# KRF

**Knowledge Representation Framework Project**
Department of Computer and Information Science, Linköping University,
and Unit for Scientific Information and Learning, KTH, Stockholm

Erik Sandewall

# The Goals of Artificial Intelligence Research

## A Brief Introduction

This series contains technical reports and tutorial texts from the project on the Knowledge Representation Framework (KRF).
The present report, PM-krf-014, can persistently be accessed as follows:

Project Memo URL: www.ida.liu.se/ext/caisor/pm-archive/krf/014/
AIP (Article Index Page): http://aip.name/se/Sandewall.Erik.-/2010/004/
Date of manuscript: 2010-08-08


Related information can also be obtained through the following www sites:

KRFwebsite: http://www.ida.liu.se/ext/krf/
AIP naming scheme: http://aip.name/info/
The author: http://www.ida.liu.se/~erisa/

# Chapter 1

# Children Scenarios that Exemplify Intelligence

## 1.1 On Artificial Intelligence

The goal of research in Artificial Intelligence is to develop a technology whereby computers behave intelligently, for example:

- Intelligent behavior by simulated persons or other actors in a computer game
- Mobile robots in hostile or dangerous environments that behave intelligently
- Intelligent responses to user requests in information search systems such as Google or Wikipedia
- Process control systems that supervise complex processes and that are able to respond intelligently to a broad variety of malfunctioning and other unusual situations
- Intelligent execution of user requests in the operating system and the basic services (document handling, electronic mail) of standard computers.

The words "intelligent" and "intelligence" are then used in the same sense as they have in ordinary language. Several other scientific disciplines, such as psychology and philosophy, have developed fairly precise notions as well as considerable internal disagreement about what intelligence "really is" or what is the appropriate meaning of the term. This has influenced artificial intelligence research a bit, but not very much. Similarly, when an aeroplane designer says that a particular aircraft "flies," he or she realizes of course that the performance of the aircraft differs in fundamental ways from the flying of a bird or a bee, and he or she is not concerned about what is the "true meaning" of the concept of flying.

It is important to keep in mind that intelligence is a graded concept: a person may be more or less intelligent, and it is not possible to divide the world's population into one group of people that are intelligent and others that are not. Just like "hot" and "cold" describe the two ends of a temperature scale, similarly "intelligent" and "stupid" describe the two ends of the intelligence scale. The behavior of conventional computer programs is manifestly at the stupidity end of that scale, and the goal of artificial intelligence is to move software away from pure stupidity and in the direction of being more intelligent. This is a gradual process and one should not think of intelligence as a magical spirit or mystical power that can be placed in computers once and for all.

The present chapter will characterize the kind of intelligence that artificial intelligence research is concerned with by describing three scenarios involving a small child, referred to as Ronnie, who exhibits a certain degree of intelligence. These are not invented stories, they are actual scenarios that I have observed in a small child (except for the modified freezer door scenario), but they have been selected since they are good examples of the type of capabilities that artificial intelligence research addresses. The next chapter will proceed to show how these capabilities that one can observe when intelligence emerges in young children, are also relevant and needed when a certain level of intelligence is embedded in computer software.

## 1.2 The Freezer Door Scenario

Ronnie and her older brothers like icecream, and their parents allow them to go and get icecream bars by themselves in the kitchen freezer but only after having received permission for this on each particular occasion. This time the brothers are away and Ronnie is allowed to pick up an icecream bar herself. The freezer door is too tall for her to reach the handle at the top of the door, so she runs to the nearest chair in the kitchen, pushes it to a position in front of the refrigerator that is next to the freezer, climbs onto the chair, and opens the freezer door. This is her own idea at this occasion, since usually it is one of the brothers that opens the freezer door without needing a chair and that is able to get icecream bars for all the children.

This scenario involves several aspects of intelligence. First of all, Ronnie understood that she needed an *instrument* that made it possible for her to open the freezer door. She was able to select an appropriate instrument and to operate it for the intended purpose.

Moreover, Ronnie made a smart move by placing the chair in front of the refrigerator door, and not in front of the freezer door which was after all at the center of her attention. If she had put it in front of the freezer door then it would have blocked that door so that she could not open it. Similarly, if she had put it to the right of the freezer door, instead of to the left, then she would maybe have been able to open that door, but she would not have been able to reach the icecream bars.

