

DEEP GREEN

Industry Day

COL John “Buck” Surdu, PM
Information Processing Technology Office

5 December, 2007

Deep Green Supercharges the Commander



“Crystal Ball” updates assessments with information from ongoing operations

Subordinates execute decisions

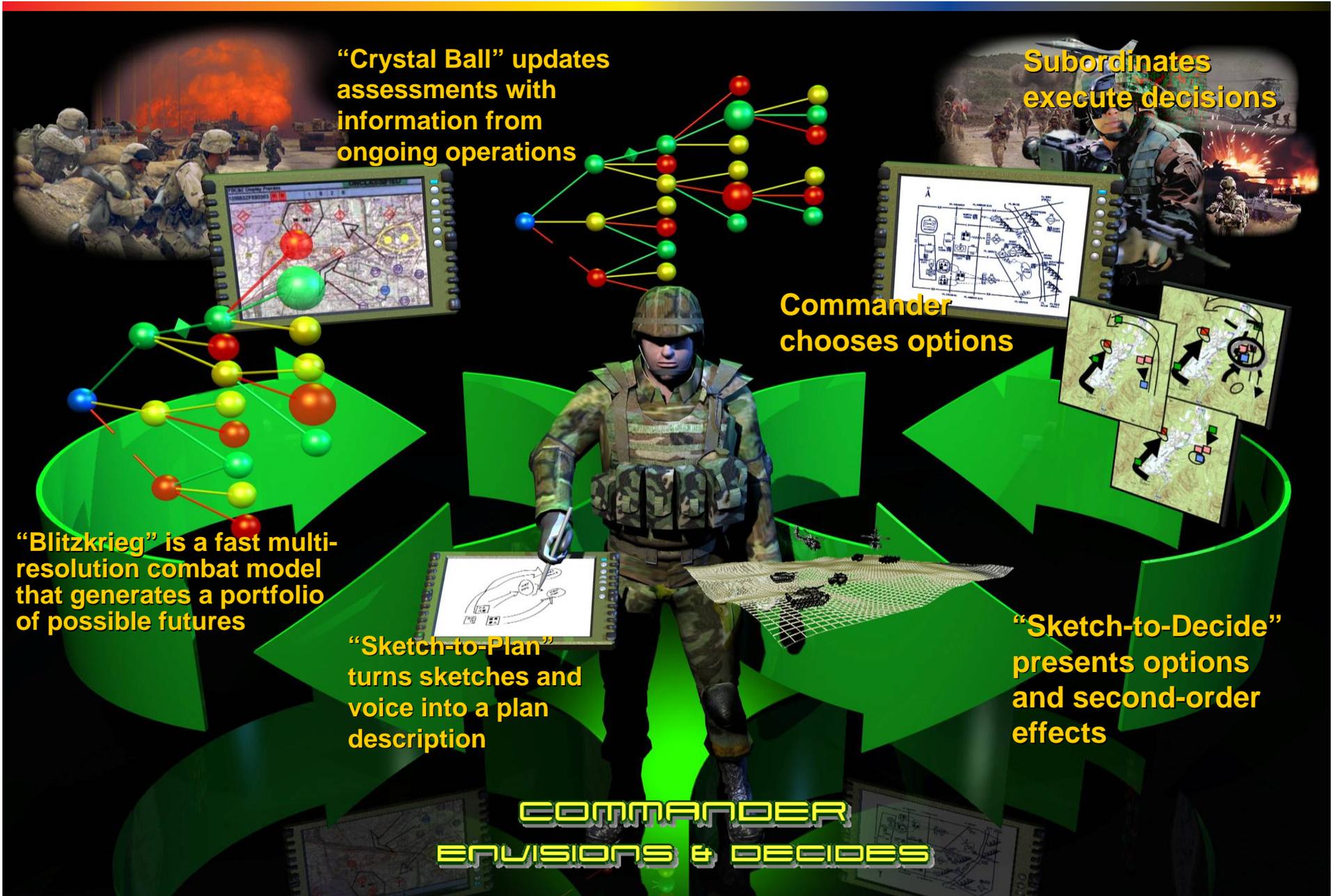
Commander chooses options

“Blitzkrieg” is a fast multi-resolution combat model that generates a portfolio of possible futures

“Sketch-to-Plan” turns sketches and voice into a plan description

“Sketch-to-Decide” presents options and second-order effects

COMMANDER
ENVISIONS & DECIDES



The Deep Green Vision



- Enhances commanders' option generation and decision processes
 - *Plan understanding* from multi-modal human sketching and speech
 - Intuitive presentation of decision points and second- and third-order effects of decisions
- Avoids “black swans” (unanticipated plan breakage) by predicting likely futures and building options before they are needed
- Generates a broad spectrum of possible futures with an instant-response, multi-resolution combat modeler
- Streamlined transition to the battle command community

“All current battle command efforts provide incremental improvements over existing capabilities. The only thing that provides something I don't have today is Deep Green.”

