SHIFT OF BIAS IN LEARNING FROM DRUG COMPOUNDS: THE FLEMING PROJECT

L. DI PACE (*) F. FABROCINI (*) G. BOLIS (**)
(*) IBM Rome Scientific Center, AI Group, via Giorgione 159, 00147 Rome. Italy
(**) Farmitalia Carlo Erba srl Erbamont Group, R&D/CAMD, via dei Gracchi 35, 20146 Milan. Italy

Abstract

In this paper we will illustrate the results of a machine learning application concerning drug design. Dynamic bias management, in this context, will be presented as a critical mechanism to deal with complex problems in which good representations are unavailable even to human experts. A number of domain-dependent and domain-independent operators which allow automatic bias adjustment will be discussed with the mechanisms used to decide when and how to vary bias. Finally, we will summarize the results that a system named FLEMING adopting these techniques has obtained on the domain of the inhibitors of the thermolysin enzyme.

Keywords: bias management, constructive learning, learning from discovery, computer-aided molecular design.

Introduction

A fascinating aspect of machine learning techniques is their use to discover knowledge unavailable even to human experts. We will describe a system named FLEMING that discovers explanatory theories from a set of experimental observations concerning compounds synthesized during the drug design process. Therefore, our investigation falls into the category of the task-oriented studies; even so the need to deal with a real-world problem suggested a number of theoretically-oriented considerations. This paper is intended to be a report on our experience in applying machine learning techniques to a complex problem.

Firstly, we will briefly outline the learning problem; the limits of the approaches traditionally adopted to face this problem will be also illustrated. Then, we will sketch the general methodology used to deal with the learning task previously described. The remainder of the paper will concentrate on FLEMING's approach to dynamic bias management. We will illustrate, as regards our application, where the need for automatic bias adjustment comes from and the solutions adopted to decide when and how to vary bias. Specifically, dynamic bias management will be presented as a necessity in real-world problems which typically exhibit a large number of distinct disjuncts (see Rendell & Cho, 1990). Finally, a summary of FLEMING's results on the system of the inhibitors of the thermolysin enzyme will be presented and discussed.

The domain problem

The study of the correlations between pharmacological activity and molecular structure is a central issue in the drug design process. Such a study is based on the concept that a biological (or pharmacological) effect caused by a given molecule (drug) is a function of structural or electronic properties of that molecule. Conventional approaches to the Structure-Activity Relationship (SAR) problem make use of statistical techniques in order to relate the activity of the compound to substructures and/or properties used to describe the compound itself (see, for instance, Martin, 1978). The great difficulty in all the SAR studies is the selection of the molecular descriptors as their number is very large. Such a choice introduces a strong bias on the results of the analysis insofar as it abstracts secondary objects to be the "real" objects from which the learning system generalizes. The bias comes from an assumption behind this choice which says that no considerations useful to capture the target concept have been omitted from the language used to represent the instances.

Klopman's CASE program is presented as an attempt to overcome such difficulty (see Klopman, 1984). CASE is capable of manipulating a molecular structure in order to generate all the fragments that can be formed by breaking the molecule. In this way, the selection of the descriptors would not be constrained by the prejudices of the investigators. Once the fragments have been collected, they are analyzed statistically to discover those fragments relevant to the activity of the compounds. CASE, however, suffers from other drawbacks common to most of the SAR studies. Firstly, the functional dependence existing betweeen the experimental observations and the pharmacological activity of the compound is assumed to be the sum of independent contributions of the most relevant activating/deactivating fragments automatically selected by the program. Yet, context cannot be ignored when talking about molecular structures since the relevance of the compound fragments is heavily dependent on the relative positions in the compound itself. Moreover, the properties characterizing the fragments play a fundamental role in determining the binding between compound and receptor. In this sense, statistical frequency does not appear a meaningful approximation of the relevancy of the fragments. Indeed, medicinal chemists while visually analyzing a set of compounds for determining their structure-activity relationship do not constrain the investigation at the identification of a number of fragments deemed to be invariant in the active compounds. Instead, they attempt to build a theoretical model suitable to explain the nature of the binding between active compounds and biological receptor. A set of notions that actually make the basic knowledge of a medicinal chemist about molecules and their properties such as, for instance, hydrophobicity, polarity, hydrogen bonding, etc. play an important role during the explanation process. Our attempt was to devise a computational tool which could support medicinal chemists to reduce the complexity of the problem without suffering from the drawbacks of the more traditional approaches.

Learning methodology

FLEMING's input is a database of compounds and their inhibitory activity as reported in Figure 1. Specifically, molecules are formalized in terms of a set of atoms and their connectivity. The atoms are characterized as atom types, i.e. there is an attempt to define their electronic state.

