

Prolegomenon* to Commonsense Reasoning in User Interfaces

by [Ergun M. Bicici](#)

“All our knowledge starts with the senses, proceeds from thence to understanding, and ends with reason, beyond which there is no higher faculty to be found in us for elaborating the matter of intuition and bringing it under the highest unity of thought.”

[\[14\]](#)

Introduction

Human-computer interaction experiences unbalanced talents of counterparts in the user interface, which can only be eased with the introduction of new solutions. Commonsense reasoning is a promising answer that offers formalization and computational models about how humans reason and think in a sensible way. In user interface design, assumed conventions and rules are widely followed and carried to user interfaces. For the most part, these assumptions are obvious to humans yet incomprehensible to computers. As a result, it is essential that tools are developed, capable of retrieving relevant, sensible inferences that in turn can serve as catalysts for future reasoning. Embedding this tool in user interfaces can provide many benefits including representation of assumptions and unspoken rules, the addition of useful tool abilities, and increasingly usable and accessible environments where computers and humans have extended communication capabilities to assist in understanding each other.

In this article, we present groundwork for applying commonsense reasoning to user interfaces. We start with identifying the asymmetry in human-machine communication and later focus on some approaches such as softbots and the proposed anti-mac interface. Then we identify the problems faced in user interfaces such as the correspondence problem inherited from computer vision. Next, we state our research ambition as adding commonsense reasoning functionality to user interfaces and further survey previous approaches and the state of the art. We depict the big picture we are facing and list some of the rewards to be earned by applying this technique. Later, characteristics of common sense are investigated together with examples in first-order logic. We then report various lessons learned from earlier attempts that began in physical systems concentrated on small microworld problems. Finally, we cite a sample methodology for automating commonsense reasoning and identify a number of questions that have yet to be answered.

Human-Machine Communication

Interaction between humans and computers requires mutual understanding and comparable intelligibility just as in human communication, where people seek intelligence alike and common ground while communicating with each other. Face-to-face interaction between people can provide a base model for *face-to-screen* interaction. However, there is an important difference: behavior of each participant in this communication is based on the resources provided by their circumstances and sensors, which creates an asymmetry [23]. People take advantage of a rich verbal and nonverbal set of resources whereas machines have a set of sensors that map to commands and reactions.

The resulting asymmetry limits the extent of the communication between humans and computers. Suchman [23] believes the solution can be provided by: extending the access of computers to actions and circumstances of the user, making the user aware of the computer's limits in accessing those interactional resources and finally compensating for the computer's inabilities with computational alternatives.

The designer of an interactive machine, as Suchman [23] calls, must ensure that the user gets proper response from the machine for his actions. Each interactive action assumes the intent of the actor with an adequate interpretation of the prior actions and the intent of the recipient with interpretation of the responses' implications. So, the interaction between computers and humans is dependent on each other's responses and their corresponding interpretations.

User Interfaces

In the communication and interaction medium between humans and computers, *user interfaces*, the language used is yet to evolve. In today's Mac-based user interface terminology, the vocabulary serve as the building blocks of this medium like windows, buttons, text boxes; the grammar serve as the rules between these building blocks, like a button cannot be inside another button or a window should contain a button inside itself; and idioms or expressions serve as the unspoken rules in this environment such as an expectation of an event after clicking on a button.

The anti-Mac interface proposed by Gentner and Nielsen [11] tries to take control over the actions from the user's autonomy and proposes a shared control of the environment between the user and other entities, especially computer agents and other users. In addition to shared control, they suggest richer internal representation of objects and a

more expressive interface. These suggestions, from a communication point of view, can potentially decrease the asymmetry between the human and computer mediums of interaction.

However, there is still a lot of progress to be made, especially in the current UI systems, the building blocks, the rules in between them and the unspoken conventions used vary according to the implementation technique used and the audience targeted on both sides of the communication. This makes it hard to come up with a general UI grammar.

Softbots

User interface *softbots* are intelligent software agents designed to control an interactive system through the graphical user interface. Previous detection efforts in softbots include using statistical pattern recognition techniques and rules and conventions in a Mac-based environment for finding the objects building blocks on the screen via a statistical search for more abundant forms [4].

The functionality or purpose of a labeled button on the screen can be easily guessed by a human user who understands the label's meaning without clicking on the button. On the other hand, identifying this knowledge is not as straightforward as it seems for a softbot [2] without NLP (Natural Language Processing) capabilities that try to determine which screen button can be used for opening a file, for example. For this purpose, a functional exploration of the user interface may be needed.

