Experience based reasoning system coupled with real world knowledge

Ronald L Hartung\textsuperscript{a*}, Anne Hakansson\textsuperscript{b}

\textsuperscript{a}Franklin University, 201 S. Grant Ave, Columbus, Ohio & The Design Knowledge Company Fairborn Ohio
\textsuperscript{b}KTH Royal Institute of Technology, Software and Computer Systems, Forum 100, 164 40, Kista

Abstract

Human intelligence draws its conclusions from a base of experience-generated knowledge. Beside being able to use this knowledge, it is limited to how much can be accessed at a time and this reasoning is often shown to be illogical, with respect to mathematical logics. The human’s knowledge is always growing and being modified by current experience. In addition, the humans’ processing capability appears to be severely limited. This limitation is far from being a burden; it is part of the brilliance of the solution. The human mind does a every effective job of dealing with the world. For proof, we invoke the fact that human kind as survived and thrived. The current work is a first step in exploring a reasoner than can act in a human inspired performance, which is in the direction of general artificial intelligence.

1. Introduction

When developing artificial intelligence (AI) systems for particular solutions, the systems are often built with known data, facts, information and knowledge and without the dynamic aspects of learning. Commonly, the systems are not building up their own content from scratch because there is a need of a starting point for the systems to work
properly in a domain. In contrast a general AI-system must look at the totality of the whole situation using several senses. In this paper, general AI systems are the focus; the term AI will mean general AI in the rest of the paper. The AI-system needs simplicity, meaning that it does not have to have all the knowledge about everything in a domain but it needs to have an understanding about the different parts including experience of the parts. Hence, different AI-computerized sensory systems need to be incorporated into a totality that constitutes a combination of senses.

The research in this paper, builds on earlier research in AI-systems and abductive reasoning. Building on prior work, in this paper is a synthesis of work from both separate and joint projects. The overall goal is AGI, artificial general intelligence. While that goal seems a long way off, the attempts to build it will be found generally instructive and the author hope this will lead to better and more robust AI systems along the way.

To develop a system that is self-developed with its own knowledge base, the AIC system has been proposed in Håkansson1. The AIC system is built for handling a combination of senses, as perceptions, but strives to handle emotions and, finally, thoughts. It builds up the knowledge using senses about its surroundings and draws conclusions from this obtained knowledge. To understand it surrounds, it builds on Aleksander3, consciousness of five axioms3: sense of place, imagination, directed attention, planning and decision/emotion. By developing its own knowledge base and incorporate the five axioms, the AIC system becomes a knowledgeable system. However, the AIC-system is not experience-based yet, since it needs a reasoning strategy that can distinguish what is really good valid knowledge which can be used as a relevant experience and what can be sorted out as useless information which can be stored, but not used in a particular setting, where it is not needed or wanted.

The approach followed in this research work in this paper is to build parts of the overall architecture piecemeal and test each one. The first piece is described below and was aimed at situation recognition. The part discussed here is a first step to building a reasoning system to augment the processing of situations.

There are a number of assumptions that drive the design: First of all, the assumption that human processing is limited. This is not taken as a limit on human intelligence, but rather a defining characteristic. This has driven a system that uses parsimonious and highly effective processing and representation approaches. The representation is necessarily not a complete high fidelity representation, but one that is effective in the human’s ability to deal with the real world. For example, humans have no difficulty picking up objects. An assumption made in this paper, is that situations are the key to representation. That situations are linked representations and have an ability to use chunks to both simplify and to expand the representation as needed. Chunks are used gather a part of the representation into a single unit. This reduces the load on short-term, or working, memory in the human by reducing the number of things to be stored at one time. The chunk can also be expanded, reversing the gathering of details.

Intelligent systems must be experience driven, as are human intelligences. The blank slate (tabula rasa) is often cited with respect to experience. In our work, the assumption is that human intelligence relies on experience and the slate is “rather” blank at the beginning. However, the human is given a great gift of evolution, i.e. humans have some basic abilities and neural systems that are set to work with the surrounding world. This helps develop an effective and useful intelligence in a reasonable time frame (e.g. childhood). One consequence of this is that childhood development is very informative about the processes of the human mind. As a side note, the authors believe AI researchers should spend more time looking at research in childhood development.