How did Ronnie have this insight, and how could a computer program (in a computer game, or in a household robot, for example) have the same insight? Given that this was a new situation that Ronnie had not encountered before, the most plausible explanation is that she was able to *simulate* her successive actions in her own mind before starting out, or maybe while she was already pushing the chair, and as a way of anticipating the effects of her continued actions. If she first imagined pushing the chair to the position in front of the freezer door, the simulation would show how she climbs the chair and tries to open the freezer door but fails, and then she is able to revise the plan before it has taken effect.

In artificial intelligence terms, Ronnie has showed the capability of *planning, anticipating* the effects of a proposed plan, and *plan revision* where her first plan was modified in order to pass the anticipation test, and of course also *plan execution* where the selected plan is performed.

## 1.3 The Revised Freezer Door Scenario

Now let us modify the scenario so that it applies to Ronnie's friend, Valerie, who is a little younger and not yet as smart. Valerie actually pushes the chair to the place in front of the freezer door. She climbs the chair, tries to open the freezer door but fails, climbs off the chair, moves it to the position in front of the refrigerator door instead, climbs it again, opens the freezer door, and gets the icecream bar.

In this case Valerie had not been able to anticipate the full effects of her original plan so that the execution of that plan failed. This called for another important aspect of intelligent behavior, as understood in artificial intelligence, namely for *diagnosis* which means finding out the reason for why something has gone wrong: a plan does not work, or an available device does not work properly. Diagnosis leads to making a new plan, where typically some of the effects of the original plan are undone and a new plan is made.

Notice an important difference between this scenario and the first one: In the original scenario Ronnie may first have thought of moving the chair to in front of the freezer door, and then changed to a better plan, but the revised plan *did not* consist of first moving the chair to the freezer door, climbing it, unclimbing it, and moving it to the refrigerator door. In other words, the plan did not replicate the planning process; it was a better plan that took the obstacles into account.

Now suppose Valerie had performed once according to the revised scenario, and the next day the situation is repeated: Valerie is allowed to pick up an icecream bar, and she has to do it herself. She is no smarter than the day before, so if her anticipation capability was not sufficient yesterday, there is no reason that it should be so today. However, this time Valerie makes the right move and takes the chair to the refrigerator door. Why? - because she has *learnt* from the experience of yesterday. The scenario involving the plan failure and plan revision was retained in her memory, and she was able to use it for doing the right thing. However, she certainly did not repeat yesterday's sequence of action, so she did more than just repeating a memorized sequence of actions: she *revised* yesterday's plan so as to eliminate the false start, obtaining the same plan as was obtained by Ronnie in the original scenario.

Moreover, it turned out that the second day all the chairs were out of the kitchen, so Valerie was not able to apply the revised plan. There were two possibilities: find something else to climb on, or get a chair from the next room. Valerie used the second possibility and got her icecream bar. Plan revision is therefore used not merely for simplifying away false starts from a plan in plan memory; it is also used for modifying a plan to fit the current situation.

The technique of maintaining a "knowledgebase" of problems that one has encountered before as well as their solutions, and to adapt them and apply them to forthcoming situations is an important one in artificial intelligence, and it goes by the name of *case-based* techniques.

## 1.4 The Doll Problem Scenario

This day Valerie is visiting Ronnie and the two girls are playing. They are not playing together; each of them is doing her own thing. Ronnie is putting her two dolls to bed, which involves first putting diapers on them. She is busy with one doll (Ada) when Valerie comes into the room, a comb in her hand since she has just been combing the hair of one or her dolls. Valerie notices that Beda also needs to have her hair combed, and picks her up. Father observes the situation but decides not to intervene, but to watch the scene to see what will happen.

Ronnie also notices the situation of course, frowns, and goes to the sofa and picks up Valerie's beloved little handbag, and brings it to Valerie. Valerie receives the handbag, lets go of Beda, and Ronnie proceeds to putting diapers on Beda as well.

This scenario is very similar to the freezer door scenarios in some ways, and very dissimilar in other ways. A major difference is that here we are dealing with social actions and their effects on the state of mind of persons, whereas the previous scenarios only concerned physical things: moving chairs, climbing onto them, opening freezer doors. However, in spite of these differences there are also an important similarity: both scenarios are concerned with *actions and their effects,* and with *making a plan* that is expected to achieve effects that are desired by the person in question.

