On Approximate Concept Learning

Erhard Konrad
Technical University of Berlin

Ewa Orłowska
Polish Academy of Sciences Warsaw

Zdzisław Pawlak
Polish Academy of Sciences Warsaw

TECHNISCHE UNIVERSITÄT BERLIN

FACHBEREICH 20 — INFORMATIK

1000 BERLIN 12, Straße des 17. Juni 135

On Approximate Concept Learning

Erhard Konrad
Technical University of Berlin

Ewa Orłowska
Polish Academy of Sciences Warsaw

Zdzisław Pawlak
Polish Academy of Sciences Warsaw

Bericht Nr. 81-7

Oktober 1981

Vertrieb:

Universitätsbibliothek der TU Berlin

Abteilung Publikationen

Straße des 17. Juni 135

D-1ooo Berlin 12

Abstract

In this paper the idea of approximate concept learning is introduced. It covers situations where a complete description of knowledge is not available. The problems are discussed within a rigorous formal framework. An application from computer chess is analysed which is typical for learning to play positional games.

1 Introduction

In this paper we deal with concept learning depending on an approximate description of knowledge. In the area of artificial intelligence related work was done by Banerji (/1/), Michalski (/5/), Morgan (/6/), Plotkin (/9/), /1o/), and Popplestone (/11/). The concept of an approximate description was introduced by Pawlak (/8/); an extension including inductive generalization is given in /4/.

A variety of models and languages is being applied to problems of knowledge representation (/2/). Our approach is based on a mathematical model which may be informally described as follows (/7/): The basic component of a knowledge representation system is a finite, non-empty set of objects, the universe of discourse. The knowledge concerning the objects will be expressed by characteristic features. This is modelled by a function assigning values of attributes to objects. A formalized language L is introduced such that the expressions of this language are interpreted as sets of objects. We will say that a subset Y of the universe of discourse is definable in L, if there is an expression E in L such that Y is the meaning of E. However, if Y is not a definable set, we only have an approximate description of Y in L.

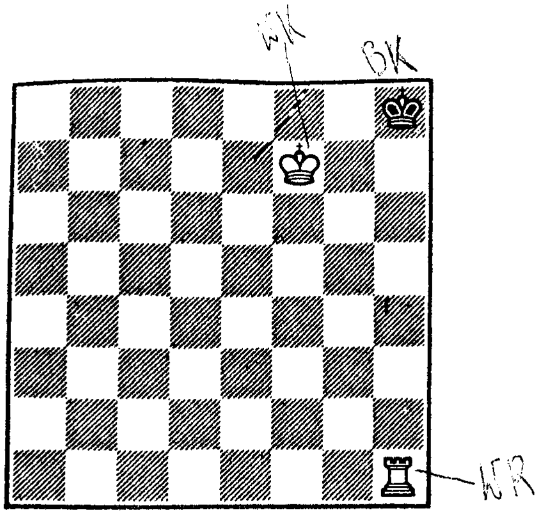
In this work concept learning is being understood as inductive generalization from examples. We consider the problem what features are inessential in order to describe a set of objects. The application we have chosen is concerned with learning to play positional games. It is shown that attributes may be omitted without losing the adequacy of the description. If there is no characteristic description of the extension of a concept, we speak of approximate concept learning.

In the following sections we will discuss these problems within a rigorous formal framework.

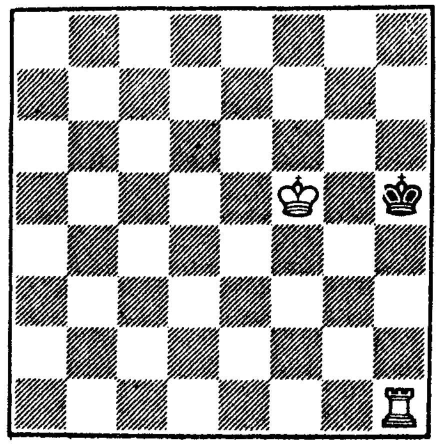
2 Concept Learning

In this section we want to explain an important type of learning. Extending the work of Banerji (/1/) it covers situations where a complete description of objects is not available. As an application we consider learning to play positional games. In order to illustrate the main idea we analyse a sample of 12 chess positions being described by a set of attributes (figure 1a,b The task for a human being or a computer is to learn the concept of checkmate.

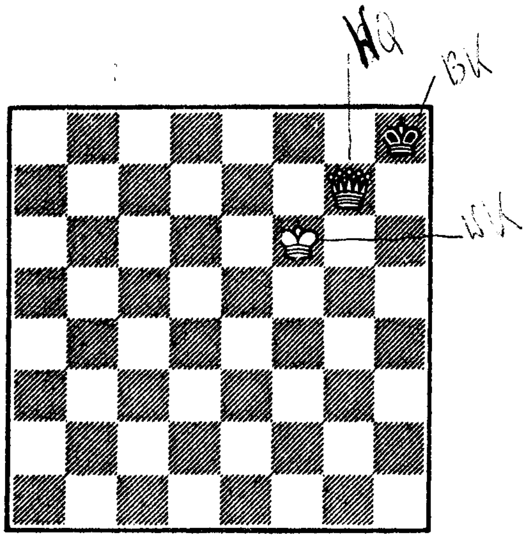
According to the laws of chess of the World Chess Federation (FIDE) checkmate is defined as follows (/3/):
"1o.1 The king is in check when the square it occupies is attacked by an enemy piece; in this case the latter is said to be 'checking the king'.
1o.2 Check must be parried by the move immediately following. If the check cannot be parried, it is said to be 'mate'.
Note to Art. 1o.2:
Check may be parried:
a) by moving the king to a square which is not threatened by an enemy piece,
b) by capturing the opponent's piece which is checking the king and



