Article

Order out of chaos: emergent patterns in soccer matches

Alessandro Ferrarini
Department of Evolutionary and Functional Biology, University of Parma, Via G. Saragat 4, I-43100 Parma, Italy
E-mail: sgtpm@libero.it, alessandro.ferrarini@unipr.it

Received 27 December 2015; *Accepted 10 February 2016*; *Published online 1 June 2016*

Abstract
A soccer match is a very complex, apparently chaotic, human event. However, order clearly emerges from such chaos if we have at hand the right tools to extract pattern configurations. In this work, a new algorithm called *Soccer-Decompiler* is presented, which is able to analyze soccer matches and extract emerging patterns from apparently chaotic sequences of events. Detecting and filtering the frequencies of events is used by *Soccer-Decompiler* to discover such patterns. The application of *Soccer-Decompiler* to a real soccer match shows that order out of chaos in complex human events can be effectively extracted by isolating highly frequent events. An applicative example is given.

Keywords chaotic patterns; complexity; emergent behaviors; frequencies filtering; order; pattern configurations; real life analysis; soccer analysis; Soccer-Decompiler.

Selforganizology
ISSN 2410-0080
URL: http://www.iaees.org/publications/journals/selforganizology/online-version.asp
RSS: http://www.iaees.org/publications/journals/selforganizology/rss.xml
E-mail: selforganizology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

"Soccer decoding" has been introduced by Ferrarini (Ferrarini, 2014) to indicate the use of scientific tools to: 1) reduce the complexity of a soccer match to its irreducible essence using a game theory algorithm, 2) simulate soccer matches by adding iteration and stochasticity to such structural essence. *Soccer-Decoder* (Ferrarini, 2012a; Ferrarini, 2015) is a math algorithm, implemented through the software *Soccer-Lab* (Ferrarini, 2012b), that simulates soccer matches by merging together 3 scientific methods: game theory, differential calculus and stochastic simulations. The rationale behind *Soccer-Decoder* is that even very complex real world events, when turned into their irreducible essence, can be understood and predicted.

In this work, a twin algorithm called *Soccer-Decompiler* (Ferrarini, 2011) is presented, which is able to analyze soccer matches and extract emerging patterns from apparently chaotic sequences of events. The application of *Soccer-Decompiler* to a real soccer match shows that order out of chaos in complex human events can be extract by detecting and filtering highly frequent events. An applicative example is given.

2 *Soccer-Decompiler*

Soccer-Decompiler (Ferrarini 2011) is both a statistical algorithm and a software to analyze soccer events. It receives data from both keyboard and vocal inputs. From a computation viewpoint, it:

a) turns input data into strings;

b) analyzes frequencies of such strings;

c) filters the most frequent events;

d) tests frequencies versus null hypotheses.

So doing, *Soccer-Decompiler* detects the most frequent patterns and tests them to compute their statistical significance. Data are inputted using an *ad hoc* code called *Soccer-encipher* (Ferrarini, 2011).

3 An illustrative example

A real soccer match (90 minutes length) has been analyzed using the *Soccer-Decompiler* algorithm/software. Only one of the two teams has been studied, the amateur team “Montebello72” that allowed to record through a digital camera and then analyze the friendly match played in the Cittadella public park in Parma on August 18th 2015 versus the team “Montebello71”. Data were inserted using vocal input.

First, *Soccer-Decompiler* detected the number of actions and the percentage of positive (shots on goal), negative (balls lost) and neutral (the action was interrupted but the ball remained to the studied team) actions (Table 1).

The analyzed team developed 110 actions during the match, of which 22 positive (shots on goal). It results clear that the first half was better for both total actions (60 vs. 50) and positive ones (13 vs. 9).

Table 1 Number and outcome of the actions developed by the studied team.

Actions	total	%	1st half	%	2nd half	%
number of actions	110	100	60	54.55	50	45.45
positive	22	20.00	13	59.09	9	40.91
negative	70	63.64	37	52.86	33	47.14
neutral	18	16.36	10	55.56	8	44.44

Second, *Soccer-Decompiler* detected action lengths (Table 2). The analyzed team developed prevalently short actions with 2 passages (25 actions) or just 1 passage (19 actions). It should be noted that short actions were sometimes decided by the opponent team that stopped the action of the analyzed team.

Table 2 Action lengths of the studied team.