The making and the execution of the plan can be thought of as *reasoning:* one thinks about different possibilities, different combinations of actions, and all aspects of the reasoning can be expressed clearly and it can even be represented as formulas in formal logic. However, this reasoning aspect is accompanied by other aspects that do not have the same rational character, and which may be considered as more or less intuitive when performed by a person. The freezer door scenario required a sense of positions and movements in the three-dimensional world of the kitchen; the doll problem scenario required Ronnie to have an understanding of Valerie's likely response to her plan. In fact, the success of the plan may well have depended on a positive understanding and a will to compromise on Valerie's side, and if Ronnie understood this then she might have responded differently if it had been one of her brothers that had taken the doll, rather than Valerie.

Is Ronnie's problem-solving competence in the doll problem scenario the *same capability* as her problem-solving competence in the freezer door scenario? Is it reasonable to refer to both of them using the same term, such as "intelligence" , or should one refer to one as "mechanical intelligence" and the other as "social intelligence," for example, and should one consider them as two different phenomena each with its own characteristics?

This question is not very important from the point of view of artificial intelligence, since both types of scenarios are included in the topic of interest of artificial intelligence as a field. The question whether it is essentially the same capability or essentially different capabilities will then be answered on technical grounds, by comparing the software solutions for different aspects of intelligence in order to see how much the differ or overlap.

If one should ask the same question for intelligence in people, then there is little chance of obtaining a definite answer. It is not possible to extract "the intelligence" from a human person in order to inspect it and analyze it, of course, so there is not direct way of saying what (human) intelligence *really is.* The use of the term intelligence is a linguistic convention, and one may advance various kinds of arguments in favor of saying that there is just one kind of intelligence, or there are several kinds of intelligence, but this requires debate and consensus; it can not be finally decided merely by experiments or other empirical observations.

## 1.5 The Bath Assistance Scenario

The third scenario takes place when Ronnie is a bit older and is able to speak, although this is only since a few months. Father is watching Ronnie as she plays with her favorite doll, Ada, and now she has taken off the doll's clothes. Guessing what she has in mind, he asks "are you going to give Ada a bath?" Ronnie answers yes; father says "are you going to take her to the bathroom then?" Ronnie answers yes again, and waits. Father waits. Ronnie says: "Come you also, I can not myself open the water." The reason for this is that although Ronnie can easily reach into the wash basin, the faucet handle is located in such a position that she can not reach it.

This scenario resembles the freezer door scenario since there is an instrumental aspect. Just as Ronnie realized that she needed a chair in order to open the freezer door, she now realizes that she needs father's help in order to "open the water." She knows that he is able to do that since she has seen him do it before, or since it is one of the things that adults in general are able to do according to her experience. She is also able to use the powerful medium of language in order to get him to move to the bathroom.

However, besides this instrumental side of the scenario, there is also a linguistic side: How did Ronnie formulate the phrase "Come you also, I can not myself open the water." It is fair to assume that there is a large portion of case-based reasoning in this case as well, since we learn language – children do, in particular – by listening to many examples of language, imitating them, and adapting them to new situations, for example by putting together sentences in new ways.

The phrase "open the water" can not be understood merely as a memorized phrase or a rearrangement, however, since it is not the normal way of expressing oneself in English. One must assume that Ronnie has a notion of the operation of "opening" that she has learnt and used in the context of opening doors, drawers and boxes, and that she can apply it by analogy to the action of allowing the stream of water to flow into the wash-basin.

Furthermore, notice that as Ronnie wanted father to come with her, she did not say exactly what she wanted; she did not say "come you also and open the water for me." Instead, she provided an explanation why father's assistance was needed, in order to explain why she could not perform the task herself. This was all she needed to say; she could assume that this explanation was sufficient for father to understand that his assistance was needed and that she wanted him to come along.

Like in the doll problem scenario, this interaction must be understood by an analysis of the actions and the reactions of both the people involved. When father hears "I can not open the water myself," and when he has understood what Ronnie means by "open the water," nominally he has just received a piece of information about the state of the world, and this information he probably knew already. If you think of linguistic utterances as means of transmitting information, questions and requests, then it was a redundant utterance. However, in the given context it was clear what was Ronnie's intention when saying this: besides being a piece of fact, it also served a purpose namely being the *motivation for* something. Understanding language is not merely understanding immediate content, it is also understanding the reasons for why phrases are being said, and those reasons are often the real meaning of the utterances.

