Forking symmetry and the Vapnik-Chervonenkis theorem: talk at La Roche meeting

Anand Pillay

25 April 2008

Invariant types I

We fix a complete theory T, and work in a monster model $\bar{M}$.

Definition 1.1

Let M be a saturated model, and $p(x) \in S_x(M)$. With some abuse of language we call p invariant if for some small (with respect to $|M|$) set A (which can be taken to be a model), p is invariant under $Aut(M/A)$, namely is fixed by all automorphisms of M which fix A pointwise.

- Among examples of invariant types are (i) $p \in S_x(M)$ which is definable, and (ii) $p \in S_x(M)$ which is finitely satisfiable in some small model.
- If $p \in S_x(M)$ is invariant, witnessed by A then for $\phi(x, y) \in L$, whether or not $\phi(x, b) \in p$, depends only on $tp(b/A)$, hence p has a kind of "infinitary definition" over A.

Invariant types II

- So for any set of parameters $B \supseteq A$ (not necessarily contained in M), we can talk about $p|B$, the complete x-type over B obtained by applying the infinitary definition to B.
- (M saturated) If $p(x) \in S_x(M)$ is invariant and $q(y) \in S_y(M)$ is arbitrary we can form $p_x \ltimes q_y \in S_{xy}(M)$ which will be $tp(a, b/M)$ where b realizes q and a realizes $p|Mb$.
- If also $q(y)$ is invariant we can form $q_y \ltimes p_x \in S_{xy}(M)$, which is $tp(a', b'/M)$ where now a' realizes p and b' realizes $q|Ma'$. (So either way, we get an xy type.)
- When both $p(x)$, $q(y)$ are invariant, so is $p_x \ltimes q_y$.
- We can also ask whether symmetry holds: $p_x \ltimes q_y = q_y \ltimes p_x$.
- If T is stable, then any type over M is invariant and we always have symmetry.

A symmetry theorem for types I

With previous notation (M saturated etc.):

Lemma 1.2

Let $p(x) \in S_x(M)$ be finitely satisfiable (in some small model) and $q_y \in S_y(M)$ be definable. Then $p_x \ltimes q_y = q_y \ltimes p_x$.

Proof.

- This is very basic and amounts to the uniqueness of heirs for definable types. But we will give a brief proof as we want to generalize it later.
- Suppose for a contradiction that $\phi(x, y) \in p_x \ltimes q_y$, but $\neg\phi(x, y) \in q_y \ltimes p_x$.
- Let $\psi(x) \in L_M$ be the $\phi(x, y)$-definition for $q(y)$. Hence $\neg\psi(x) \in p(x)$.
- Let b realize q and a realize $p|Mb$. Then $\models \phi(a, b)$, and moreover $tp(a/Mb)$ is finitely satisfiable in M.

A symmetry theorem for types II

- On the other hand we saw that $\models \neg\psi(a)$.
- By finite satisfiability, there is $a' \in M$ such that $\models \neg\psi(a') \wedge \phi(a', b)$.
- This contradicts $\psi(x)$ being the $\phi(x, y)$-definition of $q(y) = tp(b/M)$. End of proof.

- We want to generalize Lemma 1.2 to "Keisler measures".
- Before doing this let us remark that if T has NIP then a type $p(x) \in S(M)$ is invariant if and only if p does not fork (divide) over some small set.

Definition 1.3

A Keisler measure μ_x over A is a finitely additive probability measure (with values in the real unit interval) on formulas $\phi(x) \in L_A$ (A-definable sets of sort x), namely $\mu_x(x = x) = 1$, $\mu_x(x \neq x) = 0$, and for disjoint A-definable subsets X, Y of sort x, $\mu(X \cup Y) = \mu(X) + \mu(Y)$.

- A special case of a Keisler measure μ_x over A is a complete type $p(x) \in S_x(A)$
- A Keisler measure μ_x over A is the "same thing" as a regular Borel probability measure on the Stone space $S_x(A)$.