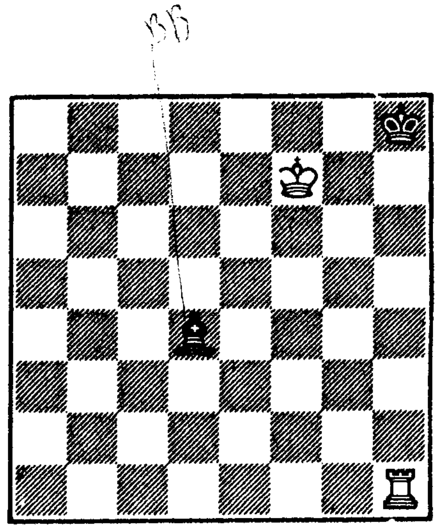
P1: CHECKMATE



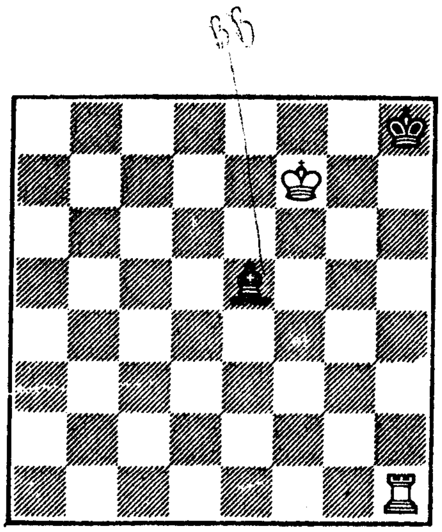
P2: CHECKMATE



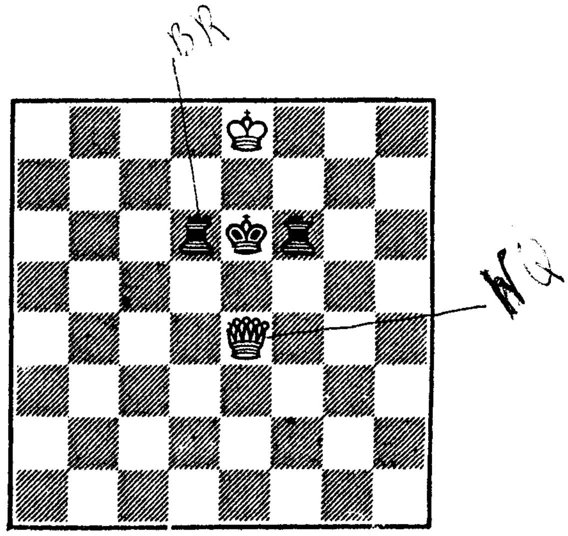
P3: CHECKMATE



P4: CHECKMATE

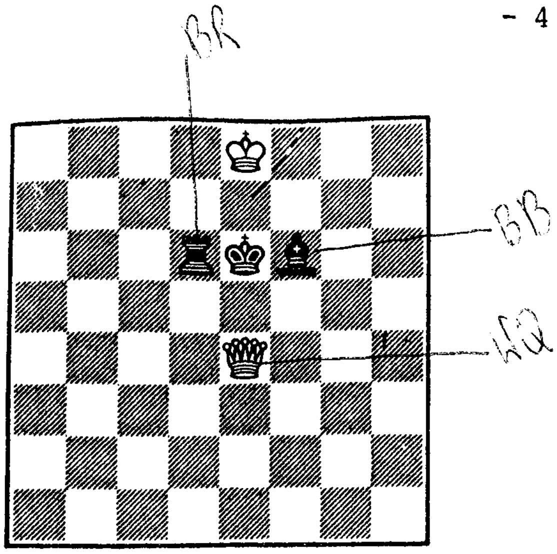


P5: NOT CHECKMATE

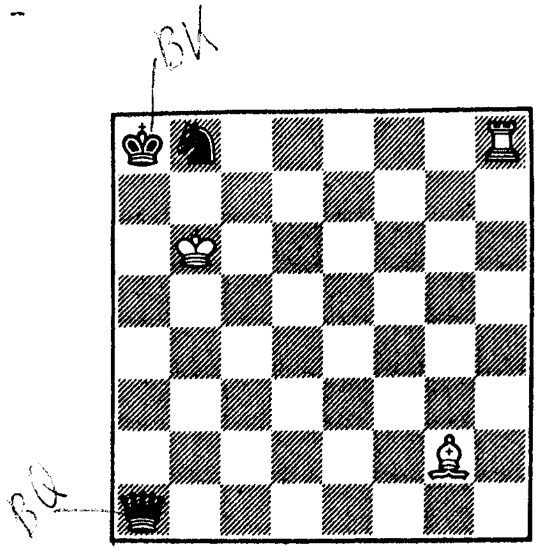


P6: CHECKMATE

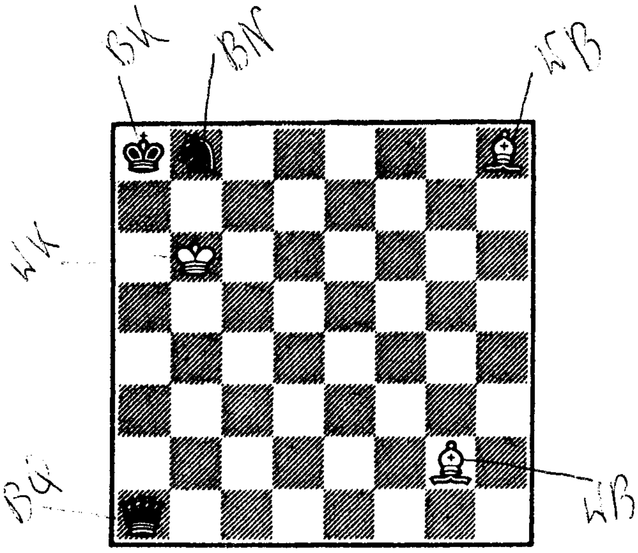
Fig. 1a: Sample of chess positions



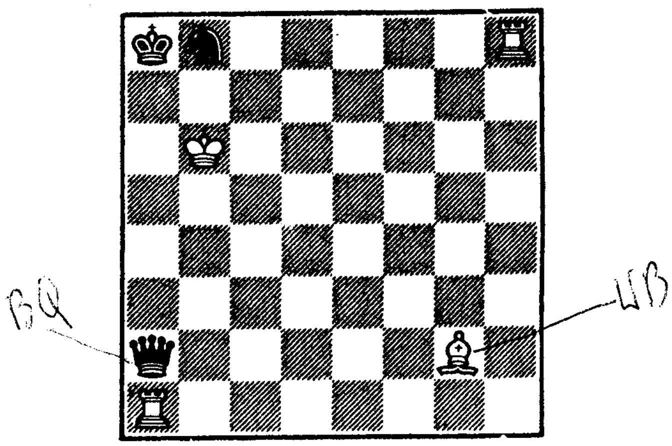
P7: NOT CHECKMATE



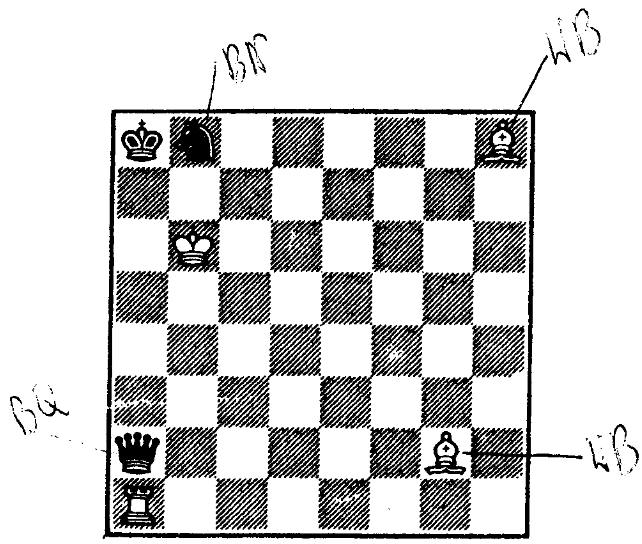
P8: CHECKMATE



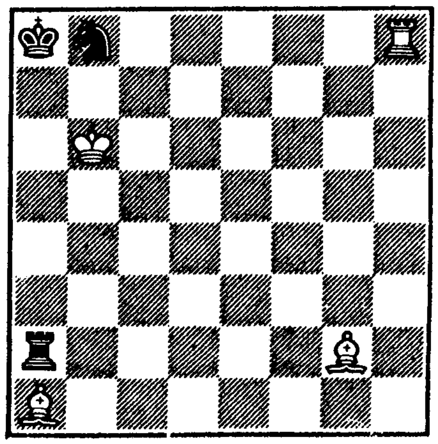
P9: NOT CHECKMATE



P1o: CHECKMATE



P11: NOT CHECKMATE



P12: NOT CHECKMATE

<u>Fig. 1b:</u> Sample of chess positions

c) by placing one of one's own pieces on one of the squares lying between the king and the attacking enemy piece. This last means of defense is evidently not possible when the check comes from the knight or in the case of a double check.

11.1 The game is won by the player who has mated his opponent's king."

In our investigation we distinguish 3 cases:

1) All essential and only essential attributes are known.

2) All essential and one inessential attribute is known.

3) One essential attribute is unknown.

The following attributes are considered:

a1) Safety of the checking piece (saf) with the values "capture" (c) and "no capture" (c').

a2) Mobility of the attacked king (mob) with the values O, 1, ..., 8.

a3) Number of pieces on the board (num) with the values O, 1, ..., 32.

a4) Possibility of interposition (int) with the values "possible" (p) and "impossible" (p').

Now let us analyse the sample of positions in figure 1a,b and explain the meaning of approximate concept learning. In every position P1 till P12 the black king is in check. The number of pieces on the board varies from 3 to 7. Chess players know that this number is not an essential feature of checkmate, but a checkmate position needs at least 3 pieces. Let us imagine we know how to move the pieces, but we are not familiar with the concept of checkmate. After having studied the 12 positions we will recognize by inductive generalization that the attribute "num" is redundant.