Reasoning is not defined by symbolic logic. This runs counter to AI’s heritage. We reject conventional symbolic logic for a number of reasons. First of all, humans display actions that are clearly illogical when viewed by those schooled in symbolic logic. The view herein is, humans do what they do, it is not wrong! Nor are humans are not limited to deductive styles of reasoning. Induction and abduction are required to achieve human like reasoning. Humans can work with contradictory beliefs and agilely shift between them; Godel’s theorems are not an issue for humans. In doing these shifts, they are not acting like non-monotonic logics, they do not back out prior deductions, and they just ignore what does not fit into the current process.

Another assumption comes from the idea of flow. The best definition for what is meant by flow comes from the writing of Pred4. This is based on ideas from William James, John Searle and A. N. Whitehead, and it deals with the idea that thought is a continuous process flowing from situation to situation. Pred4 is more concise and a easier read than James and Whitehead, so we reference his work and indirectly, through him, the prior works. The flow is directed; the situation recognized by the situation triggers what James referred to as habits. A habit is an action on the part of the human. In an AI sense, these as learned procedures. The triggering of a habit affects the system in two ways. First the situation generated provides expectation, i.e. what the system expects to see next. Second, a human
sensory system does not attend to every thing, the habit drives the sensory system by focusing the attention, for example the head moves to view the hand as an object is being grasped. Therefore, the present situation is a driver for the next situation along with the sensory input. This is important because it answers the issues around humans ignoring part of the input and being part of the process that determines attention. Also, it allows the human to “tune” out parts of the environment and maintain a simpler internal representation.

2. Overall architecture

The first step in this architecture was published at WorldComp in 2013. This was looking at the issue of how can an intelligent system be grounded in world and how situations can be stored and retrieved in a useful way. The goal of the system was to represent situations and to associatively retrieve them. The input system used the Concept Geometry of Gardenfors. The concept geometry gives a grounded approach to processing incoming perceptions and producing concepts. It is based on observations of human processing. The concepts used as a basis in this work would be produced by a concept geometry implementation, but that part is not discussed here.

The goal in this paper is to use ConceptNet as long term memory of concepts and use the NARS reasoner by Wang as short-term memory and reasoning engine. The fundamental problem here is how to map ConceptNet terms and relations into a form that the NARS engine can use, preserving the semantics between the two systems. In this work we are not adding experience to ConceptNet, we use it to establish a set of experience. This will be changed in future versions of the system to allow either learning from scratch or to add experience to ConceptNet. But the goal here is to establish the NARS reasoned and the flow model. The figure, figure 1, below is the test system for this paper.

The sensors are feed through the Concept Geometry to generate terms. ConceptNet is queried finding the terms and relations to generate a network. The network is built from nodes for terms and with edges for relations. This network is a very rich representation of the perceived situation. NARS will be used to draw conclusions from the situation. A near term change is to tie this reasoned into the rest of the systems so that the system can direct the reasoner with goals.

The diagram was kept simple by not showing the loop back to the sensors. NARS implements habits, in the sense described in Hartung, and has control of the system output (effectors) and commands the sensors to attend to the environment to meet the expectation of the habit.
3. Components

The work in this paper, is to define a reasoning system that can be used to extend the architecture in Hartung\(^5\). This was initiated by searching for a useful reasoning system that was consistent with the principles given previously. To that end, two specific aspects are the subjects of the paper: First, how do we incorporate enough real world knowledge. Second, how does the system fit into the model of flow.

The issue of logic has been with AI for a long time. Just as chess was once seen as an ultimate test, so was logic assumed to be the basis of thought. Symbolic logic is, in reality, one of the tools that the crafty ape called man invented to extend his capabilities, just as he earlier invented the lever. It is common to hear some psychologists and cognitive scientists when discussing a test like the Wason selection task\(^{12}\), that humans don’t use logic. Perhaps the issue is we are blinded by our invention, i.e. logic.

The current topic is to extend the architecture in providing a reasoner that can act like a human. When someone invokes logic in the current era, one usually envisions first order logic. It is, after all, what most reasoners are based on and what most of the AI researchers, and most scientists, have been trained in. A common comment that arises when physiological experiments are analyzed is that “the subject is not acting logically”. In fact, this means the experimenter looks that the problem analytically with the training in logic, but the subject is doing human reasoning, not first order logic.

