

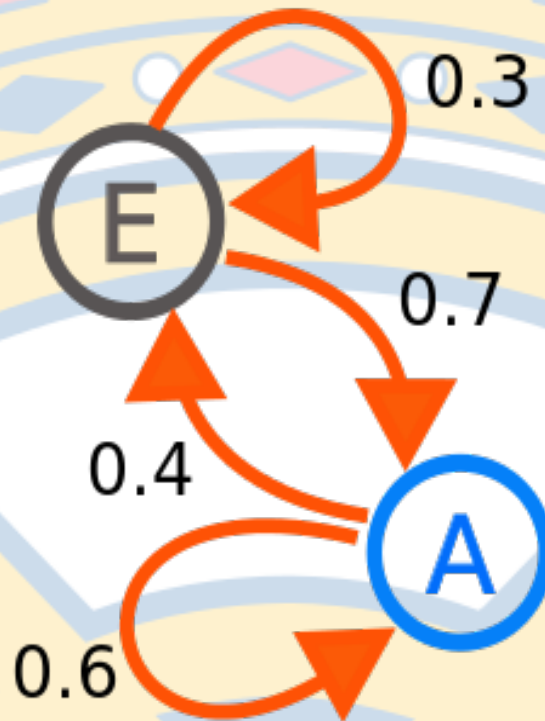
El Clásico: A Markov Approach to Real Madrid's Victory

Thomas Stephens and Sean Duggan



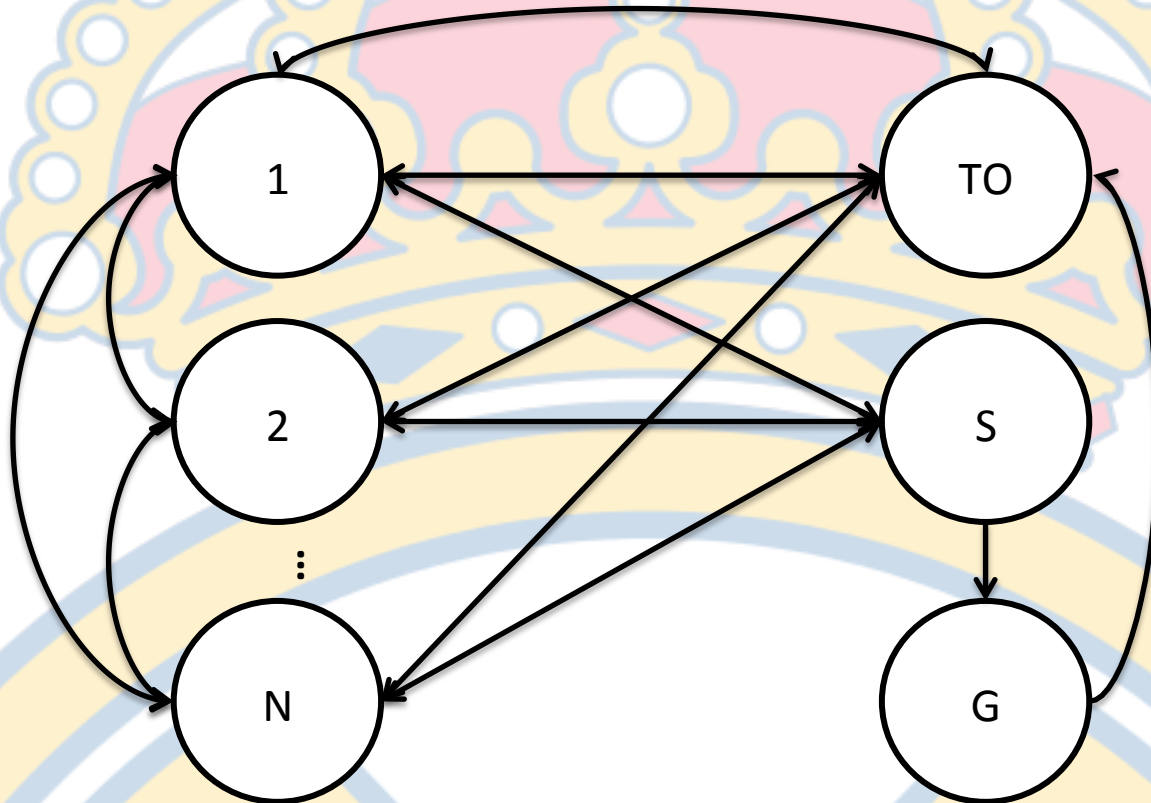
What is a Markov Chain?