BG Justice, Deputy Program Executive Officer for Command, Control and Communications Tactical (PEO C3T)

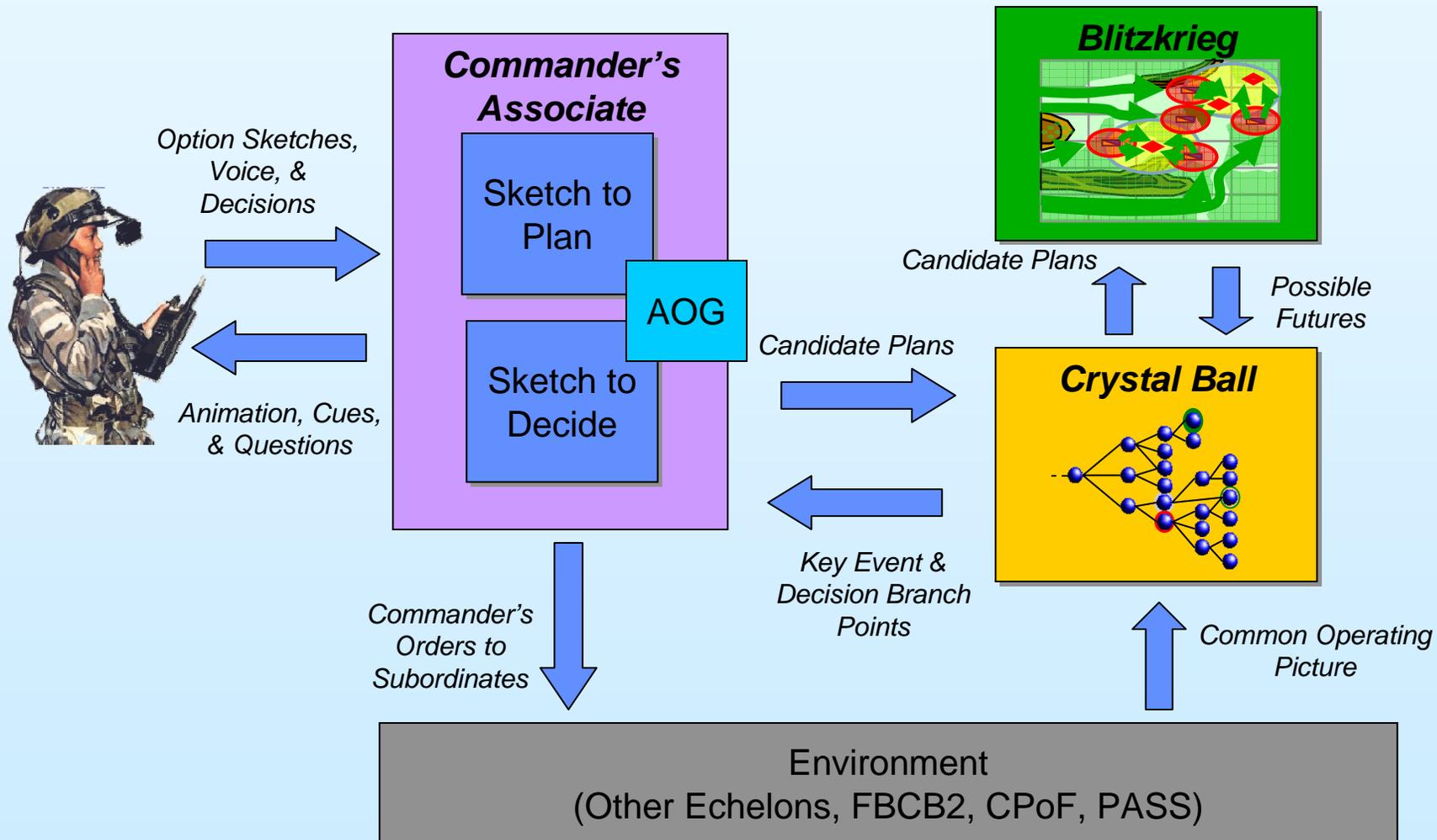
Gaps and Barriers



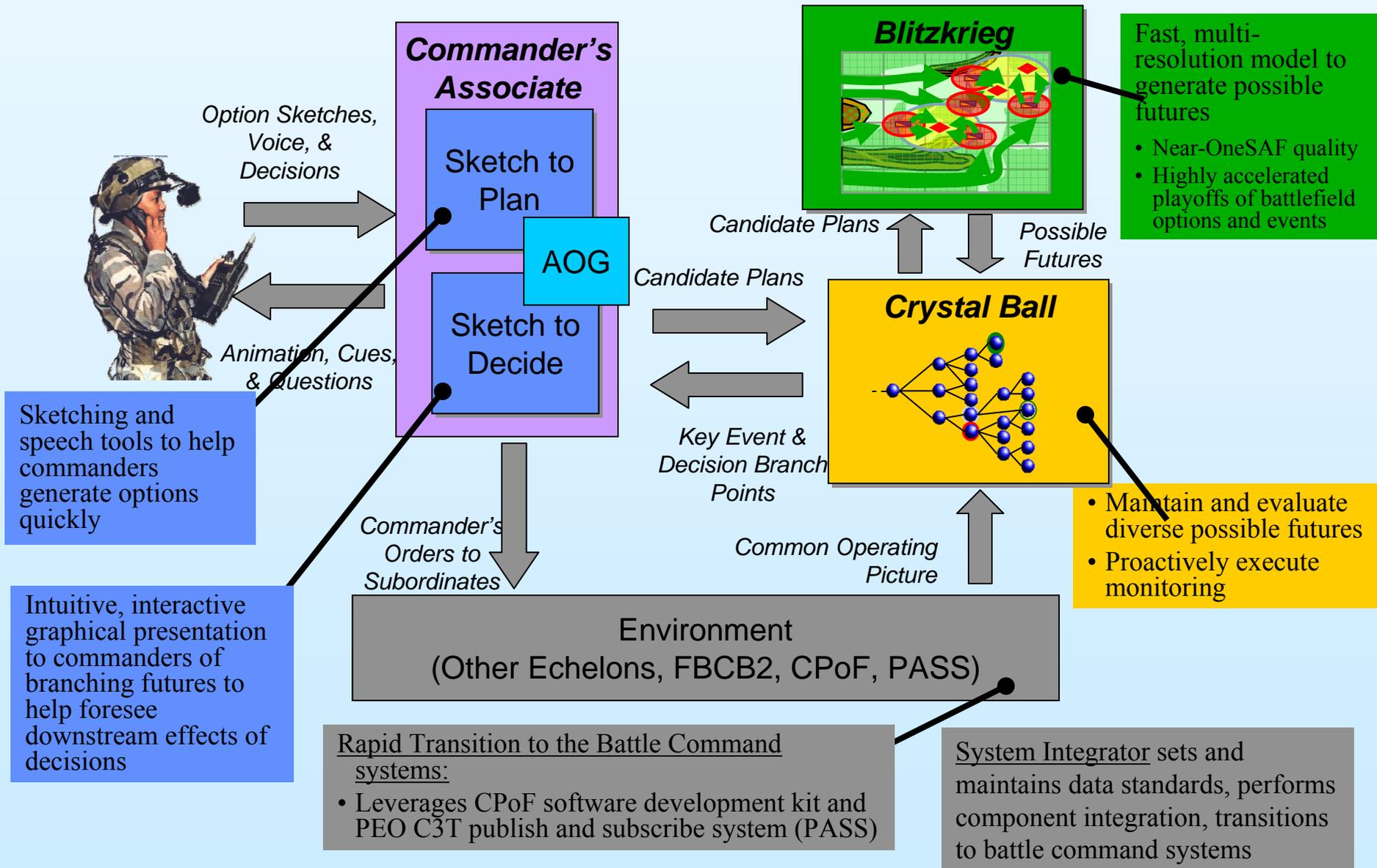
- Army written requirements for machine support for *commander-driven* planning and execution (italics are DARPA's):
 - Commanders require capabilities to enable more *rapid decision – action cycles, with much less effort* (TRADOC)
 - Capabilities to automatically present commanders with *relevant potential courses of action* (ARMY PEO C3T)
 - Enhanced cognitive *understanding of battlefield dynamics* (ARMY PEO C3T)
 - Commanders find it difficult to:
 - Construct more than a small portfolio of hand-crafted plans
 - Foresee impending plan breakage
 - Generate branches and sequels during operations *before* plan breakage
- New technology is needed for machine induction of intuitively expressed plans
 - Existing AI planning & monitoring systems focus on full automation and can be brittle
 - Current generation of combat models are too slow and require significant manual intervention

Today, commanders generate three courses of action, and then choose *one* to execute. Limited contingency planning is conducted; when the *one* plan breaks, the commander becomes reactive.