Another problem we have with reasoners is the human technique referred to as abduction, Peirce\(^{11}\). First introduced by C.S. Peirce. Abduction can leap to quick and often effective answers that are not logically justified by a first order logic. Abduction is not “valid” by first order logic.

3.1 NARS an non axiomatic reasoner

Pei Wang\(^9\) has been developing his non-axiomatic reasoner since the early 1990’s. It has a number of characteristics that are attractive to the research presented in this paper. (While a full discussion would consume too much space in this paper, it is required that some main points be presented.) A principal driver of the work is the recognition that humans have to work with limited resources and limited time.

Conclusions in NARS\(^9\) do not depend on axioms that remain true for all time, or even require the axioms and theorems to be consistent. In that sense, NARS bypasses Gödel’s theorem\(^{13}\). Of course, having so stated, NARS does not attest to truth in the sense of normal logics. Truth in NARS is based on collective experience, namely how often a fact is seen as true versus how often it is seen to be false. In addition, the system uses an event horizon into the future. The event horizon is used to estimate how likely the fact may tip to false in the near future. The expression for “truth” in NARS is a pair, (frequency, confidence). These are values that range between 0 and 1, hence NARS is a multi-valued logic, like fuzzy logic. Frequency is a measure of the evidence for a statement (positive evidence / total evidence). Confidence is the measure of stability a fact, that is, a measurement of the possibility that the past evidence will be overturned. Confidence is computed as the total evidence / (total evidence + horizon). The horizon is a constant and represents an interval into the future; Wang generally uses 1 time unit. While this may look like probabilities, they are not, they do not follow the proper axioms of probability. Likewise, they are similar to fuzzy logic, but the confidence term is outside normal fuzzy logic.

NARS supports deduction, induction and abduction, as deduction rules defined in NARS. NARS is a more general reasoner than the usual first order logic deductive only systems. The basic derivation rules are defined with respect to the NARS basic operator, i.e. inheritance (\(\rightarrow\)). This is defined as “\(S \rightarrow P\)” meaning \(S\) is a \(P\). Or alternatively, \(P\) is the generalisation of \(S\). In NARS, deduction rules include a computation of the resulting truth-values. An important rule in NARS is called choice and is used to choose between multiple contradictory alternatives. Choice is what allows the use of inconsistent statements in the reasoner. This is not a form of classical non-monotonic reasoning that depends on revoking previous derivations in the light of new input. This strongly supports our intuition that humans deal with contradictions very smoothly, often by ignoring part of the information. Note, humans don’t necessarily revoke information; they can just ignore it, or reject it, for the present situation.

Another important issue in NARS is the handling of intention and extension. NARS has explicit representations of these concepts, and they do not exactly follow the classical definitions. This is an interesting problem translating from relations from ConceptNet to correct forms for NARS. Some relations are intentional or extensional in nature. NARS defines extension and intentions based on the basic inheritance operator. The extension of a term is the set of
all x from \( x \rightarrow \text{Term} \) and the intention of a term is the set of all x where Tem \( \rightarrow x \).

NARS is implemented with a memory that includes forgetting. This is in response to the limitations of human capabilities, and it is an implementation of short-term memory. NARS statements in the working memory are erased over time by decaying when they are not accessed. This is a limitation of NARS that lead us to include ConceptNet as a long-term memory to complement the working memory of NARS. It also works with the idea that humans do not seem to consider all information as they reason, but limit themselves to what seems to be relevant in the current situation.

NARS has a large set of operators; the most basic operator is inheritance. The mapping from ConceptNet is based on defining the correct operator and deriving the truth-value from the data in ConceptNet. Event representations in NARS have a rich set of operations and goals. The truth value in the goals is similar to the standard NARS values, but represents desire with respect to the goal.

### 3.2 ConceptNet

In order to achieve a useful level of ability and still maintain the ideal of experience requires a carefully chosen short cut. ConceptNet is a crowd sourced semantic net. The terms in the net are of two basic types. They are either concepts or parts of speech (imported from WordNet). ConceptNet also has terms from multiple languages. For this work, only the English terms are used. ConceptNet defines 50 relations. As shown in the table 1 in section 4.0, some of the relations map directly other use the general NARS relation notation. The relations in ConceptNet are binary relations. Terms can have multiple occurrences, for example red is a concept, noun, adjective and has three difference terms in ConceptNet.

