Big Data and ML Workshop 2019
Centro Carlo Azeglio Ciampi, October 21-22, 2019

Weaving the Property Graph of **Company Ownerships**

Luigi **Bellomarini**
Andrea **Gentili**
Michela **Iezzi**

Divisione Ricerca sulle Tecnologie Avanzate
Dipartimento **IT**

A R T
www.bankit.art

BANCA D'ITALIA
EUROSISTEMA

Company graphs

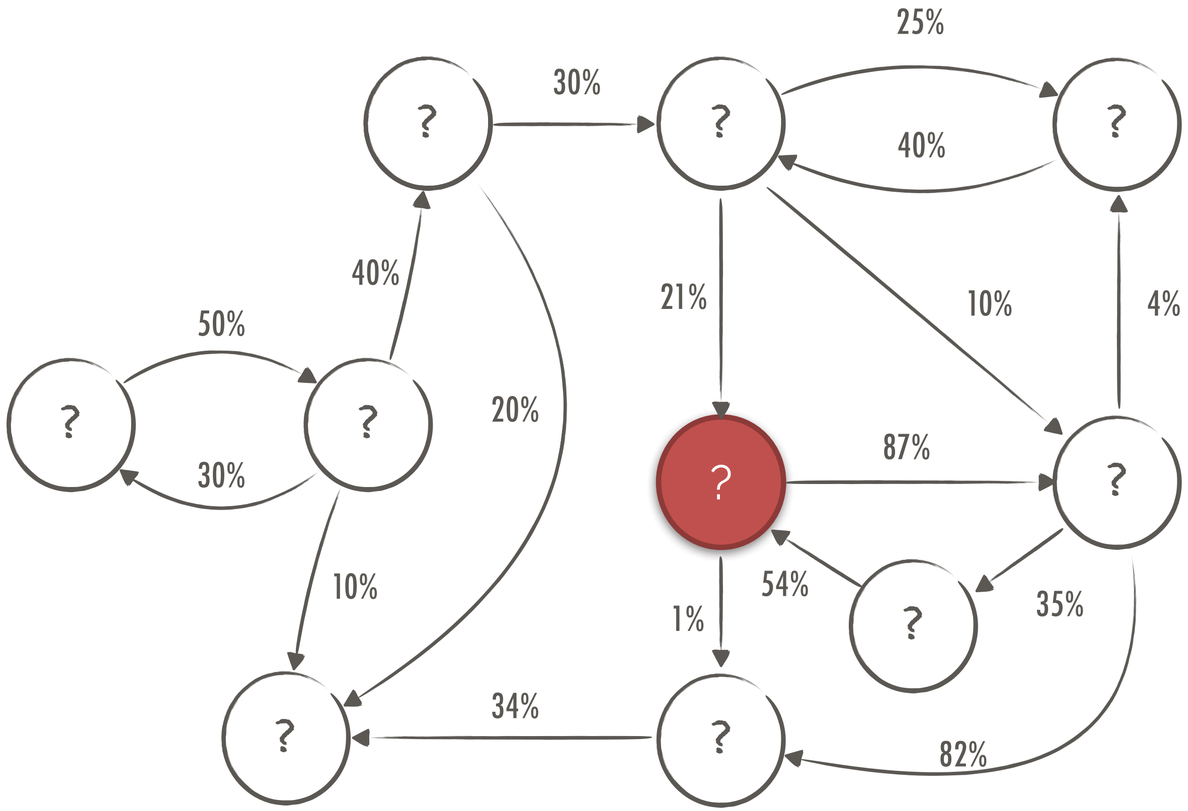
We build graphs of company networks, to:

1. reveal **power**
    1. finding **controllers**
    2. studying the **structure** of Italian market
    3. studying **dispersion** of control
    4. global **shareholding** analysis

2. detect **collusion** and do **forensics**
    1. support **AML**
    2. detecting **ultimate beneficial owners**

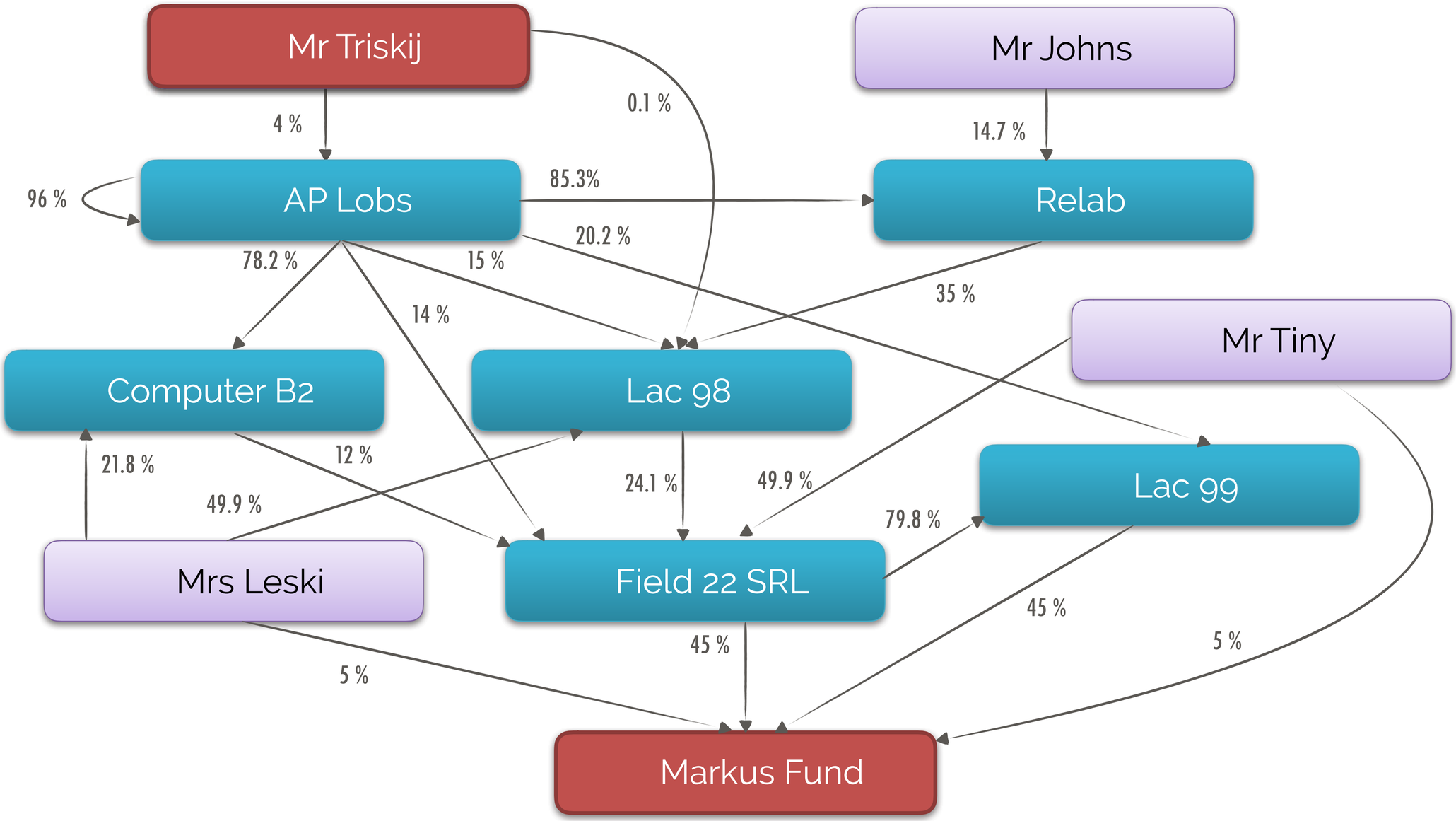
3. evaluate **risks**
4. model **propagations** (e.g., of shocks)
5. guarantee **compliance**
6. perform enhanced **due diligence**
7. understand complex **foreign shareholder structures**
8. know real **cash flows**