Deep Green is Commander-Driven



Deep Green is Commander-Driven

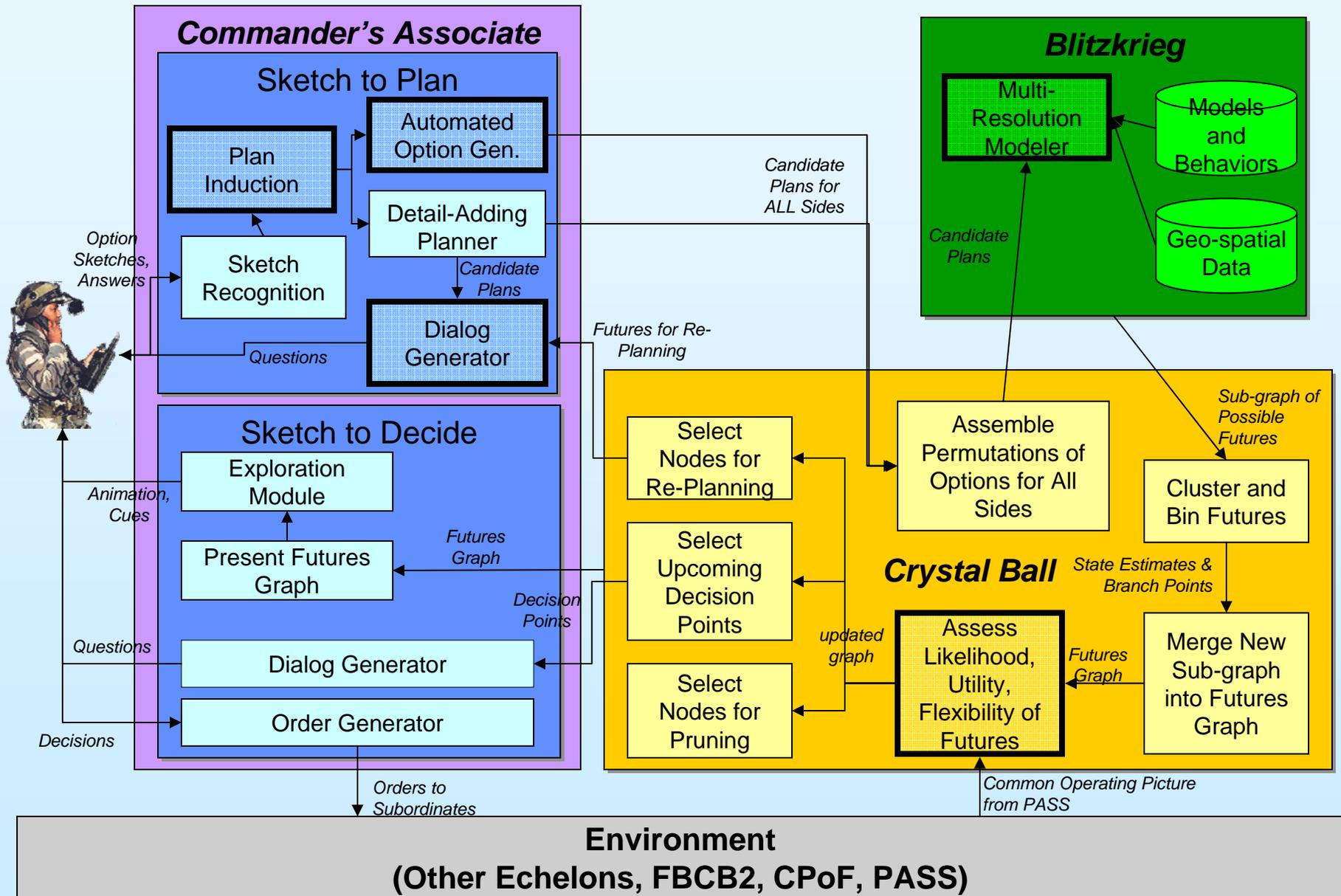


Two Separate Tasks



- Task 1: Deep Green Development
 - Commander's Associate
 - Sketch to Plan
 - Sketch to Decide
 - Automated Option Generation
 - Blitzkrieg
 - Crystal Ball
 - Integration
- Task 2: Test and Evaluation
- Teams may bid both tasks, but if you are selected for test and evaluation, you cannot win any other contracts

