A Tale of Two Tasks: Data Control and Modeling

Wayne D. Gray

Rensselaer | Cognitive Science

Cog Works

Beyond OpenData Sharing: Making Sense of Massive Data Sets

(Isn't that what MacSHAPA was really about???)

Beyond Open Data Sharing

- High bandwidth data collection with well formatted records, easy to reuse documentation, and ability to address new questions after the data is collected
- Tools that will aggregate sampled data to form meaningful units at different levels of analysis
- Visualizing and exploring data in terms of sequence, co-occurrence, and other patterns
- Newell's Dream: Automated or semi-automated protocol analyses, which enable theory-based parsing of log files to form runnable cognitive models

Then & Now

- MacSHAPA
  - MacSHAPA (1995's) - Submarine Commanders: managing complexity of verbal and action protocols
  - MacSHAPA Cognitive Metrics Profiling
- Action Protocol Tracer
  - Finite state grammars for pattern recognition in action protocol data
- SANLab
  - SANLab - tool for Stochastic Analytic Network modeling +++
  - Newell's Dream: Automating production of cognitive models from behaviorial/action protocol analysis

Experience with MacSHAPA

- Submariners (≈ 1995 to 2000)
- Tool for examining output of computational cognitive modeling

Project Nemo, or, Subgoaling Submariners

A Joint University–N Collaboration

Universit Ph.D.

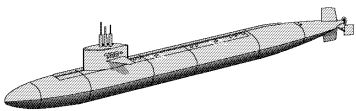
Hum ition

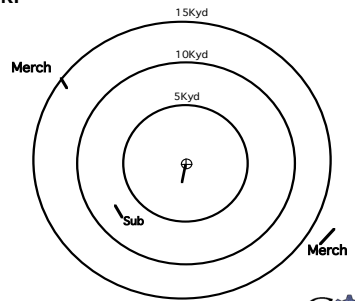
≈1996

schenbaum, Ph.D.

Warfare Center Division

Newport, RI





Rensselaer | Cognitive Science

CogWorks

Seven Phased Approach

- Phase 1: Data Collection usin[illegible]ulation (at NUWC) --**>COMPLETED<--**
- Phase 2: Encodin[illegible] . rotocols from Phase 1 **-->COMPLETED**
- Phase 3: Developme[illegible] .ied simulation (scaled world) --**>COMPLETED<--**
- Phase 4: Development of preliminary computational cognitive models ****CURRENT****
- Phase 5: Data Collection using scaled world ****CURRENT****
- Phase 6: Analysis of data and refinement of models
- Phase 7: Modifications of suite of models and scaled world as deliverables

≈1996

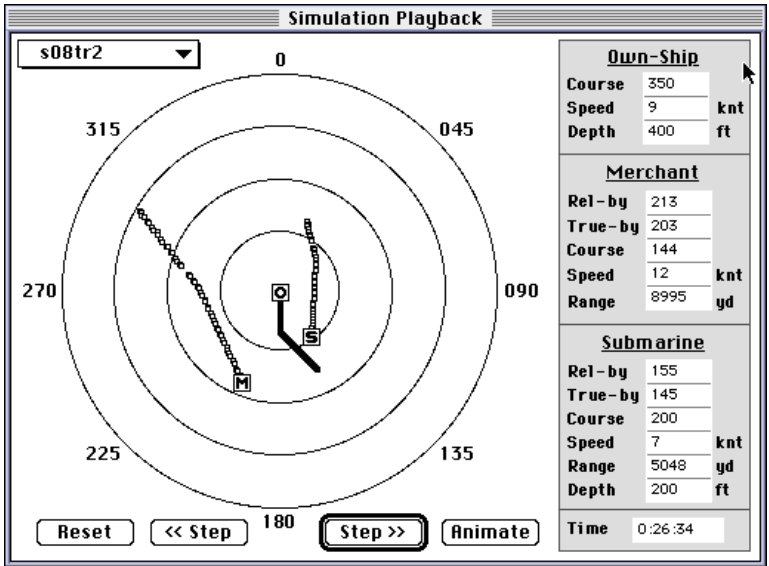
Table 2: Segment Shown in Table 1 Following Resegmentation and Encoding of Goals by the Experimenters

Time	L1	L2	L3	Operator	Info-Source	Ship	Attribute	Value	Duration
	DETECT-SUB								
62.428				DISPLAY-NAV	SONAR-NB-TOWED				
63.98				QUERY	NBT-WATERFALL				
66.82				RECEIVE	NBT-WATERFALL	SUB	ON-SONAR	NO	4.221
	POP								
	LOCALIZE-MERC								
		SET-TRACKER							
67.02				SET-TRACKER	SONAR-NB-TOWED	MERC			
68.201				RECEIVE	NBT-WATERFALL	MERC	ON-SONAR	YES	4.221
68.201				RECEIVE	NBT-WATERFALL	MERC	BEARING	BEAM	4.221
68.201				RECEIVE	NBT-WATERFALL	MERC	TRACKING	YES	4.221
	POP								
	DETERMINE-CONICAL-ANGLE								
70.063				QUERY	NBT-CONICAL-ANGLE-FIELD	MERC	CONICAL-ANGLE		
70.724				RECEIVE	NBT-CONICAL-ANGLE-FIELD	MERC	CONICAL-ANGLE	82.15	0.661
	POP								
	DETERMINE-BY								
71.111				QUERY	NBT-BEARING-FIELD	MERC	BEARING		
72.088				RECEIVE	NBT-BEARING-FIELD	MERC	BEARING	152_OR_314	0.977
	POP								
	DETERMINE-SNR								
72.393				QUERY	NBT-SNR-FIELD	MERC	SNR		
72.987				RECEIVE	NBT-SNR-FIELD	MERC	SNR	6.63	0.594
		POP							
	POP								
	EVALUATE-ARRAY-STATUS								
74.635				QUERY	NBT-ARRAY-STATUS-FIELD	OS	ARRAY		
75.601				RECEIVE	NBT-ARRAY-STATUS-FIELD	OS	ARRAY	STABLE	0.966
	POP								