The setting



- Who takes **decisions**?
- Who's the ultimate **beneficial owner**?
- Is there **collusion**?
- How does **risk** propagate?
- What are the **real cash flows**?

The setting


Mr Triskij
Mr Johns
4 %
0.1 %
14.7 %
96 %
AP Lobs
85.3%
Relab
20.2 %
78.2 %
15 %
35 %
14 %
Mr Tiny
Computer B2
Lac 98
21.8 %
12 %
24.1 %
49.9 %
Lac 99
49.9 %
79.8 %
Mrs Leski
Field 22 SRL
45 %
45 %
5 %
5 %
Markus Fund

Ownership and Control

- **Integrated ownership** is about direct and indirect, owners of a company
  - it can be seen in terms of cash flow rights

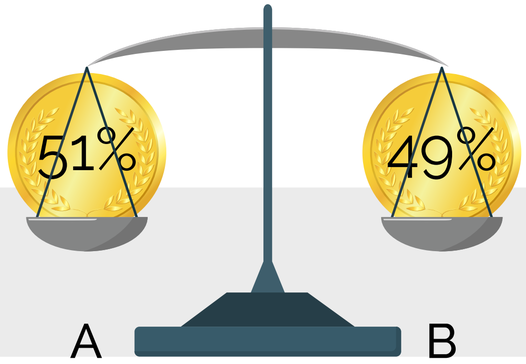
- **Control** is about voting power
  - of any direct and indirect owner of a company

Ownership vs Control

Companies A and B have both **direct** (shares) and **indirect** ownership of company C.

