APPLICATION OF CASE-BASED REASONING TO SAFETY EVALUATION OF *PROCESS CONFIGURATION*

A.-M. Heikkilä, T. Koiranen and M. Hurme
Helsinki University of Technology, Department of Chemical Technology,
P.O. Box 6100, FIN-02015 HUT, Finland

An index based method for estimating the effect of process configuration on inherent safety has been presented. Process configuration means which operations are involved in the process and how they are connected together. An inherently safe process structure is not possible to define by explicit rules, but one has to rely on standards, recommendations and accident reports. When problem solving is based on experience it is possible to use case-based reasoning. Therefore CBR was employed for determining the value of safe process structure subindex. For this purpose a casebase of accident cases and recommended designs was created. A case can be retrieved on five levels of aggregation (from process to detail) from the casebase to ensure the relevancy of retrieved information in various phases of preliminary process design.

Keywords: inherent safety, case-based reasoning, process configuration

INTRODUCTION

The most important process design decisions are made during the conceptual design phase when the process route is selected. Also the process safety - especially the inherent safety - is determined by the early design decisions. The essence of the inherent safety is to avoid hazards by proper design rather than to control them by added-on protective systems (Kletz, 1991). Thus inherent safety is related to the selection of chemicals, process conditions and operations which are used. Based on these factors Heikkilä et al. (1996) have developed a methodology and a safety index which allows safety comparison of process alternatives to be done in the conceptual design phase.

In the Inherent Safety Index the subindex of safe process structure describes the inherent safety of different process configurations. In this subindex the designed process structures are compared with known safe and unsafe process solutions. Experience based data from different process solutions can be found in safety standards, accident reports and design recommendations. While experience based data has no explicit rules, the safe process structure need to be evaluated by case-based reasoning. The advantage of case-based reasoning is its ability to present a more explicit representation of knowledge compared to rule-based reasoning, since it is based on detailed case histories rather than on their interpretation and recollection by an expert. Thus no heuristic rules which are based on generalizations are needed since case-based reasoning relies on analogies. This same approach is used mentally by practicing engineers to generate new process designs.

Extensive databases have been collected from accident reports (Anon, 1996). Also much data has been published as design recommendations of different process systems (Lees, 1996). From this data a database of good and bad designs can be collected. By using this database case-based reasoning can check, if a new design resembles known safe or unsafe design cases.

INHERENT SAFETY INDEX

In process synthesis an interactive rule-based system can be used for generating process alternatives (Hurme and Järveläinen, 1995). The comparison of generated alternatives is based on economics, safety and environmental considerations. For the safety comparison the Inherent Safety Index (ISI) is used (Heikkilä et al., 1996). ISI is formed of several subindices, which describe reactivity, flammability and toxicity of chemicals, inventory, process conditions, type of equipment and process structure.

Total Inherent Safety Index (Eq.1) is calculated for each process step separately as a sum of Chemical Inherent Safety Index (Eq.2) and Process Inherent Safety Index (Eq.3). More details of the method have been given by Heikkilä et al. (1996).

$$I_{TI} = I_{CI} + I_{PI} \qquad (1)$$

$$I_{CI} = I_{RM,max} + I_{RS,max} + (I_{FL} + I_{EX} + I_{TOX})_{max} + I_{COR,max} + I_{INT,max} \qquad (2)$$

$$I_{PI} = I_{I} + I_{T,max} + I_{p,max} + I_{EQ,max} + I_{ST,max} \qquad (3)$$

The subindices for the heat of main and side reactions, I_{RM} and I_{RS}, are related to the maximum heat release from the process. Flammability (I_{FL}), explosiveness (I_{EX}), toxicity (I_{TOX}) and corrosiveness (I_{COR}) describe the hazardous properties of the chemical substances. Subindices I_{FL}, I_{EX} and I_{TOX} are summed for each substance separately, and the maximum sum is used in the calculations. Chemical interaction (I_{INT}) describes the reactivity between the substances present in the process.

The subindex for inventory (I_I) is related to the amount of process materials present in the process. Process temperature (I_T) and pressure (I_p) reflect the maximum temperature and pressure in the process. Equipment safety (I_{EQ}) describes the safety of individual process items such as a pump or a reactor. The estimation of the Safe Process Structure Subindex (I_{ST}) is discussed in more detail in this paper.