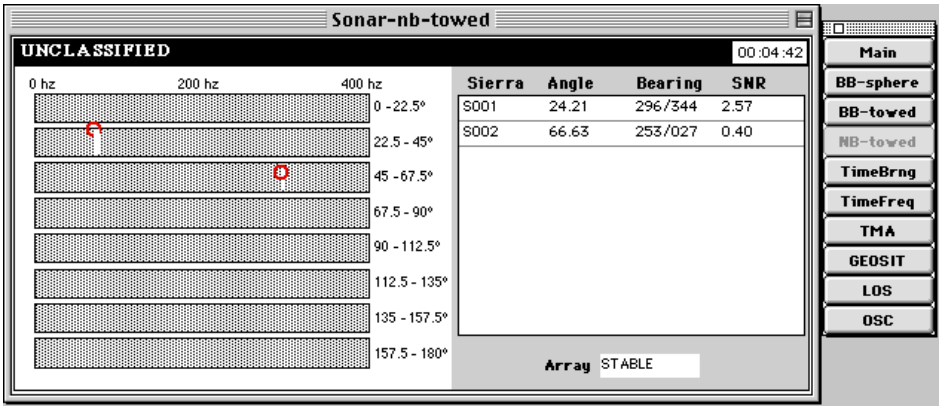
Note: The headings are the same as in Table 1 with the addition of three fields for goals and subgoals: levels 1 (L1), 2 (L2), and 3 (L3). No L3 goals are encoded in this segment.

Phase 2: Tool Development -- To facilitate

Encoding of Data we developed a Tool to playback the files collected at NUWC



Phase 3: NED



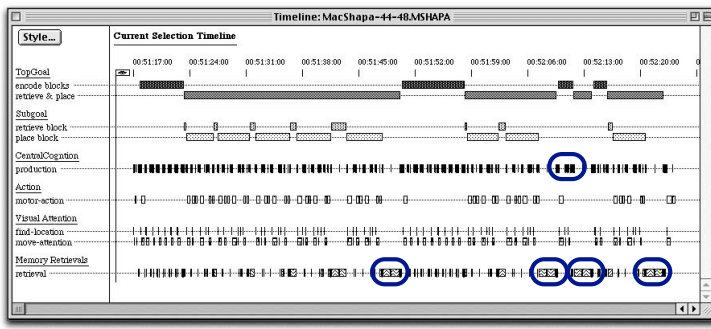
One of Ned's 10 displays that AOs use for situation assessment.
In data-collection mode, all AO interactions with Ned are recorded and time stamped at 60hz (16.67 msec); along with the current state of the simulation (truth!)

Experience with MacSHAPA

- Submariners (≈ 1995 to 2000)
- **Tool for examining output of computational cognitive modeling**

Visualizing the Output of a Process Model (ACT-R)

- ? We were asking whether we could use this approach to develop a predictions of cognitive workload by identifying tasks or subtasks where the resource demands are excessive
- Especially places where the using the system (i.e., the structure of the interactive system) consumes resources required for doing the task



Then & Now

- MacSHAPA
  - MacSHAPA (1995's) - Submarine Commanders: managing complexity of verbal and action protocols
  - MacSHAPA Cognitive Metrics Profiling
- Action Protocol Tracer
  - Finite state grammars for pattern recognition in action protocol data
- SANLab
  - SANLab - tool for Stochastic Analytic Network modeling +++
  - Automating production of cognitive models from behaviorial/ action protocol analysis

Our Focus was on Discrete Action Protocols

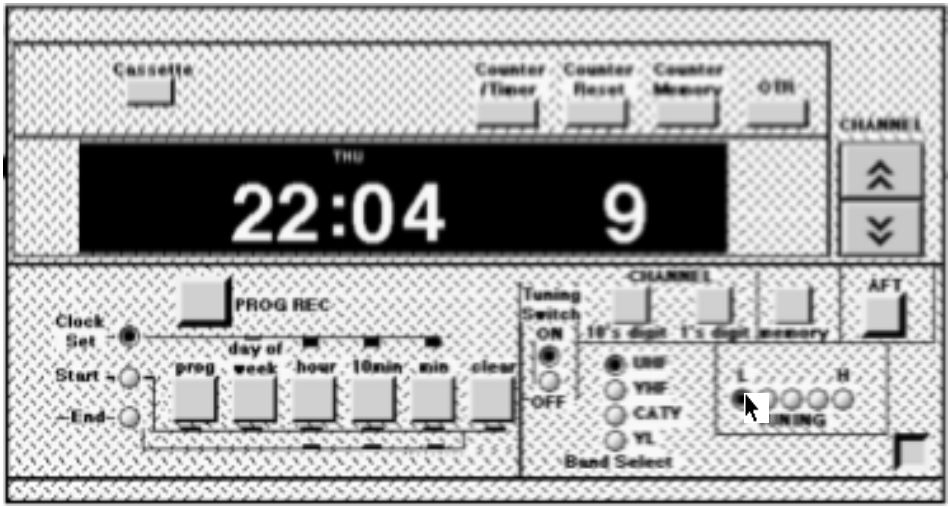
- E.g. Mouse clicks, key presses collected by a computer system
- Characteristics:
  - A large volume of protocols can be easily collected
  - High temporal resolution (e.g. 16.67 msec)
  - Constrained and easy to interpret (compared to verbal protocols)
  - Easy to aggregate across subjects

But, approach could be applied to any data process data that could be encoded in SHAPA spreadsheets

Fu, W.-T. (2001). ACT-PRO: Action protocol tracer -- a tool for analyzing discrete action protocols. *Behavior Research Methods, Instruments, & Computers, 33*(2), 149–158.