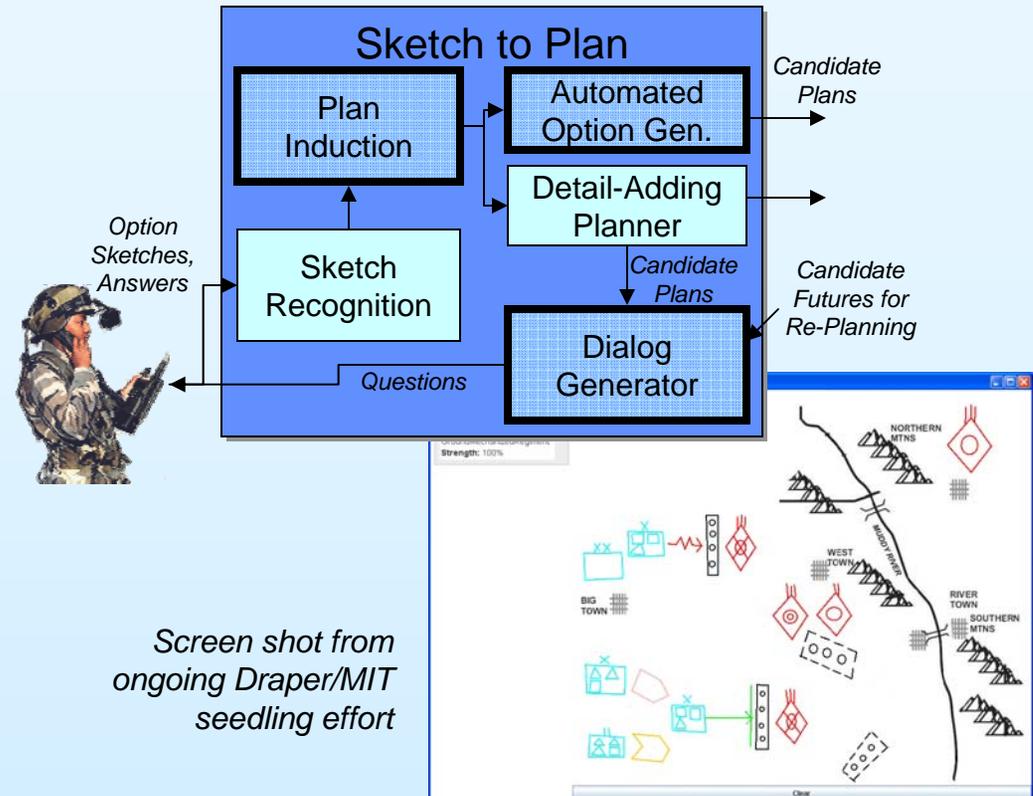
More Detailed Architecture View



“Sketch to Plan” Concept



- Commander **draws** in “free hand” and **speaks**; STP interprets the symbols, replacing them with the correct standard military symbols
- STP *accurately* induces plans from sketch and speech, fills in missing details
- STP asks clarifying questions if it doesn’t understand the sketch
- This allows commanders to specify an option at a coarse level, then move on to the next cognitive task



DARPA Hard: Inferring plan from sketch understanding, including clarifying questions between the commander and **Sketch to Plan**; automated option generation

Why We Think We Can Get There:

- Current work in military sketching (Forbus, Cohen)
- Seedling effort with Draper Labs and MIT CSAIL (Davis)
- Current work on mixed-initiative planning (Allen, best AAAI paper 2007)
- Prior work on “natural” commander’s intent language (Alphatech/BAE)
- Recent work in semi-automated plan generation (Myers, Wilkins), playbook approaches to semi-autonomous control (Goldman)

- Multi-modal (sketch and speech)
- Free-hand sketching, not pallet-driven
- Allow *COMMANDERS* and their staffs to generate options
- Detail adding planner assists, but does not generate options
- MIL STD 2525b recognition
- Provide user feedback
- Formulate clarifying questions
- Output: enhanced MSDL
- Expectation:
 - Technical approach to sketch understanding, plan induction, induction of commander's intent, and query generation

Sketch to Plan Guidance

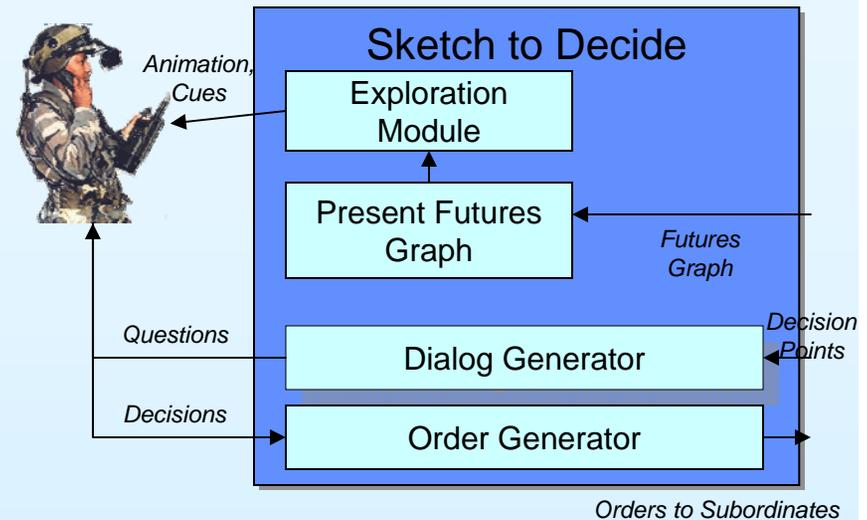


- **Commander's Intent (Army):**
 - A clear, concise statement of what the force must do and the conditions the force must meet to succeed with respect to the enemy, terrain, and desired end state. (FM 3-0)
- **Commander's Intent (Marine Corps):**
 - A commander's clear, concise articulation of the propose(s) behind one or more tasks assigned to a subordinate. It is one of the two parts of every mission statement which guides the exercise of initiative in the absence of instructions.
- **The commander must not be constrained to specifying intent in artificial, stilted language; however, he should be constrained to use of doctrinally accurate terminology.**
 - "The red zone" and "thunder run" are not doctrinal terms.