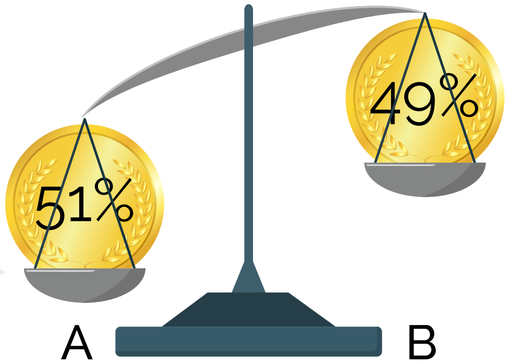
A has 51% of C
B has 49% of C

Ownership



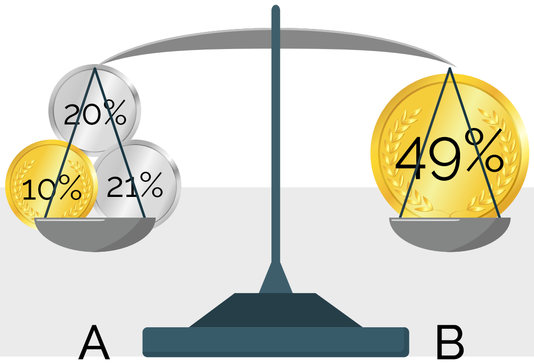
Ownership of C is comparable

Control

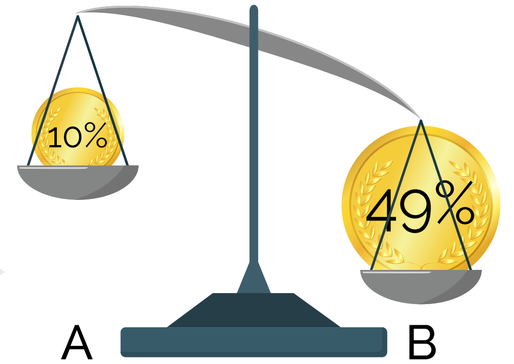


A has 100% decision power
A controls C

A has 10% and 41% of C
B has 49% of C



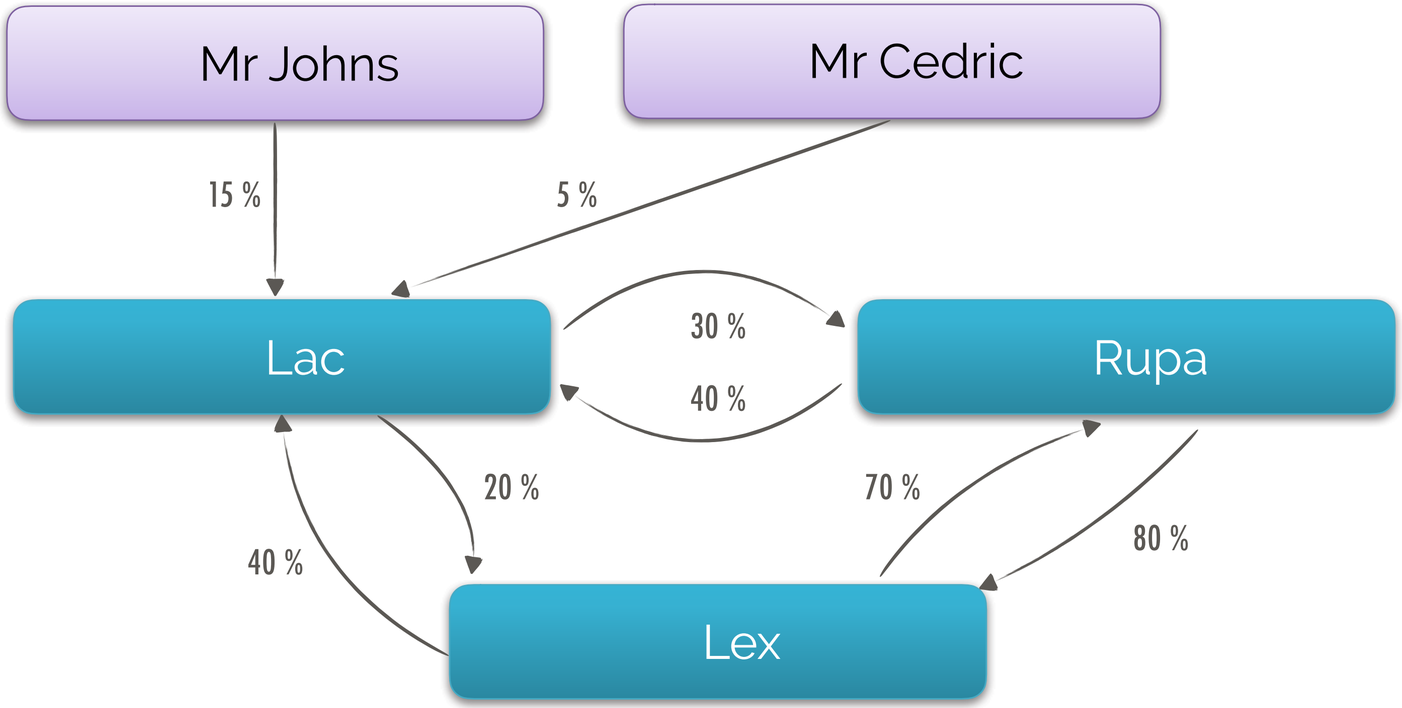
Ownership of C is comparable



A has 10% decision power
A does not control C

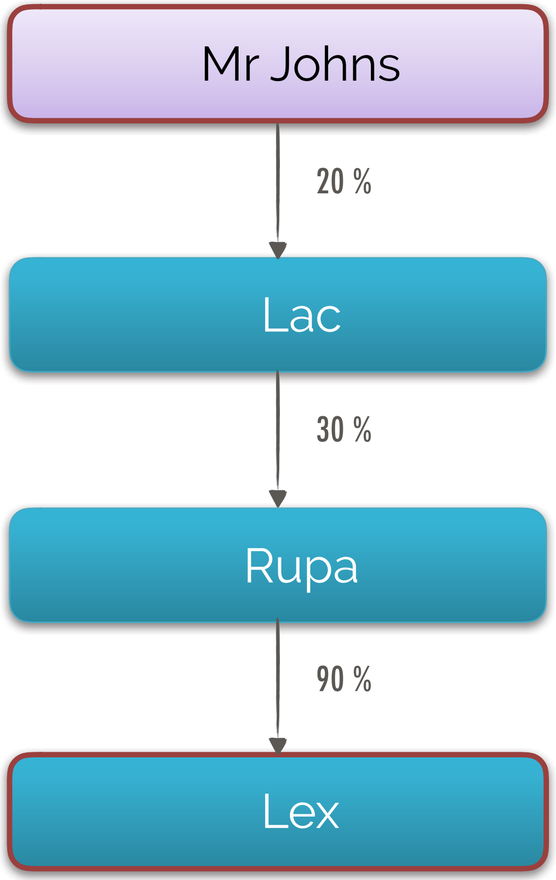
Integrated Ownership

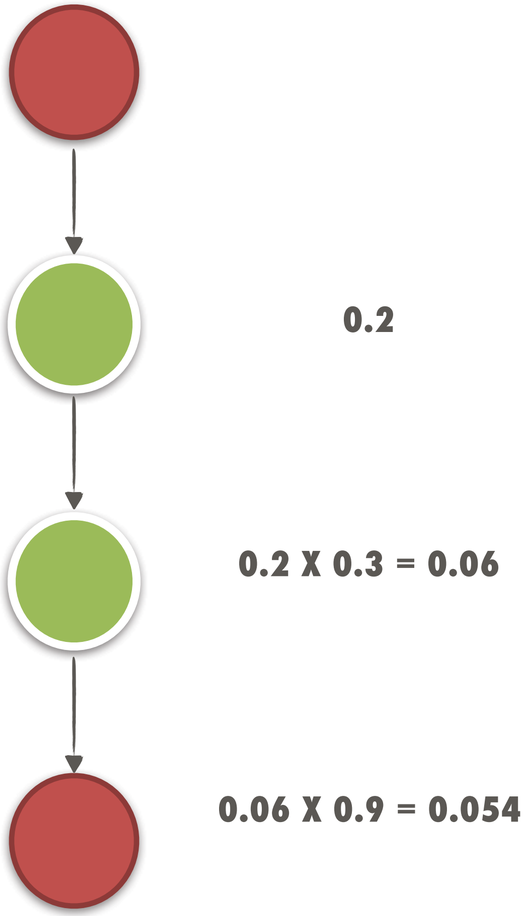
- Traditional systems only store the first-level (the closest) shareholders for a specific company