- A way of describing how a process moves from one state to another
- $x_n = P^n x$, where x = state, and P is the “Transition Matrix”
- Memoryless



Source: Wikipedia

Markov Chain State Matrix



Real Madrid Starting Lineup



Game Actions

Possession Recording Ex: Casillas → Pepe → Kroos → Benzema → Shot → T.O.

	I.C 1	Pepe 3	S.R. 4	D.C. 15	M.V. 12	L.M. 19	A.A. 17	T.K. 8	Isco 23	A.I. 24	J.R. 10	K.B. 9	S.K. 6	C.R. 7	T.O	Shots	Goals
Casillas 1	0	2	1	1	2	0	1	1	2	0	0	0	0	2	3	0	0
Pepe 3	3	1	4	3	3	2	0	7	1	2	1	1	1	5	4	1	0
Ramos 4	0	3	0	1	7	3	0	5	3	1	1	1	0	0	0	0	0
Carvajal 15	1	10	1	2	1	7	0	3	0	0	22	0	1	2	4	0	0
Marcelo 12	1	0	4	1	5	0	0	12	15	1	4	6	0	11	12	0	0
Modric 19	0	4	1	7	6	0	0	9	3	2	7	1	0	3	1	0	0
Arbeloa 17	0	0	0	0	2	0	0	1	0	0	0	0	0	0	1	0	0
Kroos 8	0	4	3	5	8	9	1	0	3	1	7	11	2	4	2	1	0
Isco 23	0	0	2	1	10	2	0	5	4	0	2	3	0	4	11	0	0
Illarramendi 24	0	1	2	0	0	0	0	1	0	0	0	0	2	1	1	0	0
James 10	0	0	1	11	4	5	1	1	4	0	5	6	0	7	12	7	0
Benzema 9	0	0	0	1	7	4	0	4	1	0	4	2	0	5	3	6	0
Khedira 6	0	2	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0
Ronaldo 7	0	0	0	1	4	4	1	3	5	0	6	3	0	13	13	5	0
Turnovers	11	12	6	19	14	7	0	7	3	0	3	2	0	0	0	0	0
Shots	0	0	0	0	0	1	0	2	0	0	1	1	0	1	11	0	3
Goals	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0

Note: Table indicates actions from Row to Column
i.e. Casillas passes to Pepe twice.