Action protocol analysis

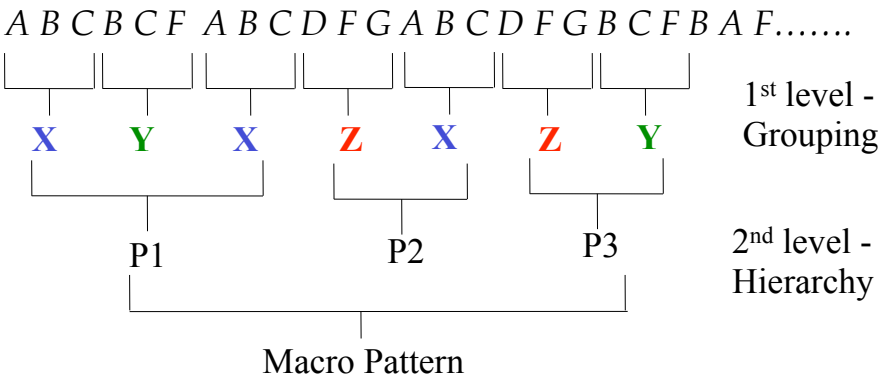
- Two approaches to do the analysis:
  - Exploratory: searching for possible patterns in the protocols
  - Confirmatory: Looking for evidence supporting the researcher's theory
  - Both approaches require some kind of pattern matching to patterns generated by the researcher
- Automatic (or semi-automatic) protocol analyzer
  - Reduce effort
  - Increase objectivity


Cassette
Counter /Timer
Counter Reset
Counter Memory
CLOCK
CHANNEL
THU
22:04
9
PROG REC
Clock Set
Start
End
prog
day of week
hour
10min
min
clear
Tuning Switch
ON
OFF
CHANNEL
10's digit
1's digit
memory
AFT
UHF
VHF
CATV
YL
Band Select
L
H
TUNING

CURTIME	WINDOW	BUTTON	REC-MODE	PR	CLOCK	PROGNUM	CHANNEL	DOW	WEEK?	POWER?	KP-HOUR	KP-10MIN	KP-MIN
141802	SHOW-INFO-WIN	Start Trial	CLICKON	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
141802	VCR-WINDOW	Start Trial	CLOCKSET-MODE	OFF	15:00	NONE	NIL	WED	NIL	ON	15	0	0
141865	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
141924	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
142047	VCR-WINDOW	START-RADIO-BUT	START-MODE	OFF	E:EE	1	15	WED	NIL	ON	E	E	E
142097	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	WED	NIL	ON	15	0	0
142202	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	THU	NIL	ON	15	0	0
142213	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	FRI	NIL	ON	15	0	0
142230	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	SAT	NIL	ON	15	0	0
142241	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	SUN	NIL	ON	15	0	0
142271	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	MON	NIL	ON	15	0	0
142307	VCR-WINDOW	DAY-OF-WEEK	START-MODE	OFF	15:00	1	15	TUE	NIL	ON	15	0	0
142453	VCR-WINDOW	HOUR	START-MODE	OFF	16:00	1	15	TUE	NIL	ON	16	0	0
142483	VCR-WINDOW	HOUR	START-MODE	OFF	17:00	1	15	TUE	NIL	ON	17	0	0
142512	VCR-WINDOW	HOUR	START-MODE	OFF	18:00	1	15	TUE	NIL	ON	18	0	0
142529	VCR-WINDOW	HOUR	START-MODE	OFF	19:00	1	15	TUE	NIL	ON	19	0	0
142550	VCR-WINDOW	HOUR	START-MODE	OFF	20:00	1	15	TUE	NIL	ON	20	0	0
142579	VCR-WINDOW	HOUR	START-MODE	OFF	21:00	1	15	TUE	NIL	ON	21	0	0
142621	VCR-WINDOW	HOUR	START-MODE	OFF	22:00	1	15	TUE	NIL	ON	22	0	0
142650	VCR-WINDOW	HOUR	START-MODE	OFF	23:00	1	15	TUE	NIL	ON	23	0	0
142730	VCR-WINDOW	10MIN	START-MODE	OFF	23:10	1	15	TUE	NIL	ON	23	1	0
142824	VCR-WINDOW	10MIN	START-MODE	OFF	23:20	1	15	TUE	NIL	ON	23	2	0
142853	VCR-WINDOW	10MIN	START-MODE	OFF	23:30	1	15	TUE	NIL	ON	23	3	0
142909	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
142980	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	14	TUE	NIL	ON	23	3	0
142993	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	13	TUE	NIL	ON	23	3	0
143066	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	12	TUE	NIL	ON	23	3	0
143076	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	11	TUE	NIL	ON	23	3	0
143084	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	10	TUE	NIL	ON	23	3	0
143092	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	9	TUE	NIL	ON	23	3	0
143099	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	8	TUE	NIL	ON	23	3	0
143111	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	7	TUE	NIL	ON	23	3	0
143132	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	6	TUE	NIL	ON	23	3	0
143141	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	5	TUE	NIL	ON	23	3	0
143150	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	4	TUE	NIL	ON	23	3	0
143178	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	3	TUE	NIL	ON	23	3	0
143194	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	2	TUE	NIL	ON	23	3	0
143220	CCL:VCR-WINDOW	CH-DOWN	START-MODE	OFF	23:30	1	1	TUE	NIL	ON	23	3	0
143255	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
143324	VCR-WINDOW	END-RADIO-BUT	END-MODE	OFF	E:EE	1	15	NIL	NIL	ON	E	E	E
143343	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
143371	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
143556	VCR-WINDOW	HOUR	END-MODE	OFF	23:30	1	1	NIL	NIL	ON	23	3	0
143707	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
143755	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
143836	VCR-WINDOW	CS-RADIO-BUT	CLOCKSET-MODE	OFF	15:00	NONE	15	WED	NIL	ON	15	0	0
143902	VCR-WINDOW	PROGREC	CLOCKSET-MODE	ON	15:00	NONE	NIL	WED	NIL	ON	15	0	0
143943	VCR-WINDOW	NIL	VCR-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
144008	SHOW-INFO-WIN	Stop Trial	CLICKON	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
144008	ERROR-TRIAL	RESET	SHOW-ET	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL
144028	SHOW-INFO-WIN	NIL	SW-LEAVE	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL

Finding patterns in data

- A sequential stream of discrete action protocol



Structure of ACT-PRO

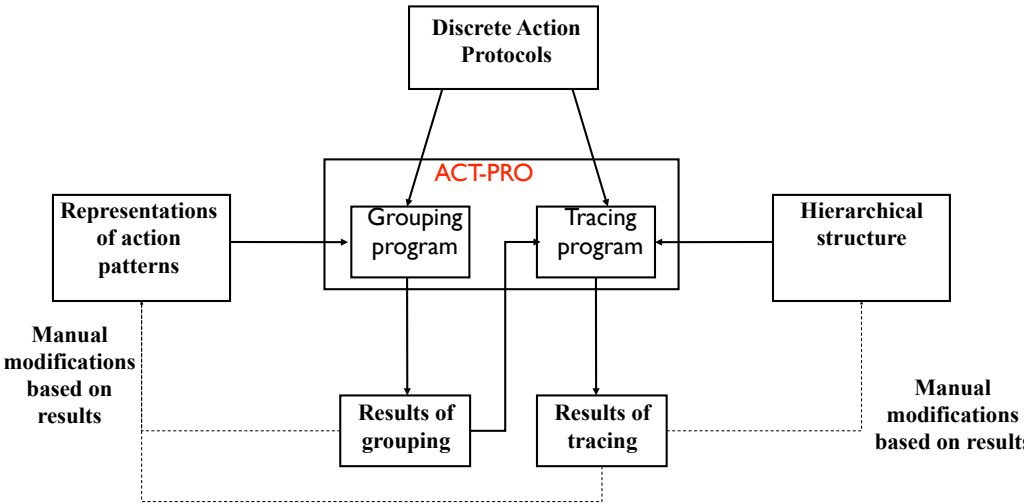


Table 2
Grammar That Captures Variations of the Action Sequence Formed by Pressing the Buttons Channel, Up-Arrow, Down-Arrow, and Enter

SET-CHANNEL: [Object1][Object2][Object3]	(1)
[Object1] → *channel*	(2)
[Object2] → *up-arrow*	(3)
[Object2] → *down-arrow*	(4)
[Object2] → *up-arrow* [Object2]	(5)
[Object2] → *down-arrow* [Object2]	(6)
[Object2] → *enter* [Object2]	(7)
[Object3] → *enter*	(8)

Table 3
An Example of the Trace and Validation Results of Using the Hierarchical Goal Structure of Figure 1

Trace	Validation results
Push goal: PROGRAM-SHOW	Push goal match
Push goal: SET-START-TIME	Push goal match
Push goal: SET-START-HOUR	Push goal match
Action: *start-hour*	
Action: *start-hour*	
Pop goal: SET-START-HOUR	Pop goal match
Push goal: SET-START-10MIN	Push goal match
Action: *start-10min*	
Pop goal: SET-START-10MIN	Pop goal match
Pop goal: SET-START-TIME	*Pop goal mismatch*
Push goal: SET-CHANNEL	Push goal match
Action: *channel*	
Pop goal: SET-CHANNEL	Pop goal match
Push goal: SET-START-TIME	Push goal match
Push goal: SET-START-MIN	Push goal match
Action: *start-min*	
...	

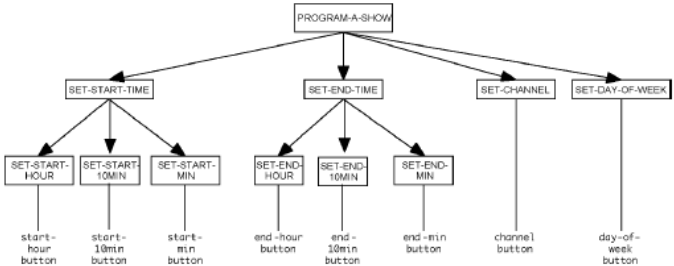
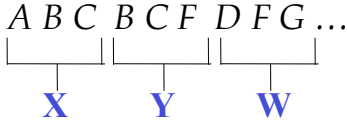


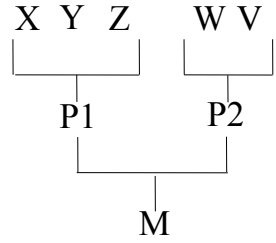
Figure 1. The simplified task-relevant hierarchical goal structure of a VCR interface (Gray, 2000; Gray & Fu, 2000).

Table 3
An Example of the Trace and Validation Results of Using the Hierarchical Goal Structure of Figure 1

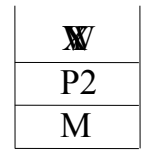
Trace	Validation results
Push goal: PROGRAM-SHOW	Push goal match
Push goal: SET-START-TIME	Push goal match
Push goal: SET-START-HOUR	Push goal match
Action: *start-hour*	
Action: *start-hour*	
Pop goal: SET-START-HOUR	Pop goal match
Push goal: SET-START-10MIN	Push goal match
Action: *start-10min*	
Pop goal: SET-START-10MIN	Pop goal match
Pop goal: SET-START-TIME	*Pop goal mismatch*
Push goal: SET-CHANNEL	Push goal match
Action: *channel*	
Pop goal: SET-CHANNEL	Pop goal match
Push goal: SET-START-TIME	Push goal match
Push goal: SET-START-MIN	Push goal match
Action: *start-min*	
...	

Output Trace and Goodness-of-fit





Trace	Validation Results
Push M	Push match
Push P1	Push match
Push X	Push match
Actions: A B C	
Pop X	Pop match
Push Y	Push match
Actions: B C F	
Pop Y	Pop match
Pop P1	Pop mismatch
Push P2	Push match
Push W	Push match
Actions: D F G	
Pop W	Pop match
⋮	



Results

- 64 subjects, 1,228 trials, 51,232 actions
- 8 grammars were constructed for each interface, each representing a structural pattern (a strategy)
- Worst-fitting trial: 81.1%; best-fitting trial: 100% Average: 95.1% of the actions were captured by the grammars
- By inspecting the results, we found change of strategies in different interfaces

- Two different hierarchies were used in the two interfaces
- We also found differences in the higher-level patterns in the two interfaces
- 15,245 higher-level patterns are parsed
- 464 (3%) of the patterns were identified as mismatches between the data and the hierarchy

AT:ST Ratio – Analysis Time to Sequence Time

- Pre-Action Protocol Tracer
  - Gray (2000) estimated as 100:1
  - Analyzed data from 9 Ss, ≈ 72 trials
- With the Action Protocol Tracer
  - For the 3 data sets described in the Fu 2001 the building of grammars, on average, took the researchers 2–3 h, and the average running time was about 1 h.
  - 1:10

Fu, W.-T. (2001). ACT-PRO: Action protocol tracer -- a tool for analyzing discrete action protocols. *Behavior Research Methods, Instruments, & Computers, 33*(2), 149–158.