K_i Values of Inhibitors of Thermolysin

No.	Compound†	K_i (μM)
1	Z-NHNH-CS-NHNH2	6700
2	Z-Agly-Leu-NHNH2	380
3	Z-Gly-Leu-NHNH2	1100
4	Ac-Ala-Aphe-Leu-NHNH2	6500
5	Ac-Ala-Ala-Aala-Leu-NHNH2	7900
6	L-Leu-NHOH	190
7	Z-L-Leu-NHOH	10
8	Z-Gly-L-Leu-NH2	21000
9	Z-Gly-L-Leu-NHOH	13
10	Z-Gly-L-Leu-N(CH3)OH	2230
11	Z-Gly-L-Leu-NHOCH3	No Inhib.
12	Z-Agly-L-Leu-NHOH	27
13	Z-Gly-Gly-NHOH	940
14	Z-Gly-Gly-L-Leu-NHOH	39
15	HONH-Bzm-OEt	20
16	HONH-Bzm-L-Ala-Gly-NH2	0.66
17	HONH-Bzm-L-Ala-Gly-OH	0.65
18	HONH-Ibm-L-Ala-Gly-NH2	0.48
19	HONH-Mal-L-Ala-Gly-NH2	1100
20	HO-Bzm-L-Ala-Gly-NH2	420
21	CHO-HOLeu-L-Ala-Gly-NH2	3.8
22	Ac-HOLeu-L-Ala-Gly-NH2	3400
23	P-NH-Et	No Inhib.
24	P-Leu-NH2	1.3
25	P-Phe-OH	73
26	P-Ala-Ala-OH	88
27	P-Ile-Ala-OH	0.36
28	P-Leu-Phe-OH	0.019
29	Z-Phe-Gly-NH2	350
30	Z-Phe-Gly	4500
31	Phe-Gly-NH2	10900
32	Phe-Gly	10300
33	Z-Leu-Gly-NH2	3070
34	Z-Leu-Gly	4030
35	Leu-Gly-NH2	8300

† Z, benzyloxycarbonyl; Agly, -NHNHCO-; Aala, -NHN(CH_3); Aphe, -NHN($CH_2C_6H_5$)CO-; P, phosphoryl group $(HO)_2PO$-; Bzm, benzylmalonyl -COCH($CH_2C_6H_5$)CO-; Ibm, Isobutylmalonyl -COCH($CH_2CH(CH_3)_2$)CO-; Mal, malonyl; Ac, Acetyl; Et, ethyl.

Figure 1. THE SET OF COMPOUNDS. THE ACTIVITY IS EXPRESSED IN TERMS OF K_i VALUES DETERMINED BY DIXON PLOTS.

We will now briefly outline the operation of the system. It mirrors the strategy actually used by the medicinal chemist during the drug design process. Yet, because of the "cognitive overload" problem, the medicinal chemist is forced to reason locally while the nature of the problem would require spanning over a large number of variables. At the very beginning, the program identifies a set of active/inactive pairs deemed to be useful to make the learning process effective. Specifically, the

system looks for compound pairs which, on the one hand, maximize the difference in activity, $\Delta(K)$, and, on the other, minimize the difference in structure, $\Delta(S)$. Usually such a set is quite large because the medicinal chemist proceeds step-by-step via small modifications to the previously experimented compounds. Although methods for quantifying differences in structure exist, at the present time two molecules are regarded as structurally similar whenever they have at most one difference in terms of residues.

Once the pairs which appear to be more informative have been defined, each compound in each pair is matched against the other in order to generate the fragments supposed to be responsible for the $\Delta(K)$. We will refer to the fragment appearing in the active compound c_i as Required (REQ_{ij}) and the fragment in the inactive compound c_j as Forbidden (FOR_{ji}). One problem, here, was the definition of the level of abstraction more appropriate for detailing such a fragment. FLEMING automatically makes the choice of the level of representation considered to be more adequate. Whenever the difference between the compounds in each pair is made of only one residue (as it is by definition), the system will attempt at reformulating such a difference in terms of functional groups. Yet, the reformulation step will be accomplished only when satisfied the condition that the functional groups which actually make the difference also make a connected region. The same holds as concerns the functional group level. In this case, when the difference is expressed in terms of at most one functional group, the system will reformulate such a difference in terms of atom types on the condition quoted above. The description of the molecules, then, is appropriately shifted so as to allow REQ_{ij} and FOR_{ji} being encoded explicitly in the representation of the compounds c_i and c_j. What we get is an abstraction space deductively derived from the domain knowledge in which a number of clues useful to speed up the learning task have been marked inside the compound which they come from after being opportunely reformulated. This process can be regarded as a form of constructive induction in which domain-dependent knowledge is used to derive a "useful" instance representation suitable to facilitate the inductive task (see Flann & Dietterich, 1986; see also Drastal, Czako & Raatz, 1989). Specifically, because by assumption the active/inactive molecules of each pair differ at most for one fragment, we think of REQs as the fragments necessary to the activity of the compound while FORs must be considered as "forbidden" insofar as they involve a decrease in the activity.