Let us consider case 3. If the attribute "int" e.g. is not included in our language, then we will not be able to distinguish between position P4 (checkmate) and position P5 (not checkmate); the same applies to P6 and P7, P8 and P9, P1o and P11. Therefore it is impossible to learn the exact concept of checkmate, but we will show that approximations are definable. In the following sections we will analyse this situation in more details.

3 Formal Framework

In this section we summarize concepts and facts from /4/.

A knowledge representation system is a quadruple

$$S = (X, A, V, \varrho) ,$$

where X is a non-empty, finite set of objects, the universe of discourse; A is a non-empty finite set of attributes; and ϱ is a total function from the set $X \times A$ into the set V. $X \times A$ is the cartesian product of X and A. Two objects x_1 and x_2 of the universe X are called equivalent, if they are not distinguishable with respect to all attributes of S. The equivalence classes are called elementary sets. If all elementary sets of the system S are units, then S is a selective system.

A knowledge representation language L_S consists of

- attribute symbols: $a_1, a_2, \ldots, a_n;\ a_i \in A;$
- value symbols: $v_1, v_2, \ldots, v_m;\ v_i \in A;$
- operation symbols: $-, +, \cdot, \rightarrow, \leftrightarrow;$
- brackets (,) .

Attribute-value pairs (a,v) are called descriptors being the atomic expressions of the language L_S. The set of all terms of L_S is the least set containing the atomic expressions and being closed under the operations $-, +, \cdot, \rightarrow, \leftrightarrow$.