This kind of analysis of dialogs is the subject of *speech act theory* which has aspects of both linguistics, psychology and philosophy. Speech act theory is also used in artificial intelligence, both for the design of dialog software and as the basis of techniques for message exchange between autonomous agents. The latter topic is covered in our compendium "Programming Techniques for Knowledge-based Autonomous Systems."

## 1.6 Can Intelligence be Implemented?

The examples in this chapter represent relatively simple examples of some important aspects of intelligence in the sense that the term is used in artificial intelligence: planning, anticipation, plan revision, diagnosis, case-based reasoning, learning, linguistic competence. Implementing capabilities such as these is complicated, especially when one proceeds to situations that are larger in the sense of containing many more persons and aspects. Practical implementations also tend to be further complicated by the requirements of the "non-intelligence" aspects of the entire system, such implementing an "understanding" of the movements of objects in a three-dimensional world, in such a way that it can be effectively connected to the core intelligence part of the system.

However, although it is complex, this task is by no means impossible. There is no need for magic in it; all the subtasks can be approached, and in particular since we think of intelligence as a graded concept with "stupid" at one end and "intelligent" at the other, there is not need to feel overwhelmed by the daunting task of creating human adult-level intelligence; we can proceed gradually. This is what research in artificial intelligence is about.

# Chapter 2

# Grownup Scenarios that Exemplify Intelligence

The examples of intelligence in the previous chapter can be exhibited by small children. What additional aspects of intelligence do we develop under the continued years until adulthood, and to what extent can they also be implemented in computer software?

A comprehensive answer to that question can not be given here, and we shall only give one example as a very partial answer.

## 2.1 The Party Preparation Scenario

Consider the task of preparing your livingroom so that it is ready when your guests arrive, the assumption being that you will start there with a drink and then go elsewhere for the meal. Two main activities are involved: the activity of preparing the livingroom, and the activity of entertaining the guests. The latter activity is only partly known, since surprises may come up (one guest may report an allergy that you were not aware of before, another guest may fall and hurt himself, and so on) but still you have sufficient knowledge about it that you can determine what needs to be done in advance. This involves bringing a number of things to the livingroom, such as beverages, glasses, napkins, probably a corkscrew and a bottle opener, and whatever you will offer to eat or to nibble. This knowledge is sufficient for making a plan for the preparation phase and for carrying it out. Contingencies may arise, not only in the entertaining phase but also in the preparation phase, but these have to be taken care of if and when they happen since you can not plan for everything.

There are a number of similar scenarios in ordinary life, such as putting things in order in the kitchen before cooking according to a complicated recipe, putting a conference room in order for a meeting, or preparing a surgery room for an operation.

A cognitive autonomous agent, such as Mia, provides support for the preparation scenario through its ability to simulate the execution of actions, and through its concurrent use of a model world and an outside world. The following is how this task can be approached. One must assume that the following information is available to the system at the outset, besides a description of the environment as a whole:

- A description of the performance stage (e.g., the entertaining activity in the initial example) that can be executed in the model world. This will be a nondeterministic action sequence or a set of typical, deterministic action sequences. It does not span the entire set of possible developments in the performance stage, but it can generate a number of typical examples.
- A specification of the place where the performance stage is to occur.
- A specification of the objects that must be available in that place before the beginning of the performance stage, and of any location constraints for these objects within the place of performance.
- A specification of the present location of objects that may be needed, or of actions that can produce them if they are not presently available in the home. Those actions may for example include calling a delivery service.
- A specification of the actions that can be used for moving objects from their present location to the place of performance.

The following procedure may produce an initial plan hypothesis for the preparation stage:

- Identify the set of objects that will be needed in the place of performance.
- Identify their present location.
- Identify move actions that will get each required object to the place of performance.
- Aggregate the move actions, for example using actions that can take several actions at the same time, instead of one by one. This also includes e.g. using a tray or a cart in order to be able to bring several objects at the same time.

Then check out the preparation plan in the following way:

- Execute the aggregated tentative plan in lookahead mode in the model world. If it fails at some point, for example because some action is determined not to be applicable, then revise the plan and iterate.
- When the tentative plan executes correctly in the model world, then execute the performance stage as well in the model world and verify that it works correctly. If it fails at some point, then consider two possibilities: revising the preparation plan, or postponing the problem until it really occurs (if it is going to occur) in the performance stage.
- Execute the preparation plan in the outside world. If problems arise in spite of the checkout in the preceding steps then assess the situation at hand, and revise the plan.

