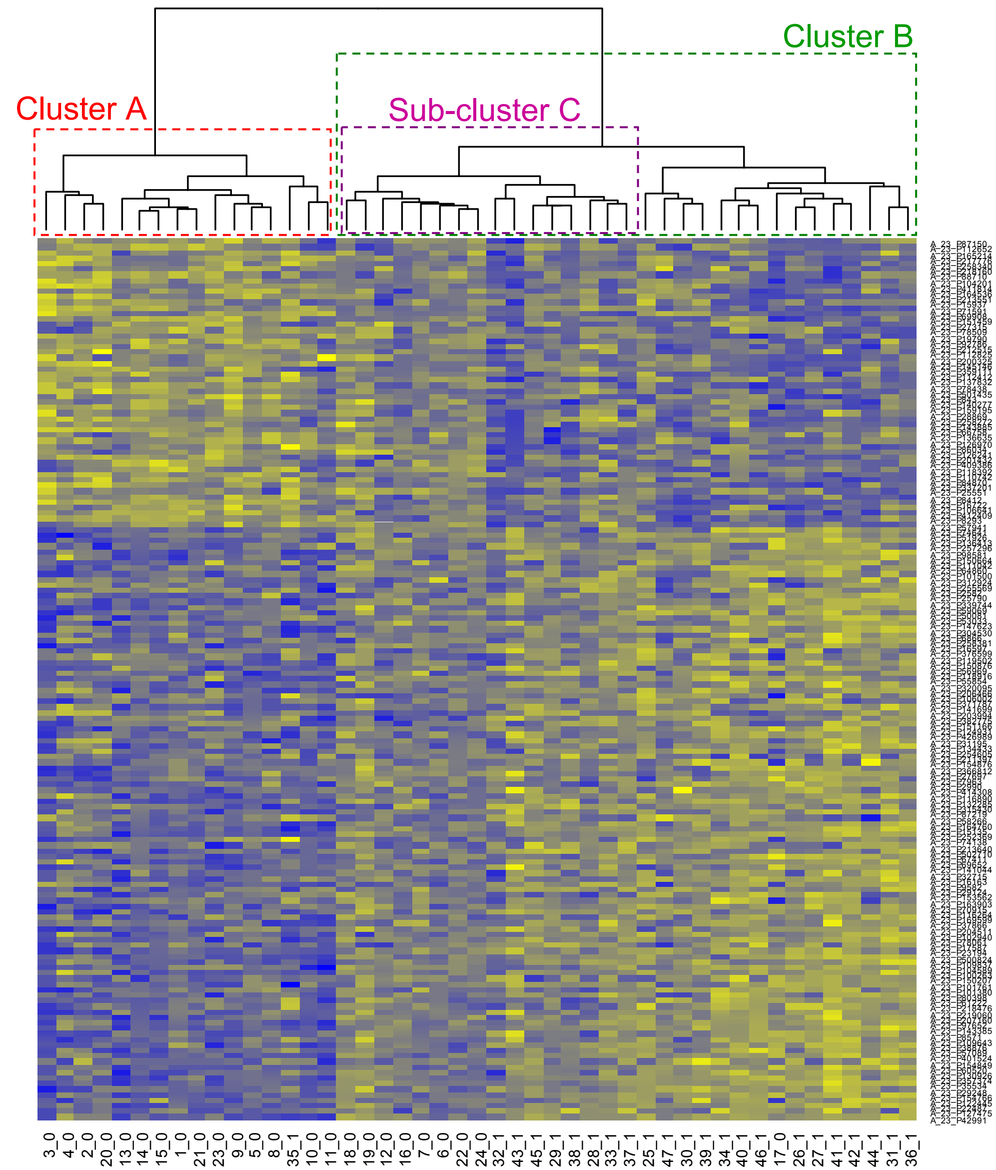
838 genes 889 genes



p-value	Genes	σ^{TP+}	σ^{TP-}
< 0.0001	20	16	4
< 0.001	33	27	6
< 0.01	159	107	52
< 0.05	434	245	189

Game theory Gene selection

- 47 biological samples (columns) and 159 genes (rows) with highest Shapley values and with un-adjusted p-value smaller than 0.01.
- yellow = high expression
- blue = low expression



Distance: *Euclidean*

Agglomerative method: *Ward*