When ConceptNet is queried, it can generate a large network of related terms form the input set. The reasoner will need to apply context to limit the network. Context is part of the flow model (section 5). This models our view that humans have large connectivity between concepts and that they limit what is considered in any context. However, they also can use this complex linking as needed. Chunking can also be applied to reduce the concept space for a given problem to produce a suitable universe of discourse. Implementing chunking is planned for a subsequent phase of our work, is not discussed in the current work.

In order to create a rich representation of the perceived input to the system, the terms generated by the concept geometry is used to drive a set of queries to ConceptNet. This produces a very large network of concepts and relations. While this is representative of all the possible knowledge that a human can draw conclusions on, it is not likely to represent what is actually invoked in the mind. This conclusion is reasonable based on the idea that human processing power is limited. The human mind appears to be selecting what a subset out of this possible network.

### 4. Translation

The table below gives a summary of the translations used between ConceptNet and NARS.

| Relation in ConceptNet | Transformation in NARS |
|------------------------|------------------------|
| IsA, HasSubEvent, InheritsFrom, InstanceOf | Inheritance | \( X \rightarrow Y \) |
| SimilarTo | Similarity | \( X \leftrightarrow Y \) |
| SimilarSize | Relation | \((\text{Prod} \ X \ Y) \rightarrow \text{SimilarSize}\) |
| NotIsA, NotHasSubEvent | Negated Inheritance | \(~(X \rightarrow Y)\) |
| Attribute, HasA, HasProperty | Property | \( X \rightarrow] Y \) |
| All others | Relation | \((\text{Prod} \ X \ Y) \rightarrow \text{relation}\) |
| Causes, Entails | Predictive Implication | \( X =/> Y \) |
| CausesDesires | Predictive Implication plus a relation | \( Z =/> X \ (\text{Prod} \ X \ Y) \rightarrow \text{desire} \) |
| ReceivesAction | Either predictive implications or | |
Extend NARS with action functions (See text)

In the above table, several operators are directly defined in NARS, such as inheritance, similarity, property and instance that are all basic operations of NARS. Only assigning truth-values is a challenge. Since ConceptNet does not represent the accumulation of experience explicitly, we have to assume some basic values from the NARS frequency and confidence terms (discussed below).

NARS representation of relations is a bit trickier. Relations are of the form (Prod term1 term2) -> R where R is the name of the relation. Prod operator is a called the product connector. The Prod operator can also be written in infix notation. It joins the terms. There are no derivation rules that operate directly on relations, but rather work by substitution. That is, when two terms can be derived as equivalent, they can be substituted into the relation.

The relation of causality is represented by an operator called predictive implication. This is also used for the implementation of habits.

Truth Value derivations are somewhat problematic with this approach. ConceptNet has a weight value, but it is not normalized. This leads to several possible solutions. One could scan the existing dataset, given that ConceptNet is treated as static, to determine the maximum value of the weight and then normalize the value. However, this would assume that the weighting is universally meaningful, which is unclear. Second, it would have to be readjusted when additions are made the set. The current system under test does remain static, however the long-term plan is to allow the system to acquire new knowledge in both concepts and relations. Also, since the work to define ConceptNet accumulates the collective knowledge of the crowd that created it, it can be assumed that the frequency value assigned to truth is not limited to the instances in ConceptNet.

The initial solution being tested in this work is to assign truth-values by the following method. The frequency is always given a high value, 0.9. The choice is based on the idea that ConceptNet is assumed to be correct. The confidence value is determined by the weight from ConceptNet. Since weight is generally 1.0, but can going less or more than 1.0, the following formula is used: confidence is given a default value of 0.7. When the weight is less than 1.0, we reduce the base c by multiplying c by with weight. When the weight is positive, we increase the confidence, computed as 0.3 multiplied by the reciprocal of the weight. Thus infinite weight will boast confidence to 1.0. This is a first cut and will be used to test the system.