Correspondence problem in user interfaces

This problem is also experienced by a human who browses a web page (or a different UI environment) that is written in an unknown language. The user will proceed by matching the previously known functional objects to the ones present on the screen by comparing their similarities, resemblances etc. In the same situation, the softbot can similarly look at its knowledge base for recognizable objects that were identified before (an expert system solution [4]) and try to match these to the current interface. Nevertheless this finite list of objects in the knowledge base will exhaust very quickly in a relatively infinite space of previously unrecognized ones. So the ability to resemble and find similar objects is crucial for a softbot with a restricted knowledge base.

Matching a known set of objects consistently to the objects we recognize on the screen is the same problem as finding the maximal clique of consistent labels in region matching

problem experienced in computer vision, which is known to be NP-complete. There are different techniques applied to cope with the complexity of this constraint satisfaction problem, such as *relaxation labeling*. The idea is if we can represent the previously known objects as a set of constraints, we can use relaxation labeling for further relaxing these constraints to match newly recognized ones. This reduces the computational complexity.

However, assumptions, rather than specific constraints, are also extensively used when we speak of user interfaces and their usability. Usability is related to the effectiveness and the efficiency of a user interface with respect to user's expectations and reactions [13]. Usable interfaces characteristically promote ease of learning and user satisfaction with presumptions about user needs. It is not clear if we can efficiently represent all of the subjects pertaining to the interaction between humans and computers as a set of constraints; most likely we cannot.

Commonsense Reasoning in User Interfaces

Using former conventions and rules is a widely used practice in user interface design. These are carried to interfaces as assumptions, which cause a point of weakness in user interface softbots [2]. Most of the time, these assumptions are clear to a human but it is hard for a computer to grasp what is obvious to human perception. *Commonsense reasoning* is helpful in this sense because it concentrates on formalizing and finding computational models for sensible human reasoning. Adding this functionality to the communication medium between humans, computers and user interfaces is the focus of this research.

McCarthy [16] was the first to propose common sense reasoning ability as a key ingredient of AI. He claimed that a program that has common sense should be able to deduce the consequences from what is told and already known. Earlier approaches originate from applying qualitative reasoning to physical systems. De Kleer [7] introduced the notion of envisionment that refers to predicting and analyzing changes in qualitative states. According to the framework he gives [8], after the topology of the system is deduced from the physical state, it is combined with the current knowledge base and envisionments are created to produce behavioral predictions and causal explanations.

De Kleer [8] uses confluences to model the behaviour of devices in his physical system. A confluence is a qualitative differential equation and a widely used modeling tool for qualitative behaviour. For example, the qualitative behaviour of a rabbit population can be expressed by the confluence $dN = B - D$, where dN is the change in the number of rabbits, B is the birth, and D is the death rate. To verify the behaviour of a device, the set

of confluences that models it must be solved. Since each confluence acts as a constraint, the trouble resolves to a constraint satisfaction problem, which is very similar to the problem of matching the constraints of the recognizable objects to the descriptions of the objects we see on the screen.

Hayes' complaints of AI's previously narrow focus on toy-worlds and his suggestion of building a large-scale formalization like formalizing everyday knowledge about the physical world [12] turned the research direction to systems that reason in a physical domain.

More “expert” commonsense reasoning

As De Kleer [8] mentions, failure of expert systems stem from their narrow range of expertise and their inability to distinguish when a problem is outside of their know-how. In qualitative reasoning [9], *resolution* involves the depth of information detail in a qualitative representation of the knowledge. It is important to know how much information will suffice in order to produce valuable inferences for a commonsense reasoner to then predict how deep it should browse through knowledge. Most of the information that is easily accessible is sparse with low resolution results such as “the bird is flying south” rather than “the bird is flying 2 degrees west in the direction of the South Pole at 30 mph.”

Commonsense reasoning cannot be dependent on a single knowledge base or an expert system. Since it is highly context dependent, it is really “common” sense rather than any previously defined, controlled, specific, predictable sense. With the introduction of new evidence, we may change or abandon previous common knowledge. One should not be surprised to receive different answers to the same question either, since common sense is *variable* and *non-monotonic* in time.

So, in its general sense, it has a *dynamic* structure. Trying to come up with general problem solvers (advice taker [16], ThoughtTreasure [19], CYC [15]) has always been attractive for people who oversimplify the path to common sense intelligence as coming up with a knowledge base containing terms, concepts, facts, and rules of thumb that involve human common sense thought. This scheme can lead to expert systems in common sense world, yet the path to commonsense reasoners is much more arduous.