Then & Now

- MacSHAPA
  - MacSHAPA (1995's) - Submarine Commanders: managing complexity of verbal and action protocols
  - MacSHAPA Cognitive Metrics Profiling
- Action Protocol Tracer
  - Finite state grammars for pattern recognition in action protocol data
- SANLab-CM
  - SANLab - tool for Stochastic Analytic Network Cognitive Modeling
  - Automating production of cognitive models from behaviorial/ action protocol analysis

SANLab-CM

- An extension of the tools used by Gray & John (1993) and Gray & Boehm-Davis (2000) & other studies
- Schweickert in numerous studies

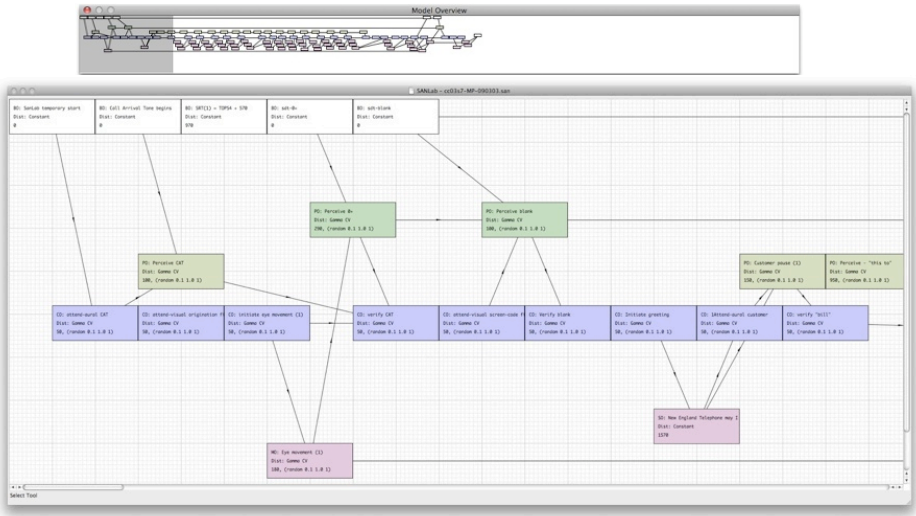
Schweickert, R., Fisher, D. L., & Proctor, R. W. (2003). Steps toward building mathematical and computer models from cognitive task analyses. *Human Factors, 45*(1), 77–103.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology, 18*(2), 105–139.

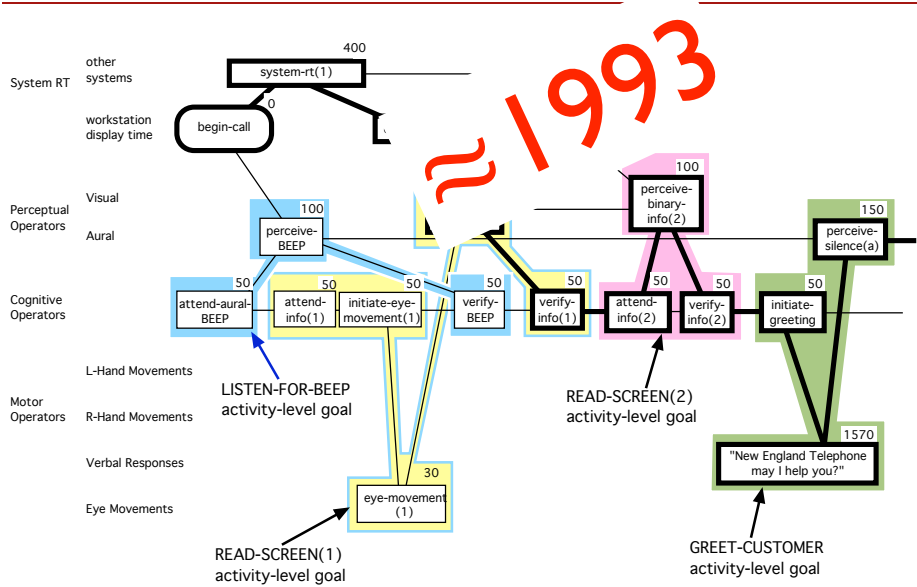
Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds Matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied, 6*(4), 322–335.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world performance. *Human-Computer Interaction, 8*(3), 237–309.

Model Window and Model Overview Window



Telephone Operator Workstation
CPM-GOMS Level

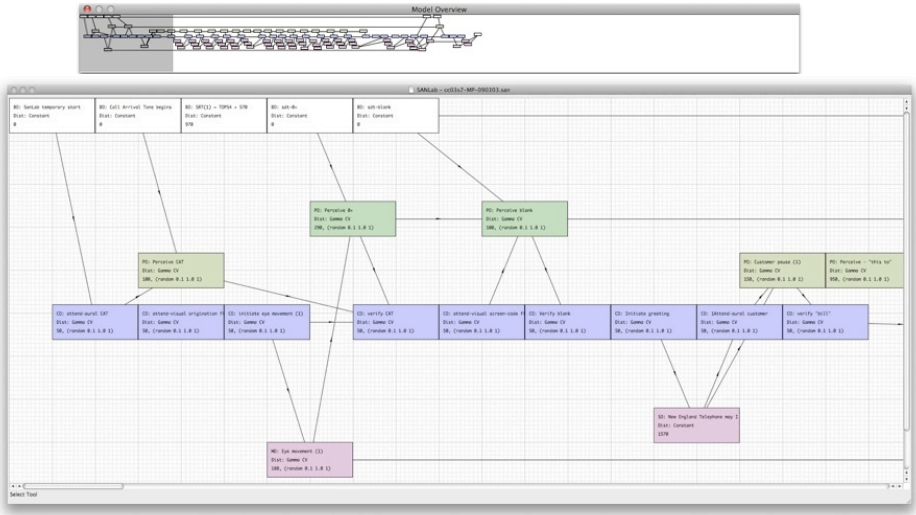


SANLab-CM

- Stochastic Activity Network Laboratory for Cognitive Modeling
- Idea inspiring SANLab-CM
  - Cognitive, perceptual, and motor processes are inherently variable
  - This variability may result in changes in workload even when load conditions are constant
- Hence, SANLab-CM is a tool for analyzing and predicting variability with and without extra workload

Patton, E. W., Gray, W. D., & Schoelles, M. J. (2009). SANLab-CM: The Stochastic Activity Network Laboratory for Cognitive Modeling. *53rd annual meeting of the Human Factors and Ergonomics Society*. San Antonio, TX: Human Factors and Ergonomics Society.

