TL

Transfer Learning

Proposer Information Pamphlet (PIP)

for

Broad Agency Announcement 05-29

Defense Advanced Research Projects Agency
Information Processing Technology Office
3701 North Fairfax Drive
Arlington, VA 22203-1714

TABLE OF CONTENTS

1 PROGRAM OBJECTIVE

The goal of the Transfer Learning Program solicited by this BAA is to develop, implement, demonstrate and evaluate theories, architectures, algorithms, methods, and techniques that enable computers to apply knowledge learned for a particular, original set of tasks to achieve superior performance on new, previously unseen tasks. This goal reflects the observation that key cognitive abilities of humans include the abilities to generalize, abstract, reuse, reorganize and apply knowledge learned in previous life experiences to novel situations. Three types of superior performance may result on the new, previously unseen tasks by the system that was trained on a set of original tasks, compared to a version of the system that was not trained on the original tasks, from transfer learning: initial performance improvement (Type 1), rate of performance improvement (Type 2), and asymptotic/maximal achievable performance increase (Type 3). These are illustrated in Figure 1.

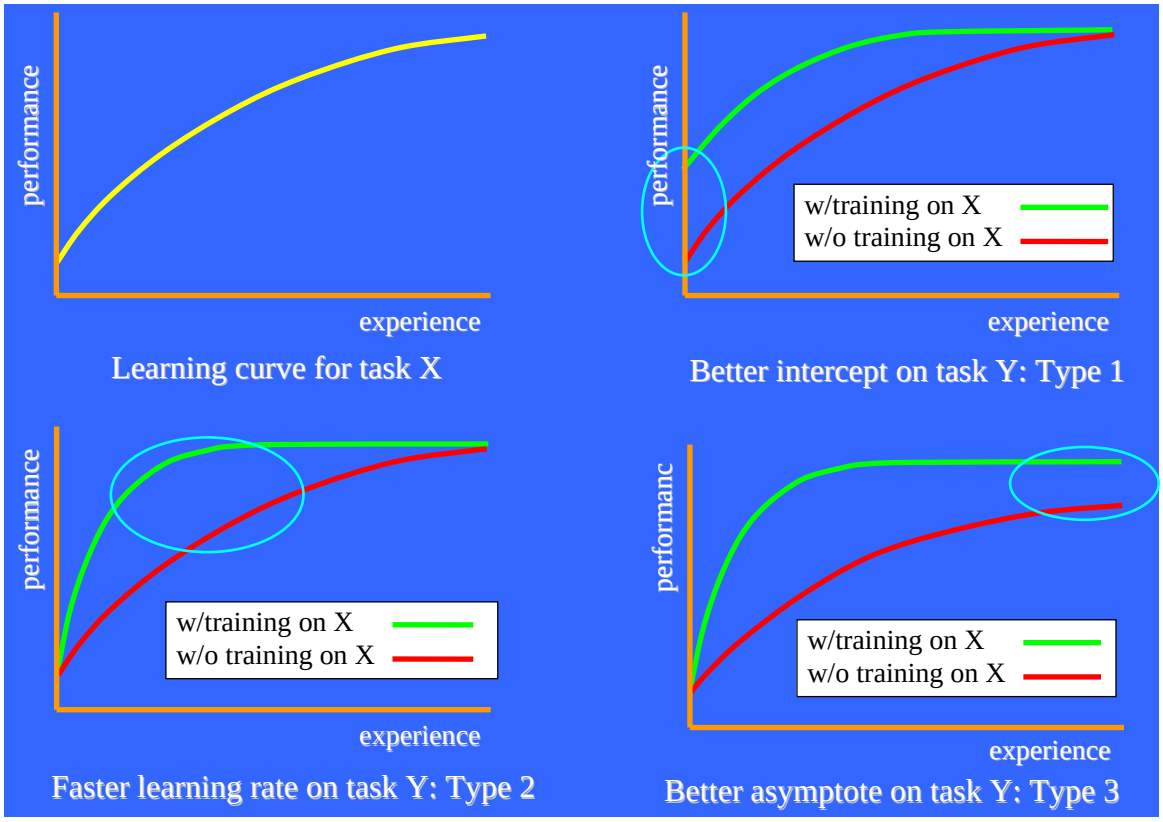

Figure 1: Types of Transfer Learning

DARPA is interested in the application of knowledge and skills gained on one set of problems to a different and novel set of problems or tasks.

2 BACKGROUND & MOTIVATION

Future military conflicts are likely to involve unanticipated situations to a great degree. The traditional military approach of training for specific situations and of designing hardware and software for these situations will need to be augmented to allow for rapid

determination of how to behave appropriately and effectively in a novel situation. This will require transferring knowledge and skills learned from a wide variety of previous situations to the current, previously unencountered situation. US Forces and systems, especially computing systems that exhibit significant cognitive capabilities (either as assistants or associates to humans or as partially or fully autonomous entities), will need to be able to recognize and adapt to novel situations and to act appropriately and effectively. They will need to compose behaviors for these novel situations from components of behaviors that were learned in previous situations that were encountered in training or in operations, to modify the resulting compositions to match unique aspects of the current novel situation, and to do this the first time the novel situation is encountered, with no or extremely limited training examples of the novel situation.

DARPA's Information Processing Technology Office (IPTO) has as its mission the creation of a new generation of computational and information systems that possess capabilities far beyond those of current systems. These *cognitive systems* – systems that know what they're doing –

- Will be able to reason, using substantial amounts of appropriately represented knowledge;
- Will learn from their experiences and improve their performance over time;
- Will be capable of explaining themselves and taking naturally expressed direction from humans;
- Will be aware of themselves and able to reflect on their own behavior; and
- Will be able to respond robustly to surprises, in a very general way.

Transfer learning is one key capability of such a cognitive system.

Learning comprises the set of techniques by which people – the only existing examples of complete functioning cognitive systems – acquire knowledge and competence. Learning enables far more than performance improvement on existing tasks; it also enables performing new tasks that are different from those performed in previous experience. Learning is what provides the necessary abilities to function in a complex, dynamic environment. Computational systems that have the ability to learn will provide not only these capabilities but also the additional capability to learn to perform tasks that are beyond the capabilities of individual people to perform or beyond the capabilities of people to specify (i.e., perception tasks such as vision, speech, and language).