SAFE PROCESS STRUCTURE SUBINDEX

The safe process structure means which operations are involved in the process and how they are connected together. Therefore the Safe Process Structure Subindex describes the safety of the process from systems engineering point of view. It describes: how well certain unit operations or other process items work together, how they should be connected and controlled together. The index describes also how auxiliary systems such as cooling, heating or relief systems should be configured and connected to the main process.

The importance of this subindex is increasing as the processes are becoming more integrated through heat and mass-transfer networks.

The Process Structure Subindex does not describe the safety of process items as such or their interaction through nonprocess route (i.e. through layout), since this is described by the Equipment Safety Subindex (Heikkilä and Hurme, 1998).

EVALUATION OF SAFE PROCESS STRUCTURE

Many different alternative process configurations can be created for a process in the conceptual design phase. In choosing the most feasible alternative safety should be one of the major evaluation criterias. Therefore information on the safety features of alternative process structures are needed on preliminary process design (Heikkilä and Hurme, 1998).

Most of the subindices of ISI are quite straightforward to estimate since they are e.g. based on the physical and chemical properties of compounds present. The process structure subindex looks at the process from a systems engineering point of view. Therefore it is much more difficult to estimate. In fact there is no explicit way of estimating the safety of the process structure but one has to rely on experience based data which is documented as standards, design recommendations and accident reports.

SOURCES OF EXPERIENCE BASED SAFETY INFORMATION

Process solutions have shown their strong and weak, safe and unsafe points in operation practice. The knowledge of practising solutions consists of the details collected during the operation and maintenance. Practising solutions reveal for instance which unit operations are preferable for certain purposes and how the units can be connected safely together. Some of the information can also be found from design standards which have been created on the basis of the experience on the operation of existing process plants (Lees, 1996).

Another source of design information is the accident reports made after an accident. They give valuable information of the possible weaknesses that can occur in unit operations, while they are used for certain purposes. In the past many of the unit operations have shown their adverse characteristics. This information is mainly collected to accident reports and included to safety standards. Accident reports tell us for example:

* which process equipment configurations have unfavourable properties
* which type of chemicals do not suit to certain unit operations
* which unit operations/ configurations are risky
* when the connection of process units should be avoided.

The difficulty in utilizing accident reports lies in the lack of accident report standards. Reports vary a lot how they document the details of the accident itself, the path to the final event, the causes, and the consequences. Still the reports can tell much experience based information which can - and should be - utilized in designing new plants. In fact a major goal in improving the design of safe process plants should be to enhance the reuse of design experience. This is important since the same mistakes are done again and again (Kletz, 1991).

A more refined form of accident reports is an accident database, where all the reports are presented in a standardized format. Extensive databanks have already been collected from accident reports (Anon, 1996). This kind of standardized format allows easier retrieval of accident information also by computerized means.

STRUCTURE OF THE DATABASE

The basis for the estimation of safe process structure lies in the integration of the two types information sources: 1) recommendations and standards how the process should be designed, and 2) accident database which describes the negative cases from which one can learn. Therefore a casebase of good and bad design cases is needed. Both of these information sources should be readily available to the design engineer through the database. A design problem can be compared with the cases in this combined databank for instance by case-based reasoning.

In this approach accident cases and design recommendations are analysed level by level. In the database the knowledge of known processes is divided into categories of process, system, subsystem, equipment and detail (Fig. 1). Process is an independent processing unit (e.g. hydrogenation unit). System is an independent part of a process such as reactor or separation section. Subsystem is a functional part of a system such as a reactor heat recovery system or a column overhead system including their control systems. Equipment is an unit operation or an unit process such as a heat exchanger, a reactor or a distillation column. Detail is an item in a pipe or a piece of equipment (e.g. a tray in a column, a control valve in a pipe).

A search for cases in the databank can be made on these levels on the basis of the nature of the design problem. If a process is being designed from beginning the first search is made for a whole process. The search is then made for those systems, subsystems and equipment, which are informable for the design. On the basis of the retrieved information the designer can evaluate the right index value for the process structure of the section under review. The input data for a database search contains information on the process level and on the raw materials and products, reaction types and their details such as catalysts and phase of reaction. As output there is information about the unfavourable process configurations, recommended configurations and accident cases.