The interpretation of terms is defined inductively by the function val_S mapping the set of all terms into the family of all subsets of the universe X:

$$val_S(a,v) = \{x \in X : \rho(x,a) = v\},$$
$$val_S(-t) = \sim val_S\ t,$$
$$val_S(t_1+t_2) = val_S\ t_1 \cup val_S\ t_2,$$
$$val_S(t_1 \cdot t_2) = val_S\ t_1 \cap val_S\ t_2,$$
$$val_S(t_1 \rightarrow t_2) = \sim val_S\ t_1 \cup val_S\ t_2,$$
$$val_S(t_1 \leftrightarrow t_2) = val_S(t_1 \rightarrow t_s) \cap val_S(t_2 \rightarrow t_1).$$

The symbols $\sim, \cup, \cap$ are denoting the set-theoretic operations of complement, union, and intersection.

A term t is called valid in the system S, if $val_S\ t = X$. We will write $\models_S t$. Then we have the following facts:

Fact 1: If $\models_S t_1 \rightarrow t_2$, then $val_S\ t_1 \subseteq val_S\ t_2$

Fact 2: If $\models_S t_1 \leftrightarrow t_2$, then $val_S\ t_1 = val_S\ t_2$

A term t is said to be elementary, if t is of the form $(a_1,v_1) \bullet (a_2,v_2) \bullet \ldots \bullet (a_n,v_n)$, where $a_1, a_2, \ldots, a_n$ are all attributes of A and $v_1, v_2, \ldots, v_n$ are values of V respectively. We note the following fact:

Fact 3: If t is an elementary term, then $val_S\ t$ is an elementary set or the empty set.

Let $t_1, t_2, \ldots, t_m$ be all elementary terms with $val_S\ t_i \neq \emptyset$. Then the term $t_S \equiv t_1 + t_2 + \ldots + t_m$ will be called the description of the system S. Then we have

Fact 4: $\models_S t_1 + t_2 + \ldots + t_m$.

Now we are going to consider the concept of definability. We will say that a set $Y \subseteq X$ is definable in the language L_S, if there is a term t of L_S such that $Y = val_S\ t$. We have

Fact 5: If the system S is selective, then every subset of the universe X is definable in L_S.

Suppose the set $Y \subseteq X$ is definable in L_S. Then the term $t_Y \equiv t_1 + t_2 + \ldots + t_k$ with elementary terms $t_1, t_2, \ldots, t_k$ such that $val_S\ t_i \neq \emptyset$ and $Y = val_S\ t_1 \cup \ldots \cup val_S\ t_k$ is called a characteristic description of Y.

If the set Y is not definable in L_S, then we may construct approximate descriptions of Y. By an upper description of a set $Y = \{x_1, x_2, \ldots, x_m\}$ we mean the term $\bar{t}_Y \equiv t_1 + t_2 + \ldots + t_k$, $k \geq 1$, where $t_1, t_2, \ldots, t_k$ are elementary terms such that $x_i \in val_S\ t_i$ holds for every i = 1, 2, ..., k. Then we have $Y \subseteq val_S\ \bar{t}_Y$. The set $val_S\ \bar{t}_Y$ is said to be an upper approximation of the set Y and will be denoted by $\bar{Y}$. Analogously, the lower description of a set $Y = \{x_1, x_2, \ldots, x_l\}$, $l \geq 1$, is established by the term $\underline{t}_Y \equiv t_1 + t_2 + \ldots + t_l$, $l \leq m$, where $t_1, t_2, \ldots, t_l$ are all the elementary terms such that $val_S\ t_i \subseteq Y$ for i = 1, 2, ..., l.

We have $val_S\ \underline{t}_y \subseteq Y$. The set $val_S\ \underline{t}_y$ is said to be a lower description of the set Y and denoted by $\underline{Y}$.

Given a system $S = (X,A,V,\rho)$, we will say that a set $B \subsetneq A$ is the reduct of A, if B is the minimal set such that $\tilde{B} = \tilde{A}$. Notice that $\tilde{A}$ and $\tilde{B}$ are the partitions on X being generated by the set of attributes A and B respectively. $\tilde{B} = \tilde{A}$ implies: if a pair of objects cannot be distinguished by attributes of B, then it cannot be distinguished by attributes of A.

A concept C will be characterized by a partition $Z = \{Y,Y'\}$, where Y is the extension of C and Y' the complement of Y with respect to the universe X. The problem we will deal with in the next sections is to find the minimal subset of a given set of attributes being necessary to describe the partition Z. As mentioned above concept learning is closely connected with the difference between essential and inessential features of objects.

4 Characteristic Description of Knowledge

We begin with the case where a characteristic description of knowledge is available. Learning consists in dropping inessential features of objects.

Let us consider the system

$$S_1 = (X,A_1,V_1,\rho_1), \text{ where}$$
$$X = \{P1,P2,\dots,P12\},$$
$$A_1 = \{saf,mob,num,int\},$$
$$V_{saf} = \{c,c'\},$$
$$V_{mob} = \{0,1,\dots,8\},$$
$$V_{num} = \{0,1,\dots,32\},$$
$$V_{int} = \{p,p'\},$$
$$V_1 = V_{saf} \cup V_{mob} \cup V_{num} \cup V_{int}.$$

The function $\wp_1 : X \times A_1 \longrightarrow V_1$ is represented by the table in figure 2. The set of attributes A_1 induces a partition A_1 which is shown in figure 3. Notice that the system S_1 is not selective: the set of positions $\{P1, P2\}$ e.g. is not definable in S_1.

