Knowledge-Based Systems 21 (2008) 551–562

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A compound framework for sports results prediction: A football case study

Byungho Min, Jinhyuck Kim, Chongyoun Choe, Hyeonsang Eom *, R.I. (Bob) McKay

School of Computer Science and Engineering, Seoul National University, San 56-1, Sillim-dong, Gwanak-gu, Seoul 151-742, Republic of Korea

ARTICLE INFO

Article history:
Received 21 February 2007
Accepted 21 March 2008
Available online 29 March 2008

Keywords:
Rule-based reasoning
Bayesian inference
Sports prediction
Football
In-game time-series approach
Simulation

ABSTRACT

We propose a framework for sports prediction using Bayesian inference and rule-based reasoning, together with an in-game time-series approach. The framework is novel in three ways. The framework consists of two major components: a rule-based reasoner and a Bayesian network component. The two different approaches cooperate in predicting the results of sports matches. It is motivated by the observation that sports matches are highly stochastic, but at the same time, the strategies of a team can be approximated by crisp logic rules. Furthermore, because of the rule-based component, our framework can give reasonably good predictions even when statistical data is scanty: it can be used to predict results of matches between teams which have had few previous encounters. Machine learning techniques have great difficulty in handling such situations of insufficient data. Second, our framework is able to consider many factors, such as current scores, morale, fatigue, skills, etc. when it predicts the results of sports matches: most previous work considered only one factor, usually the score. Third, in contrast to most previous work on sports results prediction, we use a knowledge-based in-game time-series approach to predict sports matches. This approach enables our framework to reflect the tides/flows of a sports match, making our predictions certainly more realistic, and somewhat more accurate. We have implemented a football results predictor called FRES (Football Result Expert System) based on this framework, and show that it gives reasonable and stable predictions.

1. Introduction

Predicting the results of sports matches is interesting to many, from fans to punters. It is also interesting as a research problem, in part due to its difficulty: the result of a sports match is dependent on many factors, such as the morale of a team (or a player), skills, coaching strategy, etc. So even for experts, it is very hard to predict the exact results of individual matches. It also raises very interesting questions, regarding the interaction between the highly stochastic nature of the game itself, and the highly structured nature of the rules and of the strategies that may be employed. Equally important, there is a huge amount of highly structured data available to researchers.

In this paper we present a novel framework for sports results prediction. Our framework is novel in a number of ways. First, we take a compound perspective; a rule-based reasoner [30] and Bayesian networks [21] cooperate very closely in the framework. Roughly stated, we formalize sports strategies using a rule-based reasoner, and handle uncertainty with Bayesian networks. Through this compound approach we can take into account the fact that most sports results are highly stochastic, but at the same time, the strategies of a team (or a player) can be represented by crisp logic rules.

Second, when it predicts the results of sports matches, our framework considers many factors, such as current scores, morale, fatigue, skills, etc. By contrast, most previous work considered only one factor, usually the score, or at most a few factors. We are motivated by the widely accepted assumption that the accuracy of prediction in non-trivial prediction domains (such as sports) can be improved if the many factors affecting the prediction results are properly considered. When people predict something complex they generally try to consider the many factors that affect the results or outcomes they want to predict.

Third, we propose to use an in-game time-series approach. Most sports have tides and flows, situations in the match change over time. Our approach is designed to reflect these tides and flows. In this respect, our framework can be viewed as a simulator for a sports match.

Fourth, our system is stochastic, so the results from different runs may vary. We take a Monte–Carlo approach to evaluating the overall results from the system. The system is run a number of times, and the results aggregated.

Prediction of Association Football (soccer) matches is an interesting and difficult problem for knowledge-based applications, raising new issues of representation and acquisition. Most previous

* Corresponding author. Tel.: +82 2 880 6755; fax: +82 2 871 4912.
E-mail addresses: bhmin@meslab.snu.ac.kr (B. Min), jhkim@meslab.snu.ac.kr (J. Kim), joesmith@3map.snu.ac.kr (C. Choe), hseom@cse.snu.ac.kr (H. Eom), rim@cse.snu.ac.kr (R.I. (Bob) McKay).

0950-7051/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2008.03.016

work on this problem has treated the individual match as atomic, applying statistical or machine learning techniques to generate forecasts, and hence has not needed to address these issues. By contrast, we aim to model the game process in more detail.

In this paper, we introduce a system, FRES (Football Result Expert System), which embodies our framework in the domain of soccer. In this implementation, we divide a single soccer match into ten time frames to apply our in-game time-series approach. And in each frame of a match, we attempt to model the reasoning that might be carried out by a good head coach. Both teams infer the current state of the match and derive corresponding strategies. This can be viewed as a knowledge-based in-game time-series approach; using it enables FRES to give realistic, and somewhat more accurate, predictions.

The remainder of this paper is organized as follows: Section 2 gives a summary of previous work on football prediction and their limitations. In Section 3, we describe our framework and its implementation, FRES, explaining how our rule-based reasoner and Bayesian networks interact. We present the in-game time-series approach of our framework. We detail how we applied the framework to soccer, and the relevant aspects of rule-based reasoning and Bayesian inference in FRES. In Section 4, we present our experimental evaluation and analysis. For evaluation, we use the results of the World Cup 2002, comparing two historic predictors with the predictions of FRES. Finally, conclusions and future work are given in Section 5.

2. Background

Association Football (soccer) is perhaps the World's pre-eminent sport, so it is not surprising that there has been a substantial amount of research on soccer prediction. Actually, among all sports, soccer prediction is probably the most widely and deeply researched area. We thus survey prediction-related researches for soccer, as representative of the sports in our target domain, and categorize them into a few groups. Although the studies mostly deal only with mathematical/statistical models or methods, they are important background for building a soccer result expert system.

Statistical analysis forms one of the major strands on soccer prediction. Researchers have suggested a range of models or processes to analyze the results of soccer matches. They generally aim to show that their models or methods represent the results of soccer matches well, i.e. fit well. Many models and methods, such as Poisson regression models, a logistic regression model using seed positions, and an updating process for the intra-match winning probability, were used to analyze/interpret the results of soccer matches [8,9,18,22]. While most of these studies did provide some predictions, they are more focused on statistical analysis of the results of soccer matches.

Statistical prediction of soccer results has also been heavily researched. The general outline of these approaches is as follows: models are developed by fitting them to real data – at this step, additional information, knowledge, or assumptions are often defined and used. Then the models are used to predict the results of soccer matches – actually this process is similar to many machine learning approaches. Some of the studies in this area took more statistical approaches in predicting soccer results – they use little prior knowledge/information and are heavily based on pure statistical models, such as ordered probit and Poisson models [19,20]. Koning [19] estimated the quality of soccer teams. Koning et al. [20] estimated the expected number of goals in a soccer match based on a scoring intensity measure, predicting the probability that a team wins a tournament using the estimated scoring intensities. Other work used models or methods that are more dependent on prior information or knowledge about soccer matches: a Bayesian dynamic generalized linear model to estimate and predict the time dependent skills of teams [29], an ordered probit regression model with match venue information [10,11], an independent Poisson model with a method of modeling a team's offensive and defensive strengths [7], and a ranking system based on the seasonal coefficients of variation (CVs) of the end-of-season points [12,13], have all been used to predict the results of soccer matches.

Machine learning techniques [26] and related methods have been applied to soccer prediction. Tsakonas et al. [31] separately studied fuzzy models, neural networks (NN) and genetic programming (GP). Although only limited experimental details were given, they concluded that GP out-performed the other two (though at higher computational cost). They appear to have concluded that fuzzy models performed better than NN, though we have not been able to follow how the data led them to that conclusion. Rotshtein et al. [28] integrated some of these ideas, using fuzzy logic representation and applying genetic and neural optimization methods to tune the fuzzy model. Joseph et al. [17] compared a subjective, expert-derived Bayesian Network (BN), decision tree learning, naïve Bayesian learning, a data-driven BN, and K-nearest-neighbour learning. They concluded that "the expert BN is generally superior to the other techniques". All these studies relied on previous match results, i.e. win/draw/lose or scores, as training data and forecasted the results of a league or tournament matches. The expert BN also relied on information about the playing status (presence or absence, position) of some key players, and on home-ground advantage.