Integrated Ownership: the basic math

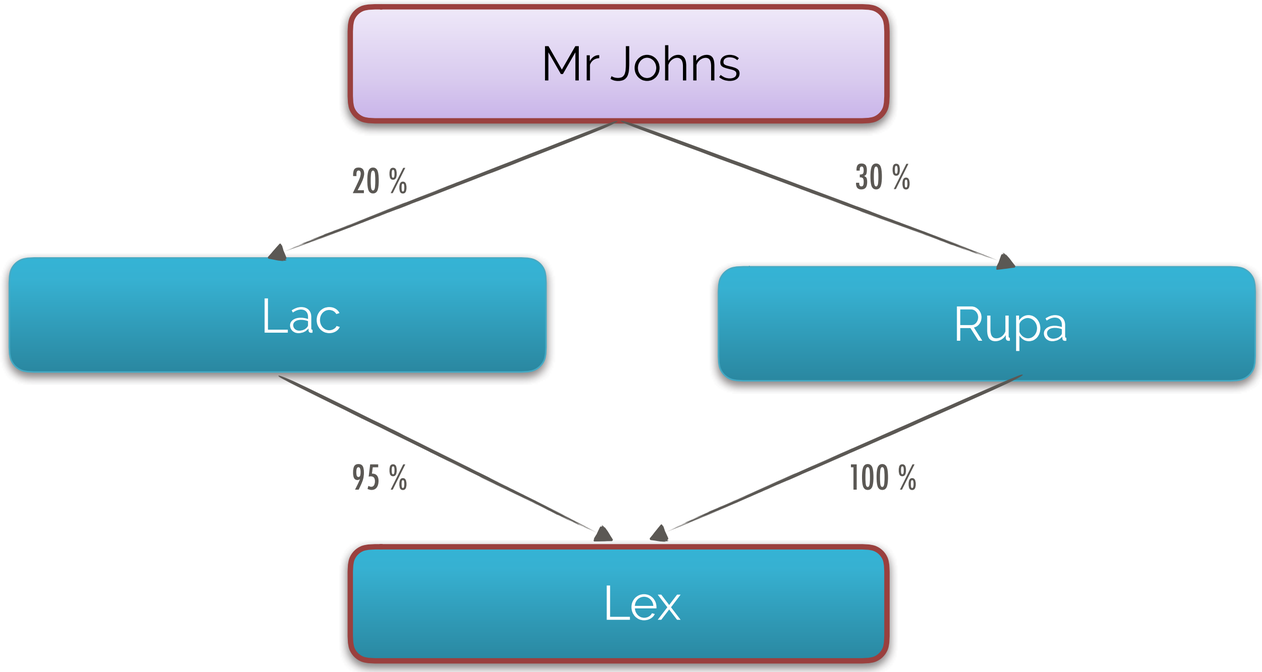
- Indirect ownership

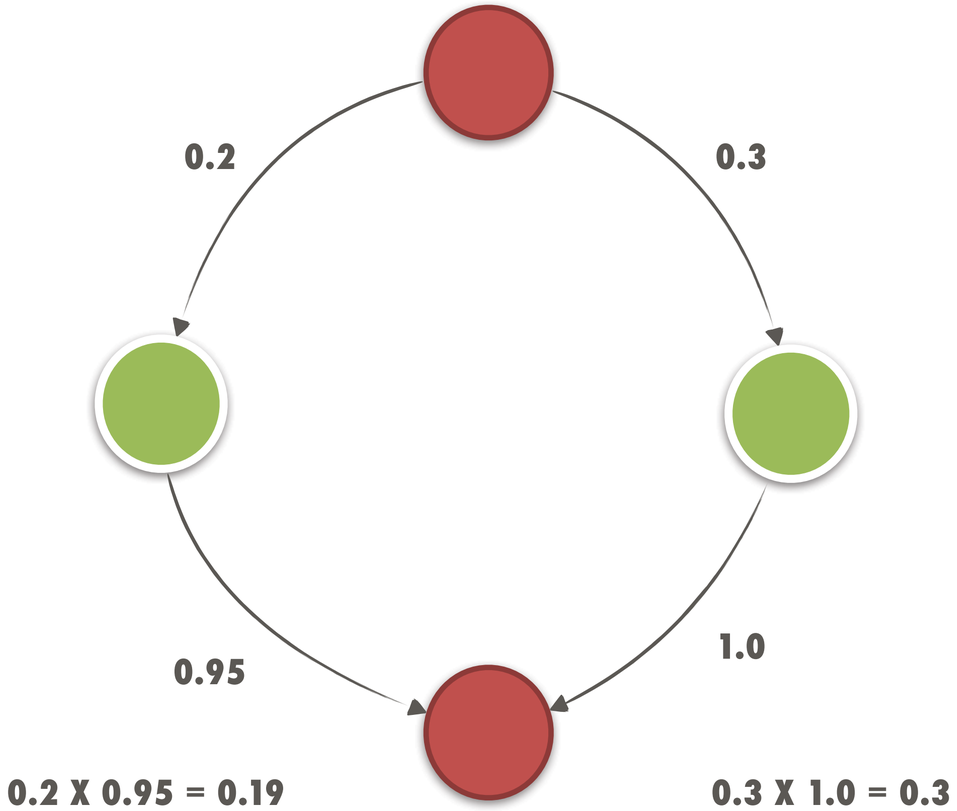




Integrated Ownership: the basic math

- Parallel ownership