POSITION	saf	mob	num	int
P1	c'	O	3	p'
P2	c'	O	3	p'
P3	c'	O	3	p'
P4	c'	O	4	p'
P5	c'	O	4	p
P6	c'	O	5	p'
P7	c'	O	5	p
P8	c'	O	6	p'
P9	c'	O	6	p
P1o	c'	O	7	p'
P11	c'	O	7	p
P12	c	O	7	p'

Fig. 2: Table of the system S_1

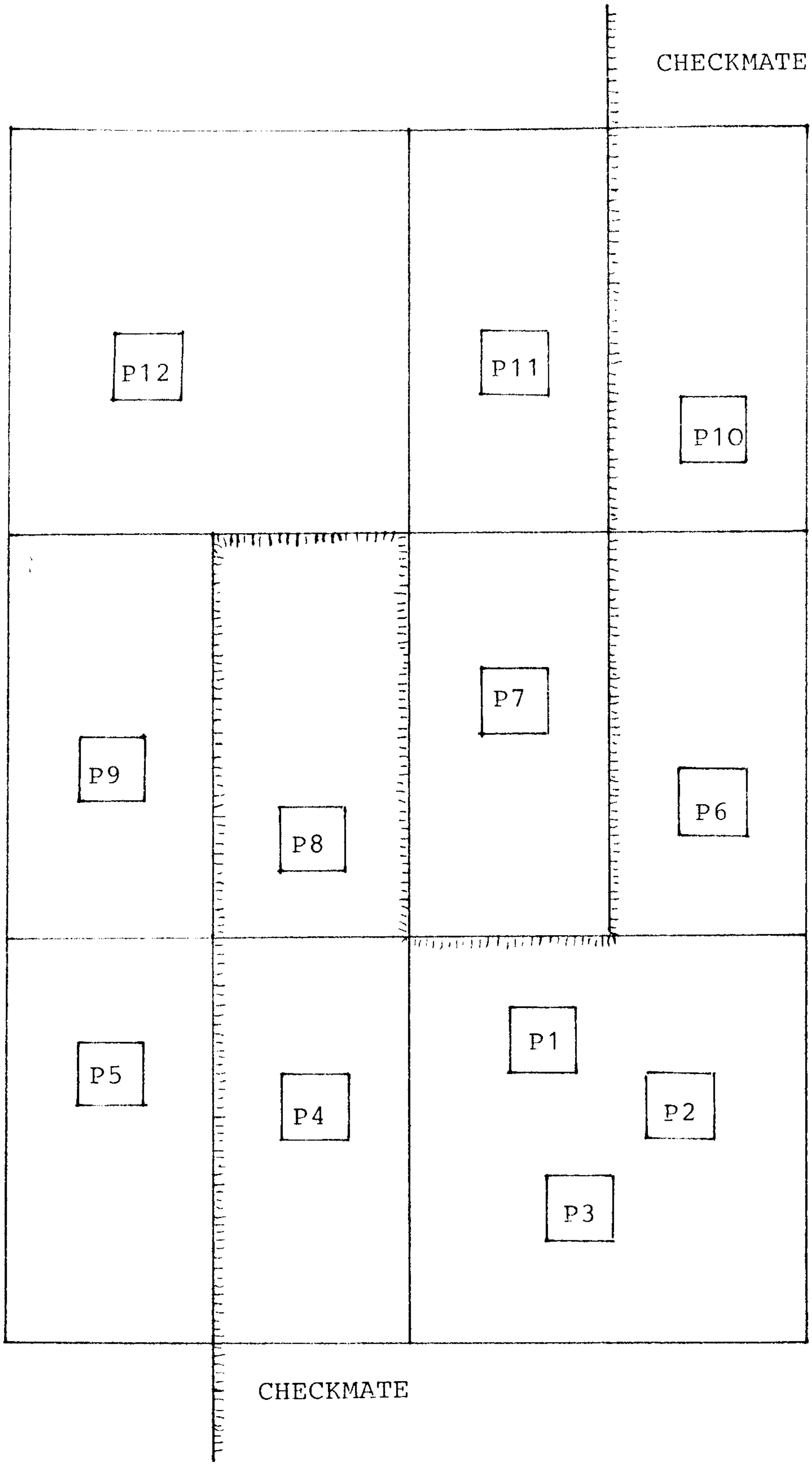


Fig. 3: Partition of the system S_1. Checkmate as union of elementary sets .

Our aim is to reduce the system S_1 with respect to the number of attributes. For this purpose we will consider the equivalence classes of the relations $\widetilde{saf}$, $\widetilde{mob}$, $\widetilde{num}$, and $\widetilde{int}$:

$$\widetilde{saf} = \{\{P1,\ldots,P11\}, \{P12\}\},$$
$$\widetilde{mob} = \{\{P1,\ldots,P12\}\},$$
$$\widetilde{num} = \{\{P1,P2,P3\}, \{P4,P5\}, \{P6,P7\}, \{P8,P9\}, \{P1o,P11,P12\}\},$$
$$\widetilde{int} = \{\{P1,P2,P3,P4,P6,P8,P1o,P12\}, \{P5,P7,P9,P11\}\}.$$

The equivalence classes of the relation

$$\widetilde{saf} \cap \widetilde{inf} = \{\{P1,P2,P3,P4,P6,P8,P1o\}, \{P5,P7,P9,P11\}, \{P12\}\}$$

cover the sets M (checkmate positions) and M' (complement of M), and none of the relations $\widetilde{saf}$, $\widetilde{mob}$, $\widetilde{num}$, and $\widetilde{int}$ have this property. From this we suppose that the attributes "mob" and "num" may be dropped for describing the checkmate positions.

Let us consider the term t_M being the characteristic description of the set M:

$$\begin{aligned} t_M \equiv\ & (saf,c') \cdot (mob,o) \cdot (num,3) \cdot (int,p') \\ + & (saf,c') \cdot (mob,o) \cdot (num,4) \cdot (int,p') \\ + & (saf,c') \cdot (mob,o) \cdot (num,5) \cdot (int,p') \\ + & (saf,c') \cdot (mob,o) \cdot (num,6) \cdot (int,p') \\ + & (saf,c') \cdot (mob,o) \cdot (num,7) \cdot (int,p') \end{aligned}$$

t_M is a conjunction of elementary terms. Obviously we have $val_{S_1}\ t_M = M$.

By abstraction from the properties "mobility of the king" and "number of pieces" we get the term

$$t'_M \equiv (saf,c') \cdot (int,p') .$$

Notice that $val_{S_1}\ t'_M = M$. Furthermore we have $\models_{S_1} t_M \Longleftrightarrow t'_M$.