Action lengths	total	%	1st half	%	2nd half	%
0 passages	14	12.73	11	18.33	3	6.00
1 passage	19	17.27	9	15.00	10	20.00
2 passages	25	22.73	12	20.00	13	26.00
3 passages	12	10.91	6	10.00	6	12.00
4 passages	10	9.09	6	10.00	4	8.00
5 passages	4	3.64	3	5.00	1	2.00
6 passages	9	8.18	3	5.00	6	12.00
7 passages	3	2.73	1	1.67	2	4.00
8 passages	8	7.27	5	8.33	3	6.00
9 passages	3	2.73	2	3.33	1	2.00
10 passages	1	0.91	1	1.67	0	0.00
11 passages	1	0.91	0	0.00	1	2.00
12 passages	0	0.00	0	0.00	0	0.00

13 passages	0	0.00	0	0.00	0	0.00
14 passages	0	0.00	0	0.00	0	0.00
15 passages	1	0.91	1	1.67	0	0.00
16 passages	0	0.00	0	0.00	0	0.00
17 passages	0	0.00	0	0.00	0	0.00
18 passages	0	0.00	0	0.00	0	0.00
19 passages	0	0.00	0	0.00	0	0.00
20 passages	0	0.00	0	0.00	0	0.00

Third, *Soccer-Decompiler* found the passages among the team divisions (Fig. 1). The analyzed team used an isotropic play where passage flows were common in all the directions. Passages from the four defenders to the three midfielders (45) and vice versa (48) were a bit more common than others. The three forwards tended to scarcely interact (only 27 passages among them), while they tended to prefer retro-passages to midfielders (39).

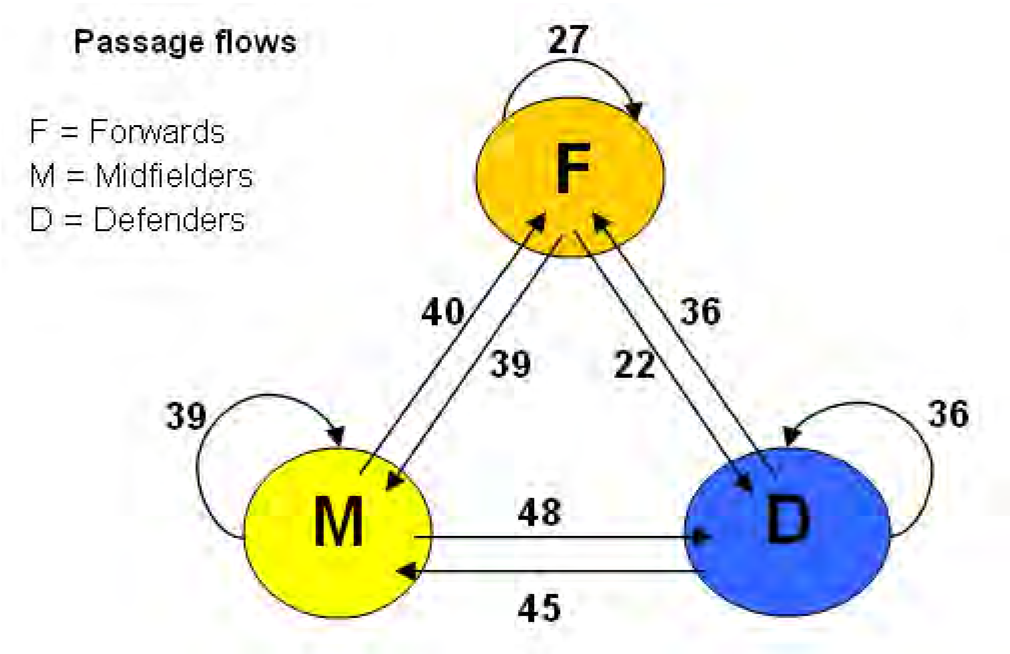

Fig. 1 Passage flows.

Fourth, *Soccer-Decompiler* found where ball was received by players (Table 3). For each player (on the row), numbers refer to the ball receipts in each zone of the field (see figure on the right). For instance, Def1 (right defender) received 21 balls in position 4 and 16 balls in position 7. Midf1, Midf3 and Forw2 were the only players without a clear position of ball receipt. Zones 5 and 8 were the dominant field zones for ball receipt.