FLEMING's learning strategy will be illustrated as follows. Let C be the set of compounds. At the beginning of each iteration, a K number of compounds c_u deemed to be useful for improving the current model *PMODEL* is selected from C . The system will evaluate a compound as useful when sufficiently similar to *PMODEL*, so as to reduce the ambiguities of matching, and when including a number of REQs and FORs, so as to maximize the information content. Whenever c_u is a maximally active compound, FLEMING generalizes on the compound as a whole. In this case, FLEMING will consider all the substructures which the molecule is made of as noise-immune and, therefore, equally relevant to the inhibitory activity. Otherwise FLEMING will take into account only the REQs and the FORs included in c_u. Specifically, the system will generalize *PMODEL* whenever encountering REQs and specialize whenever encountering FORs. A quality function is used to evaluate each of the new models. FLEMING will accord its preference to those models that minimize the predictive error. In other words, the system will prefer the models that predict, for each compound c_i, an activity value $\tilde{a}_i$ as close as possible to the real one a_i. The new models so generated are merged into a K-limited OPEN list of current partial models and only the best K models are maintained for further expansions.

Selecting the most appropriate bias

Mitchell defined bias as any information for controlling the complexity of the learning problem that can be considered as "extra-evidential" in the sense that it does not come from the objects or events to be described by the target concept (see Mitchell, 1980; see also Utgoff, 1986 and Rendell, Seshu & Tcheng, 1987). The satisfactory performance that researchers get from learning programs are often due to the bias previously "hard-coded" in the program. What they do, then, is to test the program and eventually shift the bias "by hand" to move to a better bias. This methodology allows inductive systems to control the combinatorial explosion of hypotheses; yet, in this way, the learning task is reduced to find a "trick" for discarding inconsistent hypotheses as instances are examined.

The use of a highly restricted hypothesis space is a very common method for focussing the search on the set of preferred hypotheses. This kind of bias has been named "restriction-type bias" (see Russell & Grosof, 1989). The most common way to control the dimensions of the hypothesis space is to restrict the object description to an abstraction of the complete observation (instance-language bias). Such an approach is successful when the objects are described in a form suitable to capture the target concept, but when the most appropriate language to describe the objects is not known a priori, then learning can be impossible using selective methods.

Our problem falls into this class. As we have seen above, all the traditional SAR methods presuppose some knowledge of the chemical or electronic properties that make the binding between molecule and receptor feasible. In other words, what these approaches presuppose is the very answer to the problem at hand. FLEMING's approach is to start the search from a restricted hypothesis space and eventually to shift to a less restricted one (see also Utgoff, 1986). The nature of the problem itself suggested this approach. As Rendell and Cho have stressed, real-world problems and, especially, complex problems in which even human experts lack understanding typically exhibit a very large number of distinct disjuncts (see Rendell & Cho, 1990). In our case, the low-level descriptors used to represent the compounds are not sufficient to derive the explanatory model we look for. What we need is a more theoretically-oriented language exploiting higher-level regularities suitable to compact the problem (i.e. suitable to produce class membership functions having at most few disjuncts). A number of operators which result in the introduction of new descriptors will accomplish the mapping between the problem definition in the initial space and its definition in an abstraction space more appropriate for capturing the target concept.

Shift of bias is closely related to constructive induction. Specifically, when we apply a set of operators to one or more existing descriptors to generate new descriptors intended for use in describing the target concept, we also shift the concept-language bias. One problem we faced was to identify exactly when and how to shift bias.

FLEMING makes use of three constructive operators. We distinguish:

- ***Domain-dependent operators***:
  - ***R-operators***: Allow the introduction of multiple levels of abstraction; specifically they take the form:

$$R_i\colon\ S_i \rightarrow G(\{S_{i1}, \ldots, S_{in}\},\{A_{l1}, \ldots, A_{lm}\})$$

where S_i is a generic fragment and G is the graph defined by the set of nodes $\{S_{ik}\}$ and the set of arcs $\{A_{ij}\}$ such that each S_{ik} is a substructure of S_i. The R-operators work selectively via the identification of compound subunits. Thus, the same molecule can be formalized as a hierarchy of representations that proceeds from a very detailed formalization of the compound to a coarse one as illustrated in Figure 2.