Sketch to Decide Concept



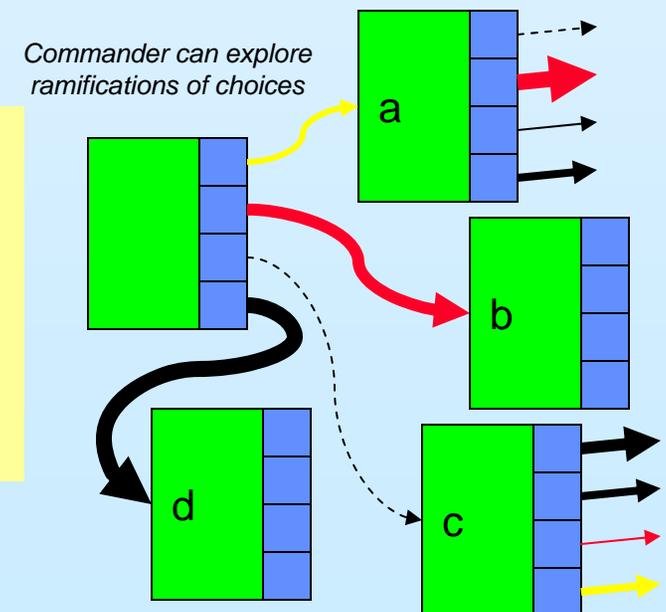
- **Sketch to Decide** provides a window into the futures that have been generated by **Blitzkrieg**.
- Allows the commander to see how options may play out by following different “flows”
- Intuitive ways to depict likelihood, goodness, and flexibility as flows are explored



Hard: Presentation of branching possible futures to the commander in an intuitive way; Supporting user understanding of branch points across multiple decision dimensions and utilities

Why We Think We Can Get There:

- Recent efforts to generate “comic graphs” to show passage of time (Forbus)
- Recent work (Han) shows temporal relationships
- Seedling effort with Draper Labs and MIT CSAIL (Davis)



Sketch to Decide Guidance



- Present multi-dimensional, decision-relevant information in a way that enhanced decision making
 - Enhance understanding of second- and third-order effects
- Exploration of the future graph
- From any “frame” of Sketch to Decide, the commander can perform Sketch to Plan
- Aids in commander’s visualization:
 - The mental process of achieving a clear understanding of the force’s current state with relation to the enemy and environment (situational understanding), and developing a desired end state which represents mission accomplishment and the key tasks that move the force from its current state to the end state (commander’s intent). (FM 6-0)
- Expectation:
 - Technical approach to presentation of multi-dimensional information that AIDS in understanding rather than increasing cognitive load
- Across the board, this was not addressed very strongly in most previous proposals

- *Modest* effort
- Not the main focus of Deep Green
- Given induction of plan and commander's intent, generate qualitatively different, relevant "mutations"
- Expectation:
 - Technical approach to
 - Creating qualitatively different mutations of the commander's plan
 - Generating alternative ways to accomplish commander's intent that are qualitatively different than specified plan(s)

Blitzkrieg Concept



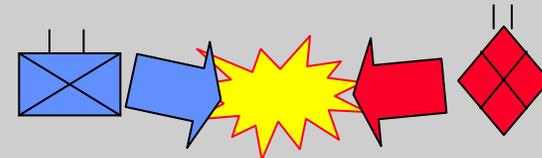
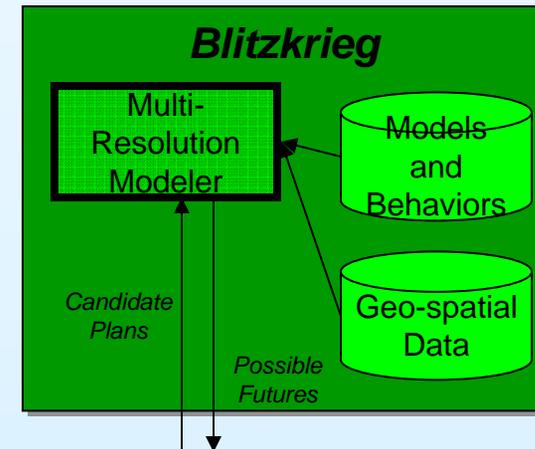
- High speed combat model links the portfolio of plans to potential battlefield outcomes
- Uses multi-resolution modeling (components of varying resolution) to achieve acceleration
- No human intervention during execution
- Uses hybrid of quantitative and qualitative methods to identify branch points
- Permits large time warping

DARPA Hard:

- Determining qualitative branch points and using quantitative techniques to assess likelihoods of each branch
- Determining when and how to dynamically adjust model resolution to answer the appropriate question
- Massive time warping with little loss of fidelity

Why We Think We Can Get There:

- Multi-trajectory simulation research (Gilmer, Sullivan)
- Hybrid of quantitative modeling and qualitative physics (Forbus)



Example: two forces collide.