Fifth, *Soccer-Decompiler* detected where ball was released by players (Table 4). For each player (on the row), numbers refer to the ball releases in each zone of the field. For instance, Def1 (right defender) released 20 balls in position 4 and 17 balls in position 7. Again, Midf1, Midf3 and Forw2 were the only players without a clear position of ball release. Zones 5 and 8 were the dominant field zones also for ball release.

Table 3 Positions of ball receipt. Numbers in bold refer to dominant values for each player.

ball receipt	Field zones (see image on the right)											
	1	2	3	4	5	6	7	8	9	10	11	12
def1	0	1	0	**21**	2	0	16	0	0	2	0	0
def2	0	0	0	2	**19**	0	0	1	0	0	1	0
def3	0	0	0	0	**12**	1	0	2	0	0	0	0
def4	0	0	1	0	2	**25**	0	1	23	0	2	1
midf1	0	0	0	6	14	0	**16**	11	0	0	0	0
midf2	0	1	0	0	**27**	3	0	19	0	0	0	0
midf3	0	0	0	0	11	2	0	15	**17**	0	0	1
forw1	0	0	0	0	8	1	0	**27**	2	4	4	3
forw2	0	0	0	1	3	5	3	**16**	15	1	1	3
forw3	0	1	0	0	3	0	1	**10**	0	2	2	0
sub1	0	0	0	0	0	0	0	**2**	0	0	0	0
sub2	0	1	0	0	**8**	0	0	1	0	0	0	0
sub3	0	0	0	0	1	0	0	**2**	1	0	1	0
TOTAL	1	6	4	34	**115**	43	43	**115**	67	19	22	20

Table 4 Positions of ball release. Numbers in bold refer to dominant values for each player.

ball release	Field zones (see image on the right)											
	1	2	3	4	5	6	7	8	9	10	11	12
def1	0	0	0	20	1	0	17	1	0	2	0	0
def2	0	0	0	2	**19**	0	0	1	0	0	1	0
def3	0	0	0	0	**12**	1	0	2	0	0	0	0
def4	0	0	1	0	2	24	0	1	24	0	2	1
midf1	0	0	0	5	11	0	15	**16**	0	0	0	0
midf2	0	1	0	0	**27**	3	0	19	0	0	0	0
midf3	0	0	0	0	11	2	0	16	16	0	1	0
forw1	0	0	0	1	6	1	1	24	6	4	4	3
forw2	0	0	0	0	4	3	3	17	15	1	1	3
forw3	0	1	0	0	3	0	0	**10**	0	2	2	0
sub1	0	0	0	0	0	0	0	**2**	0	0	0	0
sub2	0	1	0	0	**8**	0	0	1	0	0	0	0
sub3	0	0	0	0	0	0	0	3	1	0	1	0
TOTAL	1	5	4	32	**109**	40	43	**121**	71	19	23	19

Table 5 Passage flows (full match).

TO (receivers)	FROM (transmitters)														
	def1	def2	def3	def4	midf1	midf2	midf3	forw1	forw2	forw3	sub1	sub2	sub3	balls received	transmitters
def1	0	4	2	1	9	7	1	5	1	1	0	2	1	34	11
def2	6	0	2	2	1	1	1	0	0	1	0	2	0	16	8
def3	0	4	0	2	0	4	1	0	0	0	0	0	0	11	4
def4	0	4	4	0	2	9	10	6	5	0	0	0	1	41	8
midf1	12	2	2	0	0	9	5	8	3	3	0	1	0	45	9
midf2	4	1	2	6	8	0	5	4	7	1	0	1	0	39	10
midf3	1	1	3	8	5	7	0	4	7	1	0	1	1	39	11
forw1	6	0	0	7	8	8	4	0	9	4	0	0	1	47	8
forw2	1	0	0	15	3	2	5	5	0	1	1	0	0	33	8
forw3	3	0	0	2	5	0	1	5	0	0	0	0	0	16	5
sub1	0	0	0	1	0	1	0	0	0	0	0	0	0	2	2
sub2	0	1	0	1	0	2	0	1	0	0	0	0	0	5	4
sub3	0	0	0	1	1	0	2	0	1	0	0	0	0	5	4
balls passed	33	17	15	46	42	50	35	38	33	12	1	7	4		
receivers	7	7	6	11	9	10	10	8	7	7	1	5	4		

Sixth, *Soccer-Decompiler* uncovered the from-to nature of passages (Table 5). For each player (on the column; transmitter) numbers refer to the passages to player on the row (receiver). For instance, Def4 prevalently passed to Forw2, while Def1 to Midf1. Midf2 was the best transmitter (50 passages), while Forw1 was the best receiver (47 balls received).