Some activities recur several times in these steps and one can in fact identify the following three major types of activities in them:

- Initial analysis of the situation in order to identify what objects are needed and where they are presently located.
- Making a plan, and modifying an existing plan according to information about some situation where it fails.
- Robust execution of a script, i.e. a sequence of actions, which means executing it in such a way as to respond adequately if some of the actions fail. This includes both the test-execution of a plan in the model world, test-execution of the execution-phase script in the model world, and execution of these in the outside world.

These three kinds of activities have one thing in common, namely, they are based on the use of a *goal* . The first activity identifies a goal, the second activity finds or changes a plan for achieving a goal, and the third activity relies on the available goals when there is a problem, and also when there is an uncertainty about exactly how to execute an action in the script at hand.

One may object that the script for the performance stage is not accompanied by a goal: it just specifies what may happen during that stage. However, the concept of a "goal" must be understood in two ways. There can be specific goals, such as "cool these beer bottles in time for the arrival of the guests," but there can also be standing goals such as "make sure that the glasses of all guests are at least half-filled unless the guest has declined a refill." In that sense there are goals even for the performance stage.

The "Compendium of Programming Techniques for Knowledge-based Autonomous Systems, Part II: Intelligent Autonomous Agents" contains a discussion of how to approach the implementation of this scenario.

# Chapter 3

# Non-Deliberating Intelligence

The intelligence scenario in Chapter 2 must be understood as an example of *deliberation:* the person performing the party planning thinks systematically (more or less) and rationally (more or less) about the task at hand. This deliberation can to some extent be observed by introspection, and this is the natural basis for how one would implement a similar capability in a computer.

## 3.1 Limits of Deliberating Behavior

With respect to the scenarios in Chapter 1 it is not as obvious whether, or to what extent the young child actually thinks about her tasks in a deliberating way. It is hard to know, since in particular the child is not likely to be able to introspect about this and to give us an answer on that basis. We can observe however that when we arrive at similar behavior as grownups then some of it is clearly deliberative, in particular if we pick a method that does not work (analogous to putting the chair so that it blocks the refrigerator door), and otherwise we do these things "without thinking about them."

Some of the non-deliberating behavior may be analyzed plausibly as *compiled* versions of deliberating behavior. This means that at some earlier times we have performed the same or very similar tasks in a deliberating manner, and the methods that we used then have been stored away in our memory in such a way that they can be taken out and applied without us thinking about them. If such a spontaneous application of a compiled behavior fails to work then we are often able to reinvoke the deliberating-level account of the behavior and think again about how to best perform the task at hand.

The analysis in the previous section is tentative and un-scientific since it is not supported by properly done experiments. It shall be taken as a way of thinking about deliberating behavior that can give guidance to whether and how such behavior can be implemented in computers.

## 3.2 Blink Behavior

Besides the cases that were discussed in the previous section, there are also some types of cognitive behavior that can not reasonably be represented as even the indirect result of deliberation. This arises in particular when we make very rapid decisions on specific topics.

The use of *reflex actions* is one example of such rapid decisions. In that case there is an obvious rational reason why we are capable of reacting in that way, for example, it may be necessary in order to avoid imminent danger, or take take immediate advantage of a possibility that is only available for a moment.

Reflex actions are innate and should not be confused with *trained rapid response* actions which are extensively used in sports and other physical activities. (More about this here).

*Immediate judgements* represent an additional type of non-deliberating behavior. The arise when we observe a person, an object or a situation and decide immediately on a characterization of the situation. This can be a characterization chosen from a small set of categories, such as “like” or “dislike,” but it can also be obtained from a broader set. An art expert may look at a painting and not merely decide whether he believes it is fake or not; he may also make a rapid decision about the country and the century where it originates, or even about who is the painter.

A fairly recent and popularly written book, “Blink” by Xxx, describes examples of this kind of decision as well as what is known about the cognitive processes that make them possible. This book is recommended reading.

Immediate judgements or “blinks” are sometimes the basis for decisions of paramount importance in the life of a person, a company, or a country. However, small-scale judgements of the same kind are used as components for deliberating actions. In terms of the scenarios in Chapter 1, how did Ronnie decide to use a chair as an instrument for being able to open the freezer door? How did she know that she could push the chair? How did she decide on how to grip the chair with her hands so that she could climb onto it? These are a multitude of small decisions without which the deliberating behavior would not be possible, but they can not all be understood in terms of deliberating behavior, compiled or not.