Keisler measures II

- Let M be a saturated model, and μ_x a Keisler measure over M. Let A be a small subset of M.
- We say that μ is definable (Borel definable) over A if for each $\phi(x,y) \in L$ and closed $C \subseteq [0,1]$, $\{b \in M : \mu(\phi(x,b)) \in C\}$ is type-definable over A (Borel over A, explain).
- If μ_x, λ_y are Keisler measures over (saturated) M, and μ is Borel definable over some small set, then we can form $\mu_x \ltimes \lambda_y$, a Keisler measure in xy over M:
- Given $\phi(x,y)$ (over M), $(\mu_x \ltimes \lambda_y)(\phi(x,y)) =_{def} \int \mu(\phi(x,y)) d\lambda_y$. (Explain.)
- If also λ_y is Borel definable over a small set, then we can also form $\lambda_y \ltimes \mu_x$ and we can again ask about symmetry.

Keisler measures III

- For μ_x a Keisler measure over (saturated) M, we say that μ is finitely satisfiable if it is finitely satisfiable in some small model M_0, that is, any formula of positive μ-measure is realized by an element of M_0.
- Likewise for "μ does not fork (divide) over A" (if you know what that means).
- If T has NIP and μ_x is a Keisler measure over (saturated) M, and M_0 is a small submodel of M, then μ is $Aut(M/M_0)$-invariant iff μ is Borel definable over M_0 iff μ does not fork over M_0.
- Moreover in the situation of the last bullet: if $\phi(x,y) \in L$ and $tp(b/M_0) = tp(c/M_0)$ then $\mu(\phi(x,b)\Delta\phi(x,c)) = 0$.
- In any case, assuming NIP, a Keisler measure μ over (saturated) M is "invariant" iff it is Borel definable over small set, so for μ, λ invariant measures over M we can form $\mu \ltimes \lambda$ and $\lambda \ltimes \mu$.

A Symmetry Theorem

Theorem 1.4

(T has NIP.) Suppose M is saturated, μ_x, λ_y are Keisler measures over M, and μ is finitely satisfiable (in a small model) and λ is definable. Then $\mu_x \ltimes \lambda_y = \lambda_y \ltimes \mu_x$.

- The theorem yields, among other things, a quick proof of the uniqueness of translation invariant Keisler measures on a certain class of definable groups (namely, groups with fsg) in NIP theories.
- This latter uniqueness theorem is in a sense a common generalization of the uniqueness of Haar measure on compact groups, and the uniqueness of translation invariant types in connected stable groups. (Algebraic groups case?)
- Be careful! A definable group is NEVER compact, unless it is finite.
- The uniqueness result already appears in "NIP II" but with a complicated proof.

The Vapnik-Chevonenkis theorem, or inequality I

- Theorem 1.4 is proved using the VC theorem which was already used in NIP II to prove for example the Borel definability of invariant Keisler measures.
- We set up notation towards stating the VC theorem.
- Let X be a probability space; that is a set equipped with a σ-algebra Ω of subsets (or events), and a countably additive measure μ with values in $[0,1]$ such that $\mu(X) = 1$, $\mu(\emptyset) = 0$, and $\mu(E)$ is defined for any $E \in \Omega$.
- For $A \subseteq X$ an event and $\bar{p} = (p_1, .., p_k) \in X^k$, let $fr_k(A, \bar{p})$ be the proportion of p_i's which are in A.
- For $k > 0$ let μ^k be the product measure on X^k.

The Vapnik-Chevonenkis theorem, or inequality II

- A family $\mathcal{C}$ of events is said to have NIP (or finite VC-dimension), if there is some natural number d, such that for no subset F of X of cardinality d, is $\{F \cap A : A \in \mathcal{C}\}$ the full power set of F.
- We fix such a family $\mathcal{C}$ with NIP.

Theorem 1.5 (Vapnik-Chervonenkis)

For any $\epsilon > 0$, $\mu^k(\{\bar{p} \in X^k : sup_{A \in \mathcal{C}} |fr_k(A, \bar{p}) - \mu(A)| > \epsilon\}) \to 0$ *as* $k \to \infty$ *(as long as the set in brackets is* μ^k*-measurable, which for example holds if* $\mathcal{C}$ *is countable.)*

The Vapnik-Chevonenkis theorem, or inequality III

- Given T with NIP, and μ_x a Keisler measure over M, and $\phi(x,y) \in L$, we will apply Theorem 1.5 to the Stone space $S_x(M)$ equipped with the Borel probability measure (given by) μ, and with $\mathcal{C}$ the family of clopen subsets of $S_x(M)$ given by instances of $\phi(x,y)$.
- We will for now ignore the measurability proviso in Theorem 1.5. We can get around it by using compactness or working in countable models of countable reducts.