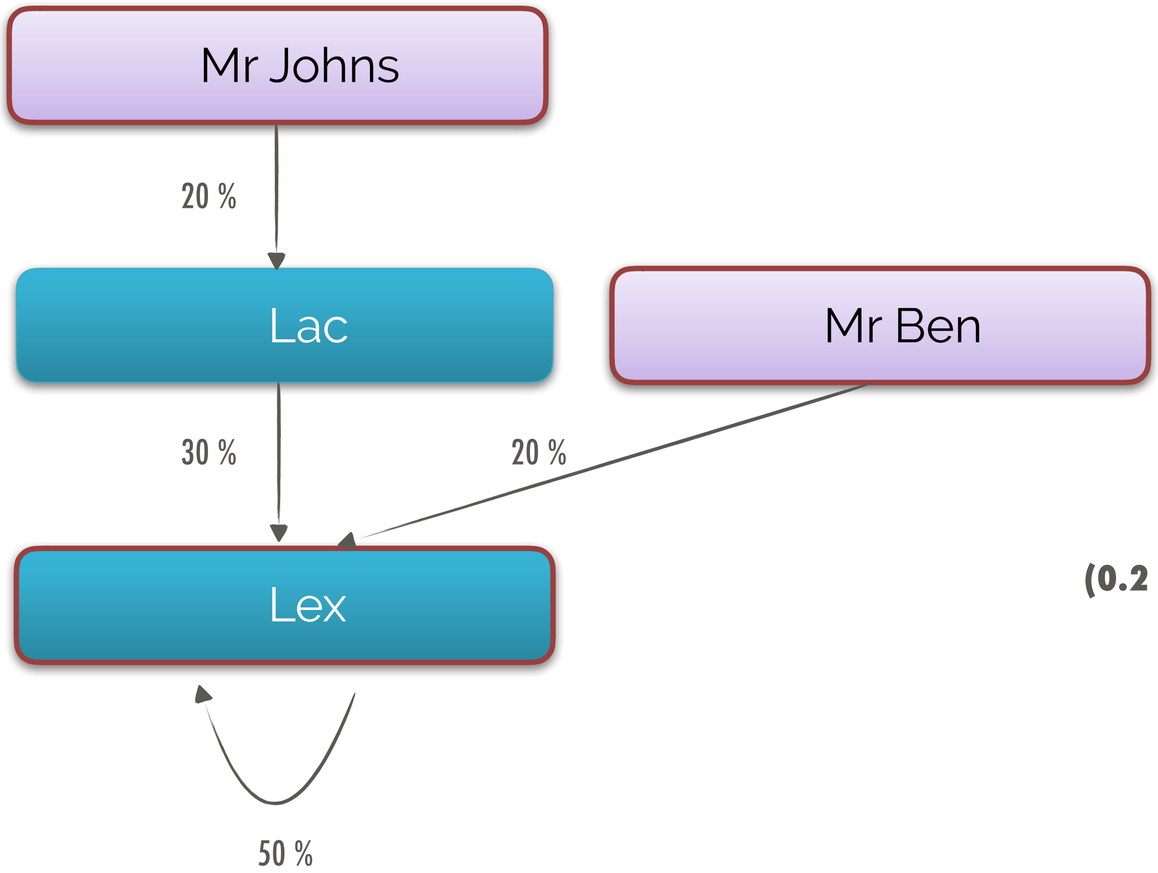


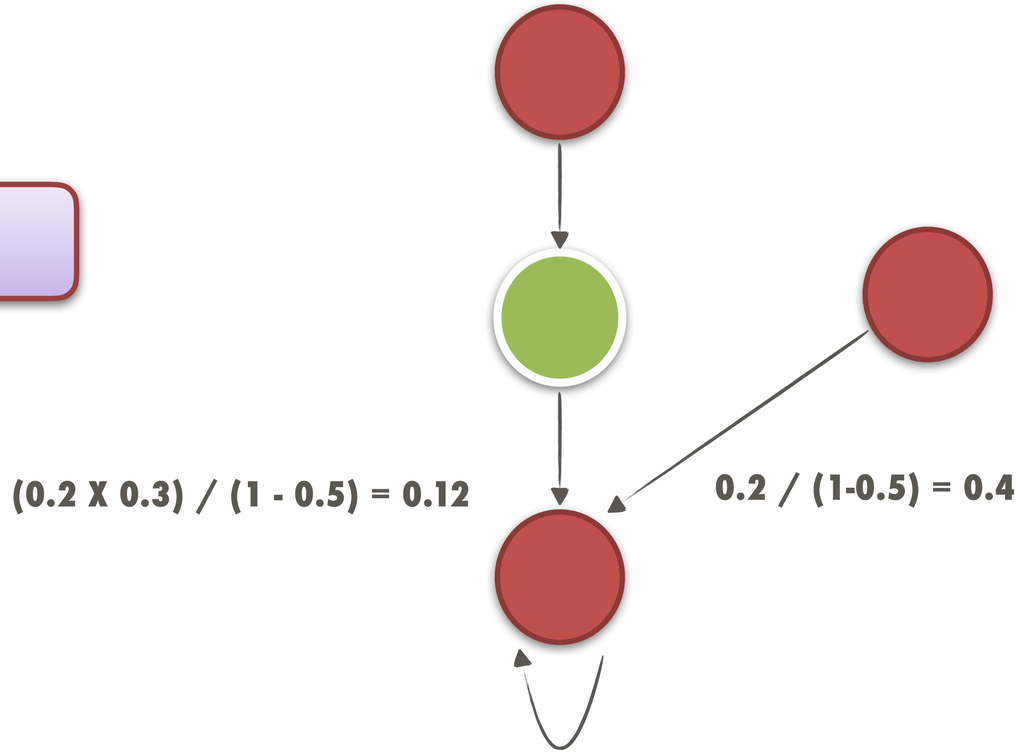


0.19 + 0.3 = 0.49

Integrated Ownership: the basic math

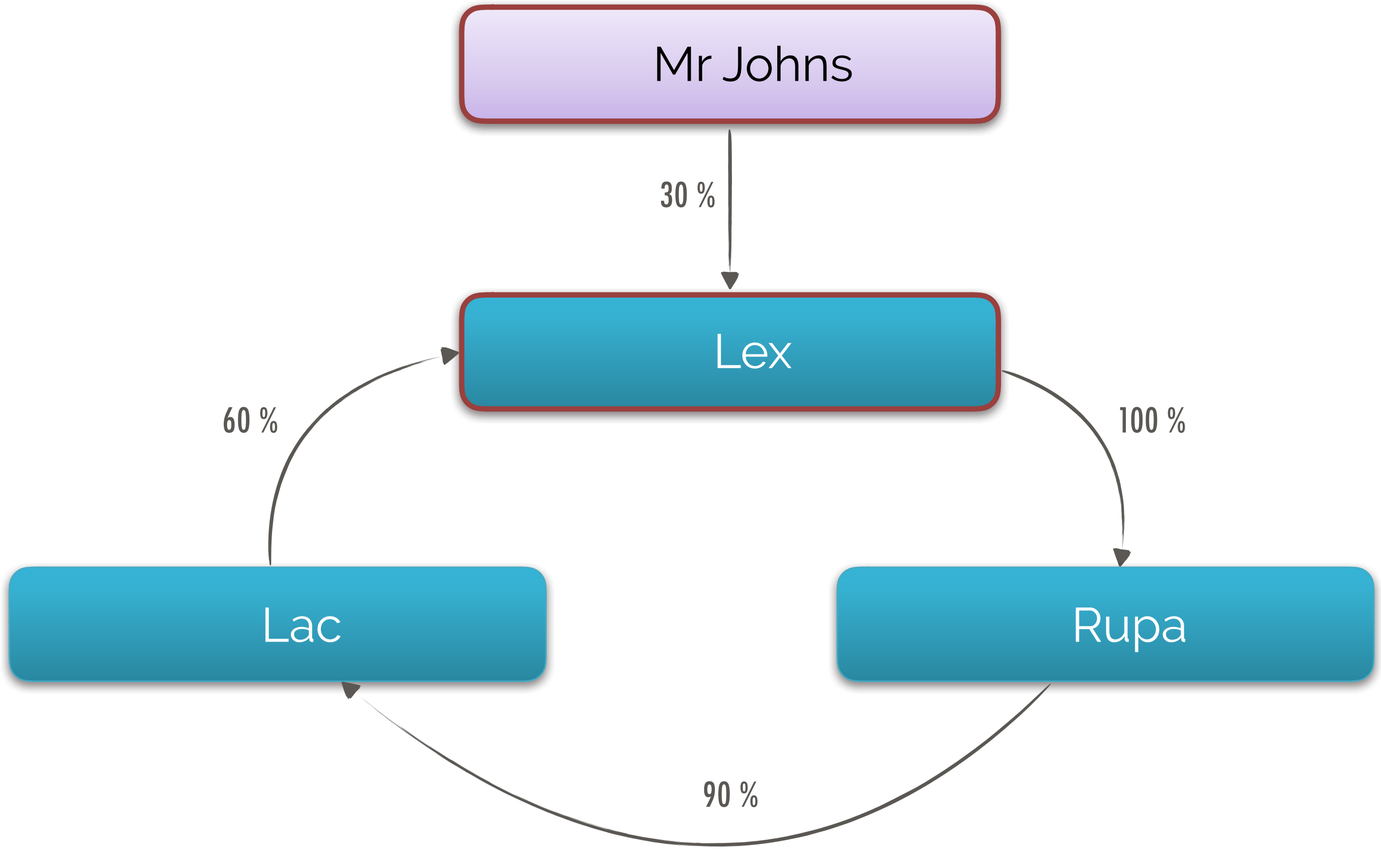
- Cycles (direct)