Qualitatively different outcomes include:

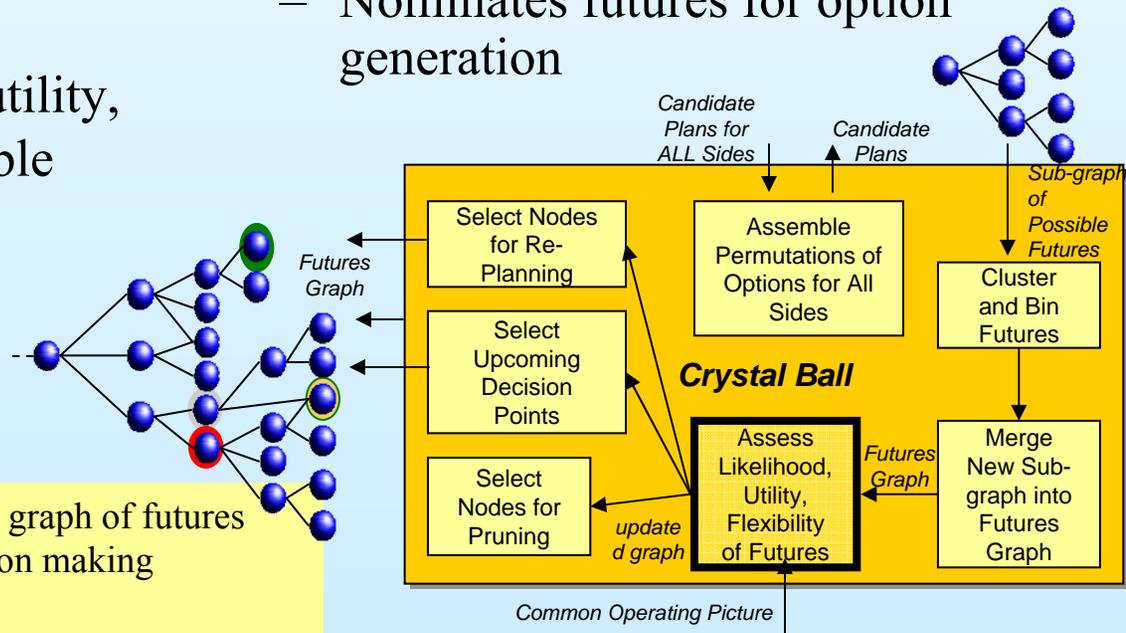
- Blue is destroyed
- Red is destroyed
- Blue begins to lose and withdraws
- Red begins to lose and withdraws
- The two forces choose not to engage
- And so on...

- Does not *do* planning; generates futures from options generated by humans
- Generates a broad set of qualitatively different futures, the “futures graph”
- May or may not have SAF-like intelligence of actors
- Probably aggregate level
- Integration of qualitative and quantitative methods
- Expectation: Offerors will describe the technical approach to building Blitzkrieg

Crystal Ball Concept



- During Option Generation
 - Clusters outputs of Blitzkrieg
 - Merges outputs from Blitzkrieg into graph of possible futures
 - Computes likelihood, utility, and flexibility of possible futures
- During Execution Monitoring
 - Monitors execution of operation
 - Updates estimates of likelihood, utility, and flexibility
 - Nominates futures for option generation



DARPA Hard: Maintaining an evolving graph of futures while identifying key branches for decision making

Why We Think We Can Get There:

- Seedling effort with ISI (Cohen)
- Recent developments in machine learning: Incremental case-based reasoning (Cox), Hybrid Bayesian networks (Pattipati), Structured Bayesian networks (Koller)

- Does not *do* planning
- Collects information from battle command systems to assess the state of the current operation
- Maintains futures graph
- Predicts likelihood of possible futures
- Identifies futures for pruning, re-planning, and decision support
- Expectation:
 - Technical approach to maintaining graph
 - Details on how various metrics associated with futures will be computed
 - Transparency to facilitate explainability of results and recommendations

- Two main purposes
 - Interoperability among Deep Green components
 - Interoperability with real battle command systems
- Concern about retaining modular architecture
 - *Independent!!!*
 - How will you ensure that in a hundred little ways local decisions don't defeat the modular intent of Deep Green?
- Establishes data standards, interfaces, APIs, etc.
 - Expected to peer review recommendations
 - Enforce decisions
- Develop overall architecture in a way that supports the multi-echelon goal of Deep Green in subsequent phases
- Desirable attributes
 - “Built-in test”
 - Spiral development with running code in an early spiral that increases monotonically in functionality throughout the phase.
 - Open architecture, interfaces, and standards required. Open source desired.