What is the big picture?

Given that we managed to design and program a commonsense reasoning system, what will this buy us in terms of user interfaces? First of all, as we mentioned in the introduction, one of our goals in user interface design is *reducing the asymmetry* that takes place in the communication between humans and computers. We believe that importing commonsense knowledge to both sides of the communication, but mostly to the computer's side, can decrease the asymmetry between the abilities of the two participants.

Secondly, we aim to *formalize the idioms and expressions*, the unspoken rules in the currently used user interfaces, with the assistance of commonsense knowledge. With this help, a softbot will be able to handle simple reasoning skills, such as when a window is opened on the screen, the objects behind it are not lost, just hidden, without the need to hardcode those naive information. Similarly, end-tools like mouse pointers, that aid users in user interfaces, can gain *interesting and useful abilities* by matching respective functionalities of {tool, object} tuples. We may have both micro and macro-tool based reasoners. For example, when a mouse pointer approaches a button (which has relevant functionalities in terms of the context) it can automatically click on it. If the action's consequences are irreversible, we will probably need confirmation as well. So, it will also help those user interfaces and softbots that have difficulties in *reaching the user's goals and intentions*.

Lastly, application of commonsense reasoning techniques will *increase the usability* of the user interface environment and help to create computers that are *more accessible* to those with disabilities and find accessing the interface is problematic. The outcome will be user satisfaction, friendly, and easy to use and learn computers across all age groups. We suspect that the future of user interfaces lies in those interfaces armed with tools caring commonsense knowledge applicable to daily life. Anthropomorphic user interfaces [\[25\]](#) and tools like HabilisDraw [\[3\]](#) will likely dominate the next generation interfaces.

Characteristics of the *common sense*

One can approach the problem of finding a model for commonsense by taking advantage of similarities [\[20\]](#); with the assumption that commonsense qualitative reasoning is a function that has components like analogical reasoning, qualitative and quantitative reasoning. However, commonsense reasoning must cover many more different approaches, as it is based on "propositional logic, the probability calculus and the concept of maximum entropy" [\[22\]](#), or on metaphor [\[5\]](#), or similarity matching [\[24\]](#).

Commonsense reasoning examples (water is wet, birds can fly, wood can burn, cars can move) convince us that it behaves a series of logical deductions where we just accept and believe in the transitions in between. This characteristic is named *jumping to conclusions* [21].

Most of the commonsense knowledge and reasonings are based on implicit assumptions and expectations, which are accepted to hold, but are constantly surrendered when new evidence contradicting those presumptions are found. In this sense, it is *non-monotonic* because when new facts are added, many deductions may no longer hold true [17].