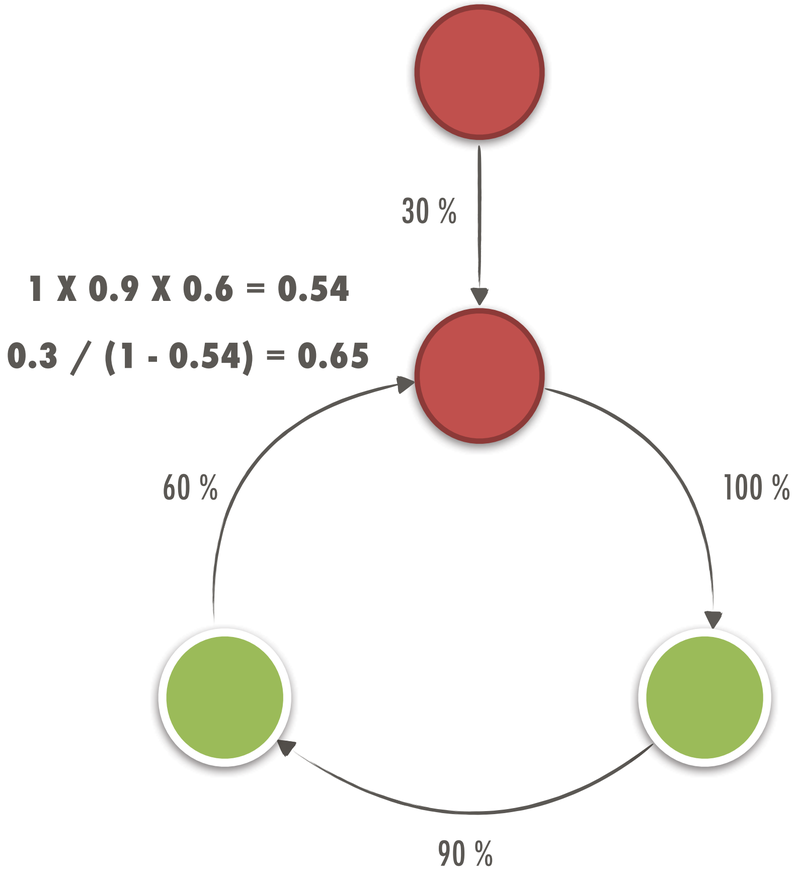




Integrated Ownership: the basic math

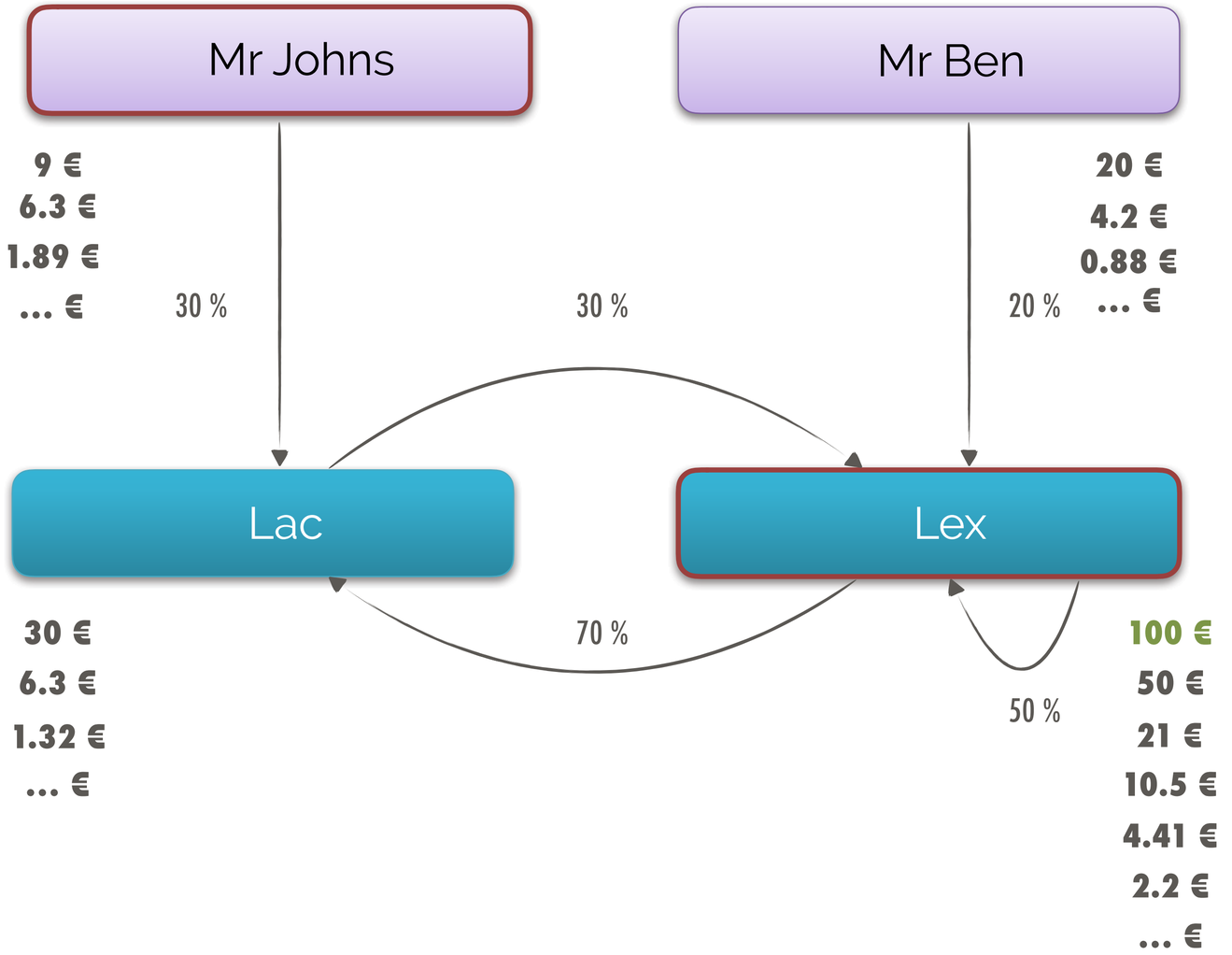
- Cycles (indirect)