Omitting the attributes "mob" and "num" we get the reduced system

$$\begin{aligned}
S_1^* &= (X, A_1^*, V_1^*, \rho_1^*) \text{, where} \\
X &= \{P1, P2, \ldots, P12\} , \\
A_1^* &= \{saf, int\} , \\
V_{saf} &= \{c, c'\} , \\
V_{int} &= \{p, p'\} , \\
V_1^* &= V_{saf} \cup V_{int} .
\end{aligned}$$

The function $\rho_1^* : X \times A_1^* \longrightarrow V_1^*$ is represented by the table in figure 4, the corresponding partition in figure 5.

The process of inductive generalization leading from S_1 to S_1^* is interpreted by us as learning. The safety of the checking piece and the impossibility of interposition are recognized as essential features of the concept of checkmate.

POSITION	saf	int
P1	c'	p'
P2	c'	p'
P3	c'	p'
P4	c'	p'
P5	c'	p
P6	c'	p'
P7	c'	p
P8	c'	p'
P9	c'	p
P1o	c'	p'
P11	c'	p
P12	c	p'

Fig. 4: Table of the reduced system S_1^*.

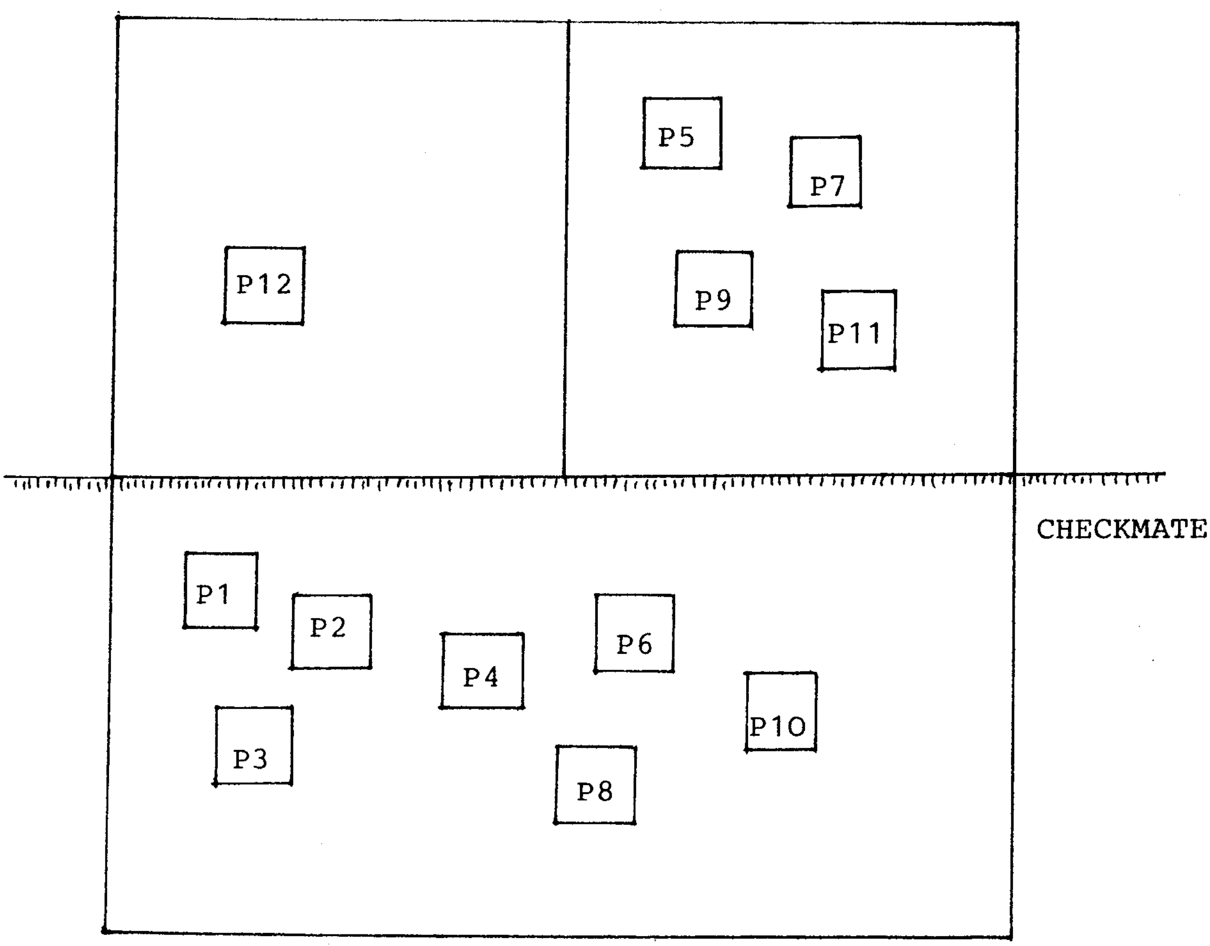


Fig. 5: Partition of the system S_1^*. Checkmate as an elementary set.

5 Approximate Description of Knowledge

Different from the system S_1 the system S_2 which we will analyse now has not incorporated all essential features of checkmate. In this case we only can get an approximate description reflecting the incomplete knowledge of the system S_2.

Let us analyse the knowledge representation system

$$S_2 = (X, A_2, V_2, \varrho_2), \text{ where}$$
$$X = \{P1, P2, \ldots, P12\},$$
$$A_2 = \{saf, mob, num\},$$
$$V_{saf} = \{c, c'\},$$
$$V_{mob} = \{o, 1, \ldots, 8\},$$
$$V_{num} = \{o, 1, \ldots, 32\},$$
$$V_2 = V_{saf} \cup V_{mob} \cup V_{num}.$$

The function $\varrho_2 : X \times A_2 \longrightarrow V_2$ is represented by the table in figure 6. The set of attributes A_3 induces the partition in figure 7. Like the system S_1 the system S_2 is not selective.

