

Seeking Explanations in Soar

Steven R. Haynes
School of Information Sciences & Technology
Penn State University
shaynes@ist.psu.edu

23rd Soar Workshop, June 27, 2003

This project is supported by the US Office of Naval Research
Award number N00014-02-1-0021

Soar Explanations

- Goal: A content theory for Soar explanations
- First, define what we mean by explanation
 - Hard, little theoretical agreement
- What does it mean to explain a Soar model?
 - Easier, but still hard
 - Analytic, theoretical approach to what might be explanatory of a Soar model
 - Then, empirical approach to what questions Soar users, modelers, and developers ask (Councill, et al., 2003)
 - Then, design-build-evaluate iterations to refine

Explanation Seeking Questions

- Two forms of explanation seeking questions for systems:
- Operational explanations – directly support using the system
 - Do not fit *why* form often associated with explanation seeking questions but found to be explanatory in empirical studies
- *Why* explanations
 - Seeking reasons, causes; more interesting but harder

Operational Explanations

- What?
 - Identity: ontological
 - Definition: description, attributes
 - Relations: how are entities and events linked
- How?
 - How do I use it?
 - How does it work?
- Where? When? (Who is like What)
- But note that these may often be re-framed as why questions, i.e., Why is it there? Why do I use it that way? (Draper, 1988)

Why Explanations

- Baseline: Deductive-Nomological Explanation
 - Explanations derive from laws (Hempel, 1965)
- Probabilistic Explanation
 - Similar to D-N but laws are statistical (e.g., Salmon, 1984)
- Functional Explanation
 - Entity or event explained by its purpose (e.g., Cummins, 1975)
- Structural Explanation
 - Existing structure imposes constraints (e.g., Little, 1991)
- Bounded-Rational Choice Explanation
 - Agency, bounded-rational actors make choices (e.g., Elster, 1985)
- Pragmatic Explanation (van Fraassen, 1988)
 - What counts as explanatory is purely interest-relative

Explanation Seeking in Soar

- Analysis of transcripts from usability study of the Soar Situation Awareness Panel (SAP) (Avraamides & Ritter, 2002)
- Novices and experts reviewing usability of the SAP
- Similar to Councill, et al., 2003, but uses explanation framework, all study subjects
- No measure (yet) of inter-rater reliability

Classes of Soar Explanation Seeking

- N = 236
- Operational Explanations
 - 88% (207)
- Why Explanations
 - 12% (29)

Operational Explanations

	Approx. %	Count
• N = 207		
• What		
– Identity	20%	(41)
– Definition	38%	(79)
– Relation	3%	(6)
– Event (what happened?)	3%	(6)
• How		
– How do I use it?	12%	(25)
– How does it work?	21%	(44)
• Where, When		
– When	0.5%	(1)
– Where	2.5%	(5)

Why? Explanations

- | • N = 29 | Approx. % | Count |
|--|-----------|-------|
| • Functional Explanation | 52% | (15) |
| <i>“...okay, and what is the ultimate goal of this?”</i> | | |
| • Pragmatic Explanation | 31% | (9) |
| (contrast classes) | | |
| <i>“why didn't he do that?”</i> | | |
| • Why – Exploratory | 3% | (1) |
| <i>“is this gulf-like for any particular reason...?”</i> | | |
| • Why – Unclassified | 14% | (4) |
| <i>“why the heck is this guy doing something stupid...?”</i> | | |

Validity/Reliability

- Data from a usability study, not a study focused on explanation requirements *per se*
- Novice SAP users may have different requirements than expert or experienced users

Implications for HLBRL

- Beginning to understand what Soar users (and users of other cog. architectures) want to know:
 - New users may be satisfied with relatively superficial what and how (operational) explanations
 - Relatively simple entity and event lists (state displays) and definitions may cover many explanation requests
 - Expert users may require deeper, why explanations
- Functional explanations require designer intent
- Contrast classes require design rationale

Open Questions

- What explanation content is available in a cognitive model or in the Soar architecture?
 - Addressing this with analytic decomposition of the Soar architecture and model structure
- To what extent does a Soar explanation require external knowledge, i.e., domain or contextual knowledge not resident in the model or architecture?
- Where does the external knowledge required to service these questions come from, if not from the model or architecture?

Discussion

References

- Avraamides, M. N. & Ritter, F. E. (2002). Using multidisciplinary expert evaluations to test and improve model interfaces. *Proceedings of the 11th Computer Generated Forces Conference*, 553-562. Orlando, FL: U. of Central Florida.
- Councill, I. G., Haynes, S. R., & Ritter, F. E. (2003). Explaining Soar: Analysis of existing tools and user information requirements. In *Proceedings of the Fifth International Conference on Cognitive Modeling*. 63-68. Bamberg, Germany: Universitats-Verlag Bamberg.
- Cummins, R. (1975). Functional Analysis. *The Journal of Philosophy*, 72, 741-765.
- Draper, S. W. (1988). What's going on in everyday explanation? In C. Antaki (Ed.), *Analyzing Everyday Explanation*, London: Sage Publications, pp. 15-31.
- Elster, J. (1985). The Nature and Scope of Rational-Choice Explanation. In E. LePore and B. McLaughlin (Eds.), *Actions and Events: Perspectives on Donald Davidson* (pp. 60-72). Oxford, UK: Blackwell
- Hempel, C.G. (1965). Aspects of Scientific Explanation. In *Aspects of Scientific Explanation and Other Essays*. New York: Free Press.
- Little, D. (1991). *Varieties of Social Explanation : An Introduction to the Philosophy of Social Science*. Boulder, CO: Westview Press.
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
- van Fraassen, B. C. (1988). The Pragmatic Theory of Explanation. In J. Pitt (Ed.), *Theories of Explanation* (pp. 135-155). Oxford, UK: Oxford University Press.