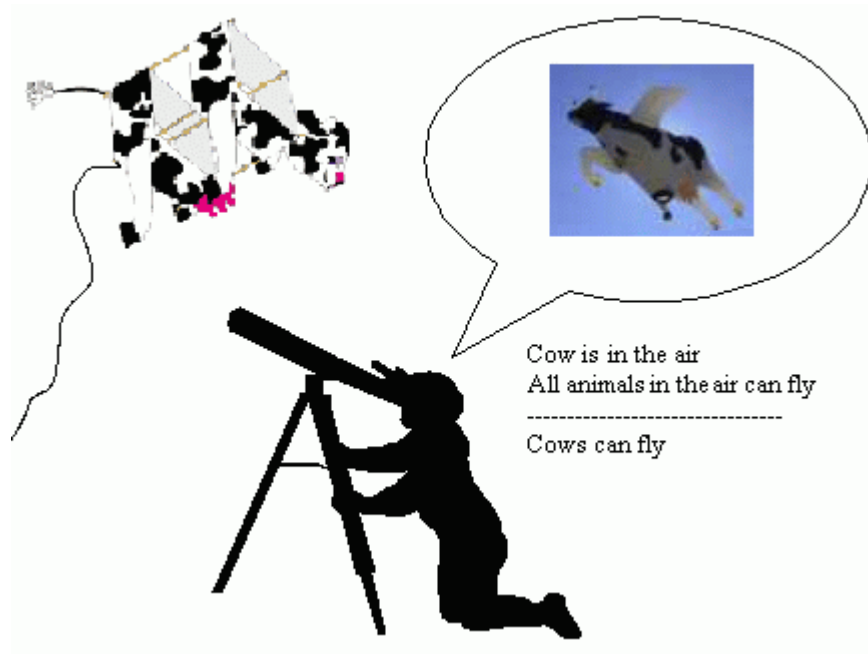


Figure – 1

At the same time, we should be careful about distinguishing between formerly known information and deducible information if we want to find new, previously unknown information in our inference mechanism. In this sense, deduction should be *monotonic*. However, humans sometimes err by forgetting to make this distinction. A child who sees a paper kite in the shape of a cow (Figure – 1) will probably think that cows can fly and assume that it is true that cows actually fly. If not told otherwise, the child will recall this information during the next encounter with cows. The same mistake is involuntarily experienced by adults in similar circumstances. To overcome such problems, first-order logic is a solution because of its monotonic formation. In the current CYC system, an extended version of first-order predicate calculus (FOPC), CycL [15] is used.

So, with commonsense reasoning, we are actually trying to reason monotonically with non-monotonic data. For instance, we can infer that cars move with the following sentence in first-order logic:

$$cars \textit{ HAVE } wheels \wedge \dots \wedge wheels \textit{ CAN } rotate \wedge \dots \wedge rotation \textit{ CAN } move \Rightarrow cars \textit{ CAN } move$$

We could re-solve this equation and deduct this long list of inferences every time we look for a vehicle to go somewhere (we also need to infer that we can move with cars while they move). But instead, in daily life, we just assume that this long chain of deductions is true and say what is “relevant” or “important” to us:

$$cars \textit{ CAN } move,$$

without questioning why and how cars can actually move. This *information hiding* is very important in commonsense reasoning and when we start constructing knowledge bases storing this type of knowledge, we will eventually realize that the transitions in between are not actually hidden, but rather lost. This provides evidence to the importance of monotonic format of commonsense reasoning. For example, once we come up with commonsense knowledge nodes like this, the nodes of information we care about, we cannot derive backwards:

$$cars \textit{ HAVE } windows \wedge cars \textit{ CAN } move \Rightarrow windows \textit{ CAN } move$$

However, FOPC is not enough since mathematical logic deals with *how people should think* rather than *how people actually do think* [18]. Also, humans don't utilize logic to store and represent their experiences [18], which pushes us forward to identify new formalisms for inference methods that currently use general logical deduction (modus ponens/tolens, universal and existential quantification) [15]. On the other hand, McCarthy also argues [18] that an intelligent logical program needs only monotonic and nonmonotonic reasoning abilities and mechanisms for entering and leaving contexts. The rest can be managed by specific functions and predicates.

On a different note, *spatial reasoning*, which is believed to have many qualitative aspects, is used in formalizing commonsense knowledge and it is claimed to be ubiquitous in human problem solving [10].

Relevancy Analysis

In the light of these characteristics, commonsense reasoning can be redefined as: “Retrieving only the relevant or sensible deductions that can serve as a springboard for future reasonings.” In data network analogue, we can claim that these points of deduction are the nodes where hot spots occur (and become the bottleneck of the system; we need to know this information to overcome bottlenecks) or in road network analogue, these are the roads where there is highly condensed traffic (so we need to know how to drive in those paths). So, a commonsense knowledge learner may need to conduct a relevancy analysis to find these important nodes of inference in relation to the context or problem domain.

Relevancy analysis lies at the core of the above-described commonsense knowledge learner. However, as McCarthy points out [18], formalizing relevancy is difficult.

Commonsense Reasoning in Physical Systems

Can we envision commonsense reasonings with a program? Can we reach the same reasoning abilities of humans? For instance, can we come up with commonsense reasoning about physical world like “Iron sinks in water” by using a program like NEWTON [7], which searches the knowledge representation and reasoning methodology for physical domains by using quantitative knowledge to clarify ambiguities? What will be the structure of the program that makes those envisions? Will it use abstract entities, principles and laws of physics for representing and reasoning [1]?

Commonsense reasoning in physical systems is different from reasoning with the laws of the nature since individuals usually have their own naive assumptions about the theories of nature. It happens to be the case that the guesses developed by different individuals are all different forms of the same central hypothesis, which is highly inconsistent with the basic principles of classical physics [1].

So, commonsense reasoning does not necessarily come up with the fundamental laws of nature that govern the bodies in the physical world; but rather, it helps us envision how different individuals would think and solve problems where these fundamental laws are not present.