5. NARS reasoning and flow

ConceptNet will easily generate large semantic networks in response to input. The role of NARS is to provide conclusions and actions over the representation. In doing so, it must act to reduce the representation to a manageable level. This is an essential aspect of the limited compute power of a human like intelligence. And our philosophical position, abet unproven, is that this limitation is used to build the robustness and effectiveness of general artificial intelligence. It works by reducing the representation to the only relevant terms and further by filtering through expectation and focusing the attention of the sensors. As this section will show, the concept of “flow” is the cycle of the system.

NARS provides the basic facilities to construct a solution. That solution is based on two constructs from NARS. These are operations and events/goals.

NARS defines operations. An operation is defined as an atomic operation with arguments. As NARS is written in JAVA, operations can be added through extension by tying them to JAVA method calls. The effectors and sensor focusing controls (the reason for focusing will be shown below) required by the solution domain, they define the vocabulary of operations needed by the system under construction. While the goal of this research is general AI, the test systems we construct are limited to a specific environment and purpose. In this paper operators are written as (op:name , op1, op2, op3…). In NARS, special symbols are used, but this notation of the operators is typographically easier.

The second operator is the goal. A goal is a NARS statement with a desire value instead of a truth-value. A goal is written as <statement>!<desire value>. The “!” syntax denotes the goal. A desire value is a pair of real numbers in the range 0 to 1. The terms are desirability and plausibility. The terms are used to select the possible plans that
NARS will select from the current situation and library of plans and actions. The situation is the current semantic network. There is an additional component of internal state, that is, the set of “habits” that have been selected and operationalized.

The term “habit” is chosen here from the work of William James\(^5\). He used the term habit to describe the idea of a human’s “pre-compiled” scripts for actions. These are built from the basic operations available to the system. Since the set of basic operations are fixed, the system can learn new habits by creating a set of NARS statements and adding them to the memory of the system. It is important to state the nature of habits in order to understand the operation about to be described. However, the learning portion of the system is not presented here, although it is both necessary and very central to general AI systems. The learning system will also include a level below the flow process, the supposed type 1 as in Hartung\(^5\).

In NARS, a habit is a set of production rules using the predictive implication operator to construct sequences. In general, this is a set of statements of the form (condition, operator) \(\Rightarrow\) consequence. Thus a habit is a sequence of these statements, linked in sequence by the condition produced as part of the consequence of the prior statement. (As a side note, in a human this is implemented in the neural systems of the brain, and probably has a more natural form. The logic rules work in our model, but we do not claim a neuro-physical basis.) The simple form of rules can also become compiled into a tighter form by replacing the single operator with a sequence of operators that work atomically. Atomically in this case means without NARS reasoning interceding between steps. This represents the idea that humans learn habits that operate below the level of conscious reasoning (James’s type 1 system). Being able to reach for objects is a habit in the Jamesian sense\(^5\), it is a common sequence that requires no conscious thought. Reaching for a glass reduces the field of view and it also causes the visual system and early processing stages to ignore some parts of the environment. This keeps a simplified view of the environment and also limits the things perceived by the system while the habit is running on the system. In the human, at the lower levels of the type 1 system that James described, events can occur that will interrupt a habit, no matter how pre-compiled it is.

The system operating cycle, based of flow, can now be presented. This is inspired by process philosophy found in Hartung\(^5\). The sensors are active components, not just in transduction of external stimuli into system input, but also as being actively influenced by the operation of the system. The system focuses the sensors by the actions it is undertaking. An example from humans is when you reach for something; the eyes commonly turn to view the object and focuses on it. But there is more, the situation being generated in the mind is strongly influenced by the current habits. Whitehead used the terms aim and trajectory to describe this directedness of the system. In essence, the system is primed to expect situations to evolve in ways that are learned experience.

The cycle generates the next situation from the sensors and the prior situation. This is achieved by the execution of habits. These habits limit and control the situation in two ways. As previously stated, they focus the sensors on the input the habit is looking for. The other aspect is a set of expected conditions the habit is looking to achieve as goals. This becomes a filter on the next situation by choosing the concepts that match the expectations and ignoring the ones that don’t match. The result of this is a system that presents a dynamic of planned action. It generates a limited situation representation based on the expectation from the habit.