If we consider separately the two halves of the match, passage flows are described in Tables 6 and 7. *Soccer-Decompiler* can produce chi squared tests to test if:

- matrix A (first half) is significantly different from matrix B (second half);
- there are significant (higher than by chance) lines of passage within each matrix.

Table 6 Passage flows (first half).

	FROM (transmitters)														
TO (receivers)	def1	def2	def3	def4	midf1	midf2	midf3	forw1	forw2	forw3	sub1	sub2	sub3	balls received	transmitters
def1	0	3	2	1	6	5	1	3	0	1	0	0	0	22	8
def2	6	0	2	0	0	1	1	0	0	0	0	0	0	10	4
def3	0	4	0	2	0	4	1	0	0	0	0	0	0	11	4
def4	0	1	4	0	0	5	5	4	4	0	0	0	0	23	6
midf1	6	2	2	0	0	6	3	2	2	3	0	1	0	27	9
midf2	2	1	2	3	4	0	3	2	5	1	0	0	0	23	9
midf3	1	0	3	4	4	4	0	3	2	0	0	0	0	21	7
forw1	1	0	0	5	6	4	3	0	5	3	0	0	0	27	7
forw2	1	0	0	8	3	1	1	4	0	0	1	0	0	19	7
forw3	3	0	0	1	3	0	1	3	0	0	0	0	0	11	5
sub1	0	0	0	1	0	1	0	0	0	0	0	0	0	2	2
sub2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
sub3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
balls passed	20	11	15	25	26	31	19	21	18	8	1	1	0		
receivers	7	5	6	8	6	9	9	7	5	4	1	1	0		

Table 7 Passage flows (second half).

	FROM (transmitters)														
TO (receivers)	def1	def2	def3	def4	midf1	midf2	midf3	forw1	forw2	forw3	sub1	sub2	sub3	balls received	transmitters
def1	0	1	0	0	3	2	0	2	1	0	0	2	1	12	7
def2	0	0	0	2	1	0	0	0	0	1	0	2	0	6	4
def3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
def4	0	3	0	0	2	4	6	2	3	0	0	0	1	21	7
midf1	6	1	0	0	0	3	2	6	2	0	0	1	0	21	7
midf2	2	0	0	3	4	0	3	2	2	0	0	1	0	17	7
midf3	0	1	0	5	1	4	0	1	6	1	0	1	1	21	9
forw1	5	0	0	2	2	4	2	0	4	1	0	0	1	21	8
forw2	0	0	0	7	1	1	4	1	0	1	1	0	0	16	7
forw3	0	0	0	1	2	0	0	2	0	0	0	0	0	5	3
sub1	0	0	0	1	0	1	0	0	0	0	0	0	0	2	2
sub2	0	1	0	1	0	2	0	1	0	0	0	0	0	5	4
sub3	0	0	0	1	1	0	2	0	1	0	0	0	0	5	4
balls passed	13	7	0	23	17	21	19	17	19	4	1	7	4		
receivers	3	5	0	9	9	8	6	8	7	4	1	5	4		

In addition, *Soccer-Decompiler* can export its results to the free software Netdraw (Borgatti 2002) in order to produce the network graphs of the gameplay (Fig. 2). Many properties of the team's tactic become evident if only the most frequent events are isolated. For instance, Def2, Def3 and Forw3 scarcely took part to the play. Def4 to Forw2 was a very common play (15 times). Midf2 preferred the right side of the field, Midf3 preferred an easy backward play towards Def4, Forw2 prevalently sought Forw1 etc.