Expert knowledge, despite the positive findings of [17], does not seem to have been widely used; the only other system based clearly on expert knowledge we have encountered is [4], which incorporates knowledge from many experts, using evidence theory as an integrating framework.

Prediction-related research is not the only direction in soccer. A number of studies concentrate on how to determine strategies [3,5,15,27]. [3] described an interactive stochastic simulation model intended to assist collegiate soccer strategists in the design and evaluation of various offensive strategies. [5] discussed the home-field effect in professional team sports and provided further evidence of home advantage in the English Premier League by suggesting a match-based production function. [27] developed a game-theoretic model of a soccer match in which the optimal strategy of a team depends on the current state of the game and concluded that teams' skills, current score, and home advantage are significant explanatory variables of the probability of scoring. [15] proposed a game theoretic approach to modeling tactical changes of formation in a soccer match and demonstrated that the decisions of each team's head coach affects the probability of winning the match, using data from the Japan professional soccer league.

A few previous studies specifically addressed tournament-specific issues and/or directly modeled soccer tournaments [2,9,12,13,20,22]. Of course, other sports rather than soccer, such as American football and major league baseball, have been the target of several studies [1,23–25,32].

Among the work discussed, the approach of [28] is most similar to ours. They use logic and rules to predict the results of soccer matches. They differ from us in using fuzzy logic, whereas our framework uses crisp rule-based reasoning, but more importantly, in only considering matches at the whole-game level. The latter leads to the primary distinction between our framework and the other systems detailed above, in the use of non-score attributes. Most of the previous studies consider score information as the only factor influencing the results of future soccer matches. Even where additional information or knowledge is used, the score information is always taken as the most important and dominant determinant of results. But soccer games generate huge amounts of readily-accessible data, not just score data. Our framework and FRES are unique among the above systems in making use of these additional data as the primary factors that influence the results of soccer matches. This has two major advantages. Scores in soccer are generally low, and highly stochastic. What is more, data about the relative performance of teams may be quite sparse – there is a huge amount of playing data available, but meetings between particular teams are still infrequent events. Moreover the composition of teams is constantly changing, hence the relevance of data ages relatively rapidly. Making use of this additional data is an important contributor to the effectiveness of FRES. Another large distinction is the use of an in-game time-series approach. As mentioned earlier, our framework can reflect tides/flows of a sports match by using this approach. These distinctions are detailed in the following sections.

To summarize, we note first (following Joseph et al. [17]) that expert knowledge can significantly out-perform machine learning systems in predicting football results. However their expert approach was hamstrung by the volatility of the knowledge they incorporated: even by the time of publication, the relevant players had left the team, so that the system could no longer predict correctly. Our aim is to base our system on more generic knowledge which might be more robust to changing circumstances. Second, following Hirotsu et al. [15] we emphasise the critical importance of the strategic decisions of the head coach. Finally, we follow the vast bulk of these systems in incorporating uncertainty, though in a different way to these previous systems.

None of the systems mentioned above take full advantage of the information available about soccer games. By and large, they also ignore the characteristics of the particular sport: there are many important factors in soccer that affect the result of a soccer match. In particular, the strategies that a head coach decides are significant factors. Furthermore, soccer is a very dynamic sport, i.e. the situations and tides of a soccer match change dynamically over time, and these dynamics figure crucially in determining strategies, and hence winners and losers. Lastly, soccer results are highly stochastic. Our framework is designed to reflect these characteristics of soccer.

3. Overview of FRES

In this section, we give an overview of our system, FRES. FRES concentrates on a particular kind of soccer forecasting, namely tournament prediction. It infers the results of all soccer matches in a tournament, and shows its full inferences. We chose to address tournament prediction because of the enormous importance of the World Cup tournament in soccer, the consequent huge interest, and the related availability of high-quality data. FRES is evaluated against World Cup tournament data. We use Jess [16] and JavaBayes [14] together to represent knowledge explicitly. These modules and an overview of our implementation are given in more detail in Appendix C.

The architecture of FRES is shown in Fig. 1. The user runs the system, and receives the results of the tournament – i.e. which team wins the tournament, which teams advance to semi-finals, which to quarter-finals, etc. Jess and JavaBayes communicate via Java, and Java records all the outputs of each module in a file (see Fig. 2).

The detailed modelling has two consequences on knowledge representation and acquisition.

First, we need to model the uncertainty of soccer. Many approaches are available; we have chosen to use Bayesian modelling. Bayesian methods offer the particular advantage of readily incorporating both human knowledge about uncertainty (in the form of priors) and evidential knowledge learnt from data. While we take only very limited advantage of this at present, it is our intention to extend this aspect in future.

Second, we need to model the strategic decisions of the soccer coach. These decisions have a huge impact on the course of the game. In principle, we could try to model the whole process of the coach's reasoning. We don't believe this is likely to be successful. It would lead us into deep issues of human commonsense reasoning, far beyond the scope of this work. Even if these problems could be solved, they would only lead into even more insuperable issues of knowledge acquisition. We would need to model, not just general coaching knowledge, but the detailed knowledge of each individual coach – knowledge which is likely largely unconscious, and moreover, knowledge which the individual coach has strong motivation not to reveal. As a result, we have chosen to model knowledge at the level of the generally knowledgeable soccer fan, and to encode it into the relatively simple but accessible form of predictive production rules.

3.1. Knowledge acquisition

Given our pragmatic decision, to limit the knowledge modelling to the level of a knowledgeable fan, there was a limit to how sophisticated knowledge acquisition strategy was needed. In the event, we chose to base the rules primarily on interviews with a renowned soccer expert, Mr. Hyeongwook Seo, a well-known Korean

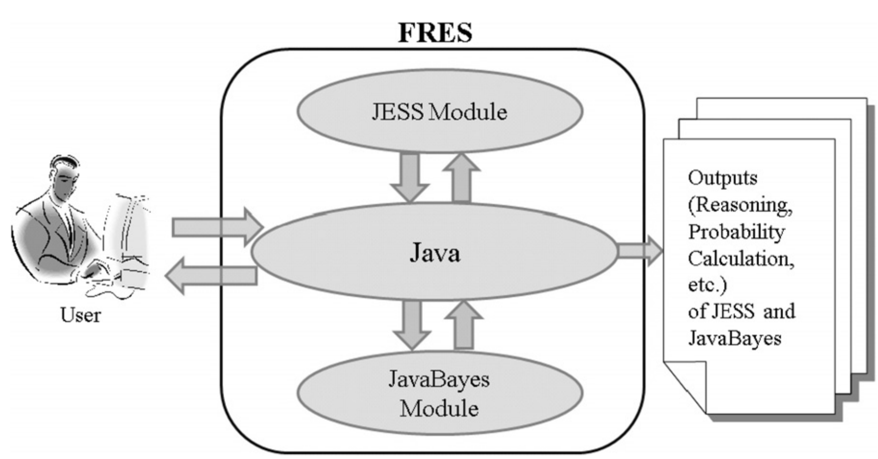

Fig. 1. Architecture of FRES.