Current machine learning technology has achieved significant successes but is still far from what is needed. Traditional approaches to machine learning have provided many achievements, especially in the area of techniques for classification of large data sets based on a reasonable number of labeled training examples and with respect to selection of actions in limited domains. Recent advances have enabled the learning of relational knowledge. However, many human learning capabilities have not yet been addressed. Many domain independent algorithms exist, but a human engineer is typically needed to formulate and encode a particular problem in a representation suitable for the algorithm, to provide a significant amount of training data, to select the form of the hypothesis to be learned, and to evaluate the results. These techniques do not rapidly learn knowledge in one domain and use that knowledge to learn more quickly and/or perform better in

another domain. The representation is limited to a particular problem and the learned knowledge is limited to a particular application. Overcoming these limitations to achieve the vision of cognitive learning requires acquisition of knowledge elements and problem solving strategies that can be generalized, abstracted, represented, retained, extended, recognized, selected and re-composed dynamically for purposes unanticipated during acquisition and system design. The Transfer Learning Program aims to create technologies to attain this vision.

Cognitive scientists have developed computational theories and architectures that model human cognitive functions, including learning. These theories and architectures have been able to simulate human behavior in many areas; however, they do not account for the wide variety of learning methods exhibited by people nor do they admit a variety of representations. In particular, while most of today's cognitive architectures support some learning capabilities, they lack sufficient robustness and flexibility for applying learned knowledge on tasks in unfamiliar domains. Also, these architectures have repeatedly adopted only a few representational frameworks (e.g., production systems, plans), and have largely ignored others (e.g., frames, description logics). Thus, this keeps them from transferring learned knowledge for problems and domains whose appropriate representations differ. Attaining the goal of broadly capable cognitive architectures, and in particular those that can transfer learned knowledge effectively, requires frameworks that can encode both abstract and episodic knowledge in multiple formalisms, relate them to each other, and employ meta-cognitive, analogical, and other reasoning processes. The Transfer Learning Program aims to spur the development of cognitive theories and architectures to encourage these developments.

Learning is viewed as an emergent behavior of an ongoing integrated cognitive architecture rather than solely as a collection of algorithms that explore a space (structural and/or parameterized) of models with the goal of identifying the model with the best performance on a particular task based on data drawn from a single distribution. It is anticipated that these integrated cognitive architectures will contain and exhibit both learning and problem-solving capabilities. DARPA is interested in capabilities that exhibit a wide range of learning behaviors appropriate to a range of tasks and domains. Rather than a one-shot learner on a specific data set, DARPA envisions an ongoing learner that operates on a series of situations, incorporating feedback from each into the next, and acquiring and reorganizing knowledge to enable solution of a wider range of problems. Key ideas that may contribute to this vision include the ability to create and then reuse abstract or general concepts, the ability to create and use hierarchical or otherwise-structured representations, the ability to recognize what knowledge is applicable to a task, the ability to incorporate advice or guidance from an experienced problem solver (i.e., teacher or coach), the ability to recognize and use similarly structured problems, the ability to decompose problem solutions into component parts and reassemble new solutions from combinations of these components, the ability to recognize what must be learned (i.e., formulate a new learning problem appropriate for a novel task), the self-awareness needed to enable recognition of one’s own problem-solving capabilities and limitations, and many others.

It is anticipated that ideas that will enable achievement of this vision will require new concepts and syntheses of approaches from the fields of artificial intelligence, cognitive science, and perhaps others.

3 SCOPE

3.1 Phasing

This Program consists of three one-year phases, with decision points between each phase. These decisions will be based on the amount of progress made against the project goals specified in the proposals. This BAA is for all three phases. DARPA anticipates awarding contracts for Phase 1 with two one-year options for Phases 2 and 3.

3.2 Organization

The Program solicits research contributions drawn from all applicable fields, particularly artificial intelligence and cognitive science. It is anticipated that no single approach will provide the desired capability; rather, integrated approaches based on a synthesis of new and existing ideas from multiple research areas implemented as integrated systems based on a coherent theory or model are likely to be needed.

This Program and solicitation has two areas: the primary area, which aims to develop integrated transfer learning capabilities, and the evaluation area, which will develop challenge problems and execute evaluations appropriate to measure to what degree the objectives have been achieved (or exceeded). Offerors may bid on the primary area; all offerors must also bid on the evaluation area, as described below.

- Area A (primary) - this area seeks to make dramatic advances in transfer learning technology. It will result in software that can learn the knowledge and skills necessary to perform a task in one area and then demonstrate superior performance (relative to software that has not learned the original task) in another area by taking advantage of the learned knowledge and skills.

- Area B (evaluation) - this area seeks to establish evaluation frameworks and conduct evaluations of the software developed in Area A. The purpose of these evaluations is to measure progress against the goals of Area A. The evaluation frameworks will consist of specifications and implementations of the tasks to be solved and the necessary knowledge and skills to be acquired as well as the experimental protocols to be followed that will provide useful, meaningful, fair, unbiased, and statistically significant measurements. Note that these evaluation frameworks are envisioned as testbeds rather than as data sets in order to enable the evaluation of ongoing integrated transfer learners.

Every team awarded work under Area A will be expected to be evaluated in three independent domains as follows:

- One of the two domains developed by an independent Area B contractor, to be selected by the Area A contractor and mutually agreed to by DARPA

- The domain proposed by the Area A contractor in their proposal for Area B
- A third domain to be selected from either the second of the two domains developed by independent Area B contractors or from one of the Area B domains developed by another Area A contractor.

The purpose of having the transfer learning capabilities developed under Area A evaluated in three distinct domains is to ensure that the capabilities that are developed are general and do not depend on unique features of a particular domain. However, recognizing that not all approaches to transfer learning will be applicable to all possible evaluation domains, we do not require that every capability be evaluated against all the evaluation domains. (Note that the learned knowledge and skills in each domain will be different, so the evaluation on the transfer task will use a different trained instantiation of the transfer learning software, although the untrained instantiations for all three domains will be identical.)

Both of these areas are discussed in detail in Section 4, below.

3.3 Awards

The dollar amount of awards will be determined by the quality of proposals and funds available. Amounts below are shown only to illustrate the relative emphasis among the different Program elements.

Area A – DARPA anticipates awards to several integrated teams with alternative approaches to achieving the desired technical capabilities for Area A. Teams may have any organizational structure capable of performing the work; examples of potential team structure are as follows: an integrator and multiple research groups from different institutions, peer-collaborations among multiple research groups from different institutions, or multiple research groups from within a single institution. However a team is structured, a specific individual and institution with overall responsibility for team performance, including provision of an integrated transfer learning capability and development of the evaluation framework under Area B, must be identified. Annual funding for each team is anticipated to be around $1-4M, to include the work devoted to Area B, although larger awards may be possible for larger teams.