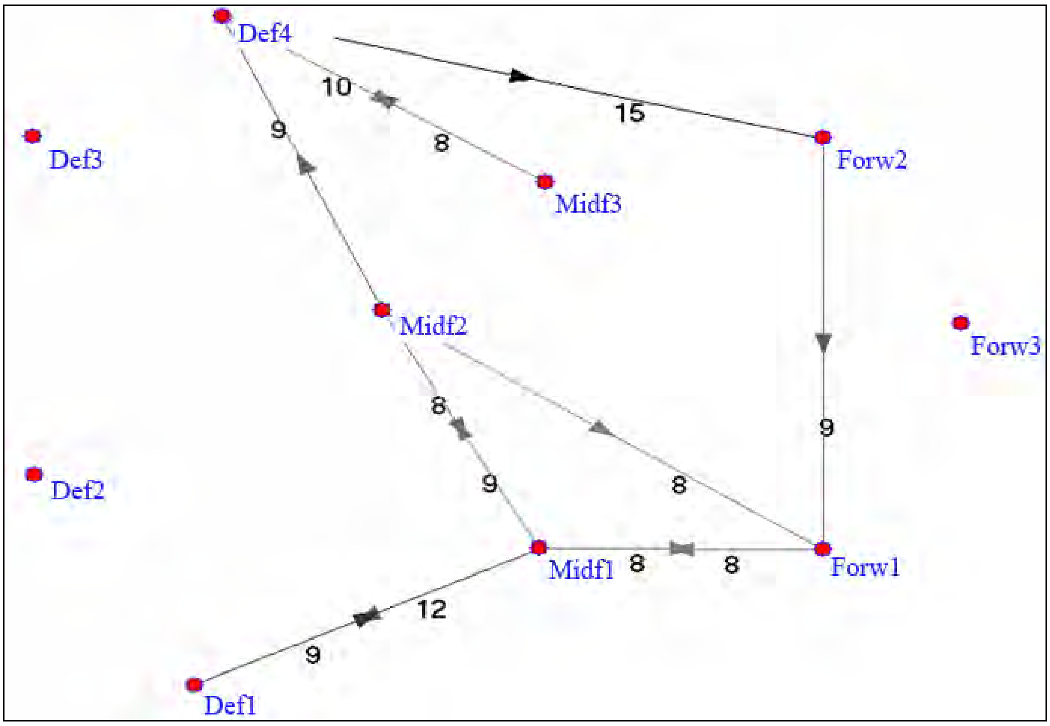

Fig. 2 Passage flow during the full match. Only links with 8 or more passages are showed.

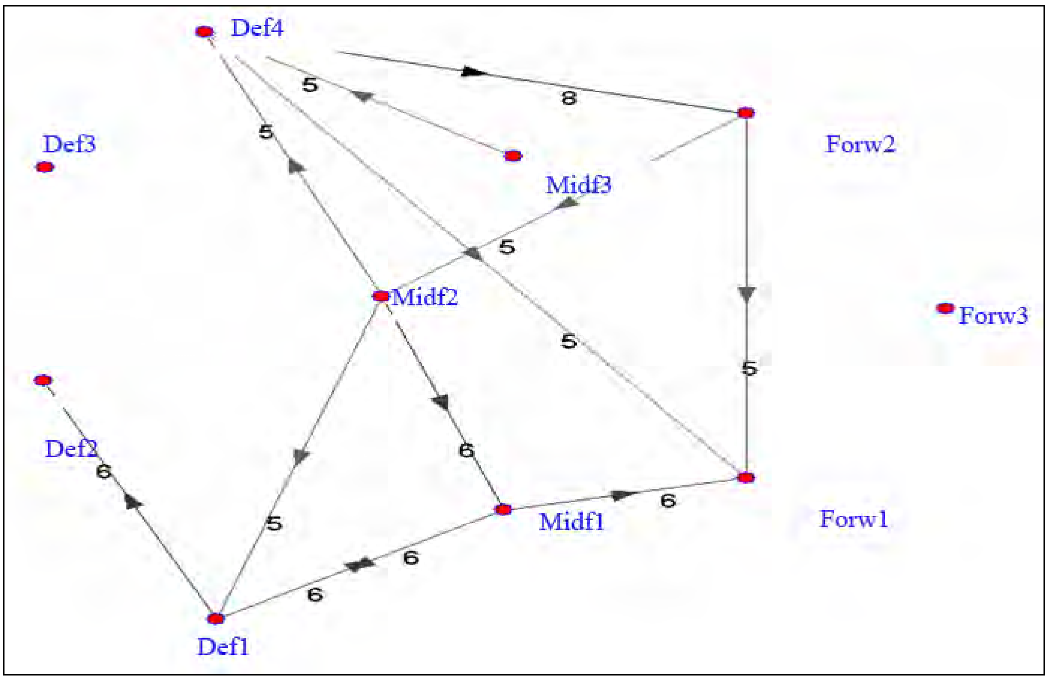

Fig. 3 Passage flow in the first half of the match. Only links with 5 or more passages are showed.