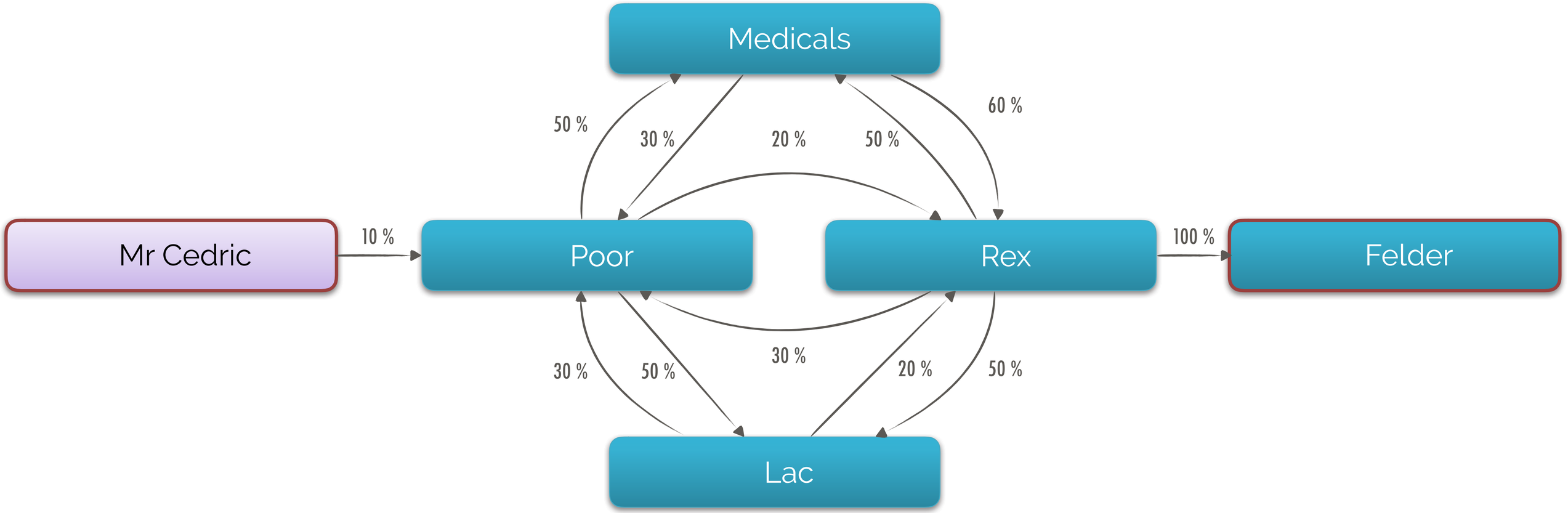
Integrated Ownership: the basic math

- Cycles (nested)



Integrated Ownership: real cases

- Cycles (nested)



Integrated Ownership: our model

The weight of a path P in an ownership graph is $w(P) = \Pi_{(p_i,p_j)\in P} w(p_i, p_j)$

An ϵ**-Baldone path** P from s to t is a sequence $[s, p_1, \ldots, p_n, t]$ such that $s \neq p_i$, for $i = 1,\ldots, n$ and $w(P) > \epsilon$ with $\epsilon \in \mathbb{R}^+$

The ϵ**-Baldone ownership** of a company s on a company t in an ownership graph G is a function $\mathcal{O}_\epsilon^G(s,t) : (s,t) \to \mathbb{R}$ defined as $\sum_B w(P_i)$ where B is the set of all possible ϵ**-Baldone** paths from s to t.

The **Baldone ownership** of a company s on a company t in an ownership graph G is a function $\mathcal{O}^G(s,t) : (s,t) \to \mathbb{R}$ defined as $\lim_{\epsilon\to 0}\mathcal{O}_\epsilon(s,t)$ where B is the set of all possible ϵ**-Baldone** paths from s to t.

Integrated Ownership: computation methods

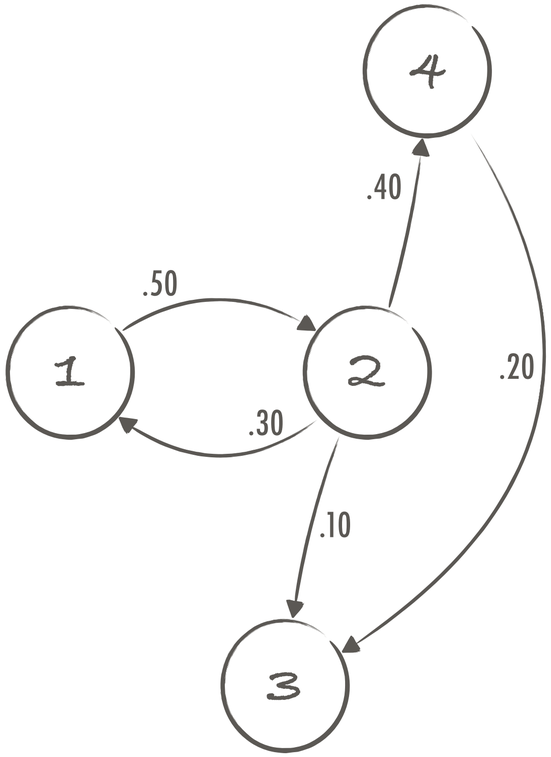
- **Problem complexity**

  Computing "all-to-all" Baldone ownerships can be solved in polynomial time in the number of companies. Conjecture: $n^y, y \in [2,3]$

  - Our approaches:

    - **Closed-form expression** (let's see how it works)
    - **Pure Reasoning** (for approximated but efficient results)
    - **Ad-hoc algorithm** (ongoing, for top-level performance)