POSITION	saf	mob	num
P1	c'	O	3
P2	c'	O	3
P3	c'	O	3
P4	c'	O	4
P5	c'	O	4
P6	c'	O	5
P7	c'	O	5
P8	c'	O	6
P9	c'	O	6
P1o	c'	O	7
P11	c'	O	7
P12	c	O	7

Fig. 6: Table of the system S_2

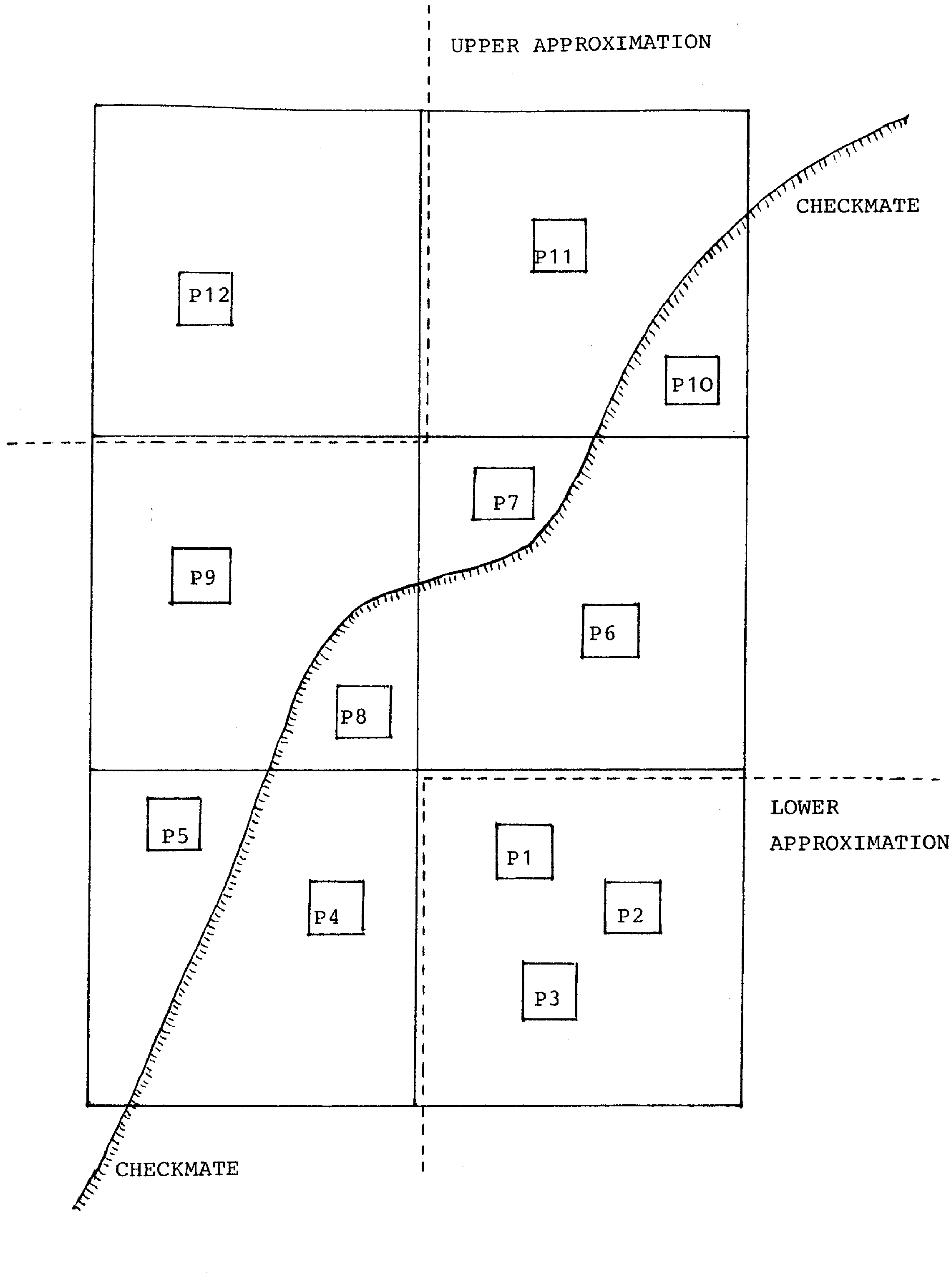


Fig. 7: Partition of the system S_2. Lower and upper approximation of checkmate.

In order to reduce the number of attributes we consider the equivalence classes for the relations $\widetilde{saf}$, $\widetilde{mob}$, and $\widetilde{num}$:

$$\widetilde{saf} = \{\{P_1,\ldots,P11\}, \{P12\}\}$$

$$\widetilde{mob} = \{\{P1,\ldots,P12\}\},$$

$$\widetilde{num} = \{\{P1,P2,P3\}, \{P4,P5\}, \{P6,P7\}, \{P8,P9\}, \{P1o,P11,P12\}\}$$

It follows

$$\widetilde{\{saf,mob,num\}} = \widetilde{saf} \cap \widetilde{mob} \cap \widetilde{num} = \widetilde{saf} \cap \widetilde{num} = \widetilde{\{saf,num\}}$$

Therefore the set of attributes $\{saf,mob,num\}$ can be reduced to $\{saf,num\}$.

Figure 7 shows that the set of checkmate positions $M = \{P1,P2,P3,P4,P6,P8\}$ cannot be represented as a union of elementary sets. $\{P1,P2,P3\}$ is a lower approximation of M, and $\{P1,\ldots,P11\}$ an upper approximation of M. Thus we have two approximate descriptions for the concept of checkmate.

a) Lower description of checkmate:

$\underline{t}_M \equiv (saf,c') \cdot (num,3)$

It follows $val_{S_2}\ \underline{t}_M \subset M$.

In this description the number of pieces comes out as an essential property. So the positions P4, P6, and P8 cannot be recognized as checkmate.

b) Upper description of checkmate:

$$\bar{t}_M \equiv (saf,c') \cdot (num,3)$$
$$+ (saf,c') \cdot (num,4)$$
$$+ (saf,c') \cdot (num,5)$$
$$+ (saf,c') \cdot (num,6)$$
$$+ (saf,c') \cdot (num,7)$$

It follows $M \subset val_{S_2}\ \bar{t}_M$.