Artificial intelligence systems must therefore also contain aspects or components that correspond to “blink behavior.” The use of *decision trees* is one of the techniques for this purpose.

With respect to terminology, should “blink behavior” be considered as an integrating aspect of intelligence, or as another kind of intelligence besides ordinary intelligence, or as a separate phenomenon that does not go by the name of intelligence at all? This is largely a matter of definition, since there is no objective way of extracting the intelligence out of a person and inspecting it for its contents. Some people have strong opinions to the effect that one should take about several different kinds of intelligence, and there may be certain empirical arguments in favor of one or another position in this respect. However, from the point of view of Artificial Intelligence this question does not matter much; one must merely realize that A.I. systems must include methods for both immediate judgements and compilation of deliberating behavior.

# Chapter 4

# Scenarios for A.I. Systems

The previous chapter defined a computer implementable notion of intelligence. Let us now proceed to some scenarios for how intelligence in the same sense can be integrated and used in computer software.

## 4.1 The Trip Preparation Scenario

The following scenario is inspired by the CALO project (Cognitive Assistant that Learns and Organizes), `http://www.ai.sri.com/project/CALO`, and by the Party Preparation Scenario. The scenario concerns one use of a software system that assists its users with practical tasks such as planning forthcoming activities, obtaining information of relevance for those tasks and communicating such information to others, and so forth. The system contains the information for, and interfaces to familiar services such as calendar, email and information search, but it is autonomous and is able to use those services by its own decision, and not merely under direct command by the user.

In this scenario, Vilhelm and Jennifer intend to make a long trip through several countries in South-East Asia using ground transportation, in particular buses, but also trains. They have a few well-defined stops, such as persons they want to see and places they want to visit, but this will only determine a few aspects of the plan. They also have a number of criteria for what they consider "of interest" and "convenient," and "practical."

Planning such a trip well may require considerable effort, but in this case Vilhelm and Jennifer rely on the assistance of a system called SAPA, for Scenario-Assumed Personal Assistant. This system performs two main tasks for them: *planning their itinerary in outline* and *identifying what needs to be done before departure.* Planning the itinerary makes use of detailed knowledge about the geography of the countries involved, knowing about means of transportation, knowing about what may be possible places to visit and from what point of view they may be of interest, and of course knowing about the preferences of Vilhelm and Jennifer. Their preferences are by the way only partly the same, but not identical. At first sight this may seem to be a straightforward optimization problem: consider the space of possible itineraries, assign a merit value to each of them according to Vilhelm's and Jennifer's preferences, and pick the best one.

There are a number of problems, however. There is a problem of *incomplete information* since the knowledgebase in the SAPA system is not able to contain all information about local transportation options, especially since this concerns what will be available two to four months into the future, and not at the present time. Representating complete absence of knowledge is trivial, but in order to be useful the knowledgebase must also be able to represent partial and unreliable information, which poses particular demands on both the representation language and the computational processes for planning the itinerary.

Another problem concerns the terminology. Vilhelm has told his SAPA system that he is interested in "cathedrals" but not so interested in "monasteries," but what does this imply with respect to Buddhist temples – are they more like European cathedrals or more like European monasteries from his point of view? Both the general knowledgebase and the specifications of Jennifer's and Vilhelm's preferences must be expressed using terms in ordinary language, which means that the SAPA system must have access to an *ontology,* that is, a structure which specifies how different concepts are related for example as synonyms, or one being a special case of another, or one having a variety of meanings that apply according to context.

Identifying what needs to be done before departure is also more complex than one may think at first. Some basic things are easy: identifying visa requirements and applying for visas, identifying health requirements, scheduling vaccination and purchasing protection against malaria and possibly other things. The difficult thing is however to *foresee possible problems* and to prepare for them. What happens if I lose my passport, what shall I do then? What happens if there is a problem with my bank account so that I can not obtain additional cash while travelling? What happens if there is a natural disaster or a civil disturbance that causes me to miss the expiration date for my visa, will I be in trouble then and what shall I do? What should I do if I get arrested? The list can be continued for a long time.

The SAPA system needs to have several capabilities with respect to these exceptional circumstances and events. It must be able to generate a list of things that can go wrong, in line with the examples just mentioned. This requires that it has a capability for anticipation, just like in the Freezer Door scenario, but in this case the anticipation should cover not only the "normal," problem-free execution of the travel plan, but also a number of alternative executions where something goes wrong.