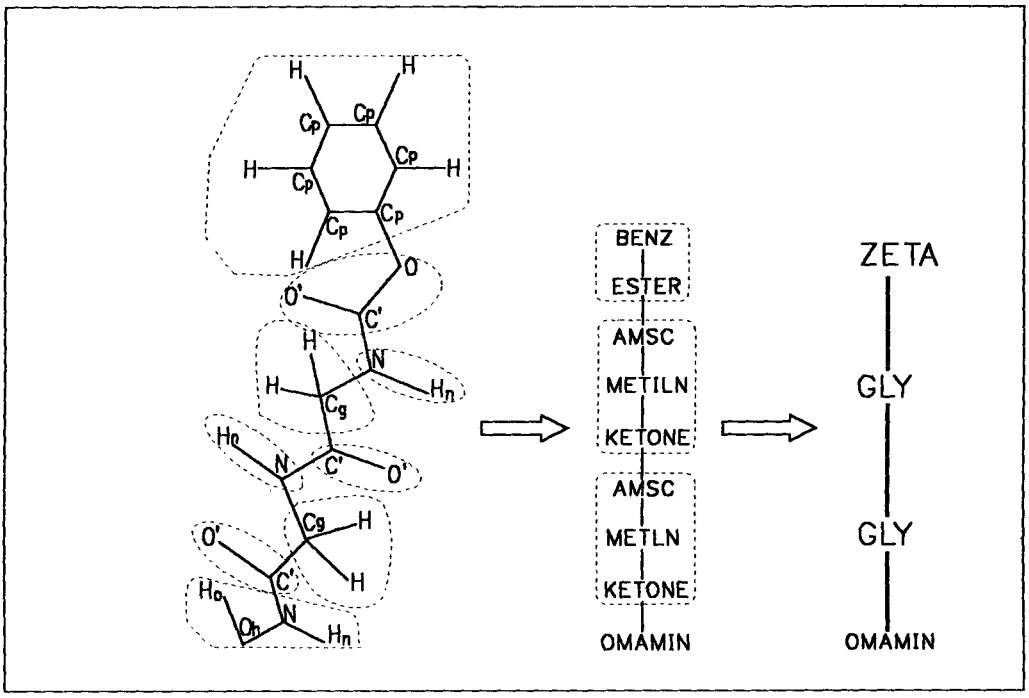


Figure 2. THE HIERARCHICAL REPRESENTATION OF THE COMPOUNDS

Specifically, atomic entities, as we have just said above, are used to represent the compound at the lower level; in the following one, the molecule is represented in terms of larger fragments named functional groups (for instance, a carboxil group); at the next level up, the compound is described in terms of residues (taken in a broad sense, for example, an amino acid in a peptide); finally it is always possible to represent the molecule as a whole with some properties associated with the compound such as, for instance, the molecular weight.

- *I-operators*: Allow switching from the observation language used to describe the compounds to a more theoretically-oriented language of chemical properties associated with the hierarchy of molecular substructures S_i; I-operators have been formalized as follows:

$$I_i\colon\ S_i \rightarrow P_{i1}, P_{i2}, \ldots, P_{in}$$

Thus, the system, for instance, is told that a hydroxyl group is a hydrophobic group with the property of being electron donor. Some of these properties such as, for instance, volumes and partial charges are computed via a set of domain-dependent procedures.

Both the R-operators and the I-operators make the prior knowledge used by the system for guiding the inductive process.

- ***Domain-independent operators***:

  - ***U-operators***: Such operators are defined as boolean combinations of other terms previously defined in the system; thus, for instance, FLEMING makes use of the following operator:

$$U_1\colon\ S_i = G(\{S_{ik}\}\{A_{ij}\}) \cup S_w = G(\{S_{wk}\}\{A_{wj}\}) \rightarrow S_z = G(\{S_{ik}\} \cup \{S_{wk}\},\{A_{ij}\} \cup \{A_{wj}\})$$

Thus, the fragment R_1 is a *new* descriptor previously unknown to the system resulting from the application of the operator U_1 to F_1 and F_2 (see Figure 3). R_1 will also inherit the properties common to F_1 and F_2.