Let us give a couple of consequences of Theorem 1.5, the first immediate, the second requiring a bit of work. We will be vague regarding approximations, now and for the rest of the slides.

The Vapnik-Chevonenkis theorem, or inequality IV

Lemma 1.6

(T has NIP.) Let μ_x be a Keisler measure over M. Fix $\phi(x,y) \in L$. Then for sufficiently large k, there are many (in the sense of μ^k on $S_x(M)^k$) , sequences $p_1,..,p_k \in S_x(M)$, each p_i being random for μ, such that for any $b \in M$, $\mu(\phi(x,b))$ is approximated by the proportion of p_i' containing $\phi(x,b)$.

Lemma 1.7

(T has NIP). Let μ_x, λ_y be Keisler measures over saturated M, both Borel definable over small M_0 (in fact Borel definability of λ suffices). Fix $\phi(x,y) \in L$, and suppose that $(\mu_x \times \lambda_y)(\phi(x,y)) = r$. Then for large k, and "many" $q_1,..,q_k \in S_y(M_0)$, and any realizations $b_1,..,b_k$ of q_1,..,q_k, for "many" $p_1,..,p_k \in S_x(M)$, and any realizations $a_1,..,a_k$ of $p_1,..,p_k$, the proportion of (a_i,b_j) which satisfy $\phi(x,y)$ approximates r.

"Approximation" to proof of Theorem 1.4 I

- Let μ_x over M be finitely satisfiable (in M_0), and λ_y over M be definable (over M_0, and $\phi(x,y) \in L$, and suppose for a contradiction that $(\mu_x \ltimes \lambda_y)(\phi(x,y)) = s$, but $(\lambda_y \ltimes \mu_x)(\phi(x,y)) = r$ and $r \neq s$.
- Let us simplify things by supposing that the integral $\int \lambda_y(\phi(x,y))d\mu_x = r$ is a finite sum (of course it IS approximated by such).
- Namely there are $r_1, .., r_n > 0$ and $t_1, .., t_n > 0$ such that $t_1 + .. + t_n = 1$, $t_1 r_1 + t_2 r_2 + .. + t_n r_n = r$, $\theta_i(x)$ (over M_0) defines $\{a \in M : \lambda(\phi(a,y)) = r_i\}$, $\mu(\theta_i(x)) = t_i$ and the θ_i partition x-space. (So all this is really true up to some approximation.)

"Approximation" to proof of Theorem 1.4 II

- By Lemmas 1.6 and 1.7, for some large k, choose $q_1, .., q_k \in S_y(M_0)$ and realizations $b_1, .., b_k$, then $p_1, .., p_k \in S_x(M)$ and realizations $a_1, .., a_k$ such that:
- (i) The conclusion of Lemma 1.7 holds.
- (ii) $q_1(y), .., q_k(y)$ are "good" for $\phi(x, y)$ and parameters from M_0.
- (iii) $p_1, .., p_k$ are "good" for the finite set of formulas $\theta_1(x), .., \theta_n(x)$, and
- (iv) $tp(a_1, .., a_k/M)$ is finitely satisfiable in M_0.
- (iii)' So by (iii) for each $i = 1, .., n$, approximately $t_i k$ of the a_j's satisfy $\theta_i(x)$.
- By (iv) we can find $a'_1, ..., a'_n \in M_0$ such that:
- (v) for each $i = 1, .., n$ and $j = 1, .., k$, we have that $\models \theta_i(a_j) \leftrightarrow \theta_i(a'_j)$, and
- (vi) for each $i, j = 1, .., k$ we have $\models \phi(a_i, b_j) \leftrightarrow \phi(a'_i, b_j)$.

"Approximation" to proof of Theorem 1.4 III

- By (i) and (vi), we have
- (vii) $|\{(a'_i, b_j) :\models \phi(a'_i, b_j)\}|$ is approximately $k^2 s$.
- By (ii) whenever $\theta_i(a'_j)$ then approximately $r_i k$ of the $b_1, .., b_n$ satisfy $\phi(a'_j, y)$.
- By (v) and (iii)' for each $i = 1, .., n$, approximately $t_i k$ among $a'_1, .., a'_k$ satisfy $\theta_i(x)$.
- The last two items imply that $|\{(a'_i, b_j) :\models \phi(a'_i, b_j)\}|$ is approximately $k^2 r$, which is a contradiction to (vii).