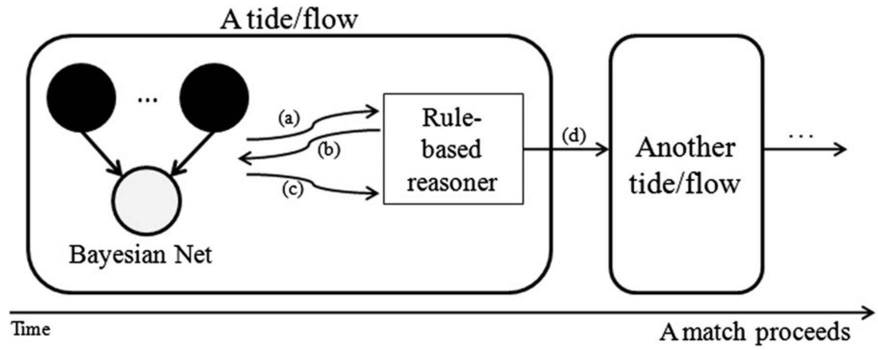

Fig. 2. Overview of the proposed framework: (a) Bayesian network calculates probabilities of each factor (black) node; at this step, the grade (white) node is fixed with the highest value. Then the values of each factor are passed to rule-based reasoner. (b) Rule-based reasoner reasons and decides strategies according to the values of each factor. Then the determined strategies are passed back to Bayesian network. (c) Bayesian network calculates probabilities of the goal node's values using the determined strategies that include the values of each factor (black) node. Then the grade value is passed to rule-based reasoner. (d) Rule-based reasoner reasons the results of current tide/flow.

football commentator who received his football industries MBA from the University of Liverpool graduate school; he is, especially, an expert on global football events, such as the World Cup and UEFA Champions league. The knowledge obtained from these interviews was subsequently validated and adapted based on a number of more accessible sources:

- a survey (Table 1) of a group of ten people; these people have strong interest in, and knowledge about, soccer
- articles and books on soccer
- web pages and data bases on soccer
- common knowledge of soccer (primarily based on the knowledge of the first three authors)

3.2. In-game time-series approach

Most previous work on soccer result prediction has taken a time-series approach using a match base, i.e. the results of previous matches themselves are used as the primary factor in predicting a match. Actually, in most systems, the previous results, i.e. win/draw/lose, scores, or difference in scores, are the only factors they consider. Our framework, however, considers tides/flows in a match. In doing so, we can divide a match into smaller time frames, and see for each of them whether there is any change in tides/flows. For example, if a team scores a goal in a certain time frame of a soccer match, the team's morale can be raised and the change can affect the subsequent tides/flows of the match. We can detect and utilize such changes in tides/flows by checking each time frame.

If a game is to be modeled in temporal detail, then we need to determine an appropriate level of granularity. At one extreme, we could simply take the formal game divisions (two halves in Association Football, four quarters in Australian Football, etc.), but this granularity is unlikely to be fine enough: coaches make strategic decisions far more often, and these decisions have significant impact on the progress of the match. At the other extreme, we could model the match at the level of individual plays. This level of granularity would require modelling the decisions and skills of individual players, a task beyond our current capabilities. We address the issue of modelling granularity in more detail later in this section.

FRES takes an in-game time-series approach and realizes the tides/flow by dividing a single match into ten time frames. The number of time frames was chosen based on several experiments. However this was non-trivial exercise. Because of the dependencies between components in the in-game time-series approach, once the number of time frames changes, the Bayesian networks and rules must be changed correspondingly. For example, we have a rule which encodes knowledge that more goals are scored near the beginning and the end of a match. This kind of rule requires modification when the time scale changes. In fact, the rules may not merely change in detail. We might need to add or subtract rules when the number of frames changes, e.g. we may need more detailed rules about substitution to get better (or at least equivalent) performance when we divide a match into more time frames, because only three substitutions are permitted in soccer. However this would lead to a virtually impossible problem, of finding an appropriate rule set. We thus took a pragmatic approach, of building a reasonable-sized rule set based on expert knowledge, and then changing its parameters only as we varied the model granularity.

In more detail, we tested a few sets with different time frames, rules, and Bayesian networks. These preliminary results are shown in Table 2 (details of the error measurements are given later in Section 4). Accuracy differences among the three sets ranges from about 10 to 15%. We selected the best based on performance, giving ten frames within each game.

At each frame, the rule-based reasoner and Bayesian networks decide each team's strategies and fatigue for the next frame, from the head coach's perspective. This approach can reflect the dynamics of soccer well – thus FRES can be more realistic than previous works, and using this approach, FRES can produce a long and meaningful reasoning chain for a single match. Furthermore, by

Table 1
Survey for knowledge acquisition

Relationship between formation and stamina/fatigue	(Strong/normal/weak)
Does overlapping cause fatigue?	(Strong/normal/weak)
Relationship between the number of offenders and that of defenders	(Strong/normal/weak)
Does offensiveness derive more goals?	(Strong/normal/weak)
Is Finishing more important than having chance?	(Strong/normal/weak)
Relationship between weather and fatigue	(Strong/normal/weak)
Relationship between weather and match situation	(Strong/normal/weak)
Does better morale derive better play?	(Strong/normal/weak)
Does high concentration induce better play?	(Strong/normal/weak)
Are long pass and heading harmonious when used at the same time?	(Strong/normal/weak)
Does better teamwork cause less fatigue?	(Strong/normal/weak)
Is man-marking safer than zone defense?	(Strong/normal/weak)
Does pressing induce fatigue?	(Strong/normal/weak)
Does closeness between offenders and defenders affect match playing?	(Strong/normal/weak)
Is good stamina more important than good technique?	(Strong/normal/weak)
Does home team have advantage?	(Strong/normal/weak)

Table 2
Error comparisons among different number of time frames, rules, and Bayesian networks

	No. of Frames	Relative rank error	Relative rank error ratio (set 2/ Each)	Standard deviation	RMS error	RMS error ratio (set 2/ each)	Standard deviation
Set 1	5	6.593	0.898	0.429	6.108	0.869	2.740
Set 2	10	5.920	1	0.308	5.306	1	2.359
Set 3	15	6.809	0.869	0.355	6.358	0.835	2.601

predicting the result of a match using this simulation-like time-series approach, FRES can estimate the final scores of the match. Finally, together with the rule-based reasoner and Bayesian networks, our time-series approach can improve the accuracy of the predictions of FRES. We now discuss the overall structure.

3.3. Frame interactions

Fig. 3 shows the sequence of interactions between Jess and JavaBayes in a single frame out of the ten frames of a match. In each frame, the model for each team decides its strategy according to its current state – i.e. it reasons and infers its strategy based on its fatigue, score, morale, etc. For example, assume that team A is ahead by one goal at the eighth frame of a match. The opposite team, team B, can decide/infer its strategy as follows (team B's perspective):

We are trailing by one.
This match is almost over (it is the eighth frame!).
Therefore in order to win this match,
  We should be extremely offensive from now on and do overlapping.
  We should change our formation to offensive one, such as 3-5-2.
  And so on...
However,
Our players are very tired.
Our players' skills are worse than those of team A's players.
Consequently,
  Substitute some players to recover our team's fatigue.
  Change the formation to 3-5-2.
  Do overlapping and long passes so that we can overcome our inferiority in skills.

The strategy determined in this way is applied to the ninth frame of the match. Subsequently, after simulating the effect on the state of the match, the same procedure is applied to the remaining frames.

If two teams are tied after ten frames, they go to extra periods. In FRES, there are three extra frames reserved to cover the extra periods. If they are tied even after the extra periods, the two teams go into a penalty shoot-out in order to finish the match. After the simulation of a match is completed, the result is displayed on the screen with scores, and prediction of the next match in the tournament begins. This process is illustrated in Figs. 4 and 5.