Normalized Game Actions

	I.C 1	Pepe 3	S.R. 4	D.C.15	M.V. 12	L.M. 19	A.A. 17	T.K. 8	Isco 23	A.I. 24	J.R. 10	K.B. 9	S.K. 6	C.R. 7	T.O	Shots	Goals
Casillas 1	0.00	0.13	0.07	0.07	0.13	0.00	0.07	0.07	0.13	0.00	0.00	0.00	0.00	0.13	0.20	0.00	0.00
Pepe 3	0.08	0.03	0.10	0.08	0.08	0.05	0.00	0.18	0.03	0.05	0.03	0.03	0.03	0.13	0.10	0.03	0.00
Ramos 4	0.00	0.12	0.00	0.04	0.28	0.12	0.00	0.20	0.12	0.04	0.04	0.04	0.00	0.00	0.00	0.00	0.00
Carvajal 15	0.02	0.19	0.02	0.04	0.02	0.13	0.00	0.06	0.00	0.00	0.41	0.00	0.02	0.04	0.07	0.00	0.00
Marcelo 12	0.01	0.00	0.06	0.01	0.07	0.00	0.00	0.17	0.21	0.01	0.06	0.08	0.00	0.15	0.17	0.00	0.00
Modric 19	0.00	0.09	0.02	0.16	0.14	0.00	0.00	0.20	0.07	0.05	0.16	0.02	0.00	0.07	0.02	0.00	0.00
Arbeloa 17	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
Kroos 8	0.00	0.07	0.05	0.08	0.13	0.15	0.02	0.00	0.05	0.02	0.11	0.18	0.03	0.07	0.03	0.02	0.00
Isco 23	0.00	0.00	0.05	0.02	0.23	0.05	0.00	0.11	0.09	0.00	0.05	0.07	0.00	0.09	0.25	0.00	0.00
Illarramendi 24	0.00	0.13	0.25	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.25	0.13	0.13	0.00	0.00
James 10	0.00	0.00	0.02	0.17	0.06	0.08	0.02	0.02	0.06	0.00	0.08	0.09	0.00	0.11	0.19	0.11	0.00
Benzema 9	0.00	0.00	0.00	0.03	0.19	0.11	0.00	0.11	0.03	0.00	0.11	0.05	0.00	0.14	0.08	0.16	0.00
Khedira 6	0.00	0.40	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.00
Ronaldo 7	0.00	0.00	0.00	0.02	0.07	0.07	0.02	0.05	0.09	0.00	0.10	0.05	0.00	0.22	0.22	0.09	0.00
Turnovers	0.13	0.14	0.07	0.23	0.17	0.08	0.00	0.08	0.04	0.00	0.04	0.02	0.00	0.00	0.00	0.00	0.00
Shots	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.10	0.00	0.00	0.05	0.05	0.00	0.05	0.55	0.00	***
Goals	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00

Note: Table indicates actions from Row to Column. All Rows add to 1.

Markov Customization

Problem: it's rare that everyone scores in one match.