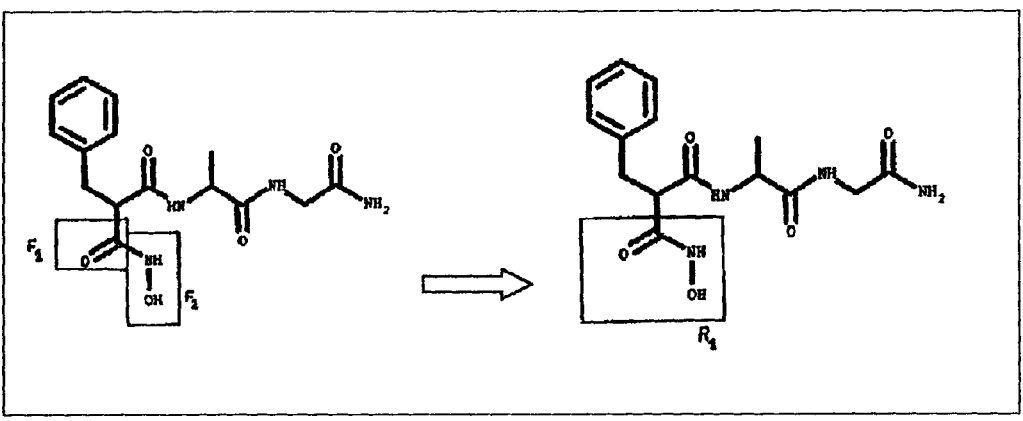


Figure 3. AN APPLICATION OF THE U-OPERATOR: (F1,F2) = > R1

Rendell has identified four aspects considered to be inherent to the constructive induction problem (see Matheus & Rendell, 1989; see also Wrobel, 1988):

1. detection of when construction is required;
2. selection of the constructive operators;
3. generalization of the new descriptors;
4. evaluation of the new features so derived.

1. ***Detection***: Detection, in our approach, is triggered via a set of evaluation functions. Because of their nature, these functions introduce a meta-reasoning mode dedicated to control the adequacy of the representational level. Specifically, constructive induction is performed either in a *demand-driven* fashion whenever the model building process faces an impasse as when the system

does not appear successful in finding a matching between c_u and *PMODEL* or in a *tentative* way as when FLEMING applies the R-operators in the attempt at finding the level of abstraction more adequate for detailing REQs and FORs.

2. ***Selection***: our system makes use of a two-step approach to select the useful constructors:

   - firstly FLEMING uses the domain knowledge as a set of heuristics for focussing on the class of useful operators;
   - then it completes the selection process with the selection of the appropriate operator by using information in the data set.

   We are now ready to discuss one of the detection mechanisms used by FLEMING and the selection procedure adopted to choose the most appropriate constructor among the set of candidate operators. The function used in this example is triggered as a latter resource to overcome an impasse whenever the R-operators have been already exhaustively applied to the description of the compound (the attempt, here, is at finding a match between an c_u and *PMODEL*). We call G_u the graph used to represent c_u and G_m the graph representing *PMODEL* where $G_u = (\{S_u\},\{A_u\})$ and $G_m = (\{S_m\},\{A_m\})$. Then let us define $P_{k\,i} = \{p_{k1}, \ldots, p_{kn}\}$ as the set of properties appended to the i -th node $S_{k\,i}$ of G_k by the application of the I-operators. Moreover, let $MATCH(G_k,G_z)$ be the function matching the graph G_k onto the graph G_z. $MATCH(G_u,G_m)$ produces as output a set of pairs $(S_{u\,i}, S_{m\,j})$. Then, the evaluation function can be stated as follows:

$$EF_1\colon\ \forall(S_{u\,i}, S_{m\,j}) \ni (S_{u\,i} \neq S_{m\,j}) \wedge (P_{u\,i} \cap P_{m\,j}) = \emptyset \rightarrow U_t(S_{u\,i}, x)$$

   in which the condition-part asserts that whenever a pair of fragments $S_{u\,i} \neq S_{m\,j}$ do not also match in terms of properties, the set of the U-operators must be selected. Now FLEMING must select the fragment x more useful to construct the new descriptor when joined with $S_{u\,i}$. The construction of the new descriptor should finally allow the generalization between c_u and *PMODEL*. Specifically, FLEMING will prefer those descriptors and, therefore, those operators that maximize the number of common properties between $(S_{u\,i}, x)$ and the corresponding fragment $(S_{m\,j}, y)$. A simple inspection of the data at hand will finally produce the selection of the appropriate operator.

3. ***Generalization***: The application of the standard selective methods such as dropping condition, turning constants into variables, closing interval, etc., allows the generalization of the new descriptors (Michalski, 1983). Yet, constructive induction, in our case, is successful only insofar as FLEMING encodes domain knowledge which works as selection bias. FLEMING, actually, utilizes domain heuristics to prune the space of the U-operators (the boolean operators) to a candidate set far more tractable. Such heuristics, specifically, will narrow the search to those operators which generate new descriptors consisting of adjacent fragments.