3.4. Rule-based reasoning and Bayesian inference applied to soccer

In the real world, the results of the strategic choices are not completely predictable – soccer is a highly stochastic game, so the results are uncertain. Bayesian networks provide an effective framework in which to model these complex problems by allowing stochastic generation of different outputs for the same set of inputs. We use them for the core part of our inference engine. However the determination of strategy is sufficiently complex that it

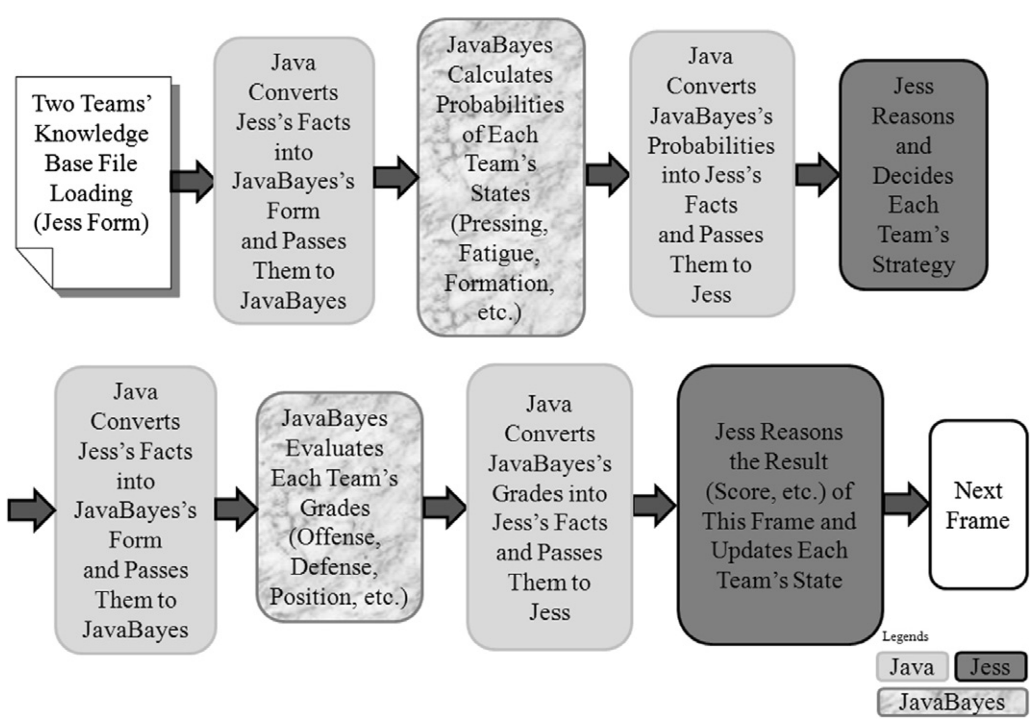

Fig. 3. Interaction between Jess and JavaBayes in a single frame of a soccer match.

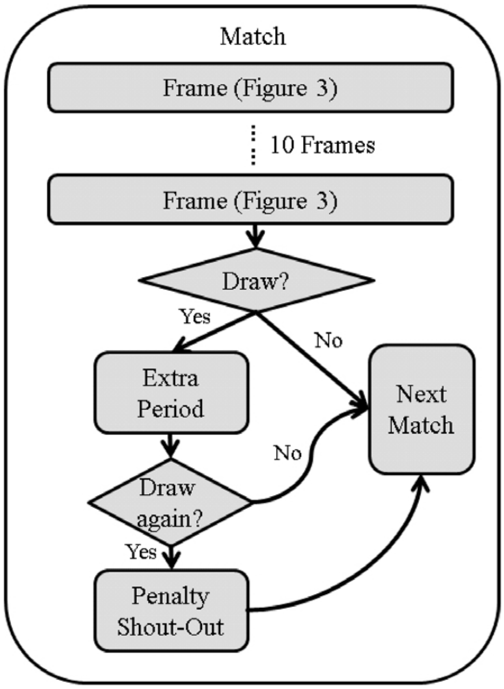

Fig. 4. Procedure of a match.

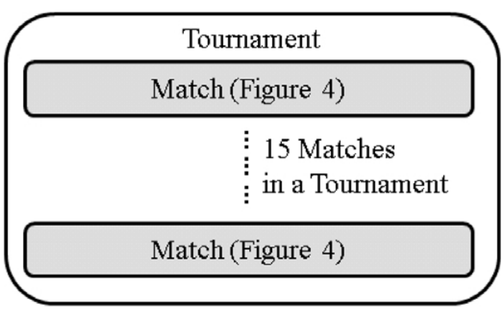

Fig. 5. Procedure of a tournament.

would be difficult, if indeed possible at all, to build a suitable Bayesian network for this component; hence we used Jess as an appropriate tool for this component. In this section, we outline the detailed operations of each module. Jess is responsible for rule-based reasoning and JavaBayes is responsible for uncertain reasoning.

3.4.1. Rule-based reasoning

Our Jess code has two parts: a reasoning part and a non-reasoning part, handling mainly knowledge input and task-oriented sequential work. In this section, we focus on the reasoning part, which consists of rules based mainly on interviews with our soccer expert.

The reasoning can be divided into two stages, strategy-making and result-calculating. Strategies include overlapping, man-marking, pressing, position, and passing. The results from Bayesian networks form the bases for these decisions. Each team is assumed to have its own particular characteristics, such as work rate, aggressiveness, pass length, etc. The total knowledge base is given in Appendix A. Jess takes all these facets and the outputs of JavaBayes into consideration to determine a strategy. As well as play-making strategies, the system also reasons about higher-level decisions such as substitutions and formation changes. The result-calculating part models the actual flow of a match. It models such aspects as scoring, the effect of goals on morale, the effect of reputations, relative scores, and locations on the state of the players. The state changes throughout the match and these changes are encoded as rules in Jess. For example, suppose a team's morale is very good at one moment; if nothing special happens for a long time then their morale can be expected to converge to normal. Scoring rules is a more intuitive example of how to make result-calculating rules; for example, if the offensive grade of a team, say Brazil, is 'A' and the team's strategy is to be very offensive at a frame, while the defensive grade of the opposite team, say Senegal, is 'D' and the number of defenders it has is too small, Brazil scores at this frame.

3.4.2. Bayesian inference

We use four networks to model four important features of soccer matches. They model a team's offense, defense, possession of the ball, and the level of a team's tiredness. Because FRES is designed to act from the head coach's perspective, it tries to choose a way to make attack, defense and possession better while it tries to find ways to limit the increase in fatigue. The results from the four networks are passed to the rule-based reasoner, i.e. Jess, to decide an appropriate combination of strategies. The results passed to Jess are probabilities of the values of overlapping, passing, position,

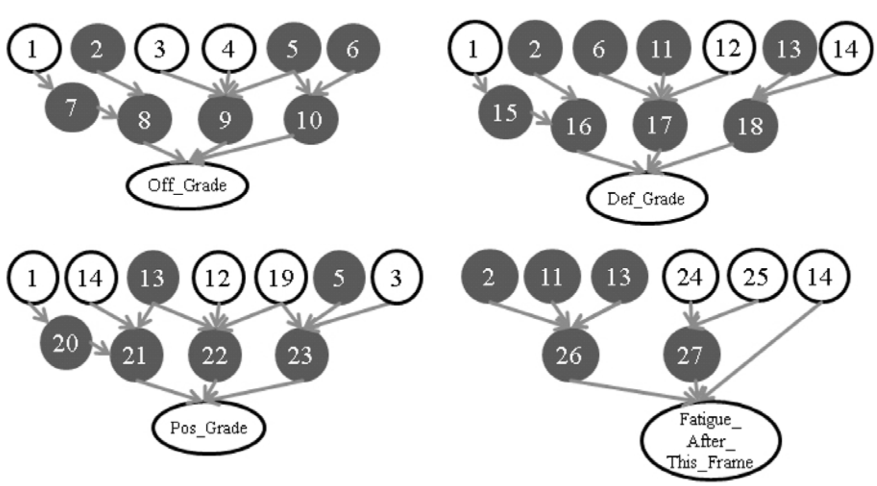

Fig. 6. Four Bayesian networks for deciding a team's strategies: white nodes and dark (black) nodes represent fixed values and probabilistic values which can be queried to get optimized probabilities, respectively.

man-marking, and pressing nodes. The four networks are shown in Fig. 6. The relationships shown in Fig. 6 were built through a knowledge acquisition process detailed in the Section 3.1; the four Bayesian networks are detailed below:

3.4.2.1. Network for offense strategy. In a real soccer team, learning multiple formations is difficult, costing a great deal of training time; in general, it has been found ineffective compared with focusing on a small number of specialist formations. So in FRES, each team may have two formations – the main formation and the sub formation. The team's concentration changes over time. But it is not a part of the strategy, and so it is set to a specific fixed value by FRES. Offensive grade is always set to the aim, 'A.' Formation and overlapping are directly related to the number of offensive players. The number of chances that could lead to a score, and how well each team scores when the chances occur, are treated separately, each having its own node in this network. For example, Korea is quite good at making chances, but traditionally weak at scoring. If a team moves forward a lot and tries long passes, then many chances may be made. On the other hand, the chances are likely to be better for scoring when made from short passes. Skills and concentration also play an important role in finishing.