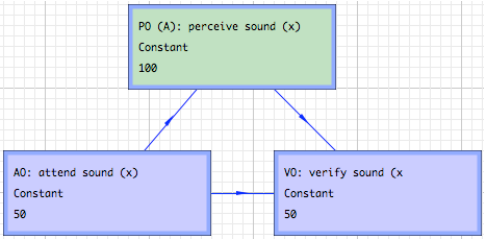
Model Window and Model Overview Window



Example 1: Constructing a very simple CPM-GOMS model in SANLab

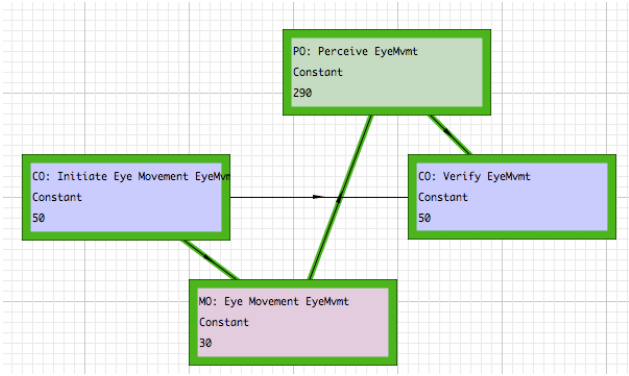
- Parts
- Interleaving
- Stochasticity
- Comparison of very simple models

Building a Preliminary CPM-GOMS Model
CPM-GOMS Templates

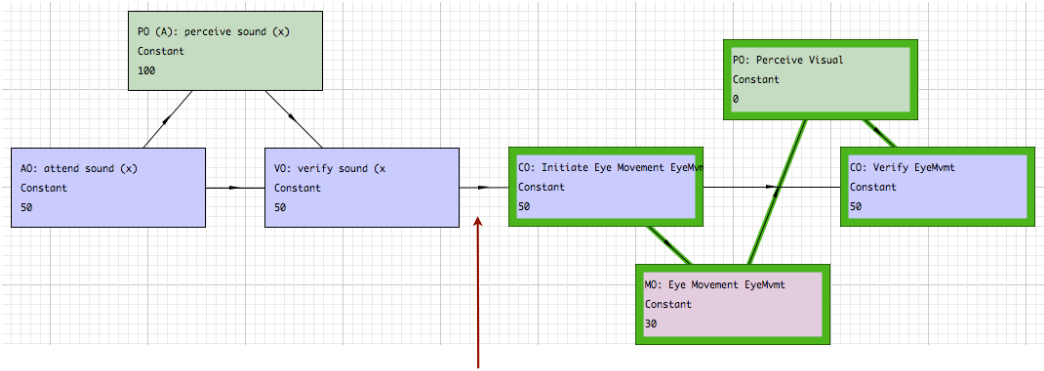


Perceive Simple Sound

Perceive Visual Information With Eye Movement



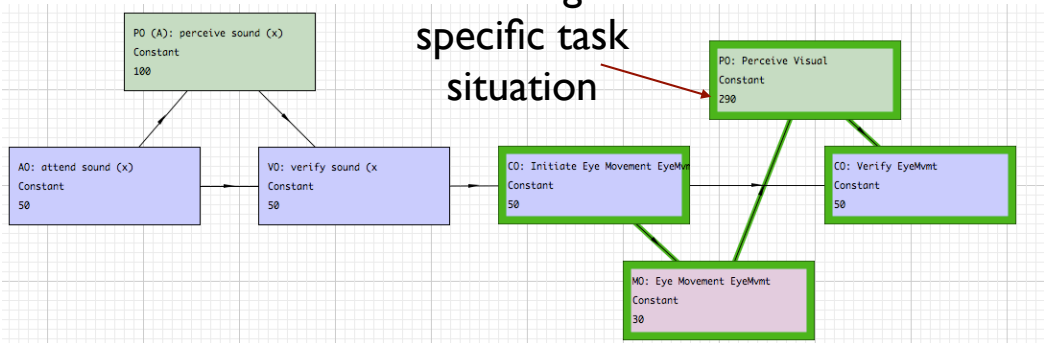
Building a Preliminary CPM-GOMS Model
Cut & Paste & String Together



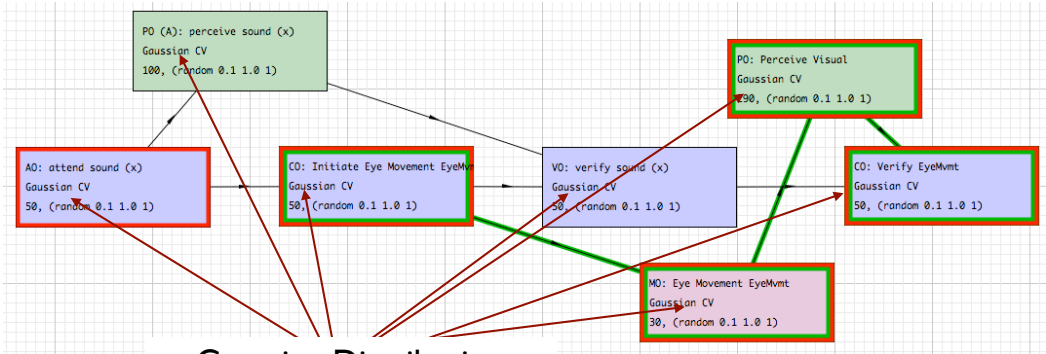
Inserted dependency line

Building a Preliminary CPM-GOMS Model
Insert Operator Durations

Inserted duration according to specific task situation

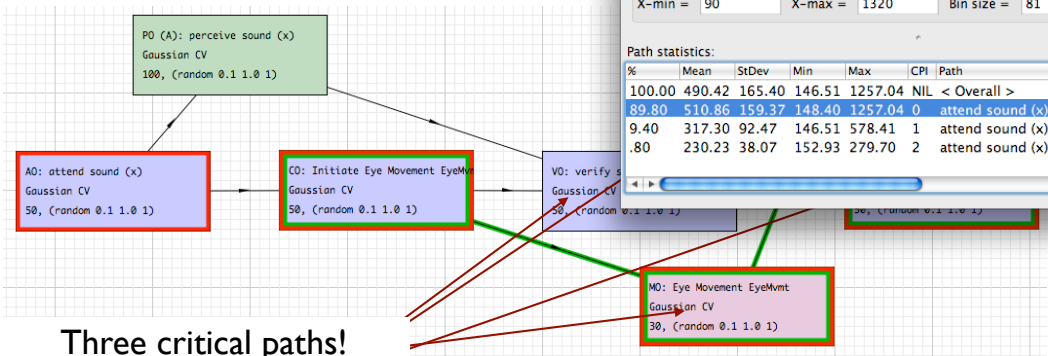


Building a Preliminary CPM-GOMS Model
Interleave Operators + Stochastic Operation Times