Area B – DARPA anticipates awards to two teams to develop and execute independently of each other the evaluations under Area B. In addition, each team awarded work under Area A is expected to develop their own evaluation framework and to conduct a fair and unbiased evaluation of their own work. We anticipate that these awards will be in the range of $1-2M per year.

All Program participants are expected to share all intellectual property developed under Area B with each other as needed to conduct the evaluations. Further, all intellectual property needed to conduct Area B evaluations shall be provided to the Government and made publicly available for use by any non-participating organization upon their request and with their willingness to reimburse reasonable costs of executing the evaluation to the evaluation framework developer and executor.

4 TECHNICAL APPROACH

An ongoing integrated transfer learning capability will require combinations of diverse approaches drawn from traditionally distinct fields of research. DARPA desires to achieve progress against the goal of an ongoing integrated transfer learning capability that is applicable to diverse tasks and domains. The areas described in this BAA are designed to promote the formation of teams across disciplinary boundaries.

Key ideas that contribute to the ability to achieve the desired transfer learning will include modular knowledge representations, dynamically composable knowledge modules, and mapping between knowledge representations. Recognizing the applicability of existing knowledge to novel problem-solving situations will be as important as the ability to reuse the existing knowledge in the novel situation. The difficulty of this recognition task will depend, of course, on the amount of overlap of content as well as on the similarity of representation.

Some approaches that may be of interest include learning by analogy, incremental or cumulative learning, chunking, learning by doing, and learning by taking advice. This list is meant to be suggestive, rather than exhaustive; ideas not mentioned here are at least as desirable as those that are. DARPA believes that a combination of existing and new and even radical ideas is needed to achieve the desired transfer learning capabilities.

Learning by analogy involves the recognition and reuse of similar abstract structures in models of the domain. These abstract structures enable the identification of similarities between novel problems and previously encountered problems. New problem instances may consist of modifications or combinations of structures or substructures developed in previous problem instances. Problem solving is a matter of recognizing which previously learned structures or substructures are identical or similar to those in the new problem and then combining and/or modifying these structures/substructures into a proposed solution to the new problem. Learning consists of efficiently acquiring and retaining broadly accurate and applicable substructures. A key technical challenge is determining the mapping between existing and new structures.

Incremental or cumulative learning involves developing a hierarchy (or multiple hierarchies) of concepts and/or problem-solving techniques. More complex problems may be solved by composing solutions to simpler problems. These techniques are well suited for tasks involving transfer of ever increasing complexity and for tasks that can be subdivided into non-interacting or weakly interacting components. Learning can operate on problem-solving traces, observations of another agents' behavior, or on direct instructions.

Chunking creates a single piece of knowledge or concept as a substitute for a reuseable set of more complex pieces of knowledge or concept. As a concept is learned and described, it can be reused to learn higher-level concepts. The reduction in complexity obtained by the specification of a concept is what enables the effective learning of new concepts on which it depends. The complete hierarchy of concepts is available for problem solving so for instances not matching exactly the higher-order concepts the

lower-level concepts can be invoked when needed. This type of learning allows for complex problems to be decomposed into a set of subproblems that can be solved. It enables transfer by decreasing the search space for new tasks and by increasing the domain coverage due to the higher level of abstraction.

Learning by doing leverages the observation that people incrementally build models of problem solving in new situations by starting with common situations and then adding exceptions and complexity as they arise. This type of learning provides an initial basic capability and evolves to more specialized capabilities over time. It enables the representation to develop in a manner that focuses on the most common and reusable knowledge rather than on a collection of cases.

Advice-taking is motivated by the observation that much human learning occurs with the assistance of a teacher. The teacher can provide examples of correct problem-solving or behavior or can provide explicit directions and explanations. Advice may be generally rather than always true; it may guide search rather than constrain it, resulting in overall increases in efficiency without ruling out the ability to address rare situations. Advice typically will make search vastly more efficient and apply to a variety of situations more general than the particular problem instance.

4.1 Area A – Transfer Learning Technology Development

The objective of this research area is to develop ongoing integrated transfer learning capabilities. Transfer learning is defined as the ability to use knowledge or skills learned for an original set of tasks in an original domain to solve problems on a different set of tasks and/or in a different domain. Since no two situations are exactly alike, all learning involves transfer of some sort. However, the amount of transfer that is achieved depends on many factors, primary of which is the degree of similarity (or "distance") between the original and novel task sets and domains. Note that this "distance" depends on how the domains and tasks are represented; learners that can reduce this distance are those that are likely to be successful at transfer learning. Existing approaches adequately address transfer to new examples that are assumed to arise from the same distribution or generating process as the training examples; some existing techniques also address gradual or abrupt shifts in the distribution or process. What is of interest in this new Transfer Learning Program is neither of these two limited types of transfer; rather, DARPA is interested in the application of knowledge and skills gained on one set of problems to a different and novel set of problems or tasks.

Several dimensions of transfer have been identified. These include *content* (how different is the specific knowledge needed to solve different problems?), *representation* (how differently is the knowledge represented?), *task objective* (what problem-solving task is being addressed? – e.g., classification, configuration, planning, inference, etc.) and *problem characteristics*, which is analyzed in more detail below. While it is helpful to think of these dimensions separately, they are not intended to suggest a complete or comprehensive theory or taxonomy of transfer learning capabilities at this point. It may also be useful to distinguish between lateral and vertical transfer: lateral transfer is between problems of similar complexity or difficulty but differs in task and/or domain,

while vertical transfer solves more complex problems by combining solution elements from simpler problems in a particular domain. Transfer may also be analyzed according to the type of knowledge that is re-used between the original and novel tasks/domains. A comprehensive theory of transfer that guides a research approach would have to account for all these dimensions; a research project that responds to this solicitation would incrementally achieve capabilities, exhibiting ever increasingly ambitious types of transfer in each year.

An initial model of levels of transfer tasks, emphasizing characteristics of the types of problems to which the transfer applies, rather than the representations or semantic content of the knowledge needed for solutions, is presented below. This model should be used to characterize the anticipated progress for the approaches provided in response to this area of this solicitation. These levels are in anticipated order of increasing difficulty; however, the ordering is not exact and modifications may be necessary. It is not assumed that lower levels are always prerequisites to higher levels, although it is likely that this may frequently be the case. Further, the incremental difficulty of achieving each level compared to the previous level may vary significantly. Offerors are free to suggest extensions or modifications to the model to make it more accurate and useful. It should be recognized that the real goal of the Program is to achieve the final level of transfer (i.e., differing) and that this level is anticipated to be significantly more difficult to attain than all the previous levels. Approaches that incrementally achieve Levels 0 through 9 but do not clearly identify the breakthroughs that will enable achievement of Level 10 will not be considered adequate; however, approaches that can not identify how Levels 0 through 9 are achieved will not be considered feasible.