3.4.2.2. Network for defense strategy. Fatigue is one of a team's states. So formation and fatigue are fixed inputs in FRES. Each team always aims for defensive grade 'A.' Again, formation and overlapping are directly related to the number of defensive players. Defensive stability is affected by strategies such as whether to man-mark or not, whether to emphasize offense or defense, and the teamwork of each team. When a part of the defense system is harassed by the opponent's attack, there should be support from other parts; the covering node embodies this concern. Good covering can be created by effective pressing, but this requires a lot of work.

3.4.2.3. Network for possession. Teamwork and skills are intrinsic abilities of teams which are fixed for most of the time, while morale changes over time. But morale is not part of the strategy, so it is also fixed during the operation of this network. Possession of the ball is mainly affected by mid-fielders, rather than attackers or defenders of both teams; and the team with more players around the ball may have more chance of gaining possession. Intercepting the opponent's possession, and retaining possession of the ball are also important measures. Pressing increases both the numbers of players around the ball, and the possibility of interception. However short passes are harder to intercept than long passes.

3.4.2.4. Network for fatigue. Location plays an important role in the network for fatigue; for example, the time-difference between a team's home and the playing field has a large effect through jet-lag, while the supporting and opposing crowd sizes are also significant factors for fatigue. Stamina is an intrinsic characteristic of a team, and so is handled separately from fatigue. The change in stamina is determined both by how hard a team is working – activity-level – and how well it can endure the work – endurance. Overlapping, man-marking and pressing are all stamina-consuming strategies.

3.4.2.5. Grading offense, defense, possession, and fatigue. After deciding which combination of strategies to use for one frame by rule-based reasoning, the strategies should be put into the networks in Fig. 6 to query and to determine the grades for offense, defense, possession and fatigue level after the given frame. This time, therefore, only the top-most nodes of each network become white (fixed) nodes with the values of the strategies determined in the previous reasoning step, and all the other nodes are set to black (probabilistic). Then the probabilities of the second level are inferred from the top-most nodes, and based on the probabilities, the probability of each grade is inferred.

3.5. Example of FRES's operation

Let us look at how this works in practice. In the first stage, the four Bayesian networks from Fig. 6 are run to determine desirable strategies. Each network has a preferred output state. For example, for the 'offence' network, the desired output is 'aggressive.' This node is set to this value, as are the fixed input nodes – in the case of offence, the 'formation,' 'skills,' and 'concentration' nodes. These values induce a distribution on the other (black) nodes – in the case of offence, on 'overlapping,' 'passing,' and 'position'. These distributions are sampled, to generate discrete values for each attribute. This generates a kind of naïve strategy. However it can not be used as is, because the four networks are almost inevitably in conflict – the strategies which produce aggressive offence and strong defense are also likely to generate fatigue. So when inferring backwards from the desired grades, conflicting attribute values are almost inevitably generated.

This is where the Jess knowledge base comes in. Rules in the knowledge base may change the values of the attributes. In general, the rules act to reduce conflict. For example, one of the rules says (in simplified form)

> *Passing Rule 1:* If the value of 'pressing' for a team is 'yes,' and the team is aggressive, then the value of 'passing' for the team becomes 'long-pass.'

Once the strategy values have been determined by the Jess reasoner, the networks are run again. But this time, the values of both the white input nodes and the black (strategy) nodes are fixed, only the output nodes are inferred. For example, the values in the 'offense' network might be 'overlapping' = 'yes,' 'passing' = 'long-pass,' and 'position' = 'offensive.' The distributions yielded by this process are sampled to yield the outcome of this frame of the game.

Table 3
Legend for Fig. 6 (Full explanation for these terms is given in Appendix B)

1	Formation
2	Overlapping
3	Skills
4	Concentration
5	Passing
6	Position
7	No. of offenders
8	Average number of offenders
9	Finishing
10	Chances
11	Man-marking
12	Teamwork
13	Pressing
14	Fatigue
15	Number of defenders
16	Average number of defenders
17	Defensive stability
18	Covering
19	Morale
20	No. of midfielders
21	Average number of players near the ball
22	Intercepting
23	Keeping
24	Location
25	Stamina
26	Activity level
27	Endurance

Note that this process introduces stochastic behaviour in two places:

- When the Bayesian network is used for reasoning backward from the desired goal, the induced distribution is sampled to determine naïve strategies
- When the Bayesian network is used for reasoning forward from the derived strategy, the induced distribution is sampled to determine the final outcome of the current frame.

4. Evaluation

Ideally, we would compare FRES with some of the previously-published systems. However FRES aims to predict the results of the World Cup, and the World Cup is a tournament. Most previous work deals with the results of a single match or of league matches, so comparison with FRES is inappropriate. In the case of the few studies which predicted tournaments, the available comparison data was based on old data, such as that of the World Cup in 1994 or 1998, which would unfairly hobble FRES, since some of the data it relies on is not available for these earlier tournaments. So, we have not been able to use these as appropriate comparators for FRES. Instead, we evaluate FRES with the results of the World Cup 2002 (the most recent World Cup at the time this research was in progress). We built a historic predictor (HP) and a discounted historic predictor (DHP), which can be used as benchmarks for comparison with FRES.

The historic predictor predicts the result of a match based on the number of goals each team had scored during the entire World Cup history (1930–2002). The data provided to the HPs include the results of the World Cup 2002 (so in principle, the HPs have sufficient data to make perfect predictions – except that they do not have the knowledge that what they are predicting are the results of the World Cup 2002). The reasons we include the results of the World Cup 2002 are twofold: primarily, we need to provide it with some score information for the teams which advanced to the World Cup for the first time in 2002, because without this information, it has no basis to make predictions for these teams. Second, although we aimed not to unfairly incorporate knowledge of the results of the World Cup 2002 in our knowledge base, construction of the knowledge base is a subjective activity, hence it is not possible to guarantee that our system does not incorporate such knowledge. Providing this knowledge directly to the HPs should counterbalance any such effect.

The discounted historic predictor is same as HP except that it incorporates weight decay, i.e. the older information is, the less influence it has. In doing so, we used an exponential decay weight ($w = e^{-kt}$). Because k is a significant parameter, i.e. the results are sensitive to this parameter, we had to find an optimal value. We tested several values ranging from 0.5 to 0.001, and found that 0.01 is the best value for k. A few k values are given in Table 4. This method is quite helpful in improving HP not only because more recent information has more significance but also because it may reflect the trend that the number of goals has been decreasing in soccer in recent years.

Table 4
Performance of discounted historic predictor with variable k

k	Relative rank error	Standard deviation	RMS error	Standard deviation
0.02	8.825	0.580	7.449	4.949
0.01	7.420	0.614	7.306	6.292
0.005	8.302	0.663	7.706	7.434

If we wish to compare three predictors, we need metrics on which to base our comparisons. We introduce a metric, 'point,' to compare the three predictors. It is similar to the FIFA ranking and calculated as follows (champion gets the highest point, 2nd gets second-highest point, etc.):

$$\text{Point} = 5 \times \#\text{ of champion} + 4 \times \#\text{ of 2nd} + 2 \times \#\text{ of semi-final} + \#\text{ of quarter-final}$$

Our evaluation results against this metric are shown in Table 5. Nations are sorted according to the rankings of the World Cup 2002.

As can be seen from the evaluation results, HP works well. But as it primarily considers each team's offensive ability – the goals a team scores – offensive teams tend to be awarded high points, and thus are ranked at relatively high positions. For example, Denmark (137) gets higher points than England (68); Denmark has a higher proportion of scoring more than two goals in a match than that of England. In fact, Denmark is ranked at fifth position in the results of HP. Considering that Denmark was actually ranked at tenth position in the World Cup 2002, it is an anomalously high position. Moreover Turkey is impregnably at the top, its point score is 180, while that of the actual winner, Brazil, is somewhat behind at 170. This also is due to Turkey's relatively highly multi-scoring matches. These tendencies can be seen in the results of DHP. Although it shows better predictions than the historic one, Denmark and Turkey are still ranked at very high positions – 6th and 4th, respectively. Of course, as we mentioned earlier, it shows better predictions, e.g. Brazil is the top-ranked team and Germany is ranked at the second position.