4. ***Evaluation***: as other systems such as STAGGER (see Schlimmer, 1987) and CITRE (see Pagallo, 1989), FLEMING is completly autonomous in the evaluation of the quality of new descriptors. This methodology must be contrasted with DUCE's oracle-based approach in which the evaluation is delegated to the user (see Muggleton, 1987). In FLEMING's case, the quality of the new descriptors is assessed with the results of the generalization process itself. Yet, such an approach

results in an increase of complexity for the system insofar as it also carries out an increase in the number of models to be considered by the system. We are studying, at the moment, more powerful evaluation mechanisms to deal with this problem.

Results and Discussion

In order to test FLEMING's performance we have chosen the system of the inhibitors of the thermolysin enzyme for which 3D-structural information of the complex enzyme-inhibitor is available from the Brookhaven Protein Data Bank. Thermolysin is a thermostable metalloprotease involved in several important physiological processes and, like other metalloproteases, contains a zinc ion essential for activity. It is also known that thermolysin is very specific for hydrophobic aminoacids. The data concerning the inhibitory activity of a number of thermolysin inhibitors were taken from literature and reported in Figure 1. The activity is expressed in terms of K_i values determined by Dixon plots (see, for further details, Bolis, Di Pace & Fabrocini, forthcoming).

To start with, FLEMING reformulates the description of the compounds so as to allow REQs and FORs being encoded explicitly. In this way, the instance space is trasformed into a more abstract space which exploits higher-order regularities in order to improve concept concentration. Thus, the pair (16)/(20) is selected from the compound database. Such a pair satisfies the similarity constraint we adopted, insofar as compound (16) differs from compound (20) only because of one fragment (i.e. the fragment named HONH). In other words, given the same context (i.e. * - BZM - L - ALA - GLY - NH2, where the star denotes the variable), then the substitution of the HO fragment (20) with the HONH fragment (16) causes a major change in the activity of the compound. At this point, FLEMING attempts to detail as much as possible such a difference. In this case, the system will finally generate as difference between compounds (16) and (20) the atom type ***n'*** (see Figure 4).

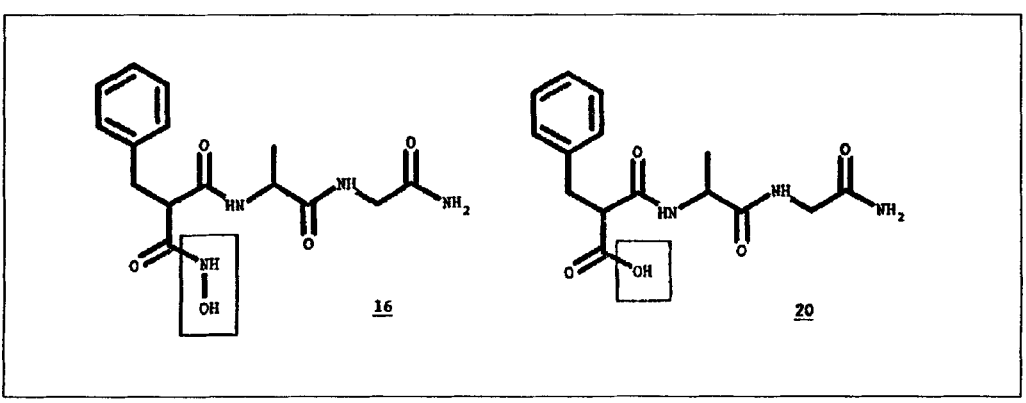
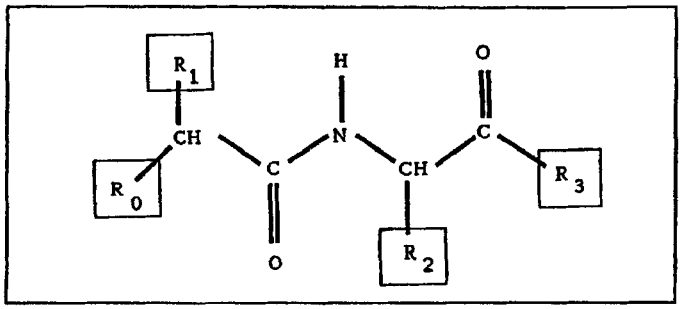


Figure 4. A CASE OF REFORMULATION. THE STRUCTURAL DIFFERENCE EVIDENTIATED BY FLEMING BETWEEN COMPOUNDS 16 AND 20.

A schematic representation of the explanatory model generated by FLEMING, as concerns our problem, is illustrated in Figure 5 where the inhibitor is depicted as composed by a number of generalized fragments named as $R_0, \ldots, R_n$.