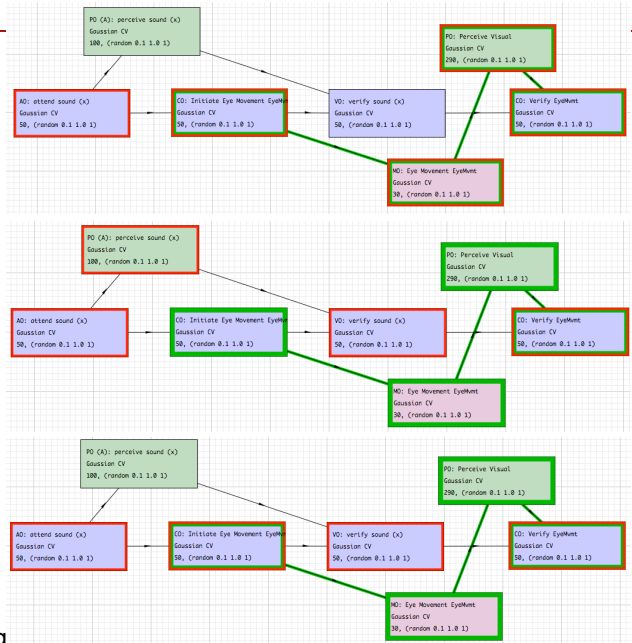
Gaussian Distribution
(randomly sampled on each model run)

Building a Preliminary CPM-GOMS
Interleave Operators + Stochastic C



Three critical paths!
Different mean times!

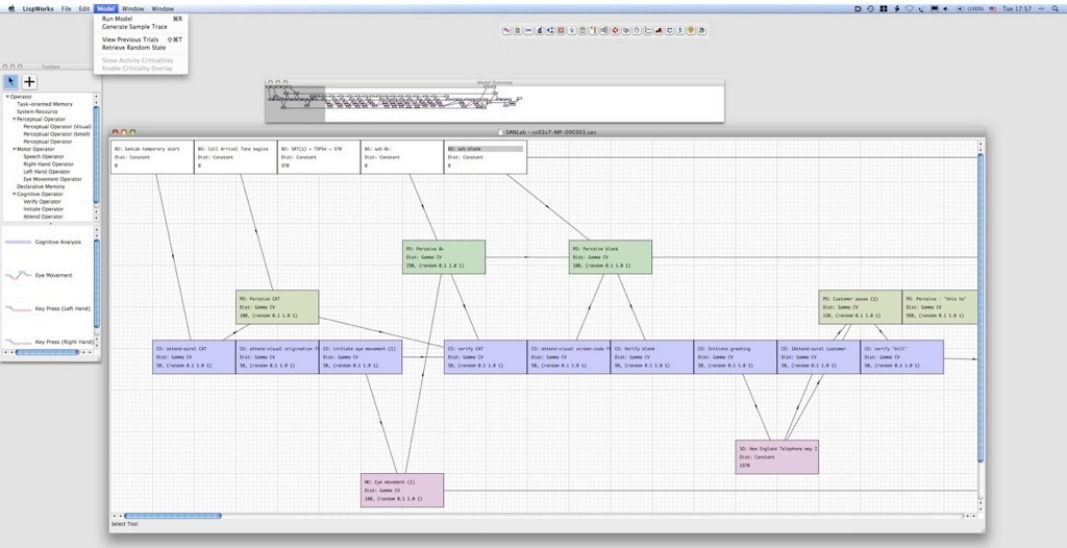
Three Critical Paths



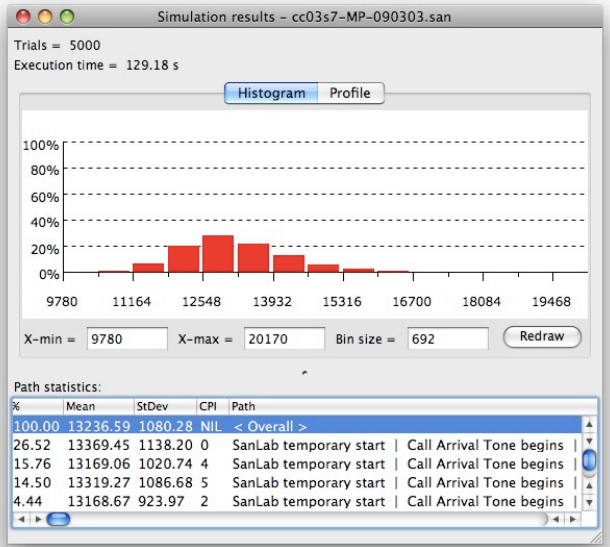
Very Simple Model: Summary

Interleaving	Fixed/ Stochastic	Critical Path	Predicted Times
No Interleaving	Fixed	One	620 ms
Interleaving	Fixed	One	470 ms
Interleaving	Stochastic	Average	490 ms
Interleaving	Stochastic	90%	511 ms
Interleaving	Stochastic	9%	317 ms
Interleaving	Stochastic	1%	230 ms