By contrast, FRES shows more accurate and stable results, because it considers many aspects of soccer and tournaments. Especially, FRES considers the relationships between the two teams in a match. For example, England is a great team, but when it advances to the quarter-final, it will encounter Brazil with high probability. So the points FRES allocates to England are somewhat low (98). The points for Turkey are relatively high, on the same basis, because Turkey has relatively good luck with the draw of matches – it had a round-of-16 match with Japan, and quarter-finals with either Senegal or Sweden.

FRES predicts six countries out of the actual top eight countries of the World Cup 2002, while HP and DHP predict five – the three predictors are relatively close in performance in this sense. Table 6 shows the differences between each predictor's predictions and the real rankings. FRES shows more stable and realistic predictions than HP, while DHP shows similar performance to FRES in terms of the number of errors. But when we consider all errors (Table 6 only deals with eight teams), FRES outperforms DHP. We can see these relationships directly in Table 7. The relative rank error and the RMS (Root Mean Squared) error are calculated as described below; FRES gives substantially lower errors than the other two predictors. This indicates that FRES's predictions are more stable and accurate than those of HP and DHP.

$$\text{Relative rank error} = \sum_{\text{teams}} \frac{|\text{rank}_{\text{actual}} - \text{rank}_{\text{predicted}}|}{\text{rank}_{\text{actual}}}$$

$$\text{RMS error} = \sqrt{\sum_{\text{teams}} \frac{(\text{rank}_{\text{actual}} - \text{rank}_{\text{predicted}})^2}{\text{rank}_{\text{actual}}}}$$

To confirm these results, we performed T-tests on the hypothesis that FRES is significantly better than the other two predictors. Table 7 shows relative rank errors and RMS errors based on the rankings from 100 runs of each predictor. All six results (relative rank errors and RMS errors for each predictor) passed the Jarque-Bera normality test. The T-test results are given in Table 8. The T-test we performed is heteroscedastic, unpaired, and

Table 5
Evaluation results of 100 tournaments: Nations are sorted according to the rankings of the World Cup 2002

Nation	Historic predictor							Discounted HP (k =0.01)							FRES							Rankings of the World Cup 2002
	Ch	2nd	SF	QF	16	Pts	Rank	Ch	2nd	SF	QF	16	Pts	Rank	Ch	2nd	SF	QF	16	Pts	Rank	
Brazil	12	9	23	28	28	170	2	15	10	15	27	33	172	1	17	3	10	23	47	140	1	1
Germany	16	10	11	24	39	166	3	14	10	14	21	41	159	2	13	10	5	22	50	137	2	2
Turkey	14	14	12	30	30	180	1	10	7	18	28	37	142	4	5	12	13	19	51	118	5	3
Korea	5	4	7	14	70	69	13	2	4	9	20	65	64	14	10	6	3	26	55	106	7	4
Spain	6	6	14	28	46	110	7	8	9	16	31	36	139	5	8	5	22	30	35	134	3	5
England	1	4	11	25	59	68	14	4	5	8	25	58	81	11	4	3	18	30	45	98	11	6
Senegal	1	8	11	26	54	85	10	3	5	11	22	59	79	12	4	10	10	33	43	113	6	7
USA	3	10	19	38	30	131	6	4	6	12	32	46	100	8	3	3	15	25	54	82	15	8
Japan	3	7	7	22	61	79	12	3	3	9	22	63	67	13	3	6	19	23	49	100	9	9
Denmark	11	8	10	30	41	137	5	10	7	13	28	42	132	6	4	5	12	24	55	88	13	10
Mexico	5	2	15	19	59	82	11	4	5	10	28	53	88	10	4	5	18	27	46	103	8	11
Ireland	0	3	9	18	70	48	15	1	3	9	23	64	58	16	1	6	7	21	65	64	16	12
Sweden	13	8	11	27	41	146	4	7	7	18	28	40	127	7	3	9	6	25	57	88	14	13
Belgium	0	2	9	17	72	43	16	2	3	8	20	67	58	15	3	9	12	23	53	98	10	14
Italy	5	2	17	28	48	95	8	8	9	19	26	38	140	3	13	4	15	23	45	134	4	15
Paraguay	5	3	14	26	52	91	9	5	7	11	19	58	94	9	5	4	15	26	50	97	12	16

Table 6
Comparison of rankings between FRES and the other two historic predictors

	Real rank	Historic predictor		Discounted HP		FRES	
		Rank	Diff.	Rank	Diff.	Rank	Diff.
Brazil	1	2	1	1	0	1	0
Germany	2	3	1	2	0	2	0
Turkey	3	1	2	4	1	5	2
Korea	4	13	9	14	10	7	3
Spain	5	7	2	5	0	3	2
England	6	14	8	11	5	11	5
Senegal	7	10	3	12	5	6	1
USA	8	6	2	8	0	15	7
Total Diff.		28		21		20	

Table 7
Relative rank error and RMS error of FRES and the other two historic predictors

	Relative rank error	Standard deviation	RMS error	Standard deviation
FRES	5.920	0.308	5.306	2.359
Historic predictor	9.651	0.545	7.310	5.277
Discounted HP	7.420	0.614	7.306	6.292
Ratio (FRES/HP)	0.613	0.565	0.726	0.447
Ratio (FRES/DHP)	0.798	0.502	0.726	0.375

one-tailed. The heteroscedastic setting is obvious, because the standard deviations are not equal. Unpaired T-test means that there is no dependence between two sample sets. And we used a one-tailed T-test, as our hypothesis was that FRES is better than HP and DHP – not just different from them. As can be seen from Table 8, we found very small p-values for relative rank error (we show it as <0.0001, because it is enough to show significance, but the actual p-values were 3.341×10^{-111} between FRES and HP, 7.072×10^{-48} between FRES and DHP). We also obtained very

Table 9
Relationships among some teams shown in FRES's predictions (wins out of 100 matches): there are too many relationships to show the whole table, so we include only part

	No. of wins (historic predictor)	No. of wins (discounted HP)	No. of wins (FRES)
Denmark (vs. England)	53	62	42
Denmark (vs. Belgium)	58	64	69
England (vs. Belgium)	56	52	34
Denmark (vs. Brazil)	42	36	48
Denmark (vs. Senegal)	63	69	51
Brazil (vs. Senegal)	64	78	87
England (vs. Brazil)	35	26	47
England (vs. Italy)	50	46	36
Brazil (vs. Italy)	63	71	73
England (vs. Brazil)	35	26	47
England (vs. Sweden)	42	41	49
Brazil (vs. Sweden)	59	67	65
Sweden (vs. Mexico)	60	68	72
Germany (vs. Mexico)	60	74	51

small p-values for RMS error (0.0000539 between FRES and HP, 0.000189 between FRES and DHP). We can conclude that FRES is significantly better than the other two predictors.