At the same time, the system can not be allowed to delve forever on all possible problems and unlikely combinations of unlikely problems. It must be able to assign levels of improbability and of dangerousness, and give priority to thinking about the less improbable and the very dangerous alternatives first.

For each case that it considers, the system must think of possible ways to recover, that is, it must consider what are the options for plan revision. However, it need not always *produce* a revised plan, in many cases it is sufficient for it to have enough evidence (i.e., to believe) that *there will be* a revised plan if and when the problem situation shall arise. On the other hand, if the system realizes that in particular type of problem situations it will be important to have access to a particular piece of information (for example, your bank's emergency telephone number) or a particular object, then the system should arrange for this to be available when the trip starts.

Besides anticipating problems and having solutions ready for them, the system will also have to interact with its users. Vilhelm and Jennifer will certainly also think about things that can go wrong; they will need to ask the system for its answers to some of the problems, and they may also think of some problems that the system did not identify. The Bath Assistance scenario in the previous chapter described how in order to understand a dialog, you must be able to understand the reasons why particular phrases are being uttered by the other party. The same requirement will hold in the dialog between Vilhelm and Jennifer and their SAPA system.

## 4.2 The Major Disruption Scenario

Suppose a major disruption occurs in a place where a large number of people are present at the same time, such as a large railway station or an airport. We will not think of physical disasters like major earthquakes or major terrorist attacks, but merely of situations where the normal routines are completely disrupted, for example the cancellation of most flights due to an ash-cloud problem, or a situation where trains under the English Channel have stopped in the midst of the tunnel and are forced to stand still for many hours with passengers trapped in them.

In such situations, a large number of people will wish to communicate with the outside in order to send or obtain information, and many will also wish to have access to services, for example for changing travel reservations, or for changing appointments that were on their agendas. The most obvious way of doing this in the world of today is using mobile phones. This sometimes works fine, but there are two limitations that must be taken into account: will the mobile phone system be able to accomodate the large number of calls, and will the requested services have a sufficient capacity?

A possible answer to both of those concerns is to restrict communication to SMS messages, and to arrange that only those voice calls or data calls that are accompanied by a special priority license will be serviced during severe disruptions. SMS messages pose less requirements on the communication network, and it would be possible to channel them to a dedicated software system that reads and interpret these messages and takes appropriate action on them.

Consider therefore a scenario where a national authority that is responsible for disruption and disaster management has built a powerful software system that is always available on a standby basis, and that is able to receive a large flow of SMS messages that may arise in a disaster situation, and to service them in reasonable ways. Such a system will have to communicate in natural language, but its task can be expected to be much simpler than the general case of natural-language dialog since SMS messages are forced to be so short, and since one can expect that a large proportion of the messages fall within a limited number of topics. Moreover, when there are incoming messages that the system does not understand, it will always be able to forward those to people that also participate in the assistance and rescue effort.

Nonetheless, the kinds of dialog behavior that were exemplified above can be expected to occur even in such SMS dialogs, for example, people's natural habit of stating their essential message implicitly and not explicitly. If anything, the compact nature of SMS messages will rather encourge the use of such concise ways of communicating.

Practical problem-solving with the same general character as moving the chair to the freezer can of course apply in a disaster scenario as well. The contribution that the computer system can make, besides what people are able to do with their own common sense, is that the system may have knowledge of hidden or non-obvious resources of various kinds that may be present in a railway car, for example, and in this way it may be able to find a problem solution that the people having the problem would not be able to think of.

## 4.3 Size of the Knowledgebase

The two scenarios in this chapter have been grossly simplified in order to make it possible to describe them in a short text, but anyway they demonstrate two important things: (1) Many of the intelligence-related capabilities that were described in the children scenarios of Chapter 1 are also applicable, and in fact necessary for practical scenarios such as those described here; (2) Practical scenarios will often require large knowledgebases containing substantial collections of information about the world where the system is used, or is going to be used. This applies even to the scenario examples, and even more to real-life systems.

The list of important research topics in artificial intelligence therefore includes not only methods for performing these various capabilities, such as planning, diagnosis and learning, but also methods for *representation of knowledge,* namely, the knowledge that is required in order for those capabilities to operate in a meaningful way.