Level Title Description

0. Memorization: New problem instances are identical to those previously encountered during training. The new problems are solved more rapidly because learning has occurred.
1. Parameterization: New problem instances involve the same components in the same configurations as those previously encountered during training, but with different parameter values. Parameters are chosen “within-range” to ensure that the quantitative differences do not require qualitatively different solutions.
2. Extrapolating: New problem instances involve the same components in the same configurations as those previously encountered during training, but with different parameter values that may cause qualitatively different solutions to arise. These qualitatively different solutions may result in new concepts that are useful for a wider range of problems than those encountered on the training set.
3. Restructuring: New problem instances involve the same sets of components but in different configurations from those previously encountered during training.
4. Extending: New problem instances involve a greater number of components than those encountered during training, but are chosen from the same sets of components.
5. Restyling: New problem instances involve the same set and number of components as those encountered during training but may be formulated

differently, therefore requiring recognition of the commonalities by the problem solver.

6. Composing: New problem instances consist of combinations of components from distinct component sets encountered during training.
7. Abstracting: New problem instances are similar to training instances only when a particular abstraction is applied. For example, the components may be completely different but the network diagram showing their interactions or the set of differential equations describing their behavior may be identical (except, of course, for notational choices).
8. Generalizing: New problem instances involve principles drawn from a superset of the distribution or generating process of the training distribution or generating process.
9. Reformulating: New problem instances are recognized as similar to training instances only when they are reformulated through a well-specified transformation.
10. **Differing: New problem instances bear minimal apparent overlap with the training instances; however, the problem solutions use some common knowledge or problem-solving strategies. Note that this level represents the goal of the Transfer Learning Program; the previous levels are meant to suggest steps along the way that may be useful milestones and without which Level 10 would not be possible. Most of the effort of a responsive proposed research project should be focused at attaining this level of transfer.**

An instantiation of the above model to a particular domain may be useful to illustrate what is meant. Consider the domain of mechanics, as taught in college-level freshman physics classes. Memorization is self-explanatory and is not really transfer. Parameterization, extrapolating, and restructuring would be typical of the problems often found at the end of a chapter. Extending might be typical of the "extra-credit" or "more-difficult" problems that would be found at the end of a chapter. Restyling might be demonstrated by solving problems in the style of one textbook's formulation after having been trained on one or more other formulations from different textbooks. Composition might combine knowledge about rotational motion with knowledge about momentum. Abstracting might allow for the solution of problems involving angular momentum in addition to the previously encountered solutions involving momentum. Generalization might be the realization, given an understanding of conservation of momentum, that conservation laws may apply to many quantities. Reformulation is the type of transfer needed to solve problems using Hamiltonian equations instead of Newtonian equations. Differing would cover a wide range of areas that would be taught in other classes, ranging from electromagnetism as taught in a separate semester of freshman physics all the way through perhaps chemistry, which is taught in a different department.

A second example instantiation concerns the domain of real-time strategy games, which is a popular genre of computer games that involve decision making in a simulated environment that is populated by adversaries. Typical decisions involve issues concerning resource management, movement, adversarial plan recognition, combat, and diplomacy. Memorization concerns repeating a game with the same initial state (e.g.,

same map, units, starting locations, resources, number and type of adversaries). An example of Parameterization is slightly different starting locations and amounts/locations of resources. Extrapolating could be among problems with different world sizes and starting units. Restructuring is exemplified by problems with different climates, which could require different prioritization of goals and attention (e.g., to irrigation, resource acquisition). Modifying the number of adversaries, so that they can collaborate against you, is an example of Extending. Restyling could involve problems starting on a different map, but otherwise under the same scenario. Composing might require problem solving in a scenario requiring both warfare and economic management skills when training involved at most one of these skills. Starting in a different era, and thus with different unit capabilities and research advancements, exemplifies Abstracting. Generalizing might involve opponents who, instead of using deception only in combat, might also use deception during diplomacy, movement, and other activities. Reformulating could involve problems from different games, but in the same genre and with similar starting scenarios. Finally, Differing could involve games in different genres (e.g., turn-based or individual real-time strategy), requiring the application of perhaps both common knowledge and problem-solving strategies.

Specific sub-areas in this technical area include the following:

4.1.1 Area A1 -- Theory

The successful offeror will specify a computational theory of transfer learning that will form the basis for the offeror's proposed research project. This theory shall be capable of being implemented in software. This theory shall comprise an architecture for learning and problem-solving and shall include all necessary components required to achieve the proposed capability. Proposals should identify all components by function and should illustrate how transfer learning is achieved by their operation and interaction. Proposals should clearly state the current status of each component (e.g., complete, prototype, designed, planned, precursor exists, etc.); the planned status of each component over the life of the proposed research, and the integration plan. Claims as to the types and degree of transfer learning that will be enabled by an implementation of the theory shall be specified. A theory of transfer should among other things specify what is leaned and how it is transferred (e.g., facts, cases, problem solving strategies, domain models, etc.), how the transferred knowledge/ability is modified to be applicable to the novel task/domain, how the applicability to new tasks and/or domains is recognized, how knowledge is represented and retained for reuse and transfer, as well as the factors identified in Section 4.1. It should identify algorithms, the overall architecture, and knowledge representations.

4.1.2 Area A2 -- Implementation

The successful offeror shall implement all components identified in the proposed theory of transfer learning. The implementations shall result annually in a complete software product that contains all necessary components to allow for an evaluation of the claims of the theory. A complete integrated version of the software must be available in month 10 of each year of the Program so evaluations can be conducted in month 11.

Proposals should specify the following information about their approach:

1. The specific measurable quantities (i.e., metrics) against which progress will be evaluated.
2. A clear and specific explanation of how the measurable quantities reflect learning (as distinct from other possible causes).
3. The specific values of these measurable quantities to be achieved in each year of the research. These values may be specified as a range with a minimum performance level and a maximum goal.
4. The precise methods for providing initial knowledge to the software.
5. The precise methods for training the software.
6. A precise description of any tailoring that may be required to use the trained software on a previously unseen domain or task.
7. A precise interface description of the representation (i.e., language) needed for the training examples, the new evaluation examples, and the new task description.