FRES has other advantages relative to the historic predictors (and by extension, to the previous score-based systems that resemble them). FRES can be readily tuned with up-to-date information, so that it is more flexible and realistic, whereas the other two predictors take a team's offensive ability as a static value. Notably, HP and DHP can become unrealistic because information about a team, such as the squad, offensive and defensive ability, etc., are time variant. Although DHP uses weight decay, it still only considers goal information. Moreover, as FRES considers each team's characteristics, it can show ranking relationships between the teams. For example, as shown in Table 9, though Mexico is ranked at a higher position than Sweden in Table 5 (FRES's rankings), Sweden wins more when they have matches. By contrast, the

Table 8
T-test results between FRES and the other two historic predictors

	Relative rank error					RMS error			
	HP	FRES	DHP	FRES		HP	FRES	DHP	FRES
Mean	9.677	5.957	7.443	5.957	Mean	7.841	5.601	7.842	5.601
Standard deviation	0.525	0.294	0.609	0.294	Standard deviation	5.092	2.381	5.661	2.381
p-value		<0.0001		<0.0001	p-value		<0.0001		0.000189

HPs' predictions always follow the rankings, i.e. the stronger team always wins more. This phenomenon is clearer in the results of DHP. Furthermore, we can see that DHP always considers that aggressive teams – e.g., Brazil – are much stronger than defensive ones – e.g., England. This tendency, of course, stems from the fact that it only considers the offensive power of a team. Systems which also look at goal differences may perform better in this respect. Other interesting examples are listed in Table 9. For example, Denmark wins more when it has matches with Belgium; and Belgium wins more when it has matches with England. Yet When Denmark and England play each other, England usually wins.

5. Conclusions and future work

5.1. Applicable domains

FRES is simply an instance of a more general framework, applied to soccer. What other sports could it cover? In this section, we consider the possibilities for extension.

Even though the framework can in principle be adapted to a wide range of sports domains, it cannot be used in domains which have insufficient expert knowledge. Of course, most knowledge-based systems suffer from this limitation. Knowledge-based systems usually require knowledge of relatively good quality while most machine learning systems need a huge amount of data to get good predictions.

Some sports are just too simple to justify our complex framework. In most track and field sports, a large part of the framework becomes unnecessary. Most field and track sports are not so stochastic, and the strategy and the physical ability of a player (or a team) are often the only dominant determinant of results (usually records in these cases). Thus the Bayesian network components would not be needed in these sports. Suppose an athlete doing pole vault. He has three trials so he needs a strategy. And the strategy may be changed after each trial. There can be stochastic processes but they are small enough to ignore. So our framework may be inefficient in such sports. However, these sports are usually not so complicated that a result predictor is needed at all; our main interest is more complex sports, in which a predictor can be useful and helpful.

5.2. Summary

We have taken a compound framework in predicting sports results – rule-based reasoning and Bayesian inference – and combined it with an in-game time-series approach for more accurate and realistic predictions. This framework is implemented in Java with Jess and JavaBayes. Essentially, the rule-based reasoner determines teams' strategies and the direct results of those strategies – change in score, substitution, etc.; in effect, it simulates the role of the head coach. The Bayesian networks probabilistically sample the stages of the game progression. Evaluation shows that our system, FRES, gives reasonable predictions without machine learning; it uses only a relatively small amount of initial knowledge – knowledge which can be readily obtained. Through evaluation, we show that FRES gives reasonable and stable predictions, and that it can reflect each team's characteristics, and predict the relationships among teams.

FRES takes a novel approach, so there are many opportunities to investigate alternatives and possible improvements. First of all, we plan to build a more formalized knowledge acquisition process. We plan to use machine learning techniques to build the initial team knowledge. We plan to use parameter learning methods to tune FRES automatically. Tournament-specific issues, such as the effect of differences in rest periods, and the effect of the results of previous matches on a team's morale, will be incorporated in future versions of the system.

Our in-game time-series approach also has potential for improvement. Currently, FRES uses a static and uniform 10 time frames. Ultimately, however, we need to develop dynamic methods which can reflect tides/flows of a match more directly.

Finally, the framework can be readily applied to other sports, such as baseball, basketball, etc., so long as appropriate rules and Bayesian networks can be built. That is, our approach can effectively be applied to many kinds of sports. We aim to build a number of other sports result expert systems by applying the approach in the future.

Appendix A. Initial team knowledge

See Table A.1

Appendix B. Detailed explanations for Fig. 6 and Table 3

See Tables B.1–B.4

Appendix C. Overview of the implementation

We use Jess [16] and JavaBayes [14] together to represent knowledge explicitly. Jess is a rule engine and scripting environment written in Java. It is based on CLIPS [6] – Jess originated as a Java version of CLIPS. Like CLIPS, Jess is a fairly traditional expert systems shell. In FRES, Jess is responsible for rule-based reasoning. JavaBayes is a system that handles Bayesian networks. It calculates

Table A.1
Team knowledge used in FRES: these values can be tuned easily

	Location	Reputation	Skills	Teamwork	Squad-depth	Stamina	Main formation	Sub formation	Hardworking	Aggressiveness	Passlength
Korea	3	45	40	85	4	90	3-4-3	5-4-1	80	30	50
Italy	7	90	90	85	2	90	4-4-2	0	70	30	75
Spain	9	75	80	65	3	75	4-4-2	4-3-3	50	50	50
Ireland	9	50	60	80	2	80	4-4-2	4-3-3	70	50	50
USA	1	50	55	75	2	80	4-4-2	0	75	45	50
Mexico	1	70	60	65	3	75	4-5-1	4-4-2	55	55	40
Paraguay	2	65	35	90	1	80	3-5-2	0	70	25	75
Germany	6	90	85	90	2	95	4-4-2	0	75	60	40
Japan	3	35	35	75	3	75	4-4-2	0	75	30	50
Turkey	5	40	80	80	2	80	4-4-2	0	85	40	50
Sweden	8	65	70	70	2	75	4-4-2	0	50	40	70
Senegal	6	35	85	65	1	85	4-4-2	0	80	60	30
Brazil	2	100	100	75	5	100	4-4-2	5-3-2	75	75	35
Belgium	6	60	55	80	2	70	4-4-2	0	60	45	40
Denmark	8	55	70	75	2	70	4-4-2	0	50	40	80
England	8	85	85	75	4	75	4-4-2	0	75	55	55

Table B.1
Detailed explanations for the offensive grade network

Formation	Formation for a team 5-4-1 means 5 defenders, 4 mid-fielders and 1 attacker
Overlapping	Soccer has a position named 'wingback', or in simpler terms, 'side defender'. Modern soccer strategy uses 'overlapping' which encourages these 'wingbacks' to go forward to attack when they have possession of the ball. This can put more power into the attack, but it can also cause fatigue, and it may lead to weaker defense
Skills	How good are skills of players of a team
Concentration	How much of their attention players give to the match
Passing	Which kind of passes players prefer (long or short passes)
Position	Overall position of whole members
Number of offenders	Basic number of attackers
Average number of offenders	Expected average number of defenders in the right position during the match (overlapping will be likely to increase this number)
Finishing	How well a team scores when chances occur
Chances	How many chances a team can make
Off_grade	Overall grade for offense

Table B.2
Detailed explanations for the defensive grade network

Formation	Formation for a team 5-4-1 means 5 defenders, 4 mid-fielders and 1 attacker
Overlapping	Soccer has a position named 'wingback', or in simpler terms, 'side defender.' Modern soccer strategy uses 'overlapping' which encourages these 'wingbacks' to go forward to attack when they have possession of the ball. This can put more power into the attack, but it can also cause fatigue, and it may lead to weaker defense
Position	Overall position of whole members
Man-marking	Defenders can move either man-based or region-based. Under man-based, defenders tend to follow their opponent's attackers man to man, while under region-based, defenders tend to keep their relative positions to each other
Teamwork	How good is the teamwork of a team
Pressing	A couple of players near the opponent player who has the ball will bear down on him together. It is a stable strategy but players need to move very much more (and so get more fatigued)
Fatigue	How tired is a team as a whole
Number of defenders	Basic number of defenders
Avgerage number of defenders	Expected average number of defenders in right position during the match. Overlapping will be likely to make this number lower
Defensive stability	How stable is the defense of a team
Covering	How well a team recovers from a break in the defense system
Def_Grade	Overall grade for defense

marginal probabilities and expectations, produces explanations, performs robustness analysis, and allows the user to import, create, modify and export Bayesian networks. JavaBayes is, of course, responsible for the Bayesian inference in FRES. But building JavaBayes as a module in a Java program was not a trivial task. It was necessary to rebuild JavaBayes for use as a module. Jess and JavaBayes being both written in Java, Java is a natural choice for the underlying framework of FRES. That is, these modules are combined using Java. And all the outcomes of Jess and JavaBayes are recorded in a single output file. This greatly simplifies exposition of the system, since Jess and JavaBayes are relatively familiar and well-documented tools; the behaviour of the system is largely