A plant is divided into inside and offsite battery limit areas. The configurations of ISBL and OSBL areas differ considerably. Generally the size of equipment, the amount of chemicals and also the spacings are larger in OSBL area. The safety of the process structure is also affected by these factors. Therefore this aspect is included also into the database.

The database does not always contain information which is directly related to the process under review. Therefore it is important to be able to use analogies. In general much of the design of new processes relies on analogies. For example most hydrogenation processes have similar features, most tanks of liquefied gases have similarities etc. For that reason information has been included into the database on the type of materials in incident (e.g. liquefied gas), the type of the reaction (e.g. oxidation), the thermal nature of the reactor (e.g. exothermic), the phase of the reaction and the type of catalyst.

To learn from the accident cases it is essential to indicate the type of incident which happened (e.g. explosion), the direct cause of the incident (e.g. static electricity), the reason why this could take place (e.g. filling through the gas phase) and finally - most important - the lesson how this can be avoided (e.g. fill the tank through the bottom).

CASE-BASED REASONING

When problem solving is based on experience which is difficult to define as explicit rules, it is possible to apply case-based reasoning (CBR). CBR uses directly solutions of old problems to solve new problems. The functional steps in CBR are (Gonzalez and Dankel, 1993):

1. New problem presentation.
2. Retrieval of the most similar cases from case-base.
3. Adaptation of the most similar solutions for generating a solution for a current problem.
4. Validation of the current solution.
5. Learning from the problem cases by adding the verified solution into the case-base.

A data table of a case-base can be divided into input and output sections. Input parameters are retrieval parameters and output parameters are design specification parameters. The problem is characterized as input data to the system. In the retrieval phase a set of retrieval parameter values of all cases in the case-base are compared to the input data. The most similar cases are then selected and ranked based on the comparison.

In the case of string data types suitability is simply:

$$X_i = C_{i,j} \Rightarrow Y_{i,j} = 1 \qquad (4)$$
$$X_i \neq C_{i,j} \Rightarrow Y_{i,j} = 0 \qquad (5)$$

where X_i is the input value of parameter i, $C_{i,j}$ is the value of parameter i of case j, and $Y_{i,j}$ is the suitability of a parameter i for the case j.

The quality of reasoning increases, if the importance of selection parameters can be altered. The user should determine the importances of selection parameters for the topic under study. Weighted suitability R_{ij} can be expressed:

$$R_{ij} = W_i Y_{i,j} \qquad (6)$$

where W_i is weight factor of a selection parameter i evaluted by user. Overall suitability can be calculated for the case j based on the number of parameters N and parametric suitabilities R_{ij}:

$$S_j = \frac{\sum_{i=1}^{N} R_{ij}}{N} \qquad (7)$$

Case-based reasoning has earlier been used for instance for equipment design. Koiranen and Hurme (1997) have used case-based reasoning for fluid mixer design and for the selection of shell-and-tube heat exchangers. They have included an estimation of design quality for the case retrieval beside technical factors.

Chung and Jefferson (1997) have combined the IChemE Accident Database (Smith et al., 1997) with case-based reasoning to create an automatic data retrieval for designers' and operators' use. They intend to develop an intelligent system, which takes for example the term 'electrical equipment failure', works out all the related terms and retrieves the relevant information automatically. The method should be integrated with computer tools used by designers, operators and maintenance engineers so that appropriate accident reports can be automatically presented to the user. The employed IChemE database contains much information on accident causes. The aim of the system presented by Chung and Jefferson (1997) is to find all relevant causes of past accidents to improve processes, whereas our CBR system is intended for reasoning on the structure of a process and its favourable and unfavourable characteristics for preliminary process design purposes. The database used by Chung and Jefferson (1997) is an accident database, whereas our database contains also design recommendations. On the other hand our CBR system is intended specifically for the use of process designers, but the system of Chung and Jefferson (1997) is developed for wider use from chemical plant designers and operators to maintenance teams.

DESCRIPTION OF PROTOTYPE APPLICATION

Prototype CBR application has been implemented on MS -Excel spreadsheet. The program has been organized on several sheets. A database of cases was created which consists of accident cases collected from literature (e.g. Lees (1996) and Loss Prevention Bulletin) and of design recommendations. The application program includes retrieval functions which are used to retrieve the most suitable cases from the database.