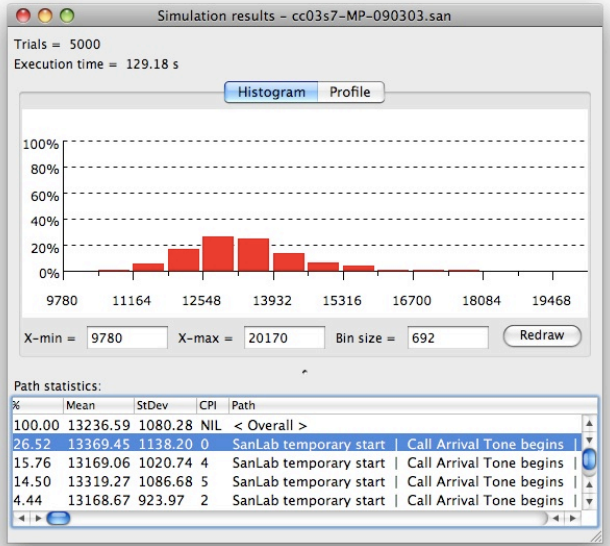
Running a Model 5,000 Times



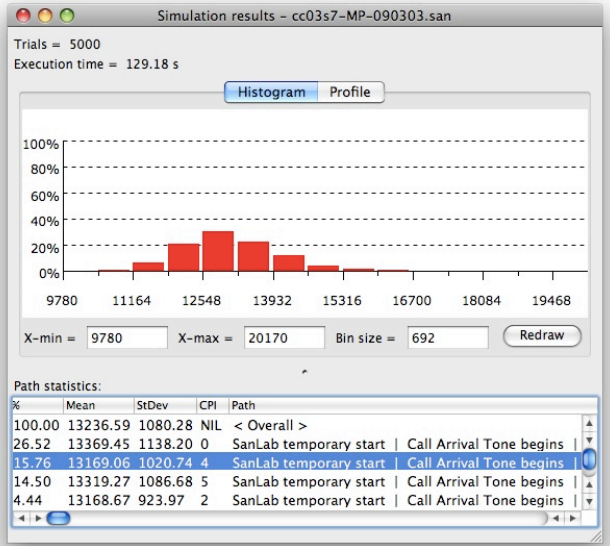
Histogram: Runtime Distribution of 5000 model runs – min ≈ 10s, max ≈ 16s



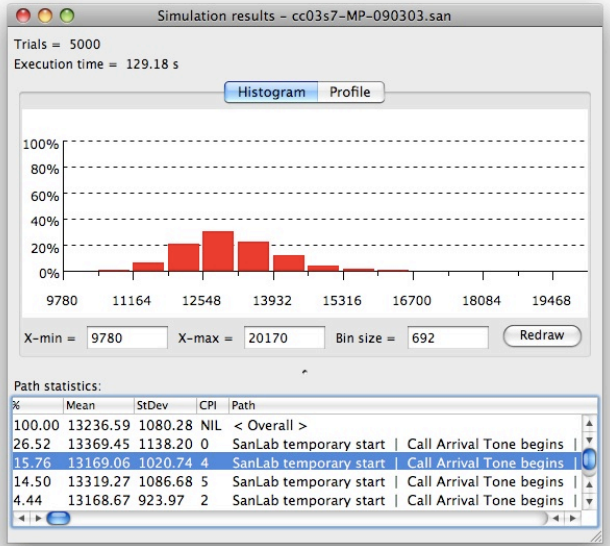
Most Frequent Critical Path Accounts for 27% of Runs



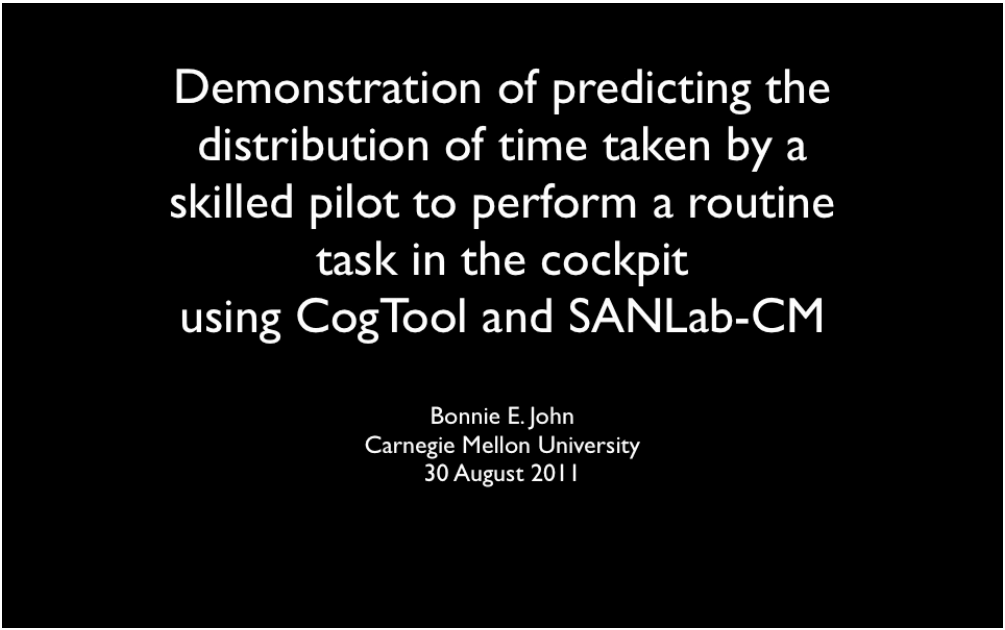
2nd Most Frequent Critical Path Accounts of 16% of Runs



2nd Most Frequent Critical Path Accounts of 16% of Runs



CogTool to SANLab



Newell's Dream

- CogTool to SANLab is an important but limited step
- How about the ability to go from log files of people performing tasks directly to modeling?
- Newell's dream of an automatic protocol analyzer

Newell's Dream

- SANLab+
  - Requires cognitive architectures that encompass
    - Control of cognition
    - Cognition
    - Perception
    - Action
  - Ability to swap out architectural assumptions
    - For example, ACT-R, Soar, EPIC
  - Initial data sets will be taken from people performing three different paradigms

Newell's Dream

- SANLab+
  - Initial data sets will be taken from people performing three different paradigms
  - PRP – psychological refractory period
    - Behaviorally this is a very simple response time paradigm
  - NavBack – a dual-task paradigm
    - Continuous motor movement
    - Eye movements
    - Working memory maintainance
  - DMAP – Decision Making Argus Prime
    - Complex visual search and decision making task

Then & Now

- MacSHAPA
  - MacSHAPA (1995's) - Submarine Commanders: managing complexity of verbal and action protocols
  - MacSHAPA Cognitive Metrics Profiling
- Action Protocol Tracer
  - Finite state grammars for pattern recognition in action protocol data
- SANLab
  - SANLab - tool for Stochastic Analytic Network modeling +++
  - Automating production of cognitive models from behaviorial/ action protocol analysis


Thank You!