By abstraction from the property "number of pieces" being inessential in this case we have

$$\bar{t}'_M \equiv (saf,c')$$

Because of $\models_{S_2} \bar{t}_M \longleftrightarrow \bar{t}'_M$ we conclude that the attribute "num" is redundant for describing the upper approximation of the concept of checkmate.

By the definitions $\underline{M} = val_{S_2}\ \underline{t}_M$ and $\bar{M} = val_{S_2}\ \bar{t}_M$

we have

$$\underline{M} \subset M \subset \bar{M} .$$

Thus we have shown that approximate concept learning delivers the following classes of positions (figure 8):

a) Surely checkmate:

$$\underline{M} = \{P1,P2,P3\} .$$

b) Possibly checkmate:

$$\bar{M} = \{P1,P2,P3,P4,P5,P6,P7,P8,P9,P1o,P11\} .$$

c) Impossibly checkmate:

$$X - \bar{M} = \{ P12 \} .$$

Syntactic methods for deriving abstractions and recognizing approximations are developed in /4/. A machine implementation of these techniques will be finished in the near future.

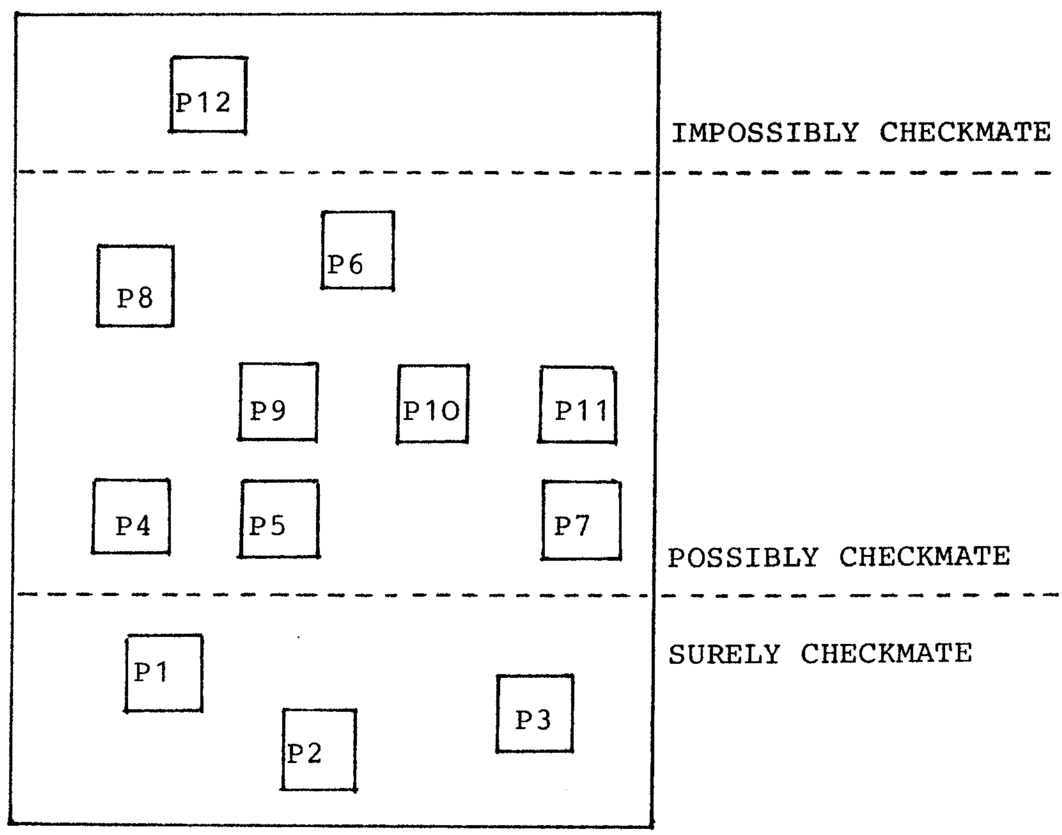


Fig. 8: Approximate concept of checkmate.

6 Conclusion

A method for learning approximate concepts from examples is described. The problem has been tackled from a semantical point of view giving deeper insight into the conceptual foundations. Current work is dedicated to an implementation of syntactic techniques.

References

/1/ Banerji, R.B.: Artificial Intelligence - A Theoretical Approach, North Holland 198o.

/2/ Brachman, R.J.; Smith, B.C.: Special Issue on Knowledge Representation, SIGART Newsletter 7o (198o).

/3/ Kažić, B.M.: The Chess Competitor's Handbook, Batsford London 198o.

/4/ Konrad, E.; Orłowska, E.; Pawlak, Z.: Knowledge Representation Systems - Definability of Informations, ICS PAS Report 433, Warsaw 1981.

/5/ Michalski, R.S.: Pattern Recognition as Rule-Guided Inductive Inference, IEEE Transactions on Pattern Analysis and Machine Intelligence, No. 4, 198o

/6/ Morgan, C.G.: Hypothesis Generation by Machine, Artificial Intelligence 2 (1971), 179-187.

/7/ Pawlak, Z.: Mathematical Foundations of Information Retrieval, CC PAS Report 1o1, Warsaw 1973.

/8/ Pawlak, Z.: Classification of Objects by Means of Attributes - An Approach to Inductive Inference, CC PAS Report, Warsaw 198o.

/9/ Plotkin, G.D.: A Note on Inductive Generalization, Machine Intelligence 5 (197o), 153-163.

/1o/ Plotkin, G.D.: A Further Note on Inductive Generalization, Machine Intelligence 6 (1971), 1o1-124.

/11/ Popplestone, R.J.: An Experiment in Automatic Induction, Machine Intelligence 5 (197o), 2o3-215.