Input and output parameters

The scope of a database search is defined by using categories of process, system, subsystem, equipment and detail as input parameters. This hierarchy is used for clarifying the process structure and for making the use of process analogies more feasible in reasoning. E.g. a condenser has certain safety characteristics undependent on the process it is located. Beside the process structures input parameters include the raw materials and products and some reaction details. The importance of the parameters may be evaluated by using weighting factors.

Output parameters contain the input parameters plus information on the safety characteristics of the process and information on accidents and their causes. Specific design recommendations are included in the output. On the accidents the output describes e.g. following information:

* what kind of incidents have happened
* what is the actual cause of the incident
* what are the contributing factors or circumstances of the incident
* how to improve the application for better safety

All stored cases are validated on the basis of the Safe Process Structure Subindex. The validations are given for every case and included in the output. Further information on the cases is given as appendices, which describe the case in more detailed.

Retrieval of cases

In this work the cases in the database are stored on their own MS-Excel worksheets. The stored cases are copied on a retrieval calculation data sheet during the retrieval phase. All retrieval parameters in this application are textual string parameters. Thus the comparison between casebase and input problem is simple. When the input value is equal with the case value, the distance is 1, otherwise the distance is 0. The weighted suitability of parameters is then calculated by Equation 6. The weighting factors are introduced by the user. Overall suitability is calculated by Equation 7. Cases are ranked according to their overall suitability and the five nearest cases are shown for the user on an output worksheet.

The retrieval of cases can be done in several steps. The first step is the evaluation of the process with the stored cases. This way can be seen, if the process is safer or unsafer than the alternative processes. The second step is the safety evaluation of specific process systems, subsystems or pieces of equipment. The database contains improvement recommendations to avoid the same accidents happening again. The evaluation of processes can be extended to detailed level. Also the equipment details or safety valves etc. can be checked on this level.

INHERENT SAFETY INDEX OF SAFE PROCESS STRUCTURE

All included processes and their subprocesses in the database are evaluated according to the Inherent Safety Index. Process structures are divided into six groups of scores from 0 to 5 according to the knowledge of their safety behaviour in operation.

The first group is the safest group with the score 0. It consists of recommended and standardized process and equipment solutions. The second group is based on sound engineering practice, which implies the use of well known and reliable process alternatives. In the third group there are processes which look neutral, or on which there is no safety data available. The fourth group includes configurations which are probably questionable on the basis of safety even accidents have not occured yet. The fifth and sixth groups contain process cases on which documented minor or major accident cases exist.

Table 1. Values of the Safe Process Structure Subindex I_{ST}

Safety level of process structure	**Score of I_{ST}**
Recommended (safety etc. standard)	0
Sound engineering practice	1
No data or neutral	2
Probably unsafe	3
Minor accidents	4
Major accidents	5

CASE STUDY

As a case study an acetic acid process is discussed. Acetic acid is produced by the liquid-phase methanol carbonylation. The reaction is carried out at 175 degrees of Celsius and 30 bar pressure. The process diagram is shown in Figure 2.

For the safety evaluation CBR database searches were done on two levels. First level was the acetic acid process as a whole. On the second level the reactor system was studied in more detail.

Reasoning on the acetic acid process alternatives

First the acetic acid process was studied as a whole to find out if the alternative processes have differences in the safety on the conceptual (i.e. process) level. The search (Table 2) found cases for carbonylation and oxidation processes (Table 3). It can be seen that there has been explosions and fires on both types of plants. The explosion in the carbonylation plant was due to static electricity in loading of a storage vessel. This type of explosions are not specific to carbonylation plants, but

Table 2. Input data of the search for the acetic acid process

INPUT DATA

Retrieval parameters	Active	Importance	Value
raw material	TRUE	9	methanol
product	TRUE	9	acetic acid
reaction type	FALSE		
termic type of reaction	FALSE		
phase of reaction	FALSE		
catalyst	FALSE		
ISBL / OSBL	TRUE	6	isbl
system	FALSE		
subsystem	FALSE		
equipment	FALSE		
detail	FALSE		

they are possible also in many other processes. The fires and explosions on the oxidation plants were related to the chemicals present in that process. They are more likely to happen in such a plant than somewhere else. Thus the carbonylation process can be considered safer than the oxidation process based on the information from this search.