4.1.3 Area A3 -- Evaluation

The successful offeror shall provide necessary support to conduct at least annual evaluations of their transfer learning capability. This support shall consist of frequent technical interactions with the contractors selected to perform the evaluations awarded under Area B of this BAA, the purpose of which is to ensure that the evaluations thoroughly and adequately measure the capabilities that are being developed. The contractor will be required to identify at least three domains/tasks from those being developed under Area B in which the developed transfer learning capability will be evaluated. Proposals should specify the criteria to be used to select these three domains.

The Government plans to develop at least two challenge problems in distinct domains, one in the area of multi-agent adversarial games and the other in the area of standardized tests. In addition, each team awarded a contract under this BAA is responsible for developing their own challenge problem. To ensure a general approach to the problem of transfer learning, each team will be evaluated in three distinct domains, one of which will be one of the two independently developed challenge problems under Area B and one of which will be the one developed by the team. The third may be either the other one developed by the Government or one developed by another team.

4.2 Area B – Transfer Learning Evaluation Frameworks

In Area B, DARPA seeks a set of capabilities to demonstrate and evaluate the effectiveness of transfer learning. These evaluation capabilities will be used to evaluate the effectiveness of the technologies developed under Area A. It is anticipated that at least three distinct evaluation capabilities will be developed and used in this Program and that every technology developed under Area A will be evaluated in at least three distinct domains. This is to ensure that the technical approaches result in general transfer learning capabilities rather than domain-specific solutions.

The general approach to the evaluation is to use testbeds rather than datasets. A testbed may be thought of as a controlled, instrumented and interactive problem instance

generator. Unlike data sets that remain fixed and static, testbeds can generate problems over time; can generate particular instances as a series of steps that adapt to the problem solver/learner's behavior, that contain multiple agents, some of which may be non-cooperative or adversarial, that can have a lifetime much longer than a set of problem solving instances, that can intermix elements of different problems and types of problems, and that require a rich set of problem-solving behaviors that extend far beyond even incremental or real-time classification to include planning, strategizing, resource allocation, and multi-level problem solving. The testbeds must be instrumented and controlled so all interactions with the testbed can be recorded for analysis and so the space of parameter settings for the test problems can be explored systematically.

It is anticipated that of the two testbeds to be developed independently of the technology developers, one will be highly interactive and emphasize learning how to act or behave in an area such as computer-based, real-time strategy games and the other will be more descriptive and emphasize reasoning, inference, and problem-solving, likely involving extensions of existing standardized tests and using techniques widely employed for measuring performance in people.

A completed operational testbed capability must be available in month 10 of each year of the proposed effort so evaluations can be conducted in month 11. The capabilities of the testbeds may grow from year to year to reflect the increasing transfer learning capabilities that are being developed under area A of the Transfer Learning Program. An initial operational capability version of the testbed suitable for use in "dry runs" must be available by month 8 of each year. This dry-run capability shall be made available to all Transfer Learning Program technology development contractors for their use so they can prepare for the annual evaluation.

The testbeds and challenge problems should get progressively more difficult in each year of the Program; however, it is essential that each year's increased difficulty subsume the previous year's problem in a manner that clearly demonstrates improvements in transfer learning from year-to-year. There are many approaches that could meet this criterion. For example, a larger set of problems could be used in a subsequent year that contains instances of equivalent difficulty from the previous year's set. Another method would be to use problems of equivalent difficulty but require better performance.

4.2.1 Area B1 - Development of Challenge Problems and Testbed Implementations

The contractor shall develop a challenge problem and implement it in a testbed. The challenge problem must clearly evaluate progress in transfer learning technology. Isolation of transfer learning capabilities from other capabilities to the maximum extent possible is needed – i.e., success on the challenge problem should be dependent on the ability to perform transfer learning to the maximum extent and on any other factors to the minimum extent. A key criterion for a challenge problem is that the domain *not* require large amounts of pre-coded domain-specific or common sense knowledge beyond what is learned during training on the initial domain. (Of course, transfer learners may require significant amounts of general knowledge as part of the solution.) A second criterion is

that a simple language can be developed in which domain problems and tasks can be expressed. Domains with well-defined metrics of performance are preferable, especially where the performance metrics already exist for people to perform the task.

Offerors shall specify the domains and tasks they intend to use for this Area in the proposal. Proposals shall include:

1. The criteria used to select the domain and task
2. A clear explanation of how the domain and task will evaluate transfer learning.
3. A plan specifying the challenge problem and testbed capabilities for each year/phase of the Program.
4. A specification of what information will be provided to the transfer learning software for the "training" domain(s) and task(s).
5. A specification for the language in which domain knowledge and tasks will be represented.
6. A description of how the "target" domains and tasks will be selected.
7. An analysis of the commonalities between the proposed training and target domains that demonstrates why transfer learning would enable increased performance on the target domain.
8. An analysis of the degree and types of transfer required to achieve improved performance in the target domain. This analysis may be in terms of the 10-level model of transfer learning described in section 4.1.
9. A protocol for how the transfer learning technology developers with Area A contracts will be able to use the testbed for research support, training, and evaluation.

4.2.2 Area B2 – Conduct Evaluations and Report Results

This Area will involve measuring the performance of the capabilities developed in Area A on the challenge problems developed in Area B1 using the testbeds implemented in Area B1. The software developed by the technology development teams will be provided to the evaluation teams along with appropriate instruction, documentation, and consulting so it can be executed in the testbed against the challenge problems.

The envisioned high-level protocol for conducting an evaluation is as follows:

0. Specify training domain, task, representations, and metrics. Specify experimental design.
1. Train an untrained version of the transfer learning software by having it interact with the testbed on a set of problems. This can continue for a fixed amount of interactions or until the transfer learning software has achieved a specified level of performance. Measure the learning curves (i.e., performance on the task) as a function of all dependent variables.
2. Specify a novel target domain and task. The representation should be identical or minimally modified from that used in the training domain.

3. Measure learning curves of the trained version of the transfer learning software that resulted from Step 1 on the domain and task specified in Step 2.
4. Measure learning curves on an untrained version of the transfer learning software (i.e., a version identical to that available prior to Step 1) on the task domain and task specified in Step 2.
5. Compare the results from Step 3 and Step 4.
6. Conduct ablation experiments on the version of the transfer learning software used in Step 3 by selectively removing some of the knowledge learned in Step 1.
7. Analyze results. Repeat previous steps as needed to ensure that results are meaningful and valid. Ideally, of course, the experimental design determined in advance will make repetitions unnecessary.