Still, finding the mappings between empirical objects to abstract objects is necessary

since physical laws are stated over abstract entities and as state transformations [1]. Akman [1] also uses FOPC to represent these mappings in terms of predicates.

Since the principles that we deduct with our commonsense reasoning system will not necessarily be compliant with the principles that govern the physical world (or our world of context), it seems wise to divide the space into microworlds where each microworld satisfies its own consistency measures (each one is consistent in itself) and endogenous principles are drawn within each specific context. In this highly clustered space, one can still expect interesting reasonings applicable to the whole physical world (the world we get when all clusters are joined together). However, Akman [1] states that current envisioners lack the ability to switch between microworlds and macroworlds.

Methodology

The methodology for automating commonsense reasoning is given by Davis [6] as: *(i)* collect some examples of commonsense inference in a domain; *(ii)* recognize the general domain knowledge and the particular problem definition used; *(iii)* build up a formal language where this knowledge can be expressed; *(iv)* name the primitives of the language.

We believe that his scheme is helpful to researchers who are interested in further research within this promising area. We will, by following Davis' scheme, try to produce a minimal sized commonsense knowledge base in a physical microworld domain, since our aim is to focus on reasoning rather than proposing an alternative to current expert commonsense systems.



An albatros, flying

A penguin, "flying"

Figure – 2

There are still some questions left unanswered (same problematic issues with default reasoning [\[21\]](#)):

- What is a good set of commonsense knowledge to start with?
- What happens when the evidence matches the premises of two default rules with conflicting conclusions? Specificity preference tells us to prefer the rule that is more specific. CYC have extended truth-values by using two different true values: monotonic true (true with no exceptions) and default true (can have exceptions). For example, by default, birds can fly. A penguin is a bird but penguins cannot fly (Figure – 2). So, if we have a predicate like *flyPenguin* in our knowledge base, it has a default value of true but a monotonic value of false, which overrides the default value.
- How will the system clear some conclusions that contradict new findings? Truth maintenance systems can help to address these issues.

Conclusion

Commonsense reasoning is a promising technique that aims to represent how humans reason and think in a sensible way. While designing user interfaces, former conventions and rules are typically carried as assumptions. These assumptions are clear for human perception but to a computer, they may be the source of ambiguity that threatens the robustness of the system. Hence, user interfaces need a tool that will arm them with

reasoning and comprehension abilities relating to user actions, goals and assumptions.

Constructing a knowledge base containing terms, concepts, facts, and rules of thumb involving human common sense thought may suffice for expert systems in the common sense world, but building a commonsense reasoner appears to be a harder task. The non-monotonic structure of commonsense knowledge, the need for monotonic reasoning with this data, relevancy analysis required for creating these key data nodes and constraint satisfaction problems increase the complexity of any commonsense reasoning system. Conflicting with Hayes' suggestions, focusing on a minimal sized commonsense knowledge base in a physical microworld domain can postpone some of these issues that need to be addressed.

Commonsense reasoning seems to have a lot to offer to user interfaces, especially in bridging the gap between the asymmetric abilities of the two counterparts, computers and humans, and the communication taking place in this domain. Additions to such a system of user interfaces will likely provide better representation of assumptions and unspoken rules, increased abilities of tools used, and improved usability and accessibility of the environment where computers and humans communicate more efficiently.

There are still some problems to be solved before building functional and practical commonsense reasoners. However, in the future, we envision human-computer interaction media will be armed with tools that have commonsense knowledge (as in daily life) that will likely dominate the next generation interfaces.

References

1. Akman, V., and Ten Hagen, P. J. W. 1989. The power of physical representations. *AI Magazine* 10(3): 49-65.
2. Amant, R. St. and Dudani, A. An environment for user interface softbots, *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI)*. 2002.
3. Amant, R. St. and Horton, T. E. A tool-based interactive drawing environment. *ACM Conference on Human Factors in Computing Systems (CHI) Extended Abstracts*. 2002.
4. Amant, R. St. and Riedl, M. O. Toward automated exploration of interactive systems, *Intelligent User Interfaces 2002*.