# Chapter 5

# The Concept of Intelligence in Psychology

Let us now return to the question that was briefly addressed in Chapter 1: is it reasonable to consider intelligence as one concept or as several? For example, should Ronnie's ability to open the freezer door be considered as a manifestation of the same competence as her ability to resolve the doll problem with Valerie, or should these be considered as two different competences? As already stated, this question is of marginal importance for artificial intelligence, but it may be more important for cognitive psychology, that is, for that branch of the science of psychology where intelligence is a major topic of study.

This question is in fact deeply controversial, and the reader is encouraged to visit the Wikipedia page for "intelligence" to get a first impression of the divisions. Besides reading the article itself, it is also interesting to read the "discussion" section that is attached to that page, and the archived earlier discussions. The following is a brief excerpt from that page:

Intelligence derives from the Latin verb *intellegere;* per that rationale, "understanding" (intelligence) is different from being "smart" (capable of adapting to the environment). Scientists have proposed two major "consensus" definitions of intelligence:

(i) from *Mainstream Science on Intelligence* (1994), a report by fifty-two researchers:

*A very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings catching on, making sense of things, or figuring out what to do.*

(ii) from Intelligence: Knowns and Unknowns (1995), a report published by the Board of Scientific Affairs of the American Psychological Association:

*Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, [and] to overcome obstacles by taking* *thought. Although these individual differences can be substantial, they are never entirely consistent: a given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. Although considerable clarity has been achieved in some areas, no such conceptualization has yet answered all the important questions, and none commands universal assent. Indeed, when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen, somewhat different, definitions.*

These definitions emphasize what almost everyone can agree on, which means that they do not address the hard questions where opinions differ. As a framework for understanding the disagreements, notice first of all that intelligence is considered as a combination of a number of more specific capabilities *each of which can be measured by suitably designed tests* whereas for the general concept of intelligence as such it is much less clear how one would design a single test. An important question is therefore whether, and for what reasons is it appropriate to bind together a number of separate, measured capabilities and apply the single term "intelligence" to the combination of them.

Several approaches to that question may be considered. If it were possible to identify a region of the human brain that is active in all these separate capabilities and inactive otherwise, then that would be a very concrete reason for assuming the existence of a general intelligence facility in humans. Unfortunately, however, no such brain region has been identified and available evidence suggests that the contrary is true, and that different brain regions are active for different purposes. On the other hand, this does not prove that intelligence should not be considered as a single concept: it could be a single competence that is realized in a distributed way in the brain.

Another approach is to consider *co-occurrence* of the specific capabilities that are mentioned in the definition of intelligence. Using the statistical data from tests for specific capabilities in a large number of people, one can check to what extent the presence of one capability is related to the presence of another capability. A strong degree of co-occurrence of all the capabilities, or of a group of some of them, will be a plausible reason for assigning a name to that group and to study it as an entity.

In principle, strong co-occurrence between several capabilities may be taken as an indication that there is a common underlying mechanism that produces all of them. However there are also a number of other possible explanations, including neural-level explanations (maybe the average number of connections from each brain cell to neighboring brain cells is important for all of these capabilities, and maybe this is the underlying factor explaining an observed co-occurrence?), genome-level explanations (maybe particular aspects of the human genome influence the development of these capabilities in the fetus?), and nutritional and social explanations. It may or may not be possible to identify which of these are plausible explanations for observed facts, but even in the absence of an agreed explanation it can anyway be useful to have a name for a bundle of co-occurring capabilities. For example, given that intelligence as measured by IQ tests appears to be a usable predictor for "success in life," it may be useful to have it even if one does not understand the underlying mechanism.

Besides observed co-occurrence, there are also a number of other ways of studying the relationship between identifiable capabilities, in particular by studying individuals with unusual properties: individuals that are extremely skillful in some particular capability but not in others, or individuals that have suffered brain damage that strongly impairs some of the capabilities. This may shed light on to what extent there is a common basis in one way or another for several of the capabilities.

# Recommended Additional Reading

## General Aspects

http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/AIOverview – Association for the Advancement of Artificial Intelligence, Webpage for Artificial Intelligence Topics, subdivision of A.I. Overview

http://www-formal.stanford.edu/jmc/whatisai/whatisai.html – John McCarthy's answers to What is Artificial Intelligence

Wikipedia articles on Intelligence, Artificial Intelligence and Knowledge Representation

## Philosophical Aspects of Artificial Intelligence

http://www.aaai.org/aitopics/pmwiki/pmwiki.php/AITopics/Philosophy – AAAI Topics Article

http://en.wikipedia.org/wiki/Philosophy_of_artificial_intelligence – Wikipedia article