Proposals should explain the experimental protocol and design in depth sufficient to demonstrate that they will yield meaningful and valid results. Proposals should also explain the process of conducting the evaluation to include issues such as these: how much real time will elapse; when transfer learning software versions will be delivered from the Area A technology developers to the Area B evaluators; how configuration management of the software versions will be performed; how – if at all – the Area A transfer learning technology developers will be needed to assist and/or permitted to observe the Area B evaluators; how results will be collected and analyzed; where the evaluation will occur; what resources will be necessary to conduct the evaluation; and any other actions needed to enable the evaluation process.

4.2.3 Area B3 - Metrics

The evaluation contractor is responsible for developing metrics for transfer learning and specifying protocols and experiments to measure the capabilities developed by the technology developers against these metrics. These metrics should be domain-independent to the maximum extent possible, although domain-specific instantiations should also be identified. The metrics should reflect the capabilities and approaches of the transfer learning technology developers as well as the overall goals of the Transfer Learning Program.

Metrics should be designed for the three types of transfer learning identified in Figure 1. Metrics may be expressed in terms of the model of levels of transfer learning presented in Section 4.1. above, or in any other clearly specifiable and measurable framework considered suitable by the offeror.

Metrics can be absolute or relative to the performance of people. Absolute metrics that can be calibrated by comparison to how people perform on the same task are preferred over those that can not be calibrated with respect to human performance. An example of relative metrics (and milestones) for the Program is depicted in Table 1. Note that the level of detail expected in the proposals for project metrics is greater than that depicted in the table.

Year	Within-domain lateral & vertical transfer	Cross-domain lateral transfer	Cross-domain vertical transfer
Y1	80%		
Y2	85%	80%	
Y3	90%	85%	80%

Table 1: Example of Relative Metrics and Milestones

This table would be interpreted as, for example, in year 2, the transfer learning software would exhibit cross-domain lateral transfer at a level equivalent to the 80th percentile of people on the same task; i.e., given the same training and evaluation problems. Note how this example also depicts both improved performance on the same task from year to year and increased difficulty of the task set from year to year, as discussed in the final paragraph preceding Section 4.2.1.

5 GENERAL INFORMATION

Proposals not meeting the format described in this pamphlet may not be reviewed. Proposals MUST NOT be submitted by fax or e-mail; any so sent will be disregarded. This notice, in conjunction with the BAA 05-29 FedBizOpps Announcement and all references, constitutes the total BAA.

A Frequently Asked Questions (FAQ) list may be provided. If so, it will be found on the DARPA/IPTO Solicitation page at http://www.darpa.mil/ipto/solicitations/solicitations.htm.

No additional information is available, nor will a formal Request for Proposal (RFP) or other solicitation regarding this announcement be issued. Requests for same will be disregarded.

All responsible sources capable of satisfying the Government's needs may submit a proposal that shall be considered by DARPA. Small Disadvantaged Businesses, Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) are encouraged to submit proposals and join others in submitting proposals. However, no portion of this BAA will be set aside for Small Disadvantaged Business, HBCU and MI participation due to the impracticality of reserving discrete or severable areas of this research for exclusive competition among these entities.

The Government anticipates that proposals submitted under this BAA will be unclassified. In the event that a proposer chooses to submit a classified proposal or submit any documentation that may be classified, the following information is applicable. Security classification guidance on a DD Form 254 will not be provided at this time since DARPA is soliciting ideas only. After reviewing incoming proposals, if a determination is made that the award instrument may result in access to classified information, a DD

Form 254 will be issued and attached as part of the award. Proposers choosing to submit a classified proposal must first receive permission from the Original Classification Authority to use their information in replying to this BAA. Applicable classification guide(s) should be submitted to ensure that the proposal is protected appropriately.

Proposals selected for funding are required to comply with provisions of the Common Rule (32 CFR 219) on the protection of human subjects in research (http://www.dtic.mil/biosys/downloads/32cfr219.pdf) and the Department of Defense Directive 3216.2 (http://www.dtic.mil/whs/directives/corres/html2/d32162x.htm). All proposals that involve the use of human subjects are required to include documentation of their ability to follow Federal guidelines for the protection of human subjects. This includes, but is not limited to, protocol approval mechanisms, approved Institutional Review Boards, and Federal Wide Assurances. These requirements are based on expected human use issues sometime during the entire length of the proposed effort.

For proposals involving "greater than minimal risk" to human subjects within the first year of the project, performers must provide evidence of protocol submission to a federally approved IRB *at the time of final proposal submission to DARPA.* For proposals that are forecasted to involve "greater than minimal risk" after the first year, a discussion on how and when the proposer will comply with submission to a federally approved IRB needs to be provided in the submission. More information on applicable federal regulations can be found at the Department of Health and Human Services – Office of Human Research Protections website (http://www.dhhs.gov/ohrp/).

Any public release of information developed as part of any contract awarded against this BAA must be prepared in accordance with DARPA's Public Release Policy and Procedures, which are available at http://www.darpa.mil/tio/ .

6 SUBMISSION PROCESS

This BAA requires completion of an online Cover Sheet for each Proposal prior to submission. To do so, the offeror must go to **http://www.dyncorp-is.com/BAA/index.asp?BAAid=05-29** and follow the instructions there. Each offeror is responsible for printing the BAA Confirmation Sheet and attaching it to every copy. The Confirmation Sheet should be the first page of the Proposal. If an offeror intends to submit more than one Proposal, a unique UserId and password must be used in creating each Cover Sheet. Failure to comply with these submission procedures may result in the submission not being evaluated.

Proposers must submit the original and 2 copies (3 total) of the full proposal and 2 electronic copies (i.e., 2 separate disks) of the full proposal (in PDF or Microsoft Word 2000 for IBM-compatible format on a 3.5-inch floppy disk or CD). **Mac-formatted disks will not be accepted**. Each disk must be clearly labeled with BAA 05-29, proposer organization, proposal title (short title recommended) and "Copy ___ of 2". The full proposal (original and designated number of hard and electronic copies) must be submitted in time to reach DARPA by 12:00 PM (ET) April 29, 2005 in order to be

considered during the initial evaluation phase. However, BAA 05-29, Transfer Learning (TL) will remain open until 12:00 NOON (ET) March 14, 2006. Thus, proposals may be submitted at any time from issuance of this BAA through March 14, 2006. While the proposals submitted after the April 29, 2005 deadline will be evaluated by the Government, proposers should keep in mind that the likelihood of funding such proposals is less than for those proposals submitted in connection with the initial evaluation and award schedule. DARPA will acknowledge receipt of submissions. The acknowledgement receipt (sent via email) will contain a control number that should be used in all correspondence regarding your proposal.