Reasoning on the reactor heat transfer system

In the second phase searches were made on the system and subsystem level. This is needed for the design of the reactor and its heat transfer systems. Carbonylation of methanol is an exothermic reaction. Thus only the exothermal reactors were searched. The CBR search found two cases which are recommendations on the design of exothermic reactors with heat transfer systems. They are shown in Figures 3 and 4.

The case in Figure 3 represents a reactor with two different cooling systems. In the not recommended case (right) the cooling system presents a feedback loop between a reactor heat rise and the rise in the coolant temperature, which should be avoided. On the left is the recommended system, where the coolant temperature does not depend on the reactor temperature.

The case in Figure 4 shows a heat recovery system of a reactor. The not recommended case on the left shows the feed to an exothermic reactor being heated by the product. In this case the temperature rise in the reactor may lead to the temperature rise in feed. The recommended case on right is safer since the connection is broken because the heat transfer is done by generating and using medium pressure steam.

Table 3. Output data of the search for the acetic acid process

	1st Case	2st Case	3st Case
PROCESS:			
Raw material	methanol	butane	butane
Product	acetic acid	acetic acid	acetic acid
Reaction type	carbonylation	oxidation	oxidation
Thermic type of reaction	exo	exo	
Phase of reaction	liquid	liquid	liquid
Catalyst	Rh complex		
Isbl / Osbl	isbl	isbl	isbl
SYSTEM		reaction	reaction
SUBSYSTEM	intermediate storage	purging	feed
EQUIPMENT	tank	reactor	boiler
DETAIL	inlet pipe		
Incident:	explosion	fire	explosion
Cause 1:	static electricity	self-ignition of acetaldehyde	oxygen leak
Cause 2:	filling through vapor phase	methane ignited	
Recommendations:	fill through bottom		
Material:	acetic acid	acetaldehyde	butane/air
Nature of material:	organic acid	aldehyde	LPG
Safety Index (0-5)	4	4	5
Appendix:			App.1

Appendix 1: Explosions occured because pure oxygen entered a gas-fired boiler and mixed with the butane and steam used to form acetic acid. The first blast occured near a gas fired boiler and the second blast occured at a nearby reactor. (3 killed, 37 injured)

Score of the Safe Process Structure Subindex

From the reasoning on the process level we get score 2 (no data or neutral) for the carbonylation process, since the found case was not specific to this process. For oxidation process we get score 5, since a major accident has taken place.

For the recommended reactor system we can get scores 0 (recommended/standard) or 1 (sound engineering practice) depending how we value these recommendations.

The final score of the Safe Process Structure Subindex for the carbonylation process would be 2 based on this limited reasoning, since the final score of I_{ST} is chosen on the basis of the worst case. Of course in practice one should do the reasoning on all the systems and subsystems in the process.

This case study was given only to represent the principle of CBR in reasoning the value of the Safe Process Structure Subindex.

CONCLUSIONS

In this paper an index based method for estimating the effect of process configuration on inherent safety has been presented. Process configuration means which operations are involved in the process and how they are connected together. The Safe Process Structure Subindex describes the safety of the process configuration from system engineering point of view. Importance of this aspect is becoming more important since the processes are becoming more and more integrated.

Since an inherently safe process structure is not possible to define by explicit rules, information which is based on cases and engineering experience has to be used. This type of knowledge is presented as standards, engineering recommendations and accident reports.

When problem solving is based on information which is difficult to define as rules, it is possible to use case-based reasoning. The advantage of case-based reasoning is its ability to present a more explicit representation of knowledge compared to rule-based reasoning, since it is based on detailed case histories rather than their interpretation and recollection by an expert. Thus no generalizations are needed since CBR relies on analogies. This same approach is used mentally by practicing process engineers to generate new process designs.

For the estimation of the Safe Process Structure Subindex a casebase of good and bad cases was created from recommendations, standards and accident reports. The cases can be retrieved on five levels of aggregation (process, system, subsystem, equipment, detail) to ensure the relevancy of retrieved information for various phases of process design. The cases are valued by using the Safe Process Structure Subindex. The subindex has six values (0-5) representing recommended design, sound engineering practice, no data or neutral, probably unsafe, minor and major accident cases. The final score of the subindex is chosen on the basis of the worst case of different levels of the reasoning. The results can be used with other subindices for estimating the total inherent safety of process alternatives for the selection of process concept or details of the process configuration.