Table B.3
Detailed explanations for the possession grade network

Formation	Formation for a team. 5-4-1 means 5 defenders, 4 mid-fielders and 1 attacker
Fatigue	How tired is a team as a whole
Pressing	A couple of players near the opponent player who has the ball will bear down on him together. It is a stable strategy but players need to move very much more (and so get more fatigued)
Teamwork	How good is the teamwork of a team
Morale	How good are players' feelings now
Passing	Which kind of passes players prefer (long or short passes)
Skills	How good are skills of players of a team
Number of midfielders	Basic number of midfielders
Average number of Players near the Ball	Expected average number of players near the ball during the match
Intercepting	How well a team can intercept opponent's ball
Keeping	How well a team can keep its ball
Pos_grade	Overall grade for possession

Table B.4
Detailed explanations for the fatigue network

Overlapping	Soccer has a position named 'wingback', or in simpler terms, 'side defender'. Modern soccer strategy uses 'overlapping' which encourages these 'wingbacks' to go forward to attack when they have possession of the ball. This can put more power into the attack, but it can also cause fatigue, and it may lead to weaker defense
Man-marking	Defenders can move either man-based or region-based. Under man-based, defenders tend to follow their opponent's attackers man to man, while under region-based, defenders tend to keep their relative positions to each other
Pressing	A couple of players near the opponent player who has the ball will bear down on him together. It is a stable strategy but players need to move very much more (and so get more fatigued)
Location	Stands for relative location. Is this team Home team or Away team for this match? Nearer or farther from the stadium where the match is held?
Stamina	Natural or trained stamina of a team. The more stamina a team has, the less likely it is to gain fatigue
Fatigue	How tired is a team as a whole
Activity level	How much players of a team moved
Endurance	How much players of a team can endure fatigue gain
Fatigue_after_this_term	Expected fatigue level after given frame

understandable simply from an understanding of the Jess and JavaBayes outputs. FRES provides a GUI for user convenience, and is totally portable, i.e. it can run on any platform that supports JRE (Java Runtime Environment), such as Windows, Linux, Mac OS, Symbian, etc.

FRES is available at http://meslab.snu.ac.kr/bhmin/SportsPredictor/sportspredictor.htm.

References

[1] M. Amor, W. Griffiths, Modelling the Behaviour and Performance of Australian Football Tipsters, Working Paper, Department of Economics, University of Melbourne, 2003.

[2] P. Andersson, M. Ekman, J. Edman, Forecasting the Fast and Frugal Way: A Study of Performance and Information-Processing Strategies of Experts and Non-Experts when Predicting the World Cup 2002 in Soccer, SSE/EFI Working Paper Series in Business Administration No. 2003:9, Stockholm School of Economics, Stockholm, 2003.

[3] Ralph B. Bisland Jr, A stochastic CAI model for assisting in the design of football strategy, ACM SIGSIM Simulation Digest 10 (1) (1978) 28–35.
[4] A.G. Büchner, W. Dubitzky, A. Schuster, P. Lopes, P.G. O'Donoghue, J.G. Hughes, D.A. Bell, K. Adamson, J.A. White, J. Anderson, M.D. Mulvenna, Corporate evidential decision making in performance prediction domains, in: Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI'97), Brown University, Providence, Rhode Island, USA, 1997.
[5] F. Carmichael, D. Thomas, Home-field effect and team performance: evidence from English premiership football, Journal of Sports Economics 6 (3) (2005) 264–281.
[6] CLIPS, <http://www.ghg.net/clips/CLIPS.html>.
[7] M. Crowder, M. Dixon, A. Ledford, M. Robinson, Dynamic modelling and prediction of English football league matches for betting, The Statistician 51 (2002) 157–168.
[8] J.M. Falter, C. Perignon, Demand for football and intra-match winning probability: an essay on the glorious uncertainty of sports, Applied Economics 32 (2000) 1757–1765.
[9] D. Forrest, R. Simmons, Outcome uncertainty and attendance demand in sport: the case of English soccer, The Statistician 51 (2) (2002) 241–291.
[10] J. Goddard, Modelling Football Match Results and the Efficiency of Fixed-Odds Betting, Working Paper, Department of Economics, Swansea University, 2003.
[11] J. Goddard, I. Asimakopoulos, Forecasting football results and the efficiency of fixed-odds betting, Journal of Forecasting 23 (2004) 51–66.
[12] F. Halicioglu, Can we predict the outcome of the international football tournaments: the case of EURO 2000?, Dogus University Journal 6 (1) (2005) 59–69
[13] F. Halicioglu, Forecasting the professional team sporting events: evidence from Euro 2000 and 2004 football tournaments, in: Fifth International Conference on Sports and Culture: Economic, Management and Marketing Aspects, 2005.
[14] JavaBayes, <http://www.cs.cmu.edu/javabayes>.
[15] N. Hirotsu, B. Mike, Wright, modeling tactical changes of formation in association football as a zero-sum game, Journal of Quantitative Analysis in Sports 2 (2) (2006). Article 4.
[16] Jess, <http://www.jessrules.com/jess/index.shtml>.
[17] A. Joseph, N.E. Fenton, M. Neil, Predicting football results using Bayesian nets and other machine learning techniques, Knowledge-Based Systems 19 (7) (2006) 544–553.
[18] D. Karlis, I. Ntzoufras, On Modelling Association Football Data, Technical Report, Department of Statistics, AUEB, 1999.
[19] Ruud H. Koning, Balance in competition in Dutch soccer, The Statistician 49 (3) (2000) 419–431.
[20] R.H. Koning, M. Koolhaas, G. Renes, G. Ridder, A simulation model for football championships, European Journal of Operational Research 148(2).
[21] Kevin Korb, Ann E. Nicholson, Bayesian Artificial Intelligence, CRC Press, Boca Raton, FL, 2003.
[22] D. Kuonen, Statistical Models for Knock-Out Soccer Tournaments, Technical Report, Department of Mathematics, Ècole Polytechnique Federale de Lausanne, 1996.
[23] M.K.K. Lau, Artificial Intelligence in Sporting Results Prediction Problems, B.Eng. Thesis, Department of Electrical and Computer Engineering, Curtin University of Technology, Perth, WA, Australia, 2004.
[24] Joseph Martinich, College football rankings: do the computers know best?, Interfaces 32 (5) (2002) 85–94
[25] A. McCabe, An artificially intelligent sports tipper, in: Bob McKay, John K. Slaney (Eds.), Proceedings of the 15th Australian Joint Conference on Artificial Intelligence, Lecture Notes in Computer Science, 2557, Springer, Berlin, 2002, p. P 718.
[26] Tom M. Mitchell, Machine Learning, McGraw-Hill, NY, 1997.
[27] F. Palomino, L. Rigotti, A. Rustichini, Skill, Strategy, and Passion: An Empirical Analysis of Soccer, Center Discussion Paper 98129, Tilburg University, 1999.
[28] A.P. Rotshtein, M. Posner, A.B. Rakityanskaya, Football predictions based on a fuzzy model with genetic and neural tuning, Cybernetics and Systems Analysis 41 (4) (2005) 619–630.
[29] H. Rue, O. Salvesen, Prediction and retrospective analysis of soccer matches in a league, The Statistician 49 (3) (2000) 399–418;
H. Rue, O. Salvesen, Prediction and retrospective analysis of soccer matches in a league, The Statistician 49 (3) (2003) 268–276.
[30] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Second ed., Prentice-Hall, Upper Saddle River, NJ, 2002.
[31] A. Tsakonas, G. Dounias, S. Shtovba, V. Vivdyuk, Soft computing-based result prediction of football games, in: The First International Conference on Inductive Modelling (ICIM'2002), Lviv, Ukraine, 2002.
[32] Tae Young Yang, Tim Swartz, A two-stage Bayesian model for predicting winners in major league baseball, Journal of Data Science 2 (1) (2004) 61–73.