To visualize the operation of flow, a simple example follows. Given an agent in a small building, what might the agent do? The agent is standing the middle of a room, there is a sign on the wall. On perceiving the sign the agent’s aim is to read the sign and a habit is initiated to approach the sign. The effect of the habit is to focus on the sign, ignoring the other details of the room. Since the aim of the agent is to read the sign, the agent may not notice issues like marks on the nearby wall and what may be visible through doorways. The habit will stop when the sign can be read and which time the aim of the agent becomes reading the sign. When the sign is read, then the aim of the agent will be determined by what was read and possibly other perceptions. For example, a noise from another room may affect the aim.

Another possible flow would be if a load noise or a bright flash occurred during the move to the sign. This will cause the agent to choose what aim to follow. This would happen in a couple of possible ways. The agent could respond with a habit to look at the sound or light. This changes the perception and hence the internal model of the situation. This will cause the aim to be re-computed by the reasoner. In NARS, this can result in an inconsistent representation, in which case, NARS will use it choice rule to select a new choice of aim.

This basic cycle will require further extension as the move to more general AI systems is pursued. Much work needs to go into the handling of situations when the habit has produced an expectation and the sensor data fail to
show the expectation or show a denial of the expectation. This will need to trigger abduction as a solution and preplanning activity. It is unclear if NARS will naturally bridge this gap or if more explicit mechanisms will be required.

6. Conclusions

The driving goal of this work is to find a system that will respond to experience in meaningful ways. NARS shows promise in this regard by bypassing axiomatic basis for truth, supporting abduction and induction as well as deduction and supporting a tolerance for contradiction. It also facilitates the evolution of state in the system by implementing a short term memory with forgetting. As the memory forgets and adds new facts the conclusions and habits chosen will change. However, it falls short in its lack of a long-term memory. This design presented here allows us to tie an experience based long-term memory to the NARS reasoned. The fusion shown in this paper is not a total system. The concepts within Concept Net need to be grounded in the sensory system of the total system. The reference paper5 develops and approach to that grounding. The approach there will allow the system to add concepts and to apply analogical techniques to both the inclusion of new concepts and to support analogical reasoning. The testing phase of this project is just beginning, and it is uncertain if this will be part of our ultimate solution. We have put this paper out in hopes of invoking discussion into these topics.

References

1. Håkansson A, AIC-An AI-system for Combination of Senses. KES 2013: 40-49
2. Aleksander I, The World in my Mind, My Mind in the World, 2007. http://www.imprint.co.uk/books/aleksander.html
3. Aleksander I, Aleksander's five axioms of consciousness; 2003, http://www.converj.com/sites/converjed/2005/09/aleksanders_five_axioms_of_con.html
4. Pred R, Onflow, Dynamics of Consciousness and Experience, A Bradford Book, MIT press, Cambridge Mass, 2005
5. Hartung R, Weigand K, “Microgenetic Critic for Situation Assessment Supporting Abduction and Surprise”, International Conference on Artificial Intelligence, 2013
6. Gärdenfors P, “Conceptual Spaces, the geometry of thought”, A Bradford book, MIT Press, Cambridge Mass, 2000
7. ConceptNet conceptnet5.media.mit.edu/
8. Liu H, Singh P, "ConceptNet practical common sense reasoning tool-kit”, http://web.media.mit.edu/~hugo/publications/papers/BTTJ-ConceptNet.pdf
9. Pei Wang, “Non-Axiomatic Logic a model of intelligent reasoning”, World Scientifc, Singapore, 2013
10. NARS implementation http://www.cis.temple.edu/~pwang/demos.html
11. Peirce CS, "Pragmatism as the Logic of Abduction” (Lecture VII of the 1903 Harvard lectures on pragmatism), see parts III and IV. Published in part in Collected Papers v. 5, paragraphs 180–212 (see 196–200, Eprint and in full in Essential Peirce v. 2, pp. 226–241 (see sections III and IV)
12. Wason, PC (1966), "Reasoning". In Foss, B. M. New horizons in psychology. Harmondsworth: Penguin.
13. Some basic theorems on the foundations of mathematics and their implications in Solomon Feferman, ed., 1995. Kurt Gödel Collected works, Vol. III. Oxford University Press: 304-23.