The results of the database search can be presented as reports which highlight possible danger points of the process. These reports should follow the process alternatives till the end of process design and even till the operation stage of the process.

REFERENCES

Anon. (1996), Accident Database Expands, *The Chemical Engineer*, 25 April, No 610, 5.

Chung, P.W.H. and Jefferson, M., Accident Databases - Indexing and Retrieval, *EC/EPSC seminar on "Lessons Learnt from Accidents"*, Linz, 16-17 October 1997.

Gonzalez, A.J. and Dankel, D.D., (1993), *The Engineering of Knowledge-based Systems: Theory and Practice*, Prentice-Hall International, Inc., New Jersey.

Heikkilä, A.-M., Hurme, M., Järveläinen, M., (1996), Safety Considerations in Process Synthesis, *Computers chem. Engng*, **20** Suppl. S115-S120.

Hurme, M. and Järveläinen, M., (1995), Combined Process Synthesis and Simulation System for Feasibility Studies, *Computers Chem. Engng*, **19**, Suppl., S663-S668.

Kletz, T. (1991), *Plant Design for Safety*, Hemisphere, New York, 11.

Koiranen, T. and Hurme, M., (1997), Case-Based Reasoning Application in Process Equipment Selection and Design, *SCAI'97 Sixth Scandinavian Conference on Artificial Intelligence*, G. Grahne (Ed.), IOS Press, Amsterdam, 273-274.

Lees, F.P., (1996), *Loss Prevention in the Process Industries: Hazard Identification, Assessment and Control*, Butterworth-Heinemann, Oxford, UK.

Smith, M.K., Bond, J. and Mellin, B., (1997), *The Accident Database: capturing corporate memory*, SHE Department, IChemE, Rugby.

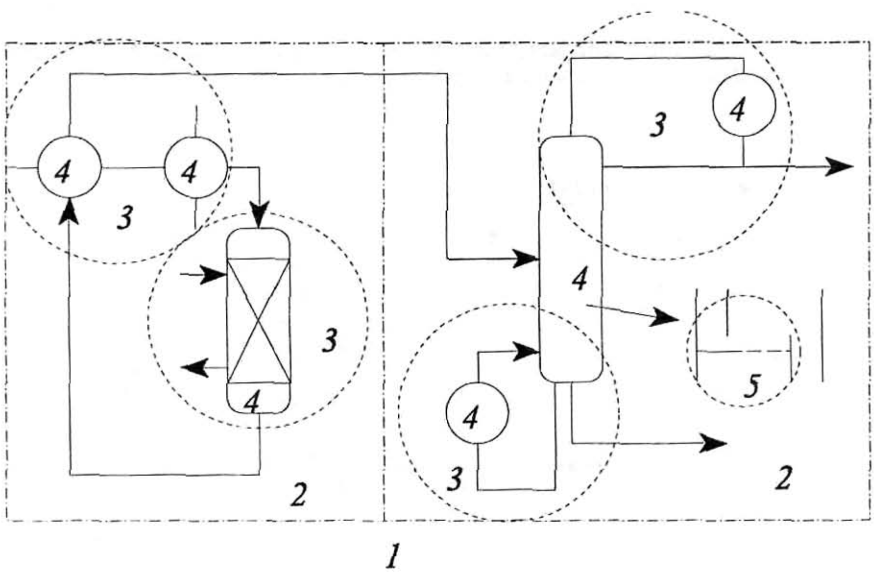

Figure 1. Example of the levels of the process as used in the CBR database. (1 = process, 2 = system, 3 = subsystem, 4 = equipment, 5 = detail)