Solution: Use their annual Shots/Goal Ratio

Implementation: $P(\text{Shot} \rightarrow \text{Goal}) =$

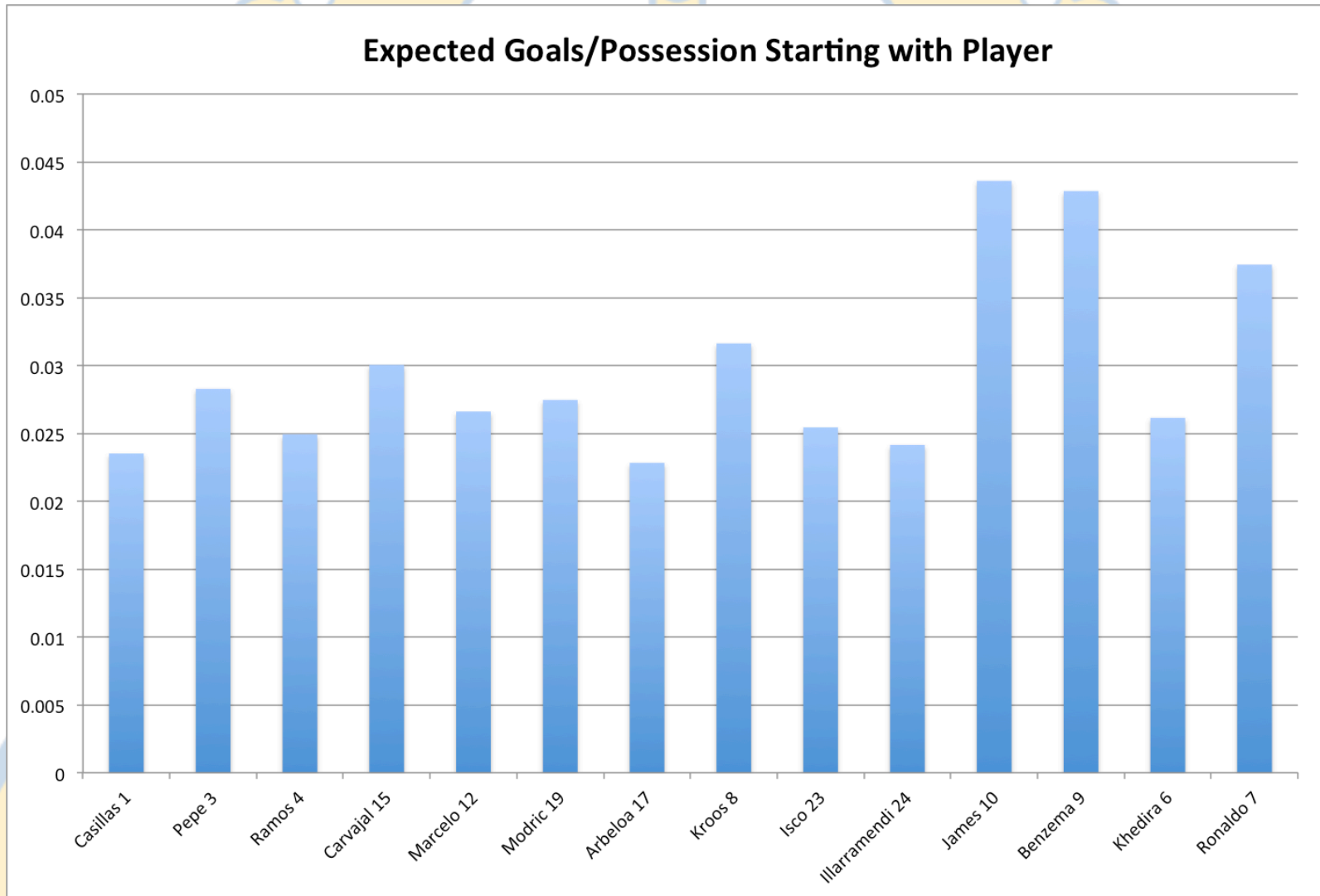
[player possession vector] x [player shooting percentages]

State 1:	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
State 2:	0	0.13	0.07	0.07	0.13	0	0.07	0.07	0.13	0	0	0	0	0.13	0.2	0	0	0
State 3:	0.04	0.06	0.05	0.07	0.15	0.07	0	0.12	0.07	0.01	0.08	0.05	0.01	0.09	0.12	0.02	0	0

Player	Player Goals/Shot
Casillas 1	-
Pepe 3	0.24
Ramos 4	0.10
Carvajal 15	0.40
Marcelo 12	0.04
Modric 19	0.03
Arbeloa 17	-
Kroos 8	0.04
Isco 23	0.18
Illarramendi 24	0.25
James 10	0.17
Benzema 9	0.32
Khedira 6	0.13
Ronaldo 7	0.28

Etc.