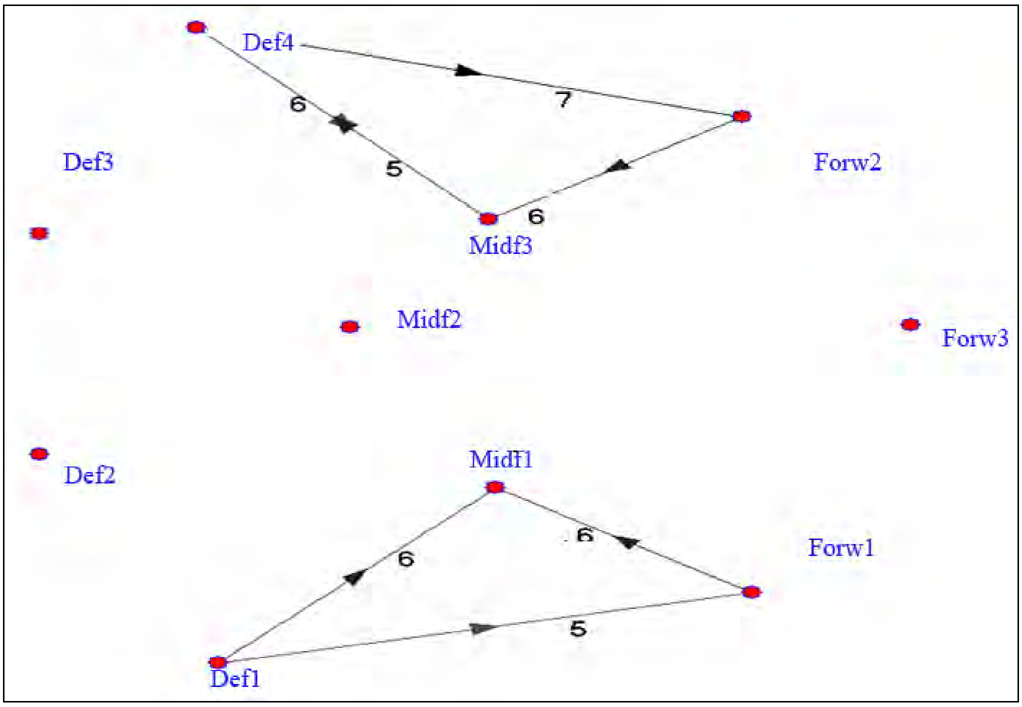

Fig. 4 Passage flow in the second half of the match. Only links with 5 or more passages are showed.

If we separately consider the two halves of the match, differences are evident. During the first half, the studied team used a complex web of passages (Fig. 3). Instead, during the second half the analyzed team applied a very simple and neat web of passages involving separately the left and right sides of the field (Fig. 4). The change in tactic was due to the fact that the studied team scored at the end of the first half, and defended its advantage in the second half.

4 Conclusions

A soccer match is a very complex, apparently chaotic human event. However, order clearly emerges from such chaos if we have at hand the right tools to extract pattern configurations.

Detecting and filtering the frequencies of events has been used here by the algorithm *Soccer-Decompiler* to discover such patterns. Results presented in this work show that this algorithm is on top of discovering emergent behaviors at both player and team level. *Soccer-Decompiler* is also a scientific software with both keyboard and vocal inputs that can be applied to any soccer event, both amateur and professional.

Acknowledgement

I thank the amateur team “Montebello72” that allowed me to record through a digital camera and analyze the friendly match played in the Cittadella public park in Parma on August 18th 2015 versus the team “Montebello71”.

References

Borgatti SP. 2002. NetDraw network visualisation. Analytic Technologies, Harvard, MA, USA

Ferrarini A. 2011. *Soccer-Decompiler*: A statistical algorithm and a software to decompile soccer matches. Manual, 84 pages (in Italian)

Ferrarini A. 2012a. *Soccer-Decoder*: A math engine for simulating soccer matches. Manual, 82 pages (in Italian)

Ferrarini A. 2012b. *Soccer-Lab*: A software simulator of soccer matches. Manual, 85 pages (in Italian)

Ferrarini A. 2014. A new game theory algorithm simulates soccer matches: Reducing complexity to its irreducible essence. Selforganizology, 1(3-4): 211-218

Ferrarini A. 2015. Connecting game theory and evolutionary network control for the computational control of soccer matches. Selforganizology, 2(1): 1-12