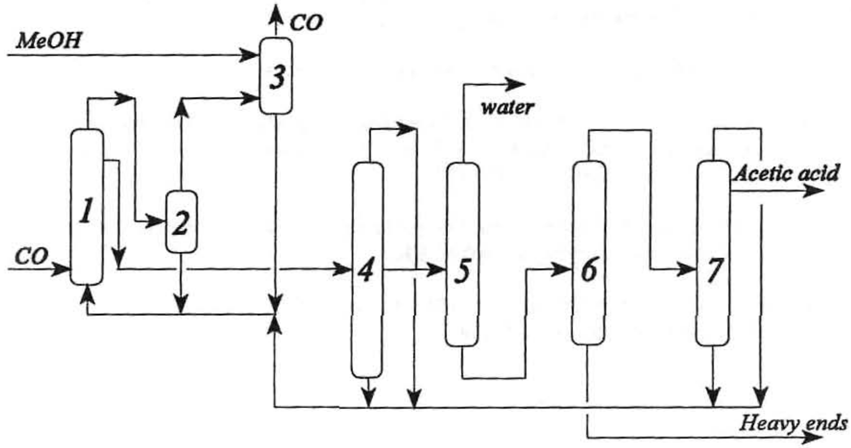

Figure 2. Flowsheet of the acetic acid process: 1) reactor, 2) separator, 3) scrubber, 4) light ends separator, 5) drying column, 6) product recovery, 7) product finishing

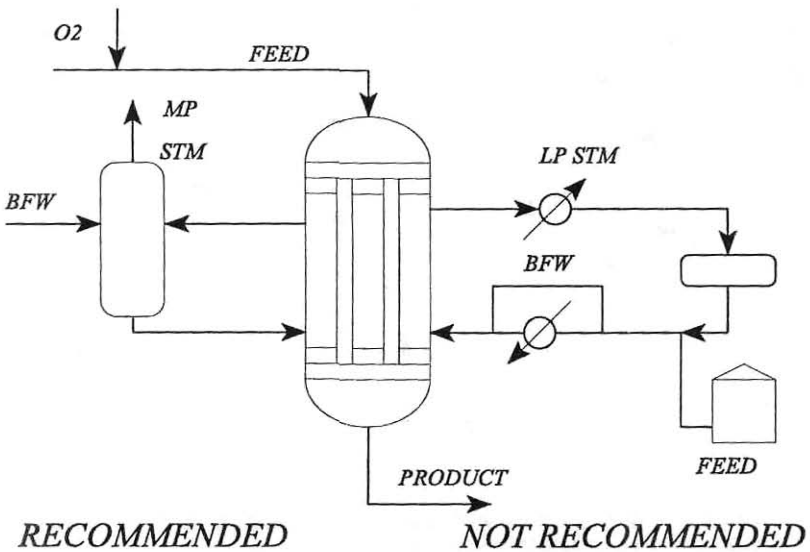

Figure 3. A recommendation to avoid the feedback loop between a reactor heat rise and a rise in coolant temperature.

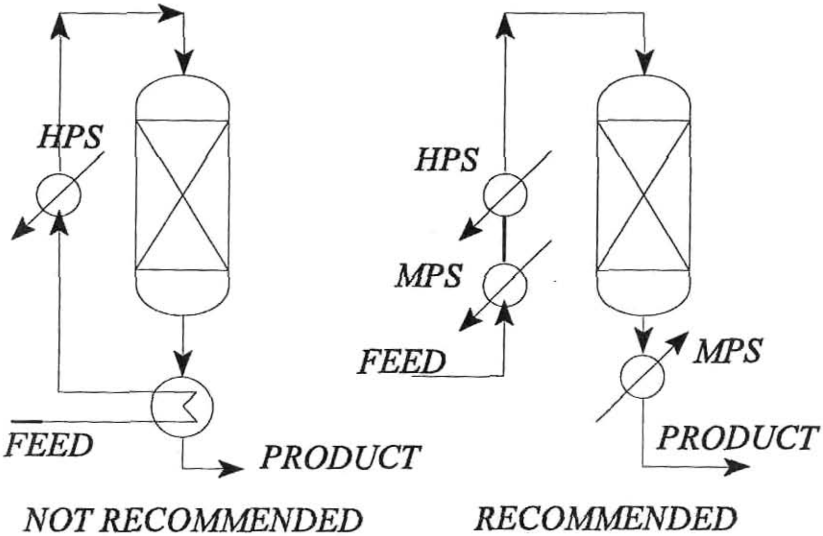

Figure 4. A recommendation for preheating the feed of an exothermic reactor.