Covering: 16,17,18,21,27,28

R0 A: 21 ≤ Volume ≤ 31
B: Hydrophobicity = No
C: dist (HB donor atom, CH) = 1,2,3
D: dist (HB acceptor atom, CH) = 2,3
E: dist (N,CH) = 1 or 2 bonds
F: dist (O [atomtype oh],CH) = 3 bonds
G: dist (O [atomtype o'],CH) = 2 or 3 bonds
H: dist (furthest heavy atom,CH) = 3

R1 A: Hydrophobicity = Yes
B: 34 ≤ Volume ≤ 60

R2 A: Hydrophobicity = Yes
B: 13.6 ≤ Volume

R3 *

Positive and negative instances of the generalized fragments

R0 : positive
1: HONHCO (16,17,18)
2: PO(OH)2NH (27,28)
3: CHONOH :
[exception to F:
dist (O,CH) = 2] (21)

negative
1: HOCO : ¬ (H) and ¬ (F) (20)
2: CH3CONOH : ¬ (A) and ¬ (B) (22)

R1 : positive
1: CH2CH(CH3)2 (18, 21, 28)
2: CH2*Benz* (16, 17, 27)

negative
1: H : ¬ (A) and ¬ (B) (19)

R2 : positive
1: CH3 (16, 17, 18, 27)
2: CH2*Benz* (28)

Lowest activity of a covered compound : 3.8
Highest activity of an uncovered compound :420

Figure 5. THE FINAL MODEL. R_0, R_1 AND R_2 ARE THE BINDING LOCATIONS DISCOVERED BY FLEMING DESCRIBED IN TERMS OF PROPERTIES (ON THE LEFT) AND FRAGMENT INSTANCES (ON THE RIGHT).

The same figure illustrates their characterization where, for each R_i , a number of properties labelled with a capital letter and a number of positive and negative instances of fragments that the model is covering are associated to the fragment. The model covers six among the most active compounds of Figure 1, namely no. 16, 17, 18, 21, 27 and 28. Wherever the inductive process did not lead to the definition of a generalized fragment, such a fragment is reported in the model as it is; in our case the CH, NH and C=O groups. Thus, in FLEMING's model, the fragment R_0 is defined as hydrophylic, hydrogen bond donor and acceptor while the fragment R_1 is described as hydrophobic with volume values between 34 and 60. Both fragments R_0 and R_1 are supposed to be essential for activity. Fragment R_2 is also hydrophobic while fragment R_3 does not seem to be very relevant for any change of activity of the compounds considered.

These results obtained via the machine learning techinques previously described are fully in accord with experimental observations made on the crystal structure of thermolysin complexed with an hydroxamic acid inhibitor (see Holmes & Matthews, 1981). In this complex (see Figure 6), the group HONHCO (corresponding to the R_0) is making hydrogen bond interactions with the active site Zn and with residues ALA 13 and GLU 143. Furthermore, the benzyl group in the malonyl moiety corresponding to R_1 is found in a hydrophobic pocket of the active site while the following ALA side chain corresponding to R_2 is found in proximity of LEU 202 and PHE 130 of the enzyme.