5. Barnden, J. A. and Lee, M. G. An implemented context system that combines belief reasoning, metaphor-based reasoning and uncertainty handling, Available:
<http://www.cs.bham.ac.uk/~jab/ATT-Meta/Papers/context99.pdf>
6. Davis, E. *Representation of Commonsense Knowledge*. Morgan Kaufmann Publishing Co., 1990.
7. De Kleer, J. Qualitative and quantitative knowledge in classical mechanics, Technical Report AI-TR-352, AI Laboratory, MIT
8. De Kleer, J. and Brown, JS. A qualitative physics based on confluences. *Artificial Intelligence* 24 (1984) 7-83.
9. Forbus, K. D. Qualitative reasoning, In A.B. Tucker, editor, *The Computer Science and Engineering Handbook*, pages 715--733. CRC Press, 1996.
10. Forbus, K. D., Nielsen, P. and Faltings, B. Qualitative spatial reasoning: The clock project, *Artificial Intelligence* 51, 1991, 417-471.
11. Gentner, D. and Nielsen, J. The anti-Mac interface, *Communications of the ACM*, 39(8), 1996, 70-82.
12. Hayes, P. The naive physics manifesto, *Expert Systems in Microelectronics age*, D. Ritchie Ed., Edinburgh University Press, 1978, 242-270
13. Hix, D. and Hartson, H. R. *Developing User Interfaces, Ensuring Usability Through Product and Process*, John Wiley and Sons, 1993, 221-222
14. Kant, I. *Critique of Pure Reason*, trans. by Kemp Smith, The Macmillan Press Ltd., London, 1950, 300
15. Lenat, D. B. From 2001 to 2001: Common Sense and the Mind of HAL, *HAL's Legacy: 2001's Computer as Dream and Reality*, ed. Stork, D. G. MIT Press, 1996 Available:
<http://www.cyc.com/halslegacy.html>
16. McCarthy, J. Programs with common sense, *Teddington Conference on the Mechanization of*

Thought Processes, 1958

17. McCarthy, J. Concepts of logical AI, 1999. Available: <http://www-formal.stanford.edu/jmc/concepts-ai/concepts-ai.html>
18. McCarthy, J. From here to human-level AI, *Knowledge Representation Proceedings*, 1996. Available: <http://www-formal.stanford.edu/jmc/human.html>
19. Mueller, E. T. *Natural language processing with ThoughtTreasure*, Signiform, New York, 1998. Available: <http://www.signiform.com/tt/book/>
20. Paritosh, P. K. and Forbus, K. D. Common sense on the envelope, *Proceedings of the Fifteenth International Workshop on Qualitative Reasoning*, San Antonio, 2001.
21. Russell, S. and Norvig, P. *Artificial Intelligence: A Modern Approach*, Prentice Hall, 1995, 458-460.
22. Schramm, M. and Ertel, W. Reasoning with probabilities and maximum entropy: the system PIT and its application in LEXMED, *Symposium on Operations Research*, 1999
23. Suchman, L. A. *Plans and Situated Actions, The problem of human-machine interaction*, Cambridge University Press, 1987, 180-181
24. Sun, R. Robust Reasoning: Integrating rule-based and similarity-based reasoning, *Artificial Intelligence* 75, 1995, 241-295.
25. Xiao, J., Stasko, J. and Catrambone, R. Anthropomorphic User Interfaces, Available: http://www.cc.gatech.edu/gvu/ii/myagent/index_old.htm

About the Author

Ergun M. Bicici is a graduate student in the Intelligent Interfaces, Multimedia, and Graphics Lab in Computer Science Department of North Carolina State University. His research interests include human-computer interaction, intelligent interfaces, computer vision, robotics and commonsense reasoning. He can be reached at: embicici@ncsu.edu.

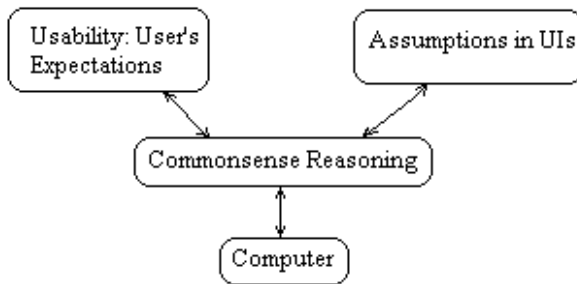
Acknowledgments

I would like to thank my advisor, Dr. Robert St. Amant, for helping me organize my thoughts, to Dr. Matthias Stallmann about his comments regarding the structure, to my editor Tony Hall for his inspiring and supportive remarks and to ACM Crossroads Editorial Board for their review.

Notes:

* Prolegomenon: A formal essay or critical discussion serving to introduce and interpret an extended work. Neuter present passive participle of *prolegein* to say beforehand, from *pro-* before + *legein* to say. (www.m-w.com)

Attachment



assumptions.gif (2k)
Original Size: 310 x 163