Expected Goals Per Possession Per Starting Player



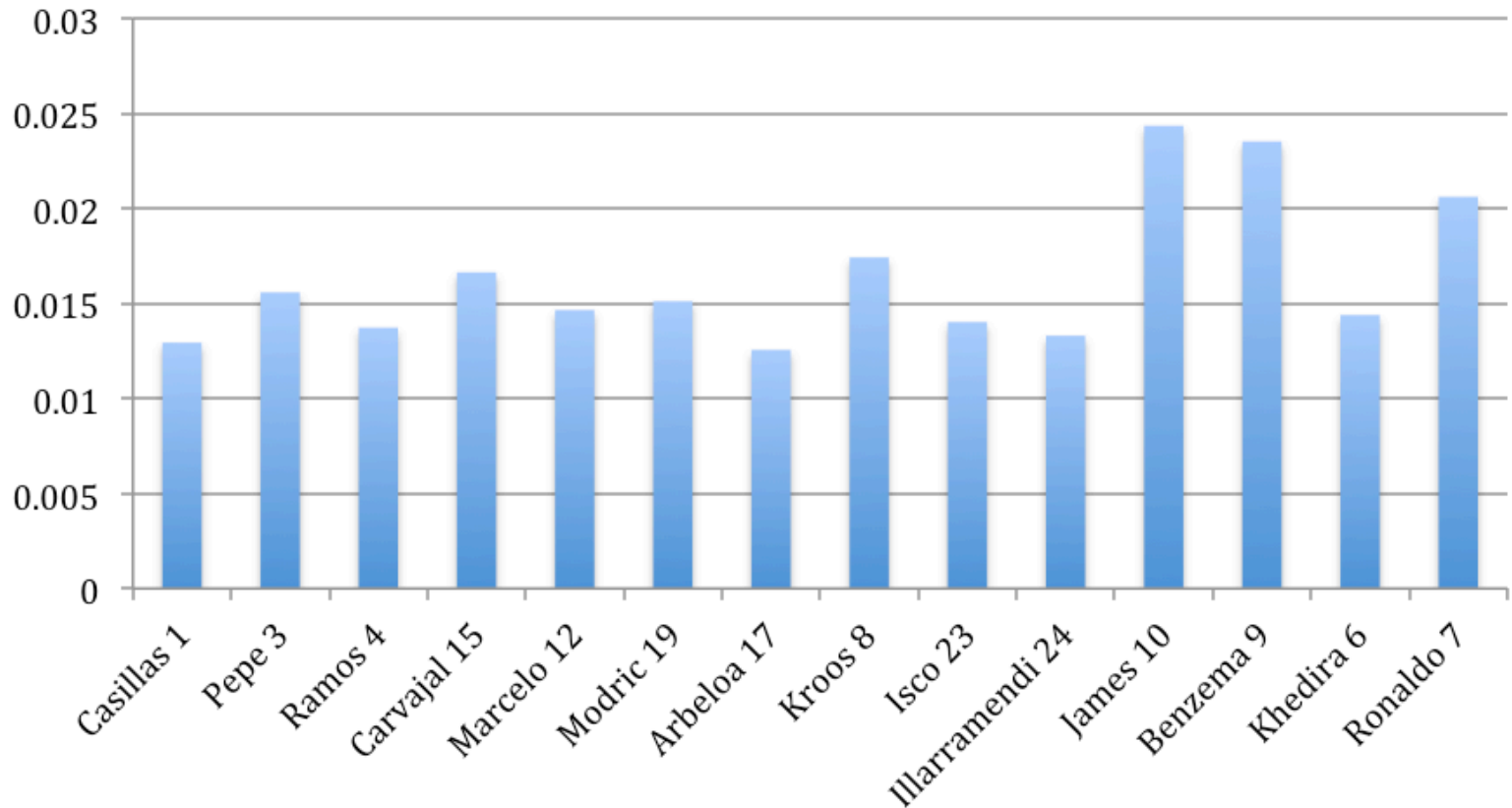
Point Findings

- **Expected Goals per Game using Markov Method: 2.37 Goals**
- Expected Amount of Goals per Possession: 0.02854
- Super Batter vs. Super Pitcher: GK Claudio Bravo vs. Real Madrid
 - Bravo's Goals per Shot rate: 0.056
 - League Average Goals per Shot: 0.1136
 - **Expected Goals per Game: 1.30 Goals**

$$\frac{B^*}{1 - B^*} = \frac{B}{1 - B} \frac{1 - B_L}{B_L} \frac{B_P}{1 - B_P}$$

Expected Goals Per Possession Per Starting Player (SB/SP)

Expected Goals/Possession Starting with Player



In Words

Cristiano Ronaldo is a ball hog--- Holds the ball 22% of his actions.

James Rodriguez was the most immediately deadly player in the game.

Dani Carvajal was the most important setup man in this game.

Sergio Ramos was the least important player offensively.

Of players, the three subs were among the bottom 5 in effectiveness.

This analysis was very interesting for soccer, but may be better used in other sports, most notably basketball, due to much shorter possession lengths.

References

- YouTube footage of El Clásico
- WhoScored.com
- FoxSports.com
- Wikipedia.com