Restrictive notices notwithstanding, proposals may be handled for administrative purposes by support contractors. These support contractors are prohibited from competition in DARPA technical research and are bound by appropriate non-disclosure requirements. Input on technical aspects of the proposals may be solicited by DARPA from non-Government consultants/experts who are also bound by appropriate non-disclosure requirements. However, non-Government technical consultants/experts will not have access to proposals that are labeled by their offerors as "Government Only." Use of non-government personnel is covered in FAR 37.203(d).

7 REPORTING REQUIREMENTS/PROCEDURES

The Award Document for each proposal selected and funded will contain a mandatory requirement for submission of DARPA/IPTO Quarterly Status Reports and an Annual Project Summary Report. These reports, described below, will be electronically submitted by each awardee under this BAA via the DARPA/IPTO Technical - Financial Information Management System (T-FIMS). The T-FIMS URL will be furnished by the Government upon award. Detailed data requirements can be found in the Data Item Description (DID) DI-MISC-81612A available on the Government's ASSIST database (http://assist.daps.dla.mil/quicksearch/).

8 PROPOSAL FORMAT

Proposals shall consist of a cover page, a technical volume, and a cost volume. The submission of other supporting materials—including bibliographies, technical papers, and research notes—along with the proposal is strongly discouraged.

A "page" is 8-1/2 by 11 inches with type not smaller than 12 point, and with text on one side only.

8.1 Cover Page

The cover page shall be a single page containing the following information.

1. BAA number
2. Proposal title
3. Lead Organization submitting proposal

4. Contractor's type of business, selected from among the following categories: "WOMEN-OWNED LARGE BUSINESS," "OTHER LARGE BUSINESS," "SMALL DISADVANTAGED BUSINESS [Identify ethnic group from among the following: Asian-Indian American, Asian-Pacific American, Black American, Hispanic American, Native American, or Other]," "WOMEN-OWNED SMALL BUSINESS," "OTHER SMALL BUSINESS," "HBCU," "MI," "OTHER EDUCATIONAL," "OTHER NONPROFIT", or "FOREIGN CONCERN/ENTITY."
5. Other team members (if applicable) and type of business for each
6. Technical point of contact to include: salutation, last name, first name, street address, city, state, zip code, telephone, fax (if available), electronic mail (if available)
7. Administrative point of contact to include: salutation, last name, first name, street address, city, state, zip code, telephone, fax (if available), electronic mail (if available)
8. Total funds requested from DARPA, and the amount of cost-share (if any)
9. Summary of the costs of the proposed research, including total base cost, estimates of base cost in each year of the effort, estimates of itemized options in each year of the effort, and cost sharing if relevant;
10. Date proposal was prepared

8.2 Volume I. Technical

This volume provides the detailed discussion of the proposed work necessary to enable an in-depth review of the specific technical and management issues. Specific attention must be given to addressing both the risk and payoff of the proposed work that make it desirable to DARPA.

The Technical Volume shall not exceed 50 pages, and shall include sections A through J, each beginning on a new page. Maximum page lengths for each section are shown in braces { } below, where applicable.

A. Innovative claims for the proposed research{1 Page}. This page is the centerpiece of the proposal and should succinctly describe the unique proposed contribution.

B. Proposal Roadmap{1 Page}. The roadmap provides a top-level view of the content and structure of the proposal. It contains a synopsis (or “sound bite”) for each of the nine areas defined below. It is important to make the synopses as explicit and informative as possible. The roadmap must also cross-reference the proposal page number(s) where each area is elaborated. The nine roadmap areas are:

1. Main goals of the proposed research.
2. Tangible benefits to future efforts (i.e., benefits of the capabilities afforded if the proposed technology is successful).
3. Critical technical barriers (i.e., technical limitations that have, in the past, prevented achieving the proposed results).
4. Main elements of the proposed approach.

5. Rationale that builds confidence that the proposed approach will overcome the technical barriers. ("We have a good team and good technology" is not a useful statement.)
6. Nature of expected results (unique/innovative/critical capabilities to result from this effort, and form in which they will be defined).
7. The risk if the work is not done.
8. Criteria for scientifically evaluating progress and capabilities on an annual basis.
9. Cost of the proposed effort for each performance year.

C. Statement of Work {3 Pages}. Detailed statement of work, written in plain English, outlining the scope of the effort and citing specific work to be performed, references to specific subcontractors if applicable, and specific contractor requirements.

D. Research Objectives {8 Pages}

1. Problem Description. Provide concise description of problem area addressed by this research project.
2. Research Goals. Identify specific research goals of this project. Identify and quantify expected performance improvements from this research. Identify new capabilities enabled by this research. Identify and discuss salient features and capabilities of developmental hardware and software prototypes. Provide a set of metrics and success criteria for the concepts proposed under each phase.
3. Expected Impact. Describe expected impact of the research project, if successful, to problem area.

E. Technical Approach:

1. Detailed Description of Technical Approach {20 Pages}. Provide detailed description of technical approach that will be used in this project to achieve research goals. Specifically identify and discuss the innovative aspects of the TL technical approach. Note: An optional technical viewgraph summary in MS Power Point format (maximum of 8 viewgraphs) may also be included as part of the Technical Volume and will not be considered as part of the volume page count.
2. Comparison with Current Technology {3 Pages}. Describe state-of-the-art approaches and the limitations within the context of the problem area addressed by this research.

F. Schedule and Milestones

1. Schedule Graphic {1 Page}. Provide a graphic representation of project schedule including detail down to the individual effort level. This should include but not be limited to, a multi-phase development plan, which demonstrates a clear understanding of the proposed research. Show all project milestones. Use absolute designations for all dates.
2. Detailed Individual Effort Descriptions {3 Pages}. Provide detailed work effort descriptions for each individual effort and/or subcontractor in schedule graphic.

G. Deliverables Description {4 Pages}. List and provide detailed description for each proposed deliverable. Include in this section all proprietary claims to results, prototypes, or systems supporting and/or necessary for the use of the research, results, and/or prototype. If there are no proprietary claims, this should be stated. The offeror must submit a separate list of all technical data or computer software that will be furnished to the Government with other than unlimited rights (see DFARS 227.) Specify receiving organization and expected delivery date for each deliverable.