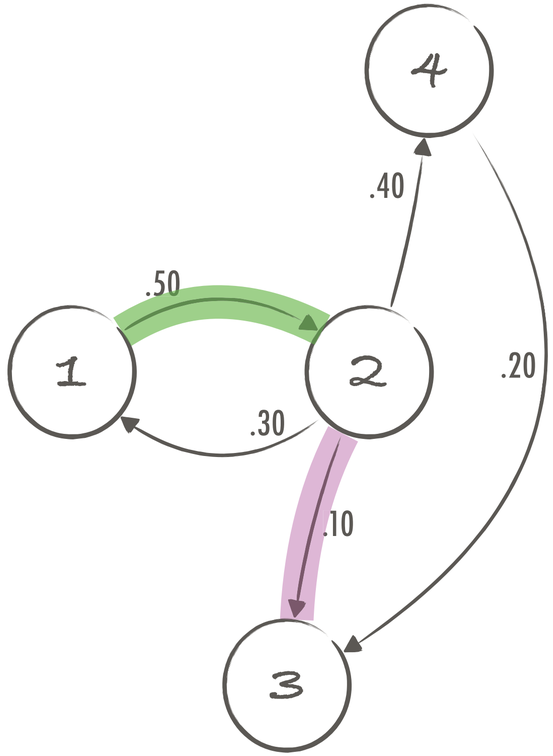
Integrated Ownership: closed form



Adjacency matrix

	1	2	3	4
1		.50	.09	.20
2	.30		.18	.40
3				
4			.20	

Integrated Ownership: closed form



$$A_{ik} = \sum_j A_{ij} \times A_{jk}$$

	1	2	3	4
1		.50		
2	.30		.10	.40
3				
4			.20	

A

	1	2	3	4
1		.50		
2	.30		.10	.40
3				
4			.20	

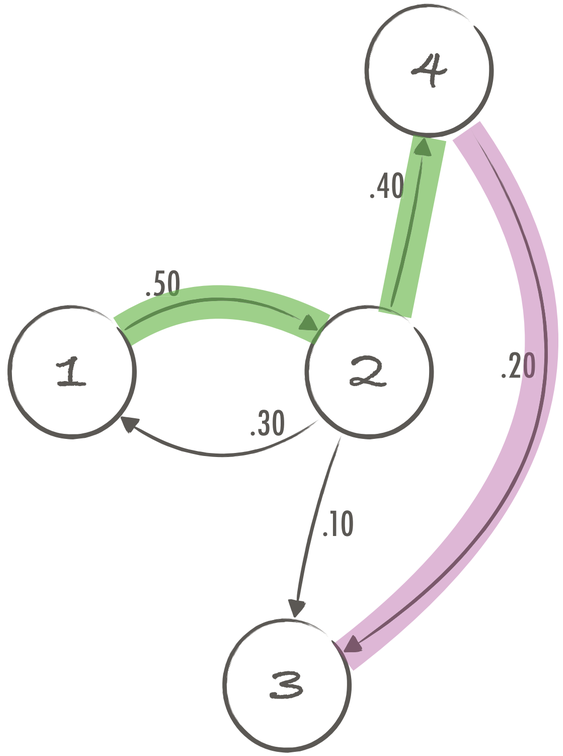
A

A2

	1	2	3	4
1	.15		.05	.20
2		.15	.08	
3				
4				

Squared Adjacency Matrix

Integrated Ownership: closed form



$$A_{ik} = \sum_j A_{ij} \times A_{jk}$$

	1	2	3	4
1	.15		.05	.20
2		.15	.08	
3				
4				

A2

	1	2	3	4
1		.50		
2	.30		.10	.40
3				
4			.20	

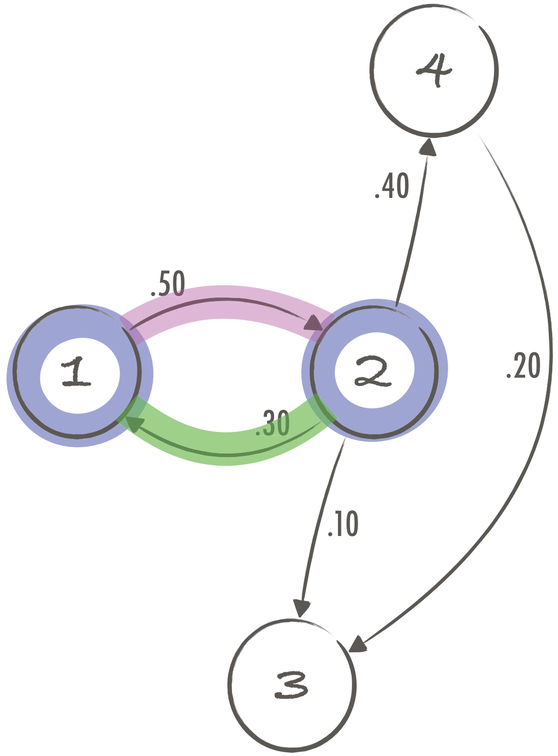
A

A3

	1	2	3	4
1		.075	.04	
2	.045		.015	.06
3				
4				

Cubed Adjacency Matrix

Integrated Ownership: closed form



$$A_{ik} = \sum_j A_{ij} \times A_{jk}$$

A2

	1	2	3	4
1	.15		.05	.20
2		.15	.08	
3				
4				

A

	1	2	3	4
1		.50		
2	.30		.10	.40
3				
4			.20	

A3

	1	2	3	4
1		.075	.04	
2	.045		.015	.06
3				
4				

Cubed Adjacency Matrix

Integrated Ownership: closed form

partial integrated ownership

	1	2	3	4
1		.50		
2	.30		.10	.40
3				
4			.20	

A

+

	1	2	3	4
1	.15		.05	.20
2		.15	.08	
3				
4				

A^2

+

	1	2	3	4
1		.075	.04	
2	.045		.015	.06
3				
4				

A^3

=

	1	2	3	4
1	.15	.575	.09	.20
2	.345	.15	.195	.46
3				
4			.20	

$A+A^2+A^3$

$$A^1 + A^2 + \ldots = \sum_{i=1}^{N-1} A^i$$

Integrated Ownership: closed form

$$\hat{W}_{st} = W_{st} + \sum_{k \neq s} \hat{W}_{sk} W_{kt}$$ which can be manipulated into:

$$\hat{W} = (I - \text{diag}(\hat{W}))\ W + \hat{W}W$$

and solved as:

$$\hat{W} = \text{diag}(V)^{-1} VW \qquad \text{with} \quad V = (I - W)^{-1}$$

In summary

- **Ownership problem characterization**
  - Theoretical study (e.g., complexity analysis)
  - Novel algorithms to compute **all-to-all Baldone ownerships**
  - Efficient and fully transparent ownership model
- **Construction of the Italian company graph**
  - all Italian companies, all links, all shareholders
    - 4.059M **nodes**, 3.960M **edges**, ~4M **SCC**, ~600K **WCC**
  - **family links** between shareholders
- **Data available soon + basic AI tools for many applications …**

Conclusions

1. reveal **power**
    1. finding **controllers**
    2. studying the **structure** of Italian market
    3. studyng **dispersion** of control
    4. global **shareholding** analysis

2. detect **collusion** and do **forensics**
    1. support **AML**
    2. detecting **ultimate beneficial owners**

3. evaluate **risks**
4. model **propagations** (e.g., of shocks)
5. guarantee **compliance**
6. perform enhanced **due diligence**
7. understand complex **foreign shareholder structures**
8. know real **cash flows**

Open
discussion