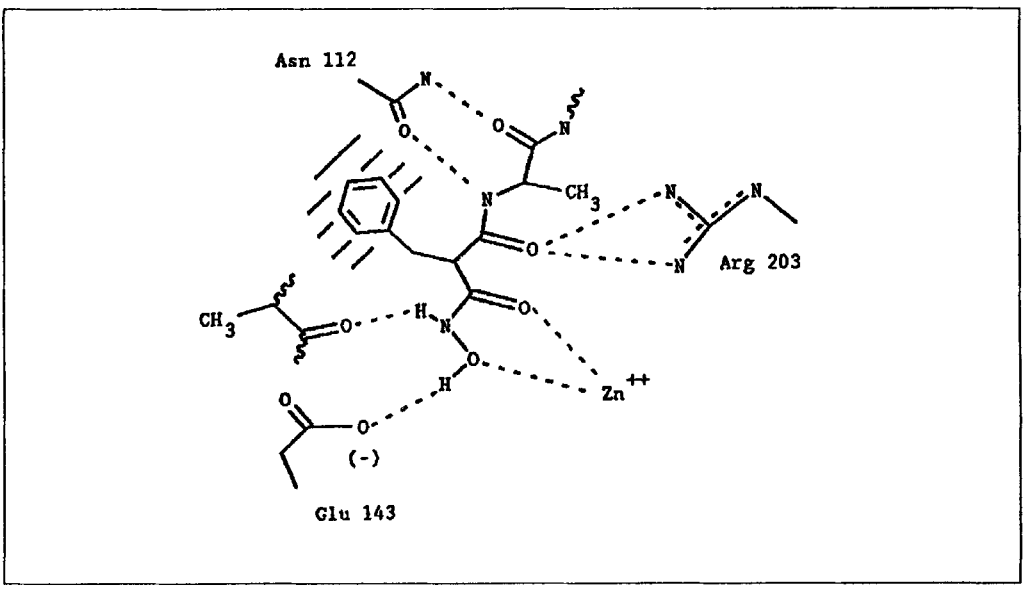


Figure 6. A SCHEMATIC VIEW OF THE ACTIVE SITE. THE ACTIVE SITE OF THERMOLYSIN WITH A FRAGMENT OF A HYDROXAMIC ACID DERIVATIVE INHIBITOR.

Conclusions

We have presented a system named FLEMING that discovers explanatory theories from a set of experimental observations concerning drug compounds. Dynamic bias management has been presented as a critical mechanism to deal with complex structured objects such as molecules. In this case, the use of low-level descriptors appears inadequate to capture the target concept while the adoption of high-level predicates introduces a too strong bias on the results of the analysis. Because of the close integration between bias management system and learning system, we could not easily apply standard induction algorithms, notably the AQ family (Michalski, 1983). FLEMING has been successfully applied to the domain of the inhibitors of the thermolysin enzyme, where it automatically generated an explanatory model containing structural relations between generalized fragments described in terms of binding properties. In the next future we plan to test the FLEMING system on a number of problems currently under study in pharmaceutical research laboratories.

References

[1] Bolis, G., Di Pace, L., Fabrocini, F., **A Machine Learning Approach to Computer Aided Molecular Design,** *in International Journal of Computer Aided Molecular Design,* (forthcoming).

[2] Drastal, G., Czako, G., Raatz, S., **Induction in an Abstraction Space: A Form of Constructive Induction,** *in Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI,* 1989.

[3] Flann, N., Dietterich, T. G., **Selecting Appropriate Representations for Learning from Examples,** *in Proceedings of the Fifth National Conference on Artificial Intelligence, AAAI,* 1986.

[4] Holmes, M. A., Matthews, B. W., *in Biochemistry, 20,* 1981.

[5] Klopman, G. J., Bendale, R. D., **Computer Automated Structure Evaluation (CASE): A Study of Inhibitors of the Thermolysin Enzyme,** *in J. Theor. Biol., 136,* 1989.

[6] Martin, Y. C., **Quantitative Drug Design,** *New York, Marcel Dekker,* 1978.

[7] Matheus, C. J., Rendell, L. A., **Constructive Induction on Decision Trees,** *in Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI,* 1989.

[8] Michalski, R. S., **A Theory and Methodology of Inductive Learning,** *in Michalski, R. S., Carbonell, J. G., Mitchell, T. M., eds., Machine Learning: An Artificial Intelligence Approach, vol. I, Palo Alto, Tioga Publishing Company,* 1983.

[9] Mitchell, T. M., **The Need for Biases in Learning Generalizations,** *Tech. Rep. CBM-TR-117, Computer Sc. Dep., Rutgers University,* 1980.

[10] Muggleton, S., **Duce, an oracle based approach to constructive induction,** *in Proceedings of the Tenth International Joint Conference on Artificial Intelligence, IJCAI,* 1987.

[11] Pagallo, G., **Learning DNF by Decision Trees,** *in Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI,* 1987.

[12] Rendell, L., Cho, H., **Empirical Learning as a Function of Concept Character,** *in Machine Learning, 5,* 1990.

[13] Rendell, L., Seshu. R., Tcheng, D., **Layered Concept-Learning and Dynamically-Variable Bias Management,** *in Proceedings of the Tenth International Joint Conference on Artificial Intelligence, IJCAI,* 1987.

[14] Russell, S. J., Grosof, N. B., **Delarative Bias: An Overview,** *in Benjamin, P., ed., Proceedings of the Philips Workshop on Reformulation and Inductive Bias, Boston, Kluwer Academic,* 1989.

[15] Schlimmer, J. C., **Incremental Adjustment of Representations in Learning,** *in Proceedings of the International Workshop on Machine Learning,* 1987.

[16] Utgoff, P. E., **Shift of Bias for Inductive Concept Learning,** *in Michalski, R. S., Carbonell, J. G., Mitchell, T. M., eds., Machine Learning: An Artificial Intelligence Approach, Los Altos, Morgan Kaufmann Publishers,* 1986.

[17] Wrobel, S., **Automatic Representation Adjustment in an Observational Discovery System,** *in Proceedings of the Third European Working Session on Learning,* 1988.