H. Technology Transition and Technology Transfer Targets and Plans {2 Page}. Discuss path for technology transition and transfer in each phase of the Program. Offerors should identify transition opportunities for appropriate technology components and information to the user community.

I. Personnel and Qualifications {3 Pages}. List of key personnel, concise summary of their qualifications, and discussion of proposer's previous accomplishments and work in this or closely related research areas. Indicate the level of effort to be expended by each person during each contract year and other (current and proposed) major sources of support for them and/or commitments of their efforts. DARPA expects all key personnel associated with a proposal to make substantial time commitment to the proposed activity.

J. Facilities {1 Page}. Description of the facilities that would be used for the proposed effort. If any portion of the research is predicated upon the use of Government Owned Resources of any type, the offeror shall specifically identify the property or other resource required, the date the property or resource is required, the duration of the requirement, the source from which the resource is required, if known, and the impact on the research if the resource cannot be provided. If no Government Furnished Property is required for conduct of the proposed research, the proposal shall so state.

K. Organizational Conflict of Interest: Awards made under this BAA may be subject to the provisions of the Federal Acquisition Regulation (FAR) Subpart 9.5, Organizational Conflict of Interest. All offerors and proposed subcontractors must affirmatively state whether they are supporting any DARPA technical office(s) through an active contract or subcontract. All affirmations must state which office(s) the offeror supports, and identify the prime contract number. Affirmations should be furnished at the time of proposal submission. All facts relevant to the existence or potential existence of organizational conflicts of interest, as that term is defined in FAR 2.101, must be disclosed, organized by task and year. This disclosure shall include a description of the action the Contractor has taken, or proposes to take, to avoid, neutralize, or mitigate such conflict.

8.3 Volume II. Cost

Cost proposals are subject to no page limits, and shall provide a detailed cost breakdown of all direct costs, including cost by Area and Sub-Area, with breakdown into accounting categories (labor, material, travel, computer, subcontracting costs, labor and overhead rates, and equipment), for the entire contract and for each calendar year, divided into

quarters. Where the effort consists of multiple portions that could reasonably be partitioned for purposes of funding, these should be identified as contract options with separate cost estimates for each.

Offerors should expect to participate in teams and workshops to provide specific technical background information to DARPA, attend semi-annual Principal Investigator (PI) meetings, and participate in numerous other coordination meetings via teleconference or Video Teleconference (VTC). Funding to support these various group experimentation efforts should be included in technology project bids

Offerors requiring the purchase of information technology (IT) resources as Government Furnished Property (GFP) MUST attach the following information:

1. A letter on Corporate letterhead signed by a senior corporate official and addressed to Ted Senator, Program Manager, DARPA/IPTO, stating that you either can not or will not provide the information technology (IT) resources necessary to conduct the said research.
2. An explanation of the method of competitive acquisition or a sole source justification, as appropriate, for each IT resource item.
3. If the resource is leased, a lease purchase analysis clearly showing the reason for the lease decision.
4. The cost for each IT resource item.

IMPORTANT NOTE: IF THE OFFEROR DOES NOT COMPLY WITH THE ABOVE STATED REQUIREMENTS, THE PROPOSAL WILL BE REJECTED.

9 EVALUATION AND FUNDING PROCESSES

Proposals will not be evaluated against each other, since they are not submitted in accordance with a common work statement. DARPA's intent is to review proposals as soon as possible after they arrive; however, proposals may be reviewed periodically for administrative reasons. For evaluation purposes, a proposal is the document described in Proposal Format. Other supporting or background materials submitted with the proposal will be considered for the reviewer's convenience only and not considered as part of the proposal.

Evaluation of proposals will be accomplished through a scientific review of each proposal using the following criteria, which are listed in descending order of relative importance:

(1) Capabilities to be Achieved: The amounts, levels, and types of transfer learning capability to be achieved in each phase must be clearly identifiable and compelling. These transfer learning capabilities to be developed should be novel and useful. The goals should be ambitious and achievable. The scientific and technical breakthroughs should be specified and significant. Offerors must specify quantitative experimental

methods and metrics for measuring progress of the effort against the proposed goals in each phase of the proposed effort.

(2) Technical Innovation and Feasibility: Offerors should apply new and/or existing technology in an innovative way that supports the objectives of the proposed effort. The specific technical approach(es) to be followed should show breadth of innovation and depth of understanding across all dimensions of the proposed solution. The technical concepts should be clearly defined and developed. The technical approach must be sufficiently detailed to support the proposed concepts, technical claims, and achievement of the proposed goals for new transfer learning capabilities.

(3) Potential Contribution and Relevance to DARPA/IPTO Mission: The offeror must clearly address how the proposed effort will meet the goals of the undertaking and how the proposed effort contributes to significant advances to DARPA/IPTO.

(4) Offeror's Capabilities and Related Experience: The qualifications, capabilities, and demonstrated achievements of the proposed principals and other key personnel for the primary and subcontractor organizations must be clearly shown. Key personnel should be clearly identified and their proposed level of effort (and their other commitments) specified.

(5) Potential to Accomplish Technology Transition: The offeror should identify specific potential military applications of the proposed technology. A plan that identifies specific deliverables available as transitionable products or capabilities and the time at which they will be available (during any of the phases, up to and including at the end of the Program) should be included.

(6) Cost Realism: The overall estimated costs should be clearly justified and appropriate for the technical complexity of the effort. Evaluation will consider the value of the research to the government and the extent to which the proposed management plan will effectively allocate resources to achieve the capabilities proposed.

The Government reserves the right to select all, some, or none of the proposals received in response to this solicitation and to make awards without discussions with offerors; however, the Government reserves the right to conduct discussions if the Source Selection Authority later determines them to be necessary. Proposals identified for funding may result in a contract, grant, cooperative agreement, or other transaction depending upon the nature of the work proposed, the required degree of interaction between parties, and other factors. If warranted, portions of resulting awards may be segregated into pre-priced options.

10 Administrative Addresses

The administrative addresses for this BAA are:

Fax: 703-741-7804 Addressed to: DARPA/IPTO, BAA 05-29

Electronic Mail: baa05-29@darpa.mil

Electronic File Retrieval: http://www.darpa.mil/ipto/Solicitations/solicitations.htm

Mail to: DARPA/IPTO
ATTN: BAA 05-29
3701 N. Fairfax Drive
Arlington, VA 22203-1714