Task 2: Test and Evaluation



- BAA provides guidance on goals and metrics
- DARPA expects offerors to propose test plans and methodologies
- OneSAF Objective System as major part of test harness
 - OOS is ITAR protected!
- External support from ARL, HRED
- Within first 90 days
 - Provide test harness to component developers for at-home testing
 - Identify needed enhancements to OOS
 - Identify tentative test scenarios for Phases I and II
- Develop test scenarios
- Test readiness reviews
- Two tests per phase

Phase I

“Functional Components”

12 months



SAF Plan
View Display



Deep Green

- Component tests
- Mid Intensity Conflict (e.g., 3rd ID during invasion of Iraq, Masr al Sharif)
- Test SMEs: Component testers
- SMEs interact directly with various components
- OneSAF used as exercise driver

Phase II

“Functional, Integrated System”

12 months



Battle
Command



Deep Green

- A series of force-on-force, human-in-the-loop experiments in a simulation center
- Counter-insurgency operations
- Test SMEs: Commander, S-3 (operations officer), and S-2 (intelligence officer)
- Staff interacts directly with the Commander’s Associate
- OneSAF used as exercise driver

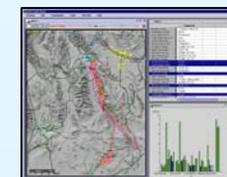
Phase III

“Integrated with Battle Command”

12 months



Battle
Command



Deep Green

CPOF with Deep Green
“Under the Hood”

- A series of force-on-force, human-in-the-loop experiments in a simulation center and a tactical environment
- Large, Three-Block War Operation (e.g., Fallujah or Basra)
- Test SMEs: Commander, S-3 (operations officer), fire support officer, engineer officer, and S-2 (intelligence officer)
- Staff interacts with Commander’s Associate through CPOF
- OneSAF used as exercise driver

Go/No Go Criteria



Task	Phase I	Phase II	Phase III
<p><u>Sketch to Plan:</u> Accurate recognition of sketched MIL STD 2525b symbols. $F = (2 \times H \times P) / (H + P) \times 100$, where #T = number of symbols that should be correctly interpreted #I = number of symbols that are interpreted H = Hit rate = number of correctly interpreted symbols / #T P = Precision = 1 - False positive rate = number of correctly interpreted symbols / #I</p>	<p>A subset of the MIL STD 2525b symbols (approx. 1/5th of the total) are commonly used when describing military operations. For this subset, $F > 90\%$</p>	<p>For a subset consisting of 80% of all MIL STD 2525b symbols, $F > 90\%$</p>	<p>For all MIL STD 2525b symbols, $F > 90\%$</p>
<p><u>Sketch to Plan:</u> Accurate machine induction of user's intended plan. Subject matter experts judge key aspects of machine-induced plans by playing them through OneSAF. $S = (\# \text{ aspects of plan judged to be correct} / \# \text{ aspects of plan entered by user}) \times 100$</p>	<p>$S > 70\%$</p>	<p>$S > 80\%$</p>	<p>$S > 90\%$</p>
<p><u>Crystal Ball:</u> Reduce blind alleys during execution. A "blind alley" occurs during execution when Blue reaches an unpredicted or ill-prepared state (less than three good options available). $A = (\# \text{ blind alleys with Crystal Ball} / \# \text{ blind alleys without Crystal Ball}) \times 100$</p>	<p>$A < 80\%$</p>	<p>$A < 50\%$</p>	<p>$A < 20\%$</p>
<p><u>Blitzkrieg:</u> Reduced time to evaluate combinations of representative Blue and Red courses of action. (Canonical test case will involve 3 Blue COAs vs. 3 Red COAs.)</p>	<p>Today: 120 min Deep Green < 30 min.</p>	<p>Today: 120 min. Deep Green < 10 min.</p>	<p>Today: 120 min. Deep Green < 3 min.</p>
<p><u>Overall System:</u> Reduce staff requirements. Measured by reduction in staff usually needed to accomplish brigade planning/execution for the milestone scenario. $T = \# \text{ Staff With Deep Green} / \# \text{ Staff Without Deep Green}$</p>	<p>$T < 80\%$</p>	<p>$T < 50\%$</p>	<p>$T < 25\%$</p>
<p><u>Overall System:</u> Commander's Performance. Measured by a single numerical score P, computed as a weighted function $f(\bullet)$ of mission accomplishment, friendly losses, enemy losses, neutral losses, time to accomplish tasks, and various human factors. $P = f(\text{with Deep Green}) / f(\text{without Deep Green})$</p>	<p>$P \geq 1$</p>	<p>$P > 1$</p>	<p>$P > 2$</p>

- To be considered responsive, address all boxes from the architecture pictures for your components
- Propose Phase I with options for Phases II and III
- SMEs will be a program asset
- Planning for three 12-month phases
- Address all BFAs for each phase, as well as COE, etc. For instance, how can Crystal Ball facilitate more dynamic intelligence collection planning?

General Notes (2 of 2)



- Use services and functions allocated to other components of Deep Green rather than create redundant capabilities
- OOS is ITAR protected, so precautions have to be taken by ALL performers
- Reports from seedlings will be provided GFI to all performers
- All software, documentation, and data will be unlimited use
- IF there are more than one performer per task, down-selection will be made at the end of Phase I.
- All components as learning agents.